



**ATLAS VERTEBRA GÖRÜNTÜLERİNİN
GÖRÜNTÜ İŞLEME İLE OTOMATİK
MORFOLOJİK ÖLÇÜMÜ VE MAKİNE
ÖĞRENMESİ CİNSİYET TAHMİNİ MODELİ**

Muhammed Çağrı ÇELİK

**2021
YÜKSEK LİSANS TEZİ BİLGİSAYAR
MÜHENDİSLİĞİ**

**Tez Danışmanı
Dr. Öğr. Üyesi Caner ÖZCAN**

**ATLAS VERTEBRA GÖRÜNTÜLERİNİN GÖRÜNTÜ İŞLEME İLE
OTOMATİK MORFOLOJİK ÖLÇÜMÜ VE MAKİNE ÖĞRENMESİ
CİNSİYET TAHMİNİ MODELİ**

Muhammed Çağrı ÇELİK

**T.C.
Karabük Üniversitesi
Lisansüstü Eğitim Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalında
Yüksek Lisans Tezi
Olarak Hazırlanmıştır**

**Tez Danışmanı
Dr. Öğr. Üyesi Caner ÖZCAN**

**KARABÜK
Aralık 2021**

Muhammed Çağrı ÇELİK tarafından hazırlanan “ATLAS VERTEBRA GÖRÜNTÜLERİNİN GÖRÜNTÜ İŞLEME İLE OTOMATİK MORFOLOJİK ÖLÇÜMÜ VE MAKİNE ÖĞRENMESİ CİNSİYET TAHMİNİ MODELİ” başlıklı bu tezin Yüksek Lisans Tezi olarak uygun olduğunu onaylarım.

Dr. Öğr. Üyesi Caner ÖZCAN

Tez Danışmanı, Bilgisayar Mühendisliği Anabilim Dalı

Bu çalışma, jürimiz tarafından Oy Birliği ile Bilgisayar Mühendisliği Anabilim Dalında Yüksek Lisans tezi olarak kabul edilmiştir. 29/12/2021

Ünvanı, Adı SOYADI (Kurumu)

İmzası

Başkan : Dr. Öğr. Üyesi Emrah ÖZKAYNAK (KBÜ)

Üye : Dr. Öğr. Üyesi Caner ÖZCAN (KBÜ)

Üye : Dr. Öğr. Üyesi Emel SOYLU (SAMÜ)

KBÜ Lisansüstü Eğitim Enstitüsü Yönetim Kurulu, bu tez ile, Yüksek Lisans derecesini onamıştır.

Prof. Dr. Hasan SOLMAZ

Lisansüstü Eğitim Enstitüsü Müdürü

“Bu tezdeki tüm bilgilerin akademik kurallara ve etik ilkelere uygun olarak elde edildiğini ve sunulduğunu; ayrıca bu kuralların ve ilkelerin gerektirdiği şekilde, bu çalışmadan kaynaklanmayan bütün atıfları yaptığımı beyan ederim.”

Muhammed Çağrı ÇELİK

ÖZET

Yüksek Lisans Tezi

ATLAS VERTEBRA GÖRÜNTÜLERİNİN GÖRÜNTÜ İŞLEME İLE OTOMATİK MORFOLOJİK ÖLÇÜMÜ VE MAKİNE ÖĞRENMESİ CİNSİYET TAHMİNİ MODELİ

Muhammed Çağrı ÇELİK

Karabük Üniversitesi

Lisansüstü Eğitim Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı

Tez Danışmanı:

Dr. Öğr. Üyesi Caner ÖZCAN

Aralık 2021, 55 sayfa

Bilgisayarlı Tomografi ile elde edilen görüntülerden cinsiyet tahmini yapılabilmesi için analiz, ölçüm ve ölçüm sonucu elde edilecek özneliklerin belirlenmesi temel süreçleri oluşturmaktadır. Atlas Vertebra kemiği üzerinde yapılan morfolojik ölçümler ile kemik analiz işlemleri gerçekleştirilmektedir. Bu ölçümler manuel olarak el ile ya da görüntüleme teknolojileri kullanılarak bilgisayar yazılımları ile ölçü alınarak yapılmaktadır. El ile yapılan ölçümler oldukça zaman alıcı işlemler iken diğer tarafta dijital ölçümler zaman maliyeti açısından daha avantajlıdır. Ancak dijital ölçümde kullanılan cihazın özellikleri, görüntüyü alacak uzmanın bilgi birikimi ve tecrübesi gibi faktörler sebebiyle görüntülerin kontrastlarının, çözünürlüklerinin, görüntünün alındığı alanın farklılıklar gösterdiği tespit edilmiştir. Böyle durumlarda ölçümden önce görüntüler üzerinde ön işlemlerin yapılması aşamaları uzmanlık gerektirmekte, insan kaynaklı bu düzeltmelerde elde edilen verilerde değişimlere sebep olmaktadır.

Bu çalışmada insan faktörü ortadan kaldırılarak kısa sürede yüksek doğrulukla sonuca ulaşmak hedeflenmiştir. Görüntüler üzerinde yapılan otomatik morfolojik ölçümler sonucu öznitelikler belirlenmiş ve makine öğrenmesi sınıflandırma algoritmalarından Naive Bayes, Karar Ağacı, Destek Vektör Makineleri, K-En Yakın Komşuluk, Rastgele Orman yöntemleri kullanarak cinsiyet tahmini yapılmıştır. Kullanılan bu algoritmalar ile deneysel karşılaştırmalı sonuçlar ortaya konularak grafik ve çizelgeler ile detaylı bir şekilde gösterilmiştir. Karşılaştırılmalı sonuçlar incelendiğinde en yüksek başarımın elde edildiği yöntemin %82,14 ile Gaussian Naive Bayes olduğu görülmüştür.

Anahtar Sözcükler : Makine öğrenmesi, görüntü işleme, OpenCV kütüphanesi, cinsiyet tahmini, morfometrik ölçüm, atlas vertebra

Bilim Kodu : 92432.

ABSTRACT

M. Sc. Thesis

AUTOMATIC MORPHOLOGICAL MEASUREMENT OF ATLAS VERTEBRA IMAGES WITH IMAGE PROCESSING AND MACHINE LEARNING GENDER PREDICTION MODEL

Muhammed Çağrı ÇELİK

**Karabük University
Institute of Graduate Programs
Department of Computer Engineering**

Thesis Advisor:

Assist. Prof. Dr. Caner ÖZCAN

December 2021, 55 pages

Analysis, measurement and determination of the features to be obtained as a result of the measurement constitute the basic processes in order to make an estimation of gender from the images obtained by Computed Tomography. Bone analysis procedures are carried out with morphological measurements on Atlas Vertebra bone. These measurements are made manually or using imaging technologies (Digital). Manual measurements are very time consuming. Digital measurements are more advantageous in terms of time costs. However, due to factors such as the features of the device used in digital measurement, the knowledge and experience of the expert who will take the image, it has been determined that the contrasts, resolutions of the images and the area where the image is taken differ. In such cases, preprocessing steps on images before measurement requires expertise, and these human-induced corrections cause changes in the data obtained.

In this study, it is aimed to reach the result with high accuracy in a short time by eliminating the human factor. As a result of automatic morphological measurements made on the images, the features were obtained and gender estimation was made using the machine learning classification algorithms, Naive Bayes, Decision Tree, Support Vector Machines, K-Nearest Neighbors, Random Forest methods. Experimental comparative results were produced with these algorithms used. Experimental comparative results with these algorithms are presented and shown in detail with graphs and charts. When the comparative results were examined, it was seen that the method with the highest success was Gaussian Naive Bayes with %82,14.

Key Word : Machine learning, image processing, OpenCV library, gender prediction, morphometric measurement, atlas vertebra.

Science Code : 92432.

TEŐEKKÜR

Bu tez alıőmasının planlanmasında ve yürütülmesinde ilgi ve desteęini esirgemeyen, bilgi ve tecrübeleri, yönlendirme ve bilgilendirmeleriyle alıőmamı bilimsel temellere dayandırmamı saęlayan sayın hocam Dr. Öğr. Üyesi Caner ÖZCAN'a teşekkürlerimi sunarım.

alıőmada kullandığım görüntüleri benimle paylaşan, alıőmam boyunca tüm problemlerde desteęi ile yön veren, bilgi ve tecrübelerini hiç esirgemeyen Karabük Üniversitesi Diő Hekimlięi Fakültesi Aęız Diő ve ene Radyolojisi Ana Bilim Dalı Dr. Öğr. Üyesi Adem PEKİNCE'ye teşekkür ederim.

Manevi desteęini her an olsun eksik etmeyen, hiçbir yardımı esirgemedен, yanımda olan sevgili aileme tüm kalbimle teşekkür ederim.

Zorlu her süreçte olduęu gibi tez dönemi boyunca da varlıęı ve desteęi ile yanımda olan, alıőmamın her anında inancı ile alıőmama ilham katan sevgili eőim Büőra ELİK'e sonsuz teşekkürlerimi sunarım.

İÇİNDEKİLER

	<u>Sayfa</u>
KABUL.....	ii
ÖZET.....	iv
ABSTRACT.....	vi
TEŞEKKÜR.....	viii
İÇİNDEKİLER	ix
ŞEKİLLER DİZİNİ.....	xi
ÇİZELGELER DİZİNİ	xiii
SİMGELER VE KISALTMALAR DİZİNİ	xiv
BÖLÜM 1	1
GİRİŞ.....	1
BÖLÜM 2	3
GENEL BİLGİLER	3
2.1. KİMLİK	3
2.2. KİMLİKLENDİRME.....	3
2.3. MORFOMETRİ VE MORFOMETRİK ÖLÇÜM.....	4
2.4. ATLAS VERTEBRA.....	5
2.5. RADYOLOJİK GÖRÜNTÜLEME YÖNTEMLERİNDE BİLGİSAYARLI TOMOGRAFİ	6
BÖLÜM 3	8
VERİ SETİ, GÖRÜNTÜ İŞLEME VE MAKİNE ÖĞRENMESİ YÖNTEMLERİ ...	8
3.1.GÖRÜNTÜ İŞLEME.....	9
3.2.MAKİNE ÖĞRENİMİ.....	11
3.2.1.Naive Bayes	12
3.2.2. Destek Vektör Makinesi.....	13

	<u>Sayfa</u>
3.2.3. K-En Yakın Komşu.....	15
3.2.4. Karar Ağaçları.....	16
3.2.5. Rastgele Orman.....	18
BÖLÜM 4	21
KONİK IŞINLI BİLGİSAYARLI TOMOGRAFİ GÖRÜNTÜLERİNİN MAKİNE ÖĞRENMESİ İÇİN HAZIRLANMASI.....	21
BÖLÜM 5	34
MAKİNE ÖĞRENMESİ İLE CİNSİYET TAHMİNİ SONUÇLARI.....	34
BÖLÜM 6	48
SONUÇLAR VE ÖNERİLER.....	48
KAYNAKLAR.....	50
ÖZGEÇMİŞ	55

ŞEKİLLER DİZİNİ

Sayfa

Şekil 2.1. Femura ait morfometrik çalışmalar, parametrelerin çizgisel ve dijital gösterimi.	4
Şekil 2.2. Atlas Vertebra'ya ait görüntü.....	5
Şekil 2.3. Konik ışımlı bilgisayarlı tomografi.....	6
Şekil 3.1. Atlas Vertebra görüntü verisinin hazırlık aşaması.....	8
Şekil 3.2. Görüntü işleme ile geliştirilmiş örnek sistemin şematik görünümü.	9
Şekil 3.3. OpenCV kütüphanesi temel yapısı.	10
Şekil 3.4. Makine öğrenmesi işleyiş adımları.	12
Şekil 3.5. K-EYK algoritma yapısı.	15
Şekil 3.6. Karar ağacı yapısı.	17
Şekil 4.1. Yöntem ve Tasarım.....	21
Şekil 4.2. KIBT ile elde edilen Atlas Vertebra görüntüleri.....	22
Şekil 4.3. Histogram eşitleme işlemi öncesi örnek görüntüler ve histogram grafikleri	24
Şekil 4.4. Histogram eşitleme işlemi sonrası örnek görüntüler ve histogram grafikleri.	25
Şekil 4.5. Atlas Vertebra tespit işlemleri	26
Şekil 4.6. Atlas Vertebra tespit sonucu.	28
Şekil 4.7. Atlas Vertebra görüntüsünün morfometrik ölçüme hazırlık aşamasına örnekler.....	29
Şekil 4.8. Atlas Vertebra çevresinin çizilmesi.	30
Şekil 4.9. Atlas Vertebra morfometrik ölçümlerinin çizim örnekleri.	31
Şekil 5.1. Karmaşıklık matrisi.....	35
Şekil 5.2. Test – eğitim veri seti bölümlenmesi örneği.....	36
Şekil 5.3. Gaussian Naive Bayes çapraz doğrulama yöntemi grafiği.	38
Şekil 5.4. Destek Vektör Makineleri çapraz doğrulama yöntemi grafiği.	40
Şekil 5.5. K en yakın komşuluk çapraz doğrulama yöntemi grafik örneği.....	41

Sayfa

Şekil 5.6. Karar ağaçları çapraz doğrulama yöntemi grafik örneği. 43
Şekil 5.7. Rastgele orman çapraz doğrulama yöntemi grafik örneği. 46

ÇİZELGELER DİZİNİ

Sayfa

Çizelge 5.1. Gaussian Naive Bayes yönteminde test – eğitim setleri bölümlemesine göre başarı sonuçları.	37
Çizelge 5.2. Gaussian Naive Bayes hata metrikleri tablosu.....	37
Çizelge 5.3. Gaussian Naive Bayes yöntemi sonucu oluşan karmaşıklık matrisi.	38
Çizelge 5.4. Destek Vektör Makineleri yönteminde test – eğitim setleri bölümlemesine göre başarı sonuçları.....	39
Çizelge 5.5. Destek Vektör Makineleri hata metrikleri tablosu.	39
Çizelge 5.6. Destek Vektör Makineleri yöntemi sonucu oluşan karmaşıklık matrisi.	40
Çizelge 5.7. K-En Yakın Komşuluk yönteminde test – eğitim setleri bölümlemesine göre başarı sonuçları.	41
Çizelge 5.8. K en yakın komşuluk hata metrikleri tablosu.	42
Çizelge 5.9. K en yakın komşuluk yöntemi sonucu oluşan karmaşıklık matrisi.	42
Çizelge 5.10. Karar ağaçları yönteminde test – eğitim setleri bölümlemesine göre başarı sonuçları.	43
Çizelge 5.11. Karar ağaçları hata metrikleri tablosu.....	44
Çizelge 5.12. Karar ağaçları yöntemi sonucu oluşan karmaşıklık matrisi.	44
Çizelge 5.13. Rastgele orman yönteminde test – eğitim setleri bölümlemesine göre başarı sonuçları.....	45
Çizelge 5.14. Rastgele orman hata metrikleri tablosu.....	46
Çizelge 5.15. Rastgele orman yöntemi sonucu oluşan karmaşıklık matrisi.....	47
Çizelge 5.16. Uygulanan makine öğrenmesi yöntemlerinin test – eğitim setleri bölümlemesine göre başarı sonuçlarının karşılaştırılması.....	47

SİMGELER VE KISALTMALAR DİZİNİ

KISALTMALAR

- KIBT : Konik Işınlı Bilgisayarlı Tomografi
CV : Computer Vision
ROI : Region of Interest
NB : Naive Bayes Algoritması
DVM : Destek Vektör Makineleri Algoritması
K-EYK : K – En Yakın Komşu Algoritması
KA : Karar Ağaçları Algoritması
RO : Rastgele Orman Algoritması
ATM : Atlas Vertebra Kemiğinin Tranvers Çıktıları Arasındaki Mesafe Uzunluğu
APAM : Atlas Vertebra Kemiğinin Ön-Arka Uzunluğu
ADA : Atlas Vertebra Kemiğinin Alanı
ACU : Atlas Vertebra Kemiğinin Çevresi
AFCU : Atlas Vertebra Kemiğinin Vertebral Forameninin Çevre Uzunluğu
AFA : Atlas Vertebra Kemiğinin Vertebral Forameninin Alanı

BÖLÜM 1

GİRİŞ

Bir insanın tanımlanabilmesi için, diğer insanlara göre ayırt edici olarak taşıdığı tüm özelliklere kimlik, yaşayan ya da ölmüş kişilerin bu özellikler kullanılarak tespit edilmesini kimlik tespiti denilmektedir. Kimliklendirilmemiş insan kalıntıları; kurbanların aileleri için hem yasal hem de duygusal boyutta çok sayıda sorun yaratabilmektedir [1]. Bu sebeple günümüzde adli tıp, arkeoloji gibi bilim dallarında yapılan antropolojik ve paleontolojik çalışmalarda insanlara ait kemiklerden çeşitli yöntemlerle kimlik tespiti yapılmaktadır. Yanmış kişilerin belirlenmesi, kayıp yakınlar veya diğer benzer olaylarda; kemikten yaş, cinsiyet vb. özellikler ile kişiye ait bilgilere ulaşılmaktadır. Adli antropologlar, bu tür araştırmalar sonucu oluşturulmuş biyolojik profil ile insan kalıntılarının tanımlanmasını kolaylaştırarak bu alanda önemli bir rol üstlenmektedirler [2,3].

Kimliklendirme denilen kimlik tespiti işlemi birçok aşamadan oluşmaktadır. Bu aşamaların ilk adımı olan cinsiyet tayini ile kimliklendirme işlemlerinde incelenen referans örneklerinin sayısı yarıya kadar düşürülebilmektedir. Bu sebeple kimlik tespiti işlemlerinde cinsiyet tayini aşamasında doğru tespit gerçekleştirmek büyük önem arz etmektedir. Yaş, cinsiyet, atasal yakınlık ve boy uzunluğu analizi gibi işlemler ile biyolojik profillerin oluşturulması, iskelet kalıntılarını kimliklendirme işlemi sırasındaki ilk ve en önemli adımlardan biri olarak kabul edilmektedir.

Biyolojik profil oluşturabilmek için yapılan yaş ve boy uzunluğu tespiti çalışmaları cinsiyete bağlı olarak yapıldığı için, cinsiyet tayini çalışmaları kimliklendirme sürecinin ilk aşamasını oluşturmaktadır [4,5]. Yetişkin kadın ve erkek bireyler arasında cinsiyet farklarını araştıran ve bu konularda birçok yöntem ortaya koyan çalışmalar mevcuttur [6-8]. Bir iskeletin cinsiyetinin pelvis ve kranium kullanarak, yüksek bir doğruluk oranı ile belirlenebileceği belirtilmektedir [9].

Bir başka çalışmada ise humerustan cinsiyet tayini ile ilgili dünyada birçok popülasyona ait çalışmalar yapılmış olduğunu fakat bununla birlikte, Türkiye’de humeruslar üzerinde cinsiyet tayini açısından bir araştırmaya rastlanılmaması sebebiyle bu yönde bir çalışma gerçekleştirmiştir [10]. Yapılan birçok araştırma sonucunda vertebral incelemelerin cinsiyet tayini sonuçları için yüksek doğruluk oranı gösterdiği belirtilmektedir [11]. İncelenen çalışmalarda sakral ve lumbal vertabraların kullanıldığı görülmektedir [12,13]. Tüm bu çalışmalarda kullanılan kemiklerin morfometrik ölçümleri manuel olarak (el ile) ya da bir bilgisayar destekli ölçüm aracı ile ilgili uzman tarafından gerçekleştirilmiştir. Morfometrik çalışmalarda manuel (el ile) ve dijital (sayısal) bilgisayar destekli ölçüm yöntemlerini karşılaştırarak antropometrik ve osteometrik çalışmalarda, manuel (el ile) yöntemler ile yapılan ölçümlerde, çalışmanın niteliği ve içeriğine göre değişen oranlarda kullanılan zaman, bilgisayar destekli dijital (sayısal) ölçüm yöntemleri ile kısalabilmekte ve ölçüm sonuçlarının daha kesin olarak alınabildiği sonucuna ulaşılmaktadır [14]. Adli vakalarda radyolojik görüntüleme bulguların görselleştirilmesinde yaygın olarak kullanılmaktadır. Adli radyoloji kimliklendirme çalışmaları ile tarihsel sürecine başlamakta ve bu süreçte şekillenmektedir [15]. 1970’lerde geliştirilen Bilgisayarlı Tomografi (BT) adli radyolojide damar ve iç organ yaralanmaları, iç kanamalar ve kemik kırıkları gibi durumları saptamada önemli bir yer tutmuş ve halen faydalı bir yöntem olarak kullanılmaktadır. Özellikle kemik dokuların görüntülenmesi açısından etkin bir yöntem olduğu bilinmektedir [16].

Tüm bu çalışmalar ışığında BT ile elde edilen Atlas Vertebra kemiği görüntüleri üzerinde manuel yöntemlerin dışında, bilgisayar destekli görüntü işleme yöntemleri ile ölçüm yapan ve bu ölçüm sonuçlarının makine öğrenmesi yöntemlerinde kullanıldığı, cinsiyet tayininde kullanılabilecek bir yöntem geliştirildi. Bu tez çalışmasında, insan bağımlılığı olmayan, doğruluk oranı yüksek, daha hızlı sonuç alınabilen bir karar mekanizması ortaya çıkarmak amaçlanmıştır.

BÖLÜM 2

GENEL BİLGİLER

Bu bölüm ilk olarak kimlik, kimliklendirme ve cinsiyet tahmininde kullanılan yöntemler anlatılmıştır. İkinci olarak morfometri ve morfometrik ölçümlerden bahsedilmiştir. Daha sonra Atlas Vertebra yapısına ve son olarak medikal görüntüleme alanında kullanılan radyolojik görüntüleme yöntemlerinden konik ışınli bilgisayarlı tomografiye değinilmiştir.

2.1. KİMLİK

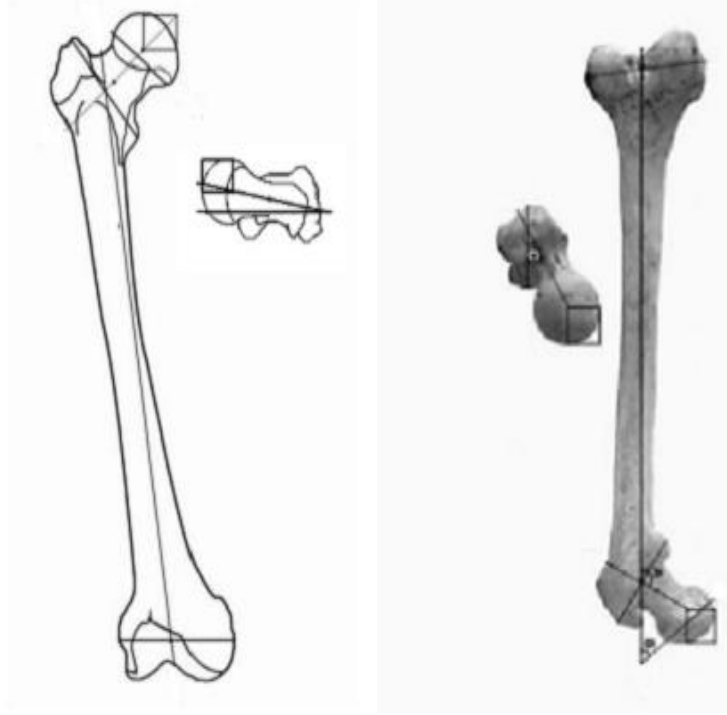
Bir insanın tanımlanmasında diğeri insanlara göre ayırt edici olarak taşıdığı tüm özelliklere kimlik denilmektedir. Kısaca bir insan için sorulan kim sorusuna verilen tüm cevaplardır. Sosyoloji, felsefe, edebiyat, antropoloji gibi birçok bilim dalında temel kavram olarak kullanılmaktadır [17]. Kimlik terimi Türkçe'de tek olmak, hangi kişi olmak olarak ifade edilmektedir [18].

2.2. KİMLİKLENDİRME

Yaşayan ya da ölmüş kişilerin kimlik olarak adlandırılan ayırt edici özellikleri kullanılarak tespit edilmesi kimliklendirme (kimlik tespiti) olarak adlandırılmaktadır. Kimliklendirme adli kimliklendirme ve tıbbi kimliklendirme olarak iki boyutta yürütülmektedir. Adli kimliklendirme yaşayan ya da ölmüş kişi üzerinden çıkan belgelerin değerlendirilmesi ile yapılmaktadır. Tıbbi kimliklendirme ise yaşayan ya da ölmüş kişilerin diğeri kişilerden ayırt edilmesini sağlayan fiziksel yapılar boy, cinsiyet, vücut ağırlığı, saç ve göz rengi, sakal, bıyık, ağız gibi yüz özellikleri, vücudundaki yara ve yanık izleri, radyolojik görüntülerden elde edilen organların büyüklüğü şekli gibi özellikleri dikkate alınmaktadır [19-21].

2.3. MORFOMETRİ VE MORFOMETRİK ÖLÇÜM

Morfometri, cisimlerin veya organların (büyüklük, şekil gibi) varyasyonlar ve değişimleri ile ilgilenen alandır. Şekil farklılıklarından doğan noktalar arası mesafe, açılar ve oranlar ile elde edilen uzunluk, genişlik gibi ölçüm sonuçlarına dayanmaktadır [22] .



Şekil 2.1. Femura ait morfometrik çalışmalar, parametrelerin çizgisel ve dijital gösterimi [14].

Morfometri cinsiyet tayini üzerine yapılan çalışmalarda sıklıkla kullanılmaktadır [10]. Kafatası ve pelvis gibi kemiklerin kimyasal veya mekanik etkenlerle hızlı bozulması sebebiyle femur ve humerus gibi dayanıklı ve uzun kemikler üzerinde yapılan morfometrik çalışmalar cinsiyet tayininde önemli rol oynamıştır [23]. Kafa iskeletinden yaş tespiti üzerine yapılan bir çalışmada sınır tabanlı geometrik morfometrik yöntemler kullanılmıştır [24] .

1962 yılında yapılan bir çalışmada femur morfoloji incelenmiş, femur üzerinde yapılan morfometrik ölçümler sonucunda ırk belirlemesi üzerinde bir sonuç çıkarımı gerçekleştirilmiştir [25]. Morfolojik farklılıklar üzerinde yapılan çalışmalarda morfometrik ölçümler sonucunda çevresel ve kalıtsal faktörlerin etkili olduğu, yetişkin erkeklerin kemiklerinin yetişkin kadınlara göre daha büyük ve sağlam olduğu bilinmektedir. Geçmişten günümüze uzanan tüm bu çalışmalar incelendiğinde morfometri alanında yapılan tüm çalışmalarda morfometrik ölçümler kullanılmıştır.

2.4. ATLAS VERTEBRA

Omurgayı oluşturan 33 kemikten her birine vertebra yani omur denilmektedir. 7 Adet boyun omurundan birinci sıradaki omur Atlas, ikinci sıradaki omur Axis olarak adlandırılmaktadır. Bu iki omur birbirine benzeyen diğer omurdan yapı itibariyle ayrılmaktadır. Atlas Vertebra baş küresini destekleyerek kafatasını taşımaktadır. Atlas ve Axis diğer omurlara göre daha geniş hareket alanı oluşturmak üzere özel bir yapıya sahiptir. Bu iki omur başın öne-arkaya, sağa-sola hareket etmesini sağlar.



Şekil 2.2. Atlas Vertebra'ya ait görüntü [14].

Atlas omuru diğer omurlara göre daha karakteristik bir yapıya sahiptir. Atlas omurunun gövdesi yoktur. Sağ ve sol yanında, yan kitle olarak adlandırılan iki kemik kitlesi mevcuttur. Solunum ve dolaşım merkezlerinin bulunduğu hayati önem taşıyan omurilik parçasını koruma görevini de üstlenmektedir.

Atlas omuru farklı varyasyonlara sahiptir. Konumu itibariyle hayati önem taşımaktadır. Bu sebeple Atlas omuru morfolojisi ile alakalı birçok çalışmaya konu olmuştur [26,27]. Kas, bağ ve kemik dokular ile sarılmış olan Atlas omuru küçük yapısı sebebiyle kafa tabanında gizlenmiş bir halde bulunmaktadır. Atlas omurunun incelenmesi için en sağlıklı yöntem radyolojik görüntüleme teknikleridir. Yaş belirleme amacıyla yapılan morfolojik çalışmalar incelendiğinde boyun omurları diğer kemiklere göre çok daha özel bir yere sahip olmuştur.

2.5. RADYOLOJİK GÖRÜNTÜLEME YÖNTEMLERİNDE BİLGİSAYARLI TOMOGRAFİ

Çene, diş ve yüzdeki çevre dokuların üç boyutlu olarak görüntülenmesini sağlayan görüntüleme aracına konik ışınli bilgisayarlı tomografi denilmektedir (KİBT). Radyasyon dozunun ve maliyetinin düşük olması, az yer kaplaması, kısa sürede üç boyutlu görüntü oluşturmaları sebebiyle yüz ve çevresinde tercih edilen daha çok diş hekimliği alanında kullanılan medikal görüntüleme teknolojisidir [28].



Şekil 2.3. Konik ışınli bilgisayarlı tomografi.

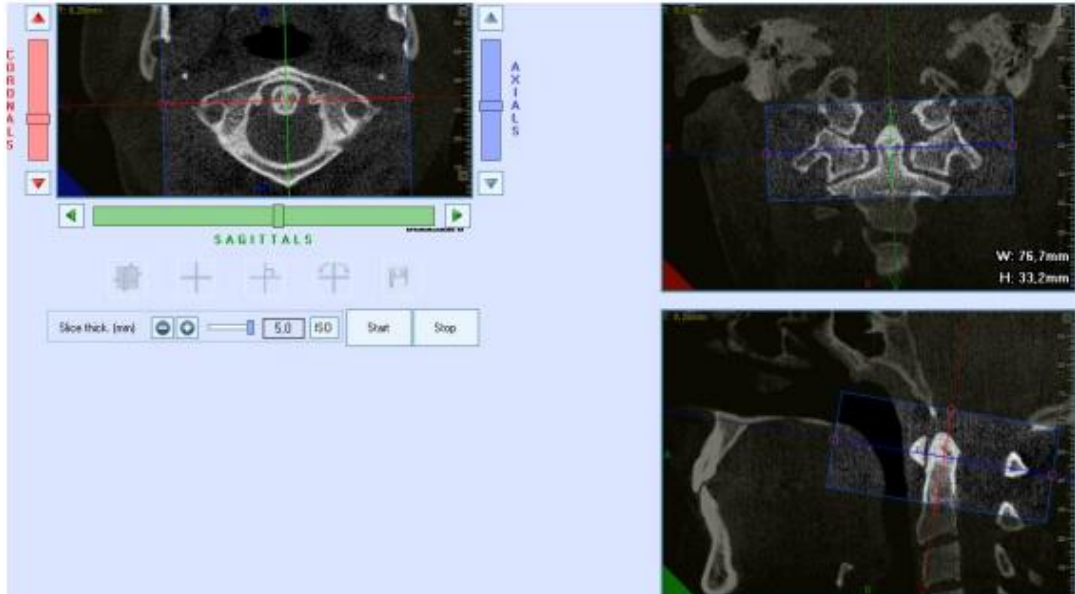
İlk olarak 1982 yılında anjiyografi için üretilen KİBT cihazının gelişimi günümüzde de devam etmektedir [29]. Bilgisayarlı tomografi ile karşılaştırıldığında daha kaliteli

görüntüler elde edilmektedir. Ham veri olarak elde edilen görüntüler bilgisayar destekli uygulamalar aracılığı ile üç boyutlu hacimsel veriye dönüştürülmektedir [30]. Böylelikle üç boyutlu görüntüler kişisel bilgisayar ortamlarına aktarılarak düzenlenmesi ve incelenmesi sağlanmaktadır.

BÖLÜM 3

VERİ SETİ, GÖRÜNTÜ İŞLEME VE MAKİNE ÖĞRENMESİ YÖNTEMLERİ

Bu çalışmada kullanılan görüntü veri seti 2017 yılında gerçekleştirilen tez çalışmasından [31] alınmıştır. Atatürk Üniversitesi Diş Hekimliği Fakültesi Ağız Diş ve Çene Radyolojisi Anabilim Dalı'nda görüntüleri alınmış olan 18 yaş ve üzeri 69 kadın ve 70 erkeğe ait görüntüler içinden 139 adet BT görüntüsünden oluşmaktadır. Atlas Vertebra volümetrik verisi hazırlanırken öncelikle aksiyal düzlemde atlasın ön-arka geniş uzunluğunun (APAM) kesilmesi sonrası aksiyal düzlem transvers çıktılarının en uç noktalardan geçecek şekilde 5mm hareket ettirildi akabinde ise kesit kalınlığı bu değerlere uygun şekilde alınarak görüntü setleri oluşturuldu.

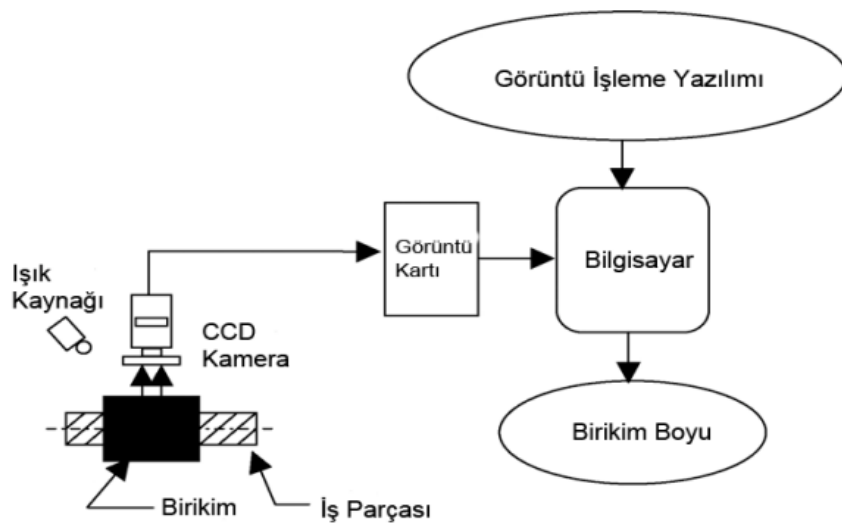


Şekil 3.1. Atlas Vertebra görüntü verisinin hazırlık aşaması [31].

3.1.GÖRÜNTÜ İŞLEME

Bir görüntüyü dijital forma çevirerek görüntü üzerinde işlemler gerçekleştirmek ve görüntüden daha gelişmiş bir görüntü ya da yararlı bilgiler elde etme yöntemlerinin tümüne görüntü işleme denilmektedir. Görüntü işleme ile görüntüler üzerinde bulanıklığın azaltıldığı, görüntüdeki netliğin artırıldığı bir ön hazırlık süreci bulunmaktadır. Ön hazırlık evresinin ardından görüntüler üzerinde düşük, orta ve yüksek seviye işlemler uygulanmaktadır. Düşük seviye işlemlerde görüntüler üzerinde filtreleme işlemlere yapılmaktadır. Görüntüler üzerindeki nesnelere tanımlama, sınıflandırma, görüntüleri bölme gibi işlemler orta seviye işlemler olarak tanımlanmaktadır. Yüksek seviye işlemlerde ise görüntünün tamamından analiz sonuçları elde edilmektedir [32]. Görüntü içerisindeki nesnelere bilgisayar ortamında detaylandırılmaktadır. Bu detaylandırma aşamalarının tamamı görüntü işleme ile gerçekleştirilmektedir [33].

Görüntü işleme yöntemleri ile geliştirilen uygulamalar günümüzde ciddi oranlarda artış göstermektedir. Güvenlik sistemleri, insansız araçlar, tarım alanında kullanılan uygulamalar, biyomedikal ve tıp, coğrafi bilgi sistemleri, tasarım ve imalat uygulamaları gibi birçok alanda görüntü işleme yöntemlerinin kullanıldığı bilinmektedir [32].

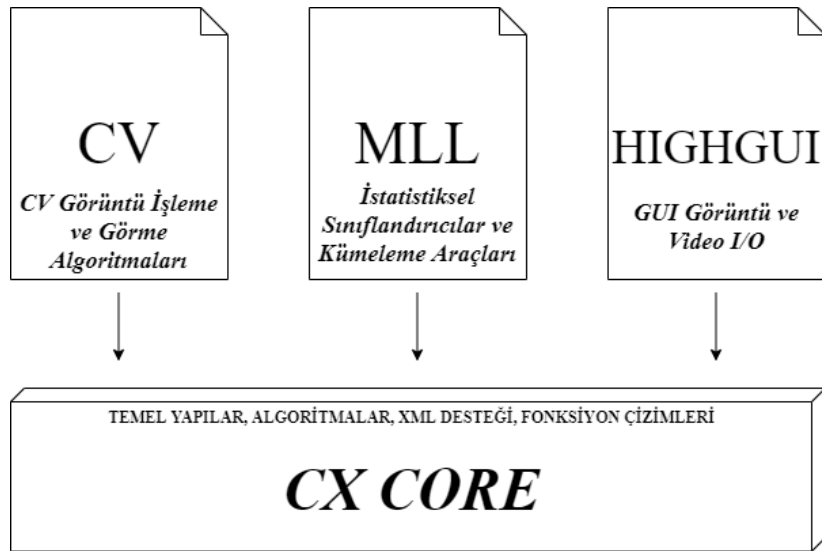


Şekil 3.2. Görüntü işleme ile geliştirilmiş örnek sistemin şematik görünümü.

Görüntü işleme tekniklerinin kullanılabilmesi için her bir yazılım diline özel oluşturulmuş kütüphaneler bulunmaktadır. SimpleCV, OpenCV, OpenFrameworks, PIL/Pillow bu kütüphanelere örnek olarak gösterilmektedir. Bu çalışmada Python yazılım dili ve bu dil için özelleştirilmiş OpenCV Kütüphanesi kullanılacaktır.

OpenCV, Intel firmasının ortaya çıkardığı sonrasında Willow Garage ve Itseez'in, günümüzde ise AMD, Nvidia, Google gibi şirket ya da toplulukların geliştirilmeye devam ettiği görüntü işleme uygulamalarında kullanılan açık kaynak kodlu bir kütüphanedir. Platform bağımsız çalışmaktadır. Berkeley Software Distribution lisanslı ile tüm projelerde ücretsiz kullanılmaktadır. Gerçek zamanlı ve verimli bir şekilde uygulama geliştirmek için tasarlanmıştır.

OpenCV 2500'den fazla görüntü işleme ve makine öğrenmesi algoritması ve 47 bin kişinin üzerinde bir geliştirme topluluğuna sahip olması nedeniyle görüntü işleme teknolojisi hakkında çok geniş bir kaynak sunmaktadır. Geliştiricilere yardımcı olacak teknik doküman sayısının fazlalığı bilişim sektöründe kullanılan tüm programlama dillerine kolayca entegre edilebilmesi sebebiyle diğer tüm kütüphanelere göre daha çok ön plana çıkmaktadır.



Şekil 3.3. OpenCV kütüphanesi temel yapısı.

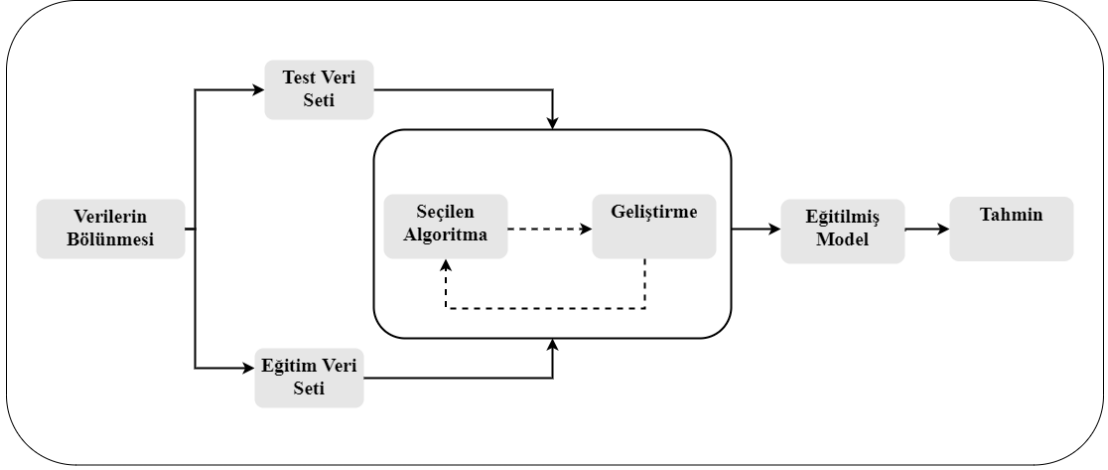
OpenCV kütüphanesi CV, MLL, HighGUI ve CXCore şeklinde isimlendirilen 4 adet ana bileşenden oluşmaktadır. İlk bileşen Görüntü işleme ve yüksek seviyeli görüntü işleme algoritmalarını içeren CV bileşenidir. MLL bileşeninde birçok istatistiksel sınıflandırıcı ve kümeleme aracı içeren makine öğrenme komutları bulunmaktadır. Video ve görüntülerin saklanması ve yüklenmesi için I/O işlemlerini ve fonksiyonlarını HighGUI bileşeni içermektedir. CXCore bileşeni, temel veri yapıları ve içeriğini, resim çizme gibi bölümleri içermektedir. OpenCV C++ başta olmak üzere C, Python, Java gibi yazılım dillerinde de kullanılmaktadır.

3.2.MAKİNE ÖĞRENİMİ

Makine Öğrenmesi Yapay Zekanın bir alt alanı olarak kabul edilmektedir. Makine öğrenmesi ilhamını bilgisayar bilimi, istatistik, biyoloji ve psikoloji gibi çeşitli akademik disiplinlerden almıştır. Makine öğrenmesi algoritmalarının temel işlevi, bilgisayarlara geçmiş deneyimlere dayanarak otomatik olarak iyi bir tahminleme ve sınıflandırma işlevi kazandırmaktır.

Belirli bir veri kümesinin her örneği, aynı özellikler kümesi kullanılarak temsil edilir. Bu özelliklerin doğası sürekli, kategorik veya ikili olabilir. Örnekler bilinen etiketlerle verilirse o zaman öğrenme şeması denetimli olarak bilinir, denetimsiz öğrenme yaklaşımında ise örnekler etiketlenmez. Bir başka makine öğrenimi türü de pekiştirmeli öğrenmedir. Burada, öğrenme sistemine çevre tarafından sağlanan eğitim bilgileri, sistemin ne kadar iyi çalıştığının bir ölçüsünü oluşturan pekiştirme sinyali biçimindedir.

Makine öğrenimi algoritmalarında, öncelikli olarak kullanılacak örneklerden oluşan örnek veri seti oluşturulur. Özellikler belirlendikten sonra uygulanacak olan makine öğrenmesi algoritması seçilir. Ardından eğitim seti ile eğitilir. Oluşturulan modelin performansı uygun görülürse kullanılan veri setinden bağımsız bir test veri seti ile tekrar model test edilir.



Şekil 3.4. Makine öğrenmesi işleyiş adımları.

3.2.1. Naive Bayes

Naive Bayes algoritması, makine öğrenmesi yaklaşımlarından biridir. Algoritma esas olarak olabirlik (olasılık) mantığına dayalı olarak çalışır. Yani her sınıfın olasılığı, belirli bir sınıfa ait olma olasılığını karşılaştırmak için hesaplanır. Ön bilgiyi alır ve verilen hipotezin olasılığını hesaplar.

Bu algoritma en iyi sınıflandırma tekniklerinden biri olarak adlandırılır; ayrıca verilerin bağımsız özelliklerini işlemekle de ünlüdür [34]. Modelin oluşturulması oldukça basittir, ancak özellik belirleyicileri arasında bağımsızlık varsayımına dayanır ve bu kısımda zorluklarla karşılaşılabilir. Tembel bir öğrenme algoritmasıdır ancak dengesiz veri kümeleri üzerinde de çalışabilir. Algoritma, bir kayıt için her bir olasılık derecesini hesaplar ve onu en yüksek olasılık değerine göre sınıflandırır. Algoritma, test veri setinde bulunan ve eğitim veri setinde olmayan veriler için tahmin yapamamaktadır. Bu duruma “Sıfır frekans” denir [35]. Literatürde problemi çözmek için Laplace tahmini gibi düzenleme teknikleri bulunmaktadır.

Temel olarak, bir sonraki olaya önceki olaylara dayanarak karar verilebilecek şekilde çalışır. Başka bir deyişle, bu mantık önceki deney temelinde çalışır. Naive Bayes algoritmasını bazı gerçek hayat problemlerinde gerçekleştirmek için yapılacak ilk adım veri setini tanımlamaktır. Veri kümesi sınıfları açıkça görülebilir ve bu nedenle sınıf soyutlaması kolayca yapılabilir. Bir olayı meydana getiren veya üreten bir faktörü

gözlemeleme olasılığı $P(L|S)$ ile adlandırılan Naive Bayes algoritmasından hesaplanacak ana olasılık terimidir. Diğer Bayes düşük özellikleri aşağıdaki gibi tanımlanabilir:

$$P(L|S) = \frac{P(S) \times P(S|L)}{P(S)} \quad (3.1)$$

$P(S)$: S olayını başka herhangi bir şeyden bağımsız olarak gözlemeleme olasılığını ifade eden bir ön olasılıktır.

$P(L)$: Olayın diğer faktörlerinden bağımsız olarak L faktörünü gözlemeleme olasılığıdır.

$P(S|L)$: Sonsal olasılık olarak adlandırılır ve çift S'nin gözlenmesinin L faktörünü üretme olasılığını temsil eder.

Algoritma esas olarak yukarıda açıklandığı üzere parametreleri ve gerekli olasılığı değerlendirmek için bunların çarpımını ve bölünmesini gerçekleştirmek için kullanılır. Daha yüksek olasılık değeri her zaman bir tahmin sonucu olarak alınır. Bu kavramı veri setine uygulayabilmek için öncelikle; veri kümesi sınıfları görünür olmalı ve açıkça tanımlanmalıdır. Sınıf sıklığı, her sınıfın oluşturulma sayısının değerlendirilmesi anlamına gelir. Bu nedenle, sınıflar frekans tablosu, Naive Bayes algoritmasındaki ilk önemli adımdır. Naive Bayes algoritmasındaki ikinci adım, frekans tablosunun soyutlanmasıdır.

3.2.2. Destek Vektör Makinesi

Destek vektör ağı olarak da bilinen destek vektör makinesi (DVM), Cortes ve Vapnik [36] tarafından ortaya konan denetimli sınıflandırma tekniklerinden biridir. DVM, dağılımı bilinmeyen veriler üzerinde öğrenme gerçekleştirerek yeni veriler hakkında tahmin ve genelleme gerçekleştiren makine öğrenme algoritmasıdır. DVM'nin temel ilkesi, iki sınıfın verilerini en iyi şekilde ayıran bir hiper düzlemin varlığına dayanmaktadır. DVM, veri setinin lineer ayrımı ve lineer olmayan ayrımına göre ikiye ayrılmaktadır [37].

DVM, karar sınırlarını belirleyen karar düzlemleri kavramına dayanmaktadır. Karar

düzlemi farklı sınıf özelliklerine sahip nesne kümesinin elemanlarını birbirinden ayırmaktadır.

Farklı sınıfsal özelliklere sahip nesnelere birbirinden ayırmak için bir ayıraç çizerek yapılan sınıflandırma işi hiper düzlemde sınıflandırma olarak bilinmektedir. DVM, özellikle bu tür işler için çok uygun bir yapıya sahiptir [38]. Ayrıca çok boyutlu bir özellik alanındaki doğrusal fonksiyonların hipotez alanını kullanan öğrenen sistemlerdir ve optimizasyon teorisinden bir öğrenen algoritma ile eğitilirler. Bu öğrenme stratejisi ilk olarak Vapnik tarafından geliştirilmiştir. Çok geniş bir uygulama alanında kullanılan güçlü bir yöntemdir. Temel kavramı hiper düzlem sınıflandırıcı ya da doğrusal ayrılabilirlik. Doğrusal ayrılabilirliği sağlamak için, destek vektör makineleri marjin maksimizasyonu ve çekirdek olarak iki temel fikir kullanılır. Girdi alanı, daha üst boyut ya da özellik alanına eşleşmektedir [39].

DVM, denetimli bir öğrenme yöntemi olup, eğitim verisi ve algoritmayı test etmek maksadıyla ayrıca test verisi gerektirmektedir. DVM ile doğrusal ve doğrusal olmayan sınıflandırma yapılabilir [40]. DVM iyi genelleme doğası ve boyutluluk sorununun üstesinden gelebilme yeteneğinden ötürü anomali saldırı tespit sistemleri için kullanılan popüler tekniklerden bir olmuştur [41]. Çok az eğitim verisinin bulunduğu durumlarda ve çok boyutlu alanlarda bile, çekirdek marifetiyle çok iyi genelleme yapabilmektedir. Destek vektör makinelerinin en önemli avantajı hızlarıdır, zira gerçek zamanlı saldırı tespiti yapabilme yeteneği çok önem taşımaktadır. Destek vektör makineleri, sınıflandırma karmaşıklığının özellik alanının boyut sayısına bağlı olmamasından dolayı, çok geniş bir örüntü kümesini öğrenebilmektedir. Ayrıca sınıflandırma esnasında yeni bir örüntü ortaya çıktığında eğitim örüntülerini dinamik olarak güncelleyebilme yeteneğine sahiptirler [41].

DVM'de sınıflamanın, doğrusal ayırıcı hiper düzlem ile yapılmasıyla birlikte, orijinal girdi alanında doğrusal ayırımın yapılamadığı durumlar da ortaya çıkabilmektedir. Bu durumlar da çekirdek adı verilen fonksiyonlar kullanılmaktadır. Maksimum hiper düzlem bulunduğu sadece hiper düzleme en yakın noktalar için $\alpha_i > 0$ olacaktır ve bunlar destek vektörü olarak adlandırılmaktadırlar. Diğer tüm noktalar için $\alpha_i = 0$ olur. Sadece hiper düzleme en yakın noktalar hipotez sınıflandırıcıyı temsil etmektedir. Bu

en önemli noktalar destek vektörü olarak hizmet ederler. Bu noktaların değerleri hipotez sınıflandırıcının güvenilirliği ile ilgili olarak bağımsız bir sınır vermek için de kullanılabilir [39]. Diğer sınıflandırma tekniklerinin aksine, DVM sınıflandırma hatasını değil beklenen hatayı minimize etmektedir. Ayrıca DVM, etkili bir hesaplama yöntemine izin veren matematiksel programlamanın ikilik teorisini kullanmaktadır [42]. DVM, kullanıcı tarafından belirlenen, hata faktörü olarak adlandırılan bir parametre de sunmaktadır. Kullanıcıların, yanlış sınıflandırılmış örnekler ile karar sınırının kalınlığı arasında seçim yapmalarına olanak sağlamaktadır [43].

3.2.3. K-En Yakın Komşu

K-En Yakın Komşu (K-EYK) makine öğrenmesi algoritmalarından biridir, denetimli makine öğrenmesi algoritmaları altında sınıflandırılır. Bu algoritma basitliği ile popülerdir ve parametrik değerlendirme ve olasılık hesaplamaları gerektirmez [44].

Yöntemde k sayısının belirlenmesi önemlidir. k pozitif tam sayı olmalıdır. Kabul edilen komşuların sayısını temsil eder. Bilinen bir veri kümesi için, verilerin ön işleme başlangıçta yapılmalıdır. Daha sonra K-EYK süreci üç ana adıma ayrılabilir:



Şekil 3.5. K-EYK algoritma yapısı.

1. Mesafenin değerlendirilmesinde mesafeler, test verileri ile eğitim verileri arasında hesaplanır. Mesafe için en yaygın ölçümler Öklid mesafesi, Manhattan mesafesi ve Hamming mesafesidir. Öklid en sık uygulananıdır.
2. Mesafe bilgisine göre en yakın komşunun belirlenmesinde mesafeler artan düzende sıralanır. Daha sonra üst k komşular tercih edilir.
3. En yakın Komşulara göre, tahmini temsil eden sonuçlar yapılır. Sıralanmış listeden en çok k uzaklık seçilir ve listenin en sık görülen sınıfına bağlı olarak test noktasına bir sınıf atanır. İki nokta arasındaki Öklid mesafesi denklem (3.2) kullanılarak hesaplanabilir.

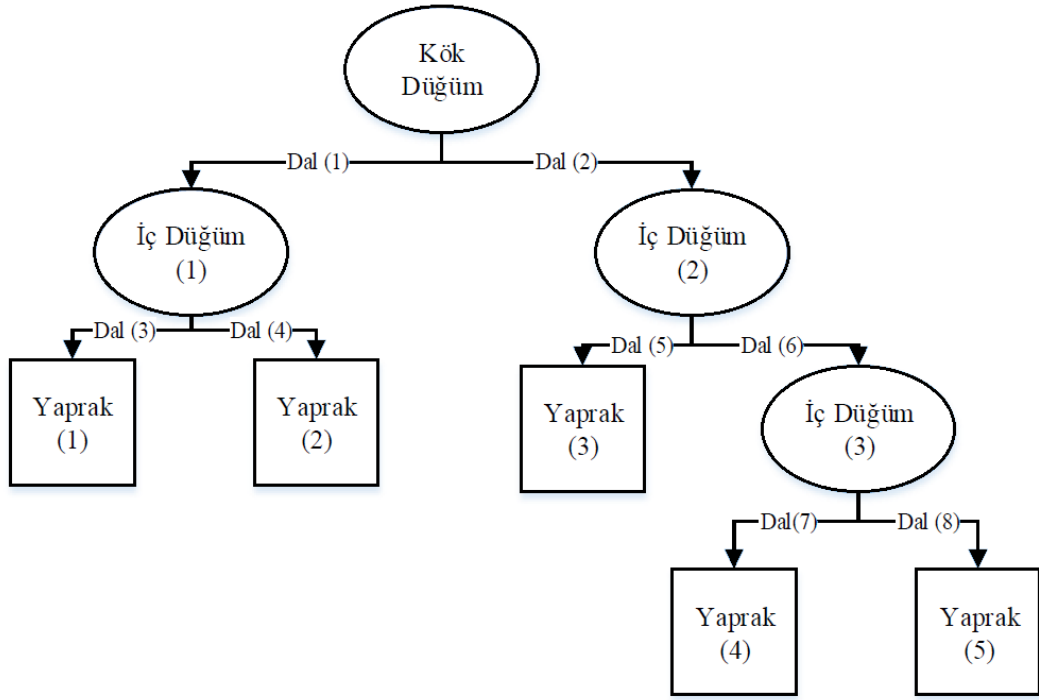
$$d(x, x') = \sqrt{(x_1 - x'_1)^2 + (x_2 - x'_2)^2} \quad (3.2)$$

Burada d , test verilerinin ögesi x ile her bir eğitim ögesi x' arasındaki mesafeyi temsil eder. Her eleman için Öklid mesafesi değerlendirilirken, sınıflar üç sınıf tipine ayrılmıştır. Her sınıf diğerinden belirli bir Öklid mesafesinden uzaktadır. Her sınıfı bir daire; daireler ise sarı, mavi ve yeşil gibi belirli bir renk kullanılarak temsil edilir.

K-EYK algoritmasının yapacağı ikinci işlem, sınıflar arasındaki en yakın mesafeyi değerlendirmektir. Mesafe değerleri artan düzende sıralanır ve en üst k mesafe sayısı seçilir. K-EYK algoritmasındaki son adım, sınıflandırmayı gerçekleştirmektir. Ancak, sınıflandırma, en yakın k komşunun eğitim veri kümesi sınıfının frekansına göre yapılacaktır. Test verilerinin sınıfı, en yakın k komşuda en sık görülen sınıfa göre atanır. Mesafenin benzerliği, test setinin bu girişinin en iyi benzerliğe sahip sınıfla ilgili olduğu kararına yol açacaktır.

3.2.4. Karar Ağaçları

Karar ağacı, örneklerin entropisine dayanan bir makine öğrenme prosedürüdür. Denetimli bir algoritma olan karar ağacı algoritması, bir sınıflandırma yöntemi veya problemle ilgili bir regresyon belirleyicisi olarak hizmet edebilir. Bir kök düğümden başlar ve entropiye dayanarak, örnekler bir yaprak düğümüne ulaşana kadar bölünür.



Şekil 3.6. Karar ağacı yapısı.

Karar ağaçları, kullanımı ve anlaşılması basit tahmine dayalı öğrenme algoritmalarıdır. Akış diyagramı benzeri ağaç yapısı ile her sonuçla ilişkili tüm olası yolları temsil ederler. Görselleştirilebilmesi ve kolay yorumlanabilmesi avantajlarındandır. Ayrıca elde edilen sonuçların güvenilirliği ve yüksek performansa sahip olması da tercih sebeplerindedir [45].

Karar ağaçları hem sınıflandırma hem de tahmin görevlerinde kullanılabilir. Karar ağaçlarını kullanarak öğrenme iki adımdan oluşur. İlk adımda, eğitim verileri kullanılarak bir ağaç oluşturulur. Ardından, her kayıt için, kaydın ait olduğu sınıfı belirlemek için ağaç geçilir. Ağaçtaki her iç düğüm bir öznitelik üzerinde bir sınımayı gösterir, her dal sınımanın sonucunu gösterir ve her yaprak düğüm bir sınıfı gösterir [46,47]. Karar ağaçlarının avantajından biri sadeliği ve verimliliğidir. Ağaç algoritmaları genellikle diğer yöntemlerden daha hızlıdır. Ağacın boyutu veri tabanı boyutundan bağımsızdır. Başka bir deyişle, karar ağaçları büyük veri tabanı ve yüksek boyutlu veriler için de etkilidir.

Bir ağaç modelinin karmaşıklığı, yaprak düğümlerinin sayısı ve ağacın bölümü ile

ifade edilebilir. Budama daha basit modeller elde etmek için kullanılabilir. Karar ağacı sonuçlarının anlaşılması ve yorumlanması kolaydır, çünkü okunabilir kurallar modellerden kolayca türetilir.

Karar ağacı algoritmalarının çoğunun karşılaştığı bazı sorunlar vardır. Ağaç, gürültü, eksik veri ve benzeri anormalliklerden etkilenir. Özellikle aykırı değerler ve bu anormallikler aşırı öğrenmeye neden olabilir. Bazı istatistiksel önlemleri kullanarak ağaçtan en az güvenilir dalları çıkarma işlemi olan ağaç budaması bu sorunu çözebilir. Ek olarak, nitelikler arasındaki korelasyonlar dikkate alınmaz, göz ardı edilir. Bölme niteliklerini seçmek bir başka önemli konudur. Bazı öznelikler değerli bilgilere veya girdi alanını verimli bir şekilde bölme yeteneğine sahip olmayabilir. Bölme sırası niteliklerde önemlidir. Bilgilendirici nitelikleri doğru seçmek için kirliliğe dayalı önlemler kullanılabilir. Örneğin, entropi ve gini indeksi [47] bu tür bölünmüş seçim önlemleridir. Bu şekilde, seçilen nitelikler safsızlık işlevini ve verileri sınıflandırmak için gereken test sayısını en aza indirir. Bölünme sayısı, belirlenmesi bir başka zor konudur. Öznitelik küçük etki alanıyla kategorik olduğunda veya algoritma ikili bir ağaç oluşturuyorsa bölme sayısı açıktır. Ancak, öznitelik sürekli ise, bu sayıyı belirlemek kolay bir iş değildir. Durdurma kriterleri, ağacın sınıflandırılmasının ve performansının doğruluğu nedeniyle dikkate alınması gereken önemli bir konudur. Ağaç oluşumu eğitim veri kümesine bağlıdır. Farklı veri kümeleri farklı sınıflandırmalara neden olabilir. Çok küçük bir veri kümesi sınıflandırmayı doğru temsil etmeyebilir. Tersine, veri kümesi çok büyük olduğunda daha sonra aşırı öğrenme oluşabilir. Çapraz doğrulama tekniği aşırı öğrenmeyi azaltmak için kullanılabilir [45].

3.2.5. Rastgele Orman

Rastgele orman sınıflandırıcısı hem regresyon hem de sınıflandırma amacıyla kullanılabilen denetimli makine öğrenme algoritmasıdır ve ayrıca öğrenmeyi birleştirmenin bir örneğidir. Genellikle sınıflandırma problemleri için tercih edilmektedir. Öğrenmeyi bir araya getirmek, sorunu çözmek için tek bir öğrenme yöntemi üzerinde çalışan erken makine öğrenme yöntemlerinin aksine, bir soruna çözüm getirmek için çeşitli öğrenme stratejilerinin tek bir sürece uygulanmasıdır [48].

Yani birden fazla karar ağacının bir araya gelmesiyle oluşan rastgele orman algoritması veri seti üzerinde yeni bir rastgele orman oluşturur ve oluşturmuş olduğu bu ormanlar için tahmin yapar. Sonuç olarak elde edilen tahmin sonuçlarının arasından en iyi çözüme ait olanı seçer. Aşırı öğrenmeyi önlemek için sonuçların ortalamasını alır ve bağımsız parametrelerle farklı karar ağaçları üzerinden değerlendirmeye aldığı için başarısı oldukça yüksektir [49].

Rastgele orman, temel öğrenenler olarak karar ağacı ile bir araya getirilmiş bir öğrenmedir. 2001 yılında Breiman, kümeleme, sınıflandırma ve tahmin gerçekleştiren sözde rastgele orman algoritmasını önermişti. Literatürde, güvenilir tahmin/kümeleme doğruluğu sağlanması ve gürültü ve kayıp değerler gibi verilerde daha fazla anormalliği tolere edebilmesi için rastgele orman algoritmasının performansını artırmak için birçok çalışma yapılmıştır. Bir algoritmadaki gürültü toleransı, algoritmanın sonuçlarını gerekli doğrulukla üretebildiği gürültü seviyesi olarak adlandırılır [50].

Rastgele orman önyüklemesi, sözde ağacı oluşturmak için orijinal örneklerden öğeleri elde etmek amacıyla bu algoritma tarafından gerçekleştirilen bir işlemdir. Ağacın, verilerin sınıfları/kümeleri olarak adlandırılan dalları oluşturmak için kendi kısıtlaması vardır. Rastgele orman, sınıflandırma görevlerini aşağıdaki adımlarla gerçekleştirir:

1. Ön işleme aşamasından elde edilen veri kümesi değerlerinin toplanması ve ana verilerden sınıf (dal bazında) örneklerin toplanmasını içeren önyükleme adı verilen eylemin ilk adımını gerçekleştirme.
2. Toplanan her örnek için, seçilen örnek ile belirli dal arasındaki korelasyon miktarını belirlemek için entropi hesaplaması başlatılabilir.
3. Orman terimi, birden fazla ağaç içeren sınıf ağaçları için ayrılmıştır. Her bir önyükleme örneği için, rastgele ormanın eğitim aşaması olarak adlandırılan M karar ağaçları oluşturulur.
4. Her bir karar ağacının eğitimi, her bir ağaçtaki veri örneklerinin yeniden düzenlenmesiyle bağımsız olarak gerçekleşir. Her seferinde, ağaçtaki veri örneğinin

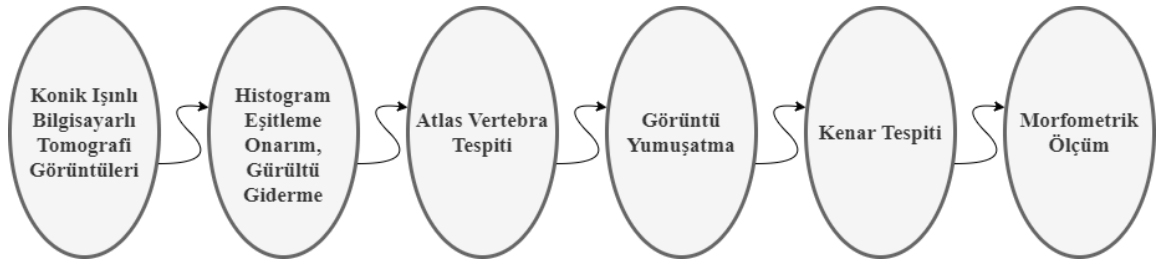
yeniden düzenlenmesi, sınıfa giriş korelasyonunu tespit etmek için entropi ölçümü yapılabilir.

5. Rastgele orman algoritmasındaki sınıflandırma performansı, doğruluk, ortalama kare hatası ve ortalama mutlak hata gibi standart performans ölçütleri kullanılarak ölçülebilir.

BÖLÜM 4

KONİK IŞINLI BİLGİSAYARLI TOMOGRAFİ GÖRÜNTÜLERİNİN MAKİNE ÖĞRENMESİ İÇİN HAZIRLANMASI

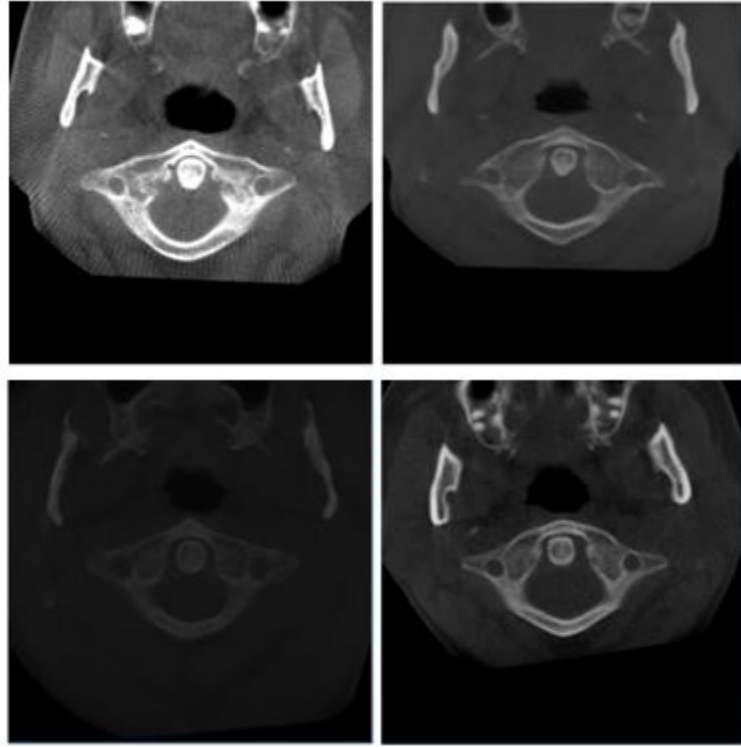
Makine öğrenmesinde öznitelik belirleme, kullanılacak olan verilerden anlamlı özniteliklerin belirlendiği en önemli aşamadır [51]. Girdi olarak konik ışınli bilgisayarlı tomografi ile elde edilen görüntüler üzerinde otomatik morfometrik ölçümler gerçekleştirilmekte ve ölçüm sonuçlarından elde edilen veriler öznitelik tablosuna eklenmektedir. Makine öğrenmesinde kullanılacak olan özniteliklerin, en doğru şekilde alınabilmesi için öncelikle görüntü üzerinde ön işlemler gerçekleştirilmektedir. Ön işlemlerde filtreleme, histogram eşitleme, tüm görüntüden atlas omurunun tespiti, kenar tespiti ve bazı dönüşüm teknikleri kullanılmıştır.



Şekil 4.1. Yöntem ve Tasarım.

KIBT ile alınmış görüntüleri ölçüme hazırlama aşamaları ve ölçüm esnasında yaş, cinsiyet, yetişkin, çocuk gibi ayırt edici hiçbir özellik kullanılmamış ve tüm görüntüler aynı şekilde değerlendirilmiştir. Ölçüm yapılacak görüntüler üzerinde öncelikle gürültü giderme, histogram eşitleme, onarım gibi işlemler yapılarak Atlas Vertebra tespiti için görüntü en uygun hale getirilmiştir. Temizlenen görüntü üzerinde ROI işlemi ile Atlas Vertebra tespiti yapılarak görüntüden kesilerek çıkarılmıştır.

Görüntüden kesilerek çıkarılan Atlas Vertebra üzerinde ölçüm yapabilmek için yumuşatma işlemleri ile kemik kenarlarının netleştirilme işlemi ve akabinde kemik kenarlarının çizdirilmesi gerçekleştirilmiştir. Son olarak çizimi yapılan Atlas Vertebra üzerinde morfometrik ölçümler yapılmaktadır. Anlatılan tüm bu işlemler insan müdahalesi olmadan açık kaynak kodlu OpenCV Kütüphanesi kullanılarak otomatize edilmiştir.

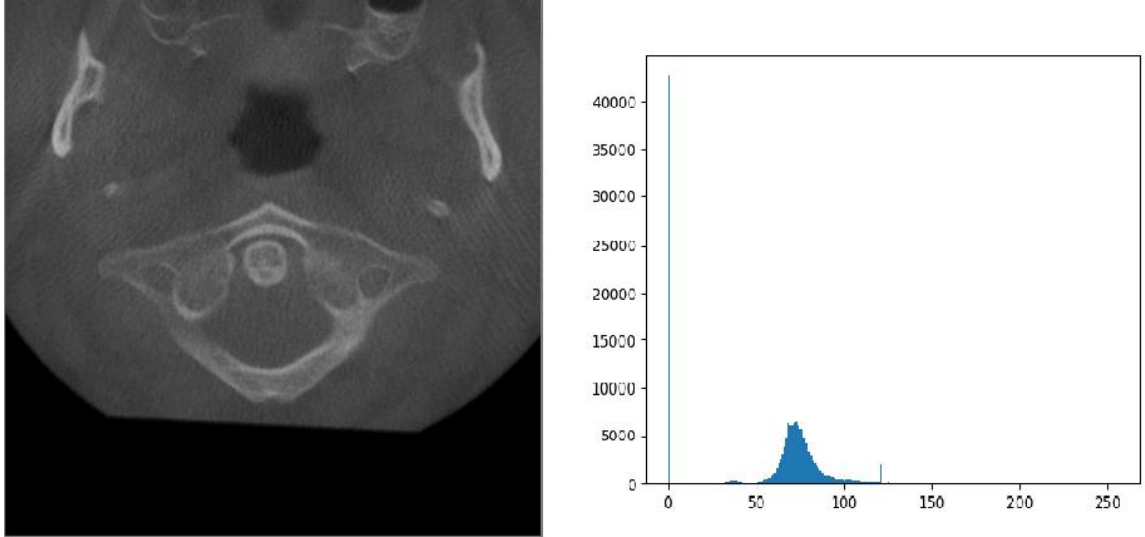


Şekil 4.2. KIBT ile elde edilen Atlas Vertebra görüntüleri.

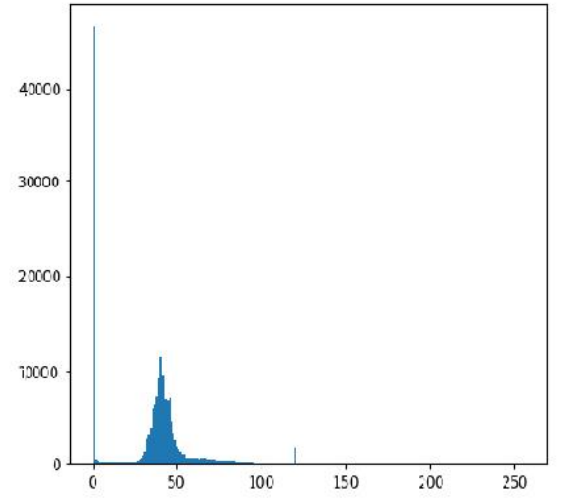
Çalışmamızda kullanılan sayısal görüntü örnekleri incelendiğinde bu görüntülerin farklı kontrastlara sahip olduğu görülmektedir. Farklı kontrastlara sahip olan görüntülerde değişik netlikler ortaya çıkmaktadır. Görüntünün elde edilmesi aşamasında ilgili uzmanların farklı şekilde çalışması, görüntü alınan görüntülüne cihazının ayarları vb. gibi durumlar kontrast farklılıklarına sebebiyet vererek tespit işlemlerini zorlaştırmıştır. Ortaya çıkan görüntülerin farklı kontrastlara sahip olma durumu sebebiyle tüm görüntüler üzerinde öncelikle kontrast eşitleme işlemi uygulanması gerçekleştirilmiştir. Görüntüler analiz edildiğinde 0 ile 255 arasında gri tonlu renk değerlerine sahip oldukları, histogramları incelendiğinde belirli bölgelerde

yığılmaların oluşması zıtlığın az olduğunu ve histogram eşitleme işlemi ile görüntüde iyileştirme işlemi yapılabileceği tespit edilmiştir ve gerekli algoritma adımları uygulanmıştır.

Görüntüler üzerinde yapılan işlemlerde histogramlar kullanılarak görüntülerin netliklerinin daha fazla artırılması, nesnelerin daha da belirgin hale gelmesi sağlanmaktadır [52]. Histogram grafiği incelendiğinde belirli bir tonda toplanmış eğriye sahip grafiklerde homojenlik sağlamak için renk dağılımına gidilmektedir. Bu şekilde hedeflenen doğrusal bir kümülatif histograma sahip olmaktır.

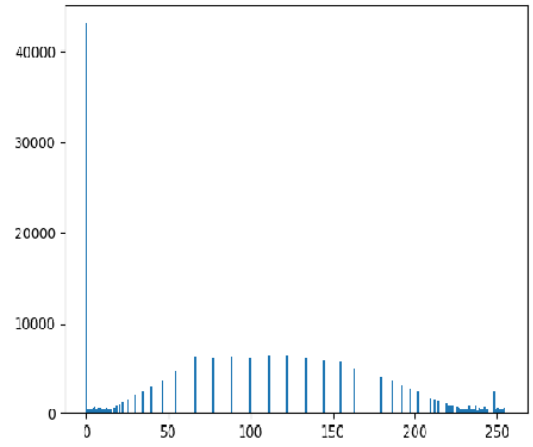


(a)

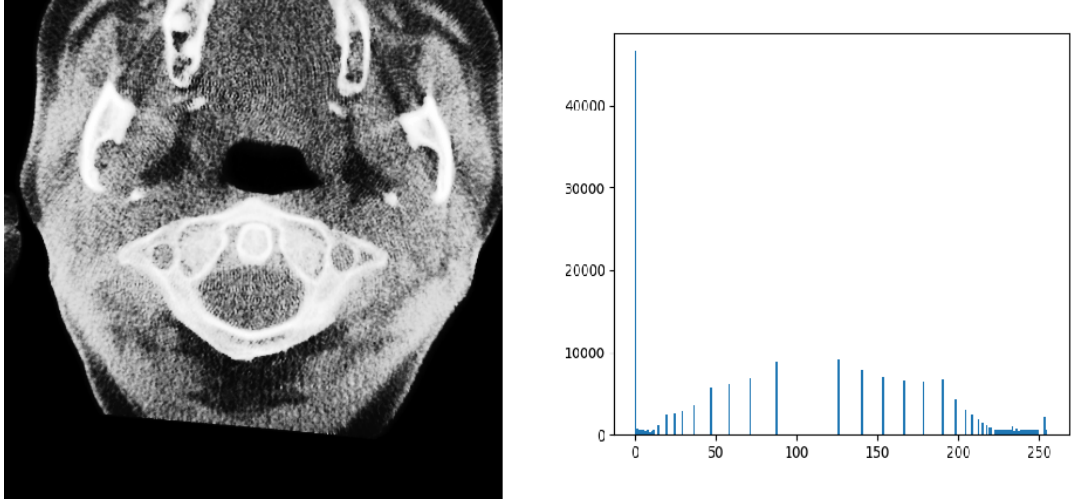


(b)

Şekil 4.3. Histogram eşitleme işlemi öncesi örnek görüntüler ve histogram grafikleri.
a) ve b) Farklı kontrastlara sahip 2 görüntünün histogram grafikleri.



(a)



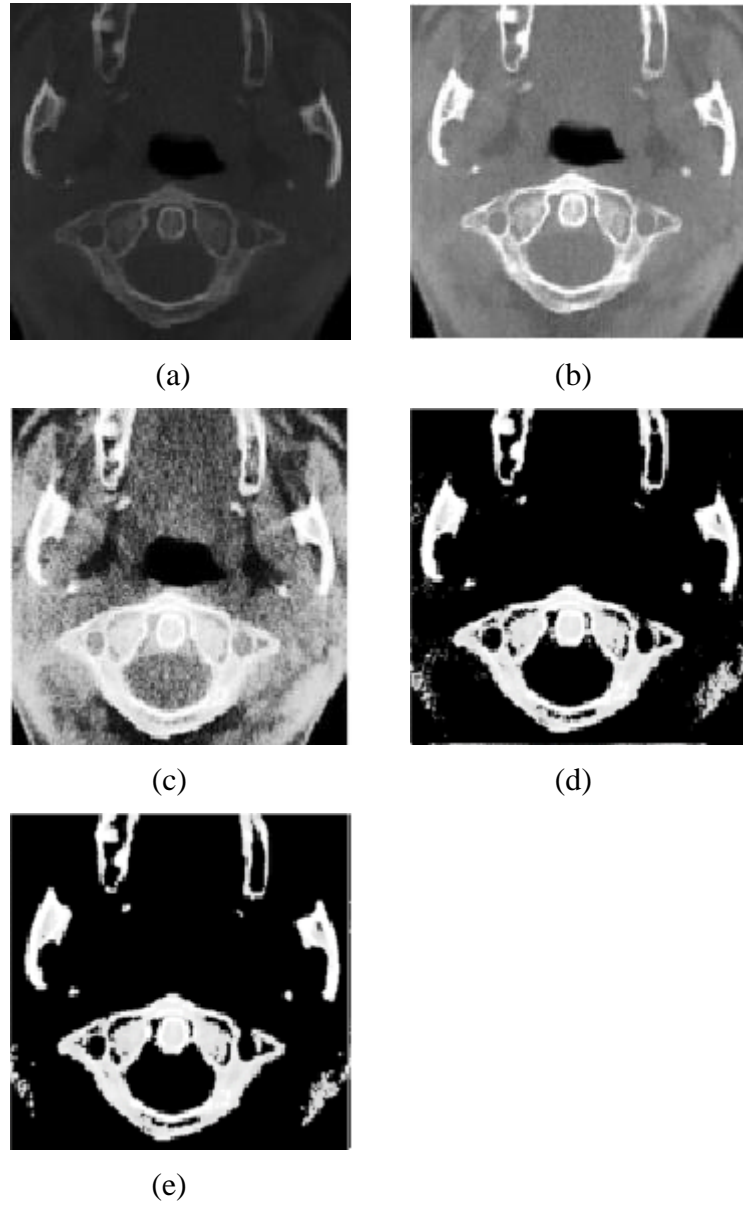
(b)

Şekil 4.4. Histogram eşitleme işlemi sonrası örnek görüntüler ve histogram grafikleri. a) ve b) Farklı kontrastlara sahip 2 görüntünün histogram grafikleri.

Şekil 4.3 ve Şekil 4.4'te örnek görüntülerin eşitlemeden önce ve sonra histogram grafikleri verilmiştir. Grafikler incelendiğinde eşitleme işlemi öncesinde gri seviyesindeki yoğunluğun belli bir bölgede toplandığı görülmektedir. Eşitleme yoğunlaşma homojen olarak dağıtılarak görüntü üzerinde optimizasyon yapılmıştır. Bu optimizasyon işlemi ile gri seviyesi dağılımında dengeleme sağlanmaya çalışılmış ve Atlas Vertebra kemiği görüntü üzerinde daha net bir hale gelmiştir. Böylelikle görüntü ROI işlemlerine hazırlanmıştır. ROI işlemleri görüntü üzerinde istenilen bir alanı temsil etmektedir. Bu çalışmada görüntünün tamamından istenilen alan olarak Atlas Vertebra kemiği çıkarılacak ve çıkarılan parçada ölçüm yapılacaktır. Bu sebeple Atlas Vertebra tüm görüntü içinden tespit edilmektedir. Görüntü işleme yöntemlerinde istenilen alanın tespiti aşamasında görüntülerin karşıtlıklarının ve parlaklıklarının ayarlanması, gürültü giderimi, görüntü üzerinden yapay şekillerin çıkarılması gibi işlemler yapılmaktadır. Bizim çalışmamızda kullanılan görüntülerde çıkarılması gereken yapay şekiller bulunmamaktadır. Bu sebeple görüntü iyileştirme işlemleri sırasında yapay şekillerin tespiti ve görüntüden çıkarılması gibi işlemlere ihtiyaç duyulmamıştır.

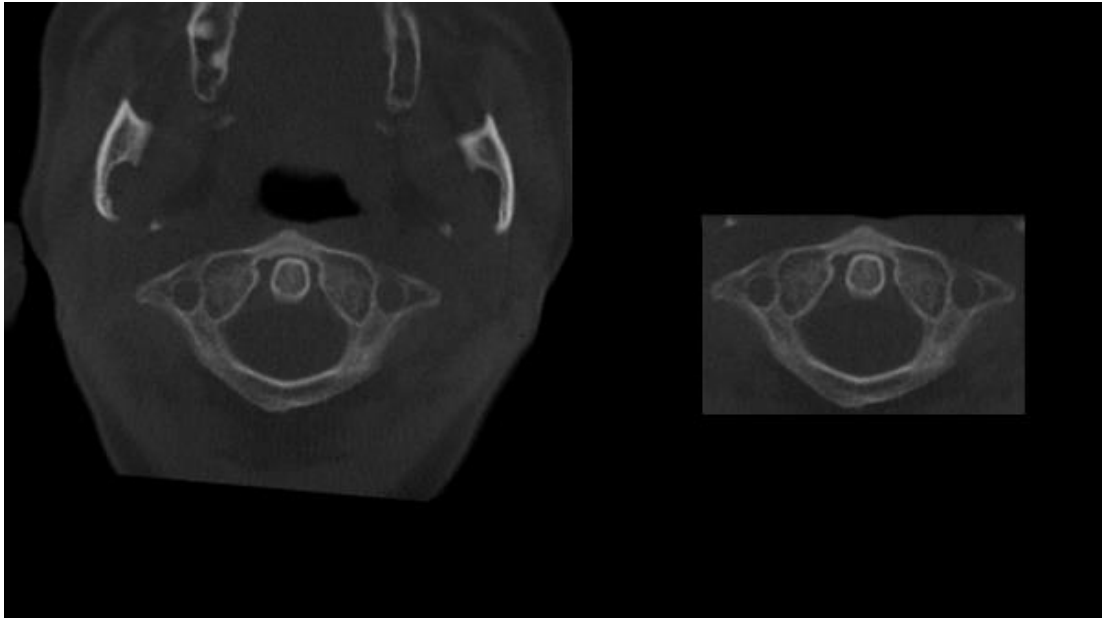
Görüntü iyileştirme işlemleri adımları;

1. Farklı kontrasta sahip KIBT görüntüleri üzerinde otomatik kontrast eşitleme,
2. Histogram eşitleme,
3. Görüntü üzerinden tüm piksellerin okunarak ortalama piksel değerinin belirlenmesi,
4. Ortalamanın altında kalan piksel değerlerinin sıfır değerine çekilmesi,
5. Morfolojik dönüşüm işlemleri.

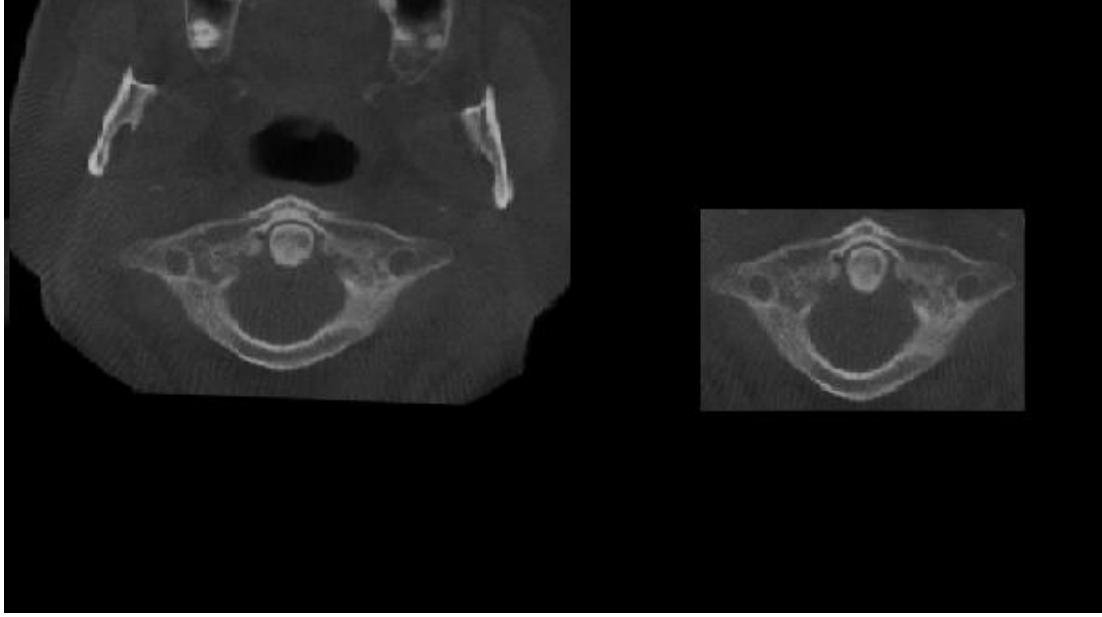


Şekil 4.5. Atlas Vertebra tespit işlemleri a) Okunan görüntü b) Histogram eşitleme c) Otomatik parlaklık ve kontrast düzeltme işlemi d) Ortalama piksel değerinin altında kalan tüm piksellerin sıfır değerine çekilmesi e) Morfolojik işlemler.

Şekil 4.5'te gösterilen görüntü iyileştirme işlemlerinde amaçlanan morfometrik ölçüm işlemlerinde en temiz görüntü ile en doğru sonucu gerçekleştirmektir. Bu sebeple Atlas Vertebra kemiğinin tüm görüntü içinden tespit edilmesi ve çıkarılması amaçlanmıştır. Bu duruma örnek vermek gerekirse göz tanıma sistemlerinde insanın bedeninin tamamından işlem yapmak yerine öncelikle yüz tespiti gerçekleştirilir ve tespit edilmiş yüz üzerinden göz tanıma yapılmaktadır. Bizim çalışmamızda amaçlanan da baş boyun bölgesinden alınan KIBT görüntüsü içerisinde öncelikle Atlas Vertebra tespiti yaparak göz tanıma sistemlerindeki gibi performans ve zaman kazanımı konusunda fayda sağlamaktır. ROI işlemleri sonucu Atlas Vertebra kemiği tespit edilerek tüm görüntüden çıkarılmış ve morfometrik ölçüm işlemleri aşaması için hazır hale getirilmiştir.



(a)

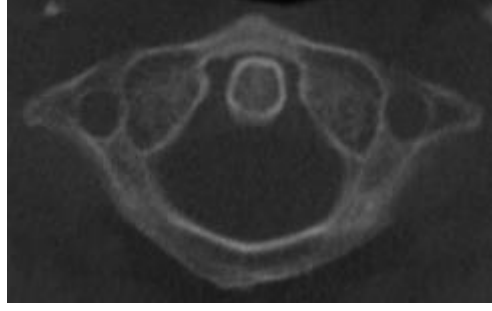


(b)

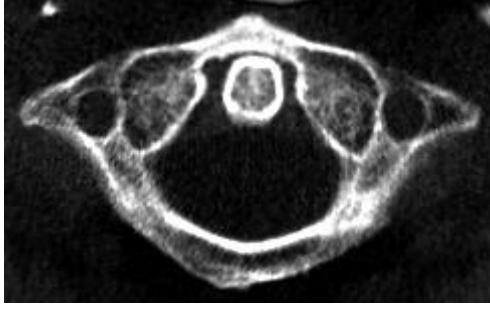
Şekil 4.6. Atlas Vertebra tespit sonucu. a) ve b) KIBT görüntüsünden Atlas Vertebra tespiti ve kesilerek çıkarılması örnekleri.

KIBT görüntüsü içinden kesip çıkarılan Atlas Vertebra görüntüsü incelendiğinde görüntü üzerindeki gürültülerin kontrastı azalttığı, ölçüm aşamasında olumsuz etkiye sebep olarak hatalara yol açtığı tespit edilmiştir. Bu nedenle morfometrik ölçümler için yapılacak kenar tespiti öncesi görüntü üzerinde bulunan gürültüler görüntüden çıkarılarak, görüntüye yumuşatma işlemi uygulanıp görüntünün ölçüme hazır hale getirilmesi gerekmektedir.

Gauss, medyan, ortalama ve geometrik filtreleme gibi yöntemler gürültü azaltmada kullanılmaktadır. Bu filtrelerde temel amaç, görüntü üzerindeki kenarlar korunurken noktasal unsur olarak bulunan gürültüyü mümkün olduğunda azaltmak ve böylece görüntü kalitesini artırmaktır. Fakat gürültü azaltmada kullanılan bu yöntemler karmaşık yapıdaki gürültüyü azaltma konusunda istenilen düzeyde başarılı olamadığı için sınıflandırma, bölge tespiti, şekil analizi gibi uygulamalarda ihtiyacı karşılama konusunda eksik kalmaktadır [53–55]. Bu nedenle çalışmada, gürültü gidermede başarılı sonuçlar sunan seyreklik tabanlı gürültü giderme filtresi kullanılmıştır [56].



(a)



(b)



(c)

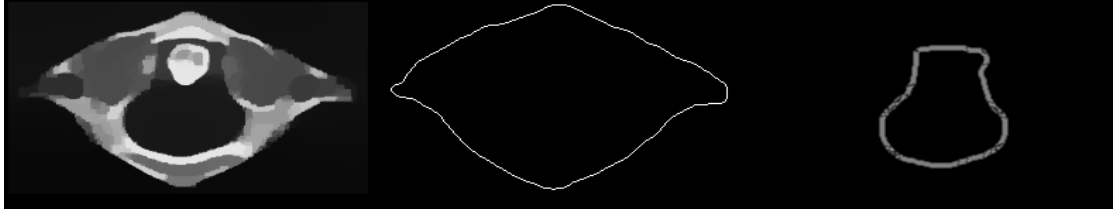
Şekil 4.7. Atlas Vertebra görüntüsünün morfometrik ölçüme hazırlık aşamasına örnekler a) Okunan görüntüden kesip çıkarılan Atlas Vertebra b) Görüntü işleme işlemleri sonucu c) Morfometrik ölçüm öncesi yumuşatma işlemi yapılan görüntü.

Seyreklik tabanlı yumuşatma işlemi uygulanarak elde edilen Atlas Vertebra görüntüsü üzerinde morfometrik ölçümler yapılabilmesi için yapısal şeklin çıkarılması gerekmektedir. Bu yapısal şeklin çıkarılmasında görüntü işleme yöntemlerinde bulunan Canny kenar tespit yöntemi kullanılmıştır. Kenarların çizilebilmesi için görüntü üzerindeki onarım ve keskinleştirme işlemleri yapılarak en doğru kenar çizimi amaçlanmıştır. Kenar tespiti için işlemler;

1. Segmentasyon işlemleri,
2. Kalan gürültülerin temizlenmesi ve yumuşatma işlemleri,
3. Kenar çizimi için kullanılacak eşik değeri tespiti,
4. Kenarların çizimi.



(a)



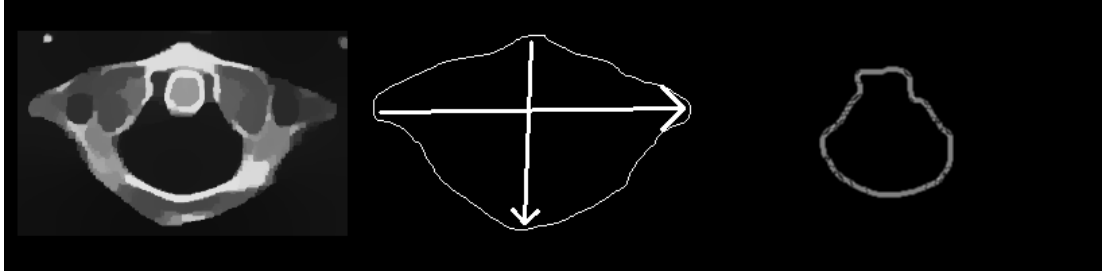
(b)

Şekil 4.8. Atlas Vertebra çevresinin çizilmesi. a) ve b) Farklı kontrastlara sahip iki görüntü üzerinde Atlas Vertebra çizimi örnekleri

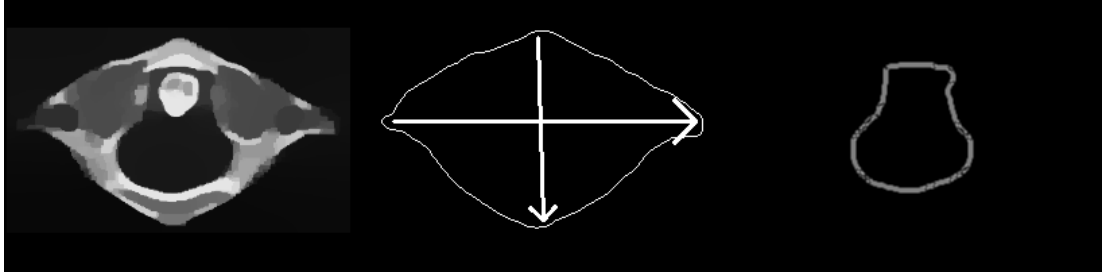
Morfometrik ölçüm işlemi için öncelikle Atlas Vertebra görüntüsü 4 eşit parçaya bölünmüştür. Her parça içinde piksel değerleri okunarak, piksel geçişleri kontrol edilmiş ve her bir parçaya ait eşik değeri belirlenmiştir. Her parçada belirlenen bu eşik değerine göre kenar çizimleri gerçekleştirilmiş ve tüm parçalar tekrardan birleştirilerek görüntü ölçüme hazır hale getirilmiştir. Son adım olarak çizimi yapılan Atlas Vertebra görüntüsü üzerinde ölçüm işlemlerine geçilmiş ve hesaplama işlemleri yapılmıştır. Bu çalışmada Atlas Vertebra görüntüleri incelenerek elde edilen görüntüler üzerinden en hızlı şekilde ölçüm yapılabilecek alanlar belirlenmiştir. Aynı zamanda ölçümü yapılan bu alanların hata oranlarının, Atlas Vertebra üzerindeki diğer alanlara göre daha az olması sebebiyle bu altı farklı ölçüm üzerinde çalışma yapılmasına sebep olmuştur. Ölçüm yapılan alanlar;

1. ATM: Atlas Vertebra kemiği transvers çıkıntıları arasındaki mesafe uzunluğu (Atlas kemiği üzerinde enine yapılan ölçüm),
2. APAM: Atlas Vertebra kemiği ön-arka mesafe uzunluğu,
3. ACU: Atlas Vertebra kemiği çevre uzunluğu,

4. ADA: Atlas Vertebra kemiği alanı,
5. AFCU: Atlas Vertebra kemiği forameninin çevre uzunluğu,
6. AFA: Atlas Vertebra kemiğinin forameninin alanı.



(a)



(b)

Şekil 4.9. Atlas Vertebra morfometrik ölçümlerinin çizim örnekleri. a) ve b) Farklı kontrastlara sahip iki görüntü üzerinde Atlas Vertebra ölçüm çizimi örnekleri

Atlas Vertebra yapısal çizimi gerçekleştirilerek oluşturulan görüntü üzerinde tepe, alt, sağ, sol olmak üzere 4 uç nokta tespit edilmiştir. Sağ-sol uç noktaları arasında kalan mesafe Öklid uzaklık hesaplaması yapılarak ATM ölçüm sonucuna ve tepe-alt noktaları arasında kalan mesafe yine Öklid uzaklık hesaplaması yapılarak APAM ölçüm sonucuna yazılmıştır. Öklid uzaklık formülü şu şekildedir:

$$\begin{aligned}
& \text{öklid uzaklığı} = d(p, q) = d(q, p) \\
& = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2} \\
& = \sqrt{\sum_{i=1}^n (q_i - p_i)^2}
\end{aligned} \tag{4.1}$$

ACU ve AFCU ölçümlerinde çizimi yapılan Atlas Vertebra görüntüsü üzerinde çevre hesaplamaları OpenCV kütüphanesine ait kontür işlemleri içinde bulunan arcLength metodu kullanılmıştır. ADA ve AFA ölçümlerinde ise yine çizimi yapılan Atlas Vertebra görüntüsü üzerinde OpenCV kütüphanesine ait contourArea metodu kullanılmıştır.

Morfometrik ölçümler ile alan, çevre ve uzaklık ölçümleri yapılan Atlas Vertebra kemiğinin tüm ölçüm sonuçları otomatik olarak kayıt altına alınarak denetimli öğrenmede kullanılmak üzere hazır hale getirilmiştir.

Çizelge 4.1. Ölçüm sonuçlarının örnek görüntüsü.

SIRA	ATM	APAM	ACU	ADA	AFCU	AFA
1	223,000	132,004	557,696	16873,000	271,137	4764,000
2	274,046	140,057	668,080	22181,500	304,877	6200,000
3	221,002	131,061	552,867	17063,000	278,250	5083,000
4	233,002	137,015	580,666	18823,000	278,149	5396,500
5	239,075	137,091	594,423	19022,500	284,836	5372,500
6	230,020	132,306	571,110	17590,500	293,765	5640,000
7	244,008	152,161	618,607	22091,000	323,463	6929,500
8	259,002	155,048	647,921	23106,000	310,635	6108,500
9	232,054	138,521	583,009	19581,000	316,350	6934,500
10	235,009	146,055	600,323	19696,000	279,037	5248,500
11	244,008	151,053	619,536	21067,500	335,806	7648,500
12	218,009	128,472	543,210	16282,000	297,321	5895,500
13	301,239	150,652	733,879	25899,500	341,563	7635,000
14	252,002	140,228	623,495	19834,000	290,836	5377,500
15	248,002	149,405	621,779	21918,000	302,693	6054,500
16	243,051	138,177	603,252	20215,500	323,522	6862,500
17	223,009	153,000	576,264	19479,000	314,693	6410,500
18	217,002	130,096	542,867	16706,000	280,794	5206,000
19	224,143	139,004	580,139	18985,500	316,392	6705,000
20	228,020	144,014	578,222	19029,500	301,220	5834,000
21	204,002	132,306	523,210	16544,000	292,836	5690,500
22	227,002	142,127	585,696	20163,000	310,392	6930,000
23	226,000	132,095	567,796	17226,500	277,137	5237,000
24	278,065	144,347	679,980	24154,000	338,291	7516,500

BÖLÜM 5

MAKİNE ÖĞRENMESİ İLE CİNSİYET TAHMİNİ SONUÇLARI

Cinsiyet tahmini yapabilmek için ilk olarak veri elde edilmeli, ikinci aşama olarak bu verilerden öznitelikler çıkarılmalı ve son olarak bir tahmin yapacak makine öğrenmesi algoritması seçilmelidir. Morfometrik ölçüm sonucunda KIBT ile alınan Atlas Vertebra görüntüsü üzerinde gerçekleştirilen otomatik morfometrik ölçümler ile veriler toplanmaktadır. Toplanan bu veriler öznitelik olarak dosyaya kaydedilerek veri seti oluşturulmaktadır. Oluşturulan bu veri setinin tamamı etiketli verilerden oluşmaktadır. Son aşamada ise farklı makine öğrenmesi algoritmaları kullanılarak tahminler gerçekleştirilmektedir. Cinsiyet tahmini gibi tahmin ve sınıflandırma işlemleri makine öğrenmesinin temel görevlerinden olduğu bilinmektedir [57]. Bu çalışmada makine öğrenmesi algoritmalarından Naive Bayes (NB), Destek Vektör Makineleri (DVM), K En Yakın Komşuluk (KEYK), Karar Ağaçları (KA), Rastgele Orman (RO) kullanılmaktadır.

Cinsiyet tahmini aşamasında kullanılacak olan makine öğrenmesi algoritmalarının başarı, performans durumlarının literatürdeki diğer örneklerine göre karşılaştırmak adına doğruluk (accuracy), duyarlılık (recall), kesinlik (precision), F1 skor (F1 score), ortalama karekök hata (rmse), ortalama kare hata (mse) değerleri seçilmiştir. Her yöntemin hata matrisi olarak bilinen o yonteme ait performansın görselleştirilmesini sağlayan karmaşıklık matrisi çizilmiştir.

	<i>Tahminlenen</i>	
<i>Gerçekleşen</i>	<i>Gerçek Pozitif (GP)</i>	<i>Yanlış Negatif (YN)</i>
	<i>Yanlış Pozitif (YP)</i>	<i>Gerçek Negatif (GN)</i>

Şekil 5.1. Karmaşıklık matrisi.

$$\text{Doğruluk} = \frac{GP + GN}{GP + YP + GN + YN} \quad (5.1)$$

$$\text{Kesinlik} = \frac{GP}{GP + YP} \quad (5.2)$$

$$\text{Duyarlılık} = \frac{GP}{GP + YN} \quad (5.3)$$

$$\text{F1 skor} = 2 * \frac{\text{Kesinlik} * \text{Duyarlılık}}{\text{Kesinlik} + \text{Duyarlılık}} \quad (5.4)$$

$$\text{MSE} = \frac{1}{n} \sum_{t=1}^n e_t^2 \quad (5.5)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n e_t^2} \quad (5.6)$$

Bu çalışmada kullanılan makine öğrenmesi yöntemlerinde yapılan deneysel çalışmalarda test ve eğitim kümelerinin veri seti boyutları farklı şekillerde bölünerek en başarılı sonuçlar alınmaya çalışılmıştır. Alınan sonuçlar karmaşıklık matrisi ile

gösterilmiştir. Şekil 5.1’de gösterilen karmaşıklık matrisinde GP ve GN değerleri doğru değer sayısını vermektedir. YP ve YN değerleri ise yanlış değer sayılarını vermektedir. Şekil 5.2’de gösterilen kod parçasında eğitim seti ve test seti bölünmesi yapılmıştır.

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeClassifier
from sklearn.metrics import confusion_matrix
from sklearn.metrics import accuracy_score
from sklearn.metrics import recall_score
from sklearn.metrics import f1_score
from sklearn.metrics import precision_score
from sklearn.metrics import mean_squared_error
from sklearn.metrics import r2_score
import math
import numpy as np
from sklearn.model_selection import cross_val_score

data = pd.read_excel('sample1.xlsx')
X = data.drop(columns=['CINS', 'SIRA', 'ADSOYAD'])
y = data['CINS'].values

x_train, x_test, y_train, y_test = train_test_split(X,y,test_size=0.35,random_state=42, stratify=y)
```

Şekil 5.2. Test – eğitim veri seti bölünmesi örneği.

Şekil 5.2’de verilen kod parçasında rastgele durum (random_state) parametresi test ve eğitim veri kümelerini rastgele biçimde bölmektedir. Modelin kalitesinde önemli etkisi olduğu bilinmektedir. Bu değer belirtilmediği durumlarda eğitim ve test kümeleri kodlar her çalıştığında farklı şekillerde bölünmekte ve modelin her defasında farklı bir sonuç verdiği görülmektedir. Rastgele durum değerine sabit bir tamsayı tanımlandığında eğitilen veriler her zaman sabit olmakta ve hata ayıklama işlemi kolaylaşmaktadır.

Eğitim ve test verilerinin bölünmesinde kullanılan katmanlaştırma parametresi (stratify) değeri, tahmin edilecek sütun değeri ile eşitlenerek y değişkenindeki 0 ve 1 değerlerine sahip ikili kategorik değişkenlerin bölünmesinin paralel olmasını sağlamaktadır. Test boyutu (test size) parametresi ise eğitim ve test verilerinin bölünmesi sırasında tüm veri seti içinde test yapılacak veri oranını belirtmektedir.

Bu çalışmada kullanılan ilk yöntem sınıflandırma algoritmalarından olan makine öğrenmesi yöntemi Naive Bayes’tir. Gaussian Naive Bayes, Multinomial Naive Bayes, Bernoulli Naive Bayes olmak üzere 3 türde bu yöntem kullanılabilir. Veri setimizdeki özelliklerin sürekli değerlere sahip olduğunu düşünülerek, normal

dağıtıldığı varsayılmış ve Gaussian Naive Bayes yöntemi uygulanmıştır. Gaussian Naive Bayes yöntemi farklı test ve eğitim veri kümelerine ayrılarak her bölünmedeki başarı oranları Çizelge 5.1’de verilmiştir.

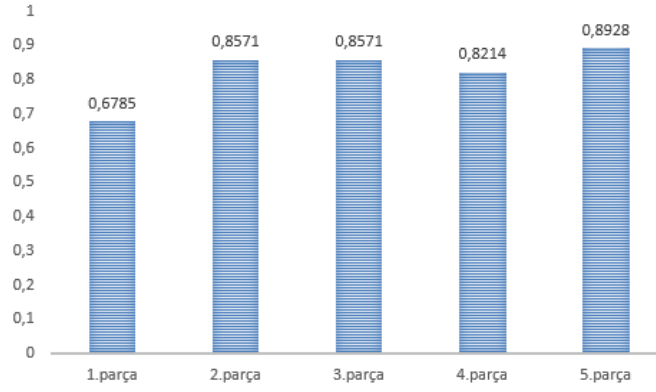
Çizelge 5.1. Gaussian Naive Bayes yönteminde test – eğitim setleri bölümlenmesine göre başarı sonuçları.

Test Verisi (%)	Eğitim Verisi (%)	Duyarlılık Oranı (%)	Kesinlik Oranı (%)	F1 skor	Doğruluk Oranı (%)
20	80	89,28	89,48	89,27	89,28
25	75	91,42	91,55	91,41	91,42
30	70	90,47	90,84	90,45	90,47
35	65	87,75	88,75	87,65	87,75
40	60	87,5	87,93	87,46	87,5

Makine öğrenmesi yöntemlerinde modelin görmediği veriler üzerindeki performansını objektif olarak ve en doğru şekilde bulmak için kullanılan değerlendirme yöntemine çapraz doğrulama denilmektedir. Bu yöntem ile istatistiksel olarak yeniden örnekleme yapılmaktadır.

Çizelge 5.2. Gaussian Naive Bayes hata metrikleri tablosu.

Test Verisi (%)	Eğitim Verisi (%)	Rmse Değeri	Mse Değeri
20	80	0,3273	0,1071
25	75	0,2927	0,0857
30	70	0,3086	0,0952
35	65	0,3499	0,1224
40	60	0,3535	0,1250



Şekil 5.3. Gaussian Naive Bayes çapraz doğrulama yöntemi grafiği.

Bu çalışmada Gaussian Naive Bayes yöntemi kullanılarak elde edilen en yüksek başarı olan %91,42 sonucunun alındığı model çapraz doğrulama yöntemi ile tekrardan test edilerek elde edilen grafik Şekil 5.3'te verilmiştir. Grafik incelendiğinde çaprazlama yöntemi ile veri seti 5 parçaya ayrılarak model test edilmiştir. En yüksek başarı 0,8928 ile 5. parçada bulunmuştur. Modelin ortalama başarısı 0,8214 olarak belirlenmiştir. Modelin yapmış olduğu tahmin sonuçlarına göre en yüksek başarının alındığı modelin oluşan karmaşıklık matrisi Çizelge 5.3'te verilmiştir.

Çizelge 5.3. Gaussian Naive Bayes yöntemi sonucu oluşan karmaşıklık matrisi.

GERÇEK	TAHMİN	
	VAR	YOK
VAR	15	2
YOK	1	17

Bu çalışmada ikinci yöntem sınıflandırma algoritmalarından olan makine öğrenmesi yöntemlerinden destek vektör makineleridir. İki sınıfa ait verileri birbirinden en uygun şekilde ayırabilmek amacıyla kullanılan bu yöntemde karar sınırları düzlemler belirlemektedir. Destek Vektör Makineleri veri setinin doğrusal ayrılabilme ya da ayrılamama durumlarına göre Doğrusal Destek Vektör Makineleri ve Doğrusal

Olmayan Destek Vektör Makineleri olmak üzere ikiye ayrılmaktadır. Çalışmamızda Doğrusal Destek Vektör Makineleri yöntemi kullanılmaktadır.

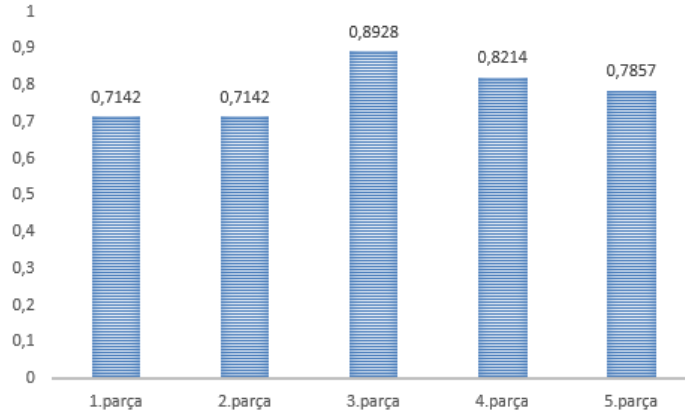
Çizelge 5.4. Destek Vektör Makineleri yönteminde test – eğitim setleri bölümlenmesine göre başarı sonuçları.

Test Verisi (%)	Eğitim Verisi (%)	Duyarlılık Oranı (%)	Kesinlik Oranı (%)	F1 skor	Doğruluk Oranı (%)
20	80	96,42	96,66	96,42	96,42
25	75	91,42	91,55	91,41	91,42
30	70	83,33	83,4	83,32	83,33
35	65	89,79	91,49	89,66	89,79
40	60	91,07	92,42	90,99	91,04

Model tam başarısının ölçümü için çapraz doğrulama yöntemine başvurulmuş ve 5 parçaya ayrılan veri setinde her parça için alınan başarı oranları Şekil 5.4'te gösterilmiştir.

Çizelge 5.5. Destek Vektör Makineleri hata metrikleri tablosu.

Test Verisi (%)	Eğitim Verisi (%)	Rmse Değeri	Mse Değeri
20	80	0,1889	0,0357
25	75	0,2927	0,0857
30	70	0,4082	0,1666
35	65	0,3194	0,1020
40	60	0,2988	0,0892



Şekil 5.4. Destek Vektör Makineleri çapraz doğrulama yöntemi grafiği.

Destek Vektör Makineleri yöntemi kullanılarak elde edilen başarı oranının %96,42 olduğu görülmüştür. Farklı test ve eğitim verileri üzerinde yapılan testler sonucunda alınan başarı oranlarının çapraz doğrulama yöntemi ile tekrardan test edilmesi ile elde edilen grafik Şekil 5.4'te verilmiştir. Veri seti 5 parçaya ayrılarak model test edilmiş ve en yüksek başarı 0,8928 ile 3. parçada bulunmuştur. Modelin ortalama başarı oranı 0,7857 olarak belirlenmiştir. Şekil 5.4'te gösterilen test sonuçları incelendiğinde en yüksek başarıya sahip olan modelin karmaşıklık matrisi Çizelge 5.6'da verilmiştir.

Çizelge 5.6. Destek Vektör Makineleri yöntemi sonucu oluşan karmaşıklık matrisi.

	<i>TAHMİN</i>		
	VAR	YOK	
<i>GERÇEK</i>	VAR	14	0
	YOK	1	13

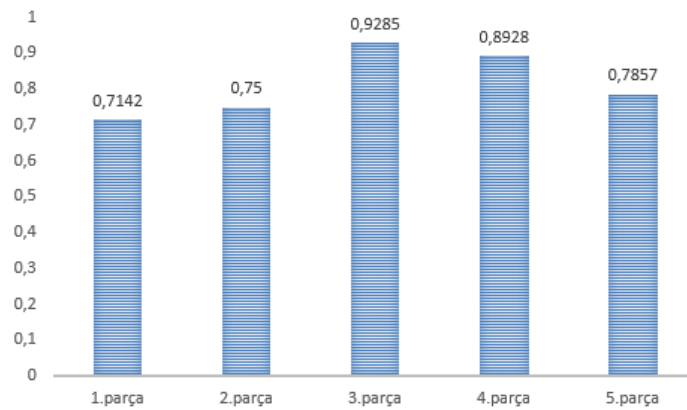
Çalışmadaki üçüncü yöntem sınıflandırma algoritmalarından olan makine öğrenmesi yöntemlerinden k en yakın komşuluktur. Sınıflandırma ve regresyon problemlerinde kullanılan K-EYK uygulama konusunda en kolay gözetimli öğrenme yöntemlerinden biri olarak bilinmektedir. Bu çalışmada kullanılan K-EYK yönteminde komşuluk değeri 15 olarak alınmıştır. Bu parametre bir noktadaki en yakın komşu sayılarını

belirtmektedir. Komşuluklar arasındaki mesafe ölçümünde öklid yöntemi seçilmiştir. Öklid uzaklık formülü 3.1 numaralı denklemde belirtilmiştir. Bu çalışmada K-EYK yöntemi farklı test ve eğitim veri kümelerine ayrılarak her bölünmedeki başarı oranları Çizelge 5.7’de verilmiştir.

Çizelge 5.7. K-En Yakın Komşuluk yönteminde test – eğitim setleri bölümlenmesine göre başarı sonuçları.

Test Verisi (%)	Eğitim Verisi (%)	Duyarlılık Oranı (%)	Kesinlik Oranı (%)	F1 skor	Doğruluk Oranı (%)
20	80	92,85	93,75	92,82	92,85
25	75	88,57	89,11	88,55	88,57
30	70	85,71	86,04	85,68	85,71
35	65	81,63	81,7	81,63	81,63
40	60	82,14	82,3	82,12	82,14

Çizelge 5.7’de gösterilen model doğruluk oranları incelendiğinde aşırı öğrenme gibi durumlar göz önüne alınarak model gerçek başarısı tespiti için çapraz doğrulama yöntemine başvurulmuştur. Şekil 5.5’te ilgili grafik belirtilmiştir.



Şekil 5.5. K en yakın komşuluk çapraz doğrulama yöntemi grafik örneği.

Çizelge 5.8. K en yakın komşuluk hata metrikleri tablosu.

Test Verisi (%)	Eğitim Verisi (%)	Rmse Değeri	Mse Değeri
20	80	0,2672	0,0714
25	75	0,3380	0,1142
30	70	0,3779	0,1428
35	65	0,4285	0,1836
40	60	0,4225	0,1785

Bu çalışmada K-EYK yöntemi kullanılarak elde edilen en yüksek başarı oranı %92,85 olduğu görülmüştür. Çapraz doğrulama yöntemi ile model doğruluğu tekrardan test edilerek elde edilen grafik Şekil 5.5'te verilmiştir. Grafik incelendiğinde çapraz doğrulama yöntemi ile veri seti 5 parçaya ayrılarak model test edilmiştir. En yüksek başarı değerinin 0,9285 modelin ortalama başarısının 0,8142 olduğu tespit edilmiştir. Çizelge 5.7 incelendiğinde en yüksek başarıya sahip modelin karmaşıklık matrisi Çizelge 5.9'da verilmiştir.

Çizelge 5.9. K en yakın komşuluk yöntemi sonucu oluşan karmaşıklık matrisi.

GERÇEK	TAHMİN	
	VAR	YOK
VAR	14	0
YOK	2	12

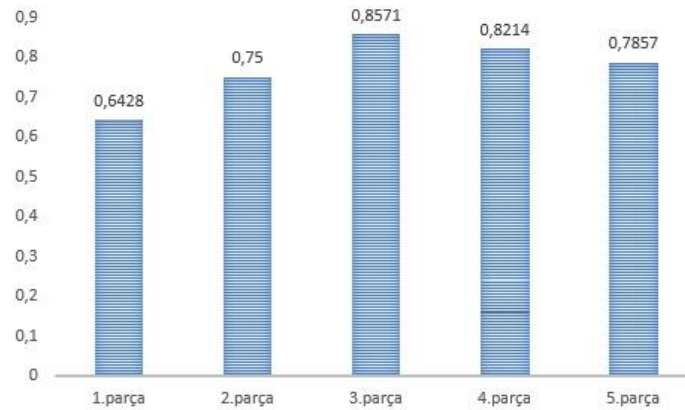
Çalışmadaki dördüncü yöntem sınıflandırma algoritmalarından olan makine öğrenmesi yöntemlerinden karar ağacıdır. Denetimli bir sınıflandırma algoritması olan Karar Ağacı yöntemi genel olarak tüm sınıflandırma ve regresyon problemlerinin çözümlerine uyarlanabilmektedir. Az oranda veri hazırlığına ihtiyaç duyması, sayısal ve kategorik verileri işleyebilmesi, çok çıktılı sorunları ele alabilmesi gibi avantajlarından dolayı deneysel çalışmalar içinde kullanılmıştır. KA yöntemi farklı

test ve eğitim veri kümelerine ayrılarak her bölünmedeki başarı oranları Çizelge 5.10'da verilmiştir.

Çizelge 5.10. Karar ağaçları yönteminde test – eğitim setleri bölümlemesine göre başarı sonuçları.

Test Verisi (%)	Eğitim Verisi (%)	Duyarlılık Oranı (%)	Kesinlik Oranı (%)	F1 skor	Doğruluk Oranı (%)
20	80	85,71	85,71	85,71	85,71
25	75	85,71	85,8	85,69	85,71
30	70	85,71	86,04	85,68	85,71
35	65	83,67	83,67	83,67	83,67
40	60	85,71	85,71	85,71	85,71

Çizelge 5.10'da gösterilen model doğruluk oranları incelendiğinde model başarısının doğruluğunun kontrol edilebilmesi için çapraz doğrulama yöntemi gerektiği öngörülmüştür. Yapılan çapraz doğru işleminde veri seti 5 parçaya bölünerek modele uygulanmıştır.



Şekil 5.6. Karar ağaçları çapraz doğrulama yöntemi grafik örneği.

Çizelge 5.11. Karar ağaçları hata metrikleri tablosu.

Test Verisi (%)	Eğitim Verisi (%)	Rmse Değeri	Mse Değeri
20	80	0,3779	0,1428
25	75	0,3779	0,1428
30	70	0,3779	0,1428
35	65	0,4040	0,1632
40	60	0,3779	0,1428

Karar ağaçları yöntemi kullanılarak elde edilen en yüksek başarının %85,71 olduğu görülmüştür. Çapraz doğrulama yöntemi ile model doğruluğu tekrardan test edilerek elde edilen grafik Şekil 5.6'da verilmiştir. Grafik incelendiğinde çapraz doğrulama yöntemi ile veri seti 5 parçaya ayrılarak model test edilmiştir. En yüksek başarı değerinin 0,8214 modelin ortalama başarısının 0,8571 olduğu tespit edilmiştir. 20-80 test-eğitim bölümlenmesi yapılarak elde edilen sonucun en yüksek elde edildiği modelin karmaşıklık matrisi Çizelge 5.12'de verilmiştir.

Çizelge 5.12. Karar ağaçları yöntemi sonucu oluşan karmaşıklık matrisi.

	<i>TAHMİN</i>		
	VAR	YOK	
<i>GERÇEK</i>	VAR	12	2
	YOK	2	12

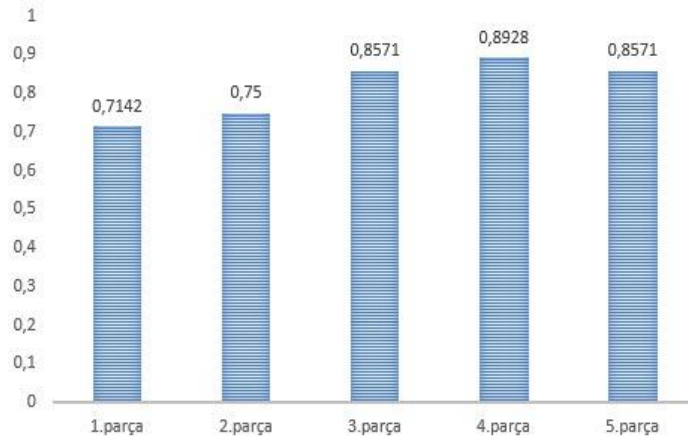
Çalışmadaki beşinci ve son yöntem sınıflandırma algoritmalarından olan makine öğrenmesi yöntemlerinden rastgele ormandır. Denetimli bir sınıflandırma algoritması olan rastgele orman yöntemi basit anlamda rastgele bir orman oluşturmakta ve ormandaki ağaç sayısı ile doğru orantılı olarak sonuç vermektedir. Ağaç sayısının artış sonuçlardaki kesinliği de artırmaktadır. Aşırı uyum problemini ortadan kaldırması, sınıflandırıcı da eksik olan değerli ele alabilmesi hem sınıflandırma hem de regresyon

görevlerinde kullanılabilmesi sebebiyle bu yöntemle deneysel çalışmalar içinde yer verilmiştir. RO yöntemi farklı test ve eğitim veri kümelerine ayrılarak her bölünmedeki başarı oranları Çizelge 5.13'te verilmiştir.

Çizelge 5.13. Rastgele orman yönteminde test – eğitim setleri bölümlemesine göre başarı sonuçları.

Test Verisi (%)	Eğitim Verisi (%)	Duyarlılık Oranı (%)	Kesinlik Oranı (%)	F1 skor	Doğruluk Oranı (%)
20	80	92,85	93,75	92,82	92,85
25	75	91,42	91,57	91,42	91,42
30	70	90,47	90,47	90,47	90,47
35	65	83,67	83,67	83,67	83,67
40	60	89,28	89,28	89,28	89,28

Çizelge 5.13 incelendiğinde eğitim-test bölümlemesine göre %20-80'lik oranda oluşturulan modelin en başarılı sonucu verdiği görülmektedir. Alınan tüm sonuçlarda n_estimators parametresi 200, max_depth parametresi 4 olarak alınmıştır. Burada n_estimator değeri oluşturulan ormandaki bulunması gereken ağaç miktarını temsil etmektedir. Max_depth parametresi ise oluşturulan ağaçların derinliğinin ne kadar olması gerektiğini belirtmektedir. Bu testler sonucunda modelin gerçek başarısını öğrenmek için çapraz doğrulama işlemi yapılmaktadır. Çapraz doğrulama yapılırken tüm veri seti 5 parçaya bölünerek oluşturulan modelin her parçadaki değerine bakılmakta ve ortalama ağırlık değeri modelin başarısını vermektedir.



Şekil 5.7. Rastgele orman çapraz doğrulama yöntemi grafik örneği.

Çizelge 5.14. Rastgele orman hata metrikleri tablosu.

Test Verisi (%)	Eğitim Verisi (%)	Rmse Değeri	Mse Değeri
20	80	0,2672	0,0714
25	75	0,2927	0,0857
30	70	0,3086	0,0952
35	65	0,4040	0,1632
40	60	0,3273	0,1071

RO yöntemi kullanılarak elde edilen en yüksek başarının %92,85 olduğu görülmüştür. Çapraz doğrulama yöntemi ile model doğruluğu tekrardan test edilerek elde edilen grafik Şekil 5.7’de verilmiştir. Çapraz doğrulama yöntemi ile 5 parçaya ayrılan veri setinde model test edilerek en yüksek başarı değerinin 0,8928 ile 4. parçada olduğu görülmektedir. Modelin ortalama başarısının 0,8142 olduğu tespit edilmiştir. 20-80 test-eğitim bölümlenmesi yapılarak elde edilen sonucun en yüksek elde edildiği modelin karmaşıklık matrisi aşağıda verilmiştir.

Çizelge 5.15. Rastgele orman yöntemi sonucu oluşan karmaşıklık matrisi.

GERÇEK	TAHMİN		
		VAR	YOK
	VAR	14	0
	YOK	2	12

Görüntü işleme yöntemleri ile yapılan ölçüm işlemleri ile oluşturulan veri seti üzerinde yukarıdaki 5 adet makine öğrenmesi yöntemleri kullanılarak cinsiyet tahmini işlemleri gerçekleştirilmiştir. Buna göre çapraz doğrulama işlemi sonrasında en yüksek başarı oranı K-EYK yöntemi ile elde edilmiştir.

Çizelge 5.16. Uygulanan makine öğrenmesi yöntemlerinin test – eğitim setleri bölümlenmesine göre başarı sonuçlarının karşılaştırılması.

Makine Öğrenmesi Yöntemi	Test Verisi (%)	Eğitim Verisi (%)	Çapraz Doğrulama (%)	Doğruluk Oranı (%)	RMSE	MSE
Destek Vektör Makineleri	20	80	78,57	96,42	0,18	0,03
K en yakın komşu	20	80	81,42	92,85	0,26	0,07
Karar Ağacı	20	80	78,57	85,71	0,37	0,14
Rastgele Orman	20	80	81,42	92,85	0,26	0,07
Gaussian Naive Bayes	25	75	82,14	91,42	0,29	0,08

Çalışmada görüntü işleme yöntemleri kullanılarak yapılan otomatik morfolometrik ölçüm işlemleri sonucunda hazırlanan veri seti 5 farklı makine öğrenmesi yöntemi kullanılarak cinsiyet tahmini gerçekleştirme işlemi yapılmıştır. Model başarı oranlarının %81 ile %96 arasında değişkenlik gösterdiği yapılan çapraz doğrulama işlemlerinde ise ağırlıklı başarı ortalamasının %78 ile %82 arasında değişkenlik gösterdiği tespit edilmiştir.

BÖLÜM 6

SONUÇLAR

Bu çalışmada biyomedikal görüntüleme araçlarından konik ışınli bilgisayarlı tomografi kullanılarak elde edilen görüntülerden cinsiyet tahmini gerçekleştirme işlemi yapılmıştır. Kimliklendirme çalışmalarında önemli bir yere sahip olan cinsiyet tespiti işlemi manuel işlemler ile insan faktörü ve diğer dış etkenler neticesinde çok uzun sürelere ihtiyaç duymaktadır. Ölçüm sırasında yapılan insani hatalar, görüntülerin elde edilmesi sırasındaki işlemler gibi faktörler tespit sonucunu etkilemektedir. Bu çalışmada görüntülerin elde edilme şekli, kontrast farkı, çözünürlük farkı, insan ve çevre hatalar gibi sonucu etkileyecek tüm hataların önüne geçilmek amaçlanmıştır. Atlas Vertebra görüntüsünden cinsiyet tahmini işlemi görüntü işleme ve makine öğrenmesi yöntemleri ile saniyeler sonunda tespit sonucu çıkarmaktadır.

Bu çalışmada 139 adet KIBT görüntülerinden öncelikle Atlas Vertebra kemiği tespit edilmiştir. Tespiti yapılan Atlas Vertebra kemiği tüm görüntüden çıkarılarak morfolometrik ölçüme hazırlanmıştır. Morfolometrik ölçüm işlemleri ile Atlas Vertebra üzerinde yapılan ölçümler ile öznelik çıkarımı yöntemi kullanılarak veri seti hazırlanmıştır. Hazırlanan veri seti üzerinde makine öğrenmesi yöntemleri kullanılarak Atlas Vertebra'dan cinsiyet tahmini yapılmaktadır. Çalışmada önerilen çözüm kimliklendirme çalışmalarında cinsiyet tespiti işlemlerinde yüksek performans sunmaktadır. Elde edilen sonuçlara dayanarak, çalışmada kullanılan yöntemlerin içinden en yüksek başarı oranına sahip çapraz doğrulama sonucu 0,8142 ile k-en yakın komşuluk yöntemi olduğu görülmüştür.

Atlas Vertebra'dan makine öğrenmesi yöntemleri ile cinsiyet tahminine dair çalışmalar incelendiğinde, Öztürk'ün yapmış olduğu çalışmada, bir bilgisayar yazılımı kullanılarak uzman tarafından görüntüler üzerinde etiketleme yapılmıştır. Etiketlenen görüntüler üzerinde makine öğrenmesi yöntemleri kullanılmıştır. Yapılan bu çalışmada Öztürk, %86-89 aralığında tahmin başarısı elde etmiştir [58]. Bu tez çalışmasında insandan bağımsız, tam otomatik morfolometrik ölçüm ile elde edilen

veriler makine öğrenmesinde kullanılmış ve çapraz doğrulama sonucunda %78 – 82 aralığında tahmin başarısı elde edilmiştir.

Bu alanda yapılabilecek çalışmalar düşünüldüğünde daha fazla görüntü üzerinde derin öğrenme yöntemleri kullanılarak cinsiyet tahmini gerçekleştirme işlemleri yapılabilir. Çalışmada modeller üzerinde aşırı öğrenme olduğu varsayılarak aşırı öğrenmeye karşı modelleri optimize etme ve performans artırımı, öznitelik çıkarımı yönteminde kullanılan özniteliklerin artırımı ve daha fazla makine öğrenmesi yöntemi kullanılarak başarı oranında artış hedeflenmektedir.

KAYNAKLAR

1. Gülhan, Ö., “Pelvis’ten radyolojik yöntemler ile cinsiyet tayini: Türkiye örnekleme”, *Antropoloji*, 36, 53-69 (2018).
2. Gill, G.W., “Racial variation in the proximal and distal femur: Heritability and Forensic utility”, *Journal of Forensic Sciences*, 46 (4): 791-799 (2001).
3. Kranioti, E.F., Vorniotakis, N. , Galiatsou, C. , Işcan, M.Y., Michalodimitrakis, M., “Sex identification and software development using digital femoral head radiographs”, *Forensic Science International*, 189 (1-3): 113.e1-e7 (2009).
4. Thompson, S., Black, T., “Forensic human identification: An introduction 1st edition.”, *CRC Press.*, Boca Raton,(2006).
5. Srivastava, R., Saini, V., Rai, R.K., Pandey,S., Tripathi,S.K., “A study of sexual dimorphism in the Femur among north Indians”, *Journal of Forensic Sciences*, 57 (1): 19-23 (2011).
6. Gonzalez, P.N., Bernal, V., Perez, S.I., “Geometric morphometric approach to sex estimation of human pelvis”, *Forensic Science International*, 189 (1-3): 68-74 (2009).
7. Biwasaka, H., Aoki, Y. , Sato, K., Tanijiri, T., Fujita, S., Dewa, K., Yoshioka K., Tomabeche, M., “Analyses of sexual dimorphism of reconstructed pelvic computed tomography images of contemporary Japanese using curvature of the greater sciatic notch, pubic arch and greater pelvis”, *Forensic Science International*, 219 (1-3): 288.e1-e8 (2011).
8. Karakas, H.M., Harma,A., Alicioglu, B., “The subpubic angle in sex determination: Anthropometric measurements and analyses on Anatolian Caucasians using multidetector computed tomography datasets”, *Journal of Forensic and Legal Medicine*, 20 (8): 1004-1009 (2013).
9. Otağ, İ. , Çimen, M., C. “Femurdan morfometrik yöntemlerle cinsiyet tayini”,*Cumhuriyet Üniversitesi Tıp Fakültesi Dergisi*, 25 (4): 165-170 (2003).
10. Metin Tellioglu, A., Karakaş, S., “Humerus’tan morfometrik yöntemlerle cinsiyet tayini”, *Fırat Üniversitesi Sağlık Bilimleri Tıp Dergisi*, 27 (2): 75-79 (2013).
11. Hora, M., Sladek, V., “Population specificity of sex estimation from vertebrae”, *Forensic Science International*, 291: 279.e1-e12 (2018) .
12. Ostrofsky, K.R., Churchill, S.E., “Sex Determination by discriminant function analysis of Lumbar Vertebrae”, *Journal of Forensic Sciences*, 60 (1): 21-28 (2015).

13. Tsubaki, S., Morishita, J., Usumoto, Y., Sakaguchi, K., Matsunobu, Y., Kawazoe, Y., Okumura, M., Ikeda, N., “Sex determination based on a thoracic vertebra and ribs evaluation using clinical chest radiography”, *Legal Medicine*, Tokyo, 27: 19-24 (2017).
14. İlker, M., Kafa, İ., Ari, U., “Morfometrik çalışmalarda manuel (el ile) ve dijital (sayısal)-bilgisayar destekli ölçüm yöntemlerinin karşılaştırılması”, *Uludağ Üniversitesi Tıp Fakültesi Dergisi*, 30 (3): 141-144 (2004).
15. Teke, H.Y., Duran, S., Bilge, Y., “Radyolojik metotların adli antropolojide kullanımı”, *Türkiye Klinikleri J Foren Med-Special Topics*, 3 (1): 108-12 (2017).
16. Torimitsu, S., Makino, Y., Saitoh, H., Sakuma, A., Ishii, N., Yajima, D., Inokuchi, G., Motomura, A., Chiba, F., Yamaguchi, R., Hashimoto, M., Hoshioka, Y., Iwase, H., “Sexual determination based on multidetector computed tomographic measurements of the second cervical vertebra in a contemporary Japanese population”, *Forensic Science International*, 266: 588.e1-e6 (2016).
17. Dalbay, R.S., “Kimlik ve toplumsal kimlik kavramı”, *Journal of Süleyman Demirel University Institute of Social Sciences Year*, 31 (2): 161-176 (2018).
18. Aydoğdu, H., “Modern kimlikte öznenin ölümü”, *Atatürk Üniversitesi Kazım Karabekir Eğitim Fakültesi Dergisi*, 10: 115-147 (2010).
19. Zeyfeoglu, Y., Hancı, İ. H., “İnsanlarda kimlik tespiti”, *Türk Tabipleri Birliği Sürekli Tıp Eğitimi Dergisi*, 10: 375-377 (2001).
20. Bilge, Y., “Adli Tıp Kitabı 1. baskı”, *Üçbilek matbaa*, (2005).
21. Aşirdizer, M., Yavuz, M., Zeyfeoglu, Y., “Adli Tıp Staj Ders Notları”, Celal Bayar Üniversitesi Tıp Fakültesi, Manisa (2005).
22. Aytek, A. İ., “Geometrik morfometri”, *MASROP E-Dergi*, 11 (17): 1-7 (2017).
23. Gilbert, B.M., “Anterior femoral curvature: Its probable basis and utility as a criterion of racial assessment”, *American Journal of Physical Anthropology*, 45 (3): 601-604 (1976).
24. Abdel Fatah, E.E., Shirley, N.R., Jantz, R.L., Mahfouz, M.R., “Improving sex estimation from crania using a novel three-dimensional quantitative method”, *Journal of Forensic Sciences*, 59 (3): 590-600 (2014).
25. Stewart, T., “Essentials of forensic anthropology 1st edition”, *Thomas, Washington* (1979).
26. Gosavi, S. N., Vatsalaswamy P., “Morphometric study of the Atlas Vertebra using manual method”, *Malaysian Orthopaedic Journal*, 6 (3): 18-20 (2012).
27. Arinci, K., Elhan, A., “Anatomi, 4. baskı”, *Güneş Tıp Kitabevi*, Ankara (2006).

28. Tozođlu, Ü., “Konik ışınli bilgisayarlı tomografinin alıřma prensibi ve kullanım alanları”, *Türkiye Klinikleri Ađız Diř ve ene Radyolojisi - Özel Konular*, 2 (1): 9-13 (2016).
29. Robb, R.A., “The dynamic spatial reconstructor: an X-Ray video-fluoroscopic CT scanner for dynamic volume imaging of moving organs”, *IEEE Transactions on Medical Imaging*, 1 (1): 22-33 (1982).
30. Özdede, M., Paksoy, C.S., “Konik ışınli bilgisayarlı tomografi: teknik, alıřma ilkeleri ve görüntü oluşumu”, *Türkiye Klinikleri Ađız Diř ve ene Radyolojisi - Özel Konular*, 5, 1-6 (2019).
31. Pekince, A., “Konik ışınli bilgisayarlı tomografi ile bař ve boyun bölgesinde yapılan eřitli morfometrik ölçümlerin cinsiyet tayininde kullanılabilirliklerinin deđerlendirilmesi”, Doktora Tezi, *Atatürk Üniversitesi Sađlık Bilimleri Enstitüsü*, Erzurum (2017).
32. Samtař, G., Gülesin, M., “Sayısal görüntü işleme ve farklı alanlardaki uygulamaları”, *Ejovoc (Electronic Journal of Vocational Colleges)*, 2 (1): 85-97 (2012) .
33. Jahne, B., “ Digital image processing 6th edition”, *CRC Press*, Berlin (2005).
34. Kumar, N.K., Sindhu, G.S., Prashanthi, D.K., Sulthana, A.S., “Analysis and prediction of cardio vascular disease using machine learning classifiers”, *6th International Conference on Advanced Computing and Communication Systems*, 15-21 (2020).
35. Wu, J., Cai, Z., Zhu, X., “Self-adaptive probability estimation for Naive Bayes classification”, *The 2013 International Joint Conference on Neural Networks*, 1-8 (2013).
36. Cortes, C., Vapnik, V., “Support-vector networks”, *Machine Learning*, 20 : 273-297 (1995).
37. Kudoh, T., Matsumoto, Y., “Use of support vector learning for chunk identification”, *Proceedings of the 2nd workshop on learning language in logic and the 4th conference on Computational natural language learning*, 7, 142-144 (2000).
38. Demidova, L., Egin, M., Sokolova, Y., “SVM prediction of the attestation success on the base of the poll results”, *ITM Web Conference*, 18 : 7-14 (2018).
39. Masud, M., Gao, J., Khan, L., Han, J., Thuraisingham B.M., “Classification and novel class detection in concept-drifting data streams under time constraints”, *IEEE Transactions on Knowledge and Data Engineering*, 23 (6): 859-874 (2011).
40. Takami, M., Bell, P., Ommer, B., “Offline learning of prototypical negatives for

- efficient online Exemplar SVM”, *IEEE Winter Conference on Applications of Computer Vision*, 377-384 (2014).
41. Jha, J., Ragha, L., “Intrusion detection system using support vector machine.”, *International Journal of Applied Information Systems*, ICWAC (3) ,25-30 (2013).
 42. Balakrishnan, B. K. P. D., Dahnil, M. I., Yi, W. J., “The impact of social media marketing medium toward purchase intention and brand loyalty among generation Y”, *Procedia – Social and Behavioral Sciences*, 148, 177-185 (2014).
 43. Tsai, C.F., Hsu, Y.F., Lin, C.Y., Lin, W.Y., “Intrusion detection by machine learning: A review”, *Expert Systems with Applications*, 36 (10): 11994–12000 (2009).
 44. Simsekler, M.C.E. , Qazi, A. , Alalami, M.A. , Ellahham, S. , Ozonoff, A., “Evaluation of patient safety culture using a random forest algorithm”, *Reliability Engineering System Safety*, 204: 107186 (2020).
 45. Metz, M., Lesnoff, M., Abdelghafour, F., Akbarinia, R., Masegla, F., Roger, J.M., “A big-data algorithm for KNN-PLS”, *Chemometrics and Intelligent Laboratory Systems*, 203 (2): 104076 (2020).
 46. Indurkha, N., Weiss, S., “Estimating performance gains for voted decision trees”, *Intelligent Data Analysis*, 2 (4): 303-310 (1998).
 47. Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., “Classification and regression trees 1st edition”, *CRC Press*, Boca Raton (2017).
 48. Biau, G., Scornet, E., “A random forest guided tour”, *Test*. 25,197–227 (2016).
 49. Huo, J., Shi, T., Chang, J., “Comparison of random forest and SVM for electrical short-term load forecast with different data sources”, *7th IEEE International Conference Software Engineering Service Science*, 1077-1080 (2016).
 50. Breiman, L., “Random Forests”, *Machine Learning*, 45 (1): 5–32 (2001).
 51. Açıl, T., “Güncel makine öğrenmesi teknikleri ile iris görüntülerinden cinsiyet analizi”, Yüksek Lisans Tezi, *İskenderun Teknik Üniversitesi Mühendislik ve Fen Bilimleri Enstitüsü*, Hatay (2019)
 52. Pinar, Z., Kaya Gülağiz, F., Altuncu, M. A., Şahin, S., Kocaeli Üniversitesi, “Denim kumaşlarda görüntü işleme ile hata tespiti”, *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi* ,9 (4): 1609-1620 (2020).
 53. Nar, F., Demirkesen, C., Okman, O.E., Cetin, M., “Region based target detection approach for synthetic aperture radar images and its parallel implementation”, *Proceedings of SPIE – The International Society for Optical Engineering*, 8394, 201-214 (2012).

54. Sevilmiş, B., Okman, O.E., Nar, F., Demirkese, C., Çetin, M., “A robust nonlinear scale space change detection approach for SAR images”, *Image and Signal Processing for Remote Sensing XIX*, 8892 ,889215-1 (2013).
55. Karvonen, J.A., “Baltic sea ice SAR segmentation and classification using modified pulse-coupled neural networks”, *IEEE Transactions on Geoscience and Remote Sensing*, 42, 1566-1574 (2004).
56. Ozcan, C., Sen, B., Nar, F., “Sparsity-driven despeckling for SAR images”, *IEEE Geoscience Remote Sensing Letters*, 13 (1): 115-119 (2016).
57. Lewis, N.D.C., “Machine learning made easy with R: an intuitive step by step blueprint for beginners”, *CreateSpace Independent Publishing Platform*, California, (2017).
58. Öztürk, M., “Birincil servikal vertebranın antropometrik ölçümler ile makine öğrenme algoritmaları kullanılarak cinsiyet tayini üzerine bir çalışma”, Yüksek Lisans Tezi, *Karabük Üniversitesi Lisansüstü Eğitim Enstitüsü* Karabük (2021).

ÖZGEÇMİŞ

Muhammed Çağrı ÇELİK ilk ve orta öğrenimini Adana'da tamamladı. Adana Piri Reis Anadolu Lisesinden mezun olduktan sonra 2010 yılında Karabük Üniversitesi Bilgisayar Mühendisliği bölümünde üniversite eğitimine başladı. 2016 yılında üniversite mezuniyeti sonrası yazılım sektöründe tecrübe kazanmak için Ankara'da çalışma hayatına başladı. Özel sektörde 2 senelik yazılım geliştirme uzman yardımcısı, Ar-Ge mühendisi ve bilgisayar mühendisi gibi görevlerde çalıştıktan sonra 2017 Kasım ayında yazılım geliştirme uzman yardımcısı olarak EGM Siber Suçlarla Mücadele Daire Başkanlığında göreve başladı. Şu an yazılım geliştirme uzmanı olarak görevine devam etmektedir.