



**YÜZ GÖRÜNTÜLERİNDEN GEOMETRİK
ÖZNİTELİKLER ÇIKARILARAK ÇOCUKLARDAKİ
DİKKAT EKSİKLİĞİ VE HİPERAKTİVİTE
BOZUKLUĞUNUN DENETİMLİ ÖĞRENME
ALGORİTMALARIYLA TESPİTİ VE ANALİZİ**

Satuk Buğrahan ÖZTÜRK

**Yüksek Lisans Tezi
Bilgisayar Mühendisliği Ana Bilim Dalı
Dr. Öğr. Üyesi Barış ÖZYER
2022**

(Her hakkı saklıdır.)

T.C.
ATATÜRK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANA BİLİM DALI

**YÜZ GÖRÜNTÜLERİNDEN GEOMETRİK ÖZNİTELİKLER ÇIKARILARAK
ÇOCUKLARDAKİ DİKKAT EKSİKLİĞİ VE HİPERAKTİVİTE BOZUKLUĞUNUN
DENETİMLİ ÖĞRENME ALGORİTMALARIYLA TESPİTİ VE ANALİZİ**

(Detection and Analysis of Attention Deficit and Hyperactivity Disorder in Children with
Supervised Learning Algorithms by Extracting Geometric Features from Facial Images)

YÜKSEK LİSANS TEZİ

Satuk Buğrahan ÖZTÜRK

Danışman: Dr. Öğr. Üyesi Barış ÖZYER

Erzurum
Ocak, 2022

T.C.
ATATÜRK ÜNİVERSİTESİ
Fen Bilimleri Enstitüsü Müdürlüğü

TEZ KABUL VE ONAY TUTANAĞI

**YÜZ GÖRÜNTÜLERİNDEN GEOMETRİK ÖZNİTELİKLER ÇIKARILARAK
ÇOCUKLARDAKİ DİKKAT EKSİKLİĞİ VE HİPERAKTİVİTE BOZUKLUĞUNUN
DENETİMLİ ÖĞRENME ALGORİTMALARIYLA TESPİTİ VE ANALİZİ**

Dr. Öğr. Üyesi Barış ÖZYER danışmanlığında, Satuk Buğrahan ÖZTÜRK tarafından hazırlanan bu çalışma, 03/01/2022 tarihinde aşağıdaki jüri tarafından Bilgisayar Mühendisliği Anabilim Dalı Bilgisayar Mühendisliği Bilim Dalı'nda yüksek lisans tezi olarak **oybirliği (3/0)** ile kabul edilmiştir.

Jüri Başkanı: Doç. Dr. Akif DURDU

Konya Teknik Üniversitesi

Ashı Islak İmzalıdır.

Danışman: Dr. Öğr. Üyesi Barış ÖZYER

Atatürk Üniversitesi

Ashı Islak İmzalıdır.

Jüri Üyesi: Dr. Öğr. Üyesi Mete YAĞANOĞLU

Atatürk Üniversitesi

Ashı Islak İmzalıdır.

Enstitü Yönetim
Kurulunun/.../.... tarih
ve sayılı kararı.

Bu tezin Atatürk Üniversitesi Lisansüstü Eğitim ve Öğretim Yönetmeliği'nin ilgili maddelerinde belirtilen şartları yerine getirdiğini onaylarım.

Prof. Dr. Saltuk Buğrahan CEYHUN

Enstitü Müdürü

Ashı Islak İmzalıdır.

ETİK BİLDİRİM VE İNTİHAL BEYAN FORMU

Yüksek Lisans Tezi olarak Dr. Öğr. Üyesi Barış ÖZYER danışmanlığında sunulan “Yüz Görüntülerinden Geometrik Öznitelikler Çıkarılarak Çocuklardaki Dikkat Eksikliği ve Hiperaktivite Bozukluğunun Denetimli Öğrenme Algoritmalarıyla Tespiti ve Analizi” başlıklı çalışmanın tarafımızdan bilimsel etik ilkelere uyularak yazıldığını, yararlanılan eserlerin kaynakçada gösterildiğini, Fen Bilimleri Enstitüsü tarafından belirlenmiş olan Turnitin Programı benzerlik oranlarının aşılmadığını ve aşağıdaki oranlarda olduğunu beyan ederiz.

Tez Bölümleri	Tezin Benzerlik Oranı (%)	Maksimum Oran (%)
Giriş	0	30
Kuramsal Temeller	2	30
Materyal ve Yöntem	3	35
Bulgular	2	20
Tartışma	1	20
Tezin Geneli	7	25

Not: Yedi kelimeye kadar benzerlikler ile Başlık, Kaynakça, İçindekiler, Teşekkür, Dizin ve Ekler kısımları tarama dışı bırakılabilir. Yukarıdaki azami benzerlik oranları yanında tek bir kaynaktan olan benzerlik oranlarının %5'den büyük olmaması gerekir.

Beyan edilen bilgilerin doğru olduğunu, aksi halde doğacak hukuki sorumlulukları kabul ve beyan ederiz.

Tez Yazarı (Öğrenci)	Tez Danışmanı
Satuk Buğrahan ÖZTÜRK	Dr. Öğr. Üyesi Barış Özyer
3.1.2022	3.1.2022
İmza: Aslı Islak İmzalıdır.	İmza: Aslı Islak İmzalıdır.

* Tez ile ilgili YÖKTEZ'de yayınlamasına ilişkin bir engelleme var ise aşağıdaki alanı doldurunuz.

☐ Tezle ilgili patent başvurusu yapılması / patent alma sürecinin devam etmesi sebebiyle Enstitü Yönetim Kurulunun/....../.... tarih ve sayılı kararı ile teze erişim 2 (iki) yıl süreyle engellenmiştir.

☐ Enstitü Yönetim Kurulunun/....../.... tarih ve sayılı kararı ile teze erişim 6 (altı) ay süreyle engellenmiştir.

TEŞEKKÜR

Tezimin başlangıcından bitimine kadar destek ve katkılarını esirgemeyen danışmanım Sayın Dr. Öğr. Üyesi Barış ÖZYER'e, tez aşamasında yol gösteren Sayın Dr. Öğr. Üyesi Gülşah TÜMÜKLÜ ÖZYER'e, tezime fikirleriyle katkı sağlayan Sayın Arş. Gör. Ömer Çağrı YAVUZ'a teşekkür ederim.

Son olarak tezimin her aşamasında yanımda olan ve desteklerini esirgemeyen başta babam Ahmet Emin ÖZTÜRK olmak üzere, sevgili aileme teşekkür ederim.

Satuk Buğrahan ÖZTÜRK



ÖZET

YÜKSEK LİSANS TEZİ

YÜZ GÖRÜNTÜLERİNDEN GEOMETRİK ÖZNİTELİKLER ÇIKARILARAK ÇOCUKLARDAKİ DİKKAT EKSİKLİĞİ VE HİPERAKTİVİTE BOZUKLUĞUNUN DENETİMLİ ÖĞRENME ALGORİTMALARIYLA TESPİTİ VE ANALİZİ

Satuk Buğrahan ÖZTÜRK

Danışman: Dr. Öğr. Üyesi Barış ÖZYER

Amaç: Tezin amacı yüz görüntülerinden geometrik öznitelikler çıkarılarak çocuklardaki dikkat eksikliği ve hiperaktivite bozukluğunun denetimli makine öğrenme algoritmaları ile tespiti ve literatürde yaygın olarak kullanılan sınıflandırma algoritmalarının başarı oranlarının analizidir. Bu amaç doğrultusunda yüz görüntü verilerinin toplanacağı ve analizlerin yapılacağı bilgisayarlı görü tabanlı mobil uygulama sistemi geliştirilmiştir.

Yöntem: Bu çalışma kapsamında ilk olarak yüz görüntülerinden geometrik öznitelikler çıkarılmıştır. Yüz üzerindeki farklı bölgelerde bulunan iki boyutlu noktalar FaceNet modeli kullanılarak elde edilmiştir. Sonrasında yüz hizalama ağı kullanılarak iki boyutlu noktalar yüz görüntülerinden 3 boyutlu yüz görüntüleri ve yüz üzerindeki noktalar elde edilmiştir. Elde edilen noktalar arasında uzunluk ve oranlar öznitelik vektörleri olarak belirlenmiştir. K-Best, rasgele karar ormanları, modelden seçim ve T-Test yöntemleri ile öznitelik seçimi gerçekleştirilmiştir. K en yakın komşu, naive bayes ve destek vektör makineleri yöntemleri sınıflandırma aşamasında kullanılmıştır. İnternet ortamından elde edilen yüz görüntülerinden veri seti oluşturulmuştur. Bilgisayarlı görü tabanlı algoritmaların ve makine öğrenmesi yöntemlerinin kullanıldığı bir mobil uygulama geliştirilmiştir.

Bulgular: Öznitelik vektörleri arasında en yüksek başarı yüzdesi, yüz üzerindeki tüm uzunlukların kullanıldığı 2 boyutlu öznitelik vektörünün rastgele karar ormanları öznitelik seçimi yöntemi ile beraber destek vektör makineleri algoritmasının uygulandığı analiz sonucunda %80 doğruluk elde edilmiştir. 3 boyutlu öznitelik vektörü için destek vektör makineleri ile %65 doğruluk elde edilmiştir. Öznitelik sayısının düşük olduğu veri setlerinde k en yakın komşu ve modelden seçim algoritmalarının beraber kullanılması, öznitelik sayısının yüksek olduğu veri setlerinde ise destek vektör makineleri ve rastgele karar ağaçları algoritmalarının beraber kullanımının başarı sonucunu artıracığı söylenebilir.

Sonuç: Sınıflandırma işlemlerinden elde edilen veriler doğrultusunda öznitelik seçiminin sınıflandırma işlemine olumlu etkilerinin olduğu görülmüştür. Yüzden çıkarılan geometrik öznitelikler ile dikkat eksikliği ve hiperaktivite bozukluğu hastalığı arasında bir korelasyon olduğu gözlemlenmiştir.

Anahtar Kelimeler: yüz tanıma, dikkat eksikliği, hiperaktivite bozukluğu, makine öğrenmesi, sınıflandırma

Ocak 2022, 58 sayfa

ABSTRACT

MASTER'S THESIS

DETECTION AND ANALYSIS OF ATTENTION DEFICIT AND HYPERACTIVITY DISORDER IN CHILDREN WITH SUPERVISED LEARNING ALGORITHMS BY EXTRACTING GEOMETRIC FEATURES FROM FACIAL IMAGES

Satuk Buğrahan ÖZTÜRK

Supervisor: Asst. Prof. Dr. Barış ÖZYER

Purpose: The aim of the thesis is to determine the attention deficit hyperactivity disorder in children with supervised machine learning algorithms by extracting the geometric features from the facial images and to analyze the success rates of the classification algorithms commonly used in the literature. For this purpose, a computer vision based mobile application system has been developed to collect and analyze facial image data.

Method: In this study, firstly, geometric features were extracted from face images. Two-dimensional points located in different regions on the face were obtained using the FaceNet model. Then, using the face alignment mesh, 3D face images and points on the face were obtained from two-dimensional dot face images. The lengths and ratios between the obtained points were determined as feature vectors. Feature selection was carried out using K-Best, random decision forests, model selection and T-Test methods. K nearest neighbor, naive bayes and support vector machine methods are used in the classification phase. A data set was created from face images obtained from the internet. A mobile application has been developed in which computer vision-based algorithms and machine learning methods are used.

Findings: The highest percentage of success among the feature vectors, the random decision forest feature selection method of the 2-dimensional feature vector in which all the lengths on the face are used and the support vector machines algorithm were applied together with 80% accuracy. For the 3D feature vector, 65% accuracy was obtained with the support vector machines. It can be said that the use of k nearest neighbor and model selection algorithms together in data sets with low number of features, and the use of support vector machines and random decision tree algorithms together in datasets with high number of features will increase the success result.

Results: According to the data obtained from the classification processes, it has been seen that feature selection has positive effects on the classification process. It has been observed that there is a correlation between geometric features extracted from the face and attention deficit hyperactivity disorder disease.

Keyword: face detection, attention deficit, hyperactivity disorder, machine learning, classification

January 2022, 58 pages

İÇİNDEKİLER

KABUL VE ONAY TUTANAĞI.....	i
ETİK BİLDİRİM VE İNTİHAL BEYAN FORMU	ii
TEŞEKKÜR	iii
ÖZET	iv
ABSTRACT	v
İÇİNDEKİLER.....	vi
TABLolar DİZİNİ.....	viii
ŞEKİLLER DİZİNİ	ix
KISALTMALAR VE SİMGELER DİZİNİ.....	x
GİRİŞ.....	1
Tezin Amacı ve Katkısı.....	2
KURAMSAL TEMELLER.....	4
MATERYAL VE METOT.....	13
Sistemin Geliştirilmesi	13
Veri Seti	15
Öznitelik Çıkarımı.....	15
İki boyutlu görüntülerden yüz noktalarının tespiti.....	16
Üç boyutlu görüntülerden yüz noktalarının tespiti.....	18
Yüz Noktalarının İşlenmesi.....	19
Öznitelik Seçimi.....	23
K-Best	24
Rastgele Karar Ormanları (Random Forest)	24
Modelden Seçim (Select From Model)	25
T-Test	25
Sınıflandırma.....	26
K En Yakın Komşu	26
Naive Bayes	26
Destek Vektör Makineleri	28
Doğruluk ve Değerlendirme Metrikleri	29
ARAŞTIRMA BULGULARI VE TARTIŞMA.....	32
Sınıflandırma Algoritmalarının Sonuçları	32
Öznitelik Seçimi Sınıflandırma Sonuçları	34
3 Boyutlu Yeniden İnşa Sonucu Elde Edilen Uzunlukların Sınıflandırma Sonuçları	38

SONUÇLAR VE ÖNERİLER	40
KAYNAKÇA	42
ÖZGEÇMİŞ.....	46



TABLÖLAR DİZİNİ

Tablo 1. Çalışmada Kullanılan Veri Seti	15
Tablo 2. Önerilen 16 nokta ve koordinat bilgileri.....	23
Tablo 3. Sınıflandırma İşlemi İçin Veri Setinden Elde Edilen Noktaların Uzunluklarından Elde Edilen Öznitelikler	24
Tablo 4. Öznitelik 1'in Sınıflandırma Sonuçlarının Doğruluk Değerleri	32
Tablo 5. Öznitelik 1'in Sınıflandırma Sonuçları	32
Tablo 6. Öznitelik 2'nin Sınıflandırma Sonuçlarının Doğruluk Değerleri	33
Tablo 7. Öznitelik 2'nin Sınıflandırma Sonuçları.....	33
Tablo 8. Öznitelik 3'ün Sınıflandırma Sonuçlarının Doğruluk Değerleri	34
Tablo 9. Öznitelik 3'ün Sınıflandırma Sonuçları.....	34
Tablo 10. Öznitelik 1'in Öznitelik Seçimi Sonrası Sınıflandırma Sonuçlarının Doğruluk Değerleri	35
Tablo 11. Öznitelik 1'in Öznitelik Seçimi Sonrası Sınıflandırma Sonuçları.....	35
Tablo 12. Öznitelik 2'nin Öznitelik Seçimi Sonrası Sınıflandırma Sonuçlarının Doğruluk Değerleri	36
Tablo 13. Öznitelik 2'nin Öznitelik Seçimi Sonrası Sınıflandırma Sonuçları.....	36
Tablo 14. Öznitelik 3'ün Öznitelik Seçimi Sonrası Sınıflandırma Sonuçlarının Doğruluk Değerleri	37
Tablo 15. Öznitelik 3'ün oranlarının öznitelik seçimi sonrası sınıflandırma sonuçları.....	38
Tablo 16. Öznitelik 4'ün Sınıflandırma Sonuçlarının Doğruluk Değerleri	39
Tablo 17. Öznitelik 4'ün Sınıflandırma Sonuçları.....	39

ŞEKİLLER DİZİNİ

Şekil 1. Sistemin çalışma diyagramı	14
Şekil 2. Mobil uygulama sistem diyagramı.....	15
Şekil 3. CNN çalışma mantığı.....	16
Şekil 4. 4x4 matrisin 2x2 matrise indirgenmesi	17
Şekil 5. 2 boyutlu görüntüde işaretlenmiş yüz noktaları	19
Şekil 6. 2 boyutlu görüntünün 3 boyutlu olarak yeniden inşa edilmesi sonrası işaretlenen yüz noktaları	20
Şekil 7. Verilerin Firestore veri tarabanında saklanması	20
Şekil 8. Önerilen 16 uzunluğun model üzerinde gösterimi	22
Şekil 9. Rastgele karar ormanları çalışma diyagramı.....	25
Şekil 10. KNN çalışma diyagramı.....	28
Şekil 11. Destek vektör makinesi sınıflandırma örneği. A, B destek vektörleri ve C maksimum kenar hiper düzlemi.	29
Şekil 12. Karışıklık matrisi.....	31

KISALTMALAR VE SİMGELER DİZİNİ

API	: Uygulama Programlama Arayüzü
CNN	: Evrişimli Sinir Ağı
CUDA	: Compute Unified Device Architecture
DEHB	: Dikkat Eksikliği ve Hiperaktivite Bozukluğu
DSM	: Diagnostic and Statistical Manual of Mental Disorders
DVM	: Destek Vektör Makineleri
EEG	: Elektroensefalografi
GPU	: Graphics Processing Unit
JSON	: JavaScript Object Notion
MR	: Manyetik Rezonans
SDK	: Yazılım Geliştirme Kiti

GİRİŞ

Yüz görüntülerinden insanların fiziksel ve ruhsal durumları analiz edilebilmektedir. İnsanların mutlu, üzüntülü, kızgın vb. duyguları yüzdeki ağız, burun ve gözlerin şekillerinden ve hareketlerinden çıkarılabilmektedir (Tarnowski et al., 2017). Yüz özniteliklerinden cinsiyet, yaş, ırk ve güzellik verilerinin sınıflandırılması (Zhao et al., 2018), suçlu tespiti ve takibi (Amato et al., 2018) gibi konular araştırmacıların oldukça ilgisini çekmektedir.

Yüz görüntüleri aynı zamanda uzun yıllardan beri tıp alanında teşhis ve tanı koyma işlemlerinde kullanılmaktadır (Kosilek et al., 2015). Gözetimli öğrenme algoritmaları aracılığıyla yüz fotoğraflarından anomali tespiti (Wang & Luo, 2016), cilt hastalıklarının tespiti (Wirdayanti et al., 2020), otizm spektrum bozukluğunun tespiti (Liu et al., 2016) yapılabilmektedir. Yüz görüntüleri çoğunlukla klinik ortamlarda uzmanlar tarafından çekilen video veya fotoğraflar üzerinden elde edilmektedir. Teknolojinin gelişmesiyle beraber mobil cihazların kameraları çözünürlüğü ve hassasiyeti de gelişmiştir. Bu sayede telefon üzerinden alınan veriler görüntü işleme konularında ve sınıflandırma işlemlerinde kullanılabilmektedir. Bu gelişmelerle birlikte mobil cihazlar tıp alanında da hastalık tespiti için önemli bir yere sahip olmuştur (Sim, 2019).

Bu tez kapsamında Dikkat Eksikliği ve Hiperaktivite Bozukluğu (DEHB) olan çocuk hastaların yüz görüntüleri incelenmiştir. DEHB vakası bulunan çocuklarda çeşitli psikolojik ve davranış bozuklukları gözlemlenmektedir. Örneğin; DEHB hastalığı bulunan çocuklar arkadaşlarıyla iletişim kurma güçlüğü çekerler, belirli bir işe veya göreve odaklanamazlar, genelde eğitim hayatlarında başarılı değildirler, düşünmeden davranırlar ve kurallara uymakta zorluk çekerler. Bu durumlar çocuğun sosyal hayatında problemlere yol açmaktadır. Hekimler DEHB tanılı hastaların erken yaşta tespit edilmesi, çocuğun sosyal hayata adapte olabilmesi için önemli olduğunu vurgulamaktadır. Erken tanının konulabilmesi için bilinen fiziki muayene bulgusu ya da laboratuvar testi bulunmamaktadır. Bu durumda hekimin hastasını uzun süren seanslarla takip etmesi ve gözlemlemesi gerekmektedir (Öner ve diğerleri, 2003). Hastalık teşhisinin koyulması için uzun süreli gözlemlere ve testlere ihtiyaç duyulmaktadır. Görüşmeler esnasında bozukluğun şiddeti belirlenir ve çocuğun ebeveynleriyle olan iletişimi gözlemlenmektedir. Tüm bunların dışında teşhis için kullanılan ölçekler vardır. Fakat bu ölçekler tanı koymak için tek başlarına yetersizdirler (Wilens & Spencer, 2010). Hekimin

durumu yanlış analiz etmesi ya da tecrübesizliği hastalığın tespitinin gecikmesine yol açmaktadır.

Hastalık tespitinin hızlandırılması amacıyla klinik tanı alanına görüntü işleme ve sinyal işleme yöntemleri dahil edilmiştir. Bu yöntemlerin kullanılmaya başlamasıyla beraber psikolojik bir hastalık olan DEHB'in tespiti için de çalışmalar yapılmıştır. Hastalığın tespiti için yapılan çalışmalar genel olarak çocuğa verilen görev sonrasında elektroensefalografi (EEG) sayesinde beyin dalgalarındaki değişimler izlenmesi, manyetik rezonans görüntülerinin karşılaştırılması ya da göz takip cihazı aracılığıyla dikkat eksikliğinin tespitine yöneliktir. Fakat EEG kullanımında, MR görüntülerin çekiminde ve göz takip cihazının kullanımında çocuğun fazla hareket etmemesi gerekmektedir. Çocuklarda hiperaktivite bozukluğu olabileceği için aksaklıklar yaşanabilir. Hastalığın tespiti için yapılan çalışmalar göz önüne alındığında tezin temel motivasyonu hekimlere DEHB tanısını koymasına yardımcı olabilecek bilgisayarlı görü ve yapay zekâ tabanlı bir sistem geliştirmektir.

Bu tez kapsamında geliştirilen sistem iki aşamalı olarak çalışmaktadır. İlk aşamada mobil uygulama aracılığıyla çocuğun fotoğrafı çekilmektedir. Çekilen fotoğraf üzerinde yüzdeki belirgin noktalar çıkarılarak çevrimiçi veri tabanına kaydedilmektedir. Elde edilen bu veriler ile eğitim veri seti de oluşturulmaktadır. İkinci aşamada bilgisayar ortamında kaydedilen noktaların birbirlerine olan uzunlukları ve elde edilen uzunlukların birbirlerine olan oranları hesaplanarak geometrik öznitelikler çıkarılmıştır. Daha sonra iki boyutlu olan görüntü üç boyutlu olarak tekrardan inşa edilmektedir. Üç boyutlu görüntü üzerindeki belirgin noktalar alınarak, birbirlerine göre uzunlukları hesaplanmaktadır. Bu sayede dört adet öznitelik elde edilmektedir. Öznitelik veri setleri üzerinde sınıflandırma algoritmaların başarımları test edilmektedir. Son olarak öznitelik seçim algoritmaları kullanılarak başarımlarının artırılması amaçlanmaktadır.

Tezin Amacı ve Katkısı

Tezin amacı makine öğrenmesi içinde yer alan gözetimli öğrenme algoritmaların hastalık tespitinde kullanılabileceği bir sistem geliştirmektir. Bu amaç doğrultusunda tezin katkısı aşağıda açıklanmaktadır.

1. Hekimlerin DEHB tanısını koymasına yardımcı olabilecek bilgisayarlı görü tabanlı mobil uygulama sistemi geliştirilmiştir. Geliştirilen sistem sayesinde hastalık teşhisi için çocuğun ön yüz fotoğrafının alınması yeterli olacaktır.

2. Uygulama aracılığıyla çekilmiş yüz görüntülerinden elde edilen noktalar üzerinde üç farklı sınıflandırma algoritması aracılığıyla sınıflandırma işlemi gerçekleştirilmiştir.
3. Sınıflandırma işleminin doğruluğunu artırmak için öznitelik seçim algoritmaları kullanılmıştır.
4. Yapılan çalışmada sınıflandırma algoritmalarının başarımları analiz edilmiştir.
5. DEHB teşhisine sahip çocukların yüz görüntülerinden veri seti oluşturulmuştur. Yüz noktalarının çıkarılması sonucunda uzunluk ve oranlar alınarak 4 farklı öznitelik elde edilmiştir.



KURAMSAL TEMELLER

Literatürde ilk olarak dikkat eksikliği ve hiperaktivite bozukluğu hastalığının semptomları ve tanı yöntemleri incelenmiştir. Hastalık hakkında fikir sahibi olduktan sonra mobil uygulamalarla ilgili literatür taraması yapılmıştır. Hastalığın yüz görüntülerinden tespiti amaçlandığı için yüz algılama ve öznitelik çıkarma yöntemleri incelenmiştir. Son olarak makine öğrenmesi ve sınıflandırma yöntemleri incelenmiştir.

Dikkat Eksikliği ve Hiperaktivite bozukluğunun tarihi hakkında yapılan çalışmalar incelendiği zaman bu hastalığa en yakın tanım ilk olarak 1775 yılında Alman Doktor Melchior Adam Weikard ait bir kitapta yapılmıştır. Kitabın bir bölümde dikkat hacmi anlamına gelen *Attentio Volubilis* başlığı kullanılmıştır. İlgili bölümde dikkatli bir insanın belli bir konuda daha fazla konsantre olması gerektiği ve diğer kişilerden daha ısrarlı olması gerektiği konusuna değinilmektedir. Kitapta verilen bir örneğe göre genç bir papazın meditasyon yaparken gelen sinek sesi ya da gölgeler gibi faktörlerden dikkati dağılmaktadır (Barkley & Peters, 2012). Bu örneğin dikkat eksikliği için bir tanımlama olduğu varsayılmaktadır.

Dikkat eksikliği tanımının yapılmasından sonra 1798 yılında İskoç bilim insanı Sir Alexander Crichton akıl hastalıkları üzerinde çalışmalar yapmıştır. Yaptığı çalışmalardan biri de dikkate bağlı hastalıklar üzerinedir. Crichton'a göre dikkate bağlı hastalıklar iki başlık altında toplanabilmektedir. İlki herhangi bir nesneye olan ilgisizlik, ikinci ise beyin üzerindeki faaliyetlerin tamamen askıya alınmasıdır. Crichton'a göre bu tip hastalıklar kişinin doğumuyla birlikte gelmektedir ya da ilerleyen zamanlarda başka hastalıkların etkisiyle ortaya çıkmaktadır ve erken yaşta ortaya çıkan bu hastalık engellenmezse çocuğun eğitimini engelleyebilmektedir (Crichton, 2008).

1844 yılında ise Doktor Heinrich Hoffmann, Fidgety Phil adında resimli bir hikâye yayınlamıştır. Hikâyede akşam yemeği için sofraya oturmuş bir aileden bahsedilmektedir. Hikâyenin başlangıcında Philip'in babasının "Philip biraz beyefendi olabilir mi bir bakayım; Bakalım masada bir kez olsun kıpırdamadan oturabilecek mi?" ifadesi yer almaktadır (Hoffman, 2012). Bu hikâyede DEHB hastalığı tam anlamıyla tanımlanmamış olsa da babasının Philip'in aşırı hareketli olmasından rahatsız olması ve Philip'in masada uslu duramayıp yere düşerken sofrayı devirmesi hiperaktivite bozukluğuna işaret etmektedir. Bu hikâye günümüzde DEHB için alegori olarak kullanılmaktadır. Hoffmann'ın yazmış olduğu bir diğer hikâyede ise Johnny isimli bir çocuğun dikkatsizliği anlatılmaktadır. Hikâyeye göre

Johnny'in dikkati kolayca dağılmaktadır, hatta hikâyeye göre bir köpekle çarpışarak, nehre düşmektedir (Lange et al., 2010).

1902 yılında Sir George Frederick Still, ahlaki karakter bozukluğu kavramından bahsetmiştir. İncelemiş olduğu çocuklardan bir tanesinin farkında olmadan birçok kez iyi geceler dediğini, diğer bir çocuğun genellikle farkında olmadan ayakkabılarını ters giydiğini, diğer bir çocuğun ise oyun oynarken bile uzun süre boyunca dikkatini veremediğini gözlemlemiştir. Frederick'e göre incelemiş olduğu çocuklar her ne kadar normal ve akıllı çocuklar gibi gözükseler de bu durum eğitimin hayatlarını ve notlarını da etkilemiştir (Still, 2006).

Dikkat eksikliğiyle ilgili yapılan çalışmalarından ardından 1932 yılında Franz Kramer ve Hans Pollnow 42 vakada konsantrasyon eksikliği, artan dikkat dağınıklığı, amaçsızca dolaşma ve dolaşım sırasında etrafta bulunan nesnelere dokunma gibi temel hareket bozuklukları gözlemlenmiştir (Neumärker, 2005). Bu durum literatürde hiperkinetik dürtü bozukluğu olarak tanımlanmıştır. 1957 yapılan çalışmada hiperkinetik dürtü bozukluğu olan çocuklarda beyin hasarı değil de beyinde işlevsizlik olabileceği düşüncesi ele alınmıştır. Bu sebepten hastalığın minimum beyin hasarı olarak adlandırılması gerektiği savunulmuştur (Laufer et al., 1957). 1962 senesinde Oxford Uluslararası Çocuk Nörolojisi Çalışma Grubu düzenlemiş olduğu konferansta beyin hasarının sadece sorunlu davranışlarla alakalı olmadığı ve minimal beyin disfonksiyonun daha doğru bir terim olduğu savunulmuştur (Lange et al., 2010). 1968 yılında Amerikan Psikiyatri Birliği yayınlamış olduğu kitapçıkta minimal beyin disfonksiyonun yüzeysel bir terim olduğunu ve hastalığın çocukluğun hiperkinetik reaksiyonu (DSM II) olarak tanımlanması gerektiği söylemiştir. Kitapçıkta ayrıca bu hastalığa sahip çocukların aşırı hareketli, dikkatlerinin dağınık, huzursuz olduğu vurgulanmıştır (Mannuzza et al., 1998).

1980 yılında ise bu hastalığa DSM III etiketi verilmiştir ve dikkat eksikliği bozukluğu olarak isimlendirilmiştir. DSM III hastalığının iki çeşidinin olduğu savunulmuştur; hiperaktivite olan ve hiperaktivite olmayan. 1987 yılında ise hastalık DSM IIIR olarak etiketlenmiştir ve dikkat eksikliği ve hiperaktivite bozukluğu olarak adlandırılmıştır. Sonrasında yapılan 6 yıllık bir çalışmada tanı ölçütlerinde ve hastalığın açıklamasında birçok değişiklik yapılmıştır. Böylelikle 1994 yılında hastalık için DSM IV etiketini kullanılmaya başlanmıştır. 2000 yılında kurulan çalışma grubunun yapmış olduğu çalışmalarla birlikte DSM V etiketi ortaya çıkmıştır (Anonymous, DSM History, 2013).

Günümüzde yapılan çalışmalar incelendiği zaman dikkat eksikliği ve hiperaktivite bozukluğu genellikle çocukları etkileyen psikolojik bir hastalık olduğu gözlemlenmiştir. DEHB

hastalığı ayrıca yetişkinleri de etkilemektedir. Hastalık çocuklarda %8 oranında, ergenlerde %6 oranında ve yetişkinlerde ise %4 oranında görülmektedir (Mercan ve diğerleri, 2016). DEHB hastalığı erkeklerde kadınlara göre üç kata daha fazla gözlemlenmektedir (Palmer & Finger, 2001). DEHB hastalığının başlıca belirtileri arasında dikkatsizlik, odaklanamama, aşırı hareket ve düşünmeden ortaya çıkan aceleci davranışlar bulunmaktadır. DEHB'in sebebi kesin olarak bilinmese de genetik faktörlerden ve dopamin salgısındaki dengesizliklerin DEHB'e neden olabileceği düşünülmektedir. (Biederman & Faraone, 2005). DEHB'e neden olan faktörlerin incelendiği bir çalışmada ise erken doğan bebeklerin DEHB hastalığına yakalanma oranlarının 2,6 ile 4 kat daha fazla olduğunu belirtilmiştir. Bunun sebebi kortikal genişlemenin azalması olarak gösterilmektedir. DEHB'e neden olan bir diğer faktör ise çocuklarda görülen yüksek kan konsantrasyonu olarak belirtilmektedir (Núñez-Jaramillo et al., 2021).

Dikkat eksikliği ve hiperaktivite bozukluğu hastalığının tespiti için dikkat edilmesi gereken semptomlar dikkatsizlik, hiperaktivite ve dürtüsellik olarak iki başlık altında incelenmektedir (Anonymous, Symptoms and Diagnosis of ADHD, 2021).

Dikkatsizlik: 16 yaşına kadar olan çocuklarda aşağıdaki belirtilerden en az altı tanesi gözlenmelidir.

- Genellikle ayrıntılara dikkat etmezler ve günlük hayatlarında sıklıkla hatalar yaparlar.
- Genellikle yaptıkları işlerde ya da oyunlara dikkat etmekte güçlük yaşarlar.
- Kendileriyle konuşulduğu zaman dinlemiyormuş gibi görünürler.
- Çoğu zaman yönergeleri takip etmezler. Okulda verilen ödevleri, iş yerinde verilen görevleri yapmazlar.
- Genellikle görevleri ve etkinlikleri organize etmekte güçlük çeker.
- Uzun süre zihinsel çaba göstermeleri gereken aktiviteleri sevmezler ve bunlardan kaçınırlar.
- Genellikle eşyalarını kaybederler.
- Genellikle kolayca dikkatleri dağılır.
- Sıklıkla unutkanlık yaşarlar.

Hiperaktivite ve Dürtüsellik: 16 yaşına kadar olan çocuklarda aşağıdaki belirtilerden en az altı tanesi gözlenmelidir.

- Genellikle oturdukları zaman kıpır kıpır olurlar ve ellerine veya ayaklarına dokunurlar.
- Çoğu zaman oturmaları gereken ortamlarda oturmaktan kaçınırlar.

- Genellikle uygun olmayan ortamlarda koştururlar veya tırmanırlar.
- Çoğu zaman oyun oynarken sesler çıkarırlar ve sessiz olması gereken aktivitelerden kaçınırlar.
- Genellikle aşırı konuşurlar.
- Genellikle kendilerine bir soru sorulduğu zaman, soru bitmeden cevap verme eğilimi gösterirler.
- Çoğu zaman sıralarını beklemekte zorlanırlar.
- Sıklıkla konuşma esnasında veya oyun oynarken araya girerler.

DEHB teşhisi için yapılan çalışmalar temel olarak iki alanda incelenebilir. Tıbbi alanda yapılan çalışmalarda genellikle uzun süreli gözlemler ve sorulara dayanmaktadır. DEHB tanısı koyabilmek için klinik ortamında kullanılan ölçekler bulunmaktadır. Aile ve hastayla görüşme esnasında semptomlar, bozukluğun şiddeti ve aile öyküsü belirlenir. Pediatrik görüşmelerle çocuğun davranışları ve ebeveyniyle olan etkileşimi gözlemlenir. Bunun dışında Vanderbilt Ölçeği, Conners Ölçeği, DEHB Derecelendirme Ölçeği (DEHB-RS-V), Swanson, Nolan ve Pelham (SNAP) Ölçeği gibi ölçekler tanı koymak için kullanılmaktadır. Fakat bu ölçütlerin tanı koymak için tek başına yeterli olmadığı belirtilmektedir (Wilens & Spencer, 2010). Bu sebeple araştırmacılar hastalığın tespiti için bilgisayar yöntemleri kullanmaya başlamıştır. Örnek olarak yapılan bir çalışmada beyin dalgalarındaki değişikliklerin DEHB hastalığıyla bağlantısının incelenmesi amacıyla çocuklara belli başlı görevler verilmiştir. Görevler gerçekleştirilirken EEG aracılığıyla beyin dalgaları incelenmiştir. Beyin dalgasına ek olarak çekilen MR görüntüleri aracılığıyla beyindeki değişiklikler saptanmaya çalışılmıştır (Sridhar et al., 2017).

Yapılan bir diğer çalışmada ise ADHD200 veri seti kullanılarak 3 boyutlu evrişimli sinir ağı eğitilmiştir. MR görüntülerinden çıkarılan öznetelikleri sınıflandırma amacıyla destek vektör makineleri, k en yakın komşu ve doğrusal diskriminant analizi kullanılmıştır. Yapılan çalışma sonucunda doğrusal diskriminant analiziyle %74,93'lük bir başarı oranı elde edilmiştir (Sahar & Saniee Abadeh, 2020).

Chen ve ekibi makine öğrenmesi yöntemleriyle DEHB hastalığını tespit etmeyi amaçlamıştır. Bu amaç doğrultusunda 69 katılımcıya sorular sorulmuştur. Sorulardan elde edilen cevaplar sınıflandırma işlemlerinde kullanılmıştır. Çalışma kapsamında karar ağaçları algoritmasıyla %82,6'lık başarı elde edilmiştir (Chen et al., 2021).

Bir diğer çalışmada ise makine öğrenmesi yöntemleri, Minnesota Çok Yönlü Kişilik Envanteri (MMPI) testi üzerinde kullanılarak DEHB hastalığının tespiti amaçlanmıştır. Bu amaç doğrultusunda yapılan MMPI testine 5726 öğrenci katılmıştır. Testten elde edilen

sonular sınıflandırma algoritmalarında kullanılmıştır. Sınıflandırma sonucunda en yüksek başarıyı %93,1 ile KNN algoritması göstermiştir (Kim et al., 2021).

DEHB hastalığının MR görüntülerinden tespitinin amaçlandığı bir diğer çalışmada toplama tabanlı sinerjik yapay sinir ağı (SSANN) oluşturulmuştur. SSANN iki adet üç boyutlu evrişimli sinir ağı içermektedir. Sinir ağlarından bir tanesi fonksiyonel MR görüntülerinden öznitelik çıkarırken, diğeri yapısal MR görüntülerinden öznitelik çıkarmaktadır. Elde edilen veriler birleştirildikten sonra sınıflandırma işlemi gerçekleştirilmektedir. Sınıflandırma işlemi sonucunda ADHD-200 veri seti üzerinde %72,89'luk başarı elde edilmiştir (Peng et al., 2021).

Tıp alanında hastalık teşhisi için beyin dalgası sinyalleri ve MR görüntülerinin yanında teşhis ve tedavi sürecinin etkin ve verimli yürütülebilmesi amacıyla çeşitli mobil uygulamalar ve görüntü işleme teknikleri de kullanılmaktadır. Örnek olarak gerçek zamanlı yüz tespitinin yapıldığı bir çalışmada otizm spektrum bozukluğu olan çocuklarda öğrenme güçlüklerinin önüne geçmek amacıyla “World of Kids” isimli mobil oyun serisi oluşturulmuştur. Çocuğun öğrenirken rahat etmesi için Viola-Jones algoritması aracılığıyla gerçek zamanlı yüz ifadesi tespit edilerek en uygun oyun önerilmiştir. Çocuğun öğrenme aşaması kolaylaştırmak için uygulama aylık analizler sunmaktadır. Bu sayede veliler veya eğiticiler istedikleri zaman öğrenme kavramlarını oyuna dahil etmektedir (Heni & Hamam, 2016).

Tıp alanında yapılan bir diğer çalışmada ise yüz noktaları kullanılarak koroner kalp hastalığının tespiti amaçlanmıştır. Yapılan çalışmada ön yüzde 68 nokta belirlenerek 6 alan seçilmiştir. Sağ ve sol kulağın tespiti için önceden eğitilmiş Faster-RCNN ağı kullanılmaktadır. Sonrasında karar ağaçları ve rastgele karar ormanları sınıflandırma teknikleri kullanılmaktadır. Yapılan çalışmada 226 erkek 83 kadının bulunduğu bir veri seti kullanılmaktadır. Bu veri setinde 309 kişinin 195'i koroner kalp hastalığına sahiptir. Derinlik 3'ten 6'ya kadar olduğunda rastgele karar ormanları için %72,73, karar ağaçları için de %70,45 başarı oranı elde edilmiştir (Lin et al., 2019).

Yapılan bir diğer çalışmada ise yüz tanıma yöntemleri ile psikolojik bozuklukların saptanması amaçlanmıştır. Çalışmada evrişimli sinir ağları tabanlı LeNet mimarisi kullanılmıştır. Yüz tanıma veri seti olarak 213 görüntü ve 7 yüz ifadesinden oluşan JAFFE veri seti, 4900 görüntü ve 7 yüz ifadesinden oluşan KDEF veri seti ve 140 görüntü 7 yüz ifadesinden oluşan özel bir veri seti kullanılmıştır. Veri setlerinde görüntü işleme yöntemi olarak Haar Cascade kütüphanesi kullanılmıştır. Çalışmada 7 farklı yüz ifadesinin sınıflandırılması için %96,43 doğruluk ve %91,81 başarı oranları elde edilmiştir (Ozdemir et al., 2019).

Ek olarak farklı alanlarda yapılan çeşitli uygulamalarda yüz tespiti ele alınmaktadır. Örnek olarak yüz görüntülerine ait simetri notunun hesaplanarak güzellik açısından ilgili sınıfa

atama amacıyla geliştirilen bir mobil uygulamada OpenCV kütüphanesi aracılığıyla yüz üzerindeki noktalar tespit edilmiştir. Daha sonra simetri notunun hesaplanması için bir formül kullanılmıştır. Doğruluğunu test etmek için 5 farklı senaryo uygulanmıştır. İlk olarak hayvan fotoğrafları denenmiştir. Maymunun yüz yapısı insana benzediği için uygulamanın çalıştığı gözlemlenmiştir. Sonra birden fazla yüzün olduğu fotoğraflar denenmiştir ve uygulama sadece bir yüzü tespit edebilmiştir. Üçüncü olarak Zhao ve ekibi tarafından güzel kabul edilen 5 insan yüzü denenmiştir ve ortalama simetri notu olarak 99 bulunmuştur. Dördüncü olarak yine ekip tarafından ortalama kabul edilen 5 yüz denenmiştir ve ortalama simetri notu olarak 95 bulunmuştur. Son olarak çekici olmayan 5 yüzü denenmiştir ve ortalama simetri notu olarak 93 bulunmuştur. Sonuç olarak yüzdeki simetri notuna göre yüzleri güzel, çekici değil şeklinde ayırmanın bu yöntemle mümkün olmadığını söylenmiştir (Zhao et al., 2018).

Yapılan bir diğer çalışmada gerçek zamanlı yüz tanıma ile suçlu tespit etme, suçlu takibi ve fabrika ziyaretçilerinin takibi gibi birçok alanda yarar sağlanması amaçlanmıştır. Bu işlem için düzenlenmiş tek atış dedektörü (SSD), yüz işaretleri ve MobileNet V2 modeli kullanılmıştır. Yapılan çalışma için SSD mimarisi olarak SqueezeNet ve yüz işareti olarak 2 göz, burun ve 2 dudak kenarı olmak üzere 5 adet işaret seçilmiştir. Görüntü boyutu 250x250 olan, 13000 görüntüden ve toplam 5749 kimlikten oluşan bir model kullanılmıştır. Kullanılan sistem birden fazla yüzde %99,01 tanımlama başarısı göstermiştir (Teoh et al., 2021).

Yapılan benzer bir çalışmada da 2 boyutlu karikatürden yüz tanımlama ile 3 boyutlu modellerin çıkarılması amaçlanmıştır. Bu işlem için 2 boyutlu karikatürler ve bu karikatürlere karşılık gelen 3 boyutlu modellerinden bir veri seti oluşturulmuştur. Sonrasında 3 boyutlu karikatürlerde oluşan bozulmalar için köşe tabanlı bir parametrik model oluşturulmuştur. 2 boyutlu karikatürleri modellemek için 8000 karikatürden ve ağız, burun gibi 68 farklı işaret, işaret pozisyonları için Dlib kütüphanesi kullanılmıştır ve optimizasyon temelli 3 boyutlu bir karikatür veri seti oluşturulmuştur. 3 boyutlu karikatürleri iyileştirmek konusunda oluşan bozulmalar için kodlayıcı-kod çözücü yapı iskeleti kullanılmıştır. Yapılan kapsamlı deneyler yöntemin çeşitli karikatürlerde iyi çalıştığını göstermiştir (Cai et al., 2021).

Yüz üzerindeki işaretlerin daha hızlı tespit edilmesinin amaçlandığı başka bir çalışmada YOLOv3 kullanılmıştır. Hızı ve doğruluğu artırmak için EfficientNet-B0 mimarisi tercih edilmiştir. Veri seti olarak 122450 veriden oluşan ve 68 yüz noktasının işaretlendiği 300W-LP veri seti kullanılmıştır. Modelin ortalama normalleştirme hatası 2.32 pikseldir. Çalışmada ortalama yüz algılama hassasiyeti olarak %99 elde edilmiştir. Çalışmayı doğrulamak için 64 yerine 29 yüz işareti kullanan ve normalleştirme hatası 2,56 piksel olan COFW veri tabanı kullanılmıştır. Bu çalışmanın vaat edilen performansı verdiğini doğrulamıştır (Kim et al., 2021).

Yüz ve yüz noktalarının tespitinden elde edilen verilerin işlenmesi, incelenmesi ve anlamlandırılması da önem arz etmektedir. Bu sebeple makine öğrenmesinden faydalanılmaktadır. Makine öğrenmesi yapay zekanın alt kümesidir (Sharma et al., 2021). Veri ve algoritmaları kullanarak insanların öğrenme biçimini taklit etmektedir. İstenilen bir görevi gerçekleştirmek için algoritmalar ve belirli modeller kullanılmaktadır. İnternette arama yapmak için kullanılan arama motorlarında makine öğrenmesi kullanılmaktadır. Bu sayede aranan veriye karşılık gelen sayfaları alakalı bir şekilde sıralayabilmektedir (Mahesh, 2020). Makine öğrenmesi ayrıca görüntü işleme, tahmine dayalı analitik gibi işlemlerde de kullanılmaktadır. Temelde dört öğrenim yaklaşımına sahiptir (Sarker, 2021).

Öğrenim yaklaşımları aşağıdaki gibi sıralanabilmektedir;

- Gözetimli Öğrenme
- Gözetimsiz Öğrenme
- Pekiştirmeli Öğrenme
- Yarı Gözetimli Öğrenme

Bu çalışmada gözetimli öğrenme yaklaşımı kullanılmaktadır. Gözetimli öğrenme yaklaşımında algoritma eğitim verileri üzerinde iteratif olarak tahminlerde bulunmaktadır ve verilen işaretlerle karşılaştırmaktadır. Bu işlem doğru sonuç tahmin edilene kadar devam etmektedir. Gözetimli öğrenme sınıflandırma ve regresyon olarak iki sınıfa ayrılmaktadır. Sınıflandırma işleminde girdi olarak verilen veriler işaretlerine göre sınıflara ayrılmaktadır. Regresyonda ise çıktı değişkeninin maaş veya ağırlık gibi gerçek veya sürekli bir değer olması durumudur (Brownlee, Supervised and Unsupervised Machine Learning Algorithms, 2016).

Sınıflandırma işlemi hakkında yapılan bir çalışmada makinelerin de insanlar gibi nesneleri, renkleri, kokuları ayırt etmesini ve farklı özelliklere göre nesneleri sınıflandırılması üzerine çalışılmasından bahsedilmiştir. 2000 yılından sonraki çalışmaların incelendiği çalışmada sınıflandırma ve regresyon için en iyi makine öğrenmesi destek vektör makineleri olduğu vurgulanmıştır (Chandra & Bedi, 2018).

Yapılan bir diğer çalışmada ise yeni bir K en yakın komşu sınıflandırma yönteminin performansının artırılması amaçlanmıştır. Bu kapsamda caKD+ algoritması geliştirilerek kümeleme ve k en yakın komşu algoritmalarıyla beraber kullanılmıştır. Yapılan araştırmalar ve gözlemler sonucunda ortaya çıkan sonuçların klasik k en yakın komşu yöntemine göre daha az kaynak kullanmaktadır ve daha hızlı çalışmaktadır (Gallego et al., 2022).

Diğer çalışmalardan farklı olarak domates yaprağı hastalıklarının tespit edilmesi için bir mobil uygulama geliştirilmesi amaçlanmıştır. Derin evrişimli sinir ağlarının (CNN) hesaplama

yükünün fazla olması sebebiyle cep telefonunda çalışabilen CNN tabanlı bir model önerilmiştir. Model geliştirilirken MobileNet CNN'den esinlenilmiştir. Uygulamada 7176 tane domates yaprağı görüntüsünden oluşan bir veri seti kullanılmıştır ve makine öğrenmesi bu veri seti aracılığıyla sağlanmıştır. Geliştirilen uygulama ile 10 domates yaprağı hastalığının tanımlanabildiği vurgulanmıştır (Elhassouny & Smarandache, 2019).

Amato ve diğerleri tarafından yapılan çalışmada yüz doğrulamanın güvenlik sistemlerde kullanılan önemli bir yöntem olduğundan bahsedilmiştir. İki yüzün aynı kişiye ait olmasıyla ilgili adli tıpta yüz noktaları arasındaki kritik noktalar kullanılmaktadır. Amato ve diğerleri yüz üzerindeki noktaların yüz doğrulama için yetersiz olduğunu düşünmektedir. Bu sebeple yüz doğruluğunun tespiti için yüz noktaları yerine yüz görüntüsünde derin öğrenme metotların kullanılmasını önermektedir. Yapılan çalışmalar doğrultusunda yüz doğrulamada yüz görüntüsünün doğrudan kullanımının diğer yöntemlere göre yüksek başarımlı gösterdiği vurgulanmıştır (Amato et al., 2018).

Yukarıda bahsedilen sınıflandırma yöntemleri aynı zamanda hastalıkların erken teşhisi için de kullanılmaktadır. Örnek olarak yapılan bir çalışmada termal kamerayla çekilmiş göğüs görüntüleri işlenerek kanser belirtileri aranmıştır. Bunun için görüntüye filtreler uygulayıp son görüntü oluşturulmuştur. Daha sonra DMI veri seti kullanılarak eğitilen evrişimli sinir ağıyla öznetelik çıkarımı yapılmıştır. Son olarak sınıflandırma işlemi yapılarak, 140 görüntü üzerinden %98,95 başarı elde edilmiştir (Ekici & Jawzal, 2020).

Hastalık teşhisi için yapılan bir diğer çalışmada beyin tümörünün tespiti için yeni bir CNN olan BrainMRNet önerilmiştir. 155 tümörlü, 98 normal olmak üzere toplam 253 MR görüntüsünden oluşan bir veri seti üzerinde çalışılmıştır. Görüntüler BrainMTNet içinde işlendikten sonra evrişimli sinir ağıları kullanılmıştır. Yapılan sınıflandırma sonucu %96,05 başarı elde edilmiştir (Toğaçar et al., 2020).

Yapılan bir diğer çalışmada alzheimer hastalığının günümüzde yaşlılarda çok sık rastlanan bir hastalık olduğundan bahsedilmiştir. Yapılan araştırmalarda erken dönemde belirlenmesinin alzheimer hastalarının yaşam kalitesini artırdığını ve ilerlemenin yavaşlatılmasında önemli bir etken olduğu gözlemlenmiştir. Yapay zekâ algoritmaları olan destek vektör makineleri, rastgele karar ormanları, evrişimli sinir ağıları ve k-means kümeleme yöntemi kullanılarak alzheimer hastalığının erken yaşta tanımlanması amaçlanmıştır. MRI ve EKG başta olmak üzere insanların biyomedikal verilerini kullanarak yapılan testlerde alzheimer hastalığının erken dönemde tespitinde %70 üstü başarılı sonuçlar alınmıştır (Mirzaei & Adeli, 2022).

Son olarak yapılan bir alıřmada dzensiz solunum hastalıęının tespiti amalanmıřtır. Bu ama doęrultusunda 19 hastadan alınan basınlı akıř, EKG, basınlı horlama, SpO₂ verileri yapay sinir aęları, destek vektr makineleri, rastgele karar ormanları, naive bayes, k en yakın komřu, karar aęları ve lojistik regresyon yntemleri kullanılmıřtır. Elde edilen verilerin sınıflandırılması sonucunda uykuda dzensiz solunum hastalıęının tespitinde %76,3 oranla doęru sonu elde edilmiřtir. Hipopne verisi ele alınmadan yapılan hesaplarda ise doęruluk oranı olarak %86,6 elde edilmiřtir (Balci et al., 2022).



MATERYAL VE METOT

Bu tez çalışmasında, dikkat eksikliği ve hiperaktivite bozukluğunda tanı süresinin kısaltılması ve makine öğrenmesi yöntemleriyle geliştirilecek çalışmaları desteklemek için veri seti oluşturulmasına yardımcı olacak bir mobil uygulama geliştirilmiştir. Bu uygulama aracılığıyla çekilen çocuk fotoğraflarından yüz üzerindeki noktalar tespit edilerek noktaların koordinatları çevrimiçi veri tabanına kayıt edilmiştir. Yüz üzerinde elde edilen bu noktalardan, 3 farklı geometrik öznitelik oluşturulmuştur. Derinlik bilgisinin incelenebilmesi için, yüz görüntüleri bilgisayar ortamında 3 boyutlu olarak yeniden inşa edilerek yüz noktaları tespit edilmiştir. Bu noktalar kullanılarak 4. geometrik öznitelik oluşturulmuştur. Elde edilen öznitelikler üzerinde sınıflandırma işlemleri gerçekleştirilmiştir. Öznitelik verilerinde bulunan gürültüleri engellemek adına öznitelik seçimi algoritmaları kullanılarak sütun sayısı azaltılmıştır. Elde edilen verilerle öznitelik çıkarma algoritmalarının sınıflandırma işlemi üzerindeki etkisini görmek için sınıflandırma işlemi tekrardan gerçekleştirilmiştir. Şekil 1’de sistemin çalışma diyagramı gösterilmiştir.

Sistemin Geliştirilmesi

Yüz görüntülerinin toplanabilmesi ve yüz üzerindeki noktaların tespit edilebilmesi için Android Studio ortamında bir mobil uygulama geliştirilmiştir. Şekil 2’de uygulamanın çalışma diyagramı gösterilmiştir.

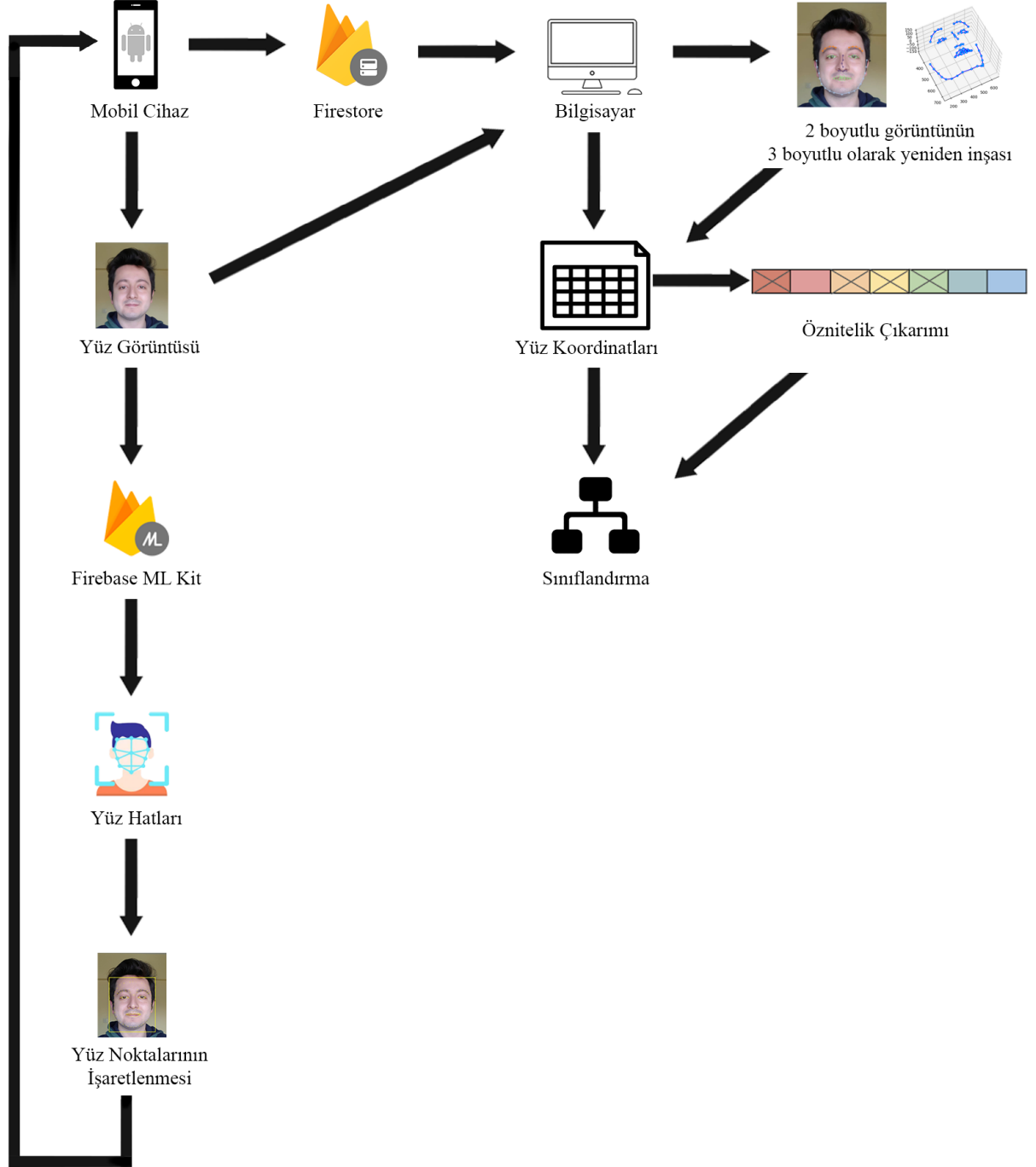
Yüz üzerindeki noktaların tespiti için Firebase ML Kit kullanılmıştır. Firebase ML Kit mobil uygulamalar için Google tarafından tasarlanmış bir makine öğrenimi yazılım geliştirme kitidir. Firebase ML kit SDK’sı içinde Cloud Vision API, TensorFlow Lite ve Neural Network API’yi bulundurmaktadır. Bu sayede mobil uygulamalarda makine öğrenmesi algoritmalarının çalıştırılması sağlamaktadır.

Firebase ML kit içerisinde önceden eğitilmiş öğrenme modelleri sunmaktadır. Bu sayede yüz algılama ve yüz üzerinde nokta tespiti yapılabilmektedir.

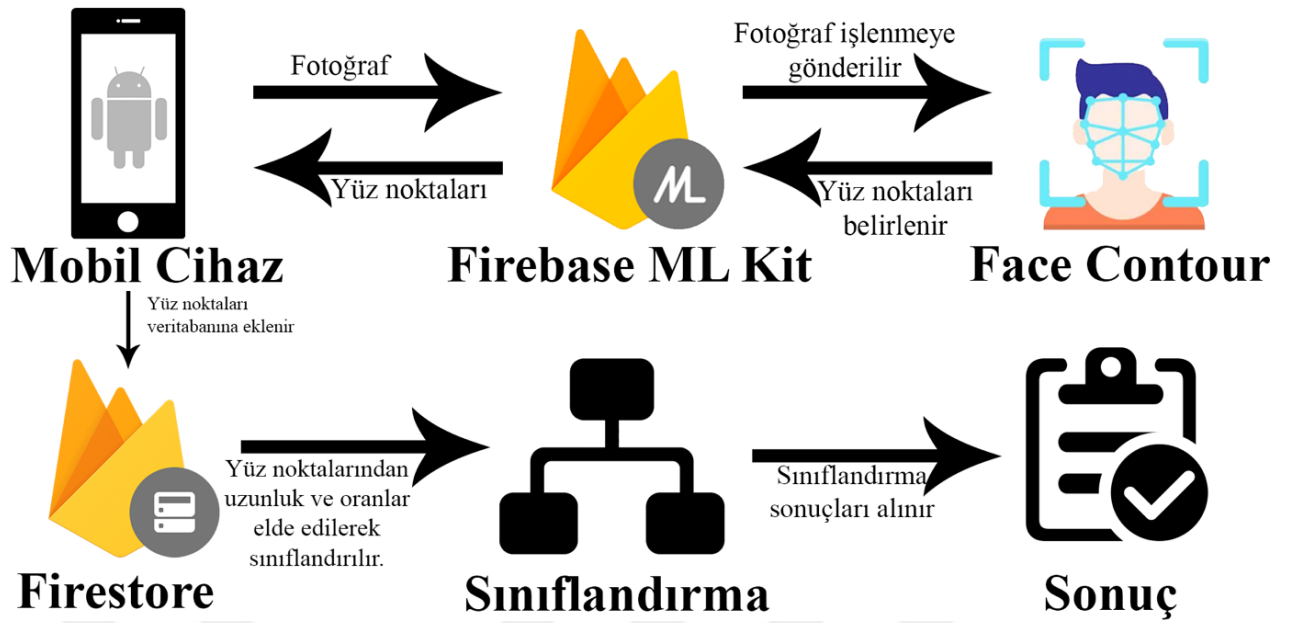
İçerisinde bulunan Cloud Vision API sayesinde hem bulut üzerinde hem de cihaz üzerinde işlem yapabilmektedir. API ağ bağlantısından bağımsız olarak çalışabilmektedir (Singh & Bhadani, 2020). Bu sayede yapılan işlemler internet bağlantısı sağlanana kadar telefon üzerinde muhafaza edilebilmektedir.

TensorFlow Lite ise mobil cihaz üzerinde makine öğrenimi için optimize edilmiştir. Sunucuyla cihaz arasında gecikme yaşanmamaktadır, veriler şifreli olarak taşındığı için

güvenlidir, gerekli durumlarda çevrimdışı olarak çalıştığı için internet bağlantısına her zaman ihtiyaç duymamaktadır, telefonda az yer kapladığı ve güç tüketimi az olduğu için çoğu marka telefonda çalışabilmektedir (Anonymous, TensorFlow Lite, 2021).



Şekil 1. Sistemin çalışma diyagramı



Şekil 2. Mobil uygulama sistem diyagramı

Veri Seti

Çalışma için kullanılan veri seti internet ortamından elde edilmiş yüz görüntüleri içermektedir. Yüz görüntüleri DEHB'li çocuklar ve kontrol grubu için yüzün ön tarafı görünecek şekilde alınmıştır. 2 boyutlu görüntüden 3 boyutlu yeniden inşa işlemi işlemci gücü kullanıldığı için alınan görüntülerde dosya boyutunun 1mb'dan büyük olduğu fotoğraflarda, dosya boyutu düşürülmüştür. Böylece dönüştürme işlem süresi kısaltılmıştır. Bu veri setinde Tablo 1'de görüldüğü üzere 40 tanesi erkek 11 tanesi kız olmak üzere toplamda 51 tane DEHB teşhisi konulmuş çocuk bulunmaktadır. Ayrıca 25 tanesi erkek 15 tanesi kız olmak üzere toplamda 40 tane de kontrol verisi bulunmaktadır.

Öznitelik Çıkarımı

Çekilen 2 boyutlu yüz görüntülerinden öncelikle yüz noktalarının tespiti sağlanmıştır. Sonrasında 2 boyutlu görüntüler 3 boyutlu olarak yeniden inşa edilerek, yüz noktalarının tespiti sağlanmıştır. Yapılan işlemler ayrıntılarıyla aşağıda anlatılmıştır.

Tablo 1. Çalışmada Kullanılan Veri Seti

	Erkek	Kız	Toplam
DEHB	40	11	51
Sağlıklı	25	15	40

İki boyutlu görüntülerden yüz noktalarının tespiti

İki boyutlu görüntülerden yüz noktalarının tespiti için FaceNet modeli kullanılmıştır. FaceNet modelinin amacı yüz üzerindeki önemli noktaların hatlarını belirlemektir. FaceNet modeli öznitelik çıkarımı için öncelikli olarak görüntü üzerinde yüz bölgesini tespit ederek yüz görüntüsünü arka plandan ayırmaktadır ve yüz noktalarının tespiti için evrişimli sinir ağlarını (CNN) kullanmaktadır (Nyein & Oo, 2019). Evrişimli sinir ağları aracılığıyla alınan yüz görüntüsü üç boyutlu matrislere çevrilmektedir. Daha sonra giriş matrisinin uzaysal boyutu küçültülmektedir. Bu sayede görüntü küçük parçalara ayrılarak derin öğrenme algoritmaları aracılığıyla sınıflandırılmaktadır. Sınıflandırma işlemi için matris çarpımı kullanılmaktadır.

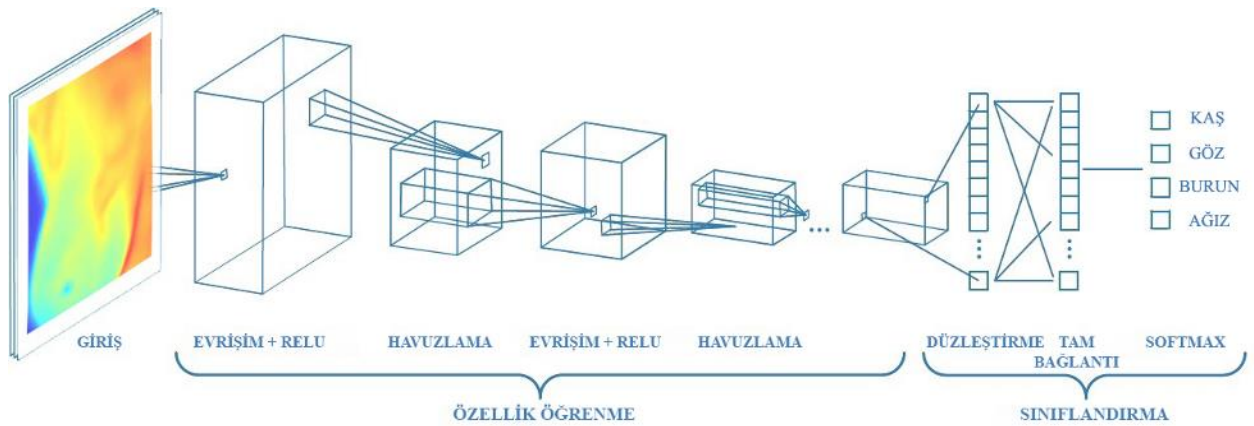
CNN eğitime rastgele ağırlıklarla başlamaktadır. Eğitim sırasında CNN'e işaretli bir veri seti verilmektedir. CNN, her görüntüyü rastgele işleyerek ve kendi bulduğu çıktıyla verilen görüntünün etiketini karşılaştırmaktadır. Çıktı ve etiketinin eşleşmemesi durumunda doğru sonucu bulana kadar ağırlıkları değiştirerek işlemleri tekrar yapmaktadır. CNN beş farklı katmandan oluşmaktadır ve bu katmanların işleyişini gösteren diyagram Şekil 3'de verilmiştir.

Giriş katmanı

Giriş katmanı yüz görüntü verilerini içermektedir. Görüntü verileri üç boyutlu matrisler ile temsil edilmektedir. Bu matrisin tek sütuna şekillendirilmesi gerekmektedir.

Evrişim katmanı

Görüntünün özelliklerinin ayıklandığı katmandır. Özellik çıkarıcı katman olarak da bilinmektedir. Her şeyden önce, görüntünün bir kısmı evrişim katmanına bağlanmaktadır ve özellik çıkarılmaktadır. Ardından filtreyi aynı giriş görüntüsünün bir sonraki alıcı alanı üzerinde bir adım kaydırılarak aynı işlem tekrar yapılmaktadır. Görüntünün tamamını gözden geçirene kadar bu işlem devam etmektedir. Çıktı havuzlama katmanının girdisi olacaktır.



Şekil 3. CNN çalışma mantığı (Fiori , 2020).

Havuzlama katmanı

Havuzlama katmanı girdi görüntüsünün uzamsal hacmini azaltmak için iki evrişim katmanı arasında kullanılmaktadır. Böylelikle iş yükü azaltılıp, işlem hızlandırılmaktadır. Şekil 4'deki gibi 4x4 boyutlu girdi görüntüsü, 2x2 boyutuna indirgenmektedir.

Havuzlama katmanında 2 tane hiperparametre bulunmaktadır. Bunlar Filtre (F) ve Adım (S)'dir. Eğer giriş boyutu E_1 (E_1) x Yükseklik₁ (Y_1) x Derinlik₁ (D_1) ise, aşağıdaki hesaplamalara göre

$$E_2 = (E_1 - F) \div S + 1 \quad (1)$$

$$Y_2 = (Y_1 - F) \div S + 1 \quad (2)$$

$$D_2 = D_1 \quad (3)$$

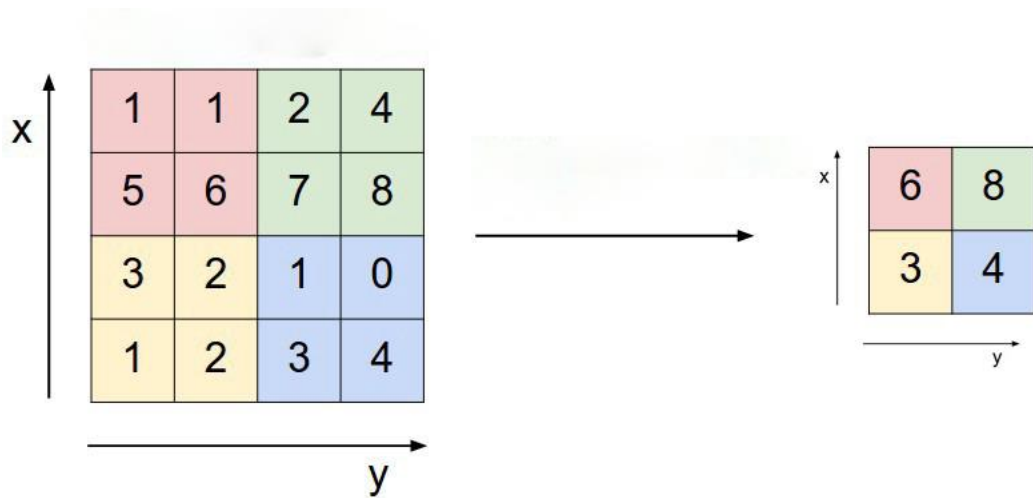
çıktının ölçüleri $E_2 \times Y_2 \times D_2$ olmaktadır.

Tam bağlı katmanı

Bir katmanda bulunan nöronları başka bir katmandaki nöronlara bağlar. Bu sebeple görüntüleri farklı kategoriler arasında sınıflandırmak için kullanılmaktadır.

Softmax katmanı

Softmax veya Lojistik katmanı CNN'nin son katmanıdır. Sınıflandırma işlemleri için kullanılmaktadır. Bu sayede yüz üzerinde bulunan 131 noktanın koordinatları tespit edilmektedir.



Şekil 4. 4x4 matrisin 2x2 matrise indirgenmesi (Shahid, 2019).

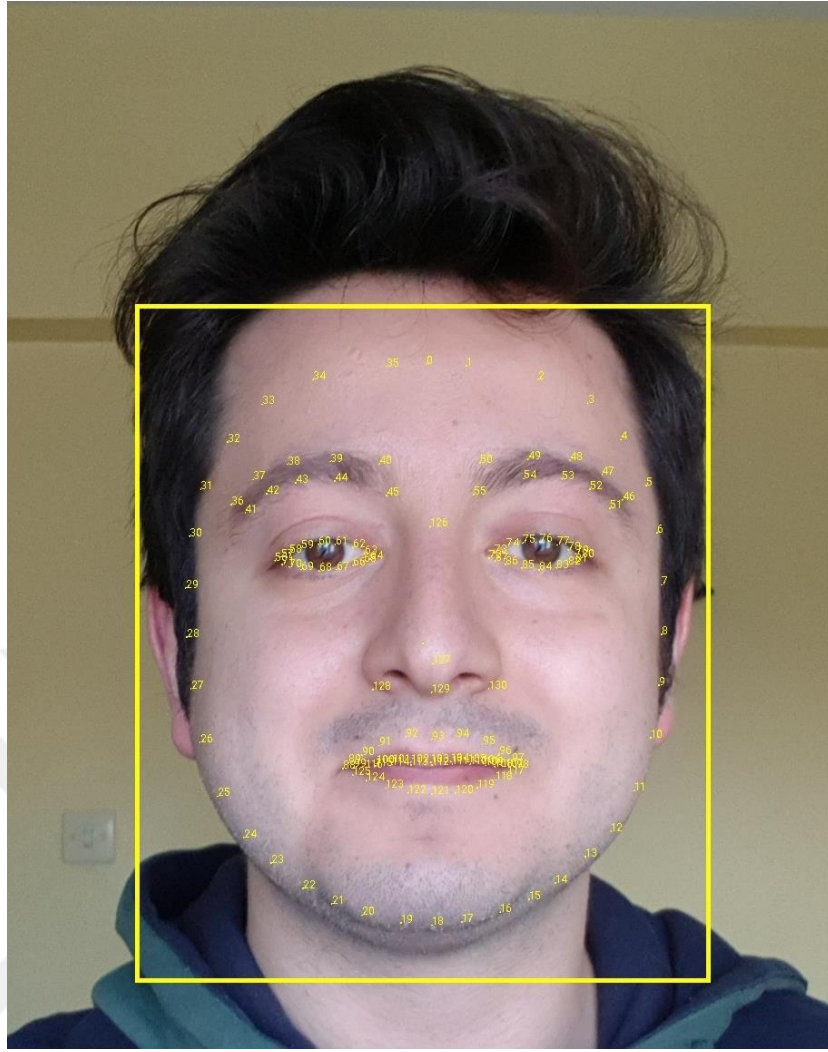
TensorFlow Lite kütüphanesi aracılığıyla tespit edilen koordinatlar sunucudan mobil uygulamaya gönderilmektedir. Mobil uygulama gelen koordinatları işleyerek fotoğraf üzerinde Şekil 5'deki gibi işaretleme yapmaktadır. Noktaların oluşturduğu yüz hatları ve işaretlenen nokta sayıları aşağıdaki gibidir;

- Yüz çevresi – 36 nokta
- Sol kaşın altı – 10 nokta
- Sol kaşın üstü – 10 nokta
- Sağ kaşın altı – 10 nokta
- Sol kaşın üstü – 10 nokta
- Sol göz – 16 nokta
- Sağ göz – 16 nokta
- Üst dudağın altı – 9 nokta
- Üst dudağın üstü – 11 nokta
- Alt dudağın altı – 9 nokta
- Alt dudağın üstü – 9 nokta
- Burun köprüsü – 3 nokta
- Burun alt ve üst noktaları – 2 nokta

Bu sayede nokta tespitinin doğru yapılıp yapılmadığı hakkında fikir sahibi olunabilmektedir. Nokta tespitinin yanlış yapılması sınıflandırma sonuçlarını etkileyeceği için gerekli durumlarda tekrardan fotoğraf çekimi yapılabilmektedir.

Üç boyutlu görüntülerden yüz noktalarının tespiti

Fotoğrafın 3 boyutlu görüntüye çevrilmesi için öncelikli olarak 2 boyutlu görüntü üzerindeki yüz noktaları tespit edilmiş, sonrasında yüz hizalama ağı oluşturulmuştur. Oluşturulan ağ 300-W-LP veri setiyle eğitilmiştir. 300-W-LP veri setinin kullanılmasının amacı 2 boyutlu yüz görüntülerindeki noktaların dışında 3 boyutlu yüz görüntülerinin noktalarını içermesidir. Oluşturulan ağ aracılığıyla yüz üzerindeki noktaların tespiti sağlanmıştır. Daha sonra işaretlenen noktaların yakınlarına rastgele şekilde bir nokta daha eklenmiştir. Eklenen noktalarda standart sapmanın 1 piksel olmasına dikkat edilmiştir. Bu noktalar 3 boyutlu görüntünün yüz noktalarının izdüşümleri olarak varsayılarak yeni noktalar oluşturulmuştur (Bulat & Tzimiropoulos, 2017). Bu sayede 2 boyutlu görüntüden 3 boyutlu görüntü elde edilmiştir.



Şekil 5. 2 boyutlu görüntüde işaretlenmiş yüz noktaları

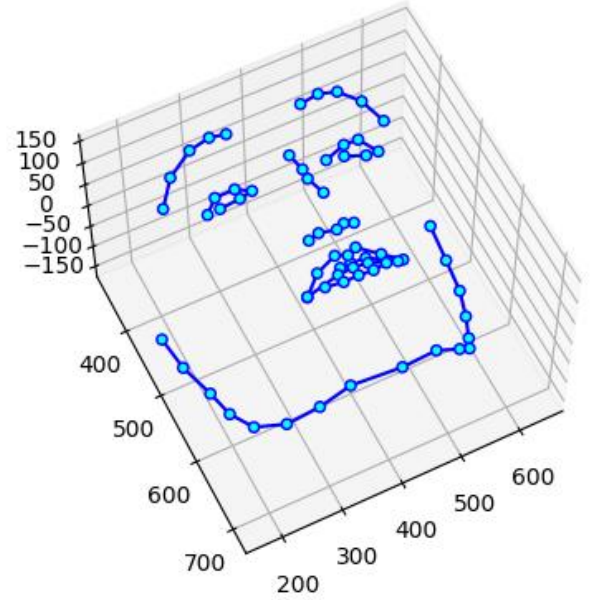
Yeniden inşa işlemi sırasında CUDA mimarisi kullanılmıştır. CUDA mimarisi aracılığıyla yazılan kod GPU üzerinde çalıştırılmıştır. Bu sayede işlem süresi kısaltılmıştır. 3 boyutlu inşa işleminden sonra yüz üzerinde 102 nokta işaretlenmiştir. 2 boyutlu görüntünün 3 boyutlu olarak yeniden inşa edilmesinin örnek görüntüleri Şekil 6'da verilmiştir. 3 boyutlu görüntülerden nokta tespiti için açık kaynaklı face-alignment kütüphanesi kullanılarak veri setinde bulunan görüntüler 3 boyutlu olarak yeniden inşa edilmiştir.

Elde edilen koordinatlar dizi içerisinde alındıktan sonra bilgisayar ortamında noktaların birbirlerine göre uzaklıkları hesaplanıp tablo olarak kaydedilmiştir. 2 boyutlu görüntünün 3 boyutlu olarak yeniden inşa edilmesinin amacı fotoğraflardaki derinlik faktörünün sınıflandırma işlemine etkilerinin incelenebilmesidir.

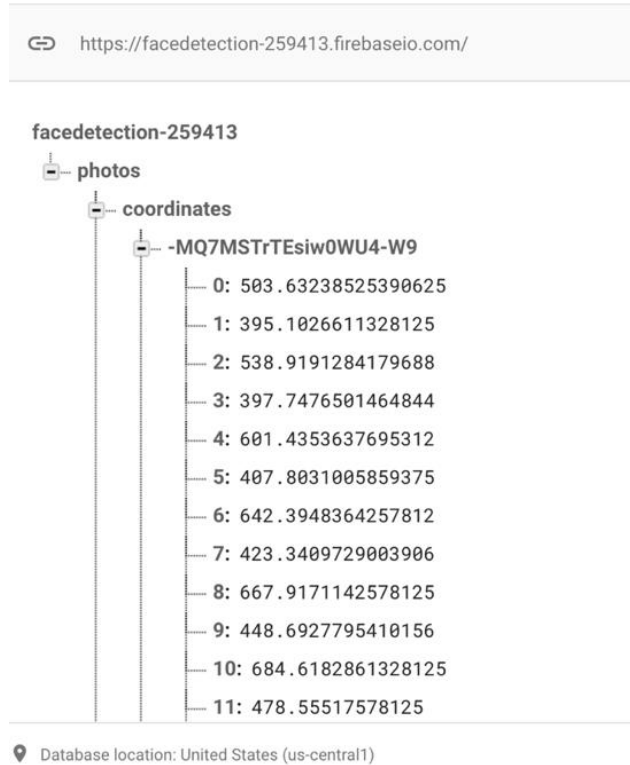
Yüz Noktalarının İşlenmesi

Mobil uygulama aracılığıyla alınan koordinatlar Firestore veri tabanına kaydedilmektedir. Şekil 7'de veri tabanının örnek görüntüsü verilmiştir. İnternetin olmaması

durumunda ise çevrimdışı olarak yüz tespiti yapılarak yüz üzerindeki noktalar işaretlenmektedir. Noktaların koordinatları JavaScript Object Notion (JSON) formatında telefonda saklanmaktadır. Telefon internet erişimi sağladığı zaman Firestore veri tabanına kaydedilmektedir.



Şekil 6. 2 boyutlu görüntünün 3 boyutlu olarak yeniden inşa edilmesi sonrası işaretlenen yüz noktaları



Şekil 7. Verilerin Firestore veri tarabanında saklanması

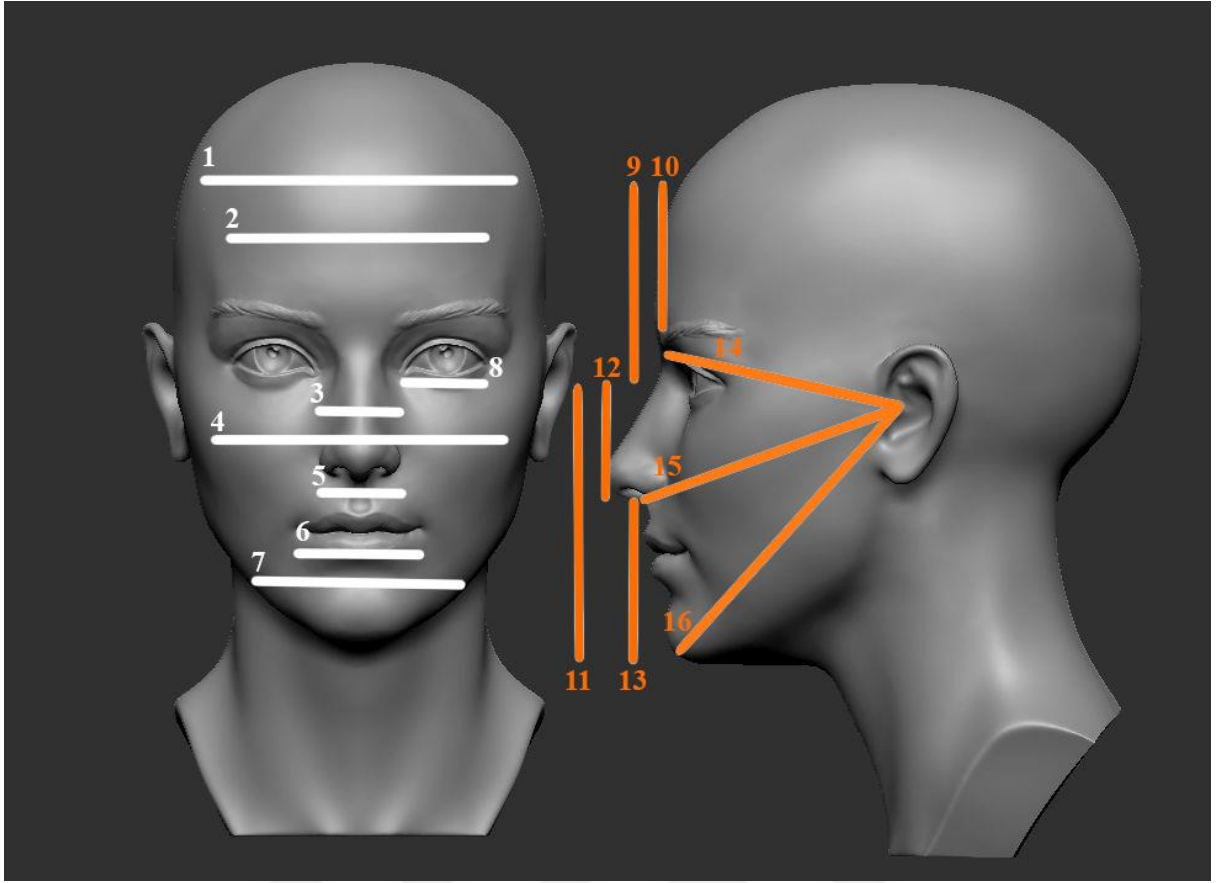
Koordinat verileri bilgisayar ortamı aracılığıyla Firestore veri tabanından alınarak tablo biçiminde bilgisayara kaydedilmektedir. Aktaş ve diğerlerin yapmış olduğu çalışmada çekilen fotoğraflardan manuel olarak elde edilen yüz üzerindeki 16 uzunluğun dikkat eksikliği ve hiperaktivite bozukluğu üzerindeki etkileri izlenmiştir. Aktaş ve diğerleri tarafından yapılan çalışmada önerdiği 16 uzunluk Şekil 8’de verilmiş ve tanımlamaları Tablo 2’de yapılmıştır (Aktaş et al., 2020). Bu çalışmada da bu işlemin görüntü işleme algoritmaları aracılığıyla yapılması planlanmıştır. Bunun için öncelikle 16 uzunluğu elde etmek için hangi noktaların alınacağı tespit edilmiştir. Sonrasında alın genişliği, burun genişliği, yüz yüksekliği gibi uzunlukların hesaplanmasında ikinci bir koordinata gerek olmadığı için genişliklerde $x = 0$ varsayılarak y koordinatlarının uzunlukları, yükseklikler için de $y = 0$ varsayılarak x koordinatlarının uzunlukları aşağıdaki gibi hesaplanmıştır.

$$(x_2 - x_1) \text{ veya } (y_2 - y_1) \quad (4)$$

Üst yüz derinliği, orta yüz derinliği ve alt yüz derinliği için x , y koordinatlarının uzunlukları öklid teoremine göre aşağıdaki gibi hesaplanmıştır.

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (5)$$

İşaretlenmiş noktalar saç ve alının kesiştiği bölgeden geçmemektedir. Bu sebeple alın yüksekliğinin hesaplanabilmesi için yüz görüntüleri üzerinde manüel ölçümler yapılmıştır. Yapılan ölçümlere göre göz bebeğinin merkezi hizasında olan yüz çevresi noktası (34. Nokta) ile kaşın üst hattının orta noktası (38. nokta) arasındaki dikey uzaklığın iki katının, kaşın üst hattının orta noktasının (38. nokta) gözün sol köşe noktasına (56. nokta) dikey uzaklığıyla toplamından alın yüksekliği elde edilmektedir. Alın yüksekliği aşağıdaki formüle göre hesaplanmıştır.



Şekil 8. Önerilen 16 uzunluğun model üzerinde gösterimi (Aktaş et al., 2020).

$$(Y_{34} - Y_{38}) \times 2 + (Y_{38} - Y_{56}) \quad (6)$$

Bilgisayar ortamında ilgili formüller kullanılarak 16 uzunluk hesaplanmış ve excel tablosu oluşturulmuştur. 16 uzunluğun yanı sıra tüm noktaların birbirlerine olan uzunlukları hesaplanarak, excel tablosuna şeklinde kaydedilmiştir. Algoritma 1’de uzunlukları bulmak için kullanılan algoritma açıklanmıştır. Algoritmada kullanılan x ve y koordinat noktalarını temsil etmektedir.

Algoritma 1. Noktaları arası uzaklığı bulmak için kullanılan algoritma

Girdi: Yüz üzerindeki noktaların koordinatlarının bulunduğu dizi (K),

Çıktı: Noktalar arası uzaklık

- 1: **for** x = 0 to length of K, increase x by 2 **do**
- 2: **for** y = x + 2 to K, increase x by 2 **do**
- 3: uzaklık = square root of $(y - x)^2 + ((y + 1) - (x + 1))^2$
- 4: **append** uzaklık to tablo satırı

Tablo 2. Önerilen 16 nokta ve koordinat bilgileri.

#	Koordinat	Alınan uzunluk
1	X4 – X32	Alın genişliği
2	X56 – X80	Göz dışlarının genişliği
3	X64 – X72	Göz içlerinin genişliği
4	X36 – X46	Üst yüz genişliği
5	X128 – X130	Burun genişliği
6	X88 – X98	Ağız genişliği
7	X12 – X24	Alt yüz genişliği
8	X72 – X80	Göz genişliği
9	(Y34 – Y38) * 2	Alın yüksekliği 2
10	(Y34 – Y38) * 2 + (Y38 – Y56)	Alın yüksekliği 1
11	Y18 – Y56	Yüz yüksekliği
12	Y126 – Y129	Burun uzunluğu
13	Y18 – Y129	Alt yüz uzunluğu
14	(X9, Y9) – (X126, Y126)	Üst yüz derinliği
15	(X9, Y9) – (X129, Y129)	Orta yüz derinliği
16	(X9, Y9) – (X18, Y18)	Alt yüz derinliği

Yukarıdaki işlemler yapıldıktan sonra sınıflandırma işleminde kullanılmak üzere 4 farklı öznitelik oluşturulmuştur. İlk öznitelik olarak 16 uzunluk kullanmaktadır. İkinci öznitelik olarak yüz noktalarının birbirlerine göre uzunlukların alındığı, 8515 uzunluğu kullanmaktadır. Üçüncü öznitelik için 16 uzunluğun birbirlerine göre oranlarını alınmıştır. Son öznitelikte ise 2 boyutlu görüntülerden oluşturulmuş 3 boyutlu görüntüler üzerindeki yüz noktaları tespit edilerek birbirlerine göre uzaklıkları alınmıştır. Tablo 3’te elde edilen öznitelikler verilmiştir.

Öznitelik Seçimi

Öznitelik seçimi, sınıflandırma işlemine en çok katkıda bulunan verilerin seçildiği bir yöntemdir. Bu sayede sınıflandırma sonucunu olumsuz etkileyen faktörler ortadan kaldırılmaktadır. Öznitelik seçiminin avantajları aşağıdaki gibi sıralanabilir (Brownlee, 2016);

- Sınıflandırma işlemi sırasında öznitelik sayısını azaltarak, sınıflandırma işleminin başarı oranını düşürecek yanıltıcı verileri ortadan kaldırmak.
- Daha az veri kullanarak algoritmalarının hızlı çalışmasını sağlamak.

Yapılan çalışma kapsamında yüz noktalarıyla yapılan çalışmalarda ağaç yapısının kullanıldığı öne sürülmüştür (Li et al., 2018). Bu sebeple bu çalışmada ağaç yapısı kullanan K-Best, rastgele karar ormanları ve modelden seçim algoritmaları kullanılmıştır. Ek olarak istatistiksel bir öznitelik seçimi yöntemi olan T-Test algoritması kullanılmıştır.

K-Best

K-Best öznitelik seçimi için SelectKBest metodu kullanılmıştır. Score_func parametresine verilen değere göre sınıflandırma verisi ve regresyon verisi için kullanılabilir. Sınıflandırma işlemi için parametre değeri olarak chi2 verilmektedir. Öznitelik seçimi sonrası oluşacak öznitelik sayısı için k parametresi tanımlanabilmektedir. SelectKBest veri seti içerisinde sınıflandırma sonuçlarına en çok katkıda bulunan verilerinin seçilmesini sağlamaktadır. SelectKBest sınıfı bu özellikleri en yüksek k puanına göre seçmektedir. Eğitim verisinde çok fazla sütun olması halinde en iyi özellikleri seçerek eğitim süresini kısaltmaktadır. K-Best öznitelik seçimi için Python programlama dilinde makine öğrenmesi için geliştirilen Scikit-learn kütüphanesi kullanılmaktadır.

Tablo 3. Sınıflandırma İşlemi İçin Veri Setinden Elde Edilen Noktaların Uzunluklarından Elde Edilen Öznitelikler

	Öznitelik 1	Öznitelik 2	Öznitelik 3	Öznitelik 4
Öznitelik	16 uzunluk	Tüm uzunluklar	16 uzunluğun oranı	3D görüntüden elde edilen uzunluklar
Öznitelik sayısı	16	8515	120	5151

Rastgele Karar Ormanları (Random Forest)

Rastgele karar ormanları birçok karar ağacının bir araya gelmesinden oluşmaktadır. Öznitelik seçimi için Gini önem formülü kullanılmaktadır. n_{ij} j düğümünün önemi, w_j j düğümüne ulaşan ağırlık sayısı, C_j j düğümüne ulaşan yetersiz ağırlık sayısı olmak üzere Gini önem formülü aşağıda verilmiştir.

$$n_{ij} = w_j C_j - w_{l(j)} C_{l(j)} - w_{r(j)} C_{r(j)} \quad (7)$$

Karar ağacında bulunan her bir özneliğin önemi f_i i özneliğinin önemi, n_{ij} j nodunun önemi olmak üzere aşağıdaki şekilde hesaplanmaktadır.

$$f_i = \frac{\sum_{j:j \text{ nodunun } i \text{ özneliği üzerindeki dağılımı}} n_{ij}}{\sum_{k \in \text{tüm nodlar}} n_{ik}} \quad (8)$$

Rasgele karar ormanlarında her ağaç performansı artırmak için Şekil 9'daki gibi sınıf tahmini yapmaktadır. Ağaç sayısının fazla olması doğruluk oranını artırmaktadır. Karar ağaçlarından çıkan sonuçların çoğunu alarak sonucu iyileştirmektedir. Karar ağaçlarında kullanılan veriler öznitelik seçimi olarak kullanılmaktadır.

Modelden Seçim (Select From Model)

SelectFromModel metodunda öznitelik seçimi sonrası oluşacak veri seti boyutu için parametre verisi girilmemektedir. Metot veri setinde bulunan tüm verilerin öznitelik önem değerlerini kontrol etmektedir. Ortalamanın altında kalan değerleri veri setinden çıkararak sınıflandırma başarısı yüksek bir veri seti oluşturmaktadır. SelectFromModel metodu için Python programlama dilinin Scikit-learn kütüphanesi kullanılmıştır.

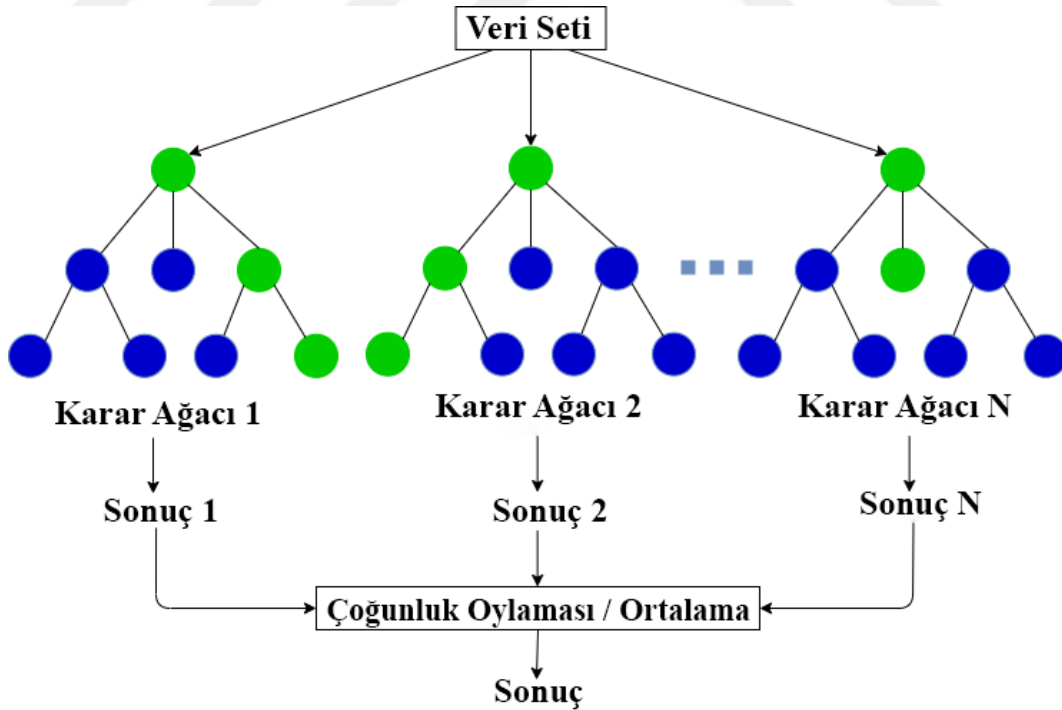
T-Test

T-Test iki verinin arasında belirli özelliklerle ilişkili anlamlı bir ilişki olup olmadığını belirlemek için kullanılır. T-Test'i uygulamak için üç adet değere ihtiyaç duyulmaktadır.

- Her veriden elde edilen ortalama değerler arasındaki fark
- Her veri grubunun standart sapması
- Her veri grubunun veri değerlerinin sayısı

T-Test ile öznitelik seçimi yapabilmek için aşağıdaki denklem kullanılmaktadır;

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left(s^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right)}} \quad (9)$$



Şekil 9. Rastgele karar ormanları çalışma diyagramı (Ampadu, 2021).

9 numaralı denklemde x_1 ve x_2 karşılaştırılan iki grubun elde edilen ortalama değerleri arasındaki farkı, s iki veri grubunun standart hatasını ve n_1 ve n_2 her bir gruptaki veri değerlerinin sayısını temsil etmektedir.

Sınıflandırma

Sınıflandırma verilen girdilerin kategorilerine göre ayrılması olarak adlandırılmaktadır. Yapılan sınıflandırma işlemleri sayesinde makine öğrenme işlemi gerçekleştirilmektedir. Bu sayede çalışan algoritma sonradan gelecek yüz görüntülerinin hastalıklı bir birey mi yoksa sağlıklı bir birey mi olacağını tahmin edebilecektir. Yapılan çalışmada üç farklı sınıflandırma işlemi kullanılmaktadır. Çalışmada kullanılan sınıflandırma algoritmaları aşağıdaki gibidir;

- K En Yakın Komşu
- Naive Bayes
- Destek Vektör Makinesi

K En Yakın Komşu

K En Yakın Komşu (KNN) algoritması, sınıflandırma işlemi gerçekleştirmek için en yakın noktaları kullanmaktadır. Veri seti yüklendikten sonra kullanıcı tarafından K değeri belirlenmektedir. Aşağıdaki öklid formülü kullanılarak bulunan uzaklıklar listeye eklenmektedir.

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (10)$$

Daha sonra seçilen K değerine göre test verisine en yakın eğitim verisi noktaları tespit edilir ve bu noktalar içerisinde en çok geçen etiket, test verisinin etiketi olarak atanır (Zhang & Zhou, 2007). Algoritmanın çalışma diyagramı Şekil 10'da verilmiştir.

Naive Bayes

Naive Bayes sınıflandırması makine öğrenmesinde gözetimli öğrenme kullanan algoritmalardan bir diğeridir. Algoritma Bayes teorisine göre çalışmaktadır. Algoritmanın sözde kodu Algoritma 2'de verilmiştir. Algoritma çalışırken veri setinde bulunan her eleman için durumların olasılığını hesaplamaktadır ve olasılık değeri en yüksek veriyi seçip, o veriye göre sınıflandırma işlemi gerçekleştirilmektedir.

Naive Bayes teorisine göre B olayı gerçekleştiğinde A olayının gerçekleşme olasılığı A olayı gerçekleştiğinde B olayının gerçekleşme olasılığıyla A olayının gerçekleşme olasılığının çarpımının B olayının gerçekleşme olasılığına bölümünden elde edilir. Bu durumda aşağıdaki formül ortaya çıkmaktadır;

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)} \quad (11)$$

Algoritma 2. Naive Bayes algoritmasının sözde kodu (Saputra et al., 2018).

Girdi: Eğitim veri seti T,

$F(f_1, f_2, f_3, \dots, f_n)$ // test veri setinde tahmin değişkeninin değeri

Çıktı: Test veri seti sınıfı

1: T eğitim veri setini oku;

2: Her sınıftaki tahmin değişkenlerinin ortalamasını ve standart sapmasını hesapla;

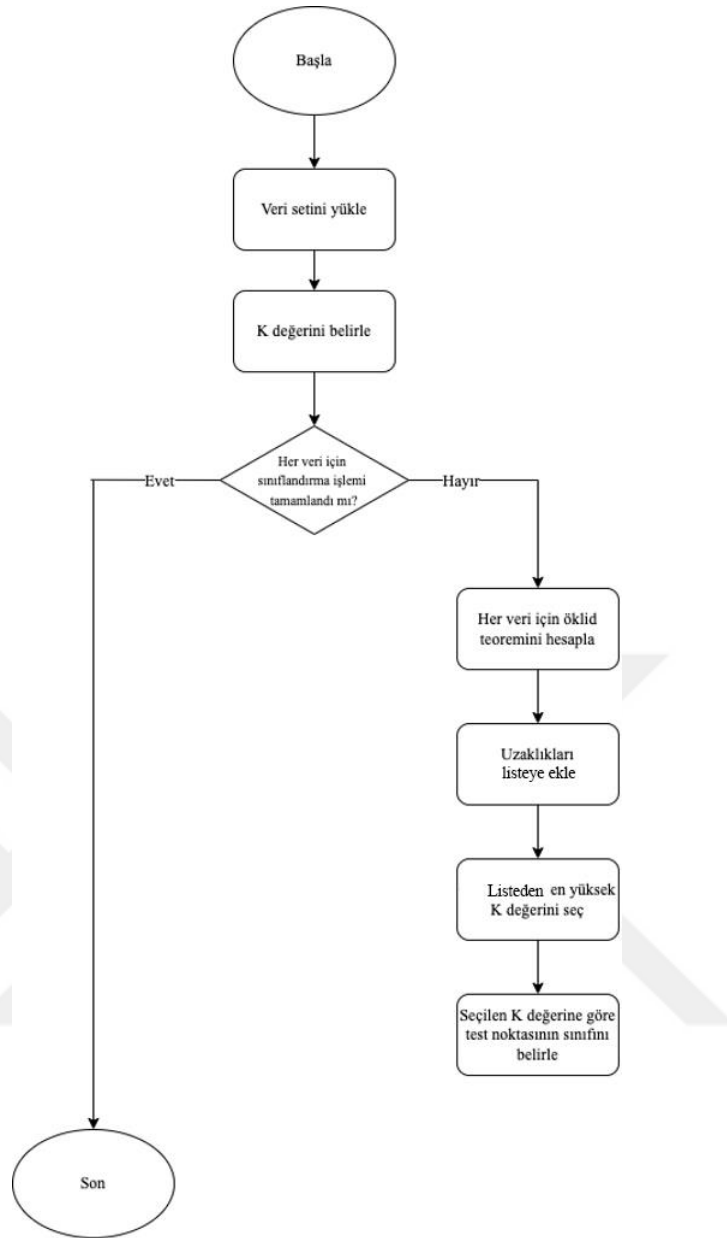
3: Tekrarla

Her sınıfta gauss yoğunluk denklemini kullanarak f_i olasılığını hesapla;

Tüm tahmin değişkenlerinin $(f_1, f_2, f_3, \dots, f_n)$ olasılığı hesaplanana kadar.

4: Her sınıf için olasılığı hesapla;

5: En yüksek olasılığı al;



Şekil 10. KNN çalışma diyagramı (Sam , 2019).

Destek Vektör Makineleri

Destek Vektör Makineleri (DVM) gözetimli öğrenme problemlerinde kullanılmaktadır. Küçük ve ortak ölçekli veri setlerinde daha iyi performans göstermektedir. DVM'nin amacı veri setinde bulunan verileri çok boyutlu uzay üzerinde işaretlemektir ve bu çoklu boyutta hiper düzlemler oluşturarak sınıflandırma işlemi yapmaktadır. Şekil 11'de yıldızlar ve daireler olmak üzere iki farklı sınıf bulunmaktadır. Sınıfların yerleri çok boyutlu uzay üzerinde işaretlenmiştir. Sonradan gelecek verinin hangi sınıfa ait olduğunu tespit etmek için öncelikli olarak iki sınıfı ayıran C doğrusu çizilmiştir. Bu doğrunun ± 1 'ine ise A, B destek vektörleri çizilmiştir. A ve B destek vektörleri arasındaki mesafe ne kadar fazla olursa sınıflandırma işleminin yapılması kolaylaşmaktadır.

Sınıflandırma işleminde kullanılan destek vektörleri ve maksimum kenar hiper düzlemi Vapnik teorisine göre çizilmektedir (Chandra & Bedi, 2018). Eğer girdi veri seti x_i , $m \times n$ boyutlu, nöron ağırlığı w ve karşılık gelen sınıf $y_i \in \{-1, +1\}$ ise;

$$w^T \phi(x_i) + b = 1 \quad (12)$$

Hiper düzlem her iki sınıf içinde aşağıdaki şekilde oluşturulmaktadır;

$$w^T \phi(x_i) + b \geq +1 \quad y_i = +1 \quad (13)$$

$$w^T \phi(x_i) + b < +1 \quad y_i = -1 \quad (14)$$

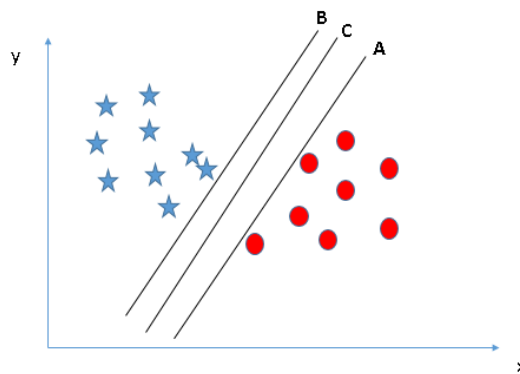
12. ve 13. denklem kullanılarak 15. denklem oluşturulmaktadır.

$$y_i(w^T \phi(x_i) + b) \geq +1 \quad i = 1 \dots N \quad (15)$$

Bu durumda yeni gelen veri için çıkan sonuç denklem 11'deki gibi 1'den büyükse veri yıldız olarak sınıflandırılmaktadır. Fakat çıkan sonuç denklem 3'teki gibi 1'den küçükse veri daire olarak sınıflandırılmaktadır.

Doğruluk ve Değerlendirme Metrikleri

Çalışmada doğruluk ve değerlendirme metriği olarak K katlamalı çapraz doğrulama algoritması kullanılmıştır. Algoritmanın amacı başarı oranının test edilmesi için, veri seti içerisinde bulunan verilerin eğitim ve test veri seti olarak ayırmaktır. Bu ayırma işlemi istenilen oranlarda yapılabilmektedir. Ayırma işlemi gerçekleştirildikten sonra eğitim veri setiyle makine öğrenmesi gerçekleştirilip, sonrasında test veri setiyle makine öğrenmesi test edilmektedir.



Şekil 11. Destek vektör makinesi sınıflandırma örneği. A, B destek vektörleri ve C maksimum kenar hiper düzlemi (Alay, 2020).

K katlamalı çapraz doğrulama algoritmasında tüm bunların dışında bir k değeri seçilmektedir. Veri seti k değeri kadar eşit parçalara bölünmektedir. Sistemin $1/k$ kadarı test için kullanılır, geri kalan veriler ise eğitim için kullanılmaktadır. Daha sonra sınıflandırma

işlemi için bu yöntem k kere çalışmaktadır ve her sınıflandırma işlemi için bir sonuç değeri elde edilmektedir. Sistemin genel başarısı ve genel hata payı elde edilen sonuç değerlerinin ortalamasından elde edilmektedir. Bu durumda aşağıdaki denklem elde edilmektedir;

$$\frac{\sum_{i=0}^k T_s E_s(t_i, VK - t_i)}{k} \quad (16)$$

Denklemden T_s test veri setinin sınıflandırma fonksiyonunu, E_s eğitim veri setinin sınıflandırma fonksiyonunu, VK veri kümesini, k kaç adet katlama kullanıldığı ve t ise seçilen test veri kümesini ifade etmektedir.

K katlamalı çapraz doğrulama algoritmasının sonucunda karışıklık matrisi elde edilmektedir. Karışıklık matrisi Şekil 12’de verilmiştir. Karışıklık matrisinden el edilen veriler doğrultusunda doğruluk, kesinlik, duyarlılık ve F1 skoru değerleri hesaplanmaktadır.

Doğruluk değeri veri setinde doğru tahmin ettiğimiz verilerin tüm veri setine oranı ile hesaplanmaktadır. GP gerçek pozitif, YN yanlış negatif, YP yanlış pozitif ve GN gerçek negatif olmak üzere doğruluk değerinin denklemi aşağıdaki gibidir;

$$\frac{GP + GN}{GP + GN + YP + YN} \quad (17)$$

Kesinlik değeri pozitif olarak tahmin etmemiz gereken verilerin gerçekte ne kadarının pozitif olduğunu gösteren metriktir. Hesaplanması için kullanılan denklem aşağıdaki gibidir;

$$\frac{GP}{GP + YP} \quad (18)$$

Duyarlılık ise pozitif olarak tahmin edilmesi gereken verilerin ne kadarının pozitif tahmin edildiğini gösteren metriktir. Duyarlılık hesaplaması için kullanılan denklem aşağıdaki gibidir;

$$\frac{GP}{GP + YN} \quad (19)$$

F1 skoru ise kesinlik ve duyarlılık değerlerinin harmonik ortalamasını göstermektedir. Harmonik ortalama kullanılmasının amacı veri setinde bulunan çok düşük ve çok yüksek değerleri ortalamaya dahil etmemektir. F1 skorunun hesaplanması için kullanılan formül aşağıda gösterilmiştir.

$$F1 = 2 \times \frac{Kesinlik \times Duyarlılık}{Kesinlik + Duyarlılık} \quad (20)$$

		Tahmin edilen deęer	
Gerçek deęer		Gerçek Pozitif	Yanlış Negatif
		Yanlış Pozitif	Gerçek Negatif

Şekil 12. Karışıklık matrisi

ARAŞTIRMA BULGULARI VE TARTIŞMA

Bu çalışma kapsamında çocuk yüz fotoğraflarında çeşitli noktalar yardımıyla dikkat eksikliği ve hiperaktivite bozukluğunun tespitinde gözetimli öğrenme algoritmalarının uygulanabilirliğini ortaya koymak amaçlanarak k en yakın komşu, naive bayes, destek vektör makineleri kullanılmıştır. Yüz görüntülerinden elde edilen öznitelikler üzerinde sınıflandırma işlemi yapılmıştır. Yapılan çalışmada veri setinde bulunan değerlerin %20'si test verisi, %80'i ise kontrol verisi olarak seçilmiştir. Ek olarak öznitelik seçimi işlemi yapılmış ve sınıflandırma işlemi tekrarlanmıştır. Araştırma bulguları üç başlık altında toplanmıştır. Bunlar;

- Sınıflandırma algoritmalarının sonuçları,
- Öznitelik seçimi sonrası sınıflandırma algoritmalarının sonuçları,
- 3 boyutlu yeniden inşa sonrasında sınıflandırma algoritmalarının sonuçlarıdır.

Sınıflandırma Algoritmalarının Sonuçları

Öznitelik 1 için yapılan ilk uygulamada elde edilen doğruluk değerleri Tablo 4'te, kesinlik, duyarlılık ve F1 skoru değerleri Tablo 5'te verilmiştir. KNN ve SVM algoritmaları için %63,2'lik başarı elde edilmiştir. Her iki algoritma için K-Fold değerleri incelendiği zaman %10'luk bir sapma gözlemlenmektedir. Bu duruma bağlı olarak aşırı öğrenmenin gerçekleştiği sonucuna varılabilir. Algoritmalar için kesinlik ve duyarlılık sonuçları incelendiği zaman DEHB grubunu tahmin etmekte zorlandıkları fakat kontrol grubunu %75'in üzerinde bir oranla tahmin edebildikleri görülmüştür.

Tablo 4. Öznitelik 1'in Sınıflandırma Sonuçlarının Doğruluk Değerleri

Sınıflandırma Yöntemi	Doğruluk	K-Fold Çapraz Doğrulama
KNN	0,632	0,570
Naive Bayes	0,579	0,460
SVM	0,632	0,530

Tablo 5. Öznitelik 1'in Sınıflandırma Sonuçları

Sınıflandırma Yöntemi	İşaret	Kesinlik	Duyarlılık	F1 Skoru
KNN	DEHB	0,540	0,880	0,670
	KONTROL	0,830	0,450	0,590
Naive Bayes	DEHB	0,710	0,450	0,560
	KONTROL	0,500	0,750	0,600
SVM	DEHB	0,600	0,900	0,720
	KONTROL	0,750	0,330	0,460

Elde edilen performans değerlerinin artırılması amaçlanarak veri seti yeniden örneklendirilmiştir. Bu doğrultuda öznitelik 1’den farklı olarak yüz üzerindeki görüntülerden elde edilen 131 noktanın birbirlerine olan uzaklıkları içeren öznitelik 2 oluşturulmuş olup, k en yakın komşu, naive bayes ve destek vektör makineleri algoritmaları uygulanmıştır. Uygulama sonucunda elde edilen doğruluk değerleri Tablo 6’da, kesinlik, duyarlılık ve F1 skoru değerleri Tablo 7’de verilmiştir. Öznitelik 2 için en yüksek başarı %78,9’luk doğruluk oranıyla KNN algoritmasıyla elde edilmiştir. Çapraz doğrulama sonucuna bakıldığında zaman %13’lük bir sapma görülmektedir. Kontrol grubu için duyarlılık değeri incelendiğinde ise kontrol grubu tahminlerinde başarının düştüğü gözlemlenmektedir.

Tablo 6. Öznitelik 2’nin Sınıflandırma Sonuçlarının Doğruluk Değerleri

Sınıflandırma Yöntemi	Doğruluk	K-Fold Çapraz Doğrulama
KNN	0,789	0,654
Naive Bayes	0,684	0,530
SVM	0,737	0,740

Tablo 7. Öznitelik 2’nin Sınıflandırma Sonuçları

Sınıflandırma Yöntemi	İşaret	Kesinlik	Duyarlılık	F1 Skoru
KNN	DEHB	0,770	0,910	0,830
	KONTROL	0,830	0,620	0,710
Naive Bayes	DEHB	0,640	0,780	0,700
	KONTROL	0,750	0,600	0,670
SVM	DEHB	0,690	0,900	0,780
	KONTROL	0,830	0,560	0,670

Öznitelik 2 ile yapılan uygulamada performans ve doğruluk değerlerinde artış görülmüştür. Bu amaçla önerilen 16 uzunluğun sınıflandırma sonuçlarını iyileştirebilmek adına 16 uzunluğun birbirlerine göre oranlarını içeren öznitelik 3 oluşturulmuş olup, k en yakın komşu, naive bayes ve destek vektör makineleri algoritmaları uygulanmıştır. Uygulama sonucunda elde edilen doğruluk değerleri Tablo 8’de, kesinlik, duyarlılık ve F1 skoru değerleri Tablo 9’da verilmiştir. Öznitelik 3 için en yüksek doğruluk değeri %73,7 ile KNN için tespit edilmiştir. Öznitelik 1 ile karşılaştırıldığında zaman öznitelik sayısındaki artış doğruluk değerinin %10 artmasına neden olduğu sonucu çıkarılabilir. KNN algoritması için kesinlik ve duyarlılık sonuçları incelendiğinde zaman kontrol grubu tahminlerinde azalma olduğu görülmüştür.

Tablo 8. Öznitelik 3'ün Sınıflandırma Sonuçlarının Doğruluk Değerleri

Sınıflandırma Yöntemi	Doğruluk	K-Fold Çapraz Doğrulama
KNN	0,737	0,605
Naive Bayes	0,632	0,530
SVM	0,630	0,630

Tablo 9. Öznitelik 3'ün Sınıflandırma Sonuçları

Sınıflandırma Yöntemi	İşaret	Kesinlik	Duyarlılık	F1 Skoru
KNN	DEHB	0,800	0,730	0,760
	KONTROL	0,670	0,750	0,710
Naive Bayes	DEHB	0,750	0,550	0,630
	KONTROL	0,550	0,750	0,630
SVM	DEHB	0,800	0,670	0,730
	KONTROL	0,560	0,710	0,630

Öznitelik Seçimi Sınıflandırma Sonuçları

Özniteliklerin sınıflandırma işlemi sırasında öznitelik sayısını azaltılması ve sınıflandırma işleminin başarı oranını düşürecek gürültülerin ortadan kaldırılması amacıyla yukarıda elde edilen üç farklı öznitelige öznitelik seçimi algoritmaları olan K-Best, rastgele karar ormanları, modelden seçim, T-Test uygulanmıştır. Yapılan uygulamalarda k en yakın komşu, naive bayes ve destek vektör makineleri algoritmaları kullanılmıştır.

Öznitelik 1 için yapılan uygulamalar sonucunda elde edilen doğruluk değerleri Tablo 10'da, kesinlik, duyarlılık ve F1 skoru değerleri Tablo 11'de verilmiştir. Elde edilen sonuçlar incelendiği zaman öznitelik çıkarma işleminin sınıflandırma işlemlerine olumlu katkısı olduğu görülmüştür. K-Best, rastgele karar ormanları, modelden seçim, T-Test öznitelik seçimi sonrasında her algoritma için sırasıyla 5 ,10, 5, 2 öznitelik elde edilmiştir. En yüksek başarı öznitelik seçimi yöntemi olan modelden seçimin kullanıldığı KNN algoritmasında görülmüştür. T-Test uygulanan SVM algoritmasının kontrol grubu tahminlerinde düşük performans sergilediği görülmüştür.

Tablo 10. Öznitelik 1'in Öznitelik Seçimi Sonrası Sınıflandırma Sonuçlarının Doğruluk Değerleri

Öznitelik Çıkarma Yöntemi	Sınıflandırma Yöntemi	Doğruluk	K-Fold Çapraz Doğrulama
K-Best	KNN	0,632	0,678
	Naive Bayes	0,632	0,485
	SVM	0,684	0,530
Rastgele Karar Ormanları	KNN	0,579	0,638
	Naive Bayes	0,474	0,484
	SVM	0,684	0,580
Modelden Seçim	KNN	0,737	0,614
	Naive Bayes	0,474	0,539
	SVM	0,684	0,530
T-Test	KNN	0,684	0,624
	Naive Bayes	0,368	0,430
	SVM	0,526	0,470

Tablo 11. Öznitelik 1'in Öznitelik Seçimi Sonrası Sınıflandırma Sonuçları

Özellik Çıkarma Yöntemi	Sınıflandırma Yöntemi	İşaret	Kesinlik	Duyarlılık	F1 Skoru
K-Best	KNN	DEHB	0,670	0,730	0,700
		KONTROL	0,570	0,500	0,530
	Naive Bayes	DEHB	0,500	0,860	0,630
		KONTROL	0,860	0,500	0,630
	SVM	DEHB	0,700	0,700	0,700
		KONTROL	0,670	0,670	0,670
Rastgele Karar Ormanları	KNN	DEHB	0,640	0,640	0,640
		KONTROL	0,500	0,500	0,500
	Naive Bayes	DEHB	0,710	0,380	0,500
		KONTROL	0,330	0,670	0,440
	SVM	DEHB	0,670	0,800	0,730
		KONTROL	0,710	0,560	0,630
Modelden Seçim	KNN	DEHB	0,750	0,820	0,780
		KONTROL	0,710	0,620	0,670
	Naive Bayes	DEHB	0,420	0,620	0,500
		KONTROL	0,570	0,360	0,440
	SVM	DEHB	0,670	0,800	0,730
		KONTROL	0,710	0,560	0,630
T-Test	KNN	DEHB	0,620	1,000	0,770
		KONTROL	1,000	0,330	0,500
	Naive Bayes	DEHB	0,670	0,290	0,400
		KONTROL	0,230	0,600	0,330
	SVM	DEHB	0,530	1,000	0,690
		KONTROL	0,000	0,000	0,000

Öznitelik 2 için yapılan uygulamalar sonucunda elde edilen doğruluk değerleri Tablo 12’de, kesinlik, duyarlılık ve F1 skoru değerleri Tablo 13’de verilmiştir. K-Best, rastgele karar ormanları, modelden seçim ve T-Test algoritmalarının uygulanması sonucunda veri setinden sırasıyla 10, 30, 526 ve 32 öznitelik seçilmiştir. Elde edilen özniteliklerin sınıflandırma işlemi sonrasında rastgele karar ormanlarıyla birlikte uygulanan SVM sınıflandırıcısının başarı oranı %80 olarak tespit edilmiştir.

Tablo 12. Öznitelik 2’nin Öznitelik Seçimi Sonrası Sınıflandırma Sonuçlarının Doğruluk Değerleri

Öznitelik Yöntemi	Çıkarma	Sınıflandırma Yöntemi	Doğruluk	K-Fold Çapraz Doğrulama
K-Best		KNN	0,684	0,531
		Naive Bayes	0,632	0,515
		SVM	0,579	0,570
Rastgele Karar Ormanları		KNN	0,737	0,706
		Naive Bayes	0,632	0,545
		SVM	0,800	0,800
Modelden Seçim		KNN	0,789	0,707
		Naive Bayes	0,632	0,552
		SVM	0,737	0,790
T-Test		KNN	0,684	0,639
		Naive Bayes	0,632	0,556
		SVM	0,632	0,530

Tablo 13. Öznitelik 2’nin Öznitelik Seçimi Sonrası Sınıflandırma Sonuçları

Özellik Çıkarma Yöntemi	Sınıflandırma Yöntemi	İşaret	Kesinlik	Duyarlılık	F1 Skoru
K-Best	KNN	DEHB	0,690	0,920	0,790
		KONTROL	0,670	0,290	0,400
	Naive Bayes	DEHB	0,800	0,400	0,530
		KONTROL	0,570	0,890	0,700
	SVM	DEHB	0,560	0,900	0,690
		KONTROL	0,670	0,220	0,330
Rastgele Karar Ormanları	KNN	DEHB	0,730	0,920	0,810
		KONTROL	0,750	0,430	0,550
	Naive Bayes	DEHB	1,000	0,360	0,530
		KONTROL	0,530	1,000	0,700
	SVM	DEHB	1,000	0,670	0,800
		KONTROL	0,670	1,000	0,800

Tablo 13. (Devamı)

Modelden Seçim	KNN	DEHB	0,790	0,920	0,850
		KONTROL	0,800	0,570	0,670
	Naive Bayes	DEHB	0,830	0,450	0,590
		KONTROL	0,540	0,880	0,670
	SVM	DEHB	0,670	1,000	0,800
		KONTROL	1,000	0,440	0,620
T-Test	KNN	DEHB	0,670	0,910	0,770
		KONTROL	0,750	0,380	0,500
	Naive Bayes	DEHB	0,750	0,330	0,460
		KONTROL	0,600	0,900	0,720
	SVM	DEHB	0,620	0,800	0,700
		KONTROL	0,670	0,440	0,530

Öznitelik 3 için yapılan uygulamalar sonucunda elde edilen doğruluk değerleri Tablo 14’de, kesinlik, duyarlılık ve F1 skoru değerleri Tablo 15’te verilmiştir. K-Best, rastgele karar ormanları, modelden seçim ve T-Test algoritmalarının uygulanması sonucunda veri setinden sırasıyla 5, 30, 45 ve 81 öznitelik elde edilmiştir. Modelden seçim algoritması uygulanan naive bayes, T-Test algoritması uygulanan KNN ve naive bayes algoritmaları %68,4’lük başarı sağlamıştır. Kesinlik ve duyarlılık değerlerine bakıldığı zaman T-Test uygulanmış KNN algoritmasının tahmin etme konusunda daha başarılı olduğu görülmüştür.

Tablo 14. Öznitelik 3’ün Öznitelik Seçimi Sonrası Sınıflandırma Sonuçlarının Doğruluk Değerleri

Öznitelik Çıkarma Yöntemi	Sınıflandırma Yöntemi	Doğruluk	K-Fold Çapraz Doğrulama
K-Best	KNN	0,368	0,461
	Naive Bayes	0,579	0,570
	SVM	0,526	0,580
Rastgele Karar Ormanları	KNN	0,579	0,597
	Naive Bayes	0,632	0,584
	SVM	0,526	0,530
Modelden Seçim	KNN	0,632	0,650
	Naive Bayes	0,684	0,682
	SVM	0,632	0,630
T-Test	KNN	0,684	0,653
	Naive Bayes	0,684	0,652
	SVM	0,632	0,680

Tablo 15. Öznitelik 3'ün oranlarının öznitelik seçimi sonrası sınıflandırma sonuçları.

Özellik Çıkarma Yöntemi	Sınıflandırma Yöntemi	İşaret	Kesinlik	Duyarlılık	F1 Skoru
K-Best	KNN	DEHB	0,290	0,670	0,400
		KONTROL	0,600	0,230	0,330
	Naive Bayes	DEHB	0,840	0,420	0,560
		KONTROL	0,460	0,860	0,600
	SVM	DEHB	0,530	0,900	0,670
		KONTROL	0,500	0,110	0,180
Rastgele Karar Ormanları	KNN	DEHB	0,500	0,880	0,640
		KONTROL	0,800	0,360	0,500
	Naive Bayes	DEHB	0,670	0,600	0,630
		KONTROL	0,600	0,670	0,630
	SVM	DEHB	0,530	0,900	0,670
		KONTROL	0,500	0,110	0,180
Modelden Seçim	KNN	DEHB	0,530	1,000	0,700
		KONTROL	1,000	0,360	0,530
	Naive Bayes	DEHB	0,700	0,700	0,700
		KONTROL	0,670	0,670	0,670
	SVM	DEHB	0,640	0,700	0,670
		KONTROL	0,620	0,560	0,590
T-Test	KNN	DEHB	0,630	0,890	0,730
		KONTROL	0,830	0,500	0,620
	Naive Bayes	DEHB	1,000	0,450	0,620
		KONTROL	0,570	1,000	0,730
	SVM	DEHB	0,640	0,700	0,670
		KONTROL	0,620	0,560	0,590

3 Boyutlu Yeniden İnşa Sonucu Elde Edilen Uzunlukların Sınıflandırma Sonuçları

2 boyutlu görüntülerin 3 boyutlu olarak yeniden inşa edilmesi sonucunda elde edilen 102 noktanın birbirlerine olan uzaklıklarından elde edilen öznitelik 4'e k en yakın komşu, naive bayes ve destek vektör makineleri algoritmaları uygulanmıştır. Uygulama sonucunda elde edilen doğruluk değerleri Tablo 16'da, kesinlik, duyarlılık ve F1 skoru değerleri Tablo 17'de verilmiştir. Elde edilen veriler incelendiği zaman en yüksek doğruluk oranı %65 ile SVM algoritmasıyla elde edilmiştir.

Tablo 16. Öznitelik 4'ün Sınıflandırma Sonuçlarının Doğruluk Değerleri

Sınıflandırma Yöntemi	Doğruluk	K-Fold Çapraz Doğrulama
KNN	0,409	0,410
Naive Bayes	0,511	0,520
SVM	0,650	0,614

Tablo 17. Öznitelik 4'ün Sınıflandırma Sonuçları

Sınıflandırma Yöntemi	İşaret	Kesinlik	Duyarlılık	F1 Skoru
KNN	DEHB	0,460	0,740	0,570
	KONTROL	0,140	0,050	0,070
Naive Bayes	DEHB	0,510	1,000	0,680
	KONTROL	0,000	0,000	0,000
SVM	DEHB	0,640	0,700	0,670
	KONTROL	0,670	0,600	0,630

SONUÇLAR VE ÖNERİLER

2 boyutlu görüntü üzerinde yapılan çalışma kapsamında üç farklı öznitelik oluşturulmuştur. Öznitelik sayıları sırasıyla 16, 120 ve 8515'tir. Veri setindeki öznitelik sayısının artışına bağlı olarak naive bayes ve k en yakın komşu algoritmalarının başarımının arttığı görülmüştür. Bu doğrultuda benzer çalışmalarda veri setindeki kayıt sayısının artırılması durumunda naive bayes ve k en yakın komşu algoritmalarının kullanımının katkı sağlayacağı söylenebilir. Ek olarak öznitelik sayısındaki artış destek vektör makineleri algoritmasının başarı oranında artışlara ve azalmalara neden olmuştur.

Öznitelik 1 için uygulanan öznitelik seçimi algoritmaları sonrasında üç farklı sınıflandırma işleminden elde edilen doğruluk değerlerinin ortalamaları alınmıştır. Sırasıyla K-Best için 0,649, rastgele karar ormanları için 0,579, modelden seçim için 0,632, T-Test için 0,526 başarı oranları elde edilmiştir. Benzer şekilde öznitelik 2 için uygulanan öznitelik seçimi algoritmaları sonrasında üç farklı sınıflandırma işleminden elde edilen doğruluk değerlerinin ortalamaları alınmıştır. Sırasıyla K-Best için 0,632, rastgele karar ormanları için 0,723, modelden seçim için 0,719, T-Test için 0,649 başarı oranları elde edilmiştir. Son olarak Öznitelik 3 için uygulanan öznitelik seçimi algoritmaları sonrasında üç farklı sınıflandırma işleminden elde edilen doğruluk değerlerinin ortalamaları alınmıştır. Sırasıyla K-Best için 0,491, rastgele karar ormanları için 0,579, modelden seçim için 0,649, T-Test için 0,667 başarı oranları elde edilmiştir. Bu kapsamda Aktaş ve diğerleri tarafından önerilen 16 uzunluğun DEHB tespiti için yetersiz olduğu söylenebilir. Ek olarak öznitelik 2 için rastgele karar ormanları yönteminin destek vektör makineleri algoritmasına uygulanması sonucunda %80'lik başarı oranı elde edilmiştir. Benzer şekilde öznitelik 3 için de modelden seçim yönteminin uygulandığı k en yakın komşu algoritmasında %73,7 başarı oranı elde edilmiştir. Bu bağlamda öznitelik sayısının düşük olduğu veri setlerinde k en yakın komşu algoritmasıyla yapılan uygulamalarda özellik çıkarma yöntemi olarak modelden seçim yönteminin kullanılması, öznitelik sayısının yüksek olduğu veri setlerinde ise destek vektör makineleri algoritmasının kullanıldığı rastgele karar ormanları yönteminin kullanımı sonucunda başarı oranının artırılacağı söylenebilir.

2 boyutlu görüntülerin yanında 2 boyutlu görüntüden 3 boyutlu olarak yeniden inşa edilen görüntülerden oluşturulan verilerin başarı oranlarının test edilmesi amaçlanarak 3 boyutlu verilerden oluşan öznitelik oluşturulmuş ve üç farklı sınıflandırma algoritması uygulanmıştır. K en yakın komşu, naive bayes ve destek vektör makineleri algoritmaları için

sırasıyla 0,409, 0,511, 0,650 başarı oranları elde edilmiştir. Bu başarı oranlarının diğer uygulamalara göre ciddi bir farklılık göstermemesi sebebiyle öznitelik seçimi algoritmaları uygulanmamıştır. 3 boyutlu olarak yeniden inşa edilen görüntülerden ziyade 2 boyutlu görüntülerin öznitelik sayısının artırılması başarı oranlarının yükselmesine daha çok katkı sağlamıştır. Bu durumun 3 boyutlu görüntü oluşturulurken gerçek verilerden ziyade 2 boyutlu görüntülerdeki yüz noktalarını iz düşüm noktası olarak kabul edilmesinden kaynaklandığı söylenebilir.

Bu çalışma kapsamında öznitelik 1 ile yapılan uygulamalarda üç farklı algoritma için elde edilen başarı oranları 0,632, 0,579 ve 0,632'dir. Yüz görüntülerinden hastalık tespitine yönelik yapılan benzer bir çalışmada Lin ve diğerleri koroner kalp hastalığının tespiti için 1528 yüz görüntüsü kullanarak oluşturulan özniteliğe rastgele karar ormanları ve karar ağacı algoritmalarını uygulayarak sırasıyla %72,73, %70,45 başarı oranları elde etmişlerdir (Lin et al., 2019). Elde edilen başarı oranlarının yüksek olmasının sebebi olarak veri setinde kullanılan yüz görüntüsü sayısının fazla olması söylenebilir. Lin ve diğerleri tarafından yapılan çalışmadan farklı olarak bu çalışmada üç farklı sınıflandırma algoritmasının başarı oranları denenmiştir. Benzer çalışmalarda da kullanılan algoritma sayısının artırılması başarımların denemesinde ve performansının artırılmasında katkı sağlayabilir.

Amato ve diğerleri yüz doğruluğunun tespiti için yüz noktaları yerine yüz görüntüsünde derin öğrenme metotların kullanılmasını önermektedir (Amato et al., 2018). Yapılan çalışma kapsamında yüz görüntüsünde derin öğrenme metotları kullanarak %97 başarı oranı elde etmişlerdir. Bu çalışmada elde edilen en fazla başarı oranı %80'dir. Sınıflandırma işlemi için hazırlanan öznitelikler için yapılan uygulamalarda 91 yüz görüntüsü yer almaktadır. Veri setinde bulunan yüz görüntülerinin sayısı artırılarak derin öğrenme metotlarının uygulanması elde edilecek başarı oranını artırılabilir. Ancak kullanılan mevcut veri setindeki yüz görüntüsü sayısının az olması sebebiyle derin öğrenme metotlarının uygulanması başarı oranını artırmayabilir.

Veri setinde bulunan veri sayısının az olmasına karşın sınıflandırma sonuçlarının kabul edilebilir doğruluk oranlarına sahip olduğu söylenebilir. Daha yüksek boyutlu veri setlerine uygulanan makine öğrenmesi yöntemleriyle hastalık teşhisine yönelik sınıflandırma problemlerinin çözümünde önerilen sınıflandırma algoritmalarının kullanımı alana katkı sağlayacaktır.

KAYNAKÇA

- Aktaş, H., Esin, İ. S., & Dursun, O. B. (2020). Is it possible to recognize children diagnosed with ADHD from their facial anthropometric measures? A case-control study. *Medical Hypotheses*, 140.
- Alay, S. (2020, 06 22). *AlgoRithm:Destek Vektör Makineleri*. Data Science for The Earth: <https://www.datascienceearth.com/algorithmdestek-vektor-makinelerisupport-vector-machinesr-kod-ornekli/> adresinden alındı
- Amato, G., Falchi, F., Gennaro, C., & Vairo, C. (2018). A Comparison of Face Verification with Facial Landmarks and Deep Features. *The Tenth International Conference on Advances in Multimedia*. Athens: MMEDIA.
- Ampadu, H. (2021, 05 01). *Random Forests Understanding*. Retrieved from AI Pool: <https://ai-pool.com/a/s/random-forests-understanding>
- Anonymous. (2013). *DSM History*. Retrieved from American Psychiatric Association: <https://www.psychiatry.org/psychiatrists/practice/dsm/history-of-the-dsm>
- Anonymous. (2021, 04 24). *TensorFlow Lite*. Retrieved from TensorFlow: <https://www.tensorflow.org/lite/guide>
- Anonymous. (2021, 09 23). *Symptoms and Diagnosis of ADHD*. Retrieved from Centers for Disease Control and Prevention: <https://www.cdc.gov/ncbddd/adhd/diagnosis.html>
- Balci, M., Tasdemir, S., Ozmen, G., & Golcuk, A. (2022). Machine Learning-Based Detection of Sleep-Disordered Breathing Type Using Time and Time-Frequency Features. *Biomedical Signal Processing and Control*, 73.
- Barkley, R. A., & Peters, H. (2012). The Earliest Reference to ADHD in the Medical Literature? Melchior Adam Weikard's Description in 1775 of "Attention Deficit" (Mangel der Aufmerksamkeit, Attentio Volubilis). *Journal of Attention Disorders*, 623–630.
- Biederman, J., & Faraone, S. V. (2005, Temmuz). Attention-deficit hyperactivity disorder. *Lancet*, 366, 237-248.
- Boyko, N., Basystiuk, O., & Shakhovska, N. (2018). Performance Evaluation and Comparison of Software for Face Recognition, Based on Dlib and Opencv Library. *IEEE Second International Conference on Data Stream Mining & Processing (DSMP)*. Lviv: IEEE.
- Brownlee, J. (2016, 03 16). *Supervised and Unsupervised Machine Learning Algorithms*. Retrieved from Machine Learning Mastery: <https://machinelearningmastery.com/supervised-and-unsupervised-machine-learning-algorithms/>
- Brownlee, J. (2016, 05 20). *Feature Selection For Machine Learning in Python*. Retrieved from Machine Learning Mastery: <https://machinelearningmastery.com/feature-selection-machine-learning-python/>
- Bulat, A., & Tzimiropoulos, G. (2017). How far are we from solving the 2D & 3D Face Alignment problem? (and a dataset of 230,000 3D facial landmarks). *IEEE International Conference on Computer Vision*. Venice: IEEE.
- Cai, H., Guo, Y., Peng, Z., & Zhang, J. (2021). Landmark Detection and 3D Face Reconstruction for Caricature using a Nonlinear Parametric Model. *Graphical Models*, 115.

- Canu, S. (2019, 03 12). *Face landmarks detection – Opencv with Python*. Retrieved from Pysource: <https://pysource.com/2019/03/12/face-landmarks-detection-opencv-with-python/>
- Chandra, M. A., & Bedi, S. S. (2018). Survey on SVM and their application in image classification. *International Journal of Information Technology*, 13(2).
- Chen, T., Antoniou, G., Adamou, M., Tachmazidis, I., & Su, P. (2021). Automatic Diagnosis of Attention Deficit Hyperactivity Disorder Using Machine Learning. *Applied Artificial Intelligence*, 35(9), 657-669.
- Crichton, A. (2008, Kasım). An inquiry into the nature and origin of mental derangement: on attention and its diseases. *Journal of Attention Disorders*, 12(3), 200-204.
- Ekici, S., & Jawzal, H. (2020). Breast cancer diagnosis using thermography and convolutional neural networks. *Medical Hypotheses*, 137.
- Elhassouny, A., & Smarandache, F. (2019). Smart mobile application to recognize tomato leaf diseases using Convolutional Neural Networks. *2019 International Conference of Computer Science and Renewable Energies (ICCSRE)*. Agadir, Morocco: IEEE.
- Fiori, L. (2020, 03 18). *What is like inside a Convolutional Neural Network?* Retrieved from Medium: <https://medium.com/@luigi.fiori.lf0303/what-is-like-inside-a-convolutional-neural-network-a5de21e27883>
- Gallego, A., Rico-Juan, J., & Valero-Mas, J. (2022). Efficient k-nearest neighbor search based on clustering and adaptive k values. *Pattern Recognition*, 122.
- Heni, N., & Hamam, H. (2016). Design of Emotional Educational System Mobile Games for Autistic Children. *2nd International Conference on Advanced Technologies for Signal and Image Processing - ATSIP'2016*. Monastir: IEEE.
- Hoffman, H. (2012, Ekim 14). *The Story of Fidgety Phil*. Retrieved from Dr.Kenny: <https://www.drkenny.com/fidgety-phil>
- Kim, S., Lee, H.-K., & Lee, K. (2021). Can the MMPI Predict Adult ADHD? An Approach Using Machine Learning Methods. *Diagnostics*, 6.
- Kim, T., Mok, J., & Lee, E. (2021). Detecting Facial Region and Landmarks at Once via Deep Network. *Sensors*, 21(16).
- Kosilek, R. P., Frohner, R., Würtz, R., Berr, C., Schopohl, J., Reincke, M., & Schneider, H. (2015). Diagnostic use of facial image analysis software in endocrine and genetic disorders: review, current results and future perspectives. *European Journal of Endocrinology*, 173(4), M39–M44.
- Lange, K. W., Reichl, S., Lange, K. M., Tucha, L., & Tucha, O. (2010, Kasım). The history of attention deficit hyperactivity disorder. *Atten Defic Hyperact Disord*, 2(4), 241–255.
- Laufer, M. W., Denhoff, E., & Solomons, G. (1957). Hyperkinetic Impulse Disorder in Children's Behavior Problems. *Psychosom Med*, 38-49.
- Li, J., Cheng, K., Wang, S., Morstatter, F., Trevino, R., Tang, J., & Liu, H. (2018). Feature Selection: A Data Perspective. *ACM Computing Surveys*, 50(6), 1-45.
- Lin, S., Zhuang, X., Huang, S., Liu, Y., Shen, L., & Liao, X. (2019). Face Analysis for Coronary Heart Disease Diagnosis. *2019 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*. Suzhou: IEEE.
- Liu, W., Li, M., & Yi, L. (2016). Identifying children with autism spectrum disorder based on their face processing abnormality: A machine learning framework. *Autism research : official journal of the International Society for Autism Research*, 9.

- Mahesh, B. (2020). Machine Learning Algorithms - A Review. *International Journal of Science and Research*, 9(1), 381-386.
- Mannuzza, S., Klein, R. G., Bessler, A., Malloy, P., & LaPadula, M. (1998). Adult Psychiatric Status of Hyperactive Boys Grown Up. *The American Journal of Psychiatry*, 493-498.
- Mercan, N., Bozkurt, Ö. Ç., & Oyur, E. (2016, Ekim). İş Yerinde Dikkat Eksikliği ve Hiperaktivite Bozukluğuna (DEHB) Yönelik Bir Araştırma. *İş ve İnsan Dergisi*, 3(2), 109-118.
- Mirzaei, G., & Adeli, H. (2022). Machine learning techniques for diagnosis of alzheimer disease, mild cognitive disorder, and other types of dementia. *Biomedical Signal Processing and Control*, 72.
- Neumärker, K.-J. (2005). The Kramer-Pollnow syndrome: a contribution on the life and work of Franz Kramer and Hans Pollnow. *History of Psychiatry*, 435–451.
- Núñez-Jaramillo , L., Herrera-Solis, A., & Herrera-Morales, W. V. (2021). ADHD: Reviewing the Causes and Evaluating Solutions. *Journal Of Personalized Medicine*, 11.
- Nyein, T., & Oo, A. N. (2019). University Classroom Attendance System Using FaceNet and Support Vector Machine. *2019 International Conference on Advanced Information Technologies (ICAIT)*. IEEE.
- Ozdemir, M. A., Elagoz, B., Alaybeyoglu, A., Sadıghzadeh, R., & Akan, A. (2019). Real Time Emotion Recognition from Facial Expressions Using CNN Architecture. *2019 Medical Technologies National Congress*. Aydın: TIPTEKNO.
- Öner, P., Öner, Ö., & Aysev, A. (2003). Dikkat Eksikliği Hiperaktive Bozukluğu. *Sürekli Tıp Eğitimi Dergisi*, 3(12), 97-99.
- Palmer, E. D., & Finger, S. (2001). An Early Description of ADHD (Inattentive Subtype): Dr Alexander Crichton and ‘Mental Restlessness’. *Child Psychology and Psychiatry Review*, 66 - 73.
- Peng, J., Debnath, M., & Biswas, A. (2021). Efficacy of novel Summation-based Synergetic Artificial Neural Network in ADHD diagnosis. *Machine Learning with Applications*, 6.
- Sahar, A., & Saniee Abadeh, M. (2020). Brain MR Image Classification for ADHD Diagnosis Using Deep Neural Networks. *Iranian Conference on Machine Vision and Image Processing*. Tehran: MVIP.
- Sam , E. (2019, 07 20). *Introduction to kNN algorithm by experiment on Khmer Handwriting classification using Java 8*. Retrieved from Towards Data Science: <https://towardsdatascience.com/introduction-to-knn-machine-learning-algorithm-by-experiment-on-khmer-handwriting-classification-66a64652a02c>
- Saputra, M. F., Widiyaningtyas, T., & Wibawa, A. (2018). Illiteracy Classification Using K Means-Naïve Bayes Algorithm. *JOIV : International Journal on Informatics Visualization*, 2, 153.
- Sarker, I. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, 2.
- Shahid, M. (2019, 02 24). *Convolutional Neural Network*. Retrieved from Towards Data Science: <https://towardsdatascience.com/covolutional-neural-network-cb0883dd6529>
- Sharma, N., Sharma, R., & Jindal, N. (2021). Machine Learning and Deep Learning Applications-A Vision. *Global Transitions Proceedings*, 2(1), 24-28.
- Sim, I. (2019). Mobile Devices and Health. *New England Journal of Medicine*, 956-968.

- Singh, A., & Bhadani, R. (2020). In *Mobile Deep Learning with TensorFlow Lite, ML Kit and Flutter* (p. 27). Packt Publishing.
- Sridhar, C., Bhat, S., Acharya, U. R., Adeli, H., & Bairy, G. M. (2017). Diagnosis of attention deficit hyperactivity disorder using imaging and signal processing techniques. *Computers in Biology and Medicine*, 88, 93-99.
- Still, G. F. (2006). Some Abnormal Psychological Conditions in Children. *Journal of Attention Disorders*, 126-136.
- Tarnowski, P., Kołodziej, M., Majkowski, A., & Rak, R. (2017). Emotion Recognition Using Facial Expressions. *Procedia Computer Science*, 108, 1175-1184.
- Teoh, S. K., Wong, Y., Leong, C., & Tan, L. (2021). Face Detection and Face Re-identification System Using Deep Learning and OpenVINO. *2nd International Conference on Artificial Intelligence and Data Sciences (AiDAS)*. IPOH: IEEE.
- Toğaçar, M., Ergen, B., & Cömert Zafer. (2020). BrainMRNet: Brain tumor detection using magnetic resonance images with a novel convolutional neural network model. *Medical Hypotheses*, 134.
- Wang, K., & Luo, J. (2016). Detecting Visually Observable Disease Symptoms from Faces. *EURASIP Journal on Bioinformatics and Systems Biology*, 1.
- Wilens, T. E., & Spencer, T. J. (2010). Understanding Attention-Deficit/Hyperactivity Disorder From Childhood to Adulthood. *Postgrad Med*, 122(5), 97-109.
- Wirdayanti, Mahmudi, I., Ahsan, A. C., Kasim, A. A., Nur, R., Basalamah, R., & Septiarini, A. (2020). 6th International Conference on Science in Information Technology (ICSITech). *6th International Conference on Science in Information Technology (ICSITech)*. Palu: IEEE.
- Zhang, M.-L., & Zhou, Z.-H. (2007). ML-KNN: A lazy learning approach to multi-label learning. *Pattern Recognition*, 40(7), 2038-2048.
- Zhao, S., King, S., & Kar, D. C. (2018). A Tool for Detection and Analysis of a Human Face for Aesthetical Quality Using Mobile Devices. *Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition*. Athens.

ÖZGEÇMİŞ

Kişisel Bilgiler	
Adı Soyadı:	Satuk Buğrahan ÖZTÜRK
Doğum tarihi:	
Doğum Yeri:	
Uyruğu:	
Adres:	
E-mail:	
Eğitim	
Lise:	Özel Bilkent Lisesi
Lisans:	Anadolu Üniversitesi Bilgisayar Mühendisliği
Yabancı Dil Bilgisi	
İngilizce:	Çok İyi