Rowan University

Rowan Digital Works

2-13-2024

MRI IMAGE REGRESSION CNN FOR BONE MARROW LESION VOLUME PREDICTION

Kevin Yanagisawa
Rowan University

Recommended Citation

2-13-2024

MRI IMAGE REGRESSION CNN FOR BONE MARROW LESION VOLUME PREDICTION

Kevin Yanagisawa

MRI IMAGE REGRESSION CNN FOR BONE MARROW LESION VOLUME PREDICTION

By

Kevin Yanagisawa

A Thesis

Submitted to the
Department of Biomedical Engineering
College of Engineering
In partial fulfillment of the requirements
For the degree of
Master of Science in Biomedical Engineering
at
Rowan University
August 14, 2023

Thesis Chair: Erik Chrisopher Brewer, Ph. D., Professor, Department of
Biomedical Engineering

Committee Members:
Mary Staehle, Ph. D., Professor, Department of Biomedical Engineering
Nidhal Bouaynaya, Ph.D., Professor, Department of Electrical and Computer
Engineering
Shen-Shyang Ho, Ph.D., Professor, Department of Computer Science

Dedications

To my parents, family, close friends, and swim team, for their constant support.

Abstract

Kevin Yanagisawa
MRI IMAGE REGRESSION CNN FOR BONE MARROW LESION VOLUME
PREDICTION
2022-2023
Erik Christopher Brewer, Ph.D.
Master of Science in Biomedical Engineering

Bone marrow lesions (BMLs), occurs from fluid build up in the soft tissues inside your bone. This can be seen on magnetic resonance imaging (MRI) scans and is characterized by excess water signals in the bone marrow space. This disease is commonly caused by osteoarthritis (OA), a degenerative join disease where tissues within the joint breakdown over time [1]. These BMLs are an emerging target for OA, as they are commonly related to pain and worsening of the diseased area until surgical intervention is required [2]–[4]. In order to assess the BMLs, MRIs were utilized as input into a regression convolutional neural network (CNN). Initial experimentation handled the MRI using each individual slice in a 2D convolutional neural network as a baseline model, eventually progressing to a tensor stacked input into a 3D convolutional neural network. The viability and effectiveness of the models were evaluated using mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), and R^2. The 3D regression CNN models were observed to perform better than the 2D regression CNN models.

Table of Contents

Table of Contents (Continued)

Table of Contents (Continued)

List of Figures

List of Tables

Chapter 1

Introduction

1.1. Research Motivation

Osteoarthritis of the knee is well known for the cause of pain and disability that affects millions of patients at some point in their lifetime. This results in a decreased quality of life and economic impacts in terms of health care costs and lost productivity [5]. Along with these effects, bone marrow lesions (BMLs) are commonly found in patients with knee osteoarthritis [6]. These BMLs can cause further inflammation and discomfort for the patients and have potential cause in the approximate 700,000 total knee replacements performed annually in the United States [7]. In 2008, Scher found that there was a nine-fold increase in the rate of progression to total knee arthroscopy within three years when there was an associated BML identified on MRI compared to patients without an associated BML [4]. Total knee arthroplasty is a major surgery, so other methods needed to be developed to prevent BMLs.

In recent years, a subchondroplasty procedure was developed to treat BMLs by injecting a calcium phosphate bone substitute into the diseased areas of subchondral bone under the guidance of a fluoroscope [8]. This joint-preserving treatment that helps reverse the progression of pain and immobility was used in patients with bone marrow lesions in the knee. To identify the diseased area, MRI scans were typically used. These BMLs can be described as high T2 and low T1 signal intensity detection and low T1 signal on

magnetic resonance (MR) fluid-sensitive sequences [9]. These areas can be visually identified when looking at the MRI scans, but the volume of the BML must be estimated by the clinician. During this procedure, since the volume of the BML is not known precisely based on the MRI, there can be potential pitfalls such as not directing the calcium phosphate into the center of the BML and not abutting the subchondral plate [10]. This would be known as underfilling and may not remedy the symptoms caused by BMLs. On the other hand, overfilling and rapid delivery of the calcium phosphate may occur, leading to postoperative pain [10], [11]. With complications revolving around knowing how much calcium phosphate injection is needed, there is an ensuing need for technology that can utilize the MRI scans and have a predictive outcome for the volume of injection needed to treat the BML.

The aim of this research is to develop a model that can use MRI scans of patients with BMLs as its input and provide predictions on the volume of injection needed to treat the BML.

<center>**Chapter 2**</center>

<center>**Background**</center>

2.1. Prevalence and Scope of Bone Marrow Lesion

Bone marrow lesions (BMLs) can be found across all age groups and all genders. In a study involving 328 younger adults aged 31-41, researchers set out to describe the prevalence and environmental, structural, and clinical correlates of BMLs [12]. Through this, the overall prevalence of BML was 17% and positively correlated with increasing age and previous knee injuries. In another study observing healthy middle aged women, researchers sought to determine the risk factors for BMLs [13]. There, 176 women with no prior history of knee injury or clinical care had MRI scans done on their dominant knee; 13% of them had knee BMLs. A similar study took MRI scans of 266 Japanese women without evidence of osteoarthritis (OA) and found that there was a 35.3% prevalence of BMLs [14]. For patients that do not have healthy knees, there is an even greater potential for BMLs. A study evaluated the prevalence of BMLs in 255 subjects with knee pain and found that 11% of subjects with no (OA) had BMLs, 38% of subjects with preradiographic OA had BMLs, and 71% of subjects with radiographic OA had BMLs. With this prevalence across all genders and ages, there is a need for proper diagnosis and treatment of this disease.

<center>3</center>

2.2. Treatment Options

Treatment options for BMLs have commonly been focused on conservative care and pain management, but in more extreme cases, surgical intervention may be appropriate.

2.2.1. Conservative Care

Conservative care commonly includes symptomatic treatment or pharmacological treatment. Symptomatic nonoperative treatment involves anti-inflammatory drugs, removable walker boot, partial weightbearing or nonweight-bearing, physiotherapy, and message therapy[15]. Recovery times using these conservative care treatment options can be anywhere from 6 to 12 months, with no definable duration time of use. Pain management options include different medications such as prostacyclins, bisphosphonates and polysulfated polysaccharides, as well the use of pulsed electromatic fields[15]. Depending on the medication used and the duration of usage, they can provide instant pain relief or eventual improvement after months of usage. Additionally, side effects vary dependent on the medication used and the dosage. It has also been seen that low vitamin D levels can be cofactor in the development and aggravation of many orthopedic diseases.

2.2.2. Surgical Intervention

When conservative care is not effective, surgical intervention can be called for, and this is typically in the form of core decompression or subchondroplasty. Core decompression involves drilling a hole in the affected area to reduce pressure and allow for increased blood supply. This method is not commonly used, but may be considered for a rapid   decrease   in   symptoms   and   improvement.

This has a low morbidity and can be done as an outpatient procedure, but the procedure requires partial weightbearing and physiotherapy. Even with this, in a study of 38 patients with hip bone marrow edema, core compression compared similarly, if not equal or worse, than the medication iloprost [16]. Subchondroplasty is a relatively new procedure which involves a localized injection of calcium pyrophosphate bone substitute into the bone marrow lesion cavity to help preserve the joint [17]. In order to inject the calcium pyrophosphate cement, MRI scans are needed to identify and localize the diseased area. Currently, MRIs and ultrasounds are the only technology used to identify these locations.

The subchondroplasty procedure is the preferred surgical intervention method because of its minimally invasive properties while still being effective [7]. Patients who need this procedure are typically evaluated using different scoring methods including the visual analog scale (VAS) to measure pain, International Knee Documentation Committee (IKDC), Knee Injury and Osteoarthritis Outcome Score (KOOS), or the Western Ontario and McMaster Universities Arthritis Index (WOMAC) [18]. By using these grading scales, improvement can be quantified post-operation. Although this method is effective and minimally invasive compared to other surgical interventions, it still comes with its own complications. Reported complications include extravasation of calcium phosphate into the surrounding soft tissues or joint [19], [20],[21], postoperative pain [8], [19], [22], and drainage at the injection site[8]. The extravasation of calcium phosphate is also known as overfilling. This overfilling occurs because the amount of calcium phosphate to use for each patient is not automatically determined, it is up to the discretion of the doctor. To

estimate the appropriate volume of injection, the doctors utilize MRI scans of the patient's knee.

2.3. CNN for Medical Imaging

2.3.1. Magnetic Resonance Imaging

Standard MR modalities include fat-saturated T2-weighted and T1 weighted MRIs. T1 weighted MRIs enhance the signal of the fatty tissue and suppress the signal of water. T2 weighted MRIs enhance the signal of water. This causes T1 weighted MRIs to show only fat as the brighter areas in the scan and T2 weighted MRIs to show fat and water as brighter sections of the scan. For bone marrow lesions, the T1 weighted MRI will show an abnormally low signal at the point of disease and the T2 weighted MRI will show an abnormally high signal [23, 24]. The most validated MRI image biomarker is T2 mapping, which bases its measurements on T2 relaxation time [25].

MR viewing planes include axial, coronal, and sagittal planes. The axial plane makes cross sections of the knee from top to bottom. The coronal plane takes slices through the knee from front to back. The sagittal plane looks through the knee from side to side.

2.3.2. Convolutional Neural Network

A convolutional neural network (CNN) is a deep-learning algorithm designed to use different interconnected layers for learning patterns or features of raw input data. For this experiment, these layers included convolutional layers, pooling layers, flattened layers, and fully connected layers. Within a convolutional layer, a filter or kernel is applied to the input data to extract features. These features produce a feature map that highlights the

patterns within the data. The pooling layer down-samples the dimensions of a feature map, effectively reducing the size while preserving important information. This helps reduce the complexity of the computations and reduce overfitting. Flattened layers are used to convert the multidimensional feature map into a one dimensional vector. This is in preparation for use in the fully connected, or dense layer. The fully connected layer connects every layer of the model and is used as the final layer. With each dense layer, an activation function is used to introduce non-linearity into the neural network. This allows for complex and non-linear relationships to be learned from the data.

   2.3.2.1. CNN for Medical Imaging. CNNs are commonly used in medical imaging tasks such as image classification, object detection, and image segmentation. In previous studies, CNNs were utilized to locate and classify brain tumors [26–28]. Muhammad [27], for example, goes into detail about surveying which deep learning based method is best for brain tumor classification. Additionally, CNNs have been used to predict ages based on MRI images. Pardakhti [29] used a 3D regression CNN to train a brain age estimation system.

Chapter 3

Research Aims

The main objective of this research is to develop a model based on pre-operative MRI scans that can predict the volume of filling needed during a subchondroplasty procedure. Typically, when treating a patient for their bone marrow lesion (BML) the surgeon will fill the affected area with a calcium-phosphate based bone cement. The problem arises due to the surgeon not being able to tell the exact volume of filling necessary solely based off the MRI scans. The amount of filling is determined by the surgeon when they feel a push back from the syringe. Without this measurement, there is no way to prevent overfilling or underfilling.

We plan on doing this by developing MRI imaging techniques that utilize both a deep learning algorithm with a regression CNN model and cropping methods to provide surgeons with accurate injection volumes. Others have used regression CNN models for age estimation, density estimation, or identification of areas within an MRI, but ours is unique because it works with novel pre-operative MRI scans of BMLs. Therefore, preoperative MRIs of each patient with accompanying volumes of injection may similarly be a novel way to guide clinicians on the treatment of BMLs. But to date, no group has attempted this with the subchondroplasty procedure.

In addition, the output from an MRI scan are multiple slices through the body part being inspected. By utilizing tensors, the MRI scans can be read into the model not as individual images, but as a singular stack of images. With this ability, viewing planes of

the MRI and modalities of the MRIs can be combined and read as one patient. Our objective is to create an informed delivery strategy for percutaneous fixation of BMLs.

The specific aims of this work are:

Specific Aim #1: Create a baseline regression CNN model that utilizes various viewing planes from BML MRI scans.

Specific Aim #2: Create a 3D regression CNN model that utilizes tensor stacked MRI viewing plane images

Specific Aim #3: Utilize 3D regression CNN model with combinations of viewing plane inputs.

Chapter 4

Base Regression CNN Model

4.1. Introduction

Convolutional neural networks (CNNs) have been utilized in medical imaging studies involving classification, segmentation, or prediction models [30–32]. When an MRI scan is completed, a three-dimensional image of the scanned area gets split into multiple sections called slices. Each one of these slices had the same height, width, and thickness, so all together they show the desired area as a three-dimensional image. A common usage of these MRI scans is to use each slice as an input for a CNN [33–35]. Using each slice of the MRI scan as an input can be utilized in detection of brain tumors [34], segmentation [30,33], and regression [35],31]. For the regression studies, determination of the best model was based on the RMSE, MSE, and MAE metrics. A study that predicted brain age using slices of an MRI compared multiple CNN models [31]. In order to make a comparison between these CNN models, the authors utilized the RMSE and MAE values. This allowed them to look across the models and compare to see which one had the lowest RMSE and MAE result. Through this, they found that their 2D Slice model performed better than their 3D models. With these models, it is also important to observe the convergence of the loss function to ensure it reaches its minimum [36].

Another element of an MRI scan is the viewing plane. The viewing planes are the axial, coronal, and sagittal planes, and each one of these shows the same area on a patient,

but from a different view. There have been studies to determine the best viewing plane to use when looking at a knee MRI [37, 38], but in these cases, they were not looking for bone marrow lesions. In addition, radiologists were the people to review the MRI scans, so it was all manually determined by the radiologists. With a lack of studies comparing the different viewing planes, it is important to determine if the best viewing planes are similar between the manual method by radiologist and the use of a CNN.

The models for each of these CNNs can vary, but most of them contain similar elements including convolution layers, max pooling layers, activation layer, and a fully connected (dense) layer. In addition, models may be pretrained using transfer learning and use various CNN architectures, but for this model, training from scratch, based on the model from Bellary [32], performed better.

Therefore, the purpose of this specific aim is to develop a model based on these prior slice by slice studies and compare their metrics with commonly accepted values to determine the validity of the model. Furthermore, we will compare different viewing planes to determine which is best when being inputted into the CNN. Since the MRI scans are inputted into this model one slice at a time and not one full MRI at a time, it is understood that a combination of all the slices creates a 3D representation of the scanned body part. Therefore, we will selectively compare the predicted output of the model on individual slices to their neighboring slices.

4.2. Methods

4.2.1. Pre-Processing

Before loading the images into the model, preprocessing steps were taken to enhance the information fed into the model. After downloadeding, MRIs were typically of shape 256 x 256 voxels or 512 x 512 voxels and contained 20-30 images per patient per viewing plane. To encapsulate the largest bone marrow lesion, an 80mm x 80mm bounding box was drawn around the center of the bone marrow lesion. This step was completed using ImageJ. Once the size of the image was reduced, slices were selected to be used in the model. An example of slices can be seen in Figure 1. The largest bone marrow lesions spanned over 10 slices, so in order to encapsulate only the bone marrow lesion for this patient, the minimum number of slices needed was 10. The bounding box of 80mm x 80mm and slice removal until only 10 slices remained was done on all images.

MULTIPLE SLICE IMAGING

0 200 400 600 800 1000 mSec

Excitation Sequence
←—1 Cycle Time (TR)—→
Repeat for each cycle

Sprawls
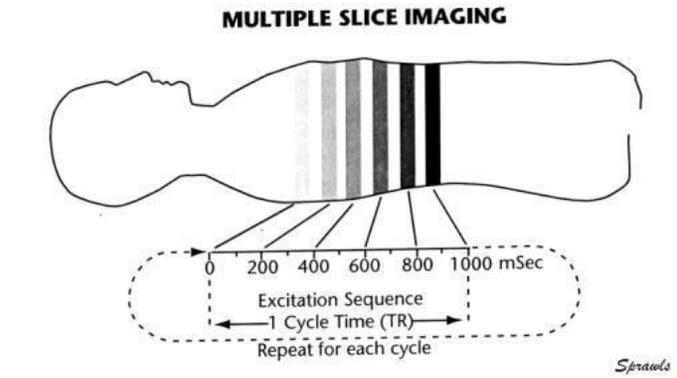
Figure 1. Example of multiple slices from an MRI scan

4.2.2. Architecture

The input of the slice by slice model were individual slices of an MRI scan, and each slice was assigned the true value of injection volume as used by physician. Once these images were fed into the model, they went through a convolution layer, max pooling, convolution layer 2, max pooling layer 2, a flatten, 2 RELU activation dense layers, and finally a linear activation dense layer. This architecture can be seen in Table 1.

Table 1

Slice by Slice Model CNN Architecture

Layer	Output shape	Parameter Values
Conv1	254 x 254 x 64	1,792
MaxPool1	127 x 127 x 64	0
Conv2	125 x 125 x 64	36,938
MaxPool2	62 x 62 x 64	0
Flatten	246016	0
Dense1	64	15,745,088
Dense2	64	4160
Dense3	1	65

4.2.3. Dataset

This study focuses on three magnetic resonance (MR) sequences of 35 preoperative knees from Dr. Sean McMillan Orthopedics & Sports Medicine and Rothman Orthopedic Institute. All patients have been affected by and treated for their bone marrow lesions through a subchondroplasty procedure in their femoral or tibial condyle. Before treatment, a preoperative MRI scan was taken so the areas of infection can be identified. Postoperative injection volume of calcium phosphate compound was recorded and exists as the ground truth. This model included 10 images per patient, 350 images total, for each viewing plane (coronal, sagittal, and axial). Subject data selected for the study are required to have all sequences acquired.

4.2.4. Experimental Conditions

The Keras library with its 'ImageDataGenerator' was utilized to apply augmentation techniques, aiming to enhance diversity and robustness of the training dataset. A rescale at 1./255, rotation of 20 degrees, shear by a factor of 0.2, zoom into the image by a factor of 0.2, shifts in the horizontal and vertical directions by 20% of its width and height, horizontal flip, and brightness transformations from -0.5 to 0.5 were implemented as experimental conditions. The model was run for 100 epochs to ensure convergence.

4.2.5. Metric for Evaluation of Regression Model

To understand the performance of a model, it is important to understand the metrics for evaluation. This allows for understanding of the performance and allows others to comprehend how well your model is doing. For this regression model, three metrics will be observed for evaluation. This includes R square (R^2), mean absolute error (MAE), mean square error (MSE), and root mean square error (RMSE). Such metrics can be seen in other regression models [39] .

4.3. Results and Discussion

The primary objective of this body of work was to develop a model capable of predicting subchondroplasty injection volumes based on bone marrow lesions from MRI scans. Table 2 summarizes the results from the slice by slice model for the axial, coronal, and sagittal views on the testing dataset. Each model was trained and evaluated using either 12 patient (120 slices) for training, 3 patients (30 slices) for validation, and 5 patients (50

slices) for testing or 27 patients (270 slices) for training, 3 patients (30 slices) for validation, and 5 patients (50 slices) for testing. Each of the models exhibited convergence, as demonstrated by the decrease in the mean squared error loss function over the training epochs seen in Figure 2. Models were trained over 100 epochs to ensure the convergence of the model. At this point, additional training does not improve the model, so it is around the final loss value. This means the model is around as good as it will be and the data potentially fits the model well.

| Axial | Coronal | Sagittal |
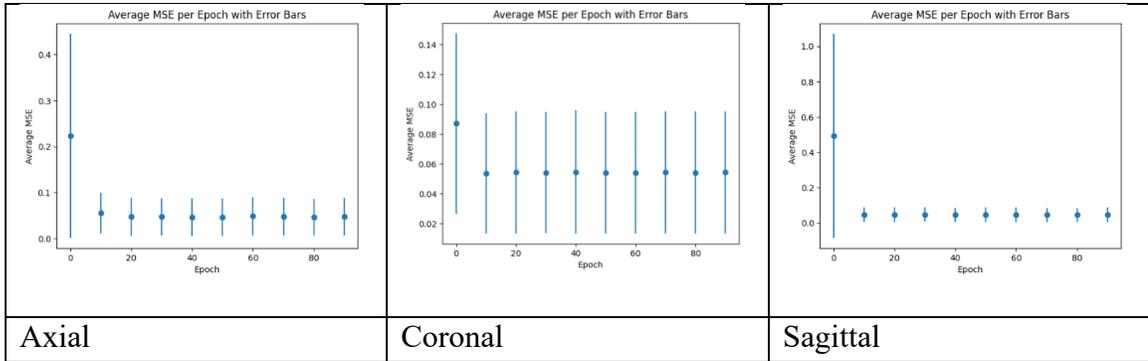
Figure 2. Average MSE (over 10 trainings) per epoch. A convergence in MSE value can be seen in the Axial, Coronal, and Sagittal models.

The best MAE and MSE was shown in the coronal 27 patient (270 slices) dataset with a value of 0.249561 and 0.139518, respectively. The coronal 27 patient (270 slice) dataset also showed the best RMSE and R^2 values at 0.37352 and -0.45820, respectively. A study looking at brain age estimation used similar metrics, MAE and RMSE, to the ones used here [29], and MSE can additionally seen being used in a study predicting age from

structural brain images[40]. Within these studies, they were looking for the model that produced the lowest MAE, MSE, and RMSE. This model seems to have MAE, MSE, and RMSE metrics similar to the previously listed journal articles, but the R^2 result is poor. In lieu of this poor R^2 result, it is important to note that the data set size used in this model was low, 35 patients, where as in similar studies such as a brain age prediction [31], the data set was around 10,000 patients. Despite this lack of patients, there was a decrease in MSE, decrease in MAE, decrease in RMSE, and increase in R^2, when going from 12 patients (120 slices) to 27 patients (270 slices). In a study using GAN based image augmentation to increase the size of their dataset, they found that with more data, their classification accuracy increased[41]. In an additional study, which fed additional image features into their model, their classification performance improved[42]. With this, even though the initial results are poor, this increase in patient count shows potential for obtaining better results as the patient count increases.

Table 2

Metric Evaluations From Slice by Slice CNN Model Using 12 and 27 Patients. MSE, MAE, RMSE, and R2 Evaluations are From the Test Dataset

	12 Patient, 120 Slice Training	27 Patient, 270 Slice Training
Axial MAE	0.27688	0.26610
Axial MSE	0.15667	0.15308
Axial RMSE	0.39582	0.39126
Axial R2	-0.63751	-0.59997
Coronal MAE	0.29887	0.249561
Coronal MSE	0.18814	0.139518
Coronal RMSE	0.43376	0.37352
Coronal R2	-0.60264	-0.45820
Sagittal MAE	0.29536	0.294169
Sagittal MSE	0.17353	0.169953
Sagittal RMSE	0.41658	0.41225
Sagittal R2	-0.81374	-0.77629

The secondary objective of this specific aim was to compare the three different viewing angles from the MRI scan and evaluate if one provided a better depiction of the bone marrow lesion when being fed into a CNN. From this, it appears that the coronal view produced the most appropriate BML volume of injection model. In a study which evaluated the MRI images for anterolateral ligament of the knee using experienced radiologist found that the coronal viewing plane was best for observing and characterizing the anterolateral ligament[37]. Such results are also supported in a study on a similar topic. The authors compared the coronal to the sagittal view to determine whether coronal imaging allowed for better detection when compared to the sagittal view[38]. Through this, the experiments found that coronal magnetic resonance images allow for better detection than sagittal

images alone. Additionally, another found that sagittal and coronal are better for visualizing

the areas of disease[43]. This coincides from our own experiment where we found coronal

to be the best performing. Additionally, it is important to note the negative R^2 values. This

means that the model does not match the given data well.

Figure 3. Identification of Bone Marrow Lesion on MRI scans as depicted by the area
circled in red.

Figure 4 depicts each slice of an MRI scan for one patient. It was observed that the

predicted values are not close to the true value of the bone marrow lesion. Additionally,

when observing the MRI scans, they all look different. Some of the slices have more visible

bone marrow lesion than in other slices. Examples of identified bone marrow lesions from

MRI scans can be seen in *Figure 3*. With this differing amount of bone marrow lesion per

slice, it does not make sense to use the same one true value for each slice. The true value

represents the total injection volume after being completely filled. With this observation, it

seems that the volume predictions should be different from one another, based on how

much bone marrow lesion is visible in the MRI. This can be a potential explanation for the negative R^2 values.

When examining Figure 4, certain slices have more observable bone marrow lesion and certain slices have greater prediction values. BML can be identified by areas that look white or lighter in color on the mid portion of the lateral femoral condyle and the posterior lateral tibial plateau [44]. This white portion can be seen best in slices 11 and 12 of Figure 4. These slices potentially show the bone marrow lesion best since initial cropping included slice removal centralized around the bone marrow lesion. This peaking of volume can be seen in Figure 4. When looking at slices 15 and 16 from Figure 4, a greater presence of white can be seen, but not in these areas or within the bone at all. These additionally have greater prediction values. With this increased presence of white in the slices, having a greater prediction value seems reasonable. The white areas seen here are not bone marrow lesion diseased area, but instead, an area of fat or water. Since the cropping standards were meant to bound the greatest sized bone marrow lesions, in terms of how many slices and cropped area, there are potentially too many scans for this patient. Instead of only the bone marrow lesion being inputted into the model, inaccurate slices without any bone marrow lesion were also used in the learning, which could lead to poor training of the model.

Slice	7	8	9	10	11
Prediction in cc	4.434067	4.40012	4.47432	4.54963	4.59267
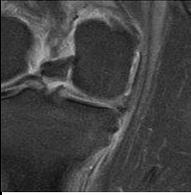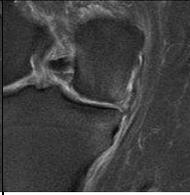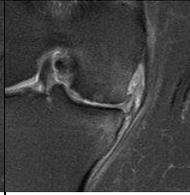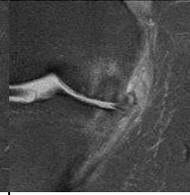
Slice	12	13	14	15	16
Prediction in cc	4.6276	4.49937	4.516747	4.6479	4.685696
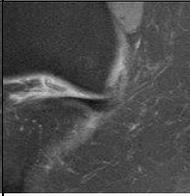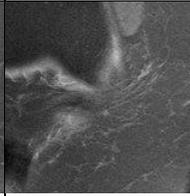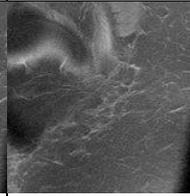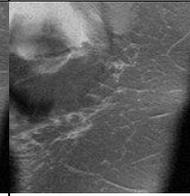
Figure 4. Each slice of one patient's MRI scans. Numbered from 7-16 is the slice number in the full series before slicing. The numbers below them are the predicted volumes for each scan. The true value for this patient was 5 cc.

It is also important to note that this prediction was based on the training of MRI images from Dr. McMillian and Dr. Cohen, and volume of treatment by the respective physician. This poses a complication with the novelty of the dataset. The total patient count was relatively low –35 patients. A greater patient dataset will allow for a wider grasp of bone marrow lesions at different volumes of injection.

4.4. Conclusion

The slice by slice model provides an unviable measurement for the prediction of bone marrow lesion volumes due to the relationship between each slice of the MRI and the total injected volume. Even though this is an unviable measurement for the prediction, a

baseline model for comparison is set and can be used in comparison with other models. With how the images are loaded into the model being the primary concern, a new method for reading in the images was required where each set of MRI slices can be given a true value instead of each slice individually.

Chapter 5

3D Regression CNN Model With Tensor Stacked Images

5.1. Introduction

Specific aim 2 sets out to improve upon complications with the slice by slice model, primarily how the MRI scans can be utilized with the true value being the entire bone marrow lesion and the negative R^2 score, discussed in Section 4.3. This will be accomplished by using a 3D Regression CNN model.

The main difference between this experimental model versus the previously employed one is the use of tensors, or the stacking the images together. This multi-dimensional array, consisting of number of patients, number of images, and the size of the images, creates a method so every slice from the MRI scan can be read as one patient. This singular patient becomes one data point that gets assigned their true value. Use of a 3D regression CNN can be seen by Masaru, Pardakhti, or Sturmfels in age estimation studies [29], [39], [40]. Within these, they were able to use brain MRI scans as input to their CNN, allowing them to estimate the age of the patient or the brain age of the patient. In another article from Bellary, in order to input their MRI scans of the knee for segmentation analysis, they utilized tensors [32]. These tensors had the form picture numbers x image height x picture width x picture depth. Through using this tensor input, they were able to achieve a classification model with 99% accuracy. Using a stacked tensor can potentially allow for better results and a more accurate representation of the data and injection volume.

The effectiveness of using these stacked images in a 3D regression CNN will determined by certain metrics -- MAE, MSE, RMSE, and R^2. MAE and RMSE can be seen used in similar research such as the estimation of ejection fraction from Inomata [45]. Within their study, MAE and RMSE were used as metrics to compare the results from the image regression model. MSE can also be seen in a 3D pose regression CNN study as a minimizer for their model [46], as well as a metric for evaluation in a model to predict clinical disability using FLAIR MRIs [47]. R^2 was additionally seen in the prediction of brain age using 3D regression CNNs [48]. This allowed for a comparison between different networks being tested.

Therefore, the purpose of this specific aim is to improve upon preexisting models by using tensor stacked MRI images and determine the validity of the model.

5.2. Methods

5.2.1. Preprocessing

MRI scans were preprocessed in the same way as in Section 4.2.1. After downloading, MRIs were typically of shape 256 x 256 voxels or 512 x 512 voxels and contained 20-30 images per patient per viewing plane. An 80mm x 80mm bounding box was drawn around the center of the bone marrow lesion to capture the largest sized bone marrow lesion possible. This step was completed using ImageJ. Once the size of the image was reduced, slices were selected to be used in the model. The largest bone marrow lesions spanned over 10 slices, so in order to encapsulate only the bone marrow lesion for this

patient, the minimum number of slices needed was 10. The bounding box of 80mm x 80mm

and slice removal until only 10 slices remained was done on all images.

5.2.2. 3D CNN Architecture

The input of the patient by patient model were tensor stacked slices of an MRI scan.

Each of these tensors represents one patient, and each tensor was assigned its true value of

injection volume as used by physician. Once these images were fed into the model, they

went through a 3D convolution layer, 3D max pooling, 3D convolution layer 2, 3D max

pooling layer 2, a flatten, 2 RELU activation dense layers, and finally a linear activation

dense layer. This architecture can be seen in Table 3.

Table 3

Patient by Patient 3D CNN Architecture

Layer	Output shape	Parameter Values
Conv1	8 x 254 x 254 x 64	5248
MaxPool1	4 x 127 x 127 x 64	0
Conv2	2 x 125 x 125 x 64	110,656
MaxPool2	1 x 62 x 62 x 64	0
Flatten	246016	0
Dense1	64	15,745,088
Dense2	64	4160
Dense3	1	65

5.2.3. Dataset

This study focuses on three magnetic resonance (MR) sequences of 35 preoperative knees from Dr. Sean McMillan Orthopedics & Sports Medicine and Rothman Orthopedic Institute. All patients have been affected by and treated for their bone marrow lesions through a subchondroplasty procedure in their femoral or tibial condyle. Before treatment, a preoperative MRI scan was taken so the areas of infection can be identified. Postoperative injection volume of calcium phosphate compound was recorded and exists as the ground truth. This model included 10 images per patient, 350 images total, for each viewing plane (coronal, sagittal, and axial). Subject data selected for the study are required to have all sequences acquired.

Two patient totals, 1 from the training dataset and 1 from the testing dataset were temporarily removed from the overall dataset. As seen in Figure 5, there are large discrepancies between the bone marrow lesion volumes of the two patients. These differences could skew the datasets and produce inaccurate results. With this, the resultant patient pool is 33 patients total; 26 patients belonged to the training dataset and 7 patients belonged to the testing dataset. Before training, all data from the training and testing were normalized so the true volume of bone marrow lesion was between 0 and 1, just as done in the slice by slice model.
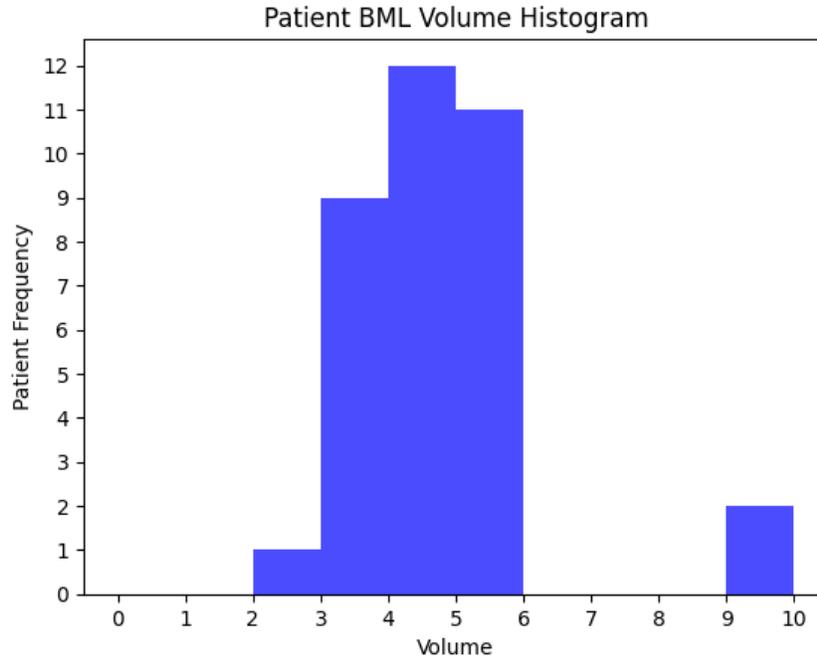
Figure 5. Histogram of entire dataset (training and testing datasets), showing the number of patients with a certain bone marrow lesion volume

5.2.4. Experimental Conditions

The Keras library with its 'ImageDataGenerator' was utilized to apply augmentation techniques, aiming to enhance diversity and robustness of the training dataset. A rescale at 1./255, rotation of 20 degrees, shear by a factor of 0.2, zoom into the image by a factor of 0.2, shifts in the horizontal and vertical directions by 20% of its width and height, horizontal flip, and brightness transformations from -0.5 to 0.5 were implemented as experimental conditions.

5.2.5. Metrics for Evaluation

Through using the 3D CNN regression architecture, as shown in Table 3, all of the slices from one MRI scan were used as the input. One tensor was created for each patient and their MRI scan. By doing this, all of the slices from that particular viewing plane are read together as one instead of as separate slices. Since each tensor corresponds to one patient's MRI scans, each true value for the bone marrow lesions can be assigned to one tensor. Through this method, all slices from a patient are considered per scan and aids in the resolution of the previous complication where each slice was being assigned an incorrect true value.

5.3. Results and Discussion

The primary objective of this specific aim was to develop a 3D regression CNN model that was able to take in tensor stacked MRI images as its input, allowing for the prediction of subchondorplasty procedure injection volumes. The neural network completed predictions on the five patients in the test dataset. This can be seen in Table 4. There were three patients with 5 cc of actual injection volume and the other two patients had 3.5 cc and 10 cc. With this, it seems that many of the output predictions are close to 5 cc, regardless if it was 3.5 or 10 cc. Even though the predictions are close together, there are slight deviations between the predictions. When looking at patient 3, they have the lowest actual injection volume from the dataset. This patient having the lowest prediction volume can be seen when utilizing the coronal or sagittal models. A similar observation can be made about the patient with the greatest injection volume, patient 4. The coronal

and sagittal models once again have the greatest patient prediction for the patient with largest actual injection volume. In order to better understand why some of these are poor predictions, it is important to look at the loss function convergence. This can help determine if a model is overfitting [49].

Table 4

Actual vs Predicted Volume of Injection for Axial, Coronal, and Sagittal Viewing Planes

	Patient 1	Patient 2	Patient 3	Patient 4	Patient 5
Actual Injection Volume	5 cc	5 cc	3.5 cc	10 cc	5 cc
Axial Predicted Volume	4.58867	4.49495	4.79698	4.57310	4.40498
Coronal Predicted Volume	3.96252	4.86794	3.81790	4.94882	3.8499
Sagittal Predicted Volume	5.43437	5.50570	5.40366	5.64669	4.58240
Note. The actual values for each patient were set based on the value that the clinician injected during the procedure.					

When observing the MSE over epoch graphs, Figure 6, for each of these models, the loss function converged. This was monitored using early stopping implementations with a patience of 5, so if no improvement in training performance was observed for 5 consecutive epochs, training would be finished and the best model from the ones trained

would be used. Such implementations help prevent the model from overly learning the training data to the point where it cannot generalize on unseen data, otherwise known as overfitting.

| Axial loss graph | Coronal loss graph | Sagittal loss graph |
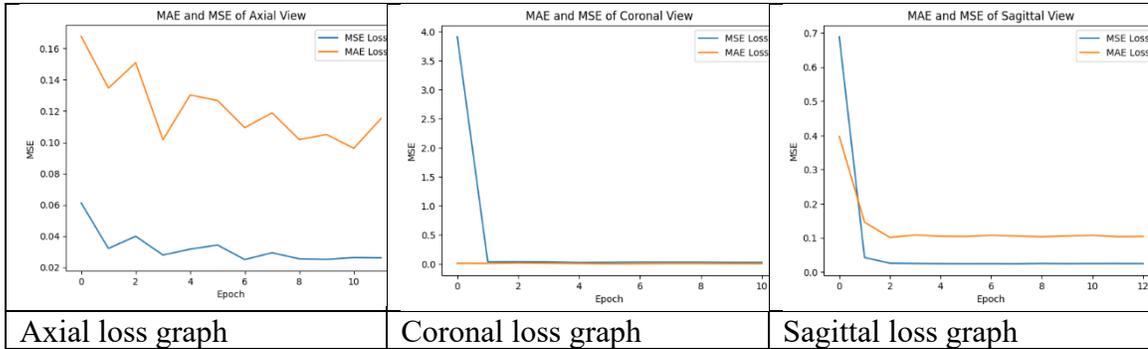
Figure 6. MSE over Epoch graphs for Axial, Coronal, and Sagittal viewing planes

Table 5 summarizes the results from the patient by patient regression CNN model for the axial, coronal, and sagittal views from the test dataset, showing the MAE, MSE, RMSE, and R^2 outputs [45]–[48]. The lowest MAE, MSE, and RMSE was shown in the axial patient by patient model. The greatest R^2 value was additionally shown in the axial patient by patient model. These outputs suggest that the model using tensor stacked axial MRI images performed the best. This differs from other studies [38], [43], where the coronal viewing plane was determined as the suggested viewing plane. All patient by patient models had improved performance compared to the slice by slice model, regardless of the viewing plane. The MSE, RMSE, and R^2 from the patient by patient model were improvements from the slice by slice model(Table 4 & Table 5). Even with the

30

improvement seen, these two different models cannot be directly compared. Additionally, the resultant R^2 values from the 3D regression model are still negative. Although nearing zero, the R^2 value for the coronal, sagittal, and axial viewing planes were negative. This means that the model performed worse than a baseline model that guesses the average of the pooled R^2 output. Figure 7, Figure 8, and Figure 9 show the actual bone marrow lesion injection volume against the predicted bone marrow lesion volume. Optimally, a linear relationship will have been formed and a rough linear line will be depicted on the graph [50]. The coronal test, shown in Figure 7, has datapoints in a rough linear line. The other two results, Figure 8 and Figure 9, have greater dispersion of results and less formed linearity. With a greater number of testing datapoints, stronger relationships may be formed, increasing the R^2 value. Overall, this shows promise for switching to a 3D regression model, since there are overall improvements compared to the slice by slice model.

Table 5

MAE, MSE, RMSE, and R^2 of Patient by Patient Model

Metric		Patient by Patient Model
Axial	MAE	0.22875
	MSE	0.12311
	RMSE	0.35088
	R2	-0.28676
Coronal	MAE	0.23417
	MSE	0.13049
	RMSE	0.36124
	R2	-0.36385
Sagittal	MAE	0.21151
	MSE	0.08948
	RMSE	0.29914
	R2	0.06476

Note. Metric evaluations from the patient by patient CNN model. The MSE, RMSE, and R^2 come from the test dataset.
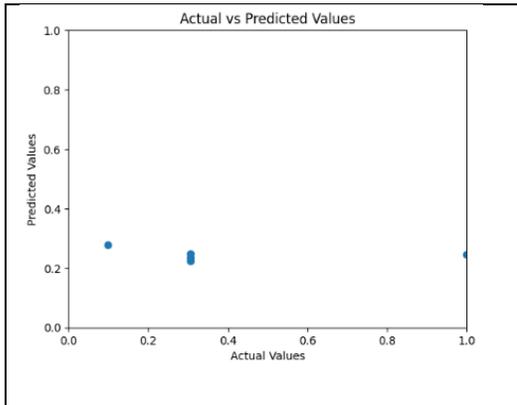
Figure 7. Axial actual vs predicted test dataset values
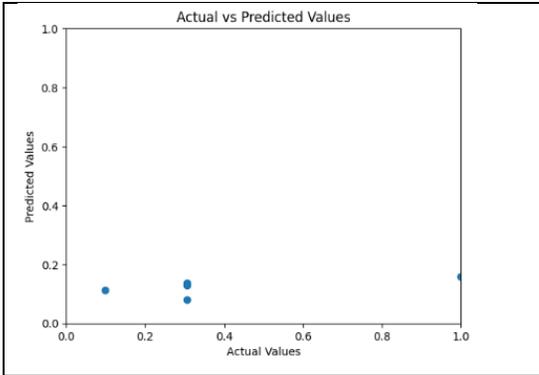
Figure 8. Coronal actual vs predicted
test dataset values
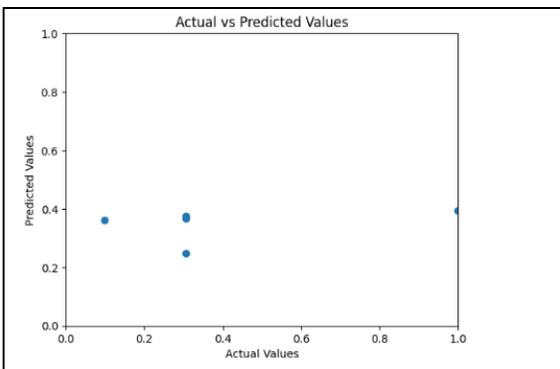
Figure 9. Sagittal actual vs predicted test
dataset values

For further testing, 2 patients were temporarily removed from the total dataset due

to the possibility of producing inaccurate results. This removal was only temporary since

more data could be acquired and have bone marrow lesion volumes similar to theirs. Figure

5 shows the discrepancy between these two patients and the other patients. In addition, to

have more patients for the testing dataset, the validation dataset was removed. Table 6

shows the output comparison using the 35 patient data pool to using the 33 patient modified

data pool. Through this new method, most metrics saw improvement. Only the axial and

sagittal MAE declined after using the 33 patient model. The MSE and RMSE both

decreased from 35 patients to 33 patients for all three viewing planes. In addition, the R^2

increased for all three viewing planes as well. The positive R^2 value can be seen on the

predicted versus actual graphs shown in Figure 7, Figure 8, and Figure 9. These improved

results can be indicative of an improved model.

Table 6

Metric Evaluations From the Patient by Patient CNN Model When Using Input With and Without Potential Outlying Patients

	With Outlying Patients (35 patients)	Without Outlying Patients (33 Patients)
Axial MAE	0.22875	0.257228
Axial MSE	0.12311	0.083363
Axial RMSE	0.35088	0.28873
Axial R2	-0.28676	0.03558
Coronal MAE	0.23417	0.173305
Coronal MSE	0.13049	0.0404458
Coronal RMSE	0.36124	0.20111
Coronal R2	-0.36385	0.53209
Sagittal MAE	0.21151	0.2374108
Sagittal MSE	0.08948	0.066409
Sagittal RMSE	0.29914	0.25770
Sagittal R2	0.06476	0.23172
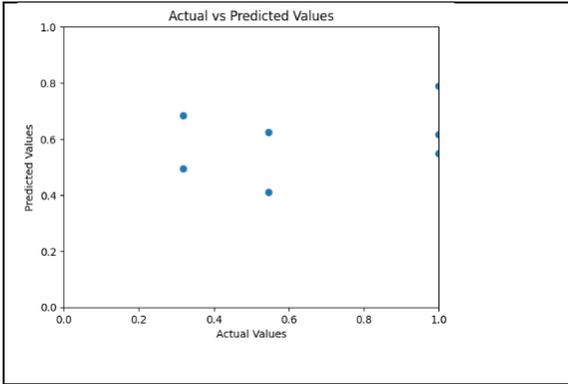
Figure 10. Axial actual vs predicted test dataset values
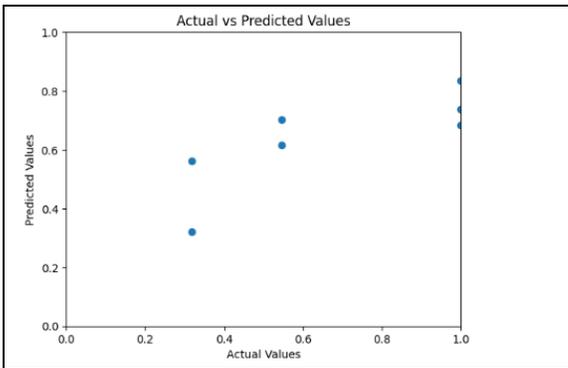
Figure 11. Coronal actual vs predicted test dataset values
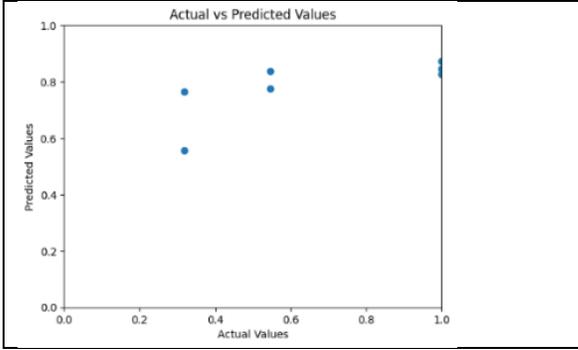
Figure 12. Sagittal actual vs predicted test
dataset values

5.4. Conclusion

The tensor stacked patient by patient model establishes a different way to process

these MRI images, and this resulted in improved prediction output from the model.

Through this, it appears that the coronal viewing plane allows for the best model creation

and prediction. Even though these results are promising and have improvement from the

initial slice by slice model, there are improvements that can be implemented such as a

multi-plane input into the model. This provides multiple viewing planes as the input for

one CNN.

Chapter 6

3D Regression Using Multiplanar Tensor Stacked Images

6.1. Introduction

Using tensor stacked images from one viewing plane in a 3D regression CNN allowed for promising results that improved upon the baseline model. With this, even better results are sought and can potentially be accomplished by utilizing multiple viewing angles in a singular CNN. In a knee injury detection article, the authors compared a single plane analysis to a multi-plane analysis and discovered that using multiple planes improved their area under the curve results [51]. Therefore, specific aim 3 utilizes the existing 3D regression model by combining multiple viewing planes within each tensor. Since a tensor is a multi-dimensional array, multiple images can be used per tensor, as we saw in Chapter 5. With this concept, it opens the ability of combining multiple viewing planes into one tensor. Similar multiplanar tensors have been used, such as in the article from Khan [52]. Within this, they sought to segment knee tissue from MRI images, and by using this tensor constructed MRI representation, they were able to achieve high accuracy scores. Even though the purpose of the study was for knee segmentation purposes, adapting the input to apply towards a 3D regression model may allow for a better resultant model.

In order to determine the validity and performance of the 3D regression model with multiple viewing plane stacked tensor input, the same metrics will be observed from Chapter 4 and Chapter 5. These metrics include the MSE, MAE, RMSE, and R^2. During

training, the convergence of the loss function will also be monitored to ensure that training does not overfit.

Therefore, the purpose of this specific aim is to improve upon preexisting models by using tensor stacked MRI images from various viewing planes and compare the resultant metrics to determine the validity of the model.

6.2. Methods

6.2.1. Preprocessing

MRI scans were preprocessed in the same way as seen in Section 4.2.1. After downloading, MRIs were typically of shape 256 x 256 voxels or 512 x 512 voxels and contained 20-30 images per patient per viewing plane. An 80mm x 80mm bounding box was drawn around the center of the bone marrow lesion to capture the largest sized bone marrow lesion possible. This step was completed using ImageJ. Once the size of the image was reduced, slices were selected to be used in the model. The largest bone marrow lesions spanned over 10 slices, so in order to encapsulate only the bone marrow lesion for this patient, the minimum number of slices needed was 10. The bounding box of 80mm x 80mm and slice removal until only 10 slices remained was done on all images.

6.2.2. 3D CNN Architecture

The input of the patient by patient model were tensor stacked slices of an MRI scan. Each of these tensors represents one patient, and each tensor was assigned its true value of injection volume as used by physician. Once these images were fed into the model, they went through a 3D convolution layer, 3D max pooling, 3D convolution layer 2, 3D max

pooling layer 2, a flatten, 2 RELU activation dense layers, and finally a linear activation dense layer. This architecture can be seen in Table 3.

6.2.3. Dataset

This study focuses on three magnetic resonance (MR) sequences of 35 preoperative knees from Dr. Sean McMillan Orthopedics & Sports Medicine and Rothman Orthopedic Institute. All patients have been affected by and treated for their bone marrow lesions through a subchondroplasty procedure in their femoral or tibial condyle. Before treatment, a preoperative MRI scan was taken so the areas of infection can be identified. Postoperative injection volume of calcium phosphate compound was recorded and exists as the ground truth. This model included 10 images per patient, 350 images total, for each viewing plane (coronal, sagittal, and axial). Subject data selected for the study are required to have all sequences acquired.

Through using the 3D CNN regression architecture, as shown in Table 3, all of the slices from one MRI scan were used as the input. One tensor was created for each patient and their MRI scan. By doing this, all of the slices from that particular viewing plane are read together as one instead of as separate slices. Since each tensor corresponds to one patient's MRI scans, each true value for the bone marrow lesions can be assigned to one tensor. Through this method, all slices from a patient are considered per scan and aids in the resolution of the previous complication where each slice was being assigned an incorrect true value.

6.2.4. Experimental Conditions

The Keras library with its 'ImageDataGenerator' was utilized to apply augmentation techniques, aiming to enhance diversity and robustness of the training dataset. A rescale at 1./255, rotation of 20 degrees, shear by a factor of 0.2, zoom into the image by a factor of 0.2, shifts in the horizontal and vertical directions by 20% of its width and height, horizontal flip, and brightness transformations from -0.5 to 0.5 were implemented as experimental conditions. In addition, an early stopping mechanism was implemented to help prevent overfitting during training. This early stopping utility had a patience of 5 epoch. The models were trained with a batch size of 1 and Adam optimization with a learning rate of 0.01.

6.2.5. Metrics for Evaluation

Similar to previous chapters, MAE, MSE, RMSE, and R^2 were used as metrics for evaluation.

6.3. Results and Discussion

The primary objective of this aim was to determine if combinations of viewing planes result in better predictions from the CNN model. Table 7 depicts the actual vs predicted volumes of injection for each of the different models. Observations show that some predictions are close to the actual value, and others are drastically off. This results in relatively poor prediction volumes. When looking at the output from the axial and coronal tensor model, many of its predictions are of similar values. This can also be seen for the

axial and sagittal tensor model. A potential reason for these poor results can be due to the complexity of the model. If this model was too simple, it might not be able to pick up small patterns or features. This would cause underfitting of the model. In this case, the easiest way to optimize the prediction is to output a fixed average value for all input. This low complexity can come from the type of learning, number of trainable parameters, and the extracted features [53]. On the other hand, too complex of a model can lead to overfitting [54]. Underfitting and overfitting can be better seen when observing the MSE over epoch graphs and the actual versus predicted graphs.

Table 7

Actual vs Predicted Volumes of Injection for Combinations of Viewing Planes

	Patient 1	Patient 2	Patient 3	Patient 4	Patient 5	Patient 6	Patient 7
Actual Injection Volume (cc)	5	3.5	5	4	3.5	5	4
Axial, Coronal, and Sagittal Tensor Volume Prediction	3.368621	3.323306	3.661082	3.791838	4.056524	3.548677	3.85672
Axial and Coronal Tensor Volume Prediction	3.051375	3.049541	3.007726	3.019025	3.022358	3.038955	3.004276
Axial and Sagittal Tensor Volume Prediction	3.001431	3.015057	2.962816	3.004389	2.957187	3.004678	3.003724
Coronal and Sagittal Volume Prediction	3.41688	3.843759	4.183746	4.813856	3.717685	4.490603	3.802804

Figure 13 shows the MSE and MAE graphs over epochs. With this, we can see that for every model, the graph properly converged following the early stopping implementations. This convergence signifies that the model will not improve any further if allowed to go additional epochs.

Axial, Coronal, and Sagittal Graph	Axial and Coronal Graph	Axial and Sagittal Graph	Coronal and Sagittal Graph
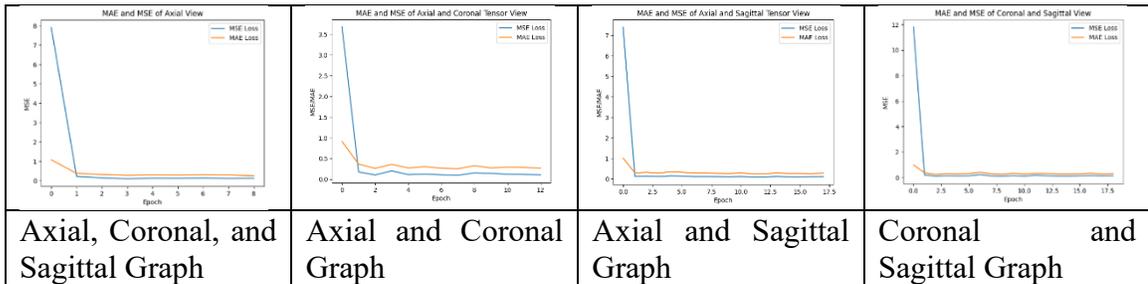
Figure 13. MSE and MAE graphs over epoch for different combinations of viewing planes.

Figure 14 shows the actual versus predicted volumes of injection for each of the different combinations of viewing planes. With linear regression, the optimal model would have predictions that follow a diagonal line across the graph from the bottom left corner to the top right. These graphs additionally aid in determining overfitting or underfitting by depicting how well the data points follow the optimal line. The axial, coronal, and sagittal plane model and the coronal and sagittal model appear to have predictions that could possibly be on the optimal line. The other two models have very poor predictions, and these predictions are all under the optimal line. This could possibly mean that the model was underfit for the axial and coronal and axial and sagittal models. Additionally, when observing the MAE, MSE, RMSE, and R^2, the axial and coronal and the axial and sagittal

43

models both performed the worst. This can be seen in Table 8. Out of the four models, the coronal and sagittal viewing planes model performed the best since it had the lowest MAE, MSE, RMSE, and the highest R^2. Although, additionally on this table there is the coronal only viewing plane model for comparison. This model performed the best when initially testing the 3D regression CNN, and the coronal only viewing plane had the best results.

Axial, Coronal, and Sagittal Planes	Axial and Coronal Planes	Axial and Sagittal Planes	Coronal and Sagittal Planes
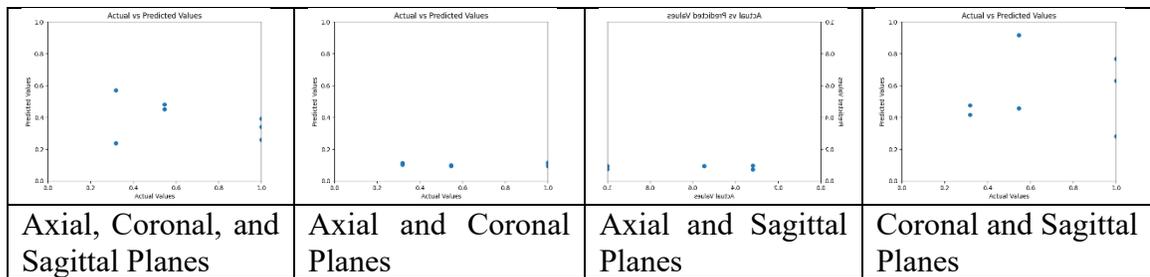
Figure 14. Actual vs predicted results from different tensor combinations of viewing planes its input. All predictions are made on the testing dataset

Table 8

MAE, MSE, RMSE, and R^2 of Models using Different Combinations of Viewing Planes. The Coronal Output is From the Previous Model

	Coronal	Ax, Cor, Sag	Ax, Cor	Ax, Sag	Cor, Sag
MAE	0.173305	0.20558	0.41312	0.43209	0.12688
MSE	0.0404458	0.35755	0.57186	0.58770	0.29099
RMSE	0.20111	0.45342	0.64275	0.65734	0.35621
R^2	0.53209	-1.37837	-3.77940	-3.99880	-0.46789

6.4. Conclusion

The usage of the 3D regression CNN model with multiplanar inputs allowed for combinations of multiple viewing planes. From the combinations tested, the model which combined the coronal and sagittal views performed the best but was still worse than when only using the coronal viewing plane. Even though the models which combined viewing planes performed worse, using multiple viewing planes opens the possibility of using combinations of images as the input.

Chapter 7

Project Summary and Future Work

7.1. Project Summary

The aims of this research were to create a regression CNN that would take preoperative MRI scans of patients with BMLs as input, and output predicted volumes of injection for that BML. There is a need for the prediction of BML volume of injection because of further pain that can come from overfilling or underfilling during the subchondroplasty procedure [8], [11]. The regression CNN model viability was measured through the actual versus predicted volumes of injection, the convergence of the loss function and the regression metrics MAE, MSE, RMSE, and R^2.

Results from testing 2D and 3D CNN models with varying viewing planes as inputs revealed that utilizing 3D regression CNNs with a tensor stacked image input from the coronal view provided the best model. This model provided the best predictions, resulting in the lowest MAE, MSE, RMSE, and the highest R^2.

7.2. Future Work

These results demonstrate the potential for training a 3D regression CNN with novel preoperative MRI scans of patients with BMLs in order to create a prediction algorithm. The methods outlined in this research can be used to evaluate and compare against future iterations of the regression model. Going forward, potential avenues for testing include using T1 modality MR scans [55], increasing the number of patients being

put into the training and testing datasets, and increasing the complexity of the model [40]. The addition of T1 modality MR scans can potentially allow the CNN to learn more, simply by having access to more and potentially better information about each MRI scan. Increasing the dataset may allow the model to draw better generalizations about the training set, allowing for better results and predictions. This increase in data can also allow us to increase the complexity of the model, but may increase the cost of computation and slow down the performance [56]. Overall, the impact of this research allows for preliminary volume of injection predictions for patients with BMLs and may aid the doctors in patients by preventing overfilling and the related pain.

References

[1]   M. Cross *et al.*, "The global burden of hip and knee osteoarthritis: Estimates from the Global Burden of Disease 2010 study," *Ann. Rheum. Dis.*, vol. 73, no. 7, pp. 1323–1330, 2014, doi: 10.1136/annrheumdis-2013-204763.

[2]   D. T. Felson *et al.*, "Bone Marrow Edema and Its Relation to Progression of Knee Osteoarthritis," *Ann. Intern. Med.*, vol. 139, no. 5 I, 2003, doi: 10.7326/0003-4819-139-5_part_1-200309020-00008.

[3]   E. Yusuf, M. C. Kortekaas, I. Watt, T. W. J. Huizinga, and M. Kloppenburg, "Do knee abnormalities visualised on MRI explain knee pain in knee osteoarthritis? a systematic Review," *Ann. Rheum. Dis.*, vol. 70, no. 1, pp. 60–67, 2011, doi: 10.1136/ard.2010.131904.

[4]   C. Scher, J. Craig, and F. Nelson, "Bone marrow edema in the knee in osteoarthrosis and association with total knee arthroplasty within a three-year follow-up," *Skeletal Radiol.*, vol. 37, no. 7, pp. 609–617, 2008, doi: 10.1007/s00256-008-0504-x.

[5]   H. A. Oliver *et al.*, "Enhanced Subchondroplasty Treatment for Post-Traumatic Cartilage and Subchondral Bone Marrow Lesions in a Canine Model," *J. Orthop. Res.*, vol. 38, no. 4, pp. 740–746, 2020, doi: 10.1002/jor.24508.

[6]   M. R. Klement and P. F. Sharkey, "The Significance of Osteoarthritis-associated Bone Marrow Lesions in the Knee," *J. Am. Acad. Orthop. Surg.*, vol. 27, no. 20, pp. 752–759, 2019, doi: 10.5435/JAAOS-D-18-00267.

[7]   N. M. Krebs, J. L. Kehoe, M. J. Van Wagner, and C. Rios-Bedoya, "The Efficacy of Subchondroplasty for the Treatment of Knee Pain Associated with Bone Marrow Lesions," *Spartan Med. Res. J.*, vol. 4, no. 2, pp. 1–7, 2020, doi: 10.51894/001c.11767.

[8]   S. B. Cohen and P. F. Sharkey, "Subchondroplasty for Treating Bone Marrow Lesions," *J. Knee Surg.*, vol. 29, no. 7, pp. 555–563, 2016, doi: 10.1055/s-0035-1568988.

[9]   S. K. Tanamas *et al.*, "Bone marrow lesions in people with knee osteoarthritis predict progression of disease and joint replacement: A longitudinal study," *Rheumatology*, vol. 49, no. 12, pp. 2413–2419, 2010, doi: 10.1093/rheumatology/keq286.

[10] B. J. Rebolledo, K. M. Smith, and J. L. Dragoo, "Hitting the Mark: Optimizing the Use of Calcium Phosphate Injections for the Treatment of Bone Marrow Lesions of the Proximal Tibia and Distal Femur," *Arthrosc. Tech.*, vol. 7, no. 10, pp. e1013–e1018, 2018, doi: 10.1016/j.eats.2018.06.006.

[11] O. A. Brimmo *et al.*, "Subchondroplasty for the treatment of post-traumatic bone marrow lesions of the medial femoral condyle in a pre-clinical canine model," *J. Orthop. Res.*, vol. 36, no. 10, pp. 2709–2717, 2018, doi: 10.1002/jor.24046.

[12] B. Antony *et al.*, "Correlates of knee bone marrow lesions in younger adults," *Arthritis Res. Ther.*, vol. 18, no. 1, pp. 1–9, 2016, doi: 10.1186/s13075-016-0938-9.

[13] E. Guymer *et al.*, "A study of the prevalence and associations of subchondral bone marrow lesions in the knees of healthy, middle-aged women," *Osteoarthr. Cartil.*, vol. 15, no. 12, pp. 1437–1442, 2007, doi: 10.1016/j.joca.2007.04.010.

[14] S. Ota *et al.*, "Symptomatic bone marrow lesions induced by reduced bone mineral density in middle-aged women: A cross-sectional Japanese population study," *Arthritis Res. Ther.*, vol. 21, no. 1, pp. 1–9, 2019, doi: 10.1186/s13075-019-1900-4.

[15] R. A. Ghasemi, S. Sadeghi, N. Rahimee, and M. Tahmasebi, "Technologies in the Treatment of Bone Marrow Edema Syndrome," *Orthop. Clin. North Am.*, vol. 50, no. 1, pp. 131–138, Jan. 2019, doi: 10.1016/J.OCL.2018.08.008.

[16] N. Aigner *et al.*, "Bone marrow edema syndrome of the femoral head: Treatment with the prostacyclin analogue iloprost vs. core decompression: An MRI-controlled study," *Wien. Klin. Wochenschr.*, vol. 117, no. 4, pp. 130–135, 2005, doi: 10.1007/s00508-005-321-3.

[17] K. Chua *et al.*, "Subchondroplasty for Bone Marrow Lesions in the Arthritic Knee Results in Pain Relief and Improvement in Function," *J. Knee Surg.*, vol. 34, no. 6, pp. 665–671, 2021, doi: 10.1055/s-0039-1700568.

[18] L. N. Nairn, M. Subramaniam, S. Ekhtiari, D. E. Axelrod, J. A. Grant, and M. Khan, "Safety and early results of Subchondroplasty® for the treatment of bone marrow lesions in osteoarthritis: a systematic review," *Knee Surgery, Sport. Traumatol. Arthrosc.*, vol. 29, no. 11, pp. 3599–3607, 2021, doi: 10.1007/s00167-020-06294-w.

[19] J. Farr and S. B. Cohen, "Expanding applications of the subchondroplasty procedure for the treatment of bone marrow lesions observed on magnetic resonance imaging," *Oper. Tech. Sports Med.*, vol. 21, no. 2, pp. 138–143, 2013, doi: 10.1053/j.otsm.2013.03.006.

[20] S. B. Cohen and P. F. Sharkey, "Surgical treatment of osteoarthritis pain related to subchondral bone defects or bone marrow lesions: Subchondroplasty," *Tech. Knee Surg.*, vol. 11, no. 4, pp. 170–175, 2012, doi: 10.1097/BTK.0b013e318276e1fe.

[21] M. B. Bonadio, P. N. Giglio, C. P. Helito, J. R. Pécora, G. L. Camanho, and M. K. Demange, "Subchondroplasty for treating bone marrow lesions in the knee – initial experience," *Rev. Bras. Ortop. (English Ed.*, vol. 52, no. 3, pp. 325–330, 2017, doi: 10.1016/j.rboe.2017.04.003.

[22] W. K. Conaway, R. Agrawal, M. R. Nazal, J. W. Stelzer, and S. D. Martin, "Changing MRI after subchondroplasty with partial meniscectomy for knee osteoarthritis," *Clin. Imaging*, vol. 56, no. February, pp. 13–16, 2019, doi: 10.1016/j.clinimag.2019.02.016.

[23] D. Kawahara and Y. Nagata, "T1-weighted and T2-weighted MRI image synthesis with convolutional generative adversarial networks," *Reports Pract. Oncol. Radiother.*, vol. 26, no. 1, pp. 35–42, 2021, doi: 10.5603/RPOR.a2021.0005.

[24] S. Mittal, G. Pradhan, S. Singh, and R. Batra, "T1 and t2 mapping of articular cartilage and menisci in early osteoarthritis of the knee using 3-tesla magnetic resonance imaging," *Polish J. Radiol.*, vol. 84, pp. e549–e564, 2019, doi: 10.5114/PJR.2019.91375.

[25] M. T. Nieminen *et al.*, "Quantitative MR microscopy of enzymatically degraded articular cartilage," *Magn. Reson. Med.*, vol. 43, no. 5, pp. 676–681, 2000, doi: 10.1002/(SICI)1522-2594(200005)43:5<676::AID-MRM9>3.0.CO;2-X.

[26] S. Chakrabarty, A. Sotiras, M. Milchenko, P. Lamontagne, M. Hileman, and D. Marcus, "MRI-based identification and classification of major intracranial tumor types by using a 3D convolutional neural network: A retrospective multi-institutional analysis," *Radiol. Artif. Intell.*, vol. 3, no. 5, 2021, doi: 10.1148/ryai.2021200301.

[27] K. Muhammad, S. Khan, J. Del Ser, and V. H. C. D. Albuquerque, "Deep Learning for Multigrade Brain Tumor Classification in Smart Healthcare Systems: A Prospective Survey," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 32, no. 2, pp. 507–522, 2021, doi: 10.1109/TNNLS.2020.2995800.

[28] E. J. van Kempen *et al.*, "Accuracy of machine learning algorithms for the classification of molecular features of gliomas on mri: A systematic literature review and meta-analysis," *Cancers (Basel).*, vol. 13, no. 11, pp. 9638–9653, 2021, doi: 10.3390/cancers13112606.

[29] N. Pardakhti and H. Sajedi, "Brain age estimation based on 3D MRI images using 3D convolutional neural network," *Multimed. Tools Appl.*, vol. 79, no. 33–34, pp. 25051–25065, 2020, doi: 10.1007/s11042-020-09121-z.

[30] V. Rajinikanth, S. Kadry, and Y. Nam, "Convolutional-neural-network assisted segmentation and svm classification of brain tumor in clinical mri slices," *Inf. Technol. Control*, vol. 50, no. 2, pp. 342–356, 2021, doi: 10.5755/j01.itc.50.2.28087.

[31] P. K. Lam *et al.*, "Accurate brain age prediction using recurrent slice-based networks," p. 32, 2020, doi: 10.1117/12.2579630.

[32] M. Zakir Bellary, F. Fameeza, and D. B. A. Musthafa, "The Mri Knee Pain Classification Using Cnn Algorithm And Segmentation Using Clustering Algorithm," *Turkish J. Comput. Math. Educ.*, vol. 12, no. 10, pp. 306–315, 2021.

[33] M. El Adoui, S. A. Mahmoudi, M. A. Larhmam, and M. Benjelloun, "MRI breast tumor segmentation using different encoder and decoder CNN architectures," *Computers*, vol. 8, no. 3, 2019, doi: 10.3390/computers8030052.

[34] S. Sarkar, A. Kumar, S. Chakraborty, S. Aich, J.-S. Sim, and H.-C. Kim, "A CNN based approach for the detection of brain tumor using MRI scans," *Test Eng. Manag.*, vol. 83, no. June, pp. 16580–16586, 2020.

[35] C. De Vente, P. Vos, M. Hosseinzadeh, J. Pluim, and M. Veta, "Deep Learning Regression for Prostate Cancer Detection and Grading in Bi-Parametric MRI," *IEEE Trans. Biomed. Eng.*, vol. 68, no. 2, pp. 374–383, 2021, doi: 10.1109/TBME.2020.2993528.

[36] Ö. İnik, M. Altıok, E. Ülker, and B. Koçer, "MODE-CNN: A fast converging multi-objective optimization algorithm for CNN-based models," *Appl. Soft Comput.*, vol. 109, pp. 4–17, 2021, doi: 10.1016/j.asoc.2021.107582.

[37] C. P. Helito *et al.*, "Evaluation of the anterolateral ligament of the knee by means of magnetic resonance examination," *Rev. Bras. Ortop. (English Ed.*, vol. 50, no. 2, pp. 214–219, 2015, doi: 10.1016/j.rboe.2015.03.009.

[38] T. Magee and D. Williams, "Detection of meniscal tears and marrow lesions using coronal MRI," *Am. J. Roentgenol.*, vol. 183, no. 5, pp. 1469–1473, 2004, doi: 10.2214/ajr.183.5.1831469.

[39] M. Ueda *et al.*, "AN AGE ESTIMATION METHOD USING 3D-CNN FROM BRAIN MRI IMAGES Graduate School of Information Sciences , Tohoku University , Japan . South China University of Technology , China . Institute of Divelopment , Aging and Cancer , Tohoku University , Japan .," *2019 IEEE 16th Int. Symp. Biomed. Imaging (ISBI 2019)*, no. Isbi, pp. 380–383, 2019.

[40] P. Sturmfels, S. Rutherford, M. Angstadt, M. Peterson, C. Sripada, and J. Wiens, "A Domain Guided CNN Architecture for Predicting Age from Structural Brain Images," *Proc. Mach. Learn. Res.*, vol. 85, pp. 295–311, 2018.

[41] M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification," *Neurocomputing*, vol. 321, pp. 321–331, 2018, doi: 10.1016/j.neucom.2018.09.013.

[42] S. Zhang *et al.*, "An investigation of CNN models for differentiating malignant from benign lesions using small pathologically proven datasets," *Comput. Med. Imaging Graph.*, vol. 77, p. 101645, 2019, doi: 10.1016/j.compmedimag.2019.101645.

[43] T. W. Hash, "Magnetic Resonance Imaging of the Knee," *Sports Health*, vol. 5, no. 1, pp. 78–107, 2013, doi: 10.1177/1941738112468416.

[44] M. Marcacci, L. Andriolo, E. Kon, N. Shabshin, and G. Filardo, "Aetiology and pathogenesis of bone marrow lesions and osteonecrosis of the knee," *EFORT Open Rev.*, vol. 1, no. 5, pp. 219–224, 2016, doi: 10.1302/2058-5241.1.000044.

[45] S. Inomata, T. Yoshimura, M. Tang, S. Ichikawa, and H. Sugimori, "Estimation of Left and Right Ventricular Ejection Fractions from cine-MRI Using 3D-CNN," *Sensors*, vol. 23, no. 14, pp. 1–13, 2023, doi: 10.3390/s23146580.

[46] S. Mahendran, H. Ali, and R. Vidal, "3D Pose Regression using Convolutional Neural Networks 6D Task : given a single 2D image , estimate 6D object pose."

[47] P. Roca *et al.*, "Artificial intelligence to predict clinical disability in patients with multiple sclerosis using FLAIR MRI," *Diagn. Interv. Imaging*, vol. 101, no. 12, pp. 795–802, 2020, doi: 10.1016/j.diii.2020.05.009.

[48] H. Jiang *et al.*, "Predicting Brain Age of Healthy Adults Based on Structural MRI Parcellation Using Convolutional Neural Networks," *Front. Neurol.*, vol. 10, no. January, 2020, doi: 10.3389/fneur.2019.01346.

[49] S. Gupta, R. Gupta, M. Ojha, and K. P. Singh, "A comparative analysis of various regularization techniques to solve overfitting problem in artificial neural network," *Commun. Comput. Inf. Sci.*, vol. 799, pp. 363–371, 2018, doi: 10.1007/978-981-10-8527-7_30.

[50] A. Schneider, G. Hommel, and M. Blettner, "Lineare regressionsanalyse - Teil 14 der serie zur bewertung wissenschaftlicher publikationen," *Dtsch. Arztebl.*, vol. 107, no. 44, pp. 776–782, 2010, doi: 10.3238/arztebl.2010.0776.

[51] N. Belton *et al.*, "Optimising Knee Injury Detection with Spatial Attention and Validating Localisation Ability," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 12722 LNCS, no. Ml, pp. 71–86, 2021, doi: 10.1007/978-3-030-80432-9_6.

[52] S. Khan, B. Azam, Y. Yao, and W. Chen, "Deep collaborative network with alpha matte for precise knee tissue segmentation from MRI," *Comput. Methods Programs Biomed.*, vol. 222, 2022, doi: 10.1016/j.cmpb.2022.106963.

[53] D. P. F. Möller, "Machine Learning and Deep Learning," *Adv. Inf. Secur.*, vol. 103, pp. 347–384, 2023, doi: 10.1007/978-3-031-26845-8_8.

[54] L. Alzubaidi *et al.*, *Review of deep learning: concepts, CNN architectures, challenges, applications, future directions*, vol. 8, no. 1. Springer International Publishing, 2021.

[55] K. Usman and K. Rajpoot, "Brain tumor classification from multi-modality MRI using wavelets and machine learning," *Pattern Anal. Appl.*, vol. 20, no. 3, pp. 871–881, 2017, doi: 10.1007/s10044-017-0597-8.

[56] J. H. Cole *et al.*, "Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker," *Neuroimage*, vol. 163, pp. 115–124, 2017, doi: 10.1016/j.neuroimage.2017.07.059.