



Covid19 Tanısı: X-Işını ve Kan Ölçüm Verileri Arasında Karşılaştırmalı Yaklaşım

Covid19 Diagnosis: Comparative Approach Between Chest X-Ray and Blood Test Data

Ali Emre ÖZTAŞ

Elektronik ve Haberleşme Mühendisliği
Bölümü
İstanbul Teknik Üniversitesi
İstanbul, Türkiye
oztas18@itu.edu.tr
ORCID: 0000-0003-1323-8176

Dorukhan BONCUKÇU

Elektronik ve Haberleşme Mühendisliği
Bölümü
İstanbul Teknik Üniversitesi
İstanbul, Türkiye
boncukcu17@itu.edu.tr
ORCID: 0000-0003-0393-5605

Ege ÖZTEKE

Elektronik ve Haberleşme Mühendisliği
Bölümü
İstanbul Teknik Üniversitesi
İstanbul, Türkiye
ozteke18@itu.edu.tr
ORCID: 0000-0003-4865-9325

Mahir DEMİR

Elektronik ve Haberleşme Mühendisliği
Bölümü
İstanbul Teknik Üniversitesi
İstanbul, Türkiye
demirma18@itu.edu.tr
ORCID: 0000-0003-4845-5074

Arzu MİRİCİ

Göğüs Hastalıkları Tıp Fakültesi
Çanakkale Onsekiz Mart Üniversitesi
Çanakkale, Türkiye
arzumirici@comu.edu.tr
ORCID: 0000-0002-7189-9258

Pinar MUTLU

Göğüs Hastalıkları Tıp Fakültesi
Çanakkale Onsekiz Mart Üniversitesi
Çanakkale, Türkiye
pinarmutlu@comu.edu.tr
ORCID: 0000-0002-7496-0026

Öz

Covid-19 virüsü dünya üzerinde büyük bir etki bırakmıştır ve yayılmaya devam etmektedir. Daha fazla yayılmasını engellemek için koronavirüs hastalarına erken tanı koymak oldukça önemlidir. Her ne kadar akciğer X-Işını görüntüsü tanısı ile çözüm en hızlı ve en kolay yöntem olsa da ortalama bir radyoloğun X-Işını verilerini kullanarak tanı koymadaki doğruluğu tamamen mesleki deneyimine dayanmaktadır. Yani, daha deneyimsiz radyologların hata yapma olasılığı daha fazladır. Bu nedenle tutarlı sonuçlar verebilen bir yapay zekâ modeli üretilmesi istenmektedir. Çalışmamızda göğüs X-Işını görüntüleri ve sıradan kan ölçüm verileri kullanılarak sınıflandırma yapılmış ve sonuçları karşılaştırılmıştır. X-Işını verileri hem açık kaynak çalışmalardan hem de yerel bir hastaneden anonim olarak toplanmıştır ve yaklaşık 7200 görüntüye sahiptir. Kan ölçümü sonuçları da yine aynı yerel hastaneden toplanmıştır. Göğüs X-Işını verilerinin tanısı için yaygın olarak kullanılan evrişimsel sinir ağı algoritmalarından ResNet, SqueezeNet, DenseNet ve VGG kullanılmıştır.

Sonuçlar, SqueezeNet modelinin daha yüksek AUC değeri vermesiyle birlikte, diğer algoritmaların da %85 üstünde bulma ve tutturma değeri sağladığını göstermektedir. Covid-19'un kan ölçümlerinden tanısı için ise çok katmanlı yapay sinir ağı ve destek vektör makinası kullanılmıştır. Kan ölçüm verileri kullanarak sınıflandırma kısıtlı bir veri kümesi üzerinde yapılmış olsa da yapay sinir ağı ve destek vektör makinası için doğruluk oranları sırasıyla %76 ve %82 olarak bulunmuştur. Genelleme yapılırsa X-Işını yoluyla tanının kan ölçümü yoluyla yapılan tanıdan daha uygulanabilir olduğu ve Covid tanısında yapay zekânın insanlardan daha doğru sonuç çıkardığı sonucuna ulaşılmıştır.

Anahtar Sözcükler: Covid-19, tanı, gözetimli öğrenme, derin öğrenme, göğüs X-Işını, kan ölçümü, aktarmalı öğrenme]

Abstract

The Covid-19 virus has made a major impact on the world and is still spreading rapidly. A reliable solution to prevent further damage, early diagnosis of coronavirus patients is incredibly important. Although the diagnosis with lung X-Ray image is the fastest and easiest method, the accuracy of an average radiologist in diagnosing using X-Ray data is completely based on his/her professional experience. That is, less experienced

radiologists are more likely to make mistakes. Thus, it is desirable to achieve an accurate artificial intelligence model for this. Throughout this study, chest X-Ray data and blood routine test data are utilized and compared. X-Ray data was collected anonymously from both open-source studies and a local hospital and has approximately 7200 images. The results of the blood test were also collected from the same local hospital. Popular convolutional neural network algorithms ResNet, SqueezeNet, DenseNet, and VGG were used for the diagnosis of chest X-Ray data. The results show that while the SqueezeNet model gives a higher AUC value, other algorithms also provide sensitivity and specificity above 85%. A multilayer artificial neural network and support vector machine was used for the diagnosis of Covid-19 from blood test data. Although classification using blood test data was done on a limited data set, an accuracy rates for ANN and SVM are respectively %76 and %82. If generalization is made, it has been concluded that diagnosis via X-Ray is more applicable than diagnosis via blood test, and artificial intelligence is more accurate than humans in diagnosing Covid-19.

Keywords:Covid-19, diagnosis, supervised learning, deep learning, chest X-Ray, blood test, transfer learning

1. Giriş

2019 yılında korona virüsün ortaya çıkması ve tüm dünyaya hızlı biçimde yayılması ile hiçbir ülkenin sağlık kurumunun bu düzeydeki bir salgın için hazır olmadığı görülmüştür. Bu makalenin yazıldığı tarihe kadar salgın nedeni ile 269 milyon olay yaşanıp 5,29 milyon insan hayatını kaybetmiştir. Her ne kadar bazı ülkeler bu duruma karşı daha katı ve dikkatli önlemler alıp salgından dolayı artan olay sayısında daha büyük bir kontrole sahip olsa da nüfusu çok olan veya gelişmemiş ülkelerde sağlık çalışanları sürekli olarak fazla çalıştırılmaktadır. Dahası, Covid-19 vakalarının sayısı arttıkça ölçümlerin maliyeti hem hükümetlere hem de vatandaşlarına ağır bir yük olmaktadır.

Şu aşamada, ölçümler X-ışını taramaları, bilgisayarlı tomografi ve PCR-ICT ölçümleri ile yapılmaktadır. Bu ölçüm yöntemleri arasında PCR-ICT ölçmesi basitliğinden, hızından ve doğruluğundan dolayı en fazla kullanılandır. Ancak, bu ölçmenin tutturma değeri %95 ila %99,7 arasında, duyarlılığı ise %70 ila %98 arasında sınaama araçlarına göre değişmektedir. DSÖ'ye göre, bulma değeri %80 ve tutturma %90 olan ölçümler kabul edilebilir; bulma değeri %90 ve özgüllüğü %99 olan ölçümler ise istenilen seviyede olarak belirlenmiştir [5].

Dahası, gerçek dünyadaki radyologların deneyimleri ile ilgili yanılma payı göz önüne alındığında, tomografi ile yapılan tanıları doğruluk oranı daha yüksek olduğu için, hastaları X-ışını ölçümlerine göre daha yüksek oranda radyasyona maruz bırakmasına karşın tercih edilmektedir. Sadece X-ışını göğüs taramaları göz önüne alındığında, deneyimi 10 yıldan fazla olan radyologların 10 yıldan az olan radyologlarla karşılaştırılması hastalığı doğru okuma oranlarının sırasıyla %83,7 ve %76 olduğunu göstermektedir [2].

Sonuç olarak, günümüzde çoğu araştırmacı hem sağlık çalışanlarının iş yükünü azaltmak amacı ile hem de tam

anlamda bir kontrol mekanizmasına gereksinim duymadan hastaları doğru tanılabilmek için bir derin öğrenme ağı tasarlamaya çalışmaktadır.

Covid19'u tanılabilmek amacı ile derin öğrenme içeren tıbbi görüntü işleme teknikleri kullanılmıştır. Genel olarak, derin öğrenme ile görüntü işleme uygulamaları için veri kümesi oldukça büyük bir sorundur. Bu sorunu çözebilmek amacıyla aktarımlı öğrenme ve önceden eğitilmiş CNN modelleri sıklıkla kullanılır. Yakın zamandaki araştırmacılar da bu yöntemleri kullanmış ve aşağıda özetlenen umut verici sonuçlarla karşılaşmışlardır.

Abbas ve ark. önceden eğitilmiş CNN modelleri kullanarak "Decompose (Ayrıştır), Transfer (Aktar), Compose (Birleştir)" (DeTraC) adı verilen, görüntü işleme için kullanılan bir CNN modeli geliştirmişlerdir. Bu modelin kullanılmasının nedeni homojen olmayan ve düzensiz verilerin deneme ve eğitim işlemlerinin daha doğru bir şekilde gerçekleştirilmesini sağlamaktır. Kullandıkları modeller arasında, en iyi sonucu VGG19 modeli ile elde etmişlerdir (%93,1 doğruluk, %100 duyarlık ve %85,18 tutturma). Ancak, kullandıkları veri kümesi çok küçük olduğu için, (105 Covid | 80 Covid olmayan | 11 SARS) sonuçların aşırı uyumlu olma olasılığı vardır [1].

Wang ve ark. yaptıkları araştırmada kendi geliştirdikleri açık kaynaklı derin öğrenme sinir ağı tasarımı ile iyi bir doğruluk oranına sahip sonuçlar elde etmişlerdir [16].

Narin ve ark. önceden eğitilmiş 5 farklı CNN modelini kullanmış ve en iyi sonucu ResNet50 ile elde etmişlerdir. Araştırmalarında, 3 farklı veri kümesi için 3 farklı model kullanmışlardır. Kullandıkları modeller Covid olan hastalar ile Covid olmayan hastalar arasında bir ayırım yapmamaktadır. Bir başka deyişle, yaptıkları araştırma kullandıkları modellerin ikili sınıflandırma yapmasından dolayı tıbbi tanıya uygun değildir. Ayrıca, kullandıkları modellerden bazıları %100 duyarlık ve/veya özgüllüğe sahiptir. Bu durum aşırı uyumlu olma olasılığını düşündürmektedir [13].

Alazab ve ark. X-ışını tanısı için VGG16 modelini kullanmışlar ve %99 F değeri elde etmişlerdir [3].

Minaee ve ark. önceden eğitilmiş 4 farklı CNN modeli kullanmışlar ve en doğru sonuçları ResNet18 ve SqueezeNet modellerinden elde etmişlerdir. Ayrıca, olası bulaş bölgelerini saptayan ısı haritalarını da çalışmalarında üretmişlerdir. Yaptıkları çalışmadaki eksiklik, Covid hastaları için toplamış olukları verileri tek bir kaynaktan almış olmalarıdır. Bu durum genelleştirme yapmayı engeller. Kendi çalışmamızda Minaee ve ark. yaptığı çalışmadaki yaklaşımı temel aldık. Bunun nedeni öbür çalışmalar ile karşılaştırıldığında yönteminin basit olması ve oldukça iyi doğruluk oranları sağlamasıdır [12].

Ayrıca sıradan kan ölçümleri Covid-19'un ilk tanısı için bir yöntem olarak kullanılabilir. PCR ölçümleri, sonuçlarının alınmasının 3-4 saat sürdüğü ve daha pahalı olduğu için hemogramlara oranla zaman açısından verimsizdir. Ayrıca gelişmiş donanıma ve belirli malzemelere sahip laboratuvarlarda, nitelikli personele gerek duyulmaktadır [17]. Öte yandan, sıradan kan ölçümleri bir hematoloji analizörü ile 30 dakika ile 2 saat arasında hızlı bir şekilde iletilir [25].

Bu nedenle bazı arařtırmacılar, Covid-19 tanısı için sıradan kan ölçümlerinin uygulanabilirliđi üzerine çalıřıyor.

Brinati ve ark. dođrulukları %82-86 ve duyarlılıkları %92-95 arasında olan iki farklı model geliřtirdi [17].

Banerjee ve ark. hastane servislerindeki hastaları ve hastaneye kabul edilmeyen hastaları karşılařtıran ađlar modellediler. İlk model için yaklaşık %90 kesinlik ve %82 ile %92 arasında duyarlık elde ettiler. İkinci model ise kesinlikleri %82-87 ve duyarlılıkları %43-65 olarak tahmin etti [18].

Kullandıkları Albert Einstein Hastanesinden alınan veri kümesinde çok fazla eksik deđer olmasına karşı Aljame ve ark. topluluk modelleri (ensemble model) ile %99,88 kesinlik elde etmişlerdir. Dolayısıyla bu modellerin güvenilirliklerinin dođrulanması açısından yeterli boyutta bir veri kümesi ile sınanmasının gerekli olduđu düşünölmektedir [19].

Bu çalıřmanın dört temel katkısından söz edebiliriz. Öncelikle, bu çalıřma Covid tahmini için X-Iřını verileri ya da kan ölçüm verilerini kullanan sınıflandırıcıların başarımlarını karşılařtırmayı göstermektedir. İkinci olarak, kan tanısı için yeni bir model geliřtirilmiş ve mevcut modeller düzenlenerek daha iyi sonuçlar elde edilmiştir. Dahası, veri kümesi farklı kaynaklardan veri toplanarak oluşturulmuş, sonuçların daha genel ve güvenli olması sağlanmıştır. Son olarak, X-Iřını görüntüleri 2 farklı veri kümesine ayrılarak iki sınıf için de ayrı evriřimsel sinir ađı modelleri eğitilmiştir. İlk veri kümesi sınıfı Covid olan hastalar ile hem tamamen sağlıklı hem de Covid dışında başka hastalıkları da olan hastaları sınıflandırmaktadır. Kullanılan ikinci veri kümesi sınıfı ise sadece tamamen sağlıklı olan insanlar ile sadece Covid hastalarını sınıflandırmaktadır.

Bu makale ařađdaki gibi düzenlenmiştir. Veri içerikleri ve verinin hazırlanması 2. Bölümde açıklanmıştır. Kullandığımız yöntemlerin anlaşılması için gerekli olan teorik bilgi 3. kısımda açıklanmıştır. Deneysel sonuçlar ile deđerlendirme sonuçları 4. kısımda gösterilmiş ve yorumlanmıştır. 5. kısım makaleyi sonuçlandırmaktadır.

2. Veri Kümesi

2.1 Göđüs X-Iřını Veri Kümesi

X-Iřını görselleri için Kaggle platformunda açık kaynak olarak paylaşılan 3616 Covidli X-Iřını akciđer görseli içeren veri kümesi kullanılmıştır [20, 21]. Covidli olmayanlar için Shervin Minaee'nin kullanmış olduđu Chex-Pert veri kümesi kullanılmıştır [12]. Bu veri kümesinde hem tamamen sağlıklı insanlar hem de 13 sınıfa bölünmüş olmak üzere çeřitli hastalıkları olan insanların X-Iřını görüntüleri mevcuttur. Toplamda bu veri kümesinde Covidli olmayan 5000 X-Iřını görüntüsü mevcuttur. řekil-1 sırasıyla Covid hastasına ve sağlıklı bir insana ait olan X-Iřını görüntülerini göstermektedir. Veri kümesinde dengeyi sağlamak amacıyla bunlardan 3600 tanesi rastgele seçilerek kullanılmıştır. Ayrıca, kişilerin bilgileri tamamen anonim olmak üzere Çanakkale Onsekiz Mart Üniversitesinden 151 Covidli ve 35 Covidsiz X-Iřını görüntüsü elde edilmiştir. Çizelge-1 elde edilen verilerin eğitim, deneme ve dođrulama amaçlı ayrılan miktarlarını belirtmektedir.

Çizelge-1: X-Iřınında kullanılan hastaların sağlık durumunu ve verilerin ne amaçla kullanıldığını belirten çizelge

	Deneme	Eđitim	Dođrulama
Covid	753	2636	377
Covid Olmayan	757	2649	379
Tamamen Sağlıklı	720	2520	360



řekil-1: Soldaki bir Covid hastasına sağdaki sağlıklı bir insana ait olmak üzere iki X-Iřını görüntüsü

Verilerimiz iki farklı kaynaktan elde edildiđinden dolayı, veri formatları ve boyutları farklıydı. Daha tutarlı bir eğitim aşaması için verilerin ön işlemeden geçirilmesi gerekiyordu. Bu nedenden dolayı, öncelikle tüm görselleri görüntü işleme araçları kullanarak 224x224 beneklik bir boyuta getirdik. 224x224 boyutunun seçilme nedeni kullanılan derin öğrenme modellerinin 224x224 boyutunda girdi kabul etmesidir. Özgün resimlerdeki en boy oranları resimden resime deđiřtiđi için en boy oranı korunmamıştır. Ayrıca, bazı resimler bazı belirteçler ve sağlık terimleri içeriyordu. Yanlıř sınıflandırmayı engellemek için bunlar kaldırıldı. Son olarak, topladıđımız bütün Covidli ve Covidsiz hastaların veri kümesini birleřtirdik ve yazdıđımız Python kodu yardımıyla %70 eğitim, %20 test, %10 dođrulama kümesi olmak üzere rastgele bölümlere ayırdık.

2.2 Kan Testi Veri Kümesi

Sıradan kan ölçüm verileri üzerinden Covid tanısının uygulanabilirlik arařtırmasını ve de göđüs X-Iřını tanısı ile karşılařtırılmasını yapabilmek için, Çanakkale Onsekiz Mart Üniversitesi Göđüs Hastalıkları Bölümü'nden etiketlenmiş bir veri kümesi alınmıştır. Veri kümesi koronavirüs ölçümü yaptırmış olan hastaların sıradan kan ölçümü verilerini içermektedir. Veri kümesi hastaların kişisel verilerini içermemektedir. Kan ölçümleri üzerinden koronavirüs tanısı yapabilmek için, bu veri kümesinden gereksiz olan parametreler ve hatalı hasta verileri çıkarılmıştır. Çizelge-2 veri kümesinin son halini modellemektedir.

Çizelge-2: Kan testi veri kümesi

	Covid Hastası	Covid Olmayan	Toplam
Hasta sayısı	53	78	131

Hastaneye başvuru anında hastaların genel durumu hastalığın gelişimi ile ilgili bilgi verebilmektedir. Normalde dakikada 16-20 olan solunum sayısının bu hastalarda artması hastalıktan etkilenmenin fazla olduğunu göstermektedir.

Hastalığın gelişimini izleyebilmek ve uygun tedaviyi uygulayabilmek için bazı kan değerleri bakılmaktadır. Lökositler kanın enfeksiyonla savaşan hücreleridir. Lökositlerin alt grupları olan nötrofileozinofil ve lenfositler enfeksiyonun nedenine ve yoğunluğuna göre farklı sayı ve oranlarda görülebilirler. Bu değerler hastalığın ağırlığı ve kaynağı ile ilgili önemli bilgiler vermektedir. Lenfositler genellikle virüs kaynaklı enfeksiyonlarla savaşta önemli rol oynamaktadır. Viral enfeksiyonlarda artması beklenen lenfosit sayılarının düşük olması Covid enfeksiyonuna özel gibi görünmekle beraber hastalığın ağırlığı ile yakın ilişkili bulunmuştur. Sedimantasyon, crp ve ldh gibi değerler pnömönide mikrobik enfeksiyonun yoğunluğunu gösteren değerlerdir. Bakteriyel pnömönilerde arttığı gibi Covid pnömönisinde de artış görülebilir [19].

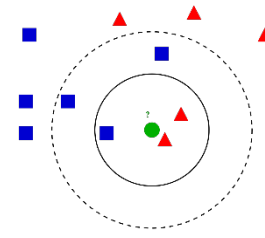
Trombositler kanın pıhtılaşmasını sağlayan hücrelerdir. Fibrinojen ve ddimer ise kanın pıhtılaşmaya yakınlığını gösteren değerlerdir. Sarscov2 virüsünün pıhtılaşma sistemini bozarak pıhtılaşmaya yakınlık yaptığı çalışmalarla gösterilmiş olup bu değerlerin yüksek olması halinde kan sulandırıcı tedaviler ile beyin, akciğer gibi önemli organlara pıhtı atması önlenmeye çalışılır. Ayrıca bu değerlerdeki artışlar hastalığın kötüleştiğini de göstermektedir [19].

Sarscov2 virüsü solunum yolu ile akciğer tutulumu yaparak pnömöniiye neden olmaktadır. Pnömöniiye neden olan diğer nedenlerden farklı olarak akciğer dışında böbrek kalp karaciğer bağırsaklar damar duvarı gibi birçok organa ve hücreye de tutunarak bozukluklar yapabilir. Alt ast karaciğer hücre hasarında artış gösteren enzimlerdir. Üre kreatinin böbreklerin fonksiyonunu göstermede önemli değerlerdir. Troponin kalp dokusu hasarlandığı zaman artan bir enzimdir. Bu değerlerdeki artışlar ilgili organın Covid ile ilişkili hasarı hakkında bilgi verebilir. Tüm bu organları etkileyerek vücudun mineral dengesini bozabilmektedir. O yüzden kan sodyum ve potasyum değerleri hastanın gelişinde ve takibinde kontrol edilir [19].

Anlatılan bilgiler doğrultusunda kan ölçme modelini eğitmek için 24 parametre seçtik. Bunlardan bazıları dakikadaki solunum sayısı ile nötrofil, eozinofil, lenfosit, fibrinojen, d-dimer, ALT, AST, üre, kreatin, sodyum ve potasyum değerleridir.

2.2.1 k-NN Veri Ataması

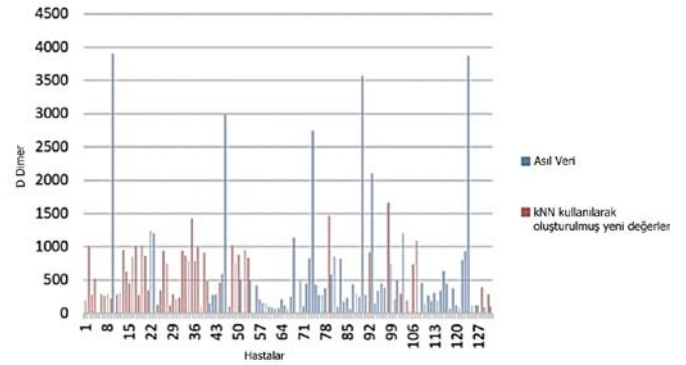
Hastaneden alınan veri kümesinin ilk halinde bazı hastalarda eksik veriler olduğundan, bunu gidermek için k-NN (k-Nearest Neighbours) algoritmasını kullandık. k-NN genellikle örnekler arasındaki ilişkileri bulmak için hızlı ve kolay bir yaklaşım sağlar. Her örneğin eksik verileri, o belirli noktanın Öklid uzaklığı ile hesaplanan en yakın k tane komşusunu bularak ve ortalamalarını alarak oluşturulur. Bu algoritmanın bir dezavantajı, k komşuları için tüm veri kümesini araması ve daha büyük veri kümeleri için uygun olmamasıdır [16]. Şekil-2'de bu algoritmanın çalışması görselleştirilmiştir.



Şekil-2: k-NN algoritması

$$d(p, q) = \sqrt{\sum_{i=1}^n (q_i - p_i)^2} \quad (1)$$

Veri kümesindeki 131 örnek için, bazı öznitelikler 20'den fazla eksik hücreye sahipti ve modelin genel davranışını bozmayacak yeni veriler oluşturmak için kNN eksik veri atamasını kullandık. Şekil-3, ilk kayıp değerleri ile örnekler üzerinde çizilen ve daha sonra k=5 ve k=10 olduğunda kNN yöntemiyle oluşturulan D-Dimer özniteliği verilerini göstermektedir. Çoğu model için genellikle iyi bir değer olarak kabul edildiğinden çalışmamıza k=10 ile devam etmeyi seçtik ve bizim durumumuzda da daha iyi çalıştığını gözlemledik. Uç değerlerin üretilen verileri etkilemediği de görülebilir.



Şekil-3: k-NN ile eksik değer ataması

3. Önerilen Modeller ve Eğitim

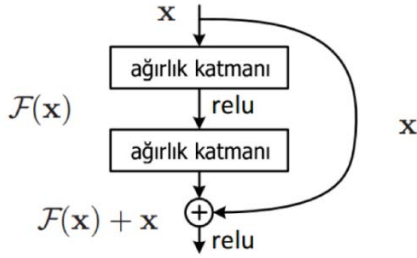
3.1 Göğüs X-Ray Tanısı Modelleri

Kullanılabilecek modeller hakkında kaynak taraması yapıldıktan sonra, daha önceki çalışmalarda en iyi sonuç veren modellerden ResNet18, Resnet50, SqueezeNet, DenseNet ve VGG modellerin kullanılması kararlaştırılmıştır. Eğitim için aktarmalı eğitim modeli kullanılmıştır. Bu yaklaşım ile daha önce ImageNet veri kümesinde ön eğitimden geçmiş olan modellerin ağırlık ve eğilim parametreleri ince ayardan geçirilmiştir.

3.1.1 ResNet

ResNet algoritması için biz hem ResNet18 hem de ResNet50 kullandık. Bu modeller günümüzde "residual" adı verilen önceki katmanlardan kısa yol bağlantı içeren modellerden en bilineni olan 18 ve 50 katmanlı uyarlamalarıdır. Bu model kaybolan gradyanlar problemine bir çözüm getirmektedir. Derin öğrenme algoritmalarında ağırlığı güncelleme yöntemi gradyan indirgeme yöntemi olarak adlandırılır. Bu yöntem zincir kuralını kullanarak hata fonksiyonun model parametrelerine göre kısmi türevini alarak parametreleri günceller. Bazı durumlarda aktivasyon fonksiyonun türevi -1

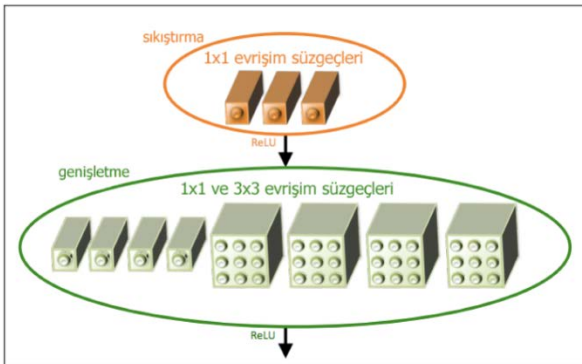
ile 1 arasında olmaktadır. Bu nedenden dolayı zincir kuralı kullanımı ile türevlerin çarpılması giderek oranı azaltır ve sonunda parametreler hiç güncellenemez ve model eğitimi durur. Bu kaybolan gradyenler problemi olarak adlandırılır. ResNet algoritması bu problemi çözmek için başlangıç katmanlarından daha sonraki katmanlara bağlanan bir kestirme bağlantı oluşturan özel bir yöntem kullanır. Bu yapı, gradyenlerin aşırı artmasını veya kaybolmasını engelleyerek öğrenme sürecine katkıda bulunur. Bu kestirme yollara "Identity Mapping" adı verilmiştir. Şekil-4'te bu işlem gösterilmiştir. ResNet50 ve ResNet101 de aynı yöntemi daha fazla katman kullanarak uygular [10].



Şekil-4: ResNet18'in kimlik eşleme katmanları

3.1.2 SqueezeNet

SqueezeNet, modelin boyutunu küçültmek amacıyla girdi değerlerini sıkıştıran bir algoritmadır. SqueezeNet'in makalesinde yazılan bilgiler, bu sıkıştırma işlemine rağmen doğruluk oranının korunduğu ve AlexNet ile yaklaşık olarak aynı değerde olduğunu belirtmektedir. Bu doğruluk oranı AlexNet'e göre 50 kat daha az parametre ile gerçekleştirilmiştir ve SqueezeNet'in toplam boyutu 0,5 MB'den küçüktür. SqueezeNet sıkıştırma için 3 ana yöntem kullanmaktadır. Öncelikle 3×3 'lük konvolüsyon çerçevelerini 1×1 olanlarla değiştirir. Daha sonra, girdi kanallarının sayısını düşürür. En son olarak, daha geniş aktivasyon haritaları elde etmek için alt örnekleme yöntemlerini kullanır. Bu işlemler "Fire Module" adı verilen bir kapsülde gerçekleşir. Aşağıdaki şekil "Fire Module"nin çalışma yöntemini göstermektedir. Şekil-5 SqueezeNet'in genel gösterimi içerir [9].



Şekil-5: SqueezeNet model mimarisi

3.1.3 DenseNet

DenseNet giriş ve çıkış katmanlarına yakın olan katmanlar arasına kısa yollar ekleyerek evrişimsel sinir ağlarının daha derin, tutarlı ve verimli olacağını göstermiştir. Bunu başarmak için L katmanlı bir ağda her katman ile ileriki katmanlar arasında Şekil-6'da gösterildiği gibi toplamda $C(L,2)$ tane

doğrudan bağlantı kurulmuştur. Sonuç olarak her katman çıkış üretmek için kendi girişiyle birlikte önceki katmanlardan çıkarılan öznelik haritasını kullanır. DenseNet'in avantajları kısa yol bağlantı içerdiğinden kaybolan gradyanlar problemine dayanıklı olması, güçlü öznelik yayılımı, özneliklerin tekrar kullanılması ve parametre sayısının azaltılmasıdır.



Şekil-6: DenseNet model mimarisi

3.1.4 VGG

Klasik evrişimsel sinir ağı modellerinden biri olan VGG, derin bir yapıya sahiptir. Bu ağda, 3×3 lük, göreceli olarak küçük bir evrişim filtresi kullanıldığı için 16-19 derinliğine kadar ulaşılabilmiştir. Biz çalışmamızda 16 katmanlı olan türünü kullandık.

3.1.5 Model Eğitimi ve Parametreler

Kullandığımız ağların çıktı katmanlarında çıktıları olasılıksal değerlere dönüştüren (2)'de verilen "Softmax Regression" yöntemi kullanılmıştır. Bu nedenden dolayı eğitim esnasında hedef olasılık değerleri ile hesaplanan olasılık değerleri arasındaki hatayı hesaplayan (3)'de verilen "Cross Entropy Loss Function" kullanıldı. Bu hata fonksiyonunu en küçük değere ulaştırmak için (4)'de verilen Stokastik Gradyan İnişi optimizasyon algoritması kullanılmıştır. Bu algoritma, ağırlık ve eğilimleri hata fonksiyonun kendilerine göre kısmi türevini alarak günceller. Bu güncelleme değeri "Öğrenme hızı" denilen bir parametreye bağlıdır. Daha önceki çalışmalarda bu hız 0,001 olarak belirlendiği için bu değer bizim çalışmamızda da kullanılmıştır.

$$f(s)_i = \frac{e^{s_i}}{\sum_j e^{s_j}} \quad (2)$$

$$CE = -\sum_i t_i \log f(s)_i \quad (3)$$

$$\theta_j = \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta) \quad (4)$$

Gradyan inişi yönteminde en düşük olan kayba yakınsarken, hata fonksiyonunu ekstremum noktasından uzaklaştıracak güncelleme değerleri ile karşılaşabiliriz. Bu problemi çözmek için, Şekil-7'de gösterildiği gibi daha önceki ağırlık ve eğilimlerin güncelleme değerlerini de göz önüne alıyoruz. Bunu yapmak için SGD'nin her bir aşamasında, daha önceki gradyan iniş değerini, "momentum katsayısı" adlı bir katsayı ile çarpıp şu anki değer ile topladığımızda elde edilen sonuca göre model parametrelerini güncelliyoruz. Öğrenme oranındaki sebeplere benzer olarak, momentum katsayısı 0,9 olarak alınmıştır.

Yakınsamaya kadar tekrar {

$$\nu \leftarrow \boxed{\eta * \nu_j} - \alpha \nabla_{\omega} \sum_1^m L_m(\omega)$$

$$\omega_j \leftarrow \nu_j + \omega_j$$

}

Şekil-7: SGD'de momentum kullanımını gösteren görsel

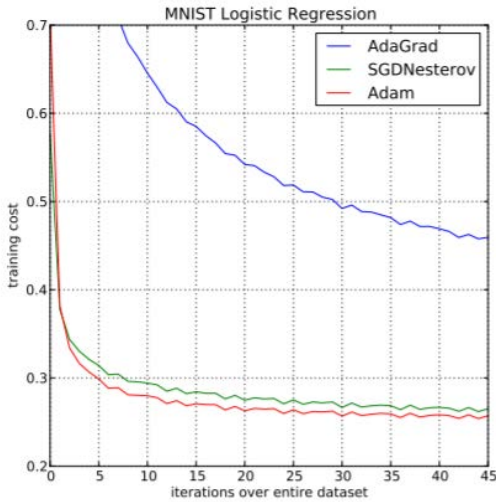
Modelimizi 20 yığın (batch) ve 100 devir (epoch) olacak şekilde eğittik. Modeller Google Colab platformunda Tesla K80 GPU kullanılarak CUDA teknolojisi ile eğitildi ve test edildi. Eğitim esnasında doğruluğu tespit etmek için modelin eğitim esnasında kullanmadığı doğrulama kümesindeki görseller kullanılarak doğrulama yapıldı.

3.2 Kan Ölçümleri ile Tanı

3.2.1 Yapay Sinir Ağı (ANN)

Hâlihazırda sayısal olan veri yöntemi ve gerekli işlem gücünün düşük miktarı göz önüne alındığında, bu sınıflandırma problemi için çok katmanlı yapay sinir ağı modeli kullanılmıştır.

Modelin oluşturulma aşamasında, aykırı değerlerin öğrenme sürecini etkilemesini engellemek ve de aşırı öğrenme sorunundan kaçınmak için gizli katmanlar 8 ve 4 birim içerecek şekilde belirlenmiştir. Hedef ve tahmin edilen sonuçlar arasındaki hatanın hesaplanması için, daha önce açıklanan çapraz entropi kaybı fonksiyonu kullanıldı. Buna ek olarak, geri yayılım aşamasında ağırlıkların güncellenmesi için düşük dereceli momentlerin tahminlerini oluşturan bir algoritma olan Adam optimizer kullanılmıştır. Şekil-8'de açıkça görülmektedir ki Adam algoritması diğer algoritmalarından daha etkilidir [11].



Şekil-8: Adam iyileştirici (optimizer) karşılaştırması [11]

3.2.2 Destek Vektör Makinesi (SVM)

İki gruba sınıflandırma problemlerinde kullanılabilecek olan Destek Vektör Makinesi makine öğrenmesi yöntemi, giriş verilerini çok büyük boyutlu öznelik uzayına eşler [25]. Bir DVM, bir veya daha fazla hiper-yüzey oluşturur. Destek vektörleri hiper-yüzeğe daha yakın olan veri noktaları olup, hiper-yüzeyin oryantasyonunu ve konumu belirler. Gamma parametresi bir eğitim örneğinin etkisinin ne uzağa ulaştığını ters orantılı olarak belirler. [26] Problemden doğrusal bir ayırım yapılamayacağı için RBF (Radial Basis Function) çekirdeği (kernel), dağılımını etkileyecek olan γ değeri 0,02 alınarak sınıflandırmada kullanıldı.

$$K(x, y) = e^{-\gamma|x-y|^2} \quad (5)$$

4. Deneysel Sonuçlar

4.1 Göğüs X-ışını

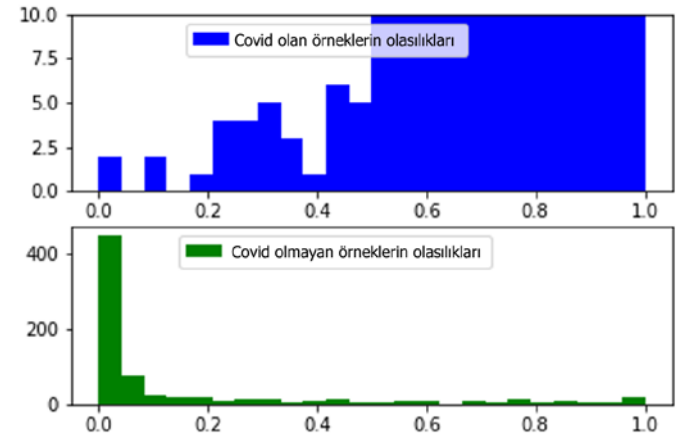
Çalışmamızın bu kısmını iki parçaya ayırdık. İlk kısım Covid olan hastalar ile hiçbir hastalığa sahip olmayan insanların X-ışını görüntülerinin tahminlerinin karşılaştırılmasını içermektedir. İkinci kısım ise Covid olan hastalar ile Covid dışı hastalıkları olan hastaların X-ışını görüntülerinin tahminlerinin karşılaştırılmasını içermektedir. Bu ayırmanın yapılmasının sebebi Covid dışındaki diğer hastalıkların X-ışını görüntülerinin, Covid olan hastaların X-ışını görüntülerine oldukça benzemesi ve bu durumun yapay sinir ağının yanlış sonuçlar üretmesine neden olmasıdır.

Değerlendirme ölçütü olarak bulma, tutturma ve Eğrinin Altındaki Alan (AUC) değerlerini kullandık.

$$\text{Bulma} = \frac{\text{Doğru tahmin edilen Covid-19 vakaları}}{\text{Bütün Covid-19 vakaları}} \quad (6)$$

$$\text{Tutturma} = \frac{\text{Doğru tahmin edilen Covid-19 olmayan örnekler}}{\text{Bütün Covid-19 olmayan örnekler}} \quad (7)$$

Öncelikle Covid hastalarının ve Covid bakımından sağlıklı olasılık değerlerinin histogramlarını elde ettik. Şekil-9'deki histogramlardan görülebileceği gibi kullandığımız modeller Covid olan ve Covid olmayanların X-ışını görüntülerini ayırt edebilmektedir.



Şekil-9: Covid-Covid değil sınıflandırmasında SqueezeNet modelinin olasılıksal çıktısının histogram değeri

Histogramın sol tarafında Covid olmayan örneklerin sağ tarafında ise Covid olan örneklerin varlığı modelin ayırım yapabildiğini göstermektedir. Her bir modelin her bir veri kümesi için en uygun olan eşik değeri bulunmuş ve bu değer için bulma ve tutturma değerleri elde edilmiştir. Çizelge-3'te bu sonuçlar görülebilir.

Modellerin test sonuçlarına göre 1. veri kümesi için ResNet50, SqueezeNet ve DenseNet %91 civarı bulma ve tutturma sergileyerek en başarılı sonucu vermişlerdir. 2. veri kümesi için SqueezeNet %91,6 bulma ve %91,8 tutturma ile en başarılı sonucu vermiştir.

Çizelge-3:Sınıflandırma modellerinin bulma ve tutturma değerleri

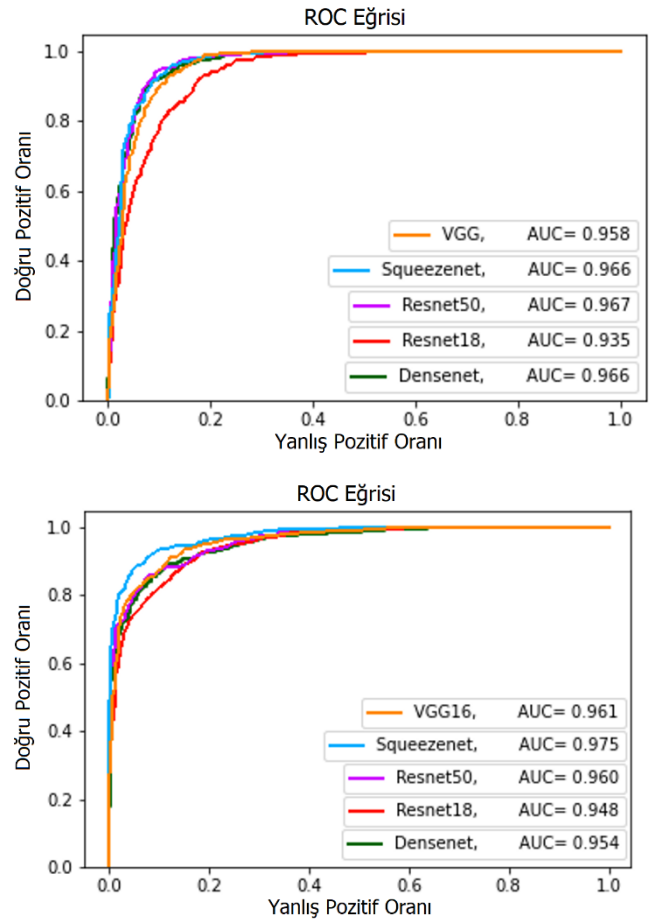
Sınıflandırma	Model	Bulma	Tutturma
Covid / Sağlıklı	ResNet18	%86,2	%86,9
	ResNet50	%88,3	%87,6
	DenseNet	%88,2	%88,6
	SqueezeNet	%91,6	%91,8
	VGG	%88,8	%88,6
Covid / Covid olmayan	ResNet18	%85,4	%85,7
	ResNet50	%91,8	%91,8
	DenseNet	%91,5	%91,4
	SqueezeNet	%91,1	%91,4
	VGG	%89,8	%90

Daha önceki çalışmamızda hem daha az veri kümesi hem de sadece ResNet18 ve SqueezeNet modelleri kullanılmış olduğundan dolayı tamamen sağlıklı insanlar ile Covid sınıflandırması daha başarılı çıkmıştı. Bu çalışmamızda yeni veriler ile yeni modeller eklenmiş ve yapılan ayrımın yeterince büyük bir fark üretmediği gözlenmiştir. Ayrıca önceki sonuçlarla uyumlu bir şekilde ResNet18 ve SqueezeNet modelleri sağlıklı insanlar ile Covid hastaları arasındaki sınıflandırmada daha iyi sonuç verdi.

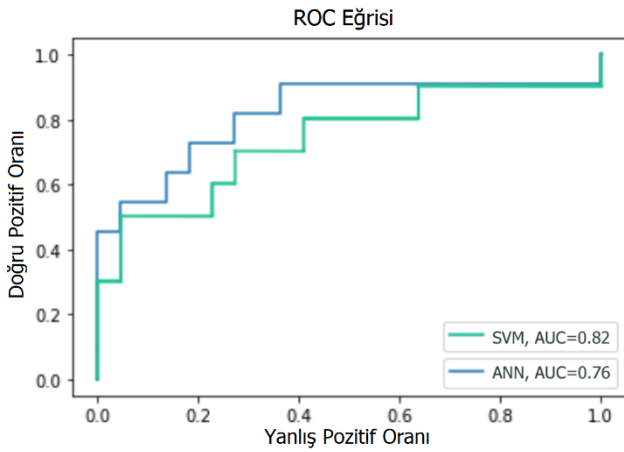
Bir modelin performansını sadece bulma ve tutturma değerlerine göre ölçmek yeterince doğru değildir. Farklı eşik değerlerine göre bu sonuçlar farklılık gösterebilir. Bu sebepten dolayı doğruluğu eşik değerinden bağımsız bir biçimde ölçmek gerekmektedir. Bu problemi çözmek için çeşitli yöntemler bulunmakla beraber, bu makalede "Receiver Operating Curves" (ROC) kullanılmıştır. Bu eğri, doğru pozitif örneklerin (True Positive) hatalı negatif örnekler (False Positive) oranının değişimini göstermektedir ve tüm eşik değerleri için modelin performansını değerlendirdiğinden eşik değerinden bağımsız bir performans ölçüsüdür. ROC eğrisinin sol alt köşesi özgülüğün 1, bulma değerinin 0 olduğu noktadır. Sağ üst köşesi ise bulma değerinin 1, özgülüğün 0 olduğu noktadır. ROC eğrilerinde eğrinin xy düzleminin sol üst bölgesine yakınlığı yüksek bulma ve yüksek tutturmanın bir arada olduğunu gösterir, başarılı modeller bu tür eğrilere sahiptir. Bazı ROC eğrileri birbirleriyle kesişmektedir ve hangisinin daha iyi olduğunu eğrilere bakarak söylemek yanılgıya sebep olmaktadır. Bu nedenle eğrinin altındaki alanı ifade eden AUC metriğine bakılmaktadır ve AUC değeri 1'e yakın olan modeller daha başarılıdır. Bu ölçüte göre Şekil-10'da görüldüğü üzere Covid ile Covid olmayan insanlar arasındaki sınıflandırmada en yüksek başarıyı 0,966 AUC değeriyle SqueezeNet ve DenseNet modelleri gösterdi. Covid ve sağlıklı insanlar arasındaki sınıflandırmada ise en yüksek başarıyı 0,975 ile SqueezeNet gösterdi. Bu sonuçlara göre SqueezeNet iki sınıflandırma için de en yüksek başarıyı gösteren modeldir.

4.2 Kan Ölçümü

Veri kümemiz sadece 131 kişinin kan testi verisini içerdiğinden, modelin daha iyi öğrenebilmesi için eğitim aşaması için mümkün olduğunca çok örnek kaydetmemiz gerekiyordu, aynı zamanda modelimizin kesinliğini doğru bir şekilde değerlendirebilmemiz için sınama kısmına yeterince örnek bırakmamız gerekiyordu. Rastgele dağıtarak verilerimizin ¼'ünü eğitim için ve ¼'ünü test için kullandık. Veri kümemiz çoğunlukla Covid olmayan vakalardan oluştuğu için bu dengesizlik, eğitim ve sınama veri kümeleri arasında eşit olmayan bir şekilde bölünürse modelin çıktısını da etkiler. Bu nedenle orta derecede ayrılmış bir veri kümesini kullandık. Şekil-9'daki ROC grafiğinde de görüleceği üzere kullandığımız ANN modeli ile 0.76 AUC ve de SVM modeli ile 0.82 AUC değerine ulaştık. AUC değerine göre SVM daha iyi sonuç vermiştir ancak bunun nedeni eğrinin sağ üst köşesinde eşik değeri 0'a yakın olduğunda modelin bulma değerinin ANN modeline göre daha çok artmasıdır. Pratikte seçilen eşik değeri eğrinin orta bölgesine denk gelecektir ve ANN bu bölgelerde bulma değerinin ve tutturma olarak SVM'den daha iyidir. Bu sebeple AUC değeri daha düşük olmasına rağmen ANN daha başarılıdır.



Şekil-10: Covid - Covid olmayan ve Covid - Sağlıklı sınıflandırmalarının ROC grafikleri



Şekil-11: Kan testi modelinin ROC grafiği

Bir diğer yaklaşım ise, medikal veriler ile uğraştığımız için eksik verileri k-NN ile oluşturmak yerine direkt silmek olacaktır. Bu durumda veri kümesindeki hasta sayısı ve de öznelik uzayı küçülmekle birlikte, aynı modelleri denediğimizde sonuçlarda dramatik bir değişiklik gözlenmedi ancak daha kötü sonuçlar elde edildi.

5. Sonuç

X-ışını sonuçları göz önüne alındığında, SqueezeNet diğer modellere göre belirli bir miktar daha iyi çalışmaktadır. Önceki çalışmamızda, Covid hastası olan hasta miktarının tüm veri kümesi karşılaştırıldığında sınırlı olması sınıflandırmayı dengesiz hale getirmekteydi. Bu çalışmada X-ışını veri kümesi dengelenip daha çok örnek ile modeller eğitildiği için daha güvenilir bir sonuç elde edilmiştir. Önceki çalışmamızda SqueezeNet ve ResNet18 modellerinin Covid-Sağlıklı sınıflandırması için daha iyi sonuç vermesi bize diğer hastalıkların modelin doğru karar vermesini engellediğini düşündürmüştü. Ancak bu çalışma yeterli veri, dengeli veri kümesi ve çeşitli modeller ile eğitim yapıldığında sonucun sınıflandırma türünden çok etkilenmediğini gösterdi. Kan testi modelinde ise kullanılan SVM ve ANN modellerinin birbirine karşı dramatik bir baskınlığı gözlenmemiştir ancak ANN modeli ROC eğrisine göre daha iyi sonuç vermiştir.

İki farklı yaklaşımının sonuçları karşılaştırıldığında, göğüs X-ışınları üzerine çalışmamızın sonuçlarının, rutin kan testi verileri ile çalışmamızın sonuçlarına kıyasla çok daha kullanılabilir olduğu görüldü. Bu proje yüksek doğruluk gerektiren medikal çalışmalarda kullanılması amacıyla geliştirildiğinden, özellikle kan testi modelinin sonuçlarının medikal alan için yeterli olmadığı da söylenebilir. Dahası, tam ayrıntılı bir çalışma yapmak için kan verisi yetersiz olsa da X-ışını üzerine tercih edilebilir bir yöntem olmadığı, ancak yine de araştırmacılara fikir vermesi açısından denenebilir olduğu sonucu çıkarılabilir. Sonuç olarak çeşitli makine öğrenmesi yöntemleri kullanılarak Covid-19 tanısı yapılması olasıdır ve modern tıp uygulamalarında deneysel olarak kullanılabilir.

6. Gelecek Çalışmalar

Çalışmamızda, hastaların Covid olup olmadığını X-ışını verileri ya da kan verilerini kullanarak sınıflandırma yapan modeller ile tespit ettik. X-ışını sonuçları ve kan sonuçları aynı

hastalardan gelmediğinden dolayı kullanılan modeller ve veriler her iki veri kümesini birden tahmin eden bir model üzerinde birleştirilemedi. Ayrıca, [12,16] da kullanılmış olan ısı haritası yöntemi projemizin zaman kısıtlamalarından dolayı kullanılamadı. Eğer bu geliştirmeler gelecek çalışmalarda kullanılabilir ise makine öğrenmesi modelleri sağlık çalışanlarının bilgisini ve bakış açılarını son derece yüksek bir oranda geliştirebilir.

7. Bilgilendirme

Göğüs X-ışını ve kan testi sonuçları, Çanakkale Onsekiz Mart Üniversitesi'nden etik komitesinin 2011-KAEK-27/2020-E.2000063714 izin numarası ile elde edilmiştir. Araştırma sırasında hastalar anonim bırakılmış ve kişisel bilgiler kullanılmamıştır.

Bu çalışma 6. Uluslararası Bilgisayar Bilimleri ve Mühendisliği Konferansında (UBMK 2021) sunulan [27] numaralı çalışmanın genişletilmiş halidir.

Kaynakça

- [1] Abbas, A., Abdelsamea, M. M., & Gaber, M. M. (2020). Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. *Applied Intelligence*, 1-11.
- [2] A. Cozzi, S. Schiaffino, F. Arpaia, G. Della Pepa, S. Tritella, P. Bertolotti, et al. Chest x-ray in the COVID-19 pandemic: radiologists' real-world reader performance *Eur J Radiol*, 132 (2020), Article 109272, 10.1016/j.ejrad.2020.109272
- [3] Alazab, M., Awajan, A., Mesleh, A., Abraham, A., Jatana, V., & Alhyari, S. (2020). COVID-19 prediction and detection using deep learning. *International Journal of Computer Information Systems and Industrial Management Applications*, 12, 168-181.
- [4] Scikit-learn: Machine Learning in Python, Pedregosa et al., *JMLR* 12, pp. 2825-2830, 2011.
- [5] Asai, T. (2020). COVID-19: accurate interpretation of diagnostic tests—a statistical point of view.
- [6] Cohen, J. P., Morrison, P., Dao, L., Roth, K., Duong, T. Q., & Ghassemi, M. (2020). Covid-19 image data collection: Prospective predictions are the future. *arXiv preprint arXiv:2006.11988*.
- [7] Haykin, S. S. (2008). *Neural Networks and Learning Machines*, third edition, Pearson.
- [8] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).
- [9] Iandola, F. N., Han, S., Moskewicz, M. W., Ashraf, K., Dally, W. J., & Keutzer, K. (2016). SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size. *arXiv preprint arXiv:1602.07360*.
- [10] Irvin, J., Rajpurkar, P., Ko, M., Yu, Y., Ciurea-Ilcus, S., Chute, C., ... & Ng, A. Y. (2019, July). Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 33, No. 01, pp. 590-597).
- [11] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- [12] Minaee, S., Kafieh, R., Sonka, M., Yazdani, S., & Soufi, G. J. (2020). Deep-Covid: Predicting Covid-19 from chest x-ray

- images using deep transfer learning. *Medical image analysis*, 65, 101794.
- [13] Narin, A., Kaya, C., & Pamuk, Z. (2020). Automatic detection of coronavirus disease (Covid-19) using x-ray images and deep convolutional neural networks. arXiv preprint arXiv:2003.10849.
- [14] Maguolo, G., & Nanni, L. (2020). A critic evaluation of methods for Covid-19 automatic detection from x-ray images. arXiv preprint arXiv:2004.12823.
- [15] Monard, Maria-Carolina. (2002). A Study of K-Nearest Neighbour as an Imputation Method.
- [16] Wang, L., Lin, Z. Q., & Wong, A. (2020). Covid-net: A tailored deep convolutional neural network design for detection of Covid-19 cases from chest x-ray images. *Scientific Reports*, 10(1), 1-12.
- [17] Brinati, D., Campagner, A., Ferrari, D. et al. Detection of COVID-19 Infection from Routine Blood Exams with Machine Learning: A Feasibility Study. *J Med Syst* 44, 135 (2020)
- [18] Abhirup Banerjee, Surajit Ray, Bart Vorselaars, Joanne Kitson, Michail Mamalakis, Simonne Weeks, Mark Baker, Louise S. Mackenzie, Use of Machine Learning and Artificial Intelligence to predict SARS-CoV-2 infection from Full Blood Counts in a population, *International Immunopharmacology*, Volume 86 (2020)
- [19] AlJame, M., Ahmad, I., Imtiaz, A., & Mohammed, A. (2020). Ensemble learning model for diagnosing COVID-19 from routine blood tests. *Informatics in Medicine Unlocked*, 21, 10
- [20] -M.E.H. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M.A. Kadir, Z.B. Mahbub, K.R. Islam, M.S. Khan, A. Iqbal, N. Al-Emadi, M.B.I. Reaz, M. T. Islam, "Can AI help in screening Viral and COVID-19 pneumonia?" *IEEE Access*, Vol. 8, 2020, pp. 132665 - 132676.
- [21] -Rahman, T., Khandakar, A., Qiblawey, Y., Tahir, A., Kiranyaz, S., Kashem, S.B.A., Islam, M.T., Maadeed, S.A., Zughailer, S.M., Khan, M.S. and Chowdhury, M.E., 2020. Exploring the Effect of Image Enhancement Techniques on COVID-19 Detection using Chest X-ray Images.
- [22] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700-4708).
- [23] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
- [24] Republic of Turkey Ministry of Health, SARS-CoV-2 Infection Adult Patient Treatment-Scientific Advisory Board Study (7 May 2021 -ANKARA)
- [25] Cortes, C., Vapnik, V. Support-vector networks. *Mach Learn* 20, 273–297 (1995). <https://doi.org/10.1007/BF00994018>
- [26] RBF SVM Parameters. scikit. (n.d.). Erişim tarihi: 23 Mayıs, 2022, Erişim adresi: https://scikit-learn.org/stable/auto_examples/svm/plot_rbf_parameters.html
- [27] A. E. Öztaş, D. Boncukçu, E. Özteke, M. Demir, A. Mirici and P. Mutlu, "Covid19 Diagnosis: Comparative Approach Between Chest X-Ray and Blood Test Data," *2021 6th International Conference on Computer Science and Engineering (UBMK)*, 2021, pp. 472-477, doi: 10.1109/UBMK52708.2021.9558969.