



COVID-19 Detection from Lung Tomography Images Using Deep Learning and Machine Learning Methods

Resul Bütüner and M. Hanefi Calp

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

August 27, 2020

Derin Öğrenme ve Makine Öğrenmesi Yöntemleri Kullanılarak Akciğer Tomografi Görüntülerinden COVID-19 Tespiti

Resul BÜTÜNER¹, M. Hanefi CALP²

1. Bilgisayar Mühendisliği Bölümü
Konya Adil Karaağaç MTAL
resul.butuner@ogr.erbakan.edu.tr

2. Yönetim Bilişim Sistemleri Bölümü
Karadeniz Teknik Üniversitesi
mhcalp@ktu.edu.tr

Öz

Koronavirüs (COVID-19), tüm dünyada çok kısa sürede yayılan ve ölümcül sonuçları olan bir salgın hastalıktır. Bu tür bulaşıcı hastalıkların insanlara zarar vermeden veya minimum zararlarla doğru bir şekilde tespit edilmesi ve gerekli tedavinin erken süreçte başlatılması gerekmektedir. Bu çalışmada, yapay zekâ ve makine öğrenmesi teknikleri ile farklı akciğer tomografisi görüntüleri içerisinde (COVID-19, viral zatürre, bakteriyel zatürre ve normal) COVID-19 olan görüntünün tespit edilmesi amaçlanmıştır. Bu bağlamda, yapay zekânın güncel tekniklerinden derin öğrenme yaklaşımları içerisinde Evrişimli sinir ağları (CNN) ve Derin sinir ağları (DNN) ile birlikte makine öğrenmesi algoritması olan K-En Yakın Komşuluk (KNN) yöntemi kullanılmıştır. Ayrıca, CNN yönteminde farklı optimizasyon ve aktivasyon fonksiyonları ve nöron sayılarının kombinasyonlarıyla modeller oluşturularak sonuçlar test edilmiştir. Böylece, CNN, DNN ve KNN yöntemlerinin görüntü tanıma ve sınıflandırma konusunda uygulamadaki potansiyeli görülmüş ve önerilen modelin başarısı elde edilen bulgularla ortaya konulmuştur.

Anahtar Kelimeler: Derin Öğrenme, Makine Öğrenmesi, CNN, DNN, KNN, COVID-19, Tomografi, Yapay Zekâ

Abstract

Coronavirus (COVID-19) is an epidemic disease that spreads all over the world in a very short time and has fatal consequences. Such infectious diseases must be detected correctly without harming people or with minimal harm, and the necessary treatment must be initiated early. In this study, it was aimed to detect the COVID-19 image from different lung tomography images (COVID-19, viral pneumonia, bacterial pneumonia and normal) with artificial intelligence and machine learning techniques. In this context, K-Nearest Neighborhood (KNN) method, which is a machine learning algorithm, has been used together with Convolutional Neural Networks (CNN) and Deep Neural Networks (DNN), which are among the current techniques of artificial intelligence, deep learning approaches. In addition, in the CNN method, the results were tested by creating models with combinations of different optimization and activation functions and neuron numbers. Thus, the application potential of CNN, DNN and KNN methods in image recognition and classification was seen and the success of the proposed model was demonstrated with the findings obtained.

Keywords: Deep Learning, Machine Learning, CNN, DNN, KNN, COVID-19, Tomography, Artificial Intelligence

1. Giriş

COVID-19, 2019 yılı Aralık ayında Çin'in Wuhan kentinde ortaya çıkmış, yarasalardan insanlara geçmiş, çok kısa zamanda tüm dünyaya yayılarak pandemiye dönüşmüş ve birçok ülkenin sağlık sistemini çökertmiş ciddi bir salgın hastalıktır. Şiddetli akut solunum sendromu COVID-19, bulaştığı canlılarda önemli düzeyde solunum yetmezliğine sebep olmakta ve hastalığın ilerlemesi halinde ise ölümcül sonuçlarla karşılaşmaktadır. Hastalığın en belirgin semptom ve enfeksiyon belirtileri; ateş, kuru öksürük, boğaz ağrısı, baş ağrısı, halsizlik, kas ağrısı, ishal ve nefes darlığı olarak bilinmektedir. Daha ileri durumlarda, şiddetli zatürreye yol açarak oksijen farklılığından kaynaklı akciğerlerde iltihaplanmalara ve çoklu organ yetmezliğine neden olmaktadır. Özellikle bu hastalık, kronik rahatsızlıkları olanlar, direnci veya bağışıklık sistemi zayıf olanlar, sigara kullananlar ve yaşlılar için çok daha tehlikeli ve yıpratıcı etkilere sahiptir [1-4].

COVID-19, genellikle virüsü taşıyan kişilerle nefes, el veya mukoza teması yolu gibi fiziksel temas ile bulaşmaktadır [5]. Hastalığın bulaşması sonrasında tedavi edilmesinde, genel olarak antibiyotik, sıtma ilacı, ateş düşürücü, öksürük ilacı ve ağrı kesici kullanılmaktadır. Virüse bulaşmış hastanın, hastalık düzeyine, şikâyetlerine ve ciddiyetine göre hastaneye yatırılması kararlaştırılmaktadır [3,4]. Dünya genelinde, COVID-19 virüslü hasta sayısı her geçen gün artmaktadır. ABD, İtalya ve İspanya gibi güçlü ülkeler bile bu virüsten yeteri kadar kendini koruyamamış ve büyük ölçüde etkilenmişlerdir. Bütün bu bilgiler ışığında, sağlık sistemini de dikkate alarak pandemiye tamamen önlemek veya en azından virüsün vereceği muhtemel zararları minimize etmek için hastalığın erken teşhisi çok büyük önem arz etmektedir. Yani, en azından şüpheli vakaların düşük hata ile doğru ve hızlı bir şekilde tespiti gerekmektedir [6].

Hâlihazırda COVID-19 tanı ve teşhisi amacıyla, gerçek zamanlı bir ters transkripsiyon-polimeraz zincir reaksiyonu (RT-PCR) tekniği yaygın olarak kullanılmaktadır. Erken evrelerde COVID-19 tanısını koymak için Bilgisayarlı Tomografi (BT) ve X-ışını gibi göğüs radyolojik görüntüleme yöntemleri tercih edilmektedir [1]. Söz konusu hastalığın çok hızlı yayılması ve birçok ülkede ölüm oranlarının artması etkili bir tedavi yönteminin geliştirilmesi gerektiğini ortaya koymaktadır. Bu nedenle teşhis, erken karantina ve takip dâhil olmak üzere hastalığın kontrolü zorunlu hale gelmektedir. Bu noktada, yapay zekâ, yukarıdaki perspektiflere katkıda bulunabilir [7]. Ciddi bir uzman eksikliğiyle, geleneksel zatürre(pnömoni) ile büyük COVID-19 benzerliklerine sahipken, yapay zekâ (AI) destekli otomatik algılama modeli, test süresinin önemli ölçüde azaltılmasına yönelik önemli bir kilometre taşı olabilir [8]. Bu bağlamda, karmaşık problemlerin çözümünde etkili olan yapay zekâ (AI) tabanlı çözümler

gündeme gelmektedir. Söz konusu çözüm yöntemi, COVID-19 ve benzeri hastalıklar için hem daha düşük maliyetli hem de daha doğru bir tanı-teşhis-tedavi imkânı sunmaktadır. Günümüzde, özellikle tıbbi alanda, retina görüntüsü, göğüs röntgeni ve beyin MRG gibi görüntü veri seti ile derin öğrenme tekniklerini kullanarak başarılı sonuçlar elde edilmektedir [2,9] Derin Öğrenme ve Makine öğrenimi teknikleri; verilerden kalıpları çıkarmak, analiz etmek ve tanımak için birçok uygulamada kullanılmaktadır [7].

Bu çalışmada, yaygın olarak kullanılan derin öğrenme yaklaşımlarından Evrişimli Sinir Ağları (CNN) ve Derin Sinir Ağları (DNN) teknikleri ile makine öğrenmesi algoritması olan KNN yöntemi kullanılarak dört farklı durum (COVID-19, viral zatürre, bakteriyel zatürre ve normal) içerisinde COVID-19 olan hastaların tanı ve tespiti amaçlanmıştır. Çalışmanın ikinci bölümünde, literatür taramasına; üçüncü bölümünde yöntem ve tekniğin ayrıntılarına, dördüncü bölümde deneysel bulgulara ve son olarak beşinci bölümde ise çalışmadan elde edilen sonuç ve önerilere yer verilmiştir.

2. Literatür Taraması

COVID-19 virüslü hasta tanısı ve teşhisi amacıyla yapay zekâ teknikleri kullanılarak literatürde birçok çalışma gerçekleştirilmiştir. Örneğin, Öztürk ve arkadaşları tarafından yapılan bir çalışmada, derin sinir ağları kullanarak, X-Ray görüntülerinden COVID-19 tespiti yapılmıştır. Radyolojik görüntüleme ile yapay zekâ teknikleri kullanılarak, COVID-19 tespiti için ikili sınıflandırma modeli doğru teşhis yapmak için geliştirilmiştir. Oluşturulan model, ikili sınıflar için %98.08 ve çok sınıflı durumlar için %87.02 başarı elde edilmiştir [1].

Toğaçar ve arkadaşları, derin öğrenme modelleri kullanarak X-Ray görüntüleri üzerinde veri sınıflarını yeniden yapılandırmıştır. Bulanık renkler kullanarak optimizasyon ve göğüs röntgeni görüntüleri ve istifleme yaklaşımları ile COVID-19 tespiti için %99.27 doğruluk oranı elde etmişlerdir. Söz konusu virüsün tespitinde önerdikleri modelin etkin bir şekilde kullanılabileceği sonucuna varmışlardır [5].

Ioannis ve arkadaşları tarafından yapılan çalışmada 3905 adet yedi sınıfa ait X-Ray görüntüsünden yararlanarak derin öğrenme yönteminden CNN ile COVID-19 hastalığı tespit edilmiştir. Yedi sınıfın genel sınıflandırma doğruluğu % 87.66'ya ve COVID-19'un saptanmasında %99.18 Doğruluk, %97.36 Hassasiyet ve %99.42 Özgüllük-Belirginlik elde etmişlerdir. Çalışma sonucunda, derin öğrenme yaklaşımının modelin başarısında etkili olduğunu ortaya koymuşlardır [7].

Salman ve arkadaşları, yapay zekâ ile COVID-19 hastalığının algılanması üzerine bir çalışma gerçekleştirmişlerdir. Yaptıkları çalışmada, "kaggle" ve "github" kaynağından temin edilen 260 adet X-Ray görüntüleri veri setini oluşturmuştur. COVID-19, MERS, SARS, ARDS ve Normal X-Ray görüntülerinden yararlanılarak Evrişimli Sinir Ağına dayalı model oluşturulmuştur. Bu model, COVID-19 virüsünün olup olmadığını (sağlıklı veya sağlıklı) tespit ederek, %100'lük bir doğruluk oranı elde edilmiştir [10].

Jaiswal ve arkadaşları tarafından yapılan çalışmada COVID-19 ile enfekte hastaların sınıflandırılması için DenseNet201 tabanlı derin transfer öğrenmeyi kullanmışlardır. DenseNet201 tabanlı derin transfer öğrenme (DTL), hastaları COVID ile enfekte olarak sınıflandırmak için önerilmiştir. Önerilen model, COVID-19 (pozitif) veya COVID-19 (negatif) diye çıktı vererek, %97 doğruluk oranı ile DenseNet201 tabanlı CNN'nin önemli ölçüde performans gösterdiğini ortaya çıkarmıştır [11].

Rahimzadeh ve arkadaşları tarafından İran'da yapılan COVID-19'u tespit edilmesi üzerine yapılan çalışmada, 48260 CT (tomografi) tarama görüntüsü içeren yeni bir veri seti kullanılmıştır. ImageNet veri kümesi üzerinde eğitilmiş olan 50 katmanlı bir ağ olan ResNet50V2 ile görüntü sınıflandırma ile CT taramasında %98.49 doğruluk elde edilmiştir [12].

Nour ve arkadaşları, COVID-19 enfeksiyonunun tespiti için yaptıkları çalışmada, X-ray görüntülerinden faydalanarak Evrişimli Sinir Ağı (CNN) ve Bayes optimizasyon algoritması kullanmışlardır. Sonuç olarak, en verimli bulgular Destek Vektör Makinesi tarafından sağlanarak, %98.97 doğruluk, % 89.39 duyarlılık, % 99.75 özgüllük ve F skoru ile sınıflandırıcı ise %96.72 olarak etmişlerdir. Böylece, COVID-19 için ucuz, hızlı ve güvenilir bir yapay zekâ uygulaması sağlanmıştır [13].

Pathak ve arkadaşları tarafından yapılan çalışmada, bilgisayarlı tomografi görüntülerinden derin transfer öğrenmeye dayalı COVID-19 tespiti yapılmıştır. Önerilen model, sırasıyla %96.22 ve %93.01'a varan eğitim ve test doğruluğu sağlamıştır. Bu sağlanan sonuçlarla önerilen modelin COVID-19 test kitlelerine alternatif olarak kullanılabileceği ortaya çıkmıştır [14].

3. Materyal ve Yöntem

Bu bölümde, çalışma sürecinde önerilen modelin ve gerçekleştirilen adımların tüm ayrıntıları sunulmuştur. Öncelikle çalışmada kullanılan derin öğrenme yöntemleri ve makine öğrenmesi algoritmasından kısaca bahsedilmiş, daha sonra veri setinin elde edilmesi, gürültüden arındırılması ve hazırlanması süreçleri açıklanmış, son olarak da modelin oluşturulma sürecine yer verilmiştir. Modeli oluşturmak için Python açık kaynak yazılımı kullanılmıştır. Veri setindeki akciğer görüntülerinin sınıflandırılması için çalışmada CNN, DNN ve KNN algoritmalarından yararlanılmıştır.

Bu noktada, CNN ve DNN teknikleri, derin öğrenme yaklaşımlarındandır. Derin öğrenme, birden fazla katmanı kullanarak doğrusal olmayan problemler için çözüm sunan bir yöntemdir. Genellikle görüntü sınıflandırma, video analizi, konuşma tanıma, bilgi erişimi, nesne tanıma ve doğal dil öğrenme gibi alanlarda kullanılmaktadır [15-18].

KNN algoritması ise makine öğrenmesi tekniğidir. Makine öğrenmesi, bir problemi o probleme ait veriye göre modelleyen algoritmalar. Bu algoritmalar tahmin, kestirim, kümeleme ve sınıflandırma vb. yapabileceği yeteneğine sahiptir [19].

3.1. CNN (Evrişimsel Sinir Ağları)

CNN, çok katmanlı algılayıcıların bir türü olup bir veya daha fazla tamamen bağlı katmandan oluşur. Evrişim işleminde temel olarak iki boyutlu veriye uygulanacak olan filtrenin x ve y eksenine göre simetriği alınmaktadır. Filtre görüntü üzerinde adım uzunluğuna bağlı olarak gezdirilirken her adımda çakışan değerler eleman, eleman çarpılmakta ve tüm değerlerin toplamı çıkış matrisinin ilgili elemanı olarak kaydedilmektedir [20,21]. CNN, görüntüleri sınıflandırmak ve sınıflandıran de isimlendirerek kümelemek (fotoğraf arama) ve sahnelerde nesne tanıma yapmak için kullanılan bir tekniktir [22].

3.2. DNN (Derin Sinir Ağları)

DNN, insan beyninin bilgi işleme yönteminden esinlenerek geliştirilen ve nöronlardan oluşturulan yapay sinir ağlarının çok katmanlı durumudur. DNN, bir giriş katmanı ile bir çıktı katmanı arasında çok sayıda doğrusal olmayan gizli katmana sahip ve delta öğrenme yöntemiyle her katmanın ağırlığı

ayarlanmaktadır. DNN, denetimli ve denetimsiz öğrenme problemlerinde kullanılmakta, doğrusal olmayan ilişkileri modelleyebilmekte ancak öğrenme süreçleri çok yavaş gerçekleşmektedir. Bununla birlikte, genel olarak sınıflandırma ve regresyon amacıyla kullanılmakta ve yüksek performansta sonuçlar elde edilebilmektedir [18,23-25].

3.3. KNN (K-En Yakın Komşuluk Algoritması)

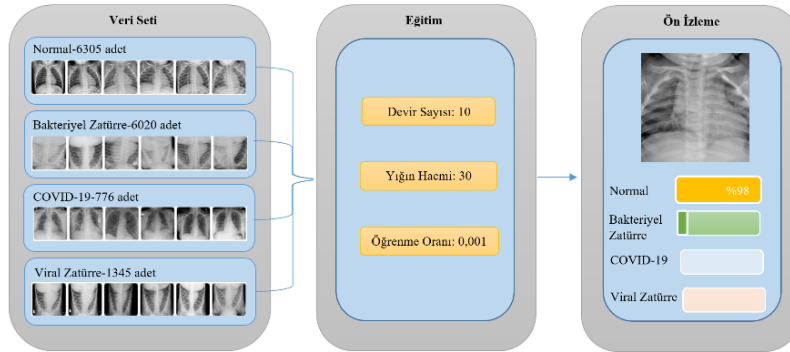
KNN algoritması, uygulaması kolay gözetimli öğrenme algoritmalarındandır. Hem sınıflandırma hem de regresyon problemlerinin çözümünde kullanılıyor olmakla birlikte endüstride çoğunlukla sınıflandırma problemlerinin çözümünde kullanılmaktadır. KNN, komşular arasındaki uzaklıkları hesaplar, en yakın komşuları bularak verileri sınıflara göre etiketlendirir. KNN, eski, basit ve gürültülü eğitim verilerine karşı dirençli olması sebebiyle en popüler makine öğrenme algoritmalarından biridir. Fakat bunun yanında dezavantajı da mevcuttur. Örneğin, uzaklık hesabı yaparken bütün durumları sakladığından, büyük veriler için kullanıldığında çok sayıda bellek alanına gereksinim duymaktadır. Sınıflandırmada temel amaç, nesnelerin sahip olduğu özelliklere bakılarak nesnelerin hangi sınıfa ait olduğunun belirlenme işlemidir [26].

3.4. Verilerin Elde Edilmesi ve Hazırlanması

COVID19 tespiti için herkese açık platform olan "https://www.kaggle.com/" web adresinden güncel ve gerçek veriler (X-Ray görüntüleri) alınmıştır. COVID-19'un X-ışını (PA-CXR) görüntülerinden elde edilen bu veri seti bakteriyel zatürre, viral zatürre, COVID-19 (enfekte olmuş, doğrulanmış vakalar) ve normal olmak üzere dört kategoriden oluşmaktadır. Modeli oluşturmak için toplam 14.446 adet veri kullanılmıştır. Tablo 1'de kategori adı, eğitim ve teste girecek veri sayıları başlıklarıyla tüm verilere yer verilmiştir.

Tablo 1: Kategoriler ve veri sayıları

Sıra No	Kategori Adı	Eğitim (%85)	Test (%15)	Toplam
1	BacterialPneumonia	5116	904	6020
2	COVID19	660	116	776



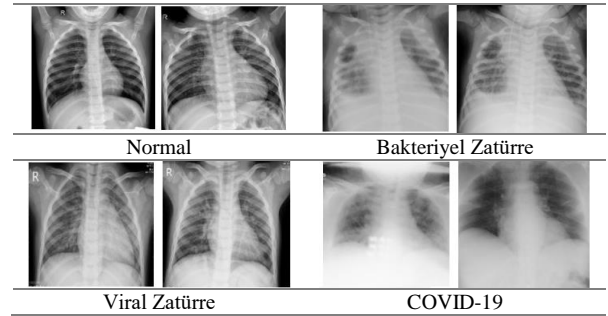
Şekil 1: CNN modelinde verilerin etiketlenmesi ve eğitilmesi süreci

Bu süreç, Python 3.7 sürümünün derin öğrenme kütüphanesi olan Keras üzerinden gerçekleştirilmektedir. Katmanlar üzerinde optimizasyon fonksiyonları (Adam, Adadelta, Adamax ve SGD), aktivasyon fonksiyonu (Relu, Softmax) ve nöron sayılarının kombinasyonları kullanılarak modeller oluşturulmuştur. Sistemin genel algoritmasına göre öncelikle X-Ray görüntüleri alınmakta, ön işlemlerden geçmektedir. Daha

Sıra No	Kategori Adı	Eğitim (%85)	Test (%15)	Toplam
3	Normal	5360	945	6305
4	ViralPneumonia	1143	202	1345
Toplam		12279	2167	14446

Modeller oluşturulmadan önce, veri setindeki X-Ray görüntüleri 220x220 boyutunda yeniden ölçeklendirilmiştir. Farklı piksel oranlarındaki görüntüleri standardize etmek, bozulmayı önlemek ve tam bir dönüşüm elde etmek amacıyla görüntüleri 200x200 piksel boyutunda siyah arka plan eklenmiştir. Düşük kontrastlı görüntüler veya tüm torasik X-Ray taramasının bölümlerini içeren görüntüler hariç tutulmuştur. Tablo 2'de, kategori bazında örnek akciğer tomografisi görüntüleri verilmiştir.

Tablo 2: Kategorilerine göre akciğer tomografisi görüntüleri

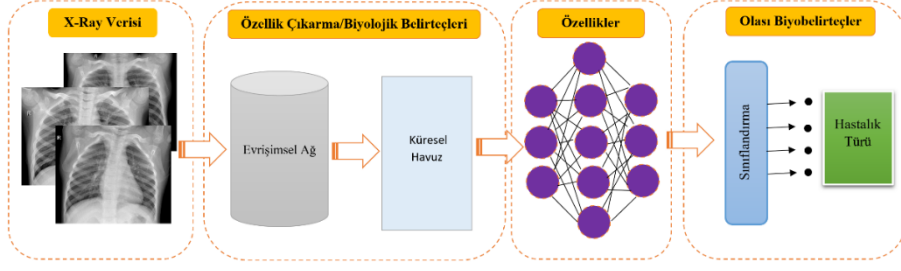


3.5. Modellerin Oluşturulması

3.5.1. CNN Modeli

Veriler gürültüden arındırıldıktan sonra modeli oluşturmak amacıyla öncelikle eğitim sürecine dâhil edilmiştir. Şekil 1'deki gibi kategorilerin isimleri yazılarak akciğer görüntüleri 4 adet sınıfa ait klasörler oluşturulmuş ve eğitim işlemi başlatılmıştır. Eğitim sürecinde devir sayısı (Epoch)=10, Yığın hacmi (Batch size) =30 ve Öğrenme oranı (Learning Rate) =0.001 olarak belirlenmiştir.

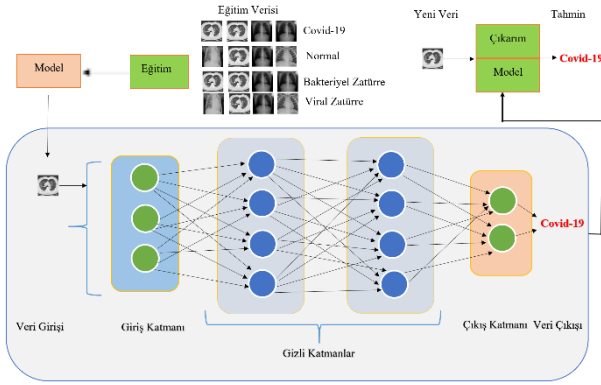
sonra, bu görüntülerden X-Ray görüntü olup olmadığı önceden eğitilmekte ve veritabanındaki görüntülerle karşılaştırılarak belirlenmektedir. Devamında ise CNN algoritması kullanılarak sınıflandırma yapılmakta ve sonuç olarak tespit edilen X-Ray görüntüsünün etiketi belirlenmektedir. Şekil 2'de, X-Ray görüntülerin özellikleri çıkarılırken sınıflandırma aşamaları biyobelirteçlere göre etiketleme işlemi görülmektedir.



Şekil 2: Sistemin özellik çıkarma işlemi

3.5.2. DNN Modeli

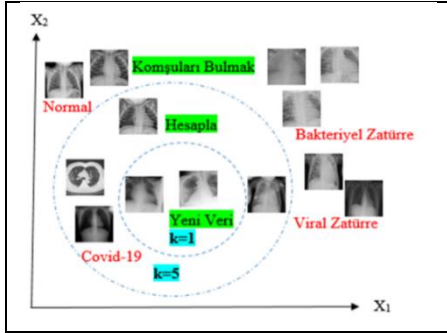
Derin sinir ağına gelen veriler; giriş, gizli ve çıkış olmak üzere üç katmandan geçerek sınıflandırma yapılmakta ve akciğer tomografisi görüntüsünün COVID-19 olduğu tespit edilmektedir. Şekil 3'te, derin sinir ağı mimarisi verilmiştir.



Şekil 3: DNN modeli ve sınıflandırma

3.5.3. KNN Modeli

KNN algoritmasında, yeni gelen akciğer verisine göre önce uzaklıkları hesaplama işlemi yapılmaktadır. Daha sonra en yakın komşuları bularak model uygulanmakta ve verinin COVID-19 olup olmadığı sınıflandırılmaktadır.



Şekil 4: X-Ray görüntülerine göre KNN modeli

4. Deneysel Sonuçlar ve Tartışma

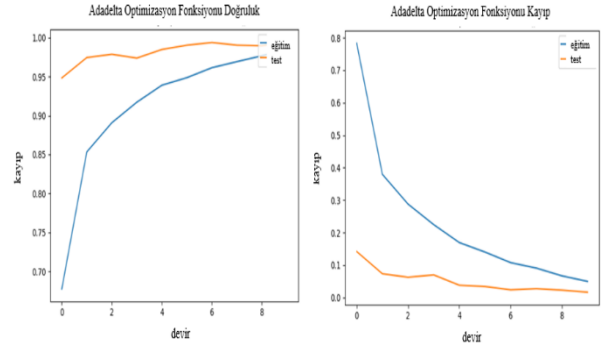
4.1. CNN modelinden elde edilen sonuçlar

Bu bölümde, 4 farklı kategoriye ait sınıflandırmanın karşılaştırılması yapılmıştır. Keras kütüphanesi kullanılarak elde edilen sonuçları toplam değil, son devir sayısı baz alınmıştır. Katmanlarda (çıkış hariç) relu fonksiyonu kullanılmıştır.

Tablo 3: Optimizasyon fonksiyonlarının etkisi

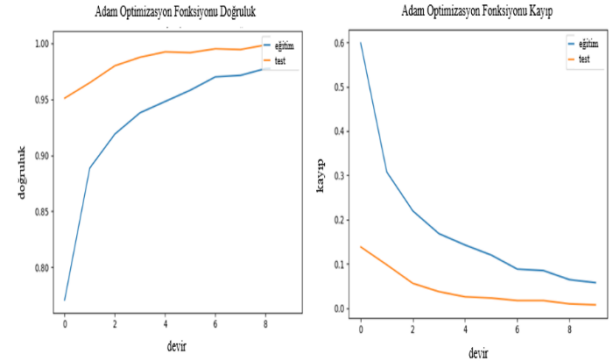
Devir Sayısı:10 Katman Sayısı:5 Yığın Hacmi:30				
Optimizasyon Fonksiyon Aracı	Eğitim		Test	
	Kayıp	Eğitim Doğrulama	Kayıp	Test Doğrulama
Adadelta	0.016	0.994	0.049	0.984
Adam	0.017	0.991	0.057	0.978
SGD	0.205	0.931	0.316	0.876
Adamax	0.033	0.986	0.108	0.960

Şekil 5'te, Adadelta optimizasyonu için doğrulama ve kayıp grafiği verilmiştir. Grafiğe göre devir değeri 10, yığın hacmi=30 alınarak, eğitim oranı %99, test oranı ise %98 olarak elde edilmiştir.



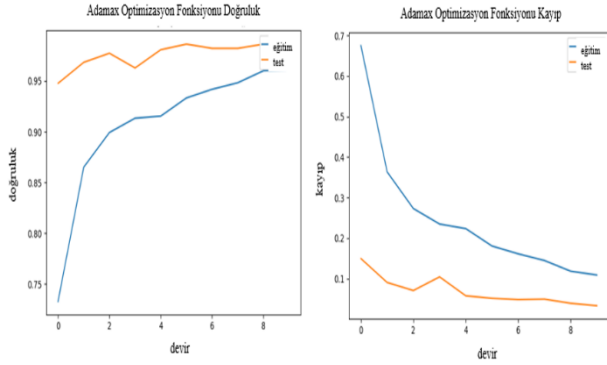
Şekil 5: Adadelta için doğruluk ve kayıp grafiği

Şekil 6'da, Adam optimizasyonu için doğrulama ve kayıp grafiği verilmiştir. Grafiğe göre, eğitim oranı %99, test oranı %97 olarak elde edilmiştir.



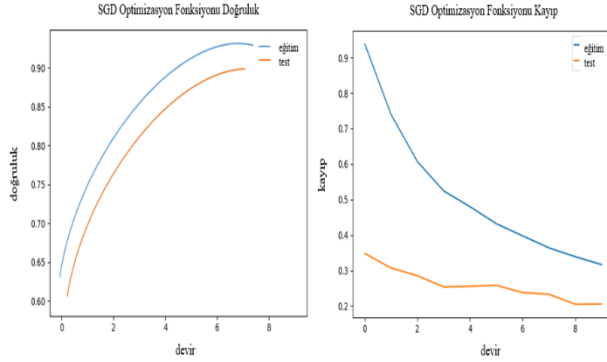
Şekil 6: Adam için doğruluk ve kayıp grafiği

Şekil 7’de, Adamax optimizasyonu için doğrulama ve kayıp grafiği verilmiştir. Grafiğe göre, eğitim oranı %98, test oranı %96 olarak elde edilmiştir.



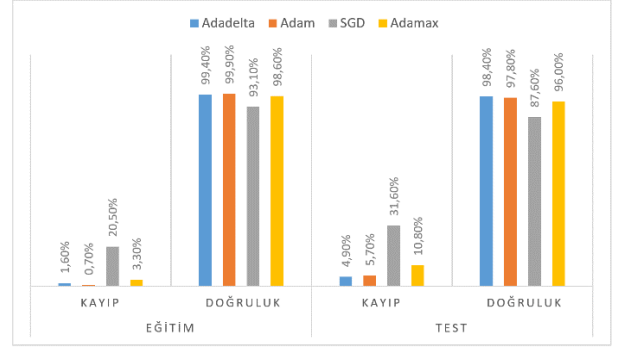
Şekil 7: Adamax için doğrulama ve kayıp grafiği

Şekil 8’de SGD optimizasyonu için doğruluk ve kayıp grafiği verilmiştir. Grafiğe göre, eğitim oranı %93, test oranı %86 olarak elde edilmiştir.



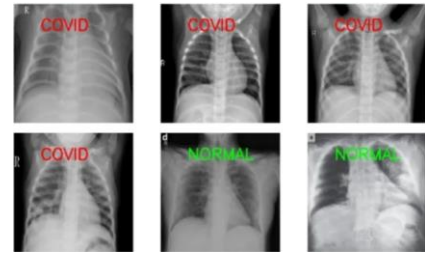
Şekil 8: SGD için doğrulama ve kayıp grafiği

CNN tekniğinde, tüm fonksiyonlardan elde edilen sonuçların karşılaştırılmasını gösteren grafiğe Şekil 9’da yer verilmiştir. Grafiğe göre en iyi sonuçlar; eğitim oranı %99, test oranı %98 ve kayıp oranı eğitim 0,016 ve test için 0,049 olmak üzere “Adadelata Optimizasyon Fonksiyonu” ile elde edilmiştir. En düşük oranlar ise, %93’lük eğitim ve %87 test oranları olmak üzere “SGD Optimizasyon Fonksiyonu” ile elde edilmiştir.



Şekil 9: Optimizasyon fonksiyonlarının performans grafiği

Tüm bu testlerden sonra, X-ray görüntülerine sınıflandırma ile etiketleme işlemi gerçekleştirilmiş (Şekil 10) ve her bir modelin performans sonuçları elde edilmiştir.



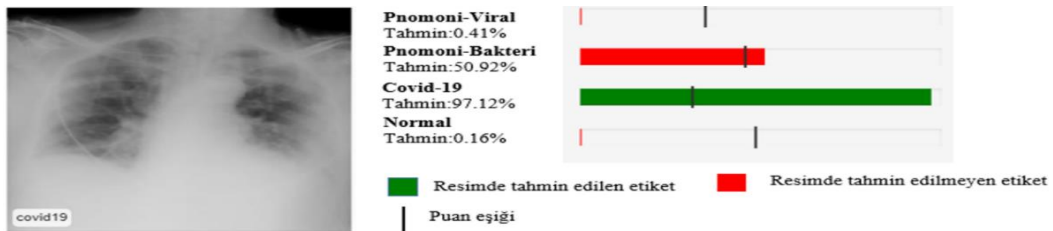
Şekil 10: Sınıflandırma ve etiketleme işlemleri

4.2. DNN modelinden elde edilen sonuçlar

X-Ray görüntülerine ait 4 kategoride eğitim ve test işlemi yapılmıştır. Derin öğrenme yöntemlerinden derin sinir ağı (DNN) modelinde veri seti, %85 eğitim ve %15 test oranları olarak iki gruba ayrılmıştır. Sınıflara göre test başarı oranı ve doğruluk oranı Tablo 4’te görülmektedir. En yüksek test başarı oranı COVID-19 sınıfında olup %98,3’lük bir orana sahiptir.

Tablo 4: DNN yöntemi başarı oranları

Sınıf	Test	Duyarlılık	F1 Skoru	Ortalama Hassasiyet	ROC Doğruluk
Pnömoni -Viral	93.6%	73.1%	82.1%	92.4%	99.2%
Pnömoni -Bakteri	91.3%	79.4%	85.0%	93.7%	94.9%
COVID-19	98.3%	50.0%	66.3%	87.6%	98.9%
Normal	94.8%	86.7%	90.5%	97.5%	98.0%



Şekil 11: DNN algoritmasına göre örnek bir sınıflandırma işlemi

Şekil 11’de, test için verilen örnek bir akciğer tomografisinin DNN yöntemine göre %97.12’lik bir oranla COVID-19 olarak etiketlenmiştir. Ayrıca sınıflandırmada puan eşliğini en fazla geçen sınıf hangisi olursa test verisinin sınıflandırması da o etiket olmaktadır.

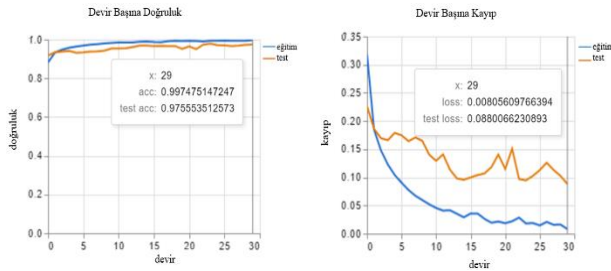
4.3. KNN modelinden elde edilen sonuçlar

KNN modeli için tüm X-Ray görüntüleri; %85 oranında ise eğitim verisi, %15 oranında ise test verisi olmak üzere iki gruba ayrılmıştır. X-Ray görüntülerine ait 4 kategoride eğitim işlemi yapılmıştır. Sınıflara göre kullanılan veri sayısı ve doğruluk yüzdeleri Tablo 5’te verilmiştir. Tablo 5’e göre; Normal %98, Bakteri %97, Viral %96 ve COVID-19 %97 şeklinde başarı oranları elde edilmiştir.

Tablo 5: Kategorilerine göre doğruluk oranları ve veri sayısı

Kategori	Doğruluk	Örnek Sayısı
Normal	% 98	946
Bakteri	% 97	903
Viral	% 96	202
COVID19	% 97	117

Şekil 12’de, KNN algoritmasına göre tüm sınıfların eğitim ve test aşamasındaki doğruluk ve kayıp grafiği verilmiştir. Grafiğe göre eğitim için %99, test için ise %97’lik başarı oranları elde edilmiştir. Dönem başına kayıp olarak da eğitim verilerinde %0,8 test verisinde ise %8,8 oranı çıkmıştır.



Şekil 12: KNN’ e göre doğruluk ve devir başına kayıp grafiği

4.4. Bulguların Karşılaştırılması

Çalışmadan elde edilen bulgular incelendiğinde (Tablo 6), tanıteşhis anlamında en yüksek performansın derin öğrenme algoritmalarından CNN ile elde edildiği görülmektedir. Bu noktada, en yüksek eğitim oranı KNN olsa da burada önemli olan eğitim oranıyla birlikte test oranının da yüksek olmasıdır. Tablo 6’ya göre bu durum CNN tekniği ile elde edilmiştir.

Tablo 6: Algoritma sonuçlarının karşılaştırılması

Sıra No	Eğitim (%)	Eğitim Kayıp (%)	Test (%)	Test Kayıp (%)	Algoritma
1	99,56	1,60	98,90	4,90	CNN (Evrışimli Sinir Ağı)
2	93,70	6,30	93,40	6,60	DNN (Derin Sinir Ağı)
3	99,74	0,80	97,55	8,80	KNN (K-En Yakın Komşuluk)

Son olarak, önerilen CNN, DNN ve KNN modelleri ile gerçek veriler kullanılarak 20 (yirmi) adet farklı test gerçekleştirilmiştir. Bu testlerden elde edilen deneysel sonuçlar Tablo 7’de verilmiştir.

Tablo 7: Modellerden elde edilen deneysel sonuçlar

No	CNN	Hata (CNN)	DNN	Hata (DNN)	KNN	Hata (KNN)
1	99,57	0,43	96,40	3,60	99,09	0,91
2	99,12	0,88	91,41	8,59	99,00	1,00
3	98,67	1,33	90,14	9,86	96,70	3,30
4	99,34	0,66	86,60	13,40	98,50	1,50
5	98,26	1,74	91,15	8,85	99,30	0,70
6	99,10	0,90	91,96	8,04	89,43	10,57
7	99,70	0,30	81,60	18,40	95,46	4,54
8	95,40	4,60	81,60	18,40	46,16	53,84
9	98,67	1,33	90,40	9,60	96,23	3,77
10	99,35	0,65	96,60	3,40	99,08	0,92
11	98,55	1,45	96,47	3,53	98,19	1,81
12	94,29	5,71	85,80	14,20	85,70	14,30
13	96,95	3,05	85,04	14,96	99,39	0,61
14	99,50	0,50	91,96	8,04	99,37	0,63
15	99,18	0,82	95,01	4,99	98,93	1,07
16	99,20	0,80	89,96	10,04	95,72	4,28
17	98,27	1,73	94,67	5,33	96,60	3,40
18	96,78	3,22	92,33	7,67	71,98	28,02
19	99,90	0,10	84,80	15,20	97,78	2,22
20	98,23	1,77	93,57	6,43	71,90	28,10
ORTALAMA	98,40	1,59	90,37	9,62	91,72	8,27

Deneysel sonuçlara göre (Tablo 7), en iyi performans ve en düşük hata oranı CNN tekniği ile elde edilmiştir.

5. Sonuç ve Öneriler

COVID-19 pandemisinin tüm dünyaya hızlı yayılması ve insanlar üzerindeki olumsuz etkileri, konunun ayrıntılı bir şekilde irdelenmesi gerektiğini ve pozitif vakaların erken evrelerde tespit edilerek hızlı ve doğru müdahalenin yapılmasını açıkça ortaya koymaktadır. Bu bağlamda çalışmada, derin öğrenme (CNN, DNN) ve makine öğrenmesi (KNN) yöntemleri kullanılarak bir model önerilmiştir. Söz konusu model ile farklı akciğer tomografisi görüntüleri içerisinde COVID-19 olan hastaların tanı ve tespiti gerçekleştirilmiştir. Sonuçlar, CNN tekniğinin DNN ve KNN yöntemlerine göre daha yüksek performans değerlerine sahip olduğunu göstermiştir. Ayrıca, modelin performansı veri sayısı ile doğru orantılıdır. Daha fazla verinin kullanılması performans üzerinde olumlu etkiye sahiptir. Elde edilen bulgular, önerilen CNN modelinin hibrit yaklaşımlarla biraz daha geliştirilerek sağlık sistemlerinde rahatlıkla kullanılabilirliğini ortaya koymaktadır.

Çalışmanın bir sonraki aşamasında, daha yüksek başarı oranları elde etmek için farklı algoritma veya teknikler de kullanılabilir. Ayrıca, çalışmanın konusu sağlık olması nedeniyle bu alandaki uzmanların da görüşleri alınarak elde edilen sonuçlar hakkında daha doğru yorumlar alınabilir. Sonuç olarak, önerilen modelin kısa sürede, düşük hata oranı ile ve minimum maliyetle tanı veya tespit yapılabilmesi, yapay zekâ tekniklerinin sağlık alanında kullanıma potansiyelini ortaya koymaktadır.

6. Kaynakça

- [1] Öztürk, T., Talo, M., Yıldırım, E. A., Baloğlu, U. B., Yıldırım, Ö., & Acharya, U. “Automated detection of COVID-19 cases using deep neural networks with X-ray images”. *Computers in Biology and Medicine*, 1-11, 2020.
- [2] Chen, Y. T., de Gracia, M. M., Diez-Tascón, A., Alonso-González, R., Rodríguez-Fuertes, P., Parra-Gordo, M. L.,

- ... & Fuertes, R. L. "Correlation between Chest Computed Tomography and Lung Ultrasonography in Patients with Coronavirus Disease 2019 (COVID-19)". *Ultrasound in Medicine & Biology*, 2020.
- [3] Brunese, L., Mercaldo, F., Reginelli, A., & Santone, A. "Explainable deep learning for pulmonary disease and coronavirus COVID-19 detection from X-rays". *Computer Methods and Programs in Biomedicine*, 2020.
- [4] Pereira, R. M., Bertolini, D., Teixeira, L. O., Silla Jr, C. N., & Costa, Y. M. "COVID-19 identification in chest X-ray images on flat and hierarchical classification scenarios". *Computer Methods and Programs in Biomedicine*, 2020.
- [5] Toğaçar, M., Ergen, B., & Cömert, Z., "COVID-19 detection using deep learning models to exploit Social Mimic Optimization and structured chest X-ray images using fuzzy color and stacking approaches". *Computers in Biology and Medicine*, 1-12, 2020.
- [6] Panwar, H., Gupta, P. K., Siddiqui, M. K., Morales-Menendez, R., & Singh, V. "Application of Deep Learning for Fast Detection of COVID-19 in X-Rays using nCOVnet". *Chaos, Solitons & Fractals*, 2020.
- [7] Apostolopoulos, I., Sokratis, A., & Mpesiana, T., "Extracting possibly representative COVID-19 Biomarkers from X-Ray images with Deep Learning approach and image data related to Pulmonary Diseases". *Journal of Medical and Biological Engineering*, 462-469, 2020.
- [8] Mahmud, T., Rahman, M. A., & Fattah, S. A., "CovXNet: A multi-dilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization", *Computers in biology and medicine*, 2020.
- [9] Khan, A. I., Shah, J. L., & Bhat, M. M.. "Coronet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images", *Computer Methods and Programs in Biomedicine*, 2020.
- [10] Salman, F. M., Abu-Naser, S. S., Alajrami, E., Abu-Nasser, B. S., & Ashqar, B. A. "COVID-19 Detection using Artificial Intelligence", *International Journal of Academic Engineering Research*, 18-25, 2020.
- [11] Jaiswal, A., Gianchandani, N., Singh, D., Kumar, V., & Kaur, M., "Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning", *Journal of Biomolecular Structure and Dynamics*, 1-9, 2020.
- [12] Rahimzadeh, M., Attar, A., & Sakhaei, S. M. "A Fully Automated Deep Learning-based Network For Detecting COVID-19 from a New And Large Lung CT Scan Dataset", Retrieved from The preprint server for health sciences: <https://www.medrxiv.org/content/10.1101/2020.06.08.20121541v2>, (4 Temmuz, 2020).
- [13] Nour, M., Cömert, Z., & Polat, K., "A Novel Medical Diagnosis model for COVID-19 infection detection based on Deep Features and Bayesian Optimization", *Applied Soft Computing Journal*, 1-13, 2020.
- [14] Pathak, Y., Shukla, P., Tiwari, A., Stalin, S., Singh, S., & Shukla, P., "Deep Transfer Learning Based Classification Model for COVID-19, Disease. *Innovation and Research in BioMedical engineering*, 1-6, 2020.
- [15] Özkan, İ. & Ülker, E., "Derin Öğrenme ve Görüntü Analizinde Kullanılan Derin Öğrenme Modelleri", *Gaziosmanpaşa Bilimsel Araştırma Dergisi*, 6(3), 85-104, 2017.
- [16] Şeker, A., Diri, B., & Balık, H. H. "Derin öğrenme yöntemleri ve uygulamaları hakkında bir inceleme". *Gazi Mühendislik Bilimleri Dergisi (GMBD)*, 3(3), 47-64, 2017.
- [17] L. Deng and D. Yu, "Deep Learning: Methods and Applications," *Found. Trends® Signal Process.*, vol. 7, no. 3-4, pp. 197-387, 2014.
- [18] Ravi, D., Wong, C., Deligianni, F., Berthelot, M., Andreu-Perez, J., Lo, B., & Yang, G. Z., "Deep learning for health informatics". *IEEE journal of biomedical and health informatics*, 21(1), 4-21, 2016.
- [19] Atalay, Muhammet, and Enes Çelik. "Büyük Veri Analizinde Yapay Zekâ Ve Makine Öğrenmesi Uygulamaları-Artificial Intelligence and Machine Learning Applications in Big Data Analysis." *Mehmet Akif Ersoy Üniversitesi Sosyal Bilimler Enstitüsü Dergisi* 9.22 (2017): 155-172.
- [20] Kızrak, M. A., & Bolat, B. "Derin öğrenme ile kalabalık analizi üzerine detaylı bir araştırma". *Bilişim Teknolojileri Dergisi*, 11(3), 263-286, 2018.
- [21] Doğan, F., & Türkoğlu, İ., "Derin öğrenme algoritmalarının yaprak sınıflandırma başarımlarının karşılaştırılması", *Sakarya University Journal of Computer and Information Sciences*, 1(1), 10-21, 2018.
- [22] Amidi, S. "Afshine Amidi ve Shervine Amidi tarafından yazılmış, Ayyüce Kızrak ve Yavuz Kömeçoğlu tarafından çevrilmiştir. Evrişimli Sinir Ağları. Stanford University: Erişim adresi: <https://stanford.edu/~shervine/l/tr/teaching/cs-230/cheatsheet-convolutional-neural-networks>, , 2018.
- [23] Küçük, D., & Arıcı, N., "Doğal Dil İşlemede Derin Öğrenme Uygulamaları Üzerine Bir Literatür Çalışması, Uluslararası Yönetim Bilişim Sistemleri ve Bilgisayar Bilimleri Dergisi, 2(2), 76-86, 2018.
- [24] Chollet F, Allaire JJ. *Deep Learning with R: Manning Publications Co.*; 2018.
- [25] Patterson J, Gibson A. *Deep Learning: A Practitioner's Approach*. Beijing: O'Reilly; 2017.
- [26] Kilinç, D., Borandağ, E., Yücalar, F., Tunali, V., Şimşek, M., & Özçift, A., "KNN algoritması ve R dili ile metin madenciliği kullanılarak bilimsel makale tasnifi", 2016.