Automatic Image Analysis for Bone Age Determination

By

ANDUALEM WUBE

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Biomedical Engineering

Center of Biomedical Engineering

Addis Ababa Institute of Technology

Addis Ababa University

Advisor: Dawit Assefa Haile (PhD)

Co-Advisor: Dr. Asfaw Atinafu

Addis Ababa, Ethiopia

December 2022

Declaration

I, the undersigned, declare that this thesis is my original work. It has never been presented for a degree in any other institution and that all sources of materials used in it have been duly acknowledged.

Name: _____

Signature: _____

Date: _____

This MSc. thesis has been submitted for examination with my approval as an advisor.

_____

Dawit Assefa Haile (PhD)

School of Graduate Studies

Certificate of Examination

This is to certify that the thesis prepared by Andualem Wube entitled "Automatic Image Analysis for Bone Age Determination" submitted in partial fulfillment of the requirements for the degree of Master of Science in Biomedical Engineering (Bioinstrumentation & Imaging) complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the examining committee

Examiner _____Signature _____ Date _____
(External)

Examiner _____Signature _____ Date _____
(Internal)

Examiner _____Signature _____ Date _____

Advisor _____Signature _____ Date _____

_____

Chief of Department or Graduate Program Coordinator

Acknowledgement

It gives me great pleasure to thank everyone who helped make this thesis possible. I want to primarily express my gratitude and special appreciation to my adviser, Dr. Dawit Assefa (PhD), who has been a fantastic mentor to me. I want to express my gratitude for supporting my research and letting me approach the situation from many angles. Your perseverance, inspiration, excitement, vast knowledge, and invaluable counseling regarding my profession as well as my research have been invaluable. Additionally, I want to thank Dr. Asfaw Atinafu, my co-advisor, for his support ever since I started developing the thesis idea.

I am also grateful to Dr. Masreshaw Demelash (PhD) for all his help and encouragement. A special thanks to my mom, brothers, and sisters, for their countless support, love, and patience. Finally, I am grateful to all my friends who encouraged me to strive towards my goal.

Abstract

Bone age evaluation is commonly performed through radiological assessment of the skeletal development of the left-hand, and then is compared with the chronological age. However, despite the time taking process, such image based automated processing remains incredibly challenging to implement. An accurate method is important to study the development of different wrist and pelvis structures which can predict age effectively. The proposed age determination scheme is composed of two major steps: a watershed based segmentation scheme used to generate segmented images which are used as inputs to a deep learning algorithm that is able to determine the age of a given subject. The deep learning scheme utilized the InceptionV3 architecture for its implementation. The model was trained, validated, and tested on radiological hand images acquired from the RSNA database. The model resulted in a mean square error of 4.4 months when compared against the available ground truth information. Overall, the results showed that the proposed bone age determination scheme comes with great promises.

Table of Contents

List of Figures

List of Tables

List of Acronyms

2D	Two Dimensional
3D	Three Dimensional
AE	Age Estimation
AS	Age Assessment
BAA	Bone Age Assessment
CAD	Computer Aided Diagnosis
CT	Computed Tomography
CRC	Convention on the Rights of the Child
EDPh	Epiphysis Distal Phalanx
EMPhx	Epiphysis Middle Phalanx
EPRPL	Epiphysis Proximal Phalanx
EMCRM	Epiphysis Metacarpals
DICOM	Digital Imaging and Communications in Medicine
G&P	Greulich and Pyle Method
MRI	Magnetic Resonance Image
PACS	Picture Archiving and Communication System
ROI	Region of Interest
RUS	Radius, Ulna and Short bones
PROI	Phalangeal Region of Interest
CROI	Carpal bone Region of Interest
TW	Tanner and Whitehouse Method
UNHCR	UN High Commissioner for Refugees
US	Ultrasound
AI	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
DCNN	Deep Convolutional Neural Networks
RSNA	Radiological Society of North America

Operational Definitions and Terms

Algorithm: set of clear rules or instructions for resolving a conflict or producing an outcome from a particular set of inputs.

Boundary: The edge of a region or item; the line created by uniting two picture regions with different light intensities.

Convolution: Superimposing an m x n operator (mask) over an area of an image, multiplying the points together, summing the results to replace the original pixel with the new value. This operation is often performed on the entire image to enhance edges, features, remove noise and perform other filtering operations.

Edge: A change in pixel values that exceeds a predetermined threshold. The boundaries between areas on an object or in a scene are referred to as edges.

Filter: A device or process that selectively transmits frequencies. In optics, the material either reflects or absorbs certain wavelengths of light, while passing others.

Gradient: The rate of change of pixel intensity (first derivative).

Gray Level: A quantized measurement of image irradiance (brightness), or other pixel properties typically in the range between pure white and black.

Gray Scale Image: An image consisting of an array of pixels which can have more than two values. Typically, up to 256 levels (using 8 bits) are used for each pixel.

Intensity: The relative brightness of a portion of the image or illumination source.

Kernel/mask: setting a consistent value to certain areas of an image and its neighbors; an operator for convolution that uses a filter matrix; a logical or physical element positioned in an optical system to block the transmission or viewing of information in a specific frequency or spatial range.

Median Filter: A method of image smoothing which replaces each pixel value with the median gray scale value of its immediate neighbors.

Noise: Irrelevant or meaningless data resulting from various causes unrelated to the source; Random, undesired video signals.

Orientation: The angle or degree of difference between the object coordinate system major axis relative to a reference axis as defined in a 30-measurement space.

Pixel: An acronym for “picture element”, it is the smallest distinguishable and resolvable area in an image; The discrete location of an individual photo-sensor in a solid-state camera.

Region: Area of an image. Also called a region of interest for image processing operations.

Resolution: The number of rows and columns of pixels in an image.

RGB: An acronym for the Red-Green-Blue color space. These three primary color systems are used for color representation.

Segmentation: The process of dividing a scene into several individual objects or contiguous regions, differentiating them from each other and the image background.

Shape: An object characteristic, often referring to its spatial contour.

Thresholding: The process of converting gray scale image into a binary image. If the pixel's value is above the threshold, it is converted to white. If below the threshold, the pixel value is converted to black.

Chapter 1: Introduction

1.1 Background of the Study

As humans age, their bodies grow and mature, especially during childhood and adolescence. Comparing physical maturity with age can be a valuable measure, but due to individual variations in skeletal and dental development, the methods used are often uncertain. Furthermore, the reference populations used for comparison are often not particularly relevant or representative. Although there is a wealth of literature on the subject of age estimation, few studies exist that offer an overview of the optimal methods to use. It is important to understand that each method has its own degree of uncertainty and precision, and this can vary greatly.

Age estimation can be done through various methods, depending on the field and context in which it is being used. In law enforcement, age estimation can involve sophisticated methods of analyzing physical features such as teeth, bones, and soft tissue, as well as assessing growth and development. Furthermore, in anthropology, age estimation is an invaluable tool for determining the age of skeletal remains, providing vital insights into the individual's life history and health. In medicine, age estimation may involve analyzing physical and developmental milestones, such as tooth eruption, bone growth, and pubertal changes. Also, age estimation can help to identify health risks and guide appropriate treatment plans based on a patient's age and stage of development. Accurate methods of age estimation are crucial to ensure that the correct information is obtained and that appropriate decisions are made in various fields. For instance, inaccurate age estimation in law enforcement can lead to unfair treatment and sentencing of individuals, while inaccurate age estimation in medicine can lead to incorrect medical treatment and potentially harmful outcomes.

No matter the methods used, age estimation will always involve some degree of biological variation and uncertainty. From a methodological standpoint, a combination of various measurement techniques, along with a more deliberate use of mapping and statistical methodology, could offer more accurate estimates and improved quantification of related levels of uncertainty. As it is not possible to precisely estimate the age of individuals with 100% accuracy, decisions must be made regarding the acceptable level of error when the degree of uncertainty is understood.

In both diagnostic and therapeutic studies of endocrinological issues, pediatric development abnormalities, and even genetic illnesses, the assessment of skeletal maturity (or "bone age") plays a particularly essential role [2]. In most cases, the skeletal development of the left hand is radiographically examined to determine the bone age, which is then compared to the chronological age. Pediatric patients routinely undergo bone age assessment to check growth, as well as to detect and treat a variety of endocrine diseases and pediatric syndromes. Skeletal maturity is a gauge of growth that takes into account the size, shape, and level of mineralization of bones to determine how close it is to full maturity.

The emergence of digital imaging has pushed researchers to strive for the development of innovative image-processing techniques that effortlessly extract the vital morphological characteristics of ossification in bones, thus allowing for a more accurate and unbiased approach to skeletal maturity assessments. Nevertheless, the crafting of effective computer algorithms to automatically calculate bone age has proven

to be a challenging task due to the intricate variations in bone mineralization tempo, shape, and size, as well as the massive amount of ossification centers in the hand and wrist. The traditional approach to skeletal bone age assessment (BAA) relies on identifying changes in the radiographic appearance of maturity indicators in a hand-wrist radiograph by contrast with a reference data set made up of series of radiographs divided by sex and age.

The bone age is calculated based on a radiological assessment of the skeletal growth of a left-hand wrist, and the results are then compared to the chronological age. If there is a difference between these two numbers, aberrant skeletal development is taking place. Radiologists do not recognize any technique as a standard, nevertheless. Greulich and Pyle's atlas matching method is the most popular approach (76%). A left-hand wrist radiograph is compared to atlas patterns using this technique. The pattern that most closely reflects the clinical image on the surface is chosen. The selection evaluates the bone age because each pattern is linked to a specific year of age. The error in determining the age is still a source of concern, though. The subjective nature of the analysis conducted by many observers with varying degrees of training is a drawback of this method. Another way uses the Tanner and Whitehouse (TW2) method, which is used by a skilled observer. This method employs a thorough evaluation of every bone, which results in a description of the bone in terms of scores.

Due to a number of variables, including sex, nutrition, hereditary factors, geographic location, and race, there was some debate regarding the timing of the formation of ossification centers and the fusing of epiphyses. Age estimation has a special practical significance when assessing the age of specific people who do not have birth certificates and who are enrolling in school, getting married, signing up for the military draft, determining their criminal culpability, or performing an illegal abortion. The aforementioned arguments suggest that age verification is crucial in many circumstances.

1.2 Overview of Age Assessment Methodologies

The materials that are currently available on the most popular techniques for performing chronological age assessments are reviewed in this section. The approaches to determining age make use of pre-existing local information as well as a variety of medical, physical, and psycho-social exams. Evidence suggests that the majority of experts concur that age assessment only ever provides an indicator of skeletal or developmental maturity from which judgments regarding chronological age may be deduced and is not a reliable method for determining age. According to the literature, skeletal assessment, external estimation, dental analysis, and social and psychological examination are the four 'pillars' of age estimation.

1.2.1 Social and Psychological Evaluation

A professionally trained clinician or social worker will undertake a Social and Psychological Evaluation to assess the mental, rather than physical, maturity of the subject. Through a slow process of inquiry and understanding, the practitioner will explore various areas of the individual's life history, assessing both their attitude towards past events and the way they responded at the time. It is essential that a trusting relationship is fostered between the individual and their assessor, so as to receive an objective opinion. Upon completion of the evaluation, it may be determined that no further physical estimation of age is necessary [5].

1.2.2 Physical Examination

A physical examination is an essential part of age determination. This include assessing physical measurements such as body type, weight, height, and body mass index (BMI) in addition to determining the outward indications of sexual maturity. It is also necessary to perform a general physical evaluation to detect any signs of pathological conditions which could affect the rate of maturation. Of the forensic methods used to determine age, assessing physical traits is one of the least accurate. Accurately assessing sexual maturity is notoriously challenging and should only be used to estimate age in combination with analyzing skeletal and tooth development. To further refine the process, there have been multiple studies which measure the correlation between physiological characteristics and chronological age in different demographics. In the end, a physical examination is still necessary to find any pathological problems that can affect physical growth [6].

1.2.3 Skeletal Age Assessment

Skeletal bone age is calculated by evaluating the stages of bone development, which are evaluated based on the maturation or fusion of particular bones.

• *Hand (wrist):* is one of the most widely employed methods for evaluating the hand wrist radiographs. It involves assessing the shape and size of the bone elements, as well as the level of epiphyseal ossification. To determine the development stage of a given image, professionals generally compare it to standard images of the same age and gender, as found in a radiographic atlas. Alternatively, the degree of maturity can be assessed for each bone individually and then combined to form an overall maturity stage. The Greulich and Pyle (G&P) atlas has become the go-to reference for the first approach, while the TW2 approach (in three editions) reigns supreme for the second. Although one might assume the TW2 method is more trustworthy, there is no hard evidence to back this up, and it is often more laborious and time-consuming. On average, the bones in the hand reach full maturity at age 17 for girls and 18 for boys (see Fig. 1) [6].

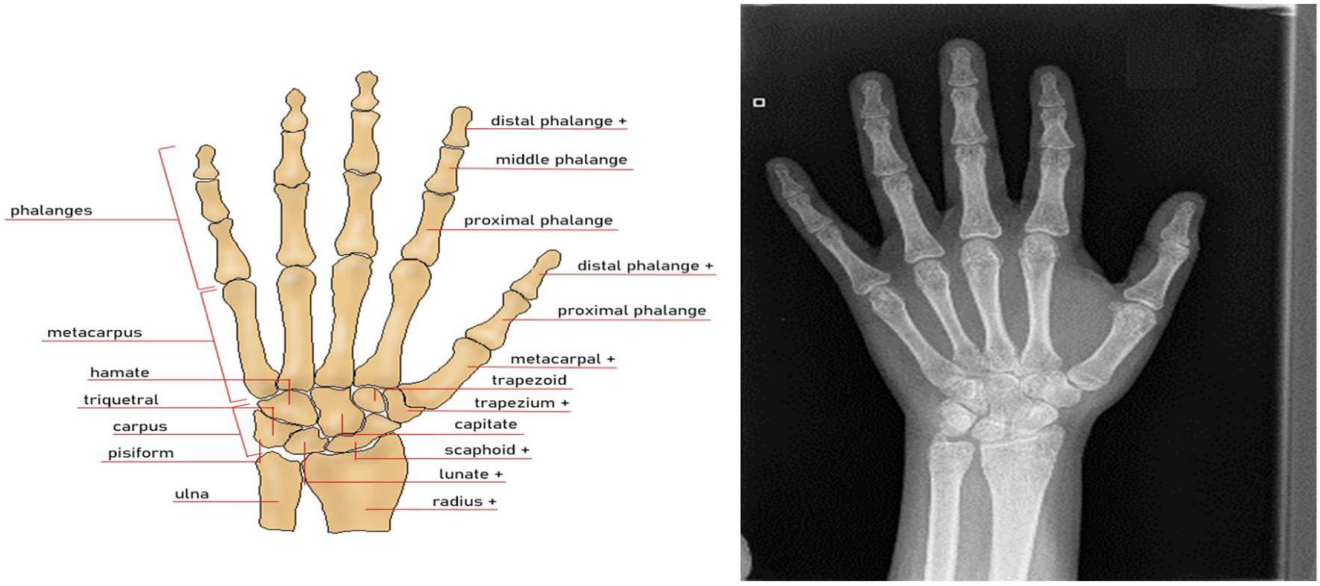

Figure 1: Hand (wrist) image [7].

• *Clavicle (collar bone):* The ossification of the cartilage at the sternal end of the clavicle (see Fig. 2) can be evaluated to provide a reliable estimate of the age of individuals aged 18 and above, since all other developmental systems examined have reached their full growth by this age. According to conventional classification systems, there are four stages of clavicle ossification: stage 1, where the ossification center is not ossified, stage 2, where the ossification center is ossified but the epiphyseal plate is not ossified, stage 3, where the epiphyseal plate is partially ossified, and stage 4, where the epiphyseal plate is completely ossified. Some literatures divide the last stage into two additional stages: stage 4 (where the epiphyseal plate is fully osseous and the epiphyseal scar is visible) and stage 5 (where the epiphyseal plate is fully osseous and the epiphyseal scar is no longer visible) [5, 6]. It is possible to accurately assess the age of an individual based on the fusion of their epiphyses and the presence of an epiphyseal scar. Females are typically at least 20 years old when their epiphyses have fused, while males require an additional year, reaching the age of 21 before their epiphyses have fused. The earliest age at which total fusion and disappearance of the scar have been noted in both sexes is 26 years.

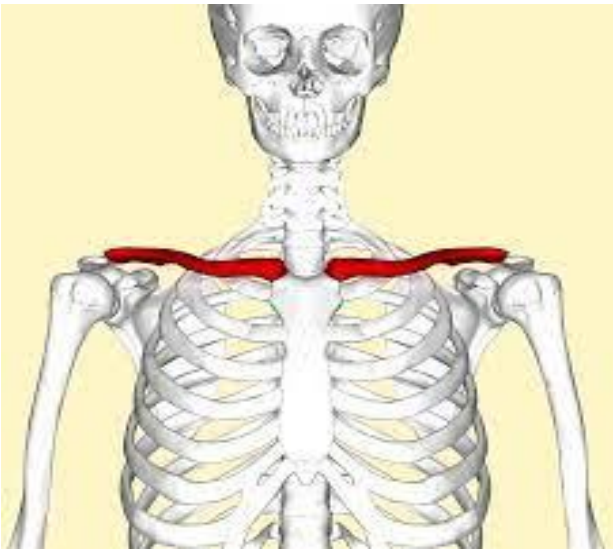

Figure 2: Clavicle (collar bone) image [8].

• *First rib:* Recent research has revealed a reliable technique for age assessment in the living by analyzing the ossification of the first rib (depicted in Fig. 3). This technique can be utilized in conjunction with the analysis of clavicle ossification, as both processes can be derived from the same set of X-ray radiographic images. Digital thorax X-rays radiography enabled researchers to grade the ossification of the costal cartilage of the first rib according to four stages, and the results were then analyzed in light of the known age and sex of the subjects. The findings suggest that the ossification of the first rib could potentially be used as an effective way to determine age in those around 21, yet further studies on more populations are essential to validate this. Earlier studies have highlighted that respiratory stress is the driving factor behind the ossification of the first rib. Therefore, individuals who are exposed to respiratory stress due to leisure activities or work-related duties could be more likely to experience an accelerated development than others [5, 6].

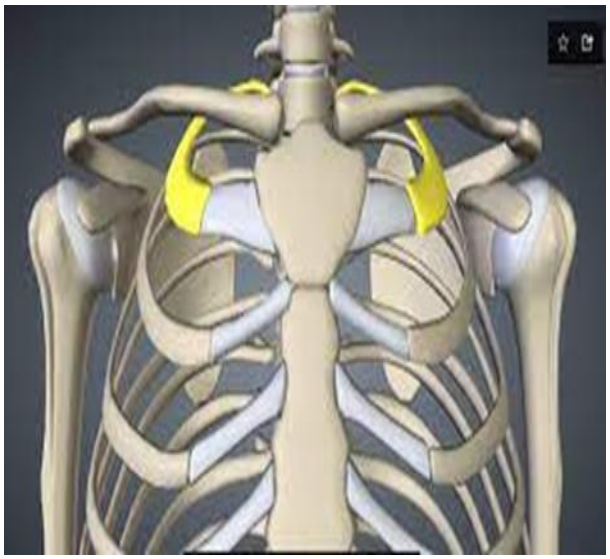

Figure 3: First rib image [8].

• *Cervical vertebrae (vertebrae of the neck):* Recent research has elucidated the development of the cervical vertebrae (see Fig. 4) through examination of cephalometric radiographs in tandem with the growth and emergence of the third molars. This method has recently been proposed as a potential way to estimate the age of the living. The efficacy of various grading methods was assessed, and the two most accurate ones were used in conjunction with third molar development phases. When used to assess the age of subjects under 14, the combined approaches yielded a remarkable increase in accuracy. Despite the fact that no or a minimal improvement in precision was witnessed for older age groups (late third molar development stages), studies still advocate this approach as supplementary to third molar information when the root of the third molar is less than ¾ developed. This recommendation comes as cephalometric radiographs need less radiation dose than chest and hand wrist radiographs [5, 6].

Figure 4: Cervical vertebrae (vertebrae of the neck) image [8].

• *Hip (iliac crest):* The iliac apophysis of the pelvis is an indicator of skeletal age, easily identified on a lateral pelvic X-ray. As a person nears adulthood, this apophysis progresses towards the spine, a phenomenon known as *Risser's sign*. This sign is used by specialists to measure the growth left in the spine, with 5 stages defined. Generally, these stages can be observed in girls aged between 14 and 16, while boys typically present them between 15 and 18. Ultrasound examinations have recently been suggested as a viable method for age estimation in the living. To assess its applicability, a pilot study was conducted to analyze the development stages of the iliac crest apophysis of 23 male and 16 female subjects, with age ranging from 11 to 20 years. Using the sonographic staging of clavicular ossification as a reference, the results of the study uncovered promising insights into the use of ultrasound for age estimation. At a minimum age of 15.7 years, the male subjects had entered ossification stage 1. For boys, ossification stage 2 was detected at a minimum age of 14.1 years, while for girls it occurred at 11.7 years. The earliest observation of ossification stage 3 was observed at 16.2 years for males and 15.2 years for females. Although the earliest stage of ossification, stage 4, was observed in male test persons at least 18.0 years of age and in female test persons at least 17.1 years of age, the approach can still be seen as a valid and efficient method. However, to further confirm its efficacy, larger studies are necessary [5, 6].

• *Knee:* This depends on fusion of growth plate maturation or ossification of the knee. Studies have devised a pioneering magnetic resonance imaging (MRI) staging system for evaluating the maturation of the knee growth plate's epiphyseal fusion, and examined its dependability and accuracy for age assessment from 10-30 years old. Notably, five distinct MRI stages were pinpointed, revealing a high correlation with age and remarkable inter- and intra-observer consistency. Though, further studies are required to validate the approach [5, 6]

1.2.4 Dental Age

Skeletal methods come with certain pitfalls when it comes to the variable nature of bone maturation which can be shaped by diet and other environmental components. Uncertainties concerning nutrition and its impact on dental development have been dispelled as it has been determined that the calcification rate is largely driven by genetic factors rather than environmental ones. The accurate determination of dental age is essential in many cases, and can be assessed by either the eruption of teeth or through examining radiographs to grade the development of the tooth crown and roots (see Fig. 5). Nowadays, Orthopantomograms, which provide a comprehensive overview of all teeth, are the most widely used, and the introduction of digital images has drastically mitigated the exposure to X-ray radiation. Through the stages of development, from initiation (a.k.a. crypt stage) to completion (closure of the dental root apex), the changes in the chosen teeth are recorded. As the roots of all teeth complete their growth around the age of 20, other degenerative changes in the teeth must be taken into account.

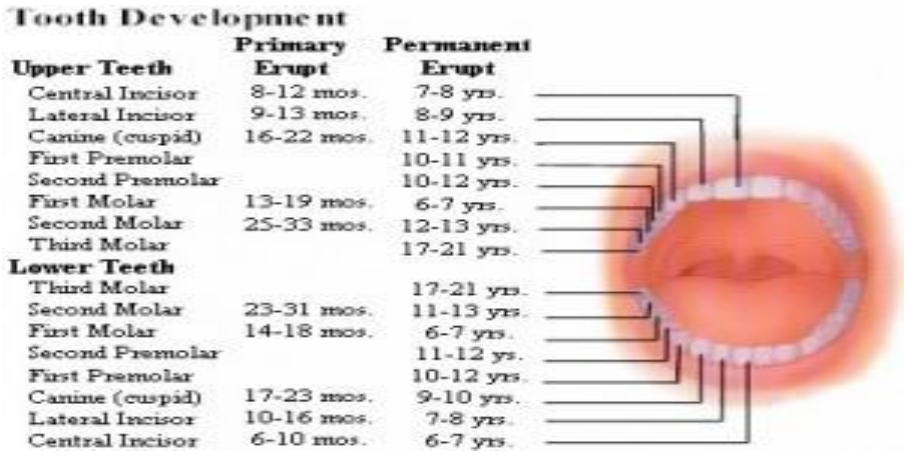

Figure 5: Dental age image [8].

1.2.5 Ethical Concerns Regarding Age Assessment Methodologies

The radiologists community has recommended its fellows and members that it was improper to perform a radiograph examination with the intention of estimating age. In addition to concerns regarding the precision of bone and dental age assessments, there are significant ethical concerns expressed by both the medical community and external critics regarding the use of these medical testing.

Although there are ethical issues about exposing children to any level of radiation, there is little radiation exposure during an X-ray procedure in relation to age assessment. The European Council's Directive 97/43/Euratom on protecting people's health from the risks of ionizing radiation in regard to medical exposure states that for residents of the European Union, medical exposure is the main source of exposure to man-made sources of ionizing radiation. The net benefit to an individual must outweigh the dangers to that individual's disadvantage, according to Article 3 of this directive. It is challenging to imagine how radiation exposure as part of an age assessment technique might be advantageous to the person. Further, Article 3 states that special consideration should be given to the justification of radiation exposure in cases where there is no immediate improvement in the person's health. The directive specifies that Member States shall utilize appropriate radiological equipment, practice techniques, and equipment when exposing minors to radiation. It also urges additional caution to be taken in this situation. In any case, exposure should only be carried out with the subject's agreement and only after they have been made aware of the hazards involved in the operation.

The autonomy, welfare, and consent of patients are essential ethical values in the dentistry and medical professions in the UK (United Kingdom), according to Crawley. The government contracts and pays a small number of doctors and dentists to perform dental and bone age tests, which violates all three of these ethical precepts. In 2007, Crawley (and also Ranta (2003)) explained that when age assessment is performed for forensic purposes, there are rarely any issues with the motivations behind the assessment or its outcomes, but when tests are performed on living individuals, ethical issues come into play. According to her, issues should be raised, such as: Whose interests are served by testing? Is it in the best interests of the particular child or the society to which the child has been sent? Like that.

Any physical examination conducted as part of the age assessment process must consider a number of ethical concerns. Given their age, maturity, and comprehension of the procedure, has the child provided

informed consent for the procedure? The ethical justification for the examination would be called into question if the practitioner conducting the assessment has any concerns that the kid might not completely understand the procedure or has been threatened, pressured, or motivated in some other way to agree to the procedure. A medical checkup might exacerbate any trauma or discomfort experienced by youngsters. Any medical expert performing a non-diagnostic operation should undoubtedly consider why they are doing it, who will profit from the test, and whether the child will get anything from it. While a skilled practitioner's physical examination may unintentionally reveal medical conditions in a child that could then be treated, such occurrences do not in and of themselves justify a physical examination. Assessments must never infringe on a child's CRC rights in any situation. The main rights in relation to the ethics of age assessment methodologies include that the best interests of the child must be the first consideration in all actions involving children (Art. 3), that the opinions of children must be given due consideration in relation to their age and maturity, that children must have opportunities to be heard in all proceedings affecting them (Art. 12), and that children have the right to be protected from arbitrary or unlawful interference with their privacy [6,7].

In situations where it is unclear how old a child is, age evaluations are carried out. They must be a part of a thorough evaluation that takes into account both the person's outward appearance and their psychological maturity. Such evaluations must be conducted in a way that is secure, considerate of children and gender, and respectful of human dignity. All age-assessment techniques have a margin of appreciation, which must be implemented in a way that, in the event of a tie, the subject will be regarded as a child. Age is not measured or valued equally everywhere, thus it is important to use caution when drawing unfavorable conclusions about believability from cultural or national standards that appear to lower or raise a child's age. In a language they can understand, children must be provided explicit explanations of the procedure's goal and steps for age assessment. A certified independent guardian must be appointed before an age assessment procedure is conducted to counsel the child [6, 7].

The current methods of bone age assessment have been primarily derived from a white population, raising the question of their generalizability on children of other ethnicities. This raises a concerning issue of fairness, as our lifestyle and genetic makeup can differ greatly from other regions of the world. Thus, there is evidence to suggest that these methods may be inadequate in assessing the bone age of those from other ethnic back grounds.

If either of the methods is employed, particularly the G&P-method, the person being evaluated runs the potential danger of being erroneously identified as an adult. It is important to note that a child's bone age and their chronological age may not always match. An advanced bone age is observed in healthy, tall children, while healthy, short children tend to have a delayed bone age [9].

Table 1 outlines a range of conditions, syndromes, and diseases which can have a considerable effect on the age assessment, paving the way for an inaccurate or unjust assessment. These issues may contribute to either an advanced or delayed bone development.

Table 1: Potential causes of delayed and advanced bone age.

S.no	Delayed Bone Age	Advanced Bone Age
1	Deficiency of growth hormone that is untreated	Childhood Obesity
2	Different Chronic diseases	Leydig cell hyperplasia
3	Inflammatory bowel disease	McCune-Albright syndrome
4	Celiac disease	Sotos and Beckwith-Wiedemann syndromes
5	Cystic fibrosis	
6	Anorexia	
7	Depression	

1.3 Statement of the Problem

In developing countries, particularly in Ethiopia, age estimation/reading is an issue that radiologists struggle with due to the absence of automated techniques, inadequate image interpretation, and a lack of properly trained personnel for age estimation interpretation. Moreover, the deficiency of compiled research and databases, alongside the dearth of well-ordered central options, further impede the process of age estimation reading.

The left-hand images from CT scans, MRIs, and digital X-rays can be manually segmented and analyzed by radiologists to estimate age. While this method is accurate, it is undoubtedly tedious, time-consuming, prone to subjectivity, and impractical in today's medical imaging diagnosis, where many images are taken for a single patient or person. Accurately segmenting images from the left-hand is vital for straightforward screening and interpretation of skeletal maturity of each morphology; it is instrumental in enhancing clinical outcomes and is the only developmental indicator that is available from birth to adulthood. It provides a better source of information for studying the various levels being an excellent estimate of the growth, development, health, and nutrition of a child. This makes skeletal age assessment invaluable for diagnostic and therapeutic decisions in pediatrics, endocrinology, orthopedics, and orthodontics. In addition, in medico-legal cases, skeletal age is the most trusted method of estimating age and is a crucial element to determine criminal culpability. For nations with many young asylum seekers, football teams, military recruitments, driver licensing and athletic competitions, determining age is vital for ensuring equal opportunities and fair play. In countries like Ethiopia, where birth records are often absent, this challenge is particularly acute. Skeletal age is typically estimated through comparing a single hand-wrist radiograph (left by convention) with a standard reference.

The process of segmenting and interpreting data is hence in high demand. The accurate and reproducible segmentation and characterization of morphologies using intelligent algorithms remain challenging and difficult tasks despite the numerous efforts and encouraging results in the medical imaging community. This is due to the wide range of shapes or locations, image intensities of various types of imaging modalities, as well as societal diversities. The images produced by the imaging devices are also susceptible to a number of abnormalities, including noise, poor image contrast, existence of anatomical

structures with pixels that are strongly connected with those of other structures, variable lighting, and hand movements during image acquisition.

The ability to utilize data from medical images in novel and ground-breaking ways has been made possible by recent developments in digital imaging and computational technology. The skeletal maturation stage and surrounding morphologies of the photos should be easily distinguishable by such a method. The automated interpretation systems boost productivity, give accurate reading, and free doctors from tedious labor. The effectiveness of this kind of age estimation method depends on precise image acquisition, and particularly on precise and trustworthy image processing methods for recognizing the borders.

In this regard, the current thesis work focuses on development of automated age estimation using watershed segmentation systems and deep learning models. By enhancing the visualization of left hand-wrist from X-ray acquired images, the proposed system can potentially improve the accuracy of age estimation by providing clearer and more detailed images for analysis. Additionally, deep learning training models can potentially improve the accuracy of age prediction by using algorithms to identify patterns and relationships in the images that may be difficult to detect by human observers alone.

The long-term research objectives are to create, verify, and promote the utilization of the most reliable techniques in determining the age of children and adolescences. These methods provide impartial, consistent outcomes, and ensure repeatable results. This research must adhere to realistic financial and practical restrictions, as well as abide by established ethical guidelines. To pursue this mission, it is essential to evaluate the benefits and drawbacks of the current methods in an organized way. Additionally, exploring the potential of combining various methods must be thoroughly explored. Carrying out a systematic review of studies of reference populations is a crucial step. To further progress in this field, new and intricate reference data, with precise age information, should be gathered through intercontinental collaboration. Furthermore, alternative methods to efficiently exploit better existing reference data should be explored from local sources [1].

1.4 Objectives of the Thesis

1.4.1 General Objective

The general objective of the thesis is to develop an automated age determination method using X-ray wrist image segmentation and prediction system.

1.4.2 Specific Objectives

The study intends to accomplish the following specific objectives.

- To acquire a dataset of hand wrist X-ray images from secondary source.
- To perform image segmentation on the collected wrist X-ray images to extract relevant features for age estimation.
- To design and develop a deep learning model architecture that can effectively estimate the age of the client based on the X-ray image of the hand wrist.
- To collect local wrist X-ray images from hospitals to expand the dataset.
- To evaluate the performance of the proposed method.

- To estimate age from the x-ray image

1.5 Significance of the Study

The gap in this area that demonstrates that only 50% of children under the age of five in underdeveloped countries have their births registered is the reason for the necessity for age assessment. About 64% of births in sub-Saharan Africa and 65% of births overall in South Asia are unregistered (UNICEF, 2010). Children may suffer greatly as a result. Their vulnerability and the likelihood that rights breaches will go unreported are both increased by their official invisibility. Children are more susceptible to underage recruitment into the armed forces, exposure to dangerous jobs, and early marriages, for instance, when they lack the necessary documentation to verify their age. Additionally, they are more likely to face adult treatment in criminal proceedings than juvenile or child treatment when requesting international protection as asylum seekers [5].

Moreover, this thesis aims to solve the problem in the Ethiopian Hospitals age determination systems by designing an automated approach to improve the manual image based segmentation and analyses of age estimation by radiologists. The manual procedure is no doubt tedious, time-consuming, highly subjective, and impractical in today's medical imaging diagnosis. Accurate segmentation of the images is critical for easy screening and interpretation of skeletal maturity of each morphology which is helpful to improve clinical outcomes being the only developmental indicator that remains at hand from birth to adulthood. It provides a better source of information for studying the various levels and an excellent estimate of the growth, development, health, and nutrition of a child. Skeletal age assessment is therefore crucial for pediatric, endocrinology, orthopedic, and orthodontic diagnostic and treatment decisions. Additionally, in medico-legal cases, skeletal age is thought to be the most reliable indicator of age; its estimation plays a significant role in determining criminal responsibility. Additionally, immigration authorities in several nations face difficulties when determining the ages of young asylum seekers, and athletic and football competitions are organized by age groups to ensure equal opportunities in the spirit of "fair play", particularly in developing nations like Ethiopia. Skeletal age is commonly calculated by comparing a single hand-wrist radiograph (usually taken with the left hand as per protocol) to a standard reference. The current study may identify gaps for researchers who would like to conduct detailed and comprehensive studies either in public or private institutions.

1.6 Organization of the Thesis

The remaining portion of the thesis has been set out as follows: The literature review in Chapter 2 looks at some of the most recent techniques for determining age. The framework for age detection is presented in the third chapter. In this chapter, the first two modules of the proposed framework—the preprocessing and post-processing stages for age determination are briefly discussed. Special attention is given to the preferred watershed segmentation of bone morphologies, which have a major impact on age due to their progressive changes over developmental years. The proposed deep learning method is discussed in detail, providing insight into the hybrid system and techniques used to determine age estimations. Chapter four presents the experimental results and discussion of the proposed framework. Finally, Chapter five wraps up the thesis and presents ideas for future work and improvements to the proposed system.

Chapter 2: Skeletal Image Age Assessment and Related Works

2.1 Bone Anatomy Development - Overview

Skeletal maturity is a gauge of growth that takes into account the shape and size, and level of mineralization of bones to determine how close it is to full maturity. The evaluation of skeletal maturity necessitates a thorough investigation of numerous variables and a comprehensive understanding of the various mechanisms by which bone develops. The process of endochondral ossification causes the long bones of the extremities to develop longitudinally. In contrast, skeletal tissue develops straight from fibrous membrane, increasing the diameter of the bones. The latter is the method through which ossification of the calvarium, the flat pelvic bones, the scapulae, and the body of the mandible occurs, and early calcification starts close to the middle of the shaft of long bones in an area known as the primary ossification center.

Despite the fact that many flat bones, including the carpal bones, ossify exclusively from this primary center, all long bones have secondary centers that appear in the cartilage of the bone's extremities. Similar to how it does in primary centers, these centers continue to mature through the ossification of cartilage and invasion of osteoclasts and osteoblasts. The diaphysis and epiphysis are the two bones that ossify from the primary and secondary centers, respectively. The cartilage is gradually replaced by bone when the secondary center ossifies, leaving only the epiphyseal plate to divide the diaphyseal bone from the epiphysis. The metaphysis, which refers to the growing end of the bone, is the portion of the diaphysis that touches the epiphysis. The diaphysis and epiphysis both continue to develop if the epiphyseal cartilage plate is still there, but eventually the osteoblasts stop reproducing and the epiphyseal plate becomes ossified. That is when the osseous structures of the diaphysis and epiphysis are fused and growth ceases [10]. A schematic of endochoridral bone formation is depicted in Fig. 6.

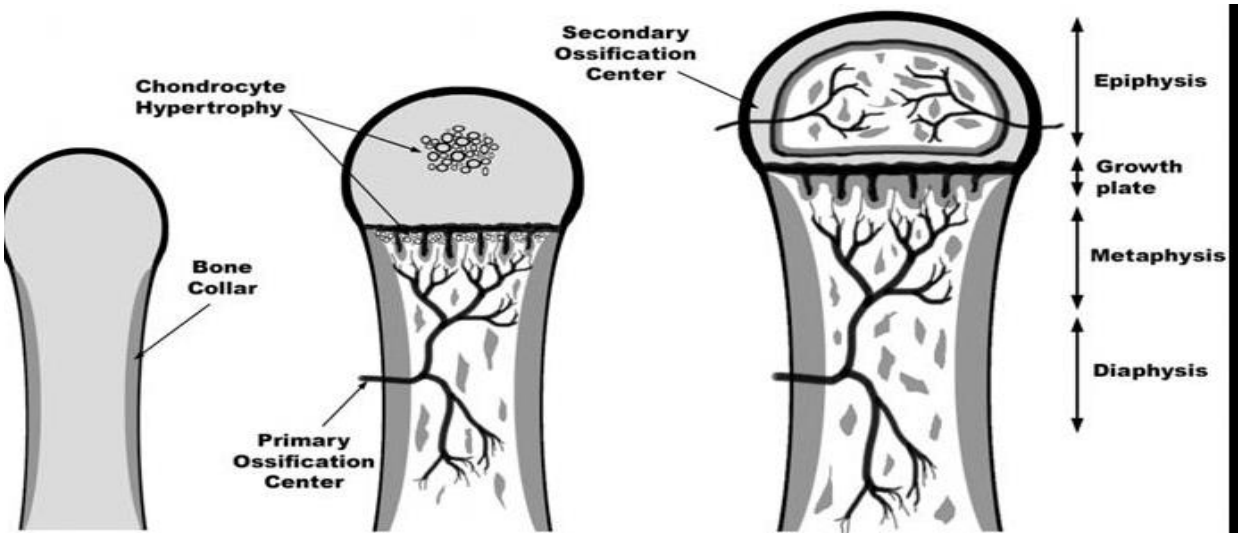

Figure 6: Schematic representation of endochondral bone formation.[10]

The degree of secondary ossification centers' growth and ossification in the epiphysis serves as a measure of skeletal maturity. The main focus on skeletal development throughout the fetal stage of life is connected to the preterm diagnosis. The calcification event, which starts at 8 or 9 weeks (approximately 2 months), signifies the end of the embryonic stage and the beginning of the fetus. The majority of the principal centers of the tubular bones have evolved into diaphyses by the thirteenth fetal week, and at birth, all of the diaphyses are fully ossified while the majority of the epiphyses are still cartilaginous. The distal femoral epiphysis, a secondary center that osseointegrates in the last two months of pregnancy, is present in the majority of full-term infants. Similar to this, the proximal epiphysis of the humerus' ossification center typically develops around the 40^{th} week of pregnancy. On the other hand, it's possible that the centers for the proximal epiphyses of the femur and tibia do not develop until the first few months of life [10].

At skeletal maturity, the epiphyses fuse with the main body of the bone after gradually ossifying in a predictable order after birth. The evaluation of skeletal maturity, often known as "bone age" or "skeletal age", is based on comparing the degree of epiphyseal maturation to typical age-related norms. Although the exact causes of a normal maturational pattern are unknown, it is known that genetics, environmental factors, and hormones like thyroxine, growth hormone, and sex steroids play significant roles. Studies carried out in patients with mutations of the gene for the estrogen receptor or for aromatase enzyme have proved that it is estrogen that is primarily responsible for ultimate epiphyseal fusion, although it seems unlikely that estrogen alone is responsible for all skeletal maturation [4,5,10].

Various models have been generated before based on rigorous analyses of the maturation of ossification centers in the hands and wrists of healthy children, and created virtual images incorporating the average development for each ossification center in each age group. Age is also decisive for the punishment of delinquents in a court of law, for military, for driving license. In order to ensure equal chances and uphold the idea of "fair play", sports, in particular athletic and football events, have been designed according to age groupings.

2.1.1 Clinical Applications for Skeletal Determinations

When combined with other clinical symptoms, a single reading of skeletal age can help a doctor distinguish between a patient's relative maturity at a given point in time and that of an advanced or retarded patient. The direction of the kid's development and/or the progress the youngster is making with treatment are indicated by subsequent skeletal age readings. Bone age should be within 10% of chronological age in healthy persons. In children who are obese or who begin puberty early, there is a greater discrepancy between skeletal age and chronological age because their skeletal age is accelerated [10].

2.1.2 Conventional Techniques for Skeletal Determinations

Variations in maturation pace are inadequately described by chronological age in assessments of children's physical development. As a result, for many years scientists have been seeking better ways to evaluate the level of development from birth to full adulthood. Despite being strongly related to biological maturation, measurements of height, weight, and body mass are insufficiently reliable because of the vast

ranges in body size. Similar to how dental age cannot be used as a general indicator of maturation due to the wide differences in dental development, other clinically validated methods have limited use. For instance, while the age at menarche (the first period) is an important biological indicator, it only applies to half the population, and the Tanner classification, while a very helpful clinical tool, can only be used to determine the stage of sexual development during adolescence. Unfortunately, the majority of maturational 'age' scales in use have particular purposes and tempos that do not always match.

The most often used indicator of a growing person's biological maturation is their skeletal/bone age, which is determined by looking at the various phases of skeletal growth as seen in hand-wrist radiographs. The only development indicator currently accessible that covers the full growth period from infancy to maturity is this method, which is utilized by pediatricians, orthopedic surgeons, physical anthropologists, and everyone else interested in the study of human growth. The expansion of the ossifying region and the calcium deposition in that region are two characteristics that determine the degree of skeletal maturity. The development of these two characteristics follows a clear pattern and timetable from infancy to adulthood, even though they may not always occur together. This approach uses radiographs to offer a useful standard for determining typical and atypical growth and maturation [5,6].

In addition to the clinical examination, which includes looking at things like weight, height, teeth, and evidence of physiological puberty, the radiological examination provides a quick and accurate way to estimate the age in people who are 25 years of age or younger. Due to the complete emergence of the ossification center and the fusion of the epiphyses in adults over the age of 25, it is impossible to determine an individual's age with precision through radiological examination. Only a few age-dependent features remain to be used for age estimation by morphological methods at the end of skeletal growth and development, which leads to a less precise estimation of age [4]. These features include the development of third molars and the bones in the hand and wrist.

By comparing a series of radiographs from the hand and wrist with a reference data set made up of radiographs categorized by age and sex, the classical method of skeletal bone age assessment (BAA) identifies changes in the radiographic appearance of the maturity indicators. The G&P atlas is the most often used reference standard. The atlas has not altered since it was first published, and it is frequently used in clinical settings to determine the bone age of kids of African American, Caucasian, Asian, Hispanic and other ancestries [11].

The TW2 method and the G&P method are the two primary clinical techniques for skeletal bone age assessment. The two approaches diverge in a number of ways. The strategy that is most frequently employed is G&P. The fundamental reason for this is that the G&P approach is quicker and more user-friendly than the TW2 method. However, studies have shown that the two approaches yield different estimates of skeletal age, and that these variations matter in clinical settings. In this sense, the reliability of an automated method contributes to more accurate age determination. Although the two methods have different theoretical underpinnings, both theories are based on the identification of maturity indicators, such as changes in the radiographic appearance of the tubular bones' epiphyses from the earliest stages of ossification until fusion with the diaphysis, or changes in flat bones until they take on an adult shape.

Unquestionably the most widely used technique, the standards established by G&P are made up of two series of standard plates taken from hand-wrist radiographs of white, upper middle-class boys and girls who participated in the Brush Foundation Growth Study from 1931 to 1942. The "central tendencies" or modal levels of maturity within chronological age groups, are depicted in the G&P atlas. Each standard has a skeletal age associated with it that is the same age as the kids that served as its foundation. By comparing the radiograph to be evaluated with a series of standard plates, the G&P method determines the child's skeletal age by assigning the age to the standard plate that fits the radiograph the closest. When determining the proper age for a radiograph, it is frequently convenient to interpolate between two standards. This atlas is the most widely used standard of reference for skeletal maturation worldwide since determining a skeletal age appears to be quick and easy.

The foundation of the G&P atlas is the assumption that in healthy children, skeletal development is uniform, all bones have the same skeletal age, and the appearance and subsequent growth of body centers follow a predetermined pattern. The ossification patterns of the numerous bones of the hand and wrist, however, exhibit a wide range of normal variation that is genetically dictated, according to a lot of data. In reality, the majority of criteria in the atlas comprise bones with varying degrees of maturity [12].

G&P did not expressly suggest any particular method for using their atlas. It is simple to conduct the examination since it only requires one view of the left hand, which includes all pertinent regions of interest in the hand and wrist. The final height of a person is frequently predicted using bone age findings. The evaluation is subjective because the radiologist evaluates each hand and wrist bone separately to determine its average age before combining the data to obtain the best fit with the atlas reference radiographs. Additionally, they advised atlas users to create their own technique based on their personal preferences. However, Pyle et al. did recommend the laborious method of assigning a bone-specific bone age to each ossification center and then computing the average of the ages. The distal centers should be given more weight when there is a difference between the carpal bones and them since they often correspond better with future growth.

Regarding bone age, there are a number of crucial considerations that must be taken into account. First, to improve inter- and intra-observer reproducibility, expertise in skeletal maturity assessments and a similar analytical approach are required. The inclusion of experienced readers who employ similar evaluation techniques in clinical investigations and trials using bone age as an end metric is beneficial. In addition, there is ethnic diversity in the normal pace of skeletal maturation between boys and girls. Last but not least, children with skeletal dysplasia, endocrine problems, or a number of other causes of growth retardation may not necessarily be able to use these references. However, the method is largely qualitative and unreliable [13].

Numerous initiatives have been taken to improve the bone age assessment's accuracy and eliminate racial and ethnic disparities in skeletal maturation. The TW2 method works by giving each bone in the hand and wrist a unique score. The bone age is then calculated by dividing the overall score by the number of bones. Additionally, by applying the Fourier transform to radiographs that had been digitally captured by a video camera, the scoring method was computerized. The process requires a lot of time and effort to complete.

The phalanx, epiphyses, and carpal bone areas are the regions of interest (ROI) that are most important in determining bone age. They then use a range of image processing techniques to extract the quantitative information related to these ROIs. The Tanner maturity index, race, age, and sex are related with these computer-extracted parameters, offering a useful method of quantitative assessments that serves as the foundation for computer-assisted bone age assessment [13].

Table 2 summarizes the Chronological Age Determining Reference Parameters [10].

Table 2: Chronological age group versus age range.

S.no	Reference Images	Age range for Females	Age range for Males
1	Infancy	0 - 10 Months	0 -14 Months
2	Toddlers	10 Months - 2 Years	14 Months - 3 Years
3	Pre-puberty	2 Years - 7 Years	3 Years - 9 Years
4	Mid and Early puberty	7 Years - 13 Years	9 Years - 14 Years
5	Late Puberty	13 Years - 15 Years	14 Years - 16 Years
6	Post-puberty	15 Years - 17 Years	17 Years - 19 Years

2.1.3 Computer Assisted Techniques for Skeletal Determinations

To quantify radiographic observations that might support the determination of bone age, computer-aided methods have also been developed [13]. With the development of digital imaging, a number of researchers have attempted to provide an accurate computer-assisted measure for determining bone age and have created image processing methods from reference databases of healthy children that automatically extract important details from hand radiographs. To date, however, efforts to create automated image analysis methods that can extract quantitative measurements of the morphological traits illustrating skeletal maturity have been hampered by the inability to take into account the wide variation in the growth and ossification of the numerous bones in the hand and wrist. Automated methods that mainly rely on measurements from a few selected ossification centers, such as those of the epiphyses, are being developed to get around these problems. The reference digital atlas was created using an alternate strategy to avoid the difficulties involved in creating software that incorporates all morphological parameters. They created artificial, idealized, sex- and age-specific illustrations of skeletal development that took into account the various stages of development of each ossification center in the hand and wrist. The idealized image was created from a composite of numerous hand radiographs taken of healthy kids and teenagers who were chosen as the ideal average for each ossification center in each age group [6, 10].

2.1.4 Indicators of Skeletal Maturity in Children and Adolescents

This section explains which hand and wrist bones, during the various stages of postnatal development, serve as the best markers of skeletal maturity. The carpal, metacarpal, and phalangeal bones ossify in a predetermined order in most healthy children that is astonishingly consistent and the same for both sexes. The capitate is typically the first ossification center to be seen on hand and wrist radiographs, while the sesamoid of the thumb's adductor policies is typically the last one [10]. The distal radius is the first epiphyseal center to form, and it is then followed by the proximal phalanges, metacarpals, middle

phalanges, distal phalanges, and, lastly, the ulna. However, there are two main exceptions to this rule: the thumb's distal phalanx and the metacarpals' epiphyses frequently appear at the same time, and the middle phalanx of the fifth finger is frequently the last to ossify. The viewer should pay particular attention to the ossification centers that most accurately represent skeletal development for the subject's chronological age because the predictive value of these centers varies and shifts as a person grows.

As listed below, the skeletal development is broken down into six primary categories in order to make bone age evaluations easier. For each group, the specific ossification sites that are the strongest predictors of skeletal maturity are underlined in parentheses:

a) Infancy (the radial epiphyses and carpal bones);

b) Toddlers (the numbers of epiphyses visible in the long bones of the hand);

c) Pre-puberty (the sizes of the phalangeal epiphyses);

d) Early and Mid-puberty (the sizes of the phalangeal epiphyses);

e) Late Puberty (the degree of epiphyseal fusion), and

f) Post-puberty (the degree of epiphyseal fusion of ulna and the radius).

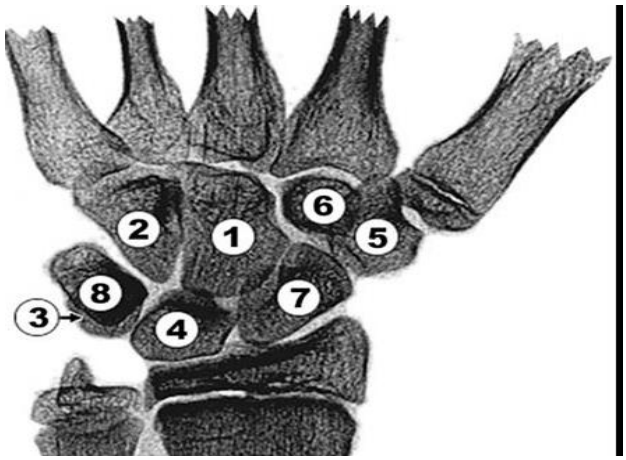

Figure 7: A representation of how each carpal appears in the order shown [10].

The standard arrangement is as follows: capitate (1), hamate (2), triquetral (3), lunate (4), trapezium (5), trapezoid (6), navicular or scaphoid (7), and pisiform (8), as seen in Fig. 7. Before the triquetrum and the pisiform, the distal epiphysis of the radius and ulna respectively ossify.

Ossification of the wrist, carpus, and hand is depicted in Fig. 8: Capitate (four months), Hamate (four months), Triquetral (three years), Lunate (four to five years), Trapezium (six years), Trapezoid (six years), Scaphoid (six years), Pisiform (eleven years), (I) Metacarpals, 10^{th} week of pregnancy; (J) Proximal phalanges, 11^{th} week of pregnancy; (K) Middle phalanges, 12^{th} week of pregnancy; (L) Distal phalanges,

9^{th} week of pregnancy; and (M) Middle phalanx of 5^{th} digit, 14^{th} week of pregnancy [14]. Although these divisions are arbitrary, decision was made to group pubertal stages together because osseous development more closely matches the level of sexual development than it does chronological age. The characteristics that distinguish these several stages of skeletal development are shown in schematic drawings that show how they appear in the posterior anterior roentgenograms of the hand and wrist from the data set. Features of the ossifications in the wrist, carpus, and hand areas are shown in Fig. 8.

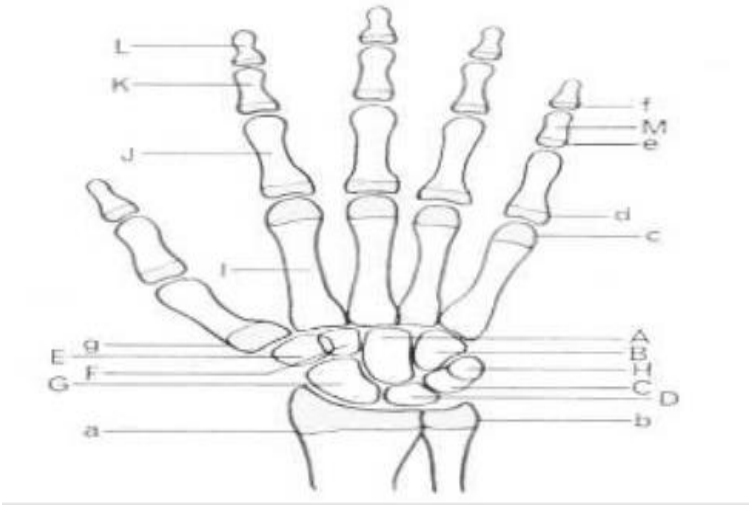

Figure 8: Ossification of the wrist, carpus, and hand [10].

2.2 Hand Bone Age Assessment Methods

Basic ideas of various age estimation approaches including the manual system of age estimation/interpretation method, hand bone image segmentation techniques, and automated ways to evaluate bone age using a digital hand atlas are reviewed in this section. The fundamental ideas behind these approaches have been examined and contrasted. The section aims to demonstrate to readers why present methodologies are insufficient to solve the hand bone segmentation issues because of a number of limitations and need for the development of novel methods for segmenting the hand bone and use the segmented hand bone in skeletal age scoring systems.

It has proven difficult to locate literatures that discuss the idea of age evaluation. It was attempted to find literatures in a variety of fields, such as anthropology, forensic medicine, and international regulations, as well as in a variety of thematic areas, including juvenile justice, migration, child labor, football players, athletics, trafficking, early marriage and recruitment into the armed forces, all of which depend on an official record of children's ages for the purpose of enforcing minimum age laws [10]. Attempts are made in this part to discuss some of the various envisioning modalities and age-estimating algorithms that have been reported in the literature.

Bone age assessment is a method that is regularly carried out in pediatric radiology. The evaluation is typically carried out by radiologically examining the left-hand's skeletal growth; age is measured, and then is compared with the chronological age. A difference between these two numbers unquestionably

indicates improper skeletal development. In addition to being often utilized in the management and diagnosis of endocrine problems, the method can also be used to gauge how well a medication is working [13]. It can show the pace of a patient's growth (whether speeding or slowing down). The results of this test are frequently used to determine whether to administer growth hormones to a patient. The social sector has a meaningful application as well. In order to boost their chances of being granted a residency visa, a significant percentage of asylum seekers in European nations assert that they are kids. Since these individuals frequently lack identification documents, skeletal maturity analysis is used to determine an individual's actual age. Due to its ease of use, low radiation exposure and accessibility of various ossification centers for the assessment of maturity, this examination is widely used. Automatic skeletal age evaluation may shorten the time spent reviewing the image and improve the accuracy of the analysis [15].

There are many indicators that show how humans develop chronologically, such as height and dental age. However, bone age measurement is frequently used because it is practical and reliable for identifying inherited diseases and endocrine disorders, and because it can indicate how well a treatment is working because it takes into account the non-dominant hand's different developmental stages. The benefits of this test, including its simplicity, low radiation dose, and accessibility to several ossification foci for the assessment of bone maturity, make it commonly utilized.

Even though Tanner and Whitehouse (TW2 or TW3) and the G&P methods are the two that are most frequently used, there is no set clinical procedure for determining bone age. Both approaches use X-Ray images obtained from the left hand, although they differ in a number of ways. The G&P method is the strategy that is most frequently utilized (by 76% of radiologists, particularly in the Netherlands), mostly because it is quicker and simpler to use than the TW2 or TW3 methods because it just compares the entire hand with a reference atlas. The G&P approach's main drawback is the variability of the analysis conducted by different observers with different levels of training, as studies have shown that inter-observer differences range from 0.07 to 1.25 years and intra-observer differences range in average from 0.11 to 0.89 years (roughly 10 and a half months). The typical case reading takes between 2 and 5 minutes, depending on the radiologists' clinical experience. The TW2/TW3 approach, which is particularly popular in the United States, differs from the G&P method in that it is based on a collection of bones whose standard maturity fluctuates depending on the population's age [16].

2.2.1 Clinical Methods

The G&P [17] and TW2 techniques are the primary clinical approaches for skeletal bone age assessment. The two approaches diverge in a number of ways. The G&P method is more expedient and user-friendly than the TW2 method. However, studies have shown that the two approaches yield different estimates of skeletal age, and that these variations matter in clinical settings. Bull claims that among the two, the TW2 approach is potentially more accurate and more reproducible, despite the fact that, in his opinion, it has never been proven to be more accurate. The left hand radiography is used in both techniques. A good understanding of the anatomy of the hand is required prior to doing tests utilizing these two techniques [15].

The traditional method for identifying developmental stages is manual analysis carried out by medical professionals. However, subjectivity-related variability and variations between and within observers will always exist in manual approaches. Estimates are not continuous, and it might be challenging to allocate around stage borders, therefore approaches based on stage assignment may contribute to this. The TW2 technique is therefore anticipated to be more trustworthy for estimating bone age because it is based on many stage estimates rather than just one. It hasn't been demonstrated to provide an improvement that is sufficiently significant in actual applications because it is more difficult and expensive to employ. The estimation of dental age is supported by similar arguments.

2.2.1.1 The Greulich and Pyle (G&P) Method

The left hand and wrist X-ray is compared using the G&P method to a standard atlas of radiographs, which is made up of samples of left hand and wrist X-ray images that have been divided into the corresponding age groups. The subject's gender matters because, as was already mentioned, females' bones mature at the appropriate stages earlier than those of males. As a result, the G&P method's atlas is divided into portions for men and women [18].

Preliminary research was initiated in 1929 at the Ohio's Western Reserve University School of Medicine. These investigations served as the foundation for a thorough examination into the growth and development of people. Numerous kids of various ages participated in the study. The left shoulder, elbow, hand, hip, and knee of these kids were radiographed. Every three months throughout the first postnatal year, an examination was done. Those aged twelve months to five years underwent examinations every six months, followed by yearly examinations. The investigation included the years 1931 to 1942 in total. Todd published an atlas titled "Atlas of Skeletal Maturation of the Hand" in 1937 [10]. The Todd atlas served as a foundation for the G&P atlas. They were able to utilize all of the radiographs gathered during the initial investigation because their atlas was first published in 1950. Between two and twenty-one hand radiographs taken at subsequent examinations of each of the 1000 youngsters were available to them in total. They offer a thorough explanation of the many maturation stages for each of these bones in their method. The descriptions are not instructions on how to rate a bone; rather, they are more of a general reference to the development of each hand bone. Most institutions employ a potentially less accurate, more quickly changed replica of the original. Below is a description of that variation. Using the atlas they have created, one can calculate the skeletal age using the modified G&P method. Because females develop more quickly than males, the patient's gender is among the most crucial details. The atlas is split into two sections, one for the patients who are male and one for the patients who are female. Standard radiography images of a child's left hand organized by chronological age are included in each section. The initial step in an analysis is to compare the provided radiograph with the atlas image that most closely matches the patient's chronological age. The next step is to compare it to nearby pictures of both younger and older kids. A doctor should use specific characteristics as maturity indications when comparing the radiograph to a picture in the atlas [19].

Based on the child's age, these characteristics change. When a child is younger, the presence or lack of specific carpal or epiphyseal ossification centers can give a doctor clue about the child's skeletal age.

Skeletal age can be accurately determined in older children by the configuration of the epiphyses and the degree of fusion with the metaphysis. After identifying the atlas image that most closely reflects the radiograph, the doctor should perform a more thorough inspection of the individual bones and epiphyses. The skeletal age is printed at the top of the page, which doctors can access after they are certain that the corresponding radiograph has been located [15].

2.2.1.2 Tanner and Whitehouse (TW2) Method

The TW2 technique is based on a group of bones that are typically mature for each age cohort rather than an age-based scale. Twenty ROIs located in the primary bones are taken into account by the TW2 approach when determining the bone age. Some of the interesting bones are shown in Fig. 9.

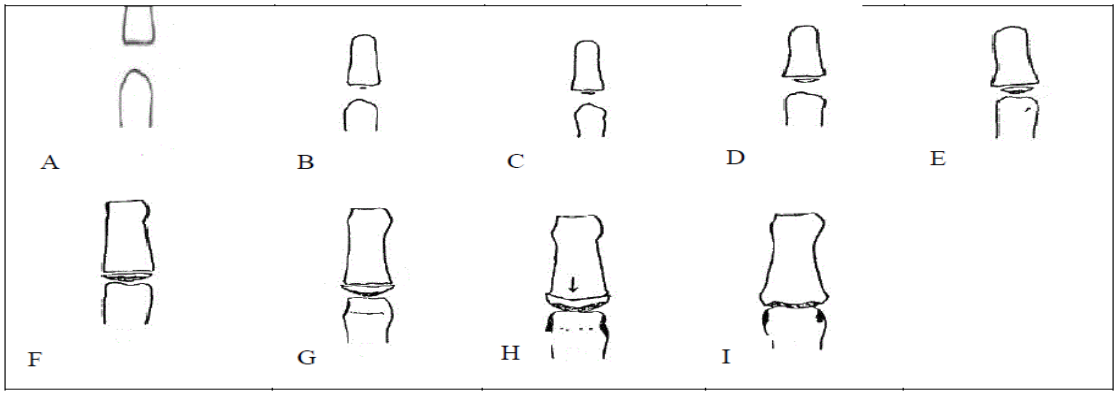

Figure 9: Stages of Phalanx bone in TW2 method [15].

The diaphysis, the metaphysis, and the epiphysis make up each ROI. It is feasible to spot these various ossification centers close to one another, particularly in young people. According to Fig. 9, each ROI's development is broken down into nine separate and unique stages, each of which is assigned an alphabetical letter (A to I).

Stage A – Absent

Stage B – Single deposit of calcium

Stage C – The center is distinct in appearance

Stage D – The maximum diameter is half or more the width of metaphysis

Stage E –The border of the epiphysis is concave

Stage F – Epiphysis is as wide as metaphysic

Stage G – Epiphysis caps the metaphysis

Stage H – Fusion of epiphysis and metaphysis has begun

Stage I – Epiphyseal fusion completed

Each stage of each bone is assigned a number score. An overall maturity score is created by averaging the scores of all ROIs. Using a functional relationship, the correlation between this score and bone age varies between males and females. Because of its modular design, the TW2 approach can be automated. Three scoring systems have been created for the TW2 method:

TW2 20 Bones: The TW2 20 Bones method consists of twenty bones, including the carpal bones, the first, third, and fifth finger bones.

RUS: the RUS method only takes into account the carpal bones, and

CARPAL: focuses solely on the carpal bones.

Although the TW2 method produces an estimate that is more accurate than the G&P method (the average inter observer spread for TW2 is 0.74 years (about 9 months) compared to 0.96 years (about 11 and a half months) for the G&P method [12]), it is less popular due to its high complexity, even though its modular structure makes it suitable for automation [16].

In general, the literature contains different algorithms for automatically determining the age of skeletal bones. The carpal ROI, ulna, and radius are not taken into account by the majority of these algorithms, which are based on the Epiphyses - Metaphases ROI (EMROI) extraction [15].

2.3 Related Literatures

The literature on age estimation was examined and addressed from an Ethiopian viewpoint on both a global and regional scale. The following search method was used to search three online electronic databases: Science Direct, PubMed, and Google Scholar. “Age Estimation Techniques” OR “BAA” were used as search terms to search databases. Additionally, the same search terms were used again, but this time with the addition of a few particular nations. Direct Google searches were also used occasionally when necessary.

A digital hand atlas and computer-aided diagnostic (CAD) system were developed by Cao et al. as part of their research to propose a system concept and preliminary implementation for Web-based bone age evaluation. The existing picture archiving and communication system (PACS) and recent advancements in internet technology are the foundations upon which the CAD system was constructed. A hand atlas database, a CAD module, and a Java-based Web user interface make up this system. The clinically normal hand photographs of many different ethnic groups that make up the digital atlas form a sizable new collection. Hand photographs, their quantitative attributes retrieved, and patient data are organized using a relational image database system. The G&P atlas is currently out-of-date, therefore the digital atlas eliminates such drawbacks while enabling computerization of the bone age evaluation. Users can interact with the database of hand images through browsers with the Java-based Web user interface. Users can upload a clinical hand image for use in determining bone age using a Web browser to push the image to the CAD server. To determine the bone age, quantitative features from the analyzed image are extracted, which reflect skeletal maturity. These features are then compared to patterns from the atlas database. An alternative to the conventional method for a quantitative, accurate, and cost-effective evaluation of bone age is such a digital atlas method, which is based on open system internet technology [13].

Using the G&P atlas, A. M. Zafar attempted to determine the accurate skeletal age (SA) in clinical and medico legal choices and to determine its applicability across various demographics in Karachi. The results pointed out that the G&P Atlas cannot be used to accurately assess SA in Pakistani children [20]. In another study, M. Mansourvar and R. G. Raj investigated the use of the BAA to assess children's growth status in order to identify hormonal imbalances and genetic diseases. Through a radiological examination of the hand-wrist skeletal region, skeletal maturity is assessed. They unveiled a brand-new method for BAA that makes use of a histogram-based comparison method. This method is implemented as a web-based system and makes use of a library of images as well as similarity tests based on content-based picture retrieval. The study also sought to address the shortcomings of earlier, frequently erroneous techniques for estimating human age. Up to the age of 18, the system can predict an individual's age from X-ray scans of their hands and wrists. The results of the system evaluation indicated that the method is reliable with error rates of 0.170625 years (about 2 months) compared with Bon expert system that has an error rate of between ±0.46 to ±0.37 years (about 4 and a half months) [21].

K.-H. Chiang created an atlas of skeletal maturation using information from kids born in the USA between 1917 and 1942, and it is commonly used for skeletal maturity assessment. The purpose of the study was to determine whether the G&P method is adequate for determining the skeletal age of children in Taiwan. The results showed that the bone age was delayed before puberty, increased during puberty, and then advanced towards the end of puberty. Additionally, in some age groups, there is a discrepancy of more than one year between the measured bone age and the chronological age. According to the study, the G&P atlas needs to be modified in some way in order to improve the accuracy, reliability, and consistency of skeletal maturation determination [22].

In a different study, Dvorak et al. evaluated the worth of football players who are arranged in age-related tournaments for men and women to ensure equitable opportunities within the game for all the various age groups. Other methods of determining age must be available in order to prevent participation in the wrong age group and because registration at birth is optional in some Asian and African nations. For many years, standard radiographs of the left wrist have been used to determine the skeletal age of an individual. But in a sporting environment, this is unethical. The study also planned to investigate if it could be able to estimate the age of healthy adolescent football players using MRI, which carries no radiation risk. The study discovered that among male teenagers between the ages of 14 and 19, MRI of the wrist offers an option as a non-invasive way of age assessment. The International Atomic Energy Agency's conventional radiographic grade no longer poses any risks, according to the authors, who claimed that the grading system they presented clearly detects the skeletal maturity by perfect fusion in all MRI slices [23].

On the basis of the integration of two systems, D. Giordano presented an automatic system for determining bone age in accordance with the clinical technique of TW2 in his work. The first system processes the bones of the fingers (EMROI - Epiphyses/Metaphyses ROI), while the second one processes the bones of the wrist (CROI - Carpal ROI). For males and females between the ages of 0 and 7, the technique guarantees an accurate assessment of bone age. Novel segmentation strategies were put forth for both approaches. Anatomical information of the hand and trigonometric principles were combined to extract the bones for the CROI analysis, while Gradient Vector Flow (GVF) Snakes and the derivative

difference of Gaussian (DrDoG) filter were combined to assign the TW2 stage. The Trapezium and Trapezoid, two of the largest wrist bones, are frequently united even in very young individuals, which presents one of the major challenges in determining bone age based on carpal bones. An extraordinarily efficient technique that separates the identified bones using a curvature function and evaluates the compactness of the bones identified solves this challenge. The use of geometric feature analysis and image processing methods based on difference of Gaussian (DoG) was suggested for EMROI analysis. A total of 106 X-ray images were used in the evaluation, and the system assigned the phases of the bones with a 90% success rate. It was reported in the paper that the technique allegedly outperforms other efficient approaches and is incredibly reliable. The mean error rate was approximately 0.37 years (roughly 4 and a half months), which is equivalent to the clinical reliability, where the error has been calculated to be 0.6 years (almost 7 months) [16].

In another study, Wavelet transforms were presented by V. Sridhar as a CAD system for BAA. The work involved measuring distinctive properties and performing bone edge recognition using wavelets. The BAA was carried out using these measurements. The Digital Hand Atlas of the Image Processing and Informatics lab provided the library of bone images used in this experiment. The study asserted that the wavelet-based CAD approach produces a ground-breaking and cost-effective technique for real-time automatic BAA to aid radiologists in the medical area [24].

In a different work, H. Y. Chai proposed an automated CAD BAA system that starts with a pre-processing step using anisotropic diffusion to reduce non-uniformity in the bone and soft tissue. The adaptive clustering algorithm (Hum Yan Chai1, 2013 #5) was then used to process the images [13]. The proposed algorithm was then applied to the system's output in order to recover any information that was lost and remove any information that was unnecessary. The epiphyseal is removed following the ROI's acquisition [25].

To aid radiologists in determining the bone age of pediatric patients, E. Pietka introduced a computer-aided categorization technique. The phalangeal region of interest (PROI) and the CROI, two areas of interest on Computed Radiography (CR) left hand wrist images, served as the basis for the classification. Preprocessing, feature extraction and classification were the three main components of the approach. A fuzzy classifier for both regions was created. Features were processed to produce a matrix that maps the collection of features to a year of age within the prescribed range after creating a membership function for each region. The membership function values in the interval were used to describe the membership grades. The matrix used to determine the age of the bones is processed by a classification method based on a max-sum operator. Due to the independent analysis of the two locations, two estimates of bone age were produced and the maturities of the carpal and phalangeal bones were assessed separately. The results showed that the difference between the two assessments in pathological instances reached up to two years [17].

Another study produced a digital hand atlas-based automated approach to determine children's bone ages. There are two parts to the hand atlas. The first part of the system is a database that contains 1,400 left-hand radiographs that were evenly distributed and digitized from children of Caucasian (CA), Asian (AS),

African American (AA), and Hispanic (HI) origin, aged 1- to 18 years old. The database also contains pertinent patient demographic information and pediatric radiologists' readings of each radiograph. The second element is a CAD module that determines a child's bone age from the data gathered. The seven ROIs used to create the CAD technique were the carpal bone ROI and six phalangeal PROIs. The three middle fingers' distal and middle regions are among the six PROIs. To determine a child's bone age, eleven category fuzzy classifiers were trained using these features: one for each race and gender, as well as one for females and one for males. A PACS is being connected with the digital hand atlas to validate its application in clinical settings [11].

Using the idea of object-based ROI, Jian Liu created new algorithms to enhance the reliability, accuracy, and usability of automatic bone age assessment (ABAA). The TW3 approach was used to assign seven carpal ROIs and thirteen RUS (radius, ulna, and short finger bones) ROIs. Particle swarm optimization (PSO) was used to extract five features from each ROI, including size, morphologic traits, and the fusional or neighboring stage. These features were then fed into Artificial Neural Network (ANN) classifiers. To process RUS and carpal features, respectively, feed-forward multilayer networks and back-propagation algorithm rules were used to train ANNs. A random sample of 1046 digital left hand-wrist radiographs were employed, half for training and the remaining half for ABAA following manual reading using the TW3 method [26].

For use in pediatric endocrinology to accurately estimate growth and pubertal maturation, J. Seok investigated bone age determination utilizing radiological scans of the left hands and wrists. The study suggested a completely automated feature extraction and machine learning classifier-based G&P atlas bone age estimation system. The paper's original contributions were (i) a modified GP atlas with photos evenly spaced at three-month intervals made using commercially available morphing tools, and (ii) a novel feature extractor based on Singular Value Decomposition (SVD) that produced a feature vector. Scale Invariant Feature Transform (SIFT) was utilized in the study to extract features from the images, and SVD was employed to decompose the feature vectors. Following that, the generated feature vectors were used to train the neural network classifier. The preliminary findings showed that even with a limited number of training data sets, positive outcomes are still possible [27].

Indicators of children's growth status, such as bone age development, have been demonstrated by C.-W. Hsieh as determining bone age is a tedious and heuristic task for pediatricians. Based on the investigation of geometric characteristics of carpal bones, a computerized approach for estimating bone age was established in the study. By computerized shape and area description, geometric features of the carpals were extracted and analyzed to determine the bone age of children. The classification of bone age was done using four classifiers: linear, closest neighbor, back-propagation neural network, and radial basis function neural network. To increase assorting accuracy, principal component and discriminant analyses were used. It was discovered that the bone area has a greater discriminating capacity to determine bone age. It was found that Taiwanese and the atlas of the GP technique have quite different trapezium and trapezoid bone ossification sequences. These findings also suggest that neural network classification can be used to accurately and practically assess the carpal bones [28].

N. A.-W. Abdullah investigated whether there was a variation in age estimation between clinical and radiographic exams. The study's secondary objective was to compare the average age of the ossification centers and epiphyseal union in the bones of Yemeni subjects to earlier studies carried out by both eastern and western specialists [4].

R. Camerierea's study sought to determine adult age using techniques based on alterations to the human skeleton that occur with advancing years. An innovative technique for the automatic calculation of pulp and tooth area was used in the study to analyze pictures from Periapical X-rays and determine the dental age of participants. A sample of 70 canines, 40 of which were male and 30 of which were female, representing subjects ranging in age from 20 to 70 (age in human years), were used to evaluate the algorithm [29].

E. Hillewig prospectively examined and contrasted plain radiography with 3T MRI of the clavicle to determine the forensic bone age. The study came to the conclusion that 3T MRI allows for a more precise diagnosis of bone age than plain radiography by providing high resolution, cross-sectional pictures of the maturation of the clavicle without ionizing radiation in a noticeably short amount of time [30].

A. M. Mughal investigated many methods of estimating bone age and showed that each one yields a different set of data and has a different range of applicability for various ethnic groups [31]. Another study examined the texture of people's fingerprints to ascertain their age and gender, as well as the relationship between RTVTR (Ridge Thickness to Valley Thickness Ratio) and ridge count on gender detection. The purpose of the study was to evaluate how well human age and gender might be ascertained using physical biometrics (thumbprint). A system was developed to collect the fingerprints of the sampled population using a fingerprint scanner device connected to the computer system through the Universal Serial Bus (USB), store the fingerprints in a Microsoft SQL Server database, and train the stored fingerprints using a back propagation neural network. The study's particular goals were to: construct a model and fingerprint-based identification system to identify individuals' age and gender; use fingerprint sensors to gather various individual fingerprints together with their age and gender; and evaluate the produced system. A fingerprint is a pattern of interlaced ridges and valleys that serves as a depiction of the epidermis of a finger [32].

Biological age (BA) estimation from radiologic data was investigated by D. Stern as a crucial area in clinical medicine. A study proposed the use of deep convolutional neural networks (DCNN) for automatic BA estimation from hand MRI volumes, and that encouraged radiologists to visually estimate age using established staging schemes that correspond to physical maturation [33].

Large-scale studies have been conducted in the past to establish a Chinese bone age reference. S.-Y. Zhang in one of his studies pointed out that because of the possibility for variation among raters around the world, it has been challenging to compare the maturity of Chinese children with populations abroad. To establish a Chinese bone age reference and compare the pace of maturation in the Chinese population with that of other populations, the study used an automated bone age rating system to analyze the radiographs from a large survey of healthy Chinese youngsters. In five Chinese cities, X-rays from 2883 boys and 3143 girls between the ages of 2 and 20 were used, and the Bone Xpert automated approach was

employed to evaluate them. The results showed that Chinese children matured at the same age as previously examined Asian children from Los Angeles, but 0.6 years (about 7 months) before Caucasian children in Los Angeles. A new bone age scale called BX-China05 was developed by adapting the G&P bone age method to the Chinese population. Boys aged 8 to 14 had a 1.01-year standard deviation and females aged 7 to 12 had a 1.08-year standard variation between BX-China05 and chronological age. The study therefore came to the conclusion that the automated technique provided a stable and effective standard for determining bone age in China by dropping more variability [34].

In order to compare his method with automated and manual bone age assessment methods based on G&P, Ian Pan created a deep learning approach to bone age evaluation using a training set of developmentally normal pediatric hand radiographs [35].

By utilizing various object identification techniques to identify and segment bones that are anatomically significant for the age evaluation and then using these segmented bones to train deep learning models to predict bone age, Erik Westerberg sought to enhance and accelerate bone age assessments in his research. The study employed a dataset of 12811 X-ray hand scans of individuals, spanning in age from infancy to 19 years old [9].

C. Spampinato et al. suggested and tested a number of deep learning methods to determine the age of skeletal bones automatically; the outcomes revealed an average difference between manual and automatic evaluation of roughly 0.8 years. Additionally, this is the first automated skeletal bone age assessment work that has been tested on a public dataset, for all age ranges, races, and genders, and for which the source code is available, serving as a thorough baseline for further research in the area. Along with the specific application scenario, the paper aimed to address broader issues surrounding deep learning on medical images, such as how to train a CNN with a small number of images, how to compare deep-learned features to manually created ones, and how to use deep learning techniques trained on general imagery for medical problems [36].

The RSNA Informatics Committee and volunteers created, coordinated, implemented, and evaluated the challenges, as Safwan S. et al. explained in their article. The data set for this challenge included pediatric radiographs and assessments of the related bone ages from multiple expert assessors. One of the main focuses of the RSNA is to advance superior patient care and healthcare delivery via technological innovation, research, and teaching. The RSNA mission's research and technological innovation goals are aligned with machine learning (ML) competitions [37].

According to L. Eikvil, determining the age of young asylum seekers presents a problem for immigration authorities in a number of nations. In these situations, candidates may not be aware of their actual date of birth, may be traveling without valid identification, or may be carrying documents with questionable origins. Occasionally, it is believed that asylum seekers are making up their age in order to increase their chances of getting a residence permit or other benefits. Children who have been the victims of human trafficking may benefit from age assessment techniques that are safer and more accurate. It is thought that automatic image-based techniques can aid in resolving these issues. There are still not many of these solutions, therefore further research is required. However, a few systems that perform this analysis have

been developed, all of which are designed to analyze the age of the bones in the hand and wrist [1]. Another study used periapical X-ray pictures of the tooth to automatically calculate the pulp and tooth area in order to determine dental age (DA). Then Cameriere's formula [38] is used to determine the dental age.

Table 3 summarizes the major techniques that are reported in the literature for use in bone age assessment.

Table 3: Summary of techniques use in bone age assessment.

Method	Bone Used	Atlas Type	Advantages	Disadvantages
Greulich and Pyle Method (G&P)	All the finger joint bones and carpal bones	Hand Atlas	By simply comparing the X-ray with the atlas, a straightforward and reliable procedure can be used; Once the age group has been identified, the observer must look for older and younger stages to determine the exact age.	Taking a long time; 22 joints need to be treated; To estimate the age, the photograph must be quite clear; Could result in misclassification; A professional observer is required.
Tanner and Whitehouse Method (TW)	Bone joints including carpal, thump, middle and last finger	Hand Atlas	It is based on the maturity of the skeleton's bones, unlike G&P; RUS, Carpal, or Phalangeal bones (20) individually can be used to estimate age based on numerical scores assigned for each bone (rather than merely shape); More precise than the G&P approach.	Taking a long time; Complex technique (each bone's maturity score is 8); in this case, the image must also be sufficiently clear for categorization; A professional observer is required.
FELS Method	All the bone joints	Digital Atlas	An approach to age assessment using the maturity scores for all joints bones; Easy to understand and implement as a program.	For each bone, too many points (more than 130) are chosen; Researchers are unable to test and verify since the software/package is unavailable.
Automated Methods	Carpal alone, Phalangeal bones alone, and both carpal and Phalangeal	Digital Atlas	Accessible online as a free or licensed product for testing; The methods of image processing are frequently employed to make the image clear; It takes less time; Enables precise measurement; The user does not require any prior information.	Still require improvement for improved outcomes; The observer must be proficient with computers.
Manual Methods	Skull, Pelvis, rib, Spine, Knees, femur, Hand	Hand Atlas	Simple to use and validate; Age evaluation is not limited to teens; Some of these bones can help determine gender.	Taking a long time; Measurement is depending on lab equipment; An expert must be the observer; Works only on the deceased.

Chapter 3: Proposed System for the Hand Bone Age Assessment

3.1 Imaging Modalities for Hand Bone Age Estimation

X-rays and, in some situations, CT have typically been utilized to obtain the images required for the various analyses. The fact that X-ray and CT imaging involve the use of ionizing radiation is a drawback of both procedures. The possibility of radiation injury is a defense against using radiation, especially for purposes other than diagnosis or treatment. In pediatric radiology, radiation exposure is a particular worry. Being accessible and affordable, ultrasound offers an edge over other techniques. Portable systems make it simple to apply as well. MRI is a different radiation-free method, although the equipment is more expensive, larger, and less accessible [1].

Numerous studies have demonstrated that not every change can be observed using every modality and that frequently it is necessary to designate modality-specific growth stages. Different modalities view the phenomenon differently, even in situations where the same stages are visible. Because of this, the development stage as perceived by one modality may seem advanced when perceived by a different modality. Today's imaging techniques, including X-ray, CT, MRI, ultrasound, and optical imaging, may successfully image anatomical structures. Furthermore, a number of MRI derivative protocols, such as diffusion-tensor MRI and functional magnetic resonance imaging (fMRI), nuclear medicine (single photon emission computed tomography, or SPECT, and positron emission tomography, or PET), ultrasound, optical fluorescence, and others can provide details about physiological structures with regard to metabolism and/or functions. For a pathological investigation to be successful, the right medical imaging modality must be chosen in order to get the desired information. For instance, if knowledge of the cardiac volumes and functions connected to a beating heart is needed, one must ascertain the specifications and constraints regarding the spatial and temporal resolution for the target set of images. It's crucial to remember the kind of pathology that the imaging test is looking for. A particular medical imaging modality, such as PET, or a combination of several modalities, such as stress-PET and ultrasound, can be chosen depending on the inquiry, such as metabolism of heart walls or measures of the opening and closing of the mitral valve.

An energy source is necessary for the fundamental process of picture generation in order to gather data about the thing that will be represented by an image. To learn about tissue's or organ's distinctive properties, radiation of some kind, such as optical light, X-ray, gamma-ray, RF, or acoustic waves should interact with the target. The energy source may be internal (SPECT, PET), external (X-ray radiography, mammography, CT, ultrasound), or a combination of both internal and external (MRI), in which case the proton nuclei present in body tissue produce electromagnetic RF energy-based signals in the presence of an external magnetic field and a resonating RF energy source. The term "functional imaging" is frequently used to refer to contemporary techniques that use functional MRI, PET, or SPECT to image physiological processes. A comparison of various imaging techniques for age assessment is shown in Table 4 below.

Table 4: Comparison of the different imaging modalities for age estimation.

	Age range	X-ray	CT	MR	US
Dental stages	→ ~19	X			
Hand/wrist	→ ~18	X		X	X
Clavicle	→ ~25	X	X	X	X
First rib	→ ~25	X	X		
Cervical vertebrae	→ ~18	X			
Iliac crest	→ ~18				X
Knee	→ ~18			X	

3.2 Steps in Medical Image Analysis

It takes more than one step to process an image. To extract the relevant information from the seen picture, several processes must be carried out sequentially. A hierarchy in the processing steps will be evident, for example, enhancement will come before restoration, which will come before analysis. These are frequently completed in order, but more complex tasks call for feedback, i.e., advanced processing steps pass parameters back to earlier steps so that the processing includes a number of iterative loops.

Data acquisition, signal conditioning, feature extraction, and decision-making are the four main stages of processing imaging signals in general. This is also illustrated in Fig. 10. The fields of computer sciences, physics, medicine, biology, and engineering all intersect in the discipline of medical image analysis, which is very interdisciplinary. Biomedical image analysis, in its most basic sense, is the application of image processing methods to biological or medical issues. However, a number of other disciplines, including anatomy, physiology, and instrumentation are important in biomedical image analysis. In addition, a number of computer science subfields, most notably artificial intelligence (AI) and computer modeling, exist on top of image processing and are relevant to biological image analysis. Artificial neural networks, fuzzy logic, evolutionary computing, and computer learning are some of the AI techniques that are used in biomedical picture processing. The depiction of time-course dynamics and sophisticated segmentation methods both heavily rely on computer models.

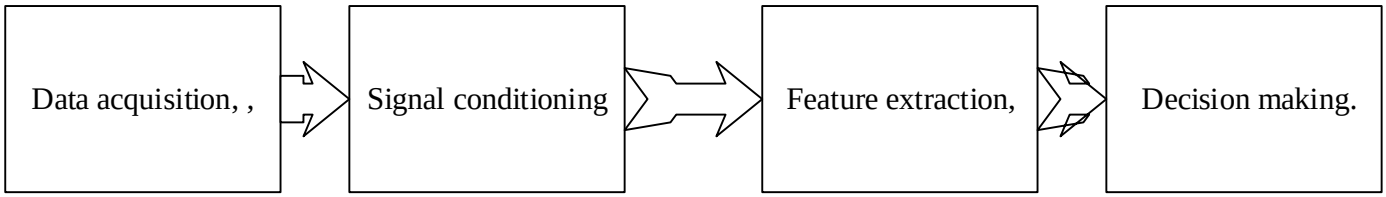

Figure 10: Basic medical image analysis flow diagram.

Processing X-rays of the human hand to identify age is the main goal of the current thesis work. To complete the assignment, images must be preprocessed, backgrounds must be removed, orientations must be corrected, images must be segmented, and training must be done to determine the clients' ages. Michael created the initial semi-automated method sometime in the late 1980s. The technique, according to the author, can automatically segment the bones in a hand radiograph, but extensive testing was not

carried out. In order to make the subsequent segmentation phase more reliable, the image must first undergo preprocessing to normalize the gray-scale. The approach initially uses a threshold procedure to separate the full hand—bones and flesh—from the backdrop. Next, the hand's bones are located using a model-based methodology. This technique makes use of information about the relative placements of the hand's bones in relation to one another and to the shape of the hand. The CAD systems that have been suggested in various works of literature are built using methods from the fields of AI, image processing, and computer vision. The primary steps in a typical CAD scheme are image preprocessing, segmentation, feature analysis (extraction, selection, and validation), and classification, which is used to either describe or minimize false positives (FPs) [39, 40, 41, 42].

3.2.1 Preprocessing

This step emphasizes the undesired aspects of the image while downplaying the subtle features of interest. As a result of the enhancement process, the objects of interest are better described, increasing the sensitivity of the detection system and enabling improved characterization in the case of diagnosis. In this stage, the contrast of the ROIs is improved, the borders of the abnormalities are sharpened, and noise is suppressed. There are several of these strategies that have been published in the literature, including those that rely on traditional image processing methods, region-based algorithms, and methodologies that call for changing the original image into a different feature space. To account for the specificity of distinct picture areas, either global processing or local altering enhancing parameters can be utilized [43]. One such technique used to improve images is histogram equalization. When comparing many photos that might have been taken in a variety of situations and attempting to standardize histograms, it is used. The goal of histogram equalization is to increase the amount of information that can be given through an image, and the modified image should appear better as a result [44].

3.2.2 Image Enhancement

Image enhancement mainly targets processing of image in a way that makes it easier for us to examine and evaluate the visual information it contains. It is arbitrary since it much depends on the precise data the user is trying to derive from the picture. The main prerequisite for image improvement is that the image must already include the information that we wish to extract, accentuate, or restore. Fundamentally, nothing can be created from nothing, and the intended information must not be completely obscured by background noise in the image. To put it simply, the processed image should be better suited than the original one for the necessary task or purpose. This is the truest and broad statement we can make regarding the purpose of image enhancement. Because of this, evaluating image enhancement is by its very nature subjective, making it challenging to evaluate its effectiveness outside of its particular field of application [41]. Some of the image enhancement techniques include filtering [45], histogram stretching and histogram equalization, to mention a few. Another important preprocessing technique is image restoration. The main difference between restoration and enhancement is that the former is more objective and requires a good understanding of the image degradation process. On the other hand, image enhancement is a subjective process, and the results depend on the observer who is doing the assessment.

3.3. Image Segmentation

The fundamental purpose of segmentation is to divide the image into parts that are mutually exclusive so that we can later apply meaningful labels to those sections. The image's background is typically referred to as everything outside of the segmented items as the foreground. We are unable to identify a single, 'correct' segmentation for any given image. Rather, the types of objects or regions that we are interested in identifying have a significant impact on the proper segmentation of the image.

The two main categories of segmentation techniques are supervised and unsupervised. The class membership probabilities, the histogram models, and their parameters that describe the distribution of the measured features are estimated in supervised approaches using a sample of labeled images (i.e., images for which the pixel class memberships are known). As compared to the supervised approach, unsupervised pixel analysis does not rely on prior knowledge of class membership [41].

Only ad hoc procedures, whose performance is frequently assessed indirectly based on the performance of the broader system to which they belong, exist as an underlying theory for picture segmentation. Despite having the same end in mind, image segmentation methods can differ significantly depending on the type of image (such as binary, gray-scale or color), the mathematical framework chosen (such as morphology, image statistics or graph theory), the features used (such as intensity, color, texture and motion), and the approach used (such as top-down, bottom-up or graph-based). In the meantime, there is no widely used taxonomy for categorizing image segmentation algorithms [48].

The object of interest must be distinct from background in some way, such as by a difference in picture intensity, by a delineating boundary, or by a difference in texture, in order to successfully execute image segmentation. A known shape or other a priori knowledge can occasionally aid with segmentation. Either a mask or an outline may be the end result of this method. In a mask, one pixel value (typically 1) represents an object pixel and another pixel value (typically 0) represents the background. A polygonal approximation of an object's shape is one example of a parametric curve or group of curves that can be used as an outline. The complexity of segmentation methodologies seems to have no upper bound, and there is unquestionably a tendency toward more complicated segmentation techniques being more application-specific.

Generally, complete autonomous segmentation is one of the most challenging issues in the construction of computer vision systems and is still a hot topic in machine learning and image processing. Because segmentation is frequently the crucial first step that must be successfully completed before subsequent tasks, such as feature extraction, classification, and description, segmentation plays a particularly important role in image processing. Segmentation has many different purposes, including extracting ROIs, detecting edges, segmenting regions, delineating specific objects, and segmenting textures [45].

3.4 Automatic Skeletal Bone Age Assessment methods

Manual bone age calculations do contain some interpretation variability. This causes issues with its clinical application, subject comparisons, and patient or customer follow-up. Theoretically, a computerized, automatic system for determining bone age would be an answer, but in practice, it is exceedingly challenging to develop an automated system that could reliably assess the changes in size,

form, and mineralization in various ossification centers in the left hand and wrist bones. Over the past thirty years or so, bone age from wrist radiographs has been computerized. Digital radiography or a scanner's digitalization followed by multiple processes is how radiographs are produced. The backdrop is eliminated, the image's orientation is adjusted, and the image is usually normalized to gray-scale (if the original is color) during pre-processing so that significant portions of the image can later be retrieved. Desired sections of the bones and soft tissue are isolated from the background in the following phase, segmentation. The image is then examined by taking into consideration specific areas of interest in order to estimate bone age using the G&P Atlas or to calculate bone age using the TW method. Older automated image processing techniques that identify ossification characteristics in hand bones demonstrated a substantial difference from bone age determined manually. However, a more recent technique for determining bone age, known as Bone Expert, has been developed. It reconstructs the edges of 15 bones in hand radiographs that are of interest and uses this data to determine bone age using both the G&P and TW techniques. The algorithm has been verified for use with several racial groups [31].

By creating a mapping from input to output, machine learning (ML) techniques offer a mechanism to automatically carry out image-based conclusions. One advantage of the ML is the opportunity for computerized and automated processing as well as quantitative measures of similarity. In therapeutic settings, little human interaction lowers labor costs and increases productivity. A technique for estimating bone age based on ML, Bone expert was first introduced in 2009.

In a variety of medical information processing and analysis sectors, deep learning techniques are finding more and more applications. The subject of medical image analysis has been a good fit for utilizing deep learning methods' image-related cognitive and inferential efficacy. A model is typically created that converts an input into the intended kind of output. The X-ray image serves as the input and the estimated age serves as the output for the problem of estimating the age of hand bones. It is necessary to establish the right weighting for the edges and place the hand bones at various edge nodes and edges. Through the process of training, a set of numerical values characterizing the relationship between input and output are formed.

In order to map an input image to the appropriate type of output, it must pass through a series of processing steps. Traditional techniques to ML often started with a collection of feature representations that were established by humans, which may not always be the best choice for the problems being addressed. From the standpoint of the training procedure, deep learning offers researchers and developers the advantage of end-to-end training, where the neural network is brought to a configuration that is reasonably close to optimal with little human intervention. High-learning capacity neural networks are typically used in deep learning and an array of processing units are tiled together to form a layer. To produce output, the training process sends inputs across the network. The difference, or error signal, is then compared to the desired output that was intended to be produced. In order to correct the problem, edge weights connecting the network's individual nodes are backward propagated. The deep learning process raises some problems. One is the extreme disparity between the quantity of edges and the usual amount of data that can be used during training. Because of this, the model will become over fitted,

becoming highly tailored to the details of the training data while failing to generalize effectively to data instances that are not included in the training dataset.

In order to address the imbalance and over-fitting problem, some strategies, including the dropout layer, are frequently used. Normally, a dropout layer is positioned between two layers that are fully connected. It helps prevent the coadaptation phenomenon, which is known to happen when nodes and weights are trained to comparable configurations when the dropout layer is not in use. It does this by dynamically switching on and off the edges bridging the two layers. Throughout the training phase, it serves as a utility layer, and throughout the test and deployment phases, it is eliminated [49].

3.4.1 Deep Learning Methods

The family of AI techniques includes deep learning. It takes its cues from the capabilities and structure of the neuron in a living cell. ANNs are so-called because they take an input, analyze it, and output the result. ANN is the foundation of deep learning. Massive volumes of data are what power deep learning. Data analysis keeps getting more difficult as vast amounts of data are produced. Traditional ML algorithms can't effectively process large amounts of data, therefore deep learning addresses this issue. Deep learning can be used to analyze any kind of data, including text, images, and more. Deep Belief Networks (DBN) and CNN are the two deep learning algorithms that are most frequently employed and ideal for image data. In various domains, such as medicine, where medical images can be used to detect a tumor and identify its type, or to assist a robot in navigating by recognizing obstacles, computer vision analysis employing CNN delivers a variety of use cases such as detection and recognition from the images [50, 51]. When used to determining bone age, Fig. 11 depicts the general architecture of a deep learning technique.

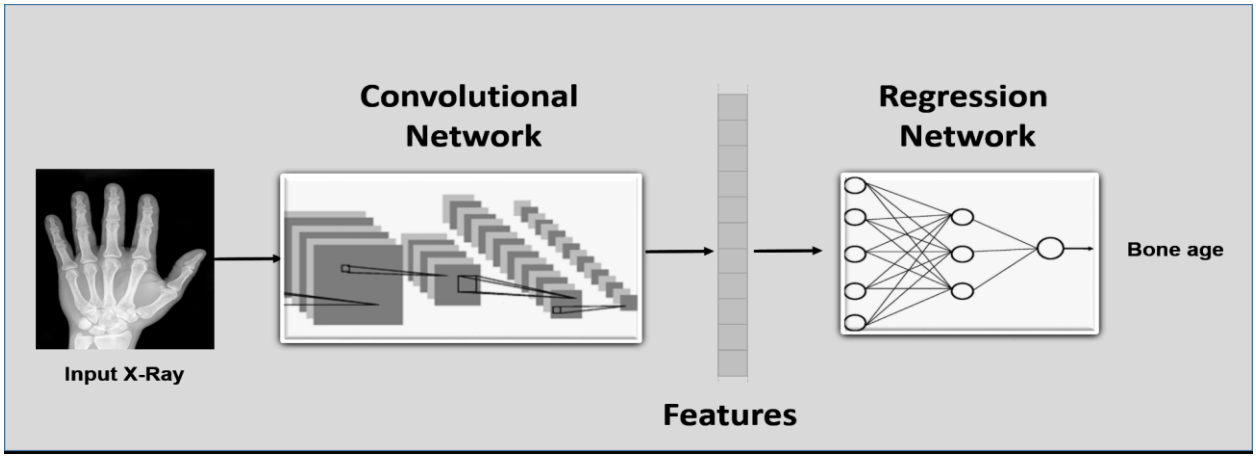

Figure 11: General architecture of deep learning methods for bone age assessment.

Deep learning uses deep neural networks with more layers and a lot more data than typical ML techniques; as a result, it requires larger models and more processing power. The performance of deep learning algorithms is directly correlated with the quantity and variety of data, but the performance of classical ML algorithms can't be improved beyond a certain point even if the amount of data is increased [27].

3.5 The Proposed Automatic Skeletal Bone Age Assessment Scheme

This thesis proposes a framework for an automated edge-based segmentation-based age determination. The proposed segmentation framework is made up of three key parts: watershed segmentation for clear visualization, edge detection, and picture preprocessing. The intensity inhomogeneity correction, background noise removal, and removal of tissue structures not required for age calculation, such as flush and fat from left hand pictures, are all done in the preprocessing step. The second part of the framework performs image enhancement and segmentation of the X-ray hand images collected from hospital data sources in addition to the reference images from the web used in this thesis. In the third phase, watershed segmentation is used to each image to identify the Phalanx, Carpal, Radius, and Ulna planes of the hand and to compare them to the left hand's reference images. The left-hand X-ray scans of participants aged 0 (1 month) to 19 are the starting point for the suggested computational model.

The proposed method begins with a preprocessing stage for background identification. This is done utilizing both local and global statistics. Following this initial step, the first, second, third, fourth, and fifth fingers are extracted along with the carpal appearances, ulna, and radius fusions. The same method is employed to take advantage of the morphological variations between each finger, including the thumb and carpal. The look and ossification state of each finger are examined. Then using a derivative operator, the algorithm identifies the bone borders, used to extract in our case the four Epiphysis ROIs (EROI). On each of the ROIs, a set of filters are applied followed by a threshold step. The system then finds the four largest areas (epiphysis distal phalanx, epiphysis middle phalanx, and epiphysis proximal phalanx and epiphysis metacarpals), the 8 carpals and Ulna and radius. To create a geometric representation of the 30 bones, including the 15 phalanges, the distal ends of the ulna and radius, eight carpals, and five metacarpals, a convex hull is applied to each of these shapes. Features are taken from each convex hull that can be compared to the features taken from the online photos. As a result, all of the retrieved EMROIs, including the carpal, ulna, and radius, are staged, and a final stage is calculated as the average of all the stages. The processing algorithm's functional diagram is shown in Fig. 12, and the overall age determination algorithm's schematic is shown in Fig. 13.

Automatic age detection and segmentation of left-hand wrist X-ray pictures provide a way to avoid the tediousness of manually segmenting and interpreting various data sets. Reproducibility, which is impacted by both inter- and intra-observer variability, is another promise made. Automated systems, however, face considerable obstacles in reaching these goals. Due to the nature of left-hand wrist photos and the presence of a sizable number of artifacts and intensity inhomogeneity in our situation, the main issues are that pixel intensities contradict the independent and identically distributed assumption within and between images. Automation must therefore take these issues into account in order to produce reliable segmentation results, and research is still ongoing to develop such widely used automated techniques.

One of the algorithms that is considered state-of-the-art in the literature included image preprocessing, background removal, edge extraction/detection, watershed-based image segmentation and features analysis as a backbone. The segmentation of the wrist bone is a crucial step in the automatic determination of skeletal age. The technique consists of two steps. The original wrist bone picture is first preprocessed using methods like a median filter and an adaptive histogram modification strategy. The wrist-bone image

is next subjected to additional filters and watershed technique segmentation. The fundamental idea behind the watershed approach is to add a force that takes into account the local information in order to increase resilience and more effectively distinguish homogeneous zones. All of the wrist bones' X-ray pictures are used in experiments. The experimental findings demonstrated that accurate and reliable segmentation may be obtained by including region qualitative information into the level set approach. A method that exclusively works on background pixels—defined as pixels that meet certain criteria discovered through a study of the image histogram statistics—is employed for background separation. Low gray level and low contrast are typical characteristics of background pixels [2].

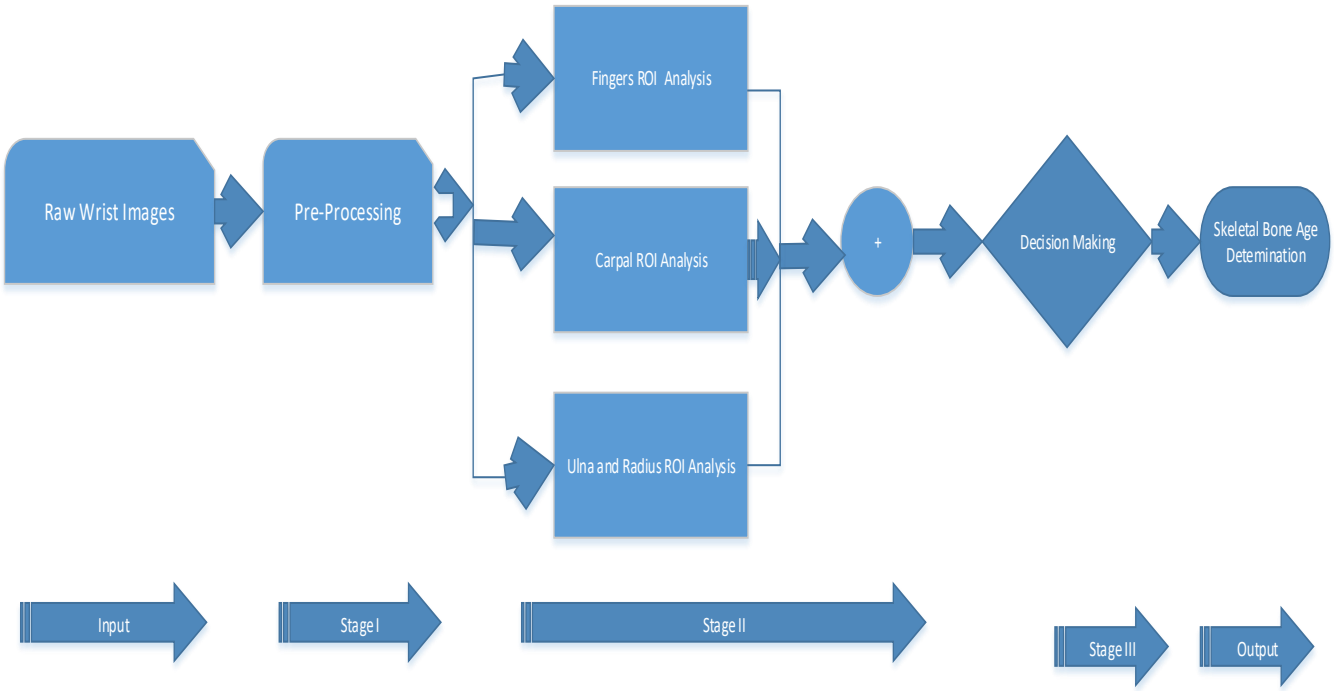

Figure 12: Functional diagram of the proposed algorithm.

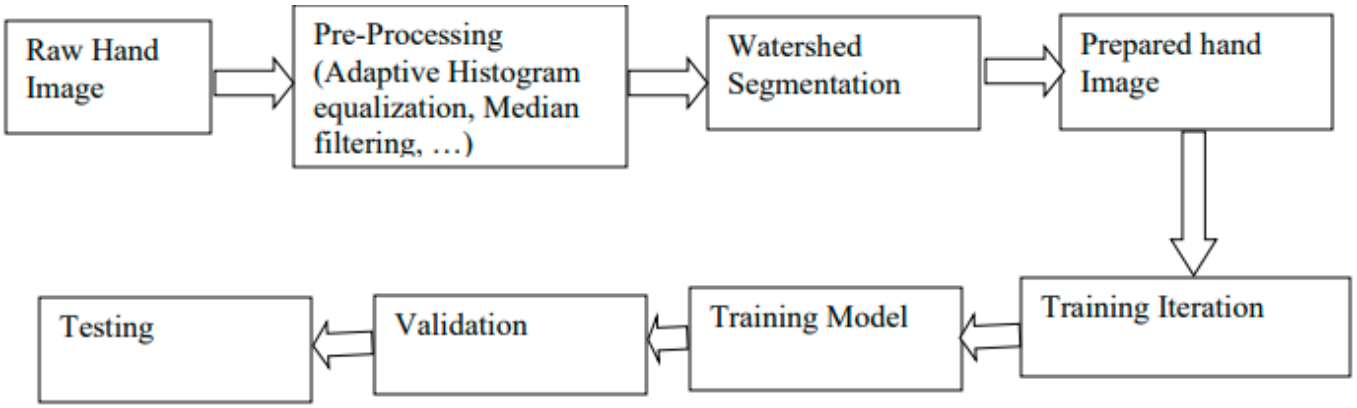

Figure 13: Overall proposed age estimation framework.

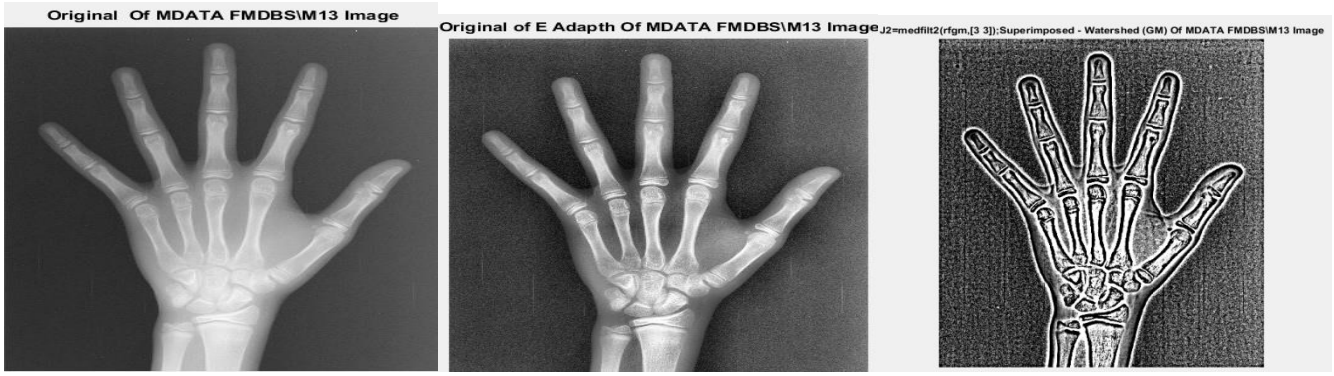

Figure 14: Raw image (left), adaptive histogram (middle) and application of different filters and watershed segmentation outputs (right).

The algorithm developed in the current study followed a similar path. A schematic of the algorithm is presented in Fig. 13. It involves adaptive histogram equalization, filtering as well as image enhancement steps. The rest of the algorithm involves ROI selection (EMROI, CROI, UROI, and PMROI) as well as watershed segmentation. Figure 14 presents typical outputs following the different enhancement steps.

3.5.1 ROI Appointments

In the current investigation, two independent bone age assessment systems were used: the PMRUS system, which consists of 21 bones (including the metacarpal, radius, ulna, and short finger bones), and the Carpal system, which consists of eight carpal bones. According to some claims, the carpal bones provide different information regarding the look and maturation process than the finger bones do, and the distinction between RUS and carpal bone assessments may be important for differentiating between diseases. According to these bones, 29 ROIs were created, 21 of which belonged to the PMRUS system and the remaining 8 to the Carpal system. Radius, ulna, metacarpals I, II, III, IV, V; proximal phalanges I, II, III, IV, V; middle phalanges II, III, IV, V; and distal phalanges I, II, III, IV, V were the PMRUS ROIs that were labeled from 1 to 21, in that order. Except for the middle phalanx that is absent from the thumb, each of the phalanges (proximal, middle, and distal) has five portions.

Let’s assume that each ROI consists of two components, the target region and the reference area, both of which are comprised of images. For our chosen age groups, the Phalangeal/Epiphysis, Metacarpals, Radius, and Ulna were designated as target areas, while the other bones served as reference frames. The proposed age estimation procedure is based on the maturity of the bone standard. When determining bone maturity using the TW method, bone joint locations were used as ROIs. As illustrated in Fig. 15, the epiphysis, diaphysis, and metaphysis of each ROI are further separated into three sections. Among these three, the epiphysis ossifies from age zero to teen age and later gets joined with diaphysis. So, the age assessment of the TW and G&P methods is only up to 19 years. These techniques have been modified by several authors. One of them is TW2 (20 bones), a modified version of the original TW1 method that was released in 1975. For bone analysis, 20 ROIs are used, comprising the carpal bones, first, middle, and fifth fingers. The second approach, RUS (Radius, Ulna and Short bones), is similar to TW2 but leaves out the carpal bones. The third one, called CARPAL, only makes use of the carpal bones. Until the age of 9, these bones ossify. Therefore, the age range for this method of age assessment is 9 years.

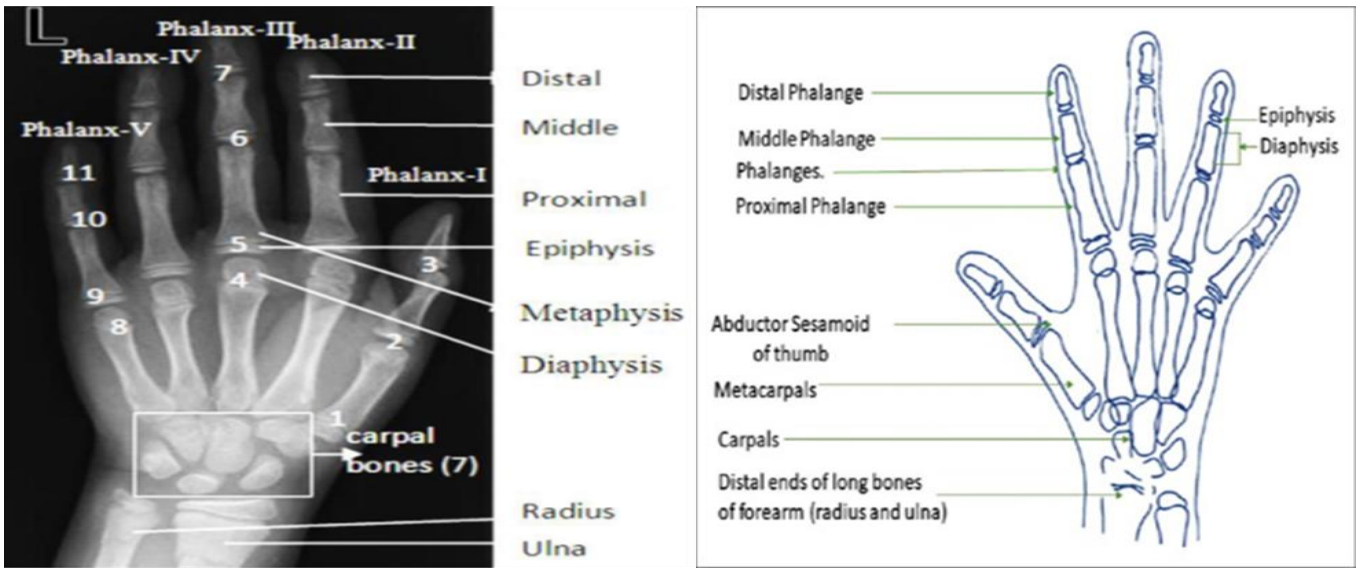

Figure 15: Left hand radiograph with URROIs, EMROIs and CROIs [46].

The foundational approaches for determining human age are the G&P and TW2 methods, along with First Atlas and look-up tables. The hand atlas must be used to locate and physically check the epiphysis-ROI or carpal-ROI index stage in the radiological image. Many researchers attempted to create fully automated systems for (digitally) confirming and verifying their results using these two techniques [47]. On the two regions separately, image-based feature extraction is carried out. Epiphyseal and metaphyseal measurements are delivered from the PROI, which contains both. The capitate, hamate, triquetrum, lunate, scaphoid, trapezoid, pisiform, and trapezium are the carpal bones that make up the CROI, which are distinguished and classified in terms of features. A bone age assessment is applied to all automatically retrieved features.

3.5.2 PROI (Proximal, Middle, Distal Phalanges) Analysis System

The suggested approach is grounded in clinical findings that specific regions are susceptible to skeletal maturation. The epiphyses' size and shape do, in fact, represent the stage of development they have reached before they reach the margin of the metaphysis. As a person ages, the two bones become more fused together (also see Fig. 16 and Fig. 17). The thumb, second, third, fourth, and fifth fingers are localized and extracted as the first step in the PMROI analysis technique. The suggested algorithm extracts each finger because of the variations in anatomy. The gray level profiles of the identified fingers are then taken into consideration to extract the EMROIs of the proposed approach. Finally, PMROI enhancement using various filters and stage assignment for the extraction of geometrical features are carried out.

PROI Extraction: The first, second, third, fourth, and fifth fingers are retrieved with a high degree of accuracy because they are simple to distinguish and because the suggested method takes advantage of their anatomical characteristics well. The metaphysis, epiphysis, and diaphysis, which are employed to extract the PROIs, have all been considered while doing the PROIs extraction, and these borders also dictate the bone edges.

PROI Enhancement: After the PROIs have been extracted, a second filter—referred to as the 'medium' filter in this thesis—is used to make it easier to extract the features. We can recognize the soft tissue (which normally appears as a smoothed area) using the medium filter. It employs a Laplacian of the Gaussian (LoG) filter with an appropriate standard deviation (0.35 in our case) and a Gaussian function (smoothing filter) with an appropriate standard deviation (0.02 in our case). One can create a new image with improved contours by subtracting the two output images produced by these two filters, while uniform areas settle to an almost uniform gray value.

PROI Features Analysis: Last but not least, we enhance the quality of the filtered picture by filling in the gaps and using a segmentation technique to more precisely define the contours of the previous filtered image in order to obtain the characteristics required for stage assignment to each PROI.

We extract and quantify geometrical features, motivated by [14], to characterize each detected EMROI, and we compare them to features extracted from the TW2 stage classification model.

Figure 16: Division of a ROI into metaphases, diaphises and epiphises (left) and epiphysis region of ossification (right).

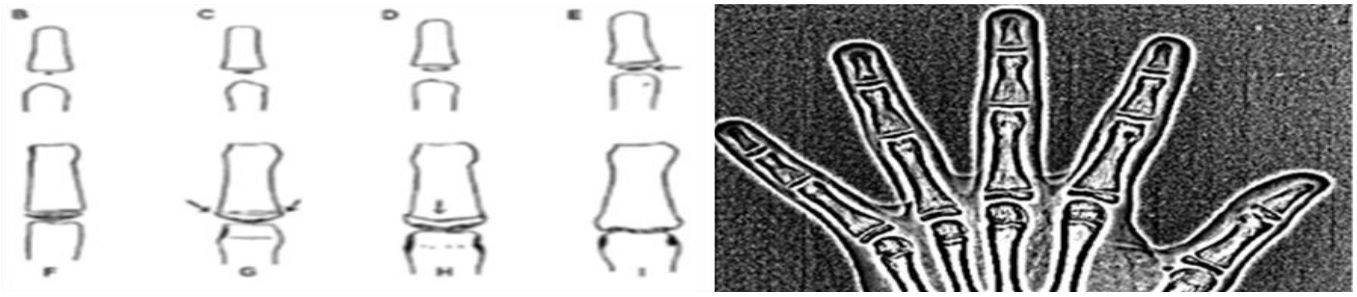

Figure 17: Stages of phalanx bone in TW (left) and identification using the proposed method (right).

3.5.3 CROI and UR System

The eight carpal bones, Ulna, radial, and other bones make up the complicated anatomical region of the wrist, which also has a large number of articular surfaces, ligaments, tendons, and neurovascular

structures. Figure 18 shows how the carpal bones looked in two distinct X-rays taken at two different ages.

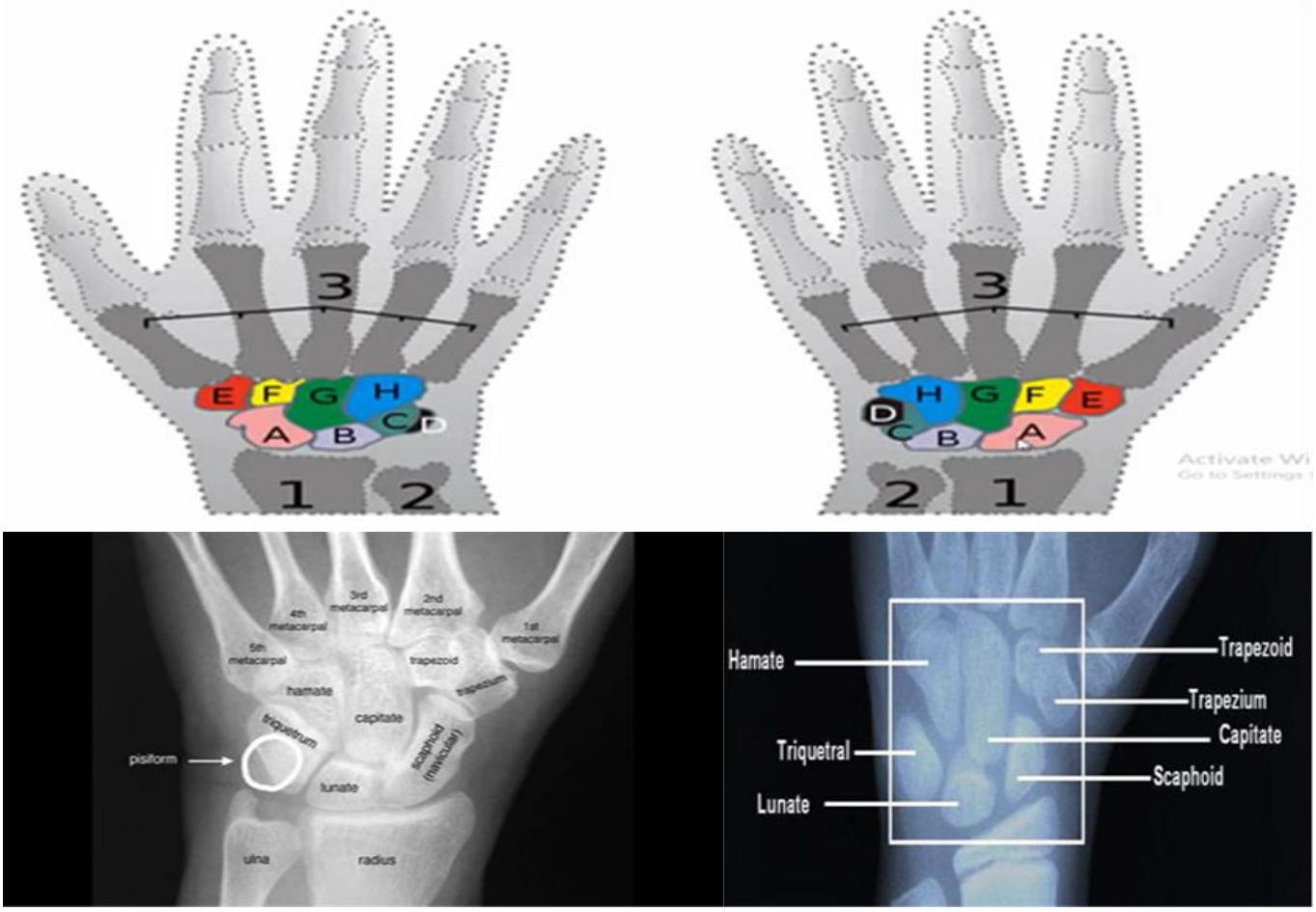

Figure 18: Carpal bones in two distinct X-ray images taken at two different ages.

On X-ray images, the carpal bones can be seen as a collection of dense pinpoints in the early stages. People grow bigger as they age until they acquire their ideal size and distinctive shape. The following is the order in which the carpal bones appear: Capitate, Hamate, Triquetral, and Lunate are often placed in that order, then the Scaphoid, Trapezium, and Trapezoid. The carpal bones are typically used in clinical treatments, but few automatization techniques have been offered because the trapezium and trapezoid typically fuse after seven years for males and five years for females. Table 5 shows the average ages at which the ossification centers first appear in the carpal bones. In the current thesis study, a system for the CROI processing is suggested that can control the fusing of these bones. The CROIs process includes the following steps: i) localization of the area containing the carpal bones; ii) augmentation and extraction of the eight carpal bones, and iii) treatment of the trapezium/trapezoid fusion.

The extraction of the carpal bones from the complete hand is the first stage in a CROI analysis, and it is carried out by properly putting the preceding procedures into practice. The soft tissue connection between the second finger and the thumb is located at the top-right portion of the carpal ROI in Fig. 18. From this point, it is simple to locate the other point, and using these two points together, we define a rectangle that designates the area known as the carpal ROI [16].

Table 5: Average age of the carpal bone' ossification centers [53].

Bone	**Female**	**Male**
Hamate	2 Months	4 Months
Capitate	2 Months	4 Months
Pyramidal	2 Years	3 Years
Lunate	3 Years	4 Years
Trapezium	3 Years	4 Years
Trapezoid	4 Years	6 Years
Scaphoid	4 Years	6 Years
Pisiform	9 Years	12 Years

3.6 Materials and Methods

The reliability of different ROIs in different age groups and genders is an important factor to consider when assessing bone age. Table 6 presents data on the degree of reliability of the different ROIs considered. The reliability of each ROI is dependent not only on the type of ROI but also on the age group and gender. For example, the reliability of the distal radius ROI is higher in younger age groups than in older age groups. This is likely since the distal radius is more developed in younger age groups, making it easier to accurately measure the bone age. The reliability of the distal ulna ROI is also more reliable in men than in women. This is most likely because measuring the bone age accurately is made simpler in men as their distal ulnas are greater than those of women.

Table 6: Comparison of the reliability with respect to ROIs.

Age group / ROI	Phalangeal ROIs	Carpal Bone ROI	Wrist Joint ROI
0 – 5 (female) 0 – 7 (male)	Feature analysis of epi-metaphysis - **NOT Reliable**	Size & shape analysis of carpal bones - **Reliable**	
6 – 13 (female) 8 – 15 (male)	Feature analysis of epi-metaphysis - **Reliable**	Degree of overlapping of carpal bones - **NOT Reliable**	
14 – 18 (female) 16 – 18 (male)	Features are - **NOT Sufficient**		Feature analysis of epi-diaphysis - **Reliable**

The reliability of the proximal humerus ROI is higher in both younger and older age groups than in the middle age group. This is likely since the proximal humerus is more developed in both younger and older age groups, making it easier to accurately measure the bone age. Similarly, the reliability of the proximal radius ROI is higher in males than in females. This is likely since the proximal radius is larger in males than in females, making it easier to accurately measure the bone age.

Overall, the reliability of different ROIs in different age groups and genders is an important factor to consider when assessing bone age. The data presented in Table 6 provides valuable insight into the degree

of reliability of the different ROIs considered. By considering the age group and gender of the patient, clinicians can better assess the reliability of the ROI and make more accurate bone age calculations.

Data synthesis in the current study utilized the open-source frameworks that are available on GitHub. Using these frameworks, object detection models are created, which are then tested and trained on the detection algorithm. The open-source frameworks provide wide range of tools and libraries that can be used to develop and deploy object detection models. These frameworks are designed to be easy to use and provide a wide range of features that can be used to customize the model to the specific application. The open-source frameworks provide a wide range of datasets that can be used to train and test the object detection models. These datasets are typically labeled with the objects that need to be detected, which makes it easier to train the model. The datasets can also be augmented with additional data to improve the accuracy of the model. Additionally, the open-source frameworks provide a wide range of tools and libraries that can be used to monitor, optimize and to make adjustment to the model to improve its accuracy. These frameworks are designed to be easy to use and provide a wide range of features that can be used to customize the model to the specific application and provide a wide range of datasets that can be used to train and test the object detection models. This makes it easier for researchers to develop and deploy accurate object detection models.

MATLAB was used as a tool for preprocessing data and implement watershed segmentation on the images. Watershed segmentation technique is used to separate objects in an image by using a 'watershed' line to separate them. This is especially useful for medical imaging, where it can be used to identify and segment different parts of the body. The combination of MATLAB 14B and the open source frameworks available on Github have been used to create the overall model for the detection scheme. The data can be preprocessed with watershed segmentation using MATLAB 14B, and the detection models can be trained and tested using free source frameworks. The bounding boxes that each model generates around the observed items are used to assess how well the predictions were made.

The final stage of the proposed algorithm does age classification based on deep learning. For this purpose, an end-to-end learning using InceptionV3 architecture was used with custom fully connected layers in the final stage for regression. Transfer learning from ImageNet pre-trained models was used to initialize the weights of InceptionV3. Experiments were done with other architectures as well and InceptionV3 was found to be performing better among those. Mean Absolute Error (MAE) is the model's chosen evaluation metric. The pre-trained InceptionV3 deep learning model is utilized to train and evaluate the assessment of bone age. A customized solution written in Python 3.9.12 is constructed utilizing Keras, Tensor Flow, and other relevant libraries.

3.6.1 Data Sets

In the current thesis work, the RSNA data sets have been used for bone assessment (BA) composed of 5778 female and 6833 male subjects (see also Fig. 19) with maximum age of 228 months (19 years) and minimum 1 month. That makes it a total of 12611 X-ray hand images [37]. The dataset is available for download from Kaggle's website, and it is a great resource for anyone interested in developing ML models to accurately identify the age of a child from an X-ray of their hand. The dataset is well-structured

and easy to use, and it provides a great starting point for anyone looking to develop a model for age assessment based on X-ray images.

As part of its initiatives to promote the development of AI tools for radiology, the Radiological Society of North America (RSNA) organized the RSNA Pediatric Bone Age Machine Learning Challenge 2017 to evaluate BA from pediatric hand radiographs. The objective of the RSNA 2017 Machine Learning Challenge was to create an algorithm that, when applied to a validation set of pediatric hand radiographs, can most accurately calculate BA. By calculating the mean difference and mean absolute difference (MAD) between the performance of each system and the average of all reviewers' assessments, the findings were reviewed [54].

Each image in the collection is an unsegmented radiograph, which means that it provides a lot of information that is not anatomically necessary for determining age. The current thesis work aimed to determine the most effective segmentation technique for this issue and to assess the impact that segmentation has on the age estimation accuracy of several pre-trained models [37].

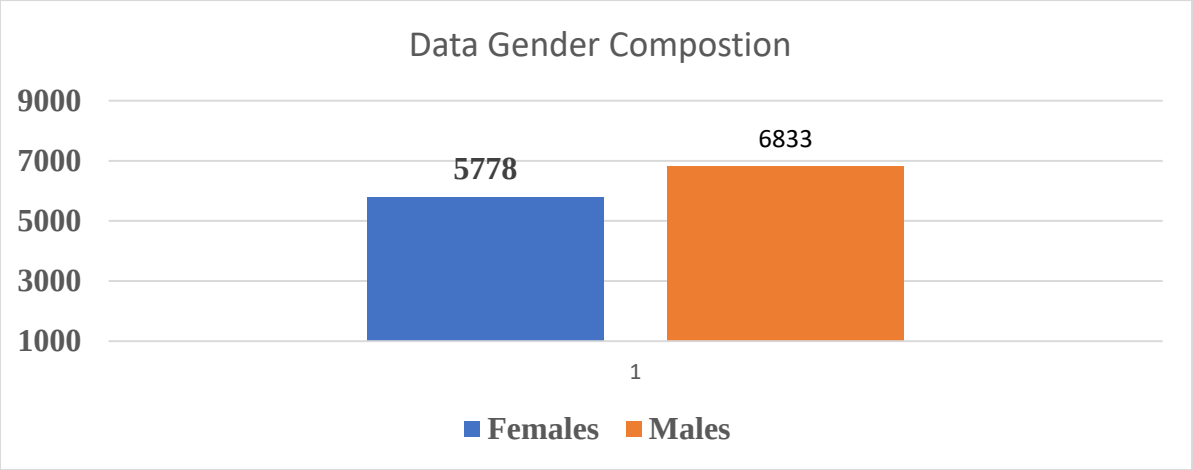

Figure 19: Data gender composition.

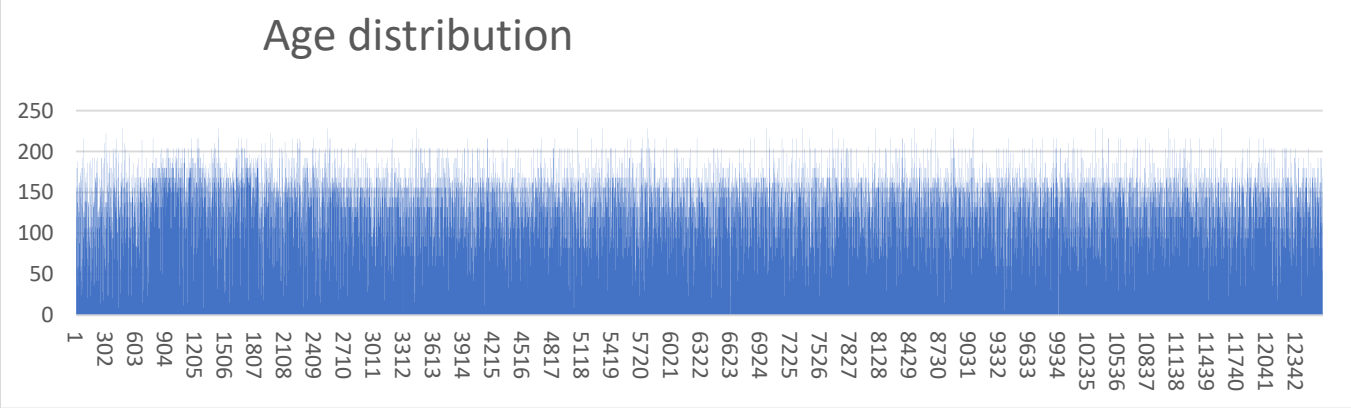

Figure 20: Age distribution of the data.

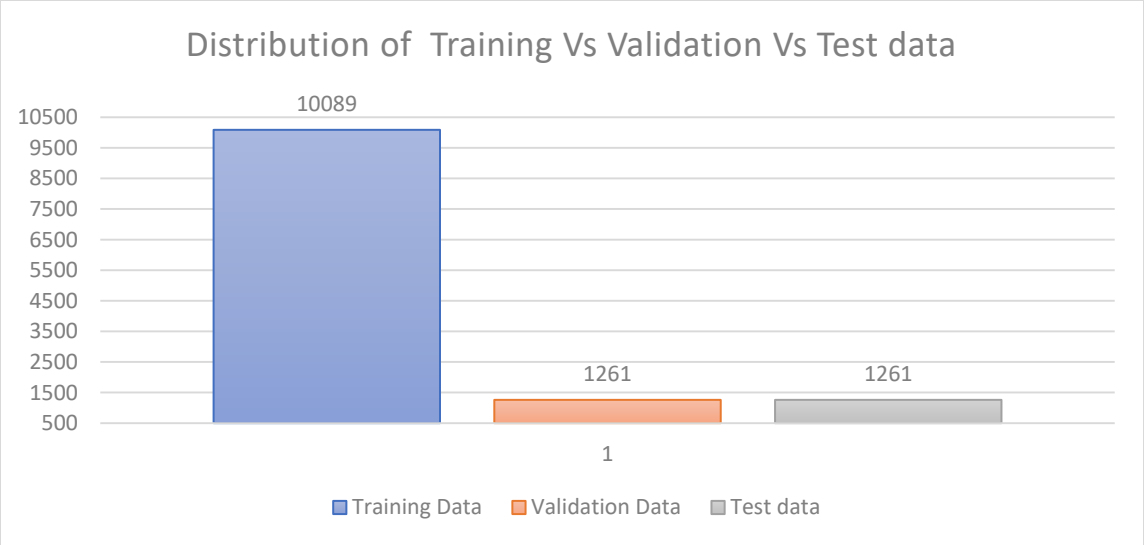

Figure 21: Distribution of Training Vs Validation Vs Test data,

Figure 19 and Fig. 20 display the age distribution and gender composition of the data with the 12611 participants (minimum 1 month and maximum 228 months). This data was decomposed into 80% for training, 10% for validation and 10% for testing as shown in Fig. 21. The percentages were chosen to ensure that the model was able to generalize well. Labeling of the images was carried out, with skeletal age estimates and sex from the accompanying clinical radiology report provided with the data source.

3.6.2 Watershed Segmentation

The discontinuity and similarity features that may be retrieved from pixel values, or a combination of them, are the foundation of the majority of picture segmentation methods. In spite of numerous proposed and realized picture segmentation approaches and algorithms over the past 40 years, the segmentation problem is still unresolved outside of comparatively 'easy' images. One of the more challenging image processing procedures is the separation of touching objects, and watershed segmentation is a method that has gained popularity in recent years. For the purpose of watershed segmentation, we consider the 2D, grey-scale image to be a topological surface or 'landscape' where the position is determined by the x, y image coordinates and the height corresponds to the image intensity or grey-scale value. Compared to the other methods we have seen in this chapter, segmentation using watershed methods offers several benefits. Contrary to edge detection-based methods, watershed methods produce closed contours to define the boundary of the objects, which is a significant advantage [55].

A segmentation function is a function of the original image (i.e., a different image derived from the original) whose characteristics are such that the watershed catchment basins exist within the objects we intend to divide. Watershed segmentation is successful when appropriate segmentation functions are computed, which is not always easy to do. The first preprocessing step in watershed segmentation frequently involves the use of gradient pictures for the straightforward reason that the gradient magnitude is high at object edges and low elsewhere. The watershed ridges ought to be located along the margins of the objects, in a perfect world.

Chapter 4: Results and Discussion

The edge detection and segmentation frameworks described in Chapter 3 generate the input images to the deep learning framework. All 12611 subjects had accurate birth certificates showing their actual ages and these have been used as the ground truth information. All of the subjects were given clinical and X-ray examinations to ensure that the data was accurate and reliable.

4.1 Experimental Results

4.1.1 Segmentation Results

Figure 22 shows a (largely) successful watershed segmentation of selected hand images of given subjects and the outputs of the watershed algorithm. It is interesting to see that the watershed algorithm (preceded by the pre-processing steps) was able to enhance edges in the raw radiographs. These segmented outputs are the once used as inputs for the deep learning scheme (Inception V3 in our case).

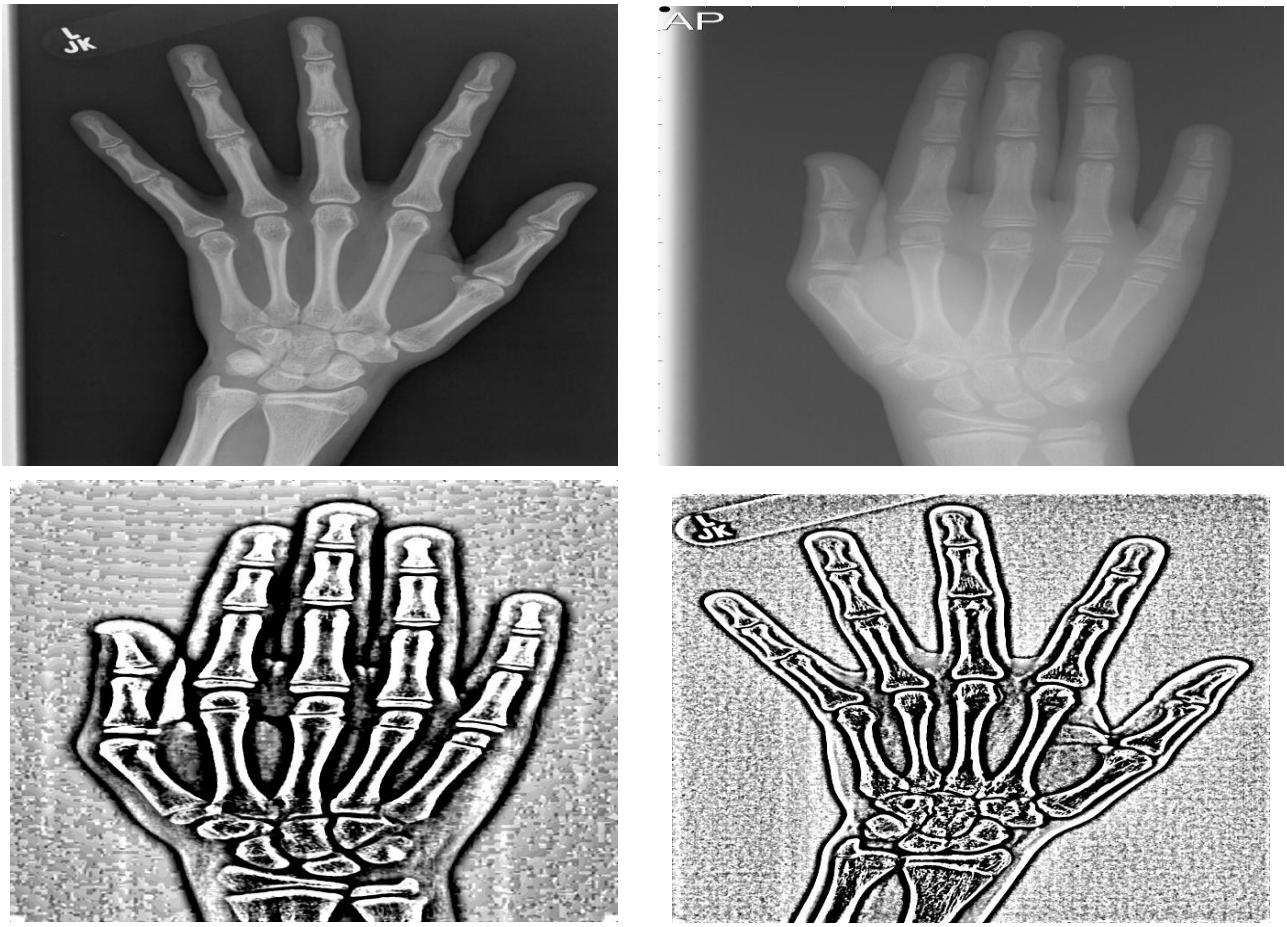

Figure 22: Input raw hand images (top) and watershed segmented images (bottom).

4.1.2Deep Learning Outputs

As mentioned in Chapter 3, the deep learning architecture utilized in the current thesis work was Inception V3. Different deep learning model parameters (including batch size, epoch, learning rate, etc…) were checked for their efficacy in classifying the subjects under consideration into different age groups.[56,57]The activation function used was Relu (Rectified Linear Unit) while Adam's was used as the optimizer.

The best results were found with a batch size of 32, epoch of 30 and default learning rate of 0.001. The best evaluation matric is found to be the mean absolute error, as also described in Chapter 3. Accordingly, the best training result was achieved with a MAE of 4.42months, which is much lower than MAE results reported in previous literatures [9, 35]. This can be considered as a very promising result demonstrating the good performance of the deep learning scheme when it is feed with a pre-processed and watershed segmented images instead of raw radiological images. The validation and testing accuracies were calculated to be 14.08 and 12.25, respectively, in terms of mean absolute error. Note that 80% of the data was used for training, 10% for validation and the rest 10% for testing. The testing result showed that on average there is about a 1 year difference between the chronological ages and the ages predicted by the proposed model. This result can still be considered encouraging but at the same time showing the fact that there are still rooms to improve the performance of the deep learning scheme. Table 7 below shows the training results with 30 epochs while the data was labeled as 80% for training (10089), 10% for validation (1261) and 10% for testing (1261), making a total of 12611 images.

Table 7: Training, validation and testing results obtained using the proposed deep learning scheme.

Epoch	Package	MAE (Training)	MAE (Validation)
1/30	316/316	18.2139	15.7053
2/30	316/316	13.9603	14.0425
3/30	316/316	12.4225	12.6981
4/30	316/316	11.0405	11.5727
5/30	316/316	9.7503	12.0645
6/30	316/316	8.7918	11.1863
7/30	316/316	8.1239	13.7378
8/30	316/316	7.6637	12.7039
9/30	316/316	7.7615	12.6125
10/30	316/316	6.5600	10.8293
11/30	316/316	7.5763	17.5336
12/30	316/316	6.7244	12.1115
13/30	316/316	5.8010	11.6304
14/30	316/316	8.8804	10.1791
15/30	316/316	7.3563	13.2543
16/30	316/316	7.5441	19.4891
17/30	316/316	8.4995	23.6113
18/30	316/316	6.4732	10.3850
19/30	316/316	5.1342	10.9953
20/30	316/316	4.7570	11.1538

21/30	316/316	4.6134	10.9069
22/30	316/316	4.7575	14.5782
23/30	316/316	5.1998	11.7560
24/30	316/316	4.9637	10.5641
25/30	316/316	4.9402	10.5766
26/30	316/316	5.0278	11.4847
27/30	316/316	4.9764	10.2043
28/30	316/316	4.4214	14.0863
29/30	316/316	4.9260	13.4748
30/30	316/316	5.2085	11.8408
	MAE (Testing): **12.2526**		

InceptionV3 is considered a potent deep learning model. Our proposed system uses deep learning to forecast the age range of the input hand image and extract high-level complicated age-related hand wrist characteristics. In 2015, Google created the DCNN architecture. The ImageNet dataset was used to train the model, which is based on the Google Neural Network architecture. The InceptionV3 model is more powerful than other models because it uses a combination of convolutional layers, pooling layers, and fully connected layers to create a deep learning architecture that can accurately classify images. The model also uses a technique called inception modules, which are a series of convolutional layers that are used to detect different features in an image and are designed to capture different levels of abstraction in the image. This allows the model to detect more complex features in an image than other models [9, 35]. The network is able to learn from a variety of different types of images, including those with different lighting conditions, angles, and backgrounds. This allows the network to be more robust and accurate when making predictions. In addition to its powerful processing capacity, InceptionV3 also has a very efficient architecture. It is able to process images quickly and accurately while using fewer resources than other architectures. This makes it an ideal choice for the hand age estimation task, as it can process a large number of images quickly and accurately.

4.2 Discussion

Using pre-trained CNNs can be a useful way to improve the performance of age estimation. The RSNA X-ray hand dataset is a common dataset used for evaluating age estimation models, and there are several pre-trained CNNs that have been evaluated on this dataset. Some of the most used pre-trained CNNs for age estimation on the RSNA X-ray hand dataset include:

VGG-16: This is a deep CNN architecture with 16 layers that was originally developed for image classification. It has been shown to perform well on the RSNA X-ray hand dataset for age estimation.

ResNet-50: This is a deep CNN architecture with 50 layers that was specifically designed to address the problem of vanishing gradients in deep networks. It has also been shown to perform well on the RSNA X-ray hand dataset for age estimation.

Mobile Net: This is a lightweight CNN architecture that is designed for use on mobile devices with limited computational resources. While it may not perform as well as other CNN architectures on the RSNA X-ray hand dataset, it can be useful when working with limited computational resources.

InceptionV3: This is the CNN architecture adopted in the current study that uses a combination of convolutional layers and inception modules to extract features from images. It has also been shown to perform well on the RSNA X-ray hand dataset for age estimation.

However, it has been shown in previous review papers that InceptionV3 outperforms the other architectures, particularly in areas of age estimation, though there were rooms for improvement. That was the main intent of the current study to improve the performance of the model by playing with the input images before imputing it to the deep neural network. Instead of using the raw radiological images as input, watershed segmented (preceded by useful pre-processing tools) images are inputted to the model. The results clearly showed that the approach has significantly improved the results. This is particularly so during training. The validation and testing results, however, needs improvement.

The size and complexity of the model, the computational resources available for training and inference, and the specifics of the dataset, such as the age ranges represented and the number of images available for training and validation, should all be taken into account when choosing a pre-trained CNN for age estimation on the RSNA X-ray hand dataset.

In the current thesis study, deep learning was applied to medical pictures, specifically for automated bone age assessment utilizing X-ray images. The dataset, which comprises of 12611 X-ray pictures of children's hands aged 0 to 19, demonstrated that deep learning methods, even those trained on generic imagery, are capable of handling all potential scenarios for automated skeletal bone age assessment. During the inception stage of the current thesis work, it was envisaged that the developed deep learning model would be checked for its efficacy when applied on data acquired from local clinics/hospitals. That was only partially successful. A total of 14 radiological hand image samples were acquired from Tikir Anbessa Specialized Hospital (TASH) with ages determined using two sources: the word of mouth from the 14 clients and that of the report by physicians. The actual chronological ages of the 14 samples were not available. The performance of the proposed deep learning scheme was checked using this samples and it was found that there is a deviation (MAE) of 27.7 months from the clients' information. The deviation from the physicians report was found to be much higher at MAE of 49.41 months. These results were supposed to be checked against the correct chronological ages, information that was not available for any of the 14 samples.

It is important to note that while automated image processing techniques for bone age determination have shown promises, they should be used as a supplementary tool to aid medical professionals in their diagnoses rather than as a replacement for clinical expertise. Additionally, it is important to consider ethical and cultural implications when implementing such technologies in sensitive contexts like Ethiopia. Before deploying an automated image processing system for bone age determination in Ethiopia, it is important to ensure that the invention is culturally appropriate and respectful of local values and beliefs. This may involve consulting with local communities and healthcare professionals to understand their

needs and concerns and ensure that the technology is implemented in a way that is acceptable to all stakeholders.

Additionally, it is important to consider the potential for bias and errors in automated image processing systems. These systems rely on algorithms that may not be accurate for all individuals or populations, which could result in inaccurate or unfair age determinations. Careful validation and testing of the system should be conducted to ensure that it is accurate and fair for all individuals, regardless of their age or background.

Overall, age estimation from hand images can be a challenging task, and it's important to carefully consider all the factors that could be affecting a given model's performance, including the quality and quantity of the data, the choice of preprocessing techniques, and the suitability of the model architecture.

4.3 Uncertainty

Age estimation seeks to accurately determine a person's chronological age when that person's actual age is unknown. However, the word 'estimation' defines the actual boundaries of this knowledge. The exact chronological age of a human being cannot currently be determined by any medical test or collection of tests. There will always be some uncertainty surrounding the estimate, and accurately communicating this uncertainty is just as crucial as accurately presenting the estimate itself. The following factors affect age estimation uncertainty:

Individual biological variations: Age-related changes occur at different rates in different individuals. This can be due to genetic factors, environmental factors, or a combination of both. As a result, a person's biological age may not correspond exactly with their chronological age. There will always be a significant difference in how people develop while being the same age chronologically.

Reference datasets: If the reference standards are built improperly, mistakes may result. Reference standards might not always accurately reflect the whole range of human variability in aging because the human body exhibits natural variance. The distribution of the reference resources may not be equal across all age groups. Reference datasets frequently provide standard deviation estimates. However, these only apply to the reference dataset itself, not to a population. Finally, not all populations have access to reference datasets, which may vary depending on the population.

Inter-/intra observer variations: The measurements made by the observers may differ between and across observers, and the perception of developmental phases is no exception.

Methodological errors and uncertainties: This will depend on the method. For instance, stage assessment-based approaches often have more uncertainty in the stage assignment near stage borders. We need to find strategies that can help reduce uncertainty whenever possible. The inherent variability will always exist despite the fact that the changes brought on by reference data, observers, and methodologies will never be zero. Therefore, it's critical to comprehend, express, and quantify uncertainty in the most effective manner possible.

Incomplete or inaccurate information: In many cases, information about a person's medical history, developmental milestones, or other factors that may affect age estimation may be incomplete or inaccurate.

There is the idea of the image database: which photos should be chosen to test an algorithm. This has to do with the variety and complexity of the chosen photos, the number of databases used in the selection procedure, and the importance of the images to the segmentation task.

It is also important to recognize these uncertainties when conducting age estimation and to clearly communicate the limitations of any estimates to stakeholders, including legal authorities and healthcare providers. This can help to prevent misunderstandings and promote accurate and ethical use of age estimation techniques. Finally, the use of any automated age estimation system should be accompanied by appropriate measures to protect the privacy and confidentiality of individuals whose images are used in the system. This may involve obtaining informed consent from individuals, ensuring that the images are securely stored and processed, and implementing appropriate data protection measures.

Chapter 5: Conclusion and Recommendations

5.1 Conclusion

Age estimation is the process of determining the age that can be used in various fields for different purposes, such as determining the age of a person and object. In medicine, age estimation is often used to determine the appropriate treatment plan for a patient, as the age of a patient can affect the way their body responds to treatment. In sports and athletics, age estimation is used to ensure fair play, as some sports have age restrictions or rules that apply to different age groups. In criminal investigations, age estimation can be used to identify suspects or victims, as well as to determine the legal consequences of a crime.

This thesis work focused on current research challenges in age assessment of humans using radiographs of hand wrist bones and image processing. The research aimed to develop a reliable and accurate method for age estimation by analyzing the radiographs of hand wrist bones of individuals of different age groups and using image processing techniques to extract features from the radiographs. The extracted features were then used to develop a model for age estimation. The research was conducted in two phases. In the first phase, the radiographs of hand wrist bones of individuals of different age groups were collected and analyzed. The radiographs were pre-processed to remove noise and enhance the features. After pre-processing, the radiographs were segmented using watershed algorithm. In the second phase, the segmented images were used as input to develop a deep learning scheme that could be used in effective age determination. The model was trained using a supervised learning approach and tested on a secondary dataset obtained from a known data source. Efforts were also excreted to check the efficacy of the method with respect to its performance when applied on locally acquired dataset.

The results showed that the model was able to accurately estimate the age of individuals from the radiographs of hand wrist bones in good agreement with the available ground truth information. The proposed watershed-based image segmentation framework was designed to improve the accuracy of the final age determination of the deep learning scheme utilized in the thesis work. The segmentation framework was mainly designed to ease the determination of the fusion status of the ROIs. Compared to other similar works proposed in the literature, the proposed work resulted in age estimation of superior performance. There are however rooms for improvement.

Such automated schemes, if proved to be robust, are believed to reduce the workload of medical practitioners who rely on their manual judgment when assessing ages. The manual methods are often time consuming, error prone, subject to issues of observer variability and at the same time being hardly repetitive. It has been shown in the current work that deep learning trained models can be used to work around these issues.

Checking the performance of the deep learning scheme when applied on hand images captured using other imaging modalities, such as MRI could be interesting. In principle, such schemes can be extended to include images taken through different modalities, other than X-rays used in the current study.

The proposed method might not be free from the limitations and uncertainties inherent in any age estimation system. These may include variations in bone development among different populations, the potential for errors or biases in the system, and ethical considerations related to privacy and data protection. Therefore, it is important to approach the use of such systems with caution and to carefully evaluate their performance and potential impact in various contexts.

Various methods are used for age estimation, including dental examination, skeletal analysis, and medical imaging. However, it is important to note that no method is entirely accurate, and there may be some margin of errors and the estimates may vary depending on the method used and the individual being examined. Additionally, certain factors, such as health status, nutrition, and environmental factors, may affect the accuracy of age estimation methods. As such, age estimation should always be considered as an approximation rather than a precise determination.

5.2 Recommendations

It is worth noting that age estimation is not always precise and may be affected by various factors such as genetics, environmental factors, and health conditions. Therefore, it is often used in conjunction with other methods and factors to make a more accurate determination of age. Additionally, there are ethical considerations when it comes to age estimation, particularly in cases where the individual is not able to provide consent or when it may impact their rights and privileges. It is important to consider the ethical implications of age estimation, particularly when it comes to determining the age of individuals who may be vulnerable or marginalized. In some cases, inaccurate or inappropriate age estimation methods may lead to discrimination, stigmatization, or other negative outcomes. It is therefore important to approach age estimation with care and sensitivity, and to use methods that are as accurate, reliable, and objective as possible, while also considering the potential social and cultural factors that may impact the accuracy of the estimation. It is also important to consider the privacy and confidentiality of individuals who are undergoing age estimation, and to ensure that any information obtained through age estimation is used only for appropriate and legitimate purposes.

There are several methods of age estimation that can be used in different fields. In medicine, age estimation can be done through a physical examination of the patient, including assessing their height, weight, and the development of their physical features such as their teeth and bones. X-rays and other medical imaging techniques are also used to estimate a subject’s age. In law enforcement, age estimation can be done through a variety of methods such as physical examination, dental examination, and radiographic imaging. However, it is important to note that these methods have limitations and potential inaccuracies and should not be relied upon as the sole basis for legal decisions. Regardless of the specific method used, age estimation is an important tool for researchers, practitioners, and investigators in a wide range of fields. There is no single, foolproof method for accurately determining the age of a person or object, and estimates may be subject to significant variation depending on the method used and the context of the estimation. Additionally, age estimation methods may be subject to biases and inaccuracies due to factors such as demographic differences, cultural factors, and variations in biological development. It is therefore important to approach age estimation with care and consideration, and to use appropriate methods and techniques to minimize potential errors and inaccuracies.

The long-term research challenges of accurately determining the age of individuals in different sectors is a complex and multifaceted issue. It requires collaboration between different sectors that require scientific age information, such as courts for young criminals, early marriages, rape, and family inheritance conflicts, as well as sectors that require accurate ages of individuals such as football players, athletes, and people joining the military. The first step in addressing this challenge is to develop a comprehensive understanding of the different sectors and the age-related issues they face. These include understanding the legal and social implications of age determination in each sector, as well as the different methods used to determine age. For example, in the court system, age is often determined through birth certificates, school records, or other documents. In the military, age is often determined through physical examinations or medical records. In sports, age is often determined through physical tests or medical records. Once a comprehensive understanding of the different sectors and their age-related issues has been developed, the next step is to develop a unified approach to age determination. This approach should take into account the different methods used to determine age in each sector, as well as the legal and social implications of age determination. It should also consider the accuracy of the different methods used to determine age, as well as the cost and time associated with each method.

Finally, the unified approach should be implemented in each sector. This could involve the development of new laws or regulations, the implementation of new technologies, or the creation of new systems for age determination. It is also important to ensure that the unified approach is implemented in a way that is fair and equitable for all individuals, regardless of their age. Furthermore, the integration of additional clinical and demographic data into the automated age estimation system could potentially improve its accuracy and applicability to different populations and contexts. For example, including information about ethnicity, nutrition, and environmental factors could help to identify specific factors that may impact bone development and age estimation accuracy.

References

1. Eikvil Line et al., (2012), Age estimation in youths and young adults: A summary of the needs for methodological research and development, Norway, Norwegian Computing Center.

2. D. Giordano et al., Epiphysis and Metaphysis Extraction and Classification by Adaptive Thresholding and DoG Filtering for Automated Skeletal Bone Age Analysis, Italy, University of Catania.

3. PietkaEwa et al., (2001) Computer Automated Approach to the Extraction of Epiphyseal Regions in Hand Radiography, Journal of Digital Imaging Vol 14.

4. Wahab Abdullah Nagi Abdul, Age estimation by using radiological methods and its Comparison with clinical examination. – Yemen, Faculty of medicine - Aden University

5. Smith Terry et al., (2011), Laura Age assessment practices: a literature review & annotated bibliography, New York by United Nations Children's Fund (UNICEF).

6. Ariane Maggio, (2014), Age estimation using the hand-wrist: morphological assessment of skeletal development in Western Australia

7. Jang Hyung Lee et al, (2018), Applying Deep Learning in Medical Images: The Case of Bone Age Estimation.

8. Wikipedia source images of Clavicle, First rib, Clavicle vertebrae and dental

9. Westerberg Erik,(2020) AI-based Age Estimation using X-ray Hand Images A comparison of Object Detection and Deep Learning models

10. Gilsanz Vicenteet al., (2005), Hand Bone Age A Digital Atlas of Skeletal Maturity, New York, Springer. Germany, Springer-Verlag Berlin Heidelberg Printed in Germany.

11. Gertych Arkadiusz, et al. (2007), Bone Age Assessment of Children using a Digital Hand Atlas, Gliwice Poland, Compute Med Imaging Graph.; 31(4-5): 322–331

12. M. V. Bramhananda et al., (2012), Bone Age Assessment using Wavelet Transform, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 6.

13. Caoa, F. et al., (2000), Digital hand atlas and web-based bone age assessment: system design and implementation, USA, Computerized Medical Imaging and Graphics, 297-307.

14. Sutton David, (2003), Textbook of Radiology and Imaging, London, UK, Noordanesh Medical Publication.

15. S. Concetto, Skeletal bone age assessment, Italy, University of Catania.

16. Giordano Daniela, et al., (2009), An Automatic System for Skeletal Bone Age Measurement by Robust Processing of Carpal and Epiphysial/Metaphysial Bones, Italy, University of Catania, Catania

17. PietkaEwa (1995), Computer-assisted bone age assessment based on features automatically extracted from a hand radiograph, Geneva. Switzerland, Computerized Medical Imaging and Graphic, Vol. 19 No. 3, pp: 251 -259.

18. Chiang Kuo-Hsien et al., (2005), The Reliability of Using Greulich-Pyle Method to Determine Children's Bone Age in Taiwan, Taoyuan, Taiwan, Chang Gung University

19. MansourvarMarjan et al., (2012), Automated web-based system for bone age assessment using histogram technique, Kuala Lumpur, Malaysian Journal of Computer Science. Vol. 25 (3).

20. Benjavongkulchai S., (2018), Age estimation methods using hand and wrist radiographs in a group of contemporary Thais,

21. Zafar Abdul Mueed, et al., (2010), An appraisal of Greulich-Pyle Atlas for skeletal age assessment in Pakistan, Karachi, Pakistan. Original Article, Vol. 60, No. 7

22. Chiang Kuo-Hsien et al., (2005), The Reliability of Using Greulich-Pyle Method to Determine Children's Bone Age in Taiwan, Taoyuan, Taiwan, Chang Gung University,

23. Dvorak Jiri, et al.,(2007), Age determination by magnetic resonance imaging of the wrist in adolescent male football players, Zurich, Switzerland Br J Sports Med:45–52.

24. M. V. Bramhananda et al., (2012), Bone Age Assessment using Wavelet Transform, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 6

25. Chai Hum Yan, et al., (2011) An artifacts removal post-processing for epiphyseal region-of-interest (EROI) localization in automated bone age assessment (BAA) Malaysia, Chai et al. Biomedical EngineeringOnline

26. Liu Jian et al., (2008), Automatic bone age assessment based on intelligent algorithms and comparison with TW3 method, China, Computerized Medical Imaging and Graphics, 678–684.

27. Seok Jinwoo et al., (2012), Automated Classification System for Bone Age X-ray Images, Seoul, Korea, IEEE International Conference on Systems, Man, and Cybernetics

28. Hsieh Chi-Wen et al., (2007), Computerized geometric features of carpal bone for bone age estimation, China, Chinese Medical Journal.

29. R. Cameriere et al., (2014), Automatic age estimation in adults by analysis of canine pulp/tooth ratio: Preliminary results, Italy, Journal of Forensic Radiology, and Imaging.

30. Elke Hillewig, (2010), Magnetic resonance imaging of the medial extremity of the clavicle in forensic bone age determination: a new four-minute approach, Ghent, Belgium, European Society of Radiology.

31. Mughal et al., (2014), Bone age assessment methods: A critical review, Karachi, Pakistan. Pak J Med Sci.

32. E.O. Omidiora, (2012), Analysis, Design, and Implementation of Human Fingerprint Patterns System “Towards Age & Gender Determination, Ridge Thickness to Valley Thickness Ratio (RTVTR) & Ridge Count on Gender Detection”, Nigeria, (IJARAI) International Journal of Advanced Research in Artificial Intelligence, Vol. 1.

33. Darko. Stern et al., (2013), Automated Age Estimation from Hand MRI Volumes using Deep Learning, Austria, Biotech Med-Graz.

34. Shao-Yan Zhang et al., (2013) Automated Determination of Bone Age in a Modern Chinese Population, China, Hindawi Publishing Corporation

35. Ian Pan, et al., (2020), A Deep Learning Approach to Pediatric Bone Age Assessment Using Pediatric Trauma Hand Radiographs, Radiology: Artificial Intelligence

36. C. Spampinato, et al., (2016), Deep Learning for Automated Skeletal Bone Age Assessment in X-Ray Images

37. Safwan S. et al.,(2019), The RSNA Pediatric Bone Age Machine Learning Challenge

38. M.Z.Lazarus, et al., (2013), Age classification: based on wrinkle analysis, India, Andhra Pradesh, international Journal on Recent and Innovation Trends in Computing and Communication, Volume: 1 Issue: 3

39. John L. Semmlow, (2004), Bio signal and Biomedical Image processing, MATLAB Based Applications, USA, Marcel Dekker, Inc

40. C. Gonzalez Rafael, (2004), Digital Image Processing Using MATLAB, Library of Congress Cataloging-in-Publication Data on File.

41. Chris Solomon, et al., (2011), Fundamentals of Digital Image Processing: A Practical Approach with Examples in Mat lab, UK, John Wiley & Sons, Ltd.

42. C. Gonzalez Rafael, (2004), Digital Image Processing Using MATLAB, third edition, Library of Congress Cataloging-in-Publication Data on File.

43. Lena Costaridou, (2005), Medical image analysis methods, Greece, University of Patras.

44. Geoff Dougherty, (2009) Digital Image Processing for Medical Applications California State University, Channel Islands, United States of America by Cambridge University Press.

45. Cristobal Gabriel et al., (2011), Optical and Digital Image Processing Fundamentals and Applications, WILEY-VCH Verlag GmbH & Co.

46. Bakthula Rajitha, et al, (2014,) Automated Human Bone Age Assessment using Image Processing Methods - Survey

47. Emma Mun Oz-moreno, Automatic Detection of Anatomical Landmarks and Image Registration Applied to Bone Age Assessment, Spain, University of Valladolid.

48. Oge Marques, (2011), Practical Image and Video Processing using MATLAB, Florida Atlantic University, John Wiley & Sons, Inc.

49. Jang Hyung Lee1 et al., (2020), Bone age estimation using deep learning and hand X ray images ·

50. Muhammad Waqas Nadeem et al.,(2020), Bone Age Assessment Empowered with Deep Learning: A Survey, Open Research Challenges and Future Directions

51. Burla Nur Korkmaz et al.,(2022),DNN for bone age prediction by hand X-rays" How effective is utilizing gender information?

52. Savita K Shetty et al.,(2019),Deep Learning Algorithms and Applications in Computer Vision

53. Vincenzo De Sanctis,et al.(2014),Hand X-ray in pediatric endocrinology: Skeletal age assessment and beyond

54. Monika Prokop-Piotrkowska et al.,(2020),Traditional and New Methods of Bone Age Assessment-An Overview

55. Yan Chai Hum, (2013), Segmentation of Hand Bone for Bone Age Assessment, Malaysia, Springer Singapore.

56. Dirk P. Kroese, et al.,(2022),Data Science and Machine Learning Mathematical and Statistical Methods

57. Jason Brownlee,(2016),Machine Learning Mastery with Python Understand Your Data, Create Accurate Models andWork Projects End-To-End

Annex 1: MATLAB Code

```
clear all; close all; clc
fileFolder ='C:\Users\INTuser\Desktop\Train\boneage-training-dataset\boneage-test-dataset';
dirOutput = dir(fullfile(fileFolder,'*.jpg'));
fileNames = {dirOutput.name}';
numFrames = numel(fileNames);
fileNames = fullfile(fileFolder,fileNames);
%%% output folder
fileFolderOut ='C:\Users\INTuser\Desktop\Train\boneage-training-dataset\boneage-test-preprocess';
fileNamesOut = {dirOutput.name}';
fileNamesOut = fullfile(fileFolderOut,fileNamesOut);
parfor n = 1:15
A =imread(fileNames{n});
%B = rgb2gray(A);
B =  A;
B1= A(:,:,1);
B=B1(:,90:end);
I=im2double(B);
D=adapthisteq(I);
J=medfilt2(D,[3 3]);                    %Apply to original image
Isp=imnoise(J, 'salt & pepper',0.03);%Apply to original image
Ispf=medfilt2(Isp,[3 3]);               %Apply to salt and pepper image
Ig=imnoise(J,'gaussian',0.02);
Igf=medfilt2(Ig,[3 3]);                 %Apply to Gaussian image
k=fspecial('log', [40 40],0.35);                %Create LoG filter
%k=fspecial('laplacian');                       %Generate 3 X 3 Laplacian filter
%B=imfilter(J,k);                                       %Filter image with Laplacian kernel
%C=imsubtract(I,Ig);  %Subtract Laplacian from original.
IEl=imfilter(double(J),k,'replicate');          %Laplacian edges
C=imsubtract(J,IEl);%J   %Subtract Laplacian from original.
J1=medfilt2(C,[3 3]);
H1=1-C;
H=1-C./max(abs(C(:)));
L=medfilt2(J1,[3 3]);
%f=C;%RELATIVELY QUALITY IMAGE VIEW1
%f=J1;%RELATIVELY QUALITY IMAGE VIEW
f=H1;%QUALITY IMAGE VIEWS
%f=H;%LIGHT IMAGE VIEW
%f=L;%RELATIVELY QUALITY IMAGE VIEW2
bw1=im2bw(f,graythresh(f));
bw=medfilt2(bw1,[3 3]);
bwc=~bw;
dst=bwdist(bwc);
ws=watershed(-dst);
w=ws==0;
rf=bw&~w;
% figure,imshow(A),title('Original  Of DIGITAL FERIHIWOT AYALE_T-16-P-20.jpg Image');
% figure,imshow(B),title('Original  Of FERIHIWOT AYALE_T-16-P-20.jpg Image');
% figure,imshow(D),title('Original of Adapth Of FERIHIWOT AYALE_T-16-P-20.jpg Image');
% figure,imshow(J),title('Original of MEDIAN Of FERIHIWOT AYALE_T-16-P-20.jpg Image');
% figure,imshow(Isp),title('Isp=imnoiseJ, salt & pepper,0.03');
% figure,imshow(Ispf),title('Ispf=medfilt2(Isp,[3 3] Image');
% figure,imshow(Ig),title('Ig=imnoise J,gaussian,0.02 Of  Image');
% figure,imshow(Igf),title('Igf=medfilt2(Ig,[3 3] F18 Gaussian  Image');
% figure,imshow(k),title('k=fspecial log, [40 40],0.35)LoGfilte Of FERIHIWOT AYALE_T-16-P-20.jpg');
```

```
% figure,imshow(IEl),title('IEl=imfilter(double(J),k,replicate');
% figure,imshow(C),title('C=imsubtract(J,IEl),J  Subtract Laplacian from original');
% figure,imshow(J1),title(' J1 =medfilt2(C,[3 3])Image');
% figure,imshow(L),title('L=medfilt2(J1,[3 3])Image');
% figure,imshow(f),title('f=J1 medfilt2(C,[3 3])Image');
% figure,imshow(bw1),title('bw=im2bw(f,graythresh(f))Negative Image');
% figure,imshow(bw),title('bw=medfilt2(bw1,[5 5]) Image');
% figure,imshow(ws),title('Watershed - Distance Transform');
% figure,imshow(rf),title('Superimposed - Watershed and original image');
 h=fspecial('sobel');
%h2=fspecial('prewit');
%h=h1+h2;
%h=fspecial(f,'canny',0.2);
%h= edge(f, 'canny', [], 0.2);
% h=edge('canny');
%fd=f;
fd=im2double(f);
%figure,imshow(fd),title('fd=im2double(f) Of FERIHIWOT AYALE_T-16-P-20.jpg');
E1=adapthisteq(fd);
%figure,imshow(E1),title('fd of  Adapth Of  FERIHIWOT AYALE_T-16-P-20.jpg');
J2=medfilt2(E1,[3 3]);
%figure,imshow(J2),title('J2=medfilt2(E1,[3 3])of  FERIHIWOT AYALE_T-16-P-20.jpg ');
%sq=sqrt(imfilter(fd,h,'replicate').^2+imfilter(fd,h','replicate').^2);
sq=sqrt(imfilter(J2,h,'replicate').^2+imfilter(J2,h','replicate').^2);
sqoc=imclose(imopen(sq,ones(3,3)),ones(3,3));
wsd=watershed(sqoc);
wg=wsd==0;
rfg=f;
rfg(wg)=255;
%figure,imshow(h),title('h=fspecial SOBLE');
%figure,imshow(fd),title('fd=im2double(f)');
%figure,imshow(J2),title('J2=medfilt2(fd,[3 3])');
%figure,imshow(wsd),title('Watershed - Gradient');
%figure,imshow(rf),title('Superimposed - Watershed and FERIHIWOT AYALE_T-16-P-20.jpg ');
im=imextendedmin(f,20);
Lim=watershed(bwdist(im));
%figure,imshow(Lim),title('Watershed - Marker Controlled');
em=Lim==0;
rfmin=imimposemin(sq,im|em);
wsdmin=watershed(rfmin);
%figure,imshow(rfmin),title('Watershed - Gradient and Marker Controlled');
rfgm=f;
rfgm(wsdmin==0)=255;
%figure,imshow(rfgm),title('Superimposed - Watershed (GM) and  L=medfilt2(J1,[3 3] Of .jpg ');
J3=medfilt2(rfgm,[3 3]);
%figure,imshow(J3),title('J3=medfilt2(rfgm,[3 3]) Superimposed - Watershed (GM) Of DIGITAL FERIHIWOT AYALE_T-
16-P-20.jpg');
file_name=fileNamesOut{n};
imwrite(J3,file_name)end
```

Annex 2: Python Code for Main Training Model

```
from keras. applications. inception_v3 import InceptionV3
from keras.preprocessing import image
from keras.models import Model
from keras.layers import Flatten, Dense, Input
from keras import backend as K
# import cPickle
import _pickle as cPickle
import numpy as np
import matplotlib.pyplot as plt
#from tensorflow.keras.callbacks import ModelCheckpoint
import tensorflow as tf
import os
import cv2
import pandas as pd
os.chdir('D:/Andualem file/boneage-training-dataset')
#os.chdir('D:/Andualem file/boneage-training-dataset/boneage-training-preprocess')
batch_size = 32
epochs = 30
# Load data
print('...loading training data')
f = open('data.pkl', 'rb')
x = cPickle.load(f)
f.close()
f = open('data_age.pkl', 'rb')
y = cPickle.load(f)
f.close()
x = np.asarray(x, dtype=np.float32)
y = np.asarray(y)
x /= 255.
x_final = []
y_final = []
# Shuffle images and split into train, validation and test sets
random_no = np.random.choice(x.shape[0], size=x.shape[0], replace=False)
for i in random_no:
x_final.append(x[i,:,:,:])
y_final.append(y[i])
x_final = np.asarray(x_final)
y_final = np.asarray(y_final)
k = 1261 # Decides split count
x_test = x_final[:k,:,:,:]
y_test = y_final[:k]
x_valid = x_final[k:2*k,:,:,:]
y_valid = y_final[k:2*k]
x_train = x_final[2*k:,:,:,:]
y_train = y_final[2*k:]
print('x_trainshape:'+ str(x_train.shape))
print('y_trainshape:'+ str(y_train.shape))
print('x_valid shape:'+ str(x_valid.shape))
print('y_validshape:'+ str(y_valid.shape))
print('x_test shape:'+ str(x_test.shape))
print('y_test shape:'+ str(y_test.shape))
# Using InceptionV3 with pretrained weights from Imagenet
base_model =  InceptionV3(weights='imagenet', include_top=False)
input = Input(shape=(224,224,3))
```

```
output_vgg16 = base_model(input)
x = Flatten()(output_vgg16)
x = Dense(512, activation='relu')(x)
predictions = Dense(1)(x)
model = Model(inputs=input, outputs=predictions)
model.compile(optimizer='adam', loss='mean_squared_error', metrics=['MAE'])
# Save weights after every epoch
checkpoint = tf.keras.callbacks.ModelCheckpoint(
filepath='weights0/weights.{epoch:02d}-{val_loss:.2f}.hdf5',
save_weights_only=True,
save_best_only = True,
      monitor = 'val_loss',
      period=1, mode='min', patience=10)
history0=model.fit(x_train,y_train,batch_size=batch_size,epochs=epochs,verbose=1,validation_data=(x_valid,y_valid),
callbacks = [checkpoint])
model.save_weights("model0.h5")
score = model.evaluate(x_test, y_test, batch_size=batch_size)
print('Test loss:', score[0])
print('Test MAE:', score[1])
with open('history0.pkl', 'wb') as f:
          cPickle.dump(history0.history, f)
f.close()
# PYTHON CODE FOR DATA UTILITY
import numpy as np
import cv2
import os
import pandas as pd
from six.moves import cPickle
# For this problem the validation and test data provided by the concerned authority did not have labels, so the training data was
split into train, test and validation sets
train_dir = '/mnt/boneage-training-dataset/'
train_dir = 'C:/Users/INTuser/Desktop/Train/boneage-training-dataset/boneage-training-preprocess/'
X_train = []
y_age = []
y_gender = []
df = pd.read_csv('C:/Users/INTuser/Desktop/Train/boneage-training-dataset/train.csv') # Changed 'train.csv' by 'boneage-
training-dataset.csv'
# a = df.as_matrix()
m = df.shape[0]
path = train_dir
k = 0
print('Loading data set...')
for i in os.listdir(path):
   print(i)
y_age.append(df.boneage[df.id == int(i[:-4])].tolist()[0])
   a = df.male[df.id == int(i[:-4])].tolist()[0]
   if a:
y_gender.append(1)
   else:
y_gender.append(0)
img_path = path + i
img = cv2.imread(img_path)
img = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)
img = cv2.resize(img,(224,224))
   x = np.asarray(img, dtype=np.uint8)
X_train.append(x)
```

```
print('')
print('100% completed loading data')
# Save data
train_pkl = open('data.pkl','wb')
cPickle.dump(X_train, train_pkl, protocol=cPickle.HIGHEST_PROTOCOL)
train_pkl.close()
train_age_pkl = open('data_age.pkl','wb')
cPickle.dump(y_age, train_age_pkl, protocol=cPickle.HIGHEST_PROTOCOL)
train_age_pkl.close()
train_gender_pkl = open('data_gender.pkl','wb')
cPickle.dump(y_gender, train_gender_pkl, protocol=cPickle.HIGHEST_PROTOCOL)
train_gender_pkl.close()
### test
# For this problem the validation and test data provided by the concerned authority did not have labels, so the training data was split into train, test and validation sets
train_dir = '/mnt/boneage-training-dataset/'
train_dir = 'C:/Users/INTuser/Desktop/Train/boneage-training-dataset/boneage-test-preprocess/'
X_train = []
y_age = []
y_gender = []
df = pd.read_csv('C:/Users/INTuser/Desktop/Train/boneage-training-dataset/test.csv')
# a = df.as_matrix()
m = df.shape[0]
path = train_dir
k = 0
print('Loading data set...')
for i in os.listdir(path):
    print(i)
y_age.append(df.boneage[df.id == str(i[:-4])].tolist()[0])
    a = df.male[df.id == str(i[:-4])].tolist()[0]
    if a:
y_gender.append(1)
    else:
y_gender.append(0)
img_path = path + i
img = cv2.imread(img_path)
img = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)
img = cv2.resize(img,(224,224))
    x = np.asarray(img, dtype=np.uint8)
X_train.append(x)
  print('')
print('100% completed loading data')
# Save data
train_pkl = open('dataTest.pkl','wb')
cPickle.dump(X_train, train_pkl, protocol=cPickle.HIGHEST_PROTOCOL)
train_pkl.close()
train_age_pkl = open('data_ageTest.pkl','wb')
cPickle.dump(y_age, train_age_pkl, protocol=cPickle.HIGHEST_PROTOCOL)
train_age_pkl.close()
train_gender_pkl = open('data_genderTest.pkl','wb')
cPickle.dump(y_gender, train_gender_pkl, protocol=cPickle.HIGHEST_PROTOCOL)
train_gender_pkl.close()
#PYTHON CODE FOR TRAINING CHECK UPS FOR ALL TEST
# Check metrics using trained weight files
from keras.applications.inception_v3 import InceptionV3
from keras.preprocessing import image
```

```
from keras.models import Model
from keras.layers import Flatten, Dense, Input
from keras import backend as K
import _pickle as cPickle
import numpy as np
import matplotlib.pyplot as plt
import os
from keras.utils.vis_utils import plot_model
import cv2
import pandas as pd
from six.moves import cPickle as cPickle2
from scikitplot.metrics import plot_confusion_matrix, plot_roc
os.chdir('D:/Andualem file/boneage-training-dataset')
batch_size = 32
epochs = 30
# New test data
### test
# For this problem the validation and test data provided by the concerned authority did not have labels, so the training data was split into train, test and validation sets
#train_dir = '/mnt/boneage-training-dataset/'
#train_dir = 'D:/Andualem file/boneage-training-dataset/'
train_dir = 'D:/Andualem file/boneage-training-dataset/boneage-training-preprocess/'
#train_dir = 'D:/Andualem file/boneage-training-dataset/boneage-test-preprocess/'
#train_dir = 'D:/Andualem file/boneage-training-dataset/boneage-validation-dataset-2/'
#train_dir = 'D:/Andualem file/boneage-training-dataset/boneage-validation-dataset-2-preprocess/'
#train_dir = 'D:/Andualem file/boneage-training-dataset/Downloadboneage-test-preprocess/'
X_train = []
y_age = []
y_gender = []
#df = pd.read_csv('D:/Andualem file/boneage-training-dataset/test_with 14 physician info.csv') # Change 'train.csv' by 'boneage-training-dataset.csv'
#df = pd.read_csv('D:/Andualem file/boneage-training-dataset/test_with 14 claint info.csv') # Change 'train.csv' by 'boneage-training-dataset.csv'
df = pd.read_csv('D:/Andualem file/boneage-training-dataset/Validation Dataset.csv') # Change 'train.csv' by 'boneage-training-dataset.csv'
#df = pd.read_csv('D:/Andualem file/boneage-training-dataset/Dowloadboneage-test-dataset.csv') # no age 200 datasets
# a = df.as_matrix()
m = df.shape[0]
path = train_dir
# Load data
print('...loading training data')
f = open('data.pkl', 'rb')
x = cPickle.load(f)
f.close()
f = open('data_age.pkl', 'rb')
y = cPickle.load(f)
f.close()
x = np.asarray(x, dtype=np.float32)
y = np.asarray(y)
x /= 255.
x_final = []
y_final = []
# Shuffle images and split into train, validation and test sets
random_no = np.random.choice(x.shape[0], size=x.shape[0], replace=False)
for i in random_no:
x_final.append(x[i,:,:,:])
```

```
y_final.append(y[i])
x_final = np.asarray(x_final)
y_final = np.asarray(y_final)

k = 1261 # Decides split count                    # Change 100 to 1261
x_test = x_final[:k,:,:,:]
y_test = y_final[:k]
x_valid = x_final[k:2*k,:,:,:]
y_valid = y_final[k:2*k]
x_train = x_final[2*k:,:,:,:]
y_train = y_final[2*k:]
print('x_trainshape:'+ str(x_train.shape))
print('y_trainshape:'+ str(y_train.shape))
print('x_valid shape:'+ str(x_valid.shape))
print('y_validshape:'+ str(y_valid.shape))
print('x_test shape:'+ str(x_test.shape))
print('y_test shape:'+ str(y_test.shape))
base_model =  InceptionV3(weights='imagenet', include_top=False)
input = Input(shape=(224,224,3))
output_vgg16 = base_model(input)
x = Flatten()(output_vgg16)
x = Dense(512, activation='relu')(x)
predictions = Dense(1)(x)
model = Model(inputs=input, outputs=predictions)
model.compile(optimizer='adam', loss='mean_squared_error', metrics=['MAE'])
model.load_weights('model0.h5')
score = model.evaluate(x_test, y_test, batch_size=batch_size)
print('Test loss:', score[0])
print('Test MAE:', score[1])
print(model.summary())
#plot_model(model, to_file='model_plot01.png', show_shapes=True, show_layer_names=True)
# # plot and save confusion matrix
# fig, ax = plt.subplots(figsize=(16,12))
# plot_confusion_matrix(x_test, y_test, ax=ax)
# fig.savefig(os.path.join('performance_vizualizations', f'confusion_matrix_epoch_{epoch}'))
#############################################################################################
# test 2 1425 data
#############################################################################################
# New test data
### test
# For this problem the validation and test data provided by the concerned authority did not have labels, so the training data was split into train, test and validation sets
#train_dir = '/mnt/boneage-training-dataset/'
#train_dir = 'D:/Andualem file/boneage-training-dataset/boneage-test-preprocess/'
#train_dir = 'D:/Andualem file/boneage-training-dataset/boneage-test-dataset/'
#train_dir = 'D:/Andualem file/boneage-training-dataset/boneage-validation-dataset-2/'
train_dir = 'D:/Andualem file/boneage-training-dataset/boneage-validation-dataset-2/'
#train_dir = 'D:/Andualem file/boneage-training-dataset/boneage-validation-dataset-2-preprocess/'
#train_dir = 'D:/Andualem file/boneage-training-dataset/Downloadboneage-test-preprocess/'
X_train = []
y_age = []
y_gender = []
#df = pd.read_csv('D:/Andualem file/boneage-training-dataset/test_with 14 physician info.csv') # Change 'train.csv' by 'boneage-training-dataset.csv'
#df = pd.read_csv('D:/Andualem file/boneage-training-dataset/test_with 14 claint info.csv') # Change 'train.csv' by 'boneage-training-dataset.csv'
```

```
df = pd.read_csv('D:/Andualem file/boneage-training-dataset/Validation Dataset.csv') # Change 'train.csv' by 'boneage-training-dataset.csv'
#df = pd.read_csv('D:/Andualem file/boneage-training-dataset/Dowloadboneage-test-dataset.csv') # no age 200 datasets
# a = df.as_matrix()
m = df.shape[0]
path = train_dir
k = 0
print('Loading data set...')
for i in os.listdir(path):
    #print(i)
    #y_age.append(df.boneage[df.id == str(i[:-4])].tolist()[0])# For Our data set
    #a = df.male[df.id == str(i[:-4])].tolist()[0]
y_age.append(df.boneage[df.id == int(i[:-4])].tolist()[0]) # For the validation data type
    a = df.male[df.id == int(i[:-4])].tolist()[0]
    if a:
y_gender.append(1)
    else:
y_gender.append(0)
img_path = path + i
img = cv2.imread(img_path)
img = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)
img = cv2.resize(img,(224,224))
    x = np.asarray(img, dtype=np.uint8)
X_train.append(x)
print('')
print('100% completed loading data')
# Save data
train_pkl = open('dataTest.pkl','wb')
cPickle2.dump(X_train, train_pkl, protocol=cPickle2.HIGHEST_PROTOCOL)
train_pkl.close()
train_age_pkl = open('data_ageTest.pkl','wb')
cPickle2.dump(y_age, train_age_pkl, protocol=cPickle2.HIGHEST_PROTOCOL)
train_age_pkl.close()
train_gender_pkl = open('data_genderTest.pkl','wb')
cPickle2.dump(y_gender, train_gender_pkl, protocol=cPickle2.HIGHEST_PROTOCOL)
train_gender_pkl.close()
# End new test data
f = open('data.pkl', 'rb')
x = cPickle.load(f)
f.close()
f = open('data_age.pkl', 'rb')
y = cPickle.load(f)
f.close()
x = np.asarray(x, dtype=np.float32)
y = np.asarray(y)
x /= 255.
x_final = []
y_final = []
random_no = np.random.choice(x.shape[0], size=x.shape[0], replace=False)
for i in random_no:
x_final.append(x[i,:,:,:])
y_final.append(y[i])
x_final = np.asarray(x_final)
y_final = np.asarray(y_final)
k = 1261 # Decides split count                                    # Change 100 to 1261
x_valid = x_final[k:2*k,:,:,:]
```

```
y_valid = y_final[k:2*k]
x_train = x_final[2*k:,:,:,:]
y_train = y_final[2*k:]
# START importing own test data
f = open('dataTest.pkl', 'rb')
x = cPickle.load(f)
f.close()
f = open('data_ageTest.pkl', 'rb')
y = cPickle.load(f)
f.close()
x = np.asarray(x, dtype=np.float32)
y = np.asarray(y)
x /= 255.
x_final = []
y_final = []
random_no = np.random.choice(x.shape[0], size=x.shape[0], replace=False)
for i in random_no:
x_final.append(x[i,:,:,:])
y_final.append(y[i])
x_final = np.asarray(x_final)
y_final = np.asarray(y_final)
x_test = x_final[:,:,:,:]
y_test = y_final[:]
# END importing own test data
print('x_trainshape:'+ str(x_train.shape))
print('y_trainshape:'+ str(y_train.shape))
print('x_valid shape:'+ str(x_valid.shape))
print('y_validshape:'+ str(y_valid.shape))
print('x_test shape:'+ str(x_test.shape))
print('y_test shape:'+ str(y_test.shape))
base_model =  InceptionV3(weights='imagenet', include_top=False)
input = Input(shape=(224,224,3))
output_vgg16 = base_model(input)
x = Flatten()(output_vgg16)
x = Dense(512, activation='relu')(x)
predictions = Dense(1)(x)
model= Model(inputs=input, outputs=predictions)
model.compile(optimizer='adam', loss='mean_squared_error', metrics=['MAE'])
model.load_weights('model0.h5')
score = model.evaluate(x_test, y_test, batch_size=batch_size)
print('Test loss:', score[0])
print('Test MAE:', score[1])
print(model.summary())
#plot_model(model, to_file='model_plot02.png', show_shapes=True, show_layer_names=True)
##################################################################################################
# test 3 14 physitian
##################################################################################################
# New test data
### test
# For this problem the validation and test data provided by the concerned authority did not have labels, so the training data was split into train, test and validation sets
#train_dir = '/mnt/boneage-training-dataset/'
#train_dir = 'D:/Andualem file/boneage-training-dataset/boneage-test-preprocess/'
train_dir = 'D:/Andualem file/boneage-training-dataset/boneage-test-dataset/'
#train_dir = 'D:/Andualem file/boneage-training-dataset/boneage-validation-dataset-2/'
#train_dir = 'D:/Andualem file/boneage-training-dataset/boneage-validation-dataset-2-preprocess/'
```

```
#train_dir = 'D:/Andualem file/boneage-training-dataset/Downloadboneage-test-preprocess/'
X_train = []
y_age = []
y_gender = []
df = pd.read_csv('D:/Andualem file/boneage-training-dataset/test_with 14 physician info.csv') # Change 'train.csv' by 'boneage-training-dataset.csv'
#df = pd.read_csv('D:/Andualem file/boneage-training-dataset/test_with 14 claint info.csv') # Change 'train.csv' by 'boneage-training-dataset.csv'
#df = pd.read_csv('D:/Andualem file/boneage-training-dataset/Validation Dataset.csv') # Change 'train.csv' by 'boneage-training-dataset.csv'
#df = pd.read_csv('D:/Andualem file/boneage-training-dataset/Dowloadboneage-test-dataset.csv') # no age 200 datasets
# a = df.as_matrix()
m = df.shape[0]
path = train_dir
k = 0
print('Loading data set...')
for i in os.listdir(path):
    ##print(i)
    #y_age.append(df.boneage[df.id == str(i[:-4])].tolist()[0])# For Our data set
    #a = df.male[df.id == str(i[:-4])].tolist()[0]
y_age.append(df.boneage[df.id == str(i[:-4])].tolist()[0]) # For the validation data type
    a = df.male[df.id == str(i[:-4])].tolist()[0]
    if a:
y_gender.append(1)
    else:
y_gender.append(0)
img_path = path + i
img = cv2.imread(img_path)
img = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)
img = cv2.resize(img,(224,224))
    x = np.asarray(img, dtype=np.uint8)
X_train.append(x)
print('')
print('100% completed loading data')
# Save data
train_pkl = open('dataTest.pkl','wb')
cPickle2.dump(X_train, train_pkl, protocol=cPickle2.HIGHEST_PROTOCOL)
train_pkl.close()
train_age_pkl = open('data_ageTest.pkl','wb')
cPickle2.dump(y_age, train_age_pkl, protocol=cPickle2.HIGHEST_PROTOCOL)
train_age_pkl.close()
train_gender_pkl = open('data_genderTest.pkl','wb')
cPickle2.dump(y_gender, train_gender_pkl, protocol=cPickle2.HIGHEST_PROTOCOL)
train_gender_pkl.close()
# End new test data
f = open('data.pkl', 'rb')
x = cPickle.load(f)
f.close()
f = open('data_age.pkl', 'rb')
y = cPickle.load(f)
f.close()
x = np.asarray(x, dtype=np.float32)
y = np.asarray(y)
x /= 255.
x_final = []
y_final = []
```

```
random_no = np.random.choice(x.shape[0], size=x.shape[0], replace=False)
for i in random_no:
x_final.append(x[i,:,:,:])
y_final.append(y[i])
x_final = np.asarray(x_final)
y_final = np.asarray(y_final)
k = 1261 # Decides split count                                    # Change 100 to 1261
x_valid = x_final[k:2*k,:,:,:]
y_valid = y_final[k:2*k]
x_train = x_final[2*k:,:,:,:]
y_train = y_final[2*k:]
# START importing own test data
f = open('dataTest.pkl', 'rb')
x = cPickle.load(f)
f.close()
f = open('data_ageTest.pkl', 'rb')
y = cPickle.load(f)
f.close()
x = np.asarray(x, dtype=np.float32)
y = np.asarray(y)
x /= 255.
x_final = []
y_final = []
random_no = np.random.choice(x.shape[0], size=x.shape[0], replace=False)
for i in random_no:
x_final.append(x[i,:,:,:])
y_final.append(y[i])
x_final = np.asarray(x_final)
y_final = np.asarray(y_final)
x_test = x_final[:,:,:,:]
y_test = y_final[:]
# END importing own test data
print('x_trainshape:'+ str(x_train.shape))
print('y_trainshape:'+ str(y_train.shape))
print('x_valid shape:'+ str(x_valid.shape))
print('y_validshape:'+ str(y_valid.shape))
print('x_test shape:'+ str(x_test.shape))
print('y_test shape:'+ str(y_test.shape))
base_model =  InceptionV3(weights='imagenet', include_top=False)
input = Input(shape=(224,224,3))
output_vgg16 = base_model(input)
x = Flatten()(output_vgg16)
x = Dense(512, activation='relu')(x)
predictions = Dense(1)(x)
model= Model(inputs=input, outputs=predictions)
model.compile(optimizer='adam', loss='mean_squared_error', metrics=['MAE'])
model.load_weights('model0.h5')
score = model.evaluate(x_test, y_test, batch_size=batch_size)
print('Test loss:', score[0])
print('Test MAE:', score[1])
print(model.summary())
#plot_model(model, to_file='model_plot03.png', show_shapes=True, show_layer_names=True)
##################################################################################################
# test 4 14 client
##################################################################################################
# New test data
```

```
### test
# For this problem the validation and test data provided by the concerned authority did not have labels, so the training data was split into train, test and validation sets
#train_dir = '/mnt/boneage-training-dataset/'
#train_dir = 'D:/Andualem file/boneage-training-dataset/boneage-test-preprocess/'
train_dir = 'D:/Andualem file/boneage-training-dataset/boneage-test-dataset/'
#train_dir = 'D:/Andualem file/boneage-training-dataset/boneage-validation-dataset-2/'
#train_dir = 'D:/Andualem file/boneage-training-dataset/boneage-validation-dataset-2-preprocess/'
#train_dir = 'D:/Andualem file/boneage-training-dataset/Downloadboneage-test-preprocess/'
X_train = []
y_age = []
y_gender = []
#df = pd.read_csv('D:/Andualem file/boneage-training-dataset/test_with 14 physician info.csv') # Change 'train.csv' by 'boneage-training-dataset.csv'
df = pd.read_csv('D:/Andualem file/boneage-training-dataset/test_with 14 claint info.csv') # Change 'train.csv' by 'boneage-training-dataset.csv'
#df = pd.read_csv('D:/Andualem file/boneage-training-dataset/Validation Dataset.csv') # Change 'train.csv' by 'boneage-training-dataset.csv'
#df = pd.read_csv('D:/Andualem file/boneage-training-dataset/Dowloadboneage-test-dataset.csv') # no age 200 datasets
# a = df.as_matrix()
m = df.shape[0]
path = train_dir
k = 0
print('Loading data set...')
for i in os.listdir(path):
    #print(i)
    #y_age.append(df.boneage[df.id == str(i[:-4])].tolist()[0])# For Our data set
    #a = df.male[df.id == str(i[:-4])].tolist()[0]
y_age.append(df.boneage[df.id == str(i[:-4])].tolist()[0]) # For the validation data type
    a = df.male[df.id == str(i[:-4])].tolist()[0]
    if a:
y_gender.append(1)
    else:
y_gender.append(0)
img_path = path + i
img = cv2.imread(img_path)
img = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)
img = cv2.resize(img,(224,224))
    x = np.asarray(img, dtype=np.uint8)
X_train.append(x)
print('')
print('100% completed loading data')
# Save data
train_pkl = open('dataTest.pkl','wb')
cPickle2.dump(X_train, train_pkl, protocol=cPickle2.HIGHEST_PROTOCOL)
train_pkl.close()
train_age_pkl = open('data_ageTest.pkl','wb')
cPickle2.dump(y_age, train_age_pkl, protocol=cPickle2.HIGHEST_PROTOCOL)
train_age_pkl.close()
train_gender_pkl = open('data_genderTest.pkl','wb')
cPickle2.dump(y_gender, train_gender_pkl, protocol=cPickle2.HIGHEST_PROTOCOL)
train_gender_pkl.close()
# End new test data
f = open('data.pkl', 'rb')
x = cPickle.load(f)
f.close()
```

```
f = open('data_age.pkl', 'rb')
y = cPickle.load(f)
f.close()
x = np.asarray(x, dtype=np.float32)
y = np.asarray(y)
x /= 255.
x_final = []
y_final = []
random_no = np.random.choice(x.shape[0], size=x.shape[0], replace=False)
for i in random_no:
x_final.append(x[i,:,:,:])
y_final.append(y[i])
x_final = np.asarray(x_final)
y_final = np.asarray(y_final)
k = 1261 # Decides split count
x_valid = x_final[k:2*k,:,:,:]
y_valid = y_final[k:2*k]
x_train = x_final[2*k:,:,:,:]
y_train = y_final[2*k:]
# START importing own test data
f = open('dataTest.pkl', 'rb')
x = cPickle.load(f)
f.close()
f = open('data_ageTest.pkl', 'rb')
y = cPickle.load(f)
f.close()
x = np.asarray(x, dtype=np.float32)
y = np.asarray(y)
x /= 255.
x_final = []
y_final = []
random_no = np.random.choice(x.shape[0], size=x.shape[0], replace=False)
for i in random_no:
x_final.append(x[i,:,:,:])
y_final.append(y[i])
x_final = np.asarray(x_final)
y_final = np.asarray(y_final)
x_test = x_final[:,:,:,:]
y_test = y_final[:]
# END importing own test data
print('x_trainshape:'+ str(x_train.shape))
print('y_trainshape:'+ str(y_train.shape))
print('x_valid shape:'+ str(x_valid.shape))
print('y_validshape:'+ str(y_valid.shape))
print('x_test shape:'+ str(x_test.shape))
print('y_test shape:'+ str(y_test.shape))
base_model =  InceptionV3(weights='imagenet', include_top=False)
input = Input(shape=(224,224,3))
output_vgg16 = base_model(input)
x = Flatten()(output_vgg16)
x = Dense(512, activation='relu')(x)
predictions = Dense(1)(x)
model= Model(inputs=input, outputs=predictions)
model.compile(optimizer='adam', loss='mean_squared_error', metrics=['MAE'])
model.load_weights('model0.h5')
score = model.evaluate(x_test, y_test, batch_size=batch_size)
```

```
print('Test loss:', score[0])
print('Test MAE:', score[1])
print(model.summary())
#plot_model(model, to_file='model_plot04.png', show_shapes=True, show_layer_names=True)
##################################################################################################
```