MOHAMAD TALAL RADWAN

ARTIFICIAL INTELLIGENCE-BASED ALGORITHM FOR CERVICAL VERTEBRAE

PhD THESIS

2022

NEAR EAST UNIVERSITY

INSTITUTE OF GRADUATE STUDIES

DEPARTMENT OF ORTHODONTICS

ARTIFICIAL INTELLIGENCE-BASED ALGORITHM FOR CERVICAL VERTEBRAE MATURATION STAGE ASSESSMENT

PhD. THESIS

MOHAMAD TALAL RADWAN

Nicosia

NOVEMBER, 2022

NEAR EAST UNIVERSITY
INSTITUTE OF GRADUATE STUDIES
DEPARTMENT OF ORTHODONTICS

ARTIFICIAL INTELLIGENCE-BASED ALGORITHM FOR CERVICAL VERTEBRAE MATURATION STAGE ASSESSMENT

PhD. THESIS

MOHAMAD TALAL RADWAN

Supervisor
Assoc. Prof. Dr. LEVENT VAHDETTİN

Nicosia
November, 2022

Approval

We certify that we have read the thesis submitted by Mohamad Talal RADWAN titled "**Artificial intelligence-based algorithm for cervical vertebrae maturation stage assessment "**and that in our combined opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Educational Sciences.

Examining Committee	Name-Surname	Signature
Head of the Committee:	Assist. Prof. Dr. Beste KAMİLOĞLU	
Committee Member*:	PROF. DR. MEHMET ÖZGÜR SAYIN	
Committee Member*:	Prof. Dr. Ulaş Öz	
Committee Member*:	Assoc. Prof. Dr. Seçil AKSOY	
Supervisor:	Assoc. Prof. Dr. Levent VAHDETTİN	

Approved by the Head of the Department

7/11/2022

Assist. Prof. Dr. Beste KAMİLOĞLU

V/E Head of Department

Yrd.Doç.Dr. Beste Kamiloğlu
Ortodonti Ana Bilim Dalı

Approved by the Institute of Graduate Studies

7/11/2022

Prof. Dr. Kemal Hüsnü Can Başer

Head of the Institute

Declaration

I hereby declare that all information, documents, analysis and results in this thesis have been collected and presented according to the academic rules and ethical guidelines of Institute of Graduate Studies, Near East University. I also declare that as required by these rules and conduct, I have fully cited and referenced information and data that are not original to this study.

Mohamad Talal RADWAN

7/11/2022

Acknowledgments

First and foremost, praises and thanks to ALLAH, the Almighty, for his showers of blessings throughout my research work to complete the research successfully and for all the blessing in my life.

Secondly, I would like to thank my parents for their love and support, without them, this day would not have been possible. I would also like to thank my entire family (Lina, Mamdouh, Zelal and Layal) for the good times in my life and for not allowing me to give up when the chips were down.

I would like to thank my supervisor Dr. LeventVahdettin for his consistent support and guidance during the My gratitude to you for all you have done. I truly appreciate you and your time you spent helping me in many occasions.

I would like to thank also my precious Hoca Beste Kamiloglu for guiding me through my journey, my dear hocam this would have not been possible without you.

I would like also to thank Dr. Ulas for teaching me to trust myself and never be afraid of making mistakes, thank you for everything hocam.

Dr. Zahir Altug it was my pleasure to be your student one day, learning from you was exceptional.

I would like also to thank many friends that became family during my journey (Aseel, Beren, Yaman, Mohammad Alslakhi, Cagla, Deniz, Ismet, Serife, Khalil, Hikmet) my second family, thank you for believing in me and thanks for being in my life.

Lastly but not least I would like to thank the jury members being with me in my last journey step.

Mohamad Talal RADWAN

Abstract

Artificial intelligence-based algorithm for cervical vertebrae maturation stage assessment

Mohamad Talal RADWAN

PhD, Department of Orthodontics

November,2022, 72 pages

The Aim of our study is to develop an AI based algorithm to automatically and accurately determine the stage of Cervical vertebra Maturation CVM mainly to aid in making the diagnosis procedure less time consuming and more accurate.

1505 Lateral cephalometric that meet the criteria will be used in the training process, labelling will be done using Computer Vision Annotation Tool (CVAT), Maturation stage classification will be according to Bachetti method by, the labelled data will be split into a training group (80%), a testing group (10%) and a validation group (10%), to measure reliability intra-examiner (ICC), inter-examiner test was carried out.

Intraclass Correlation Coefficient (ICC) was 0.981 which shows high reliability of the examiner, the segmentation network achieved a global accuracy of 0.99 and average dice score overall images was 0.93. The classification network achieved an accuracy of 0.802, class sensitivity of (pre-pubertal 0.78) (pubertal 0.45) (post-pubertal 0.98) respectively per class specificity of (pre-pubertal 0.94), (pubertal 0.94), (post-pubertal 0.75) respectively.

Our developed algorithm showed ability, accuracy and reliability to detect Vertebras and determine Cervical vertebrae maturation stage, which will aid in more accurate Diagnosis process by eliminating human factor which in turn could lead to wrong decision-making procedure which might danger the outcome of the treatment plan.

Key Words: Artificial intelligence, Orthodontics, Deep learning, Cervical vertebras.

Özet

Servikal vertebra maturasyon aşamasının tespiti için yapay zeka tabanlı algoritma

Mohamad Talal RADWAN

Doktora, Ortodonti Anabilim Dalı

Kasım 2022, 72 sayfa

Çalışmamızın amacı, Servikal Vertebra Maturasyon (SVM) aşamasının teşhisini daha hızlı, dogru ve otomatik bir şekilde belirlemek için yapay zeka tabanlı bir algoritma geliştirmektir.

Bu retrospektif calismada kriterleri karşılayan 1505 lateral sefalometrik film kullanılmıştır. Filmlerdeki maturasyon aşamaları Computer Vision Annotation Tool (CVAT) programi kullanılarak etiketleme yapılmıştır. Maturasyon aşaması sınıflandırması Bachetti yöntemine göre yapılarak etiketlenen veriler eğitim grubu (%80), test grubu (%10) ve doğrulama grubu (%10) olmak üzere üç gruba bölünmüştür. Güvenilirliği ölçmek için gözlemci içi (ICC) ve gözlemciler arası ölçümler yapılmıştır.

Segmentasyon ağının doğruluğu 0,99, Dice benzerlik katsayısı 0,93 bulunmuştur. Sınıflandırma ağında 0,802 doğruluk tespit edilmiştir. Sınıfların hassasiyet değerleri ergenlik öncesi 0,78, ergenlik 0,45, ergenlik sonrası 0,98, özgüllük değerleri ise ergenlik öncesi 0,94, ergenlik 0,94, ergenlik sonrası 0,75 bulunmuştur. Gözlemci içi güvenilirlik değeri 0,981 olarak tespit edilmiştir.

Geliştirilen algoritmamız, servikal vertebraları tespit etme ve maturasyon aşamasını belirlemede yüksek doğruluk ve güvenilirlik göstermistir. Bu da, planlamada tedavinin sonucunu etkileyen yanlış kararlar vermeye yol açabilecek insan faktörünü ortadan kaldırarak daha doğru ve hızlı bir teşhis sürecine yardımcı olacaktır.

Anahtar Kelimeler: Yapay zeka, Ortodonti, Derin öğrenme, Servikal vertebral maturasyon.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

CVM:	Cervical Vertebra Maturation
LC:	Lateral Cephalometric
AI:	Artificial intelligence
DR:	Digital Radiography
CVAT:	Computer Vision Annotation Tool
HWR:	Hand-Wrist Radiographs
CVMS:	Cervical Vertebrae Maturation Stages
ANN:	Artificial Neural Networks
CNN:	Convolutional Neural Network
ICC:	Intraclass correlation coefficient
WK:	Weighted Kappa
CK:	Cohen's Kappa

CHAPTER I

1. Introduction

The best timing of orthodontic treatment has been the topic of discussion among orthodontist for a long period of time, with most of practitioners used to prefer the early treatment approach, nowadays the most preferred period for orthodontic intervention tend to be early permeant dentition or late mixed dentition depending on the specifics of each case (Fleming, 2017).

Treating in the early permeant dentition or late mixed dentition have the upper hand for being the period of growth spurt (maximal growth) allowing effective correction of skeletal and dental anomalies (DiBiase, 2002; Fleming,2008).

Growth modification is a type of orthodontic treatment meant to take advantage of the patient ongoing growth in order to correct any skeletal anomalies by redirecting the growth in a more favourable direction, in order for the growth modification treatment to be a success it has to be matched with the period of maximal growth also known as growth spurt (De Clerck, H. J., & Proffit, 2015).

Determining and utilizing the growth spurt period will ensure a success growth modification treatment, throughout the last few decades a number of indicators were proposed to determine the growth spurt with skeletal age (bone age) proved to be the most accurate and reliable method (Szemraj-Folmer, et al.,2021).

Skeletal age (bone age) is widely used method in medical field specially in orthopaedic, forensic medicine and dentistry, to this day two technique can be used to asses skeletal age (bone age), hand and wrist and cervical vertebra maturation (CVM) technique, with hand and wrist technique proves to be the most accurate but with a disadvantage of requiring extra radiation dose (Hoseini, et al.,2016).

Since its introduction to orthodontics in 1931, lateral cephalometric (LC) has been a crucial tool of the orthodontic diagnosis process, the cervical vertebras as seen on routinely taken lateral cephalograms, have been used to determine the skeletal maturity, it is well known that with continuous growth cervical vertebral bodies experience some morphological changes.

Lamparski in 1972 described that cervical vertebra were a reliable technique in assessing skeletal age to a level of reliability similar to that of hand-wrist technique. Lamparski in his study described six stages of maturation according to the morphological changes to the vertebras, during the up followed years CVM method have been the subjected to modification in order to simplify the technique, with the most common modification among orthodontist being Baccetti modification in 2002, which demonstrate 6 definitive stages of maturation according to the changes C2, C3 and C4 vertebra (Lamparski, 1972; Baccetti, et al., 2005).

Throughout history, researchers and technologists have been fascinated by the human brain., with the continuous Industrial Revolution and with the changes in human behaviour nowadays, automation concurred every aspect of our life to enhances the speed, precision and effectiveness of human efforts (Chartrand, et al., 2017).

Artificial intelligence (AI) can be defined by the computer ability to perform tasks that usually require human intelligence. The history of artificial intelligence may be traced back to Leonardo da Vinci sketches and ancient Greek machines (Antikythera mechanism).
AI was first mentioned in the 1950s but it didn’t show any success, in recent years AI reappeared and started to show some promising results (Chartrand, et al., 2017).

AI has already invaded every element of our daily lives through smartphones and automobiles, and more recently, the healthcare field. In the past decade, AI has been effectively implemented in radiology, with a high level of success in detecting cancer in its early stages.

Despite its success in other fields, AI in dentistry has remained an outsider until recently. Early research into AI in dentistry showed promise in spotting caries using intra-oral X-ray.

The implementation of AI in orthodontics started with applying an AI algorithm to identify cephalometric landmarks to be utilized in cephalometric analysis to apply in orthodontic radiological assessment procedure (Kunz, et al., 2020).

CHAPTER II

2. Literature Review

2.1 X-Ray evolution:

William Conrad Roentgen was a professor and physicist at the University of Wurzburg in Germany when he discovered x-rays. In the fall of 1895, he was performing an experiment in which he discharged high voltage into a partly evacuated Crookes tube. The tube he was using had a black cardboard cover, and he noticed that whenever he switched on the Crookes tube, that there would be a faint glow from a nearby little screen. This discovery of the glowing effect prompted further investigation into the phenomenon (Assmus, 1995).

He couldn't figure out how the glowing occurred since he couldn't see, hear, or feel anything that may be producing it. And, to add to Roentgen's fascination with this phenomenon, the Crookes tube was covered in black cardboard, ensuring that no light was escaping out. As the experiments continued, he noticed that the screen was covered with a paint containing a substance called barium platinocyanide. He made the fascinating observation that the closer he pushed the screen to the Tube, the brighter the lighting became. Despite the black cardboard covering the Crookes tube, he could see that something invisible was stimulating the coated screen. He also discovered that the screen continued to illuminate even when he placed books between the tube and the screen. He realized that this was most likely some form of "ray" or "particles" that were able to pass through the thickness of the books and light up the screen. With this finding, he initiated a series of experiments in which he would place his hand and other objects between the Crookes tube and the painted screen, and he could see a shadow of his hand's bones (Assmus, 1995).

He enlisted his wife's help, although there was little detail in this initial image, it marked the beginning of radiography.

2.2 Cephalometric evolution:

Orthodontics began the age of cephalometry in 1931, based on historical studies that brought to the orthodontic community the cephalostat, a device that allows the patient's head to be positioned in the same position at all times. It was feasible to get serial radiographs with the cephalostat, allowing for more precise examinations of human face development. The era of the 1930s is considered a landmark for Orthodontics because of the evolution of the specialty as science (Silva & Sant'Anna, 2013).

The era of computed cephalometric radiography began in the late 1960s. Different programs that compute distances and angles of the cephalometric tracing have been developed thanks to technical advancements in data processing, minimizing the amount of manual effort required for studies and, as a result, speeding up research projects that require cephalometric assessment (Silva & Sant'Anna, 2013).

2.3 Xray digitization:

The idea of digital radiography was first proposed in the late 1970s and early 1980s, but the technology was not ready at the time. With the introduction of digital imaging, a new era in radiology began: the use of digital data instead of traditional films has provided new conveniences for radiograph manipulation, transmission, storage, and viewing.

The digital representation of radiographs has a variety of benefits for the health system, including the demonstrated efficiency of internet-based second opinion systems and even the use of such digital material in education and training activities. In order to fully benefit from the advantages of digital radiography (DR), in addition to the DR itself, which has a far greater deployment cost than traditional radiography, One option is to digitize traditional radiographic films (Chen & Hollender,1995). There are presently two techniques for digitizing radiographic films: one is via the use of digital photography, and the other is through the use of a transparency scanner (Rubia- Bulle,et al .,2007).

2.4 Growth and Development:

Determining this period of growth is important for orthodontic applications. Each growth spurt has a distinct onset, acceleration phase, peak period, deceleration phase, and termination phase of the spurt. It starts earlier in women and lasts on average 3-4 years, and 4-5 years in men.

Somatic maturation is indicated by height and weight gain throughout growth and development. Height measures can show how the skeletal system is developing overall.

The onset of the breakthrough in height increase is observed at the age of 10 in women and 12 in men. The fastest growth (peak) period is seen in the following years for both sexes (Korde, et al., 2015).

Sexual maturation is determined by evaluating the secondary sex characteristics of individuals. Tanner defined sexual maturation as varying degrees from 1 (least mature, preadolescent) to 5 (most mature, adult), with different degrees of sexual maturity in males and females.

While determining the “Tanner Sexual Maturity Scale”, pubic hair growth (amount, thickness, color and location), penile length and width, scrotal development and testicular length in men, breast development (size and morphology) and pubic hair (position, color, morphology and quantity in women) is evaluated. Evaluation of regular audio recordings can be used in boys but is impractical. Menstrual pain is an important sign of maturation in women.

Skeletal maturation is determined by the size, shape and degree of ossification of bones in areas such as the feet, knees, elbows, shoulders, hips and neck. Age estimation can be made from the changes observed in the bone development process. Ossification centre’s appearing on radiography, their size and shape; The width and shape of the cartilage structure and the degree of fusion between the diaphysis and epiphyses are characteristic features evaluated on radiography. With radiographs taken at certain periods, successive ossifications in bones, changes in the shape and size of the relevant bones, together with their old images, can be evaluated visually to determine bone age.

Bone age is an objective indicator in determining the degree of maturation of the individual and is reliable and effective (Baldin, et al.,2017; Korde, et al., 2015).

2.5 Skeletal age:

The calendar day on which a person was born is used to calculate chronological age. Chronological age can be used to evaluate a child's mental maturity, physical capabilities, height, and number of teeth in the mouth, however physiological variations can be observed even among individuals of the same race, gender, and age.

Since there are significant differences between growth spurt periods and chronological age, chronological age is not considered a reliable indicator in determining an individual's level of maturation, and this assessment can be made not only by chronological age, but also by biological and physiological age.

Skeletal age is based on determining the developmental level of the child, and it play an important role in detecting deviations from normal development, planning treatment, determining the appropriate time to start treatment and prognosis.

According to their biological clock, each person grows at a distinctive rate. The process of development/maturation in each individual is characterized as biological age. The degree of maturity of the various tissue systems determines an individual's physiological age.

Information such as height and weight of systems such as the somatic, reproductive, and skeletal systems, as well as maturation phases of tissues such as teeth, are examined separately or together in the biological age concept.

Some maturity indicators include height, weight, chronological age, sexual maturation, frontal sinus, biological and physiological age, cervical vertebrae, tooth eruption, tooth calcification grades, and biomarkers (Baccetti, et al., 2005).

Evaluation of the growth processes of individuals is necessary in medicine and dentistry, in the diagnosis of endocrinological or orthopedic disorders.

Bone age determination is commonly employed in the departments of pedodontics and orthodontics, pediatrics, orthopedics and forensic medicine.

2.5.1. Skeletal Maturation Indicators:

In order for a biological indicator to be employed in a clinical setting, it is required to meet a specific standard, it should be reliable, simple, valid for both sexes, and parallels the development of facial bones. However, there is no single biological indicator that meets all these criteria (Kasimoglu & Tuna-Ince, 2016).

Indicators evaluated in determining the skeletal maturation of individuals can be classified as follows:

A. Radiological
    1. Special radiographs:
        - Use of hand and wrist films: This is the most commonly used and accepted method.
    2. Lateral cephalograms:
        - Use of cervical vertebrae in lateral cephalograms
        - Use of the frontal sinus in lateral cephalograms
    3. Orthopantomography / intraoral periapical:
        - Evaluation of different levels of tooth development

B. Biochemical

In this method, saliva and serum can be checked:

1. Insulin-like growth factor (IGF, insulin-like growth factor) growth hormone (GH, growth hormone).
2. Creatine.
3. Alkaline phosphatase (ALP) (Korde et al., 2015)

An “ideal” growth indicator to be used in the diagnosis and treatment planning process should be characterized as accurate, easy to register, consistent, predictable and doesn’t require any extra radiation.

2.5.1.1 Radiological Indicators:

Bone age can be determined by radiographic evaluation of the developmental status of some bones of the skeletal system and various structures seen in the ossification centers. Although imaging of the wrist bones is a useful and accepted method for this purpose, it should not be forgotten that the patient is exposed to extra radiation. Evaluation of cervical vertebra or frontal sinus in lateral cephalograms and evaluation of tooth development and eruption status in orthopantomography, intraoral periapical images are other radiological methods (Baldin et al., 2017).

2.5.1.1.1 Hand-Wrist radiographs (HWR):

Hand-Wrist radiographs (HWR) assessment is considered the most standardized method for skeletal measurements. The skeletal development index is determined according to the sequential changes observed in the carpal bones and some ossification events.

Among these methods used to determine skeletal age, Greulich-Pyle atlas method compares the individual's radiograph with certain images, while in Tanner and Whitehouse method, certain ossification centers (radius, ulna, and certain short metacarpals and phalanges) in the hand and wrist are evaluated. In general, the degree of ossification of the phalanges is determined according to the following definitions:

- Level 1: Epiphyseal diaphysis equality (Equality sign, =)

- Level 2: The pineal covers the diaphysis as if wearing a cap (caping)
- Level 3: Fusion of epiphysis and diaphysis

2.5.1.1.2. Panoramic/Periapical Imaging:

Dental maturation can be determined according to the eruption stage or formation level of the teeth. This evaluation is advantageous in that it is performed with panoramic radiographs taken routinely from patients, it is a simple procedure, and requires minimal radiation dose.

Since eruption of teeth can be affected by factors such as local factors, systemic diseases and nutritional deficiency, it is suggested that the level of tooth formation is more reliable than eruption stage.

There are several studies reporting correlations with skeletal age. The mineralization level of the canine tooth is more compatible than other teeth. There are methods reported by different researchers such as Demirjian, Chertkow and Fatti, Nolla and Goldstein and Tanner (Koçak, 2018; Korde et al., 2015).

2.5.1.1.3. Cervical Vertebrae Maturation (CVM):

The CVM indicator was described by Lamparski in 1972. The primary purpose of this method is to determine the skeletal maturation periods of orthodontic patients with cephalometric radiographs taken routinely at the beginning. Six different CVM levels were defined by evaluating C2-C6 vertebrae according to the CVM index. The summary of the method is as follows (Kasimoglu & Tuna-Ince, 2016; Korde et al., 2015):

I. Level 1 / Category 1
   a. C2, C3 and C4 bodies have flat bottom edges.
   b. The upper edges are markedly angled from back to front.
   c. A very serious amount of adolescent growth is expected.

II. Level 2 / Category 2

a. Concavity develops at the lower border of C2 and C3.
b. The leading-edge length of the stems has increased.
c. The shape of C3 and C4 resembles a rectangle.
d. Significant adolescent growth is expected.

III. Level 3 / Category 3
a. The lower concavities of C2 and C3 become prominent.
b. Concavity begins to form at the lower edge of C4.
c. Moderate growth is expected

IV. Level 4 / Category 4
a. The shape of C3 and C4 is almost square.
b. The concavities below C2, C3 and C4 become prominent.
c. Little growth is expected.

V. Level 5 / Category 5
a. The shape of C3 and C4 is square.
b. The concavities of C2, C3 and C4 become more prominent.
c. A vague amount of growth is expected.

VI. Level 6 / Category 6
a. The height of C3 and C4 increases and becomes larger than its width.
b. All concavities deepen.
c. Adolescent growth is completed (Korde et al., 2015).

There are six different categories in the method reported by Hassel and Farman in 1995:

I. Category 1 (Preparation Phase): Adolescent growth has just begun and approximately 80% to 100% of growth remains. The lower edges of C2, C3, and C4 are straight. The vertebrae are wedge-shaped and their upper edges are angled from back to front.

II. Category 2 (Acceleration Phase): Growth activity increases, leaving about 65% to 85% growth. Concavities begin to develop on the lower edges of C2 and C3, while the lower edge of C4 is straight. The bodies of C3 and C4 are almost rectangular.

III. Category 3 (Transition Phase): Adolescent growth still continues to accelerate, reaching its highest level, with a growth of between 25% and 65% expected behind. Concavities become evident in the lower margins of C2 and C3, while they begin to form at the lower margins of C4. The body of C3 and C4 is rectangular.

IV. Category 4 (Slowing Phase): Adolescent growth slows down dramatically at this stage and growth is expected to be around 10% to 25%. Below C2, C3 and C4 there are prominent concavities and the bodies of C3 and C4 begin to square.

V. Category 5 (Maturation Stage): The maturation of the vertebrae is almost complete with this stage, a growth of between 5% and 10% is expected.

VI. Category 6 (Completion Stage): Growth is considered complete at this stage. Little or no growth is expected. Deep concavities are observed below C2, C3 and C4. C3 and C4 have a square body shape or a vertical height greater than a square (Hassel & Farman, 1995; Korde et al., 2015)

Bacetti et al., (2002), published a method evaluating skeletal maturation in 5 stages, he updated his method in 2005 and defined 6 different cervical vertebrae maturation stages (CVMS) according to the following criteria:

I. CVMS 1 - The lower edges of all three vertebrae (C2-C4) are straight. The body of C3 and C4 is trapezoidal (upper edge angled from back to front). Mandibular growth spurt is 2 years away.

II. CVMS 2 -There is a concavity at the lower edge of cervical level 2- C2. The bodies of C3 and C4 are still trapezoidal. There is 1 year left to the mandibular growth spurt.

III. CVMS 3- There are concavities at the lower edge of C2 and C3. The body of C3 and C4 can be trapezoidal or horizontally rectangular. The breakthrough of mandibular growth takes place in the next year after this period.

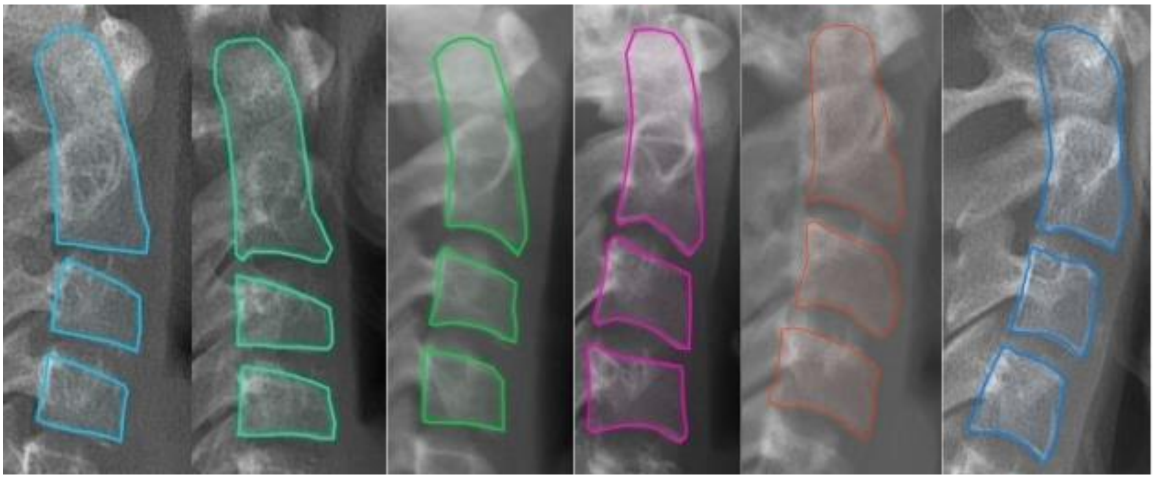
IV. CVMS 4 - C2, C3, and C4 have concavities at their lower margins. Both C3 and C4 vertebral bodies are horizontal rectangles. The breakthrough in mandibular growth was experienced a year or two before this period.

V. CVMS 5 – Concavities below C2, C3 and C4 are still present. At least one of the stems of C3 or C4 is square. If not square, the other vertebral body is horizontal rectangle. The breakthrough in mandibular growth ended at least one year before this period.

VI. CVMS 6- Concavities below C2, C3 and C4 are still present. The body of at least one of C3 or C4 is a vertical rectangle. If the vertical is not rectangular, the other vertebra is square. The breakthrough in mandibular growth ended at least two years before this period (Baccetti et al., 2005).

Figure 1.

Representation of the Stages of Cervical Vertebrae According to Bachetti Method.



Studies describing the evaluation of CVM on lateral cephalometric radiographs and their brief descriptions are in Table 1 (Baccetti et al., 2005; Baccetti et al., 2002; Caldas, Ambrosano, & Haiter Neto, 2010; Caldas, Ambrosano, & Haiter Neto, 2007; Hassel & Farman, 1995; Koukouviti, 2017; Seedat & Forsberg, 2005).

Table 1.

Brief Descriptions of CVM History.

Method	Year	Definition
Lamparski	1972	By evaluating the morphology (size and shape) of C2-C6 vertebrae, 6 different maturation degrees are determined.
Hassel & Farman	1995	6 different maturation levels by evaluating C2-C4 vertebra morphology.
Baccetti et al.	2002	5 different maturation levels by evaluating C2-C4 vertebra morphology.
Baccetti et al.	2005	6 different maturation levels by evaluating C2-C4 vertebra morphology.
Seedat et al.	2005	The formation of C3 is analyzed
Caldas et al.	2007 & 2010	Horizontal and vertical length measurements on C3-C4 vertebrae.

2.5.1.1.4 Frontal Sinus Development:

The development of the frontal sinus shows parallelism with the growth rhythm of the body and attacks during puberty. On lateral cephalometric radiographs, frontal sinus enlargement is closely related to the individual's height increase and reaches its greatest extent one year after the peak of height increase.

Researchers stated that frontal sinus development can be used as an indicator of maturation, but they added that more studies should be done on this subject (Kasimoglu & Tuna-Ince, 2016).

2.6 Artificial Intelligence (AI):

The human brain has attracted academics and scientists throughout history. Since antiquity, philosophers and scientists have tried to define the process of intellect and decision-making.
Artificial intelligence is defined differently in different disciplines, and some researchers believe it may be traced back to Leonardo da Vinci drawings and ancient Greek technologies (Antikythera mechanism) (Chartrand, et al., 2017).

In today's literature, AI is described as "science and engineering for the building of intelligent machines" .This definition covers the time period from the invention of the first usable computer after World War II through John McCarthy's first usage of the phrase "artificial intelligence" in 1956. (Chartrand, et al., 2017).

The logic of Aristotle is the foundation for today's computers and technology. Alan Turing, an English mathematician, made the first breakthrough in the history of supercomputers in 1950, when he built a system that could decipher encrypted messages.

The Turing Test, which examines whether a computer possesses intelligence by assessing how effectively it can reproduce the human mind, was invented by Alan Turing, the developer of this machine (Khanna & Dhaimade, 2017).

AI is an area of computer science that attempts to create computers that can perform functions that would normally need human intelligence.

The aim of general AI is to build machines that can completely mimic human thoughts, emotions, and logic, whereas the goal of narrow AI is to build machines that can do a specific task as well as or better than a person. Creating a machine that can perfectly mimic human brain processes, no matter how advanced technology is today, is a dream that can only be accomplished in science fiction stories (Khanna & Dhaimade, 2017).

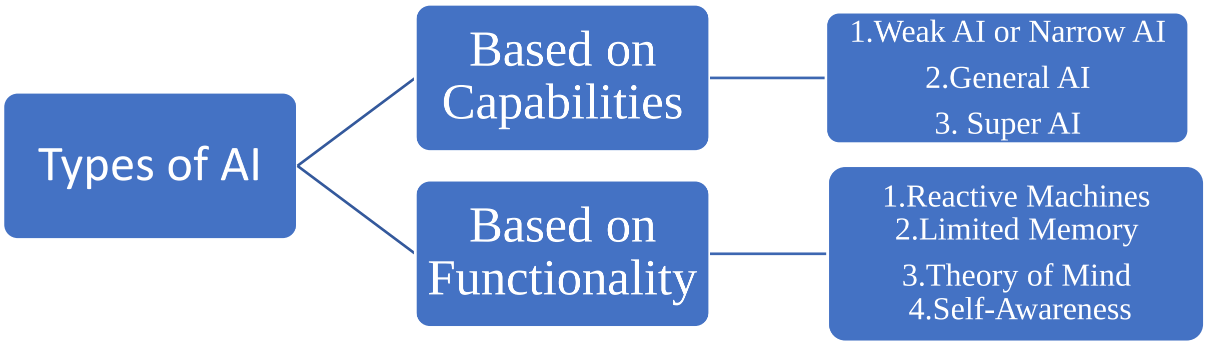
Image classification, object identification, audio interpretation, language translation, natural teeth processing, and gaming play all employ artificial intelligence technologies. Auxiliary software for breast cancer detection in mammograms, segmentation of liver metastases with computed tomography, brain tumor segmentation with magnetic resonance imaging, and classification of interstitial lung disease with high-resolution chest computed tomography has been developed using cutting-edge deep learning algorithms. Research is being conducted to create similar dental applications (Chartrand, et al., 2017).

Artificial intelligence may be divided into several types; nevertheless, there are primarily two basic categories depending capabilities and functionalities.

The following flow diagram (Figure 4) explain the basic types of AI.

Figure 2.

Types of AI.



2.6.1. AI types Based on Capabilities:

1. Weak AI or Narrow AI:

   Narrow AI is a sort of AI that can intelligently accomplish a certain task. In the area of artificial intelligence, narrow AI is the most frequent and currently accessible AI.

   Since it is exclusively trained for one function, narrow AI cannot perform outside its field or boundaries. As a result, it's also known as "weak AI." When narrow AI reaches its limits, it might fail in surprising ways.

   Apple Siri is an excellent example of Narrow AI, but it only performs a limited set of functions. IBM's Watson supercomputer, which employs an Expert system approach combined with Machine learning and natural language processing, falls under Narrow AI.
   Playing chess, purchasing suggestions on an e-commerce site, self-driving automobiles, speech recognition, and image recognition are all examples of narrow AI.

2. General AI:
   General AI is a form of intelligence able of performing any intellectual task as efficiently as a human. The goal of general AI is to create a system that can understand and think like a human on its own. There is currently no system that can be classified as general AI and perform any task as well as a person.

   Researchers all across the world are currently concentrating on developing machines with general AI. Since general AI systems are still being developed, creating such systems will take a lot of work and time.

3. Super AI:
   Super AI is a level of system intelligence at which machines may outperform humans in any task with cognitive properties. It is the product

of general AI.

Some key characteristics of strong AI include the ability to think, reason, solve puzzles, make a decision, plan, learn, and communicate on its own.

Super AI is still a futuristic Artificial Intelligence concept. Development of such systems in real is still world changing task.

2.6.2. AI types Based on functionality:

1. Reactive Machines

   The most basic kinds of Artificial Intelligence are pure reactive machines. For future actions, such AI systems do not retain memories or prior experiences.

   These robots just consider current circumstances and respond in the best way feasible.

   Reactive machines are demonstrated by IBM's Deep Blue supercomputer. AlphaGo, developed by Google, is another example of reactive machines.

2. Limited Memory

   For a limited period of time, limited memory machines can store prior experiences or certain data. These machines can only utilize stored data for a short period of time.

   One of the best examples of Limited Memory systems is self-driving automobiles. These automobiles may store information such as the speed of neighbouring cars, their distance from other cars, the speed limit, and other data to help them navigate the road.

3. Theory of Mind

   Theory of Mind AI should be capable of grasping human emotions, personalities, and beliefs, while also interact socially. This form of AI machine has yet to be developed, however researchers are working hard to enhance and construct such AI machines.

4. Self-Awareness

The future of artificial intelligence is Self-awareness AI. These machines will be extremely intelligent, with their own mind, feelings, and self-awareness. These machines will be more intelligent than the human mind. Self-Awareness AI does not exist in reality and is only a hypothetical concept.

2.7 Machine Learning:

Machine learning is a subfield of artificial intelligence in which algorithms are trained using the available data according to the generated learning model, rather than being clearly coded to perform their tasks. The amount of data required for training depends on how complex the targeted task is (Chartrand et al., 2017). Generally, two different data types are used in machine learning.

In supervised learning, each data sample is labeled by experts and the attributes in the data are determined. For example, in the analysis of an image, algorithms that an expert will program, divides the input images into their basic components, such as borders, gradients, and textures. Images are classified by statistical analysis of these components.

In unsupervised learning, data is not labeled by experts. In these systems, feature extraction is not used by experts. Instead, algorithms learn the most appropriate features to be used in clustering the available data by training examples. The challenge in this type of learning is how to learn complex features from raw data. These systems can be more successful than the features determined by the experts in the presence of sufficient training data.

Apart from these two learning methods, systems are also available. In semi-supervised learning, a large number of unlabeled data is used with less labeled data. In reinforcement learning, learning is directed by evaluating the results of the system as good or bad (Chartrand et al., 2017; Erickson, 2019).

2.8 Artificial Neural Networks (ANN):

Artificial neural networks are computer programs that emulate biological neural networks. Just like the biological neurons in the human brain and the axons that connect them, there are neurons and similar structures in the artificial neural network architecture. Neurons communicate with each other through synapses.

A synapse is the crossing between the axon of one neuron and the dendrite of another. The information processed by the nerve cell is received from the dendrites and sent to the other cells from the axons (Öztemel, 2003).

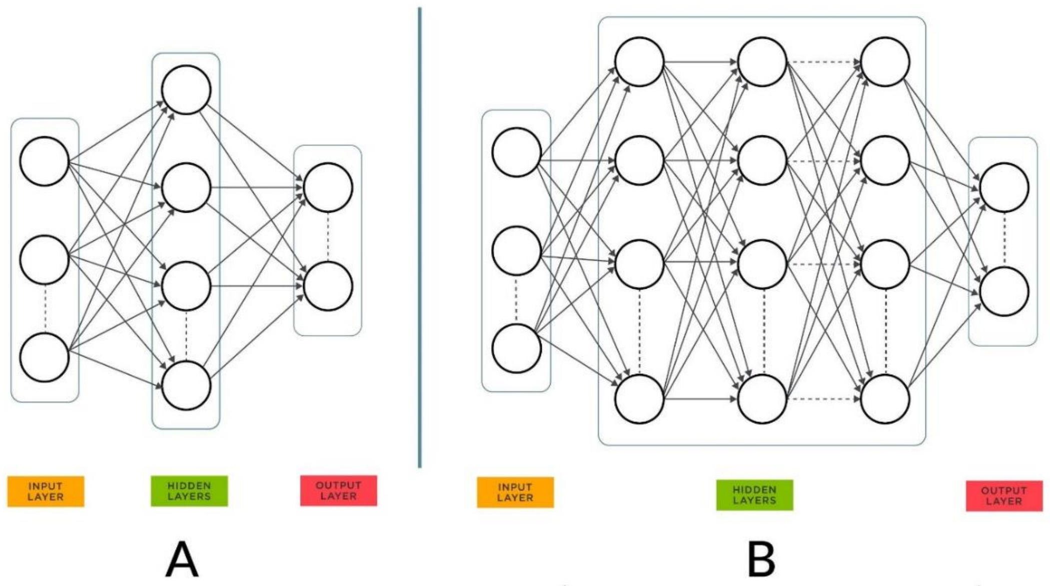
While the neural network structure is being formed, the cells are arranged in layers parallel to each other.

Neural networks consist of input layer, middleware(s) and output layer. The input layer is where the input is given and forwarded to the middleware. Intermediate layers are the layer where information is processed and decisions are made and sent to the output layer. An artificial neural network can have one or more intermediate layers (Figure 3).

The output layer processes the information coming from the middle layer and produces the output of the network (Öztemel, 2003).

Figure 3.

Artificial Neural Network Architecture, (A) Single Hidden Layer (B) Neural Network Structure with Multiple Hidden Layers.



It is not known how neural networks convert input vectors to output vectors. Due to their complexity and feature learning capabilities, the way these neural networks work is described as a "black box" (Chartrand et al., 2017).

2.8.1. Deep Learning:

Deep learning is an artificial neural network architecture with a large number of intermediate (hidden) layers. The word deep in the term deep neural network refers to a network with multiple layers. This forms the basis of deep learning.

In the first layer, simple decisions are made based on the input, in the second layer complex decisions are made based on the decisions made in the first layer. As we go deeper into the network, the extraction of increasingly complex features, representing the hierarchy of layers, takes place and more complex and abstract decisions emerge.

The biggest advantage of this technique is that the features are extracted directly from the raw data and the classification algorithm automatically trains itself according to the extracted features (LeCun, 2015).

The appropriate dataset size to adequately train deep learning models is variable and depends on the nature and complexity of the task. As a general rule, the performance of a network with more layers and more training, which is described as “deeper”, is proportionally higher (Chartrand et al., 2017).

Therefore, deep learning networks need large amounts of training data. Another solution is to modify the training data, called "data augmentation", by adding some variables. Thus, the model does not see the same inputs from the training data in the set during the training cycle (Chartrand et al., 2017; Krizhevsky et al., 2017).

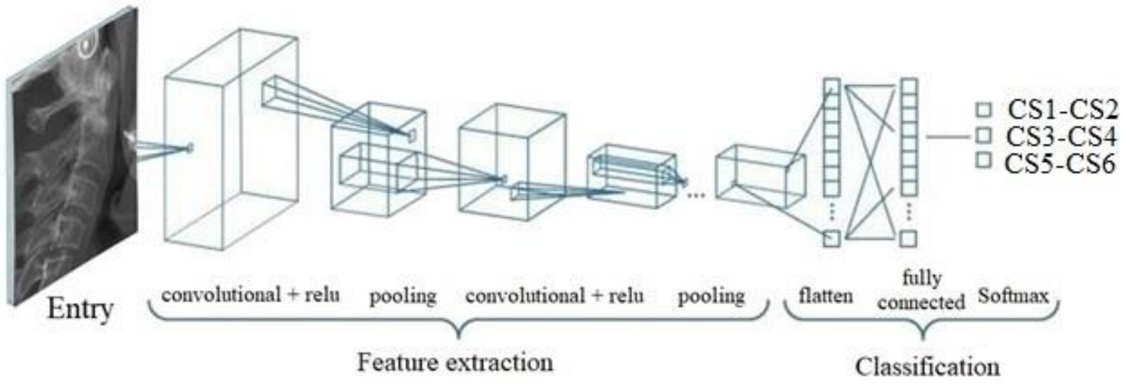
2.8.1.1. Convolutional Neural Network (CNN):

Convolutional Neural Network (CNN) is a deep learning architecture that has been proposed for solving image processing problems and is highly adaptive to images. It has the capability of recognizing, detecting and classifying objects from images such as pictures, photographs and videos with high accuracy. It has been successfully applied in many fields and medical image analysis in the health field.

Unlike other neural networks, it includes feature extraction and classification layers (Figure 7). Classical techniques care about the location of the object while searching for the object in the image, and the object must be completely there. This problem has been overcome with the CNN architecture. With the convolution process, the object is detected wherever it is in the image, thanks to the shifting of the filter on the image. It can also detect even if there is only a part of the object (Ronneberger et al., 2015; Krizhevsky et al., 2017).

Figure 4.

Example of Convolutional Neural Network Architecture.

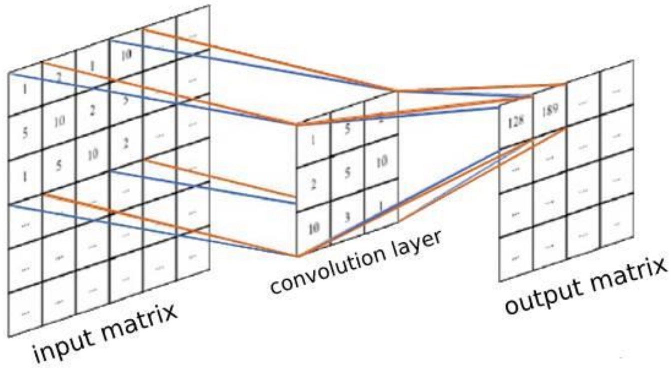


The convolution layer is the most important layer that forms the basis of the CNN architecture (Figure 4).

The convolution operation applied through the layer is the operation of circulating a filter by sliding it over the entire image. Filters create output matrices by applying this process to images from the previous layer. These created matrices are called feature maps. These maps are the regions where the features specific to each filter are discovered. Normally, all neurons are connected to each other between the previous layer and the next layer in artificial neural networks. However, by shifting the filter on the image with the convolution operation, it reduces the size of the subsequent matrix considerably, reducing the number of these connections. Thus, the network parameters are reduced much more than they should be. The low training parameter speeds up the training process and eliminates the memorization problem (LeCun, 1990; 2015; Krizhevsky et al., 2017).

Figure 5.

Convolution Matrix.



2.8.2. Training and testing of artificial neural networks:

For the training of the network, the data separated from the data set as a "train set" is used. Passing the created model over the training set once is called "loop" (epoch, iteration).

Choosing the appropriate number of loop (cycles) affects the performance of the network, and choosing less or more than necessary will reduce the performance of the network. The number of cycles also affects the duration of the training (Chartrand et al., 2017).

Training the artificial neural network is to produce the correct output by changing the weight values between neurons, that is, the network can make correct predictions.

Weight values are parameters that describe the contribution of each feature to the problem.

These weights, which are randomly assigned at the beginning, change many times during training to find the correct output (Chartrand et al., 2017; Erickson et al., 2017).

During training, an error value is calculated to represent the difference between the estimated output class of the network and the labeled data class. The weight values are updated according to the error value, so that the network produces the correct output.

During the update process, which is the most important stage of the training process, the weight values in each layer change by propagating backwards from the output layer.

These processes occur in each cycle, and the training process is terminated when either the maximum number of cycles or the minimum error value is reached (LeCun et al., 2015).

After the completion of the training process, the network is tested to determine the performance of the network. In order to test the network, samples that are not shown to the network during training and that are separated from the data set for testing purposes are used. This data set is called the "test set". The weight values of the network are not changed in the test process. The accuracy of the outputs obtained here determines the performance of the network. In addition to the training and test set, the "validation set" is used to monitor the performance of the model as the parameters change during the training. This dataset tests how much the network has learned in each cycle (Erickson et al., 2017).

One of the common problems when learning networks is overfitting (memorization, overfitting). Overfitting is when the model overfits the training data and loses its ability to generalize. The model learned so much about the training data, including noise, that it could not capture the underlying general information in the next dataset (Erickson et al., 2017).

2.9. Segmentation:

A typical use of convolutional neural networks is classification, which is a single class label of an image output. In the results produced by the object classification algorithms, only the information about whether the object exists or not is produced. The output does not contain any location information (Krizhevsky et al., 2017).

However, many visual tasks, especially biomedical image processing, must include output localization, ie a class label must be assigned to each pixel. Image segmentation is the pixel-based classification of an image. During this process, a class label is assigned to each pixel according to its properties.

Long et al. introduced ESA-based FCN for semantic segmentation in 2015, and after this study, the course of image processing studies has changed.

This network is designed and trained for pixel-based segmentation. In the semantic segmentation process, objects or regions with similar properties are labeled with the same label value. Neighborhoods and properties of each pixel are examined. If a tag value was previously assigned to a pixel group close to these properties, the same tag value is assigned to this new pixel. Thus, the location of the object is determined along with the information in the output (Long et al., 2015).

2.10 U-Net architecture:

U-Net, proposed by Ronneberger et al. in 2015 for better segmentation in medical images, is a kind of convolutional neural network approach. It takes its name from its U-shape-like architecture in which the layers are lined up. The first part of the network is usually called the encoder, and the second part is called the decoder.

In the first part of the structure, the convolution layers extract features from the image, and the image dimensions slowly decrease. In this section, the image properties are learned and the properties of the objects in the image are determined.

In the second part of the structure, deconvolution and convolution layers were used. The image dimensions are slowly increasing, like a mirror of the first half. Thus, the structure of the image is moved to its original dimensions. It is very popular today because of its structure that transfers the basic features of its architecture to the last layers in semantic segmentation problems. (Ronneberger et al., 2015).

2.11 Alex-Net:

Alex-net, which was developed in 2021 to classify images for the ImageNet LSVRC-2010 competition, is a deep neural network aimed at mimicking the biological process of human vision. It consists of 8 layers which is designed in a way to stand out from any other network architectures (Krizhevsky, et al.,2012).

CHAPTER III

3.Methodology

3.1 Ethical approve:

In accordance to the regulations of the 1964 Helsinki Declaration on medical research ethics, a request was submitted to and approved by the Near East University Scientific Research Evaluation Ethics Committee (decision date and number: YDU/2021/93-1382).

3.2 Collection of Dataset:

Data was collected from the Oral and maxillofacial radiology Department Archive at Near East University (2015-2021). Lateral LCs were taken by the multiple X-ray machines, in total 4 machines were used. The collected data constituted a wide range of age groups (7-25 years) in order to ensure the evaluation of all maturation stages of CVM.

3.3 The exclusion criteria were:

- Any presence of an artifact (Beam-hardening artefact, motion artefact).
- Images without full body of C2 C3 C4 in field of view. (Figure7)
- Presence of any syndrome that would affect maxillofacial structures.
- Improper position of the patient's head.

Figure 6.

LC Taken With Earrings Which Lead To C2 Vertebra Body Not Being Fully Visible.

Figure 7.

Example of Excluded LC, Full Body of Vertebras Are Not Visible.

In total 1505 LCs qualified to become part of the research and all lateral cephalometric cases were pseudonymized before proceeding to labelling.

After examining the collected data, we noticed a major difference in sample sizes between the different maturation stages. Thus, to reduce the effects of this limitation, we simplified the classification procedure and merged two sequential classes into a single class. In brief, the stages of CVM were divided into 3 categories as:

1. Pre-pubertal stage (CVM1 and CVM2).
2. Pubertal stage (CVM3 and CVM4).
3. Post-pubertal stage (CVM5 and CVM 6).

The data was given to the main observer (orthodontic resident) to start the process of tracing and labeling them, a process which was carried out using Computer Vision Annotation Tool (CVAT), and once the labeling was finished, each sample was immediately classified according to Bachetti method of CVM.

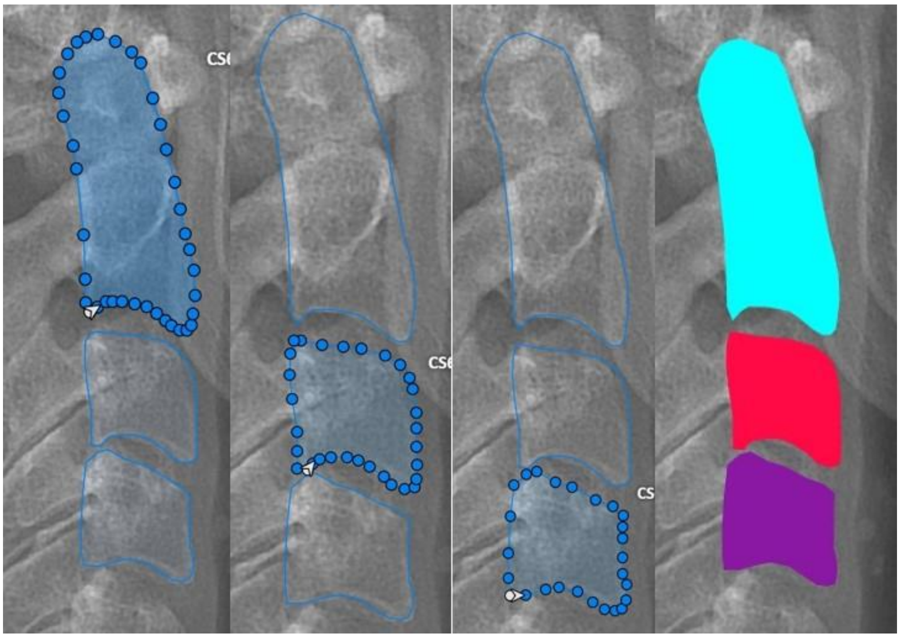
3.4 Tracing and labelling process:

The entire process of labelling and tracing was carried out using Computer Vision Annotation Tool (CVAT)(Figure 9).

CVAT is a web-based image and video annotation tool that is free and open source. It is used to label data for computer vision algorithms. CVAT, which was created by Intel, is intended for usage by a professional data annotation team, and features a user interface tailored for computer vision annotation jobs.

Figure 8.

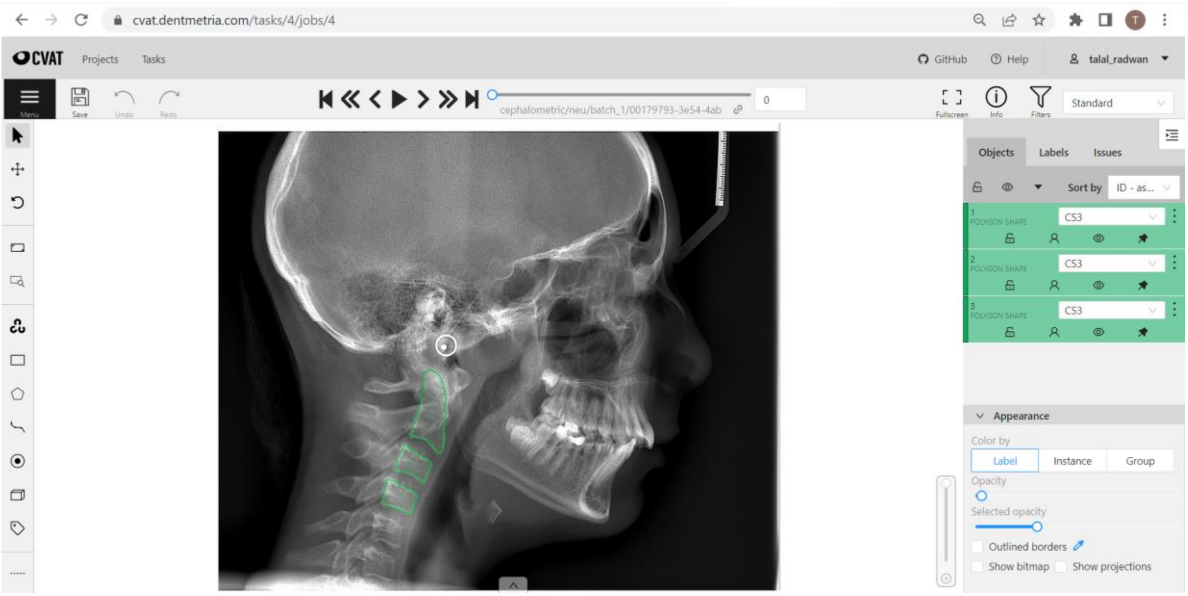
Tracing Process of Vertebras (a) Tracing of C2 (b) Tracing of C3 (c) Tracing of C4 (d) The Outcome of The Tracing Process is 3 Individual Shape That Will Be Used In The Labelling Process.



(a) (b) (c) (d)

Figure 9.

Classification Process of Maturation Stage, In this Case Patient is Classified as CS3 Which Can Be Seen In The Right Side.



Tracing the border of each vertebra was done by the same observer (orthodontic resident with 4 years' experience) (Figure 8).
once the tracing and labelling of each LC was done, the observer started the classification process immediately, classification was carried out according to Bachetti Method of CVM.

3.5 Data distribution:

To examine reliability, the primary observer re-classified 20% of the data separately two weeks after the first annotation was completed. Additionally, the same data was presented to a second observer for re-classification in order to assess the reliability between the two orthodontists. The selected data which was used for the re-evaluation was randomly generated by the computer.

When the preparation of the dataset was completed, the labelled data were randomly split into a training set (80%), a testing set (10%) and a validation set (10%).

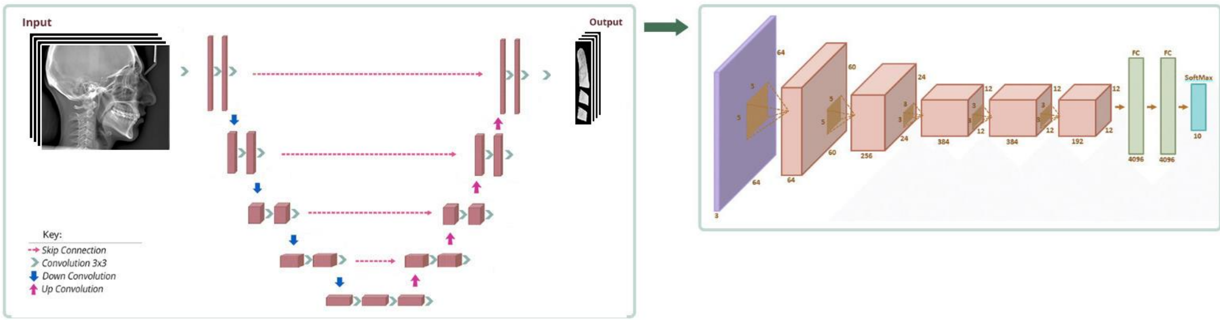
3.6 Model Pipeline:

Our study approach contained the following steps:

1. Each lateral cephalometric x-ray was resized to 1024 by 768.
2. Using Res-U Net architecture each lateral cephalometric X-Ray was segmented.
3. From each segmented Lateral cephalometric the C2 C3 C4 was isolated.
4. A patch centred on the Vertebras was cropped.
5. The data was then run through a classification network based on Alex-Net that resulted in the classification of CVM stages. (Fig 10).

Figure 10.

The workflow for the segmentation and the classification processes of CVM stage. The first schema represents the encoder-decoder style network of U-Net. The output of this network feeds the Res-Net architecture which is seen in the second diagram.



3.7 Implementation and Classification:

The study algorithm was based on the Python implementation of U-Net and Alex-Net.
All training and experiments were done using NVIDIA® GeForce® RTX 2080 Ti GPU. For segmentation, the network was trained using Adam optimizer. The cross-entropy was implemented as a loss function. The batch size was set to 1. The learning rate was set to 10−4. The model with the best validation loss was chosen for testing. For the classification, the network was trained using Adam optimizer. As a loss function cross-entropy was used. The batch size was set to 1. The learning rate was set to 10−4. The model with the best validation loss was chosen for testing.

3.8 Statistical Analysis:

Two weeks after the first labelling and classification stage, 20% of the dataset was reclassified by the same examiner to measure intra-observer reliability and by a second observer in order to measure the agreement level of the measures obtained by the two observers (inter-observer reliability).

Intraclass correlation coefficient (ICC) and Cohen's kappa were conducted using SPSS version 21.0 (SPSS 21.0 Software Package Program, Inc., Chicago, IL, USA), to measure the intra-and inter-reliability of the observers.

In research with two or more raters, an intraclass correlation coefficient (ICC) is used to assess the reliability of ratings. An ICC can have a value between 0 and 1, with 0 indicating no reliability among raters and 1 representing full reliability among raters. (Table 2).

In layman's words, an ICC is used to assess if items (or subjects) can be reliably rated by multiple raters. In our results ICC was 0.998 indicating excellent reliability for the primary observer.

Table 2.

Interpretation of ICC Values

ICC	Interpretation
0 - 0.39	Poor reliability
0.4 - 0.74	Modest reliability
0.75 - 1	Excellent reliability

Cohen's kappa coefficient is a statistic used to assess inter-rater (and intra-rater) reliability for qualitative (categorical) items. It is often regarded as a more robust metric than a simple percent agreement calculation since it considers the probability of the agreement occurring by coincidence, The Kappa statistic varies from 0 to 1 (Table 3).

The weighted Cohen's kappa standard error was 0.875 ± 0.026. This result indicates high reliability of the observers and proves that there is a good level of agreement between them (Landis & Koch, 1977).

Table 3.

Kappa Value Interpretation According to Landis & Koch (1977).

Kappa value	Interpretation
< 0	Poor agreement
0.01 - 0.20	Slight agreement
0.21-0.40	Fair agreement
0.41-0.60	Moderate agreement
0.61-0.80	Substantial agreement
0.81 - 1.00	Almost perfect agreement

3.9 Performance measurements:

When measuring algorithm performance, several criteria are available to evaluate the reliability of the algorithm. To quantify the performance of the classification model, the accuracy, specificity, sensitivity metrics and f1 scores were computed.

Another method used to evaluate the performance of the network is the Confusion Matrix, which shows correct and incorrectly predicted classes. The Confusion matrix, which is also known as error matrix, is a special table with a layout to visualize the performance and success of the algorithm. Confusion matrix factors for classification in our study (Figure 11):

- **True Positives (TP):** is a result in which the model predicts the positive class correctly.
- **True Negatives (TN):** is a result in which the model predicts the negative class correctly.
- **False Positives (FP):** is an outcome in which the model predicts the positive class incorrectly.
- **False Negatives (FN):** is an outcome in which the model predicts the negative class incorrectly.

Figure 11.

Confusion Matrix.

Prediction

Actual		1	
	1	True Positive (TP)	False Negative (FN)
	0	False Positive (FP)	True Negative (TN)

- **Accuracy** is a metric that is used to describe the performance of a model across all classes and is calculated as the ratio between the number of correct predictions to the total number of predictions.

Figure 12.

Accuracy Formula.

$$Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions}$$

- **Sensitivity** (True Positive Rate) refers to the probability of a positive test, conditioned on truly being positive.

- **Specificity** (True Negative Rate) refers to the probability of a negative test, conditioned on truly being negative.

F1-score is an important evaluation metrics in machine learning. It efficiently sums up the predictive performance of any model by combining two otherwise competing metrics — precision and recall.

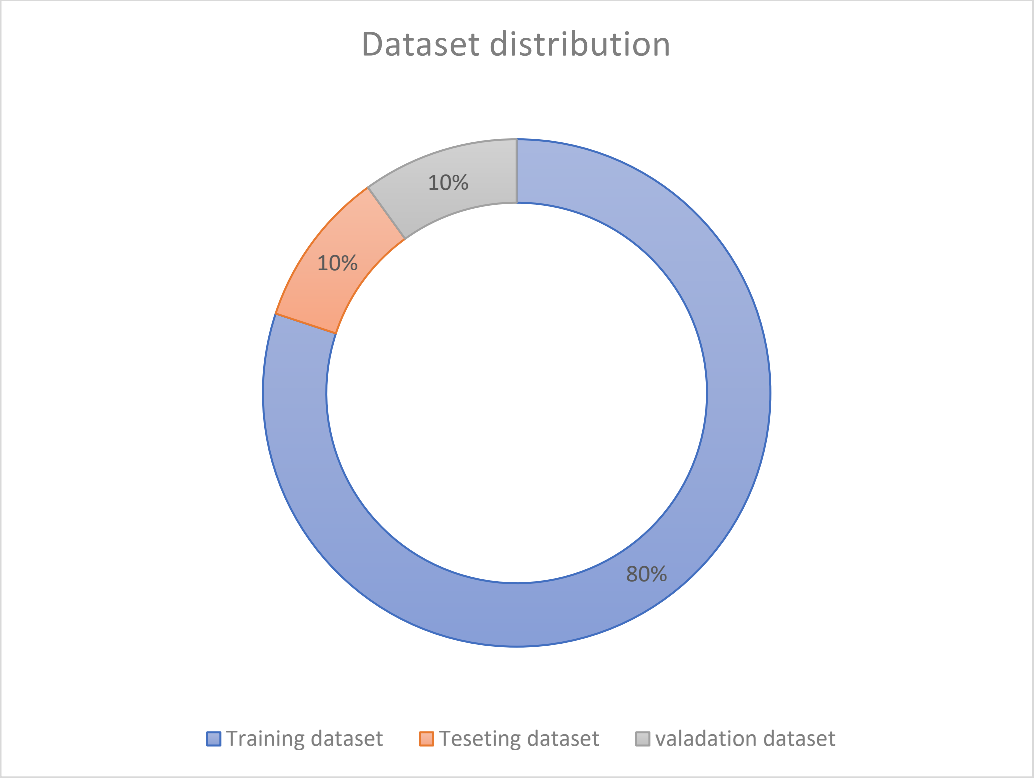
CHAPTER IV

4. Findings and Discussion

The dataset contained 1505 LCs, the training dataset of U-net contained 1205 images, and the testing dataset contained 150 images with validation group images which were counted at 150.

Figure 13.

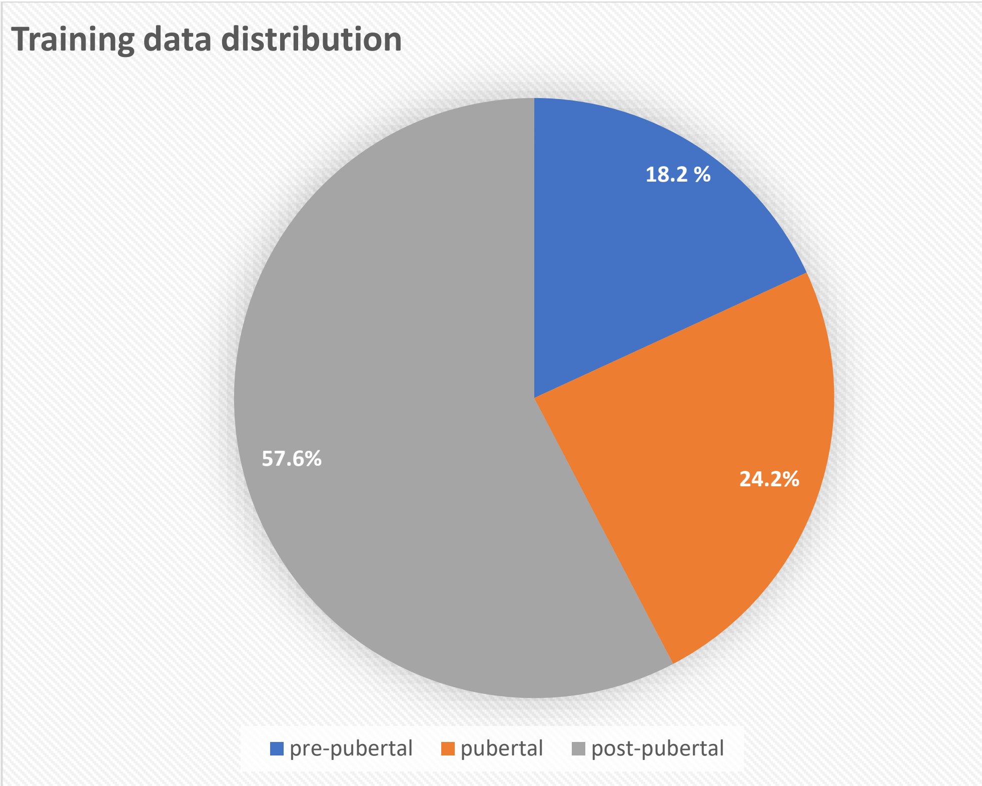
Dataset Distribution Percentage.



For the classification, the data distribution used in the Alex-Net was as follows: 18.2 % in pre-pubertal stage (CS1, CS2), 24.2% in pubertal stage (CS3, CS4) and 57.6% in post-pubertal stage (CS5, CS6). (Chart 2)

Figure 14.

Training Dataset Distribution



Sensitivity and specificity mathematically describe the accuracy of a test that reports the presence or absence of a condition.

Performance of segmentation algorithm:

Our segmentation network based on U-Net achieved a dice score of 0.931 and an accuracy of 0.998 for the test set. Our classification network based on Alex-Net achieved an overall accuracy of 0.802 for the test dataset. The confusion matrix of the classification model is shown in (Figure 15).

Figure 15.

The Confusion Matrix of The Classification Model.

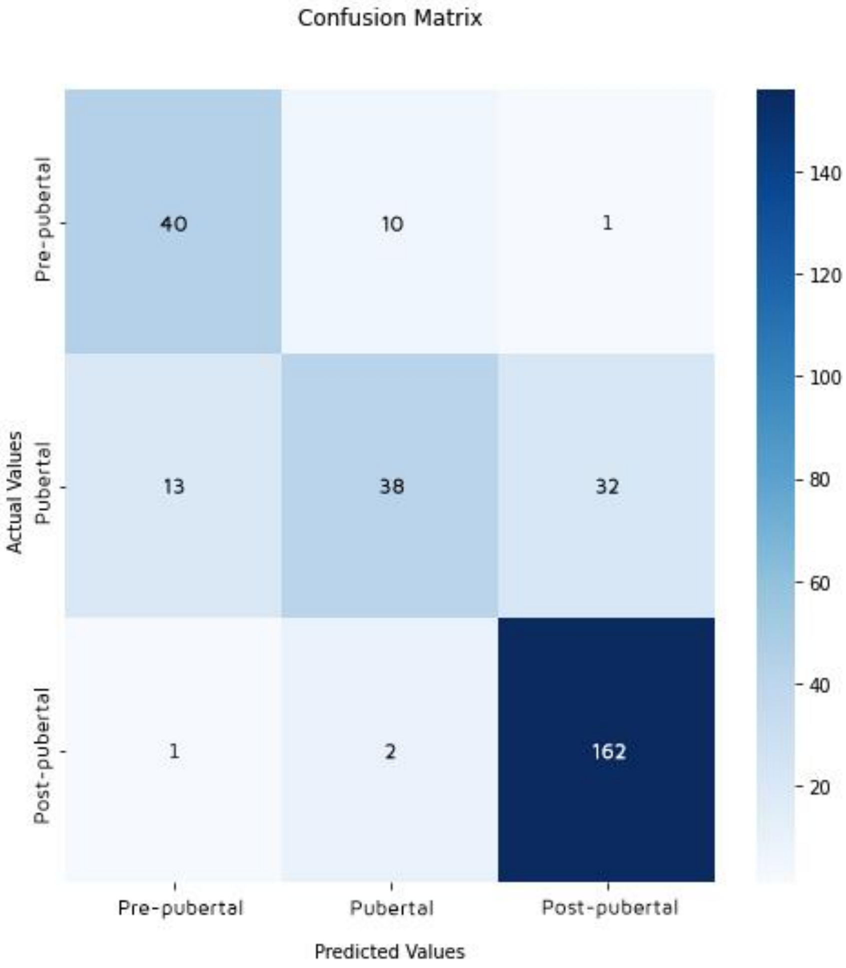


Table 4.

Performance of The Algorithm in Metric Values.

Table 1	Accuracy	Sensitivity	Specificity	F1 score
Pre-pubertal	0.9164	0.7843	0.9435	0.7619
Pubertal	0.8094	0.4578	0.9444	0.5714
Post-pubertal	0.8796	0.9818	0.7537	0.9000

We achieved per class sensitivity of 0.78, 0.45, 0.98, per class specificity of 0.94, 0.94, 0.75, and in addition, we achieved an f1 score of 0.76, 0.57, 0.90, respectively (Table 2).

CHAPTER V

5.Discussion

The optimal time of orthodontic intervention has prompted significant controversy over the years, with a majority of specialist dedicated to consistently providing 'early' treatment. (Fleming et al.,2017)

Most Authors suggest late mixed or early permanent dentition as a starting point for orthodontic treatment. This phase usually corresponds with a period of maximum development (growth spurt), allowing for effective repair of growth-related occlusal abnormalities. (Fleming et al.,2017)

The growth process is different for each individual and can be affected by several different factors. Understanding growth activities is a key element in diagnosis and treatment planning in orthodontics (Baldin et al., 2017; Korde et al., 2015).

Bone age determination is performed clinically in order to determine the facial and body development period of the individual in the treatment planning of malocclusions in the field of orthodontics.

Unless surgical intervention is performed after puberty, significant skeletal changes cannot be achieved with orthodontic treatment. It is known that extraoral forces and functional appliances are effective especially if applied just before pubertal growth spurt and during active periods of growth.

Different stages of puberty can be used for growth modification. There are studies showing that functional appliances used in the treatment of Class II malocclusions give more effective results during puberty. It has been reported that the first phase is effective in the early prepubertal period in biphasic orthodontic treatments performed in class II cases, while palatal expansion performed before late puberty in class III cases is effective and stable.

The growth status of the individual is taken into account in the decisions of application of extraoral forces, functional appliances, treatment with and without extraction, or orthognathic surgery. Bone age is used to determine the growth spurt to decide the optimal timing in growth modification treatments (Kasimoglu & Tuna-Ince, 2016; Korde et al., 2015)

Relative intrusion of the implants can be observed as a result of the eruption of the teeth over time, in implant applications made before the growth process is completed in order to eliminate congenital or traumatic tooth deficiencies in young individuals. For this reason, bone age assessment is used in the timing of implant treatment in individuals in the growth and development process.

Replantation of avulsed teeth gives very successful results when done under appropriate conditions. However, in most cases, ideal conditions cannot be achieved and ankylosis develops in the relevant region. In other words, the bone will replace the tooth over time. The rate of root resorption as a result of ankylosis is related to the rate of remodeling in the body and slows down after puberty.

The effects of ankylosis in a growing child are discussed as a controversial issue. Different treatment recommendations have been presented according to at which period of puberty the tooth will be replanted (Kasimoglu & Tuna-Ince, 2016).

In dentistry practice, different maturation indicators can be used to determine the maturation level of individuals (Altan, Kocasaraç, Sinanoğlu, & Mutaf, 2014).

For this purpose, using hand and wrist method is a popular, reliable, easy and error-free method (Altan et al., 2014; Türkoz et al., 2017). The diagnostic efficiency of the Hand and Wrist method is accepted and it is a valid method (Altan, Kocasaraç, Sinanoğlu, & Mutaf, 2014).).

The effective dose in the HWR method is stated as 0.16 microsievert (μSv). The absence of vital organs such as thyroid in the imaged area can be considered as an advantage (Kasimoglu & Tuna-Ince, 2016). Despite this, the need for an additional x-ray from the wrist region creates a disadvantage in terms of radiation dose (Altan, Kocasaraç, Sinanoğlu, & Mutaf, 2014).; Kasimoglu & Tuna-Ince, 2016; Türkoz et al., 2017).

X-rays of the wrist to determine pubertal growth spurt have been stated as contraindicated by the British Orthodontic Society (Altan, Kocasaraç, Sinanoğlu, & Mutaf, 2014).; Wong, Alkhal, Rabie, & Orthopedics, 2009).

There are authors who state that due to the complex structure of the hand-wrist bones, it contains limited information about the pubertal growth age and may lead to interpretation errors (Kasimoglu & Tuna-Ince, 2016).

In the GP Atlas method, there may be difficulties in matching when comparing the guide images and the image of the individual. It may not be possible to apply the method in judicial cases where the wrist area is not protected (Türkoz et al., 2017).

It is known that the lateral view of the cervical vertebral body shapes changes as the individual grows (Mito, Sato, Mitani, & orthopedics, 2002; Türkoz et al., 2017). As an alternative method to the HWR method, the evaluation of skeletal age with CV on lateral cephalometric radiographs routinely used in orthodontics attracts the attention of researchers (Kama, Aslan, Darı, & Özer, 2006).

In their study, Franchi and Baccetti found that there is a positive relationship between CVM and mandibular growth in the pubertal period (Baldin et al., 2017).

Evaluation of CVM in lateral cephalometric radiographs can reduce the need for additional radiography, reduce imaging costs and reduce the radiation dose to the patient (Joshi et al., 2010; Kama, Aslan, Darı, & Özer, 2006.; Türkoz et al., 2017).

The rare occurrence of events that may cause structural changes such as injury, inflammation, tumor, and deformation in the cervical vertebral bodies increases the usability of the method Kama, Aslan, Darı, & Özer, 2006; Mito, Sato, Mitani, & orthopedics, 2002).

It has been reported that C2, C3 and C4 vertebrae can be seen in the use of thyroid protectors (Baccetti et al., 2005). While scanning the archives in our retrospective study, it was observed that the radiopaque projection of the protective collar on the C4 and C5 vertebrae was superposed in rare cases.

Since CVM evaluations are made visually, they are considered subjective (Baccetti et al., 2005; Türkoz et al., 2017). It has been reported that the evaluations made using Atlas will not be objective and will not contain detailed information, and that the person performing the evaluation should be experienced for the correct diagnosis (Baldin et al., 2017; Mito, Sato, Mitani, & orthopedics, 2002).

CVM assessments are not sensitive except during the pubertal breakthrough period. Correlation coefficients may be affected as pubertal growth moves away from the breakthrough period (Wong, Alkhal, Rabie, & Orthopedics, 2009).

CVM stage analysis in lateral cephalograms with horizontal plane (yaw) and frontal plane (roll) positioning problems may be inaccurate. Rotational errors in the horizontal plane (yaw) will almost certainly result in an underestimating of maxillomandibular linear measurements, whereas rotational errors in the vertical plane (pitch) would almost certainly result in an overestimation. Errors in the vertical plane (roll) will very certainly cause the measurements to be overestimated. (Mehta et al,2020)

Positional errors in the sagittal plane (pitch) are unlikely to have a substantial impact on the accuracy of CVM measurements. CVM stages and maxillomandibular parameters in lateral cephalograms obtained with multiplane rotations errors may be more inaccurate than measurements relating to single plane rotations error. (Mehta et al,2020)

When compared to lateral cephalograms with a mild rotation of 5°, lateral cephalograms with a larger degree of rotation 10° may result in more inaccuracies in CVM stages. (Mehta et al,2020)

While conducting our study, errors in the frontal plane (roll) caused a true challenge in estimating the stage of CVM due to the fact that an error in roll will lead to creating two shadow of the same vertebra which in turn will cause problem to notice any morphological changes especially to the base of vertebras.

Different results have been reported in studies on the usability of CVM assessment instead of HWR.

In a systematic review published by (Cericato, Bittencourt, & Paranhos, 2015) it was concluded that the indices reported by Hassel and Farman (1995) and Baccetti (2002) for the evaluation of CVM are reliable enough to replace HWR (Cericato, Bittencourt, & Paranhos, 2015).

Demirturk et al., in their study with 116 Turkish patients in 2017, reported a good correlation between mandibular third molar mineralization, spheno-occipital synchondrosis fusion, chronological age and CVM level (Demirturk et al., 2017).

Turkoz et al., in their study with 324 Caucasian patients, determined that the formula they developed to evaluate CVM could be used instead of HWR (Türkoz et al., 2017).

In the study of Alhadlaq et al., which aimed to determine the bone age by applying step-by-step multiple regression analysis to the measurements made on the images of 122 Saudi boys, a high correlation was found between the method they developed and the HWR (Alhadlaq & Al-Maflehi, 2013).

Joshi et al., in their study in 2010, reported that there was no significant difference between CVM and HWR for both genders when calculating skeletal age, and that lateral cephalometric images could be reliable in detecting pubertal growth spurt (Joshi et al., 2010).

Wong et al., in their study with the northern Chinese population, showed a high correlation between CVM and HWR method, and it was reported that the CVM method could be used instead of HWR (Wong et al., 2009).

Kama et al., in their study evaluating the images of 150 male individuals, found that the CVM method was as reliable as the HWR (Kama, Aslan, Darı, & Özer, 2006).

Mito aimed to find a new formula to evaluate CVM. As a result of a study conducted with a total of 242 Japanese female patients, the assessment of CVM was found to be as reliable as HWR (Mito, Sato, Mitani, & orthopedics, 2002).

Hoseini et al., in their study of CVM and HWR analysis with 133 radiographs, reported that the agreement between the two methods was low and that these two methods could not be used interchangeably (Hoseini et al., 2016).

Szemraj et al. published a systematic review on the usability of the CVM method instead of the HWR method. In eight of the ten literatures with appropriate criteria, the CVM method can replace the HWR, which is accepted as the "gold standard"; In two articles, it is reported that the CVM method is useful but can only be used as an additional method to other methods. In the aforementioned literature, the results of the interobserver agreement analysis were found to be the lowest 0.616 and the highest 0.937 (Szemraj, Wojtaszek-Słomińska, & Racka-Pilszak, 2018).

In the study of Shah et al. with 60 lateral cephalometric images, the consensus between the two observers was expressed with a kappa coefficient of 0.491, and it was recommended to use other biological maturation indicators in determining the growth status (Shah et al., 2016).

In their study with 10 observers, Predko-Engel et al. found that the CVM method and interobserver agreement were questionable, and that this method should be supported by other biological indicators in determining the growth spurt (Predko-Engel et al., 2015).

Nestman et al., in their study with 10 observers, suggested that although there is a high consensus in determining the presence of concavity, the reproducibility of the CVM method is low due to low agreement in determining vertebral body morphologies and that it should be supported by other methods in orthodontic treatment planning (Nestman et al., 2011).

In our thesis study the intra-observer agreement level was expressed by ICC which measured at 0.998 indicating excellent reliability for the primary observer and a good level of reproducibility of CVM staging process while using Bachetti method.

Also, in our thesis study the agreement between our two observers (interobserver) was also expressed by kappa coefficient of 0.875 ± 0.026 which suggest a good agreement level between the two different observers and prove that the CVM technique can be used in order to gain a general idea about the patient growth stage, but if an accurate diagnosis and a precise growth stage determination is needed the CVM technique it should be supported by other method most preferably hand and wrist technique.

When the literature on the subject was evaluated, it was concluded that CVM could be used instead of HWR, but the method is subjective in nature and there may be differences between observers.

In the evaluation made by clinicians with visual method, it was thought that providing software support could increase standardization.

There are different methods proposed by different researchers to evaluate CVM (Koukouviti, 2017).

In their 2016 review Durka-Zając et al. It was concluded that the CVM detection method modified by Tiziano Baccetti et al., 2005, can be used instead of HWR and is the most frequently mentioned method in the literature (Durka-Zając et al 2016).

Jaqueira et al. compared three different methods commonly used to determine CVM. In this study, which was conducted with four different observers, one of whom was a radiologist and the other an orthodontist, the best results were obtained by Baccetti et al., 2005 (Jaqueira et al.,2010).

In this thesis study, cervical vertebra maturation level analyzes on lateral cephalometric images were performed by Baccetti et al. by the method reported in 2005 (Tiziano Baccetti et al., 2005).

It was planned to develop an algorithm to assist physicians while performing this visual analysis. For this purpose, sample size of 1501 was collected, the images were transferred into a web-based software and tracing were made to determine the cervical vertebra morphologies.

Tracing and labelling process was carried out using CVAT, the tracing process was carried out by tracing the border of the C2, C3, C4 vertebras individually, once tracing was done to each image, the observer will assign a class for the tracing. Both process was carried out by the same main observer.

The created data set was divided into 3 groups, training group, testing group and validation group, by machine learning algorithms. While creating the first data set, the features of C2, C3, and C4 reported in the literature (Baccetti et al., 2005) were selected. In the evaluations made with the visual method, it is recommended to evaluate factors such as age and gender in cases of uncertainty (McNamara & Franchi 2018).

In this thesis study, AI based algorithm were developed to learn vertebral morphology in order to automatically classify CVM.

In the segmentation made for CVM level determination, U-Net achieved a dice score of 0.931 and an accuracy of 0.998 for the test set, classification network based on Alex-Net achieved an overall accuracy of 0.802 for the test dataset.

The confusion matrix tables show the performance of the algorithm when applied to the testing group for classification, it is noteworthy that the misclassified samples tend to be in neighboring classes. Being undecided between two neighboring classes is also a situation that occurs in analyzes made by people using the visual method.

Nestman et al., 2011 described the reproducibility of CVM method to be poor and stated that the main problem is C3 and C4 vertebral bodies are difficult to identify as trapezoidal, rectangular horizontal, square, or rectangular vertical which might led to misclassification of neighbouring classes the same situation we faced in our thesis study.

Baptista et al., in their study with 188 digital lateral cephalometric images, aimed to determine the level of CVM according to the method reported by (Tiziano Baccetti et al., 2005). As a result of the classification using Naive Bayes algorithms, the weighted kappa coefficient was reported as 0.861. The weighted kappa result was reported as 0.992 when adjacent levels were considered acceptable (Baldin et al., 2017).

Padalino et al., in his study with 100 lateral cephalometric images, aimed to compare the feasibility of CVM by comparing a dedicated software with the manual analysis done on an acetate sheet, it was reported that there was a 94% agreement between computer analyzes and manual analyzes (Padalino et al., 2014).

Santiago et al. aimed to classify 236 digital lateral cephalometric images into four categories with the multinomial logistic regression model. In the study in which the results were compared with the hand-wrist radiographs, the weighted kappa analysis result was reported as 0.832 (Santiago et al., 2014).

Dzemidzic et al. developed a software called Cephalometar HF V1 in their study with 99 digital lateral cephalometric images. The Cohen kappa coefficient of agreement between the software's classification according to the method reported by Hassel & Farman (Hassel & Farman, 1995) and the researcher's analysis was reported as 0.985 (Dzemidzicet al ,, 2015).

The weighted kappa analysis result applied to the results obtained in our study was 0.870 ± 0.027. The values calculated to determine compliance in our study are consistent with those reported in the literature.

The education and experience level of the observer has also been shown to have a great impact on both intra- and inter-observer reliability (Predko-Engel et al., 2015). Specialists who routinely use (experienced) the CVM method in the clinical setting showed the highest reliability level in the study.

However, the reliability demonstrated by those who do not routinely use CVM method (Inexperienced) is significantly weaker (Predko-Engel et al., 2015Therefore, clinicians who do not routinely use CVM may need training and more practical practice to obtain highly reliable measurements. This can be difficult as more time will be required for practitioners.

For these reasons, automatic segmentation enables CVMS assessment, eliminating the need for manual intervention, which is a laborious and time-consuming process for routine clinical practice and is subject to subjectivity among practitioners.

Our study developed for the first time an algorithm to automatically segment and stage CVM according to Bachetti using AI and LC images.

while going throw the literature, automated segmentation showed reliable segmentation results, and also reduced random human error during measurement to zero.

In this thesis, deep learning architecture was used for the segmentation and classification process of CVM. For this, U-Net, which is the most suitable architecture for image processing was used.

In the literature U-Net architecture proved to be one of the most reliably and highly regarded architecture for image segmentation (Ronneberger et al., 2015). Regarding classification, Alex-Net is ranked first among all architecture used for this purpose, hence it was used in this thesis study.

Previous attempt to utilize AI into staging CVM are limited and have drawbacks in nature, the first real attempt to annotate the CVM staging process was in 2019. Kök et al,2019 in her study on 300 subjects, evaluated the accuracy of 7 algorithms of AI frequently used in classification process, in the result of her study ANN network was observed to be the second most accurate algorithm was the most stable algorithm. (Kök et al.,2019).

The outcome of the previously mentioned study corresponds with the finding of our study where Alex-Net which is a type of ANN showed high accuracy in the classification process in pre and post pubertal stages, and it felt short in the pubertal stage regarding accuracy.

Amasya et al., in 2020 designed an ANN algorithm to stage CVM and validated the result by comparing them to a human observer as a gold standard, sample size of 647 subject were collected and manually labelled, morphological features of the vertebras were fed to the computer by marking landmarks (points) on each vertebra, a total of 26 points were placed on each LC, the CVM stage set by the main observer was set to be the gold standard in order to measure the reliability and accuracy of the algorithm.

In order to measure the intra-observer and interobserver agreement of CVM staging process, the study designed included 4 observers, and ANN reliability was measured by comparing the result of the main observer and the algorithm assigned stage.

Weighted kappa (wk) and Cohen's kappa (ck) was performed as a statistical analysis, The wk coefficients for the intra-observer agreement were almost perfect (wk 5 0.92-0.98). Interobserver agreement, including the ANN model, ranged from substantial to almost perfect (wk 5 0.76-0.92). The ck results for the intra-observer agreement were substantial to almost perfect (ck 5 0.65-0.85). (ck 5 0.4- 0.65) among the observers, including the model.

The result of Amasya et al.,2020 study demonstrated satisfactory repeatability and reproducibility of the CVM method to evaluate skeletal development. The results also demonstrated that ANN model can perform close to, if not outperform clinicians in CVM analysis.

Kok et al.,2020 did a second study in 2020 regarding AI in CVM staging process, in her second article on this topic, Kok et al.,2020 compared the performance of ANN algorithm in determining CVM stage and compared the result to Hand and wrist radiograph which was consider to be the gold standard, the result showed high correlation between hand-wrist maturation level, CVS stage determined by ANN algorithm.

A point worth mention is that even in this mentioned study landmarks points were placed on each vertebra to determine the morphological changes in C2, C3 and C4 vertebras, similar to the technique used by Amasya et al.,2020 with the difference being in the number and location of the landmarks.

In 2021, Zhou et al., in this study with a sample of 1080 LC, he aimed to evaluate the accuracy of labelling process between AI and human, AI performance in

classification process was also evaluated in the study. In the results, the mean labelling error among human examiners was 0.48 0.12 mm. Between AI and human examiners, the mean labelling error was 0.36 0.09 mm. The agreement between AI findings and the gold standard was generally good, with intraclass correlation coefficient (ICC) values reaching up to 98%. Furthermore, CVM staging accuracy was 71%.
Just like the predecessors studies reference points and landmarks were used to determine the morphological changes to each vertebra.

All of the four previously discussed study opted for a supervised model of AI learning.
Seo et al., in 2021 published his article which aimed to compare the performance of 6 CNN algorithms in order to automatically stage CVM on LC radiographs. A total of 600 LC was included in the study, semi-automatic AI training model was applied by using gradient-weighted class activation map (Grad-CAM) technology.

Seo et al.2021, is consider to be the first to implement semi-automatic AI learning model which didn’t use landmarks or reference points in order to learn vertebras morphological features, morphological features were learned using Grad-CAM technology. In the result of this study All deep learning models showed more than 90% accuracy.

The previously conducted studies regarding automatic CVM staging process is limited in number and have a few limitations.

3 out of the 4 mentioned study decided the changes to the vertebra morphology by placing land marks on each vertebra and measure the depth of the vertebra base ,which is a technique not really used in our daily practice.

In our daily practice, morphological changes and shape of vertebras are determined visually.
In our thesis study we opted to fed the vertebra morphology as a shape to the algorithm in order to simulate the visual task done in our daily practice.

A sample larger than any of the previous studies was collected in an attempt to achieve more accuracy by enlarging the training set.

Also, in order to tune up the performance of the algorithm a validation set was added which can round up and shortcut in the algorithm performance.

Despite the attempt to collect a large sample size, we ended up with unequal destitution of the training data.

The smallest group training data was the pubertal stage group (CS3, CS4) which can be related due to the fact of growth spurt being short and quick it can be easily missed.
In general our algorithm performance could be described as good but not excellent, with a drawback being the misclassification of CS3, CS4 with neighbouring stages.

After evaluating the design of the previously conducted studies in this filed, and keeping in mind the previous limitation, we designed our study.

CHAPTER VI

6. Conclusion and Recommendations

Conclusion

Our developed algorithm for this study proved the ability to determine the CVM stage with a good level of success. Furthermore, measurement correlation showed almost a perfect agreement between observers which indicates a high reliability and reproducibility.

AI based algorithms can be used for accurate determination of the CVM stage using Lateral cephalogram in our everyday clinical practice.

Recommendations

A future study with an equally distributed dataset among all maturation ages could provide a more accurate algorithm.

7.References

Alhadlaq, A. M., & Al-Maflehi, N. S. (2013). New model for cervical vertebral bone age estimation in boys. *King Saud University Journal of Dental Sciences*, *4*(1), 1–5. https://doi.org/10.1016/j.ksujds.2012.11.001

ALTAN, B., Kocasaraç, H. D., SİNANOĞLU, E., & MUTAF, H. (2015). Türk bireylerde çeşitli maturasyon indikatörleri arasındaki ilişkilerin değerlendirilmesi. Cumhuriyet Dental Journal, 18(3), 235-248.

Amasya, H., Cesur, E., Yıldırım, D., & Orhan, K. (2020). Validation of cervical vertebral maturation stages: Artificial intelligence vs human observer visual analysis. *American Journal of Orthodontics and Dentofacial Orthopedics: Official Publication of the American Association of Orthodontists, Its Constituent Societies, and the American Board of Orthodontics*, *158*(6), e173–e179. https://doi.org/10.1016/j.ajodo.2020.08.014

Assmus, A. (1995). Early history of X rays. Beam Line, 25(2), 10-24.

Baccetti, T., Franchi, L., & McNamara, J. A. (2005). The Cervical Vertebral Maturation (CVM) Method for the Assessment of Optimal Treatment Timing in Dentofacial Orthopedics. *Seminars in Orthodontics*, *11*(3), 119–129. https://doi.org/10.1053/j.sodo.2005.04.005

Baldin, C. C., Kitt, M., Costa, A. L. F., Yasuda, C. L., Cendes, F., & Nahás-Scocate, A. C. R. (2017). Evaluation of the Skeletal Maturation of Cervical Vertebrae with Magnetic Resonance Imaging: a piloty study. *Brazilian Journal of Oral Sciences*, *16*, 1–8. https://doi.org/10.20396/bjos.v16i0.8650501

Caldas, M. de P., Ambrosano, G. M. B., & Haiter Neto, F. (2007). New formula to objectively evaluate skeletal maturation using lateral cephalometric radiographs. *Brazilian Oral Research*, *21*(4), 330–335. https://doi.org/10.1590/s1806-83242007000400009

Caldas, M. de P., Ambrosano, G. M. B., & Haiter Neto, F. (2010). Computer-assisted analysis of cervical vertebral bone age using cephalometric radiographs in Brazilian subjects. *Brazilian Oral Research*, *24*(1), 120–126. https://doi.org/10.1590/s1806-83242010000100020

Cericato, G. O., Bittencourt, M. A. V., & Paranhos, L. R. (2015). Validity of the assessment method of skeletal maturation by cervical vertebrae: a systematic review and meta-analysis. *Dentomaxillofacial Radiology*, *44*(4), 20140270. https://doi.org/10.1259/dmfr.20140270

Chartrand, G., Cheng, P. M., Vorontsov, E., Drozdzal, M., Turcotte, S., Pal, C. J., Kadoury, S., & Tang, A. (2017). Deep Learning: A Primer for Radiologists. *RadioGraphics*, *37*(7), 2113–2131. https://doi.org/10.1148/rg.2017170077

Chen, S. K., & Hollender, L. (1995). Digitizing of radiographs with a flatbed scanner. *Journal of Dentistry*, *23*(4), 205–208. https://doi.org/10.1016/0300-5712(95)91183-n

De Clerck, H. J., & Proffit, W. R. (2015). Growth modification of the face: A current perspective with emphasis on Class III treatment. *American Journal of Orthodontics and Dentofacial Orthopedics*, *148*(1), 37–46. https://doi.org/10.1016/j.ajodo.2015.04.017

Demirturk Kocasarac, H., Altan, A. B., Yerlikaya, C., Sinanoglu, A., & Noujeim, M. (2016). Correlation between spheno-occipital synchondrosis, dental age, chronological age and cervical vertebrae maturation in Turkish population: is there a link? *Acta Odontologica Scandinavica*, *75*(2), 79–86. https://doi.org/10.1080/00016357.2016.1255352

DiBiase, A. (2002). The Timing of Orthodontic Treatment. *Dental Update*, *29*(9), 434–441. https://doi.org/10.12968/denu.2002.29.9.434

Durka-Zając, M., Mituś-Kenig, M., Derwich, M., Marcinkowska-Mituś, A., & Łoboda, M. (2016). Radiological Indicators of Bone Age Assessment in Cephalometric Images. Review. *Polish Journal of Radiology*, *81*, 347–353. https://doi.org/10.12659/PJR.895921

Dzemidzic, V., Sokic, E., Tiro, A., & Nakas, E. (2015). Computer Based Assessment of Cervical Vertebral Maturation Stages Using Digital Lateral Cephalograms. *Acta Informatica Medica*, *23*(6), 364. https://doi.org/10.5455/aim.2015.23.364-368

Erickson, B. J., Korfiatis, P., Akkus, Z., & Kline, T. L. (2017). Machine Learning for Medical Imaging. *RadioGraphics*, *37*(2), 505–515. https://doi.org/10.1148/rg.2017160130

Erickson, B. J. (2019). Deep Learning and Machine Learning in Imaging: Basic Principles. *Artificial Intelligence in Medical Imaging*, 39–46. https://doi.org/10.1007/978-3-319-94878-2_4

Fleming, P. (2017). Timing orthodontic treatment: early or late? *Australian Dental Journal*, *62*, 11–19. https://doi.org/10.1111/adj.12474

Fleming, P. S., Johal, A., & DiBiase, A. T. (2008). Managing Malocclusion in the Mixed Dentition: Six Keys to Success Part 1. *Dental Update*, *35*(9), 607–613. https://doi.org/10.12968/denu.2008.35.9.607

Hassel, B., & Farman, A. G. (1995). Skeletal maturation evaluation using cervical vertebrae. *American Journal of Orthodontics and Dentofacial Orthopedics*, *107*(1), 58–66. https://doi.org/10.1016/s0889-5406(95)70157-5

Hoseini, M., Zamaheni, S., Bashizadeh Fakhar, H., Akbari, F., Chalipa, J., & Rahmati, A. (2016). Comparative Evaluation of the Efficacy of Hand-Wrist and

Cervical Vertebrae Radiography for the Determination of Skeletal Age. *Iranian Journal of Radiology*, *13*(3), e21695. https://doi.org/10.5812/iranjradiol.21695

Jaqueira, L. M. F., Armond, M. C., Pereira, L. J., Alcântara, C. E. P. de, & Marques, L. S. (2010). Determining skeletal maturation stage using cervical vertebrae: evaluation of three diagnostic methods. *Brazilian Oral Research*, *24*(4), 433–437. https://doi.org/10.1590/s1806-83242010000400010

Joshi, V. V., Iyengar, A. R., Nagesh, K. S., & Gupta, J. (2010). Comparative study between cervical vertebrae and hand-wrist maturation for the assessment of skeletal age. *Archives of Oral Research*, *6*(3). https://doi.org/10.7213/aor.v6i3.23157

Kama, J. D., Aslan, S. G., Darı, O., & Özer, T. (2006). Erkek bireylerde servikal vertebra kemik yaşının kronolojik ve iskelet yaş ile karşılaştırılması. Dicle Tıp Dergisi, 33(1), 36-41.

Kasımoğlu, Y., & Tuna-İnce, E. B. (2015). Diş hekimliğinde kemik yaşı tayininde kullanılan yöntemler: derleme. *Acta Odontologica Turcica*, *33*(2). https://doi.org/10.17214/aot.65918

Khanna, S. S., & Dhaimade, P. A. (2017). Artificial intelligence: transforming dentistry today. Indian J Basic Appl Med Res, 6(3), 161-167.

Korde, S. J., Daigavane, P., & Shrivastav, S. (2015). Skeletal Maturity Indicators-Review Article. Int. J. Sci. Res, 6, 361-370

Koukouviti, M.-M. C. P., Ioanna M; Zafeiriadis, Anastasios A; Chatzigianni, Athina. (2017). Hand-wrist radiograph and lateral cephalometric radiograph as assessment tools of skeletal maturation. A literature review. HELLENIC ORTHODONTIC REVIEW, 20(1 & 2).

Koçak, N. (2018). ADLİ DİŞ HEKİMLİĞİNDE YAŞ TAHMİNİNDE KULLANILAN MORFOLOJİK, BİYOKİMYASAL VE RADYOGRAFİK YÖNTEMLER. In İ. Uzel (Ed.), Diş Hekimliği: Akademisyen Kitabevi.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 25.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. Communications of the ACM, 60(6), 84-90.

Kunz, F., Stellzig-Eisenhauer, A., Zeman, F., & Boldt, J. (2019). Artificial intelligence in orthodontics. *Journal of Orofacial Orthopedics / Fortschritte Der Kieferorthopädie*, *81*(1), 52–68. https://doi.org/10.1007/s00056-019-00203-8

Kök, H., Acilar, A. M., & İzgi, M. S. (2019). Usage and comparison of artificial intelligence algorithms for determination of growth and development by cervical vertebrae stages in orthodontics. *Progress in Orthodontics*, *20*(1), 41. https://doi.org/10.1186/s40510-019-0295-8

Kök, H., Izgi, M. S., & Acilar, A. M. (2020). Determination of growth and development periods in orthodontics with artificial neural network. *Orthodontics & Craniofacial Research*. https://doi.org/10.1111/ocr.12443

LAMPALSKI, D. (1972). Skeletal age assessment utilizing cervical vertebrae. Master of Science Thesis, University of Pittsburgh.

Richard, L. J., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Le Cun, Y., Jackel, L. D., Boser, B., Denker, J. S., Graf, H. P., Guyon, I., Henderson, D., Howard, R. E., & Hubbard, W. (1990). Handwritten Digit Recognition: Applications of Neural Net Chips and Automatic Learning. *Neurocomputing*, 303–318. https://doi.org/10.1007/978-3-642-76153-9_35

Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. https://doi.org/10.1109/cvpr.2015.7298965

McNamara, J. A., & Franchi, L. (2018). The cervical vertebral maturation method: A user's guide. *The Angle Orthodontist*, *88*(2), 133–143. https://doi.org/10.2319/111517-787.1

Mehta, S., Dresner, R., Gandhi, V., Chen, P.-J., Allareddy, V., Kuo, C.-L., Mu, J., & Yadav, S. (2020). Effect of positional errors on the accuracy of cervical vertebrae maturation assessment using CBCT and lateral cephalograms. *Journal of the World Federation of Orthodontists*, *9*(4), 146–154. https://doi.org/10.1016/j.ejwf.2020.09.006

Mito, T., Sato, K., & Mitani, H. (2002). Cervical vertebral bone age in girls. *American Journal of Orthodontics and Dentofacial Orthopedics*, *122*(4), 380–385. https://doi.org/10.1067/mod.2002.126896

Nestman, T. S., Marshall, S. D., Qian, F., Holton, N., Franciscus, R. G., & Southard, T. E. (2011). Cervical vertebrae maturation method morphologic criteria: Poor reproducibility. *American Journal of Orthodontics and Dentofacial Orthopedics*, *140*(2), 182–188. https://doi.org/10.1016/j.ajodo.2011.04.013

Padalino, S., Sfondrini, M. F., Chenuil, L., Scudeller, L., & Gandini, P. (2014). Reliability of skeletal maturity analysis using the cervical vertebrae maturation method on dedicated software. *International Orthodontics*, *12*(4), 483–493. https://doi.org/10.1016/j.ortho.2014.10.003

Predko-Engel, A., Kaminek, M., Langova, K., Kowalski, P., & Fudalej, P. S. (2015). Reliability of the cervical vertebrae maturation (CVM) method. *Bratislava Medical Journal*, *116*(04), 222–226. https://doi.org/10.4149/bll_2015_043

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *Lecture Notes in Computer Science*, 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

Rubia-Bullen, I. R. F., Escarpinati, M. C., Schiabel, H., Vieira, M. A. da C., Rubira, C. M. F., & Lauris, J. R. P. (2007). Digitizing radiographic films: a simple way to evaluate indirect digital images. *Journal of Applied Oral Science*, *15*(1), 14–17. https://doi.org/10.1590/s1678-77572007000100004

Santiago, R. C., Cunha, A. R., Júnior, G. C., Fernandes, N., Campos, M. J. S., Costa, L. F. M., Vitral, R. W. F., & Bolognese, A. M. (2014). New software for cervical vertebral geometry assessment and its relationship to skeletal maturation—a pilot study. *Dentomaxillofacial Radiology*, *43*(2), 20130238. https://doi.org/10.1259/dmfr.20130238

Seedat, A. K., & Forsberg, C. D. (2005). An evaluation of the third cervical vertebra (C3) as a growth indicator in Black subjects. SADJ: journal of the South African Dental Association= tydskrif van die Suid-Afrikaanse Tandheelkundige Vereniging, 60(4), 156-158.

Seo, H., Hwang, J., Jeong, T., & Shin, J. (2021). Comparison of Deep Learning Models for Cervical Vertebral Maturation Stage Classification on Lateral Cephalometric Radiographs. *Journal of Clinical Medicine*, *10*(16), 3591. https://doi.org/10.3390/jcm10163591

SHAH, A., UL HASSAN, F. A. I. Z. U. L., Hussain, U., & ZAHRA, F. T. (2016). INTER-OBSERVERS LEVEL OF AGREEMENT IN CERVICAL VERTEBRAL MATURATION STAGING. Pakistan Oral & Dental Journal, 36(2)

Silva, M. B. G. da, & Sant'Anna, E. F. (2013). The evolution of cephalometric diagnosis in Orthodontics. *Dental Press Journal of Orthodontics*, *18*(3), 63–71. https://doi.org/10.1590/s2176-94512013000300011

Szemraj-Folmer, A., Wojtaszek-Słomińska, A., Racka-Pilszak, B., & Kuc-Michalska, M. (2020). Assessment of the duration of the pubertal growth spurt in patients with skeletal open bite. *Journal of Orofacial Orthopedics / Fortschritte Der Kieferorthopädie*, *82*(2), 92–98. https://doi.org/10.1007/s00056-020-00262-2

Turkoz, C., Kaygisiz, E., Ulusoy, C., & Ates, C. (2017). A practical formula for determining growth. *Diagnostic and Interventional Radiology*, *23*(3), 194–198. https://doi.org/10.5152/dir.2016.16334

Wong, R. W. K., Alkhal, H. A., & Rabie, A. B. M. (2009). Use of cervical vertebral maturation to determine skeletal age. *American Journal of Orthodontics and Dentofacial Orthopedics*, *136*(4), 484.e1–484.e6. https://doi.org/10.1016/j.ajodo.2007.08.033

Zhou, J., Zhou, H., Pu, L., Gao, Y., Tang, Z., Yang, Y., You, M., Yang, Z., Lai, W., & Long, H. (2021). Development of an Artificial Intelligence System for the Automatic Evaluation of Cervical Vertebral Maturation Status. *Diagnostics (Basel, Switzerland)*, *11*(12), 2200. https://doi.org/10.3390/diagnostics11122200

Öztemel, E. (2003). Yapay sinir ağlari. PapatyaYayincilik, Istanbul.

8.Appendices

YAKIN DOĞU ÜNİVERSİTESİ
BİLİMSEL ARAŞTIRMALAR ETİK KURULU

ARAŞTIRMA PROJESİ DEĞERLENDİRME RAPORU

Toplantı Tarihi : 29.07.2021
Toplantı No **:** 2021/93
Proje No **:**1382

Yakın Doğu Üniversitesi Diş Hekimliği Fakültesi öğretim üyelerinden Doç. Dr. Levent Vahdettin'in sorumlu araştırmacısı olduğu, YDU/2021/93-1382 proje numaralı ve **"Artificial intelligence-based cephalometric landmark detection and cervical vertebrae maturation stage assessment"** başlıklı proje önerisi kurulumuzca online toplantıda değerlendirilmiş olup, etik olarak uygun bulunmuştur.

Y

Prof. Dr. Rüştü Onur
Yakın Doğu Üniversitesi
Bilimsel Araştırmalar Etik Kurulu Başkanı

Appendix X

Turnitin Similarity Report

thesis

ORIGINALITY REPORT

15%	10%	10%	7%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.mdpi.com Internet Source	2%
2	d.researchbib.com Internet Source	2%
3	tutorialforbeginner.com Internet Source	1%
4	Hakan Amasya, Emre Cesur, Derya Yıldırım, Kaan Orhan. "Validation of cervical vertebral maturation stages: Artificial intelligence vs human observer visual analysis", American Journal of Orthodontics and Dentofacial Orthopedics, 2020 Publication	1%
5	www.researchgate.net Internet Source	1%
6	Shivam Mehta, Rebecca Dresner, Vaibhav Gandhi, Po-Jung Chen, Veerasathpurush Allareddy, Chia-Ling Kuo, Jinjian Mu, Sumit Yadav. "Effect of positional errors on the accuracy of cervical vertebrae maturation	1%

	assessment using CBCT and lateral cephalograms", Journal of the World Federation of Orthodontists, 2020 Publication	
7	Hakan Amasya, Derya Yildirim, Turgay Aydogan, Nazan Kemaloglu, Kaan Orhan. "Cervical vertebral maturation assessment on lateral cephalometric radiographs using artificial intelligence: comparison of machine learning classifier models", Dentomaxillofacial Radiology, 2020 Publication	<1%
8	pubs.rsna.org Internet Source	<1%
9	Submitted to Kyungpook National University Student Paper	<1%
10	Submitted to Yakın Doğu Üniversitesi Student Paper	<1%
11	Submitted to Universiti Sains Malaysia Student Paper	<1%
12	Yuguo Zhou, Tong Mu, Zhong-Hua Pang, Changbing Zheng. "A survey on hyper basis function neural networks", Systems Science & Control Engineering, 2019 Publication	<1%
13	Submitted to University of Sydney Student Paper	<1%

14	Submitted to University of Huddersfield Student Paper	<1%
15	Çağla Sin, Nurullah Akkaya, Seçil Aksoy, Kaan Orhan, Ulaş Öz. "A Deep Learning Algorithm Proposal to Automatic Pharyngeal Airway Detection and Segmentation on CBCT Images", Orthodontics & Craniofacial Research, 2021 Publication	<1%
16	Submitted to Piri Reis University Student Paper	<1%
17	Franchi, L.. "Phases of the dentition for the assessment of skeletal maturity: A diagnostic performance study", American Journal of Orthodontics & Dentofacial Orthopedics, 200803 Publication	<1%
18	Submitted to National College of Ireland Student Paper	<1%
19	Submitted to Noida Institute of Engineering and Technology Student Paper	<1%
20	Rabia Saleem, Jamal Hussain Shah, Muhammad Sharif, Mussarat Yasmin, Hwan-Seung Yong, Jaehyuk Cha. "Mango Leaf Disease Recognition and Classification Using	<1%

Novel Segmentation and Vein Pattern Technique", Applied Sciences, 2021
Publication

21	ema.cri-info.cm Internet Source	<1%
22	mdpi-res.com Internet Source	<1%
23	Rachid Hedjam, Reza Farrahi Moghaddam, Mohamed Cheriet. "Text extraction from degraded document images", 2010 2nd European Workshop on Visual Information Processing (EUVIP), 2010 Publication	<1%
24	Seda Arslan Tuncer, Hakan Ayyıldız, Mehmet Kalaycı, Taner Tuncer. "Scat-NET: COVID-19 diagnosis with a CNN model using scattergram images", Computers in Biology and Medicine, 2021 Publication	<1%
25	www.ncbi.nlm.nih.gov Internet Source	<1%
26	"Bildverarbeitung für die Medizin 2018", Springer Nature, 2018 Publication	<1%
27	"Maxillofacial Cone Beam Computed Tomography", Springer Science and Business Media LLC, 2018	<1%

Publication

28	Submitted to University of Technology Student Paper	<1%
29	pdfs.semanticscholar.org Internet Source	<1%
30	repository-tnmgrmu.ac.in Internet Source	<1%
31	www.irjmets.com Internet Source	<1%
32	openaccess.altinbas.edu.tr Internet Source	<1%
33	Submitted to Liverpool John Moores University Student Paper	<1%
34	hdl.handle.net Internet Source	<1%
35	Sang-Cheol Seok, Elizabeth McDevitt, Sara C. Mednick, Paola Malerba. "Global and non-Global slow oscillations differentiate in their depth profiles", Frontiers in Network Physiology, 2022 Publication	<1%
36	Submitted to University of Illinois at Urbana-Champaign Student Paper	<1%

37 Submitted to Asia Pacific Instutute of Information Technology
Student Paper
<1%

38 Rodrigo César Santiago, Luiz Felipe de Miranda Costa, Robert Willer Farinazzo Vitral, Marcelo Reis Fraga et al. "Cervical vertebral maturation as a biologic indicator of skeletal maturity", The Angle Orthodontist, 2012
Publication
<1%

39 www.scribd.com
Internet Source
<1%

40 neteducate.org
Internet Source
<1%

41 Carol Lynn Curchoe, Adolfo Flores-Saiffe Farias, Gerardo Mendizabal-Ruiz, Alejandro Chavez-Badiola. "Evaluating predictive models in reproductive medicine", Fertility and Sterility, 2020
Publication
<1%

42 Chen, J.. "Correlation between dental maturity and cervical vertebral maturity", Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology, 201012
Publication
<1%

43 Xipeng Pan, Dengxian Yang, Lingqiao Li, Zhenbing Liu, Huihua Yang, Zhiwei Cao, Yubei He, Zhen Ma, Yiyi Chen. "Cell detection in
<1%

pathology and microscopy images with multi-scale fully convolutional neural networks", World Wide Web, 2018
Publication

44	open.uct.ac.za Internet Source	<1 %
45	www.slideshare.net Internet Source	<1 %
46	"Online Only Abstracts", American Journal of Orthodontics and Dentofacial Orthopedics, 2020 Publication	<1 %
47	"Recent Advances on Soft Computing and Data Mining", Springer Science and Business Media LLC, 2017 Publication	<1 %
48	Hae Ok Lee, John P. Heller. "Carbon Dioxide—Foam Mobility Measurements at High Pressure", American Chemical Society (ACS), 1988 Publication	<1 %
49	Khaoula Telahigue, Imen Chetoui, Imen Rabeh, Med Salah Romdhane, M'hamed El Cafsi. "Comparative fatty acid profiles in edible parts of wild scallops from the Tunisian coast", Food Chemistry, 2010 Publication	<1 %

50	Mohand Tuffaha. "Adoption Factors of Artificial intelligence in Human Resource Management", Universitat Politecnica de Valencia, 2022 Publication	<1%
51	Narayan Kabra, Gianluca Serra, Miguel Lozano, Matteo Tommasini, Karthik Sankaranarayanan, Nisha Rani Agarwal. "Classification of Cannabinoid Spectra Using Machine Learning", American Chemical Society (ACS), 2022 Publication	<1%
52	Saeed Mohsen, Ahmed Elkaseer, Steffen G. Scholz. "Industry 4.0-Oriented Deep Learning Models for Human Activity Recognition", IEEE Access, 2021 Publication	<1%
53	www.javatpoint.com Internet Source	<1%
54	"Image Analysis and Processing - ICIAP 2017", Springer Science and Business Media LLC, 2017 Publication	<1%
55	Cansu Buyuk, Nurullah Akkaya, Belde Arsan, Gurkan Unsal, Secil Aksoy, Kaan Orhan. "A Fused Deep Learning Architecture for the Detection of the Relationship between the	<1%

9. CV

1. Name Surname: MOHAMAD TALAL RADWAN

2. Date of Birth: 27.01.1993

3. Title: Dentist

4. State of Education: PhD

5. Current Institution: Near East University, Faculty of Dentistry, Department of Orthodontics

Degree	**Department**	**University**	**Date**
Bachelor' s Degree	Faculty of Dentistry	Kalamoon University	2010-2015
PhD	Orthodontics Department	Near East University	2017-2022

6. Publication:

Radwan, M. T., Sin, Ç., Akkaya, N., & Vahdettin, L. (2022). Artificial Intelligence-Based Algorithm for Cervical Vertebrae Maturation Stage Assessment. *Orthodontics & Craniofacial Research*.