PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
Ministry of Higher Education and Scientific Research
University of Mohamed Khider – BISKRA
Faculty of Exact Sciences, Science of Nature and Life
Computer Science Department

Ordre N° : GLSD/M2/2022

Thesis

Submitted in fulfilment of the requirements for the Masters degree in

Computer science

Option: software engineering and distributed system.

Pediatric Bone Age Assessment from Hand X-ray using Deep Learning Approach

By:
ZERARI Achouak

Members of the jury :

Ouaar Hanane	MCB	President
Laid Kahloul	PROF	Supervisor
Mohammedi Amira	MAA	Examiner

Session 2022

Acknowledgements

First and foremost, praises and thanks to **ALLAH** the most graceful and merciful, and his strength that contained me throughout my research work for its successful completion.

A special first thanks to my deceased grandmother, may God have mercy on her soul, who bought me this computer that I worked on and did not let me down during my work, thanks to God.

I would like to express my special thanks of gratitude to my Professor **L. Kahloul** who believed in me and gave me the golden opportunity to do this wonderful project **Paediatric Bone Age Assessment from Hand X-ray using Deep Learning Approach**, as well as Dr. Oussama Djedidi, Romeo Carlo, and Dr. Ikram Remadna who helped me doing this research and learning many new things. I am really thankful to them. I extend my appreciation to all teachers who helped me in my study life.

Also, I would like to mention the support and consideration of my family who have always been in my life.

Last but not least, my thanks also go to my precious friends, and classmates, having such great people around me had a huge impact on my life and pushed me forward to always give my best. To everyone I love and care about, thank you.

ZERARI Achouak

Abstract

Bone age assessments are methods that doctors use in pediatric medicine. They are used to assess the growth of children by analyzing X-ray images. This work focuses on the development of a deep learning model to estimate bone age from X-ray images. Such a model would avoid the fallacies of subjective methods and raise the accuracy of the assessment. In our work, the model is based on convolutional neural networks (CNN) and is composed of two steps: a preprocessing step generating image masks, and a prediction step that uses these masks to generate the assessment. The model is trained and tested using a public Radiological Society of North America(RSNA) bone age dataset[1]. Finally, experimental results demonstrate the effectiveness of the proposed approach compared to similar works in the literature.

Keywords: *Bone age assessment, Deep learning, Preprocessing, Machine learning, Preprocessing, Image Processing, Convolutional Neural Networks*

[1] Radiological Society of North America(RSNA).https://www.rsna.org/education/ai-resources-and-training/ai-image-challenge/rsna-pediatric-bone-age-challenge-2017

Résumé

L'évaluation de l'âge osseux est une méthode que les médecins utilisent en médecine pédiatrique. Elle est utilisée pour évaluer la croissance des enfants en analysant des images radiographiques. Ce travail porte sur le développement d'un modèle d'apprentissage profond pour estimer l'âge osseux à partir d'images radiographiques. Un tel modèle éviterait les erreurs des méthodes subjectives et augmenterait la précision de l'évaluation. Dans notre travail, le modèle est basé sur des réseaux de neurones convolutifs (CNN) et est composé de deux étapes: une étape de prétraitement générant des masques d'images, et une étape de prédiction qui utilise ces masques pour générer l'évaluation. Le modèle est entrainé et testé à l'aide du dataset "Radiological Society of North America (RSNA)" [1]. Enfin, les résultats expérimentaux démontrent l'efficacité de l'approche proposée par rapport à des travaux similaires dans la littérature.

Mots clés : *Évaluation de l'âge osseux, Apprentissage en profondeur, Prétraitement, Apprentissage automatique, Prétraitement, Traitement d'images, Réseaux de neurones convolutifs.*

[1] Radiological Society of North America (RSNA). formation/ai-image-challenge/rsna-pediatric-bone-age-challenge-2017

List of Figures

2.1 AI-assisted robotic surgery example. 6
2.2 Example of Magnetic Resonance Imaging 6
2.3 Unsupervised Learning in ML. 7
2.4 Reinforcement Learning in ML. 8
2.5 Supervised Learning in ML. 8
2.6 Classification and Regression [29]. 9
2.7 The difference between the ML and the DL. 10
2.8 Biological Neuron. 11
2.9 The artificial neuron structure. 12
2.10 RELU Function curve. 13
2.11 ELU Function curve. 14
2.12 Sigmoid Function curve. 14
2.13 Softmax Function curve. 15
2.14 Calculation process of a filter in the convolutional layer. 17
2.15 Example of zero padding added to an image. 18
2.16 Stride in CNN. 18
2.17 A practical example of max-pooling layer. 19
2.18 Convolutional Neural Network Architecture [12]. 20
2.19 Skip-connection. 21
2.20 The encircled windows correlate with bone age [27]. 22

3.1 System design. 26

3.2 Estimated bone age distribution as assessed by a trained physician. Bone ages are not evenly divided. 26
3.3 U-net architecture [35]. 28
3.4 Poor quality images examples. 28
3.5 Preprocessing pipeline. 29
3.6 The architecture of the proposed model. (a) Composition of the VGG blocks (b) The proposed VGG-style neural network architecture for regression. 30

4.1 Python 32
4.2 OpenCv 33
4.3 TensorFlow 33
4.4 Keras 33
4.5 Numpy 34
4.6 Mathplotlib 34
4.7 kaggle 34
4.8 Spyder 35
4.9 Colab 35

5.1 Evolution of the loss function described by the Mean Squared Error (MSE) over the training of the model. 37
5.2 Clustering of the predictions around every value of the BAA. 38

List of Tables

5.1 Results of the evolution of the proposed model through diferent versions. 36
5.2 Results of the model on the test set. 37
5.3 Comparison with work from the literature using the RSNA bone age data. 38

Contents

Acknowledgements i

Abstract iii

Résumé iv

List of Figures **vii**

List of Tables **viii**

1 General Introduction **1**

1.1 Introduction 1

1.2 Organisation of the dissertation 2

2 Technical Background **4**

2.1 Introduction 4

2.2 Health Care definition 4

2.2.1 Health Care types 4

2.2.2 Artificial intelligence and healthcare 5

2.3 Machine learning 7

2.4 Deep Learning 9

2.4.1 The differences between the Machine Learning and the Deep Learning . . . 10

2.4.2 Artificial neural networks 10

2.5 Convolution Neural Network 15

2.5.1 CNN definition 16

2.5.2 CNN Architecture 16

2.5.3 CNN layers configuration 19

2.5.4 Some CNN architectures 20
2.6 Bone age assessment 21
2.6.1 Bone Age Assessment Methods 22
2.7 Related work 23
2.8 Conclusion 24

3 Design and implementation of a deep learning architecture for Bone Age Assessment 25
3.1 Introduction 25
3.2 The proposed approach 25
3.3 Data set description 26
3.4 Preprocessing phase 26
3.4.1 Image masking using U-net architecture 27
3.4.2 Edges enhancement 29
3.5 CNN-based regression model 30
3.6 Conclusion 31

4 Frameworks, tools and libraries 32
4.1 Introduction 32
4.2 Frameworks, tools and libraries 32
4.3 Conclusion 35

5 Results 36
5.1 Introduction 36
5.2 Obtained results and discussion 36
5.2.1 Results of test phase 37
5.3 Conclusion 38

6 Conclusion and Perspectives 39
6.1 Conclusion and Perspectives 39

References 40

Chapter 1

General Introduction

1.1 Introduction

The recent advancements in information technology, signal and image processing have drastically improved decision making and diagnosis in the medical field. New technologies such as Medical IoT, Cyber-physical systems, and Big Data have become a crucial part of healthcare. They gave birth to health 4.0 [8, 43]; a significant transformation from the traditional medical systems (paper-based records) to more intelligent—and data-intensive—electronic health records (EHRs).

The emergence of big medical data continues to fuel a great deal of research and development. Amongst the applications of data-driven techniques in the medical fields, we find diabetic retinopathy detection [33] and pneumonia detection [32], as well as Bone Age Assessment (BAA) [36]. The latter—the subject of this dissertation —is a method used by physicians to monitor the healthy growth of children. Traditionally, doctors would take x-ray images of the left hand of the subject. Then, they would use either the Greulich and Pyle [11] or the Tanner-Whitehouse [41] methods to assess the image and estimate the age of the bone. However, both of these techniques, while they vary in precision, rely on the judgment of the practitioner, and take a relatively long time to produce results.

To overcome these inconveniences, researchers developed novel and automatic algorithms and models to assist the doctors with their BAA estimation. For instance, BoneXpert [34] is a software that uses feature-extraction techniques to determine bone age. Other existing systems include industry veterans such as HANDX [26] and CASAS [40].

Recent advancements in machine learning and signal processing also permitted researchers to leverage new techniques and build new models that rely on deep learning such as works in [18, 35]. Nevertheless, while deep learning-based models were generally regarded as the best performing models [20], they were not widely adopted. These models require large amounts of data and are time-consuming to train. Furthermore, when compared to industry standards—such as Bonexpert—little to no difference can be found [20].

This work builds upon existing studies using deep learning for BAA. It is the first step in a study aiming to both enhance the results of the deep learning models, alleviate the aforementioned concerns, and eventually use these models to study the local population. The model presented in this work contains two stages: a segmentation stage and a CNN-based model. The first stage uses U-Net architecture for a segmentation process, whereas the second stage implements our proposed CNN model. To train the model, we use a public dataset provided by the Radiological Society of North America (RSNA) [28].

The proposed architecture allows us to leverage the power of U-Net architecture in both segmentation and edge enhancement. It also allows us to build a competitive and open model that can be trained in a timely manner and used by practitioners with no access to paid tools. Furthermore, using U-net and CNN will allow us to adapt the model to the features of the local population in future works, as well as compare them to the ones extracted from the global dataset, we currently use for training. Results of this work are sent, accepted and presented as an oral presentation in [1].

1.2 Organisation of the dissertation

The dissertation is organized as follows:

Chapter 2: Technical Background.

That presents the technical background of this work which is deep learning that is a part of machine learning in particular convolutional neural network and its popular architectures.

Chapter 3: Design and implementation of a deep learning architecture for Bone Age Assessment.

Explains the proposed approach for pre-processing phase and the proposed model which is based on CNN.

Chapter 4: Frameworks, tools, and libraries.

Shows all the frameworks and tools as well as libraries used in this work.

Chapter 5: Results.

Results of this work are presented and compared to related work.

Conclusion and Perspectives.

This concludes the main goal of this paper and it gives a point of view for future work.

Chapter 2

Technical Background

2.1 Introduction

Artificial Intelligence is one of the most important fields in Computer Science. Its goal is to make the machine capable of intelligent behavior, especially its effect in the healthcare domain where it gives a good result in a short time with less human power.

In this chapter, we will present the definition of healthcare and we will explain machine learning, deep learning, and we will describe the Bone Age Assessment method.

2.2 Health Care definition

In the medical dictionary, healthcare is "The prevention, treatment, and management of illness and the preservation of mental and physical well-being through the services offered by the medical and allied health professions" [25].

2.2.1 Health Care types

There are three types of healthcare depending on the type of the problem in health.

- **Primary healthcare:** Is essential healthcare made universally accessible to individuals and families in the community. Primary healthcare is important to touch on a person's health needs in their life, physically, mentally, and socially well-being focused on people-centered rather than disease-centered. Primary healthcare includes health promotion, disease prevention, treatment, rehabilitation, and palliative care [30].

- **Secondary healthcare:** This is when people need special healthcare. This means the problem is taken care of by an expert who has more specific details and information on your health problem. For example, the cardiologist is a specialist who focuses on the heart and its pumping system. Endocrinologists focus on hormone systems and some specialize in diseases like diabetes or thyroid disease [6].

- **Tertiary healthcare:** Specialized care that offers a service to those referred from secondary care for diagnosis or treatment, where the diagnostic or treatment facilities are scarce or require scarce combinations of resources, or which remain essentially the subject of research. These facilities are commonly found in medical schools and teaching hospitals [31].

2.2.2 Artificial intelligence and healthcare

Artificial intelligence (AI) touches many life areas. However, the largest impact of AI has been in the field of healthcare. AI simplifies the lives of patients, doctors, and hospital administrators, by performing tasks that are typically done by humans, in less time and at a fraction of the cost. Due to the availability of medical data in the form of medical history and medical images. The medical data makes the application of AI easier. Also, the development of algorithms such as deep learning which is capable to analyze the height dimension data.

Applications of Artificial Intelligence (AI) in healthcare:

Applications of artificial intelligent systems in the healthcare field can be classified into three categories [5] which are: patient-oriented AI, Clinician-oriented AI, and administrative and Operational-oriented AI.

There are many examples of artificial intelligence applications in healthcare such as:

- **AI-assisted robotic surgery:** A method of surgery with smart robotic tools that can work on complex surgical procedures with greater precision, control, and flexibility than other surgery methods. Figure 2.1 shows an example of an AI-assisted robotic.

Figure 2.1: AI-assisted robotic surgery example.

- **AI in medical diagnosis:** One of the important applications of AI in medical diagnosis is the Magnetic Resonance Imaging (MRI) scan. AI is actually revolutionizing the image diagnosis field in medicine. Figure 2.2 shows an example of Magnetic Resonance Imaging.

Figure 2.2: Example of Magnetic Resonance Imaging

2.3 Machine learning

Machine learning (ML) [47] is a branch of AI. It resolves problems such as forecasting, prediction, and classification, through data analysis and the production of algorithms. Thanks to this learning we do not need to code all the rules of the environment we are examining.

Types of Learning:

There are four types of learning [4], each of which requires particular input data:

- **Unsupervised Learning:** It can only feed the algorithm with data and it has to learn by itself the common patterns within the data. Examples of this type of algorithms: K-Nearest Neighbour(KNN), and K-means. Figure 2.3 shows an example of unsupervised learning in ML.

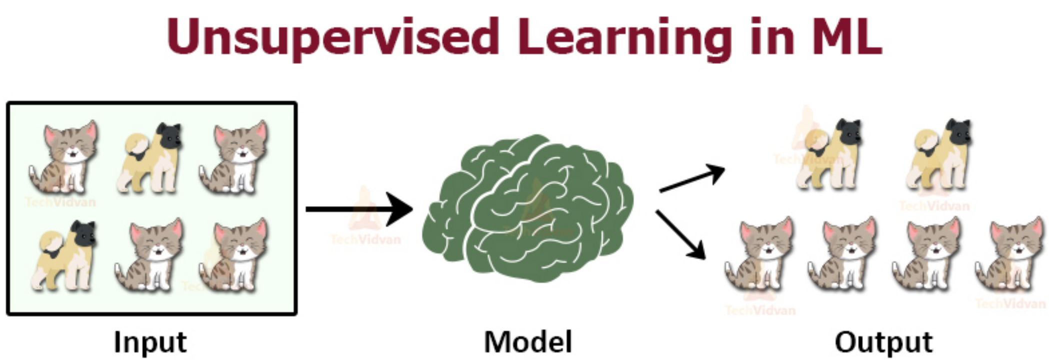

Figure 2.3: Unsupervised Learning in ML.

- **Semi-supervised Learning:** Semi-supervised Learning is able to learn using a few labeled data examples, and learn the rest of the data set without external control.

- **Reinforcement Learning:** Reinforcement Learning is a State of the art learning technique. Here we have an agent who examines an unknown environment, performing random actions to understand which action is good and which is bad by achieving rewards or penalties. The agent aims to optimize the long-term rewards. Figure 2.4 shows an example of reinforcement learning in ML.

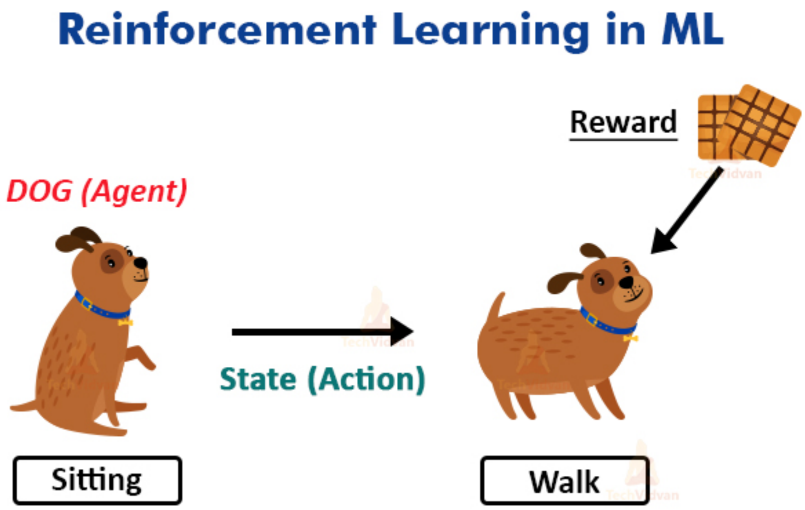

Figure 2.4: Reinforcement Learning in ML.

- **Supervised Learning:** Is the simplest way to learn something new for each entry of the dataset. Supervised Learning uses labeled datasets for a particular output, in order to train algorithms to classify data or predict outcomes accurately. This is the type of learning that we have applied in our work. Examples of this type of algorithms: SVM, Random decision forest. Figure 2.5 shows an example of supervised learning in ML.

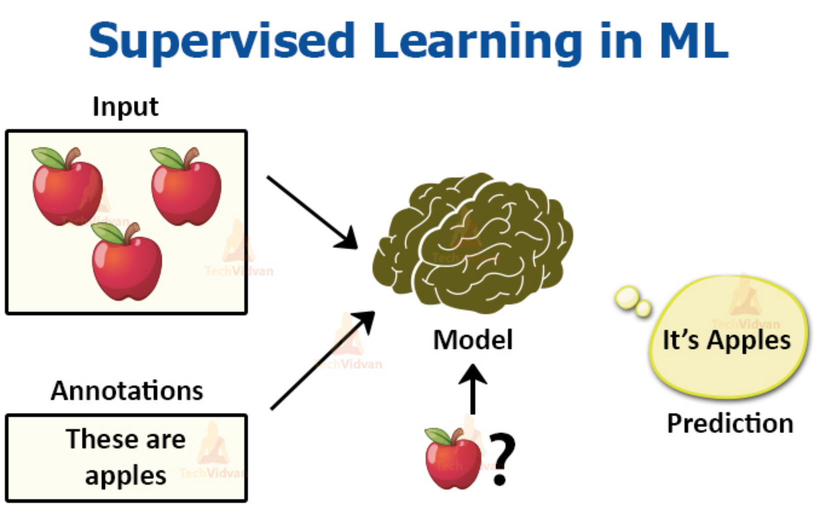

Figure 2.5: Supervised Learning in ML.

In supervised learning, there are two tasks that we are able to achieve: Regression and classification.

- **Regression:** Aims to predict a value as an output. The output could be of any type: string, numeric, date, ...etc.
- **Classification:** Aims to categorize input data into two or more classes. There is no prediction to be done, each input data will be classified to the class it belongs to.

Figure 2.6 shows the difference between the classification and regression task.

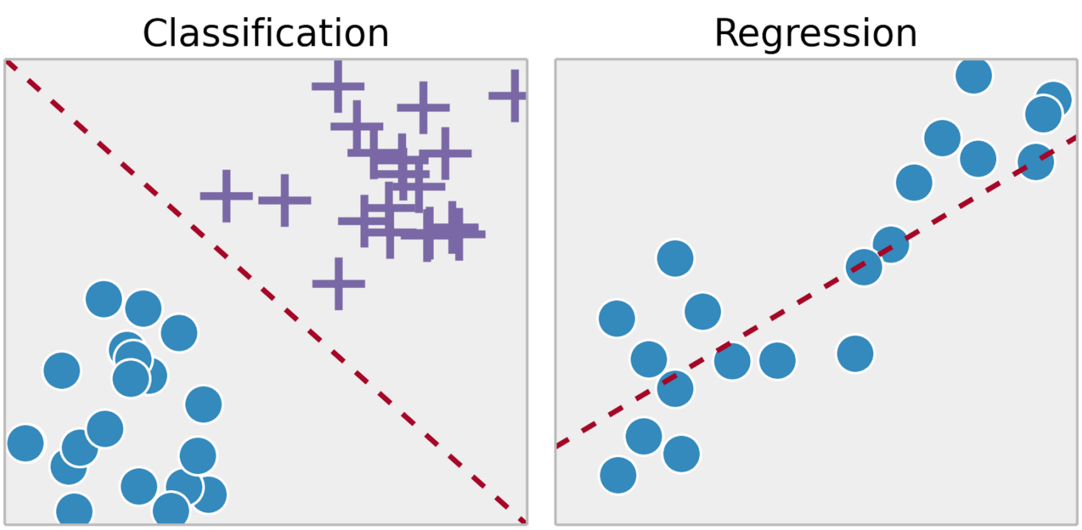

Figure 2.6: Classification and Regression [29].

2.4 Deep Learning

Deep Learning (DL) is a subfield of Machine Learning based on Neural Networks for the resolution of tasks. DL develops day by day thanks to the biggest companies in the technology sector. These companies are constantly looking for new architectures and implementations for the real world. Deep Learning is the most powerful tool that we have for dealing with Artificial Intelligence problems. Moreover, DL almost always offers better performance than its Machine Learning counterpart.

2.4.1 The differences between the Machine Learning and the Deep Learning

The differences between the Machine Learning and the Deep Learning approach are based on the algorithms [49]. Machine Learning proposed a variety of algorithms that differ from each another. Such as SVM, Linear Regression, Decisional Trees, Random Forests...etc. All of these approaches do not require as much data as required by Deep Learning.

Deep Learning offers only one algorithm, the Neural Network (NN). There are many types of NNs but the main concept is the same. Figure 2.7 shows the difference between the ML and the DL.

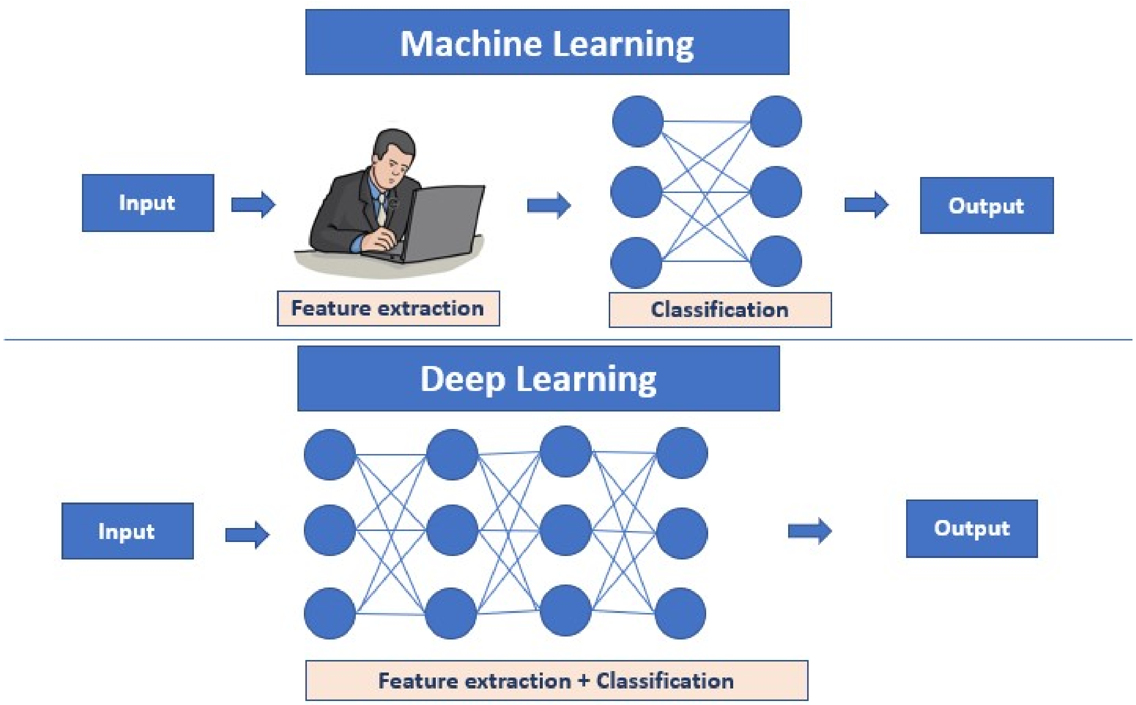

Figure 2.7: The difference between the ML and the DL.

2.4.2 Artificial neural networks

To understand artificial neural networks it is better to mention some basic notions of biological neural networks. Therefore, the variations are created within the artificial neural network model.

Biological neuron:

The neuron is a type of cell that makes up the neural tissue. The neuron is able to receive, process, and transmit signals to other neurons. Thanks to this behavior, the nervous system is created. The nervous system is the interface that allows us to interact with the environment through actions, sensations, and thoughts [21].

The power of neurons is based on the property of non-linearity. Computation made by groups of neurons or entire parts of the brain, can not produce the extraordinary results that are well known if the neuron only performed a linear combination of inputs.

Through the dendrites, the neuron receives signals coming from neighbors. After that, inside the nucleus, all the signals are combined linearly. If the results of the linear combination exceeded the activation threshold, the neuron will forward a signal to the subsequent neurons. The process of neuron activation is called firing the neuron. The Myelin sheath will modulate the intensity of the produced signal, empowering it or reducing its voltage. This behavior will affect the importance of the signal once it is received by other neurons.

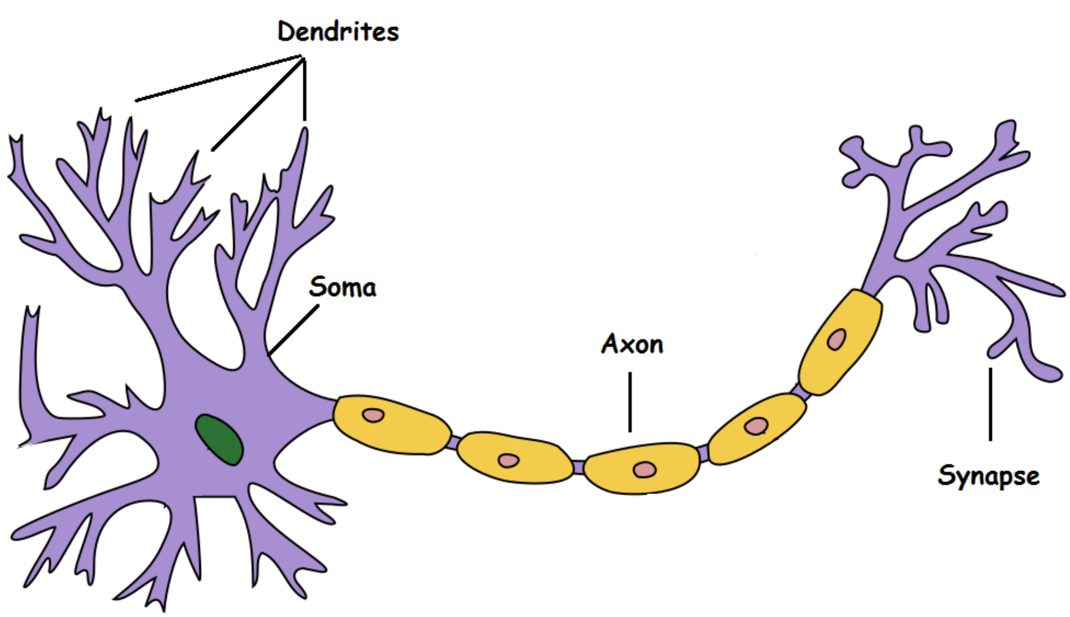

Figure 2.8: Biological Neuron.

Artificial Neuron:

Artificial neuron as Figure 2.9 shows, has the same role as the biological neuron, each artificial neuron is an elementary processor. It takes as input a number of variables in the input layer, and each input of the artificial neuron has a weight representing the value of the connection. An activation function transforms the sum of the input variables and their weights. Then, there is the output layer which receives the sum value to compare it with a threshold value. After that, the output layer will make an output response.

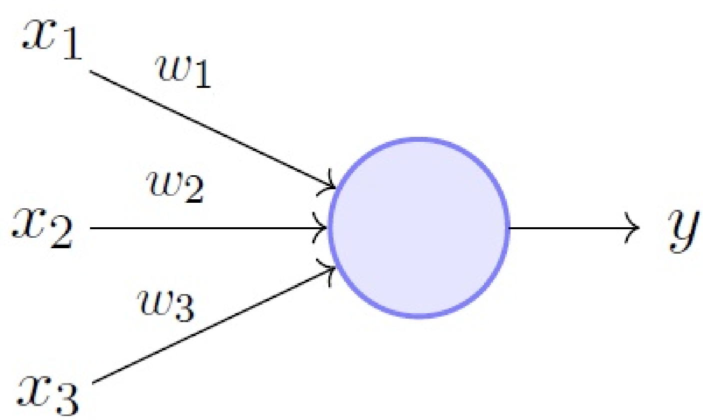

Figure 2.9: The artificial neuron structure.

Artificial Neural Network:

The artificial neural network consists of three types of layers of neurons. These layers are: the input layer, a hidden layer that helps to identify complex details within the input, and an output layer that contains a number of output neurons depending on the task that will be solved. If we should predict the income value then one output neuron. If we should classify, the output layer neuron size will be equal to a number of classes.

Activation function: It makes a decision whether a neuron should be activated or not [37]. It will decide if the input of the neuron is important or not in the process of prediction, this action uses simpler mathematical operations.

There are several forms of activation functions, each one is used in a specific context, we will mention the most used ones.

- **Rectified Linear Unit Function:** ReLU is the most commonly used activation function in neural networks to convert all input values to positive numbers.

  Defined by the following equation:

  $Relu(z) = max(0, z)$

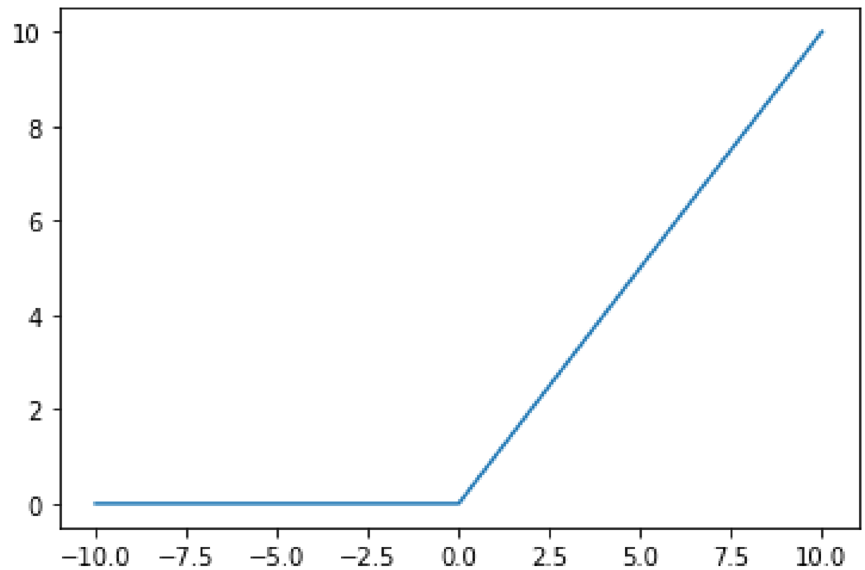

Figure 2.10: RELU Function curve.

- **Exponential Linear Unit Function:** ELU is a function that heads to converge cost to zero faster and gives more accurate results.

  Defined by the following equation:

$$ELU = \begin{cases} x & \text{if } x >= 0 \\ \alpha(\exp(x) - 1) & \text{if } x < 0 \end{cases} \tag{2.1}$$

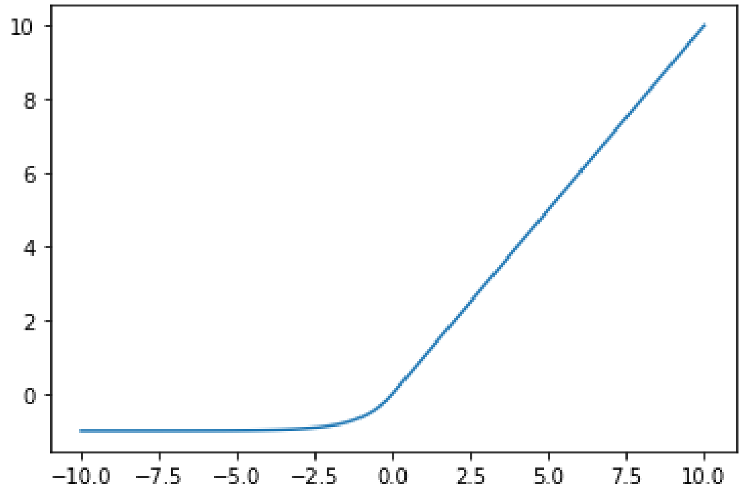

Figure 2.11: ELU Function curve.

- **Sigmoid:** It processes the input into a value between 0.0 and 1.0.

  Defined by the following equation:

  $\sigma(z) = \frac{1}{1+e^{-z}}$

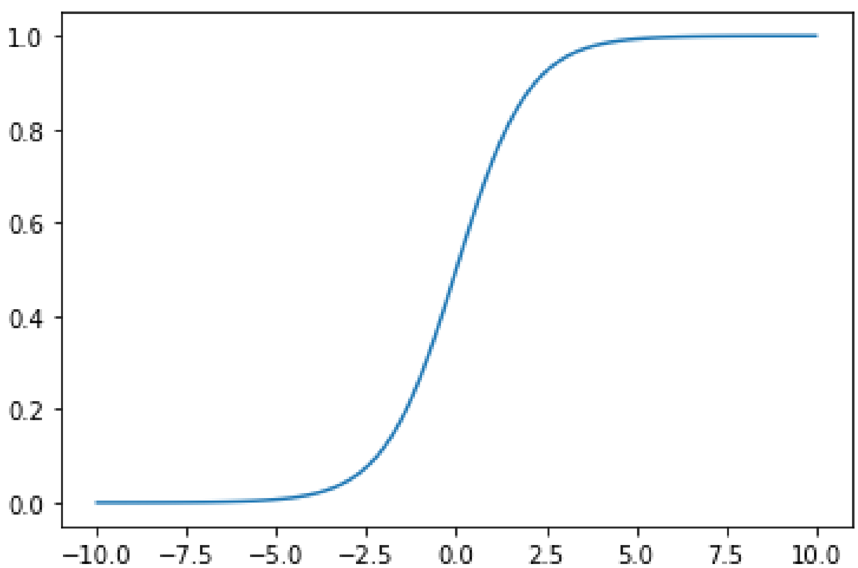

Figure 2.12: Sigmoid Function curve.

- **Softmax:** It changes a vector of numbers into a vector of probabilities, where the probabilities of each value are commensurate to the relative. It is used in machine learning as an activation function in the neural network which is configured to output N values, one for each class in the classification task. It is used to normalize and transfer the inputs from weighted sum values into probabilities their sum equal to one. Each value in the output of the Softmax function is translated as the probability of membership for each class.

Defined by the following equation:

$\sigma(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \quad for\ i = 1, 2, \ldots, K$

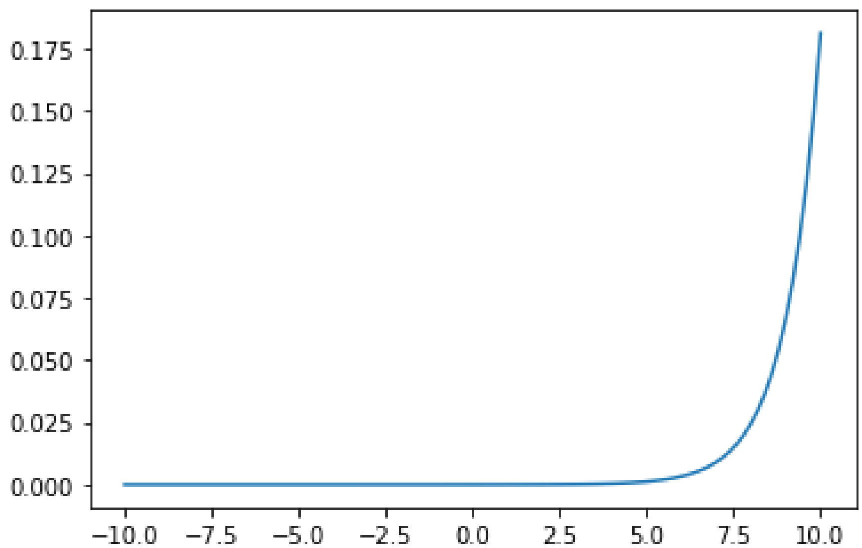

Figure 2.13: Softmax Function curve.

2.5 Convolution Neural Network

Deep Learning (DL) has improved many fields of science, and it makes possible to access great power through in-depth analysis of incoming data. One of the biggest improvements was received by Computer Vision. As the name suggests, Computer Vision performs operations on visual input to recreate the real environment. The aim pursued is the emulation of human sight. Many technologies have been implemented over the years trying to create an efficient analysis process. One of these is obviously the Dense Neural Networks, which have been shown to be unsuitable for Computer Vision tasks, due to their architecture.

2.5.1 CNN definition

Dense Neural Networks (DNNs) have a huge number of connections that require very high computing power. Therefore, DNNs are able to analyze big images. The idea of existing computer vision comes when Convolutional Neural Network was born. Many studies have been conducted on the visual cortex of humans (also studying the cortex of cats, which is similar to ours). The visual signal has been found to bounce back and forth between several parts of the brain. Each visual signal performs a different type of processing. CNN does exactly the same [2]. The signal is processed by many layers, and at each step, the output contains the most important features (characteristics) of the input.

2.5.2 CNN Architecture

Typically, the architecture of CNNs is simple [2]:

- Input Layer: receives the images (1 to 3 dimensions).
- Convolutional Layers: During the training, steps divide the input image into virtual regions. Each region will be analyzed by a filter (kernel). The filter iterates over these regions and calculates which weights should be used to represent the most important information. But how are these weights calculated?, the answer to this question represents the true power of CNN. The network tries to minimize the loss between input and predictions, and it will automatically calculate the right weights.
- Flatten Layer: receives the output matrix of the last convolution and flattens it to create a one dimension vector.
- Dense Network: receives the one dimension vector and performs all the computations.

How Convolution Works:

The convolution operation is the main function of CNN. It allows calculating the weights that must be used to recognize a particular feature within the image.

Convolution involves the initial input (original image) and the current filter. Each filter of the feature map (the group of filters for the current convolutional layer) will iterate over the image. Convolution is the group of sum operations between the multiplication of each element of the image matrix by the corresponding element of the filter matrix. Figure 2.14 shows how the filter works in the convolution layer.

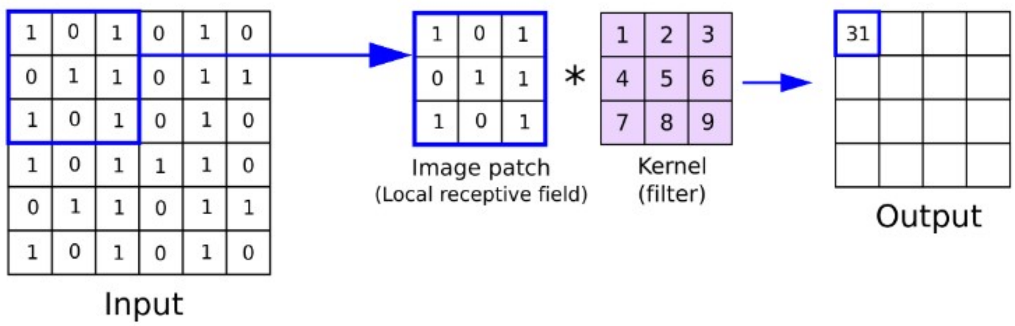

Figure 2.14: Calculation process of a filter in the convolutional layer.

Filter:

Is the main actor of convolution. Its content will determine what features will be identified in the image. If we have values only in the column shape within the filter, the vertical lines of the input will be highlighted. Conversely, row values will emphasize horizontal lines. We need to use many convolutional layers because the complexity of shapes recognized by the network will increase along with the depth of the network.

Padding:

The convolution output matrix has a smaller shape than the original. This behavior can lead to loss of information, which could drive to the poor final performance. To work around this problem, the padding operation is applied. The Padding consists of a frame of zeros surrounding the original image. Figure 2.15 shows how to add zeros padding to an image.

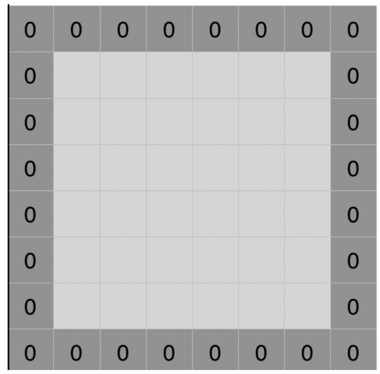

Figure 2.15: Example of zero padding added to an image.

Stride:

The default choice is that filter will only move one position at a time. But this is not a constraint. If it is necessary, you can move the filter by many positions as you like by changing the stride hyper-parameter. Figure 2.16 shows the stride controls how the filter convolves over the input image.

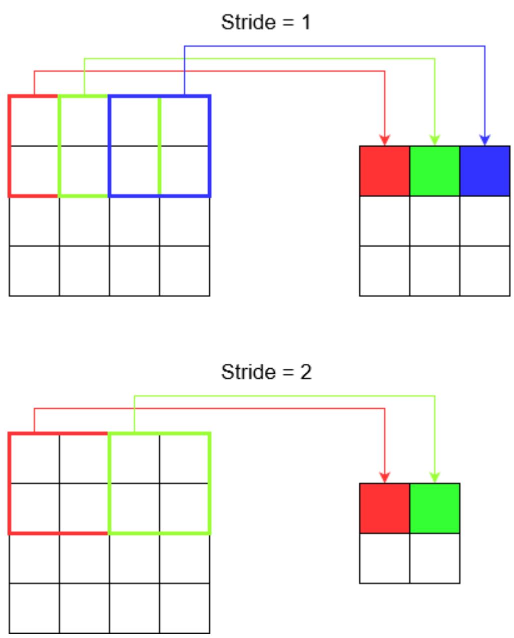

Figure 2.16: Stride in CNN.

The Pooling Layers:

The goal of CNN is to manage images and help achieve the goal of finding the most important features within the image, and keep the weight computational cost low. The Pooling layer helps reduce the amount of information passed to subsequent layers by keeping the most important information. The Pooling Layer iterates over the images doing some operations to extract relevant features, and at each step, it reduces the dimensions of the incoming image. Figure 2.17 shows how max-pooling layer selects maximum element from the feature map.

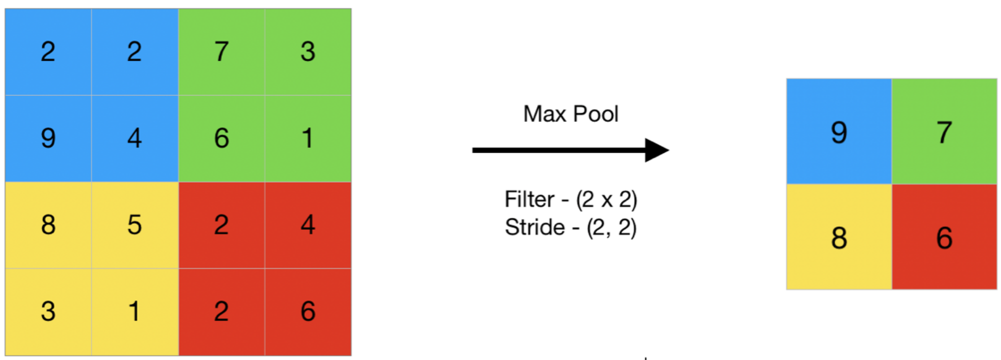

Figure 2.17: A practical example of max-pooling layer.

There are three types of pooling layers as follows:

- **Max:** It considers the most important information as the maximum value inside the iteration window.
- **Average:** The information between pixels is mitigated by the average operation, in order to reduce information but still consider all the values within the iteration window.
- **Global:** It calculates the average value within each feature map and also helps to reduce the connections with the last layers. The power of this layer relies on weight reduction.

2.5.3 CNN layers configuration

The CNN process flow runs as follows, the input image is analyzed by the first convolutional layer, which detects the basic features within the image. Usually, larger filters are used at this point to capture general information. Subsequent layers will detect more complex information, and the amount of information captured will increase. Instead, the size will contract more and more at each level. Hence, it will create a dense representation of the input. As we can see in Figure 2.18, the last portion of the CNN architecture consists of a Dense Neural Network (DNN), to perform classification. At this point, the number of information provided to the DNN is a crucial parameter. Due to the dense connections, the number of weights can skyrocket. Therefore, CNN aims to decrease the width and height while increasing the depth of the feature maps.

The GlobalAveragePooling layer performs better work in some cases because we can skip the entire DNN portion, using just one Dense Layer as the output.

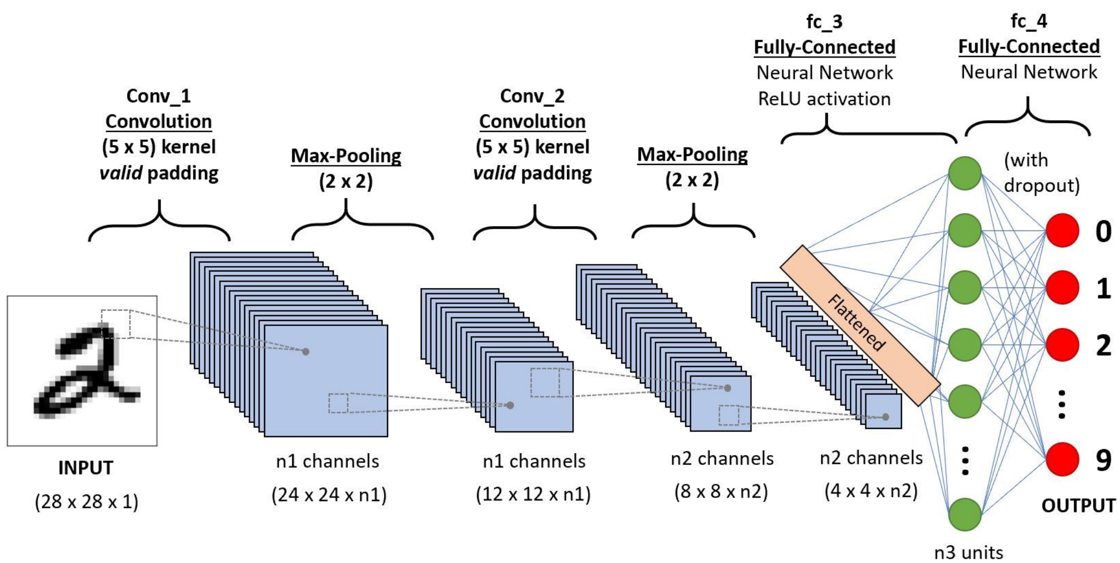

Figure 2.18: Convolutional Neural Network Architecture [12].

2.5.4 Some CNN architectures

Through the years, several powerful CNN architectures had been released. LENET [22] is one of the first proposed CNN for handwritten digit identification, ALEXNET [15] is a large-scale CNN model used in computer vision, and the architecture Network in Network [24] is suitable for simple data sets. Besides these basic CNN, other powerful CNN were developed: (i) VGG Net [38] which is the most popular CNN architecture due to its small-sized convolutional kernels and its simplicity, (ii) GoogleNet [39] which is the first popular complex architecture, and (iii) Resnet [13] which contains an important advantage of the residual network architecture which is the identity skip connections in the residual blocks. Figure 2.19 shows the structure of skip-connection in Resnet. Resnet can train very deep CNN architectures easily, and when researchers combine its power with GoogleNet it gave a new architecture called ResNeXt. By using skip connection, researchers had found a new technique to empower the training phase, reducing the amount of total computation. Skip connection is the core idea of the DenseNet [16] architecture that allows to complement complex connections between the layers.

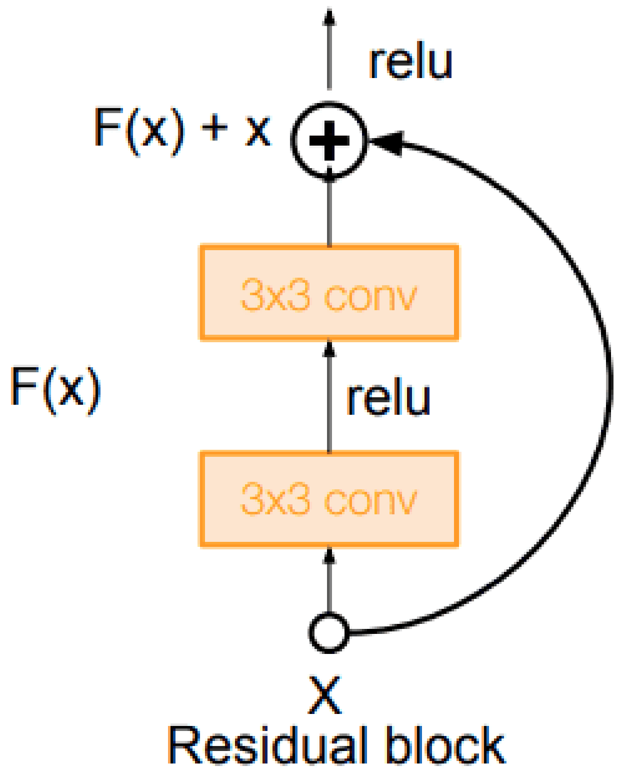

Figure 2.19: Skip-connection.

2.6 Bone age assessment

Bone age assessment is commonly performed in pediatric endocrinology and in orthodontics and pediatric orthopedics. Bone age assessment is a powerful sign to determine the timing of diagnosing various diseases and the treatment. The main goal of this method is to evaluate growth and diagnose pediatric problems. Therefore, the accuracy of bone age assessment could be very important. Although, manual bone age assessment methods have been used for an extended time, the main trouble with these techniques is inter and intra-observer variability. Recently, several computerized systems for bone age assessment have been developed such as BoneXpert which is [34] is software that provides automated age analysis from standard radiographs of the hand. Figure 2.20 shows the correlation between the highlighted places on the hand, and the bone age assessment.

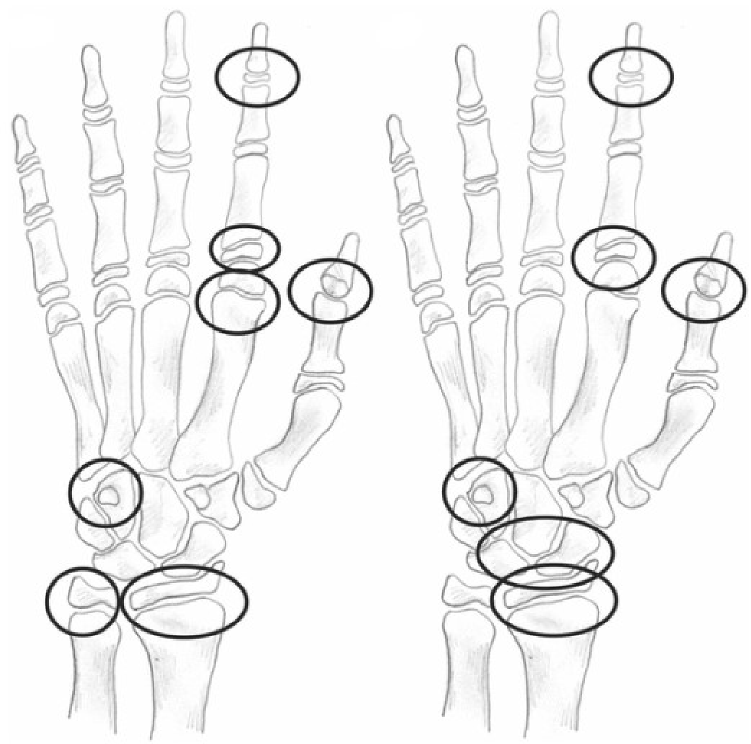

Figure 2.20: The encircled windows correlate with bone age [27].

2.6.1 Bone Age Assessment Methods

The most frequently used bone age assessment methods are the Greulich-Pyle (GP) [17] and Tanner-Whitehouse (TW2) [42] methods. These methods need the image of left-hand radiographs, because it contains many bones. There are several reasons for using left-hand radiographs for bone age assessment. The common one is that most people are right-handed, which means the right hand is more tolerable being wounded than the left hand. Also, it was determined that physical measurements have to be measured on the left side better than on the right side of the body in the early 1900s at the physical anthropologist's conferences.

- The GP method is an atlas technique where bone age is assessed by comparing the radiograph of the patient with the standard radiograph in the atlas. The GP method was improved using radiographs of upper-middle-class Caucasian children in Cleveland, Ohio, United States. It was recently reported that secondary sex characteristics in current boys and girls, start earlier than they did several decades ago in the United States [36, 14]. Therefore, probably it is complex to estimate bone age strictly in current children using the GP method.

- Tanner-Whitehouse (TW) method published in several countries, and this method has changed the link between the total bone growth score and bone age, to make the relation favorable for each generation and ethnic group. The standard deviation of bone age was approximately 1 year from the age of 5 years in both sexes to 14 yo in girls and 16 yo in boys, it was calculated using the radius-ulna-short bones (RUS) method to evaluate the thirteen long or short bones (the radius, ulna, and short bones of the first, third, and fifth fingers)[42].

2.7 Related work

There are many deep learning-based architectures that have been built to automatically assess bone age. Authors in [18] built a neural network model combining U-Net [35] to segment the dataset, and VGG convolutional neural nets (CNN) [38] to build the prediction models.

A second model presented by authors in [7], used Non-subsampled Contourlet Transform [9] on the X-Ray images.

Then, they too relied on the VGG architecture to build a regression model. Other models use heterogeneous information such as race and age [44], or heterogeneous architectures such as CNN and Support Vector Regression (SVR) [3], or even transfer learning [23, 10].

Lately, new works even used the new attention architecture in order to study its effects on the results [46].

2.8 Conclusion

In this chapter, we have explained some knowledge which are important as background in our work. We have described the different use cases of AI in healthcare. After that, we have detailed the basics of machine learning and deep learning. Then, we have explained the bone age assessment method and its different methods. Finally, we have mentioned some previous related works.

Chapter 3

Design and implementation of a deep learning architecture for Bone Age Assessment

3.1 Introduction

After presenting the CNN bases in the previous chapter, our goal now is to build a CNN-based architecture for the bone age assessment method. In this chapter, we discuss our proposed CNN architecture and the required preprocessing steps before feeding the proposed model to achieve our prediction goal.

3.2 The proposed approach

The approach we described hereafter is composed of two steps: preprocessing and prediction. A key aspect of our approach is the segmentation of images. Its goal is to keep the Region of Interest (RoI) focused on the most important information for each image, avoiding feeding the model with noisy or uninformative data. Once this phase is completed, the resulting dataset was divided into three parts: 70% for training, 20% for validation, and 10% for test.

In the following, we describe firstly the preprocessing pipeline then we detail the CNN-based architecture fed by the results of the preprocessing phase.

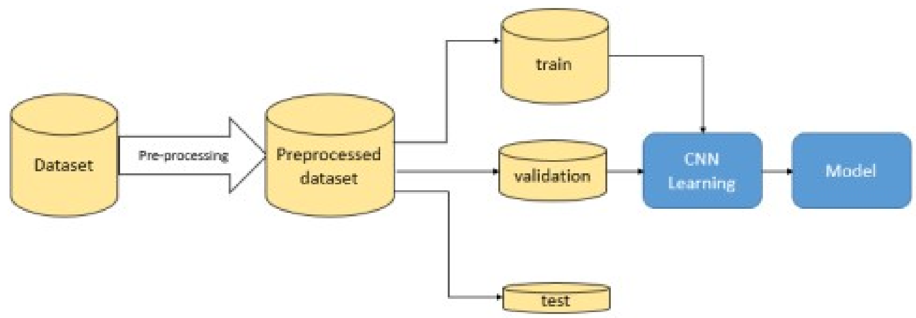

Figure 3.1: System design.

3.3 Data set description

The model is trained and tested using data provided by the Radiological Society of North America (RSNA) [28]. This publicly available dataset contains 12600 X-ray images of the left hand of young subjects. The images are all labeled with the bone age and sex of the subject. The bone age is indicated in months ranging from 1 month (the youngest) to 228 months (the oldest).

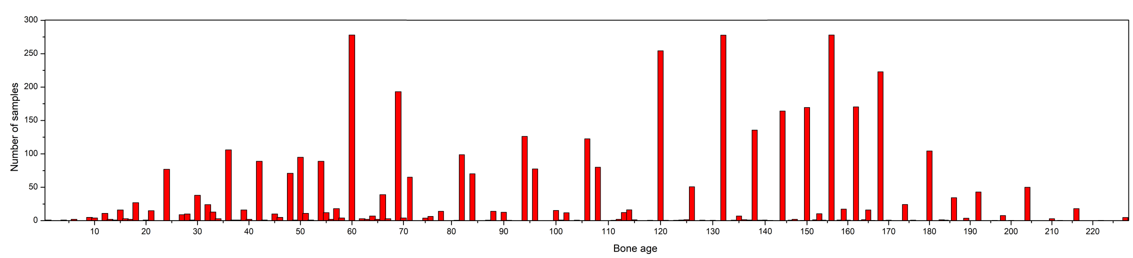

Figure 3.2: Estimated bone age distribution as assessed by a trained physician. Bone ages are not evenly divided.

3.4 Preprocessing phase

Raw data can not be fed directly to the neural network due to some crucial problems, such as: the shape of images being not fixed, images contain extra data that are uninformative with respect to BAA, and finally most images have to be preprocessed to enhance the RoI. Fig. 3.5 shows a step-by-step evolution of the X-ray images during the preprocessing pipeline. These steps are detailed in the following paragraphs.

3.4.1 Image masking using U-net architecture

To focus the RoI on the bones of each hand, we proceeded to remove the background from the images. To do so, firstly, we used the Topaz AI tool to manually label 100 masked images [45]. Admittedly, this process was arduous and time-consuming. Nevertheless, the results from the U-net layer with these images were encouraging, and labeling 100 images is feasible compared to the total number of images.

U-net architecture:

Our choice of using the U-net architecture was motivated by their results compared to traditional texture algorithms [19]. It is composed of a down-sampling and an up-sampling components [35]. The down-sampling part is based on the repetition of two 3×3 convolution layers, with a Rectified Linear Unit (ReLU) activation function after each of them. The convolution layers are followed by a max-pooling operation with a 2×2 stride.

The cropped output of the down-sampling block is then combined with the input data to feed the up-sampling block. This block is of a similar composition to the down-sampling one. Finally, at last, a 1×1 convolution is used to create the final mask by converting each of the 64-component feature vectors to a fixed dimension output.

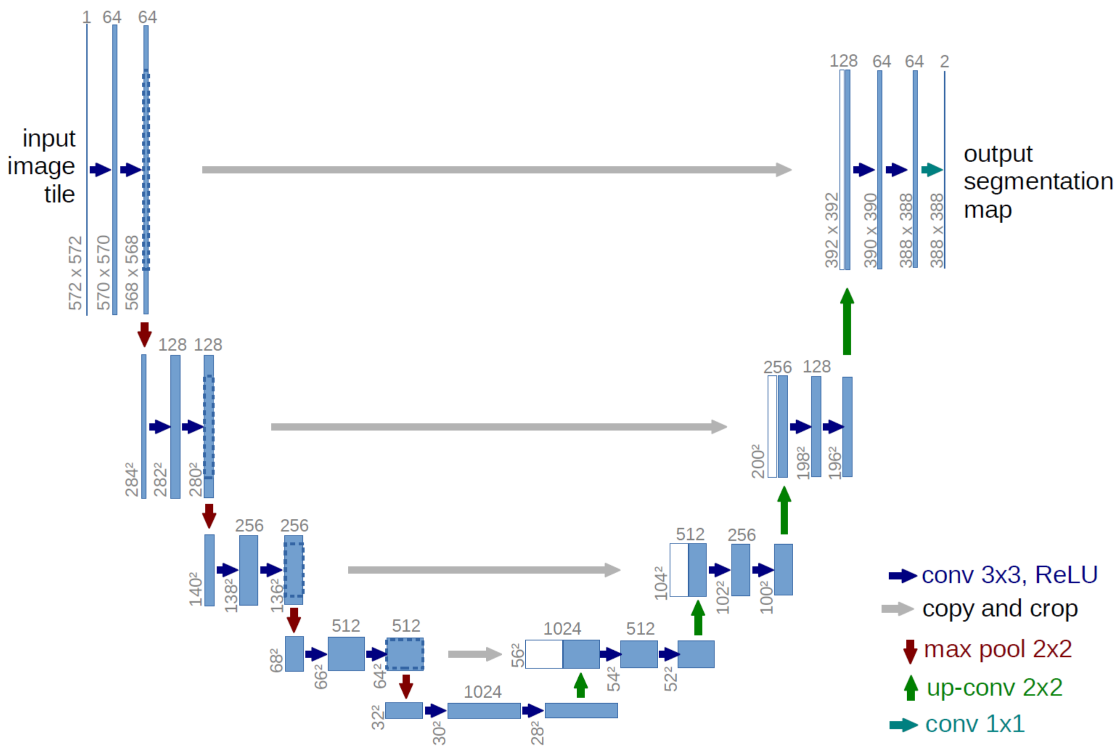

Figure 3.3: U-net architecture [35].

Once the model is trained on the manually labeled masks, it then applied to the rest of the data to create new masks. From these masks only high-quality ones were kept to retrain the U-Net model in order to achieve better performance. This process has been repeated five times. Figure Fig. 3.4 shows sample of poor quality images that have been deleted.

Figure 3.4: Poor quality images examples.

When we obtained binary masks from all of the data, we cropped the region of interest from raw data by using a simple python program.

3.4.2 Edges enhancement

To further enhance the RoI, we used an omnidirectional kernel M (Eq. 3.1), to highlight the bone within each image.

$$M = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{bmatrix} \tag{3.1}$$

Finally, in order to have a consistent input, the images are rescaled and rotated before being fed to the CNN-based model.

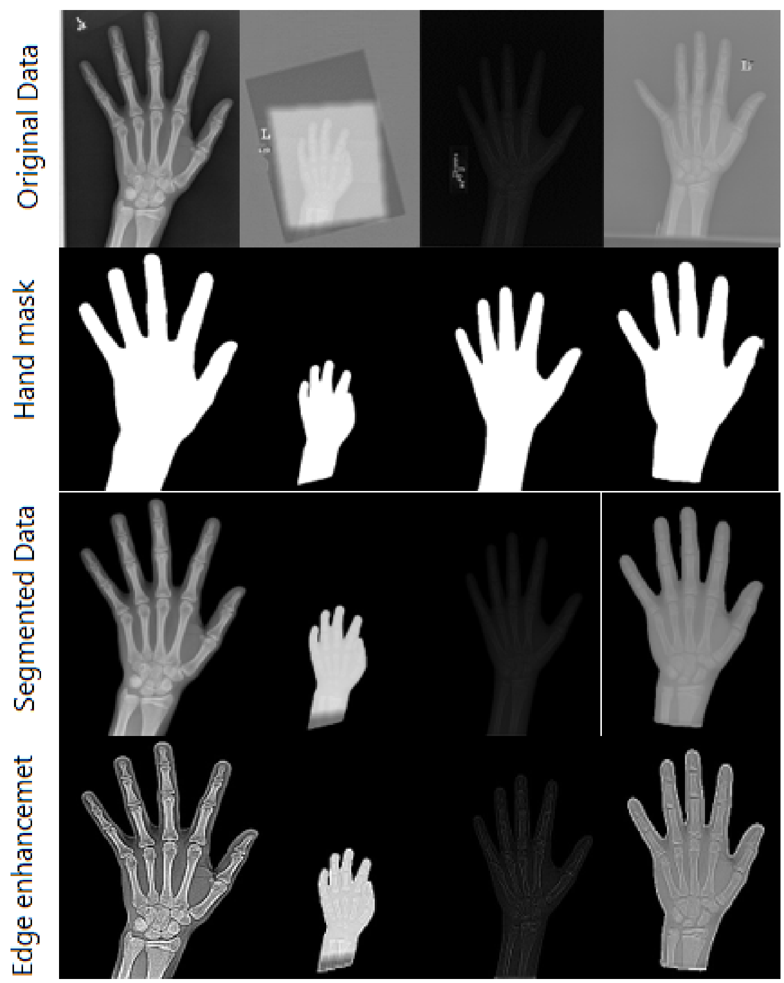

Figure 3.5: Preprocessing pipeline.

3.5 CNN-based regression model

To estimate bone age, we implement a VGG-style regression model. This model contains 6 VGG blocks consisting of 32, 64, 128, 128, 256, and 384 convolution layers, respectively. Each VGG block is formed by two convolution layers: the first one with a 3×3 size kernel and a second one with 1×1 size kernel, with the Exponential Linear Unit (ELU) as an activation function. Additionally, in each block, the two convolution layers are followed by a batch normalisation layer, and a max-pooling layer (Fig. 3.6 (a)).

Following the six VGG blocks, we add a 64 convolution layer (with an ELU activation function), a batch normalization layer, and a max-pooling layer. Then, to avoid over-fitting, we use dropout regularisation in combination with a global average pooling layer. Finally, we add a dense layer that returns the output of the regression (Fig. 3.6 (b)).

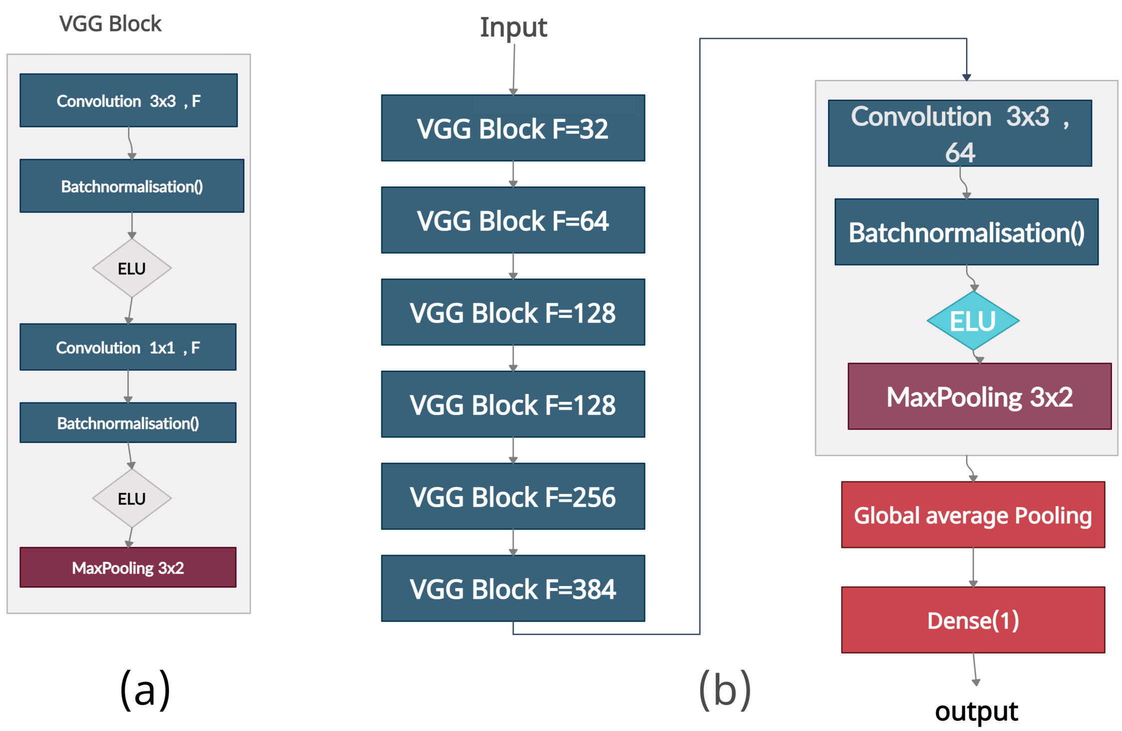

Figure 3.6: The architecture of the proposed model. (a) Composition of the VGG blocks (b) The proposed VGG-style neural network architecture for regression.

Before we move on to train the model, some parameters and functions need to be explained due to their importance:

- Optimizers: Is a function or an algorithm that modifies the attributes of the neural network, such as weights and learning rate. In our work, we choose Adam as an optimizer.

- Batch size: The total number of training or validation samples in a single batch, our batch size is set to 32 in our experiment.

- Step per epoch: Is calculated by dividing the number of samples in the training subset by the batch size.

- Validation Steps: It is calculated by dividing the number of samples in the validation subset by the batch size.

- Learning rate: Determines how fast the model fits. We used the default value.

- ModelCheckpoint: Is used in conjunction with training using model.fit() to save a model.

- EarlyStopping: The main objective of this method is to avoid overfitting. Once the model performance stops improving, it stops the training phase.

- model.fit: This function will train our model.

3.6 Conclusion

In this chapter, we have described the dataset which we used. Then, we have shown the preprocessing pipeline that we have followed. After that, we have detailed also our CNN-based model. In the next chapter, we will mention the used tools, libraries, and frameworks.

Chapter 4

Frameworks, tools and libraries

4.1 Introduction

After presenting our CNN-based architecture in the previous chapter, the objective of this chapter is to mention tools and frameworks that help us to implement our code.

4.2 Frameworks, tools and libraries

- Python: Python is a versatile, general-purpose, interpreted, and high-level computing programming language, developed by Guido van Rossum and originally published in 1991.

Figure 4.1: Python

- Open Source Computer Vision Library (OpenCv): is a library for computer vision and machine learning programs, it gives an important infrastructure for computer vision applications.

Figure 4.2: OpenCv

- TensorFlow: is an open-source framework developed by Google to perform machine learning, deep learning, and other statistical and predictive analytics workloads.

Figure 4.3: TensorFlow

- Keras: is an interface for the TensorFlow library, is an open-source software library suitable for artificial neural networks.

Figure 4.4: Keras

- NumPy: its creator is Travis Oliphant. It is a Python fundamental scientific computing library used for working with arrays. It also has functions used in the linear algebra domain, fourier transform, and matrices.

Figure 4.5: Numpy

- Matplotlib: is a python library available as a component of NumPy to make plots, it is a resource for a big data numerical treatment. Matplotlib has an object-oriented API to establish plots in Python applications.

Figure 4.6: Mathplotlib

- Kaggle: is a website that allows users to find and publish data set and share ideas, learn new information, and find a lot of coding tricks, as well as observe different examples of real-world data science applications.

Figure 4.7: kaggle

- Spyder: is a cross-platform which is an open-source integrated development environment (IDE) for scientific programming in the Python language.

Figure 4.8: Spyder

- Colab: was developed by Google to provide free access to GPU's and TPU's appropriate for building a machine learning or deep learning model. In our work, we used Google Colab pro.

Figure 4.9: Colab

4.3 Conclusion

In this chapter, we have mentioned all the libraries and frameworks, and tools that we need in this domain. In the next chapter, we will discuss different obtained results.

Chapter 5

Results

5.1 Introduction

In the previous chapter, we have shown all tools and libraries, and frameworks that are important to work. In this chapter, we will discuss our experiment and the obtained results. Finally, we will compare our obtained results with some of the previous related works.

5.2 Obtained results and discussion

The model presented in this work went through several iterations hese iterations allowed us to compare the approaches used in the literature and enhance our modeling process. The versions of these iterations are presented in Table 5.1.

Table 5.1: Results of the evolution of the proposed model through diferent versions.

Version	Architecture	MAE
I	fine-tuned DenseNet201 (with edge enhancement)	13.37
II	fine-tuned VGG16 (with edge enhancement)	14.37
III	Our work CNN-based (no edge enhancement)	10.97
Final	Final version CNN-based (with edge enhancement)	**9.46**

Firstly, we have used transfer learning with two architectures: a fine-tuned VGG16 [38] and a fine-tuned DenseNet [16], versions I and II, respectively, in Table 5.1). As transfer learning didn't yield satisfactory results, we developed the model described in the previous section. In the first version, we used the data without edge enhancement (Version III Table 5.1). Then, in the final version we enhanced our results by implementing edge enhancement.

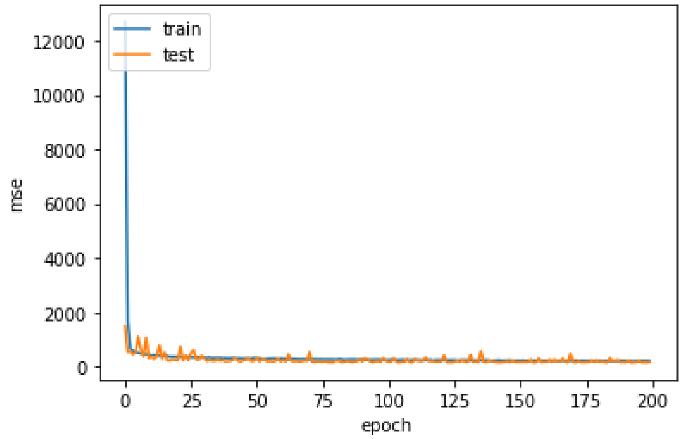

Figure 5.1: Evolution of the loss function described by the Mean Squared Error (MSE) over the training of the model.

Table 5.1 shows how the Mean Absolute Error (MAE) continues to decrease on the test with each subsequent version, showcasing the current model with edges enhancement as the best one.

5.2.1 Results of test phase

Table 5.2 details the results of the model on the 1262 images from the test set. These results are highly encouraging as the prediction fits very well to the data with relatively low errors. It also presents no overfitting, as shown in Fig. 5.1. Fig. 5.2 shows the clustering of the predictions around the corresponding values of bone age.

Table 5.2: Results of the model on the test set.

Metric	**Result**
Mean absolute Error (MAE)	**9.46**
Mean squared error (MSE)	**149.46**
Root mean square error (RMSE)	**12.22**
Max of the error	**51.29**
The mean value of the error (μ_e)	**-0.75**
The mean value of the relative error (μ_ε)	**0.09**

Comparing our work to those established in the literature (Table 5.3), we find that the results obtained by our approach are very promising and surpassed most of the established works in terms of MAE. Our results are marginally behind those obtained by [18] but remain competitive.

Figure 5.2: Clustering of the predictions around every value of the BAA.

Table 5.3: Comparison with work from the literature using the RSNA bone age data.

Work	Method	Results (MAE)
Our work	**CNN-based**	**9.46**
[48]	CNN plus machine learning: *VGG*	14.78
[48]	CNN plus machine learning: *Mobile net*	17.09
[18]	CNN-based	8.08
[46]	Transfer learning on transformers	9.99

We believe tuning of the model as well as testing other preprocessing methods would further enhance the performance of our approach.

Mean Absolute Error(MAE)= $\sum_{i=1}^{D} |x_i - y_i|$

5.3 Conclusion

In this chapter, we have explained all experimentations, and we have given results of different metrics also through the plot of clustering of the predictions around every value of the BAA for the test phase. The results obtained represent particularly good performances that encourage the improvement of our model architecture.

Chapter 6

Conclusion and Perspectives

6.1 Conclusion and Perspectives

The Deep Learning approaches have proved to be suitable in several applications. We have developed a Bone Age Assessment model on CNN and trained using the public RSNA Bone Age data set [28]. The first step was to preprocess the X-ray images using a U-net neural network. The preprocessing greatly boosted the performance of our approach as indicated by the test results.

In the second step, we implemented and tuned a VGG-based regression model to predict bone age. The model gave highly encouraging results competitive with those present in the literature.

In Future works, our objectives are twofold. On the modeling aspect, we aim to explore different preprocessing and augmentation techniques. In the use of the model, we aim to adapt to the local population to better predict bone growth in local subjects and study discrepancies with the global data set.

References

[1] Zerari Achouak, Djedidi Oussama, Kahloul Laid, Carlo Romeo, and Remadna Ikram. Paediatric bone age assessment from hand x-ray using deep learning approach. In *the 5th Conference on Computing Systems and Applications, CSA'2022', Algiers (Accepted and presented)*, pages 1–10. Springer, 2022.

[2] Saad Albawi, Tareq Abed Mohammed, and Saad Al-Zawi. Understanding of a convolutional neural network. In *2017 International Conference on Engineering and Technology (ICET)*, pages 1–6. Ieee, 2017.

[3] Mariette Awad and Rahul Khanna. Support vector regression. In *Efficient learning machines*, pages 67–80. Springer, 2015.

[4] Taiwo Oladipupo Ayodele. Types of machine learning algorithms. *New advances in machine learning*, 3:19–48, 2010.

[5] Kanadpriya Basu, Ritwik Sinha, Aihui Ong, and Treena Basu. Artificial intelligence: How is it changing medical sciences and its future? *Indian Journal of Dermatology*, 65(5):365, 2020.

[6] Fiona M Burns, Anne M Johnson, James Nazroo, Jonathan Ainsworth, Jane Anderson, Ade Fakoya, Ibidun Fakoya, Andy Hughes, Eva Jungmann, S Tariq Sadiq, et al. Missed opportunities for earlier hiv diagnosis within primary and secondary healthcare settings in the uk. *Aids*, 22(1):115–122, 2008.

[7] Xun Chen, Chao Zhang, and Yu Liu. Bone age assessment with x-ray images based on contourlet motivated deep convolutional networks. In *2018 IEEE 20th international workshop on multimedia signal processing (MMSP)*, pages 1–6. IEEE, 2018.

[8] Marek Ćwiklicki, Mariusz Duplaga, and Jacek Klich. *The Digital Transformation of Healthcare: Health 4.0*. Routledge, 2021.

[9] Arthur L Da Cunha, Jianping Zhou, and Minh N Do. The nonsubsampled contourlet transform: theory, design, and applications. *IEEE transactions on image processing*, 15(10):3089–3101, 2006.

[10] Yunyuan Gao, Tao Zhu, and Xiaohua Xu. Bone age assessment based on deep convolution neural network incorporated with segmentation. *International Journal of Computer Assisted Radiology and Surgery*, 15(12):1951–1962, 2020.

[11] William Walter Greulich and Sarah Idell Pyle. *Radiographic atlas of skeletal development of the hand and wrist*. Stanford university press, 1959.

[12] guandi1995.github. Cnn architecture. https://guandi1995.github.io/Classical-CNN-architecture/, May 2022.

[13] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Identity mappings in deep residual networks. In *European conference on computer vision*, pages 630–645. Springer, 2016.

[14] Marcia E Herman-Giddens, Jennifer Steffes, Donna Harris, Eric Slora, Michael Hussey, Steven A Dowshen, Richard Wasserman, Janet R Serwint, Lynn Smitherman, and Edward O Reiter. Secondary sexual characteristics in boys: data from the pediatric research in office settings network. *Pediatrics*, 130(5):e1058–e1068, 2012.

[15] Geoffrey E Hinton, Nitish Srivastava, Alex Krizhevsky, Ilya Sutskever, and Ruslan R Salakhutdinov. Improving neural networks by preventing co-adaptation of feature detectors. *arXiv preprint arXiv:1207.0580*, 2012.

[16] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708, 2017.

[17] G Iannaccone. Ww greulich and si pyle: Radiographic atlas of skeletal development of the hand and wrist. i volume-atlante di 256 pagine. stanford university press, stanford, california, 1959. *Acta geneticae medicae et gemellologiae: twin research*, 8(4):513–513, 1959.

[18] Vladimir I Iglovikov, Alexander Rakhlin, Alexandr A Kalinin, and Alexey A Shvets. Paediatric bone age assessment using deep convolutional neural networks. In *Deep learning in medical image analysis and multimodal learning for clinical decision support*, pages 300–308. Springer, 2018.

[19] Cefa Karabağ, Jo Verhoeven, Naomi Rachel Miller, and Constantino Carlos Reyes-Aldasoro. Texture segmentation: An objective comparison between five traditional algorithms and a deep-learning u-net architecture. *Applied Sciences*, 9(18):3900, 2019.

[20] David B Larson, Matthew C Chen, Matthew P Lungren, Safwan S Halabi, Nicholas V Stence, and Curtis P Langlotz. Performance of a deep-learning neural network model in assessing skeletal maturity on pediatric hand radiographs. *Radiology*, 287(1):313–322, 2018.

[21] Simon B Laughlin and Terrence J Sejnowski. Communication in neuronal networks. *Science*, 301(5641):1870–1874, 2003.

[22] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

[23] Hyunkwang Lee, Shahein Tajmir, Jenny Lee, Maurice Zissen, Bethel Ayele Yeshiwas, Tarik K Alkasab, Garry Choy, and Synho Do. Fully automated deep learning system for bone age assessment. *Journal of digital imaging*, 30(4):427–441, 2017.

[24] Min Lin, Qiang Chen, and Shuicheng Yan. Network in network. *arXiv preprint arXiv:1312.4400*, 2013.

[25] medical dictionary the free dictionary. Healthcare care definition. https://medical-dictionary.thefreedictionary.com/health+care, 2022.

[26] David J Michael and Alan C Nelson. Handx: a model-based system for automatic segmentation of bones from digital hand radiographs. *IEEE transactions on medical imaging*, 8(1):64–69, 1989.

[27] musculoskeletalkey. Assessment of skeletal maturity. https://musculoskeletalkey.com/assessment-of-skeletal-maturity/#CR8, May 2022.

[28] Radiological Society of North America (RSNA). Dataset. https://www.rsna.org/education/ai-resources-and-training/ai-image-challenge/rsna-pediatric-bone-age-challenge-2017, 2017.

[29] openclassrooms. openclassrooms courses. https://openclassrooms.com/fr/courses/4011851-initiez-vous-au-machine-learning/4020611-identifiez-les-differents-types-dapprentissage-automatiques, 2022.

[30] World Health Organization et al. *Primary health care: report of the International Conference on primary health care, Alma-Ata, USSR, 6-12 September 1978*. World Health Organization, 1978.

[31] pallipedia.org. tertiary healthcare care definition. https://pallipedia.org/tertiary-health-care/, 2022.

[32] Pranav Rajpurkar, Jeremy Irvin, Kaylie Zhu, Brandon Yang, Hershel Mehta, Tony Duan, Daisy Ding, Aarti Bagul, Curtis Langlotz, Katie Shpanskaya, et al. Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. *arXiv preprint arXiv:1711.05225*, 2017.

[33] Sebti Riad, Zroug Siham, Kahloul Laid, and Benharzallah Saber. A deep learning approach for the diabetic retinopathy detection. In *In The Sixth Smart City Applications Iinternational Conference, Accepted*. Springer, 2021.

[34] RR van Rijn and HH Thodberg. Bone age assessment: automated techniques coming of age? *Acta Radiologica*, 54(9):1024–1029, 2013.

[35] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[36] Mari Satoh. Bone age: assessment methods and clinical applications. *Clinical Pediatric Endocrinology*, 24(4):143–152, 2015.

[37] Sagar Sharma, Simone Sharma, and Anidhya Athaiya. Activation functions in neural networks. *towards data science*, 6(12):310–316, 2017.

[38] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. In Yoshua Bengio and Yann LeCun, editors, *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 2015.

[39] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9, 2015.

[40] James M Tanner and Robert D Gibbons. A computerized image analysis system for estimating tanner-whitehouse 2 bone age. *Hormone Research in Paediatrics*, 42(6):282–287, 1994.

[41] James Mourilyan Tanner. Assessment of skeletal maturity and prediction of adult height. *TW 2 Method*, pages 50–106, 1983.

[42] JM Tanner and Peter SW Davies. Clinical longitudinal standards for height and height velocity for north american children. *The Journal of pediatrics*, 107(3):317–329, 1985.

[43] Christoph Thuemmler and Chunxue Bai. *Health 4.0: How virtualization and big data are revolutionizing healthcare*. Springer, 2017.

[44] Chao Tong, Baoyu Liang, Jun Li, and Zhigao Zheng. A deep automated skeletal bone age assessment model with heterogeneous features learning. *Journal of medical systems*, 42(12):1–8, 2018.

[45] topaz Ai. download topaz ai tool. https://topazlabs.com/mask-ai/, 2022.

[46] Fei Wang, Mengqing Jiang, Chen Qian, Shuo Yang, Cheng Li, Honggang Zhang, Xiaogang Wang, and Xiaoou Tang. Residual attention network for image classification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3156–3164, 2017.

[47] H Wang, Z Lei, X Zhang, B Zhou, and J Peng. Machine learning basics. *Deep learning*, pages 98–164, 2016.

[48] Ari Wibisono, Mei Silviana Saputri, Petrus Mursanto, Joachim Rachmad, Ari Tri Wibowo Yudasubrata, Fadzil Rizki, Ernest Anderson, et al. Deep learning and classic machine learning approach for automatic bone age assessment. In *2019 4th Asia-Pacific Conference on Intelligent Robot Systems (ACIRS)*, pages 235–240. IEEE, 2019.

[49] Yang Xin, Lingshuang Kong, Zhi Liu, Yuling Chen, Yanmiao Li, Hongliang Zhu, Mingcheng Gao, Haixia Hou, and Chunhua Wang. Machine learning and deep learning methods for cybersecurity. *Ieee access*, 6:35365–35381, 2018.