Electronic Thesis and Dissertation Repository

3-3-2022 1:45 PM

The Role of Transient Vibration of the Skull on Concussion

Rodrigo Dalvit Carvalho da Silva, *The University of Western Ontario*

Supervisor: Jenkyn, Thomas R., *The University of Western Ontario*
A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Biomedical Engineering
© Rodrigo Dalvit Carvalho da Silva 2022

Recommended Citation

Abstract

Concussion is a traumatic brain injury usually caused by a direct or indirect blow to the head that affects brain function. The maximum mechanical impedance of the brain tissue occurs at 450±50 Hz and may be affected by the skull resonant frequencies. After an impact to the head, vibration resonance of the skull damages the underlying cortex. The skull deforms and vibrates, like a bell for 3 to 5 milliseconds, bruising the cortex. Furthermore, the deceleration forces the frontal and temporal cortex against the skull, eliminating a layer of cerebrospinal fluid. When the skull vibrates, the force spreads directly to the cortex, with no layer of cerebrospinal fluid to reflect the wave or cushion its force. To date, there is few researches investigating the effect of transient vibration of the skull. Therefore, the overall goal of the proposed research is to gain better understanding of the role of transient vibration of the skull on concussion. This goal will be achieved by addressing three research objectives. First, a MRI skull and brain segmentation automatic technique is developed. Due to bones' weak magnetic resonance signal, MRI scans struggle with differentiating bone tissue from other structures. One of the most important components for a successful segmentation is high-quality ground truth labels. Therefore, we introduce a deep learning framework for skull segmentation purpose where the ground truth labels are created from CT imaging using the standard tessellation language (STL). Furthermore, the brain region will be important for a future work, thus, we explore a new initialization concept of the convolutional neural network (CNN) by orthogonal moments to improve brain segmentation in MRI. Second, the creation of a novel 2D and 3D Automatic Method to Align the Facial Skeleton is introduced. An important aspect for further impact analysis is the ability to precisely simulate the same point of impact on multiple bone models. To perform this task, the skull must be precisely aligned in all anatomical planes. Therefore, we introduce a 2D/3D technique to align the facial skeleton that was initially developed for automatically calculating the craniofacial symmetry midline. In the 2D version, the entire concept of using cephalometric landmarks and manual image grid alignment to construct the training dataset was introduced. Then, this concept was extended to a 3D version where coronal and trans-

verse planes are aligned using CNN approach. As the alignment in the sagittal plane is still undefined, a new alignment based on these techniques will be created to align the sagittal plane using Frankfort plane as a framework. Finally, the resonant frequencies of multiple skulls are assessed to determine how the skull resonant frequency vibrations propagate into the brain tissue. After applying material properties and mesh to the skull, modal analysis is performed to assess the skull natural frequencies. Finally, theories will be raised regarding the relation between the skull geometry, such as shape and thickness, and vibration with brain tissue injury, which may result in concussive injury.

Keywords: Convolutional Neural Network, Skull Segmentation, Brain Segmentation, Orthogonal Moments, Geometric Moments, Cephalometric analysis, STL, MRI, CT, ICA, PCA, k-NN, k-Means, Finite Element Analysis, Modal Analysis

Summary for Lay Audience

A concussion is a traumatic brain injury usually caused by a direct or indirect blow to the head that affects brain function. As the maximum mechanical impedance of the brain tissue occurs at 450±50 Hz, skull resonant frequencies may play an important role in the propagation of this vibration into the brain tissue. The overall goal of the proposed research is to gain a better understanding of the role of transient vibration of the skull on concussion. This goal will be achieved by addressing three research objectives: I) develop an automatic method to segment/extract skull and brain from magnetic resonance imaging (MRI), II) create a novel 2D and 3D automatic method to align the facial skeleton, and III) identify the skull resonant frequencies and raise the theory of how these vibrations may propagate into brain tissue. For objective 1, 58 MRI and their respective computed tomography (CT) scans were used to create a convolutional neural network framework for skull and brain segmentation in MRI. Moreover, an invariant moment kernel was introduced to improve the brain segmentation accuracy in MRI. For objective 2, a 2D and 3D technique for automatically calculating the craniofacial symmetry midline from head CT scans using deep learning techniques was used to precisely align the facial skeleton for future impact analysis. In objective 3, several skulls segmented were tested to identify their natural resonant frequencies. Those with a resonant frequency of 450±50 Hz were selected to improve understanding of how their shapes and thickness may help the vibration to propagate deeply in the brain tissue. The results from this study will improve our understanding of the role of transient vibration of the skull on concussion.

Acknowlegements

Dedication

To my wife Sabrina and my children Sophie and Henry, who inspire me and make me so proud, without whom this thesis would never be possible

Thank you my beloved wife for your kindness and love. Your encouragement and support of me goes beyond what words can adequately express. I hope the sacrifices you have endured for me to pursue this dream will be repaid to you. You are my love and I am so thankful we did this together

My sweet Sophie for your sweetness, when I see your inspirational eyes, all the difficulties of my life become easy

My buddy Henry for your unlimited happiness and pleasure

You have made me stronger, better, and more fulfilled than I could have ever imagined. I love you to the moon and back
O amor de vocês é o maior presente que já recebi em minha vida! Eu amo tanto vocês que chega a dor aqui ...

Co-Authorship Statement

Chapter 1:
    Rodrigo Dalvit Carvalho da Silva - sole author

Chapter 2:
    Rodrigo Dalvit Carvalho da Silva - Conceptualization, Methodology, Software,Writing
    Thomas Jenkyn - Project Administration
    Victor Carranza - Review and Editing

Chapter 3:
    Rodrigo Dalvit Carvalho da Silva - Conceptualization, Methodology, Software,Writing
    Thomas Jenkyn - Project Administration
    Victor Carranza - Review and Editing

Chapter 4:
    Rodrigo Dalvit Carvalho da Silva - Conceptualization, Methodology, Software,Writing
    Thomas Jenkyn - Project Administration
    Victor Carranza - Review and Editing

Chapter 5:
    Rodrigo Dalvit Carvalho da Silva - Conceptualization, Methodology, Software,Writing
    Thomas Jenkyn - Project Administration
    Victor Carranza - Review and Editing

Chapter 6:
    Rodrigo Dalvit Carvalho da Silva - Conceptualization, Methodology, Software,Writing
    Thomas Jenkyn - Project Administration
    Victor Carranza - Review and Editing

Chapter 7:
    Rodrigo Dalvit Carvalho da Silva - sole author

Contents

List of Figures

List of Tables

List of Abbreviations

AD	Average Difference
ANN	Artificial Neural Network
CH	Channel
CNN	Convolutional Neural Network
CONC	Concatenation
CONV	Convolution
CR	Computed Radiography
cRBM	Convolutional Restricted Boltzmann Machines
CSF	Cerebro-Spinal Fluid
CSF	Cerebrospinal Fluid
CT	Computed Tomography
CTE	Chronic Traumatic Encephalopathy
DAI	Diffuse Axonal Injury
DCNN	Deep Convolutional Neural Network
DECON	Deconvolution
DNA	Deoxyribonucleic Acid
DSC	Dice Similarity Coefficient
FE	Finite Element
FLAIR	Fluid-Attenuated Inversion Recovery
FN	False Negatives
FP	False Positives
FRF	Frequency Response Function
GB	Gigabyte
GDL	Generalized Dice Loss
GM	Geometric Moment
GPa	Gigapascal
GPUs	Graphics Processing Units
HD	Hausdorf Distances
ICA	Independent Component Analysis
ICP	Iterative Closest Point
ICs	Independent Components
IQI	Image Quality Index
JSC	Jaccard Similarity Coefficient
k-NN	k-Nearest Neighbours
L	Legendre Moment Image
LM	Legendre Moments

LMSE	Laplacian Mean Square Error
MD	Maximum Difference
MP	Max Pooling
MRI	Magnetic Resonance Imaging
MSE	Mean-Square Error
MSP	Midsagittal Plane
mTBI	Mild Traumatic Brain Injury
NAE	Normalized Absolute Error
NFBS	Neurofeedback Skull-Stripped Repository
NK	Normalized Cross-Correlation
NPV	Negative Predictive Value
O	Original Image
OASIS	Open Acess Series Of Imaging Studies
PC	Personal Computer
PCA	Principal Component Analysis
PCs	Principal Components
PCS	Post-Concussion Syndrome
PPV	Positive Predictive Value
PT Or PET	Positron Emission Tomography
PZ	Pseudo-Zernike Moment Image
PZM	Pseudo-Zernike Moments
PZP	Pseudo-Zernike Polynomial
RELU	Rectified Linear Unit
ROI	Regions Of Interest
RT	Radiotherapy
RTDOSE	Radiotherapy Dose
RTPLAN	Radiotherapy Plan
RTSTRUCT	Radiotherapy Structure Set
SC	Secondary Capture
SD	Standard Deviation
SPM8	Statistical Parametric Mapping 8
SSIM	Structural Similarity Index
STL	Standard Tessellation Language
SVD	Symmetric Volume Difference
SVM	Support Vector Machine
T	Tchebichef Moment Image
TBI	Traumatic Brain Injury
TCIA	The Cancer Imaging Archive Data Collections
TM	Tchebichef Moments
TP	True Positives
VOE	Volumetric Overlap Error
VRAM	Video Random Access Memory
ZM	Zernike Moment

Chapter 1

Introduction

As any structure, the skull has several resonant frequencies that are excited when it is struck and transiently vibrates at its resonant frequencies [1–6]. The resonant frequencies correspond to specific deformation modes that are determined by the geometric shape of the skull and the distribution of bone thickness. The brain tissue has a maximum mechanical impedance (0.11 Ns/m) near 450±50 Hz [7,8], and any vibrations near this frequency are likely to propagate into the brain. If the propagating vibrations can constructively interfere and cause damage to brain tissue when they meet in different regions of the brain, then it is possible that axons crossing these junctions might experience shearing forces resulting in the axons stretching and tearing. This is a possible mechanism by which concussion injury to the brain occurs.

Concussion is a traumatic brain injury that affects brain function [9–13]. Symptoms are usually temporary but can include headaches, difficulty concentrating, loss of memory, drowsiness, dizziness, sensitivity to light, and feeling slowed down. The brain has the approximate consistency of gelatin and is protected from everyday jolts and bumps by cerebrospinal fluid inside the skull. Despite the natural protection, a significant direct or indirect blow to the head or upper body can cause the brain to move back and forth against the inner walls of the skull. As the brain moves back and forth within the skull, areas of varying density in the brain slide over each other at different velocities. This shearing effect leads to extensive deformation of the brain structure and micro-tears to the underlying axon fiber bundles. When severe, this is known as axonal shearing or diffuse axonal injury (DAI) [14]. DAI is a severe traumatic brain

injury with profound symptoms. Concussion is a much milder brain injury. However, shearing or stretching of the axons in a concussion can disrupt communication among neurons in the brain by creating gaps in the cellular membrane of the axon and leakage of the ions needed for transmitting action potentials.

Vibration of the skull after impact could be the cause of damage to the the underlying cortex. The skull deforms and vibrates, like a bell [15] for 3 to 5 milliseconds after impact [16], perhaps bruising the cortex [17]. When the skull vibrates, the vibration could be spread through the thin layer of cerebral spinal fluid and into the cortex, and as the vibrations propagate deep into the brain, concussion might occur deep within the brain or closer to the periphery. It would all depend on where the vibrations constructively interfere, so, as to increase the amplitude of tissue deformation locally.

The rationale for this thesis is to examine the role that vibration, excited by impact to the skull, could have on concussion injury to the brain. This thesis models the resonant frequencies excited by direct impact to the brain and the relationshp between skull shape, thickness, stiffness, and the frequencies and mode shapes that individual skull possess. Figure 1.1 illustrates the strategy for accomplishing this goal.

Therefore, the overall purpose of the proposed research is to determine how the vibrations of the skull resonant frequencies propagate in the brain tissue. This goal will be achieved by addressing three research objectives: I) Develop an automatic method to segment the skull and brain from MRI, II) Create a novel 2D and 3D automatic method to align the facial skeleton, and III) Determine how the vibrations of the skull resonant frequencies propagate in the brain tissue.

- **Objective I – Create an Automatic Method to Segment/Extract Skull and Brain from MRI**

  The goal of this objective is to present an automatic method to extract the skull and brain from MRI. This goal is critical since extracting the skull from CT exposes patients to a certain amount of ionizing radiation, MRI scans are much safer for the patient but
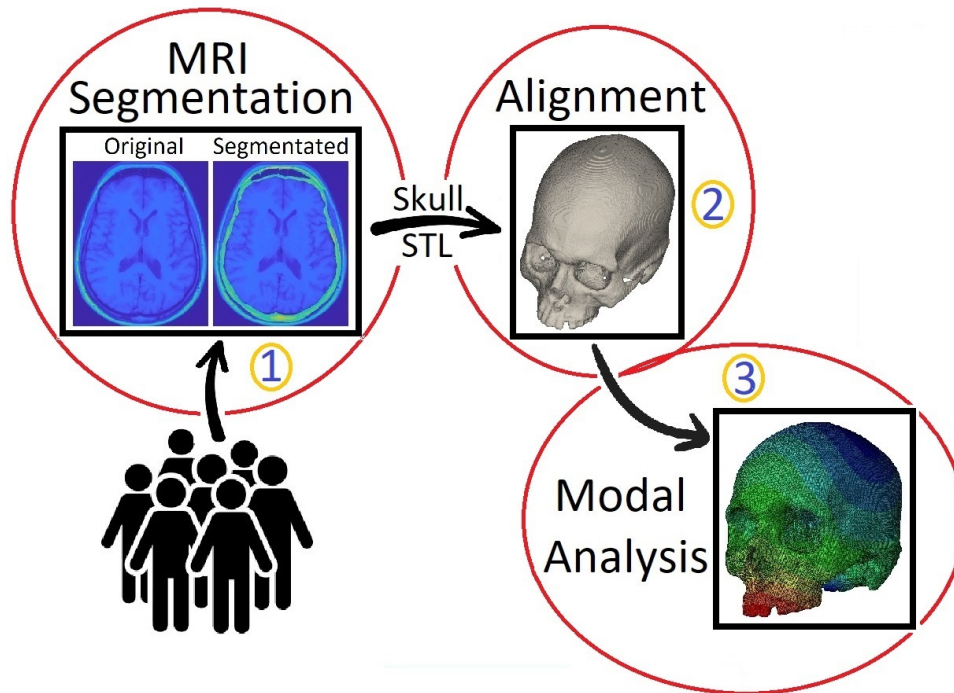
Figure 1.1: Magnetic resonance images (MRI) are taken from a group of individuals. (1) An algorithm extracts the skull from the MRI in the STL file format. (2) These skulls are then aligned using an algorithm. (3) Modal analysis assesses the natural frequency of skulls to distinguish between different skulls with similar natural frequencies.

pose their own difficulties. Due to the weak magnetic resonance signal from bones, MRI scans have difficulty in distinguishing bone tissue from other tissues. As a result, bone segmentation from MRI is a challenging task. Additionally, a strategy that enhance brain segmentation in MRI benefits the skull segmentation in MRI by reducing image information through brain removal, and also helps the improvement of neurological diseases assessment.

- **Objective II – Create a Novel 2D and 3D Automatic Method to Align the Facial Skeleton**

  The goal of this objective is to introduce a 2D and 3D technique using an invariant moment and convolutional neural network, respectively, to automatically align the transverse, sagittal, and coronal planes. Both techniques are introduced in this paper for the automatic calculation of the craniofacial symmetry midline and midplane. However,

these methods are critical for conducting a future impact analysis as the perfect model alignment is important for consistent simulations of multiple models.

- **Objective III – Determine how the skull resonant frequency vibrations propagate into the brain tissue**

    The goal of this objective is to determine how the natural frequencies of multiple skulls may propagate deep into the brain tissue causing the concussive injury. Since brain tissue has a maximum mechanical impedance at around 450 Hz, any vibrations near this frequency may easily propagate into the tissue. Therefore, from the data and results acquired in this objective, it is expected to raise a theory regarding the relation between the skull geometry (shape and thickness), skull material properties (density and Young's modulus), and its vibration with brain tissue injury.

1.1   Concussion

A concussion, also known as mild traumatic brain injury (mTBI), is a form of traumatic brain injury caused by a biomechanical force that interferes with neurologic normal function which typically occurs in the absence of detectable brain injury [18]. Headaches, difficulty concentrating, loss of memory, drowsiness, dizziness, sensitivity to light, and feeling slowed down, are some examples of clinical, functional, and cognitive symptoms that may occur as a result of mTBI [19, 20]. Depending on the severity, these symptoms last anywhere from a couple days to a few months [21]. Concussions are commonly produced by a direct or indirect contact to the person that ultimately leads to a force transmission through the neck/head region generally caused by falls, motor vehicle incidents, or sport related injury.

The probability of sustaining a concussion is influenced by linear and rotational acceleration, which causes a temporary increase in intracranial pressure and a strain effect, respectively. Initially, linear acceleration was believed to be the most significant mechanism, whereas negative pressure, cavitation, and rotational acceleration were believed to be minor or nonexistent.

For instance, Ommaya (1966) [22] suggested that rotation itself cannot generate the same levels of injury as direct impact and showed that indirect impact requires twice the rotational velocity to cause a concussion. Ono (1980) [23] performed a number of experiments with primates and discovered no association between rotational acceleration and the incidence of concussion. The concussion was determined to be caused by linear acceleration from a direct impact.

Further research found that rotation could account for approximately half of the risk for brain injury [24]. Gennarelli (1972) [25, 26] indicated that horizontal head translation resulted in predominantly local effects, such as intracerebral hematomas and well-circumscribed cerebral contusions, while DAI were only observed when a rotational component was included. This indicated that concussion could only be generated when linear and rotational motion were combined rather than linear motion alone [27].

The first to identify rotational acceleration with or without direct impact as a major mechanism in brain injury was [28], theorizing that tensile and shear strain generated by rotation alone may cause cerebral concussion and contrecoup contusion [29]. Furthermore, [25, 30–33] reported that rotational acceleration contributes more than linear acceleration to the development of concussions, diffuse axonal injuries, and subdural hematoma, based on the assumption that these injuries were caused by the shear strain induced by rotational acceleration and stated that rotational acceleration may cause most forms of brain damage. Regardless of the fact that it is typically evaluated in combination, rotational acceleration is considered to be the major force responsible for concussion and has been definitively detected in severe traumatic brain injury [34]. For linear acceleration, pressure gradient seems to be the major mechanism, while shear stress, which originates from differential motion between the brain and skull, appears to be the key mechanism for purely rotational acceleration [35]. In addition, increasing rotation duration reduces brain tolerance and causes more significant damage [34].

After a shock to the head, computer models suggest four injuries ([19]):

- Impact - impact generates a millisecond-long shockwave throughout the brain, which is transmitted via the ventricles and repeatedly reflected and refracted by the skull and

the boundaries between structures. As a result, the waves interefere so their crossings augment and cancel each other. Patches of neurons are damaged all over the brain (including the cerebellum and brainstem) where negative pressure waves coincide. As a consequence, even minor concussions are associated by extensive neuronal damage. The negative phase of the wave is the most damaging, producing microsecond bubbles that tear blood vessels and pull tissues apart, potentially causing extradural clots, and tearing or stretching structures such as proteins, cell walls, and deoxyribonucleic acid (DNA) chains.

- Deceleration - for traumatic brain damage, deceleration is common (but not essential). The corticospinal tracts and the reticular formation are stretched and torn when the cerebrum swings on the brainstem resulting in spasticity and loss of consciousness. For instance, newborns can die from brainstem traction alone by a shake, without any impact [36].

- Rotation - Brain tissue deforms more readily in reaction to shear forces than other biologic tissues due to the general highly organized nature of the brain and its physical properties. Shear forces are generated throughout the brain by rapid head rotations, hence rotational accelerations have a significant risk of causing shear-induced tissue damage [37]. Since the impact is not always centralized, rotation may occur that causes the axons crossing the hemispheres to stretch.

- Vibration - The underlying cortex is bruised by the vibration (resonance) of the skull after an impact. For 3–5 milliseconds [16], the skull deforms and vibrates like a bell [15], damaging the cortex [17]. The brain is gelatinous in texture and is cushioned from ordinary bumps and jolts by cerebrospinal fluid inside the skull. As a result of the deceleration, the frontal and temporal cortex has already swung violently against the skull, impacting the cerebrospinal fluid. Because there is no layer of cerebrospinal fluid to deflect or reduce the wave or cushion its force, skull vibration is transmitted directly to the

cerebral cortex.

After a blow to the head, neck, or upper body, the brain can deform back and forth forcefully against the inner walls of the skull. When this occur, areas of varying density in the brain, such as the white and gray matter, slide over each other at different speeds (Figure 3.1). The shearing forces experienced by axons crossing these junctions lead them to stretch and tear apart from the cell body. This is called axonal shearing or diffuse axonal injury with moderate to severe symptoms such as brief loss of consciousness, impaired long-term memory, reduced problem-solving ability, lower social inhibition, attention/perception problems, increasing headache, repeated vomiting, slurred speech, confusion, unusual behaviors, seizures, and limb weakness or numbness. This shearing action causes serious brain deformation and micro-tears in the underlying axon fiber bundles. Inertial forces applied to the head upon impact produce stretching of the deep and subcortical white matter, which is the major mechanical mechanism in DAI. This can lead to a breakdown of communication among neurons in the brain (Figure 3.2). Toxic amounts of chemicals called neurotransmitters are released into the extracellular area when the torn axons degenerate. Over the next 24 to 48 hours, many of the surrounding neurons begin to die.

Numerous mTBI patients experience severe post-concussion symptoms, that can include somatic symptoms including nausea, dizziness, headache, blurred vision, auditory disturbance, and fatigue, cognitive complaints such as memory and executive function, and emotional and/or behavioural problems the same as disinhibition and emotional lability [38]. The most common post-concussive symptom is headache, followed by dizziness, which is more commonly a sensation of disequilibrium and imbalance than objective vertigo [10].

The average recovery period is not scientifically defined and is impacted by factors such as age, gender, and previous concussion history. Concussions generally recover on their own with adequate physical and cognitive rest until symptoms diminish. Approximately 10% of athletes experience concussion symptoms that last longer than two weeks. When it comes to non-sport-related concussions, most individuals recover entirely within the first three months, however, up

Grey Matter   **White Matter**
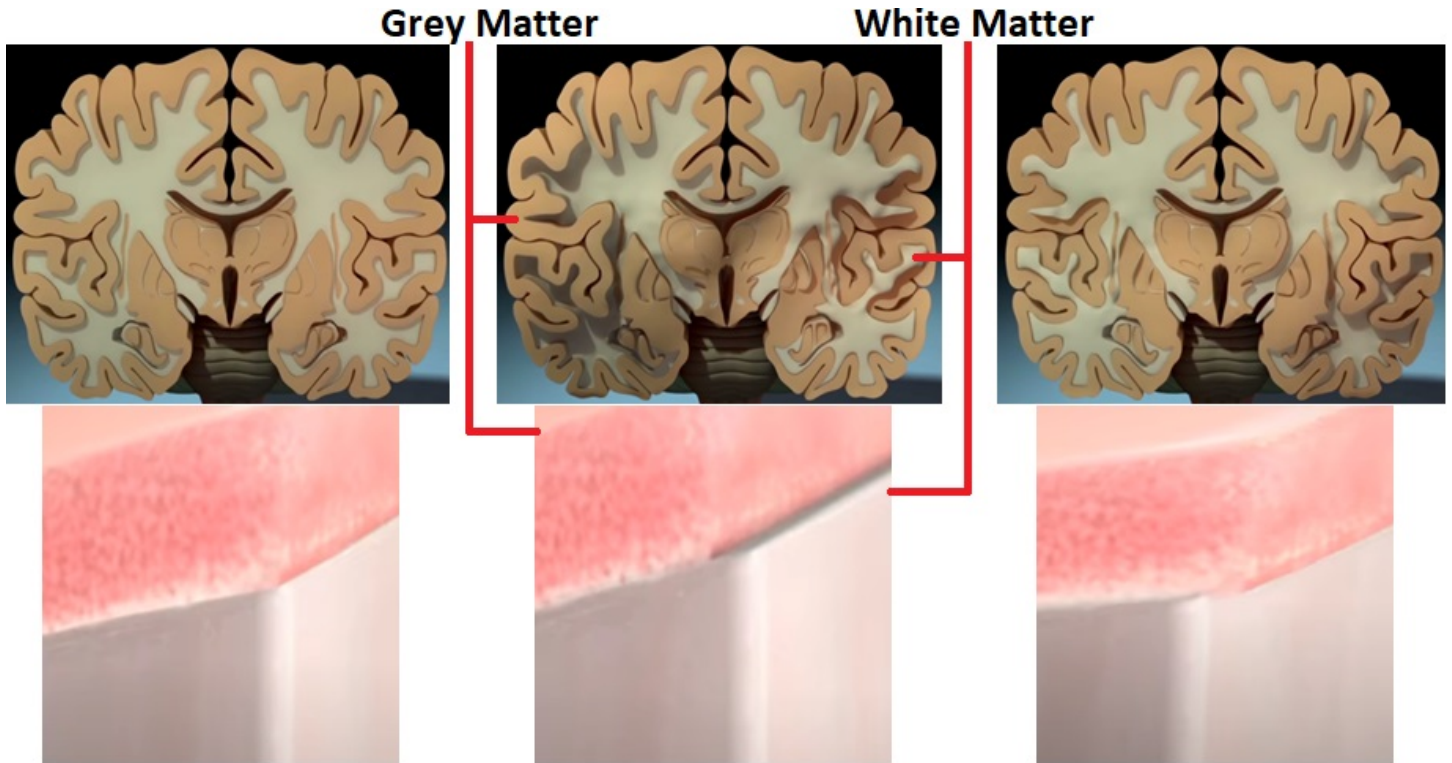
Figure 1.2: Grey and white matter in the crossing junction.

Source: Adapted from https://www.youtube.com/watch?v=55u5Ivx31og
https://www.youtube.com/watch?v=7Lfc7aZfd0Y&ab_channel=CBCNews

Dendrites

Cell Body

Shearing force

Axon

Diffuse Axonal Injury (DAI)

Breakdown of communication among neurons
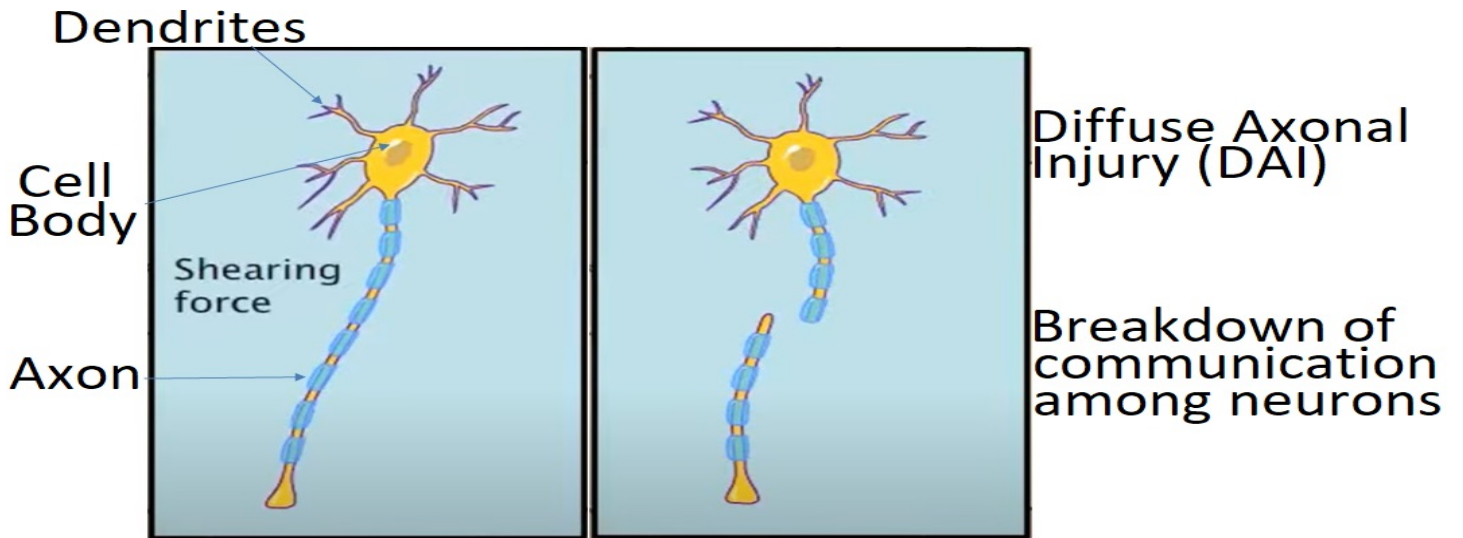
Figure 1.3: Grey and white matter in the crossing junction.

Source: Adapted from https://www.youtube.com/watch?v=k_MktbTuEdI&ab_channel=brainlinebrainline

to 33% may continue to experience symptoms beyond that [39]. The persistence of symptoms beyond the generally accepted recovery time period may indicate a sustained concussion or the development of post-concussion syndrome (PCS). An overload of subconcussive hits is linked to a degenerative brain disease known as Chronic Traumatic Encephalopathy (CTE), also known as dementia pugilistica or fistfighter's dementia, boxer's madness, or punch drunk syndrome. CTE result in symptoms more severe and may lead to dementia.

One strategy for revealing critical information of head structures is the use of medical imaging techniques such as computed tomography (CT) and MRI. When adopting medical imaging techniques to assess neurological disorders, the initial step is to isolate regions of interest (ROIs) using image processing techniques such as feature extraction. Since feature extraction is the process of transforming a large set of data into a set of feature vectors of reduced dimension, multi-dimensional datasets, such as those encountered in CT and MRI modalities, benefits from its use.

1.2   Feature Extraction

In medical imaging, feature extraction is an important step to obtain feasible information for diagnostics, classification, recognition, and detection tasks. A number of effective performance classification system relies on the feature extractor. Consequently, researchers and developers are devoting efforts to improve recognition and classification. Despite the fact that researchers have developed a number of feature extraction methods, there is still opportunity to improve both the feature extraction techniques and the classifier system [40]. A typical solution to the problem of multi-dimensional datasets is to search for a projection of the data onto a reduced number of features that preserves as much information as possible [41].

In computer vision, the data set is composed of images that, in theory, include various irrelevant information, which makes any computational calculation more time-consuming. In this context, feature extraction may be considered as a method of generating a new set of vec-

tors that represent an observation while summarizing the information contained in the original feature set. As a result of the feature extraction, a new set of data is obtained with a reduced dimensionality than the original set, thus, reducing the computational cost required in the recognition process. Furthermore, in pattern recognition problems, it is desirable to extract features that introduce high discrimination between different classes, eliminating redundant features that do not contribute to the classification process. However, decreasing the data size should not compromise the performance of the classification system. Instead, the feature extraction process should be directed so that the generated features allow the classifier to efficiently generalize the problem and obtain a high accuracy rate. In essence, the fundamental objective of a feature extractor is to reduce the complexity of a feature vector that represents an observation while maintaining the ability to discriminate across classes.

1.2.1   Geometric Moments

Invariant moments have been extensively used in feature extraction, pattern recognition, and object classification. One of the most significant features of the moments is its invariance to affine transformations, such as rotation, translation, and scale invariance. Moments are scalar quantities used to characterize a function and capture its most significant features [42]. From a mathematical point of view, moments are projections of a function on a polynomial basis.

The concept of moments was initially introduced by Hu (1962) [43], and has been widely used in the field of image analysis and pattern recognition such as in character recognition [44–46], facial recognition [47–49], fingerprint recognition [50, 51], medical image analysis [52, 53], and industrial applications [54, 55]. Hu (1962) [43] introduced 2-D geometric moments of a distribution function for an image as a structured element, called the moment invariants [56]. Hu (1962) [43] also employed algebraic function theory to create a set of Moment Invariants that describe the orthogonal invariants to linear transformation was employed, that are rotation, translation, and scaling factors. This function is also known as Geometric Moments (GM).

The general formulation of GM function m_{pq} of the order $(p + q)$ for two-dimensional

continuous function $f(x, y)$ for $p, q = 0, 1, 2, 3. \ldots$ are defined as [57, 58]

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q (x, y) dx dy. \tag{1.1}$$

Similarly, m_{pqr} denotes the three-dimensional continuous function with an order of $(p + q + r)$ defined as

$$m_{pqr} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q z^r (x, y, z) dx dy dz. \tag{1.2}$$

In digital image analysis, equations 1.1 and 1.2 can be represented as

$$M_{pqr} = \sum_x \sum_y x^p y^q I(x, y), \tag{1.3}$$

$$M_{pqr} = \sum_x \sum_y \sum_z x^p y^q z^r I(x, y, z). \tag{1.4}$$

where $x = 0, 1, ..., N\text{-}1$, $y = 0, 1, ..., M\text{-}1$, and $z = 0, 1, ..., D\text{-}1$, (x, y, z) are the image coordinates, and $I(x, y, z)$ is a digital image defined over the discrete square domain $N{\times}M{\times}D$.

This equation allows for the calculation of the center of mass of an image, and of a region by [59]

$$\bar{x} = \frac{M_{10}}{M_{00}} \quad and \quad \bar{y} = \frac{M_{01}}{M_{00}}. \tag{1.5}$$

$$\bar{x} = \frac{M_{100}}{M_{000}}, \quad \bar{y} = \frac{M_{010}}{M_{000}}, \quad and \quad \bar{z} = \frac{M_{001}}{M_{000}}. \tag{1.6}$$

The drawback of GM is that they are not orthogonal since they are based on the projection of an image's intensity function onto specific monomials, that do not form an orthogonal basis [56]. As a result, GM suffer from a high degree of information redundancy and those of a higher order are sensitive to noise [60, 61]. This disadvantage of the GM can be solved by using the orthogonal moments since their kernels are based on orthogonal polynomials.

1.2.2   Orthogonal Moments

Several research studies have been conducted in order to enhance and establish new types of moment invariants. Based on the theory of continuous orthogonal polynomials, Teague (1980) [62] introduced the Zernike and Legendre moments, and Mukundan (2001) [63] presented the Tchebichef moments. Through the property of orthogonality, the corresponding moments are given the feature of minimum information redundancy or overlapping of information between moments with different orders and repetitions. In other words, the different parts of the image are represented by distinct moment orders.

1.2.2.1   Zernike and Pseudo-Zernike Moment

The Zernike Moments (ZM) were first introduced by Teague (1980) [62], which were obtained from continuous orthogonal Zernike polynomials, and were developed to improve on traditional GM, particularly for rotation applications. ZM are more robust in the presence of noise, and reduce the invariant properties error even though the image is under the influence of independent features. Since ZM is defined using a polar coordinate system, ZM are usually applied in rotation recognition tasks which requires rotation invariance [64] such as analysing and understanding images systems [65–67], medical image analysis [68–70], recognition systems [71–73], and so forth.

   To compute ZM, the image or ROI is initially projected onto a unit circle, with the image's centre correponding to the circle's origin. Pixels that are outside the circle are ignored in the computation. The coordinates are then defined by the size of the vector from the origin to the coordinate point. To include these pixels, the disc can be expanded so that the image function *I(x,y)* is completely enclosed inside the disc (Figure 1.4). This approach can be performed by

$$x_i = -\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{N-1}x, \quad and \quad y_j = \frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{N-1}y, \qquad (1.7)$$

where *x = y = 0, 1, ..., N-1*, (x_i, y_j) are the image coordinates, and *I(x,y)* is defined over the

discrete square domain N×N.
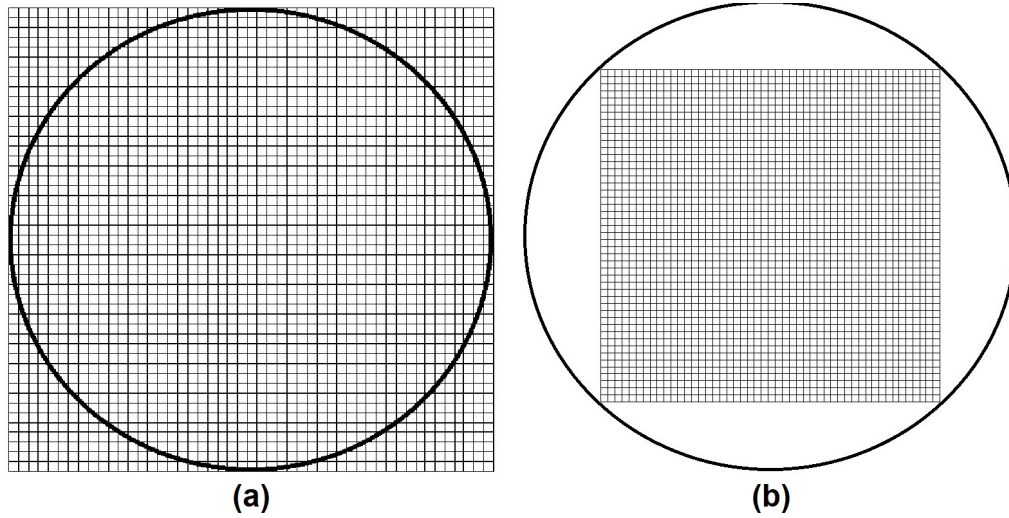
Figure 1.4: (**a**) Image mapped over and (**b**) enclosed in a unit disc.

Computing ZM for an image consists of three steps; calculating the radial polynomial, calculating the Zernike base functions, and calculating the ZM by projecting the image to the base functions. Obtaining the ZM of an image begins with the calculation of radial Zernike polynomials $R_{pq}(\rho)$ of order p and q

$$R_{pq}(r) = \sum_{s=0}^{\frac{p-|q|}{2}} c(p, q, s)\rho^{p-2s},\tag{1.8}$$

where

$$c(p, q, s) = (-1)^s \frac{(p - s)!}{s!(\frac{p+|q|}{2-s})!(\frac{p-|q|}{2-s})!},\tag{1.9}$$

where ρ is the length of the vector from the origin to the pixel (x, y)

$$\rho_{xy} = \frac{\sqrt{(2x - N + 1)^2 + (N - 1 - 2y)^2}}{N},\tag{1.10}$$

where p is a non-negative integer and q is an integer satisfying $0 \le |q| \le p$. Figure 1.5 shows the radial polynomials from 0 to 5 order in the interval $0 \le \rho \le 1$.

For the Zernike base functions, Teague (1980) [62] introduced a complex set of Zernike
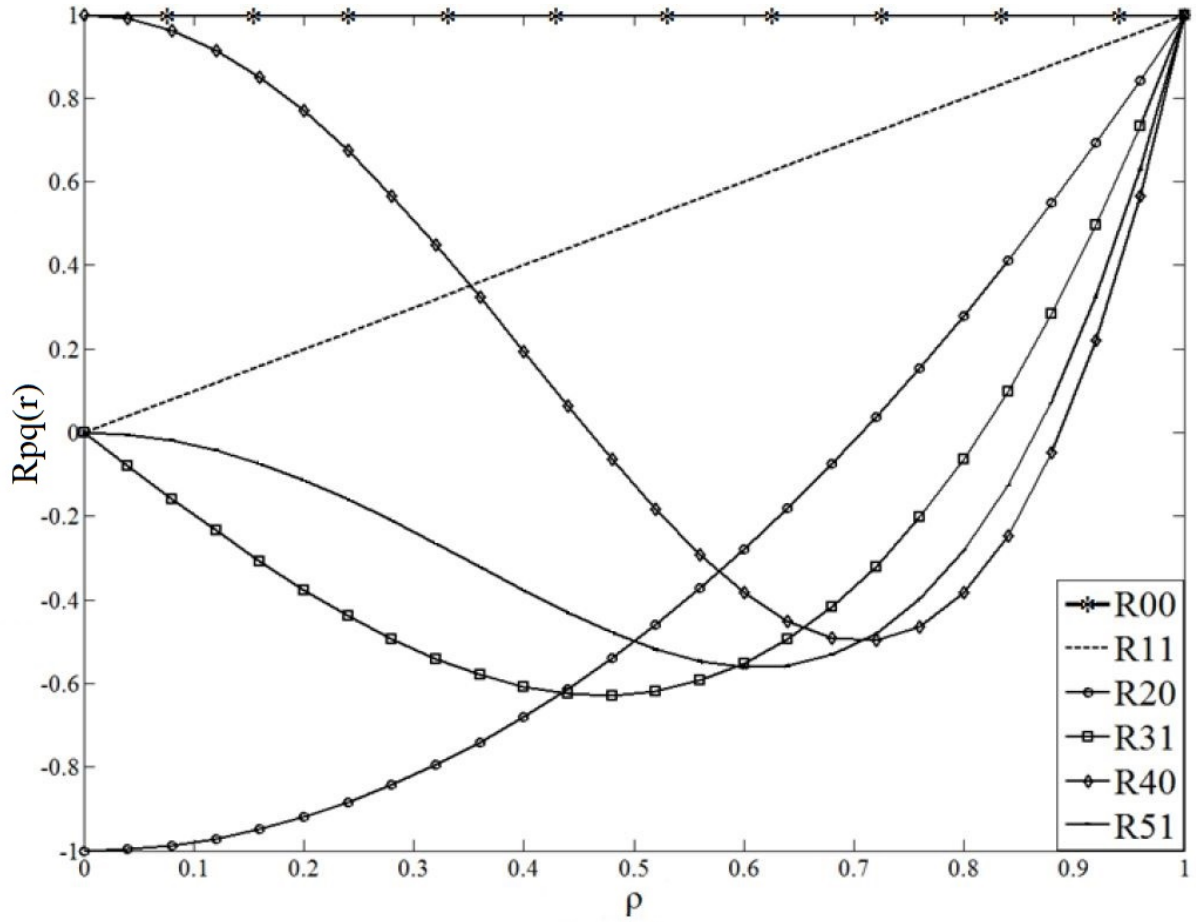
Figure 1.5: Zernike radial polynomials $R_{pq}(r)$ of order 0 to 5 and repetitions 0 and 1.

polynomials that form a complete orthogonal set on the interior of a unit circle $x^2 + y^2 = 1$,

$$V_{pq}(x, y) = V_{pq}(\rho, \theta) = R_{pq}(\rho)e^{jq\theta}, \tag{1.11}$$

where $j = \sqrt{-1}$ and θ is the angle between ρ and x axis.

The final step consists of calculating the ZM. The discrete shape of the ZM of an image size of $N \times N$ is expressed as

$$A_{pq} = \frac{p+1}{\pi} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} I(x_i, y_j)V_{pq}^*(x_i, y_j)\Delta x_i \Delta y_j, \tag{1.12}$$

where

$$x_i = -\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{N-1}x, \quad and \quad y_j = \frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{N-1}y, \tag{1.13}$$

V_{pq}^* are the complex conjugate of the Zernike polynomials (V_{pq}) in Equation(1.11) and $p - |q| = $ *even*.

Pseudo-Zernike Moments (PZM) are obtained by removing the condition $p - |q| = $ *even* from the ZM. It leads to a reduction in time complexity [74], more feature vectors since there are $(p + 1)^2$ linearly independent pseudo-Zernike polynomials of orders $\leqslant p$ as compared to $\frac{1}{2}(p + 1)(p + 2)$ of Zernike polynomials [75], less sensitive to image noise [76], and a robust orthogonality property that makes the individual order moments more independent character-istics of the image [77]. Therefore, Equation(1.11) is also applied to PZM with $q = -p, -p+1,$ ..., p. Equation(1.11) can then be rewritten as

$$R_{pq}(r) = \sum_{s=0}^{p-|q|} (-1)^s \frac{(2p + 1 - s)!}{s!(p+ |q| +1 - s)!(p- |q| -s)!} r^{(p-s)}. \tag{1.14}$$

Due to the higher computational cost of the PZM, most of the attention has been focused on ZM. Recently, instead of using the direct formula presented in Equation(1.14), efficient tech-niques to determine pseudo-Zernike polynomials have been proposed to improve the computa-tional time of the PZM. Among these methods, [78] introduced a *p-recursive* method defined by

$$R_{qq}(r) = r^q, \tag{1.15}$$

$$R_{(q+1)q}(r) = (2q + 3)r^{q+1} - 2(q + 1)r^q, \tag{1.16}$$

$$R_{pq}(r) = (k_1 r + k_2)R_{(p-1)q}(r) + k_3 R_{(p-2)q}(r), p = q + 2, q + 3, ..., p_{max}, \tag{1.17}$$

where

$$k_1 = \frac{2p(2p + 1)}{(p + q + 1)(p - q)}, \tag{1.18}$$

$$k_2 = -2p + \frac{(p + q)(p - q - 1)}{(2p - 1)}k_1, \tag{1.19}$$

$$k_3 = (2p - 1)(p - 1) - \frac{(p + q - 1)(p - q - 2)}{2}k_1 + 2(p - 1)k_2. \qquad (1.20)$$

1.2.2.2 Legendre Moments

The Legendre Moments (LM) were other moments introduced by Teague (1980) [62] and were created based on the continuous orthogonal Legendre polynomials [56]. There have been several papers describing the application of the LM in image processing [79–81]. LM are produced from the recursive relation of the p-order Legendre polynomial which is defined as [82]

$$P_p(x) = \frac{(2p - 1)xP_{p-1}(x) - (p - 1)P_{p-2}(x)}{p}, \qquad (1.21)$$

where p is a non-negative integer, $0 \le |q| \le p$, $P_0(x) = 1$, $P_1(x) = x$, and $p > 1$. The Legendre polynomial set are defined on the interval between *[-1,1]*. Figure 1.6 shows the Legendre polynomials of order 0 to 5.
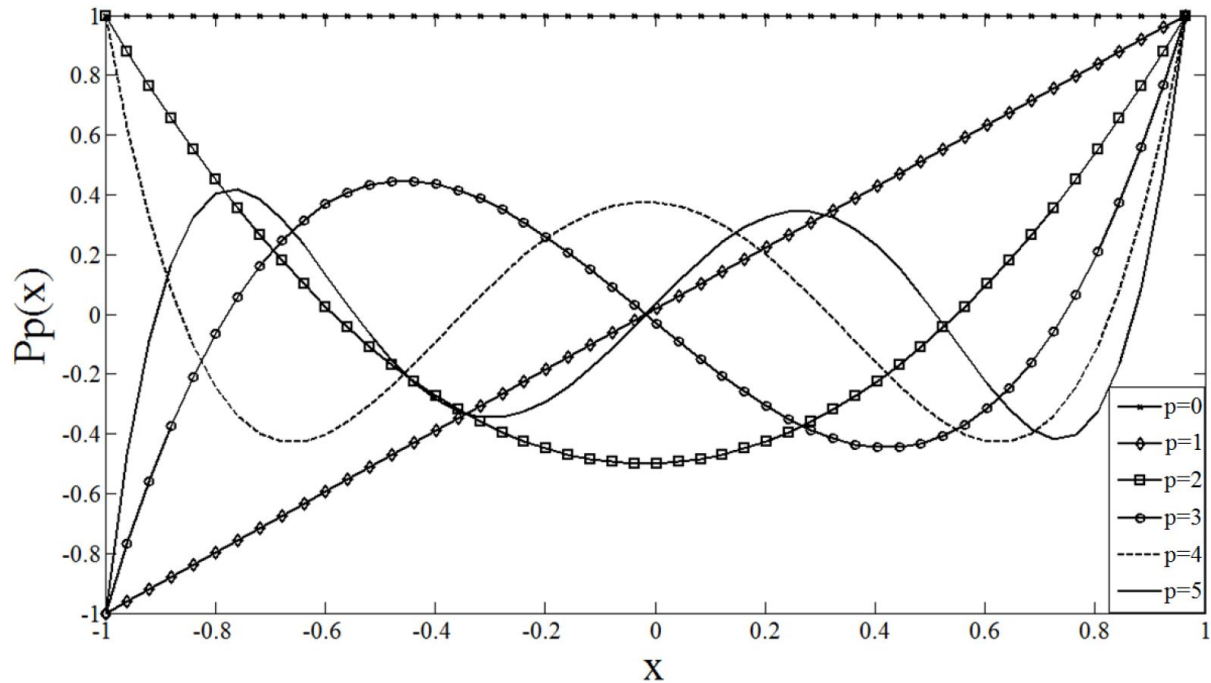
Figure 1.6: Legendre polynomials $P_p(x)$ of p-order 0 to 5.

The two-dimensional discrete Legendre moments of order *(p+q)* are expressed as

$$L_{pq} = \lambda_{pq} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P_p(x_i)P_q(y_j)I(x_i, y_j), \tag{1.22}$$

where $I(x, y)$ is defined over $N{\times}N$ size, λ_{pq} is a normalizing constant defined as

$$\lambda_{pq} = \frac{(2p + 1)(2q + 1)}{N^2}, \tag{1.23}$$

and x_i and y_j are the normalized pixel coordinates from *[-1,1]* given as

$$x_i = \frac{2i}{N - 1} - 1, \quad and \quad y_j = \frac{2j}{N - 1} - 1. \tag{1.24}$$

1.2.2.3 Tchebichef Moments

The orthogonal moments described above have the following limitations; Zernike and pseudo-Zernike polynomials are defined within a unit circle, and Legendre polynomials are only valid within the interval *[-1, 1]*. As a result, calculating these moments requires a coordinate transformation and appropriate approximations from continuous moments. Due to numerical approximations of continuous moments, this will result in a discretization inaccuracy [83]. The implementation of Tchebichef moments (TM), or Chebyshev Moments, does not involve any numerical approximation since their base set is orthogonal in the discrete domain of the image's spatial coordinates.

The discrete Tchebichef moments (TM) of order $(p+q)$ are constructed under the Tchebichef polynomials defined by Erdelyi (1953) [84]. The scaled Tchebichef polynomials were introduced in [85] as

$$t_p(x) = \frac{(2p - 1)t_1(x)t_{p-1}(x) - (p - 1)(1 - \frac{(p-1)^2}{N^2})t_{p-2}(x)}{p}, \tag{1.25}$$

where *p = 2,3,...,N-1*, $t_0(x) = 1$, and $t_1(x) = \dfrac{2x + 1 - N}{N}$. Figure 1.7 shows the Tchebichef

polynomials of order 0 to 5.
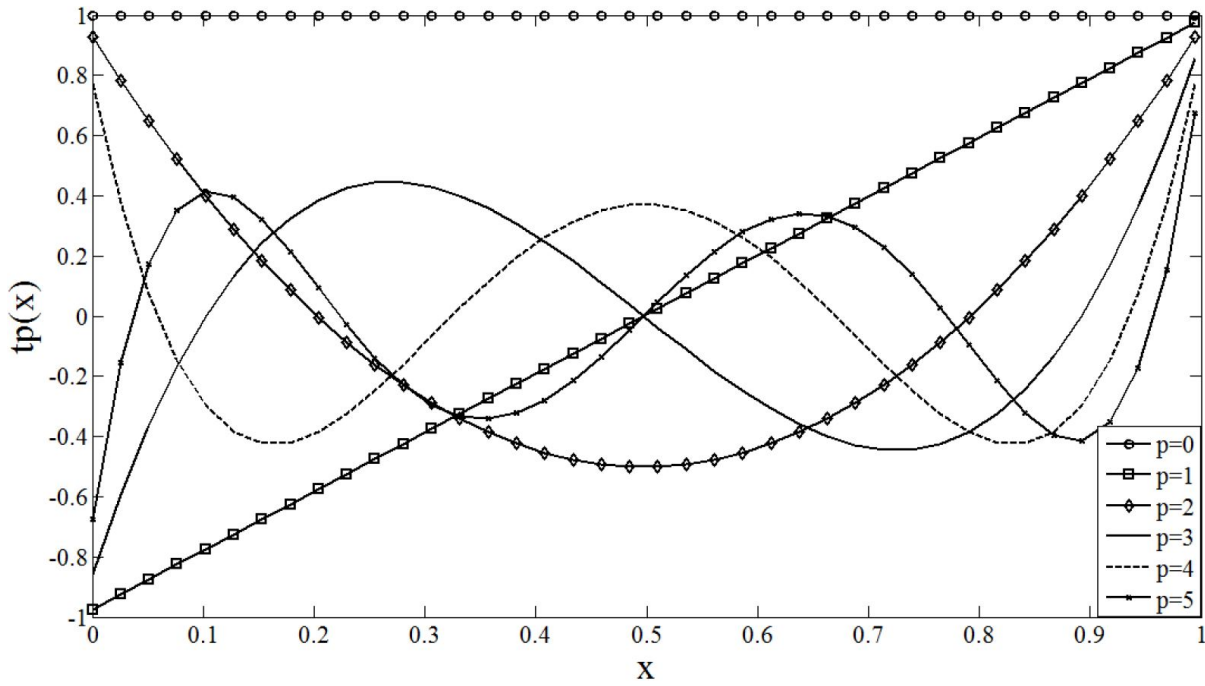
Figure 1.7: Tchebichef polynomials $t_p(x)$ of p-order 0 to 5.

The squared-norm of the scaled polynomials is then modified accordingly to

$$\rho(p, N) = \frac{N(1 - \frac{1}{N^2})N(1 - \frac{2^2}{N^2})...N(1 - \frac{p^2}{N^2})}{2p + 1}, \tag{1.26}$$

where $p = 0, 1,..., N-1$. Finally, Tchebichef moments can be constructed as

$$T_{pq} = \frac{1}{\rho(p, N)\rho(q, N)} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} t_p(x_i)t_q(y_j)I(x_i, y_j). \tag{1.27}$$

1.2.3   Principal Component Analysis

Principal Component Analysis (PCA) was introduced by Pearson (1901) [86] and indepen-dently developed and named by Hotelling (1933) [87] and in Hotelling (1936) [88]. PCA is an orthogonal linear transformation method for reducing the dimensionality of large data sets by transforming a number of variables into a reduced number of variables while preserving the essential information. It transforms the data into the new coordinate system wherein the

highest variance is positioned on the first coordinate, the second highest variance is located on the second coordinate, and so forth [89]. In summary, the objective of PCA is to reduce the number of variables in a data set while conserving as much information as possible. In order to achieve these goals, PCA computes new variables called principal components (PCs) which are formed as linear combinations of the original variables.

Since PCA is sensitive to the variances of the initial variables, the first step in calculating PCA for a data set $\mathbf{X}$ is to normalise the range of the continuous initial variables ensuring that they all contribute equally to the analysis. Mathematically, this can be achieved by

$$\mathbf{M} = \mathbf{X} - E\{\mathbf{X}\}. \tag{1.28}$$

To evaluate whether there is any correlation between the variables with regard to each other, the covariance matrix of $\mathbf{M}$ with n samples is calculated as

$$\mathbf{C_M} = \frac{1}{n-1}\mathbf{M}\mathbf{M}^T. \tag{1.29}$$

In order to determine the PCs of the data, the covariance matrix $\mathbf{C_M}$ is decomposed by eigenvalue decomposition

$$\mathbf{C_M} = \mathbf{E}\mathbf{\Lambda}\mathbf{E}^T, \tag{1.30}$$

where the eigenvectors $\{e_1, e_2, ..., e_m\}$ of $\mathbf{C_M}$ are arranged as columns in $\mathbf{E}$ and $\mathbf{\Lambda}$ is a diagonal matrix containing the eigenvalues $\{\lambda_1, \lambda_2, ..., \lambda_m\}$ of $\mathbf{C_M}$. The first $l(l<m)$ greatest columns in $\mathbf{E}$ are considered as PCs, where l denotes the number of PCs and m the number of variables in $\mathbf{M}$.

The last step aims to reorient the data from the original orientation to the orientation reflected by the PCs using the eigenvectors of the covariance matrix. Since the obtained PCs constitute the subspace, data matrix $\mathbf{M}$ can be projected into the subspace as

$$\mathbf{Y} = \widetilde{\mathbf{E}}_l^T\mathbf{M}, \tag{1.31}$$

where $\widetilde{\mathbf{E}}$ is the reordered and reduced eigenvector matrix

1.2.4   Independent Component Analysis

Independent Component Analysis (ICA) is a mathematical technique for determining components behind a set of random variables that are assumed to be mutually statistically independent and non-gaussian. In other words, it is a statistical signal processing technique whose objective is to linearly decompose a random vector into components that are not only uncorrelated, as in PCA, but also as independent as possible [90]. As a result, ICA is considered a generalization of PCA. PCA aims to develop a representation of inputs from uncorrelated variables, whereas ICA provides a representation based on statistically independent variables [91].

Given a series of random variable observations $\{x_1(t), x_2(t), ..., x_n(t)\}$, where t is the time or sample index, it is assumed that these are generated as a linear combination of independent components $\{s_1(t), s_2(t), ..., s_n(t)\}$ [92]

$$\mathbf{x} = [x_1(t), x_2(t), ..., x_n(t)]^T,$$

$$\mathbf{x} = \mathbf{A}[s_1(t), s_2(t), ..., s_n(t)]^T,$$

$$\mathbf{x} = \mathbf{As}, \tag{1.32}$$

where $\mathbf{A}$ is an unknown $m{\times}n$ matrix of full rank called the mixing matrix [93]. This model is considered a generative model meaning that it describes how observed data are generated by a process of mixing the independent components (ICs), s. ICs are latent variables which means they cannot be directly observed. As a result, the basic challenge of ICA is to estimate the underlying ICs, or equivalently to estimate the mixing matrix $\mathbf{A}$ when only $\mathbf{x}$ is observed. Moreover, the model requires that the ICs have non-Gaussian distributions, or at most one of the ICs is Gaussian. ICs can be discovered by finding a separation matrix $\mathbf{W}$ to linearly transform

x in such a way that the output components are as mutually independent as possible [90]. The problem given by equation(1.32) can then be reformulated, after estimating the matrix **A**, as

$$\mathbf{u} = \mathbf{A}^{-1}\mathbf{x} = \mathbf{Wx}, \tag{1.33}$$

such that a linear combination **u** = **Wx** is the optimized estimate of the signals from independent sources **s**. If the linear transformation **W** satisfies $\mathbf{W} \approx \mathbf{A}^{-1}$, than $\mathbf{u} \approx \mathbf{s}$.

In order to find **W**, two widely used preprocessing methods are performed to make the computation easier. First, the mean of the data **x** is subtracted to centre the data on the source so that the data **x** is a zero-mean variable, in other terms

$$\mathbf{m} = \mathbf{x} - E\{\mathbf{x}\}. \tag{1.34}$$

The second preprocessing is to whiten the observed variables. This step linearly transform the observed vector **x** into a new vector **v** with uncorrelated components and unit variances. It removes all linear dependencies in a data set and normalizes the variance across dimension, or in other words, whitening transforms the data into a spherically symmetric distribution [94]. This problem has a direct solution in terms of PCA expansion, which results in

$$\mathbf{v} = \mathbf{Vm}, \tag{1.35}$$

where **V** is the whitening matrix and **v** the whitened data. Therefore, equation(1.33) is rewritten as

$$\mathbf{u} = \mathbf{Wv}. \tag{1.36}$$

ICA applications on image processing require the random variables to be the training images. Assuming a vectorized image x_i, the training image set can be constructed with n random variables which are assumed to be the linear combination of m vectorized unknown independent components **s**, and denoted as $\mathbf{x} = \{x_1, x_2, ..., x_n\}^T$ and $\mathbf{s} = \{s_1, s_2, ..., s_m\}^T$. Each image

x_i is therefore represented as a linear combination of $s_1, s_2, ..., s_m$ with weighting coefficients $a_{i1}, a_{i2}, ..., a_{im}$, associated to the matrix $\mathbf{A}$. The features of all training images are represented by the columns of $\mathbf{A}$ or $\mathbf{A}_{train}$ and the coefficients $\mathbf{s}$ signal the presence and the amplitude of the features in the observed data $\mathbf{x}_{train}$ [90, 95]. Consequently, $\mathbf{x}_{test}$ must be multiplied by $\mathbf{s}$ to generate $\mathbf{A}_{test}$ as

$$\mathbf{A}_{test} = \mathbf{x}_{test}\mathbf{s}^{-1}. \tag{1.37}$$

Figure 1.8 illustrates the application of this method to extract image features. The FastICA [96] algorithm was chosen to compute ICA in this study since it is simple and enables for programming code modification.
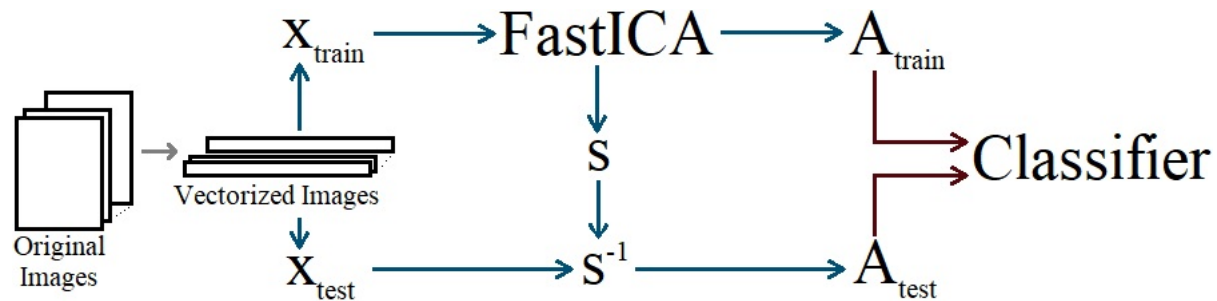
Figure 1.8: Steps of the classification process.

1.3 Classifiers

The final stage in most of the image processing systems is classification in which unknown pattern is attributed to a class. The complexity of the classification problem is determined by the variation of feature values of samples relating to the same class compared to the variation in feature values of samples relating to different classes [57]. As a result, the better the data can be represented by the features extractor in this stage, the better the classification will be. On the other hand, the influence of the classifier itself to the classification efficiency plays a important role. This capability is characterized by the ability to distinguish interclass similarities while disregarding possible intraclass differences. In this study, k-Nearest Neighbours (k-NN), k-

Means, and Convolutional Neural Network (CNN) were chosen.

1.3.1   *k*-Nearest Neighbours

The first studies of a nearest neighbour decision rules were described in a series of two publications [97] and [98] for $k \to \infty$, followed by a research in [99] for fixed values of k, and finally established in [100] a generalized classifier for multiple classes. The *k*-Nearest Neighbours (*k*-NN) method is considered to be one of the oldest, simplest and most well-known methods for pattern recognition.

The training set is formed by n-dimensional vectors and each element of this vector represents a point in n-dimensional space. To determine the class of an element that does not belong to the training set, the *k*-NN classifier searches for k elements of the training set that are nearest to this unknown element, that is, that have the smallest distance. These k elements are called *k*-nearest neighbours. The classes of these k neighbours are verified and the most frequent class is assigned to the unknown element class. In general, the *k*-NN method performs the following procedures to determine the class of an unknown test set:

- the k value is defined. Usually after a few trials, the value of k is chosen according to the results obtained;

- Distance measurement is selected (such as Minkowski, Cosine, and Correlation);

- The distances obtained are sorted in ascending order and k minimum distances are chosen;

- Classes of k nearest neighbors are identified [101].

There are several methods used to measure the distance between sets of different classes in the feature space. They will briefly be summarized below.

1.3.1.1 Minkowski Distance

Minkowski distance is a metric designed for real-valued (Normed) vector spaces, in which a distances can be represented by a positive vector. Let A and B be represented by feature vectors $A = \{x_1, x_2, ..., x_m\}$ and $B = \{y_1, y_2, ..., y_m\}$, respectively, where m is the feature space's dimensionality [102]. The distance between A and B can be calculated using

$$dist(A, B) = \left(\sum_{i=1}^{m} |x_i - y_i|^r \right)^{1/r}. \tag{1.38}$$

The value of r can be maipulated and three different distances can be calculated; Manhattan or City Block Distance - ($r=1$), Euclidean Distance - ($r=2$), and Chebychev Distance - ($r = \infty$).

1.3.1.2 Correlation Distance

The correlation distance is a scaled version of the Pearson distance, with a distance metric ranging from zero to one [103].

$$dist(A, B) = \frac{\sum_{i=1}^{m}(x_i - \overline{x})(y_i - \overline{y})}{\left(\sum_{i=1}^{m}(x_i - \overline{x})^2 \sum_{i=1}^{m}(y_i - \overline{y})^2 \right)^{1/2}}, \tag{1.39}$$

where $\overline{x}$ and $\overline{y}$ are the mean values of x_i and y_i respectively.

1.3.1.3 Cosine Distance

The cosine or Angular distance is produced from the cosine similarity, which quantifies the angle between two vectors, and is estimated by subtracting the cosine similarity from one.

$$dist(A, B) = 1 - \frac{\sum_{i=1}^{m}(x_i y_i)}{\left(\sum_{i=1}^{m} x_i^2 \sum_{i=1}^{m} y_i^2 \right)^{1/2}}. \tag{1.40}$$

1.3.1.4 Hamming Distance

Hamming distance is a distance metric that provides the number of values that diverge between two vectors.

$$dist(A, B) = \sum_{i=1}^{m} 1_{x_i \neq y_i}. \tag{1.41}$$

1.3.2 k-Means

A popular and simple unsupervised algorithm in machine learning is the k-Means [104]. k-Means is a clustering approach that separates n variables into k groups with a high level of similarity between members in one cluster and low level of similarity between members of other clusters [105]. Initially, random k observations are chosen to represent the initial cluster focal points. Second, according to the criterion of minimal distance chosen, such as Euclidean distance, the remaining observations are aggregated to its closest cluster centre. The initial classification is then obtained, and if it is unreasonable, cluster focal points are calculated repeatedly until a reasonable classification is achieved [106] (Figure 1.9).

1.3.3 Convolutional Neural Network

Convolutional Neural Network (CNN) is a class of artificial neural network (ANN) that has become a prominent approach in computer vision problems since the remarkable results presented on the object identification competition ImageNet Large Scale Visual Recognition Competition (ILSVRC) in 2012 [107]. CNN is a deep learning model for processing data with a grid pattern that was created to learn spatial hierarchies of information automatically and adaptively, from low- to high-level patterns, based on the organization of animal visual cortex [108]. CNN is typically comprised of three types of layers [109]:
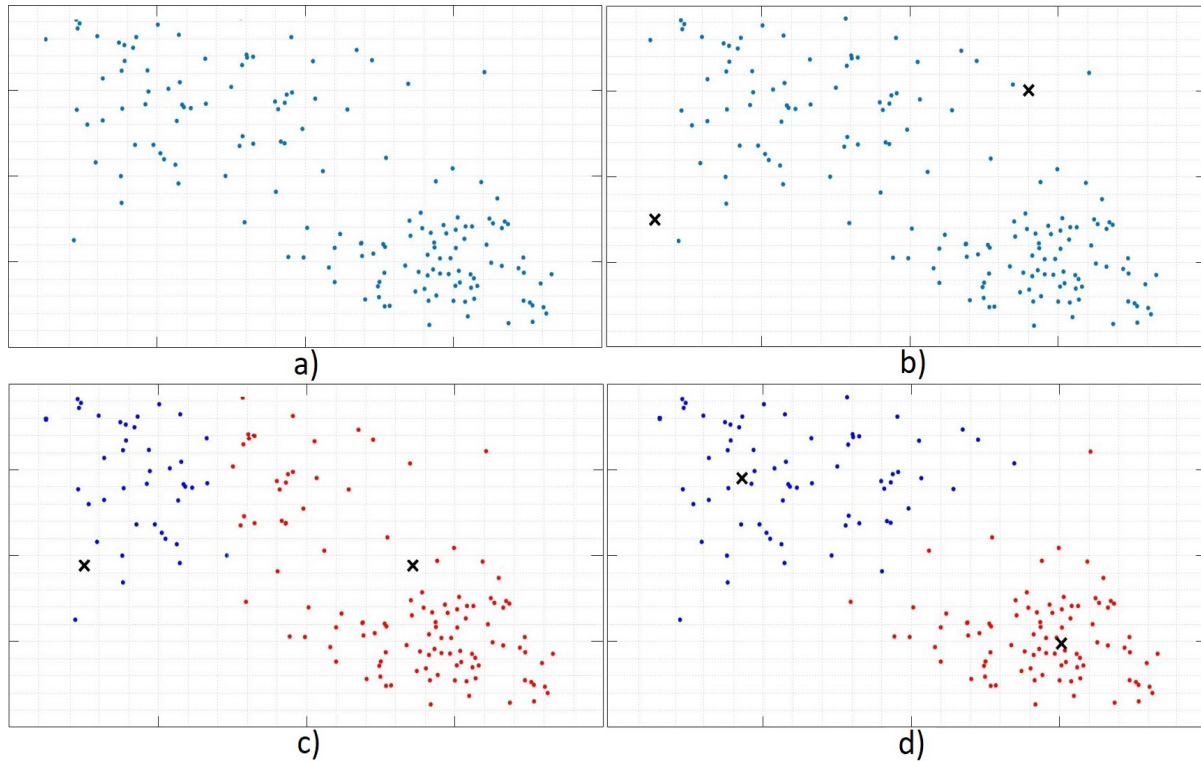
Figure 1.9: k-Means algorithm. Training observations are shown as points and cluster centroids as crosses. a) Original dataset, b) initial cluster centroids, and c) and d) are illustrations of k-means running.

1.3.3.1 Convolution Layer

Convolution is a linear operation used for feature extraction that involves applying a small array of numbers called a kernel over a tensor or array of numbers. A feature map is created by calculating a dot product between the kernel and the input tensor at each location of the tensor and combined to obtain the output value in the corresponding position of the output tensor (Figure 1.10a). This approach is repeated with several kernels to create a number of feature maps that reflect different features of the input tensors; therefore, different kernels can represented distinct characteristics (Figure 1.10b).

The spatial of the output is determined by some hyperparameters [110]:

- **Stride** - is the step each time the filter is slid or is the distance between two consecutive kernel positions;
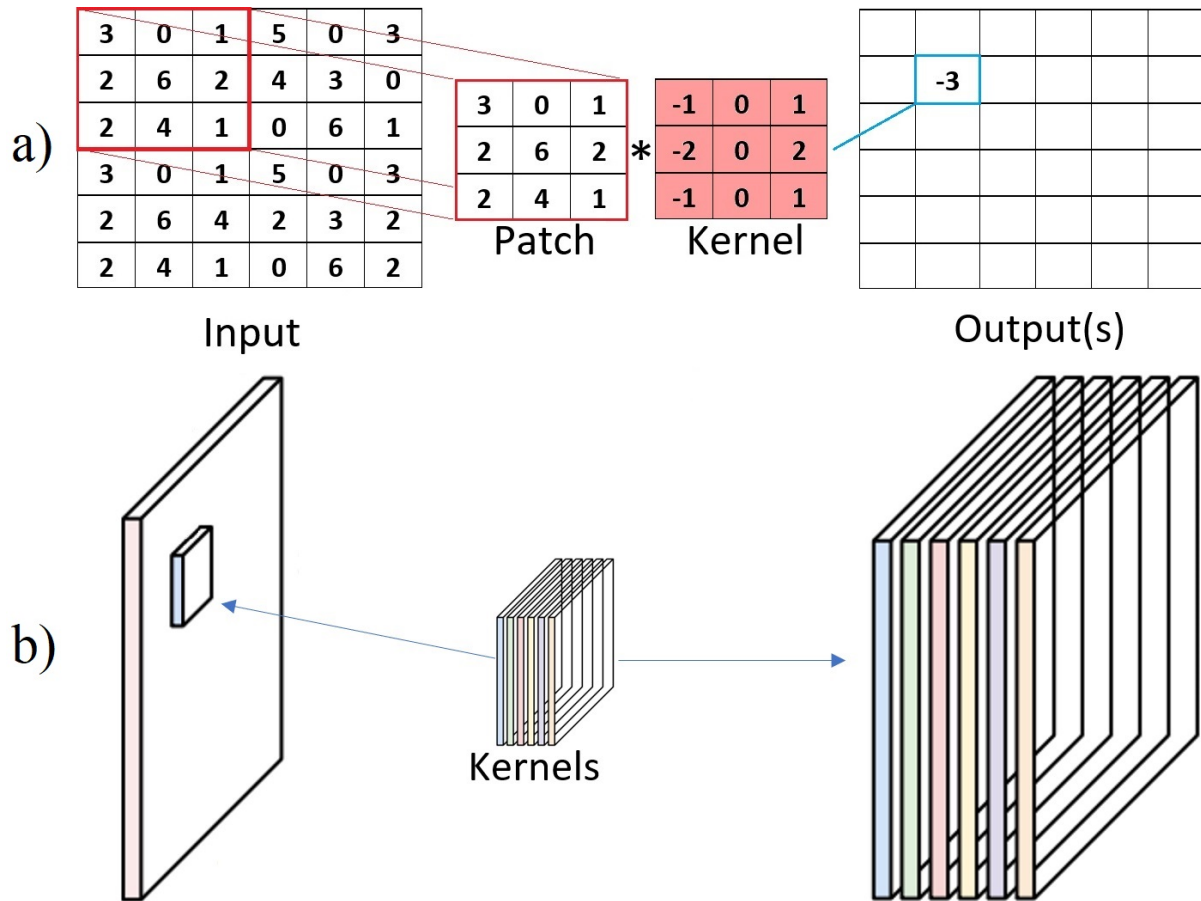
Figure 1.10: a) Example of convolution operation and b) multiple kernels creating different characteristics.

- **Padding** - the convolution process prevents the kernel from overlapping the outermost element, reducing the output height and width in comparison to the input tensor. Padding is a method for dealing with this problem that involves adding rows and columns of zeros on each side of the input tensor in order to fit the centre of a kernel on the outermost element while preserving the same input-output dimensions during the convolution process;

- **Kernel** - number and size of kernels;

- **Activation Function** - is a nonlinear activation function that receives the outputs of a linear operation and determines whether or not it should activate (Figure 1.11).
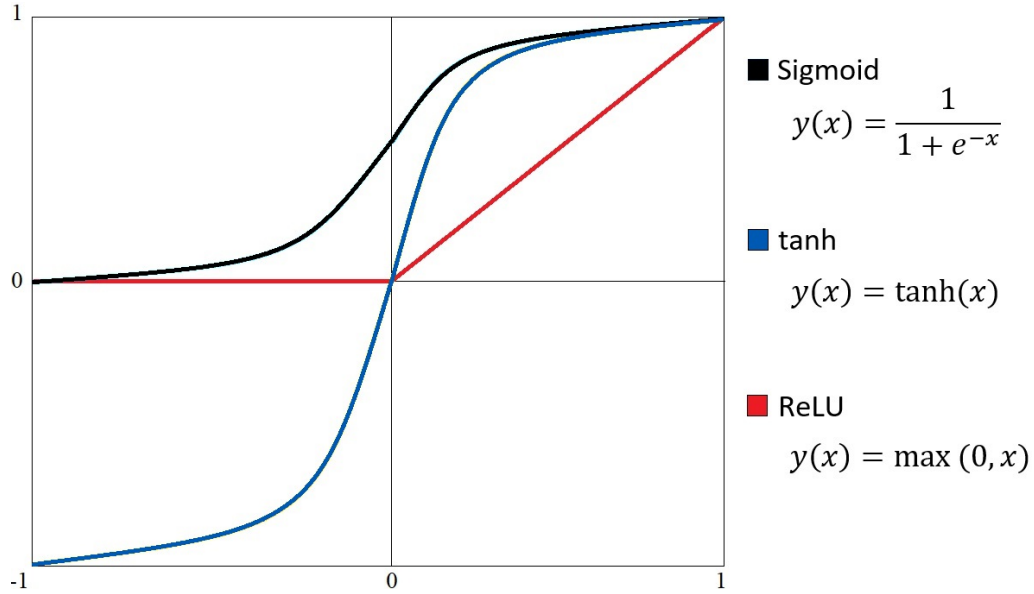
Figure 1.11: Example of nonlinear activation functions.

1.3.3.2 Pooling Layer

A pooling layer decreases the number of learnable parameters by performing a downsampling operation on the feature maps, reducing the output dimensionality and introducing translation invariance to small shifts and distortions (Figure 1.12).
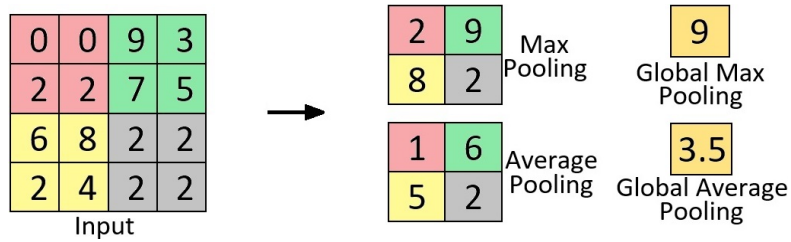
Figure 1.12: Example of a 2×2 max pooling and average pooling and 1×1 global max and global average pooling extracted from a 4×4 input tensor.

1.3.3.3 Fully-Connected Layer

The final convolution or polling layer is generally transformed into a one-dimensional vector of numbers, and connected to one or more fully-connected layers, with each input connected to each output by a learnable weight. Just the features extracted and downsampled are produced in the previous layers, they are projected to the final outputs of the network by a subset of fully

connected layers, such as the probabilities for each class in classification tasks. The number of output nodes in the final fully - connected layer is usually equivalent to the total of classes. After each fully-connected layer, a nonlinear function is applied [109].

Usually, the activation function for the final fully-connected layer is selected according to each task. Softmax is a common choice for the last layer activation function to calculate the class probabilities from the class scores. It condenses the class scores into values between 0 and 1 that sum up to 1 [111].

During the training process, kernels and weights are adjusted to minimise the discrepancies between predicted and ground truth outcomes. The backpropagation algorithm is a typical approach for training neural networks where gradient descent optimization and loss function play a crucial roles. A loss function measures the coherence between output predictions via forward propagation and assigned ground truth labels. Gradient descent is a classic optimization algorithm that iteratively adjusts the kernels and weights of the network to minimize the loss.

In addition to the aforementioned techniques, batch normalization is a method for training very deep neural network to enhance performance and speed, and deliver more reliable models. It reduces the number of training epochs in the training phase and stabilizes the learning process. During the training phase, as the parameters of the preceding layers change and, consequently, the distribution of inputs changes, the current layer needs to be updated regularly and readjusted to the current distributions. Batch normalization normalizes the output of the prior activation which result in the weights of the next layers to become no longer optimal. Thus, batch normalization adds two parameters to each layer, multiplying the normalized output by a standard deviation parameter and adding a mean parameter.

1.3.3.4   U-Net

A very popular segmentation CNN model in medical imaging is the U-Net [112]. The U-Net architecture complements a basic CNN by consecutive layers that gradually downsample

or encode, to capture anatomical context, and upsample or decode to allow for precise location, enabling the incorporation of a multiscale spatial context. To integrate various levels of information, high resolution feature representations from the contracting path are combined with upsampled feature maps, allowing the learning of context information. The upsampling path allows the network to transmit context information to higher resolution layers. As a consequence, the upsample and downsample paths produce a u-shaped architecture. The term "U-Net" derives from the similar shape to a "U". Figure 1.13 shows an example of U-Net architecture. .
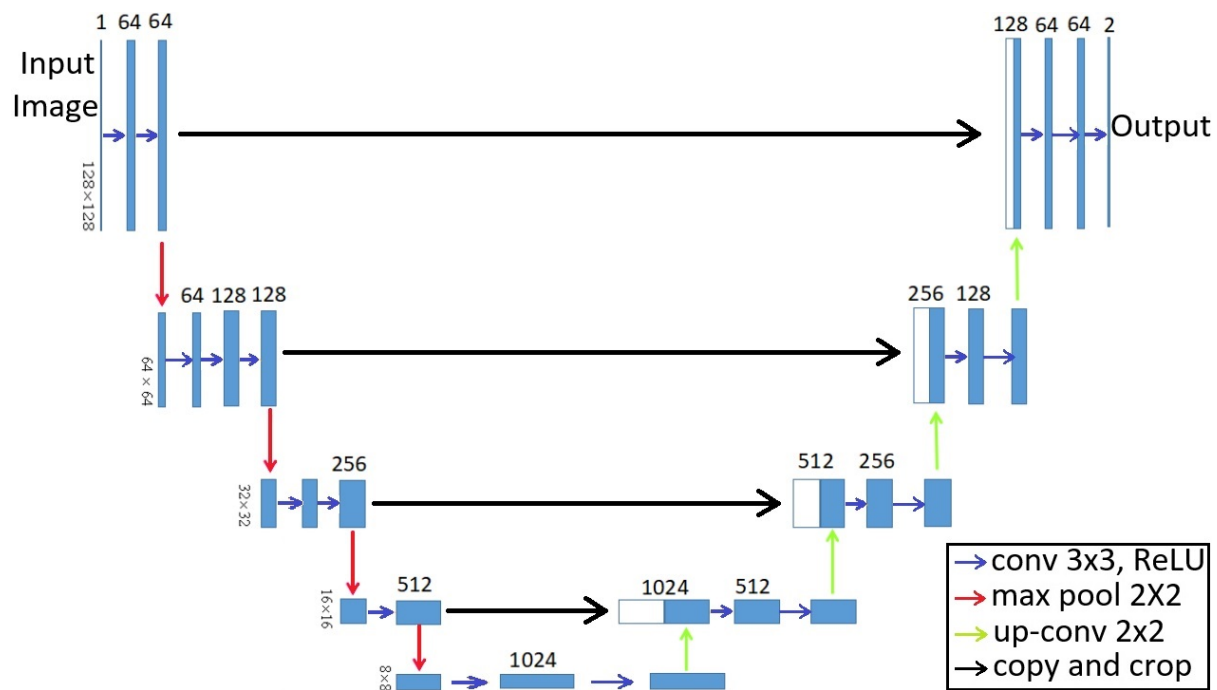
Figure 1.13: U-Net architecture.

Source: Adapted from "U-Net: Convolutional Networks for Biomedical Image Segmentation", by Ronneberger, O., et al., 2015, *In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*, pp. 234-241. doi:10.1007/978-3-319-24574-4_28

1.3.3.5   U-Net++

To improve segmentation accuracy in medical images, [113] proposed a novel segmentation framework based on nested, dense skip connections, and deep supervision called U-Net++

(Figure 1.14).
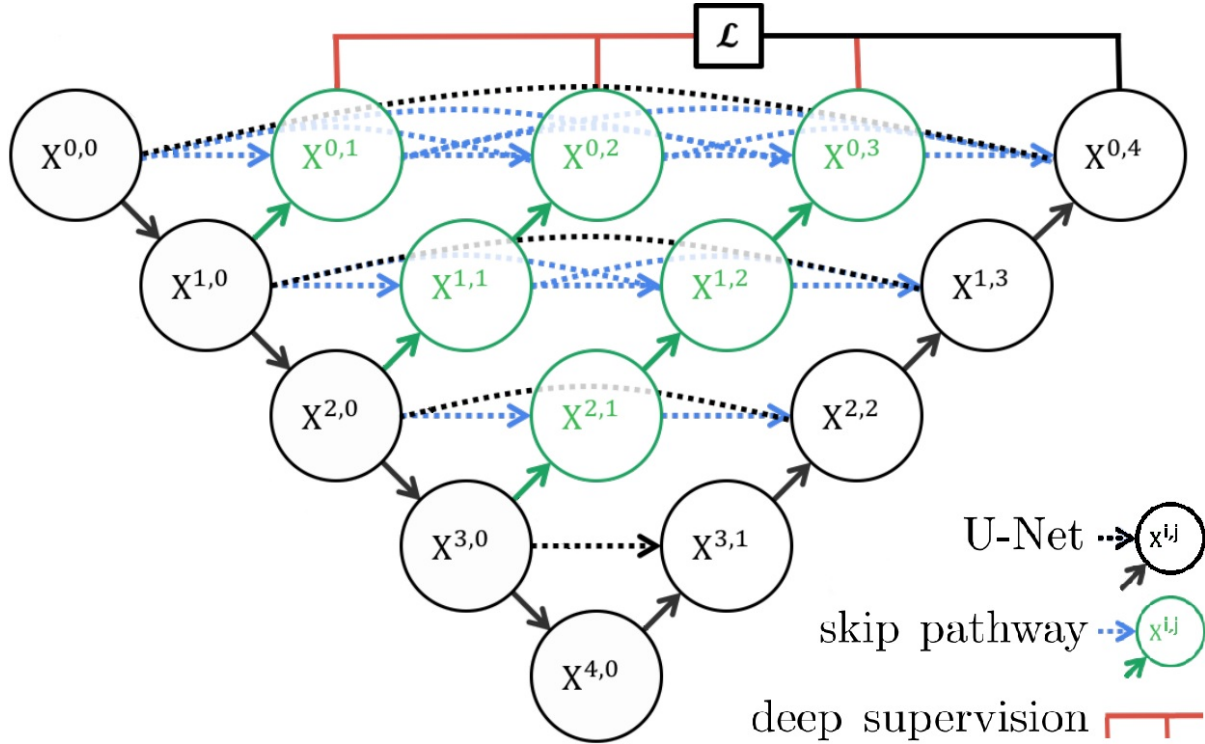
Figure 1.14: U-Net++ architecture where: black - represents a U-Net architecture, green - a redesigned skip pathways, blue dense skip connections, and red - deep supervision.

Source: Adapted from "UNet++: A Nested U-Net Architecture for Medical Image Segmentation", by Zhou, Z., et al., 2018, *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pp. 3-11. doi:10.1007/978-3-030-00889-5_1

In the skip connections employed in U-Net, the decoder receives the encoder feature maps directly resulting in a fusing of semantically different feature maps. The skip pathways proposed in U-Net++ are added to connect the semantic gap between the encoder and decoder subpaths, bringing the semantic level of the encoded feature closer to that of the feature maps waiting in the decoder. As a result, when the received encoder and the associated decoder feature maps are semantically identical, the optimizer may have an easier optimization issue to solve.

U-Net++ additionally uses a deep supervision strategy introduced in [114], which allows for an accurate mode, in which an average of all segmentation branches outputs is performed, or a fast mode, in which just one segmentation branch is used to select the final segmentation

map. Therefore, the model size and speed gain are determined by this choice.

1.3.3.6   U-Net3+

Recently, a modified version of U-Net was introduced by [115] named U-Net3+ (Figure 1.15). U-Net++ takes advantage of nested and dense skip connections, however it does not expand deep enough into full scale information. U-Net3+ combines low-level details with high-level semantics from feature maps at various scales and learns hierarchical representations from full-scale aggregated feature maps.
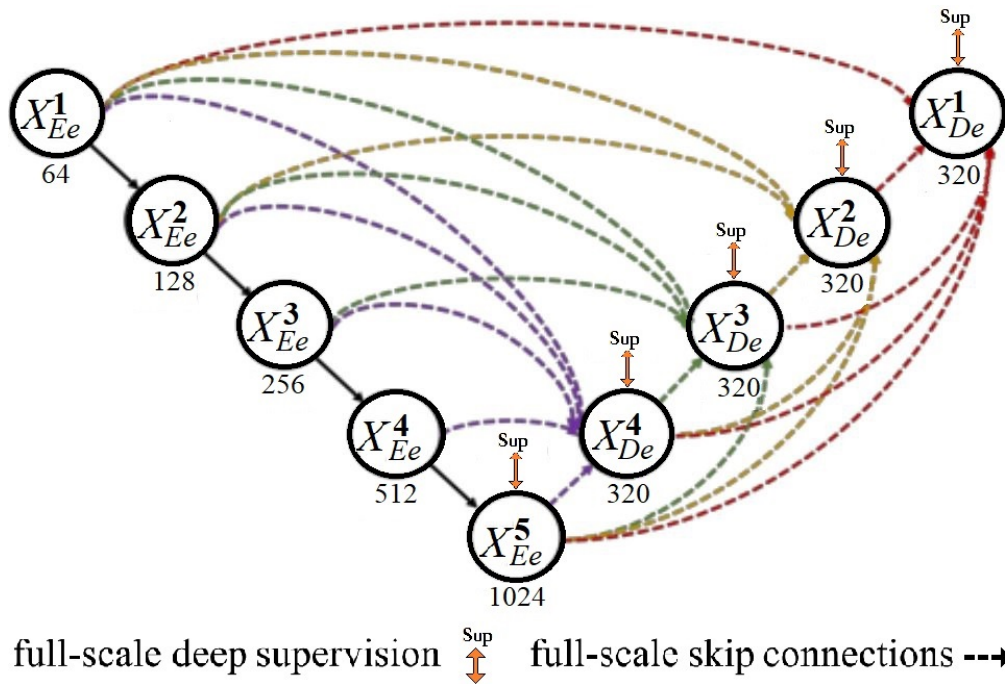
Figure 1.15: U-Net++ architecture where: black - represents a U-Net architecture, green - a redesigned skip pathways, blue dense skip connections, and red - deep supervision.

Source: Adapted from "U-Net 3+: A Full-Scale Connected U-Net for Medical Image Segmentation", by Huang, H., et al., 2020, *In Proceedings of the ICASSP 2020–2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1055–1059. doi: 10.1109/ICASSP40776.2020.9053405

U-Net3+ transforms interconnection between encoder and decoder and also intraconnection between decoder sub-networks using full-scale skip connections. Both U-Net and U-Net++ fall short of extracting substantial information from full scales, failing to learn the po-

sition and boundary of a ROI. To address this issue, each decoder layer in U-Net3+ comprises both small and regular scale feature maps from encoder and larger-scale feature maps from decoder.

Moreover, the U-Net3+ uses full-scale deep supervision to build hierarchical representations from the full-scale aggregated feature maps. Each decoder step in U-Net3+ generates a side output that is supervised by the ground truth. The last layer of each decoder stage is input into a simple 3×3 convolution layer followed by bilinear up-sampling and a sigmoid function, to provide deep supervision. To enhance boundary condition, a multi-scale structural similarity index is applied. Finally, to solve false-positives in a non-organ image, typically produced by noisy information from background, an extra classification task is added, where a 2-dimensional tensor is produced after passing a series of operations (dropout, convolution, maxpooling and sigmoid). Then, a 2-dimensional tensor is converted into a single output of 0 or 1, denoting with or without organs, using an argmax function.

1.4 Modal Analysis

When any structure is struck with a force, it will vibrat transiently until the vibration is absorbed by the structure. The vibration will occur at a set of nature frequencies that are related to the structural geometry and stiffness. A modal analysis determines the natural frequencies and the modes or shapes of the deformation.

Modal analysis plays a significant role in vibration control and damage detection in structures, bridges, and buildings, as well as in the design and optimization of health monitoring. Modal analysis uses the stiffness matrix of the structure to determine mode frequencies and shapes to assess the dynamic behaviour of structures that have been subjected to vibrational mechanical excitation. It shows the movement of distinct sections of a mechanical structure or component under dynamic loading conditions by revealing the natural frequencies, vibration characteristics, and mode shapes of the structure. [116].

Its implementation is based on the presence of specific orthogonality correlations between system eigenvectors (mode shapes), which separates the original equations of vibration into a set of independent differential equations. By solving these separated equations, it produces the response in a series of system eigenvectors called modal expansion or eigenfunction expansion [117].

When exposed to any external forces, all physical structures have natural frequencies which will tend to vibrate. The dispersion of mass and stiffness within the structure determines the frequency of these vibrations. Natural frequencies occur in all physical objects and can reso-nante under the right conditions. Resonance occurs when a dynamic force causes a structure to vibrate at its natural frequency. As a result, a small force can cause a significant vibration response when a structure is in resonance. Figure 1.16 shows the displacement of three distinct points assessed of a single degree of freedom system response.
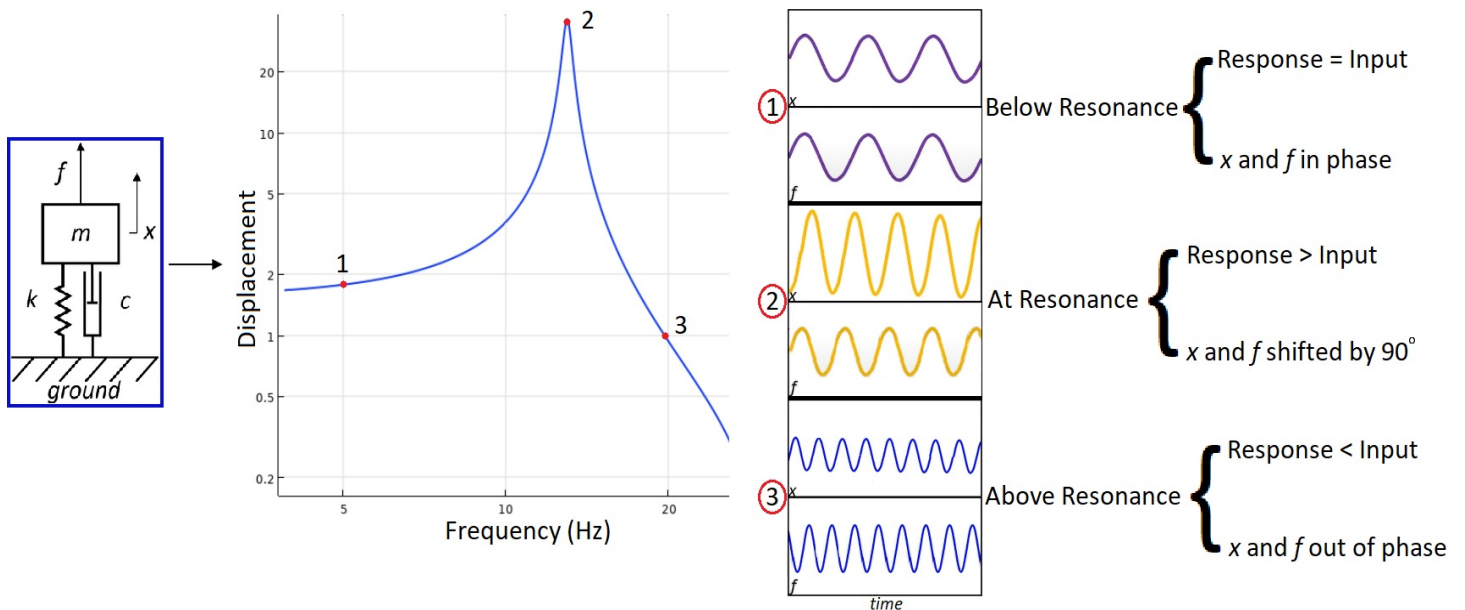
Figure 1.16: Displacement response of a mass-spring-damper system.

Source: Adapted from https://youtu.be/DyZFt3WQ3B8

To determine the natural frequencies and mode shapes, the equation of motion for a multi-

degree-of-freedom damped system is utilized [118, 119]

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{C}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{f}(t), \tag{1.42}$$

where $\mathbf{M}$, $\mathbf{C}$, and $\mathbf{K}$ are the mass, damping, and stiffness matrices respectively, and $\mathbf{x}(t)$ and $\mathbf{f}(t)$ are the nodal displacement and excitation force vectors respectively. To obtain the homogeneous eigenvalues (natural frequencies), for undamped free vibration system ($\mathbf{C} = 0$ and $\mathbf{f}(t) = 0$), hence the resultant equation can be considered in the form

$$\mathbf{x} = \mathbf{X}e^{i\omega t}, \tag{1.43}$$

where $\mathbf{X}$ is the mode shape and ω represents the correponding frequency of each eigenvector or the frequency of mode (natural frequencies). By dividing the equation (1.42) of mass $\mathbf{M}$ and considering $\omega = \sqrt{\frac{k}{m}}$, $\mathbf{C} = 0$, and $\mathbf{f}(t) = 0$, is obtained

$$\ddot{\mathbf{x}}(t) = -\omega^2 \mathbf{X}e^{i\omega t}. \tag{1.44}$$

Finally, substituting equations (1.43 and 1.44) in equation (1.42), the general eigenvalue/eigenvector problem can be written as

$$(\mathbf{K} - \omega^2 \mathbf{M})\mathbf{X}e^{i\omega t} = 0. \tag{1.45}$$

By solving equation (1.45), the natural frequencies ω and corresponding mode shapes $\mathbf{X}$ can be obtained.

1.5   Importance of Research

This work focuses on engineering principles to improve our knowledge over mild traumatic brain injuries like concussions. Though this work, novel skull and brain segmentation techniques were developed that may lead to a new research trend. Additionally, a new concept of

alignment of the coronal and transverse planes was introduced which provides a more accurate alignment method. A future sagittal alignment will be introduced, which will be expected to incorporate and improve the alignment of the coronal and transverse planes. With this additional alignment, it is expected that this method can be implemented in other bones and not just restricted to the skull. A study on the geometry and material properties of the skull was explored and the findings are expected to raise a whole new perspective on how the skull and its natural frequency might generate concussive injury. In the future, this skull can be incorporated into a head model to investigate how the vibrating transient propagates in the brain tissue. This research provides a better understanding of the mechanism that leads to concussive injury and may be used for improving current helmet designs.

Bibliography

[1] Khalil, T. B., Viano, D. C., Smith, D. L. Experimental analysis of the vibrational characteristics of the human skull, *Journal of Sound and Vibration* 1979, 63(3), 351-376. doi: 10.1016/0022-460X(79)90679-5.

[2] Håkansson, B., Brandt, A., Carlsson, P. Resonance frequencies of the human skull in vivo, *The Journal of the Acoustical Society of America* 1994, 95, 1474-1481. doi: 10.1121/1.408535

[3] Charalambopoulos, A., Fotiadis, D. I., Massalas, C. V. Free vibrations of the viscoelastic human skull, *International Journal of Engineering Science* 1998, 36(5-6), doi: 10.1016/S0020-7225(97)00087-6.

[4] Baroudi, A. E., Razafimahery, F., Rakotomanana-Ravelonarivo, L. Study of a spherical head model. Analytical and numerical solutions in fluid–structure interaction approach, *International Journal of Engineering Science* 2012, 51, 1-13, doi: 10.1016/j.ijengsci.2011.11.007.

[5] Baroudi, A. E., Razafimahery, F., Rakotomanana-Ravelonarivo, L. Three-dimensional modal analysis of an idealized human head including fluid–structure interaction effects. *Acta Mech* 2012, 223, 1899–1915. doi: 10.1007/s00707-012-0681-5

[6] Blandford, C., Neuert, M., Jenkyn, T. R. The dynamic properties of the human skull: the effect of impact location and impact energy on the vibrational response of the head, 2017 *CMBEC40 Conference*, Winnipeg MB, May 23-26, 2017.

[7] Chakroun, M., Ghozlen, H., Elloumi, I., Nicolle, S., Détermination du module d'élasticité de la matière cérébrale, *Comptes Rendus Physique* 2009, 10(2-3),pp. 236-241, doi: 10.1016/j.crhy.2009.03.010.

[8] Chakroun, M., Ghozlen, H. Mechanical Impedance of Cerebral Material, *Journal of Modern Physics* 2012, 3(3), 271-273. doi: 10.4236/jmp.2012.33037.

[9] Tator, C. H. Concussions and their consequences: current diagnosis, management and prevention, *Canadian Medical Association Journal* 2013, 185(11), 975-979. doi: 10.1503/cmaj.120039

[10] Mullally, W. J. Concussion, *The American Journal of Medicine* 2017, 130(8), 885-892. doi: 10.1016/j.amjmed.2017.04.016

[11] Meaney, D. F., Smith, D. H. Biomechanics of Concussion, *Clinics in Sports Medicine* 2011, 30(1), 19-31. doi: 10.1016/j.csm.2010.08.009

[12] Churchill, N. W., Hutchison, M. G., Graham, S. J.et al. Baseline vs. cross-sectional MRI of concussion: distinct brain patterns in white matter and cerebral blood flow, *Sci Rep* 2020, 10, 1643. doi: 10.1038/s41598-020-58073-9

[13] Anderson, E. D., Giudice, J. S., Wu, T., Panzer, M. B., Meaney, D. F. Predicting Concussion Outcome by Integrating Finite Element Modeling and Network Analysis, *Front. Bioeng. Biotechnol.* 2020, 8:309. doi: 10.3389/fbioe.2020.00309

[14] Vieira R. C. A., Paiva W. S., de Oliveira D. V., Teixeira M. J., de Andrade A. F., Sousa R. M. C. Diffuse Axonal Injury: Epidemiology, Outcome and Associated Risk Factors, *Front. Neurol.* 2016, 7:178. doi: 10.3389/fneur.2016.00178

[15] Moss, W. C., King, M. J., Blackman, E. G. Skull flexure from blast waves: a mechanism for brain injury with implications for helmet design, *Physical Review Letters* 2009; 103:108702. doi: arXiv:0809.3468v4

[16] Ganpule, S, Alai, A, Plougonven, E, Chandra, N. Mechanics of blast loading on the head models in the study of traumatic brain injury using experimental and computational approaches, *Biomech Model Mechanobiol* 2013; 12:511–531. doi: 10.1007/s10237-012-0421-8

[17] Martin, G. T. Acute brain trauma, *Ann R Coll Surg Engl* 2016, 98:6–10. doi:10.1308/rcsann.2016.0003

[18] Chancellor, S. E., Franz, E. S., Minaeva, O. V., Goldstein, L. E., Pathophysiology of Concussion, *Seminars in Pediatric Neurology* 2019, 30, 14-25. doi: 10.1016/j.spen.2019.03.004

[19] Hyden, J., Petty, B., Sideline management of concussion, *Phys Med Rehabil Clin N Am* 2016, 27(2), 395-409. doi: 10.1016/j.pmr.2015.12.004

[20] Shirley, E., Hudspeth, L. J., Maynard, J. R., Managing sports-related concussions from time of injury through return to play, *J Am Acad Orthop Surg* 2018, 26(14), e279-e286. doi: 10.5435/JAAOS-D-16-00684

[21] McQuivey, K. S., et al., Top-100 Most-Cited Sports-Related Concussion Articles Focus on Symptomatology, Epidemiology, and Demographics, Arthroscopy, Sports Medicine, and Rehabilitation, *Arthr Spor Med Rehab* 2021. doi: 10.1016/j.asmr.2021.06.016

[22] Ommaya, A., Hirsch, A., Martinez, J., The Role of Whiplash in Cerebral Concussion, *SAE Technical Pape*r 660804 1966. doi: 10.4271/660804

[23] Ono, K.; Kikuchi, A.; Nakamura, M.; Kobayashi, H.; Nakamura, N., Human Head Tolerance to Sagittal Impact: Reliable Estimation deduced from Experimental Head Injury using Subhuman Primates and Human Cadaver Skulls 1980, *Proc. 24th Stapp Car Crash Conference, SAE Technical Paper* 801302. doi: 10.4271/801303

[24] Ommaya, A.K.; Hirsch, A.E., Tolerances for cerebral concussion from head impact and whiplash in primates, *Journal of Biomechanics* 1971, 4(1), 13-31. doi: 0.1016/0021-9290(71)90011-X

[25] Gennarelli, T. A., Ommaya A. K., Thibault, L. E., Comparison of translational and rotational head motions in experimental cerebral, *Proc. 15th Stapp Car Crash Conference, SAE Technical Paper* 1972, 39, 797-803.

[26] Gennarelli, T. A., Thibault, L. E., Ommaya, A. K., Pathophysiologic responses to rotational and translational accelerations of the head, *Proc. 16th Stapp Car Crash Conference, SAE Technical Paper* 720970 1972. doi: doi.org/10.4271/720970

[27] Rowson, S., et al., Biomechanical Perspectives on Concussion in Sport, *Sports Medicine and Arthroscopy Review* 2019, 24(3), 100-107. doi: 10.1097/JSA.0000000000000121

[28] Holbourn, A. H. S., Mechanics of head injuries, *Lancet* 1943, ii, 438-441. doi: 10.1016/S0140-6736(00)87453-X

[29] King A. I, Yang K., Zhang L., Hardy W., Viano D. C., Is head injury caused by linear or angular acceleration? *IRCOBI Conference Proceedings* 2003. Lissabon, Portugal.

[30] Gennarelli T. A., Adams J. H., Graham D. I. Acceleration Induced Head Injury in the Monkey. I. The Model, Its Mechanical and Physiological Correlates, *In: Jellinger K., Gullotta F., Mossakowski M. (eds) Experimental and Clinical Neuropathology. Acta Neuropathologica Supplementum* 1981, 7, 23-25. doi: 10.1007/978-3-642-81553-9_7

[31] Gennarelli, T. A., Thibault, L. E., Adams, J. H., Graham, D. I., Thompson, C.J.; Marcincin, R.P. Diffuse axonal injury and traumatic coma in the primate, *Ann Neurol.* 1982, 12, 564-574. doi: 10.1002/ana.410120611

[32] Gennarelli, T. A., Thibault, L. E., Biomechanics of acute subdural hematoma, *The Journal of Trauma: Injury, Infection, and Critical Care* 1982, 22(8), 680-686. doi 10.1097/00005373-198208000-00005

[33] Thibault, L. E., Gennarelli, T. A., Biomechanics of diffuse brain injuries, *Proc. 29th Stapp Car Crash Conference, SAE Technical Paper* 856022 1985.

[34] Zuckerman, S. L., et al., A football helmet prototype that reduces linear and rotational acceleration with the addition of an outer shell, *Journal of Neurosurgery JNS* 2018, 130(5), 1634-1641. doi: 10.3171/2018.1.JNS172733

[35] Unterharnscheidt, F. J., Translational versus rotational acceleration: animal experiments with measured input, *Proc. 15th Stapp Car Crash Conference, SAE Technical Paper* 710880 1971. doi: 10.4271/710880

[36] Joyce T, Gossman W, Huecker M. R., Pediatric Abusive Head Trauma. 2021 Aug 26. *In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing*, 2021 Jan–. PMID: 29763011.

[37] Meaney, D. F., Smith, D. H., Biomechanics of Concussion, *Clinics in Sports Medicine* 2011, 30(1), 19-31.doi :10.1016/j.csm.2010.08.009

[38] Suzanne, P., et al. A Multidimensional Approach to Post-concussion Symptoms in Mild Traumatic Brain Injury, *Fonrtiers in Neurology* 2018, 9. doi: 10.3389/fneur.2018.01113

[39] Leddy, J. J., Sandhu, H., Sodhi, V., Rehabilitation of Concussion and Post-concussion Syndrome, *Sports Health* 2012, 4(2), 147-154. doi: 10.1177/1941738111433673

[40] Mutlag, W. K., et al., Feature Extractino Methods: A Review 2020, *Journal of Physics: Conference Series, The 15th International Scientific Conference of Al-Khwarizmi Society* (FISCAS), 1591, 012028.

[41] Hira, Z. M., Gillies, D. F., A Review of Feature Selection and Feature Extraction Methods Applied on Microarray Data, *Advances in Biominformatics* 2015, vol. 2015. doi: 10.1155/2015/198363

[42] Flusser, J., Suk, T., Zitova, B. Moments and Moment Invariants in Pattern Recognition 2009, *Wiley*, Chichester. Doi:10.1002/9780470684757

[43] Hu, M. K. Visual Pattern Recognition by Moment Invariants, *in IRE Transaction on information theory* 1962, 8(2), 179-187. doi: 10.1109/TIT.1962.1057692.

[44] Wafi, N. M., Sabri, N., Yaakob, S. N., Nasir, A. S., Nazren, A. R., Hisham, M. B, Classification of Characters Using Multilayer Perceptron and Simplified Fuzzy ARTMAP Neural Networks, *Advanced Science Letters* 2017, 23(6), 5151-5155. doi: 10.1166/asl.2017.7330

[45] Mathur, S., Sharma, P., Sign Language Gesture Recognition using Zernike Moments and DTW 2018, 2018 *5th International Conference on Signal Processing and Integrated Networks* (SPIN), 586-591. doi: 10.1109/SPIN.2018.8474179

[46] Gharde, S. S., Ramteke, R. J., Kotkar, V. A., Bage, D. D., Handwritten Devanagari numeral and vowel recognition using invariant moments 2016, 2016 *International Conference on Global Trends in Signal Processing, Information Computing and Communication* (ICGTSPICC), 255-260. doi: 10.1109/ICGTSPICC.2016.7955307

[47] Zaeri, N., Hu Li Moments for Low Resolution Thermal Face Recognition 2018, 2018 *UKSim-AMSS 20th International Conference on Computer Modelling and Simulation* (UKSim), 79-83. doi: 10.1109/UKSim.2018.00026

[48] Kaur, P., Pannu, H. S., Comparative analysis of continuous and discrete orthogonal moments for face recognition 2017, 2017 *International conference of Electronics, Communication and Aerospace Technology* (ICECA), 449-453. doi: 10.1109/ICECA.2017.8203724

[49] Gautam, G., Choudhary, K., Chatterjee, S., Kolekar, M. H., Facial expression recognition using krawtchouk moments and support vector machine classifier 2017, 2017 *Fourth International Conference on Image Information Processing* (ICIIP), 1-6. doi: 10.1109/ICIIP.2017.8313685

[50] Kahyaei, S., Moin, M. S., Robust matching of fingerprints using pseudo-Zernike moments 2016, 2016 *4th International Conference on Control, Instrumentation, and Automation* (ICCIA), 116-120. doi: 10.1109/ICCIAutom.2016.7483146

[51] Li, J., Hu, Y., Zhang, Y., Zhao, Z., Li, J., Zhou, W., Finger-vein recognition based on improved Zernike moment 2017, 2017 *Chinese Automation Congress* (CAC), 2152-2157. doi: 10.1109/CAC.2017.8243129

[52] Zhang, Y. D., et al., Alcoholism detection by medical robots based on Hu moment invariants and predator–prey adaptive-inertia chaotic particle swarm optimization, *Computers and Electrical Engineering* 2017, 63, 126-138. doi: 10.1016/j.compeleceng.2017.04.009

[53] Gornale S. S., Patravali P. U., Hiremath P. S., Automatic Detection and Classification of Knee Osteoarthritis Using Hu's Invariant Moments, *Frontiers in Robotics and AI* 2020, 7. doi: 10.3389/frobt.2020.591827

[54] Joo, J. M., Automatic classification of products in the industry via invariant boundary moments, *Industrial Data* 2007, 10(2). doi: 10.15381/idata.v10i2.6256

[55] Silva, R. D.C., Thé, G. A. P., de Medeiros, F. N. S., Geometrical and statistical feature extraction of images for rotation invariant classification systems based on industrial devices 2015, 2015 21st *International Conference on Automation and Computing* (ICAC), 1-6. doi: 10.1109/IConAC.2015.7313946

[56] Nasruding, M. W., et al., Moment Invariants Technique for Image Analysis and Its Applications: A Review, *Journal of Physics: Conference Series* 2021, 1962(1), 012028. doi: 10.1088/1742-6596/1962/1/012028

[57] Mercimek, M., Gulez, K., Mumcu, T.V., Real object recognition using moment invariants. *Sadhana* 2005, 30, 765–775. doi: 10.1007/BF02716709

[58] Xu, D., Li, H., Geometric Moments Invariants. *Pattern Recognition* 2008, 41(1), 240-249. doi: 10.1016/j.patcog.2007.05.001

[59] Li, E., Huang, Y., Xu, D., Li, H., Shape DNA: basic generating functions for geometric moment invariants 2017. *arXiv* preprint arXiv:1703.02242

[60] Fu, B., Zhou, J., Li, Y., Zhang, G., Wang, C., Image Analysis by Modified Legendre Moments, *Pattern Recognition* 2007, 40, 691-704, 2007. doi: 10.1016/j.patcog.2006.05.020

[61] Sridhar, D., Krishna, I. V. M., Face Recognition using Tchebichef Moments, *International Journal of Information and Network Security* 2012, 1(4), 243-254. doi: 10.11591/ijins.v1i4.820

[62] Teague, M. R., Image Analysis via the General Theory of Moments, *Journal of the Optical Society of America* 1980, 70(8), 920–930. doi: 10.1364/JOSA.70.000920

[63] Mukundan, R., Ong, S. H., Lee, P. A., Image Analysis by Tchebichef Moments, *IEEE Transactions on Image Processing* 2001, 10(9), 1357-1364. doi: 10.1109/83.941859

[64] Mukundan, R., Ramakrishnan, K. R., Moment Functions in Image Analysis Theory and Applications, *World Scientific* 1998, 164. doi: 10.1142/3838

[65] Bolourchi, P., Moradi, M., Demirel, H., Uysal, S., Ensembles of classifiers for improved SAR image recognition using pseudo Zernike moments, *The Journal of Defense Modeling and Simulation* 2020, 17(2), 205-211. doi: 10.1177/1548512919844610

[66] Wang, K., Zhu, T., Wang, J., Real-time Terrain Matching Based on 3D Zernike Moments, *Journal of Navigation* 2018, 71(6), 1441-1459. doi: 10.1017/S0373463318000449

[67] Khald, A., Radgui, A., Rziza, M., Omni-Zernike Algorithm for Template Matching in Catadioptric System, *Journal of Computer Science* 2020, 16(12), 1789-1795. doi: 10.3844/jcssp.2020.1789.1795

[68] Hansda, R., Nayak, R., Balabantaray, B. K., Copy-Move Image Forgery detection via Combined Pseudo-Zernike Moment Invariant 2021, *3rd International Conference on Energy, Power and Environment: Towards Clean Energy Technologies*, 1-6. doi: 10.1109/ICEPE50861.2021.9404448

[69] Ortega, R. C., et al., Zernike moment invariants for hand vein pattern description from raw biometric data, *Journal of Electronic Imaging* 2018, 28(5), 053019. doi: 10.1117/1.JEI.28.5.053019

[70] Zhang, Y. D., Jiang, Y., Zhu, W. et al., Exploring a smart pathological brain detection method on pseudo Zernike moment, *Multimed Tools Appl* 2018, 77, 22589–22604. doi: 10.1007/s11042-017-4703-0

[71] S. Li and S. Liao, "Chinese Character Recognition by Pseudo-Zernike Moments," 2019 *IEEE 4th International Conference on Image, Vision and Computing* (ICIVC), 2019, pp. 484-488, doi: 10.1109/ICIVC47709.2019.8981053.

[72] Soundes, B., Larbi, G., Samir, Z., Pseudo Zernike moments-based approach for text detection and localisation from lecture videos, *International Journal of Computational Science and Engineering* (IJCSE) 2019, 19(2). doi: 10.1504/IJCSE.2019.100231

[73] Haddadnia, J., Faez, K., Ahmadi, M., AN EFFICIENT HUMAN FACE RECOGNITION SYSTEM USING PSEUDO Zernike MOMENT INVARIANT AND RADIAL

BASIS FUNCTION NEURAL NETWORK, *International Journal of Pattern Recognition and Artificial Intelligence* 2003, 17(1), 41-62. doi: 10.1142/S0218001403002265

[74] Chong, C. W., Raveendran, P., Mukundan, R., An Efficient Algorithm for Fast Computation of Pseudo-Zernike Moments, *International journal of Pattern Recognition and Artificial Intelligence* 2003, 17(6), 1011-1023. doi: 10.1142/S0218001403002769

[75] Chong, C. W., Raveendran, P., Mukundan, R., The scale invariants of pseudo-Zernike moments, *pattern analysis and applications* 2003, 6, 176-184. doi: 10.1007/s10044-002-01835

[76] Teh, C.H., Chin, R.T., On Image Analysis by the Method of Moments, *IEEE Trans. on Pattern Analysis and Machine Intelligence* 1988, 10(4), pp.496-513.

[77] Belkasim, S. O., Pattern Recognition with Moment Invariants – A Comparative Study and New Results, *Pattern Recognition* 1991, 24(12), 1117-1138. doi: 10.1016/0031-3203(91)90140Z

[78] Al-Rawi, M.S. Fast computation of pseudo Zernike moments. *J. Real-Time Image Process.* 2010, *8*, 3–10, doi:10.1007/s11554-009-0118-0

[79] Dalvit Carvalho da Silva, R., Jenkyn, T. R., Classification of Mammogram Abnormalities Using Legendre Moments, *International Journal of Image and Graphics* 2021, 21(1), 2150010. doi: 10.1142/S0219467821500108

[80] Hjouji, A., EL-Mekkaoui, J., Jourhmane, M., Image Classification by Mixed Finite Element Method and Orthogonal Legendre Moments, *Pattern Recognit. Image Anal.* 2020, 30, 655–673. doi: 10.1134/S1054661820040185

[81] Bingcheng L., Discrete legendre polynomial based adaptive image filtering 2021, *Proc. SPIE 11729, Automatic Target Recognition XXXI*, 117290. doi: 10.1117/12.2588241

[82] Chong, C. W., Raveendran, P., Mukundan, R., Translation and Scale Invariants of Legendre Moments. *Pattern Recognition* 2004, *37*, 119-129, doi: 10.1016/j.patcog.2003.06.003.

[83] Tiagrajah, V. J., Jamaludin, O., Farrukh, H. N., Discriminant Tchebichef based moment features for face recognition 2011, 2011 *IEEE International Conference on Signal and Image Processing Applications* (ICSIPA), 192-197. doi: 10.1109/ICSIPA.2011.6144081

[84] Erdelyi, A., Magunus, W., Oberhettinger, F., Tricomi, F. G., Higher Transcendental Functions. *New York: McGraw-Hill* 1953, *12*.

[85] Mukundan, R., Ong, S. H., Lee, P. A., Image analysis by Tchebichef moments, *IEEE Transactions on Image Processing* 2001, 10(9), 1357–1364. doi: 10.1109/83.941859

[86] Pearson, K., On Lines and Planes of Closest Fit to Systems of Points in Space, *Philosophical Magazine* 1901, 2 (11), 559–572. doi:10.1080/14786440109462720

[87] Hotelling, H, Analysis of a complex of statistical variables into principal components, *Journal of Educational Psychology* 1933, 24, 417–441. doi: 10.1037/h0071325

[88] Hotelling, H., Relations between two sets of variates, *Biometrika* 1936, 28(3-4), 321–377. doi:10.2307/2333955. JSTOR 2333955

[89] Jolliffe, I. T., Principal Component Analysis, Springer Series in Statistics 2002, New York: Springer-Verlag. doi:10.1007/b98835. ISBN 978-0-387-95442-4

[90] Fan, L., et al., Applications of Independent Component Analysis to Image Feature Extraction 2002, *Second International Conference on Image and Graphics*, 4875, 471-476. doi: 10.1117/12.477183

[91] Deniz, O., et al., Face Recognition using Independent Component Analysis and Support Vector Machines, *Pattern Recognition Letters* 2003, 24(13), 2153-2157. doi: 10.1016/S0167-8655(03)00081-3

[92] Hyvarinen, A., Oja, E., Independent Component Analysis: Algorithms and Applications, *Neural Networks* 2000, 13(4:5), 411–430. doi: 10.1016/S0893-6080(00)00026-5

[93] Huang, X., Wang, B., Zhang, L., A New Scheme for Extraction of Affine Invariant Descriptor and Affine Motion Estimation based on Independent Component Analysis, *Pattern Recognition Letters* 2005, 26(9), 1244-1255. doi: 10.1016/j.patrec.2004.11.006

[94] Shlens, J., A Tutorial on Independent Component Analysis 2014, *arXiv*:1404.2986

[95] Dalvit Carvalho da Silva, R., Thé, G. A. P., de Medeiros, F., Rotation-Invariant Image Description from Independent Component Analysis for Classification Purposes 2015, *In Proceedings of the 12th International Conference on Informatics in Control, Automation and Robotics - Volume 2*: ICINCO, 210-216. doi: 10.5220/0005512802100216

[96] Hyvärinen, A., Karhunen, J., Oja, E., Independent Components Analysis, *John Wiley and Sons, Inc.* 2001. doi: 10.1002/0471221317

[97] Fix, E., Hodges, Jr., J. L., Discriminatory Analysis, Nonparametric Discrimination, Consistency Properties 1951, *USAF School of Aviation Medicine, Randolph Field, Tex., Project* 21-49-004, Rcpt. 4, Contract AF41(128)-31.

[98] Fix, E., Hodges, J. L., Discriminatory Analysis: Small Sample Performance 1952, *USAF School of Aviation Medicine, Randolph Field, Tex., Project* 21-49-004, Rept. 11.

[99] Cover, T. M., Hart, P. E., Nearest Neighbor Pattern Classification, *IEEE Transactions on Information Theory* 1967, 13(1), 21-27. doi: 10.1109/TIT.1967.1053964

[100] Patrick, E. A., Fischer III, F. P., A Generalizes k-Nearest Neighbor Rule, *Information and Control* 1970, 16(2), 128-152. doi: 10.1016/S0019-9958(70)90081-1

[101] Webb, A. R., Copsey, K. D., Statistical Pattern Recognition, 3a Edição, Wiley 2011, p. 666. doi: 10.1002/0470854774

[102] Hu, L. Y., Huang, M. W., Ke, S. W. et al., The distance function effect on k-nearest neighbor classification for medical datasets, *SpringerPlus* 2016, 5(1304). doi: 10.1186/s40064-016-2941-7

[103] Alfeilat, H. A. A., et al., Effects of Distance Measure Choice on K-Nearest Neighbor Classifier Performance: A Review, *Big Data* 2019, 7(4), 221-248. doi: 10.1089/big.2018.0175

[104] MacQueen, J., Some methods for classification and analysis of multivariate observations, *In: Proc. 5th Berkeley Symp. Math. Statistics and Probability* 1967, 1, 281-297

[105] Ahmar, A. S., et al., Using K-Means Clustering to Cluster Provinces in Indonesia, *J. Phys.: Conf. Ser.* 2018, 1028, 012006. doi: 10.1088/1742-6596/1028/1/012006

[106] Li, Y., Wu, H., A clustering method based on K-means algorithm, *Physics Procedia* 2012, 25, 1104-1109. doi: 10.1016/j.phpro.2012.03.206

[107] Krizhevsky, A., Sutskever, I., Hinton, G. E., ImageNet Classification with Deep Convolutional Neural Networks, *Advances in Neural Information Processing Systems* 2012, 25, 1097-1105.

[108] Fukushima, K., Neocognitron: a self organizing neural network model for a mechanism of pattern recognition unaffected by shift in position, *Biol Cybern* 1980, 36,193–202. doi: 10.1007/BF00344251

[109] Yamashita, R., Nishio, M., Do, R. K. G. et al., Convolutional neural networks: an overview and application in radiology, *Insights Imaging* 2018, 9, 611–629. doi: 10.1007/s13244-018-0639-9

[110] Albawi, S., Mohammed, T. A., Al-Zawi, S., Understanding of a convolutional neural network, 2017 *International Conference on Engineering and Technology* (ICET), 2017, 1-6. doi: 10.1109/ICEngTechnol.2017.8308186

[111] Sze, V., Chen, Y., Yang, T., Emer, J. S., Efficient Processing of Deep Neural Networks: A Tutorial and Survey, *in Proceedings of the IEEE* 2017, 105(12), 2295-2329. doi: 10.1109/JPROC.2017.2761740

[112] Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*; Springer: Cham, Switzerland, 2015; pp. 234–241, doi:10.1007/978-3-319-24574-4_28.

[113] Zhou, Z.; Rahman, Siddiquee, M.M.; Tajbakhsh, N.; Liang, J. UNet++: A Nested U-Net Architecture for Medical Image Segmentation. In*Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. DLMIA 2018, ML-CDS 2018. Lecture Notes in Computer Science*; Stoyanov, D., Eds.; Springer: Cham, Switzerland, 2018; Volume 11045, doi:10.1007/978-3-030-00889-5_1

[114] Lee, C., et al., Deeply-Supervised Nets, Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics, *PMLR*, 38, 562-570. doi: arXiv:1409.5185

[115] Huang, H., Lin, L., Tong, R., Hu, H., Zhang, Q., Iwamoto, Y., Han, X., Chen, Y. W., Wu, J. U-Net 3+: A Full-Scale Connected U-Net for Medical Image Segmentation. *In Proceedings of the ICASSP 2020–2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Barcelona, Spain, 4–8 May 2020, pp. 1055–1059, doi: 10.1109/ICASSP40776.2020.9053405.

[116] Uttamchandani, D., Guided Wave Optical Components and Device, *Academic Press*, 2006, pp. 353-370. doi: 10.1016/B978-012088481-0/50024-3

[117] Yang, B., Encyclopedia of Vibration, Elsevier, 2001, pp. 1290-1299. doi: 10.1006/rwvb.2001.0112

[118] Makhloufi, A., et al., Embedded Mechatronic Systems 2 (second edition), *ISTE* 2020, pp. 111-155. doi: 10.1016/B978-1-78548-190-1.50005-0

[119] Pastrama, M. I., et al., Modal analysis of nanoindentation data, confirming that reduced bone turnover may cause increased tissue mineralization/elasticity, *Journal of the Mechanical Behavior of Biomedical Materials* 2020, 84, 217-224. doi: 10.1016/j.jmbbm.2018.05.014

Chapter 2

Development of a Convolutional Neural Network-Based Skull Segmentation in MRI Using Standard Tesselation Language Models

2.1 Introduction

Image segmentation is the process of partitioning an image into multiple sections to simplify the image into something more meaningful so that we can easily locate regions of interest (ROI). In medical imaging and analysis, these ROI, identified by the segmentation process in an image scanning system, can represent various structures in the body such as pathologies, tissues, bone, organs, prosthesis, and so forth.

Magnetic resonance imaging (MRI) and computed tomography (CT) are the most common medical image scanning systems used to reveal relevant structures for automated processing of scanned data. Both techniques are excellent in providing noninvasive diagnostic images of organs and structures inside the body. However, CT is not favorable for routine anatomical imaging of the head as it exposes the patient to small doses of ionizing radiation each visit, putting the patient at risk for developing diseases such as cancer. For instance, the study in [1] pointed out that the risk to develop leukemia and brain tumors increases with the radiation exposure from CT scans. On the contrary, MRI scans have difficulty identifying different tissues

because of the low signal-to-noise ratio of MRI. Additionally, due to bones' weak magnetic resonance signal, MRI scans struggle with differentiating bone tissue from other structures. Specifically, as different bone tissues have the tendency to differ more in appearance from one another than from the adjacent muscle tissue, segmentation approaches must be robust to account for the variations in the structure [2]. Thus, bone segmentation from MRI presents a challenging problem. Current biomedical imaging segmentation methods take advantage of deep learning with convolutional neural networks (CNNs) [3], as seen in [4] where they trained a large deep CNN to classify over 1 million high-resolution images with top-1 and top-5 test set error rates of 37.5% and 17.0%, respectively, much better than the previous state-of-the-art technique.

In CNN, each layer contains various neurons fixed in subsequent layers and sharing weighted connections. During training, these layers extract features (such as horizontal or vertical edges) from the training images that allow the CNN to perform certain tasks such as segmentation by recognizing these features in subsequent images. The advantage of CNN over other techniques is that convolutional image filters are learned and adapted in an automated process for a high-level description in the finest optimization process.

With recent advances in graphical processing units and improvements in computational efficiency, CNNs have achieved excellent results in biomedical image segmentation where deep learning approaches can be performed in an efficient and intelligent way. CNN has been extensively applied in musculoskeletal image segmentation tasks such as brain and spine segmentation [5], acute brain hemorrhage [6], vessel segmentation [7], skull stripping in brain MRI [8], knee bone and cartilage segmentation [9], segmentation of craniomaxillofacial bony structures [10], proximal femur segmentation [11], and cardiac image segmentation [12].

A review on deep learning techniques has been performed by Garcia et al. [13] where the authors highlight a promising deep learning framework for segmentation tasks known as U-Net [14]. U-Net is a CNN which uses an encoding downsampling path and an upsampling decoding path for segmentation tasks to increase the resolution of the output, which has shown

high performance when applied on biomedical images [7].

Although numerous MRI segmentation techniques are described in the literature, few have focused on segmenting the skull in MRI. One approach to segment the skull in MRI is mathematical morphology [15]. The authors describe a method where they first remove the brain by using a surface extractor algorithms and mask the scalp using thresholding and mathematical morphology. During the skull segmentation process, the authors use mathematical morphology to omit background voxels with similar intensities as the skull. Using thresholding and morphological operations, the inner and outer skull boundaries are identified, and the results are masked with the scalp and brain volumes to establish a closed and non-intersecting skull boundary. Applying this segmentation method to 44 images, the authors were able to achieve mean dice coefficients of 0.7346, 0.6918, and 0.6337 for shifts CT of 1 mm, 2 mm, and 3 mm, respectively.

Wang et al. [16] utilized statistical shape information in 15 subjects, where the anatomy of interest is differentiated in the CT data by means of constructing an active shape model of the skull surfaces. The automatic landmarking on the coupled surfaces is optimized in statistical shape information by minimizing the description length that included the local thickness information. This method showed a dice coefficient of 0.7500 for one calvarium segmented. A support vector machine (SVM) combining local and global features is used in [17]. Feature vectors are constructed from each voxel in the image that is used as the first entry. The second input for this method uses a combination of intensities of a set of nearby voxels and statistical moments of the local surroundings. This feature vector is then introduced to a trained SVM that classifies the image as either a skull, or something else. By using SVM, the authors found a dice function mean of 0.7500 (0.68 minimum and 0.81 maximum) from 10 patients in a dataset of 40 patients.

The work in [18] introduced a convolutional restricted Boltzmann machine (cRBM) for skull segmentation. This technique incorporates a cRBM shape model into Statistical Parametric Mapping 8 (SPM8) segmentation framework [19, 20] applied in 23 images. This method

reached a median dice score for T1-weighted of 0.7344 and for T1-w + T2-w, 0.7446.

Most recently, the authors of [21] analyzed three methods of skull segmentation and identified multiple factors contributing to the enhancement of the standard of segmentation. Using a data set obtained from 10 patients, they concluded that improved skull segmentation was accomplished by FSL [22] and SPM12 [23], achieving a mean dice coefficient of 0.76 and 0.80 respectively.

These techniques present an effective method with a mean DSC of 0.75 for small datasets; however, for larger datasets or when extended to images collected from different MRI devices where the image suffers from noise and variation in the choice of parameter values, this effectiveness may be compromised.

Arguably, one of the most important components in machine learning and deep learning is the ground truth labels. Careful collection of data and high-quality ground truth labels that will be used to train and test a model is imperative for a successful deep learning project, but comes with a cost in computation energy and may become very time-consuming [24]. Minnema et al. [25] displayed a high overlap with gold standard segmentation by introducing a CNN for skull segmentation in CT scans. In the image processing step, a 3D surface model, which represents the label, was created in the standard tessellation language (STL) file format, a well-established method to represent 3D models [26–29]. To convert the files from CT to STL, segmentation of a high-quality gold standard STL model was performed manually by an experienced medical engineer. The results show a slight one voxel difference between the gold standard segmentation and the CNN segmentation, with a mean dice similarity coefficient of 0.9200.

Therefore, this work aims to introduce the deep learning approach, more precisely U-Net, for the skull segmentation purpose where the ground truth labels are created from CT imaging using the STL representation file format. Figure 2.1 presents the schematic overview of the proposed study. First, STL models are created from 58 CT scans. After being converted into matrices, these images are then overlapped with the MR images to create the gold standard

and the first dataset. Then, using dataset 1, the first CNN is created. To improve the accuracy, 62 MR images are used to generate brain STL models. The models are then converted into matrices to create a set of brain gold standard and a second dataset. A brain segmentation algorithm using a second CNN is created and, through this CNN model and manual corrections, the brain is removed from dataset 1. Finally, this new dataset is presented again to the first CNN topology.
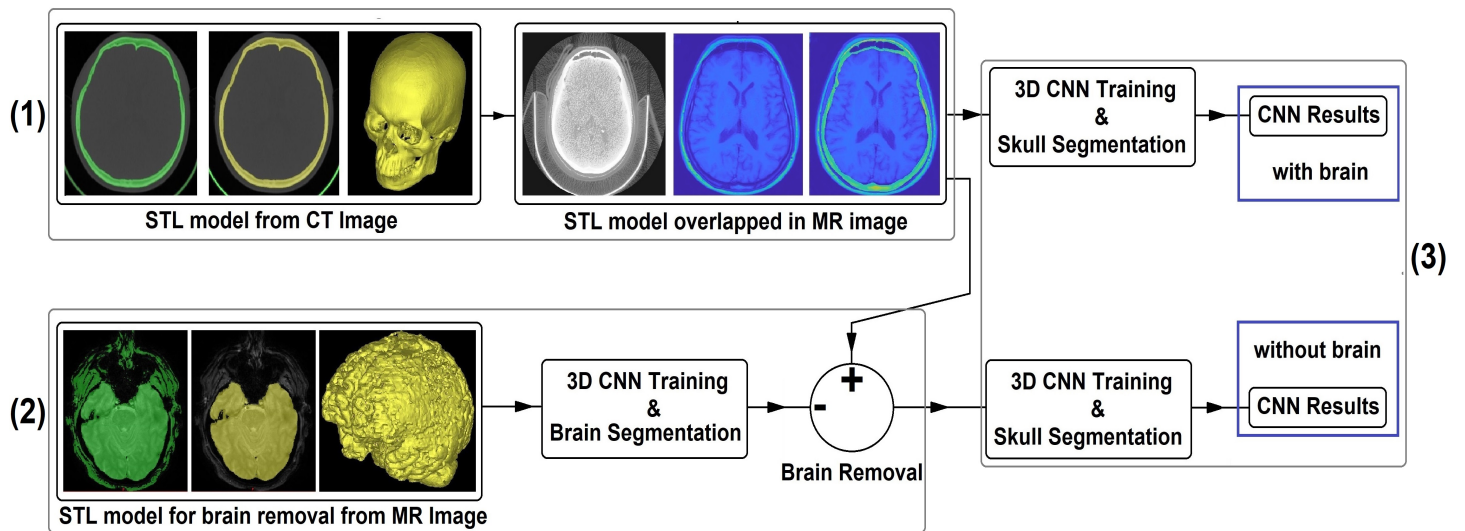
Figure 2.1: (**1**) Standard tessellation language (STL) models are produced from 58 CT scans and then overlapped with the MR images to create the first dataset. (**2**) Sixty-two MR images are used to create brain STL models, and a brain segmentation algorithm is created. The brain segmentation algorithm is combined with manual corrections to remove the brain from dataset 1 to create dataset 2. (**3**) Finally, these 2 datasets are compared using the same CNN topology.

2.2 Materials and Methods

2.2.1 Dataset

We used the cancer imaging archive data collections (TCIA) [30] to search for reliable datasets that contain CT and MRI from the same patient and a minimum variation in the coronal, sagittal, and transverse plane. Fifty-eight volumetric CT and MR images were selected from 4 datasets to meet these criteria:

- CPTAC-GBM [31]: this dataset contains collection from the National Cancer Institute's Clinical Proteomic Tumor Analysis Consortium Glioblastoma Multiform cohort. It contains CR, CT, MR, and SC imaging modalities from 66 participants, totaling 164 studies.

- HNSCC [32–34]: this collection contains CT, MR, PT, RT, RTDOSE, RTPLAN, and RTSTRUCT from 627 subjects in a total of 1177 studies.

- TCGA-HNSC [35]: the cancer genome atlas head–neck squamous cell carcinoma data collection 479 studies from 227 participants from CT, MR, PET, RTDOSE, RTPLAN, and RTSTRUCT modalities.

- ACRIN-FMISO [36–38]: the ACRIN 6684 multicenter clinical trial contains 423 studies applied in 45 participants using CT, MR, and PET modalities.

2.2.2   Data Processing I

As this study aims to use CT scans to create ground truth labels, the first step was to generate the STL models. To perform this task, CT images were imported into Mimics Medical Imaging Software (Materialise, Leuven, Belgium). By using individual global thresholding in combination with manual corrections, the 3D model mesh was built, which allowed the STL model to be constructed (Figure 2.2a–c).

Then, to convert the geometric information (STL model) into a solid domain (matrix), voxelization was performed using the method in [39] on MATLAB R2019B software (Figure 2.2d).

To generate the MRI labels, the STL models extracted from CT ground truth were overlapped into each MRI slice in 3-modal MRI (T1, T2, and FLAIR) using a combination of manual translations, rotations, and scaling. These manual alignments were followed by visual inspection and fine adjustment to ensure good quality (Figure 2.3). T2-weighted scans were included because the border between the skull and cerebrospinal fluid (CSF) can be better delineated, as CSF appears bright in T2-weighted scans and has presented good results in [18,21].
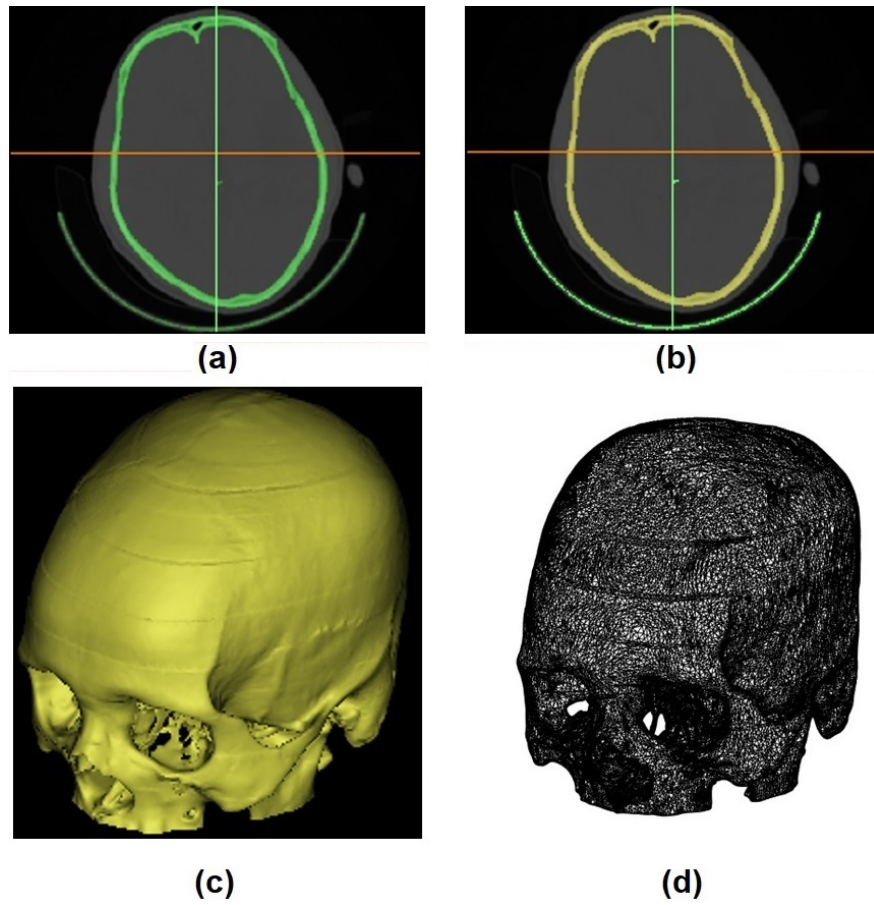
Figure 2.2: (**a**) Thresholding applied in CT scan, (**b**) region growing, (**c**) 3D mesh model (STL model), and (**d**) STL model converted into matrix.
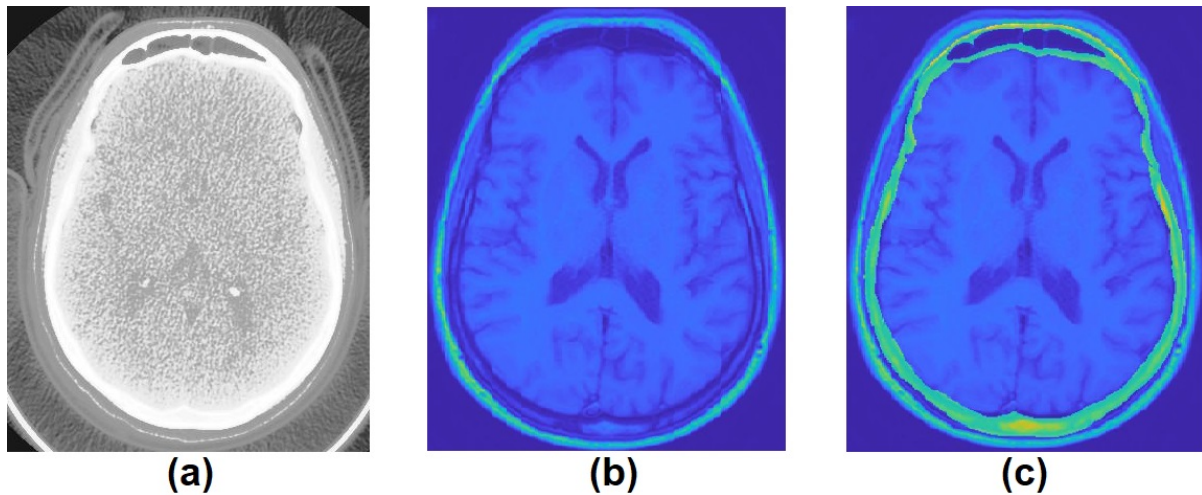
Figure 2.3: (**a**) CT scan, (**b**) MRI, and (**c**) STL model extracted from CT scan overlapped in MRI.

Finally, all images were normalized between the range of 0 and 1 and then resized to 256 × 256 using the nearest neighbor interpolation method to improve the processing time.

2.2.3   Data Processing II

To generate a second comparison, which can lead to an improvement in the accuracy, a reduction in the dataset information was performed by removing regions of non-interest. The idea is to reduce the information content in the dataset by removing the gray and white matter. To perform this task, 62 volumetric MR images were randomly chosen and, in a similar manner explained in Section 2.2.2, brain gold standard labels were created from the MR image. The creation of brain labels is easily performed in MRI as the brain is easy to identify in magnetic resonance. The processing initially starts with the application of the thresholding, followed by region growing, and then creation and extraction of the STL models from the MRI (Figure 2.4).
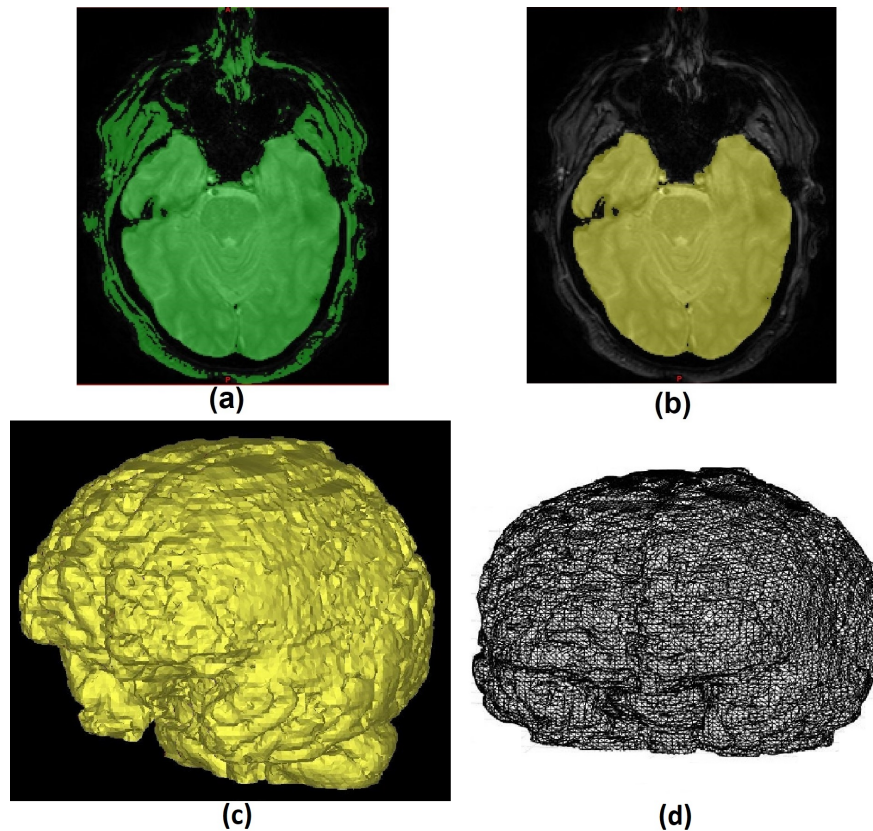
Figure 2.4: (**a**) Thresholding applied in MRI, (**b**) region growing, (**c**) 3D mesh model (STL model), and (**d**) STL model converted into matrix.

2.2.4 CNN Architecture and Implementation Details

The deep learning framework chosen in this paper was U-Net, which was introduced by Ronneberger et al. [14]. This type of CNN was chosen because it works well with very few training images, yields more precise segmentation, and has been used in a number of recent biomedical image segmentation applications [5, 9–11, 40]. This network allows for a large number of feature channels in the upsampling procedure, which contribute in the propagation of context information to the highest resolution layers. The result is a more symmetric expansive path and a U-shaped architecture.

In our implementation, we adopted a 3D U-Net initially developed for brain tumor segmentation in MRI [41]. To avoid class imbalance when using conventional cross entropy loss, a weighted multiclass dice loss function was used as the general loss of the network [42]. Table 2.1 shows the implementation parameter chosen for the skull segmentation. The parameters were chosen to avoid computational error, improve the robustness and generalization ability of the CNN, and obtain a good accuracy for the training set explored in this work.

Table 2.1: Skull segmentation implementation details.

Parameter	Value
Optimizer	Adam
Encoder Depth	4
Filter Size	3
Number of First Encoder Filters	15
Patch Per Image	1
Mini Batch Size	128
Initial Learning Rate	5×10^{-2}

The CNN model was performed on Intel i7-9700 (3.00 GHz) workstations with 64 GB of ram and two 8GB VRAM graphic cards from NVIDIA (RTX 2070 SUPER and RTX 2080). The code was implemented in MATLAB R2019B.

2.2.5   Model Performance Evaluation and Statistical Analysis

To evaluate the CNN segmentation, Dice Similarity Coefficient (DSC) [43], Symmetric Volume Difference (SVD) [44], Jaccard Similarity Coefficient (JSC) [45], Volumetric Overlap Error (VOE) [46], and Hausdorff distances (HD) [47] methods were used.

The dice similarity coefficient is a spatial overlap index that varies from the ranges 0, indicating no spatial overlap between two sets of binary segmentation results, to 1, indicating complete overlap [48]. SVD gives the symmetric difference of the shape and segmentation in terms of dice-based error metric. JSC is a similarity ratio which describes the intersection between the ground truth and the machine segmentation regions over their union. It ranges from 0% to 100% of similarity. VOE is the corresponding JSC error measure. Finally, to measure the segmentation accuracy in terms of distance between the predicted segmentation boundary and the ground truth, Hausdorff distances using Euclidean distance are used.

2.3   Results and Discussion

2.3.1   Performance Analysis

The 58 volumetric CT and MR images were randomly divided into 49 for training and 9 for validation/testing. Table 2.2 presents the statistical analysis of the 9 test images after being trained and tested 10 times for 600 epochs. DSCs, SVDs, JSCs, VOEs, and HDs are calculated from the gold standard labels and predicted labels. DSCs of the skull varies from 0.6847 to 0.8056, with a mean ± SD of 0.7300 ±0 .04, and from 0.9654 to 0.9833 for background, with a mean ± SD of 0.9740 ± 0.007.

To improve the results, by a reduction of regions of non-interest, a 3D U-Net task was performed by using the same software/equipment previously used. The differences between these two methods rely on the creation of the gold standard from different image modalities and 3D U-Net parameters. This approach does not require the CT and MRI scans to overlap in

order to create MRI gold standard labels. Instead, this approach uses it own volumetric MRI
to create the labels through the same processing presented previously.

Using the same datasets from the cancer imaging archive data collection, 62 different vol-
umetric MRI were used to create the brain dataset, where 53 were used for training, three for
validation, and six for testing purpose. Table 2.3 shows the CNN implementation details of
the brain segmentation, and the statistical analysis of the six tested images is presented in the
Table 2.4 after 10 rounds of testing. DSCs, SVDs, JSCs, VOEs, and HDs are calculated from
the brain gold standard labels and brain predicted labels.

Table 2.2: Statistical analysis of the first dataset.

DSC (Skull)	DSC (Background)	SVD (Skull)	JSC (Skull)	JSC (Background)	VOE (Skull)	HD (Skull)
0.8056 ± 0.02	0.9833 ± 0.001	0.1944 ± 0.02	0.6746 ± 0.02	0.9672 ± 0.003	0.3254 ± 0.02	18.25 ± 0.81
0.7706 ± 0.02	0.9805 ± 0.002	0.2294 ± 0.02	0.6267 ± 0.02	0.9618 ± 0.003	0.3733 ± 0.02	23.48 ± 3.44
0.7667 ± 0.02	0.9807 ± 0.002	0.2333 ± 0.02	0.6217 ± 0.02	0.9622 ± 0.003	0.3783 ± 0.02	15.41 ± 3.18
0.7366 ± 0.01	0.9779 ± 0.001	0.2634 ± 0.01	0.5830 ± 0.01	0.9568 ± 0.001	0.4170 ± 0.01	19.66 ± 2.32
0.7177 ± 0.01	0.9731 ± 0.001	0.2823 ± 0.01	0.5597 ± 0.01	0.9476 ± 0.001	0.4403 ± 0.01	40.01 ± 3.33
0.7014 ± 0.02	0.9673 ± 0.001	0.2986 ± 0.02	0.5401 ± 0.03	0.9366 ± 0.001	0.4599 ± 0.03	19.38 ± 3.29
0.6940 ± 0.01	0.9672 ± 0.001	0.3060 ± 0.01	0.5314 ± 0.01	0.9366 ± 0.001	0.4686 ± 0.01	28.41 ± 0.57
0.6917 ± 0.02	0.9654 ± 0.001	0.3083 ± 0.02	0.5287 ± 0.02	0.9331 ± 0.002	0.4713 ± 0.02	39.65 ± 3.00
0.6847 ± 0.01	0.9707 ± 0.001	0.3153 ± 0.01	0.5206 ± 0.01	0.9430 ± 0.001	0.4794 ± 0.01	34.85 ± 2.93
0.7300 ± 0.04	0.9740 ± 0.007	0.2700 ± 0.040	0.5763 ± 0.051	0.9494 ± 0.012	0.4237 ± 0.051	26.57 ± 8.98

Table 2.3: Brain segmentation implementation details.

Parameter	Value
Optimizer	Adam
Encoder Depth	3
Filter Size	5
Number of First Encoder Filters	7
Patch Per Image	2
Mini Batch Size	128
Initial Learning Rate	10^{-3}

These results show an accuracy rate of 0.9244 ± 0.04; however, the brain must be perfectly
extracted. Therefore, after CNN was tested on the 58 initial volumetric MRI, the labels gener-

Table 2.4: Statistical analysis of the brain segmentation.

DSC (Brain)	DSC (Background)	SVD (Brain)	JSC (Brain)	JSC (Background)	VOE (Brain)	HD (Brain)
0.8547 ± 0.02	0.9724 ± 0.001	0.1453 ± 0.02	0.7465 ± 0.03	0.9463 ± 0.001	0.2535 ± 0.03	22.59 ± 4.63
0.9053 ± 0.01	0.9769 ± 0.001	0.0947 ± 0.01	0.8270 ± 0.01	0.9548 ± 0.003	0.1730 ± 0.01	13.96 ± 0.40
0.9436 ± 0.01	0.9910 ± 0.002	0.0564 ± 0.01	0.8934 ± 0.02	0.9821 ± 0.003	0.1066 ± 0.02	10.91 ± 1.84
0.9464 ± 0.01	0.9894 ± 0.002	0.0536 ± 0.01	0.8984 ± 0.02	0.9790 ± 0.003	0.1016 ± 0.02	11.41 ± 0.63
0.9469 ± 0.01	0.9887 ± 0.002	0.0531 ± 0.01	0.8992 ± 0.02	0.9776 ± 0.005	0.1008 ± 0.02	17.78 ± 1.95
0.9491 ± 0.02	0.9888 ± 0.004	0.0509 ± 0.02	0.9035 ± 0.04	0.9778 ± 0.008	0.0965 ± 0.04	14.89 ± 2.34
0.9244 ± 0.04	0.9845 ± 0.008	0.0756 ± 0.04	0.8613 ± 0.06	0.9696 ± 0.015	0.1387 ± 0.06	15.26 ± 4.37

ated in this process were manually corrected using the Matlab program created in [49] in order to optimize brain removal.

After removing the gray and white matter, the modified 58 volumetric images—49 for training and 9 for validation/testing—were then presented to the CNN using the same parameters shown in Table 2.1, and the statistical analysis of the 9 tested images is shown in Tables 2.5 and 2.6.

Table 2.5: Statistical analysis of the second dataset.

DSC (Skull)	DSC (Background)	SVD (Skull)	JSC (Skull)	JSC (Background)	VOE (Skull)	HD (Skull)
0.8288 ± 0.03	0.9857 ± 0.002	0.1712 ± 0.03	0.7076 ± 0.01	0.9719 ± 0.003	0.2924 ± 0.01	09.89 ± 1.58
0.8095 ± 0.01	0.9845 ± 0.001	0.1905 ± 0.01	0.6800 ± 0.01	0.9695 ± 0.001	0.3200 ± 0.01	11.05 ± 0.50
0.8038 ± 0.01	0.9846 ± 0.001	0.1962 ± 0.01	0.6719 ± 0.01	0.9696 ± 0.001	0.3281 ± 0.01	12.32 ± 0.43
0.8052 ± 0.03	0.9839 ± 0.003	0.1948 ± 0.03	0.9684 ± 0.005	0.6739 ± 0.04	0.3261 ± 0.04	10.69 ± 2.25
0.7904 ± 0.01	0.9812 ± 0.001	0.2096 ± 0.01	0.9631 ± 0.003	0.6534 ± 0.01	0.3466 ± 0.01	18.33 ± 0.92
0.7587 ± 0.01	0.9758 ± 0.001	0.2413 ± 0.01	0.9528 ± 0.001	0.6112 ± 0.09	0.3888 ± 0.09	14.09 ± 0.31
0.7627 ± 0.01	0.9756 ± 0.001	0.2373 ± 0.01	0.9523 ± 0.003	0.6164 ± 0.01	0.3836 ± 0.01	14.50 ± 1.33
0.7532 ± 0.03	0.9753 ± 0.003	0.2468 ± 0.03	0.9518 ± 0.006	0.6042 ± 0.03	0.3958 ± 0.03	23.71 ± 1.71
0.7310 ± 0.02	0.9772 ± 0.001	0.2690 ± 0.02	0.9554 ± 0.001	0.5760 ± 0.03	0.4240 ± 0.03	24.91 ± 1.74
0.7826 ± 0.03	0.9804 ± 0.004	0.2174 ± 0.03	0.6439 ± 0.041	0.9617 ± 0.01	0.3561 ± 0.04	15.50 ± 5.28

From Tables 2.5 and 2.6, it can be stated that the reduction in the information contained in the images, such as the removal of the brain, helps in improving the segmentation of the skull bones. In fact, the DSC improvement varied from 2.31% to 7.27% for the skull and 0.24%

Table 2.6: Differences between Dataset 2 minus Dataset 1.

DSC (Skull)	DSC (Background)	SVD (Skull)	JSC (Skull)	JSC (Background)	VOE (Skull)	HD (Skull)
0.0231	0.0024	−0.0231	0.0046	0.0331	−0.0331	−08.36
0.0390	0.0040	−0.0390	0.0077	0.0533	−0.0533	−12.43
0.0371	0.0038	−0.0371	0.0074	0.0503	−0.0503	−03.09
0.0686	0.0060	−0.0686	0.0116	0.0909	−0.0909	−08.96
0.0727	0.0081	−0.0727	0.0155	0.0937	−0.0937	−21.68
0.0573	0.0086	−0.0573	0.0162	0.0711	−0.0711	−05.29
0.0687	0.0083	−0.0687	0.0157	0.0850	−0.0850	−13.91
0.0616	0.0099	−0.0616	0.0187	0.0755	−0.0755	−15.94
0.0463	0.0065	−0.0463	0.0124	0.0555	−0.0555	−09.94
0.0527	0.0064	−0.0527	0.0122	0.0676	−0.0676	−11.07

to 0.99% for the background, which demonstrates that the removal of information in images inherently affects the segmentation of the skull directly. Thus, the initial DSC mean for skull and background improved from 0.7300 ± 0.04 and 0.9740 ± 0.007 to 0.7826 ± 0.03 and 0.9804 ± 0.004, respectively, with a decrease in the standard deviation.

The results represented by DSCs from Table 2.2 in this present study are marginally lower to those reported in [15–18], who achieved mean DSCs of (0.7346, 0.6918, 0.6337), 0.75, 0.75, and (0.7344, 0.7446), respectively, that are lower than those reported in [21] of (0.76 and 0.80). However, the interpretations of the different results found in these studies must be evaluated with caution due to the differences between the databases used, computational methods, and so forth.

This distinction may be attributed to the size of the dataset used. The authors of [40, 50] reported mean dice coefficients of 0.9189 and 0.9800 for CT skull segmentation using CNN when a dataset of 195 and 199 images was used, and the authors of [50] attributed this high DSC to the dataset size and image resolution when compared to other study [25]. Thus, a change in the size of the dataset may have contributed to the improvement in the values of the DSCs. In addition, the geometric disparity, variations, and deformity between the skulls sets become more evident as the dataset increases. As the related works used small datasets, this

aspect may have led to the high DSCs reported.

Furthermore, we used 4 distinct datasets [31, 32, 35, 36] that use different CT and MRI devices with a variety of parameters. These datasets included variations in age, ethnicity, and medical history. In addition, several patients have undergone cephalic surgical treatment which may have altered the skeletal structure of the skull. In total, 40 images have part of the skull removed due to brain abnormalities. These removals may have affected the performance of the segmentation.

From Table 2.6, JSC and HD improved from the initial values (Dataset 1), while SVD and VOE decreased from the initial values. These improvements and reductions can be due to the fact that there is less overlap between the ground truth and the predicted segmentation in the brain region as there is no brain.

One drawback of the presented method occurs during the creation of the gold standard STL model. An expert manually corrected the models by edge smoothing or noise residue removal which may prompt the CNN to learn the defects the expert may have created. Furthermore, during the conversion from the STL model into labels (voxelisation), a quantity of information from the skull-voxel may have been erroneously labeled as background when converted into imaging-voxel. Another disadvantage regards the number of training images. Unfortunately, the amount of usable data that can be acquired from the same patient for both CT and MR images is small because of alignment issues and commonly limited due to ethical and privacy considerations and regulations.

The results found in this article reflect a long-standing search for the development of a technique for bone segmentation in MRI; however, the proposed method DSC (0.7826 ± 0.03) does not exceed the performance of current CT techniques, with a DSC of 0.9189 ± 0.0162 [40]), DSC of 0.9200 ± 0.0400 [25]), and DSC of 0.9800 ± 0.013 [50]).

2.3.2   Comparison between U-Net, U-Net++, and U-Net3+

Further comparison is necessary to see how the DSC behaves in various deep learning methods. U-Net was compared to U-Net++ [51], an encoder–decoder network where a series of dense skip pathways are connected in the encoder and decoder subnetworks, and U-Net3+, a deep learning approach that uses the full-scale aggregated feature map to learn hierarchical representations while, using feature maps in various scales, incorporate low-level details with high-level semantics [52]. Table 2.7 compares U-Net, U-Net++, and U-Net3+ architecture in terms of segmentation accuracy measured by dice similarity coefficient on both datasets 1 and 2. The parameters for each CNN were identical, with an encoder depth of 3, a mini batch size of 16, an initial learning rate of 0.005, and the training was carried out in 100 periods.

Table 2.7: Comparison between U-Net, U-Net++, and U-Net3+ in 4 samples.

Samples	Dataset 1			Dataset 2		
	U-Net	U-Net++	U-Net3+	U-Net	U-Net++	U-Net3+
A	0.6913 ± 0.003	0.5670 ± 0.011	0.7235 ± 0.050	0.8144 ± 0.005	0.5915 ± 0.015	0.8141 ± 0.003
B	0.6545 ± 0.013	0.6292 ± 0.016	0.6589 ± 0.005	0.7462 ± 0.004	0.6393 ± 0.008	0.7567 ± 0.003
C	0.7068 ± 0.005	0.5835 ± 0.087	0.7194 ± 0.004	0.8562 ± 0.007	0.6672 ± 0.017	0.8560 ± 0.006
D	0.6886 ± 0.005	0.5189 ± 0.091	0.6997 ± 0.002	0.7500 ± 0.003	0.6315 ± 0.008	0.7503 ± 0.005

As seen, U-Net3+ outperformed U-Net and U-Net++, obtaining an average improvement over U-Net of 1.51% and 0.26% in datasets 1 and 2. The U-Net algorithm took about an hour to train the 100 epochs, and U-Net3+ took about 2.5 times longer. Therefore, if the training time of the U-Net3+ is disregarded, this network may be used to slightly improve segmentation results.

2.4   Conclusions

This study presents a 3D CNN developed for skull segmentation in MRI where the trained labels were acquired from the same patient CT scans in standard tessellation language. This method initially demonstrated a skull DSC overlap of 0.7300 ± 0.04 and 0.9740 ± 0.007 for

background; however, after the removal of the gray and white matter, DSC reached an average of 0.7826 ± 0.03 and 0.9804 ± 0.004, respectively. Due to the limited number of datasets tested, further research may be undertaken to improve the mean DSC. In summary, the present method is a step forward in the improvement of bone extraction in MRI using CNN to achieve average DSC rates similar to those obtained in CT scans.

Bibliography

[1] Meulepas, J.M.; Ronckers, C.M.; Smets, A.M.; Nievelstein, R.A.; Gradowska, P.; Lee, C.; Jahnen, A.; van Straten, M.; de Wit, M.C.Y.; Zonnenberg, B.; et al. Radiation Exposure From Pediatric CT Scans and Subsequent Cancer Risk in the Netherlands. *JNCI J. Natl. Cancer Inst.* 2018, *111*, 256–263. doi:10.1093/jnci/djy104.

[2] Migimatsu, T.; Wetzstein, G. Automatic MRI Bone Segmentation. 2015, unpublished.

[3] Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; Van Der Laak, J.A.; Van Ginneken, B.; Sánchez, C.I. A survey on deep learning in medical image analysis. *Med. Image Anal.* 2017, *42*, 60–88. doi:10.1016/j.media.2017.07.005.

[4] Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* 2012, *25*, 1097–1105.

[5] Kolařík, M.; Burget, R.; Uher, V.; Říha, K.; Dutta, M.K. Optimized High Resolution 3D Dense-U-Net Network for Brain and Spine Segmentation. *Appl. Sci.* 2019, *9*, 404, doi:10.3390/app9030404.

[6] Ker, J.; Singh, S.P.; Bai, Y.; Rao, J.; Lim, T.; Wang, L. Image Thresholding Improves 3-Dimensional Convolutional Neural Network Diagnosis of Different Acute Brain Hemorrhages on Computed Tomography Scans. *Sensors* 2019, *19*, 2167, doi:10.3390/s19092167.

[7] Livne, M.; Rieger, J.; Aydin, O.U.; Taha, A.A.; Akay, E.M.; Kossen, T.; Sobesky, J.; Kelleher, J.D.; Hildebr, K.; Frey, D.; et al. A U-Net Deep Learning Framework for High Performance Vessel Segmentation in Patients with Cerebrovascular Disease. *Front. Neurosci.* 2019, *13*, doi:10.3389/fnins.2019.00097.

[8] Hwang, H.; Rehman, H.Z.U.; Lee, S. 3D U-Net for Skull Stripping in Brain MRI. *Appl. Sci.* 2019, *9*, 569, doi:10.3390/app9030569.

[9] Ambellan, F.; Tack, A.; Ehlke, M.; Zachow, S. Automated segmentation of knee bone and cartilage combining statistical shape knowledge and convolutional neural networks. *Med. Image Anal.* 2019, *52*, 109–118, doi:10.1016/j.media.2018.11.009.

[10] Dong, N.; Li, W.; Roger, T.; Jianfu, L.; Peng, Y.; James, X.; Dinggang, S.; Qian, W.; Yinghuan, S.; Heung-Il, S.; et al. Segmentation of Craniomaxillofacial Bony Structures from MRI with a 3D Deep-Learning Based Cascade Framework. *Mach. Learn. Med. Imaging* 2017, *10541*, 266–273, doi:10.1007/978-3-319-67389-9_31.

[11] Deniz, C.; Siyuan, X.; Hallyburton, S.; Welbeck, A.; Babb, J.; Honig, S.; Cho, K.; Chang, G. Segmentation of the Proximal Femur from MR Images Using Deep Convolutional Neural Networks. *Sci. Rep.* 2018, *8*, doi:10.1038/s41598-018-34817-6.

[12] Chen, C.; Qin, C.; Qiu, H.; Tarroni, G.; Duan, J.; Bai, W.; Rueckert, D. Deep Learning for Cardiac Image Segmentation: A Review. *Front. Cardiovasc. Med.* 2020, *7*, doi:10.3389/fcvm.2020.00025.

[13] Garcia-Garcia, A.; Orts-Escolano, S.; Oprea, SVillenaMartinez, V.; Garcia-Rodriguez, J. A review on deep learning techniques applied to semantic segmentation. *arXiv* 2017, arXiv:1704.06857.

[14] Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Proceedings of the Medical Image Computing and Computer-*

Assisted Intervention—MICCAI 2015; Springer: Cham, Switzerland, 2015; pp. 234–241, doi:10.1007/978-3-319-24574-4_28.

[15] Dogdas, B.; Shattuck, D.; Leahy, R. Segmentation of skull and scalp in 3-D human MRI using mathematical morphology. *Hum. Brain Mapp.* 2005, *26*, 273–285. doi:10.1002/hbm.20159.

[16] Wang, D.; Shi, L.; Chu, W.; Cheng, J.; Heng, P. Segmentation of human skull in MRI using statistical shape information from CT data. *J. Magn. Reson. Imaging* 2009, *30*, 490–498, doi:10.1002/jmri.21864.

[17] Sjölund, J.; Järlideni, A.; Andersson, M.; Knutsson, H.; Nordström, H. Skull Segmentation in MRI by a Support Vector Machine Combining Local and Global Features. *In Proceedings of the 22nd International Conference on Pattern Recognition*, Stockholm, Sweden, 24–28 August 2014, pp. 3274–3279, doi:10.1109/ICPR.2014.564.

[18] Puonti, O.; Leemput, K.; Nielsen, J.; Bauer, C.; Siebner, H.; Madsen, K.; Thielscher, A. Skull segmentation from MR scans using a higher-order shape model based on convolutional restricted Boltzmann machines. In *Proceedings of the Medical Imaging 2018: Image Processing, International Society for Optics and Photonics*: Bellingham, WA, USA, 2018; Volume 10574, doi:10.1117/12.2293073.

[19] Ashburner, J.; Friston, K.J. Unified Segmentation. *NeuroImage* 2005, *26*, 839–851, doi:10.1016/j.neuroimage.2005.02.018.

[20] UCL Queen Square Institute of Neurology. Statistical Parametric Mapping. 17 September 2020. Available online: https://www.fil.ion.ucl.ac.uk/spm/software/spm8/ (accessed on 01 February 2021).

[21] Nielsen, J.D.; Madsen, K.H.; Puonti, O.; Siebner, H.R.; Bauer, C.; Madsen, C.G.; Saturnino, G.B.; Thielscher, A. Automatic skull segmentation from MR images for realistic

volume conductor models of the head: Assessment of the state-of-the-art. *NeuroImage* 2018, *174*, 587–598. doi:10.1016/j.neuroimage.2018.03.001.

[22] Smith, S.M.; Jenkinson, M.; Woolrich, M.W.; Beckmann, C.F.; Behrens, T.E.; Johansen-Berg, H.; Bannister, P.R.; De Luca, M.; Drobnjak, I.; Flitney, D.E.; et al. Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage* 2004, *23*, S208–S219. doi:10.1016/j.neuroimage.2004.07.051.

[23] UCL Queen Square Institute of Neurology. Statistical Parametric Mapping. 17 February 2021. Available online: https://www.fil.ion.ucl.ac.uk/spm/software/spm12/ (accessed on 01 March 2021).

[24] Yamashita, R.; Nishio, M.; Do, R.; Togashi, K. Convolutional Neural Networks: An overview and application in radiology. *Insights Imaging* 2018, *9*, 611–629, doi:10.1007/s13244-018-0639-9.

[25] Minnema, J.; Eijnatten, M.; Kouw, W.; Diblen, F.; Mendrik, A.; Wolff, J. CT Image Segmentation of Bone for Medical Additive Manufacturing using a Convolutional Neural Network. *Comput. Biol. Med.* 2018, *103*, 130–139, doi:10.1016/j.compbiomed.2018.10.012.

[26] Ferraiuoli, P.; Taylor, J.C.; Martin, E.; Fenner, J.W.; Narracott, A.J. The Accuracy of 3D Optical Reconstruction and Additive Manufacturing Processes in Reproducing Detailed Subject-Specific Anatomy. *J. Imaging* 2017, *3*, doi:10.3390/jimaging3040045.

[27] Im, C.H.; Park, J.M.; Kim, J.H.; Kang, Y.J.; Kim, J.H. Assessment of Compatibility between Various Intraoral Scanners and 3D Printers through an Accuracy Analysis of 3D Printed Models. *Materials* 2020, *13*, 4419, doi:10.3390/ma13194419.

[28] Di Fiore, A.; Stellini, E.; Savio, G.; Rosso, S.; Graiff, L.; Granata, S.; Monaco, C.; Meneghello, R. Assessment of the Different Types of Failure on Anterior Cantilever

Resin-Bonded Fixed Dental Prostheses Fabricated with Three Different Materials: An In Vitro Study. *Appl. Sci.* 2020, *10*, doi:10.3390/app10124151.

[29] Zubizarreta-Macho, Á; Triduo, M.; Alonso, Pérez-Barquero, J.; Guinot Barona, C.; Albaladejo Martínez, A. Novel Digital Technique to Quantify the Area and Volume of Cement Remaining and Enamel Removed after Fixed Multibracket Appliance Therapy Debonding: An In Vitro Study. *J. Clin. Med.* 2020, *9*, 1098, doi:10.3390/jcm9041098.

[30] Clark, K.; Vendt, B.; Smith, K.; Freymann, J.; Kirby, J.; Koppel, P.; Moore, S.; Phillips, S.; Maffitt, D.; Pringle, M.; Tarbox, L.; Prior, F. The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository. *J. Digit. Imaging* 2013, *26*, 1045–1057, doi:10.1007/s10278-013-9622-7.

[31] National Cancer Institute Clinical Proteomic Tumor Analysis Consortium (CPTAC). Radiology Data from the Clinical Proteomic Tumor Analysis Consortium Glioblastoma Multiforme [CPTAC-GBM] Collection [Dataset]. *T Cancer Imaging Arch.* 2018, doi:10.7937/k9/tcia.2018.3rje41q1.

[32] Grossberg, A.; Elhalawani, H.; Mohamed, A.; Mulder, S.; Williams, B.; White, A.L.; Zafereo, J.; Wong, A.J.; Berends, J.E.; AboHashem, S.; et al. Anderson Cancer Center Head and Neck Quantitative Imaging Working Group HNSCC [Dataset]. *Cancer Imaging Arch.* 2020, doi:10.7937/k9/tcia.2020.a8sh-7363.

[33] Grossberg, A.; Mohamed, A.; Elhalawani, H.; Bennett, W.; Smith, K.; Nolan, T.; Williams, B.; Chamchod, S.; Heukelom, J.; Kantor, M.; et al Imaging and Clinical Data Archive for Head and Neck Squamous Cell Carcinoma Patients Treated with Radiotherapy. *Sci. Data* 2018, *5*, 180173, doi:10.1038/sdata.2018.173.

[34] Elhalawani, H.; Mohamed, A.S.; White, A.L.; Zafereo, J.; Wong, A.J.; Berends, J.E.; AboHashem, S.; Williams, B.; Aymard, J.M.; Kanwar, A.; et al. Matched computed

tomography segmentation and demographic data for oropharyngeal cancer radiomics challenges. *Sci. Data* 2017, *4*, 170077, doi:10.1038/sdata.2017.77.

[35] Zuley, M.L.; Jarosz, R.; Kirk, S.; Lee, Y.; Colen, R.; Garcia, K.; Aredes, N.D. Radiology Data from The Cancer Genome Atlas Head-Neck Squamous Cell Carcinoma [TCGA-HNSC] collection. *Cancer Imaging Arch.* 2016, doi:10.7937/K9/TCIA.2016.LXKQ47MS.

[36] Kinahan, P.; Muzi, M.; Bialecki, B.; Coombs, L. Data from ACRIN-FMISO-Brain. *Cancer Imaging Arch.* 2018, doi:10.7937/K9/TCIA.2018.vohlekok.

[37] Gerstner, E.R.; Zhang, Z.; Fink, J.R.; Muzi, M.; Hanna, L.; Greco, E.; Prah, M.; Schmainda, K.M.; Mintz, A.; Kostakoglu, L.; et al. ACRIN 6684: Assessment of Tumor Hypoxia in Newly Diagnosed Glioblastoma Using 18F-FMISO PET and MRI. *Clin Cancer Res.* 2016, *22*, 5079–5086, doi:10.1158/1078-0432.CCR-15-2529.

[38] Ratai, E.M.; Zhang, Z.; Fink, J.; Muzi, M.; Hanna, L.; Greco, E.; Richards, T.; Kim, D.; Andronesi, O.C.; Mintz, A.; et al. ACRIN 6684: Multicenter, phase II assessment of tumor hypoxia in newly diagnosed glioblastoma using magnetic resonance spectroscopy. *PLOS ONE* 2018, *13*, doi:10.1371/journal.pone.0198548

[39] Pati, S.; Ravi, B. Voxel-based representation, display and thickness analysis of intricate shapes. *In Proceedings of the Ninth International Conference on Computer Aided Design and Computer Graphics* (CAD-CG'05), Hong Kong, China, 7–10 December 2005, Volume 6, doi:10.1109/CAD-CG.2005.86.

[40] Dalvit Carvalho da Silva, R.; Jenkyn, T.R.; Carranza, V.A. Convolutional Neural Network and Geometric Moments to Identify the Bilateral Symmetric Midplane in Facial Skeletons from CT Scans. *Biology* 2021, *10*, 182. doi:10.3390/biology10030182

[41] Çiçek, Ö; Abdulkadir, A.; Lienkamp, S.; Brox, T.; Ronneberger, O. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. *In Proceedings of the Medical Image Computing and Computer-Assisted Intervention*—MICCAI 2016, Lecture Notes in Computer Science, Athens, Greece, 17–21 October 2016; Volume 9901, doi:10.1007/978-3-319-46723-8_49.

[42] Sudre, C.; Li, W.; Vercauteren, T.; Ourselin, S.; Cardoso, M. Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations. *In Proceedings of the Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: Third International Workshop*, Quebec City, QC, Canada, 14 September 2017; pp. 240–248, doi:10.1007/978-3-319-67558-9_28.

[43] Dice, L. Measures of the Amount of Ecologic Association Between Species. *Ecology* 1945, *26*, 297–302, doi:10.2307/1932409.

[44] Schenk, A.; Prause, G.; Peitgen, H.O. Efficient semiautomatic segmentation of 3D objects in medical images. *In Proceedings of the Medical Image Computing and Computer-Assisted Intervention*—MICCAI 2000, Lecture Notes in Computer Science, 2000, Pittsburgh, PA, USA, 11–14 October 2000; Volume 1935, pp. 186–195, doi:10.1007/978-3-540-40899-4_19

[45] Jaccard, P. Distribution de la flore alpine dans le bassin des Dranses et dans quelques regions voisines. *Bull. Soc. Vaudoise Des Sci. Nat.* 1901, *37*, 241–272, doi:10.5169/seals-266440.

[46] Rusko, LBekes, G.; Fidrich, M. Automatic segmentation of the liver from multi- and single-phase contrast-enhanced CT images. *Med Image Anal.* 2009, *13*, 871–882, doi:10.1016/j.media.2009.07.009.

[47] Karimi, D.; Salcudean, S. Reducing the Hausdorff Distance in Medical Image Segmentation With Convolutional Neural Networks. *IEEE Trans. Med Imaging* 2020, *39*, 499–513, doi:10.1109/TMI.2019.2930068.

[48] Zou, K.H.; Warfield, S.K.; Bharatha, A.; Tempany, C.M.; Kaus, M.R.; Haker, S.J.; Wells, W.M., III; Jolesz, F.A.; Kikinis, R. Statistical validation of image segmentation quality based on a spatial overlap index. *Acad. Radiol.* 2004, *11*, 178–189, doi:10.1016/s1076-6332(03)00671-8.

[49] Wang, G. Paint on an BW Image (Updated Version), MATLAB Central File Exchange. Available online: https://www.mathworks.com/matlabcentral/fileexchange/32786-paint-on-an-bw-image-updated-version (accessed on 10 September 2020).

[50] Kodym, O.; Španěl, M.; Herout, A. Segmentation of defective skulls from ct data for tissue modelling. *arXiv* 2019, arXiv:1911.08805.

[51] Zhou, Z.; Rahman, Siddiquee, M.M.; Tajbakhsh, N.; Liang, J. U-Net++: A Nested U-Net Architecture for Medical Image Segmentation. *In Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. DLMIA 2018, MLCDS 2018. Lecture Notes in Computer Science*; Stoyanov, D., Eds.; Springer: Cham, Switzerland, 2018; Volume 11045, doi:10.1007/978-3-030-00889-5_1

[52] Huang, H.; Lin, L.; Tong, R.; Hu, H.; Zhang, Q.; Iwamoto, Y.; Han, X.; Chen, Y.W.; Wu, J. U-Net 3+: A Full-Scale Connected U-Net for Medical Image Segmentation. *In Proceedings of the ICASSP 2020—2020 IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP), Barcelona, Spain, 4–8 May 2020; pp. 1055–1059, doi:10.1109/ICASSP40776.2020.9053405.

Chapter 3

Enhanced Pre-Processing for Deep Learning in MRI Brain Segmentation using Orthogonal Moments

3.1 Introduction

Segmentation is a very useful technique that may be used to separate and reveal the inner regions of the human body. In several medical imaging analyses, the segmentation of brain structures in Magnetic Resonance Images (MRI) is the first integral step for the assessment of neurological disorders [1, 2]. A fast, accurate, reliable, and automated segmentation method of the brain structure can improve the diagnosis, evaluation, and treatment of the neurological diseases. Furthermore, automated methods can assist specialists in manual segmentation since the computational resources for performing manual segmentation at large scale are time consuming, can lead to human errors, and demand sophisticated software. This has prompt researchers to investigate the use of deep neural networks as an approach for learning to predict the outcome of a segmentation processing task to improve accuracy while decreasing overall run time.

In recent years, Convolutional Neural Networks (CNN's) have been very successful in the field of medical image processing [11], especially in tasks involving brain segmentation [1, 4–7, 9, 15]. [10] proposed a 3D deep convolutional network (3D DCNN) for brain

segmentation that uses more layers and larger patches than current 3D DCNN's, allowing the network to model complex relationships required for fine-grained brain structures identification. For training such complex models, the proposed work uses deep learning methodology to initialize weights, adjust for internal covariate transition, and minimize overfitting. When comparing the mean dice similarity coefficient (DSC) to FreeSurfer, a commonly used brain segmentation software with a DSC of 81.90%, the 3D DCNN showed an improvement of 10.1% (92.00%). A 3D U-Net based architecture was presented in [11] for infant brain segmentation that includes DenseNet, ResNet, and a dimensionality reduction module. The architecture presents densely connected blocks in the encoder path and dimensionality reduction ResNet, as well as upsampling layers in the decoder path. The proposed method outperformed the 3D U-Net architecture as compared to other recent methods, with an average DSC of 92.77% compared to 91.59% of the 3D-U-Net, a 1.18% improvement.

A whole-brain segmentation brain imaging framework based on a neural network model that stimulates local and global competitiveness in the dense block and skip-connections, is introduced in [12]. This framework, named FastSurferCNN, incorporates efficient spatial information processing and dense blocks to enhance data retrieval and increase network connectivity. FastSurferCNN demonstrated a performance in the cortical and subcortical DSCs of 84.55% and 88.74%, respectively, indicating a slight improvement compared to the performance of 3D U-Net, with cortical and subcortical DSCs of 83.90% and 87.27%, respectively. A split-attention U-Net (SAU-Net) was introduced in [13]. This architecture was inspired by ResNet and U-Net++, where the class distinction between feature maps of the layer is enhanced by performing the split-attention module of ResNet to the output of the U-Net++ model and evolving normalization and rectified linear unit (ReLU). With this method, a small number of three-dimensional training batches is possible. This approach increased the mean DSC from 89.20% (U-Net), 88.40% (U-Net++), and 88.40% (FastSurfer), to 89.70% when compared to other methods.

One approach to increase DSC accuracy is to solve problems with discontinuous and ill-

defined boundaries of the structures, low contrast, heterogeneity of voxels in the tissue region associated with changes in the magnetic field and sensitivity during resonance scans magnetic, and low signal-to-noise ratio. This has the consequence of increasing intraclass heterogeneity while decreasing interclass distance. Additionally, the variability within patients (age, sex, ethnicity) present structures with varying shapes and sizes, resulting in repetitive and noisy features that significantly reduce efficiency. It is evident that selecting substantial features is a complex task [20]. A further approach is to use random feature kernels as initial parameters for convolution [14]. This approach is not hierarchical to adapting to the layered CNN model since the kernels are selected from the normal distribution. Invariant moments are a well-known tool for shape descriptors that can reflect details in both a global and finer form of an image. To obtain hierarchical function learning, invariant moments have a pyramidical architecture, from which the initial filter kernel parameters can be extracted by varying the moment orders as applicable to the different levels of CNN [14].

The challenge of segmenting brain becomes increasingly difficult when diverse scanners are used with different parameters during acquisition. It typically results in intensity heterogeneity, contrast fluctuations, and various forms of noise. Despite the importance of extracting the entire brain region to remove extraneous information in applications such as brain tumor or lesion segmentation, few studies have addressed preprocessing to enhance whole brain segmentation. Furthermore, unlike in CT, MRI are not calibrated in Hounsfield units and lacks a clear physical interpretation based on the absorption of X-rays in skull bones [15].

In [16], two distinct preprocessing pipelines were investigated to evaluate how they influenced deep CNN performance. The idea is to apply a basic standardization pipeline first and then compare its performance with a second registration-based pipeline. To make the network segmentation brain tissue classes robust to intensity inhomogeneity artifacts, [17] trained a network using slices containing simulated intensity inhomogeneity artifacts. The artifacts were created by combining linear gradients with random offsets and orientations on a slice without intensity inhomogeneity artifacts. Furthermore, [17] used data augmentation as input,

in addition to the given multi-modal images. The augmentation was accomplished by removing a gaussian smoothed version and performing histogram equalization using the contrast-limited adaptative histogram equalization to improve local contrast. Thereafter, the network is trained using both the generated and original volumes.Nevertheless, the real effect of this type of augmentation was not reported in their research. To enhance the diagnosis and the classification accuracy rate, [19] developed an adaptive bilateral filtering preprocessing method for removing undesired noise and image resizing while maintaining the threshold information in an MR image. While these studies demonstrate the influence of preprocessing in certain domains, the findings may not be generalizable. As there is no consensus on which of these reported preprocessing methods should be applied in medical image analysis, preprocessing must be investigated individually for each application.

Among preprocessing methods, orthogonal moments are mathematical methods that can be used to acquire important information on an object. Orthogonal moments unique properties, including orthogonality, geometric transformation invariance, and noise robustness, making them particularly useful for dealing with image noise and geometric distortions which may be benefical for the CNN since CNN are not scale and rotation invariant. Orthogonal moments are found to be less susceptible to noise and have an effective function representation capability. In addition, it presents the unique combination of low information redundancy and high discriminative capacity that makes them widely used due to minimal redundancy. Therefore, they can reflect the properties of an image with no duplication of information or redundancy across the moments.

In a variety of medical image segmentation algorithms, moment invariant features of pixels have been used as part of the feature set for training data. To assess the form of the brain structures at various scales, [20] used geometric moments as input features and the signed distance function of the structure as outputs of an artificial neural networks (ANNs). The outputs of the ANNs are then combined with the outputs of another ANN to classify the voxels. This method outperformed 4 (out of 5) approaches, [21–24], with a DSC of 88% against 82%,

78%, 85%, and 0.79%, respectively. The fifth method, evaluated its approaches on its own datasets, making the possibility of comparison difficult.

A new approach for blood vessel segmentation is presented in [25]. To determine whether pixels correspond to a real blood vessel, the feature descriptors are created using gray-level and moment invariant-based features. When accuracy is compared between 4 supervised techniques, [26–29], this method obtained an accuracy of 0.9452% against 0.9441%, 0.9417%, 0.9466%, and 0.9595% respectively, for one database. Additionally, when trained on one dataset and tested on a different databases this method outperformed [29] with 0.9448% against 0.9266% and 0.9526% against 0.9452%. The research in [30] introduce a powerful pre-processing technique to refine the blood vessels by using high-order moments to accurately differentiate thinner blood vessels from the background to create an effective training set. The technique computes an 11-D feature vector that includes both statistical and shape-based information, using Zernike moments, to perform pixel-wise classification. This approach outperformed 3 and 4 (out of 5) approaches analyzed on two datasets, with precision of 94.50% and 94.35%, respectively.

In CNN's, the initial convolutional layers learn how to represent low-level features such as edges and corners, whereas the subsequent layers learn how to represent higher-level features for the specific problem. By feeding the base-level features with considerably sophisticated information acquired by an initial filter, the obtained knowledge allows the model to converge earlier and improve the classification success of the new model over the classical model which may ultimately improve the diagnosis, evaluation, and treatment of the neurological diseases. Therefore, given the necessity for a model with initialization for transfer learning, this paper explores pre-processing convolutional neural networks by orthogonal moments, which removes the necessity for a numerical approximation and satisfies the orthogonality property. Three moments among the orthogonal moments are used: Legendre, Tchebichef, and Pseudo-Zernike moments. For each of these moments, datasets are generated from the original datasets. These datasets are then presented to a convolutional neural network to assess the performance of each

method. By improving the dataset before presenting the image to the CNN, we expect to see an improvement in DSC accuracy. Figure 3.1 presents the schematic diagram of the method proposed.
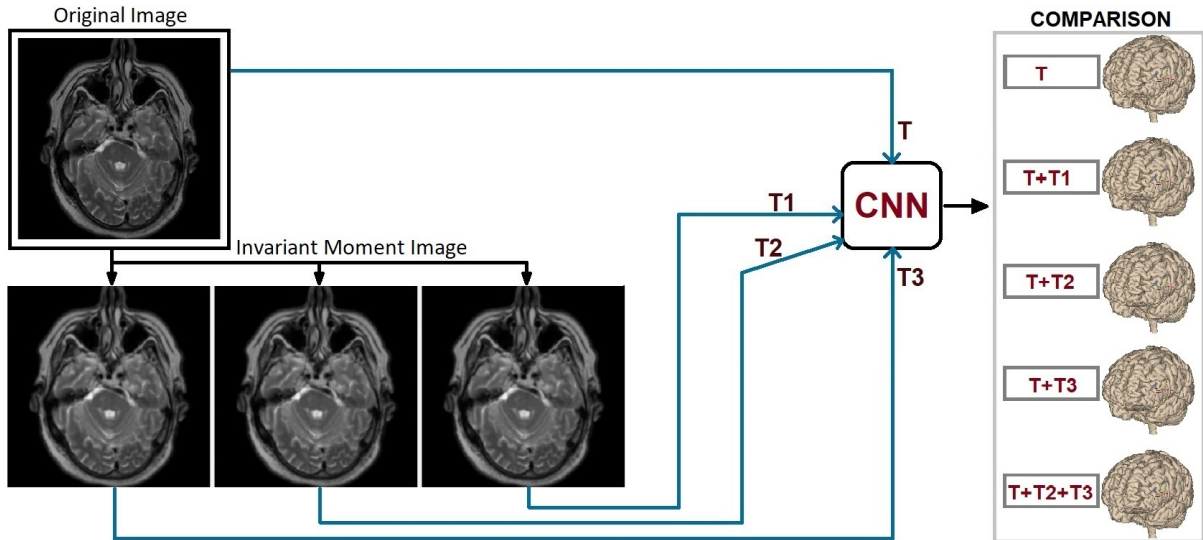
Figure 3.1: Overview of the proposed method. From the original image, 3 new image sets are created using Legendre, Tchebichef, and Pseudo-Zernike moments of first order. All image datasets (including the original image) are then presented to a convolutional neural network and, a series of statistical analysis are performed.

3.2 Materials and Methods

3.2.1 Dataset

To validate our approach, three distinct datasets were chosen:

- The Neurofeedback Skull-stripped (NFBS) repository [31] - a 125 T1-weighted MRI scan database with size of 512×512×192;

- The Open Access Series of Imaging Studies (OASIS-1) [32] - contains T1-weighted MRI scan cross-sectional collection of 416 subjects aged 18 to 96 with size of 256×256×176;

- The cancer imaging archive data collections (TCIA) [33] - this dataset contains T1, T2,

and FLAIR MRI scan modalities. We used 62 volumetric images that were explored by [34] with size of 512×512 and variable depth.

All images were normalized between the range of 0 to 1 and, to reduce the computational and processing time, all images were reduced to 256×256 using nearest neighbor interpolation method.

3.2.2 Orthogonal Moments

Moments are a projection of an image's intensity onto a suitable polynomial basis employed in a variety of image processing applications. In [35], the authors define invariant moments as features of an image that remain unchangedwhen scaling, translation, or rotation is introduced.

Orthogonal moments are moments obtained using orthogonal polynomials such as Legendre, Tchebichef, Pseudo-Zernike, and Zernike. As the orthogonal moments minimize the amount of information redundancy through approximation, they have been applied as a method of reducing image moment calculation complexity.

3.2.2.1 Legendre Moments

The construction of the Legendre moments (LMs) are performed through the calculation of the Legendre polynomial basis of order p and repetition q [36]

$$P_p(x) = \frac{(2p-1)xP_{p-1}(x) - (p-1)P_{p-2}(x)}{p}, \tag{3.1}$$

where p is a non-negative integer, $0 \leq |q| \leq p$, $P_0(x) = 1$, $P_1(x) = x$, and $p ¿ 1$. The Legendre polynomial set are defined on the interval between $[-1,1]$. The two-dimensional discrete Legendre moments of order $(p+q)$ are expressed as

$$L_{pq} = \lambda_{pq} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P_p(x_i)P_q(y_j)I_{ij}, \tag{3.2}$$

where I_{ij} is the discretized image intensity function with $N{\times}N$ size, λ_{pq} is a normalizing constant defined as

$$\lambda_{pq} = \frac{(2p + 1)(2q + 1)}{N^2},$$
(3.3)

and x_i and y_j are the normalized pixel coordinates from *[-1,1]* given as

$$x_i = \frac{2i}{N - 1} - 1, \quad and \quad y_j = \frac{2j}{N - 1} - 1.$$
(3.4)

The direct computation of LMs requires approximately $O(M^2N^2)$ additions and multiplications for an image of NxN pixels size and where M is the maximum order of LMs to be calculated [37]. The pseudo code for computing LMs can be seen in the algorithm 1.

Algorithm 1 Legendre Moments Pseudo Code

Function *LegendreMoments(p_{max}, q_{max})*
for $p = 1$ *to* p_{max} **do**
    for $i = 1$ *to* N **do**
        $x \leftarrow \frac{2i}{N} - 1$
        $P(0, i) \leftarrow 1$
        $P(1, i) \leftarrow x$
        $P(p, i) \leftarrow \frac{(2p-1)xP(p-1,i)-(p-1)P(p-2,i)}{p}$
    end
end
for $p = 0$ *to* p_{max} **do**
    for $q = 0$ *to* q_{max} **do**
        $sum \leftarrow \sum_i^{N-1} \sum_j^{N-1} P(p, i)P(q, j)I(i, j)$
        $Legendre(p, q) \leftarrow \frac{(2p+1)(2q+1)}{N^2} \cdot sum$
    end
end
return *Legendre*

3.2.2.2   Tchebichef Moments

The discrete Tchebichef moments (TMs) of order $(p+q)$ are constructed under the Tchebichef polynomials defined by [38]. The scaled Tchebichef polynomials were introduced in [39] as

$$t_p(x) = \frac{(2p-1)t_1(x)t_{p-1}(x) - (p-1)(1 - \frac{(p-1)^2}{N^2})t_{p-2}(x)}{p},$$ (3.5)

where $p = 2,3,...,N-1$, $t_0(x) = 1$, and $t_1(x) = \frac{2x+1-N}{N}$. The squared-norm of the scaled polynomials is then modified accordingly to

$$\rho(p,N) = \frac{N(1 - \frac{1}{N^2})N(1 - \frac{2^2}{N^2})...N(1 - \frac{p^2}{N^2})}{2p+1},$$ (3.6)

where $p = 0,1,...,N-1$. Finally, Tchebichef moments can be constructed as

$$T_{pq} = \frac{1}{\rho(p,N)\rho(q,N)} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} t_p(x_i)t_q(y_j)I_{ij},$$ (3.7)

where $i, j = 0,1,...,N-1$.

For a $N{\times}M$ image, the complete computation for one 2D TMs of order $(p+q)$ is $(3M-4)N$ multiplications and $(2M-3)N$ additions [40]. TMs pseudo code is shown in algorithm 2.

3.2.2.3   Pseudo-Zernike Moments

Pseudo-Zernike Moments (PZMs) [41] are a derivation of the Zernike Moments first introduced by [14]. The first step to obtain the PZMs of order p with repetition q of an image, is to calculate the Pseudo-Zernike polynomials (PZP) $V_{pq}(x_i, y_j)$ defined as

$$V_{pq}(x_i,y_j) = R_{pq}(r)e^{jq\theta},$$ (3.8)

Algorithm 2 Tchebichef Moments Pseudo Code

Function *TchebichefMoments(p_{max}, q_{max})*

$t_{(0,x)} \leftarrow 1$

$t_{(1,x)} \leftarrow \frac{2x+1-N}{N}$

for *p = 2 to p_{max}* **do**

    for *x = 1 to N* **do**

        $t(p, x) \leftarrow \dfrac{(2p-1)t_1(x)t_{p-1}(x)-(p-1)(1-\frac{(p-1)^2}{N^2})t_{p-2}(x)}{p}$

    end

end

$R \leftarrow 1$

for *p = 0 to p_{max}* **do**

    $r \leftarrow 1 - \frac{(p+1)^2}{N^2}$

    $R \leftarrow R \cdot r$

    $\rho(p + 1) \leftarrow \frac{N \cdot R}{2p+1}$

end

for *p = 0 to p_{max}* **do**

    for *q = 0 to q_{max}* **do**

        $sum \leftarrow \sum_{i}^{N-1} \sum_{j}^{N-1} t(p, i)t(q, j)I(i, j)$

        $Tchebichef(p, q) \leftarrow \frac{1}{\rho(p) \cdot \rho(q)} \cdot sum$

    end

end

return *Tchebichef*

---

where $\theta = \tan^{-1}(y/x)$, $\theta \in [0,2\pi]$, and $r = \sqrt{x_i^2 + y_j^2}$. $V_{pq}(x_i, y_j)$ is a complex function with radial polynomials $R_{pq}(r)$ and angular functions $e^{jq\theta} = (\cos\theta + j\sin\theta)^q$ polynomials in $\cos\theta$ and $\sin\theta$. Radial polynomials are defined as

$$R_{pq}(r) = \sum_{s=0}^{p-|q|}(-1)^s \frac{(2p+1-s)!r^{(p-s)}}{s!(p+|q|+1-s)!(p-|q|-s)!}. \tag{3.9}$$

For $R_{pq}(r) = R_{p,-q}(r)$, we can consider $q \geq 0$ and rewrite Equation(5.9) as

$$R_{pq}(r) = \sum_{s=q}^{p} B_{pqs} r^s, \tag{3.10}$$

where

$$B_{pqs} = \frac{(-1)^{p-s}(p+s+1)!}{(p-s)!(s+q+1)!(s-q)!}. \tag{3.11}$$

Focusing on reducing time complexity by using a hybrid approach to compute the Pseudo-Zernike radial polynomials, [30] introduced a *p-recursive* method which is defined as

$$R_{qq}(r) = r^q, \tag{3.12}$$

$$R_{(q+1)q}(r) = (2q+1)r^{q+1} - 2(q+1)r^q, \tag{3.13}$$

$$R_{pq}(r) = (k_1 r + k_2)R_{(p-1)q}(r) + k_3 R_{(p-2)q}(r), \tag{3.14}$$
$$p = q+2, q+3, ..., p_{max},$$

where

$$k_1 = \frac{2p(2p+1)}{(p+q+1)(p-q)}, \tag{3.15}$$

$$k_2 = -2p + \frac{(p+q)(p-q-1)}{(2p-1)}k_1, \tag{3.16}$$

$$k_3 = (2p-1)(p-1) - \frac{(p+q-1)(p-q-2)}{2}k_1 + 2(p-1)k_2. \tag{3.17}$$

Finally, PZMs of order p and repetition q of an image function I_{ij} over a unit disc in a discrete domain can be calculated by

$$A_{pq} = \frac{p+1}{\pi} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} I(x_i, y_j)V_{pq}^*(x_i, y_j)\Delta x_i \Delta y_j, \tag{3.18}$$

where V_{pq}^* are the complex conjugate of the PZP (V_{pq}) in Equation(5.8), $x_i = -\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{N-1}x$, $y_j = \frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{N-1}y$, $x = y = 0, 1, ..., N\text{-}1$, x_i, and y_j are the image coordinates, and I_{ij} is defined over the discrete square domain $N{\times}N$.

The computational complexity of Pseudo-Zernike moments is $O(p^2)$, as shown by [44], where the number of required multiplications for computing the radial polynomial R_{pq} is approximately $(p+1)(4p)(p\text{-}1)$, where p is the moment order.

The pseudo code for PZMs is described in algorithm 3.

3.2.3 Moment Kernel

Initially, the volumetric images acquired from datasets 1, 2, and 3 are normalized between the range of 0 to 1. Calculating higher order moments are computationally expensive. Obtaining approximations of low-order moments are less cumbersome, thus, to avoid computational complexity during moment estimation, only the first order (order-0, repetition-0) LM, TM, and PZM , which represents the center of mass of an image, is calculated.

To generate the image moment, first a 3×3 window size is extracted from the original image and the first-order moment is calculated using equations (5.1)-(3.18) or algorithms 1, 2, and 3. A 3×3 window size was chosen to reduce computational time. The result of the first-order

---

Algorithm 3 PseudoZernike Moments Pseudo Code

Function *Pseudo-ZernikePolynomials(s,p,q)*
$PZP \leftarrow 0$
for $r = 0$ to $\frac{p-q}{2}$ **do**
$\quad V \leftarrow (-1)^s \frac{(2p+1-s)!}{s!(p+|q|+1-s)!(p-|q|-s)!}$
$\quad PZP \leftarrow PZP + V * r^{(p-s)}$
end
return *PZP*
Function *PseudoZernikeMoments(p,q)*
$PZM \leftarrow 0$
$cnt \leftarrow 0$
for $y = 0$ to *N-1* **do**
$\quad$ **for** $x = 0$ to *N-1* **do**
$\quad\quad a \leftarrow \frac{-\sqrt{2}}{2} + \frac{\sqrt{2}}{N-1}x$
$\quad\quad b \leftarrow \frac{-\sqrt{2}}{2} - \frac{\sqrt{2}}{N-1}y$
$\quad\quad r \leftarrow \sqrt{a^2 + b^2}$
$\quad\quad$ **if** $r \leq 1$ **then**
$\quad\quad\quad PZP \leftarrow PseudoZernikePolynomial(s, p, q)$
$\quad\quad\quad \theta \leftarrow tan^{-1}(\frac{b}{a})$
$\quad\quad\quad PZM \leftarrow PZM + I(x, y) \cdot PZP \cdot e^{iq\theta}$
$\quad\quad\quad cnt \leftarrow cnt + 1$
$\quad\quad$ **end**
$\quad$ **end**
end
return $\frac{p+1}{cnt} \cdot PZM$

---

moment calculated is a single pixel in the moment image. This moment calculation is repeated from the center pixel of the window to each sliding window. These pixels generate a moment image that represents the first order moment extracted from each pixel of the original image by a 3×3 window size. When all slides of the original volumetric image are calculated, a volumetric moment image is created, as shown in Figure 3.2. Figure 3.3 shows the images created from a single original image.
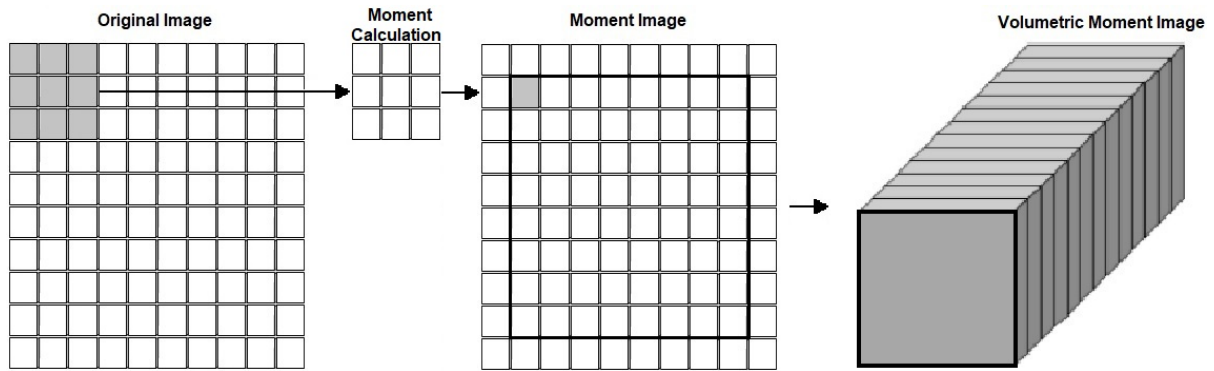
Figure 3.2: From each pixel of the original image, a 3×3 image size is used to create the moment representation. This produces a moment image for each pixel and, subsequently, a volumetric moment image.
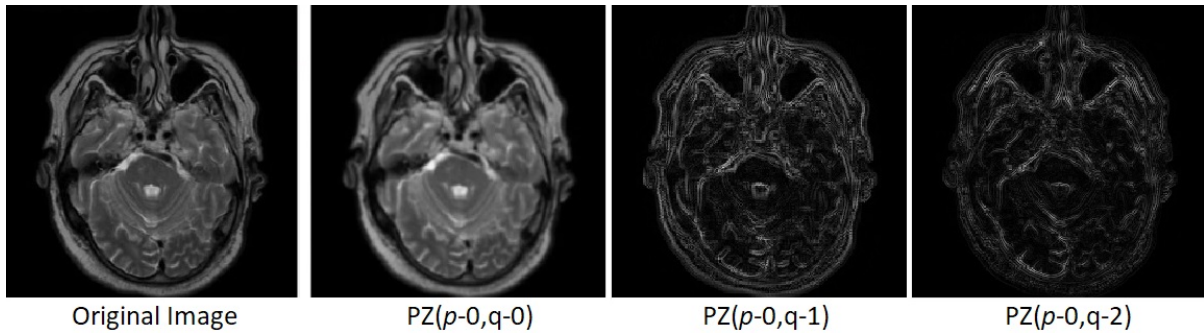
Figure 3.3: Original image and Pseudo-Zernike image (PZ) of order 0 and repetition 0 to 2 created from the TCIA dataset.

3.2.4 CNN Framework and Implementation Details

The CNN framework adopted in the first part of this work was U-Net [11]. U-Net works well with a small training set, produces precise segmentation, and has been applied in a variety of

biomedical image segmentation applications [13, 34, 45–48]. The key concept is to add layers to a traditional contracting network. In the upsampling procedure, this network allows for a large number of feature channels, which facilitate in the transmission of context information to the highest resolution layers. As a result, these layers may improve the output resolution. Furthermore, based on this knowledge, a subsequent convolutional layer will learn to assemble a precise output. Consequently, the network produces a more symmetrical expansive direction and a U-shaped architecture.

The framework adopted was initially implemented for brain tumor segmentation in MRI [16], with a MATLAB (The MathWorks, Natick, MA) version found in [50]. The model uses a patch size of 128×128 with hyper-parameters illustrated in Figure 3.4 where, CONV-convolution (5×5×5), BN-batch normalization, RELU-rectified linear units, CH-channels, MP-max pooling, DECON-transposed convolution (2×2×2), CONC-concatenation, and GDL-generalized dice loss. In total, the framework has 55 layers. Table 5.1 presents the parameters selected for the whole brain segmentation in all datasets. These parameters were chosen to improve accuracy for the training set used in this study, prevent computational error/crash, and enhance the ability and robustness of the U-Net.
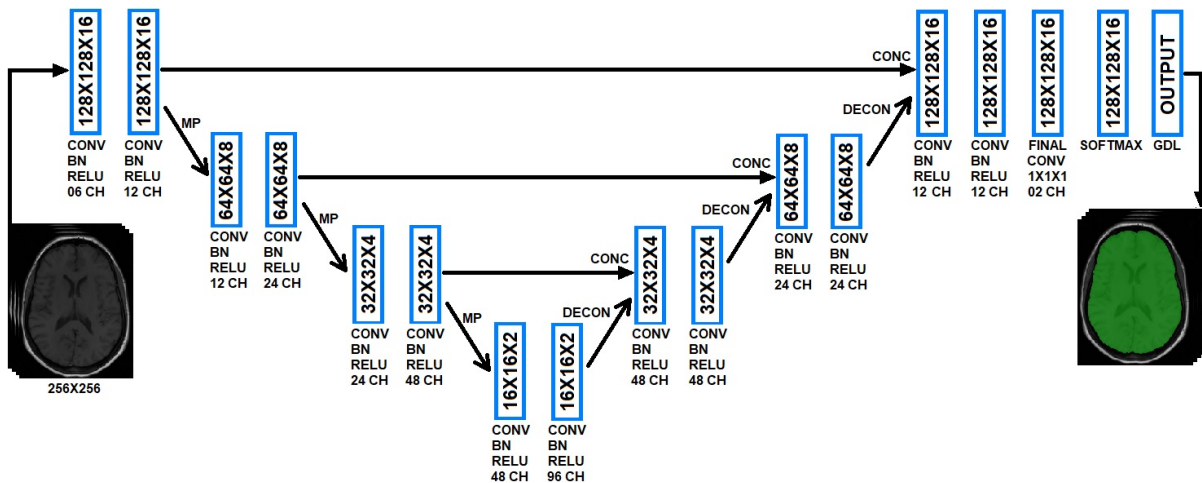
Figure 3.4: Schematic representation of the deep learning architecture used based on 3D brain tumor segmentation U-net.

The U-Net was performed on Intel i7-9700 (3.00 GHz) processor with 64 GB of ram, and

Table 3.1: Whole Brain Segmentation CNN Implementation Details.

Parameter	Value
Optimizer	Adam
Encoder Depth	3
Filter Size	5
Number of First Encoder Filters	6
Patch Per Image	1
Mini Batch Size	128
Initial Learning Rate	$1 \cdot 10^{-3}$
Epochs NFBS	125
Epochs OASIS	35
Epochs TCIA	200

two 8GB VRAM graphic cards from NVIDIA (RTX 2070 SUPER and RTX 2080). Part of the code was implemented in MATLAB R2019B.

3.2.5 Statistical Modeling

Dice Similarity Coefficient (DSC) [20], Symmetric Volume Difference (SVD) [21], Jaccard Similarity Coefficient (JSC) [53], Volumetric Overlap Error (VOE) [54], and Hausdorff distances (HD) [55] methods, were used to evaluate the CNN segmentation. DSC is a spatial overlap index between two distinct sets. It ranges from 0, to indicate no spatial overlap, to a maximum 1, that indicates complete overlap. Equation (3.19) presents the definition of DSC.

$$DSC = \frac{2TP}{2TP + FP + FN}.$$

(3.19)

where the TP is the number of true positives, FP false positives, and FN false negatives.

In terms of DSC-based error metric, SVD gives the symmetric difference of the structure and segmentation (equation (3.20)).

$$SVD = 1 - DSC.$$

(3.20)

JSC, in equation (3.21), is a similarity ratio that represents the intersection of the ground truth and the system segmentation regions over their union. It ranges from 0% to 100% of similarity.

$$JSC = \frac{TP}{TP + FP + FN},$$
(3.21)

VOE is the JSC correspondent error metric (equation (3.22)).

$$VOE = 1 - JSC.$$
(3.22)

Finally, Hausdorff distances uses Euclidean distance to calculate the segmentation precision in terms of distance between the predicted segmentation boundary and the ground truth (equation (3.23)).

$$HD(X, Y) = max(\max_{x \in X} \min_{y \in Y} \|x - y\|_2, \max_{y \in Y} \min_{x \in X} \|x - y\|_2).$$
(3.23)

3.3   Results and Discussion

3.3.1   Performance on Datasets

The NFBS dataset was randomly divided into 115 images for training and 10 images for validation and testing. These images were then trained and tested 10 times. To analyze the initial segmentation performances, the DSC metric was chosen due to its simple validation metric of reproducibility and partial overlap accuracy. Table 5.2 presents DSCs and their respective standard deviations (SD) from the test set of CNN's predicted segmentations, where, Original Image (O), Legendre Moment Image (L), Tchebichef Moment Image (T), and Pseudo-Zernike Moment Image (PZ). As O + T and O + PZ performed well, we tested a combination of O + T + PZ which obtained the best mean DSC. The respective DSCs, SVDs, JSCs, VOEs, and HDs values are shown in Table 5.3.

The OASIS dataset was randomly divided into 396 images for training and 20 images for validation and testing. The images were trained and tested 10 times. Table 5.2 presents DSCs and their respective standard deviations (SD) from the test set of CNN's predicted segmentations. As O + T and O + PZ performed well, we tested a combination of O + T + PZ. Similarly, O + T + PZ images obtained the best mean DSC. Their respective DSCs, SVDs, JSCs, VOEs, and HDs are presented in Table 5.3.

The TCIA dataset was randomly divided into 56 images for training and 6 images for validation and testing. The images were trained and tested 10 times. Table 5.2 presents DSCs and their respective standard deviations (SD) from the test set of CNN's predicted segmentations. As O + T and O + PZ performed well, we tested a combination of O + T + PZ. Again, O + T + PZ images acquired the best mean DSC. Their respective DSCs, SVDs, JSCs, VOEs, and HDs are presented in Table 5.3.

Table 3.2: Mean DSCs - Statistical Analysis on the NFBS, TCIA, and OASIS datasets.

	O	O+L	O+T	O+PZ	O+T+PZ
NFBS	0.8983 ± 0.02	0.7632 ± 0.08	0.9212 ± 0.02	0.9156 ± 0.02	0.9395 ± 0.01
OASIS	0.9370 ± 0.01	0.9066 ± 0.02	0.9547 ± 0.002	0.9358 ± 0.008	0.9561 ± 0.004
TCIA	0.9044 ± 0.01	0.8751 ± 0.02	0.9117 ± 0.002	0.9104 ± 0.008	0.9149 ± 0.004

Table 3.3: O and O + T + PZ - Statistical Analysis on the NFBS, OASIS, and TCIA datasets.

		DSC (Brain)	DSC (Background)	SVD (Brain)	JSC (Brain)	VOE (Brain)	HD (Brain)
NFBS	O	0.8983 ± 0.02	0.9844 ± 0.004	0.1017 ± 0.02	0.8173 ± 0.03	0.1827 ± 0.03	8.89 ± 1.62
	O+T+PZ	0.9395 ± 0.01	0.9914 ± 0.001	0.0605 ± 0.01	0.8864 ± 0.02	0.1136 ± 0.02	6.98 ± 0.85
OASIS	O	0.9370 ± 0.01	0.9771 ± 0.003	0.0630 ± 0.01	0.8820 ± 0.02	0.1180 ± 0.02	15.50 ± 2.35
	O+T+PZ	0.9561 ± 0.004	0.9835 ± 0.001	0.0439 ± 0.004	0.9161 ± 0.007	0.0839 ± 0.007	10.91 ± 0.86
TCIA	O	0.9044 ± 0.01	0.9797 ± 0.003	0.0956 ± 0.01	0.8281 ± 0.02	0.1719 ± 0.02	19.24 ± 2.22
	O+T+PZ	0.9149 ± 0.04	0.9827 ± 0.003	0.0851 ± 0.02	0.8448 ± 0.03	0.1552 ± 0.03	17.24 ± 2.60

For the TCIA dataset, Figure 3.5a illustrates how each set converges in 200 epochs, whereas Figure 3.5b displays the first 20 epochs. All approaches converged faster than O images, as can

be observed in Figure 3.5b. The best convergences are achieved by O+T and O+PZ, but after 10 epochs, they are outperformed by O+T+PZ, which continued to achieve the best accuracies for the remaining epochs.
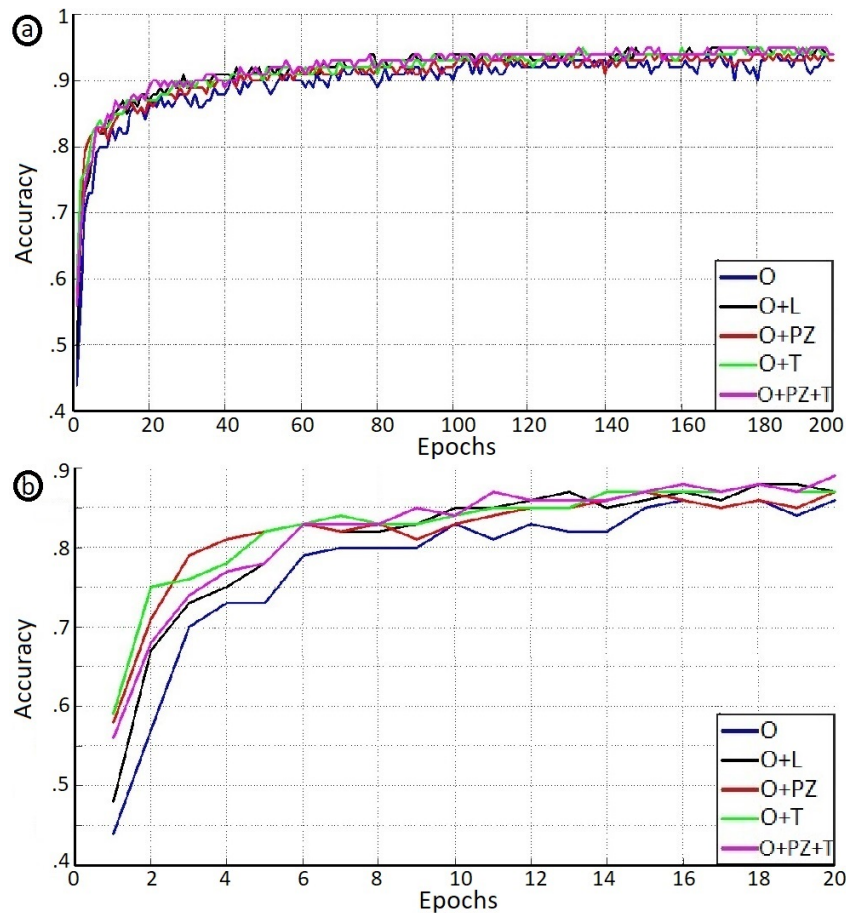
Figure 3.5: a) shows the accuracy×epochs for dataset TCIA in 200 epochs and b) presents an enlarged version for the first 20 epochs.

For LM, the computation is in general a time-consuming process. To reduce the processing, their equation is solved by closed form representations for orthogonal polynomials, and taking less care to the precision of the quadrature formulas used to approximate integrals [56]. This equation discretization may affect their image discretization process which may cause numerical errors in the computed moments, thus affecting its segmentation performance.

From Table 5.3, there was an improvement in all statistical modeling measure components. In fact, the DSCs from NFBS, OASIS, and TCIA datasets improved 4.12%, 1.91%, and

1.05% respectively. The conventional input does not provide full features to the U-Net first layer which compromise the segmentation performance due to a lack of local details whereas our proposed pre-processing method shows performance improvements. Furthermore, the differences in the performance of the dataset may be due to the dataset size and different MRI modalities such as T1, T2, and FLAIR. Factors such as multiple modalities, small dataset, variations in medical history, ethnicity, and age, make the segmentation task difficult. Moreover, this dataset is part of the cancer imaging archive and contains a variety of patients with undergoing cephalic surgery, which may change their brain structure.

As the JSC measures the similarity and diversity of sample sets, its value gives an idea of how similar the O or O + T + PZ are to the ground truth. From Table 5.3, an improvement of 6.91% for NFBS, 3.41% for OASIS, and 1.67% for TCIA were performed.

For the spatial distance based metrics HD, a 1.91%, 4.59%, and 2% reduction was observed. As this metric gives a boundary delineation information of the contour, these results show that the addition of T+PZ images provide a shape information significantly better than the single O image, which contributed to the improvement of the edge information. Furthermore, HD is generally sensitive to noise and outliers [57], which are common in medical segmentation. [58] addresses outliers by using the Hausdorff quantile method. Therefore, applying an orthogonal moments pre-processing improves the noise redundances and consequently improve edge information.

3.3.2 Inter-Dataset Analysis

An essential technique to be addressed is the transfer learning or the transferability of the trained network when processing images from different databases with an emphasis on validating the stability and robustness. To perform this experimentation, TCIA dataset was chosen to train the CNN since it comprises diverse MRI scan modalities (T1, T2, and FLAIR). Therefore, 62 volumetric images were used for training and, from these images, 6 new sets of 62 volumetric moment images were generated using Tchebichef and Pseudo-Zernike moments.

These 6 sets represents the following moment order p and repetition q - (p,q); 1-(0,0), 2-(0,1), 3-(0,2), 4-(0,3), 5-(1,0), and 6-(1,1). The identical U-Net described inthe subsection 3.2.4 and detailed in Table 5.1 was adopted.

Tables 5.4 and 5.5 illustrate the results achieved after training, using TCIA dataset, and testing 10 times on 85 images from the NFBS and OASIS datasets. When trained with order and repetition (p,q)-(0,1) and (1,1), the CNN had the best performance. As a result, an additional scenario (O+T7+PZ7 = O+T2+T6+PZ2+PZ6), was investigated.

Table 3.4: Inter-Dataset Statistical Analysis on the NFBS dataset.

	DSC (Brain)	DSC (Background)	SVD (Brain)	JSC (Brain)	VOE (Brain)	HD (Brain)
O	0.5157 ± 0.09	0.8088 ± 0.13	0.4843 ± 0.09	0.3576 ± 0.08	0.6424 ± 0.08	40.37 ± 13.1
O+T1+PZ1	0.5651 ± 0.08	0.8789 ± 0.09	0.4349 ± 0.08	0.4009 ± 0.07	0.5991 ± 0.07	36.05 ± 8.18
O+T2+PZ2	0.5809 ± 0.05	0.8921 ± 0.05	0.4191 ± 0.05	0.4149 ± 0.05	0.5851 ± 0.05	35.96 ± 5.75
O+T3+PZ3	0.5313 ± 0.04	0.8830 ± 0.04	0.4687 ± 0.04	0.3653 ± 0.04	0.6347 ± 0.04	32.45 ± 5.00
O+T4+PZ4	0.5534 ± 0.05	0.8985 ± 0.03	0.4466 ± 0.05	0.3844 ± 0.05	0.6156 ± 0.05	30.37 ± 6.23
O+T5+PZ5	0.5752 ± 0.06	0.9024 ± 0.04	0.4248 ± 0.06	0.4085 ± 0.06	0.5915 ± 0.06	29.35 ± 7.07
O+T6+PZ6	0.5881 ± 0.07	0.9028 ± 0.05	0.4119 ± 0.07	0.4215 ± 0.06	0.5785 ± 0.06	28.14 ± 7.61
O+T7+PZ7	0.6144 ± 0.05	0.9262 ± 0.02	0.3856 ± 0.05	0.4465 ± 0.05	0.5535 ± 0.05	23.89 ± 4.01

Table 3.5: Inter-Dataset Statistical Analysis on the OASIS dataset.

	DSC (Brain)	DSC (Background)	SVD (Brain)	JSC (Brain)	VOE (Brain)	HD (Brain)
O	0.6658 ± 0.07	0.7420 ± 0.15	0.3342 ± 0.07	0.5041 ± 0.08	0.4959 ± 0.08	52.23 ± 12.05
O+T1+PZ1	0.6958 ± 0.07	0.8057 ± 0.12	0.3042 ± 0.07	0.5379 ± 0.08	0.4621 ± 0.08	51.08 ± 9.88
O+T2+PZ2	0.7166 ± 0.05	0.8318 ± 0.07	0.2834 ± 0.05	0.5617 ± 0.06	0.4383 ± 0.06	50.47 ± 7.22
O+T3+PZ3	0.6649 ± 0.04	0.7884 ± 0.08	0.3351 ± 0.04	0.5002 ± 0.05	0.4998 ± 0.05	48.70 ± 6.79
O+T4+PZ4	0.6420 ± 0.05	0.7555 ± 0.09	0.3580 ± 0.05	0.4761 ± 0.05	0.5239 ± 0.05	50.98 ± 7.28
O+T5+PZ5	0.7162 ± 0.06	0.8398 ± 0.07	0.2838 ± 0.06	0.5612 ± 0.06	0.4388 ± 0.06	44.54 ± 7.98
O+T6+PZ6	0.7240 ± 0.05	0.8414 ± 0.07	0.2760 ± 0.05	0.5712 ± 0.06	0.4288 ± 0.06	45.42 ± 8.83
O+T7+PZ7	0.7434 ± 0.04	0.8725 ± 0.05	0.2566 ± 0.04	0.5944 ± 0.05	0.4056 ± 0.05	41.74 ± 6.40

According to Table 5.4, it can observed that an improvement of 7.24% in DSC accuracy occurred between O and O+T6+PZ6 with a 0.02 reduction in STD. Furthermore, when using the combination of O+T2+T6+PZ2+PZ6 (O+T7+PZ7), a greater improvement occurred of 9.87%

with a 0.04 reduction in STD. A similar behavior was found in Table 5.5 where O+T6+PZ6 improved over O by 5.82% with a 0.02 decrease in STD and O+T7+PZ7 improved by 7.76% with an STD reduction of 0.03. Not only DSC improved ,but all other metrics presented improvements and reductions when compared to O, as shown in Table 3.6.

Table 3.6: Comparison between O+T6+PZ6 and O+T7+PZ7 over O on the NFBS and OASIS datasets.

		DSC (Brain)	DSC (Background)	SVD (Brain)	JSC (Brain)	VOE (Brain)	HD (Brain)
NFBS	O+T6+PZ6	0.0723	0.0940	−0.0723	0.0639	−0.0639	−12.22
	O+T7+PZ7	0.0986	0.1174	−0.0986	0.0889	−0.0889	−16.48
OASIS	O+T6+PZ6	0.0582	0.0994	−0.0582	0.0671	−0.0671	−6.81
	O+T7+PZ7	0.0776	0.1305	−0.0776	0.0902	−0.0902	−10.49

Also, Figures 3.6 and 3.7 show the 85 images DSC accuracies of the NFBS and OASIS datasets, where the red line on each box represents the median, while the top and bottom blue margins of the box reflect the 75th and 25th percentiles, respectively. The bars extend to the most extreme data points not considered outliers, and the outliers are represented individually using the red cross marker symbol. In general, O+T2+PZ2 and O+T6+PZ6 had the best performed in terms of maximum and median accuracy when compared to O. The combination of both (O+T7+PZ7), outperformed all scenarios in both accuracy and median.

Orthogonal moments can be used to obtain pertinent information about an object. Their unique rotation, scaling, and translation properties for dealing with geometric distortions, allow them to represent the properties of an image with no duplication of information or redundancy across the moments. Additionally, other properties, such as noise robustness, orthogonality, low information redundancy, and high discriminative capacity makes them widely used for feature extraction.

This initialization process with their multilevel image structure are significant in providing both global and detailed shape characteristics of an image that make them suitable to be employed in CNN architecture. Further improvements can be implemented to address limitations
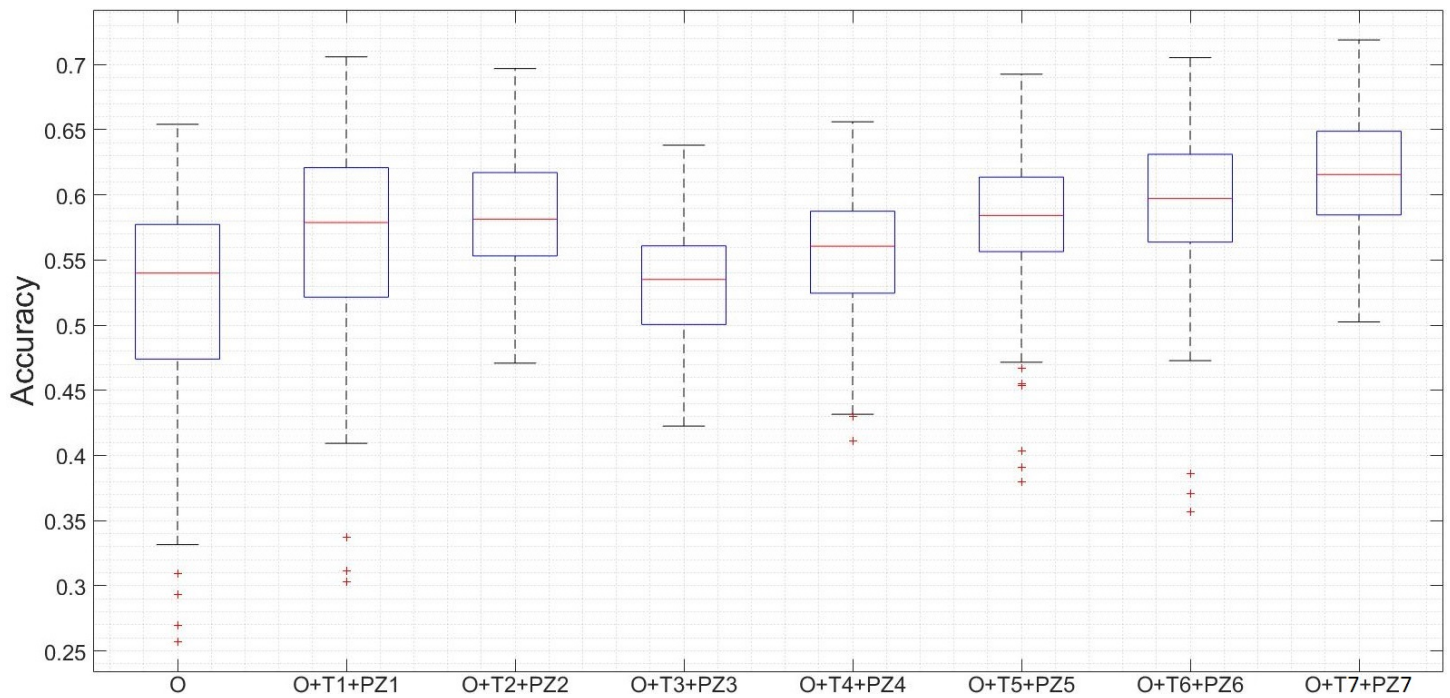
Figure 3.6: Boxchart of NFBS dataset.
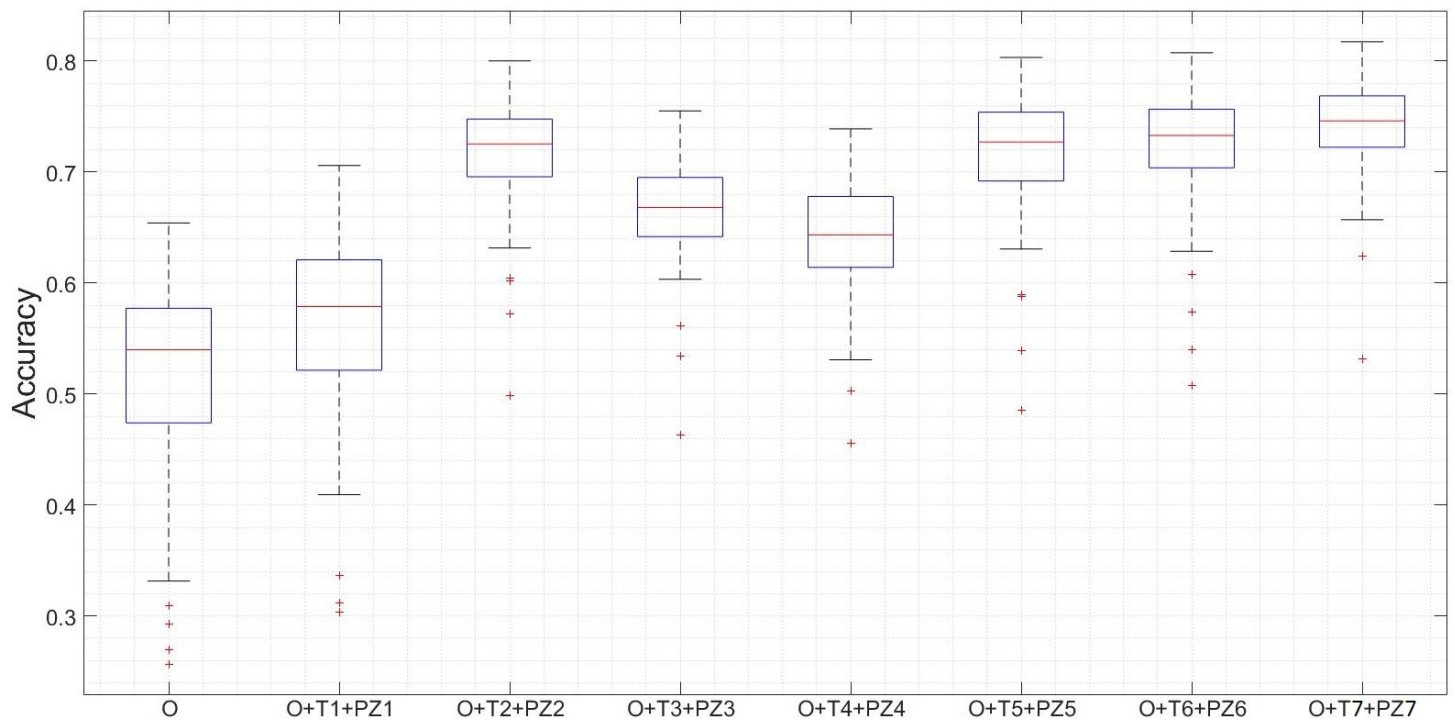
Figure 3.7: Boxchart of OASIS dataset.

such as application of 3D invariant moments, which can improve processing time and cover more image information as of its 3D kernel.

A future investigation will aim to segment distinct brain region using the approach introduced in this work. Orthogonal moments can provide image information regarding shape and robustness over scale, translation, rotational distortions, as well as noise, orthogonality, minimal information redundancy, and high discriminative capacity. By incorporating the information obtained from the orthogonal moments, it is expected to improve the CNN accuracy, based on the input information, by adding additional information acquired during the moment image pre-processing. Furthermore, the incorporation of orthogonal moments in the deep layers of the CNN is a future research also to be addressed.

To sum up, due to pre-processing transformation of the original image, that transformed the original image into a new version with greater discriminative capacity, the knowledge obtained by feeding the CNN with sophisticated information acquired by an orthogonal moment filter allowed the model to improve the classification success of the new model.

3.3.3   CNN Framework Comparisons

A final test is performed to study the differences between different frameworks when orthogonal moment images are applied. For this purpose, we included U-Net, U-Net++, and U-Net3+. U-Net++ is a U-Net based framework introduced by [59] with an inclusion of dense block and convolution layers between the encoder and decoder to improve segmentation accuracy. The theory behind this design is that when high-resolution feature maps from the encoder network are increasingly enriched prior to fusion with the corresponding semantically rich feature maps from the decoder network, the model can effectively capture fine-grained descriptions of foreground objects. [60] presents a CNN, named U-Net3+, to make full use of the multi-scale features by introducing full-scale skip connections, which combine low-level details with high-level semantics from feature maps in full sizes, but with less parameters. This approach optimizes a hybrid loss function to maximize the organ boundary by learning hierarchical representations from the full-scale aggregated feature maps. Table 3.7 compares U-Net, U-Net++, and U-Net3+ framework in terms of segmentation accuracy measured by DSC on datasets 1,2,

and 3. The parameters chosen for each CNN were identical with Adam optimizer, an encoder depth of 3, filter size of 5, mini batch size of 16, an initial learning rate of 0.005, and training was carried out in 6 epochs for NFBS and OASIS datasets, and 11 epochs for TCIA.

Table 3.7: Comparison between U-Net, U-Net++, and U-Net3+ in four samples.

		U-Net	U-Net++	U-Net3+
NFBS	O	0.9666 ± 0.006	0.8132 ± 0.034	0.9707 ± 0.009
	O+T+PZ	0.9777 ± 0.002	0.8168 ± 0.010	0.9801 ± 0.001
OASIS	O	0.9648 ± 0.032	0.5514 ± 0.023	0.9662 ± 0.005
	O+T+PZ	0.9766 ± 0.001	0.7269 ± 0.039	0.9764 ± 0.001
TCIA	O	0.9254 ± 0.022	0.7883 ± 0.027	0.9393 ± 0.006
	O+T+PZ	0.9500 ± 0.005	0.8285 ± 0.075	0.9579 ± 0.002

Table 3.7 shows that U-Net3+ outperformed U-Net and U-Net++ obtaining for O(O+T+PZ) an average DSC improvement of 0.41%(0.23%), 0.14%(-0.02%), and 1.38%(0.79%) for the NFBS, OASIS, and TCIA datasets respectively. The training time for the U-Net algorithm was half of the U-Net3+ algorithm, however, if training time is ignored, U-Net3+ can be applied to slightly improve performance.

U-Net++ improves segmentation outcomes by redesigning the dense skip connection to eliminate the gap semantic between encoder and decoder features. U-Net++ builds rough segmentation borders by using dense skip connections at various scales. Consequently, the number of parameters increases. In the literature, U-Net++ seems to perform better than U-Net, however, for certain datasets, U-Net++ performed worse than U-Net [61]. Although the causes are still being debated, we can assume that the proposed technique contributes to a decrease in the DSC generated in the dense layers of the U-Net++.

With a DSC improvement of 4.12%, 1.91%, and 1.05% compared to the original results, the findings in this study contribute to the long-standing search for the development of a pre-processing technique for whole brain segmentation in MRI. Using CNN and orthogonal moments to enhance DSC rates is a step forward in improving whole brain segmentation in MRI.

3.4   Conclusion

This study introduces a method to initialize CNN's using orthogonal moments for whole brain segmentation in MRI. This technique obtained a mean brain DSC improvement of 4.12%, 1.91%, and 1.05% for three different datasets. Other statistical analyses were performed where we can highlight the reduction in the HD values which demonstrates a decrease between the predicted and ground truth distances. A further inter-dataset analysis was conducted where Tchebichef and Pseudo-Zernike moments of various orders and repetitions (O+T6+PZ6), outperformed the regular input image (O) by an maximum improvement of 7.23% on the NFBS and 5.82% on the OASIS datasets, using TCIA dataset for training. Furthermore, the combination of O+T2+T6+PZ2+PZ6 achieved a even higher improvement of 9.86% (NFBS) and 7.76% (OASIS). A final test using three different CNN frameworks showed that U-Net3+ outperformed U-Net with a final average DSC of 0.64% for the original image and 0.33% for a combination of original, Tchebichef, and Pseudo-Zernike images.

Bibliography

[1] Yamanakkanavar, N., Choi, J. Y., Lee, B. MRI Segmentation and Classification of Human Brain Using Deep Learning for Diagnosis of Alzheimer's Disease: A Survey. *Sensors* 2020, *20(11)*, doi: 10.3390/s20113243.

[2] Segato, A., Marzullo, A., Calimeri, F., De Momi , D.Artificial intelligence for brain diseases: A systematic review, *APL Bioengineering* 2020, 4, 041503, doi:10.1063/5.0011697.

[3] Ronneberger, O., Fischer, P., Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* 2015, 234–241, doi: 10.1007/978-3-319-24574-4_28.

[4] Nguyen, D. M. H., Vu, H. T., Ung, H. Q., Nguyen, B. T. 3D-Brain Segmentation Using Deep Neural Network and Gaussian Mixture Model, *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Santa Rosa, CA, USA, 2017, pp. 815-824, doi: 10.1109/WACV.2017.96.

[5] Akkus, Z., Galimzianova, A., Hoogi, A. et al. Deep Learning for Brain MRI Segmentation: State of the Art and Future Directions. *J Digit Imaging* 2017, *30*, 449–459, doi: 10.1007/s10278-017-9983-4.

[6] Mcclure, P., et al. Knowing What You Know in Brain Segmentation Using Bayesian Deep Neural Networks. *Frontiers in Neuroinformatics* 2019, *13*, doi: 10.3389/fninf.2019.00067.

[7] Ding, Y., et al. Using Deep Convolutional Neural Networks for Neonatal Brain Image Segmentation. *Frontiers in Neuroscience* 2020, *14*, doi: 10.3389/fnins.2020.00207.

[8] Lee, B., Yamanakkanavar, N., Choi, J. Y. Automatic segmentation of brain MRI using a novel patch-wise U-net deep architecture. *PLOS ONE* 2020,*15(8)*, doi: 10.1371/journal.pone.0236493.

[9] Singh, C., Bala, A. An unsupervised orthogonal rotation invariant moment based fuzzy C-means approach for the segmentation of brain magnetic resonance images. *Expert Systems with Applications* 2021, *164*, doi: 10.1016/j.eswa.2020.113989.

[10] Wachinger, C., Reuter, M., Klein, T. DeepNAT: Deep convolutional neural network for segmenting neuroanatomy. *Neuroimage* 2018, *170*, 434-445, doi: 10.1016/j.neuroimage.2017.02.035.

[11] Qamar, S., et al. A variant form of 3D-U-Net for infant brain segmentation. *Future Generation Computer Systems* 2020, *108*, 613-623, doi: 10.1016/j.future.2019.11.021.

[12] Henschel, L., et al. FastSurfer - A fast and accurate deep learning based neuroimaging pipeline. *Neuroimage* 2020, *219*, doi: 10.1016/j.neuroimage.2020.117012.

[13] Lee, M.,et al. Split-Attention U-Net: A Fully Convolutional Network for Robust Multi-Label Segmentation from Brain MRI. *Brain Sciences* 2020, *10, 974*, doi: 10.3390/brainsci10120974.

[14] Mahesh, V. G. V., Raj, A. N. J., Fan, Z. Invariant moments based convolutional neural networks for image analysis. *International Journal of Computational Intelligence Systems* 2017, *10(1)*, 936-950, doi: 10.2991/ijcis.2017.10.1.62.

[15] Monteiro M., et al., TBI Lesion Segmentation in Head CT: Impact of Preprocessing and Data Augmentation. *In: Crimi A., Bakas S. (eds) Brainlesion: Glioma, Multiple Scle-*

rosis, Stroke and Traumatic Brain Injuries. BrainLes 2019. Lecture Notes in Computer Science, vol 11992. doi: 10.1007/978-3-030-46640-4_2

[16] Tushar, F. I., Alyafi, B., Hasan, M. K., Dahal, L., Brain Tissue Segmentation Using NeuroNet With Different Pre-processing Techniques, 2019 *Joint 8th International Conference on Informatics, Electronics and Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision and Pattern Recognition* (icIVPR), 2019, pp. 223-227, doi: 10.1109/ICIEV.2019.8858515

[17] Khalili, N., et al., Automatic brain tissue segmentation in fetal MRI using convolutional neural networks, *Magnetic Resonance Imaging* 2019, 64, 77-89. doi: 10.1016/j.mri.2019.05.020

[18] Chen, H., et al., VoxResNet: Deep Voxewise Residual Networks for Volumetric Brain Segmentation from 3D MR images, *Neuromage* 2018, 446-455. doi: 10.1016/j.neuroimage.2017.04.041

[19] SivaSai J.G., An Automated Segmentation of Brain MR Image Through Fuzzy Recurrent Neural Network, *In: Bhoi A., Mallick P., Liu CM., Balas V. (eds) Bio-inspired Neurocomputing. Studies in Computational Intelligence* 2021, 903. doi: 10.1007/978-981-15-5495-7_9

[20] Moghaddam, J., Zadeh, H. S., Automatic Segmentation of Brain Structures Using Geometric Moment Invariants and Artificial Neural Networks. *In: Prince J.L., Pham D.L., Myers K.J. (eds) Information Processing in Medical Imaging. IPMI 2009. Lecture Notes in Computer Science*, *vol 5636*. Springer, Berlin, Heidelberg. doi: 10.1007/978-3-642-02498-6_27.

[21] Akhondi-As, A R., Soltanian-Zadeh, H. Constrained optimization of nonparametric entropy-based segmentation of brain structures, *2008 5th IEEE International*

Symposium on Biomedical Imaging: From Nano to Macro, 2008, pp. 41-44, doi: 10.1109/ISBI.2008.4540927.

[22] Akselrod-Ballin, A., Galun, M., Gomori, M.J., Basri, R., Brandt, A. Atlas Guided Identification of Brain Structures by Combining 3D Segmentation and SVM Classification. *In: Larsen R., Nielsen M., Sporring J. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2006. MICCAI 2006. Lecture Notes in Computer Science*, vol 4191. Springer, Berlin, Heidelberg. doi: 10.1007/11866763_26.

[23] Jabarouti Moghaddam, M., Soltanian-Zadeh, H. Automatic Segmentation of Brain Structures Using Geometric Moment Invariants and Artificial Neural Networks. *In: Prince J.L., Pham D.L., Myers K.J. (eds) Information Processing in Medical Imaging. IPMI 2009. Lecture Notes in Computer Science*, vol 5636. Springer, Berlin, Heidelberg. doi: 10.1007/978-3-642-02498-6_27.

[24] Akselrod-Ballin, A., Galun, M., Gomori, J.M., Brandt, A., Basr,i R. Prior Knowledge Driven Multiscale Segmentation of Brain MRI. *In: Ayache N., Ourselin S., Maeder A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2007. MICCAI 2007. Lecture Notes in Computer Science*, vol 4792. Springer, Berlin, Heidelberg. doi: 10.1007/978-3-540-75759-7_15.

[25] Marín, D., Aquino, A., Gegundez-Arias, M. E., Bravo, J. M. A New Supervised Method for Blood Vessel Segmentation in Retinal Images by Using Gray-Level and Moment Invariants-Based Features, *in IEEE Transactions on Medical Imaging*, 2011, *30(1)*, 146-158. doi: 10.1109/TMI.2010.2064333.

[26] Niemeijer, N., Staal, J., Ginneken, B. V., Loog, M., Abramoff, M. D., Fitzpatrick, J., Sonka, M. Comparative study of retinal vessel segmentation methods on a new publicly available database, *in SPIE Med. Imag.*, 2004, *5370*, 648–656. doi: 10.1117/12.535349.

[27] Staal, J., Abramoff, M. D., Niemeijer, M., Viergever, M. A., Ginneken, B. V. Ridge-based vessel segmentation in color images of the retina, *in IEEE Transactions on Medical Imaging*, 2004, *23(4)*, 501-509. doi:10.1109/TMI.2004.825627.

[28] Soares, J. V. B., Leandro, J. J. G., Cesar, R. M., Jelinek, H. F., Cree, M. J. Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification, *in IEEE Transactions on Medical Imaging*, 2006, *25(9)*, 1214-1222. doi: 10.1109/TMI.2006.879967.

[29] Ricci, E., Perfetti, R. Retinal Blood Vessel Segmentation Using Line Operators and Support Vector Classification, *in IEEE Transactions on Medical Imaging*, 2007, *26(10)*,1357-1365. doi: 10.1109/TMI.2007.898551.

[30] Adapa, D,. Raj, A. N. J., Alisetti, S. N., Zhuang, Z., K. G., Naik, G.A supervised blood vessel segmentation technique for digital Fundus images using Zernike Moment based features. *PLoS ONE* 2020, *15(3)*, doi: 10.1371/journal.pone.0229831.

[31] Eskildsen, S. F., et al. BEaST: Brain extraction based on nonlocal segmentation technique. *Neuroimage* 2012, *59(3)*, 2362-2373, doi: 10.1016/j.neuroimage.2011.09.012.

[32] Marcus, D. S., Wang, T. H., Parker, J., Csernansky, J. G., Morris, J. C., Buckner, R. L. Open Access Series of Imaging Studies (OASIS): Cross-sectional MRI Data in Young, Middle Aged, Nondemented, and Demented Older Adults. *Journal of Cognitive Neuroscience* 2007, *19(9)*, 1498-1507, doi: 10.1162/jocn.2007.19.9.1498.

[33] Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., Moore, S., Phillips, S., Maffitt, D., Pringle, M., Tarbox, L., Prior F. 'The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository. *Journal of Digital Imaging* 2013, *26(6)*, 1045-1057, doi: 10.1007/s10278-013-9622-7.

[34] Dalvit Carvalho da Silva R, Jenkyn T. R., Carranza V. A. Development of a Convolutional Neural Network Based Skull Segmentation in MRI Using Standard Tesse-

lation Language Models. *Journal of Personalized Medicine* 2021, *11(4):310* , , doi: 10.3390/jpm11040310.

[35] Hu, M. K. Visual pattern recognition by moment invariants, *in IRE Transactions on Information Theory* 1962, *8(2)*, 179-187. doi: 10.1109/TIT.1962.1057692.

[36] Chong, C. W., Raveendran, P., Mukundan, R., Translation and Scale Invariants of Legendre Moments. *Pattern Recognition* 2004, *37*, 119-129, doi: 10.1016/j.patcog.2003.06.003.

[37] Yang, G. Y., Shu, H. Z., Toumoulin, C., Han, G. N., Luo, L. M. Efficient Legendre moment computation for grey level images, *Pattern Recognition* 2006, *39(1)*, 74-80. doi: 10.1016/j.patcog.2005.08.008.

[38] Erdelyi, A., Magunus, W., Oberhettinger, F., Tricomi, F. G., Higher Transcendental Functions. *New York: McGraw-Hill* 1953, *12*.

[39] Mukundan, R., Ong, S. H., Lee, P. A. Image analysis by Tchebichef moments, *in IEEE Transactions on Image Processing* 2001, *10(9)*, 1357-1364, doi: 10.1109/83.941859.

[40] Wang, G., Wang, S. Recursive computation of Tchebichef moment and its inverse transform. *Pattern Recognition* 2006, *39(1)*, 47-56, doi: 10.1016/j.patcog.2005.05.015.

[41] Chee-Way, C.; raveendran, P.; Ramakrishnan, M. An effcient algorithm for fast computation of Pseudo-Zernike moments. *Int. J. Pattern Recognit. Artif. Intell.* 2011, *17*, doi:10.1142/S0218001403002769.

[42] Teague, M. Image analysis via the general theory of moments. *J. Opt. Soc. Am.* 1980, *70*, 920–930, doi:10.1364/JOSA.70.000920.

[43] Al-Rawi, M.S. Fast computation of pseudo Zernike moments. *J. Real-Time Image Process.* 2010, *8*, 3-10, doi:10.1007/s11554-009-0118-0.

[44] Papakostas, G. A., Boutalis, Y. S., Karras, D. A. et al. Efficient computation of Zernike and Pseudo-Zernike moments for pattern classification applications. *Pattern Recognit. Image Anal.* 2010, *20*, 56–64. doi: 10.1134/S1054661810010050.

[45] Long, F. Microscopy cell nuclei segmentation with enhanced U-Net. *BMC Bioinformatics* 2020, *21(8)*, doi:10.1186/s12859-019-3332-1.

[46] Tan, M., Wu, F., Kong, D., Mao, X., Automatic liver segmentation using 3D convolutional neural networks with hybrid loss function. *Medical Physics* 2020, doi: 10.1002/mp.14732.

[47] Jimenez-Perez, G., Alcaine, A. Camara, O., Delineation of the electrocardiogram with a mixed-quality-annotations dataset using convolutional neural networks. *Sci Rep* 2021, *11*, 863, doi:10.1038/s41598-020-79512-7.

[48] Jimenez-Pastor, A., Alberich-Bayarri, A., Lopez-Gonzalez, R. et al. Precise whole liver automatic segmentation and quantification of PDFF and R2* on MR images. *Eur Radiol* 2021, doi:10.1007/s00330-021-07838-5.

[49] Çiçek, Ö., Abdulkadir, A., Lienkamp, S., Brox, T., Ronneberger, O. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. *In Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016. Lecture Notes in Computer Science* 2016, *9901*, doi:10.1007/978-3-319-46723-8_49.

[50] 3-D Brain Tumor Segmentation using Deep Learining. Available online: https://www.mathworks.com/help/vision/ug/segment-3d-brain-tumor-using-deep-learning.html;jsessionid=780267f783b8b9ec179af9e65620 (accessed on 1 November 2020).

[51] Dice, L. Measures of the Amount of Ecologic Association Between Species. *Ecology* 1945, *26(3)*, 297-302, doi: 10.2307/1932409.

[52] Schenk, A., Prause, G., Peitgen, H. O. Efficient semiautomatic segmentation of 3D objects in medical images. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2000. Lecture Notes in Computer Science* 2000, *1935*, 186-195, doi: 10.1007/978-3-540-40899-4_19.

[53] Jaccard, P. Distribution de la flore alpine dans le bassin des Dranses et dans quelques regions voisines. *Bulletin de la Societe Vaudoise des Sciences Naturelles* 1901, *37(140)*, 241-272, doi: 10.5169/seals-266440.

[54] Rusko, L. Bekes, G., Fidrich, M. Automatic segmentation of the liver from multi- and single-phase contrast-enhanced CT images. *Medical Image Analysis* 2009, *13(6)*, 871-882, doi: 10.1016/j.media.2009.07.009.

[55] Karimi, D., Salcudean, S. Reducing the Hausdorff Distance in Medical Image Segmentation With Convolutional Neural Networks. *IEEE Transactions on Medical Imaging* 2020, *39(2)*, 499-513, doi: 10.1109/TMI.2019.2930068.

[56] Di Ruberto, C., Putzu, L., Rodriguez, G. Fast and accurate computation of orthogonal moments for texture analysis, *Pattern Recognition* 2018, *83*, 498-510, doi: 10.1016/j.patcog.2018.06.012.

[57] Taha, A. A., Hanbury, A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Med Imaging* 2015, *15(29)*, doi:doi.org/10.1186/s12880-015-0068-x.

[58] Huttenlocher, D. P., Klanderman, G. A., Rucklidge, W. A. Comparing images using the Hausdorff distance. *IEEE Trans Pattern Anal Mach Intell.* 1993, *15(9)*, 850-863, doi: 10.1109/34.232073.

[59] Zhou Z., Rahman Siddiquee M.M., Tajbakhsh N., Liang J. U-Net++: A Nested U-Net Architecture for Medical Image Segmentation. *In: Stoyanov D. et al. (eds) Deep Learn-*

ing in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. DLMIA 2018, ML-CDS 2018. Lecture Notes in Computer Science, 2018, *11045*. Springer, Cham. doi: 10.1007/978-3-030-00889-5_1

[60] Huang, H., Lin, L., Tong, R., Hu, H., Zhang, Q., Iwamoto, Y., Han, X., Chen, Y. W., Wu, J. U-Net 3+: A Full-Scale Connected U-Net for Medical Image Segmentation. *In Proceedings of the ICASSP 2020–2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Barcelona, Spain, 4–8 May 2020, pp. 1055–1059, doi: 10.1109/ICASSP40776.2020.9053405.

[61] Li, C., Fan, Y. Cai, X. PyConvU-Net: a lightweight and multiscale network for biomedical image segmentation. *BMC Bioinformatics* 2021, *22(14)*. doi: 10.1186/s12859-020-03943-2.

Chapter 4

Application of a Novel Automatic Method for Determining the Bilateral Symmetry Midline of the Facial Skeleton Based on Invariant Moments

4.1 Introduction

Bilateral symmetry refers to a structure of interest having two sides, with one side being the mirror image of the other. A good example is the human skull, which has bilateral symmetry of its left and right sides. The goal of reconstructive surgery for clinical pathologies of the craniofacial skeleton, whether boney deformities, mandibular alterations, tumors, or trauma, is to restore bilateral symmetry in order to restore function. Similarly, clinical diagnosis and treatment for orthodontics [1], maxillofacial [2], cosmetic [3], and plastic reconstructive surgery [4] seek to restore craniofacial bilateral symmetry. In preceding studies, a method for finding the midline of the craniofacial skeleton was developed by the authors of [5]. This is now a generally accepted method for establishing the midline of the face. However, accurately locating the craniofacial midline is essential for correctly forming an intact template for correcting facial malformations or trauma and for the planning of procedures in the reconstruction process. To date, the most common method for locating the midline for a two-dimensional image, or the midsagittal plane (MSP) for a three-dimensional object, has been the method proposed by the

authors of [6].

Midline symmetry planes have been calculated as a perpendicular midpoint for a stack of horizontal lines crossing bilaterally across the facial skeleton containing boney landmarks or by using boney landmarks to construct a midline that best represents facial symmetry [7]. A simple way to define the midline has been to manually select a number of cephalometric boney landmarks in the dataset, either directly on the plane or at equal distances on either side of the midline plane. However, this requires great attention and care during the selection process by an expert user. Manual selection of skeletal landmarks is time-consuming and unreliable, which results in less accurate estimations of symmetry. In addition, the results have been dependent on the user's ability to find appropriate landmarks, the landmark availability and the visibility of the anatomical landmarks [8].

Methods with some amount of automation have been attempted for finding the midline plane. The authors of [9] described a semi-automatic approach for calculating the symmetry plane of the facial skeleton using principal component analysis (PCA) and the iterative closest point (ICP) alignment method along with surface models reconstructed from computed tomography image data. The initial step was to determine the precise position of the mirror plane using PCA in order to approximately align the mirrored mesh and the original mesh. The ICP algorithm was then used to get a refined registration [10]. The disadvantage of this method was the need for a central point of the image for the calculation of the symmetric plane. This point was obtained using the average of the vertices of the facial mesh. If this point was not provided or if the point was in wrong location, in case of imperfect symmetry, this method could lead to an error in the construction of the symmetric plane. A second limitation lied in the lack of self-learning or the ability to learn. Once applied to an image, the subsequent image will not learn or improve its performance based on the previous data.

Most recently, an upgraded version of a mirroring and registration technique for automatic symmetry plane detection of 3D asymmetrically scanned human faces was discussed in [11]. This work described an ICP-based method that uses particle swarm optimization.

This method starts from a 3D discrete model and evaluates the symmetry plane by a preliminary first-attempt, carried out with a PCA algorithm, which is then refined iteratively until its final estimation obtained by a Levenberg–Marquardt algorithm. This new proposed method was claimed to improve upon the limitations from the authors' previous algorithm [12]; however, the new model still has limitations and fails to incorporate self-learning to improve the model's outcome. For instance, in the case of craniofacial dysmorphism, this model requires the user to interactively segment the area to solve the asymmetry. However, if the algorithm is presented a similar image again, another intervention is required to select the area.

The aim of this study was to present a technique for the automatic calculation of the craniofacial symmetry midline using invariant moments. Figure 4.1 shows the steps of the proposed algorithm. First, based on the cephalometric landmarks, the image is rotated so that the midline passes through the center, i.e., $0°$. The image is then duplicated and a dataset of 30 images of $1°$ resolution is created from $-14°$ to $15°$. This set of images is then passed through different feature extractors and then to the k-nearest neighbors classifier. In the classification phase, pseudo-Zernike moments (PZMs) and PCA were selected after going through the independent component analysis (ICA) feature extractor. A second set of images of 0.5 degrees resolution was taken from the centered images to test the PZMs and PCA. PZMs were selected for having achieved the best accuracy. After the classifier determines the rotation degree of these images, their midpoints are found. Finally, by joining the midpoints and grades described by the classifier, the midlines can be constructed.

4.2 Materials and Methods

4.2.1 Image Creation

To create the dataset, the IdentifyMe database [13] was used. This database contains 464 skull images composed of skull and digital face images (real-world examples of solved skull identification cases) and unlabeled supplementary skull images. The database also contains several
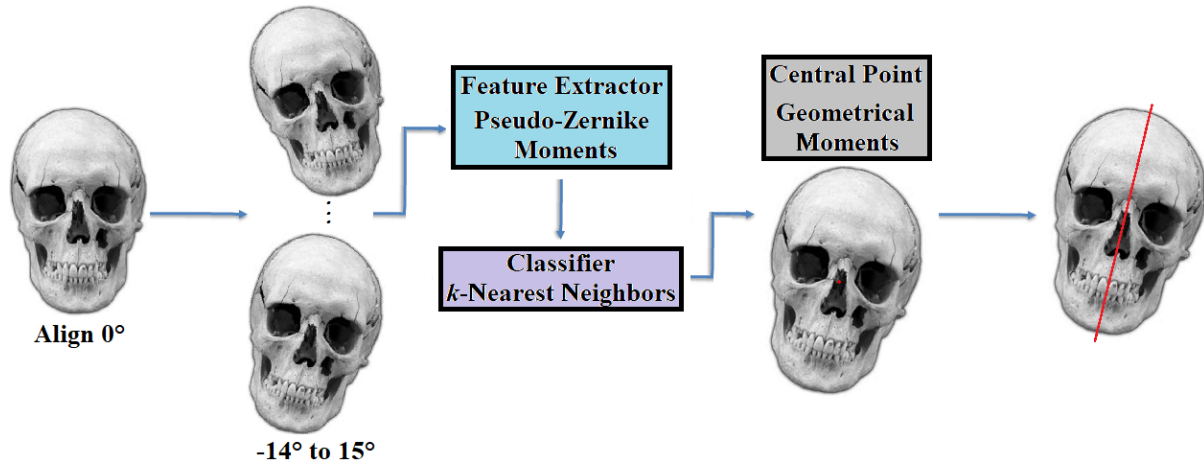
Figure 4.1: Steps of the proposed method.

unsuitable images that needed to be corrected before being added to the dataset. Additionally, this step intended to eliminate regions of non-interest through cropping and the correction of image rotation so that the processing efficiency of the system could be enhanced (Figure 4.2).
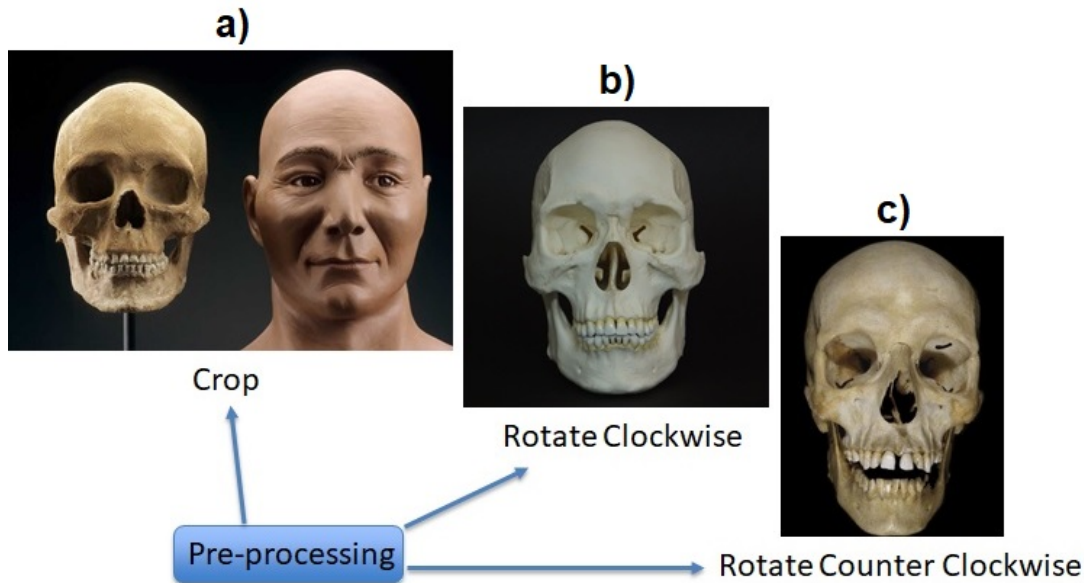
Figure 4.2: An example of images in the IdentifyMe database that required pre-processing in which the image was (a) cropped, (b) rotated clockwise, and (c) rotated counter clockwise.

As the cropping process can be easily solved, the focus was to create a method to correct the rotation, since it will be necessary for training purpose. The first step was to manually identify six cephalometric landmarks (1-Crista Galli, 2-Frontozygomatic Suture, 3-Orbitale, 4-Anterior

Nasal Spine, 5-Subspinale, and 6-Prosthion) (Figure 4.3a). Then, using a grid as reference, the images were rotated using the nearest neighbor interpolation technique so that landmarks 1, 4, 5, and 6 were vertically aligned and landmarks 2 and 3 were horizontally aligned with their respective counterpart as in Figure 4.3b. These images were used as a gold standard to create the rotated images.
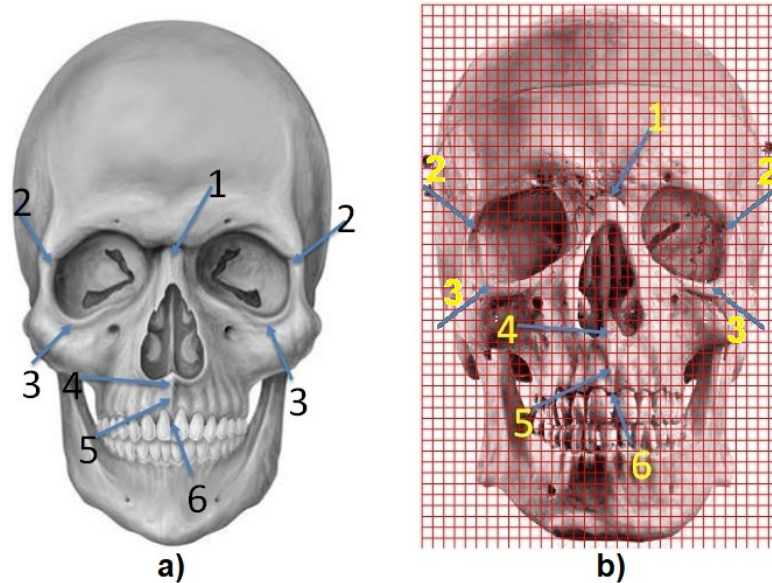
Figure 4.3: Pre-processing step to vertically align images. (a) The six cephalometric landmarks are identified. (b) A grid is then added to the image and rotated so that the landmarks can be horizontally and vertically aligned.

After the images were vertically aligned, 367 images were selected as suitable for use. Based on each image, 30 images were created with inclination angles from −14 to 15 degrees with 1 degree of variation along the sagittal plane, totaling 11,010 images (Figure 4.4) . Those angles were also used as the image labels in the classification step. Finally, the images were resized to 128 × 128 using the nearest neighbor interpolation method. A resolution of 128 × 128 was selected so as to reduce the processing time and computational energy during the creation of PCA and ICA feature vectors. Due to the limitations of the i7-8850H CPU (2.60 GHz) with 16 GB RAM computer used, a resolution of 256 × 256 resulted in the computer crashing.
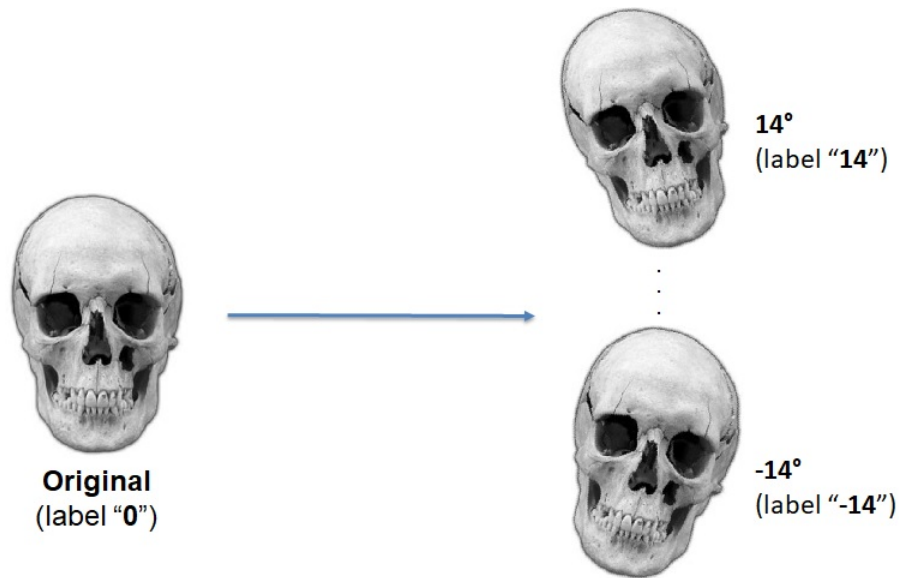
Figure 4.4: The unrotated image (original) was labeled as "0" and subsequent images were labelled based on the angle of rotation.

4.2.2 Feature Extractors

Three feature extraction methods (PZMs, ICA, and PCA) were compared to determine the method with the leading accuracy. The resultant method was then used for the algorithm to determine the midpoint to generate the final midlines.

4.2.2.1 Pseudo-Zernike Moments-PZMs

Zernike moments (ZMs) were first introduced in computer vision by the authors of [14] and are widely used to identify and highlight global features of an image in image processing and machine learning applications [15]. Zernike moments map an image into a set of complex Zernike polynomials. As these Zernike polynomials are orthogonal to each other, Zernike moments can represent the properties of an image without redundancy or overlap of information between moments. Pseudo-Zernike moments are a derivation of the Zernike moments that was shown to be more robust and less sensitive to image noise [16].

Pseudo-Zernike moments, i.e., pseudo-Zernike polynomials, are one of the best invariant image descriptors belonging to a family of circularly orthogonal moments due to their min-

imal information redundancy, robustness to image noise, provision of twice the number of moments, and having more low-order moments. In addition to being rotational invariants, they can be made into scale and translation invariants after certain geometric transformations. PZMs have been used in numerous machine learning and image analysis applications, such as cephalometric landmark detection [17], Alzheimer's disease detection ([18–20]), medical image retrieval [21], detection of tumors and brain tumors ([22–24]), facial expression [25] and facial age recognition [26], facial recognition ([27, 28]), and other industrial applications [29].

As described by the authors of [15], to calculate the PZMs, the image (or region of interest) is initially mapped into a unit disc, where the center of the image is the origin of the disc (Figure 4.5a). Pixels that are outside the disc are not used in the calculation. To include these pixels (Figure 4.5b) the disc can be expanded so that the image function $f(x,y)$ is completely enclosed inside the disc by performing the following

$$x_i = -\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{N-1}x, \quad and \quad y_j = \frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{N-1}y, \tag{4.1}$$

where $x = y = 0, 1, ...(N-1)$, x_i, and y_j are the image coordinates, and the image function $f(x,y)$ is defined over the discrete square domain $N{\times}N$.

(a)   (b)
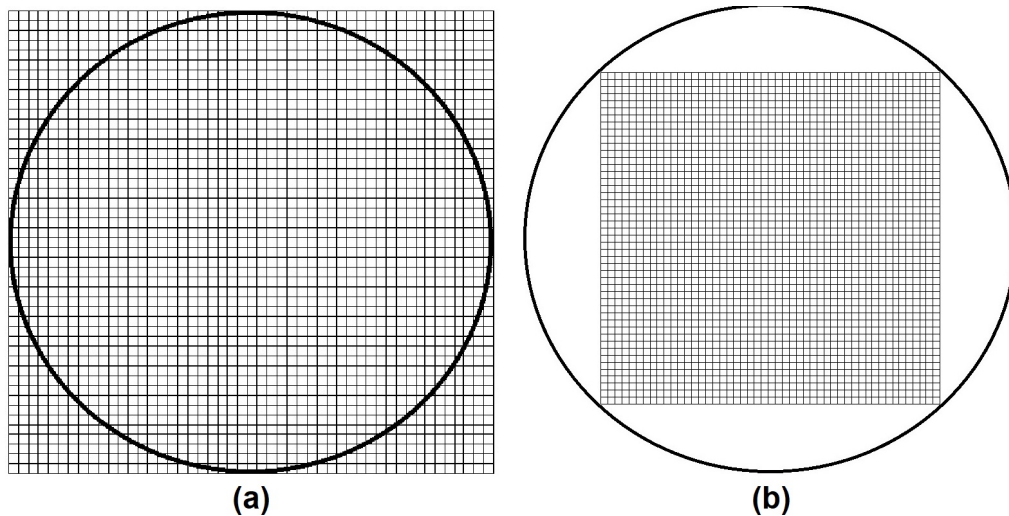
Figure 4.5: (a) Image mapped over and (b) enclosed in a unit disc.

Obtaining the PZMs [15] of an image begins with the calculation of the pseudo-Zernike polynomials (PZP) $V_{pq}(x_i, y_j)$ of order p and repetition q

$$V_{pq}(x_i, y_j) = R_{pq}(r)e^{jq\theta}, \tag{4.2}$$

where p is a non-negative integer, $0 \le |q| \le p$, $\theta = \tan^{-1}(y/x)$, $\theta \in [0, 2\pi]$, and $r = \sqrt{x_i^2 + y_j^2}$. The complex function $V_{pq}(x_i, y_j)$ has two separate parts: the radial polynomials $R_{pq}(r)$ and the angular functions $e^{jq\theta} = (\cos\theta + j\sin\theta)^q$ polynomials in $\cos\theta$ and $\sin\theta$. The radial polynomials are expressed as

$$R_{pq}(r) = \sum_{s=0}^{p-|q|} (-1)^s \frac{(2p+1-s)!}{s!(p+|q|+1-s)!(p-|q|-s)!} r^{(p-s)}. \tag{4.3}$$

Since $R_{pq}(r) = R_{p,-q}(r)$, we consider $q \ge 0$ and rewrite Equation (3) as

$$R_{pq}(r) = \sum_{s=q}^{p} B_{pqs} r^s, \tag{4.4}$$

where

$$B_{pqs} = \frac{(-1)^{p-s}(p+s+1)!}{(p-s)!(s+q+1)!(s-q)!}. \tag{4.5}$$

A *p-recursive* method focusing on reducing time complexity through the fast computing of the pseudo-Zernike radial polynomials using a hybrid method was presented by the authors of [30], where

$$R_{qq}(r) = r^q, \tag{4.6}$$

$$R_{(q+1)q}(r) = (2q+3)r^{q+1} - 2(q+1)r^q, \tag{4.7}$$

$$R_{pq}(r) = (k_1 r + k_2)R_{(p-1)q}(r) + k_3 R_{(p-2)q}(r),$$

$$p = q + 2, q + 3, ..., p_{max},$$

(4.8)

where

$$k_1 = \frac{2p(2p + 1)}{(p + q + 1)(p - q)},$$

(4.9)

$$k_2 = -2p + \frac{(p + q)(p - q - 1)}{(2p - 1)} k_1,$$

(4.10)

$$k_3 = (2p - 1)(p - 1) - \frac{(p + q - 1)(p - q - 2)}{2} k_1 + 2(p - 1)k_2.$$

(4.11)

The PZMs of order p and repetition q of an image function $f(x_i, y_j)$ over a unit disc are represented in a discrete domain by

$$A_{pq} = \frac{p + 1}{\pi} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(x_i, y_j)V_{pq}^*(x_i, y_j)\Delta x_i \Delta y_j.$$

(4.12)

4.2.2.2 Independent Component Analysis—ICA

ICA is a blind source separation or statistical signal processing technique where a given measurement is represented by a linear composition of statistically independent components (ICs) that aims to linearly decompose a random vector into components that are as independent as possible [31]. Given a set of random variables $\{x_1(t), x_2(t), ..., x_n(t)\}$, where t is time or the sample index, it is assumed to be generated as a linear mixture of ICs $\{s_1(t), s_2(t), ..., s_n(t)\}$

$$x = (x_1(t), x_2(t), ..., x_n(t))^T = A(s_1(t), s_2(t), ..., s_n(t))^T = As,$$

(4.13)

where A is an unknown mixture matrix A $R^{n \times n}$. The FastICA algorithm [31] was used to perform the task through the columns of its mixture matrix, which contain the main feature vectors.

4.2.2.3 Principal Component Analysis—PCA

PCA, first introduced by the authors of [32] and then independently developed by the authors of [33], is a technique that preserves the variation present in a large number of interrelated variables of a dataset while reducing the dimensionality. This is achieved by transforming the dataset into a new ordered set of uncorrelated variables or principal components (PCs), so that all the original variables can be represented by the first few variables that maintained the most important features [34].

4.2.3 Geometric Moments—Central Point

In current applications, the central point of the face is obtained manually. Thus, we propose the application of the geometric moments method for the automatic extraction of the central point. Geometric moments are a popular method of moments and have been used to identify the image centroid in a number of image processing tasks [35, 36].

Two-dimensional *(p+q)th* order moments of a digitally sampled $N \times M$ image that has the gray function *f(x,y), (x, y = 0,... M-1)* [37] is given as

$$M_{pq} = \sum_x \sum_y x^p y^q f(x, y),$$

$$p, q = 0, 1, 2, 3, \ldots \quad . \tag{4.14}$$

As described in [38], the mass and area of the zeroth order moment, M_{00}, of a digital image is defined as

$$M_{00} = \sum_x \sum_y f(x, y). \tag{4.15}$$

The center of mass of the image *f(x,y)* is represented by the two first moments

$$M_{10} = \sum_x \sum_y x f(x, y), \tag{4.16}$$

$$M_{01} = \sum_x \sum_y y f(x, y). \tag{4.17}$$

Thus, the centroid of an image can be calculated by

$$\bar{x} = \frac{M_{10}}{M_{00}}, \quad and \quad \bar{y} = \frac{M_{01}}{M_{00}}. \tag{4.18}$$

As best practice, the center of mass was chosen to represent the position of an image in the field of view. The centroid of the image *f(x,y)*, given by Equation (4.18), can be used to describe the position of the image in space by using the point as a reference point.

4.3 Results and Discussions

4.3.1 Classification

PZMs feature vectors were generated through the first 20 PZM orders and repetitions, based on Equations (4.1)–(4.12), totalizing 121 features. After several tests, the 24 best features were selected to represent PZMs.

Before being presented to the ICA and PCA feature extractors, the images were vectorized. For ICA, the mixing matrix A, described in Equation (4.13), was used as its feature vectors. For PCA, the eigenvectors based on eigenvalue order on the covariance matrix or its principal components were used. In total, feature vectors of size 35 were used to represent ICA and PCA.

To make a reasonable comparison among the descriptors, the feature vectors were presented uniquely to the *k*-nearest neighbors classifier (*k*-NN) using the Euclidean distance and the eight

nearest neighbors ($k=8$). In order to evaluate the performance of the feature extractors, the size of the training sets was varied from 10% to 80% of the available database and the rest was used for testing purposes. The k-NN outputs were the predicted images angles and Figure 4.6 presents the comparison of the accuracy rate among ICA, PCA, and PZMs. Accuracy was calculated using the following model

$$Accuracy = \frac{CorrectPredictions}{TotalPredictions} = \frac{TP + TN}{TP + TN + FP + FN}, \qquad (4.19)$$

where TP—true positives, TN—true negatives, FP—false positives, and FN—false negatives.

The extractor based on the ICA obtained a low accuracy rate with a maximum accuracy of 14.5% after training, making it unviable for this application. Thus, a further comparison could be carried out with PZMs versus PCA only to evaluate and finally select the best feature extractor.

Using the selected 367 images and the same inclination angles ($-14°$ to $15°$), a new set of 59 images from the original images was created with a 0.5 degree resolution, totaling 21,653 images, and the results are presented in Figure 4.7. From these results it can be seen that:

1. PCA performed well, but it required more coefficients to achieve a performance similar to pseudo-Zernike moments (Figures 4.6 and 4.7);

2. In Figure 4.6, ICA estimation achieved a bad performance in the experiment. We attributed this to the fact that it is inherently affected by the rotation of the images, which has been already explored in [39];

3. In Figures 4.6 and 4.7, pseudo-Zernike moments outperformed ICA and PCA, as PZMs maintained almost all of the images' features in a few coefficients.

An initial conclusion is that the rotation invariant feature descriptors in the image plane can be effectively developed using the pseudo-Zernike moments method, which has performed well in these scenarios. Thus, the superiority and choice of the feature extraction based on pseudo-Zernike moments in this application became obvious.
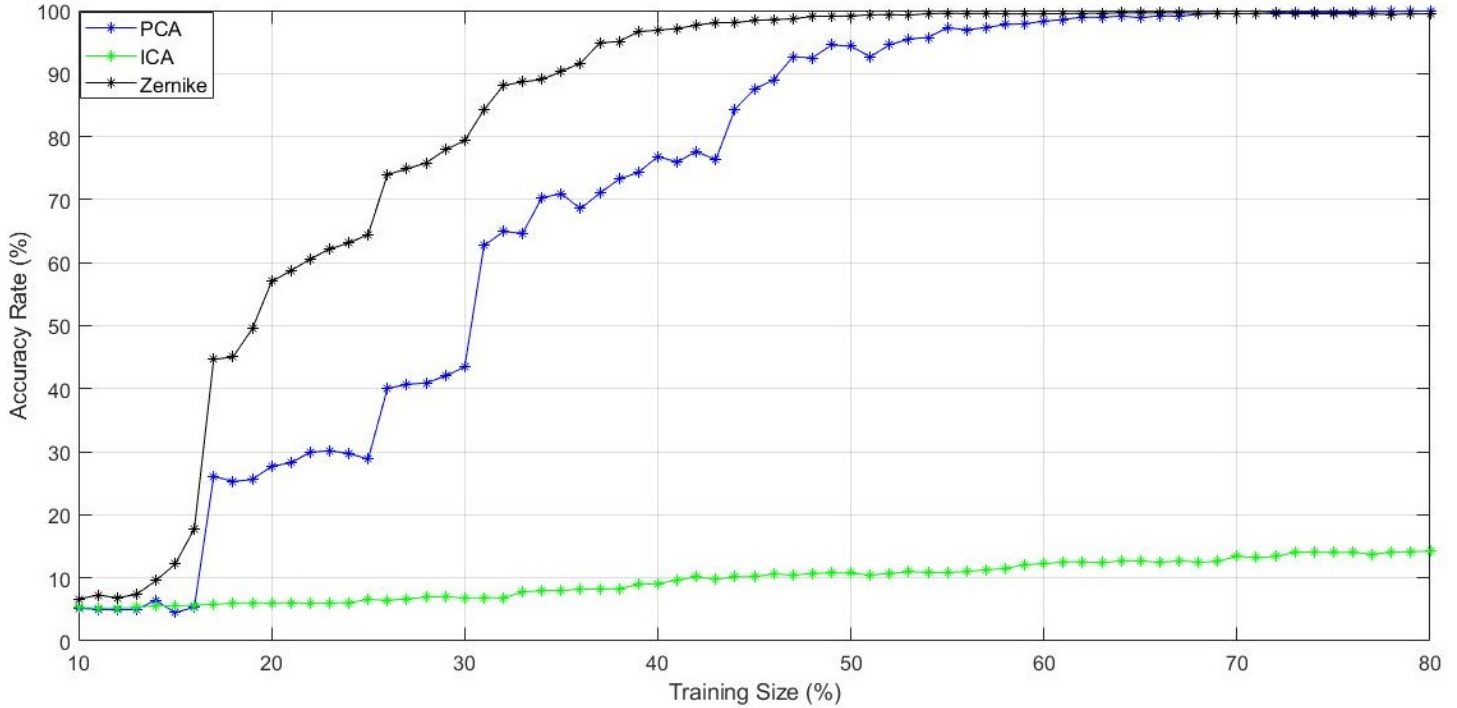
Figure 4.6: Classification accuracy for different feature descriptors, using *k*-NN and Euclidean distance for images rotated from −14° to 15° with a resolution of 1°.

4.3.2 Midpoint Calculation

By using Equations (4.15)–(4.18), the center of the image was calculated. To validate the accuracy of the midpoint technique, visual correctness was used compared with the cephalometric landmarks. However, a calculation of the image center may be performed manually and compared with the midpoint. Figure 4.8 shows three images with rotations of 0°, 3°, 6°, 9°, 12°, and 15° with centers (midpoints) represented by red stars and the angle directions obtained in the process of classifying by the pseudo-Zernike moments represented by black dots. To calculate the PZMs, the 21,653 images dataset (obtained from a 0.5 degree resolution) was divided into 80% for training and 20% for testing, with a total of 4331 testing images. The PZMs method's accuracy was 98.64%, and 1.36% of the images received wrong labels (59 images). The images with wrong labels obtained an error of 0.5° for 33 images, 1° and 1.5° for four images each, 2°, 2.5°, 10°, 13.5°, 14°, and 15° for one image each, 3.5° for three images, and 3° for two images.

Once the midpoints (red stars) and the angles predicted by *k*-NN based on PZMs (black

dots) were calculated (Figure 4.8), the symmetrical line could be easily constructed.
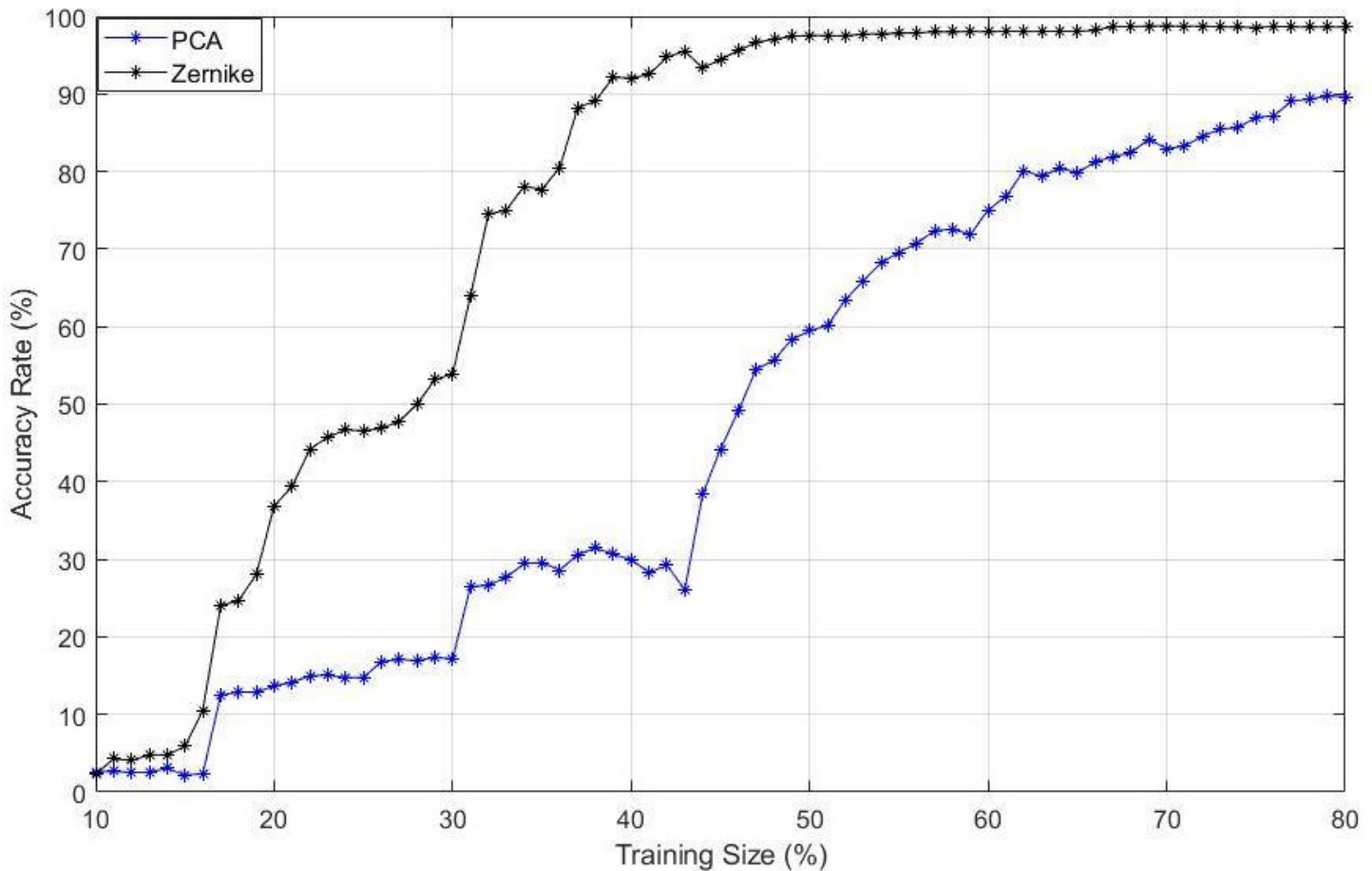
Figure 4.7: Classification accuracy for different feature descriptors, using *k*-NN and Euclidean distance for images rotated from −14° to 15° with a resolution of 0.5°.

The proposed technique presented good results in obtaining the bilateral symmetry midline of the images. However, there are a few limitations to the proposed technique. A small deviation in obtaining the center line could be seen in 59 images. This error is likely due to the fact that some images suffer from a small rotation in the sagittal plane. Additionally, if the images suffer from deformation, incompleteness, or non-uniform brightness, the image center calculus becomes difficult to perform since the moment is a quantitative measure of the function or shape of an image. Furthermore, the algorithm was not tested on non-symmetrical skull images due to a lack of non-symmetrical skull datasets. Our lab will be collecting non-symmetrical skull images to test the algorithm further. Another limitation for the proposed

Figure 4.8: Center of the images calculated using the moment technique. The symmetrical line is constructed by connecting the midpoint (red star) to the PZMs results (black dots).

method is the resolution size of the images (128×128). This resolution was selected as any resolution higher than this resulted in an error from the PC (not enough memory). A higher resolution may result in more accurate results. However, using the lower resolution of 128×128 created acceptable results.

From the viewpoint of the angles, PZMs performed with 98.64% accuracy in the angles and 1.36% (59 images) with wrong angles. However, 33 of the images (55.93%) had an error of only $0.5°$. The majority of this error could be found in $-0.5°$, $0°$, and $0.5°$, where 14 images (42.42% of the 33 images related to $0.5°$) had an error of $0.5°$.

It is possible to state that the required resolution of $0.5°$ reduces the accuracy of the feature extractor, and this error can be related to the image size, which was resized to improve the

processing of the image matrices. However, these errors can be disregarded as they would be imperceptible to the human eye.

4.4   Conclusions

This study proposed an automatic technique for determining the bilateral symmetry midline of the facial skeleton based on invariant moments. A total of 367 skull images were evaluated after preprocessing using the IdentifyMe database. A comparative study between pseudo-Zernike moments, independent component analysis, and principal component analysis as feature descriptors of images using k-nearest neighbors and Euclidean distance was performed and the study of the feature extraction step revealed that pseudo-Zernike moments for feature description had the best performance. PZMs offer an alternative to conventional landmark-based symmetry scores that depends on the general positions of cephalometric landmarks. PZMs are also an alternative to PCA-ICP techniques, which depend on the manual selection of the central point and cannot be improved. With the proposed technique, the central point could be found as the centroid of an image, and then the symmetrical midline could be constructed.

In this study, we have shown the proposed technique to be reliable and to provide the midline symmetry plane with great accuracy, which can be used to aid surgeons in reconstructive craniofacial surgeries.

Bibliography

[1] Jiang, X.; Zhang, Y.; Bai, S.; Chang, X.; Wu, L.; Ding, Y. Threedimensional analysis of craniofacial asymmetry and integrated, modular organization of human head. *Int. J. Clin. Exp. Med.* 2017, *101*, 1424–11431.

[2] Romanet, I.; Graillon, N.; Roux, M.K.L.; Guyot, L.; Chossegros, C.; Boutray, M.D.; Foletti, J.M. Hooliganism and maxillofacial trauma: The surgeon should be warned. *J. Stomatol. Oral Maxillofac. Surg.* 2019, *120*, 106–109, doi:10.1016/j.jormas.2019.01.004.

[3] Guyot, L.; Saint-Pierre, F.; Bellot-Samson, V.; Chikhani, L.; Garmi, R.; Haen, P.; Jammet, P.; Meningaud, J.P.; Savant, J.; Thomassin, J.M.; et al. Facial surgery for cosmeti purposes: Practice guidelines. *J. Stomatol. Oral Maxillofac. Surg.* 2019, *120*, 122–127, doi:10.1016/j.jormas.2019.01.012.

[4] Martini, M.; Klausing, A.; Junger, M.; Luchters, G. The self-defining axis of symmetry: A new method to determine optimal symmetry and its application and limitation in craniofacial surgery. *J. Cranio-Maxillo-Fac. Surg.* 2017, *45*, 1558–1565, doi:10.1016/j.jcms.2017.06.023

[5] Damstra, J.; Fourie, Z.; De Wit, M. A three-dimensional comparison of a morphometric and conventional cephalometric midsagittal planes for craniofacial asymmetry. *Clin. Oral Investig.* 2012, *16*, 285–294, doi:10.1007/ s00784-011-0512-4.

[6] Kim, T.; Baik, J.; Park, J.; Chae, H.; Huh, K. Determination of midsagittal plane for evaluation of facial asymmetry using three-dimensional computer tomography. *Imaging Sci. Dent.* 2011, *41*, 79–84, doi:10.5624/isd.2011.41.2.79.

[7] Roumeliotis, G.; Willing, R.; Neuert, M.; Ahluwali, R.; Jenkyn, T.; Yazdani, A. Application of a novel semi-automatic technique for determining the bilateral symmetry plane of the facial skeleton of normal adult males. *J. Craniofac. Surg.* 2015, *26*, 1997–2001, doi:10.1016/j.medengphy.2013.06.006.

[8] De Momi, E.; Chapuis, J.; pappas, I.; Ferrigno, G.; Hallermann, W.; Schramm, A.; Caversaccio, M. Automatic extraction of the mid-facial plane for cranio-maxillofacial surgery planning. *Int. J. Oral Maxillofac. Surg.* 2006, *35*, 636–642, doi:10.1016/j.ijom.2006.01.028.

[9] Willing, R.; Roumeliotis, G.; Jenkyn, T.; Yazdani, A. Development and evaluation of a semi-automatic technique for determining the bilateral symmetry plane of the facial skeleton. *Med. Eng. Phys.* 2013, *35*, 1843–1849, doi:10.1016/j.ijom.2006.01.028.

[10] Zhang, L.; Razdan, A.; Farin, G.; Femiani, J.; Bae, M.; Lockwood, C. 3d face authentication and recognition based on bilateral symmetry analysis. *Vis. Comput.* 2006, *22*, 43–55, doi:10.1007/s00371-005-0352-9

[11] Angelo, L. D.; Stefano, P. D.; Governi, L.; Marzola, A.; Volpe, Y. A robust and automatic method for the best symmetry plane detection of craniofacial skeletons. *Symmetry* 2019, *11*, 245, doi:10.3390/app8050827.

[12] Angelo, L. D.; Stefano, P. D. A Computational Method for Bilateral Symmetry Recognition in Asymmetrically Scanned Human Faces. *Comput.-Aided Des. Appl.* 2014, *11*, 275–283, doi:10.1080/16864360.2014.863487

[13] Nagpal, S.; Singh, M.; Jain, A.; Singh, R.; Vatsa, M.; Noore, A. On Matching Skulls to Digital Face Images: A Preliminary Approach. CoRR abs/1710.02866. *arXiv* 2017,

arXiv:1710.02866. Available online: http://arxiv.org/abs/1710.02866 (accessed on 01 August 2018).

[14] Teague, M. Image analysis via the general theory of moments. *J. Opt. Soc. Am.* 1980, *70*, 920–930, doi:10.1364/JOSA.70.000920.

[15] Singh, C.; Walia, E.; Pooja, ., Upneja, R. Analysis of algorithms for fast computation of pseudo zernike moments and their numerical stability. *Digit. Signal Process.* 2012, *22*, 1031–1043, doi:10.1016/j.dsp.2012.06.009.

[16] Chee-Way, C.; raveendran, P.; Ramakrishnan, M. An effcient algorithm for fast computation of pseudo-zernike moments. *Int. J. Pattern Recognit. Artif. Intell.* 2011, *17*, doi:10.1142/ S0218001403002769.

[17] Kaur, A.; Singh, C. Automatic cephalometric landmark detection using zernike moments and template matching. *Signal Image Video Process.* 2013, *9*, 117–132, doi:10.1007/s11760-013-0432-7.

[18] Shams-Baboli, A.; Ezoji, M. A zernike moment based method for classification of alzheimer's disease from structural mri. *In:Proceedings of the 3rd International Conference on Pattern Recognition and Image Analysis* Faro, Portugal, 20–23 June 2017; pp. 38–43, doi:10.1109/PRIA.2017.7983061.

[19] Prashar, A. Detection of alzheimer disease using zernike moments. *Int. J. Sci. Eng. Res.* 2017, *8*, 1789–1793.

[20] Wang, S.; Du, S.; Zhang, Y.; Phillips, P.; Wu, L.; Chen, X.; Zhang, Y. Alzheimer's disease detection by pseudo-zernike moment and linear regression classification. *CNS Neurol. Disord. Drug Targets* 2017, *16*, 11–15, doi:10.2174/1871527315666161111123024.

[21] Jyothi, B.; Latha, M.; Mohan, P.; Reddy, V. Medical image retrieval using moments. *Int. J. Appl. Innov. Eng. Manag.* 2013, *2*, 195–200, doi:10.2174/18715273156661111123024.

[22] Iscan, Z.; Dokur, Z.; Olmez, T. Tumor detection by using zernike moments on segmented magnetic resonance brain images. *Expert Syst. Appl.* 2010, *37*, 2540–2549, doi:10.1016/j.eswa.2009.08.003.

[23] Thapaliya, K.; Kwon, G. Identification and extraction of brain tumor from mri using local statistics of zernike moments. *Int. J. Imaging Syst. Technol.* 2014, *24*, 284–292, doi:10.1002/ima.22105.

[24] Nallasivan, G.; Janakiraman, S. Detection and classification of brain tumors as benign and malignant using mri scan images and zernike moment feature set with som. *J. Comp. Sci.* 2016, *10*, 24–36.

[25] Akkoca, B.; Gökmen, M. Facial expression recognition using local zernike moments. *In: Proceedings of the 21st Signal Processing and Communications Applications Conference*, Haspolat, Turkey, 24–26 April 2013; pp. 1–4, doi:10.1109/SIU.2013.6531264.

[26] Malek, M.; Azimifar, Z.; Boostani, R. Facial age estimation using zernike moments and multi-layer perception. *In: Proceedings of the 22nd International Conference on Digital Signal Processing*, London, UK, 23–25 August 2017; pp. 1–5, doi:10.1109/ICDSP.2017.8096073.

[27] Rathika, N.; Suresh, P.; Sakthieswaran, N. Face recognition using zernike moments with illumination variations. *Int. J. Biomed. Eng. Technol.* 2017, *25*, 267–281, doi:10.1504/IJBET.2018.10008622.

[28] Basaran, E.; Gokmen, M.; Kamasak, M. An efficient multiscale scheme using local zernike moments for face recognition. *Appl. Sci.* 2018, *8*, 827, doi:10.3390/app8050827.

[29] Silva, R.D.C.; Thé, G.A.P.; de Medeiros, F.N.S. Geometrical and statistical feature extraction of images for rotation invariant classification systems based on industrial devices. *In: Proceedings of the 21st International Conference on Automation and Computing* (ICAC), Glasgow, UK, 11–12 September 2015; pp. 1–6, doi:10.1109/IConAC.2015.7313946.

[30] Al-Rawi, M.S. Fast computation of pseudo zernike moments. *J. Real-Time Image Process.* 2010, *8*, 3–10, doi:10.1007/s11554-009-0118-0.

[31] Hyvarinen, A.; Karhunen, J.; Oja, E. *Independent Component Analysis*; John Wiley and Sons, Inc.: Hoboken, NJ, USA, 2001, doi:10.1002/wics.1440.

[32] Pearson, K. On lines and planes of closest fit to systems of points in space. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* 1901, *2*, 559–572, doi:10.1080/14786440109462720

[33] Hotelling, H. Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* 1933, *24*, 417–441, doi:10.1037/h0071325.

[34] Jolliffe, I. *Principal Component Analysis*, 2nd ed.; Springer: New York, NY, USA, 2002.

[35] Zhao, Y.; Ni, R.; Zhu, Z. RST transforms resistant image watermarking based on centroid and sector-shaped partition, *Sci. China Inform. Sci.* 2012, *55*, 650–662, doi:10.1007/s11432-011-4470-x

[36] Rocha, L.; Velho, L.; Carvalho, P. Image moments-based structuring and tracking of objects. *In: Proceedings of the XV Brazilian Symposium on Computer Graphics and Image Processing*, Fortaleza Ce, Brazil, 7–10 October 2002; pp. 99–105, doi: 10.1109/SIBGRA.2002.1167130.

[37] Mercimek, M.; Gulez, K.; Mumcu, T.V. Real object recognition using moment invariants. *Sadhana* 2005, *30*, 765–775, doi:10.1007/BF02716709.

[38] Xu, D.; Li, H. Geometric Moments Invariants. *Pattern Recognit.* 2008, *41*, 240–249, doi:10.1016/j.patcog.2007.05.001.

[39] Silva, R.D.C.; Thé, G.A.P.; de Medeiros, F.N.S. Rotation-invariant image description from independent component analysis for classification purposes. *In: Proceedings of the 12th International Conference on Informatics in Control, Automation and Robotics* (ICINCO), Colmar, France, 21–23 July 2015; pp. 210–216.

Chapter 5

Convolutional Neural Networks and Geometric Moments to Identify the Bilateral Symmetric Midplane in Facial Skeletons from CT Scans

5.1 Introduction

Craniomaxillofacial reconstructive surgery is a complex and difficult multidisciplinary technique due to the intricate anatomy of the skull. The aim of craniomaxillofacial reconstruction surgery is to preserve the patients appearance, strengthen facial functions, and regain the bilateral symmetry of the craniofacial. However, craniomaxillofacial reconstructive surgery may lead to complications such as bone disorders, congenital deformities, trauma, pathologies, genetic abnormalities, and cancers. A precise recognition of the bilateral symmetry facial midplane is an imperative step for pre-surgical planning and implant design techniques. For facial restoration, this midplane plays a major role when one side of the image is replicated and used as a guide to recreate the deformed or injured side. The authors of [1], have established a widely agreed approach for defining the midline of the craniofacial skeleton. To date, however, the most popular two-dimensional image application method, or the midsagittal plane (MSP) for a three-dimensional object, is the method introduced by [2].

There are a few approaches that aim to simplify the task of locating the midline plane of

the facial skeleton. One technique proposed by [3], describes a semiautomatic system that, in conjunction with surface models reconstructed from computed tomography images (CT), uses principal component analysis (PCA) and the iterative closest point (ICP) alignment method. The first step is to determine the direction of the mirror plane correctly. This was achieved by using PCA to match the replicated mesh and the initial mesh roughly. Then, the ICP algorithm was described by a refined registration. The downside of this approach was the dependency on the central point of the image for the approximation of the symmetrical plane (obtained using the average of the vertices of the facial mesh). If the central point was in the wrong position due to any external factors (such as imperfect symmetry), this approach would lead to a symmetrical plane in the wrong direction and position . In addition, this algorithm is not able to adjust and learn from previous images to improve its performance, limiting its capabilities.

Alternatively, [4] determines the midline symmetry plane by using boney landmarks to create a midline representing facial symmetry. For a stack of horizontal lines crossing bilaterally through the facial skeleton containing boney landmarks, this approach essentially measures the midline symmetry plane as a perpendicular midpoint. This approach involves the manual collection of a variety of cephalometric boney landmarks in the dataset by either specifically locating the landmarks on the plane (which requires great attention by an expert user) or by using the midline as a reference and locating the landmarks at equal distances from the midline. However, manual skeletal landmark selection is ineffective, time-consuming, and reliant on an expert operator, resulting in errors in the measurement of the symmetry plane.

[5] outlines an ICP-based process for automated symmetry plane detection of 3D asymmetrically scanned human faces that uses particle swarm optimization. This approach starts with a discrete 3D model. The symmetry plane is tested by a tentative first attempt using a PCA algorithm. The model is then refined iteratively by a Levenberg-Marquardt algorithm before its final prediction is obtained. This revised version enhances the shortcomings of [6], but the current implementation also struggles to integrate self-learning to maximize the result of the model and misses the ability to learn from previous versions.

By minimizing the error-index of the symmetry plane, an automated method based on an iterative process was recently proposed by [7]. To automatically correct the initial symmetry plane, with a significant contribution to the use of the rotation matrix derived from the registration process, this method performs analytical data analysis in 3D point sets derived from CT images. First, the plane was divided into two groups by the initial symmetry plane estimated by the PCA and the collection of skull points. Then, to match two point-sets, the ICP registration method was used.

Most recently, [8] introduced a novel automatic concept for determining the bilateral symmetry midline of the facial skeleton based on invariant moments. This technique creates a dataset from images aligned using cephalometric landmarks. The images are then rotated from $-14°$ to $15°$ with a resolution of $0.5°$ degree. Then, after comparing different feature extractors, pseudo-Zernike moments were selected for having the best accuracy using k-nearest neighbors classifier. Finally, after detecting the rotation degree of the image, the midpoint is calculated using geometric moments. However, this model still has some limitations. For instance, this method uses 2D images with an image resolution of 128×128 which become difficult in real applications on different image modalities such as computed tomography and magnetic resonance imaging (MRI). Additionally, this technique was not tested on non-symmetrical skull images which may affect its results.

Thus, this study aims to present a 3D technique for automatically calculating the craniofacial symmetry midplane from CT scans using convolutional neural network (CNN) and geometric moments. Figure 5.1 shows the overview of the proposed method. First, using 3D U-net, the skull is removed from CT images to create a dataset. Then, based on the cephalometric landmarks, the CT image is aligned in the coronal and transverse planes. The image is then duplicated and 2 datasets of 441 images, per image, of $0.5°$ resolution is created from $-5°$ to $5°$. These sets of images are presented to a 3D rotation invariant CNN. After CNN determines the rotation degree of these images in the coronal and transverse planes, the skull midpoints are calculated using 3D geometric moments. Finally, by joining the midpoints and

grades described by the CNN, the midplanes can be constructed.
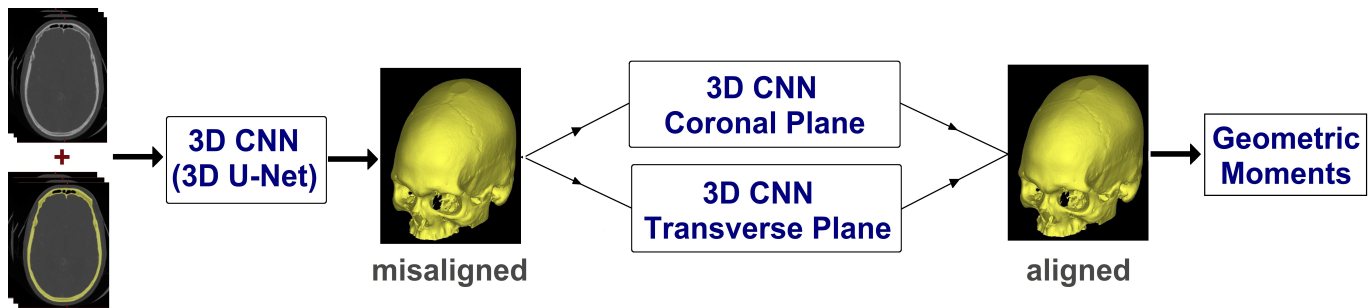
Figure 5.1: Overview of the proposed method.

5.2 Materials and Methods

5.2.1 Data Processing

The dataset used to validate the proposed method was acquired from the qure.ai CQ500 dataset [9]. From this dataset, 195 images with 512×512 and varied depths were selected for training, validation, and test purpose. To create the ground truth labels, CT images were imported into Mimics Medical Imaging Software (Materialise, Leuven, Belgium). First, individual thresholding with manual corrections was applied for each of the 3D volumetric CT images. Then, region growing was applied to create the 3D model mesh. This process allowed for the creation of the standard tessellation language (STL) file format which was converted into a matrix using voxelization method [10] so we can easily process the file in MATLAB R2019b software (Mathworks, Natick, USA) (Figure 5.2).

5.2.2 CNN Architecture and Implementation Details

5.2.2.1 CNN Framework for Biomedical Image Segmentation

The framework chosen in this paper for biomedical image segmentation was the U-Net [11]. U-Net has been used in a number of biomedical image segmentation applications such as kidney segmentation [12], prostate and prostate zones segmentation [13], brain tumor segmentation
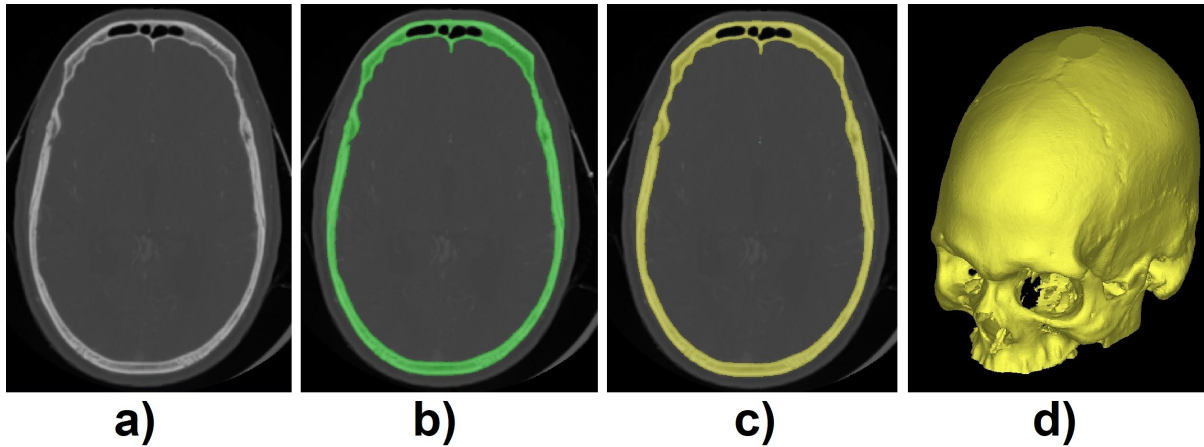
Figure 5.2: a) CT scan, b) thresholding applied, c) region growing, and d) 3D mesh model.

[14], brain segmentation [15], and so forth. Its name emerged from the idea of a U-shape architecture where in the first step, downsampling path, the spatial information is reduced while feature information is increased. In the next step, upsampling path, contracting path concatenate high-resolution features with spatial information and features. The result is a CNN that can work with few training samples and the possibility to apply large images. We adopted a 3D U-net modified version of the code [16] initially implemented for brain tumor segmentation in MRI. The parameters adopted in this work is presented in Table 5.1. These parameters were chosen to avoid computational crash and error, while obtaining a good accuracy for the training set explored in this work.

Table 5.1: Skull CT Segmentation Implementation Details.

Parameter	Value
Optimizer	Adam
Encoder Depth	4
Filter Size	5
Number of First Encoder Filter	6
Patch Per Image	4
Min Batch Size	128
Initial Learning Rate	10^{-2}

5.2.2.2 CNN Framework for Rotation Invariant

By nature, CNN are not rotation invariant, however, with a combination of convolutional, max pooling, average pooling, relu, and fully-connected layers, the CNN framework can be transformed into rotation invariant. A number of papers have exploited the rotation invariant [17–19], however, the adopted framework presented in Table 5.2 worked very well in the dataset proposed using Adam optimizer and mini-batch size of 128. Both CNN models were performed on Intel i7-9700 (3.00 GHz) computer with 64 gigabyte (GB) of ram memory, and two 8GB Video RAM graphics processing units (GPUs) from NVIDIA (one RTX 2070 SUPER and one RTX 2080). The source code was implemented and tested in MATLAB R2019b.

Table 5.2: Rotation-Invariant CNN Framework Adopted.

	Layers	Size	Number Filter	Stride
	3D Conv	1x1	3	1
1	BN + Relu	-	-	-
	3D Max pooling	5x5	-	2
	3D Conv	5x5	8	1
2	BN + Relu	-	-	-
	3D Max pooling	5x5	-	2
	3D Conv	7x7	16	1
3	BN + Relu	-	-	-
	3D Max pooling	3x3	-	2
	3D Conv	5x5	32	1
4	BN + Relu	-	-	-
	3D Max pooling	2x2	-	2
	3D Conv	5x5	64	1
5	BN + Relu	-	-	-
	3D Average pooling	2x2	-	2
6	3D Conv	1x1	128	1
	BN + Relu	-	-	-
7	FC (25 neurons)	-	-	-
	Relu	-	-	-
8	FC (50 neurons)	-	-	-
	Relu	-	-	-
9	FC (labels neurons)	-	-	-
	Softmax	-	-	-

*Conv–Convolutional, BN–Batch Normalization
FC–Fully-Connected

5.2.3 Model Performance Evaluation and Statistical Analysis

For biomedical segmentation evaluation, the Dice Similarity Coefficient (DSC) [20] is the most popular metric to evaluate segmentation models. DSC is a statistical method to gauge the similarity between two sample sets. In biomedical segmentation, DSC measures the overlap between the ground truth and the predicted segmentation where 0 represents no overlap and 1 indicates complete overlap. Equation (5.1) defines DSC, where the area of overlap is divided by the total pixels combined (TP-true positives, FP-false positives, and FN-false negatives).

$$DSC = \frac{\text{Area of Overlap}}{\text{Total Pixels Combined}} = \frac{2TP}{2TP + FP + FN}. \tag{5.1}$$

Symmetric Volume Difference (SVD) [21, 22] is the corresponding error metric.

$$SVD = 1 - DSC. \tag{5.2}$$

Hausdorff Distance (HD) is a size-based method that describes the maximum distances between the boundaries of the segmented regions and the ground truth. This metric can be defined as:

$$HD = max(h(S, GT), h(GT, S)), \tag{5.3}$$

where $h(S,GT) = max_{a \in S} min_{b \in GT} \| a - b \|$ [23].

To evaluate the performances of the CNN framework for rotation invariant, we implemented the following measures: average difference (AD), image quality index (IQI), Laplacian mean square error (LMSE), maximum difference (MD), mean-squared error (MSE), normalized absolute error (NAE), normalized cross-correlation (NK), structural content (SC), and structural similarity index (SSIM). To evaluate the classification, a sensitivity analysis was performed by using positive predictive value (PPV) and negative predictive value (NPV) defined as:

$$S\,ensitivity = \frac{TP}{TP + FN}, \tag{5.4}$$

$$S\,pecificity = \frac{TN}{TN + FP}, \tag{5.5}$$

$$PPV = \frac{TP}{TP + FP}, \tag{5.6}$$

$$NPV = \frac{TN}{TN + FN}. \tag{5.7}$$

5.2.4   hline

To avoid manual intervention, 3D geometric moments are applied for the automatic extraction of the central point. Three-dimensional *(p+q+r)th* order moments of a digitally sampled 3D image that has the gray function *f(x,y,z)* [24] is given as:

$$M_{\mathrm{pqr}} = \sum_{x}\sum_{y}\sum_{z} x^p y^q z^r f(x, y, z), \tag{5.8}$$

where *p, q, r = 0, 1, 2, 3, ...* As described in [25], the mass and area of the zeroth order moment, M_{000}, of a digital image is defined as:

$$M_{000} = \sum_{x}\sum_{y} f(x, y, z). \tag{5.9}$$

The center of mass of the image *f(x,y,z)* is represented by the two first moments:

$$M_{100} = \sum_{x}\sum_{y}\sum_{z} x f(x, y, z), \tag{5.10}$$

$$M_{010} = \sum_{x}\sum_{y}\sum_{z} y f(x, y, z), \tag{5.11}$$

$$M_{001} = \sum_x \sum_y \sum_z zf(x, y, z). \tag{5.12}$$

Thus, the centroid of an image can be calculated by:

$$\bar{x} = \frac{M_{100}}{M_{000}}, \quad \bar{y} = \frac{M_{010}}{M_{000}}, \quad and \quad \bar{z} = \frac{M_{001}}{M_{000}}. \tag{5.13}$$

As best practice, the center of mass was chosen to represent the position of an image in the field of view. The centroid of the image $f(x,y,z)$, given by Equation (5.13), can be used to describe the position of the image in space by using the point as a reference point.

5.3 Results and Discussion

5.3.1 Skull Segmentation

CT volumetric dataset and 3D mesh models were presented to the 3D U-Net with the parameters described by Table 5.1. From the 195 images, 190 were for training and 5 for validation/testing. Table 5.3 shows the DSCs, SVDs, and HDs, in terms of mean± standard deviation (SD) after being trained and tested 10 times, acquired from the testing set. When using 2 GPUs as specified, the CNN took 57 minutes in 15 epochs to converged.

Table 5.3: Skull and Background DSCs, SVDs, and HDs Values of 5 samples and their mean.

DSC-Skull	DSC-Background	SVD-Skull	HD-Skull
0.8993±0.004	0.9927±0.0003	0.1007±0.004	67.17±09.20
0.9093±0.008	0.9948±0.0005	0.0907±0.008	27.81±31.19
0.9150±0.008	0.9941±0.0006	0.0850±0.008	38.78±20.26
0.9349±0.008	0.9958±0.0005	0.0651±0.008	49.92±37.69
0.9362±0.006	0.9953±0.0004	0.0638±0.006	39.99±44.40
0.9189±0.0162	0.9945±0.0012	0.0811±0.0162	44.73±14.79

These results are close to those DSCs reported by [26] (mean DSC of 0.92), and slightly lower than the results reported by [27] (mean DSC of 0.98). Regarding HD, its discrepant values may be directly related to segmentation errors due to bright artifacts found in the original

image, which may be caused by dental filling and components of CT scan machine, Figure 5.3. However, any necessary modification was performed by manual corrections after the segmentation. These predicted labels play an important role in the coronal and transverse alignment.
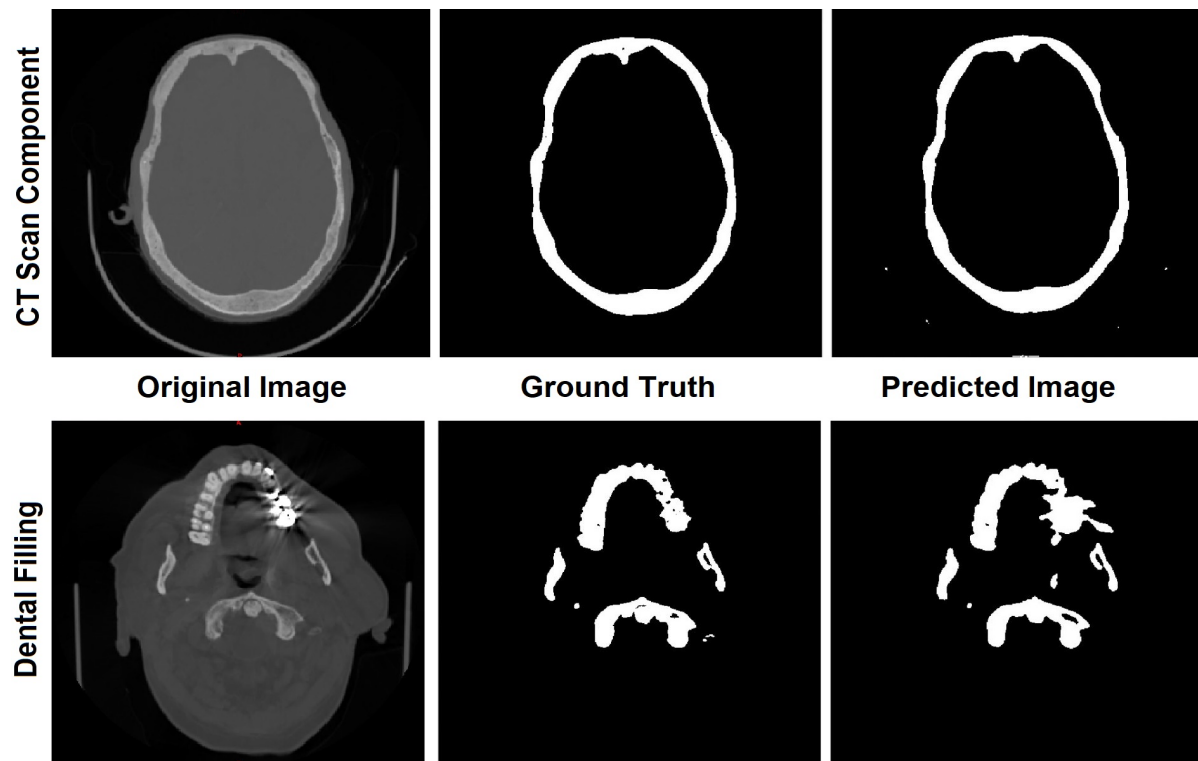
Figure 5.3: Errors in the predicted images caused by machine components and dental fillings.

5.3.2   Transverse and Coronal Angles

To identify the transverse and coronal angles through the 3D CNN, 101 volumetric images, from the 195 segmented images, were selected to create a database. The first step was to identify cephalometric landmarks to help align the predicted labels (Figure 5.3a). In the coronal plane, we selected the 1-crista galli, 2-frontozygomatic suture, and 3-orbitale while in the sagittal plane, we selected 4-lambda and 5-opistocranion. To identify these points and make the necessary alignment, two grids were generated as a reference in the transverse and coronal planes. A number of slices were verified, and the necessary adjustments were performed (Figure 5.4b and Figure 5.4c).
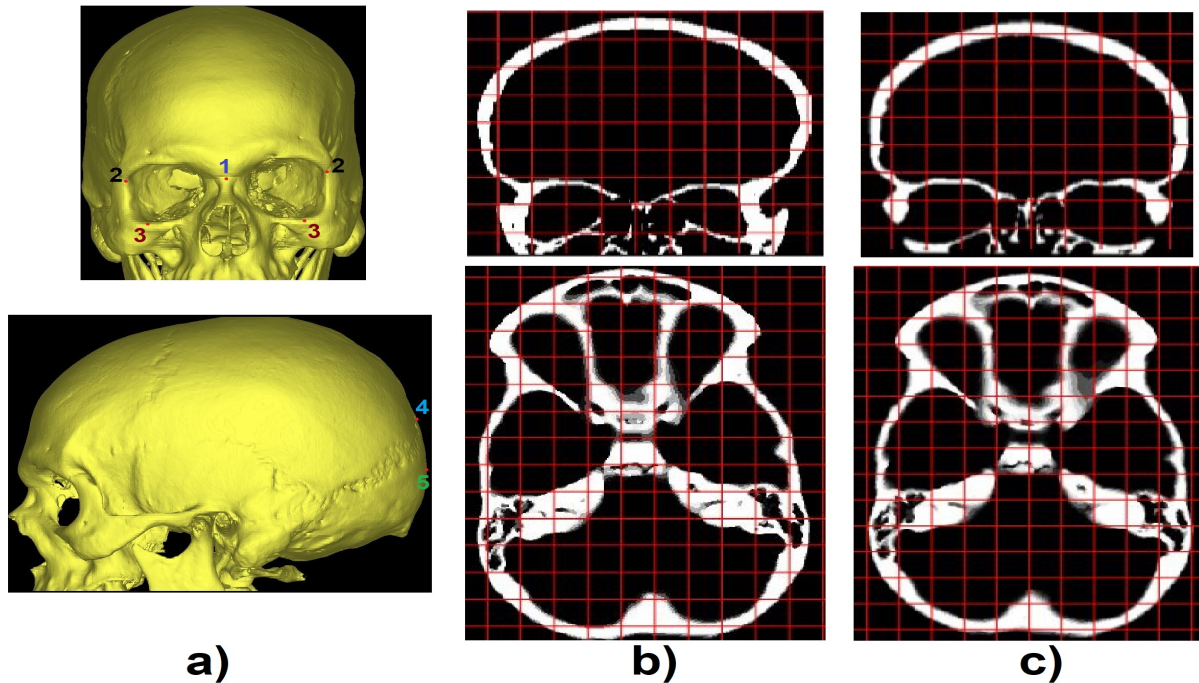
Figure 5.4: a) Identification of cephalometric landmarks in coronal (top) and sagittal plane (bottom), b) CT slices misaligned, and c) CT slices after the alignment procedure (coronal plane-top and transverse plane-bottom).

After the alignment, for each of the 101 images, a set of 441 images with inclination angles from −5 to 5 degrees, with 0.5-degree increments, along the coronal and transverse planes was created. In total, 44541 images were created and were divided in 21 labels. These labels represented the 0.5° of variation in the coronal and transverse planes from −5° to 5°, Figure 5.5.

To reduce the computational and processing time, these images were reduced to 128×128 using nearest neighbor interpolation method. Additionally, the volumetric image was divided into four rectangular sub-cubes and only one-quarter of the whole image space was used to predict the angles as shown in Figure 5.6. As this step aims to identify the coronal and transverse angles in symmetrical skulls, these steps do not affect the output image. Figure 5.7 summarizes this process.

In this phase, we used the 3D rotation invariant CNN introduced in section 2.2.2 and detailed in Table 5.2. To optimize the CNN convergence, computational time, and accuracy, 2 identical datasets were created using the 44541 images and represented by 21 labels. Thus, 21

labels represents the rotation in the transverse plane and 21 labels in the coronal plane. 90% of the dataset was used for training, 10% for validation/testing.Training and testing were performed 10 times. Table 5.4 shows the analytical performance for these 2 CNNs. It took 30 minutes in 5 epochs for the transverse and 70 minutes in 12 epochs for the coronal CNN to converge using 2 GPUs.
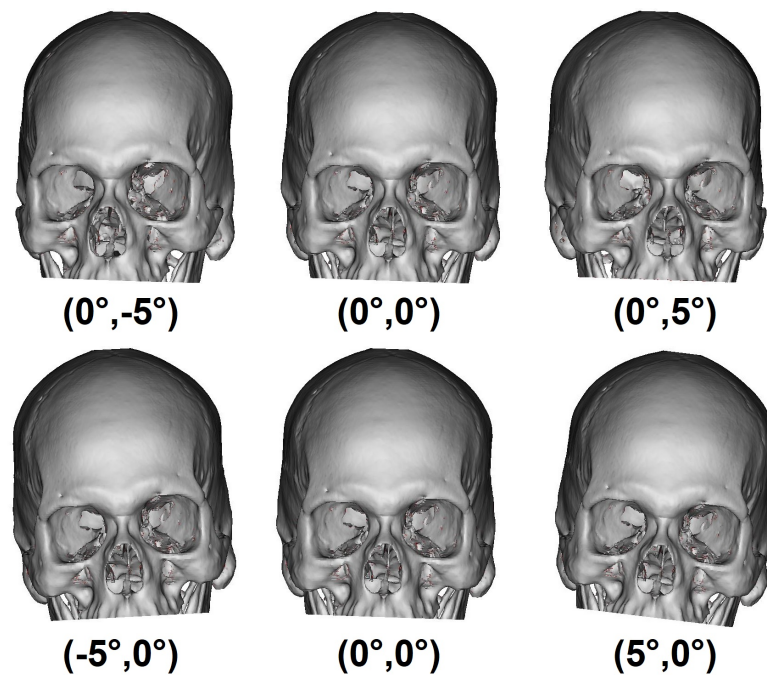
Figure 5.5: Examples of images rotated in the coronal and transverse planes where parentheses represents (Coronal,Sagittal) respectively.

As seen in Table 5.4, we can see that CNN performed well and can represent a rotation invariant image descriptor for these scenarios. In fact, the accuracy reached ≈99% with simple hyperparameter, which allows the construction of a simple 3D CNN. After the discovery of the transverse and coronal deviation angles, the original image was then rotated accordingly with these two found angles.
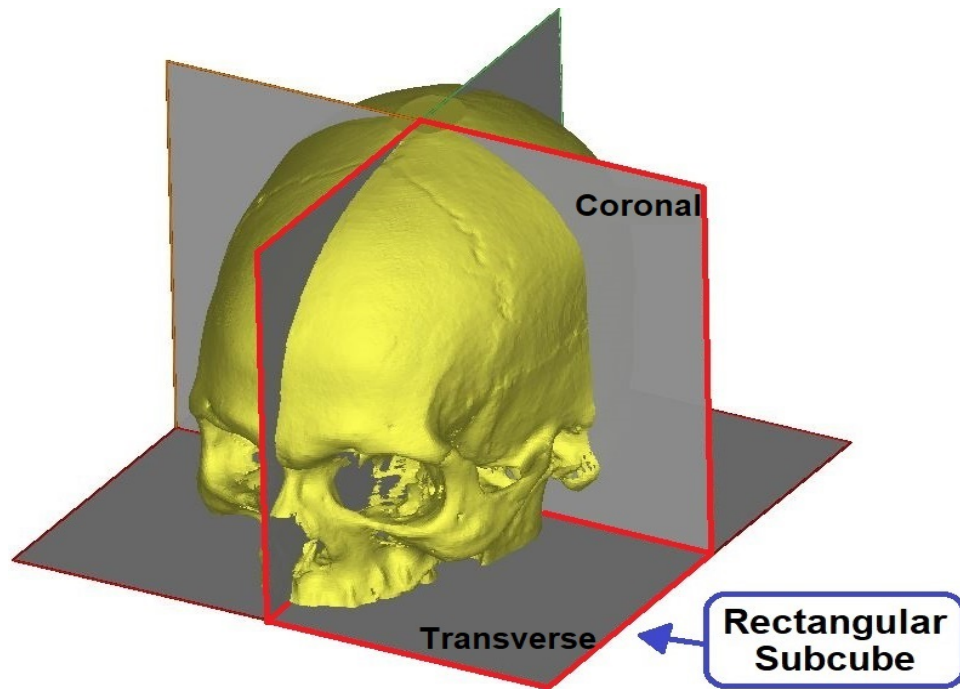
Figure 5.6: Volumetric image divided into rectagular subcubes and the selected one-quarter subcube.
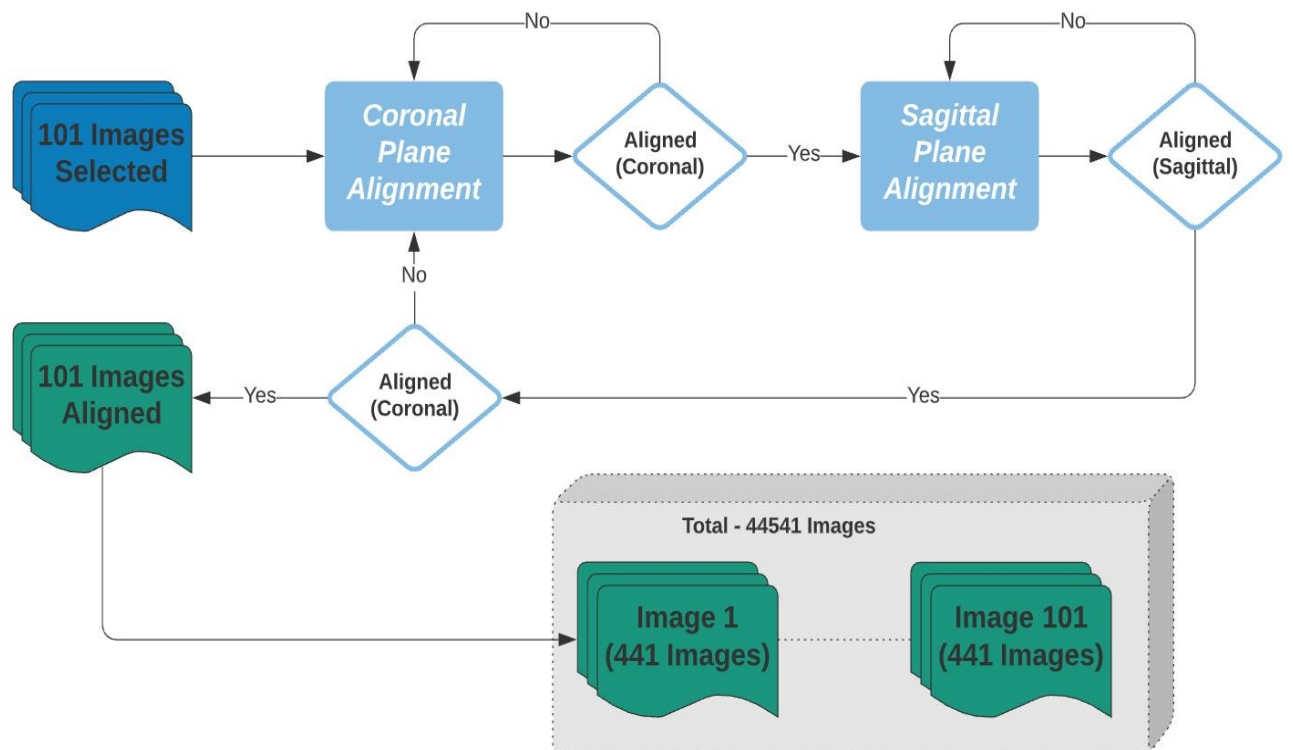
Figure 5.7: Flowchart representing the process of transversal and coronal alignment and database creation.

Table 5.4: Statistical analysis of the Coronal and Transverse CNNs.

Index	Ideal Value	Coronal CNN Value	Transverse CNN Value
Accuracy	1	0.9909±0.0038	0.9947±0.0034
Sensitivity	1	0.9811±0.0170	0.9969±0.0054
Specificity	1	0.9982±0.0012	0.9994±0.0005
PPV	1	0.9646±0.0236	0.9892±0.0106
NPV	1	0.9991±0.0009	0.9998±0.0003
AD	0	0.0124±0.0085	0.0004±0.0095
IQI	1	0.9979±0.0014	0.9994±0.0004
LMSE	0	0.9926±0.6717	0.9571±0.0814
MD	0	11.667±1.1547	9.3333±5.6862
MSE	0	0.2710±0.1883	0.0723±0.0561
NAE	0	0.0030±0.0018	0.0012±0.0007
NK	1	0.9985±0.0010	0.9998±0.0004
SC	1	0.9982±0.0008	1.0000±0.0008
SSIM	1	0.9285±0.0313	0.9523±0.0290

5.3.3 Geometric Moments - Image Centre

Finally, to calculate the center of the volumetric images, equations 5.8 to 5.13 were used. As there are no patterns to validate the accuracy of the center-point, visual evaluation was used and compared with cephalometric landmarks. Figures 5.8 and 5.9 show the cross-sectional plane created from the geometric moments. Figure 5.9a presents the perspective view and 5.9b shows the front view of one aligned sample with measured dimensions of frontozygomatic suture and orbitale in the sagittal plane displayed in Materialise MiniMagic software. We used [28] to convert from voxel into stl file format.

5.3.4 Deformed Skull Test

To validate this method in deformed images, 8 defected CT images were used from two different datasets [29–32] found in the cancer imaging archive data collections (TCIA) [33]. In the first step, STL files were generated by the 3D U-Net using the same parameters as the Table 5.1. DSCs, SVDs, and HDs result are presented in Table 5.5 and a sample is shown in
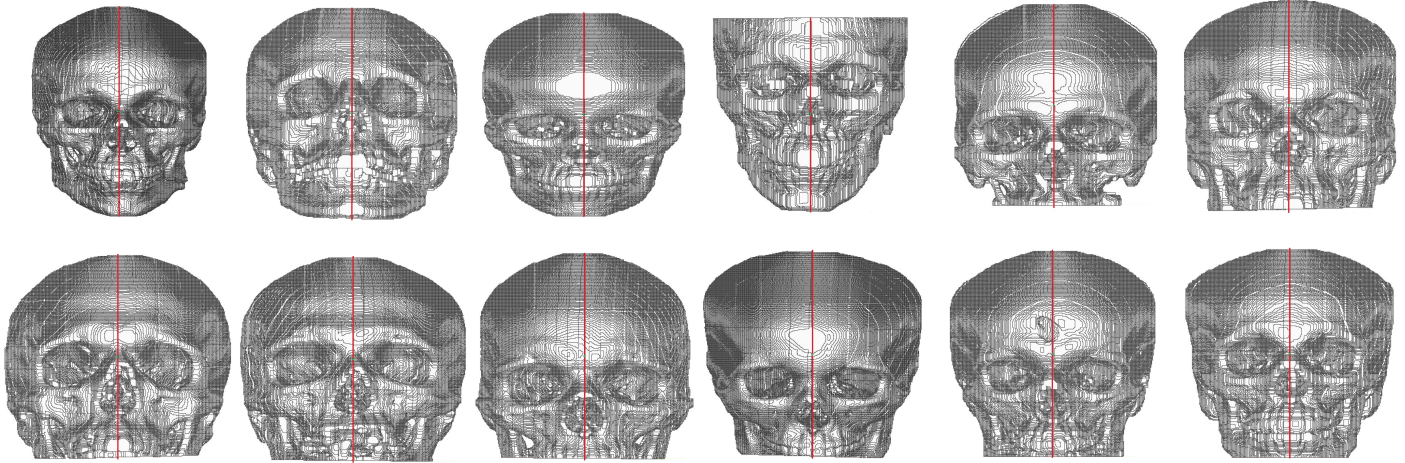
Figure 5.8: 12 samples with their respective cross-sectional plane created using geometric moments.

a)   b)
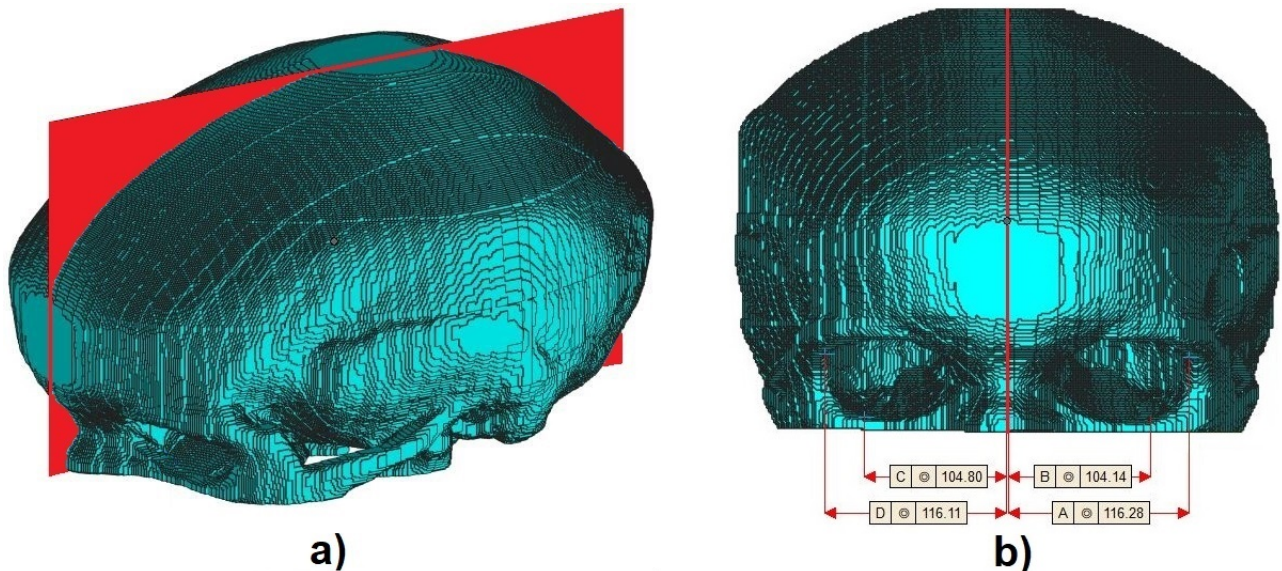
Figure 5.9: a) Perspective and b) front view.

Figure 5.10.

CQ500 database does not contain deformed images which may have caused the discrepancy in the DSCs results. Unfortunately, there are no databases of deformed skull for analysis. Furthermore, the 195 images used for training were not enough to improve the 3D U-Net prediction. [27] reported a 6% mean improvement compared to [26] and associated the improvement with the size of the image dataset for training purpose. Moreover, in the sampled image, Figure 5.10, part of the vertebral column and small segmented parts that do not belong to the

skull were segmented by mistake, which generated the worst DSC value and high HDs values. This is likely associated with the C1 vertebrae labeled during the creation of the ground truth and small bright artifacts on the outside of the skull in the CT scans. Even though it seems like an error, this can be disregarded since Figure 5.10 shows that the ground truth and predicted label to be very close.

Table 5.5: Skull and Background DSCs, SVDs, and HDs Values of the 8 defected skulls and their mean.

DSC-Skull	DSC-Background	SVD-Skull	HD-Skull
0.8206±0.080	0.9902±0.0003	0.1794±0.080	52.75±03.69
0.8114±0.005	0.9831±0.0011	0.1886±0.005	49.52±06.21
0.8294±0.012	0.9875±0.0007	0.1706±0.012	47.55±08.61
0.8625±0.016	0.9890±0.0008	0.1375±0.016	46.77±07.76
0.8973±0.011	0.9946±0.0004	0.1027±0.011	45.65±10.86
0.8302±0.011	0.9817±0.0010	0.1698±0.011	65.05±07.55
0.7888±0.010	0.9898±0.0004	0.2112±0.010	48.28±09.57
0.8360±0.014	0.9891±0.0012	0.1640±0.014	50.22±14.06
0.8346±0.0330	0.9881±0.0041	0.1654±0.0330	47.86±06.21
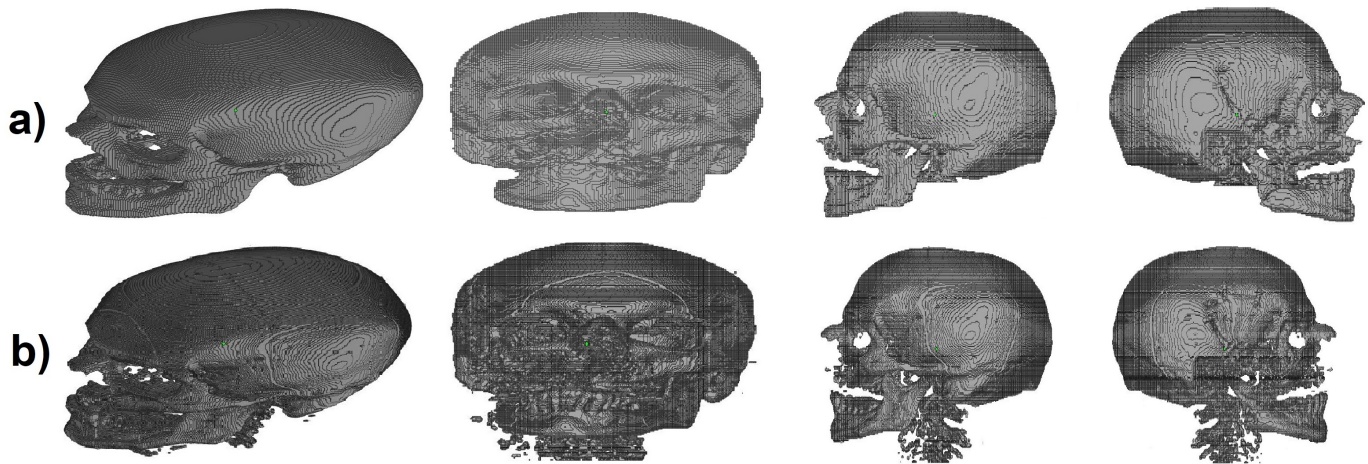
Figure 5.10: a) Ground truth and b) predicted.

Finally, using the 8 models predicted, transverse and coronal angles were calculated, and the center of the image was acquired using 3D geometric moments. Figure 5.11 shows the results for the deformed images.
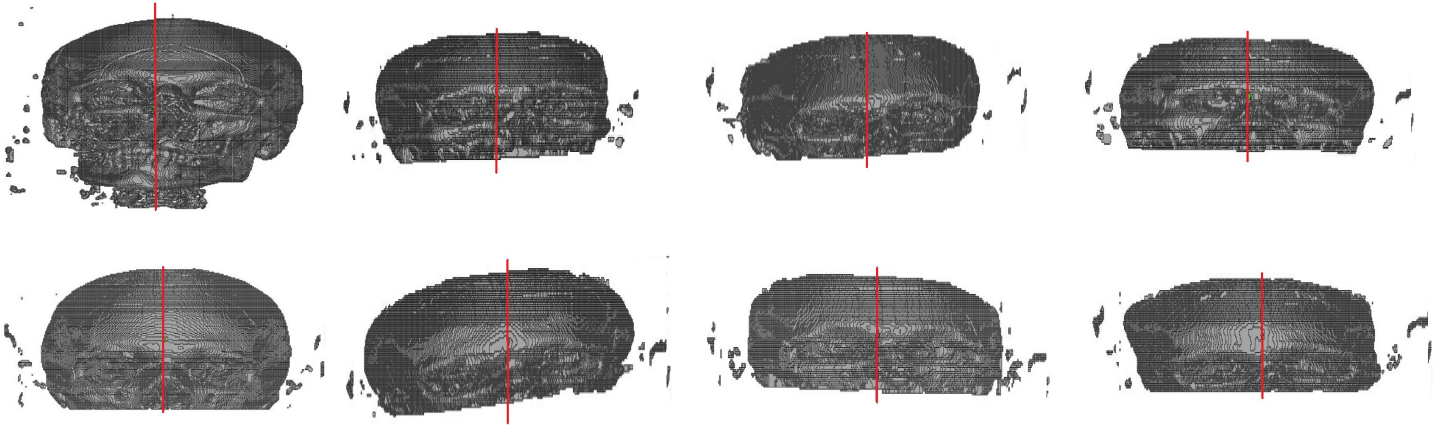
Figure 5.11: Assymetric images.

The proposed method shows good results in obtaining the bilateral symmetric midplane of regular/symmetric images. However, for deformed images, it failed in identify the rotation for some images along with the image center. These errors are likely due to some factors:

- the small database size, which is already reported in [27];

- to the best of our knowledge, there are no deformed CT database available which restricts the possibility to train the system with deformed images;

- during the ground truth segmentation process and voxelization, a few regions of interest (ROIs) may have not been incorporated in the 3D model. The first may be caused by the manual selection of the ROI, performed by an expertise, which leads to the CNN generating the defects. Secondly, a quantity of information from the skull-voxel may be lost due to the smoothing of the edges and noise residuals removal performed in the segmentation process;

- regarding the center of the 3D images, as reported by [8], when the image suffers from lack of symmetry, non-uniform brightness, deformation, interference, or incompleteness, the calculation of the image center using geometric moments becomes complex and finds some restrictions as this technique is a quantitative measure of an image's function or structure.

It is possible to state that the proposed method obtained good results from symmetric CT images datasets. However, for deformed images, an improvement is necessary to achieve better results. For this purpose, an increase in the CT database size may be performed as well as the inclusion of non-uniform and deformed CT images which may also lead to an improvement in the detection of the transverse and coronal angles. Finally, a modification in the method to identify the center of the image may be carried out by creating a 3D U-net to segment the nasal bone instead of the presented geometric moments technique. This method will allow for the definition of the center of the image using the center of the nasal bone.

5.4   Conclusions

This study aimed to introduce a 3D automatic technique for determining the craniofacial symmetry midplane from CT scans using convolutional neural network and geometric moments. A total of 195 symmetric CT images was used to evaluate this method using the CQ500 database while eight asymmetric CT images from TCIA database were used to evaluate the performance in asymmetric images. For symmetric images, this method obtained results close to 99%. However, for asymmetric images, the method needs further development to improve its results. CNNs offer an effective alternative to the pseudo-Zernike moments method and conventional landmark-based symmetry scores that depends on the general positions of cephalometric landmarks. CNNs are also an alternative to PCA-ICP techniques, which depend on the manual selection of the central point which cannot be improved. With the proposed technique, the central point could be found as the centroid of an image, and then the symmetrical midplane can be constructed for symmetric images. In this study, we have shown the proposed technique to be reliable and to provide the midplane symmetry plane with great accuracy in symmetric images. This method can be used as a tool to aid surgeons in reconstructive craniofacial surgeries.

Bibliography

[1] Damstra, J., Fourie, Z., De Wit, M., 2012. A three-dimensional comparison of a morpho-metric and conventional cephalometric midsagittal planes for craniofacial asymmetry. *Clin. Oral Invest.*, 16, 285–294. doi: 10.1007/ s00784-011-0512-4.

[2] Kim, T., Baik, J., Park, J., Chae, H., Huh, K., 2011. Determination of midsagittal plane for evaluation of facial asymmetry using three-dimensional computer tomography. *Imaging Science in Dentistry*, 41, 79–84. doi: 10.5624/isd.2011.41.2.79.

[3] Willing, R., Roumeliotis, G., Jenkyn, T., Yazdani, A., 2013. Development and evaluation of a semi-automatic technique for determining the bilateral symmetry plane of the facial skeleton. *Medical Engineering and Physics*, 35, 1843–1849. doi: 10.1016/j.ijom.2006.01.028.

[4] Roumeliotis, G., Willing, R., Neuert, M., Ahluwali, R., Jenkyn, T., Yazdani, A., 2015. Application of a novel semi-automatic technique for determining the bilateral symmetry plane of the facial skeleton of normal adult males. *The Journal of Craniofacial Surgery*, 26, 1997–2001. doi: 10.1016/j.medengphy.2013.06.006.

[5] Angelo, L. D., Stefano, P. D., Governi, L., Marzola, A., Volpe, Y., 2019. A robust and automatic method for the best symmetry plane detection of craniofacial skeletons. *Symmetry*, 11(2):245. doi: 10.3390/app8050827.

[6] Angelo, L. D., Stefano, P. D., 2014. A Computational Method for Bilateral Symmetry Recognition in Asymmetrically Scanned Human Faces. *Comput-Aided Des. Appl.*, 2014, 11(3), 275–283. doi: 10.1080/16864360.2014.863487.

[7] Noori, S. M. R., Farnia, P., Bayat, M. et al., 2020. Automatic detection of symmetry plane for computer-aided surgical simulation in craniomaxillofacial surgery. *Phys. Eng. Sci. Med.*, 43, 1087-1099. doi: 10.1007/s13246-020-00909-9.

[8] Dalvit Carvalho da Silva, R., Richard Jenkyn, T., 2020. Alexander Carranza, V. Application of a Novel Automatic Method for Determining the Bilateral Symmetry Midline of the Facial Skeleton Based on Invariant Moments. *Symmetry*, 12(9), 1448. doi: 10.3390/sym12091448.

[9] Chilamkurthy, S., Ghosh, R., Tanamala, S., Biviji, M., Warier, P., 2018. Deep learning algorithms for detection of critical findings in head ct scans: a retrospective study. *The Lancet*, 392, 2388–2396. doi: 10.1016/S0140-6736(18)31645-3.

[10] Patil, S., Ravi, B., 2005. Voxel-based representation, display and thickness analysis of intricate shapes, *In: Ninth International Conference on Computer Aided Design and Computer Graphics* (CAD-CG'05), pp. 6 pp.–. doi: 10.1109/CAD-CG.2005.86.

[11] Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation, *CoRR abs/1505.04597*. arXiv:1505.04597.

[12] Kim, T., Lee, K., Ham, S. et al. Active learning for accuracy enhancement of semantic segmentation with CNN-corrected label curations: Evaluation on kidney segmentation in abdominal CT. *Sci Rep*, 10, 366 (2020). doi: 10.1038/s41598-019-57242-9.

[13] Aldoj, N., Biavati, F., Michallek, F. et al. Automatic prostate and prostate zones segmentation of magnetic resonance images using DenseNet-like U-net. *Sci Rep*, 10, 14315 (2020). doi: 10.1038/s41598-020-71080-0.

[14] Dong H., Yang G., Liu F., Mo Y., Guo Y. 2017. Automatic Brain Tumor Detection and Segmentation Using U-Net Based Fully Convolutional Networks. In: Valdés Hernández M., González-Castro V. (eds) Medical Image Understanding and Analysis, MIUA 2017 *Communications in Computer and Information Science*, vol 723. Springer, Cham. doi: 10.1007/978-3-319-60964-5_44.

[15] Lee B, Yamanakkanavar N, Choi JY. 2020. Automatic segmentation of brain MRI using a novel patch-wise U-net deep architecture. *PLoS ONE*, 15(8): e0236493. doi: 10.1371/journal.pone.0236493.

[16] Çiçek, Özgün,. Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O., 2016. 3d u-net: Learning dense volumetric segmentation from sparse annotation, *CoRR abs/1606.06650*. arXiv:1606.06650.

[17] Chen, Y., Lyu, Z.X., Kang, X., Wang, Z.J., 2018. A rotation-invariant convolutional neural network for image enhancement forensics, *In: 2018 IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP), pp. 2111–2115. doi: 10.1109/ICASSP.2018.8462057.

[18] Chidester, B., Zhou, T., Do, M.N., Ma, J., 2019. Rotation equivariant and invariant neural networks for microscopy image analysis. *Bioinformatics*, 35, i530–i537. URL: https://doi.org/10.1093/bioinformatics/btz353, doi: 10.1093/bioinformatics/btz353.

[19] Kim, J., Jung, W., Kim, H., Lee, J., 2020. CyCNN: A Rotation Invariant CNN using Polar Mapping and Cylindrical Convolutional Layers. *arXiv*: 2007.10588.

[20] Dice, L. R., 1945. Measures of the amount of ecologic association between species. *Ecology*, 26, 297–302. doi: 10.2307/1932409.

[21] Schenk, A., Prause, G., Peitgen, HO., 2000 Efficient Semiautomatic Segmentation of 3D Objects in Medical Images. In: Delp S.L., DiGoia A.M., Jaramaz B. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2000. *MICCAI*

2000. Lecture Notes in Computer Science, vol 1935. Springer, Berlin, Heidelberg, doi: 10.1007/978-3-540-40899-4_19.

[22] Yeghiazaryan, V., Voiculescu, I. D., 2018. Family of boundary overlap metrics for the evaluation of medical image segmentation. *Journal of Medical Imaging*, 5(1), doi: 10.1117/1.JMI.5.1.015006.

[23] Vania, M., Mureja, D., Lee, D., 2019. Automatic spine segmentation from CT images using Convolutional Neural Network via redundant generation of class labels. *Journal of Computational Design and Engineering*, 6(2), pp. 224-232. doi: 10.1016/j.jcde.2018.05.002.

[24] Mercimek, M., Gulez, K., Mumcu, T.V., 2005. Real object recognition using moment invariants. *Sadhana*, 30, 765–775. doi: 10.1007/BF02716709.

[25] Xu, D., Li, H., 2008. Geometric Moments Invariants. *Pattern Recognition*, 41(1), 240-249. doi: 10.1016/j.patcog.2007.05.001.

[26] Minnema, J., van Eijnatten, M., Kouw, W., Diblen, F., Mendrik, A., Wolff, J., 2018. Ct image segmentation of bone for medical additive manufacturing using a convolutional neural network, *Computers in Biology and Medicine* 103, 130 – 139. doi: 10.1016/j.compbiomed.2018.10.012.

[27] Kodym, O., Španěl, M., Herout, A., 2019. Segmentation of defective skulls from ct data for tissue modelling. *arXiv*:1911.08805.

[28] Adam, A., 2020. Converting a 3d logical array into an stl surface mesh. https://www.mathworks.com/matlabcentral/fileexchange/27733-converting-a-3d-logical-array-into-an-stl-surface-mesh, MATLAB Central File Exchange, Retrieved: November 1, 2020.

[29] Kinahan, P., Muzi, M., Bialecki, B., Coombs, L., 2018. Data from acrin-fmisobrain. doi: 10.7937/K9/TCIA.2018.VOHLEKOK.

[30] Gerstner, E.R., Zhang, Z., Fink, J.R., Muzi, M., Hanna, L., Greco, E., Prah, M., Schmainda, K.M., Mintz, A., Kostakoglu, L., Eikman, E.A., Ellingson, B.M., Ratai, E.M., Sorensen, A.G., Barboriak, D.P., Mankoff, D.A., 2016. Acrin 6684: Assessment of tumor hypoxia in newly diagnosed glioblastoma using 18f-fmiso pet and mri, *Clinical Cancer Research* 22, 5079–5086. https://clincancerres.aacrjournals.org/content/22/20/5079,   doi:10.1158/1078-0432.CCR-15-2529.

[31] Ratai E. M. and Zhang Z. and Fink J. and Muzi M. and Hanna L. and Greco E. and et al. ACRIN 6684: Multicenter, phase II assessment of tumor hypoxia in newly diagnosed glioblastoma using magnetic resonance spectroscopy. *PLoS ONE*, 13(6): e0198548. doi: 10.1371/journal.pone.0198548.

[32] Zuley, M.L., Jarosz, R., Kirk, S., Lee, Y., Colen, R., Garcia, K., et al., 2016. Radiology data from the cancer genome atlas head-neck squamous cell carcinoma [tcga-hnsc] collection. URL: http://doi.org/10.7937/K9/TCIA.2016.LXKQ47MS, doi: 10.7937/K9/TCIA.2016.LXKQ47MS.

[33] Clark, K., Vendt, B., Smith, K., et al., 2013. The cancer imaging archive (tcia): Maintaining and operating a public information repository. *J Digt Imaging*, 26, 1045–1057. doi: 10.1007/s10278-013-9622-7.

Chapter 6

The Role of Transient Vibration and Skull Properties on Concussion: A Computational Analysis

6.1 Introduction

Concussion is a traumatic brain injury (TBI) that affects brain function [1–3]. Symptoms are usually temporary but can include headaches, difficulty concentrating, loss of memory, drowsiness, dizziness, sensitivity to light, and feeling slowed down [4,5]. The brain is protected from everyday impacts (either direct or indirect) to the head and neck by cerebrospinal fluid inside the skull. Despite the protection, this impact can cause the brain to move back and forth against the inner walls of the skull [6]. As the brain moves back and forth along the inner surface of the skull walls, the brain slides over itself at various speeds due to differences in brain density [7]. This shearing effect leads to deformation of the brain structure and stretching to the underlying axon fiber bundles. This effect is known as axonal shearing or diffuse axonal injury (DAI) when it is severe [8]. The axon stretching can lead to a breakdown of communication among neurons in the brain. This causes gaps to form in the cellular membrane of the axon and leakage of the ions needed for transmitting action potentials.

In in-vivo head impact conditions, [9] investigated the dynamic characteristics of the complex nonlinear brain-skull system of 189 head impact kinematics (187 noninjury and 2 injury

collisions). The author observed that brain's displacement modes were amplified around 30 Hz. It was also shown that combining modes with frequencies up to 33 Hz may contain 75% of the total brain displacement energy. [9] concluded that multimodal behavior occurs in greater strain areas as higher skull kinematics produce larger tissue deformations, amplify the effects of nonlinear material behavior and geometric nonlinearities, and increase the resultant oscillations in the frequency domain. Furthermore, multimodal brain activity creates local dynamics that may stimulate potentially severe oscillations in deep brain regions.

The skull has several resonant frequencies that are excited when struck [10–15]. The resonant frequencies correspond to specific deformation modes that are determined by the geometric shape of the skull and its distribution of bone thickness. Skull vibration applied in the z-axis (anterior-posterior direction) using a piezo-electrically-driven biting bar device induced a distinct shear wage propagation pattern in a ferret brain [18]. The principal displacement component of this shear wave propagated along the z-axis with a maximum magnitude of approximately 11.2 μm for 400 Hz and 7.4 μm for 600 Hz. With a lower amplitude and shorter wavelength, the displacement field at 800 Hz is qualitatively similar. The strain amplitudes at 800 Hz were found to be significantly lower than at 400 Hz.

The human brain has a maximum mechanical impedance at around 450 Hz, and vibrations near this frequency may cause a much larger propagating vibration into the brain tissue causing axon damage [16, 17]. The propagating vibrations can constructively interfere when they meet in different regions of the brain. Axons crossing these junctions experience tremendous shearing forces causing them to stretch and tear from the cell body, causing the concussive injury. After an impact, vibration resonance of the skull might bruise the underlying cortex. The skull deforms and vibrates, like a bell [19] for 3 to 5 milliseconds [20], pushing inward on the cortex [21]. As the vibrations propagate into the brain, concussion may occur deep within the brain tissue or closer to the periphery.

Given the maximum mechanical impedance of the brain tissue and the skull resonance frequencies, this study investigates the correlation between skull resonance frequencies and skull

geometry, thickness, and material properties using finite element (FE) modelling techniques. To accomplish this task, 45 segmented skulls were extracted from CT scans. Initially, one segmented skull is selected as a gold standard for identifying its diploe/cortical bone and was used to develop a specific material mapping algorithm that simplifies the material model. This was done to reduce the pre-processing time of extracting the trabecular and cortical bone, which is a time-consuming task. Once a material model is selected, the natural frequencies of all skulls were determined, and different densities and stiffness were simulated to evaluate how these parameters influence natural frequencies. Skulls with a natural frequency of 450±50 Hz were selected as a threshold to assess their thickness and conduct a statistical shape analysis. Principal Component Analysis (PCA), Independent Component Analysis (ICA), t-distributed Stochastic Neighbor Embedding (t-SNE), and *k*-means are employed in this statistical shape analysis to determine the relationship between shape and natural frequencies (Figure 6.1).
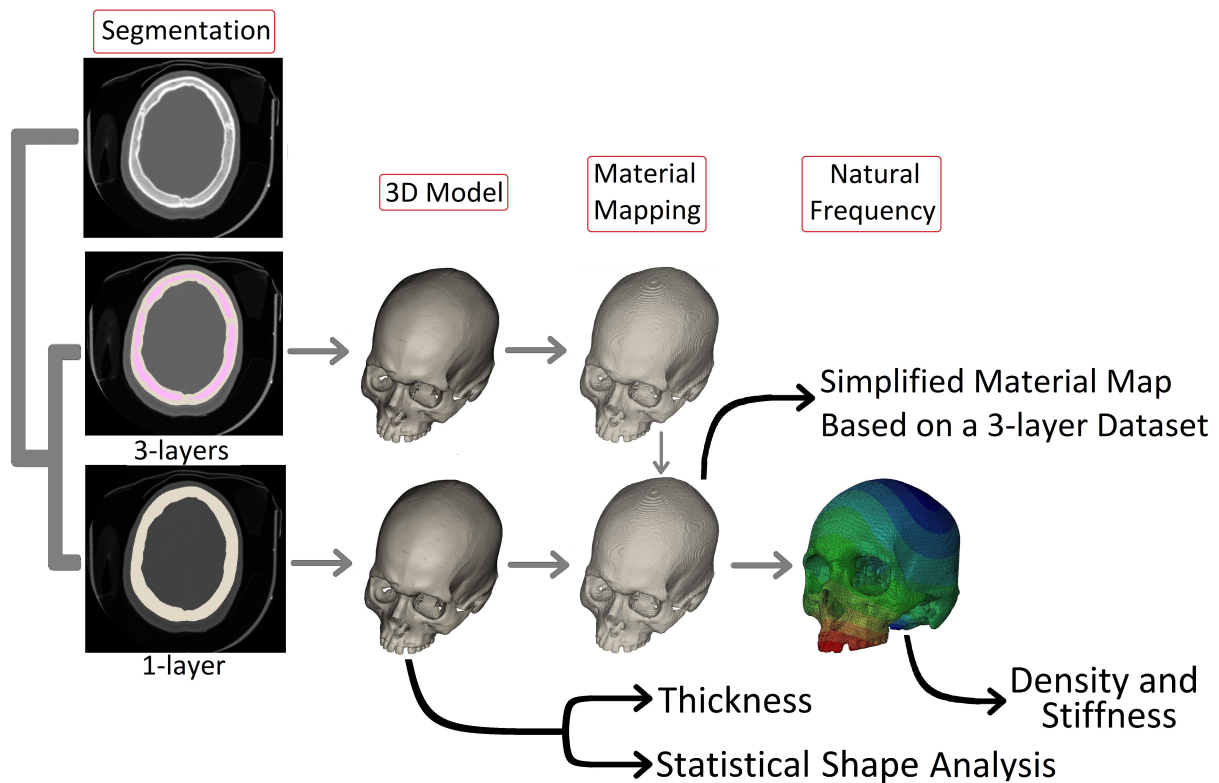

Figure 6.1: Overview of the proposed method.

6.2   Materials and Methods

6.2.1   Dataset and Preprocessing

The dataset used to validate the proposed method was acquired from the qure.ai CQ500 dataset [26]. From this dataset, 45 CT volumetric images with pixel size (width×height×depth) of 512×512 and varied depths, and spatial resolution of 0.625 mm were selected. Slice thickness, as previously reported [27], has the greatest influence on the accuracy of 3D-CT volume estimations. Thinner slices provide greater image quality details and less partial volume averaging over a series. [28]. The authors [29] quantified the tolerance of routinely used thicknesses greater than 0.625 mm and indicated that 1.25 mm slices are similarly viable for 3D rendering, and that thinner slices (1 mm) may diminish surface area measurement errors.

To create the models, CT images were imported into Mimics Research 20.0 Software (Materialise, Leuven, Belgium). First, individual thresholding with manual corrections was applied for each of the 3D volumetric CT images. Then, region growing was applied to create the 3D part model. This process allowed for the creation of the standard tesselation language (STL) file format which was converted into a matrix using voxelization method [30] to process the file in MATLAB R2019b software (Mathworks, USA). The 3D part model was then used to create a 3D mesh model with voxel grouping X, Y, and Z resolutions of 3, smoothing iteration count of 3, smoothing factor of 0.5, and meshing using hexahedral elements. The mesh was then exported as an ASCII data file (INP). The mandible was not included in the model since some images showed distortion, cuts, and a loss of mandible.

One CT image was selected as the gold standard image since it had no deformities or incompletion in the image. As such, a three-layered gold standard model was created. This model had the cortical and diploe bones segmented separately and resulted in two 3D mesh models with 227831 elements for cortical bone and 85182 elements for diploe. Single-layer models were also created for all 45 3D mesh models, with an average number of elements of 141466 (max - 241383 and min - 85927).

6.2.2   Material Properties, Modal Analysis, Steady-State Dynamics Analysis

The skull bone is a complex material consisting of a 3-layered structure [31–33]. The human skull bone has a 3-layered sandwich structure, where the middle layer can be considered as a liquid filled spongy bone (diploe), whereas the external and internal layers of this bone consist of compressed and dense bone (cortical bone).

For the FE simulation, a single-layered material model with unique material properties between models will be used. This process has been commonly adopted to help reduce pre-processing time. [22–25]. To obtain an accurate representation of a single-layered material skull model for the FE simulation, the gold standard model was used for a deep material property study. The gold standard model was replicated and two material models were explored: 1 - the three-layered material model were the material properties were averaged from the work of [33–43], and 2 - the single-layered material model were the unique material properties were used (Table 6.1). Both models were assessed on Abaqus 20 software (Simulia, Providence, RI, USA). As adopted and determined by numerical experiments in [22, 23, 44, 45], this study also considered the skulls as an isotropic homogeneous linear elastic material.

For the three-layered model, the INP files were imported into a single Abaqus model with two different parts containing the cortical bone and dipole. Two different material properties were created and assigned to cortical bone (outer and inner bone) and diploe (Table 6.1). Tie constraints were created where the cortical surface was assigned as master and diploe as slave. For the single-layered model, the same steps were followed with no constraint between the bone region (as there exist only one model).

For both material models, boundary conditions were created using symmetry/antisymmetry/encaste type and were applied on the base of the skull at the rear of the cranium, which is formed by occipital bone and rely on first vertebra of the spinal cord (Atlas - C1) (Figure 6.2). This location was fixed from all displacements and rotations. A step linear perturbation frequency was created to identify the natural frequencies from the range of 0 Hz to 900 Hz. No external load

or damping were included in this simulation. The simulation was performed on a Intel i7-9700 (3.0 GHz) processor with 64 GB of ram and took around 4 minutes to complete.
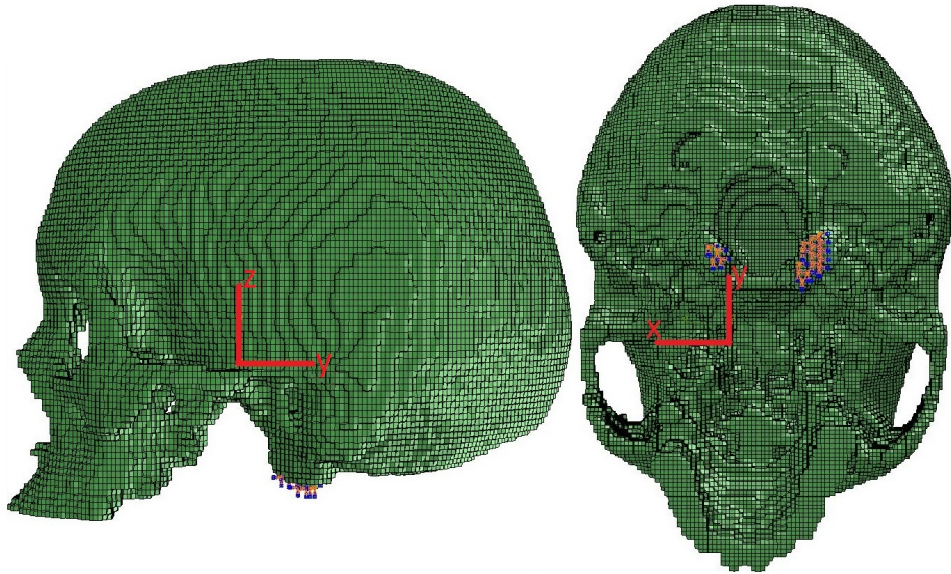

Figure 6.2: Fixed boundary condition highlighted at the rear of the cranium relying on the Atlas (C1).

From this deep material study, the 2 material models share a close natural frequency, with a maximum variance of ±34 Hz. As such, the single layer material model was determined to be acceptable and was used for the remaining 44 models (Table 6.1).

Table 6.1: Material Properties of the single and three layers skull.

Skull Model	Density (Kg/m^3)	Young's Modulus (GPa)	Poisson's Ratio
Single Layer	2000	4	0.182
Three Layers - Cortical Bone	2197	11.45	0.278
Three Layers - Diploe	1522.33	0.5842	0.182

6.2.3   Statistical Shape Analysis

Skeletal structure is determined by many factors such as subject's genetic background, environment, and exercise level in vivo. To analyze the differences of vibrational excitation between subjects, a statistical shape model analysis was performed. To evaluate the differ-

ences between thickness and natural frequencies, all 45 3D part models were imported into 3-matic Research 12.0 (Materialise, Leuven, Belgium) as STL files and one image was chosen as the gold standard for alignment purpose as this image shows a great alignment in the frontal plane. Then, taking the gold standard as a reference, cephalometric landmarks (1-Crista Galli, 2,3-Frontozygomatic Suture, 4,5-Orbitale, and 6-Subspinale) were selected in both the gold standard and in the desired image to be aligned (Figure 6.3). After the alignment, the wall thickness of all images were analyzed for thickness shape analysis.
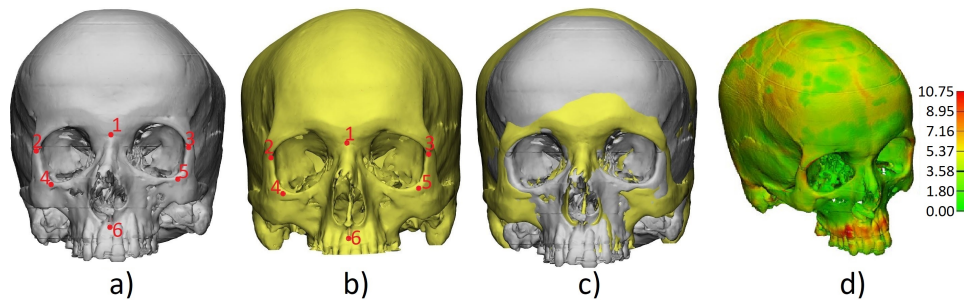

Figure 6.3: a) gold standard and b) designed image to be aligned with cephalometric landmark highlighted, c) the result of the alignment, and d) result for thickness analysis.

All final aligned STL files were imported into MATLAB and using a voxelization method [30], STL files were converted into matrices for easier processing in MATLAB. All matrices were then vectorized so PCA, ICA, and t-SNE algorithms could extract the principal components (PCs), the mixing matrix (**A**), and the components of t-SNE, respectively.

PCA, first introduced by the authors of [46] and then independently developed by the authors of [47], is a widely used technique for statistical shape analysis in the biomedical field [48–52]. It is an orthogonal linear transformation technique that maintains the variation present in a large number of associated variables of a data whilst reducing the dimensionality by transforming the data into a new coordinate system of uncorrelated variables or principal components (PCs) [53]. PCs are ordered by the eigenvalues, which reflect the relative importance such that the biggest variance is on the first coordinate or the first principal component, the second greatest variance on the second coordinate and so forth. The direction of maximum variance within the data is usually expressed by the first principal component. In the context of

anatomical data, this component usually defines size disparities between subjects. In addition, since it is orthogonal to the first principal component, the second principal component reflects the direction that maximizes the data variance. Therefore, PCA is a descriptive technique that enables a systematic investigation of shape variance analysis of a model [54].

PCA produces vectors that describe global shape variations while removing more local differences. In other scenarios, the correct contours cannot be identified locally because the model does not provide local deformations. For statistical shape purposes, information on the locality and amplitude of shape variations is extremely important. An alternative technique for statistical shape analysis is ICA [55], which has demonstrated advantages in terms of obtaining meaningfully unique features and utilizing these features for subsequent classification purposes, making ICA a feasible option to PCA in some scenarios depending upon the application [56, 57]. In order to present the inherent patterns in a linear combination, ICA attempts to linearly divide a random vector into a linear combination of several statistically independent components (ICs) [58]. Given a set of random variables $\{x_1(t), x_2(t), ..., x_n(t)\}$, where t is time or the sample index,it is assumed that is composed of a linear combination of ICs $\{s_1(t), s_2(t), ..., s_n(t)\}$:

$$\mathbf{x} = (x_1(t), x_2(t), ..., x_n(t))^T = \mathbf{A}(s_1(t), s_2(t), ..., s_n(t))^T = \mathbf{As}, \tag{6.1}$$

where $\mathbf{A}$ is an unknown mixture matrix $\mathbf{A}$ $R^{n \times n}$. The ICA algorithm chosen was the FastICA [58], where its mixture matrix $\mathbf{A}$, which contain the main feature vectors, was used to perform the statistical shape analysis.

t-SNE [59], an enhanced version of the SNE algorithm [60], is a statistical approach for transforming high-dimensional observations into a matrix of pairwise correlations. t-SNE reveals local and global structures of the high-dimensional data. t-SNE selects two similarity measures across data points, one for high-dimensional data and one for 2-dimensional embedding. The method then seeks to produce a 2-dimensional embedding that minimizes the

probability distribution between the vector of similarities between data points in the original dataset and the similarities between data points in the embedding. This is a non-convex optimization problem and t-SNE computes a reasonable solution using gradient descent with random initialization. In this work, euclidean distance is the metric adopted.

6.3 Results

6.3.1 Modal Analysis - Natural Frequency

The natural frequencies of all 45 3D mesh models were acquired using the same procedure mentioned in section 2.2 using Abaqus software. The result from 400 Hz to 900 Hz are presented in Figure 6.4. In addition, Figure 6.5 shows the effective mass and Figure 6.6 the participation factor represented in *x, y,* and *z* direction.
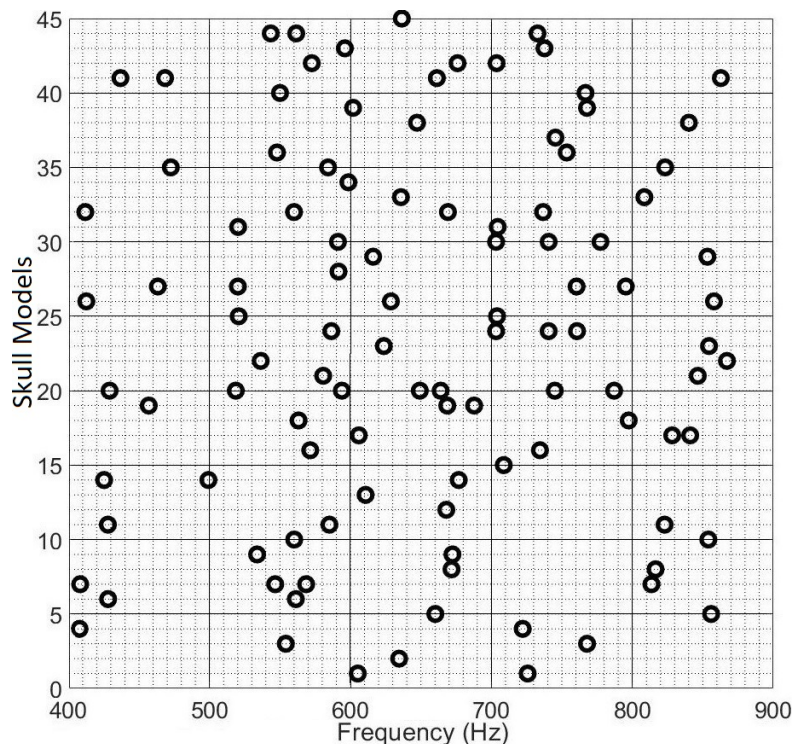

Figure 6.4: Frequency versus skull models.

Natural frequencies between 400 Hz and 500 Hz were observed in 12 skulls, whereas natu-
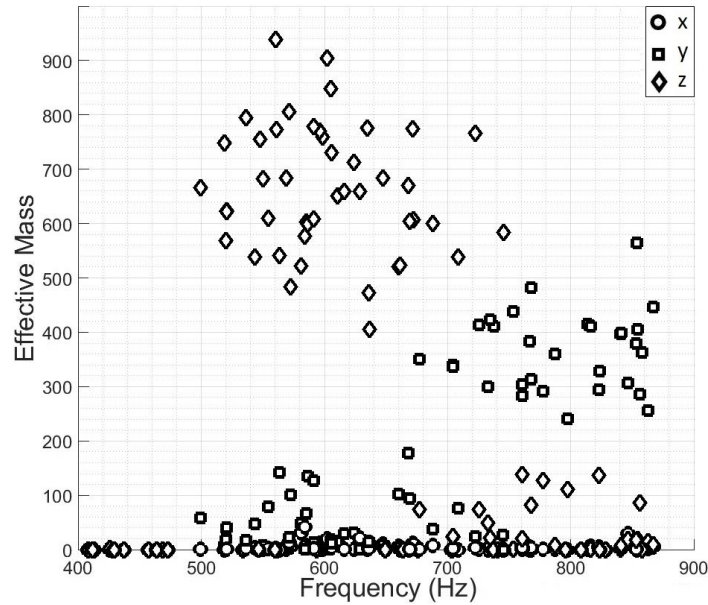
Figure 6.5: Frequency versus effective mass.
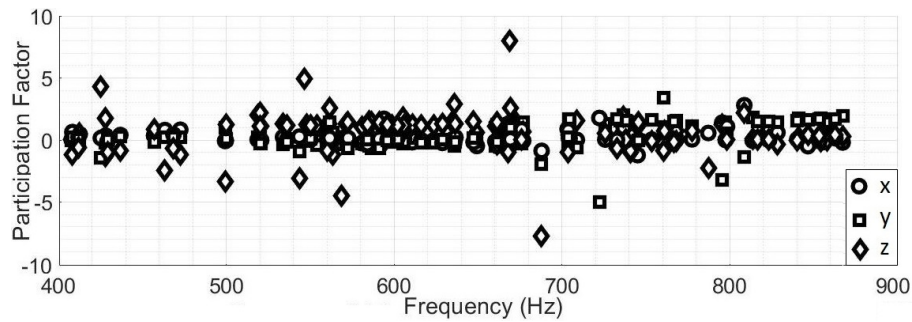
Figure 6.6: Frequency versus participation factor.

ral frequencies over 500 Hz were present in the remaining skulls. These 12 skulls were chosen

for a further test as gold standards for determining what causes the natural frequency withing

the range of 400 Hz to 500 Hz.

To analyze the skull displacement at each frequency from 400 Hz to 900 Hz, a steady-state

dynamics analysis was performed. Following the modal analysis, the steady-state dynamic step

was incorporated without structural damping and a gravity load of 9.81 kg/m^2 in the Z negative

direction was applied, based on the coordinate system defined in Figure 6.2. To analysis the

displacement response in the 12 skulls with natural frequencies between 400 Hz to 500 Hz, 5

points were selected on the skull surface at the Frontal, Occipital, Parietal, and Temporal bones.

As the temporal bone is composed of 2 bones located on each side of the skull, one point was selected for each of the temporal bones and was defined as left (X-) and right (X+). Figure 6.7 shows the displacement at each frequency from 400 to 900 Hz and Figure 6.8 presents a box chart describing the median of each displacement in the x, y, and z axes for all investigated bones from the range 400 Hz to 500 Hz.
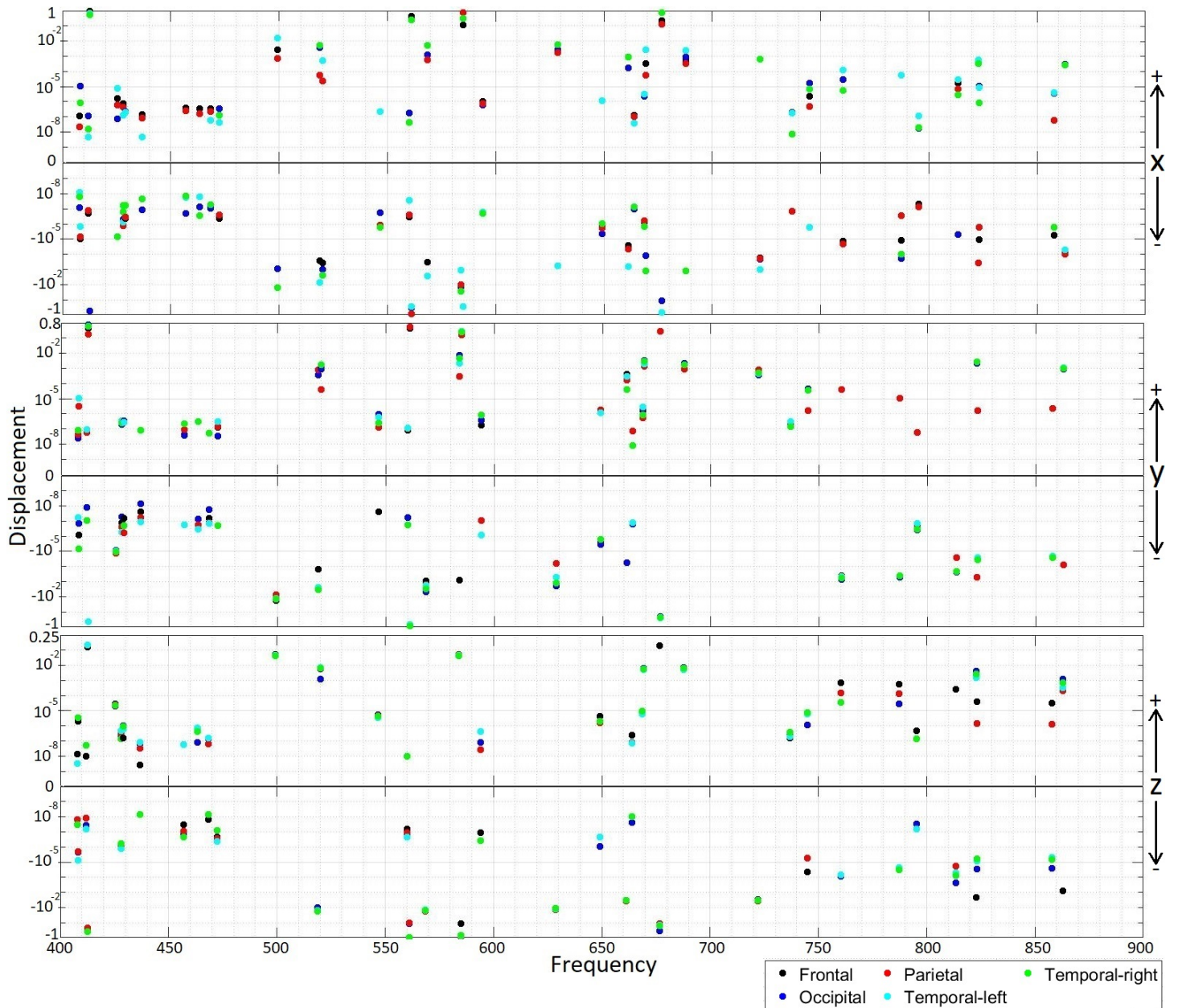
Figure 6.7: Displacement versus frequency for the 12 skulls with natural frequency ranging from 400 Hz to 900 Hz.
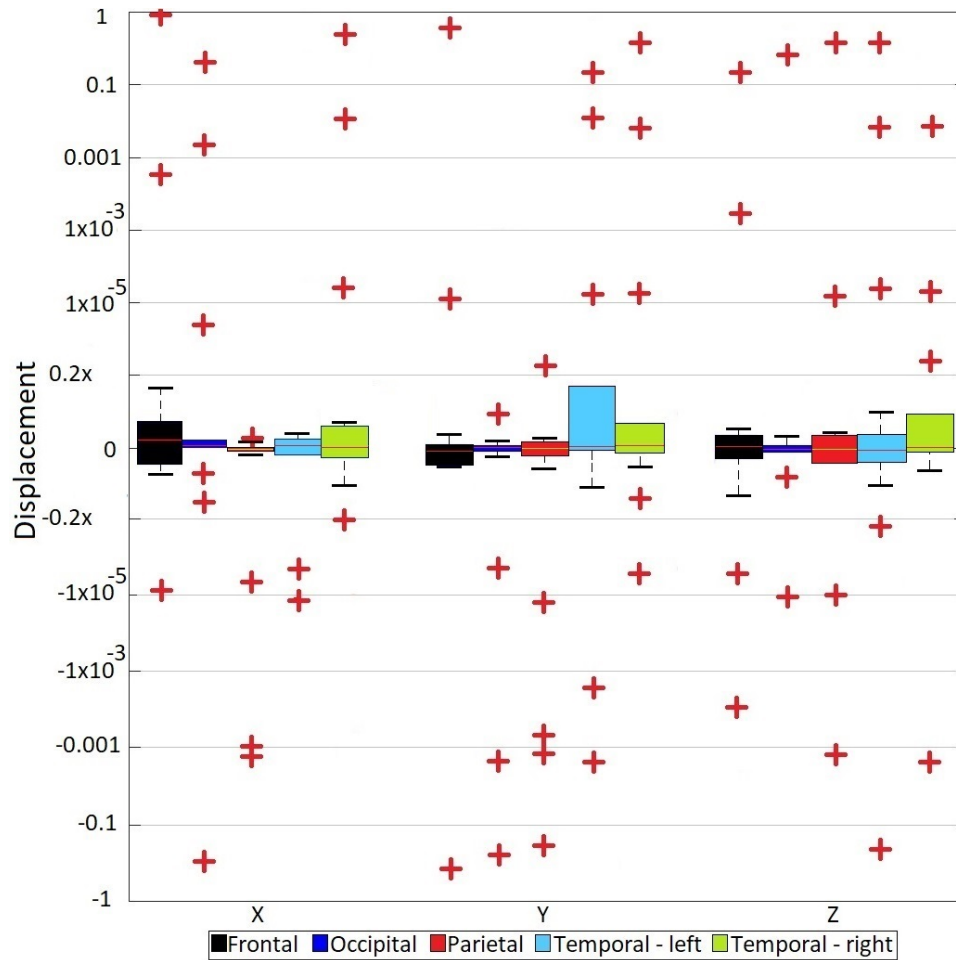
Figure 6.8: The median displacement in x, y, and z, of the 12 skulls in the frequency range of 400 Hz to 500 Hz, and the red crosses represent the outliers.

6.3.2 Density and Stiffness

The natural frequency is highly influenced by material properties and boundary conditions [61]. To better understand the behavior of different densities and stiffness on natural frequency, five 3D mesh models were selected for a modal analysis test where density and Young's Modulus were varied. As all models presented the same behavior when density and Young's modulus were varied, one 3D mesh model was selected to display the common alteration behavior. Therefore, for this 3D mesh model, density was varied from 1000 to 2000 kg/m^3 (Figure 6.9) and Young's Modulus from the range 1 to 10 GPa (Figure 6.10).
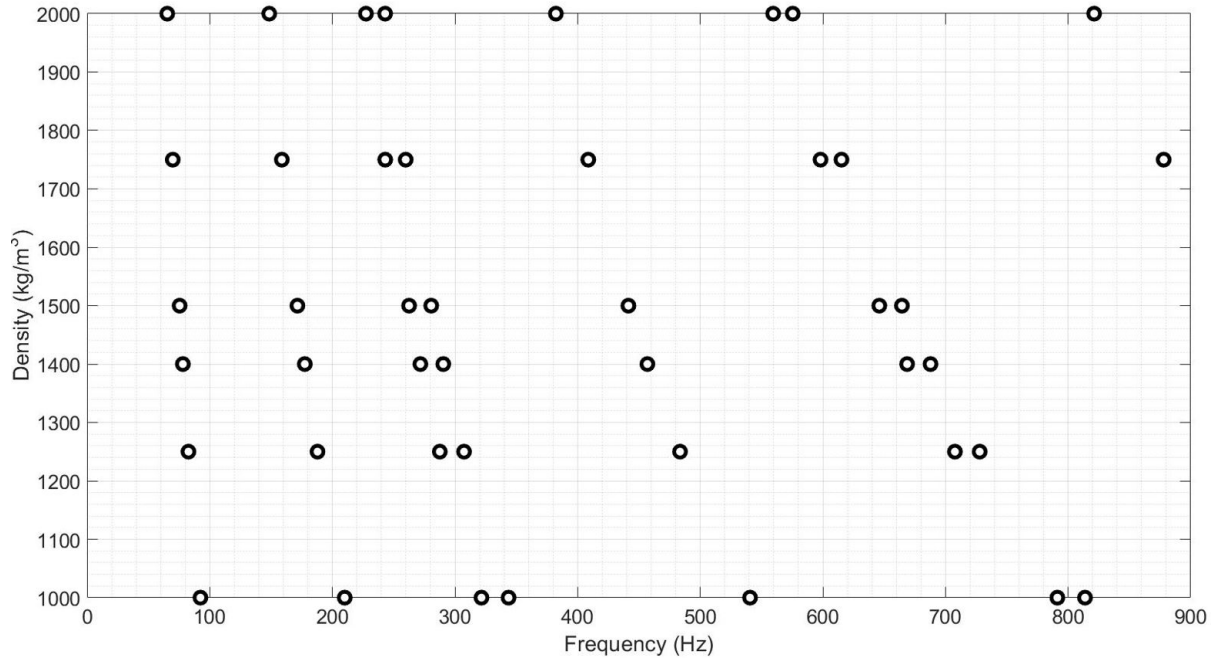
Figure 6.9: Frequency versus different densities.

Figure 6.10: Frequency versus different Young's Modulus.

6.3.3 Shape Analysis

The thickness of the 12 skulls with natural frequencies near 450±50 Hz were compared to the thickness of the remaining skulls in 3-matic software, as explained in section 2.3 (Figure 6.11).
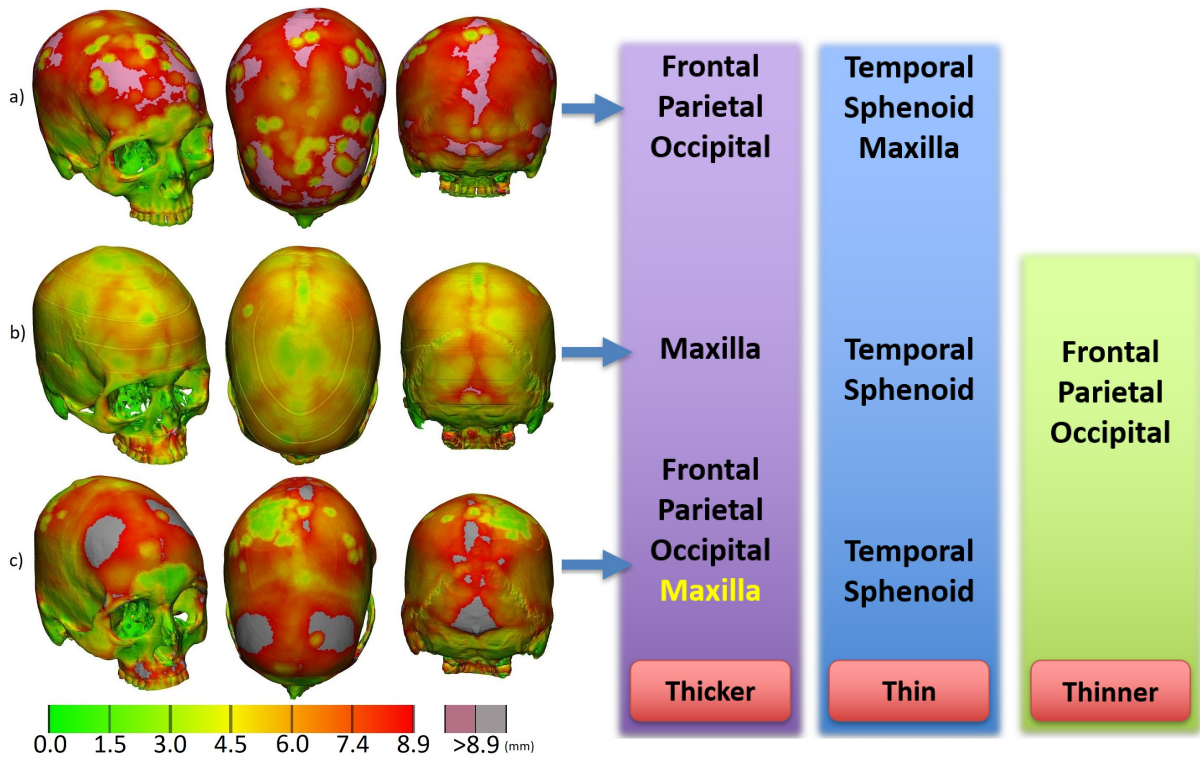
Figure 6.11: Example of skulls with a) natural frequency near 450±50 Hz, b) natural frequency greater than 500 Hz, and c) sharing the same characteristics as skull with natural frequency near 450±50 Hz but the thicker maxilla bone and natural frequencies near 600 Hz.
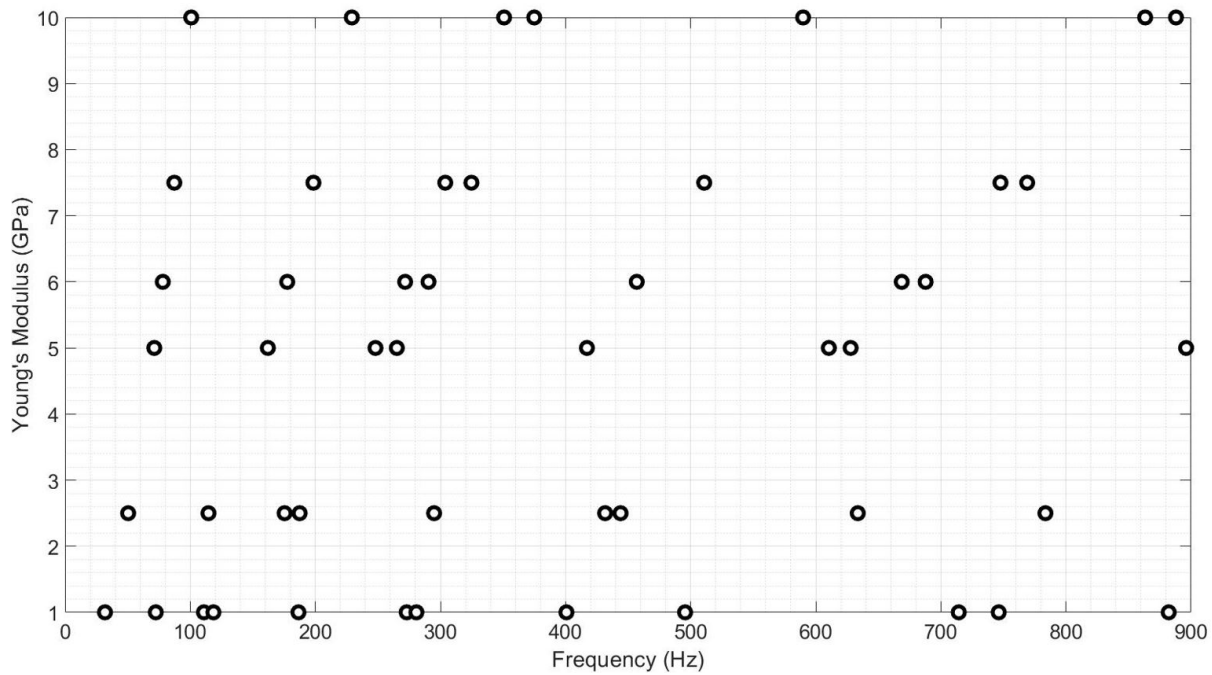
The STL skulls aligned in the thickness analysis were used to assess the statistical shape analysis in Matlab using PCA, ICA, t-SNE, and *k*-means (Figure 6.12). After being vectorized in Matlab, the 12 skulls with natural frequencies between 400 Hz to 500 Hz were used as gold standard to analyze any pattern changes in the skulls with natural frequencies greater then 500 Hz. Firstly, all skulls were presented to the PCA algorithm with cumulative variances shown in Figure 6.13. 90% and 95% of the variation can be explained by the first 21 and 29 principal components, respectively. PCs were computed with dimension 45×45, which represents the 45 samples and 45 characteristics. The PCs were then presented to the *k*-Means algorithm and Figure 6.14a, 6.14b, and 6.14c, shows the results for *k* equal to 2, 3, and 4, respectively, where the 12 skulls with natural frequency of 450±50 Hz are circled in red.

Similarly, the skulls were presented to the ICA algorithm and a 45×45 mixing matrix **A** were obtained. This matrix was then presented to the *k*-Means and Figure 6.15a, 6.15b, and

6.15c, show the results for *k* equal to 2, 3, and 4, respectively. In addition, the skulls were presented to the t-SNE algorithm, which produced a 45×2 matrix. This matrix was presented into the *k*-Means algorithm, and Figures 6.16a, 6.16b, and 6.16c show the results for *k* equal to 2, 3, and 4, respectively.

Figure 6.12: Statistical shape analysis processing.

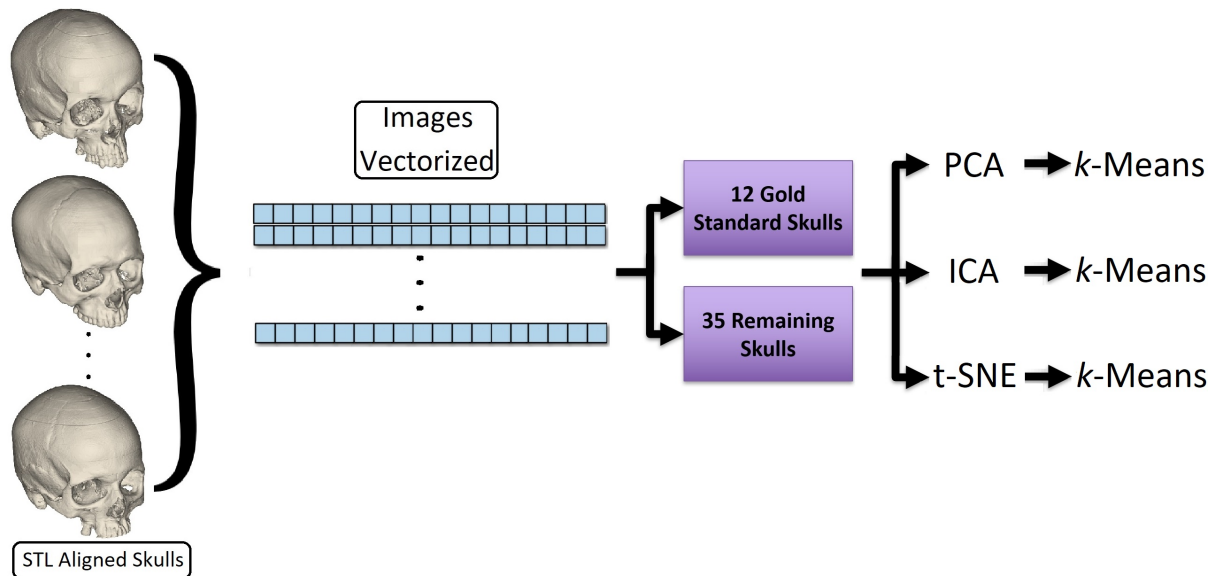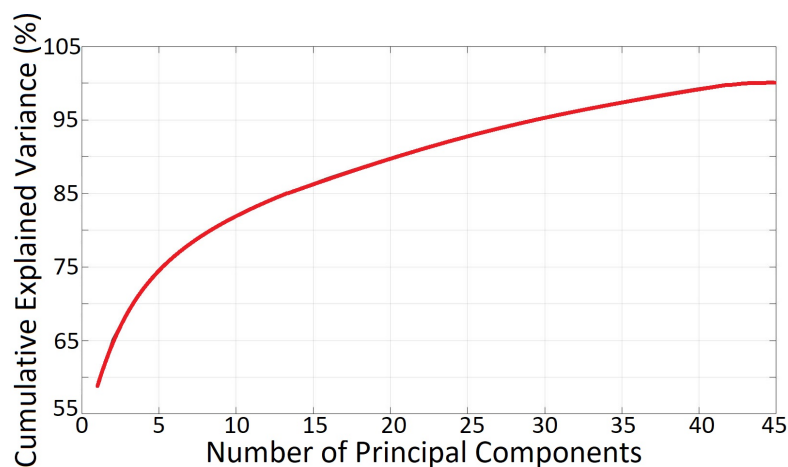Figure 6.13: Cumulative variance of the skulls by number principal components following a principal component analysis (PCA) to describe the different skeletal models.

6.4 Discussion

According to Figure 6.8, each bone can be interpreted individually based on its outliers. Considering the greatest outliers, Frontal and Occipital bones showed the most quantity of displacement in (x+, y+, y-,z+) and (x+, x-, y-, z) directions respectively, followed by the Temporal-left bone (y+, z+, z-), and Parietal (y-, z) and Temporal-right bones (x+, y+).

The modal effective mass is a metric that may be used to assess the relevance of a vibration mode. Mode shapes with high effective masses will produce high reaction forces at the base, whereas mode shapes with low effective masses will produce low reaction forces at the base. According to Figure 6.5, modes with natural frequencies greater than 500 Hz had experienced the highest effective mass, whereas modes with natural frequencies between 400 Hz to 500 Hz presented the lowest. In general, the z component showed the dominating effective mass between 500 Hz to 700 Hz, whereas the y component dominated in the range 700 Hz to 900 Hz.

Participation factor measures the amount of mass moving in a certain direction, which means a high value will excite the mode in that direction. From Figure 6.6, all modes have a similar participation factor, indicating that they all behave in a similar pattern when stimulated in different directions. Through investigation, the outliers found in Figures 6.5 and 6.6 were associated with a great displacement in the occipital bones, which is attributed to bone thickness in the region. This may be associated with subject specific factors such as ethnicity, age, gender, and external factors (such as sport, previous injury etc.).

From Figures 6.9 and 6.10, it can be observed that increasing density lowers the value of natural frequencies while increasing Young's Modulus raises the value of natural frequencies. Furthermore, it was observed that when density and Young's Modulus alter, higher frequencies exhibit more variance than lower frequencies.

As can be observed in Figure 6.11a, the frontal, parietal, and occipital bones of all 12 skulls were thicker, but the temporal, sphenoid, and maxilla bones were thin. The remaining skulls featured thinner frontal, parietal, and occipital bones, thin temporal and sphenoid bones, and

thicker maxilla bones compared to the 12 skulls (Figure 6.11b). There were few exceptions which shared a similar characteristic as the 12 skulls but featured thicker maxilla bone and presented natural frequencies of around 600 Hz (Figure 6.11c). This variance in certain bone thickness may indicate a pattern in how specific skulls present a tendency for producing natural frequencies of 450±50 Hz. However, due to the dificult to perform a bone thickness correlation analysis, further investigation is required to support this finding.

It is noticeable from Figures 6.14 and 6.15 that there is no correlation between shape and natural frequency. When performing ICA (Figure 6.15a) and t-SNE (Figurere 6.16a), it can be seen that there are at least two possible shape patterns, whereas when using PCA (Figurere 6.14a), there is no consistent pattern among skull shapes. PCA is generally used to reduce dimensionality into a smaller set of variables with less redundancy, as determined by data correlations, based exclusively on second order statistics, which lack information from higher order statistics. For ICA, which is a generalization of PCA, its goal is to separate data by transforming the input space into a basis that is maximally independent. ICA extracts hidden components from data by transforming a set of variables into a new maximally independent set, using a non-gaussianity measure to do so. Therefore, ICA is computed not just from the second order statistics, as in PCA, but also from higher order [62]. However, even using higher order statistics, ICA was unable to identify any pattern among the 12 skulls. In comparison to PCA, t-SNE is a non-linear dimensionality reduction technique that retains the local structure of high-dimensional data while presenting the data points [63]. However, even when employing local structure, t-SNE was unable to establish a pattern among the 12 skulls. To summarize, in all three scenarios (k=2, 3, and 4), PCA, ICA, and t-SNE were unable to establish a shape pattern for the 12 skulls.

The FE method presented in this study has an advantage over the experimental approach in terms of repeatability. Though this experiment, it was shown that skull properties affect the vibration within the skull. Nevertheless, since this is a numerical approximation, there are certainly some limitations to the current study. Simplifications were made in the model by us-

ing a single-layered material approach rather than segmenting a 3-layered material map which provides a more accurate approximation of real-life behavior. However, based on the deep material study performed, the single-layered model was an acceptable assumption. CQ500 database does not provide subject specifications, such as age and gender, which would be helpful to study the differences during the thickness and shape analysis. Furthermore, the experiment was conducted by only using computational methods and was not validated by in-vivo or in-vitro tests. Real models are difficult to evaluate due to the specimen limitations, which are frequently limited due to ethical, privacy considerations, and regulations. However, the findings of articles that used in-vivo or in-vitro simulations were comparable to those in this paper [22, 23, 64, 65]. Furthermore, during the facial alignment, cephalometric landmarks were selected as gold standards and if a unique image was misalignment, an error may occur in the PCA, ICA, t-SNE, and k-Means calculation resulting in an incorrect statistical shape analysis interpretation. However, this is unlikely as all procedures were cautiously carried out and 3-matic is a well-established software.

The model will be further improved in future investigations by using a comprehensive head model (including the scalp, meninges, and white and gray matter). Material properties and a validation experiment, with subject specific modelling, will be conducted to allow a more accurate analysis. The bone and soft tissues can be modeled in Mimic with an impact that induces a skull vibration near 450 Hz. A comprehensive theory on how skull vibration propagates into the brain tissue can be developed.

6.5   Conclusion

The present study aims to identify the correlation between natural frequency versus density, stiffness, thickness, and shape of skulls to better understand the mechanism that triggers a concussive injury. It was found that density and stiffness have a strong correlation with natural frequencies values, where increasing density and Young's Modulus reduce and increase

the natural frequencies, respectively. It was also identified that some skulls with natural frequency of 450±50 Hz kept a pattern regarding their thickness in certain cranial and part of the viscerocranium bones whereas skull with natural frequency over 500 Hz presented a different pattern. Statistical shape analysis was conducted using PCA, ICA, and *k*-means and no relationship was observed between skulls with natural frequency around 450±50 Hz with skull shape. In conclusion, our finding support the idea that density, stiffness, and thickness affect natural frequencies in skulls and may contribute to concussive injury.
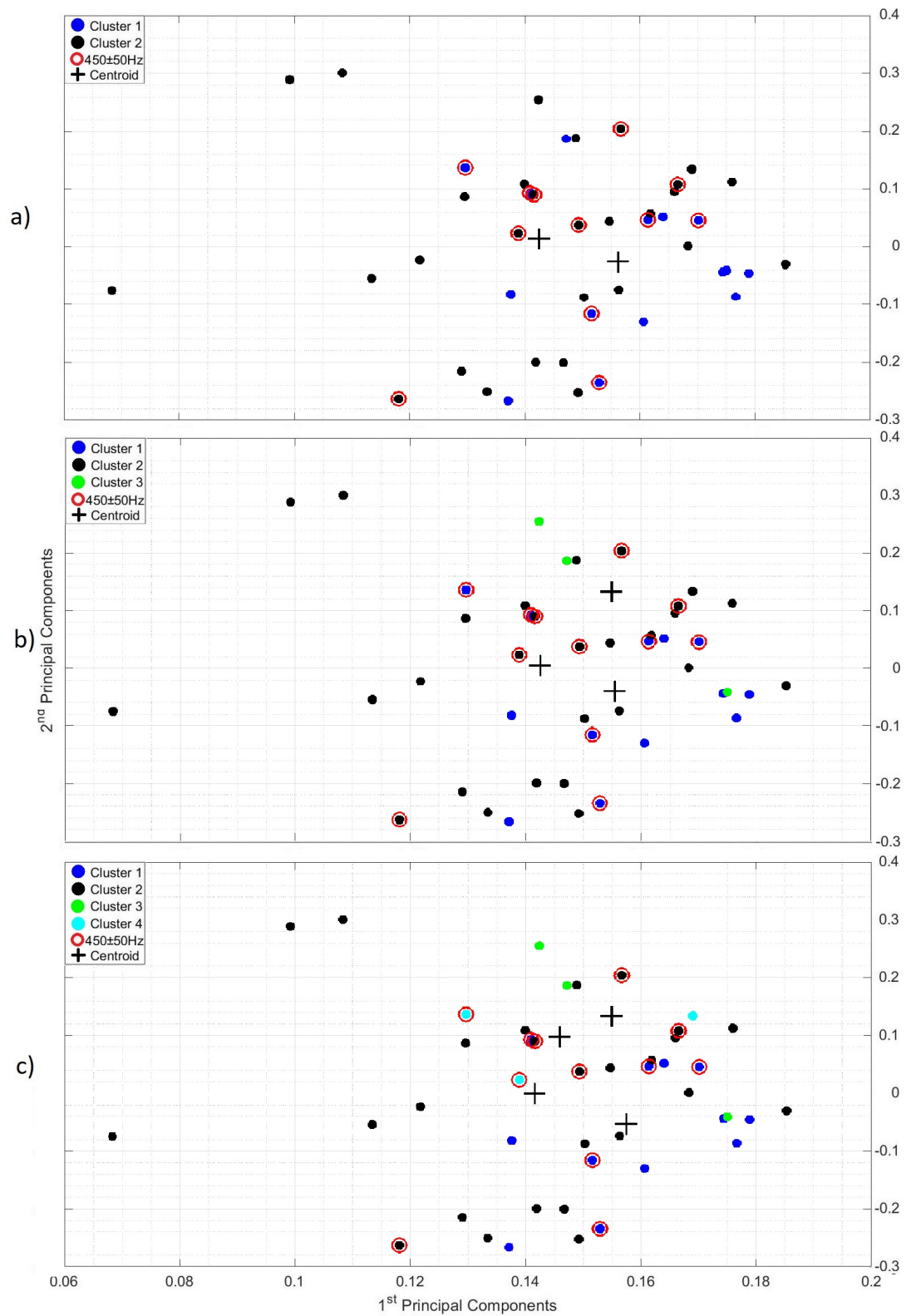
Figure 6.14: First versus second principal components (PCs) for *k* equal to a) 2, b) 3, and c) 4.
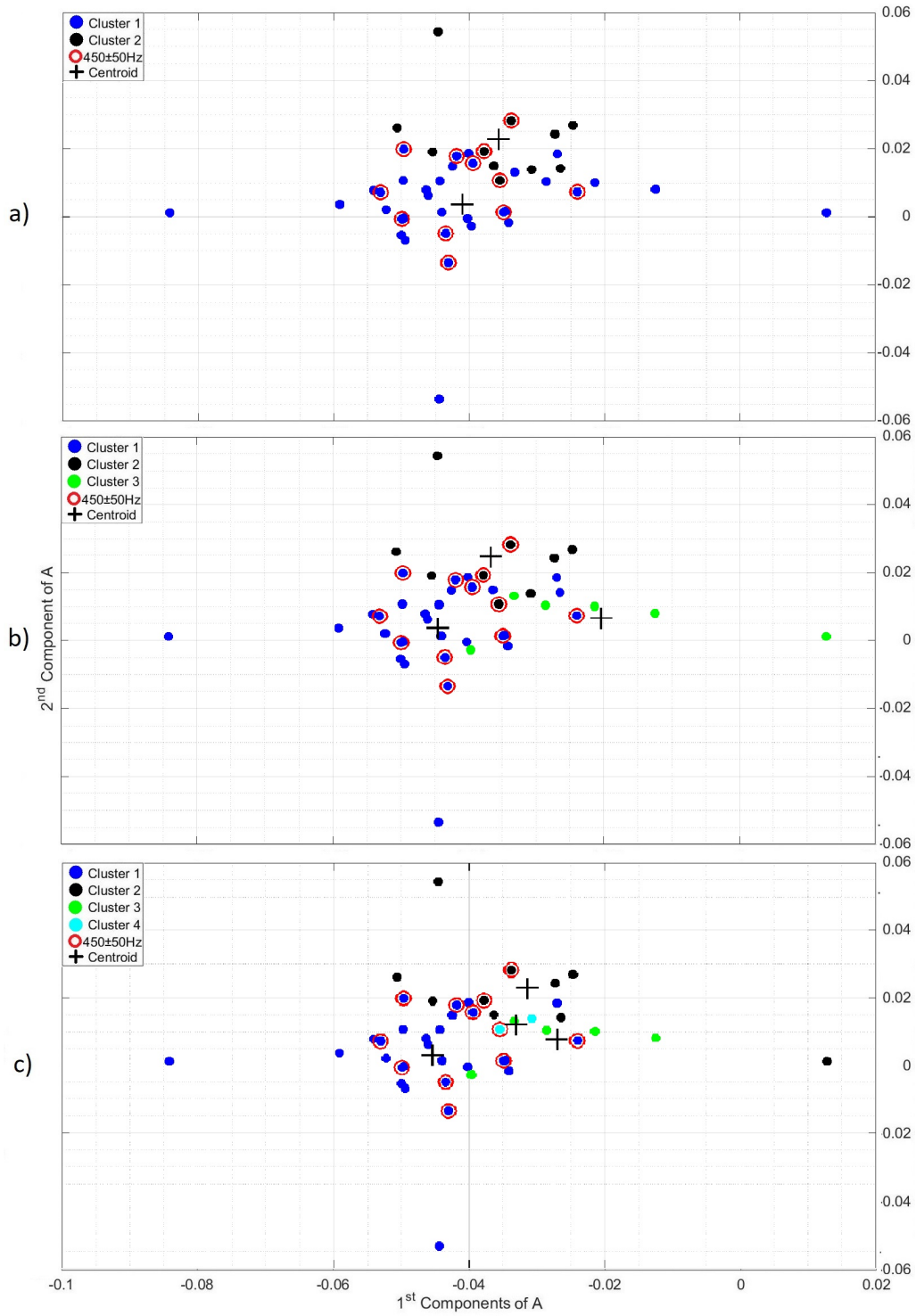
Figure 6.15: First versus second columns of mixing matrix **A** for *k* equal to a) 2, b) 3, and c) 4.
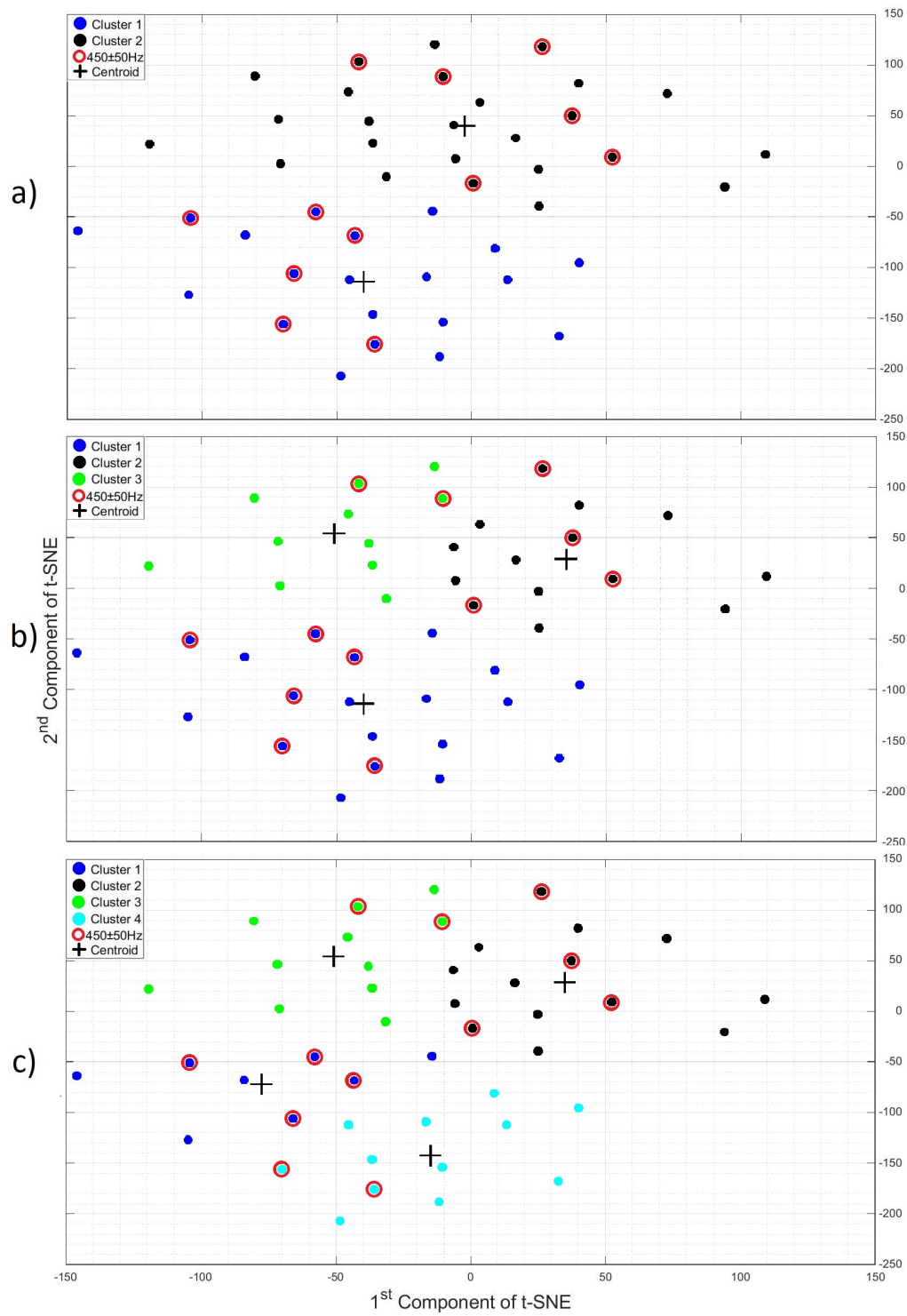
Figure 6.16: First versus second t-SNE output components for *k* equal to a) 2, b) 3, and c) 4.

Bibliography

[1] Tator, C. H. Concussions and their consequences: current diagnosis, management and prevention, *Canadian Medical Association Journal* 2013, 185(11), 975-979. doi: 10.1503/cmaj.120039

[2] Mullally, W. J. Concussion, *The American Journal of Medicine* 2017, 130(8), 885-892. doi: 10.1016/j.amjmed.2017.04.016

[3] Meaney, D. F., Smith, D. H. Biomechanics of Concussion, *Clinics in Sports Medicine* 2011, 30(1), 19-31. doi: 10.1016/j.csm.2010.08.009

[4] Churchill, N. W., Hutchison, M. G., Graham, S. J.et al. Baseline vs. cross-sectional MRI of concussion: distinct brain patterns in white matter and cerebral blood flow, *Sci Rep* 2020, 10, 1643. doi: 10.1038/s41598-020-58073-9

[5] Anderson, E. D., Giudice, J. S., Wu, T., Panzer, M. B., Meaney, D. F. Predicting Concussion Outcome by Integrating Finite Element Modeling and Network Analysis, *Front. Bioeng. Biotechnol.* 2020, 8:309. doi: 10.3389/fbioe.2020.00309

[6] Luo, Y., Li, Z., Chen, H., Finite-element study of cerebrospinal fluid in mitigating closed head injuries. Proceedings of the Institution of Mechanical Engineers, Part H: *Journal of Engineering in Medicine* 2012, 226(7), 499-509. doi:10.1177/0954411912445729

[7] Iverson, G., L., Lange, R., T., Mild Traumatic Brain Injury. In: Schoenberg M., Scott J. (eds) The Little Black Book of Neuropsychology, Springer 2011, Boston, MA. doi: 10.1007/978-0-387-76978-3_22

[8] Vieira R. C. A., Paiva W. S., de Oliveira D. V., Teixeira M. J., de Andrade A. F., Sousa R. M. C., Diffuse Axonal Injury: Epidemiology, Outcome and Associated Risk Factors, *Front. Neurol.* 2016, 7:178. doi: 10.3389/fneur.2016.00178

[9] Laksari, K., Kurt, M., Babaee, H., Kleiven, S., Camarillo, D., Mechanistic Insights into Human Brain Impact Dynamics through Modal Analysis, *Phys. Rev. Lett.* 2018, 120(13). doi:10.1103/PhysRevLett.120.138101

[10] Khalil, T. B., Viano, D. C., Smith, D. L. Experimental analysis of the vibrational characteristics of the human skull, *Journal of Sound and Vibration* 1979, 63(3), 351-376. doi: 10.1016/0022-460X(79)90679-5

[11] Håkansson, B., Brandt, A., Carlsson, P. Resonance frequencies of the human skull in vivo, *The Journal of the Acoustical Society of America* 1994, 95, 1474-1481. doi: 10.1121/1.408535

[12] Charalambopoulos, A., Fotiadis, D. I., Massalas, C. V. Free vibrations of the viscoelastic human skull, *International Journal of Engineering Science* 1998, 36(5-6), doi: 10.1016/S0020-7225(97)00087-6.

[13] Baroudi, A. E., Razafimahery, F., Rakotomanana-Ravelonarivo, L. Study of a spherical head model. Analytical and numerical solutions in fluid–structure interaction approach, *International Journal of Engineering Science* 2012, 51, 1-13, doi: 10.1016/j.ijengsci.2011.11.007.

[14] Baroudi, A. E., Razafimahery, F., Rakotomanana-Ravelonarivo, L. Three-dimensional modal analysis of an idealized human head including fluid–structure interaction effects, *Acta Mech* 2012, 223, 1899–1915. doi: 10.1007/s00707-012-0681-5

[15] Blandford, C., Neuert, M., Jenkyn, T. R. The dynamic properties of the human skull: the effect of impact location and impact energy on the vibrational response of the head, 2017 *CMBEC40 Conference*, Winnipeg MB, May 23-26, 2017.

[16] Chakroun, M., Ghozlen, H., Elloumi, I., Nicolle, S., Détermination du module d'élasticité de la matière cérébrale, *Comptes Rendus Physique* 2009, 10(2-3),pp. 236-241, doi: 10.1016/j.crhy.2009.03.010.

[17] Chakroun, M., Ghozlen, H. Mechanical Impedance of Cerebral Material, *Journal of Modern Physics* 2012, 3(3), 271-273. doi: 10.4236/jmp.2012.33037.

[18] Feng, Y., et al., Viscoelastic properties of the ferret brain measured in vivo at multiple frequencies by magnetic resonance elastography, *Journal of Biomechanics* 2013, 46(5), 863-870. doi: 10.1016/j.jbiomech.2012.12.024

[19] Moss, W. C., King, M. J., Blackman, E. G. Skull flexure from blast waves: a mechanism for brain injury with implications for helmet design, *Physical Review Letters* 2009; 103:108702. doi: arXiv:0809.3468v4

[20] Ganpule, S, Alai, A, Plougonven, E, Chandra, N. Mechanics of blast loading on the head models in the study of traumatic brain injury using experimental and computational approaches, *Biomech Model Mechanobiol* 2013; 12:511–531. doi: 10.1007/s10237-012-0421-8

[21] Martin, G. T. Acute brain trauma, *Ann R Coll Surg Engl* 2016; 98:6–10. doi:10.1308/rcsann.2016.0003

[22] Eslaminejad, A., Hosseini-Farid, M., Ziejewski, M., and Karami, G., Constitutive Properties Determination of Human Cranium by an Experimental–Computational Modal Analysis, *ASME. J. Vib. Acoust.* 2020, 142(1). doi: 10.1115/1.4045216

[23] Eslaminejad, A, Hosseini-Farid, M, Ramzanpour, M, Ziejewski, M, Karami, G., Determination of Mechanical Properties of Human Skull With Modal Analysis. Proceedings of the ASME 2018 International Mechanical Engineering Congress and Exposition, 3: *Biomedical and Biotechnology Engineering*. Pittsburgh, Pennsylvania, USA. November 9–15, 2018. doi: 10.1115/IMECE2018-88103

[24] Cao, L. B., Yang, B., Tang, Y., Sun, N., Tse, K.-M., Lee, H.-P., Modal and dynamic responses of the human head-neck complex for impact applications, *Journal of Vibroengineering* 2018, 18(7), 4743–4755. doi:10.21595/jve.2016.17457

[25] Yang B, Shi Z, Wang Q, et al., Frequency spectrum of the human head–neck to mechanical vibrations. Journal of Low Frequency Noise, *Vibration and Active Control* 2018, 37(3), 611-618. doi:10.1177/1461348417747179

[26] Chilamkurthy, S.; Ghosh, R.; Tanamala, S.; Biviji, M.; Campeau, N.G.; Venugopal, V.K.; Mahajan, V.; Rao, P.; Warier, P. Deep learning algorithms for detection of critical findings in head CT scans: A retrospective study, *Lancet* 2018, 392, 2388–2396. doi: 10.1016/S0140-6736(18)31645-3

[27] Ravenel, J. G., Leue, W. M., Nietert, P. J., Miller. J. V., Taylor. K. K., Silvestri. G. A., Pulmonary nodule volume: effects of reconstruction parameters on automated measurements–a phantom study, *Radiology* 2008, 247(2):400-8. doi: 10.1148/radiol.2472070868. PMID: 18430874; PMCID: PMC4148132

[28] Joo, I., Kim, S. H., Lee, J. Y., Lee, J. M., Han, J. K., Choi, B. I., Comparison of semiautomated and manual measurements for simulated hypo- and hyper-attenuating hepatic tumors on MDCT: effect of slice thickness and reconstruction increment on their accuracy, *Acad Radiol.* 2011, 18(5):626-33. doi: 10.1016/j.acra.2010.12.013. Epub 2011 Mar 9. PMID: 21393028

[29] Whyms, B. J., Vorperian, H. K., Gentry, L. R., Schimek, E. M., Bersu, E. T., Chung, M. K., The effect of computed tomographic scanner parameters and 3-dimensional volume rendering techniques on the accuracy of linear, angular, and volumetric measurements of the mandible, *Oral Surg Oral Med Oral Pathol Oral Radiol.* 2013, 115(5), 682-691. doi:10.1016/j.oooo.2013.02.008

[30] Patil, S.; Ravi, B. Voxel-based representation, display and thickness analysis of intricate shapes. *In Proceedings of the Ninth International Conference on Computer Aided Design and Computer Graphics* (CAD-CG'05), Hong Kong, China, 7–10 December 2005; p. 6. doi: 10.1109/CAD-CG.2005.86

[31] Motherway, J. A., Verschueren, P., Van der Perre, G., Vander Sloten, J., Gilchrist, M. D., The Mechanical Properties of Cranial Bone: The Effect of Loading Rate and Cranial Sampling Position, *J. Biomech.* 2009, 42(13), pp. 2129–2135. doi: 10.1016/j.jbiomech.2009.05.030

[32] Delille, R., Lesueur, D., Potier, P., Drazetic, P., Markiewicz, E., Experimental Study of the Bone Behaviour of the Human Skull Bone for the Development of a Physical Head Model, Int. J. Crashworthiness 2007, 12(2), pp. 101–108. doi: 10.1080/13588260701433081

[33] Cai, Z., Xia, Y., Bao, Z., Mao, H., Creating a human head finite element model using a multi-block approach for predicting skull response and brain pressure, *Computer Methods in Biomechanics and Biomedical Engineering* 2019, 22(2), 169-179. doi: 10.1080/10255842.2018.1541983

[34] Chang, Y., Kim, N., Stenfelt, S., The development of a whole-head human finite-element model for simulation of the transmission of bone-conducted sound, *The Journal of the Acoustical Society of America* 2016, 140, 1635. doi: 10.1121/1.4962443

[35] Fry, F., Barger, J., Acoustical properties of the human skull, *J. Acoust. Soc. Am.* 1978, 63(5), 1576–1590. doi: 10.1121/1.381852

[36] Peterson, J., Dechow, P. C., Material properties of the human cranial vault and zygoma, *Anatom. Rec. Pt. A: Disc. Mol. Cell. Evolut. Biol.* 2003, 274(1), 785–797. doi: 10.1002/ar.a.10096

[37] Jaasma, M. J., Bayraktar, H. H., Niebur, G. L., Keaveny, T. M., Biomechanical effects of intraspecimen variations in tissue modulus for trabecular bone, *J. Biomech.* 2002, 35(2), 237–246. doi: 10.1016/s0021-9290(01)00193-2

[38] Kleiven, S., Hardy, W. N., Correlation of an FE model of the human head with local brain motion–consequences for injury prediction, *Stapp Car Crash J.* 2002, 46, 123–144.

[39] Luo, Y., Li, Z., Chen, H., Finite-element study of cerebrospinal fluid in mitigating closed head injuries, *Proc. Inst. Mech. Eng., Part H: J. Eng. Med.* 2012, 226(7), 499–509. doi: 10.1177/0954411912445729

[40] Sahoo, D., Deck, C., Yoganandan, N., Willinger, R., Anisotropic composite human skull model and skull fracture validation against temporo-parietal skull fracture, *J. Mech. Behav. Biomed. Mater.* 2013, 28, 340–353. doi: 10.1016/j.jmbbm.2013.08.010

[41] Khalil, T. B., Hubbard, R. P., Parametric study of head response by finite element modelling, *J. Biomech.* 1977, 10(2), 119-132. doi: 10.1016/0021-9290(77)90075-6

[42] Ruan, J. S., Khalil, T. B., King, A. I., Dynamic response of the human head to impact by three-dimensional finite element analysis, *J. Biomech. Eng.* 1994, 116(1), 44-50. doi: 10.1115/1.2895703

[43] Hardy, C. H., Marcal, P. V., Elastic analysis of a skull, *ASME Trans.* 1973, 40(4), 838–842. doi: 10.1115/1.3423172

[44] Li, J., et al., Experimental and Numerical Studies on Vibration Modes and Transcranial Attenuation Characteristics in Unilateral Bone Conduction Hearing, *Shock and Vibration* 2020, Article ID 4962098, 17 pages.doi: 10.1155/2020/4962098

[45] Tse, K. M., Lim, S. P., Tan, V. B. C., Lee, H. P., A Review of Head Injury and Finite Element Head Models, *Am. J. Eng. Technol. Soc.* 2014, 1(5), pp. 28–52.

[46] Pearson, K. On lines and planes of closest fit to systems of points in space. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* 1901, *2*, 559–572, doi:10.1080/14786440109462720

[47] Hotelling, H. Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* 1933, *24*, 417–441, doi:10.1037/h0071325.

[48] Feng, Y., et al., Statistical shape modelling to analyse the talus in paediatric clubfoot, Proceedings of the Institution of Mechanical Engineers, Part H: *Journal of Engineering in Medicine* 2021, 235(8), 849-860. doi: 10.1177/09544119211012115

[49] Mei, L., Figl, M., Darzi, A., Rueckert, D., Edwards, P., Sample Sufficiency and PCA Dimension for Statistical Shape Models, *In: Forsyth D., Torr P., Zisserman A. (eds) Computer Vision – ECCV 2008. ECCV 2008. Lecture Notes in Computer Science*, 5305. doi:10.1007/978-3-540-88693-8_36

[50] Fuessinger, M. A., Schwarz, S., Cornelius, CP., et al., Planning of skull reconstruction based on a statistical shape model combined with geometric morphometrics, *Int J CARS* 2018, 13, 519–529. doi:10.1007/s11548-017-1674-6

[51] Pishchulin, L., et al., Building statistical shape spaces for 3D human modeling, *Pattern Recognition* 2017, 67, 276-286. doi: 10.1016/j.patcog.2017.02.018

[52] Quintens, L., Herteleer, M., Vancleef, S. et al., Anatomical Variation of the Tibia – a Principal Component Analysis, *Sci Rep* 2019, 9, 7649. doi: 10.1038/s41598-019-44092-8

[53] Jolliffe, I. *Principal Component Analysis*, 2nd ed.; Springer: New York, NY, USA, 2002.

[54] Audenaert, E. A., Pattyn, C., Steenackers, G., De Roeck, J., Vandermeulen, D., Claes, P., Statistical Shape Modeling of Skeletal Anatomy for Sex Discrimination: Their Training Size, Sexual Dimorphism, and Asymmetry, *Front. Bioeng. Biotechnol.* 2019, 7:302. doi: 10.3389/fbioe.2019.00302

[55] Jutten, C., Herault, J., Blind separation of sources, part I: an adaptive algorithm based on neuromimetic architecture, *Signal Process.* 1991, 24(1), 1-10. doi: 10.1016/0165-1684(91)90079-X

[56] Wu, J., et al., An implementation of independent component analysis for 3D statistical shape analysis, *Biomedical Signal Processing and Control* 2014, 13, 345-356. doi: 10.1016/j.bspc.2014.06.003

[57] Uzumcu, M., Frangi, A. F., Reiber, J. H. C., Lelieveldt, B. P. F., Independent component analysis in statistical shape models, *Proc. SPIE 5032, Medical Imaging 2003: Image Processing*. doi: 10.1117/12.481310

[58] Hyvarinen, A.; Karhunen, J.; Oja, E. *Independent Component Analysis*; John Wiley and Sons, Inc.: Hoboken, NJ, USA, 2001, doi:10.1002/wics.1440.

[59] Maaten, L. V. D., Hinton, G. E., Visualizing Data using t-SNE, *Journal of Machine Learning Research* 2008, 9(86), 2579-2605.

[60] Hinton, G. E., Roweis, S., Stochastic Neighbor Embedding, *Advances in Neural Information Processing Systems* 2003, 15.

[61] Zhu, Z., Wu, C., Zhang, W., Frequency Analysis and Anti-Shock Mechanism of Woodpecker's Head Structure, *J Bionic Eng* 2014, 11, 282–287. doi: 10.1016/S1672-6529(14)60045-7

[62] Bugli, C. Lambert, P., Comparison between Principal Component Analysis and Independent Component Analysis in Electroencephalograms Modelling, *Biom. J.* 2007, 49, 312-327. doi: 10.1002/bimj.200510285

[63] Husnain, M., Missen, M. M. S., Mumtaz, S., Luqman, M. M., Coustaty, M., Ogier, J-M., Visualization of High-Dimensional Data by Pairwise Fusion Matrices Using t-SNE, *Symmetry* 2019, 11(1):107. doi: 10.3390/sym11010107

[64] Namkeun, K., You, C., Stefan, S., A Three-Dimensional Finite-Element Model of a Human Dry Skull for Bone-Conduction Hearing, *BioMed Research International* 2014, vol. 2014, Article ID 519429, 9 pages. doi: 10.1155/2014/519429

[65] Franke, E.K., Response of the human skull to mechanical vibrations, *J. Acoust. Soc. Am.* 1956, 28 (6), 1277–1284. doi: 10.1121/1.1908622

Chapter 7

Conclusion

7.1 Summary

The main goal of this thesis was to establish a hypothesis that describes the relationship between geometric shape, material properties, and skull resonant frequencies, with concussive injury based on the brain mechanical impedance limits. Therefore, a means of supporting this theory was through the use of machine learning, deep learning, image processing, statistical analysis, and finite element analysis.

In Chapter 2, as segmentation is essential in medical imaging analysis to help extract regions of interest from different imaging modalities, a 3D CNN for skull segmentation in MRI was developed. CT scans are the most common medical imaging scanning technique for revealing bone structures; however, it is not favorable for routine anatomical imaging since CT exposes the patient to a small dosage of ionizing radiation, making MRI the ideal image modality for routinely application. The problem with MRI is the weak magnetic resonance signal of bones making it difficult to distinguish bone tissue from other structures. To overcome this issue, 58 gold standard volumetric labels were created from CT in STL models. These STL models were converted into matrices and overlapped on the 58 corresponding MR images to create the MRI gold standards labels. A CNN was trained with these 58 MR images and a DSC of 0.7300 ± 0.04 was achieved. A further investigation was carried out where the brain region was removed from the image with the help of a 3D CNN and manual corrections by using only

MR images. This new dataset, without the brain, was presented to the previous CNN which reached a new DSC of 0.7826±0.03. These findings were found to be a step forward in improving bone extraction in MRI using CNN to achieve average DSC rates comparable to those obtained in CT scans.

In Chapter 3, a method to initialize CNN using orthogonal moments for brain image segmentation in MRI was developed. To transform the original image into a modified version with orthogonal moment properties, the approach employs kernel windows based on orthogonal moments. The transformed image contains the optimal representation of the Legendre, Tchebichef, and Pseudo-Zernike moments coefficients. The method was validated on three different datasets and achieved a 4.12%, 1.91%, and 1.05% improvementl. In addition, the best image representors, a combination of Pseudo-Zernike and Tchebichef moments, were used to test alternative CNN frameworks. One particular framework, U-Net3+, demonstrated a slight improvement over the primary CNN used, U-Net, with an improvement of 0.64% when tested in the original images and 0.33% whem tested using the best image representors.

In Chapter 4, this study aimed to automatically locate the bilateral symmetry midline of the facial skeleton based on an invariant moment technique using pseudo-Zernike moments. A total of 367 skull 2D images were evaluated using a pseudo-Zernike moments, ICA, and PCA as feature descriptors of images using k-NN using Euclidean distance. The results revealed that pseudo-Zernike moments outperformed conventional landmark-based symmetry scores and PCA-ICP techniques, which rely on the general positions of cephalometric landmarks and manual selection of the central point. The technique was found to be reliable and provided good accuracy.

In Chapter 5, this study aimed to present a 3D technique using CNN and geometric moments to automatically calculate the craniofacial midline symmetry of the facial skeleton from CT scans. To perform this task, a total of 195 skull images were assessed to validate the proposed technique. For symmetric images, this approach achieved an accuracy of ≈99%. This method offers a more accurate approach to the pseudo-Zernike moments method presented in

Chapter 4, to the conventional landmark-based symmetry scores, and to PCA-ICP techniques. In the symmetry planes, the technique was found to be reliable and provided good accuracy.

In Chapter 6, the overall goal is to gain a better understanding of the relationship between transient vibration of the skull and the severity of a concussion. 45 CT scans were used to test, where an initial natural frequency investigation established the material properties of a 1-layer skull based on a 3-layer skull (cortical, diploe, and cortical bone) for FE analysis. The identified material properties were then applied to all 45 skulls and modal analysis was conducted to determine their natural frequencies. The relationship between different densities and Young's modulus versus natural frequencies were tested to identify any alteration in natural frequencies. Skull thickness was also analyzed, and shape analysis was conducted using PCA, ICA, and k-Means. A direct correlation between density, Young's modulus, and thickness with variations in natural frequency behaviour was found. No correlation was found between shape of the skull and natural frequency.

7.2 Strengths and Limitations

The work presented in this thesis provides a novel insights into the image processing, machine learning, and shape analysis of the skull with respect to its image segmentation, alignment, and relation between shape and material property with resonance frequency for concussion study purpose, incorporating probabilistic methods to account for variability in anatomical shape and material properties.

In Chapter 2, to the author's knowledge, this is the first study to apply STL models extracted from CT scans overlapped on MRI. The drawback of this work may be related to the sample size. As the dataset size increases, the various geometric disparity, variations, and deformity between the skulls sets become more evident. This aspect may have led to the high DSC reported in the related works as they used small datasets. Additionally, four different datasets were used contained differences in ethnicity, age, and medical history. Furthermore, several

patients have undertaken cephalic surgery, which may have altered the cerebrofacial skeleton structure, affecting the performance of the segmentation. Moreover, during the creation of the gold standard STL models, corrections were performed manually which may included some defects during the edge smoothing or noise residue removal. Therefore, the CNN may have learned the defect regions as ground truth leading to an error in the segmentation. Also, when transforming the STL model into matrix for execution in Matlab, a voxelisation algorithm was applied, which may have induced a quantity of information being incorrectly labeled when converted from skull-voxel to imaging-voxel.

In Chapter 3, to the authors knowledge, this was the first study to use orthogonal moments as an initialize input in CNN for MRI brain segmentation. This pre-processing transforms the original image into a new version with a robust discriminative capacity and it has been demonstrated to improve brain segmentation in MRI. The orthogonal moments unique translation, rotation, and scaling invariances allow them to describe image attributes with no duplication of information or redundancy among the moments, and deal with geometric distortions. Moreover, they are commonly employed for feature extraction since they offer orthogonality, noise robustness, high discriminative capacity, and low information redundancy. One limitation is the employment of 2D orthogonal moments, which is time consuming and only covers one image slice each interaction.

In Chapter 4, invariant moments were used to introduce a novel approach to identify the bilateral symmetry midline of the facial skeleton. This method demonstrated great potential in determining the bilateral symmetric midline of the face and allowed the algorithm to learn as additional data is collected, which was not feasible with previous methods. However, there are a few limitations. Since this method was only performed in 2D images, some of them were rotated in the sagittal plane, resulting in errors in obtaining the centre line. Furthermore, when images contain deformation, incompleteness, or non-uniform brightness, calculating the centre of the image becomes challenging since this approach uses the shape of the image as reference to calculate the centre. Furthermore, due to lack of non-symmetrical skulls, there

was no testing in non-symmetrical images. Also, to avoid computational crash/error during the moment calculus, the images were downsized to 128×128, which may result in less accurate finding as a higher resolution image offers more visual information.

In Chapter 5, deep learning was used to provide a novel 3D approach for automatically determining the craniofacial symmetry midline of the facial skeleton from CT images. This proposed method showed good results in obtaining the bilateral symmetric midline of regular/symmetric images. However, it failed to determine the rotation and the image centre for some deformed images. These inaccuracies are most likely due to a small database size, a lack of a deformed CT database which restrings the ability to train the system with deformed images, and the generation of ground truth and voxelization which may have omitted some crucial image features. Furthermore, when the image has a lack of symmetry, deformation, incompleteness, non-uniform brightness, or interference, the computation of the image center becomes complex and finds some limitations as this approach is a quantitative measure of an image's function or structure.

In Chapter 6, a new concept on the role of transient vibration and skull geometry on concussion was introduced. This study identified that the skull thickness, density, and Young's Modulus may play an important role on natural frequency to a near threshold that may trigger concussive injury. For statistical shape analysis, no relation was find among the skull that describes a shape pattern for natural frequency near 450±50 Hz. One limitation to the proposed method is the simplification of the material model. A drawback in this approach may be attributed to its model. To reduce the overall time required for simulation, a single-layered material method was adopted. A 3-layered material map would provide a more accurate representation of real life behaviour but requires more time in the segmentation of each material layer. Furthermore, cephalometric landmarks were used as gold standards during the facial alignment. If a single image was misaligned, an error could be generated in the statistical shape analysis interpretation.

7.3   Future Work

In future work, many different adaptations, testing, and experiments will be explored along with a deeper analysis of certain mechanisms. The following ideas may be tested in a future work:

Chapter 2: a strategy to improve CNN accuracy is to implement a procedure where CT and MRI would be scanned from the same patient and, through a fixture, the patients head would be perfectly aligned. As a result, all images would have the same orientation and possible scale. With a number of additional images, the CNN may be able to learn new features and improve its segmentation accuracy.

Chapter 3: a 3D orthogonal moments can be investigated to enhance accuracy and time consumption. The CNN would be able to learn features from a 3D kernel rather than a single slice in this manner. In addition, instead of applying pre-processing to the input image, this strategy might be implemented as a CNN filter.

Chapter 4 and 5: as chapter 4 and 5 share the same principal, this approach would benefit both methods. Increasing database size, as well as the addition of non-uniform and deformed images, may be advantageous. Identifying a method to calculate the precise center of the image would also be a significant improvement. For a 3D images, instead of applying the geometric moments approach, a 3D CNN might be trained to segment the nasal bone and therefore find the image center.

Chapter 6: regarding the brain mechanical impedance limits and skull resonant frequencies, there is a continuing desire to study, understand, and develop a system that can simulate an impact that triggers the skull resonant frequency near this threshold in a safe and repeatable manner. Therefore, future work aims on determining how the vibrations of the skull generated by an impact can constructively create resonant frequencies that may propagate into the brain tissue which might induce mild traumatic brain injury. This

work can be performed by using the method introduced in this thesis.

Those skulls with a resonant frequency of 450±50 Hz identified in Chapter 6, would be selected for an impact analysis. Then, using the automatic algorithm developed in chapters 4 and 5, skulls would be aligned for the impact investigation. This step has the goal to precisely simulate the same impact point in all models. These brains associate with the skull selected from Chapter 6 would be extracted, as in Chapter 3, imported into Mimics software, and material properties and mesh would be applied using descripted properties already used in the literature [1–5]. Cerebro-spinal fluid would be created from MRI by excluding the volume of the skull and brain [6]. Then, the models would be imported into Abaqus FEA software where impact simulation wouldl be assessed to identify whether the impact can generate the 450±50 Hz resonant frequency found in Chapter 6 and whether the skull vibration can propagate into the brain tissue. From the data and results acquired in this step, it is expected to raise a theory regarding the relation between the skull geometry and its vibration with the brain tissue injury, which might optimize helmet design (Figure 7.1).

These suggestions represent a narrow sample of the potential impact of the tools developed in this thesis. As the ability to further automate skull and brain segmentation in MRI, bone alignment, and the skull modal analysis continue, the true potential of the combination of computational methods and biomechanics will be realized.
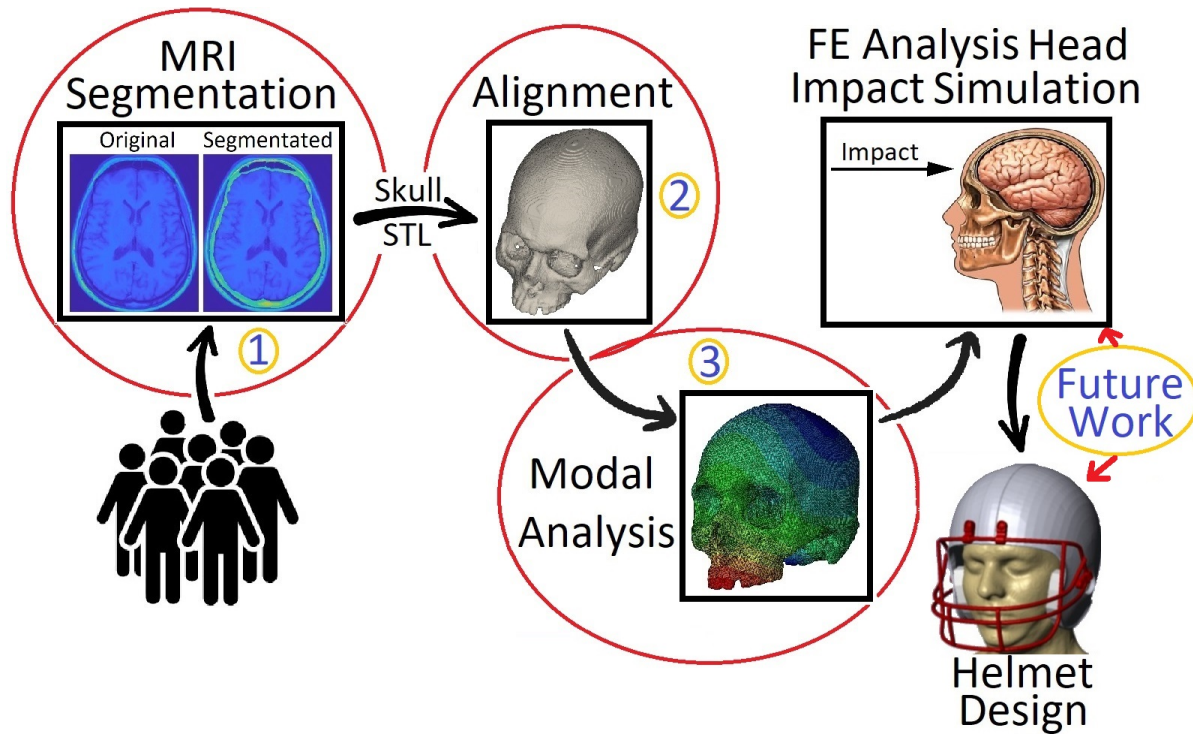
Figure 7.1: (1) An algorithm extracts the skull from the MRI in the STL file format, (2) these skulls are then aligned using an algorithm, and (3) modal analysis assesses the natural frequency of skulls to distinguish between different skulls with similar natural frequencies. Finally, impact simulations will be performed on a whole head model. Based on the data collected, helmets can be redesigned to provide better protection against concussive traumatic injuries.

Bibliography

[1] Chatelin, S., Constantinesco, A., Willinger, R. Fifty Years of Brain Tissue Mechanical Testing: From in Vitro to in Vivo Investigations, *Biorheology* 2010, 47(5-6), 255 – 276. doi: 10.3233/BIR-2010-0576

[2] Saboori, P., Sadegh, A. Material modeling of the head's subarachnoid space, *Scientia Iranica* 2011, 18(6), 1492-1499, doi: 10.1016/j.scient.2011.11.032

[3] Wei Zhao, Bryan Choate, Songbai Ji, Material properties of the brain in injury-relevant conditions – Experiments and computational modeling, *Journal of the Mechanical Behavior of Biomedical Materials* 2018, 80, 222-234, doi: 10.1016/j.jmbbm.2018.02.005

[4] Budday, S., Ovaert, T. C., Holzapfel, G. A. et al. Fifty Shades of Brain: A Review on the Mechanical Testing and Modeling of Brain Tissue, *Arch Computat Methods Eng* 2020, 27, 1187–1230. doi: 10.1007/s11831-019-09352-w

[5] Barbosa A., Fernandes F. A. O., Alves de Sousa R. J., Ptak M., Wilhelm J. Computational Modeling of Skull Bone Structures and Simulation of Skull Fractures Using the YEAHM Head Model. *Biology* 2020, 9(9):267. doi: 10.3390/biology9090267

[6] Wu, J. Z. et al. An Improved Finite Element Modeling of the Cerebrospinal Fluid Layer in the Head Impact Analysis, *Bio-Medical Materials and Engineering* 2017, 28(2), 187-199. doi: 10.3233/BME-171666

Curriculum Vitae

Name: Rodrigo Dalvit Carvalho da Silva

Post-Secondary Education and Degrees:

University of Fortaleza
Fortaleza, Ceara, Brazil
2011 B.Sc. in Control and Automation Engineering

Federal University of Ceara
Fortaleza, Ceara, Brazil
2014 M.Sc. in Teleinformatics Engineering

The University of Western Ontario
London, Ontario, Canada
2022 Ph.D. in Biomedical Engineering

Teaching and Research Work Experience:

Teaching and Graduate Research Assistant in Biomedical Engineering
The University of Western Ontario
2017 - 2021

Electrical Prototype Laboratory Assistant
Student Shop - The University of Western Ontario
February 2019 - April 2019

Dual Credit Course Professor
AEC104-CAD - St. Clair College
September 2019 - January 2020

Laboratory Professor
CNC Machining Laboratory - MACH 1144 - Fanshawe College
May 2019 - Augustl 2019

Adjunct Professor
Computer Aided Design and Engineering - CAD/CAE (Catia V5)
Machining Technology – Manual Machining
Computer Numerical Control – CNC
Department of Control and Automation Engineering - University of Fortaleza
August 2014 - June 2015

Publications

Peer Reviewed Journals:

- The work done in Chapter 2 is summarised in the following paper - DALVIT CAR-VALHO DA SILVA, R., JENKYN, T. R., CARRANZA, V. A., Development of a Convolutional Neural Network Based Skull Segmentation in MRI Using Standard Tessellation Language Models. *Journal of Personalized Medicine* 2021, 11(4), 310. doi: 10.3390/jpm11040310

- The work done in Chapter 3 is summarised in the following submitted paper - DALVIT CARVALHO DA SILVA, R., JENKYN, T. R., CARRANZA, V., A., Enhanced Pre Processing for Deep Learning in MRI Brain Segmentation using Orthogonal Moments. 2021. *Submitted - peer-reviewed journal*

- The work done in Chapter 4 is summarised in the following paper - DALVIT CAR-VALHO DA SILVA, R., JENKYN, T. R., CARRANZA, V. A., Application of a Novel Automatic Method for Determining the Bilateral Symmetry Midline of the Facial Skeleton Based on Invariant Moments. *Symmetry* 2020,12:9,1448. doi: 10.3390/sym12091448

- The work done in Chapter 5 is summarised in the following paper - DALVIT CAR-VALHO DA SILVA, R., JENKYN, T. R., CARRANZA, V. A., Convolutional Neural Networks and Geometric Moments to Identify the Bilateral Symmetric Midplane in Facial Skeletons from CT Scans. *Biology* 2021, 10(3), 182. doi: 10.3390/biology10030182

- The work done in Chapter 6 is summarised in the following submitted paper - DALVIT CARVALHO DA SILVA, R., JENKYN, T. R., CARRANZA, V. A., The role of transient vibration and skull properties on concussion. 2021. *Submitted - peer-reviewed journal*

- DALVIT CARVALHO DA SILVA, R., JENKYN, T. R., Classification of Mammogram Abnormalities using Legendre Moments. *International Journal of Image and Graphics* 2020, 21(1), doi: 10.1142/S0219467821500108

Conferences Publications:

- DALVIT CARVALHO DA SILVA, R., THE, G. A. P., DE MEDEIROS, F. N. S., Rotation-Invariant Image Description from Independent Component Analysis for Classification Purposes. In: 12th International Conference on Informatics in Control, Automation and Robotics, 2015, Colmar. doi: 10.5220/0005512802100216

- DALVIT CARVALHO DA SILVA, R., THE, G. A. P. ; DE MEDEIROS, F. N. S., Geometrical and statistical feature extraction of images for rotation invariant classification systems based on industrial devices. In: 21st International Conference on Automation and Computing, 2015, Glasgow. doi: 10.1109/IConAC.2015.7313946

- DALVIT CARVALHO DA SILVA, R., COELHO, D. N., THE, G. A. P., MENDONSA, M. R., Comparison between k-Nearest neighbors, self-organizing maps and optimum-path forest in the recognition of packages using image analysis by Zernike moments. In: 2014 11th IEEE/IAS International Conference on Industry Applications INDUSCON 2014, 2014, Juiz de Fora. doi: 10.1109/INDUSCON.2014.7059477

- DALVIT CARVALHO DA SILVA, R., THE, G. A. P., Moment Invariant based classification of objects from low-resolution industrial sensor images. In: 11th Brazilian Congress (CBIC) on Computational Intelligence, 2013, Porto de Galinhas. doi: 10.21528/CBIC2013-250

- DALVIT CARVALHO DA SILVA, R., THE, G. A. P., COELHO, D. N., Performance Analysis of Classifiers to Recognition of Objects From Low-Resolution Images Industrial Sensor. In: XI Simpósio Brasileiro de Automação Inteligente, 2013, Fortaleza. Anais do XI SBAI/DINCON 2013, 2013, 2013. v. 1. p. 1-4.

- DALVIT CARVALHO DA SILVA, R., THE, G. A. P., A Comparison Between Hough Transform and Moment Invariant to the Classification of Objects from Low-Resolution Industrial Sensor Images. In: XI Simpósio Brasileiro de Automação Inteligente, 2013, Fortaleza. Anais do XI SBAI/DINCON 2013, 2013, 2013. v. 1. p. 1-4.

- NUNES, T. M. ; DALVIT CARVALHO DA SILVA, R., ALBUQUERQUE, V. H. C., Desempenho Preliminar de uma Abordagem Vetorial sobre Metodos de Contornos Ativos. In: XXIII Congresso Brasileiro de Engenharia Biomédica - CBEB 2012, 2012, Porto de Galinhas. XXIII Congresso Brasileiro de Engenharia Biomédica - CBEB 2012, 2012. v. 1. p. 1213-1215.

- DALVIT CARVALHO DA SILVA, R., NUNES, T. M. ; PINHEIRO, G. J. B. ; ALBUQUERQUE, V. H. C., Customizacao da Transformada Imagem-Floresta para Grafos Densos Utilizando Algoritmo de Foyd-Warshall um Estudo Inicial. In: XXIII Congresso Brasileiro de Engenharia Biomédica - CBEB 2012, 2012, Porto de Galinhas, v. 1. p. 1216-1218.