



Kulak İçi Hastalıklarının Derin Öğrenme Mimarileriyle Sınıflandırılması ve Karşılaştırılması

Furkanca Demircan^{1*†}, Murat Ekinci², Zafer Cömert³

^{1*} Bilgisayar Mühendisliği, Fen Bilimleri Enstitüsü, Karadeniz Teknik Üniversitesi, Ortahisar, Trabzon, Türkiye, (ORCID: 0000-0001-8096-5731),

furkancan.demircan@gmail.com

² Bilgisayar Mühendisliği, Mühendislik Fakültesi, Karadeniz Teknik Üniversitesi, Ortahisar, Trabzon, Türkiye, (ORCID: 0000-0001-9326-8425), ekinci@ktu.edu.tr

³ Yazılım Mühendisliği, Mühendislik Fakültesi, Samsun Üniversitesi, 19 Mayıs, Samsun, Türkiye, (ORCID: 0000-0001-5256-7648), zafer.comert@samsun.edu.tr

(İlk Geliş Tarihi 26 Aralık 2022 ve Kabul Tarihi 12 Nisan 2023)

(DOI: 10.31590/ejosat.1224070)

ATIF/REFERENCE: Demircan, F., Ekinci, M. & Cömert, Z. (2023). Kulak İçi Hastalıkların Derin Öğrenme Mimarileriyle Sınıflandırılması ve Karşılaştırılması. *Avrupa Bilim ve Teknoloji Dergisi*, (51), 75-85.

Öz

Otitis media (OM), kulak zarı içerisinde oluşan akıntılı, enfeksiyonel hastalıkları tanımlamaktadır. Kulak mumu (earwax), kulak zarı içerisinde bakteri oluşumunu önleyen savunma mekanizmasının aşırı birikimi sonucunda kulakta işitme kaybı oluşmasına neden olan hastalıktır. Kulak zarı içerisinde kalsiyum birikimi sonucunda saydamlığını ve esnekliğini kaybetmesine miringoskleroz denmektedir. Bu hastalıkların tanısı Kulak Burun Boğaz (KBB) uzmanları tarafından kulak zarının otoskopi ile incelenmesi sonucunda koyulmaktadır ve hataya açıktır. Bu çalışmada, bu problemin çözümüne katkı sağlamak ve bir karar destek sistemi sunmak amacıyla derin öğrenme modelleriyle kulak zarı hastalıklarına ait görüntüler sınıflandırılmıştır. Veri seti olarak 4 sınıf ve 880 görüntünün bulunduğu Ear Imagery veri seti seçilmiştir. Sınıflandırma işlemi için AlexNet, ResNet50, ResNet101, ResNet50V2, ResNet101V2, InceptionV3, Xception ve InceptionResNetV2 derin öğrenme modelleri seçilmiştir. En yüksek başarı değeri %94 doğruluk ile InceptionResNetV2 mimarisinden ve en hızlı sonuç 438 saniye ile AlexNet mimarisinden elde edilmiştir. Bu yaklaşımla kulak zarına ait hastalıkların potansiyel uzman hatalarından arındırılarak otonom bir sistem ile gerçekleştirilebileceği gösterilmiştir. Gelecekte klinik alanda böyle bir sistemin kullanılması; uzmanların karar verme sürecini destekleyebilir ve hataya açık olan değerlendirme sürecinin daha objektif ve tekrar edilebilir bir şekilde yönetilmesini sağlayabilir.

Anahtar Kelimeler: Biyomedikal işaret işleme, Otitis media, Derin Öğrenme, Sınıflandırma.

Classification And Comparison Of Ear Diseases With Deep Learning Architectures

Abstract

Otitis Media (OM) is infectious disease with discharge in the eardrum. Earwax is a disease that causes hearing loss in the ear as a result of excessive accumulation of the defense mechanism that prevents the formation of bacteria in the eardrum. The loss of transparency and flexibility as a result of calcium accumulation in the eardrum is called myringosclerosis. The diagnosis of these diseases is made by otolaryngologists using an otoscopy examination of the eardrum and this process is prone to error. In this study, otoscopy images were classified with deep learning models to solve this problem. The Ear Imagery dataset with 4 classes and 880 images was chosen as the dataset. AlexNet, ResNet50, ResNet101, ResNet50V2, ResNet101V2, InceptionV3, Xception and InceptionResNetV2 deep learning models were selected for classification. The highest success value was obtained from InceptionResNetV2 architecture with 94% and the fastest result was obtained from AlexNet architecture with 438 seconds. With this approach, it has been shown that diseases of the eardrum can be treated with an autonomous system, freeing from expert error. In the future, such a system in the clinical field will be able to reduce labor and error.

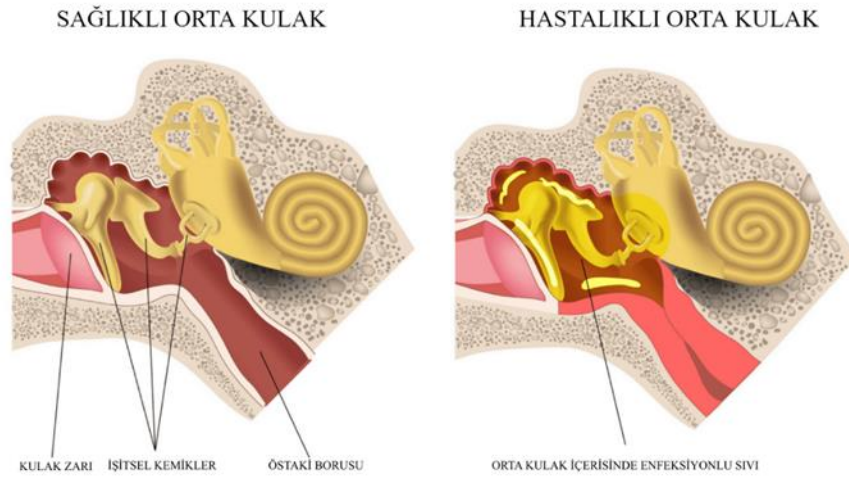
Keywords: Biomedical signal processing, Otitis media, Deep learning, Classification.

* Sorumlu Yazar: furkancan.demircan@gmail.com

†Bu araştırma makalesi, Furkanca Demircan'ın yüksek lisans tezinden türetilmiştir.

1. Giriş

Otitis media (OM), orta kulakla ilgili bir grup inflamatuvar hastalığı ifade etmektedir. Çoğunlukla çocuklarda görünen OM tedavi edilmediğinde konuşma kusurları, yüz sinir felci, işitme kaybı ve bilişsel bozukluk gibi ciddi sonuçlara neden olabilmektedir. OM ile ilgili birçok hastalık olmakla birlikte en sık görülen hastalık türleri efüzyonlu otitis media (OME), akut otitis media (AOM) ve kronik süperatif otitis media (CSOM). Bu hastalık türleri sıklıkla birbirlerini izleyebilir ya da bir arada bulunabilmektedir. OME hastalığında semptomlar hastadan hastaya değişmekle birlikte genel olarak işitme zorluğu, denge kaybı ve konuşma gelişimi zorluğu olarak göstermektedir. AOM hastalığında genel semptomlar ateş, kulak ağrısı, otore, kusma ve ishal şeklinde göstermektedir. CSOM hastalığının semptomları AOM hastalığına benzemektedir fakat şiddeti daha ağır ve süresi daha uzundur. OME hastalığında kulak içerisinde sıvı birikmesi sonucunda ödem oluşmaktadır. Ödem kulak içinde büyümesi veya dış etkenlere bağlı olarak yırtılması sonucunda dış havayla teması gerçekleşmektedir. Dış havayla temasından sonra bakteriyel oluşumlar etkisiyle OME hastalığı AOM hastalığına dönüşmektedir. AOM hastalığı tedavi edilmediğinde semptomların ve kulak içindeki enfeksiyonun artmasıyla AOM hastalığı CSOM hastalığına dönüşmektedir (Yavuz ve diğerleri, 2000). Şekil 1'de OM hastalığına bağlı olarak enfeksiyon birikimi ve östaki borusundaki daralma modellenmiştir. Ayrıca Şekil 1'de sağlıklı kulak zarı ile OM hastalıklı kulak zarı arasındaki farklar gösterilmektedir.



Şekil 1. Sağlıklı kulak zarı ile OM hastalıklı kulak zarı modeli (Figure 1. OM diseased eardrum and healthy eardrum model) (Sahling ve diğerleri, 2021)

Kulak mumu (Earwax), kulağın içerisindeki savunma mekanizmasıdır ve kulak içerisinde bakteriyel oluşumu ve kirlerin kulak kanalına geçmesini önlemektedir. Dış kulağın içerisindeki bezler ve foliküllerin ürettiği mumun fazla birikmesi sonucunda kulak mumu tıkaçları (earwax plug) oluşmaktadır. Hastalığın genel semptomları kulak ağrısı, kulakta doluluk hissi, kulak çınlaması, işitme kaybıdır (Mayo Clinic, 2022).

Kulak zarı normalde ince ve yarı saydam bir yapıdadır. Orta kulak içerisinde kronik inflamasyon sonucunda gerçekleşen değişimlerle kulak zarı içerisinde kalsiyum birikmesi sonucunda kalınlaşma, sertleşme ve esnekliğini kaybetme görülmektedir. Miringoskleroz ve timpanoskleroz orta kulağı etkileyen, kulak zarını beyazlaştıran iki benzer hastalık türüdür. İki hastalık arasındaki fark miringoskleroz hastalığında kalsiyum sadece kulak zarında depolanmaktadır. Öte yandan timpanoskleroz hastalığında kalsiyum kulak zarında ve orta kulak çeperlerinde depolanmaktadır. Bu hastalıkların oluşmasındaki genel sebep OME ve SOM hastalıkları sonucunda oluşan kronik kulak akıntısı, kronik veya tedavi edilmeyen kronik orta kulak enfeksiyonları ve cerrahi tedaviyle yerleştirilen havalandırma tüpleridir. Hastalıkların genel semptomları olmamakla birlikte timpanoskleroz hastalığında işitme kaybı görülebilmektedir (Hayes, 2022).

Wu ve arkadaşları pediatrik orta kulak iltihaplarının kolay izlenmesi ve sınıflandırılmasında, ImageNet veri setinde önceden eğitilmiş Xception ve MobileNet-V2 mimarilerini kullanan bir mobil uygulama önermiştir. Önerilen metodun eğitim ve test işlemleri üç tip otitis media hastalığını barındıran 12203 görüntü üzerinde gerçekleştirilmiştir. Yapılan çalışma sonucunda Xception ve MobileNet-V2 mimarileri sistem üzerindeki test görüntülerinde sırasıyla %97,82 ve %96,04; mobil uygulama üzerindeki test görüntülerinde ise sırasıyla %90,85 ve %88,89 doğruluk göstermiştir (Wu ve diğerleri, 2021). Alhudhaif, ve arkadaşları sekiz tip kulak zarı hastalığının sınıflandırılması için genelleştirilmiş karar mekanizması kullanan CNN mimarisini önermiştir. Önerilen mimarinin eğitim ve test işlemleri sekiz tip hastalığı barındıran 956 görüntünün bulunduğu açık erişimli veri setinde gerçekleştirilmiştir. CNN mimarisinin karar mekanizması kanal ve uzamsal model (CBAM), artık bloklar ve hiper sütun tekniğinin birleşimini kullanılarak geliştirilmiştir. Önerilen model test veri seti üzerinde %98,26 doğruluk göstermiştir (Alhudhaif ve diğerleri, 2021). Tran ve arkadaşları pediatrik iki tip otitis media hastalığı için otomatik tanı koyabilen sistem önermiştir. Sistemin sınıflandırma işlemi için ortak seyrek temsil algoritması kullanılmıştır. Önerilen modelin eğitim ve test işlemleri iki tip OM hastalığının bulunduğu 1230 görüntü üzerinde gerçekleştirilmiştir. Önerilen model test veri setinde %91,41 doğruluk göstermiştir (Tran ve diğerleri, 2018).

Bu hastalıkların tanısı Kulak Burun Boğaz (KBB) uzmanları tarafından kulak zarının otoskop ile incelenmesi sonucunda subjektif olarak koyulmaktadır (Başaran ve diğerleri, 2020). Ayrıca bu tanı şeklinin ulusal olarak kabul görmüş standartlara sahip bir kılavuzu bulunmamaktadır. Dolayısıyla bu subjektif tanı sistemi hataya açıktır. Önceki çalışmalar ışığında bu çalışmadaki ana amacımız normal ve hastalıklı kulak zarlarına ait görüntülerin otonom bir sistem ile değerlendirilmesi ve teşhis koyularak subjektif teşhisin ortaya çıkardığı hataları önlemektir. Bu çalışma kapsamında otonom sistem için en doğru teşhisi koyabilen derin öğrenme modelinin bulunabilmesi için farklı derin öğrenme modelleri karşılaştırılmıştır. Derin öğrenme modeli olarak AlexNet, ResNet50, ResNet101, ResNet50V2, ResNet101V2, InceptionV3, InceptionResNetV2 ve Xception modelleri seçilmiştir.

2. Materyal ve Metot

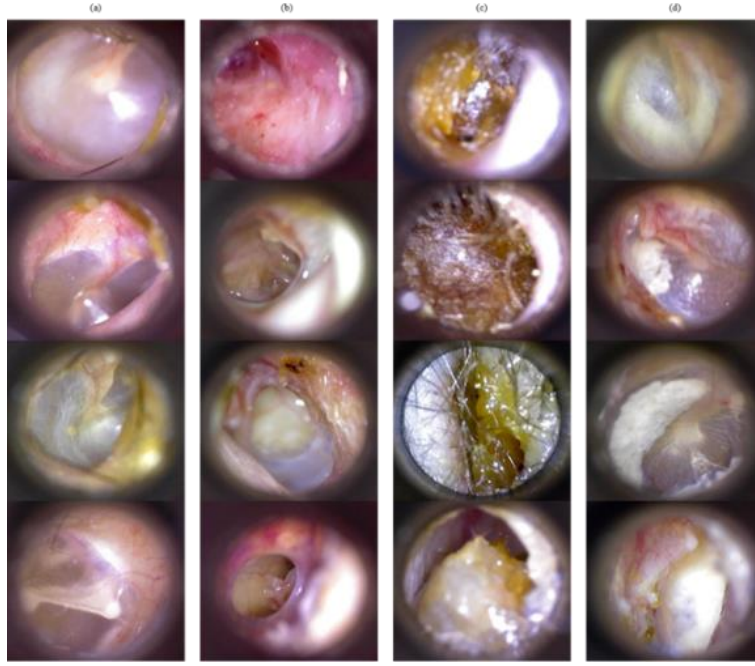
Çalışma kapsamında sınıflandırma işleminden önce tüm görüntüler ön işlem adımından geçirilmiştir. Ön işlem adımında, görüntülerin ilgili derin öğrenme modellerinde kullanılabilmesi için yeniden boyutlandırılmıştır. Orijinal görüntü boyutunda GPU hafızasının yetersiz gelmesinden dolayı görüntüler küçültülmüştür. Elde edilen işlenmiş görüntüler seçilen derin öğrenme algoritmalarında eğitim ve test işlemlerine tabi tutulmuştur. Daha sonrasında modellerden elde edilen sonuçlar değerlendirilmiştir ve ağların başarıları kendi içlerinde karşılaştırılmıştır. Şekil 2’de yapılan çalışmaya ait akış şeması verilmiştir ve akış şemasında gerçekleştirilen işlemler modellenmiştir.



Şekil 2. Yapılan çalışmaya ait akış şeması (Figure 2. Flow chart of the study)

2.1. Veri Seti

Veri seti olarak Ear Imagery veri seti kullanılmıştır (Viscaino ve diğerleri, 2020). Bu veri seti Universidad de Chile Klinik Hastahanesi Kulak Burun Boğaz bölümünden elde edilen 7 ile 65 yaşındaki 180 hastaya ait dijital otoskop altında elde edilmiş toplamda 880 görüntüden oluşturulmuştur. Veri setinde normal, CSOM, miringoskleroz ve kulak mumu tıkacı (earwax plug) olmak üzere 4 sınıf ve her sınıf için 220 görüntü bulunmaktadır. Görüntüler 420x380 piksel boyutunda ve JPG formatındadır (Viscaino ve diğerleri, 2020). Veri setine ait görüntülerden sınıf bazında örnekler Şekil 3’te verilmektedir.



Şekil 3. Veri setinde bulunan sınıflara ait görüntüler örnekleri (a) Normal, (b) CSOM, (c) Kulak Mumu ve (d) Mirinoskleroz (Figure 3. Image example of the classes in the dataset (a) Normal, (b) CSOM, (c) Earwax and (d) Myringosclerosis)

Tablo 1. Ear imageri veri setindeki verilerin dağılımı (Table 1. Distribution of data in the Ear imageri dataset)

Açıklama	Sınıf	Veri Sayısı
Sağlıklı Kulak Zarı	Normal	220
Hastalıklı Kulak Zarı	CSOM	220
	Kulak Mumu Tıkacı	220
	Mirinoskleroz	220

Tablo 1’de seçilen modellerin eğitim, doğrulama ve test işlemlerinin uygulandığı veri setinde yer alan sınıflara ve sınıflara ait veri sayılarına yer verilmiştir. Veri setinde bulunan görüntüler %82 eğitim ve %18 test işlemleri için ayrılmıştır. Eğitim ve test veri setlerinde sırasıyla 720 ve 160 görüntü bulunmaktadır. Eğitim işleminde kullanılacak veri kümesinin %20’si doğrulama işleminde kullanılmıştır. Doğrulama veri setinde 144 görüntü bulunmaktadır.

2.2. Önişleme

Veri setinde bulunan görüntülerin orijinal boyutu 420x380 pikseldir. Eğitim ve test işlemlerinde kulak hastalıklarına ait tüm görüntüler modellere uygun olarak sırasıyla AlexNet 227x227, ResNet50 224x224, ResNet50V2 224x224, ResNet101 224x224, ResNet101V2 224x224, InceptionV3 299x299, InceptionResNetV2 299x299 ve Xception 299x299 boyutunda ayarlanmıştır ve görüntüler renkli olarak kullanılmıştır. Ayrıca eğitim ve doğrulama verileri üzerinde karıştırma yapılmamıştır. Veri setine ait görüntülerde herhangi bir arttırma (augmentation) işlemi kullanılmamıştır.

2.3. Derin Öğrenme Algoritmaları

Derin öğrenme modelleri, üç veya daha fazla katmana sahip olan sinir ağı modellerine denmektedir ve makine öğrenmenin bir alt kümesidir. Derin öğrenme, çalıştığı veri türü ve öğrendiği yöntemlerle kendisini klasik makine öğrenmesinden ayırmaktadır. Derin öğrenme, makine öğrenmesinde gereken bazı ön işlem adımlarını ortadan kaldırmaktadır. Örneğin makine öğrenmesinde ilgili verilere ait özniteliklerin verilerden çıkarılarak makine öğrenmesi modeline verilmesi gerekmektedir. Öte yandan derin öğrenmede bu işlem model tarafından veriler üstünde eğitim işlemi sırasında otomatik olarak gerçekleştirilmektedir. Veriler üzerindeki örüntülerin tanınabilmesi için derin öğrenme modelleri makine öğrenmesine göre daha fazla veriye ihtiyaç duymaktadır (IBM Cloud, 2020).

Evrişimsel Sinir Ağları (CNN), farklı işlemlere sahip katmanlardan oluşan derin ağ modelleridir. Evrişim (convolution), havuzlama (pooling) ve tam bağlı (fully-connected) katman yaygın olarak kullanılan katmandan bazılarıdır (Başaran ve diğerleri, 2020).

Evrişim katmanı (CONV): Bu katman giriş olarak verilen görüntülerden öznitelik çıkartmak için kullanılmaktadır. Önceki katmanlardan elde edilen özniteliklere X_i^{l-1} dersek, bu katmanda öznitelik Denklem (1)’den elde edilmektedir (Başaran ve diğerleri, 2020).

$$X_j^l = f\left(\sum_{i \in M_j} X_i^{l-1} * k_{ij}^l + b_j^l\right) \quad (1)$$

Denklem (1)'de, k_{ij}^l öğrenilebilir kernel değerini, b_j^l eğitilebilir bias değerini, $f(\cdot)$ aktivasyon fonksiyonunu ve M_j giriş harita seçimini belirtmektedir.

Havuzlama katmanı (POOL): Havuzlama katmanları, ortalama veya maksimum gibi teknikleri kullanarak alt örnekleme işlemini gerçekleştirmektedir. Bu işlem sonucunda işlem düğümü sayısı ve aşırı öğrenmenin gerçekleşmesini azaltmaktadır. Denklem 2'de bu işlem gösterilmektedir (Başaran ve diğerleri, 2020).

$$X_j^l = \text{down}(X_i^{l-1}) \quad (2)$$

Denklem (2)'de, $\text{down}(\cdot)$ fonksiyonu alt-örneklem işlemini göstermektedir ve bu işlem sonucunda ayırt edici özellikler elde edilmektedir.

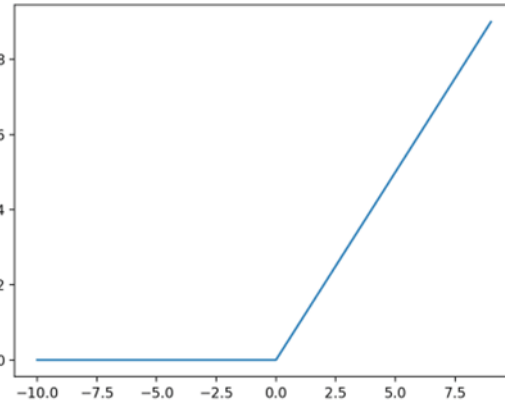
Tam-bağlı katman (FC): Tam-bağlı katmanın amacı, giriş görüntüsünü eğitim veri setine dayalı olarak sınıflandırmak için önceki katmanlardan elde edilen ayırt edici özellikleri kullanmaktır (Başaran ve diğerleri, 2020).

AlexNet mimarisi ILSVRC-2012'de tanıtılan basit, öncü derin öğrenme modelidir. Bu mimari toplamda sekiz katmandan oluşmaktadır: beş evrişimsel katman ve üç tam bağlı katman. Bazı katmanlar arasında havuzlama katmanı kullanılmaktadır. AlexNet'i diğer mimarilerden farklı kılan özellikler aşağıdaki gibidir (Yegualp, 2022).

ReLU (Rectified Linear Unit) doğrusal olmama (Nonlinearity): AlexNet çıktığı zaman popüler olan tanh aktivasyon fonksiyonu yerine ReLU aktivasyon fonksiyonunu kullanmayı tercih etmiştir. ReLU'nun avantajı eğitim süresidir; ReLU kullanan bir CNN modeli, tanh kullanan modele göre CIFAR-10 veri kümesinde %25 hata oranına altı kat daha hızlı ulaşmaktadır. ReLU aktivasyon fonksiyonu, girdinin pozitif kısmını alan ve negatif kısmı sıfırlayan aktivasyon fonksiyonudur.

$$f(x) = \max(0, x) \quad (3)$$

Denklem (3)'te x aktivasyon fonksiyonunun kullanıldığı nöronun girdisidir (He ve diğerleri, 2015). Şekil 4'te ReLU fonksiyonun grafiğine yer verilmiştir. Ağda bulunan katman sayısı artırıldığında, kısmi türevden elde edilen değer sıfıra yaklaşmaktadır ve bir noktadan sonra kaybolmaktadır. Bu probleme kaybolan gradyan problemi (vanishing gradient problem) denmektedir (Jacob, 2022). Sigmoid ve tanh aktivasyonları bu probleme yatkındır. Buna karşın, ReLU aktivasyonu kaybolan gradyan probleminden etkilenmemektedir (Dash, 2021).



Şekil 4. ReLU grafiği (Figure 4. ReLU graph) (Dash, 2021)

ResNet, ağ yapısı genel olarak VGG-19 mimarisine atlama bağlantılarının eklenmiş formatıdır. Kaybolan/patlayan gradyan problemini çözmek için bu mimari, bazı katmanları atlayarak direkt aktivasyon fonksiyonuna bağlanan atlama bağlantılarını (skip connection) kullanmaktadır. Bu yapıları Artık Bloklar (Residual Blocks) adı verilmektedir. Denklem 4'te gösterildiği üzere Artık Bloklar önceki girdi değeriyle katman sonucunda elde edilen çıktılarının toplanmasıdır. Bu yapıların avantajı, model performansını kötü etkileyen katmanlar düzenleme tarafından atlanmaktadır (Boesh, 2022).

$$y = F(x, \{W_i\}) + x \quad (4)$$

Denklem (4)'te, x ve y katmanların girdi ve çıktılarıdır. $F(x, \{W_i\})$ fonksiyonu öğrenilen haritalamayı temsil etmektedir. $F(x)+x$ işlemi atlama katmanlarında ve öge-bazında eklemelerde (element-wise addition) gerçekleştirilmektedir (He ve diğerleri, 2016).

Bazı önemli ResNet mimarileri şunlardır:

ResNet34 mimarisi; İlk oluşturulan ResNet mimarisidir. VGGNet'lerden esinlenen bu mimari daha az filtreye ve karmaşıklığa sahiptir. İki basit tasarım kuralı bulunmaktadır; katmanlar, aynı çıktı özellik haritası boyutu için aynı sayıda filtreye sahip olmalıdır ve

katman başına zaman karmaşıklığını korumak için özellik haritası boyutunun yarıya indirilmesi durumunda filtre sayısı iki katına çıkartılmalıdır.

ResNet50 mimarisi; genel olarak ResNet34 mimarisiyle benzerlik göstermesine rağmen büyük bir fark bulundurmaktadır. Yapı bloğu, katmanları eğitmek için geçen zamanla ilgili endişeler nedeniyle bir darboğaz tasarımına dönüştürülmüştür. Bu yapı önceki kullanılan 2 yerine 3 katmandan oluşan bir yığın kullanılmaktadır. Bu nedenle, ResNet34'teki 2 katmanlı blokların her biri, ResNet50 mimarisinde 3 katmanlı darboğaz bloğu (Bottleneck Block) ile yer değiştirilmiştir. ResNet101 ve ResNet152 mimarileri, 101 katmanlı ve 152 katmanlı ResNet mimarileridir ve ResNet50'de kullanılan 3 katmanlı bloklardan daha fazla kullanılarak oluşturulmaktadır.

Inception ağı, Inception modülleri olarak adlandırılan yinelenen bileşenlerden oluşturulan mimari tasarıma sahip derin öğrenme ağıdır. Inception ağının oluşturan fikirler aşağıdaki sıralanabilir (Alake, 2020);

- Yüksek performanslı derin sinir ağlarının büyük olması gerekir.
- Bir sinir ağının büyük olarak kabul edilmesi için, ağ içinde birkaç katmana ve bu katmanlar içindeki birimlere sahip olması gerekmektedir.
- Evrimsel sinir ağları, değişen ölçeklerde özneteliklerin çıkarılmasından yararlanmaktadır. Değişik boyutlarda convnet kullanılması modelin daha fazla öznetelik öğrenme potansiyeline sahiptir.
- Hebbian Prensibinin dikkate alınmasıyla; *birlikte ateşlenen nöronlar, birbirlerine bağlanmaktadır.*

Yukarıdaki fikirler birkaç dezavantajı barındırmaktadır. Büyük ağlar aşırı uymaya (overfit) eğilimlidir ve birden çok evrişim işlemi birbirine bağlamak ağın hesaplama maliyetini arttırmaktadır. Bu iki problemi ortadan kaldırmak için araştırmacılar sezgisel convnet mimarisini tasarlamıştır (Alake, 2020).

1x1 boyutunda evrişim kullanılarak, katmana giren girdinin çıktı boyutu azaltılmaktadır. Öte yandan 3x3 ve 5x5 boyutunda evrişim kullanılarak, ağın farklı ölçeklerde uzamsal örüntüleri öğrenmesi sağlanmaktadır. 1x1 boyutunda filtreler, girdinin derinliklerindeki örüntüleri öğrenmektedir. 3x3 ve 5x5 boyutundaki filtreler, girdinin tüm boyutlu bileşenlerinde uzamsal örüntüleri öğrenmektedir. Bu işlemler sonucunda, evrişim için kullanılan filtrelerin bölünmesinin daha az maliyetli olduğu ortaya çıkmıştır. Örneğin çıktı olarak 3x3 filtre yerine, 3x1 ile 1x3 filtrenin kullanılması aynı çıktıyı vermektedir ve iş yükü %33 daha azdır. Bu donanım için ağır yük oluşturan evrişim işleminin maliyetini azaltmıştır (Szegegy ve diğerleri, 2016).

InceptionResNet ve Inception V4 aynı makalede tanıtılmıştır. InceptionResNet mimarisi, ResNet performansından esinlenmektedir. InceptionResNet V1 ve V2 olmak üzere iki alt versiyondadır. Bu iki versiyon arasında küçük fark bulunmaktadır (Raj, 2018). InceptionResNetV1, InceptionV3 benzeri ve InceptionResNetV2, InceptionV4 benzeri hesaplama maliyetine sahiptir. İki mimari yapıca aynı olmasına karşın hiperparametre seçimleri farklıdır.

Xception mimarisi, Google ekibi tarafından standart Inception mimarisi baz alınarak geliştirilmiştir. Xception mimarisinde 71 katman bulunmaktadır. Bu ağ modelinde Inception mimarisinden farklı olarak Inception modülleri derinlemesine ayrılabilir evrişimlerle değiştirilmiştir. Bu evrişim modülleri derinlemesine evrişimin ardından noktasal evrişim uygulamaktadır. Derinlemesine ayrılabilir ayrışmalar normal evrişim işlemlerinden daha az maliyetlidir. Bunun sebebi bütün kanallarda evrişim yapmak yerine 1 evrişim uygulanmaktadır. Ayrıca Xception mimarisinde ResNet mimarisinde bulunan atlama bağlantıları kullanılmaktadır (Fabien, 2019).

2.4. Keras

Keras, Python programı diliyle yazılmış Tensorflow platformu üzerinde çalışan derin öğrenme API'sidir. Tensorflow, Google Brain takımı tarafından geliştirilmiş açık kaynaklı uçtan uca platformdur. Tensorflow Python ve Javascript kullanarak ön yüz API sağlamaktadır ve bu API'lerin çalıştırılmasını C++ ile gerçekleştirmektedir. Tensorflow, veri akış diyagramlarını kullanmaktadır. Bu yapılar verilerin grafiklerdeki veya işleme düğümlerindeki hareketini tanımlamaktadır (Yegualp, 2022).

2.5. Performans Metrikleri

Nesne tanıma görevlerinde genel olarak kullanılan ortalama kesinlik (average-precision) ve geri çağırma (recall) metrikleri model değerlendirmedeki performans ölçütümüzdür (Başaran ve diğerleri, 2020). Sınıflandırma işleminde genel ölçüm değerleri pozitif doğru (true positive - DP), pozitif yanlış (false positive - YP), negatif yanlış (false negative - YN) ve negatif doğru (true negative - DN) oluşan karışıklık (hata) matrisinden elde edilmiştir. Ölçüm değerlerinin bir başka ifadeyle performans kriterlerinin formülasyonu hata matrisindeki indisler yardımı ile Denklem (5) ve Denklem (9) arasında ifade edilen şekilde hesaplanmaktadır.

$$\text{Doğruluk} = \frac{DP + DN}{DP + YP + YN + DN} \quad (5)$$

$$\text{Duyarlılık} = \frac{DP}{DP + YN} \quad (6)$$

$$\text{Özgünlük} = \frac{DN}{DN + YP} \quad (7)$$

$$F - \text{Skor} = \frac{2 * DP}{2 * DP + YP + YN} \quad (8)$$

$$\text{Ortalama - Doğruluk} = \frac{DP}{DP + DN} \quad (9)$$

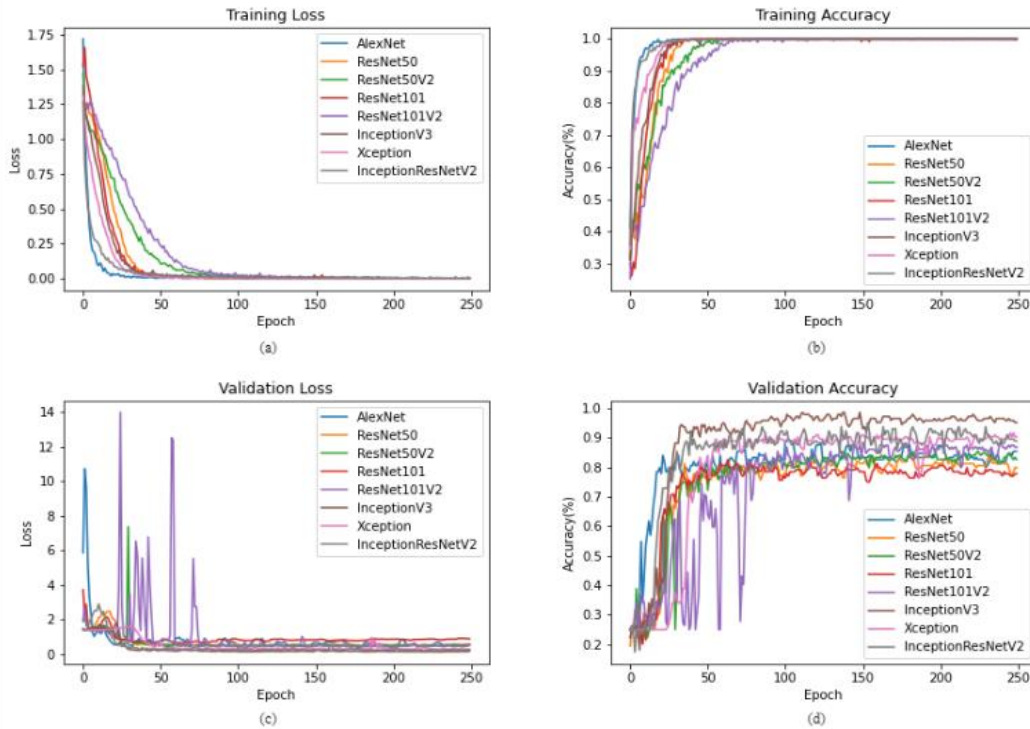
Sınıflandırma işleminde, DP ve DN doğru tanımlanmış, YP ve YN yanlış tanımlanmış hastalıklı ve sağlıklı kulak zarı görüntülerine karşılık gelmektedir. Duyarlılık metriğiyle, modelin pozitif örnekleri ne kadar iyi tespit edebildiğinin ölçüsüdür. Bu metrik doğru pozitif oran (TPR) olarak bilinmektedir. Modelin doğru sınıflandırma yapabilmesi için bu metriğin yüksek olması gereklidir. Öte yandan özgünlük metriği, modellerin doğru negatif sınıflandırma oranlarını hesaplamaktadır. Bu metrik doğru negatif oran (TNR) olarak adlandırılmaktadır. Özgünlük ile yanlış pozitif toplamı her zaman 1'e eşittir. Yüksek özgünlük değeri modelin doğru olarak negatif değerleri bulduğunu göstermektedir.

Ayrıca modellerin değerlendirilmesinde ROC eğrisi kullanılmıştır. ROC eğrisi (Receiver Operating Characteristic curve), tüm sınıflandırma eşik değerlerinde sınıflandırma modelinin performansını gösteren grafikdir. Bu grafik TN ve FN parametrelerini kullanmaktadır. Çizilen grafik eğrisi FPR (False Positive Rate) = TPR (True Positive Rate) eğrisinin üstünde ise model başarılı, altında ise model başarısız olarak varsayılmaktadır. ROC eğrisi, sınıf dağılımına bağlı değildir. Bu yüzden nadir olayları tahmin eden sınıflandırıcılar için ROC'u kullanışlı kılmaktadır. AUC (Area Under the ROC Curve), (0,0) ile (1,1) arasındaki tüm ROC eğrisinin altındaki alanı ölçmektedir. AUC, tahminlerin ne kadar iyi sıralandığını ölçmektedir (Google Developer, 2022).

3. Araştırma Sonuçları ve Tartışma

3.1. Araştırma Sonuçları

Yapılan çalışmada tüm modeller aynı hiperparametre değerlerinde çalıştırılmıştır. Öğrenme işleminde epoch sayısı 250 ve yığın boyutu (batch size) 32 olarak seçilmiştir. Öğrenme oranı 0,0001 seçilmiştir. Optimizasyon işlemi için SGD (Stochastic Gradient Descent) kullanılmıştır ve momentum 0,8 olarak ayarlanmıştır. Optimizasyon sırasında modellerin katmanlarına L2 düzenleyiciyle (L2 regularizer) 0,001 oranında penaltı uygulanmıştır. Bu işlemle modellerin aşırı öğrenmesi veya yetersiz öğrenmesi önlenmektedir. Doğrulama başarısı en yüksek olduğu durumda modelin ağırlık değerleri kaydedilmiştir. Bu sayede aşırı öğrenmenin olumsuzlukları önlenmiştir. İlgili veri setinde bulunan görüntüler mimarilerin eğitim ve test işlemlerinde kullanılabilmesi için sırasıyla %82 (720 görüntü) ve %18 (160 görüntü) olarak ayrılırken, eğitim için oluşturulan veri kümesinin %20'si (144 görüntü) de doğrulama işlemi için ayrılmıştır. Görüntüler modellere uygun olarak sırasıyla AlexNet 227x227, ResNet50 224x224, ResNet50V2 224x224, ResNet101 224x224, ResNet101V2 224x224, InceptionV3 299x299, InceptionResNetV2 299x299 ve Xception 299x299 boyutunda renkli olarak kullanılmıştır. Eğitim, doğrulama ve test işlemleri NVIDIA RTX3060 6 GB dizüstü bilgisayarda CUDA mimarisinden yararlanılarak GPU üzerinde gerçekleştirilmiştir. Mimarilerin doğruluğu "Kategorik Çapraz Entropi" (Categorical Cross-Entropy) baz alınarak karşılaştırılmıştır. Eğitim işlemi sonucunda doğruluk oranı tüm mimarilerde %100'dür.



Şekil 5. Seçilen mimariler için eğitim ve doğrulama adımları sonucunda elde edilen kayıp ve doğruluk değerleri. Mimarilerin (a) Eğitim kaybı, (b) Eğitim doğruluğu, (c) Doğrulama kaybı ve (d) Doğrulama doğruluğudur. (Figure 5. Loss and accuracy values obtained as a result of training and validation steps for selected architectures. Architectures are (a) Training loss, (b) Training accuracy, (c) Validation loss and (d) Validation accuracy.)

Şekil 5'te modellerin eğitim ve doğrulama işlemlerindeki başarısı görülmektedir. Tüm modellerin eğitim doğruluğu 100 epoch sonunda 1'e ulaşmıştır ve sabitlenmiştir. ResNet50V2 ve ResNet101V2 modelleri doğrulama işleminde aşırı öğrenmeye yaklaşmış fakat modeller doğrulama işleminin ilerleyen aşamalarında aşırı öğrenmeden kurtulmuştur. Eğitim ve doğrulama işleminde kayıp için en hızlı ivme AlexNet mimarisinde görülmüştür.

Tablo 2. Deneysel sonuçlar (Table 2. Experimental Results)

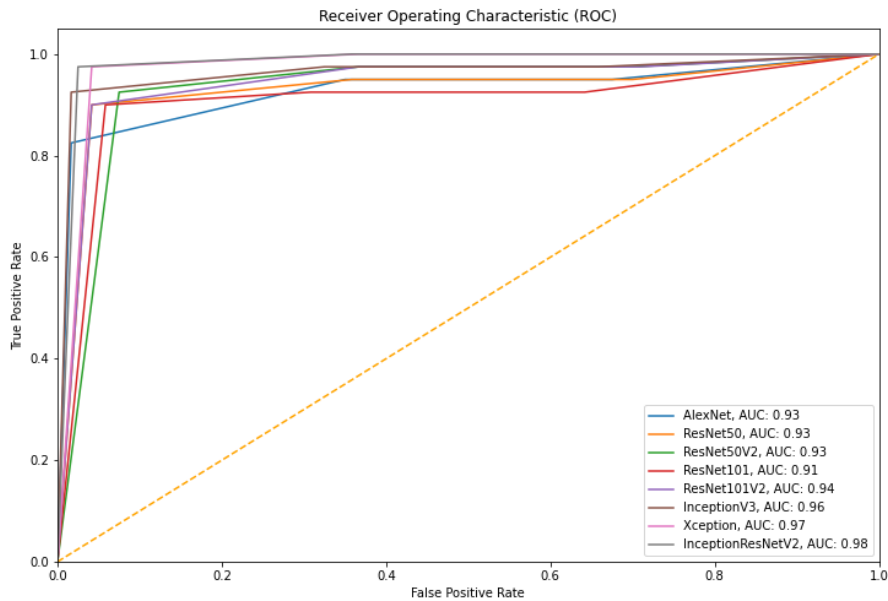
Modeller	Epoch	Öğrenme Oranı	Eğitim Zamanı (s)	Yığın Boyutu	Kayıp	Görüntü Boyutu
AlexNet	250	0.0001	438	32	0.0016	227x227
ResNet50	250	0.0001	770	32	0.0010	224x224
ResNet50V2	250	0.0001	618	32	0.0094	224x224
ResNet101	250	0.0001	1203	32	0.0109	224x224
ResNet101V2	250	0.0001	1021	32	0.0030	224x224
InceptionV3	250	0.0001	474	32	0.0077	299x299
Xception	250	0.0001	927	32	0.0005	299x299
InceptionResNetV2	250	0.0001	990	32	0.0084	299x299

Tablo 2’de eğitim işlemi için seçilen hiperparametre değerleri, modellerin eğitim süreleri ve kayıp değerleri verilmektedir. Bu sonuçlara göre eğitim aşamasında eşit hiperparametreler için en hızlı öğrenme süresi 438 saniye ile AlexNet mimarisinden elde edilmiştir. Kayıp değeri olarak en düşük değer Xception mimarisinde erişilmiştir.

Tablo 3. Seçilen mimarilere ait performans sonuçları (Table 3. Performance results for selected architectures)

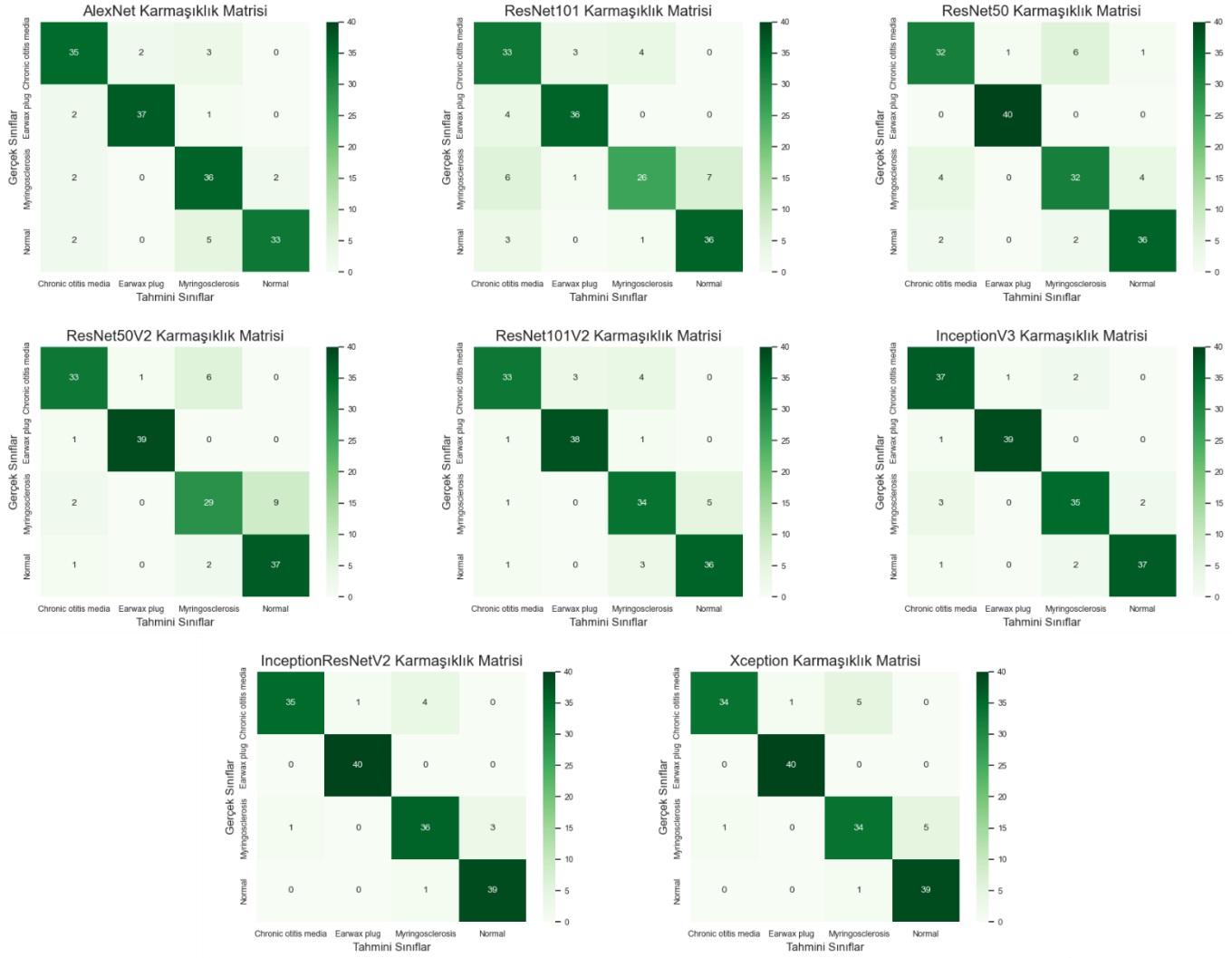
Modeller	Doğruluk (%)	Duyarlılık (%)	Özgünlük (%)	F-SCOR (%)	Ortalama - Doğruluk (%)
AlexNet	88	88	96	88	89
ResNet50	88	88	95	87	87
ResNet50V2	86	86	95	86	86
ResNet101	82	82	94	82	82
ResNet101V2	88	88	96	88	88
InceptionV3	93	93	97	93	93
Xception	92	92	97	92	92
InceptionResNetV2	94	94	98	94	94

Tablo 3’te test işlemi sonucunda elde edilen bulgular yer almaktadır. Elde edilen bulgular ışığında eşit hiperparametre değerleri için en yüksek başarı oranı %94 ile InceptionResNetV2 mimarisinden elde edilmiştir. Yine Tablo 3’te yer alan %94 duyarlılık ve %98 özgünlük değeriyle kullanılan modeller içinde InceptionResNetV2 modelinin doğru pozitif ve doğru negatif etiketleri ayırma performansının yüksek olduğu ortaya çıkmaktadır.



Şekil 6. Modeller için ROC grafiği (Figure 6. ROC graph for models)

Şekil 6’da modellerin test işlemleri sonucunda elde edilen TP ve FP değerleri üzerinden oluşturulan ROC grafiği ve AUC değerleri verilmektedir. Turuncu eğri FPR=TPR eğrisidir ve tüm modeller bu eğrinin üstündeki başarılı bölgededir. Şekilden elde edilen sonuçlara göre en yüksek AUC değeri InceptionResNetV2 modelinden 0.98 olarak elde edilmiştir.



Şekil 7. Test işlemi sonucunda elde edilen karmaşıklık matrisi (Şekil 7. The confusion matrix obtained as a result of the test process)

Şekil 7’de test işlemi sonucunda modellere ait karmaşıklık matrisleri verilmiştir. Test veri setinde toplamda 160 görüntü ve 4 sınıf bulunduğu için dolayı matris 4x4 boyutundadır ve [1,1] ile [4,4] arasındaki çizgi sınıf bazında doğru sınıflandırılan görüntü sayısını vermektedir.

3.2. Tartışma

Kullanılan veri setinde üç tip kulak hastalığı ve sağlıklı kulak zarına ait görüntüler bulunmaktadır. Öte yandan sadece otitis media hastalığına ait üç tip hastalık bulunmaktadır ve yapılan çalışmada sadece CSOM hastalığının sınıflandırılması yapılmıştır. Geniş hastalık çeşidi bulunan KBB sahasında yapılan çalışma için bu bir eksiktir. Daha geniş çaplı ve daha fazla görüntü bulunduran bir veri seti kullanımı yapılan çalışmanın KBB alanındaki kullanılabilirliğini ve başarısını artırabilmektedir.

Kullanılan modeller, içerisinde derin ağ mimarilerini barındırmaktadır. Modellerin eğitim ve test işlemleri karmaşıklık artırmaktadır ve bu durum donanımlar için daha maliyetli hale gelmektedir. Yapılan çalışma kapsamında grafik kartı hafıza kullanımını azaltmak için görüntü boyutları düşürülmüştür. Bu işlem sonucunda görüntülerde gürültüler oluşmaktadır ve modellerin performansını etkilemektedir.

Otoskop tarafından elde edilen kulak hastalıklarına ait görüntülerin bilgisayar destekli teşhisi neredeyse hatasızdır. Kulak hastalıklarının teşhisi uzmanlar tarafından da aynı prensip izlenilerek otoskop kullanılarak görüntü bazlı konulmaktadır. Bu işlem sonucunda koyulan teşhis uzmandan uzmana değişiklik gösterebileceği gibi hastaya koyulan teşhis farklı muayenelerde değişebilmektedir. Bu işlemin bilgisayar destekli hale getirilmesi kulak hastalığı tanılarındaki uzman hatalarını, yanlış koyulan teşhis sonucunda hastalığın ilerlemesini ve/veya kötüleşmesini önleyebileceği gibi doğru tedavinin erkenden uygulanmasını sağlayabilmektedir.

Tablo 4. Kulak zarı hastalıklarının teşhisi için kullanılan yöntemler arasında karşılaştırma (Table 5. Comparison between methods used for diagnosis of eardrum diseases)

Yazarlar	Metotlar	# örnekler	# sınıflar	Acc (%)
2021, Wu et al. (Wu ve diğerleri, 2021)	Xception	12203	3	97,82
2021, Alhudhaif et al. (Alhudhaif ve diğerleri, 2021)	Novel-multi level machine learning alghorithm	956	8	98,26
2020, Başaran et al. (Başaran ve diğerleri, 2020)	Faster R-CNN, pretrained CNN models, VGG16, 10-fold cross-validation	1692	2	90,48
2022, Mohammed et al. (Mohammed ve diğerleri, 2022)	Bayesian Optimization based on CNN-LSTM Architecture	880	4	100
2021, Uçar et al. (Uçar ve diğerleri, 2021)	Fused deep hypercolumn features and bidirectional LSTM	880	4	99,06
2018, Tran et al. (Tran ve diğerleri, 2018)	Multitask joint sparse representation algorithm	214	2	91,41
2021, Sungaard et al. (Sungaard ve diğerleri, 2021)	Five different loss function with InceptionV3	1336	3	85
2022, This paper	InceptionResNetV2	880	4	94

Tablo 4’de görüldüğü üzere KBB alanında farklı veri setlerine, yöntemlere ve doğruluk oranlarına sahip çalışmalar bulunmaktadır. Bu çalışmada daha önce yapılan bazı çalışmalardan farklı olarak OM hastalığıyla birlikte diğer kulak zarı hastalıkları da sınıflandırılmıştır. Aynı veri seti üzerinde Mohammed ve arkadaşlarının geliştirdiği CNN-BiLSTM mimarisi %100 ile en yüksek başarıya sahiptir (Mohammed ve diğerleri, 2022). Öte yandan en fazla kulak hastalığının sınıflandırılması Alhudhaif ve arkadaşların yapmış olduğu çalışmada görülmektedir (Alhudhaif ve diğerleri, 2021).

4. Sonuç

Bu çalışmada, kulak içi hastalıklarına ait görüntülerin farklı derin öğrenme modellerindeki sınıflandırma performansları karşılaştırılmıştır. Derin öğrenme modellerinin eğitim, doğrulama ve test işlemleri için 4 sınıf ve toplamda 880 görüntünün bulunduğu kulak zarı veri seti seçilmiştir. Sınıflandırma işlemini gerçekleştirmek için derin öğrenme modeli olarak AlexNet, ResNet50, ResNet50V2, ResNet101, ResNet101V2, InceptionV3, InceptionResNetV2 ve Xception mimarileri seçilmiştir. Sınıflandırma işlemi sonucunda derin öğrenme modelleri sırasıyla %88, %88, %86, %82, %88, %93, %94 ve %92 sınıflandırma başarıları göstermiştir. En yüksek doğruluk oranı %94 ile InceptionResNetV2 mimarisinde elde edilirken, 438 saniye ile en hızlı eğitim işlemi AlexNet mimarisinde gerçekleşmiştir. Başarı ve eğitim hızı olarak en uygun model Inception V3 modelidir. Eğitimi süresi 474 saniye ve sınıflandırma başarı değeri %93’tür.

Sonuç olarak bilgisayar destekli tanı ve karar destek sistemleri çerçevesinde kulak zarına ilişkin farklı hastalık tiplerinin otoskop imgelerine dayalı olarak otonom bir şekilde tespit edilmesinin mümkün olduğu bu çalışma kapsamında doğrulanmıştır. Bilgisayar destekli hesaplamalı sistemlerin klinik pratikte kullanılması ve yaygınlaşması, hataya açık olan ve sübjektif olarak yürütülen değerlendirme sürecinin daha objektif bir şekilde gerçekleştirilmesine, kliniksel muayenelerin geriye dönük değerlendirmesi ve bu süreçteki hataların azaltılmasına katkı sağlama potansiyeli taşımaktadır.

Kaynakça

- Alake, R. (2020, 22 Aralık). Deep Learning: Understanding The Inception Module. Erişim adresi: <https://towardsdatascience.com/deep-learning-understand-the-inception-module-56146866e652>
- Alhudhaif, A., Cömert, Z. ve Polat, K. 2021. “Otitis media detection using tympanic membrane images with a novel multi-class machine learning algorithm”, PeerJ Computer Science, 7, 405
- Başaran, E., Cömert, Z. ve Çelik, Y. 2020. “Convolutional neural network approach for automatic tympanic membrane detection and classification”. Biomedical Signal Processing and Control, 56, 101734.
- Boesh, G. (2022). Deep Residual Networks (ResNet, ResNet50) – Guide in 2022. Erişim adresi: <https://viso.ai/deep-learning/resnet-residual-neural-network/>
- Dash, A. B. (2021, 30 Haziran). Top 10 Activation Function's Advantages & Disadvantages. Erişim adresi: <https://www.linkedin.com/pulse/top-10-activation-functions-advantages-disadvantages-dash>
- Fabien, M. (2019, 20 Mart). Xception Model and Depthwise Separable Convolutions. Erişim adresi: <https://maelfabien.github.io/deeplearning/xception/#>
- Google Developer. (2022, 18 Temmuz). Classification: ROC Curve and AUC. Erişim adresi: <https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc>
- Hayes, K. (2022, 11 Nisan). An Overview of Myringosclerosis and Tympanosclerosis. Erişim adresi: <https://www.verywellhealth.com/myringosclerosis-and-tympanosclerosis-1191943#:~:text=Myringosclerosis%20and%20tympanosclerosis%20are%20similar,due%20to%20accumulated%20calcium%20deposits.>

- He, K., Zhang, X., Ren, S. ve Sun, J. 2015. "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification". In Proceedings of the IEEE international conference on computer vision, 1026-1034
- He, K., Zhang, X., Ren, S. ve Sun, J. 2016. "Deep residual learning for image recognition". In Proceedings of the IEEE conference on computer vision and pattern recognition, 770-778
- IBM Cloud. (2020, 1 Mayıs). Deep Learning. Erişim adresi: <https://www.ibm.com/cloud/learn/deep-learning#:~:text=Deep%20learning%20is%20a%20subset,from%20large%20amounts%20of%20data>.
- Jacob, T. (2022, 25 Şubat). Vanishing Gradient Problem, Explained. Erişim adresi: <https://www.kdnuggets.com/2022/02/vanishing-gradient-problem.html#:~:text=When%20there%20are%20more%20layers,this%20the%20vanishing%20gradient%20problem>.
- Mayo Clinic (2022, 12 Temmuz). Earwax Blockage. Erişim adresi: <https://www.mayoclinic.org/diseases-conditions/earwax-blockage/symptoms-causes/syc-20353004>
- Mohammed, K. K., Hassanien, A. E. ve Afify, H. M. 2022. "Classification of Ear Imagery Database using Bayesian Optimization based on CNN-LSTM Architecture". Journal of Digital Imaging, 1-15
- Raj, B. (2018, 29 Mayıs). A Simple Guide to the Versions of the Inception Network. Erişim adresi: <https://towardsdatascience.com/a-simple-guide-to-the-versions-of-the-inception-network-7fc52b863202>
- Sahling, M., Benders, T., König, O., Boll-Avetisyan, N. 2021. Building a Phoneme Inventory through Blocked Ears: The Effects of Prior Otitis Media with Effusion on Children's Phoneme Discrimination.
- Serdar Yegualp. (2022, 3 Haziran). What is TensorFlow? The machine learning library explained. Erişim adresi: <https://www.infoworld.com/article/3278008/what-is-tensorflow-the-machine-learning-library-explained.html>
- Sundgaard, J. V., Harte, J., Bray, P., Laugesen, S., Kamide, Y., Tanaka, C. ve Christensen, A. N. 2021. "Deep metric learning for otitis media classification". Medical Image Analysis, 71, 102034.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. ve Wojna, Z. 2016. "Rethinking the inception architecture for computer vision". In Proceedings of the IEEE conference on computer vision and pattern recognition, 2818-2826
- Tran, T. T., Fang, T. Y., Pham, V. T., Lin, C., Wang, P. C. ve Lo, M. T. 2018. "Development of an automatic diagnostic algorithm for pediatric otitis media". Otology & Neurotology, 39, 1060-1065.
- Uçar, M., Akyol, K., Atila, Ü. M. İ. T. ve Uçar, E. 2021. "Classification of different tympanic membrane conditions using fused deep hypercolumn features and bidirectional LSTM". IRBM.
- Viscaino, M., Maass, J. C., Delano, P. H., Torrente, M., Stott, C. ve Auat Cheein, F. 2020. "Computer-aided diagnosis of external and middle ear conditions: A machine learning approach". Plos one, 15, 0229226.
- Wu, Z., Lin, Z., Li, L., Pan, H., Chen, G., Fu, Y. ve Qiu, Q. 2021. "Deep learning for classification of pediatric otitis media". The Laryngoscope, 131, 2344-2351.
- Yavuz, H ve arkadaşları. (2000, Ekim). Otitis Media. Erişim adresi: <https://www.ttb.org.tr/STED/sted1000/2.html#top>