

**T.C.
İNÖNÜ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**COVID-19 İLİŞKİLİ İNSANİ GELİŞİM SEVİYESİNİN SINIFLANDIRILMASI
İÇİN YENİ BİR DERİN EVRİŞİMSEL SİNİR AĞI MODELİ**

YÜKSEK LİSANS TEZİ

Şeyma GÖKHAN

Bilgisayar Bilimleri Anabilim Dalı

**Tez Danışmanı: Prof. Dr. Celaleddin YEROĞLU
İkinci Tez Danışmanı: Dr. Öğr. Üyesi Gürkan KAVURAN**

Haziran 2022

T.C
İNÖNÜ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

COVID-19 İLİŞKİLİ İNSANİ GELİŞİM SEVİYESİNİN SINIFLANDIRILMASI
İÇİN YENİ BİR DERİN EVRİŞİMSEL SİNİR AĞI MODELİ

YÜKSEK LİSANS TEZİ

Şeyma GÖKHAN
36193619021

Bilgisayar Bilimleri Anabilim Dalı

Tez Danışmanı: Prof. Dr. Celaleddin Yeroğlu
İkinci Tez Danışmanı: Dr. Öğr. Üyesi Gürkan KAVURAN

Haziran 2022

TEŐEKKÜR VE ÖNSÖZ

"FYL-2021-2377" kodlu "COVID19 İliŐkili İnsani GeliŐim Seviyesinin Sınıflandırılması için Yeni Bir Derin EvriŐimsel Sinir Ađı Modeli" baŐlıklı Yüksek Lisans Tezin uygulama aŐamasında vermiŐ oldukları maddi ve manevi destekten dolayı, İnönü Üniversitesi BAP birimine,

Bu tez çalıŐmasının her aŐamasında yardım, öneri, bilgi, tecrübe ve desteklerini esirgemedi beni her konuda yönlendiren danıŐman hocalarım Sayın Prof. Dr. Celaleddin Yerođlu ve Sayın Dr. Öğr. Üyesi. Gürkan KAVURAN'a,

ÇalıŐmalarımda ayrıca tüm hayatım boyunca olduđu gibi bu çalıŐmalarım süresince benden her türlü desteklerini esirgemeyen aileme

teŐekkür ederim.



ONUR SÖZÜ

Yüksek lisans tezi olarak sunduđum “COVID-19 İliřkili İnsani Geliřim Seviyesinin Sınıflandırılması için Yeni Bir Derin Evriřimsel Sinir Ađı Modeli” bařlıklı bu alıřmanın bilimsel ahlak ve geleneklere aykırı dűőecek bir yardıma bařvurmaksızın tarafımdan yazıldıđına ve yararlandıđım bűtűn kaynakların hem metin iinde hem de kaynakada yűntemine uygun biimde gűsterilenlerden oluřtuđunu belirtir, bunu onurumla dođrularım.

řeyma Gűkhan



İÇİNDEKİLER

TEŞEKKÜR VE ÖNSÖZ	i
ONUR SÖZÜ	ii
İÇİNDEKİLER.....	iii
ÇİZELGE DİZİNİ.....	iv
ŞEKİLLER DİZİNİ.....	v
SEMBOLLER VE KISALTMALAR	vi
ÖZET	vii
ABSTRACT	viii
1. GİRİŞ.....	1
2. TEORİK TEMELLER	6
2.1 Derin Öğrenme	6
2.2 Derin Evrişimli Sinir Ağı	6
2.2.1 ResNet50	7
2.2.2 AlexNet.....	8
2.3 Makine Öğrenmesi Temelli Sınıflandırıcılar.....	8
2.3.1 Destek Vektör Makinesi	8
2.3.2 K-En Yakın Komşu Algoritması (K-NN)	10
2.3.3 Kollektif öğrenme.....	11
2.4 İnsani Gelişme Endeksi	11
3. MATERYAL VE METOT	13
3.1 Veri Setinin Tanıtımı ve Özellik Seçimi	13
3.2 SDD ve Skalogram Yöntemleri ile Verilerin Görselleştirilmesi	16
3.3 Hiper Parametre Optimizasyonu	18
4. DENEYSEL UYGULAMALAR	20
4.1 Optimize Edilmiş Hiper Parametrelili Sınıflandırıcı ile COVID-19 ve İGE Arasındaki İlişkinin Belirlenmesi	20
4.2 ResNet-50 tabanlı DESA+DVM ile COVID-19 ve İGE Arasındaki İlişkinin Belirlenmesi.....	22
5. SONUÇLAR VE TARTIŞMA.....	25
KAYNAKLAR.....	27

ÇİZELGE DİZİNİ

Çizelge 2.1: Ülkelerin gruplandırılması için İGE'deki kesme noktaları.....	12
Çizelge 3.1: COVID-19 için özellik listesi.	15
Çizelge 4.1: Elde edilen başarımların kıyaslanması.	22
Çizelge 4.2: Derin özellik kümesi için genel sınıflandırma değerleri.....	24



ŞEKİLLER DİZİNİ

Şekil 2.1: ResNet-50 Mimarisi.....	8
Şekil 2.2: SVM hiper düzlem modeli.....	9
Şekil 3.1: Çalışmanın şematik diyagramı.....	13
Şekil 3.2: Kümülatif doğrulanmış COVID-19 vakaları, 11 Ocak 2022 [50]......	14
Şekil 3.3: Özellik önem grafiği.	16
Şekil 3.4: Arnavutluk ve Demokratik Kongo Cumhuriyeti için kümülatif toplam vakaların, kümülatif toplam ölümlerin ve günlere göre günlük yeni COVID-19 testlerinin sapması.....	17
Şekil 3.5: Düşük ve Yüksek İGE'ye sahip iki farklı ülkenin IPG görselleri.	18
Şekil 4.1: a) Sınıflandırma hatasının, iterasyon sayısına göre değişimi b) Optimize edilmiş DVM için karmaşıklık matrisi	21
Şekil 4.2: a) SVMC'nin derin özniteliklerle karmaşıklık matrisi. b) Derin özellikler için tüm sınıfların AÇK eğrileri.	23

SEMBOLLER VE KISALTMALAR

İGE	: İnsani Gelişme Endeksi
DESA	: Derin Evrişimli Sinir Ağı
SDD	: Sürekli Dalgacık Dönüşümü
DVMS	: Destek Vektör Makinesi Sınıflandırıcısı
DSÖ	: Dünya Sağlık Örgütü
YZ	: Yapay Zeka
ESA	: Evrişimli Sinir Ağı
MKK	: Matthews Korelasyon Katsayısı (İng. MCC: Matthews Correlation Coefficient)
YSA	: Yapay Sinir Ağı
ResNet	: Residual Network
IPG	: İç Penetrasyon Görüntüleri
AÇK	: Alıcı Çalışma Karakteristiği (İng. ROC: Receiver Operating Characteristics)

ÖZET

Yüksek Lisans Tezi

COVID-19 İlişkili İnsani Gelişim Seviyesinin Sınıflandırılması için Yeni Bir Derin Evrişimsel Sinir Ağı Modeli

ŞEYMA GÖKHAN

İnönü Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı

31+VIII sayfa

2022

Danışman: Prof. Dr. Celaleddin Yeroğlu
İkinci Tez Danışmanı: Dr. Öğr. Üyesi Gürkan KAVURAN

Pandemi sırasında alınan önlemlerin insanların yaşamları üzerinde kalıcı etkileri olmuştur. Ama diğer taraftan da ulusal ve çok taraflı kurumların insani gelişmeyi sağlama becerisine ilişkin yeni yönelimler ve yöntemleri geliştirmesine neden olmuştur. Çünkü bu pandemi sağlığı, ekonomiyi, eğitimi, sosyal hayatı ve birikmiş kazanımları etkileyen sistemik bir insani gelişme krizi haline gelmiştir. Bu tez çalışması sağlık, eğitim ve ekonomi üzerinde birleşik etkileri olan İnsani Gelişim Endeksi'nin (İGE) ilişkisinin pandemi faktörleri bağlamında nasıl ele alınması gerektiğini göstermektedir. İlk olarak, kamuya açık ve güvenilir bir kaynaktan alınan ülkelerin COVID-19 verileri ayıklanarak kabul edilebilir bir yapıya kavuşturulmuştur. Ardından, hangi değişkenlerin İGE ile yakından ilişkili olduğunu belirlemek ve Derin Evrişimli Sinir Ağı (DESA) modelinin daha doğru sonuçlar vermesini sağlamak için istatistiksel öznitelik seçimi uygulanmıştır. Zaman serisi veri görselleştirilmesi için Sürekli Dalgacık Dönüşümü (SDD) ve scalogram yöntemleri kullanılmıştır. Her ülkenin üç farklı görüntüsü, işleme kolaylığı için birbirine nüfuz edecek şekilde tek bir görüntüde birleştirilmiştir. Bu görüntüler çeşitli ön işleme süreçlerinden geçirilerek önceden eğitilmiş bir DESA modeli olan ResNet-50 ağının girişi için uygun hale getirilmiştir. Eğitim ve doğrulama işlemlerinden sonra ağın fc1000 katmanındaki öznitelik vektörleri çizilerek Destek Vektör Makinesi Sınıflandırıcısı (DVMS) girişine verilmiştir. Özgüllük (%88,2), duyarlılık (%96,5), kesinlik (%99), F1 Puanı (%94,9) ve MKK (%85,9) toplam performans metriklerine ulaşılmıştır. Ayrıca bayes optimizasyonuna dayalı hiper parametrelili sınıflandırıcı ile COVID-19 ve İGE arasındaki ilişkinin belirlenmesi üzerine bir çalışma yapılmıştır. Yapılan çalışma DESA modeli olan AlexNet ağı kullanılarak oluşturulmuş ve özgüllük (%96), duyarlılık (%72), kesinlik (%92), doğruluk (%85) ve F1 puanı (%80) metriklerine ulaşılmıştır.

Anahtar Kelimeler: İnsani Gelişim İndeksi; Derin Öğrenme; COVID-19; Sürekli Dalgacık Dönüşümü; Yapay Zeka; Sınıflandırma

ABSTRACT

Master Thesis

A New Deep Convolutional Neural Network Model for Classification of COVID-19
Associated Human Development Level

ŞEYMA GÖKHAN

Inonu University
Graduate School of Natural and Applied Sciences
Department of Computer Science

31+VIII sayfa

2022

Supervisor: Prof. Dr. Celaleddin Yeroğlu
Second Supervisor: Asist. Prof. Dr. Gürkan KAVURAN

The measures taken during the pandemic have had lasting effects on people's lives. On the other hand it has led national and multilateral institutions to develop new directions and methods of human development capability. Because, this pandemic become a systemic human development crisis affecting health, the economy, education, social life and accumulated gains. This study shows how the relationship of the Human Development Index (HDI), which has combined effects on health, education and the economy, should be considered in the context of pandemic factors. First, the COVID-19 data of the countries taken from a public and reliable source were extracted and made into an acceptable structure. Then, statistical feature selection was applied to determine which variables are closely related to HDI and to enable the Deep Convolutional Neural Network (DCNN) model to give more accurate results. Continuous Wavelet Transform (CWT) and scalogram methods are used for time series data visualization. Three different images of each country are merged into a single image, permeating each other for ease of processing. These images were made available for the input of the ResNet-50 network, a pre-trained DCNN model, by undergoing various preprocessing processes. After the training and validation processes, the feature vectors in the fc1000 layer of the network were drawn and given to the Support Vector Machine Classifier (SVMC) input. Specificity (88.2%), sensitivity (96.5%), precision (99.9%), F1 Score (94.9%) and Matthews Correlation Coefficient (MCC) (85.9%) total performance metrics were achieved. In addition, a study was conducted to determine the relationship between COVID-19 and the human development index using the hyperparameter classifier based on Bayesian optimization. The study was conducted using the AlexNet network that is one of the DCNN model. Then, specificity (96%), sensitivity (72%), precision (92%), Accuracy (85%), and F1 score (80%) metrics were achieved.

Keywords: Human Development Index; Deep Learning; COVID-19; Continuous Wavelet Transform; Artificial Intelligence; Classification.

1. GİRİŞ

Çin'in Vuhan kentinde ortaya çıkan yeni tip korona virüsün neden olduğu COVID-19 adı verilen hastalık, dünyayı birçok yönden olumsuz etkiledi. Dünya Sağlık Örgütü (DSÖ) tarafından 11 Mart 2020 tarihinde pandemi ilan edildi. 1 Aralık 2021 tarihi itibarıyla dünya genelinde toplam vaka sayısı 262 milyon 178 bin 403, toplam ölüm sayısı ise 5 milyon 215 bin 745 oldu. Türkiye 8 milyon 795 bin 588 vaka ile dünyada altıncı sırada yer alıyor. Türkiye'de ilk doz aşılama 56 milyon 38 bin 139 kişiye yapılırken, ikinci doz 50 milyon 750 bin 190 kişiye uygulandı. Türkiye, Amerika Birleşik Devletleri, Hindistan ve Brezilya'dan sonra en fazla aşı yapılan 6. ülke oldu [1][2][3].

Yoksulluk düzeyi, sağlık hizmetleri kapasitesi, internete erişim ve sosyal koruma gibi göstergeler, COVID19 krizinin etkilerinin 189 ülkenin her birinde ne kadar ağır seyirli alacağını ortaya koymaktadır. Birleşmiş Milletler Kalkınma Programı, ülkelerin COVID-19 kriziyle baş etme ve düze çıkma kapasitelerindeki büyük farkları gösteren iki yeni gösterge tablosu yayınlanmıştır. Kalkınmanın ekonomik ve sosyal boyutlarını eşi benzeri görülmemiş biçimde etkileyen salgın, sistemsel bir insani gelişme krizi halini aldı. Bireyler ve toplumların bunun gibi şokları daha iyi atlatabilmesi ve sonrasında düze çıkabilmesi için, hem kısa hem de uzun vadede, kırılmalıkları azaltan ve krizle mücadele kapasitesini yükselten politikalar hayati önem taşımaktadır. COVID-19 pandemisinin karşı karşıya kalınan bir sağlık acil durumundan daha fazlası olduğu ve dünya ekonomisini etkileyen sistemik bir kriz olduğu açıkça ifade edilmiştir [4]. Dünyanın birçok yerinde pandemi, insani gelişmede bir krize neden oluyor. Günümüz koşulları, insani gelişmenin bazı boyutlarında en son 1980'lerin ortalarında gözlemlenen yoksunluk düzeyleriyle karşılaştırılabilir. Ancak kriz, gelir, sağlık ve eğitim dahil olmak üzere insani büyümenin tüm yönleri üzerinde önemli bir etkiye sahiptir. Salgının artan toplumsal cinsiyete dayalı şiddet üzerinde dolaylı etkileri olduğu bilinmesine rağmen henüz tam olarak belgelenmiş değildir. İnsan sermayesi, uzun vadeli refahın, yoksulluğun azaltılmasının ve müreffeh toplumların kritik bir itici gücüdür [5]. Eğitim ve öğretim, işgücü verimliliğini artıran politikalar, gelecekteki gelir seviyelerini yükselten her türlü faaliyet insan sermayesi ile bağlantılı olmalıdır. Bu bağlamda sağlıkla ilgili harcamalar da beşeri sermaye yatırımları olarak görülmektedir [6]. COVID-19 ile mücadele kapsamında geliştirilen ve Yapay Zeka (YZ) kullanılarak yapılan modeller, bu yatırımların başlangıcında hastaların erken teşhis ve tedavisini kolaylaştırmaya başlamıştır.

Yapay zeka ile ilgili çalışmaların günümüzde bilimin birçok alanında uygulanabilir olması, pratik ve sağlam bir teorik altyapıya sahip olduğunun göstergesidir. Covid-19 salgını 2019'un Aralık ayından itibaren görülmeye başladığından beri bu konuda bilimsel çalışmalarda yoğun bir şekilde yapılmaya başlanmıştır. Narin ve diğ., yaptıkları çalışmada koronavirüs pnömonisi ile enfekte olmuş hastanın göğüs röntgeni radyografileri kullanılarak saptanması için üç farklı evrişimli sinir ağı tabanlı model (ResNet50, InceptionV3 ve InceptionResNetV2) önerilmiştir. Elde edilen sonuçlara göre önceden eğitilmiş ResNet50 modelinin diğer iki modele göre % 98 doğrulukla en yüksek sınıflandırma performansını sağladığı görülmektedir. Yüksek performans sayesinde doktorların klinik pratikte karar vermesini kolaylaştıracağı öngörülmüştür [7]. Gozes ve diğ., yaptıkları çalışmada koronavirüs hastalarını hasta olmayanlardan ayırt edebilmek için ısı haritası veya 3D hacim ekranında görselleştirilmesini sağlamışlardır. Bu sayede koronavirüsün saptanmasında ve hastalık yükünün ölçülmesinde ve izlenmesinde yüksek doğruluk elde edebileceğini göstermişlerdir [8]. S. Hu ve diğ., çalışmalarında BT görüntülerini kullanarak COVID-19 enfeksiyonunun hızlı ve tam olarak tespiti ve sınıflandırılması için zayıf denetimli bir derin öğrenme çerçevesi tasarlamışlardır. Önerilen modelin sınıflandırma için yüksek doğruluk, kesinlik ve lezyon tespitleri için umut verici niteliksel görselleştirme sağladığını göstermişlerdir [9]. Hemdan ve diğ., yapmış oldukları çalışmada önerilen COVIDX-Net taslağına dayalı olarak X-ışını görüntülerinde COVID-19'u sınıflandırmak için derin öğrenme modellerinin yararlı uygulamasını göstermişlerdir [10]. Kim ve diğ., COVID-19 vakalarının sayısını tahmin ederek bu tür bir tahsise yardımcı olmak için coğrafi hiyerarşiden yararlanan Hi-COVIDNet'i önermişlerdir [11]. Uçar ve diğ., çok pratik bir derin öğrenme modeli çatısı altında girdi X-ışını görüntülerinin tanınmasını içeren YZ tabanlı bir karar verme sistemi önermişlerdir. X-ışını görüntülerinden COVID-19 teşhisi için derin Bayes-SqueezeNet karar verme sistemini uygulamışlardır [12]. Ahsan ve diğ., COVID-19 semptomları olan hastalarla diğer hastalar arasında sınıflandırma yapmak için BT taraması (400 örnek) ve göğüs röntgeni (400 örnek) görüntü veri seti dikkate alınarak kapsamlı bir çalışma gerçekleştirmişlerdir. Mevcut derin öğrenme modellerinin göğüs röntgeni görüntülerinde BT tarama görüntülerine kıyasla daha iyi performans gösterdiğini ortaya koymuşlardır. Sekiz farklı derin öğrenme modeli (VGG16, InceptionResNetV2, ResNet50, DenseNet201, VGG19, MobileNetV2, ResNet15V2 ve NasNetMobile) kullanılmış ve bu modeller içerisinde MobileNetV2 ve NasNetMobile daha iyi performans gösterdiklerini ortaya koymuşlardır [13]. Jiang ve diğ., yapılan çalışmada gerçekçi BT görüntüleri oluşturabilen cGAN tabanlı bir COVID-19 BT görüntü sentezi yöntemi önermişlerdir.

Önerilen yöntemin BT görüntüleri üretebildiğini ve yaygın görüntü sentezi yaklaşımlarıyla karşılaştırıldığında görüntü kalitesi açısından son teknoloji performansa ulaşabildiğini göstermektedir [14]. Waheed ve diğ., bu çalışmada CovidGAN adlı bir Yardımcı Sınıflandırıcı Üretken Karşıt Ağ tabanlı model geliştirerek sentetik göğüs röntgeni görüntüleri oluşturmak için bir yöntem sunmuşlardır. Tek başına Evrişimli Sinir Ağı (ESA) kullanılarak yapılan sınıflandırma, % 85 doğruluk sağlamışken CovidGAN tarafından üretilen sentetik görüntüler eklenerek doğruluk % 95'e çıkarılmıştır [15]. Nour ve diğ., çalışmalarında günlük klinik uygulamaları desteklemek için pozitif COVID-19 vakalarını otomatik olarak tespit edebilen bilgisayar destekli bir model önerilmiştir. Önerilen modelde ESA mimarisine dayanmakta ve göğüs röntgeni görüntülerinde ayırt edici özellikleri otomatik olarak ortaya çıkarabilmektedir. Bu çalışma sonucunda ucuz, hızlı, ve güvenilir teşhis aracı elde edilmiş olup % 98.97 doğruluk,% 89.39 duyarlılık,% 99.75 özgüllük ve % 95.75 F skoru sağlanmıştır [16]. L. Zhou ve diğ., farklı kaynaklardan BT taramalarında enfeksiyon bölgelerini bölümlere ayırabilen ve ölçebilen tam otomatik, hızlı, doğru ve makeden bağımsız bir yöntem önermişlerdir [17]. Panwar ve diğ., bu çalışmada göğüste bulunan görsel göstergeleri arayacak hastaların röntgenlerini analiz ederek COVID-19'u tespit etmek için kullanılacak alternatif bir hızlı tarama yöntemi olan derin öğrenme sinir ağı tabanlı bir yöntem olan nCOVnet'i önermişlerdir. Önerilen yöntem ile pozitif olan hastayı 5 saniyenin altında tespit edebilmektedir. Bu sayede hızlı bir şekilde tıbbi uzmanlara yardımcı olabileceğini öne sürmüşlerdir [18]. Ezzat ve diğ., bu çalışmada bir optimizasyon algoritması kullanılarak hibrit ESA mimarisine dayalı GSA-DenseNet121-COVID-19 adlı yeni bir yaklaşım önermişlerdir. Önerilen yöntem %98 doğruluk payı içermiştir bu sayede COVID-19 adlı yeni virüsün teşhisinde etkinliğini gösterilmiştir [19]. Chen ve diğ., çalışmalarında COVID-19 pnömonisini yüksek çözünürlüklü BT'de tespit etmek için derin öğrenmeye dayalı bir sistem kurmayı, radyologların çalışma baskısını azaltmayı ve salgının kontrolüne katkıda bulunmayı amaçlamışlardır. Bu yöntem sonucunda radyologların okuma süresi % 65 azaltılmıştır [20]. Maghdid ve diğ., yapılan çalışmada yerleşik akıllı telefon sensörlerini kullanarak COVID-19'u tespit etmek için yeni bir taslak önerilmiştir. Herkes tarafından kullanılabilmesi ve hızlı, etkili olması amaçlanmıştır [21]. Das ve diğ., çalışmalarında COVID-19 vakalarını tespit etmek için umut verici iki aşamalı bir strateji geliştirilmiş, derin öğrenme ağları kullanılarak bakteriyel pnömoni, viral pnömoni ve sağlıklı normal insanlardan göğüs röntgen görüntüleri ile ayırmak için önerilmiştir [22]. Rehman ve diğ., bu çalışma ile COVID-19 hastalığının viral pnömoni, bakteriyel pnömoni ve sağlıklı vakalardan derin öğrenme teknikleri kullanılarak ayrıştırılarak etkili bir teşhisi

geliştirmişlerdir. Farklı ESA mimarilerini karşılaştırmışlar ve bu sayede koronavirüse karşı erken tanı sağlanabilmektedir [23]. Boudrioua ve diğ., çalışmada üç evrişimli sinir ağını karşılaştırmışlardır. DenseNet 121, NASNetLarge ve NASNetMobile modellerinin hassasiyet oranlarının yüksek olması sayesinde COVID-19 enfeksiyonlarının otomatik teşhisi için yardımcı olabileceğini öne sürmüşlerdir [24]. Vaid ve diğ., çalışmada COVID-19 vakalarının doğruluğunun tespiti için derin öğrenme modeli geliştirmişlerdir. Göğüs röntgeni taramalarının kesin olarak tahmin edilmesi istenmiştir. Pozitif vaka ve negatif vakaları net bir şekilde gösterildiğini ortaya koymuşlardır [25].

Yapay zeka, nüfus taraması, tıbbi yardım, bildirim ve enfeksiyon yönetimi önererek bu virüsle mücadelede bize yardımcı olabilir [26]. Son zamanlarda, YZ uygulamalarının COVID-19 pnömonisi tanısındaki etkinliği bazı yeni çalışmalarda değerlendirilmiş ve bu alanda yüksek tanısal performansa sahip olduğu bulunmuştur [27][28][29][30][31]. Literatürde COVID-19 ile ilgili YZ ve tıbbi uygulamaların kombinasyonu ile ilgili birçok çalışma bulunmaktadır. Örneğin, Mount Sinai'deki araştırmacılar, Çinli hastalardan alınan 900'den fazla BT akciğer taraması üzerinde eğitilmiş bir algoritma kullanarak COVID-19 hastalarını teşhis etmeye yardımcı olacak bir YZ sistemi oluşturmuşlardır[32]. MIT'deki araştırmacılar tarafından geliştirilen başka bir YZ sistemi, bir hastanın öksürük sesine dayanarak COVID-19'u tespit edebilmiştir [33]. DARI algoritması kullanılarak oluşturulan DL-CRC modelinin ve ESA modelinin X-ray ile doğru bir şekilde tespit edilebildiği gösterilmiştir [34]. ESA ve ConvLSTM'yi iki derin öğrenme yöntemi olarak kullanarak, COVID-19 ve normal vakaların hem BK hem de X-ray görüntülerinde test edilebileceği gösterilmiştir. Diğer ilgili enfeksiyonlar arasında COVID-19'u tespit etmek için etkili bir teşhis konulabilmiştir [35]. ESA tabanlı transfer öğrenme-BiLSTM hibrit yapısının COVID-19'u teşhis etmede oldukça etkili olduğu gösterilmiştir. Önerilen çalışma, segmentasyon işlemini otomatik olarak gerçekleştirdiğinden, yüksek sınıflandırma doğruluğu ve kolaylık sağlamaktadır [36]. Yang ve meslektaşları ESA kullanarak 152 BT görüntüsünü analiz ettiler ve COVID-19 teşhisinde 0,89 doğruluk göstermişlerdir [37]. Ahamed ve arkadaşları, COVID-19 bulaşmış vakaları tespit etmek ve teşhis etmek için ince ayarın yapıldığı, ResNet50V2 tabanlı genişletilmiş bir derin öğrenme modeli geliştirmişlerdir [38]. Kavuran ve arkadaşları, [32]'de torasik bilgisayarlı tomografi kullanarak COVID-19 tanısı için tam otomatik bir derin öğrenme sistemini değerlendirmiştir. Birleştirilmiş katmanlardan optimize edilmiş özelliklerle genel olarak %97,7 doğruluğa ulaşmışlardır.

Bu tez çalışması kapsamında ülkelerin İGE'si ile o ülkenin COVID-19 verileri arasındaki ilişkiyi bulmak için makine öğrenmesi tabanlı bir karar destek sistemi önerilmiştir. Zamanla değişen bu veriler temel olarak günlük vaka sayısı, test sayısı, ölüm oranı, aşılama oranı ve sosyo-demografik bilgiler gibi yaklaşık 35 değişkeni içermektedir. Öncelikle bu değişkenlerden hangilerinin İGE ile yakından ilişkili olduğunu belirlemek ve DESA modelinin daha doğru sonuçlar vermesine olanak sağlamak için istatistiksel öznitelik seçimi uygulanmıştır. ANOVA f değerine göre her ülke için en yüksek üç puana sahip değişken dikkate alınarak çalışmaya başlanmıştır. Zaman serisi verilerinin görselleştirilmesi için SDD yöntemi ve skalogram yöntemi kullanılmıştır. Her ülkenin üç farklı görüntüsü, işleme kolaylığı için birbirine nüfuz edecek şekilde tek bir görüntüde birleştirilmiştir. Bu görüntüler çeşitli ön işleme süreçlerinden geçirilerek önceden eğitilmiş bir DESA modeli olan ResNet-50 ağının girişi için uygun hale getirilmiştir. Ağın öğrenilebilir evrişim katmanındaki (fc1000) ağırlık katsayıları revize edilmiş ve sınıflandırma performansı artırılmıştır. Eğitim ve doğrulama işlemlerinden sonra ağın fc1000 katmanındaki öznitelik vektörleri çizilerek diğer bir makine öğrenmesi tabanlı sınıflandırıcı olan SVMC girdisine verilmiştir. Sistem performansı, karışıklık matrisi ve Alıcı Çalışma Karakteristik (AÇK) eğrilerine dayalı olarak doğruluk, duyarlılık, özgüllük, F1 ve Matthew Korelasyon Katsayısı (MKK) gibi metrikler üzerinde ölçülmüştür. Ayrıca yeni bir çalışma da bayes optimizasyonuna dayalı hiper parametrelili sınıflandırıcı ile COVID-19 ve İGE arasındaki ilişkinin belirlenmesi üzerine yapılmıştır. Yapılan çalışma DESA modeli olan AlexNet ağı kullanılarak oluşturulmuştur. Bu çalışmada da Özgüllük, duyarlılık, kesinlik, Doğruluk ve F1 Puanı metrikleri dikkate alınarak değerlendirme yapılmıştır.

2. TEORİK TEMELLER

Bu bölümde, tez çalışmasında kullanılan teorik bilgiler kısaca özetlenmiş ve tez çalışmasındaki kullanım gerekçelerine değinilmiştir.

2.1 Derin Öğrenme

Yapay zeka, makinelerin insanlar gibi düşünebilme ve hareket edebilmesi için programlanmasını ifade eder. Makinelerin kendi kendilerine öğrenebilecekleri, durumlara uyum sağlayabilecekleri ve kendi hatalarını kendilerinin çözebileceği alanları kapsamaktadır. Makine Öğrenmesi verileri ön işlemden geçirerek tahminlerde bulunmak ve sınıflandırmaktır. Derin öğrenme ise önceden verilmiş veriyi öğrenip farklı verilerle karşılaştırabilen bir YZ uygulamasıdır.

Derin öğrenme, sinyal veya görüntü niteliklerini önceden belirlemeden YZ kullanarak matematiksel modeller geliştirme yöntemidir. Algoritmalara girdi olarak verilen eğitim verileri ile verileri tanımlayabilir, tahmin yapabilir, sınıflandırabilir ve nasıl iyileştirilmesi gerektiğini belirleyebilen makine öğrenmesinin bir alt dalıdır. Birden çok sayıda katman ve parametre içeren bir sinir ağıdır. Sinir ağı mimarileri kullanan bir çok derin öğrenme yöntemi vardır. Bu nedenle derin sinir ağları olarak da adlandırılır.

Derin öğrenme, çok sayıda doğrusal olmayan işlem birimi katmanı kullanarak özellik çıkarımı ve dönüştürme işlemi yapar. Alt katmanlar basit özellikleri öğrenirken, üst katmanlar alt katman özelliklerinden türetilen daha karmaşık özellikleri öğrenir. Derin öğrenme hem büyük miktardaki verilerden hem de farklı kaynaklardan toplanan verilerden yararlı bilgileri analiz etmek ve çıkarmak için uygun bir YZ uygulamasıdır [39].

Derin sinir ağları, tipik olarak, derinlemesine iç içe ağ mimarilerinde düzenlenen birden fazla gizli katmandan oluşur. Basit YSA'ların aksine genellikle basit bir aktivasyon fonksiyonu kullanmak yerine bir nöronda gelişmiş işlemler veya çoklu aktivasyonlar kullanabilen gelişmiş nöronlar içerirler. Bu özellikler, derin sinir ağlarının ham girdi verileriyle beslenmesine ve ilgili öğrenme görevi için gerekli olan bir temsili otomatik olarak keşfetmesine olanak tanır [40].

2.2 Derin Evrişimli Sinir Ağı

Evrişimli sinir ağı, görüntüleri analiz etmek için kullanılan en yaygın yapay sinir ağı sınıfıdır. ESA, makine öğrenimi problemlerinde mükemmel bir performansa sahiptir. Özellikle en büyük görüntü sınıflandırma veri seti olan Image Net, bilgisayarla

görme ve doğal dil işleme gibi görüntü verileriyle ilgilenen uygulamalarda iyi bir performansa sahiptir. Konvolüsyon adı verilen matrisler arasındaki matematiksel doğrusal işleminden dolayı evrişimli sinir ağı denmektedir. ESA'nın birden fazla katmanı vardır. Bunlar evrişim katmanı, doğrusal olmayan katman, havuz katmanı ve tam bağlantılı katmanlardır.

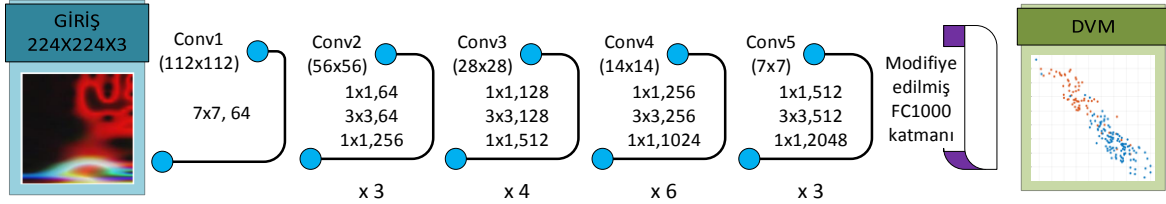
ESA, tipik olarak evrişim, havuzlama ve tamamen bağlantılı katmanlardan oluşan matematiksel bir yapıdır. Evrişim ve havuzlama katmanları, öznelik çıkarma işlemini gerçekleştirirken, tamamen bağlantılı olan katman, çıkarılan öznelikleri sınıflandırma gibi nihai çıktıyla eşler. Bir evrişim katmanı, özel bir doğrusal işlem türü olan evrişim gibi bir dizi matematiksel işleminden oluşan ESA önemli bir rol oynar. Dijital görüntülerde, piksel değerleri iki boyutlu(2B) bir ızgarada, yani bir dizi sayı içinde saklanır ve çekirdek adı verilen küçük bir parametre ızgarası, optimize edilebilir bir özellik çıkarıcı, her görüntü pozisyonuna uygulanır. Bu da ESA'ları görüntü işleme için oldukça verimli kılar, çünkü özellik görüntünün herhangi bir yerinde ortaya çıkabilir[41].

2.2.1 ResNet50

ResNet (Residual Network), Shaoqing Ren, Kaiming He, Jian Sun ve Xiangyu Zhang tarafından 2015 yılında yayınlanan makalelerinde tanıtılan ünlü derin öğrenme modellerinden biridir [42]. Artık ağ olarak bilinir. ResNet modeli, şimdiye kadar modellerden en popüler ve en başarılı derin öğrenme modellerinden biridir. ILSVRC 2015 sınıflandırma yarışmasında en üst sırayı almış ve 2015 ILSVRC & COCO yarışmalarında ImageNet tespiti ve lokalizasyonu, COCO tespiti ve segmentasyonunda da birinci olmuştur. ResNet'in farklı katmana sahip birçok çeşidi vardır.

ESA ağlarının performans düşümü (degradation) problemini ResNet ağları çözmeyi hedeflemektedir. Derin ağlar yakınsamaya başladığı zaman, bozulma problemi ortaya çıkar. Ağ derinliği arttıkça, verimliliği (accuracy) doyunluğa ulaşır ama daha sonra hızlı bir düşüş eğilimi gösterir. ResNet, girişte herhangi bir değişiklik yapmadan önceki katmandan sonraki katmana girişi sığdırmak için bir atlama bağlantısı (veya kısayol bağlantısı) sunarak, 152 katmana kadar çok derin bir ağa sahip olabilir. ResNet uygulamasında iki tür kısayol modülü vardır. Birincisi, kısayolda evrişim katmanı olmayan bir kimlik bloğudur. Bu durumda girdi, çıktı ile aynı boyutlara sahiptir. Diğeri, kısayolda evrişim katmanına sahip olan evrişim bloğudur. Bu durumda girdi boyutları çıktı boyutlarından daha küçüktür. Her iki blokta da ağın başına ve sonuna 1×1 evrişim katmanı eklenir. Bu, ağın performansını çok fazla düşürmeden parametre sayısını azaltan darboğaz tasarımı adı verilen bir tekniktir [43].

ResNet-50 modeli, her biri bir evrişim ve Kimlik bloğu içeren 5 aşamadan oluşur. Her evrişim bloğu 3 evrişim katmanına sahiptir ve her kimlik bloğu ayrıca 3 evrişim katmanına sahiptir. ResNet-50, 23 milyonun üzerinde eğitilebilir parametreye sahiptir. Resnet50, 50 sinir ağı katmanıyla çalışabilen varyantı belirtmek için kullanılır. Kısayol bağlantıları iki katman atlarken ResNet-50 için üç katman atlanmıştır [44].



Şekil 2.1: ResNet-50 Mimarisi

2.2.2 AlexNet

2012'de Alex Krizhevsky ve meslektaşları tarafından önerilen Alexnet büyük ölçekli görsel tanıma yarışmasını kazanmıştır. ImageNet LSVRC-2010 yarışmasındaki 1,2 milyon yüksek çözünürlüklü görüntüyü 1000 farklı sınıfa sınıflandırmak için büyük, derin bir evrişimsel sinir ağı eğitilmiştir.

Model Imagenet veri setinde eğitilir. Bin sınıfta yaklaşık 14 milyon görüntüye sahip olan Imagenet veri seti üzerinde eğitilmiştir. Her bir görüntü 227 X 227 X 3 boyutundaki RGB görüntülerdir. Alexnet, öğrenilebilir parametrelere sahip sekiz katmana sahiptir. Model, maksimum havuzlama ve ardından 3 tam bağlantılı katman kombinasyonu ile beş katmandan oluşur ve çıktı katmanı hariç bu katmanların her birinde ReLU aktivasyonunu kullanır. ReLU 'yu bir aktivasyon işlevi olarak kullanmanın, eğitim sürecinin hızını neredeyse altı kat hızlandırdığını keşfetmişlerdir. Ayrıca, modellerinin fazla takılmasını önleyen bırakma (dropout) katmanlarını da kullanılmıştır [45].

2.3 Makine Öğrenmesi Temelli Sınıflandırıcılar

2.3.1 Destek Vektör Makinesi

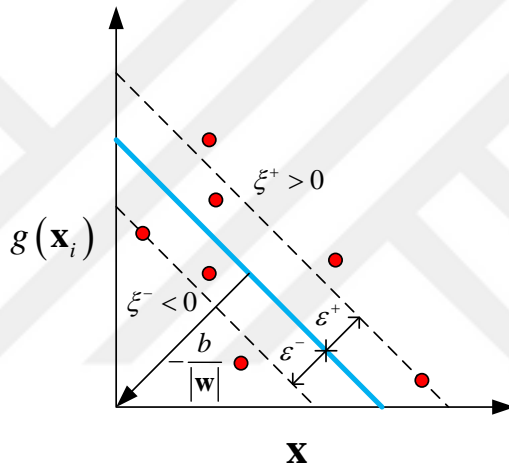
Destek Vektör Makinesi Vapnik tarafından ortaya konulan ikili sınıflandırıcı yapılarından biridir [46], [47]. DVM'nin iki ana fikri vardır, karşıt sınıfların en yakın veri noktaları arasındaki mesafeyi maksimize etmek ve doğrusal olmayan girdi vektörünü yüksek boyutlu bir özellik uzayına eşleyerek onları doğrusal olarak ayırmaktır [47]. DVM, bu amacı gerçekleştirmek için en uygun hiper düzlemleri kullanır. Örneğin, $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ $y_i \in \{-1, 1\}$, $x \in \mathbb{R}^m$ ve $1 \leq i \leq n$ olan eğitim veri setini ele

alalım. Tüm eğitim verileri için Denklem 1'deki eşitsizlikleri doğrulayacak herhangi bir w vektörü ve sabit b çifti varsa, eğitim verileri doğrusal ayrılabilir olarak kabul edilir ve Denklem 1'de verilen eşitsizlikler Denklem 2 şeklinde yazılabilir [46]

$$\begin{aligned} w \cdot x_i + b &\geq 1, & y_i &= 1 \\ w \cdot x_i + b &\leq -1, & y_i &= -1 \end{aligned} \quad (1)$$

$$y_i * (w \cdot x_i + b) \geq 1, \quad 1 \leq i \leq n \quad (2)$$

Denklem 3'te tanımlanan optimal hiperdüzlem, iki sınıfın eğitim verilerini maksimum marjla ayırır. [47]'de marjı belirleyen az sayıda eğitim verisinin optimal bir hiperdüzlem oluşturmak için yeterli olduğu bildirilmektedir. Bu eğitim verileri destek vektörleri olarak adlandırılır. Şekil 2.2 optimal hiperdüzlem ve destek vektörlerini gösterir.



Şekil 2.2: SVM hiper düzlem modeli

Eğitim verileri, herhangi bir yanlış sınıflandırılmış veri olmadan optimal bir hiperdüzlem tarafından sınıflandırılırsa, bir test verisindeki bir hata olasılığının tahmin edilebilir değeri, destek vektörlerinin sayısının beklenti değeri ile eğitim verisi sayısı arasındaki oran ile sınırlıdır. Bu nedenle, öznelik uzayı boyutları ne olursa olsun, destek vektörlerinden optimal hiperdüzlem oluşturulabilirse, genelleme yeteneği yüksek olacaktır [46], [47]

$$g(x_i) = w x_i + b \quad (3)$$

Denklem 4, Denklem (6)'da verilen Lagrange çarpanları ile SVM regresyonu için hata fonksiyonunu sunar.

$$C \sum_{i=1}^L (\xi_i^+ + \xi_i^-) + \frac{1}{2} \|w\|^2, \quad \alpha_i^+ \geq 0, \alpha_i^- \geq 0, \mu_i^+ \geq 0, \mu_i^- \geq 0 \forall_i \quad (4)$$

$$L_P = C \sum_{i=1}^L (\xi^+ + \xi^-) + \frac{1}{2} \|w\|^2 - \sum_{i=1}^L (\mu_i^+ \xi^+ + \mu_i^- \xi^-) - \sum_{i=1}^L \alpha_i^+ (\varepsilon + \xi^+ + y_i - t_i) - \sum_{i=1}^L \alpha_i^- (\varepsilon + \xi^- - y_i + t_i) \quad (5)$$

$$\max \left[\begin{array}{l} \sum_{i=1}^L (\alpha_i^+ - \alpha_i^-) t_i - \varepsilon \sum_{i=1}^L (\alpha_i^+ - \alpha_i^-) - \frac{1}{2} \sum_{i,j} (\alpha_i^+ - \alpha_i^-) (\alpha_j^+ - \alpha_j^-) \\ - \sum_{i=1}^L (\mu_i^+ \xi^+ + \mu_i^- \xi^-) - \sum_{i=1}^L \alpha_i^+ (\varepsilon + \xi^+ + y_i - t_i) x_i \cdot y_j \end{array} \right]$$

$$0 \leq \alpha_i^+ \leq C, 0 \leq \alpha_i^- \leq C, \sum_{i=1}^L (\alpha_i^+ - \alpha_i^-) = 0 \quad \forall_i \quad (6)$$

$$w = \sum_{i=1}^L (\alpha_i^+ - \alpha_i^-) x_i \quad (7)$$

Denklem 8 ve 9, i indekslerini belirleyerek destek vektörlerini ayarlamak için kullanılabilir.

$$b = \frac{1}{N_s} \sum_{s \in S} [t_i - \varepsilon - \sum_{m=1}^L (\alpha_i^+ - \alpha_i^-) x_i \cdot x_m] \quad (8)$$

$$y' = \sum_{i=1}^L (\alpha_i^+ - \alpha_i^-) x_i \cdot x' + b \quad (9)$$

2.3.2 K-En Yakın Komşu Algoritması (K-NN)

Bu yaklaşım, mevcut her örneği kaydeden ve yeni durumları bir benzerlik ölçüsüne göre sınıflandıran basit bir algoritmadır. Bu algoritma, sınıflandırmasını Manhattan, Öklid ve Minkowski uzaklık fonksiyonları gibi bir benzerlik ölçüsüne dayalı olarak gerçekleştiren örnek tabanlı öğrenmedir.

KNN yöntemi, en yakın komşularını belirlemek ve sınıflandırma sonucunu üretmek adına, test örneği ile tüm eğitim verileri arasındaki mesafeleri hesaplayarak her eğitim örneğinde doğrudan bir arama yapar. Diğer bir deyişle, diğer tüm öğrenme yöntemleri de ilk eğitim verisi olarak bir grup örnekle başladıkları için örnek tabanlıdır. Bu nedenle, örnek temelli öğrenenler, örnekleri bir kural grubu veya karar ağacı çıkarmak için kullanmak yerine, öğrenilen her şeyi temsil etmek için örnekler kullanılır.

Uzaklık fonksiyonlarının en yaygın olarak kullanılanı Öklid uzaklığıdır. Bu nedenle, bu, özelliklerin normalize edildiği ve eşit derecede önemli olduğu varsayımını verir. Eğitim grubundan ilgili öznelik ağırlıklarının türetilmesi, örnek tabanlı öğrenme türündeki temel sorunlardan biridir. Bu yöntemde, örnekler aslında verilerdeki kalıpların bir tanımını

vermez. Bununla birlikte, bir sınıfı diğerinden ayıran, bir tür doğrudan bilgi temsili olan örnek uzayında sınırları koymak için örnekler mesafe ölçüleriyle birleştirilir.

2.3.3 Kollektif öğrenme

Topluluk öğrenimi, farklı öğrenme algoritmalarını birleştiren etkili bir tekniktir. Tek bir veri setinden bir hipotezi öğrenmeye çalışan sıradan makine öğrenmesi yaklaşımlarının aksine, aynı sorunu çözmek için birden fazla alt kümenin üretildiği ve birçok hipotezin eğitildiği geniş bir makine öğrenme yöntemidir. Topluluk yöntemleri, bir dizi farklı hipotez oluşturmayı ve bunları sınıflandırma performansını iyileştirmek için birleştirmeyi amaçlar. Ampirik olarak, daha iyi tahmin performansı elde etmek için birden fazla alt küme oluşturmak, farklı özellikleri ve çeşitli modelleri birleştirmek daha uygundur. Bu, bazı öğrencilerin entegrasyonu ile daha yüksek doğruluk elde edildiği anlamına gelir. Bazı topluluk teknikleri eğitim verilerinin aşırı uydurma ve varyansı ile ilgili sorunları azaltmaktadır.

2.4 İnsani Gelişme Endeksi

İnsani Gelişme Endeksi, Birleşmiş Milletler tarafından derlenen ve bir ülkenin insan gelişiminin üç temel boyutundaki ortalama başarısını ölçmek için kullanılan bir ölçüdür. Bunlar uzun ve sağlıklı bir yaşam, bilgi ve iyi bir yaşam standardıdır. İGE, amaçların araçlardan daha önemli olduğunu vurgulayan, insan refahını anlamaya yönelik “yetenekler” yaklaşımını somutlaştırır. İGE’de, yaşam beklentisi, okuryazarlık, okul kaydı ve gelir için bileşen endeksleri, ülkeler arasında insan refahı düzeyini karşılaştırmak veya bir ülkenin zaman içindeki ilerlemesini izlemek için kullanılacak tek bir endekste birleştirilir. İGE, bir ülkenin kalkınmadaki ilerlemesini kişi başına milli gelire dayalı olarak değerlendirmeye yönelik halen yaygın olan uygulamaya bir alternatif sunar [48]. İGE, üç boyutlu indeksin [49] geometrik ortalaması olarak hesaplanır,

$$HDI = (I_{Sağlık} \cdot I_{Eğitim} \cdot I_{Gelir})^{1/3} \quad (10)$$

2014 İnsani Gelişme Raporu’nda, dört insani gelişme başarısı kategorisi için önceden belirlenmiş bir sınır kriterleri sistemi uygulanmıştır. Bu Rapor, Çizelge 2.1’de olduğu gibi 2014 Raporunda açıklanan ülkeleri sınıflandırmak için aynı İGE kesme parametrelerini vermektedir.

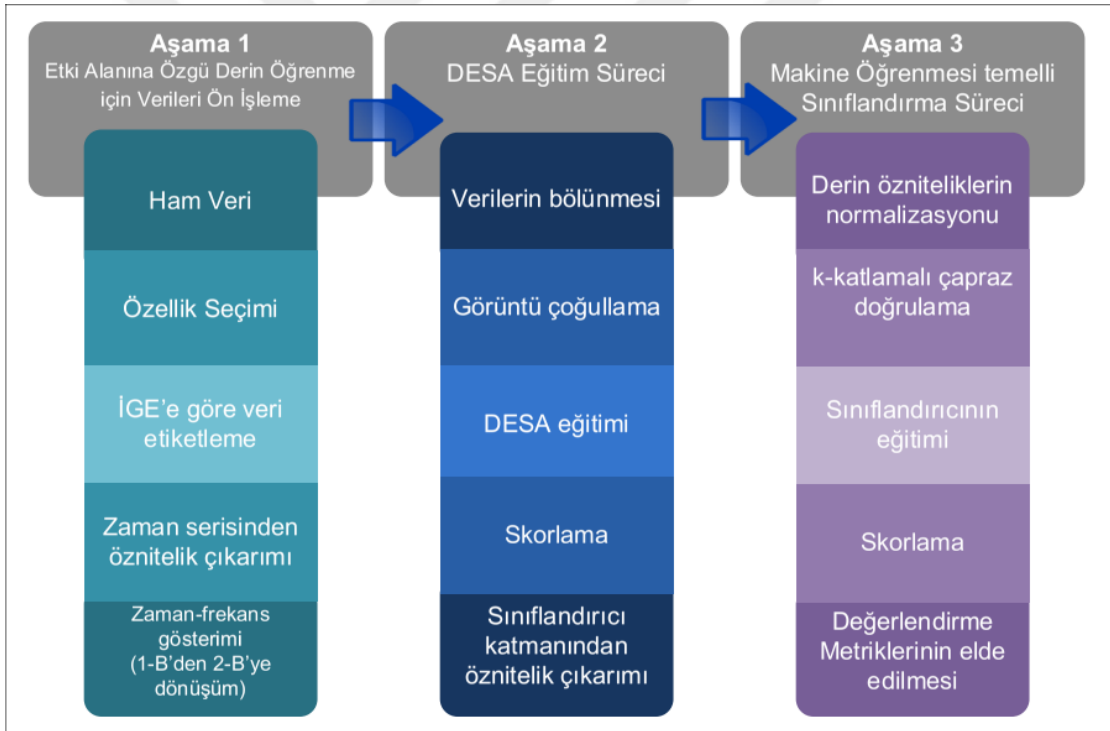
Çizelge 2.1: Ülkelerin gruplandırılması için İGE'deki kesme noktaları [49] .

Çok yüksek insani gelişme (VH)	0.800 ve üzeri
Yüksek insani gelişme (H)	0.700–0.799
Orta insani gelişme (M)	0.550–0.699
Düşük insani gelişme (L)	0,550'nin altında



3. MATERYAL VE METOT

Bu bölümde önerilen yöntemin adımları alt başlıklar halinde açıklanacaktır. Çalışma üç ana başlık altında incelenebilir. İlk aşamada, kamuya açık veri seti elde edilerek veri temizleme işlemleri gerçekleştirilmiştir. Böylece ham verilerden işlenebilir bir yapı oluşturulmuştur. Veriyi anlamlı kılan özelliklerin sayısı fazla olduğu için eleme yapılmıştır. Özellik seçimi, verilerin daha az önemli kısımlarını kaldırmamıza ve eğitim süresini kısaltmamıza yardımcı olmaktadır. Veri etiketleme için, seçilen üç özellik vektörü, ilgili ülkenin İGE'si ile ilişkilendirilmiştir. Elde edilen paketin üçüncü bileşeni olan "yeni test" vektörü, 50 birim uzunluğunda kayar pencere yöntemine göre taranmış ve standart sapması alınmıştır. Her sınıf için üç ayrı skalogram görüntüsü bir araya getirilerek İç Penetrasyon Görüntüleri (IPG) oluşturulmuştur. Böylece her üç görüntünün de yapısal özelliklerini içeren tek bir görüntü elde edilmiştir. Çalışmanın şematik gösterimi Şekil 3.1'de verilmiştir.

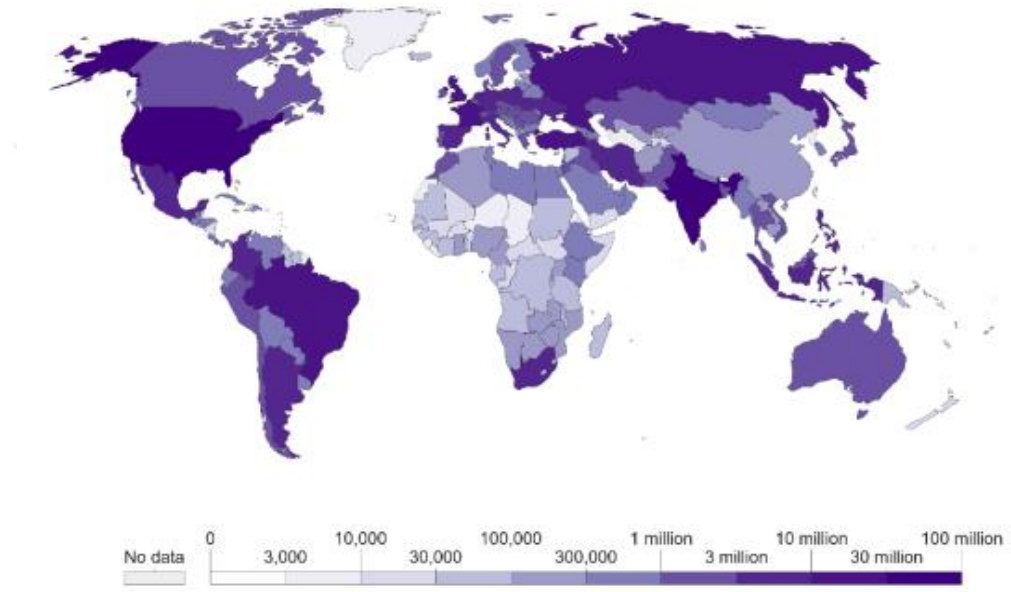


Şekil 3.1: Çalışmanın şematik diyagramı.

3.1 Veri Setinin Tanıtımı ve Özellik Seçimi

Bu çalışmada kullanılan veri seti, Johns Hopkins Üniversitesi Sistem Bilimi ve Mühendisliği Merkezi'ndeki bir ekip tarafından yönetilen, doğrulanmış vakalar ve ölümlere ilişkin verilerin raporlanmasıyla elde edilmiştir. 22 Ocak 2020'den bu yana, tüm ülkeler için teyit edilen vakalar ve ölümlerle ilgili veriler günde birkaç kez güncellenmiştir. Bu veriler

hükümetlerin ulusal ve yerel kurumlarından elde edilmiş ve her ülke için veri kaynaklarının tam listesi Johns Hopkins GitHub sitesinde yayınlanmıştır [50]. Toplam 187 ülke için oluşturulan veri setinde Çizelge 3.1’de verilen 14 Şubat 2020 ile 10 Ağustos 2021 arasındaki veriler baz alınmıştır. Her ülkenin özellik vektörlerinin, İGE ile ilişkisini kurmak için çıktı vektörü olarak İGE seçilmiştir. Şekil 3.2’de 11 Ocak 2022 itibariyle kümülatif teyit edilmiş COVID19 vakalarını göstermektedir. Teyit edilen vaka sayısı, sınırlı test nedeniyle gerçek enfeksiyon sayısından daha düşüktür.



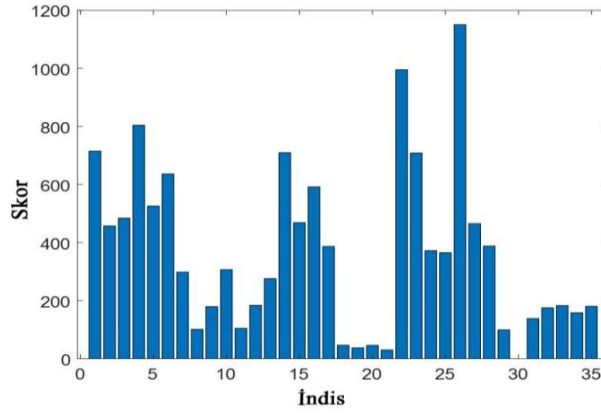
Şekil 3.2: Kümülatif doğrulanmış COVID-19 vakaları, 11 Ocak 2022 [50].

COVID-19 verilerindeki istikrarsız değişim, İGE açısından ülkeler arasında net bir sınır oluşturmamaktadır. Bu nedenle, M ve L sınıfları L altında gruplandırılmış ve H ve VH sınıfları H altında toplanmıştır. Özellik seçme yöntemi kullanılarak ele alınan 35 özellik vektöründen bazıları elenerek, DESA yapısının doğru sonuç vermesi sağlanmıştır.

Çizelge 3.1: COVID-19 için özellik listesi.

ID	Özellik	ID	Özellik
1.	Toplam vaka	19.	Milyon başına haftalık yoğun bakım hasta kabulü
2.	Yeni vaka	20.	Haftalık hastane başvuruları
3.	Düzeltilen yeni vakalar	21.	Milyon başına haftalık hastane kabulü
4.	Toplam ölümler	22.	Yeni testler
5.	Yeni ölümler	23.	Toplam testler
6.	Düzeltilen yeni vakalar	24.	Binde toplam test
7.	Milyonda toplam vaka	25.	Binde yeni test
8.	Milyonda yeni vaka	26.	Düzeltilen yeni testler
9.	Düzeltilen milyonda yeni vaka	27.	Düzeltilen binde yeni testler
10.	Milyonda toplam ölüm	28.	Pozitif oran
11.	Milyonda yeni ölüm	29.	Vaka başına testler
12.	Düzeltilen milyonda yeni ölümler	30.	Test birimleri
13.	Üreme oranı	31.	Toplam aşılar
14.	Yoğun bakım hastaları	32.	Aşılanmış insanlar
15.	Milyonda yoğun bakım hastaları	33.	Tam aşılı insanlar
16.	Hastanedeki hasta sayısı	34.	Yeni aşılar
17.	Milyonda hastanedeki hasta sayısı	35.	Düzeltilen yeni aşılar
18.	Haftalık yoğun bakım kabulleri		

Özellik seçim rutinlerini ortaya çıkarmak için Python-Scikit kütüphanesi kullanılmıştır. Çizelge 2.1'deki örnekler için ANOVA F-değeri [51] hesaplanmıştır. Bu yöntemle, en yüksek puan alan özellikler dışındaki tüm özellikleri elenmiştir. Belirli bir veri kümesinden en kullanışlı özellikleri çıkarmak için en uygun sınıfı sağlar. Bu teknik, özellikleri en yüksek k puanına göre seçmektedir. En iyi skora sahip ilk 4 özellik vektörü sırasıyla ID_26 (düzeltilmiş yeni testler), ID_22 (yeni testler), ID_4 (toplam ölümler) ve ID_1 (toplam vakalar). ID_26, ID_22'nin düzeltilmiş bir versiyonu olduğu için bu iki vektörden biri seçilmiştir. Böylece her ülke için ortalama uzunluğu 3x550 olan bir özellik vektör paketi elde edilmiştir. Özellik önem grafiği Şekil 3.3'te verilmiştir.



Şekil 3.3: Özellik önem grafiği.

Verinin karakteristiğinde zamana bağlı olarak değişen bazı özellikleri ön plana çıkarmak gerekir. Bu durum sınıflandırıcının performansında olumlu faydalar sağlamaktadır [52][53]. Bu özelliklerin analizi ve yorumlanması için bazı istatistiksel yöntemler mevcuttur. Bu yöntemlerden biri standart sapmadır. Standart sapma, verilerin ortalama etrafında yayılmasını belirlemek için istatistiksel bir analiz yaklaşımıdır. Yüksek bir standart sapma ile uğraşırken, bu, ortalamadan geniş bir alana yayılan verileri ifade eder. Benzer şekilde, düşük bir sapma, çoğu verinin ortalamayla benzer olduğunu ve kümenin beklenen değeri olarak da adlandırılabilceğini gösterir. Standart sapma, çoğunlukla veri noktalarının dağılımını belirlemek için kullanılır. Elde edilen veri paketinin üçüncü bileşeni olan “yeni_testler” vektörü, 50 birim uzunluğunda kayar pencere yöntemine göre taranmış ve standart sapması alınmıştır. Standart sapmanın formülasyonu Denklem 11'de verilmiştir.

$$\text{Standart} = \left(\frac{1}{N-1} \sum_{i=1}^N (X(i) - \bar{X})^2 \right)^{\frac{1}{2}} \quad (11)$$

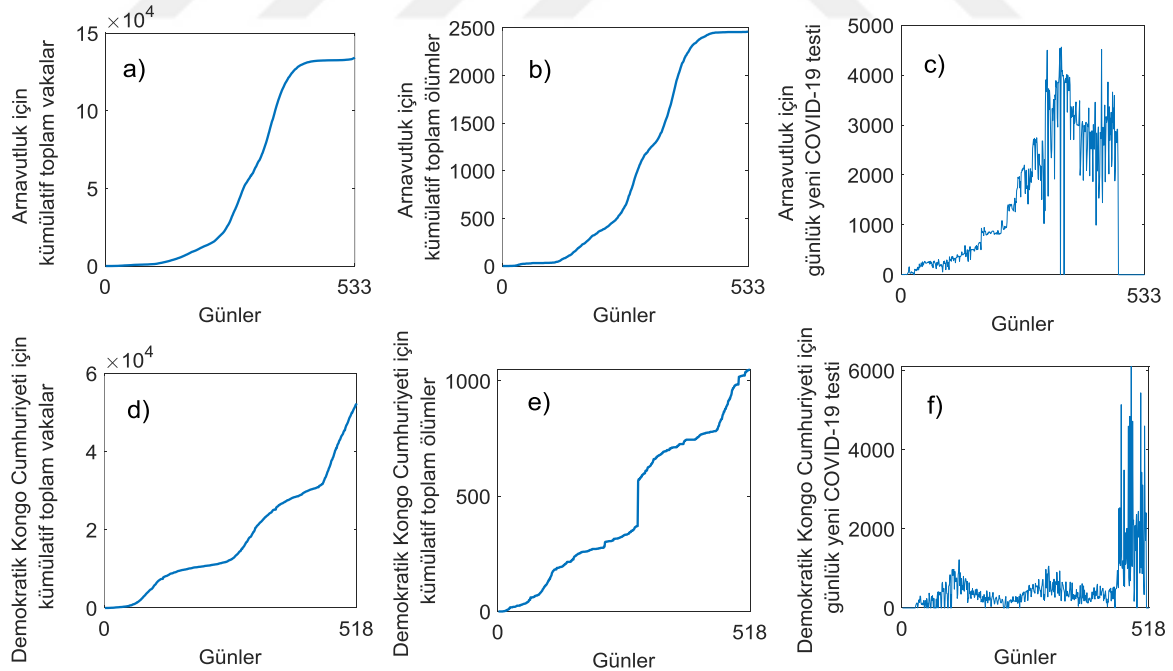
3.2 SDD ve Skalogram Yöntemleri ile Verilerin Görselleştirilmesi

Zaman-frekans analizi, giriş sinyallerini zamanla değişen spektral bileşenlerine ayırarak, sinyal özelliklerinin derinlemesine araştırılması için etkili bir araçtır. SDD, zaman-frekans verilerinin analizinde önemli bir rol oynamıştır. Dalgacıklar, sinyalleri ayrıştırmak için kullanabileceğimiz fonksiyonlardır. Fourier dönüşümü bir sinyali karmaşık sinüzoidler ailesine ayrıştırdığı gibi, dalgacık dönüşümü de bir sinyali bir dalgacık ailesine ayrıştırır. Fourier dönüşümü sayesinde sinyallerin spektrumları analiz edilebilmekte ve frekans tabanında sistemlerin ve özelliklerinin betimlenmesi sağlanabilmektedir. Dalgacık analizi, incelenen sinyallerin zamansal ve frekans dalgalanmalarına ilişkin derinlemesine bir

perspektif sağlayabileceğinden, yerel özellikleri çıkarmak için yaygın olarak kullanılır. Dalgacık fonksiyonunun kaydırılıp bir ölçek ile çarpılır ardından zaman alanı boyunca toplanmasıyla elde edilir ve aşağıdaki gibi hesaplanır.

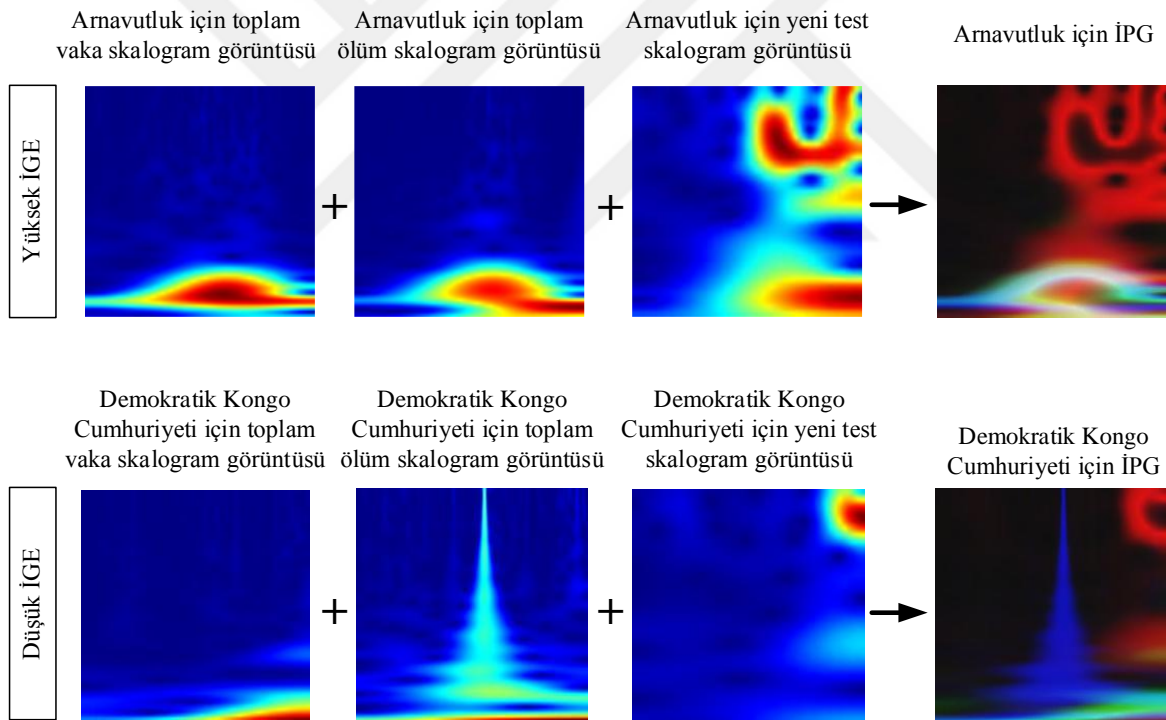
$$C(a, b, f(t), \psi(t)) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) \psi\left(\frac{t-b}{a}\right) \cdot dt \quad (12)$$

Burada $\psi(t)$ ana dalgacıktır, a ölçektir, $C(a, b)$ SDD katsayılarıdır ve b zaman kaymasıdır. SDD'deki zaman ölçeği özelliği, COVID-19 verilerinin alt özelliklerini tanımlamaya yardımcı olan yüksek zaman-frekans çözünürlüğü elde edilmesini sağlar. Dalgacık analizi, incelenen sinyallerin zamansal ve frekans dalgalanmalarına ilişkin bilgi sağlayabileceğinden, yerel özellikleri çıkarmak için yaygın olarak kullanılır. Zaman serisi COVID-19 verilerinin iki boyutlu zaman frekansı gösterimi için kullanılan SDD, COVID-19 verilerini (yeni testler, toplam ölümler, toplam vakalar) farklı zaman ve ölçeklerde analiz etmemizi sağlayan bir yöntemdir. SDD'deki zaman ölçeği özelliği, COVID-19 verilerinin alt özelliklerini tanımlamaya yardımcı olan yüksek zaman-frekans çözünürlüğü elde edilmesini sağlar. Arnavutluk ve Demokratik Kongo Cumhuriyeti için kümülatif toplam vakaların, kümülatif toplam ölümlerin ve günlük yeni COVID19 testlerinin günlere göre sapması Şekil 3.4'te verilmiştir.



Şekil 3.4: Arnavutluk ve Demokratik Kongo Cumhuriyeti için kümülatif toplam vakaların, kümülatif toplam ölümlerin ve günlere göre günlük yeni COVID-19 testlerinin sapması.

Skalogram, sinyale SDD uygulandıktan sonra elde edilen katsayıları kullanarak sinyaldeki hızlı değişiklikleri görselleştirmeye yardımcı olur. SDD skalogramı, zamanı yatay eksen ve ölçekleri dikey eksen de gösteren iki boyutlu bir çizimdir. Skalogramların zaman-frekans temsilleri, sınıflandırmak için kullanılacak derin ESA yapılarının girişine gönderilir. Filtre bankaları, COVID-19 verileriyle elde edilen skalogram görüntülerini çeşitli filtrelerden geçirerek frekans içeriğinin çıkarılmasına katkıda bulunur. Scalogram, sinyale SDD uygulandıktan sonra elde edilen katsayıları kullanarak sinyaldeki hızlı değişiklikleri görselleştirmeye yardımcı olur. IPG, işleme karmaşıklığını ortadan kaldırmak ve eğitim-doğrulama veri dağılımını doğru bir şekilde yapmak amacıyla her sınıf için üç ayrı görüntünün birleştirilmesiyle oluşturulmuştur. Böylece her üç görüntünün de yapısal özelliklerini içeren tek bir görüntü elde edilmiştir. Düşük ve Yüksek İGE'ye sahip iki farklı ülke için ID_22 (yeni test), ID_4 (toplam ölüm) ve ID_1 (toplam vaka) skalogram görüntülerinin yanısıra, iç içe nüfuz etmiş görüntüler de Şekil 3.5'te verilmiştir.



Şekil 3.5: Düşük ve Yüksek İGE'ye sahip iki farklı ülkenin IPG görselleri.

3.3 Hiper Parametre Optimizasyonu

Hiper parametre optimizasyonu, makine öğrenmesi problemlerinde tahmin ve sınıflandırmaların daha doğru, daha hassas veya özellik seçimlerinin daha başarılı yapılabilmesi için yaygın olarak kullanılan popüler araştırma konularından biridir. Bu

yöntem makine öğrenimi algoritmalarının tahmin doğruluğunda hayati bir rol oynar. Uygulamada, hiper parametreleri sürekli olarak ayarlamak ve farklı değer kombinasyonlarıyla bir dizi modeli eğitmek ve ardından en iyi modeli seçmek için model performansını karşılaştırmak gerekir. Bu nedenle, hiper parametrelerin nasıl optimize edileceği, bir makine öğrenimi algoritmasında önemli bir konu haline gelir. Hiper parametre ayarlama, optimizasyonun amaç fonksiyonunun bilinmediği veya bir kara kutu fonksiyonunun olduğu bir optimizasyon problemidir. Newton yöntemi veya gradyan inişi gibi geleneksel optimizasyon teknikleri uygulanamaz. Kullanıcıların hiper parametrelerin dağılımı hakkında ön bilgiye sahip olmaları gerekmemesi ve arkasında olasılık fikrine dayanmasından dolayı rastgele arama ve ızgara arama yöntemlerinden daha verimli sonuçlar elde etmektedir. Bayes optimizasyonu, bu tür optimizasyon problemlerini çözmek için çok etkili bir optimizasyon algoritmasıdır [12]. Bayes formülünü kullanarak fonksiyon dağılımının sonsal bilgilerini elde etmek için bilinmeyen fonksiyon hakkındaki ön bilgileri örnek bilgilerle birleştirir. Bu algoritmanın amacı, bilinmeyen bir $f(x)$ fonksiyonu için örnekleme noktasındaki maksimum değeri bulmaktır:

$$x^+ = \arg \max_{x \in A} f(x) \quad (13)$$

Burada A , x 'in arama uzayıdır.

4. DENEYSEL UYGULAMALAR

Bu bölümde tez çalışması kapsamında gerçekleştirilen iki uygulama verilmiştir. Birinci uygulamada, elde edilen İPG görüntüleri önceden eğitilmiş AlexNet ağının girişine verilmiştir. Ağın sınıflandırıcı katmanından alınan derin öznitelikler, farklı sınıflandırıcıların girişini besleyecek şekilde düzenlenmiştir. Kolektif Öğrenme (Boosted Trees), k-En Yakın Komşu (Coarse), DVM (Fine Gaussian) gibi sınıflandırıcıların performansı ayrı ayrı test edildikten sonra, en yüksek başarıyı oranına sahip DVM'nin hiper parametreleri, Bölüm 3.3'te anlatılan Bayes algoritmasıyla optimize edilmiştir. Elde edilen sonuçlar, karşılaştırmalı olarak kıyaslanmıştır. İkinci uygulamada ise, elde edilen İPG görüntüleri ResNet-50 tabanlı DESA+DVM ile COVID-19 ve İGE arasındaki ilişkinin belirlenmesinde kullanılmıştır. Tüm deneyler, AMD Ryzen 5 2600 3.4GHz CPU, 64 GB bellek ve 12 GB NVIDIA GeForce RTX 2080 TI GPU'lu bir PC üzerinde çalışan MATLAB paket programıyla gerçekleştirilmiştir.

4.1 Optimize Edilmiş Hiper Parametrelili Sınıflandırıcı ile COVID-19 ve İGE Arasındaki İlişkinin Belirlenmesi

Derin evrişimli sinir ağı olarak kullanılan önceden eğitilmiş AlexNet ağı, elde edilen İPG görüntüleri ile beslenmiştir. Öğrenme süreci boyunca bu aşamada çevrimiçi görüntü çoğullama yöntemleri kullanılarak görüntü sayısı artırılmıştır. Son aşamada, eğitilmiş ağdan alınan derin öznitelik vektörleri %75 eğitim ve %25 doğrulama için ayrılmıştır. DVM'nin hiper parametreleri Bayes optimizasyon algoritmasıyla optimize edilerek, en yüksek sınıflandırma metrikleri elde edilmiştir. Optimize edilen Hiper parametreler ve arama aralıkları DVM için şöyledir; kutu kısıtlama parametresi (0.001-1000), çekirdek fonksiyonu (Gaussian, Linear, Quadratic, Cubic), standardize (var, yok). Optimizasyon sonucunda elde edilen parametreler şunlardır; kutu kısıtlama parametresi (0.022343), çekirdek fonksiyonu (Quadratic), standardize (yok). Şekil 4.1.a'da optimizasyon sonucunda elde edilen sınıflandırma hata değerinin, iterasyon değerine göre değişimi gösterilmiştir. 30 iterasyon sonucunda, hata değeri yaklaşık %29 oranında azalmıştır. Önerilen tasarımın performansı istatistiksel değerlendirme kriterleri kullanılarak analiz edilecektir. Doğruluk (ACC), Duyarlılık (Sens.), Özgüllük (Spec.), Kesinlik (Prec.), F1 puanı (F1) ve MKK, karışıklık matrisi kullanılarak hesaplanmıştır. Denklem 14-19, kullanılan değerlendirme metriklerini tanımlamaktadır.

$$\text{Doğruluk} = \frac{N_{TP} + N_{TN}}{N_{TP} + N_{TN} + N_{FP} + N_{FN}} \quad (14)$$

$$\text{Duyarlılık} = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (15)$$

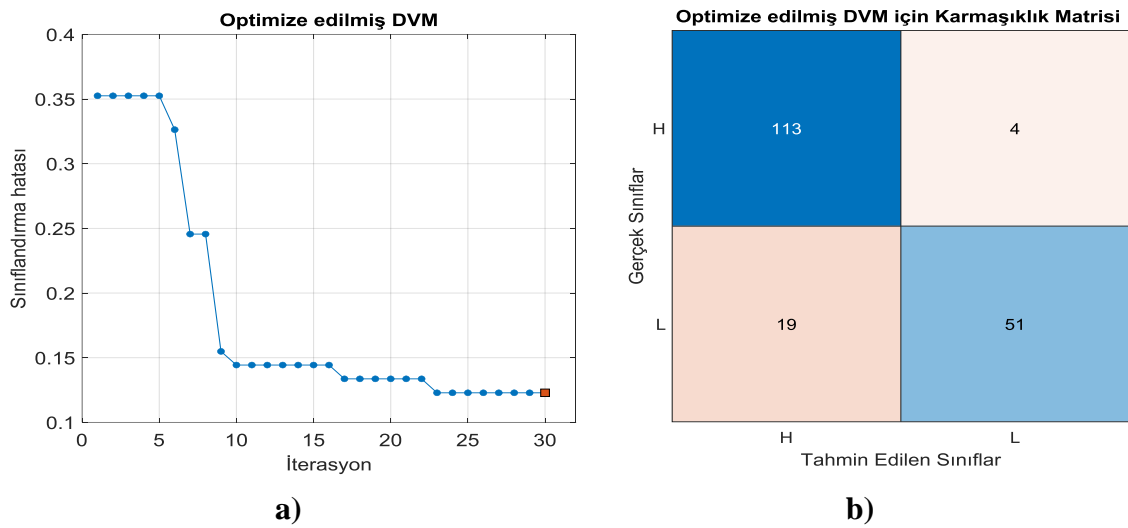
$$\text{Özgüllük} = \frac{N_{TN}}{N_{TN} + N_{FP}} \quad (16)$$

$$\text{Kesinlik} = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad (17)$$

$$F_1 = \frac{2N_{TP}}{2N_{TP} + N_{FP} + N_{FN}} \quad (18)$$

$$\text{MCC} \quad \text{MKK} = \frac{N_{TP}N_{TN} - N_{FP}N_{FN}}{\sqrt{(N_{TP} + N_{FP})(N_{TP} + N_{FN})(N_{TN} + N_{FP})(N_{TN} + N_{FN})}} \quad (19)$$

Burada N_{TP} , N_{TN} , N_{FP} ve N_{FN} sırasıyla doğru pozitif sınıfların sayısını, doğru-negatif sınıfların sayısını, yanlış-pozitif sınıfların sayısını ve yanlış-negatif sınıfların sayısını tanımlar. Şekil 4.1.b'de ise optimize edilmiş DVM'ne ait başarıyı gösteren karmaşıklık matrisi verilmiştir. Yüksek İGE'ye sahip 117 ülke içerisinde 4'ü düşük olarak sınıflandırılırken, Düşük İGE'ye sahip 70 ülke içerisinde, 19'u ise yüksek olarak sınıflandırılmıştır.



Şekil 4.1: a) Sınıflandırma hatasının, iterasyon sayısına göre değişimi b) Optimize edilmiş DVM için karmaşıklık matrisi

Çizelge 4.1’de, farklı sınıflandırıcılar ve optimize edilmiş DVM için başarımleri verilmiştir. Optimize DVM, Kolektif Öğrenmeye göre yaklaşık %7’lik bir Doğruluk iyileştirmesine sahiptir. Bu oran Duyarlılıkta yaklaşık %11, F1’de %6 ve MKK’de % 15’tir. DVM ve Optimize DVM arasındaki farkın net olduğu metrik MKK değeridir. Özellikle sayısal olarak dengesiz dağılıma sahip sınıfların ayırımında önemli bir karar metriği olan MKK, Optimize edilmiş DVM için diğerlerine göre yüksek bir değere sahiptir.

Çizelge 4.1: Elde edilen başarımleri metriklerinin kıyaslanması.

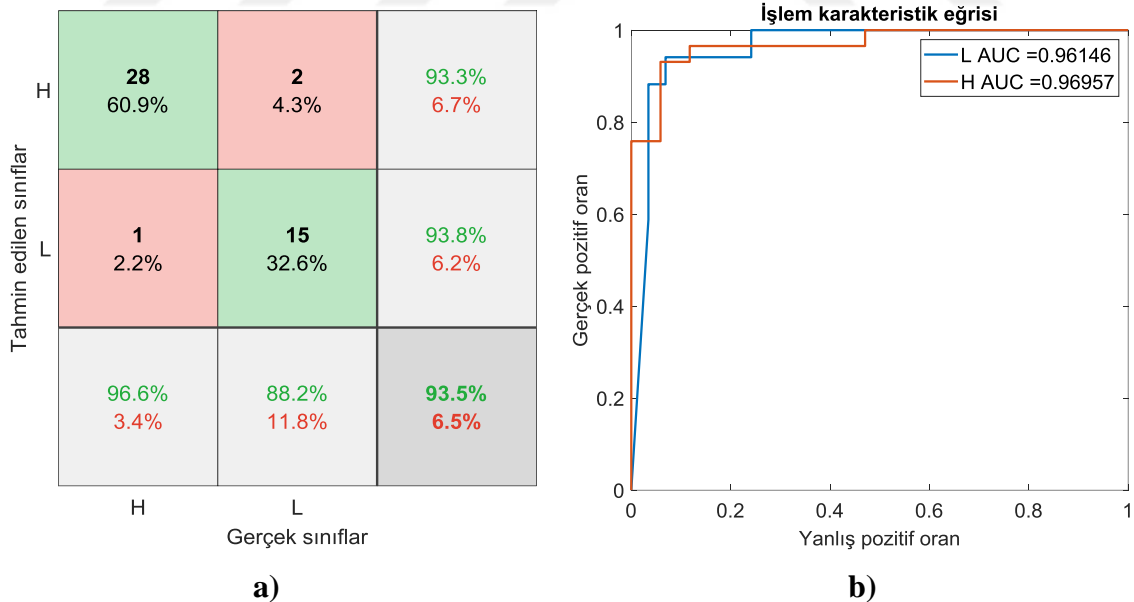
Yöntem	Başarımleri Metrikleri					
	Doğruluk	Duyarlılık	Özgüllük	Kesinlik	F1	MKK
Kolektif Öğrenme	80.75	85.47	72.86	84.03	84.75	58.68
k-En Yakın Komşu	83.96	97.44	61.43	80.85	88.37	66.15
DVM	85.56	92.31	74.29	85.71	88.89	68.74
Optimize DVM	87.7	96.5	72.8	85.61	90.7	73.75

Ağın sınıflandırıcı katmanından seçilen özellikler farklı makine öğrenmesi tabanlı sınıflandırıcıların girişine uygulanmıştır. En iyi sonuç veren mimari olan DVM’nin hiper parametreleri optimize edilerek en iyi sonuç elde edilmiştir. Böylece, Özgüllük (%72.8), Duyarlılık (%96.5), Kesinlik (%85.61), Doğruluk (%87.7) ve F1 Puanı (%90.7) metriklerine ulaşılmıştır.

4.2 ResNet-50 tabanlı DESA+DVM ile COVID-19 ve İGE Arasındaki İlişkinin Belirlenmesi

Bu bölümde, 1000 kategori ve bir milyondan fazla görüntü ile önceden eğitilmiş bir DESA ağı olan ResNet-50 modeli kullanılmıştır [54]. ResNet-50, yeni bir dizi IPG görüntüsünü tanımlamak ve önceki bölümde toplanan IPG görüntülerini kullanarak İGE’yi saptamak için eğitilmiştir. ResNet-50 ağı, bir girdi katmanı, 50 derin katman (evrişim katmanları, toplu normalleştirme katmanları ve ReLU), sınıflandırma katmanları (ortalama havuzlama katmanları, tam bağlantılı katmanlar ve SoftMax katmanları) ve çıktı katmanlarından oluşan bir yapıyı içerir. Girdi görüntüleri, ResNet-50 modelleriyle uyumlu olması için 224x224 piksel olarak yeniden boyutlandırılmıştır. Öğrenme aşamasında, her sınıf için görüntü sayısındaki dengesizliği ve aşırı öğrenme sorunlarını en aza indirmek amacıyla çevrimiçi görüntü çoğullama yöntemi kullanılarak veri sayısı artırılmıştır. Derin özellikler, modifiye edilmiş fc1000 katmanı kullanılarak DVM’ye aktarılmıştır. Tüm özellik setlerinin ortalaması z-skorlamaya göre normalize edilmiştir. Sınıflandırma aşamasında, %75 eğitim

ve %25 test verisi ile DVM kullanılmıştır. Hesaplama verimliliği ile ilgili olarak önerilen yöntemin geçerliliği ve değerlendirme kriterlerine etkisi sunulmuştur. Derin özellik seti tarafından beslenen DVM'nin çok sınıflı karışıklık matrisi Şekil 4.2-a'da gösterilmektedir. Tahmin edilen ve gerçek sınıflar, karışıklık matrisinde satırlar ve sütunlarla temsil edilir. Doğru sınıflandırılmış gözlemler çapraz hücrelerle gösterilirken, hatalı sınıflandırılmış gözlemler diğer hücreler tarafından temsil edilir. Her hücre, gözlemlerin sayısını ve toplam gözlem sayısının yüzdesini içerir. Doğru (kesinlik ölçümleri) ve yanlış olarak sınıflandırılan her bir sınıfa ait olduğu tahmin edilen tüm örneklerin yüzdeleri, grafiğin en sağdaki sütununda (yanlış tahmin oranı ölçümleri) gösterilir. Matrisin sağ alt köşesindeki hücre, toplam doğruluğu temsil eder. Karışıklık matrisine göre 46 doğrulama örneği arasında üç yanlış sınıflandırılmış örnek var. İki örnekle "L" sınıfının en fazla yanlış sınıflandırılan örneklere sahip olduğu belirlenmiştir. Yanlış pozitif oranın bir fonksiyonu olarak gerçek pozitif oranı gösteren AÇK eğrisi, aynı zamanda makine öğrenmesi algoritmalarının etkinliğini değerlendirmek için en yaygın kullanılan metriklerden biridir. Eğri altında kalan alan (İng: Area Under Curve - AUC) değerleri ile tahmin edilen sınıfların AÇK eğrileri Şekil 4.2b'de verilmiştir. AÇK eğrisinin alanı "L" ve "H" sınıfları için sırasıyla 0.961 ve 0.969 olarak elde edilmiştir.



Şekil 4.2: a) SVMC'nin derin özniteliklerle karışıklık matrisi. b) Derin özellikler için tüm sınıfların AÇK eğrileri.

DVM'nin genel sınıflandırma puanları Çizelge 4.2'de listelenmiştir. COVID-19 veri kümesi için toplam metrik değerler, tabloda gösterildiği gibi 187x2 özelliklerle önerilen tasarımda gözlemlenmiştir. COVID-19 veri kümesi için, güncellenmiş fc1000 katmanındaki derin

özellikleri kullanan modelin genel doğruluğu %93,4'tü. Diğer genel performans göstergeleri olan doğruluk, duyarlılık, özgüllük, kesinlik, F1 ve MKK sırasıyla %88,2, %96,5, %99, %94,9 ve %85,9 olarak hesaplanmıştır.

Çizelge 4.2: Derin özellik kümesi için genel sınıflandırma değerleri.

Yöntem	Başarım Metrikleri					
	Doğruluk	Duyarlılık	Özgüllük	Kesinlik	F1	MKK
ResNet-50+DVM	0.934	0.965	0.882	0.990	0.949	0.859



5. SONUÇLAR VE TARTIŞMA

Yapay zeka, tanısal görüntüleme sistemlerinde devrim yaratmış ve tıbbi görüntülemedeki en trend konulardan biri olmuştur. Derin öğrenme yaklaşımları, özellikle ESA'ların kullanımı, geleneksel makine öğrenme tekniklerine göre önemli performans iyileştirmeleri sağlamıştır. Tıbbi olmayan verilerin klinik belirleyiciler haline gelmesinin önemi literatürde kapsamlı bir şekilde gözden geçirilmiştir. Bu çalışma tıbbi verilere dayanmaktadır ve sosyoekonomik girdiler kullanılarak hesaplanan İGE arasındaki korelasyonu göstermiştir. Literatür taraması sonucunda gördüğümüz kadarıyla bu tez çalışması, COVID-19 verilerini kullanarak İGE sınıflandırması için DESA ile sinyal işleme tekniklerini kullanan yeni bir yaklaşım önermektedir ve bu açıdan literatüre katkı sağladığı söylenebilir. Bu çalışmaya göre sağlık, eğitim ve ekonomi üzerinde entegre etkileri olan İGE, pandemi nedenleri bağlamında ele alınmalıdır.

Çalışma kapsamında,

1. Dünya ülkelerinin COVID-19 verileri alınmış ve halka açık ve güvenilir bir kaynaktan elde edilen bilgiler kullanılarak kabul edilebilir bir formatta yeniden yapılandırılmıştır.
2. İstatistiksel özellik seçimini kullanarak, hangi faktörlerin İGE ile yakından ilişkili olduğu bulunmuş ve DESA modelinin daha doğru bulgular üretmesi sağlanmıştır.
3. Zaman serisi veri görselleştirmesi, SDD ve skalogram yöntemleri kullanılarak gerçekleştirilmiştir.
4. İşleme kolaylığı amacıyla, her ülkenin üç ayrı görüntüsü, tek bir görüntüde birleştirilmiştir.
 - 4.1. Bu görüntüler, farklı ön işleme adımlarından geçerek, önceden eğitilmiş bir DESA modeli olan ResNet-50 ağının girişi için uygun hale getirilmiştir.
 - 4.2. Ağın fc1000 katmanındaki öznelik vektörleri oluşturulmuş ve eğitim ve doğrulama işlemlerinden sonra DVMS girişine verilmiştir.
 - 4.3. Eğitim sonucunda ise Özgüllük (%88,2), duyarlılık (%96,5), doğruluk (%99), F1 Skoru (%94,9) ve MKK (%85,9) performans değerlerine ulaşılmıştır.

5. Çalışmanın ikinci ayağında Madde 4.’teki görüntüler, DESA modeli olan AlexNet ağının girişi için uygun hale getirilmiştir.

5.1. Ağın sınıflandırıcı katmanından seçilen öznelikler farklı makine öğrenmesi tabanlı sınıflandırıcıların girişine uygulanmıştır.

5.2. En iyi sonuç veren mimari olan DVM’nın hiper parametreleri optimize edilerek en iyi sonuç elde edilmiştir.

5.3. Böylece, Özgüllük (%72.8), Duyarlılık (%96.5), Kesinlik (%85.61), Doğruluk (%87.7) ve F1 Puanı (%90.7) metriklerine ulaşılmıştır.

Elde edilen sonuçlar, eğitim, sağlık ve gelir verilerine göre hesaplanan İGE ile COVID-19 verilerinin değişimi arasında korelasyon olduğunu ortaya koymaktadır. İncelenen skalogram görüntülerine göre, “yeni test” vektörünün sınıflandırmada baskın rol oynadığını göstermektedir. Sonraki çalışmalarda, farklı ağ yapıları kullanılarak, sınıflandırıcı katmanından çekilen özneliklerin sezgisel yöntemlerle optimize edilmesi planlanmaktadır. Ayrıca ağ analizi temelli çalışmaların yapılması da hedeflenmektedir.

KAYNAKLAR

- [1] “who,” <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/> , May 2021.
- [2] “Department of Error: Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study (The Lancet (2020) 395(10225) (689–697), (S0140673620302609), (10.1016/S0140-6736(20)30260-9)),” *The Lancet*, vol. 395, no. 10225. 2020. doi: 10.1016/S0140-6736(20)30302-0.
- [3] “who.covid19,” <https://covid19.who.int/> , May 2021.
- [4] “www.un.org/en/coronavirus/UN-response,” <https://www.un.org/en/coronavirus/UN-response>.
- [5] World Bank, “The Human Capital Index 2020 Update: Human Capital in the Time of COVID-19,” *The Human Capital Index 2020 Update*. 2021.
- [6] J. A. Hasnaoui, S. K. A. Rizvi, K. Reddy, N. Mirza, and B. Naqvi, “Human capital efficiency, performance, market, and volatility timing of asian equity funds during COVID-19 outbreak,” *Journal of Asset Management*, vol. 22, no. 5, 2021, doi: 10.1057/s41260-021-00228-y.
- [7] A. Narin, C. Kaya, and Z. Pamuk, “Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks,” *Pattern Analysis and Applications*, vol. 24, no. 3, 2021, doi: 10.1007/s10044-021-00984-y.
- [8] O. Gozes *et al.*, “Title: Rapid AI Development Cycle for the Coronavirus (COVID-19) Pandemic: Initial Results for Automated Detection & Patient Monitoring using Deep Learning CT Image Analysis Authors.”
- [9] S. Hu *et al.*, “Weakly Supervised Deep Learning for COVID-19 Infection Detection and Classification from CT Images,” *IEEE Access*, vol. 8, 2020, doi: 10.1109/ACCESS.2020.3005510.
- [10] E. El-Din Hemdan, M. A. Shouman, and M. E. Karar, “COVIDX-Net: A Framework of Deep Learning Classifiers to Diagnose COVID-19 in X-Ray Images.”
- [11] M. Kim *et al.*, “Hi-COVIDNet: Deep Learning Approach to Predict Inbound COVID-19 Patients and Case Study in South Korea,” 2020. doi: 10.1145/3394486.3412864.
- [12] F. Ucar and D. Korkmaz, “COVIDiagnosis-Net: Deep Bayes-SqueezeNet based diagnosis of the coronavirus disease 2019 (COVID-19) from X-ray images,” *Medical Hypotheses*, vol. 140, 2020, doi: 10.1016/j.mehy.2020.109761.
- [13] M. Ahsan, K. Datta Gupta, M. M. Islam, S. Sen, L. Rahman, and M. S. Hossain, “machine learning & knowledge extraction COVID-19 Symptoms Detection Based on NasNetMobile with Explainable AI Using Various Imaging Modalities”, doi: 10.3390/make2040027.
- [14] Y. Jiang, H. Chen, M. Loew, and H. Ko, “COVID-19 CT Image Synthesis with a Conditional Generative Adversarial Network,” *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 2, 2021, doi: 10.1109/JBHI.2020.3042523.
- [15] A. Waheed, M. Goyal, D. Gupta, A. Khanna, F. Al-Turjman, and P. R. Pinheiro, “SPECIAL SECTION ON EMERGING DEEP LEARNING THEORIES AND METHODS FOR BIOMEDICAL ENGINEERING CovidGAN: Data Augmentation Using Auxiliary Classifier GAN for Improved Covid-19 Detection”, doi: 10.1109/ACCESS.2020.2994762.
- [16] M. Nour, Z. Cömert, and K. Polat, “A Novel Medical Diagnosis model for COVID-19 infection detection based on Deep Features and Bayesian Optimization,” *Applied Soft Computing*, vol. 97, 2020, doi: 10.1016/j.asoc.2020.106580.

- [17] L. Zhou *et al.*, “A Rapid, Accurate and Machine-Agnostic Segmentation and Quantification Method for CT-Based COVID-19 Diagnosis,” *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, 2020, doi: 10.1109/TMI.2020.3001810.
- [18] H. Panwar, P. K. Gupta, M. K. Siddiqui, R. Morales-Menendez, and V. Singh, “Application of deep learning for fast detection of COVID-19 in X-Rays using nCOVnet,” *Chaos, Solitons and Fractals*, vol. 138, 2020, doi: 10.1016/j.chaos.2020.109944.
- [19] D. Ezzat, A. E. Hassanien Abo, and H. A. Ella, “GSA-DenseNet121-COVID-19: A hybrid deep learning architecture for the diagnosis of COVID-19 disease based on gravitational search optimization algorithm,” *arXiv*. 2020.
- [20] J. Chen *et al.*, “Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography,” 2020, doi: 10.1038/s41598-020-76282-0.
- [21] H. S. Maghded, K. Z. Ghafoor, A. S. Sadiq, K. Curran, D. B. Rawat, and K. Rabie, “A Novel AI-enabled Framework to Diagnose Coronavirus COVID-19 using Smartphone Embedded Sensors: Design Study,” 2020. doi: 10.1109/IRI49571.2020.00033.
- [22] N. Narayan Das, N. Kumar, M. Kaur, V. Kumar, and D. Singh, “Automated Deep Transfer Learning-Based Approach for Detection of COVID-19 Infection in Chest X-rays,” *IRBM*, vol. 43, no. 2, 2022, doi: 10.1016/j.irbm.2020.07.001.
- [23] A. Rehman, S. Naz, A. Khan, A. Zaib, and I. Razzak, “Improving Coronavirus (COVID-19) Diagnosis using Deep Transfer Learning,” *medRxiv*, 2020.
- [24] M. S. Boudrioua, “COVID-19 Detection from Chest X-Ray Images Using CNNs Models: Further Evidence from Deep Transfer Learning,” *SSRN Electronic Journal*, 2020, doi: 10.2139/ssrn.3630150.
- [25] S. Vaid, R. Kalantar, and M. Bhandari, “Deep learning COVID-19 detection bias: accuracy through artificial intelligence,” *International Orthopaedics*, vol. 44, no. 8, 2020, doi: 10.1007/s00264-020-04609-7.
- [26] G. Kavuran, “SEM-Net: Deep features selections with Binary Particle Swarm Optimization Method for classification of scanning electron microscope images,” *Materials Today Communications*, vol. 27, 2021, doi: 10.1016/j.mtcomm.2021.102198.
- [27] G. Kavuran, “When machine learning meets fractional-order chaotic signals: detecting dynamical variations,” *Chaos, Solitons & Fractals*, vol. 157, p. 111908, Apr. 2022, doi: 10.1016/J.CHAOS.2022.111908.
- [28] R. Vaishya, M. Javaid, I. Haleem, and A. Haleem, “Artificial Intelligence (AI) applications for COVID-19 pandemic Diabetes & Metabolic Syndrome : Clinical Research & Reviews Arti fi cial Intelligence (AI) applications for COVID-19 pandemic,” *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, vol. 14, no. 4, 2020.
- [29] L. Li *et al.*, “Using Artificial Intelligence to Detect COVID-19 and Community-acquired Pneumonia Based on Pulmonary CT: Evaluation of the Diagnostic Accuracy,” *Radiology*, vol. 296, no. 2, 2020, doi: 10.1148/radiol.2020200905.
- [30] W. Zhao, W. Jiang, and X. Qiu, “Deep learning for COVID-19 detection based on CT images,” *Scientific Reports* /, vol. 11, p. 14353, 123AD, doi: 10.1038/s41598-021-93832-2.
- [31] M. Loey and S. Mirjalili, “COVID-19 cough sound symptoms classification from scalogram image representation using deep learning models,” *Computers in Biology and Medicine*, vol. 139, 2021, doi: 10.1016/j.combiomed.2021.105020.

- [32] X. Mei *et al.*, “Artificial intelligence-enabled rapid diagnosis of patients with COVID-19,” *Nature Medicine*, vol. 26, no. 8, 2020, doi: 10.1038/s41591-020-0931-3.
- [33] M. Pahar, M. Klopper, R. Warren, and T. Niesler, “COVID-19 detection in cough, breath and speech using deep transfer learning and bottleneck features,” *Computers in Biology and Medicine*, vol. 141, 2022, doi: 10.1016/j.compbiomed.2021.105153.
- [34] S. Sakib, T. Tazrin, M. M. Fouda, M. D. Fadlullah, and M. Guizani, “DL-CRC: Deep Learning-Based Chest Radiograph Classification for COVID-19 Detection: A Novel Approach”, doi: 10.1109/ACCESS.2020.3025010.
- [35] A. Sedik, M. Hammad, F. E. Abd El-Samie, B. B. Gupta, and A. A. Abd El-Latif, “Efficient deep learning approach for augmented detection of Coronavirus disease,” *Neural Computing and Applications*, 2021, doi: 10.1007/s00521-020-05410-8.
- [36] M. F. Aslan, M. F. Unlersen, K. Sabanci, and A. Durdu, “CNN-based transfer learning–BiLSTM network: A novel approach for COVID-19 infection detection,” *Applied Soft Computing*, vol. 98, 2021, doi: 10.1016/j.asoc.2020.106912.
- [37] Y. H. Wu *et al.*, “JCS: An Explainable COVID-19 Diagnosis System by Joint Classification and Segmentation,” *IEEE Transactions on Image Processing*, vol. 30, 2021, doi: 10.1109/TIP.2021.3058783.
- [38] K. U. Ahamed *et al.*, “A deep learning approach using effective preprocessing techniques to detect COVID-19 from chest CT-scan and X-ray images,” *Computers in Biology and Medicine*, vol. 139, 2021, doi: 10.1016/j.compbiomed.2021.105014.
- [39] P. P. Shinde and S. Shah, “A Review of Machine Learning and Deep Learning Applications,” 2018. doi: 10.1109/ICCUBEA.2018.8697857.
- [40] C. Janiesch, P. Zschech, and K. Heinrich, “Machine learning and deep learning,” *Electronic Markets*, vol. 31, no. 3, pp. 685–695, Sep. 2021, doi: 10.1007/S12525-021-00475-2.
- [41] R. Yamashita, M. Nishio, R. Kinh, G. Do, and K. Togashi, “Convolutional neural networks: an overview and application in radiology”, doi: 10.1007/s13244-018-0639-9.
- [42] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016, vol. 2016-December. doi: 10.1109/CVPR.2016.90.
- [43] Q. Ji, J. Huang, W. He, and Y. Sun, “Optimized deep convolutional neural networks for identification of macular diseases from optical coherence tomography images,” *Algorithms*, vol. 12, no. 3, 2019, doi: 10.3390/a12030051.
- [44] Q. Abu Al-Haija and A. Adebajo, *Breast Cancer Diagnosis in Histopathological Images Using ResNet-50 Convolutional Neural Network; Breast Cancer Diagnosis in Histopathological Images Using ResNet-50 Convolutional Neural Network*. 2020. doi: 10.1109/IEMTRONICS51293.2020.9216455.
- [45] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Commun ACM*, vol. 60, no. 6, 2017, doi: 10.1145/3065386.
- [46] V. Vapnik, “The Support Vector Method of Function Estimation,” in *Nonlinear Modeling*, Boston, MA: Springer US, 1998, pp. 55–85. doi: 10.1007/978-1-4615-5703-6_3.
- [47] C. Cortes and V. Vapnik, “Support-Vector Networks,” *Machine Learning*, 1995, doi: 10.1023/A:1022627411411.
- [48] E. A. Stanton, “The Human Development Index: A History,” 2007.
- [49] H. Rmili *et al.*, “A Comparative Study of Fingerprint Thinning Algorithms,” *Ieee*, vol. 1, no. 5, 2020.

- [50] E. Dong, H. Du, and L. Gardner, “An interactive web-based dashboard to track COVID-19 in real time,” *The Lancet Infectious Diseases*, vol. 20, no. 5. 2020. doi: 10.1016/S1473-3099(20)30120-1.
- [51] F. J. Ferri, P. Pudil, M. Hatef, and J. Kittler, “Comparative study of techniques for large-scale feature selection,” in *Machine Intelligence and Pattern Recognition*, vol. 16, no. C, 1994. doi: 10.1016/B978-0-444-81892-8.50040-7.
- [52] X. Hu, S. Yuan, F. Xu, Y. Leng, K. Yuan, and Q. Yuan, “Scalp EEG classification using deep Bi-LSTM network for seizure detection,” *Computers in Biology and Medicine*, vol. 124, 2020, doi: 10.1016/j.combiomed.2020.103919.
- [53] D. Wang, “Deep learning reinvents the hearing aid,” *IEEE Spectrum*, vol. 54, no. 3, 2017, doi: 10.1109/MSPEC.2017.7864754.
- [54] S. R. D. Dimas, B. S. Negara, S. Sanjaya, and E. Satria, “COVID-19 Classification for Chest X-Ray Images using Deep Learning and Resnet-101,” 2021. doi: 10.1109/ICOTEN52080.2021.9493431.



ÖZGEÇMİŞ

KİŞİSEL BİLGİLER:

Ad-Soyad :Şeyma Gökhan

EĞİTİM BİLGİLERİ :

Yüksek Lisans : 2022, İnönü Üniversite, Fen Bilimleri Enstitüsü,Bilgisayar Bilimleri

Üniversite :2019, Kocaeli Üniversitesi,Mühendislik Fakültesi,Bilgisayar Mühendisliği

Lise : 2012: Malatya Cumhuriyet Anadolu Lisesi

İlköğretim : 2008: Kazım Karabekir İlköğretim Okulu

İŞ DENEYİMİ

Firma Adı: Tübitak Bilgem İleri Teknolojiler ve Araştırma Merkezi

Pozisyon: Yazılım Test ve Kalite Değerlendirme Laboratuvarı / Araştırmacı

Tarih: 2020-Devam ediyor.

Firma Adı : Tübitak Bilgem İleri Teknolojiler ve Araştırma Merkezi

Pozisyon : Araştırma Geliştirme / ARGE Stajyeri

Tarih: 2018

YÜKSEK LİSANS TEZİNDEN TÜRETİLEN ÇALIŞMALAR

- Kavuran G., Gokhan Ş., Yeroglu C., COVID-19 and Human Development: An approach for classification of HDI with Deep CNN (SCI kapsamındaki bir dergiye gönderildi ve hakem Değerlendirmesi aşamasında)
- Gokhan Ş., Kavuran G., Yeroglu C., Bayes optimizasyonuna dayalı hiper parametrelili sınıflandırıcı ile COVID-19 ve İnsani Gelişme Endeksi arasındaki ilişkinin belirlenmesi, International Conference On Emerging Sources In Science, İstanbul, Türkiye, May 26-27, 2022