

T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

ÇEKİŞMELİ ÜRETİCİ AĞ TABANLI VERİ ARTIRIMI
YAKLAŞIMI İLE OTOSKOP GÖRÜNTÜLERİNDEN KULAK
ZARI PATOLOJİLERİNİN TESPİTİ

Mustafa Furkan ESEOĞLU

YÜKSEK LİSANS TEZİ
Bilgisayar Mühendisliği Anabilim Dalı
Bilgisayar Mühendisliği Programı

Danışman
Prof.Dr. M. Elif KARSLIGİL

Temmuz, 2022

T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

**ÇEKİŞMELİ ÜRETİCİ AĞ TABANLI VERİ ARTIRIMI YAKLAŞIMI İLE
OTOSKOP GÖRÜNTÜLERİNDEN KULAK ZARI PATOLOJİLERİNİN
TESPİTİ**

Mustafa Furkan ESEOĞLU tarafından hazırlanan tez çalışması 19.07.2022 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı Bilgisayar Mühendisliği Programı **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Prof.Dr. M. Elif KARSLIGİL
Yıldız Teknik Üniversitesi
Danışman

Jüri Üyeleri

Prof.Dr. M. Elif KARSLIGİL, Danışman
Yıldız Teknik Üniversitesi

Doç.Dr. Gökhan BİLGİN, Üye
Yıldız Teknik Üniversitesi

Prof.Dr. Berrin YANIKOĞLU, Üye
Sabancı Üniversitesi

Danışmanım Prof.Dr. M. Elif KARSLIGİL sorumluluğunda tarafımda hazırlanan Çekişmeli Üretici Ağ Tabanlı Veri Artırımı Yaklaşımı ile Otoskop Görüntülerinden Kulak Zarı Patolojilerinin Tespiti başlıklı çalışmada veri toplama ve veri kullanımında gerekli yasal izinleri aldığımı, diğer kaynaklardan aldığım bilgileri ana metin ve referanslarda eksiksiz gösterdiğimi, araştırma verilerine ve sonuçlarına ilişkin çarpıtma ve/veya sahtecilik yapmadığımı, çalışmam süresince bilimsel araştırma ve etik ilkelerine uygun davrandığımı beyan ederim. Beyanımın aksinin ispatı halinde her türlü yasal sonucu kabul ederim.

Mustafa Furkan ESEOĞLU

İmza

*Sevgili eşime
ve
aileme*



TEŞEKKÜR

Tez çalışmam süresince her konuda bana destek olan ve yol gösteren değerli hocam Prof. Dr. Mine Elif KARSLIGİL'e, çalıştığımız problem üzerine sağladığı imkanlar için Prof. Dr. İsmail KOÇAK'a teşekkürlerimi sunarım. Çalışma süresinde sunduğu teknik öngörü ve yaptığımız fikir alışverişleri için Berk TAŞKIN'a teşekkür ederim. Hayatım boyunca yanımda olan aileme ve arkadaşlarıma minnettarım. Son olarak sağladığı sonsuz destek ve yeri doldurulamaz moral için sevgili eşim Elif ESEOĞLU'na teşekkürü borç bilirim.

Mustafa Furkan ESEOĞLU

İÇİNDEKİLER

SİMGE LİSTESİ	vii
KISALTMA LİSTESİ	viii
ŞEKİL LİSTESİ	ix
TABLO LİSTESİ	xi
ÖZET	xii
ABSTRACT	xiv
1 GİRİŞ	1
1.1 Literatür Taraması	2
1.1.1 GAN ile Veri Artırımı	2
1.1.2 Gözetimsiz Medikal Veri Sınıflandırma	4
1.2 Tezin Amacı	6
1.3 Hipotez	6
2 METODOLOJİ	7
2.1 Yapay Veri Üretimi	8
2.1.1 Yapay Veri Üretim Modelleri	8
2.1.2 Yapay Veri Üretimi Başarısının Ölçülmesi	16
2.2 Özellik Çıkarma (Feature Extraction)	18
2.2.1 Ön Eğitimli Evrişimli Sinir Ağı ile Öznetelik Çıkarımı	19
2.2.2 Otokodlayıcı ile Öznetelik Çıkarımı	20
2.3 Anormal Verilerin Tespiti	22
2.3.1 İzolasyon Tabanlı Karar Ağacı Ormanı ile Anormal Verilerin Tespit Edilmesi	22
2.3.2 Anormal Veri Tespiti Performans Analizi	25
3 SİSTEM TASARIMI	27
3.1 Veri Kümesi	28
3.2 Veri Ön İşleme	30

3.3	Yapay Veri Üretimi	32
3.4	Model Seçimi	33
3.5	Özellik Çıkarma	34
3.6	Patolojik Kulak Zarı Görüntüsü Tespiti	36
4	DENEYSEL SONUÇLAR	37
4.1	Deney Ortamı	38
4.2	Görsel Turing Testi ile GAN Seçimi	38
4.3	Otomatize Yöntemler ile GAN Performansının Ölçülmesi	43
4.4	Öznitelik Çıkarma Modüllerinin Eğitimi	43
4.4.1	Otokodlayıcı Eğitimi	44
4.4.2	Ön Eğitimli CNN Mimarisinde Öğrenme Aktarımı	46
4.5	İzolasyon Tabanlı Karar Ağacı Ormanı ile Patolojik Görüntülerin Tespiti	51
4.5.1	Yapay Görüntülerin Patolojik Görüntü Tespiti Başarımına Etkisi	52
4.6	Deney Sonuçlarının Değerlendirilmesi	59
5	SONUÇ VE ÖNERİLER	61
	KAYNAKÇA	63
	TEZDEN ÜRETİLMİŞ YAYINLAR	66

SİMGE LİSTESİ

μ	Veri pikselleri ortalama değeri
Σ	Veri pikselleri standart sapma değeri
D	Ayırt Edici Ağ
G	Üretici Ağ

KISALTMA LİSTESİ

AdaIN	Adaptive Instance Normalization
AUROC	Area under the ROC Curve
CNN	Evrişimli Sinir Ağı
DAGAN	Data Augmentation Generative Adversarial Networks
DCGAN	Deep Convolutional Generative Adversarial Network
FID	Fréchet Inception Distance
GAN	Generative Adversarial Networks
IAGAN	Inception-Augmentation GAN
IF	Isolation Forest
IS	Inception Score
JS	Jensen-Shannon
KBB	Kulak Burun Boğaz
KL	Kullback-Leibler
LPIPS	Learned Perceptual Image Patch Similarity
MR	Manyetik Rezonans
PGGAN	Progressive Generative Adversarial Networks
ROC	Receiver Operating Characteristic Curve
StyleGAN	Style-Based Generator Architecture for Generative Adversarial Networks
VTT	Visual Turing Test
WGAN	Wasserstein Generative Adversarial Networks
WGAN-GP	Wasserstein Generative Adversarial Networks with Gradient Penalty
YSA	Yapay Sinir Ağı

ŞEKİL LİSTESİ

Şekil 2.1	Yıllar içerisinde GAN ile alınan yapay yüz üretimi çıktıları [17] . . .	9
Şekil 2.2	GAN Sistem Mimarisi	10
Şekil 2.3	Üretici (G) ve Ayırt Edici'ye (D) yüksek çözünürlüklü katmanların eklenmesi. (a) Düşük katman eğitim sonu (b) Kademeli çözünürlük artışı (c) Bir yüksek çözünürlüğe geçişin tamamlanması [20]	13
Şekil 2.4	(a) Kademeli büyüyen GAN ile Stil tabanlı Üretici'ye sahip GAN (b) ağlarının Üretici ağları [21]	15
Şekil 2.5	CNN transfer öğrenme şeması.(a) Ön Eğitimli CNN tabanlı sınıflandırıcı. (b) İlk olarak önceden eğitilmiş YSA'nın sınıflandırıcı katmanı evrişim katmanından kopartılır. (c) Yeni tamamen bağlı sınıflandırıcı katmanı sisteme eklenir.	19
Şekil 2.6	Otokodlayıcı mimarisi	21
Şekil 3.1	Anormal kulak zarı görüntüsü tespiti sistemi tasarımı	28
Şekil 3.2	Toplanan etiketlerin veri dağılımı	29
Şekil 3.3	Toplanan patolojik etiketlerin veri dağılımı	30
Şekil 3.4	Etiketlenen verilerden örnekler. (a): Normal, (b): Otitis Media, (c): Skleroz, (d): Zar Perforasyonu	30
Şekil 3.5	Kulak zarı görüntülerinin hizalanması	31
Şekil 3.6	Çalışmada kullanılan Kademeli Büyüyen GAN mimarisi ve örnek ara çıktıları	32
Şekil 3.7	Uzmanlara sunulan Görsel Turing testinin yapısı	34
Şekil 3.8	Kullanılan otokodlayıcı yapısı	35
Şekil 3.9	CNN ile öznelik çıkarımı	35
Şekil 4.1	Kulak zarı yapısının korunabildiği yapay çıktılar	39
Şekil 4.2	Kulak zarı yapısının korunamadığı yapay çıktılar	40
Şekil 4.3	Görsel Turing Testinde gerçek olarak oylanan görüntüler	41
Şekil 4.4	Görsel Turing Testinde yapay olarak oylanan görüntüler	42
Şekil 4.5	Otokodlayıcının eğitim boyunca doğrulama verisinde aldığı kayıp grafiği	44
Şekil 4.6	Otokodlayıcının eğitim sonunda aldığı görüntüyü yeniden oluşturması	45
Şekil 4.7	EfficientNet-B2 ilk eğitim kesinlik grafiği	47

Şekil 4.8	EfficientNet-B2 ilk eğitim duyarlılık grafiği	48
Şekil 4.9	EfficientNet-B2 ilk eğitim kayıp grafiği	48
Şekil 4.10	EfficientNet-B2 ikinci eğitim kesinlik grafiği	49
Şekil 4.11	EfficientNet-B2 ikinci eğitim duyarlılık grafiği	49
Şekil 4.12	EfficientNet-B2 ikinci eğitim kayıp grafiği	50
Şekil 4.13	Otokodlayıcı vektörü ile veri artırımı F1 skoru değişimi grafiği	52
Şekil 4.14	Otokodlayıcı vektörü ile veri artırımı Kesinlik-Duyarlılık değişimi grafiği	53
Şekil 4.15	EfficientNet-B2 tabanlı CNN vektörü ile veri artırımında F1 Skoru değişimi grafiği	54
Şekil 4.16	EfficientNet-B3 tabanlı CNN vektörü ile veri artırımında F1 Skoru değişimi grafiği	55
Şekil 4.17	EfficientNet-B4 tabanlı CNN vektörü ile veri artırımında F1 Skoru değişimi grafiği	55
Şekil 4.18	EfficientNet-B5 tabanlı CNN vektörü ile veri artırımında F1 Skoru değişimi grafiği	56
Şekil 4.19	EfficientNet-B6 tabanlı CNN vektörü ile veri artırımında F1 Skoru değişimi grafiği	56
Şekil 4.20	EfficientNet-B2 tabanlı CNN vektörü ile veri artırımı Kesinlik-Duyarlılık değişimi grafiği	57
Şekil 4.21	EfficientNet-B3 tabanlı CNN vektörü ile veri artırımı Kesinlik-Duyarlılık değişimi grafiği	57
Şekil 4.22	EfficientNet-B4 tabanlı CNN vektörü ile veri artırımı Kesinlik-Duyarlılık değişimi grafiği	58
Şekil 4.23	EfficientNet-B5 tabanlı CNN vektörü ile veri artırımı Kesinlik-Duyarlılık değişimi grafiği	58
Şekil 4.24	EfficientNet-B6 tabanlı CNN vektörü ile veri artırımı Kesinlik-Duyarlılık değişimi grafiği	59

TABLO LİSTESİ

Tablo 2.1	Hata Dizeyi	26
Tablo 3.1	Yapay sinir ağlarında kullanılan çözünürlükler ($n \times n$ ikilileri)	31
Tablo 3.2	Kullanılan Mimarilerin Çözünürlük Kapasitesi	33
Tablo 4.1	Çalışmada kullanılan sistemin özellikleri	38
Tablo 4.2	Görsel Turing testi sonuçları	39
Tablo 4.3	Eğitilen GAN'ların çıktı kalitesini belirten FID skorları	43
Tablo 4.4	Eğitilen GAN'ların çeşitli çıktı üretme kapasitesini belirten LPIPS skorları	43
Tablo 4.5	CNN Mimarisi	46
Tablo 4.6	Farklı EfficientNet mimarilerinden alınan öznitelik vektör boyutları	46
Tablo 4.7	EfficientNet-B2 tabanlı modelin farklı eğitim yaklaşımlarındaki test başarımları	50
Tablo 4.8	Eğitimi tamamlanan gözetimli sınıflandırıcı başarımları	51
Tablo 4.9	Farklı modellerden çıkartılan öznitelik vektörleri kullanılarak alınan patolojik görüntü tespiti başarımları	51
Tablo 4.10	Veri artırımı ile alınan en yüksek duyarlılık başarımları	60

Çekişmeli Üretici Ağ Tabanlı Veri Artırımı Yaklaşımı ile Otoskop Görüntülerinden Kulak Zarı Patolojilerinin Tespiti

Mustafa Furkan ESEOĞLU

Bilgisayar Mühendisliği Anabilim Dalı
Yüksek Lisans Tezi

Danışman: Prof.Dr. M. Elif KARSLIGİL

Medikal örüntü tanıma sistemleri aracılığıyla doktorların tanı koymasına yardımcı olacak sistemlerin geliştirilmesi son dönemde ilgi çeken bir alandır. Yapılan çalışma kapsamında ilk olarak bir otoskop cihazından alınan 10.088 adet kulak zarı görüntüsü toplanmıştır. Toplanan görüntülerden 921 adet görüntü Kulak Burun Boğaz Uzmanı tarafından etiketlenmiştir. Etiketlenen görüntüler temizlendikten sonra normal, akut otitis media, buşon, egzositoz, eksternal otit, miringoskleroz, otomikoz, retraksiyon, seröz otitis media ve kronik otitis media perforasyon sınıflarına ait toplam 276 adet görüntü test verisi olarak ayrılmıştır. Etiketlenen görüntülerde normal sınıf haricinde kalan ve patoloji içeren görsellerin oranı daha düşük olduğu için bu patolojik görüntülerin tespiti problemine anomali tespiti şeklinde yaklaşılmıştır.

Patolojik görüntülerin tespitinde kullanmak amacıyla görüntü özniteliklerinin çıkarılması için alanda uzmanlık gerektiren yaklaşımların aksine veriyi anlamlandırarak öznitelik çıkaran derin öğrenme tabanlı modeller eğitilmiştir. Bu kapsamda gözetimsiz bir otokodlayıcı ve sınırlı miktarda etiketlenen veriyi kullanan gözetimli şekilde farklı EfficientNet modelleri eğitilmiştir. En yüksek F1 skoruna sahip EfficientNet-B6 temelli modelde etiketli test verisi üzerinde patolojik sınıf için 0.755 kesinlik puanı, 0.825 duyarlılık puanı ve 0.788 F1 puanı alınmıştır. Eğitilen modellerin alt katmanlarından öznitelik vektörleri alınmıştır.

Toplanan veri kümesindeki etiketlenmiş görüntü sayısı etiketsiz verinin yaklaşık %10'u olduğu için gözetimsiz anomali tespiti yaklaşımı ile patolojik görüntülerin

sınıflandırılması yaklaşımı tercih edilmiştir. Etiketsiz veri kümesinden yararlanmak amacıyla bir izolasyon tabanlı karar ağacı ormanı (IForest) modeli üzerinde eğitim gerçekleştirilmiştir.

Sistemin başarısını ölçmek amacıyla kesinlik, duyarlılık ve F1 başarımları dikkate alınmıştır. Patolojik görüntülerin yanlışlıkla sağlıklı olarak sınıflandırılmasını engellemenin daha önemli olmasından dolayı test başarımları izlenirken duyarlılık puanına dikkat edilmiştir. Gözetimsiz olarak eğitilen IForest modelinden alınan duyarlılık puanının, öznel çıkartımı amacıyla etikete sahip görseller ile eğitilen gözetimli sınıflandırıcı modelin duyarlılık puanından daha düşük olduğu görülmüştür. Bu başarıyı artırmak adına uygulanabilecek veri artırımı yaklaşımları incelenmiştir.

Verinin şeklinin ve renk dağılımlarının önem arz ettiği medikal görüntülerde klasik görüntü işleme yöntemleri kullanılarak veri artırımı yapmanın uygun olmadığına karar verilmiştir. Bu nedenle veri artırımını sağlamak amacıyla toplanan görüntüler kullanılarak Gradyan Cezalı Wasserstein Çekkişmeli Üretici Ağları (Generative Adversarial Networks) (GAN), kademeli büyüyen GAN ve StyleGAN mimarileri ile yapay kulak zarı görüntüleri üretilmiştir. Üretilen görüntüler uzman doktorlar tarafından bir Görsel Turing Testi kullanılarak gerçek/yapay olacak şekilde oylanmıştır.

En yüksek başarıma sahip GAN mimarisi seçildikten sonra üretilen farklı miktardaki yapay görüntüler orijinal görüntülere eklenerek gözetimsiz eğitim gerçekleştirilmiştir. Veri artırımı yapılmadan gerçekleştirilen eğitime kıyasla veri artırımı yapıldığında, duyarlılık başarımının 0.674'ten 0.895'e ve F1 puanının 0.641'den 0.684'e yükseldiği ancak kesinlik puanının 0.611'den 0.554'e düştüğü görülmüştür. Kesinlik puanındaki düşüşe kıyasla, patolojik görüntülerin doğru olarak sınıflandırılmasında önem arz eden duyarlılık başarısının 0.22 artması veri artırımı yaklaşımının etkili olduğunu göstermiştir. IForest modeli, EfficientNet-B6'ya kıyasla daha fazla yanlış alarm üretmesine rağmen patolojik görüntüleri daha yüksek başarımla yakaladığı için patolojik görüntü tespitinde doğruluk ihtiyacının daha önemli olduğu durumlarda kullanılabileceği düşünülmektedir.

Anahtar Kelimeler: Çekişmeli üretici ağlar, medikal görüntü işleme, bilgisayar destekli tanı, veri artırımı, görsel turing testi, anomali tespiti

Detection of Eardrum Pathologies from Otoscope Images with Generative Adversarial Network-Based Data Augmentation Approach

Mustafa Furkan ESEOĞLU

Department of Computer Engineering
Master of Science Thesis

Supervisor: Prof.Dr. M. Elif KARSLIGİL

There is a rising interest in computer-aided diagnosis systems which use medical images to assign a diagnosis to a patient. This study proposes a methodology for detecting pathological tympanic membrane images. To achieve this goal, 10.088 tympanic membrane images were collected by a doctor using an otoscope device. 921 of the acquired image labeled by an expert physician. After removing the labels with low image counts, the dataset of 276 labels with labels of normal, acute otitis media, ear wax, exocytosis, external otitis, myringosclerоз, otomycosis, retraction, serous otitis media, and chronic otitis media have prepared as a test set. Due to the low amount of data apart from normal labels, the problem of pathology detection is considered as anomaly detection.

Contrary to the approaches that require expertise in the field, deep learning-based models that extract features by making sense of the data have been trained and used to acquire the image features. In this context, an unsupervised autoencoder and supervisedly trained different sizes of EfficientNet models, which uses a limited amount of labeled data, are trained. The supervised model achieved a 0.755 precision, 0.825 recall, and 0.788 F1 scores for the pathological class from the labeled test data. Feature vectors of the images are taken from the lower layers of trained models.

Unsupervised image classification as a pathological image recognition approach is preferred due to the low amount of labeled data in the collected dataset. To utilize the unlabeled data, Isolation Forest (IForest) model is trained on the whole dataset.

To measure the success of the system, precision, recall, and f1 scores are selected as performance measures. To prevent the misclassification of pathological images, exceptional attention has been given to recall score. To improve the performance of the unsupervised IForest model which has lower scores across all performance metrics, different dataset augmentation approaches are looked upon.

Due to the importance of shape and color distributions of medical images, classical data augmentation approaches are omitted. To increase the data count, architectures like Wasserstein Generative Adversarial Networks (GAN) with Gradient Penalty, progressive growing GAN, and StyleGAN are trained by utilizing the training data. Then, the realness of generated images is evaluated by physicians using a Visual Turing Test.

After selecting the most realistic GAN architecture, different amounts of artificial images were added to original images for unsupervised pathology detection. In contrast to training without data augmentation using artificial eardrum images, the F1 score and recall score increased from 0.641 to 0.683 and from 0.674 to 0.895. Concerning this increase, the precision score decreased from 0.611 to 0.554. Compared to the decrease in the precision score, an increase of 0.22 in the recall score showed the success of the data augmentation strategy. As a result, the IForest model has more false positives but it is better at catching pathological images which is more important in the medical domain.

Keywords: Generative adversarial networks, medical image processing, computer aided diagnosis, data augmentation, visual turing test, anomaly detection

1 GİRİŞ

Çeşitli medikal verilerden yararlanarak bilgisayar aracılığı ile teşhis yardımı son yıllarda ivme kazanan bir alandır.[1][2] Farklı makine öğrenmesi yöntemleri ile deri hastalıkları teşhisi, farklı dokularda kanser teşhisi, diyabet teşhisi gibi alanlarda doktorların karar almasına yardımcı olacak sistemler üzerine çalışılmaktadır. Bu tür sistemler alanında uzman doktorlara erişimin kısıtlı olduğu yerlerde çalışan sağlık görevlilerinin gördükleri vakalarda karar almalarını kolaylaştırmaktadır.

Çocuklarda sıklıkla görülen Akut Otitis Media (AOM) çocuk doktorlarına yapılan ziyaretlerin en yaygın sebeplerindendir [3]. Sosyo-ekonomik açıdan halka yük bindiren bu durum potansiyel olarak koyulabilecek yanlış teşhisler sebebi ile daha vahim bir durum olmaya adaydır. Otoskopi muayenesi ile teşhisi yapılabilen bu hastalıkta eksik teşhis yapılması durumunda bireyin yaşamında sağlığa varabilen sıkıntılara yol açabilecek rahatsızlıklar ortaya çıkabileceği göz önünde bulundurulmalıdır.

Kulak ile ilgili hastalıkların teşhisinde iyi bir otoskopi muayenesi gereklidir. Kulak zarının orta kulağa açılan bir pencere olması avantajından ötürü çoğu zaman dikkatli bir otoskopi işlemi ile başarılı teşhisler koyulabilmektedir. Kulak burun boğaz uzmanları, çocuk hastalıkları uzmanları ve hatta pratisyen hekimlerin iyi bir şekilde otoskop cihazını kullanabilmeleri gerekmektedir.

Teşhise yardımcı olacak sistemler eğitimleri aşamasında yüksek miktarda veriye ihtiyaç duymaktadır. Literatürde yaygın olarak kullanılmaya başlayan derin öğrenme tabanlı yöntemler ile bu ihtiyaç daha baskın hale gelmiştir. Bu ihtiyacı karşılamak adına yapay veri üreten sistemler ile medikal görüntüler üretilip tanı sistemlerine ek veri olarak beslenmesi uygulanan bir yaklaşımdır.

Bu çalışmada çeşitli görüntü işleme yöntemleri ile patolojik kulak zarı görüntülerinin tespitini yapan bir sistem geliştirilmiştir. Normal kulak zarının dışında kalan ve patoloji içeren görüntüler anormal olarak kabul edilerek anomali tespiti

yaklaşımlarından yararlanılmıştır. Kullanılacak veri kümesindeki sınırlı etiket sayısı nedeniyle teşhis duyarlılık başarımını artırmaya yönelik yaklaşımlar incelenerek seçilen yöntem veri artırımı aracı olarak kullanılmıştır. Rastgele uzaydan medikal görüntü üreten yöntemler ve medikal görüntülerde gözetimsiz eğitilen tanı sistemlerinin literatür çalışmaları incelenmiştir. Patolojik görüntülerin otomatik olarak tespit edilebilmesi için önerilen ve gerçekleşen sistemin detayları açıklanmıştır. Çalışma boyunca kullanılan veri kümesi ile sistem performansı ölçülerek değerlendirilmiştir.

1.1 Literatür Taraması

Son yıllarda derin öğrenme alanında elde edilen gelişmelerle birlikte derin öğrenme tabanlı yaklaşımlar kullanılarak bilgisayarlı görü, görüntü tanıma, ses tanıma, doğal dil işleme ve öneri sistemleri gibi bir çok alanda çeşitli problemler ele alınmaktadır [4]. Farklı görüntüleme cihazlarından alınan medikal görüntülerde bölütleme, sınıflandırma gibi analizlerin yapılabilmesini amaçlayan sistemler de tasarlanmıştır [5].

Derin öğrenme tabanlı sistemler eğitildikleri görevde başarılı öğrenim yapabilmek için yüksek miktarlarda veriye ihtiyaç duymaktadır. Kaliteli ve yüksek miktarlarda veri elde etmenin zor olduğu medikal görüntülerde veri artırımı yöntemi olarak kullanılabilen Çekişmeli Üreci Ağlar (Generative Adversarial Networks - GAN) tabanlı yaklaşımlar incelenmiştir. Buna ek olarak medikal veride patolojik görüntülerin tespiti için uygulanan gözetimsiz sınıflandırma yaklaşımları incelenmiştir. İncelenen yaklaşımlarda patolojik görüntüler anomali olarak değerlendirilerek çözümler bu kapsamda gerçekleştirilmiştir.

1.1.1 GAN ile Veri Artırımı

Han vd. [6], DCGAN ve WGAN modellerini kullanarak yapay manyetik rezonans (MR) görüntüleri üretmişlerdir. Yapay görüntülerin üretilmesi için bu iki mimariye farklı doku biçimlerine sahip beyin görüntüleri verilmiştir. Üreticilerden alınan görüntülerin gerçekliğini ölçmek adına bir uzmana üretilen görüntüler gerçek görüntüler ile karıştırılarak bunların uzman tarafından gerçek veya yapay olarak sınıflandırılması istenmiştir. Uzmanın başarısı doğruluk puanı olarak ölçülmüştür. Gerçek ve yapay görüntülerin karıştırılarak uzmanın karşısına çıkartılmasından dolayı Üretici ağı çıktısı kalitesi arttıkça uzmanın bu iki tip görüntüyü ayırt etmesinin gittikçe zorlaştığı görülmüştür. WGAN ile üretilen görüntülerin en iyi durumda uzmanın tespit başarısını %53'e kadar düşürdüğü raporlanmıştır.

Salehinejad vd. [7], göğüs X-ray görüntülerinin sınıflandırılmasında az miktarda görüntü bulunan sınıflardaki verinin artırılması üzerine çalışmışlardır. Yapılan çalışmada kullanılan veri kümesinde yer alan 4 farklı patoloji için ayrı DCGAN modelleri eğitilmiştir. Eğitilen modellerden üretilen yapay görüntüler farklı sınıfların ham verisine eklenerek sınıflandırma başarımına katkısı F1 puanı ile ölçülmüştür. Veri artırımı olarak veri kümesindeki her sınıfa ait olan görüntü sayısı kadar yapay görüntü üretimi yapılmıştır. Veri kümesi dengeli olarak ayarlandığında f1 puanının AlexNet mimarisi için 0.37'den 0.71'e, GoogLeNet için 0.56'dan 0.69'a ve ResNet-50 için 0.44'ten 0.78'e çıktığını gözlemlenmiştir. Aynı veri artırımı yaklaşımını veri kümesinin dengesiz olması durumu için de uygulamış ve başarımların AlexNet mimarisi için 0.65'ten 0.82'ye, GoogLeNet için 0.61'den 0.80'e ve ResNet-50 için 0.74'ten 0.93'e çıktığını gözlemlenmiştir. Başarımdaki keskin artışın sebebi modellere daha çok veri beslendiğinde sistemin veri kümesinin dağılımını daha iyi öğrenebilmesi olarak değerlendirilmiştir.

Chen vd. [8] göz arkası görüntüleri üzerinde diyabetik retinopati patolojisinin tespit başarısının artırılması için bir veri artırımı yaklaşımı önermiştir. Veri artırımı yöntemi olarak hem çevirme, döndürme ve ölçekleme gibi klasik görüntü işleme tabanlı veri artırımı yöntemleri hem de patolojik sınıfa ait veri ile eğitilmiş DCGAN ağından üretilen görüntüler kullanılmıştır. İki farklı yaklaşımdaki artan verinin başarıma etkisini incelemek amacıyla yapılan deneylerde veri kümesine eklenen yapay görüntülerin üzerine klasik veri artırımı yöntemleri uygulanmıştır. Sınıflandırma için bir evrişimli yapay sinir ağından yararlanmıştır. Ham veri kümesinde diyabetik retinopati sınıflandırma başarısı %78.33 iken yapay görüntü üretimi ve klasik görüntü artırımı birleştirilip ham veri kümesine eklendiğinde diyabetik retinopati sınıflandırma başarımı %90.56'ya artmıştır.

Qin vd. [9] bir dermoskopi cihazından alınan cilt görüntülerinde patoloji sınıflandırmada yapay olarak üretilen görüntülerin sistem başarısına etkisini incelemişlerdir. Veri artırımı için StyleGAN tabanlı bir stil mimarisinin katman yapısı değiştirilerek cilt görüntüleri üretilmiştir. Yapay görüntülerin kalitesi diğer GAN mimarileri ile üretilen görüntülerle kıyaslanmıştır ve tasarlanan yaklaşımın en gerçekçi ve en çok çeşitliliğe sahip olduğu gösterilmiştir. Patolojilerin sınıflandırılması için ImageNet veri kümesinde eğitilmiş bir ResNet-50 evrişimli yapay sinir ağında transfer öğrenme uygulanmıştır. Yapılan eğitim sonucunda sınıflandırma doğruluk puanı %93.6'dan %95.2'ye artmıştır. Gerçekten patolojik olan görüntülerin başarılı olarak tespit edilip edilemediğini ölçtükleri hassasiyet puanı ise %54.4'ten %74.3'e yükselmiştir. Gözlemlenen artıştan yola çıkılarak eğitilen yapay sinir ağının bir öznetelik çıkarım yöntemi olarak da kullanılabileceğini göstermişlerdir.

Motamed vd. [10] göğüs X-ray görüntülerinden COVID-19'a sahip görüntülerin tespiti üzerine GAN tabanlı bir veri artırımı yöntemi önermişlerdir. Yapay görüntü üretimi için DAGAN tabanlı bir IAGAN mimarisi önerilmiştir. Dikkat (attention) tabanlı bu mimari ile iki farklı veri kümesi üzerinde deneyler gerçekleştirilmiştir. İlk veri kümesindeki zatürre görüntüleri ile bir adet, ikinci veri kümesindeki normal ve zatürre görüntüleri ile birer olmak üzere toplam üç adet GAN modeli eğitilmiştir. İkinci veri kümesinde bulunan COVID-19 içeren görüntüler sadece test amacıyla kullanılmıştır. Ek olarak yapay görüntü ile uygulanan veri artırımı yöntemi döndürme, yaklaştırma, kaydırma gibi klasik veri artırımı yöntemleri ile kıyaslanmıştır. Patolojilerin tespiti için kullanılan AnoGAN modeli ilk veri seti için sadece zatürre görüntülerinde, ikinci veri setinde aynı yapıdaki iki modelin birisi normal diğeri de zatürre görüntüleri üzerinde eğitilmiştir. Sistemlerden alınan anomali skorlarına (patoloji tespiti başarımı) göre son sınıflandırma kararı verilmiştir. Sistemlerin Eğri Altındaki Alan (AUC) başarımları ilk veri kümesi için 0.87'den 0.90'a; ikinci veri kümesinde 0.74'ten 0.76'ya yükselmiştir. İkinci veri kümesindeki başarımın ve veri artırımı yönteminin başarıdaki etkisinin daha düşük olmasının sebebi olarak görüntülerin farklı kaynaklardan toplanması ve dolayısıyla düzgün bir dağılıma sahip olamaması gösterilmiştir.

1.1.2 Gözetimsiz Medikal Veri Sınıflandırma

Lim vd. [11] farklı öznelik miktarlarına sahip çeşitli medikal alanlardaki patolojik örneklerin tespiti için bir veri artırımı yaklaşımı önermiştir. Patolojik örneklerin gözetimsiz olarak tespit edilmesi için ise izolasyon tabanlı karar ağacı ormanı kullanmışlardır. Verinin karakteristiğini anlamlandırmak amacıyla bir çekişmeli otokodlayıcı eğitimi yapılmıştır. Eğitilen otokodlayıcı ile veriyi içeren gizli vektör uzayı oluşturulmuştur. Oluşturulan uzaydaki normal dağılımın kenarlarında kalan gözlemler anormal veriler olarak değerlendirilmiştir ve patoloji tespiti sistemine ek örnek olarak verilmiştir. Veri artırımı sağlandıktan sonra izolasyon tabanlı karar ağacı ormanı ile patoloji tespiti işlemi yapılmıştır. Gözetimsiz olarak yapılan eğitim ile etiketli veriye olan ihtiyacın kalktığı görülmüştür. Geliştirilen sistem ile farklı tip verilerde yapılan deneylerden sonra SMOTE, INOS gibi geleneksel veri artırımı yöntemlerine kıyasla ortalama %3 iyileşme sağlandığı görülmüştür.

Ouardini vd. [12] büyük veri kümelerinde eğitilmiş evrişimli yapay sinir ağından alınan öznelik vektörleri ile gözetimsiz patoloji tespiti modellerinin eğitimi üzerine çalışmışlardır. Yapay sinir ağından alınan öznelik vektörleri gözetimsiz olarak eğitilen izolasyon tabanlı karar ağacı ormanına verilerek patoloji tespiti yapan model eğitilmiştir. Deneysel çalışmalar hem doğal görüntüler hem de medikal görüntüler üzerinde koşulmuştur. Yapılan deneyler sonucunda rastgele karar verme

yetisinden öteye geçemeyen diğer yöntemlerin aksine %20 civarında alıcı işletim karakteristiğinin altındaki alan (area under the receiver operating characteristic) (AUROC) performansı iyileşmesi gözlemlenmiştir. Patoloji modeli eğitilirken veri kümesinin dağılımındaki normal öznitelik vektörleri ile eğitim sağlanmıştır. Buna ek olarak eğitim verisinin içerisine etiketi bilinen anormal gözlemler de eklendiğinde performansın yükseldiği gözlemlenmiştir. Gözetimsiz sistemin performansı iki farklı retina veri kümesinde sırasıyla %77 ve %74.5 olarak alınmıştır. Sistem performansını daha iyi gözlemlemek amacıyla gözetimli olarak eğitilen bir evrişimli yapay sinir ağı ile de eğitim yapılmıştır ve iki farklı veri kümesinde sırasıyla %97.3 ve %94.5 AUROC başarılarının alındığı görülmüştür. Yapılan deneyler sonucunda veride etiket olmasa dahi alınan başarımın ilk aşamada tatmin edici olduğu görülmüştür. Gözetimsiz yöntemdeki performansın gözetimli yöntem ile alınan performansa yaklaşabilmesi için öznitelik vektörlerinin çıkarılması aşamasında veri kümesine daha iyi adapte olabilecek bir yaklaşım kullanılabilir.

Li vd. [13] dağılım dışı (out-of-distribution) medikal görüntülerin gözetimsiz olarak tespit edilebilmesi amacıyla bir sistem geliştirmişlerdir. 8 farklı sınıfa sahip deri lezyon görüntüsünden her aşamada bir sınıf anormal olacak şekilde ayrılıp iki sınıflı evrişimli yapay sinir ağı eğitilmiştir. Her sınıf için eğitilen ağların ara katmanlarından alınan öznitelik vektörleri ile her sınıfa özel izolasyon tabanlı karar ağacı ormanı modeli eğitilmiştir. Sisteme verilen bir görüntünün ilgili vektörleri karar ormanına verildiği zaman en yüksek anormallik skoru veren model anormal sınıfı belirtmektedir. Sistemdeki ortama doğruluk puanı başarısının güven skoru (confidence score) tabanlı bir yöntemle kıyasla %4 daha fazla olduğu görülmüştür. Her sınıf için ayrı modeller eğitilmesi çok fazla patolojik sınıfın olduğu durumda yüksek bir maliyete yol açacak olsa da, patoloji sayısının sınırlı olduğu veri kümeleri için uygulanabilir bir yaklaşım olarak düşünülebilir.

Shvetsova vd. [14] medikal görüntülerde gözetimsiz patoloji tespiti tanımını biraz daha esnekleştirerek anormal görüntülerin sisteme hiç verilmemesi yerine sınırlı sayıda anormal görüntü (hastalık içeren görüntüler) kullanımı yaklaşımını benimsemişlerdir. Yapılan deneyler göğüs X-rayleri ve boyalı histolojik görüntüler üzerinde koşulmuştur. Çözünürlüğü kademeli olarak artırılan bir otokodlayıcı mimarisi ile eğitim işlemi gerçekleştirilmiştir. Normal otokodlayıcı eğitiminden ayrı olarak kayıp metriği için girdi ve çıktı görüntülerinin büyük veri kümelerinde bir evrişimli yapay sinir ağının ara katmanından alınan öznitelik vektörleri arasındaki ayrışma hatası kullanılmıştır. Otokodlayıcının kademeli olarak büyütülmesi girdi görüntüsünün daha yüksek kalitede oluşturulabilmesini sağlamıştır. Eğitim aşamasında sisteme sınırlı sayıda anormal görüntü beslemenin otokodlayıcının normal görüntüleri daha iyi modellemesini sağladığı düşünülmüştür. Test aşamasında

otokodlayıcıdan alınan yeniden oluşturma kayıp değeri patoloji skoru olarak kullanılmıştır. Geliştirilen sistem sonuç olarak histolojik görüntülerdeki en başarılı modele göre (90.6), 94.3 ROC AUC skoru alarak daha yüksek başarımları göstermiştir. Aynı şekilde X-ray görüntülerinde de 87.3 ROC AUC skoruna kıyasla 92.6 skoru ile başarımları artışı sağlanmıştır. Çalışmada temel karşılaştırma metodu olarak kullanılan yöntemlerde de eğitim aşamasında sınırlı sayıda anormal görüntü kullanılmasının başarımları artırdığı görülmüştür. Sadece normal etiketli veriler ile yapılan eğitim aşamasına az miktar da olsa (çalışmada toplam eğitim verisinin %0.5'i) anormal veri ekleyerek son başarımları artırmak etkili bir yaklaşım olarak düşünülebilir.

1.2 Tezin Amacı

Tezin amacı kişilerden alınan kulak zarı görüntülerinden patolojik olanların tespit edilmesidir. Uzman doktorlara erişimin kısıtlı olduğu yerlerde alınan kulak zarı görüntüleri ile patolojik kulak zarlarının tespiti kolaylaşacaktır. Bu sayede doktorların muayene sırasında teşhis koymalarına yardımcı olabilecek bir sistem ortaya çıkacaktır.

Yapılan araştırmaya göre tezin hazırlandığı tarihte literatürde kulak zarı görüntüsü üreten sistem yoktur. Bu kapsamda gerçekliği uzman doktorlar tarafından onaylanmış görüntüler üretilerek eğitim gören doktorların daha fazla veriye erişimi sağlanacaktır. Buna ek olarak yapay görüntüler gözetimsiz eğitilen tanı sistemlerinde veri artırımı yöntemi olarak kullanılabilecektir.

1.3 Hipotez

Patolojik kulak zarı tespitinde normalin dışında olan kulak zarı görüntüleri tespit edilmektedir. Yapılan tespit doktorlara tanı aşamasında yardımcı olmaktadır. Bu tez çalışması kapsamında önerilen sistemde sınırlı sayıda etiketli veri ile eğitilen gözetimli modellerin görüntülerden öznelik çıkarımı yapabileceği ve etiketsiz veriler için kullanılarak gözetimsiz modellerde kullanılabileceği öngörülmektedir.

Medikal görüntülere erişimin kısıtlı ve zor olması sebebiyle yüksek miktarda veriye ihtiyaç duyan sistemler için veri artırımı mekanizmasına ihtiyaç bulunmaktadır. Bu kapsamda Çekişmeli Üretici Ağlar (GAN) ile üretilen ve uzmanlar tarafından bir sınırın üzerinde gerçekçi olarak kabul edilen görüntülerin veri artırımı yöntemi olarak kullanılabileceği düşünülmektedir. Bu çalışmada farklı GAN mimarileri ile yapay görüntüler üretilerek yüksek miktarda veriye ihtiyaç duyan sistemlerde patolojik görüntülerin tespit duyarlılığı başarımlarının artırılması hedeflenmiştir.

2 METODOLOJİ

Bu çalışmada gözetimsiz eğitilen ve patolojik kulak zarı görüntülerinin tespiti yapan bir sistem tasarlanmıştır. Buna ek olarak sistemin eğitiminde orijinal veriye eklenen yapay kulak zarı görüntülerinin başarıya etkisi incelenmiştir. Klasik görüntü işleme yöntemleri ve makine öğrenmesi algoritmalarıyla sınıflandırma yapılarak alınacak başarımlar son zamanlarda kullanımı yaygınlaşan derin öğrenme tabanlı yöntemlerin gerisinde kalmaktadır. Ancak derin öğrenme yöntemlerinin de başarılı bir şekilde eğitilebilmesi için çok miktarda veriye ihtiyaç vardır. Bu ihtiyaç doğrultusunda modelin oluşturulması için kullanılan veri kümesinde görüntü sayısının artırılması gerekmektedir.

Eğitim kümesindeki görüntü sayısını artırmak için mevcut görüntülerde döndürme, kesme gibi klasik görüntü işleme yöntemleri uygulanabilir. Ancak görüntünün oryantasyonunun ve renk dağılımının standart olduğu medikal görüntülerde bu tür yaklaşımları kullanmak yerine veri kümesinde olmayan ancak aynı karakteristiği gösteren yapay olarak üretilmiş görüntülerin kullanılmasının başarıya etkisi değerlendirilmiştir. Yapay veri üretimi için, eğitildikleri verinin karakteristiğini başarılı olarak yakalama kapasitesine sahip Çekişmeli Üretici Ağlar (Generative Adversarial Networks - GAN) yöntemi kullanılmıştır.

Sistemlerin eğitildikleri veri miktarına göre aldıkları başarımları ölçmek amacıyla, eğitim yapılan veri kümesine farklı miktarlarda eklenen yapay görüntüler ile eğitim yapıldığında sistemin performansındaki değişimler gözlemlenmiştir. Veri artırımı için yapay görüntü üreten farklı GAN mimarileri kullanılmış ve bu mimarilerden alınan çıktılar içerisinde uzman doktorlar tarafından seçilen en gerçekçi görüntüleri üreten mimarinin kullanılmasına karar verilmiştir. Patolojik görüntülerin başarılı olarak sınıflandırılabilmesi için eğitim yapılan veri kümesinin sınıf bilgisine ihtiyaç duyulmaktadır. Kullanılan kulak zarı görüntüleri veri kümesinin toplam gözlem sayısına kıyasla daha az etiket bilgisi bulunmaktadır. Bu kapsamda bütün veriden yararlanmak adına gözetimsiz anomali tespiti modelleri tercih edilmiştir. Bu bağlamda hastalık içeren patoloji görüntüleri anomaliler olarak değerlendirilmiştir.

Bu bölümde gerçekleştirilen tez için tasarlanan sistemin alt parçaları olan yapay veri üretiminde kullanılan yaklaşımlar, veriden otomatik özellik çıkartılması ve bu özellikleri kullanarak anormal verilerin tespitini yapan modül incelenmektedir.

2.1 Yapay Veri Üretimi

Derin öğrenme sistemlerinin tasarlandıkları görevlerdeki başarılarının artırılabilmesi için eğitim aşamasında yüksek miktarda veriye ihtiyaç duyulmaktadır. Aşağıdaki klasik görüntü işleme yöntemleri eğitim verisine uygulanarak veri artırımı yapılabilir [15].

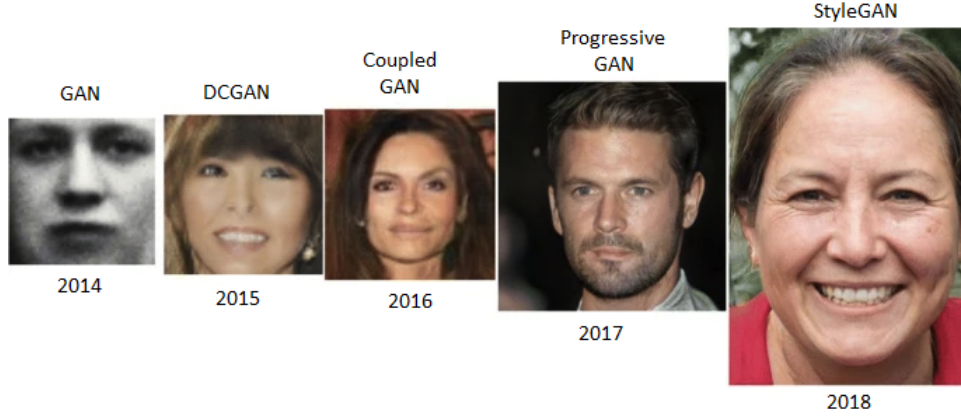
- Döndürme, çevirme gibi görüntü üzerinde yapılan geometrik dönüşümler
- Görüntünün alt renk uzayında yapılan dönüşümler
- Görüntünün farklı boyutlarda ve lokasyonlarda kırılması
- Veri kümesindeki görüntülerin farklı parçalarının birleştirilerek yeni görüntü oluşturulması
- Görüntü içerisinde rastgele bölgelerin silinmesi

Medikal görüntülerde renk ve lokasyonların standart olması, bilgi kaybı olmaması önem arz etmektedir. Bundan dolayı yukarıda belirtilen klasik görüntü işleme yöntemleri verinin yapısını bozacağından görüntülere tanı koyulması zorlaşmaktadır. Görüntünün anlamsal bütünlüğünü bozmadan sisteme yeni görüntüler eklemek amacıyla yapay veri üreten sistemlerin kullanılmasına karar verilmiştir.

2.1.1 Yapay Veri Üretim Modelleri

Goodfellow vd. tarafından sunulmasından itibaren [16] Çekişmeli Üretici Ağların çıktı kalitesi günümüze kadar büyük ilerlemeler kaydetmiştir. GAN'lar ilk ortaya çıktıklarında görüntü işleme alanında kullanılmaktayken günümüzde doğal dil işleme, ses işleme, video işleme alanlarında da kullanılmaktadırlar.

Örnek yüz üretim tekniklerinin yıllara göre gelişimi Şekil 2.1'de görülebilir. Çıktıların gerçek resimlerden ayırt edilmesinin gittikçe zorlaştığı GAN mimarileri son yıllarda makine öğrenmesi temelli sistemlerin eğitiminde veri artırma yöntemi olarak kullanılmaktadır [11]. Buna ek olarak veriye erişimin kısıtlı olduğu medikal alanlarda benzer yapıda üretilen yapay görüntüler yeni nesil doktorların eğitimi aşamasında da kullanılabilir.



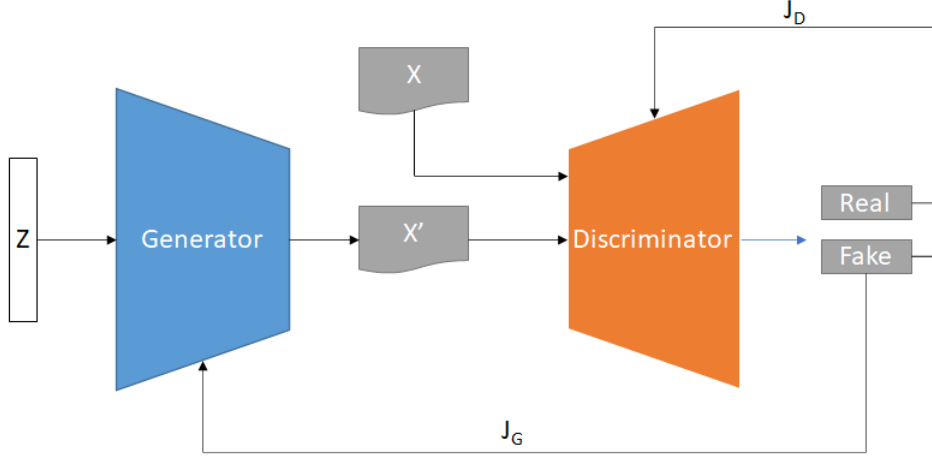
Şekil 2.1 Yıllar içerisinde GAN ile alınan yapay yüz üretimi çıktıları [17]

Yapay olarak üretilen resimlerin kalitesinin ölçülmesi bu görüntülerin kullanılacağı alanlara uygunluğunun belirlenmesi için önem arz eden bir adımdır. Bu aşamada hem otomatik sistemler hem de insanlar tarafından çıktılar kontrol edilebilmektedir. Bu tez çalışmasında kullanılan kulak zarı görüntülerinin gerçekçiliğinin ölçülebilmesi için üretilen görüntülerin alanında uzman kişiler tarafından kontrol edilmesi gerekmektedir. Gerçekçiliği düşük görüntülerin sistemlerin eğitimi aşamasında kullanılması sistemlerin tanı başarısını düşürecek veya yanlış tanı konulmasına sebep olacaktır.

2.1.1.1 Çekişmeli Üretici Ağlar (Generative Adversarial Networks)

2014 yılında Goodfellow vd. kullanılan eğitim veri kümesine benzer yapay çıktılar üretmeyi sağlayan çekişmeli üretici ağları gerçekleştirmişlerdir [16]. Esasen birbiri ile yarışan iki yapay sinir ağından(YSA) oluşan bu mimarilerde Üretici (Generator) YSA ile yapay görüntüler üretilmekte ve üretilen yapay görüntüleri ayırt etmek için bir Ayırt Edici (Discriminator) YSA'dan yararlanılmaktadır. Eğitimin sonucunda sistemde kullanılan veri kümesinin bilgi dağılımının öğrenilmesi amaçlanmaktadır. Yapay veri üretmek için öğrenilmiş olan veri dağılımından yeni bir nokta örneklenerek Üretici yardımıyla üretim yapılır. Örnek bir GAN mimarisi Şekil 2.2'de görülebilir.

Sistemin eğitimi sırasında rastgele gürültü uzayından alınan bir vektör Üretici'ye verilerek yapay bir çıktı oluşturulur. Yapay çıktı ve eğitim için kullanılan gerçek veri kümesinden alınan örnekler Ayırt Edici'ye verilerek bu ağın aldığı girdiyi gerçek veya yapay olarak sınıflandırması beklenir. Alınan sonuca göre her iki YSA'nın ağırlıkları güncellenir. Bu şekilde aynı anda eğitilen sistemin amacı temelleri oyun teorisine dayanan Nash Dengesinin sağlanmasıdır. Bu minimax oyunu Eşitlik 2.1'de verilmiştir. Denklemdaki z genellikle tekdüze (uniform) dağılımdan (p_z) örneklenmiş bir gizli



Şekil 2.2 GAN Sistem Mimarisi

değişkendir. Gerçek verinin ait olduğu dağılım p_{data} olarak ifade edilmektedir.

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (2.1)$$

Sistemde Ayırt Edici, Üretici'den aldığı girdileri en yüksek başarı ile yapay olarak sınıflandırmaya ve eğitim veri kümesinden aldığı girdileri de en yüksek başarı ile gerçek olarak sınıflandırmaya çalışır. Denklemden verilen kayıp fonksiyonundaki $D(x)$ 'i minimize etmeye çalışarak gerçek girdilerin yapay olarak sınıflandırılmasını engellemeye ve $D(G(z))$ 'yi maksimize ederek yapay girdilerin gerçek olarak sınıflandırılmasını engellemeye çalışır.

Aynı zamanda Üretici de ürettiği çıktıların Ayırt Edici tarafından gerçek olarak algılanmasını sağlamaya çalışır. Denklemden $D(G(z))$ 'yi minimize etmeye çalışarak Ayırt Edici'nin yapay girdileri tespit etmesini zorlaştırmaya çalışır. Eğitim Nash Dengesi sağlanana kadar devam eder. Denge sağlandığında Üretici gerçek veriden ayırt edilemeyen çıktılar üretir. Ayırt Edici de girdileri artık ayırt edemediği için aldığı girdilere en iyi durumda %50 olasılıkla gerçek veya %50 olasılıkla yapay diyebilir durumdadır. Her iki ağ arasındaki dengenin sağlanabilmesi eğitim sonucunda kaliteli çıktılar alınması için önemlidir. Parametreleri iyi bir şekilde ayarlanmış klasik GAN ile kaliteli çıktılar alınması mümkündür. Ancak Üretici'nin mod çöküşü'ne (mode collapse) düşme riski yüksektir. Modelde mod çöküşü meydana geldiği zaman Üretici'den alınan çıktılar Ayırt Edici'nin gerçekten ayırt edemeyeceği kalitede olacak fakat Üretici model sürekli benzer çıktıları vermeye başlayacaktır. Bu durumun aşılabilmesi için bir sonraki bölümde değinilecek olan yapı önerilmiştir.

2.1.1.2 Wasserstein GAN

2017 yılında Martin v.d. tarafından önerilen Wasserstein GAN (WGAN) [18] yöntemindeki amaç klasik GAN'da kullanılan hata fonksiyonunu değiştirerek GAN'daki eğitim dengesini daha iyi sağlamaktır. Sistem daha dengeli hale getirilerek çıktıların kalitesi ve sistemin eğitim aşamasındaki kararlılığı yükseltilmiş olur. Klasik GAN'da kullanılan Jensen-Shannon uzaklık fonksiyonu yerine Wasserstein uzaklık fonksiyonu kullanılmaktadır.

Yapay veri ile gerçek veri dağılımları arasındaki benzerliği ölçmeye yarayan Jensen-Shannon (JS) ıraksamasını (divergence) hesaplayabilmek için ilk olarak Eşitlik 2.2'de verilen Kullback-Leibler (KL) ıraksaması hesaplanır. Sürekli değişken olan gerçek veri dağılımının (p_r) yapay veri dağılımına (p_g) oranının logaritmasının her noktadaki toplamaları alınarak hesaplama yapılır.

$$D_{KL}(p_r \| p_g) = \sum_{x \in X} p_r(x) \log \frac{p_r(x)}{p_g(x)} \quad (2.2)$$

KL ıraksaması kullanılarak hesaplanan JS ıraksaması gerçek ve yapay veri dağılımlarının ortalamasını içermektedir ve KL ıraksamasının aksine simetriktir. KL ıraksaması fonksiyonunda p_r değerinin sıfıra çok yakın olduğu ancak p_g değerinin sıfırdan çok uzak olduğu durumlarda p_g değeri göz ardı edilmektedir. Her iki dağılım değerinin de önemli olduğu GAN eğitimi aşamasında dengeli bir yaklaşımı sağlayabilmek amacıyla JS ıraksaması fonksiyonu kullanılmaktadır. Eşitlik 2.3'te bu kullanılan fonksiyon verilmiştir. Eşitlikte p_r ve p_g olasılık değerlerinin ortalamaları alınarak KL ıraksaması hesaplanır.

$$D_{JS}(p_r \| p_g) = \frac{1}{2} D_{KL}(p_r \| \frac{p_r + p_g}{2}) + \frac{1}{2} D_{KL}(p_g \| \frac{p_r + p_g}{2}) \quad (2.3)$$

WGAN çalışmasında gerçek ve yapay veri dağılımları arasındaki benzerliği Wasserstein mesafesi ile hesaplamak için Eşitlik 2.4'ten yararlanılmaktadır. Burada $\Pi(p_r, p_g)$ ifadesi bütün p_r ve p_g elemanları arasındaki birleşik olasılıksal dağılımı (joint probability distribution) ifade etmektedir.

Eşitlikteki γ ifadesi p_g dağılımının p_r dağılımına benzemesi için taşınması gereken kütle miktarını ifade etmektedir. En az miktarda yük taşıyarak iki dağılım arasındaki benzerliği sağlayacak olan taşıma miktarı bulunduğunda ($\mathbb{E}_{(x,y) \sim \gamma}$) Wasserstein mesafesi hesaplanmış olur.

$$W(p_r, p_g) = \inf_{\gamma \sim \Pi(p_r, p_g)} \mathbb{E}_{(x,y) \sim \gamma} [\|x - y\|] \quad (2.4)$$

WGAN çalışmasında Wasserstein mesafesi maliyet fonksiyonu olarak kullanıldığında akıcı (smooth) olan gradyanların daha yüksek başarımlar sağlayacağı söylenmektedir. Matematiksel olarak eğitimde daha iyi bir yaklaşım olan bu kayıp fonksiyonunu hesaplamak maliyetli olduğundan dolayı Wasserstein mesafesi hesaplanırken formüle 1-Lipschitz sınırlaması (constraint) uygulanır.

WGAN eğitiminde bu sınırlamanın uygulanabilmesi için GAN'daki ağırlıkların hesaplanması aşamasında c üst parametresi kullanılır. Bu sınırlama ile model ağırlıkları $[-0.1, 0.1]$ benzeri bir dar aralığa sıkıştırılmış olur. Böylece modelin parametrelerinin aralıkları çok dağılmadan sınırlı bir uzaya yerleştirilmiş olur.

Bu şekilde eğitimin yapılması model çıktı kalitesinin artmasını sağlasa da c üst parametresinin seçimi başarılı eğitim için çok önemlidir. Yanlış parametre seçimi sonucunda model ağırlık gradyanları kaybolabilir (vanishing gradient) veya kontrolden çıkabilir (exploding gradient). Model çıktısını çok etkilemesinden dolayı Guljarani vd. Gradyan Cezalı Wasserstein GAN (WGAN-GP) [19] çalışmasında WGAN'daki sınırlama üst parametresi yerine maliyet fonksiyonuna bir cezalandırma terimi ekleyerek modelin gradyanlarını kendisinin kontrol etmesini sağlamışlardır. Bu yaklaşımla model eğitilirken alınan ağırlık gradyanlarının norm değerinin 1'e olabildiğince yakın olması zorlanmış olur.

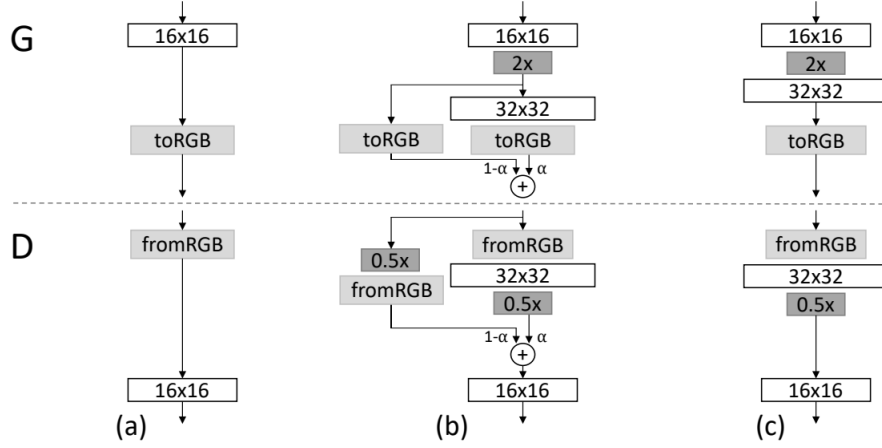
Son olarak WGAN-GP'de Ayırt Edici, Üretici'den daha fazla oranda eğitilmektedir. Bunun sebebi Ayırt Edici'ye ilk olarak gerçek verinin nasıl olduğunun öğretilmesidir. Ayırt Edici gerçek veriyi anlamlandırabilirse Üretici'yi daha doğru bir şekilde yönlendirebilecektir. Ayırt Edici'nin her 5 ağırlık güncellemesinde Üretici'nin ağırlıklarının 1 kere güncellenmesi yaygın olarak kullanılan stratejilerden birisidir.

Yüksek kalitede çıktı alınması ve eğitimin kararlılığının daha iyi olması WGAN ve WGAN-GP modellerinin yapay veri üretimi için yapılan çalışmalarda tercih edilen yöntemlerden birisi olmasını sağlamıştır.

2.1.1.3 Kademeli Büyüyen GAN (Progressive GAN)

Klasik GAN yöntemleri kullanarak 1024x1024 benzeri yüksek çözünürlükte görüntü üretebilmek zordur. Bunun sebebi eğitim zamanında Üretici'nin rastgele gürültü vektöründen başlayıp büyük bir çıktı alabilmesi için ilerlemesi gereken kayıp eğrisinin model için çok karmaşık olmasıdır. Daha yüksek çözünürlük ile üretim yapabilmek için

Karras vd. kademeli büyüyen GAN (Progressive Growing GAN)(PGGAN) mimarisini önermişlerdir[20]. Bu mimaride hedeflenen çözünürlükle eğitime başlamak yerine 4x4 gibi düşük çözünürlük ile görüntü üretimine başlayıp 8x8, 16x16 şeklinde 1024x1024'e kadar artan yüksek çözünürlüklü katmanlar sisteme kademeli olarak eklenmektedir. Düşük çözünürlükten yüksek çözünürlüğe geçiş aşamasında direkt olarak yüksek çözünürlüklü katmanın sisteme katılması yerine Şekil 2.3'te görülen katman ekleme yaklaşımı kullanılmaktadır.



Şekil 2.3 Üretici (G) ve Ayırt Edici'ye (D) yüksek çözünürlüklü katmanların eklenmesi. (a) Düşük katman eğitim sonu (b) Kademeli çözünürlük artışı (c) Bir yüksek çözünürlüğe geçişin tamamlanması [20]

Sistemin çözünürlüğünün doğrudan artırılması eğitime şok etkisi yapacağından dolayı yüksek çözünürlüğe geçiş aşaması kademeli olarak yapılmaktadır. Kademeli geçiş sırasında ara çıktıları ağırlıklandırmak amacıyla 0-1 arasında bir değer alabilen α katsayısı belirlenir. Üretici tarafında ilk olarak düşük çözünürlüklü çıktının boyutu en yakın komşu filtrelemesi ile 2 katına çıkartılır ve $1 - \alpha$ ile ağırlıklandırılır. Aynı zamanda yeni eklenen yüksek çözünürlüklü katmandan alınan çıktı da α katsayısı ile ağırlıklandırılır ve toplanarak alınan çıktı Ayırt Edici'ye gönderilir. Ayırt Edici içerisinde de alınan girdinin boyutu ortalama örnekleme (average pooling) ile yarıya düşürülür ve $1 - \alpha$ ile ağırlıklandırılır. Buna ek olarak Üretici'den alınan çıktı yeni eklenen katmana verilerek çözünürlüğü yarıya düşürülür ve α çarpanı ile ağırlıklandırılarak her iki çıktı toplanır. Toplanan çıktılar bir alt katmana gönderilerek eğitim devam ettirilir. Eğitim ilerledikçe yeni eklenen katmanların ağırlığını artırmak amacıyla α değeri 0'dan 1'e doğru yaklaştırılır.

Kademeli çözünürlük artırma yaklaşımına ek olarak sistemin eğitim kararlılığını artırmak için PGGAN çalışmasında kullanılan diğer yaklaşımlar aşağıdaki gibidir:

- Eğitim aşamasında kullanılan her mini-yığın (mini-batch) için Ayırt Edici'de

görüntülerin standart sapması hesaplanarak üretilen görüntü çeşitliliğinin artırılması

- Katmanların kendi içinde eşitlenmiş öğrenme oranı kullanması ile stabil eğitimin sağlanması
- Eğitimde kararlılığı artıran yığın normalleştirmesinin maliyetli olması nedeniyle Üretici içerisindeki ara katman özniteliklerine pikselleri tabanlı normalleştirme yapılması

Sistem, kayıp fonksiyonu olarak bir önceki bölümde bahsedilen WGAN-GP fonksiyonunu kullanmaktadır. Eğitim aşamasında WGAN'dan farklı olarak Ayırt Edici her Üretici güncellenmesi için 1 kere güncellemeye sokulur. Buna ek olarak Ayırt Edici'nin kayıp fonksiyonunun 0'dan çok uzaklaşmaması için kayıp fonksiyonuna ek girdi olarak 0.001 gibi bir sürüklenme (drift) ağırlığı da eklenmiştir. Eklenen sürüklenme ağırlığı ile kayıp fonksiyonunun son hali Eşitlik 2.5'te verilmiştir. Ayırt Edici'den alınan çıktıların kareleri alındıktan sonra ortalama değerleri hesaplanır ve sürüklenme ağırlığı ile çarpılır. Çarpımdan alınan sonuç kayıp fonksiyonuna eklenerek Ayırt Edici'nin kaybı belirlenir.

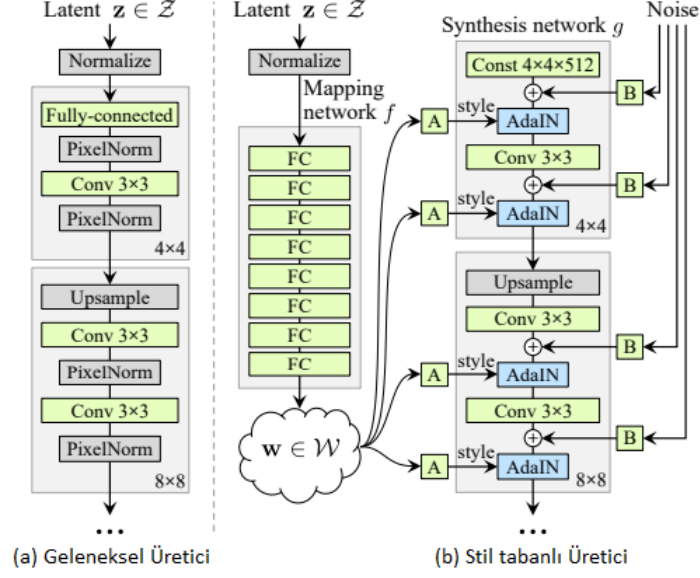
$$L' = L + \epsilon_{drift} \mathbb{E}_{x \in \mathbb{P}_r} [D(x)^2] \quad (2.5)$$

Klasik yöntemlerin aksine daha yüksek kalitede ve çözünürlükte çıktılar üretilmesine imkan sağlayan PGGAN yöntemi, derin öğrenme tabanlı sistemlerde yapay görüntü ile veri artırımı yapılabilmesi için ideal adaylardan birisidir.

2.1.1.4 Stil Tabanlı Üreticiye Sahip GAN (StyleGAN)

2018 yılında Karras vd. tarafından tasarlanan stil tabanlı Üretici'ye sahip GAN (StyleGAN) mimarisinde Üretici tarafında ve üst parametrelerde yapılan değişiklikler ile çıktı kalitesi artırılmıştır [21]. PGGAN modelinde gizli uzaydan örneklenen gizli vektörün doğrudan Üretici'ye verilmesi yerine bu vektör tamamen bağlı (fully connected) yapıdaki eşleme ağına beslendikten sonra Üretici'nin farklı katmanlarına ek girdi olarak beslenmiştir.

PGGAN benzeri kademeli büyüyen yapıda eğitimi gerçekleştirilen StyleGAN modelinde PGGAN mimarisine dört adet yenilik eklenmiştir. İlk olarak sistemin ara katmanlarından geçen görüntülerin yukarı/aşağı ölçeklenmesi için en yakın komşuluk hesabı yerine çift doğrusal örnekleme (bilinear sampling) kullanılarak ara çıktıların kalitesi artırılmıştır.



Şekil 2.4 (a) Kademeli büyüyen GAN ile Stil tabanlı Üretici'ye sahip GAN (b) ağlarının Üretici ağları [21]

Şekil 2.4'te görülebilen ikinci yenilik gizli uzaydan örneklenen vektörü direkt olarak Üretici'ye beslemek yerine bu vektörün, sekiz katmandan oluşan tamamen bağlı yapıdaki bir ağa verilerek ağın çıktısının stil vektörü olacak şekilde alınmasıdır. Bu stiller alınan gizli uzay vektörlerini ara bir gizli uzaya eşlerler. Eşlenen vektörler Üretici'nin evrişim katmanlarına verilmeden önce girdi ile beraber kullanılarak uyarlanabilir örnek normalizasyonu işleminden (Adaptive Instance Normalization) (AdaIN) geçirilir.

PGGAN mimarisinin aksine StyleGAN mimarisinde Üretici'nin giriş katmanında 4x4x512 boyutlu sabit bir girdi vektörü bulunmaktadır. Bu girdi vektörü üzerinde daha derin katmanlarda yer alan evrişim işlemlerinden önce uygulanan AdaIN işlemi Eşitlik 2.6'da verilmiştir. Burada her x_i öznetelik haritası ayrı bir şekilde normalizasyon işleminden geçirilir. Normalizasyon işlemine ek olarak eşleme açısından alınan y öznetileklerinin ilgili stil y_s ve bias y_b bileşeni kullanılarak girdinin kontrolü sağlanır.

$$\text{AdaIN}(x_i, y) = y_{s,i} \frac{x_i - \mu(x_i)}{\sigma(x_i)} + y_{b,i} \quad (2.6)$$

Üçüncü yenilik Üretici ağındaki evrişim işleminden alınan etkilenim haritalarına (activation maps) bir gürültü vektörünün AdaIN işleminden önce eklenmesidir. Eklenen gürültü stil seviyesinde çıktıların daha çeşitli olmasını sağlamaktadır. Örneğin sisteme farklı saç şekillerini göstermek amacı ile kullanılabilir.

Yapılan son yenilikte ise stil açısından alınan iki farklı vektörün ilki Üretici'nin

ara katmanlarındaki belirlenen bir AdaIN işleminden önce girdi olarak verilerek, ikinci çıktı vektörü de belirlenmiş AdaIN katmanından sonraki katmanlara eklenerek eğitimin yapılmasıdır. Bu yenilik, Üretici'nin belirli stilleri belirli katmanlara eşlemesini ve çıktı detaylarını ona göre değiştirmesini sağlar.

Sonuç olarak girdi katmanına yakın verilen stil vektörleri çıktıda büyük özellik değişikliklerine sebep olmakta, çıktıya yakın beslenen vektör ise görüntüde daha detaylı değişikliklere neden olmaktadır. Örnek olarak yüz resimleri üretilirken üst katmanlara beslenen vektörler poz, cinsiyet, yüz şekli; alt katmanlara verilen girdiler saç/göz rengi gibi durumları değiştirmektedir.

StyleGAN mimarisi ile yapılan eğitimlerde alınan çıktılar daha yüksek kalitede ve çok daha çeşitli olmaktadır. Buna ek olarak stil vektörünün etkilenebilmesi farklı görüntüler arasında stil transferi yapılmasına imkan sağlamaktadır. Bu mimari ile çok çeşitli ve kaliteli çıktılar üretmek mümkündür.

2.1.2 Yapay Veri Üretimi Başarısının Ölçülmesi

Üretilen yapay görüntülerin kalitesinin ölçülmesi, anormal görüntülerin tespiti için geliştirilecek sistemlerde hangi mimarinin kullanılabilir olduğuna karar verilmesi için önemli bir adımdır. Bu kapsamda üretilen yapay görüntülerin kalitelerine karar verilmesi için genel kabul görmüş kesin bir yöntem olmamakla beraber [22], medikal görüntü üretimi çalışmalarında yaygın olarak otomatik veya insan geri bildirimine dayalı yaklaşımlar kullanılmaktadır. Bu yaklaşımlar aşağıdaki gibidir.

2.1.2.1 Görsel Turing Testi (Visual Turing Test)

Yapay görüntülerin gerçek olanlardan ayırt edilebilmesi için insan değerlendirmesine dayanan Görsel Turing Testinden (VTT) [23] yararlanılabilir. Kullanıcılara anket şeklinde sunulan bu test içerisinde gerçek ve yapay görüntüler karıştırılarak bu kullanıcılardan gördükleri görüntüleri gerçek veya yapay olarak sınıflandırmaları istenir. Yapay görüntülere ne kadar fazla gerçek olarak oy verilirse yapay görüntü üretim kapasitesinin o kadar iyi olduğu kabul edilir.

Medikal görüntüler gibi uzmanlık gerektiren alanlarda yapay görüntülerin gerçekliğinin tespiti için bu alan dışında eğitilmiş otomatik olarak skor üreten yöntemlere bir alternatif olarak kullanılmaktadır [24].

2.1.2.2 Başlangıç Mesafesi (Inception Score)

Yapay görüntülerin gözle kontrol edilmesi zaman alan ve alan bilgisi gerektiren bir süreçtir. Bu durumu otomatize hale getirmek için kullanılan yaklaşımlardan birisi Başlangıç Mesafesi (Inception Score) (IS) metriğidir [23]. Üretilen yapay görüntüler ImageNet [25] veri kümesi üzerinde eğitilmiş InceptionV3 mimarisindeki yapay sinir ağından geçirilerek görüntülerin sistemdeki sınıflara ait olma olasılıklarından şartlı etiket dağılımları (conditional label distribution) hesaplanır ve bu dağılımların KL mesafesi hesaplanır.

Üretilen görüntülerin anlamlı nesneler içermesi dağılımın düşük entropiye sahip olmasını sağlar. Buna ek olarak üretilen görüntülerin yüksek çeşitlilikte olması gerekmektedir. Bundan dolayı üretilen nesnelerin bağlı sınıf dağılımları da yüksek entropiye sahip olmalıdır. İncelenen iki dağılım arasındaki KL uzaklığı hesaplanarak üretilen görüntülerin başlangıç skoru hesaplanır. Başlangıç skoru mesafesi metriğinin üretilen yapay görüntülerde cevaplamaya çalıştığı iki soru;

- Görüntü kalitesi: Yapay görüntüler belirli bir nesneye benziyor mu?
- Görüntü çeşitliliği: Üretilen yapay görüntüler ne kadar çeşitli?

şeklindedir. Eğer bu iki sorunun da cevabı doğruysa skor yüksek, değilse düşük çıkacaktır.

Bu skor metriği eğitildiği veri kümesi dışında medikal görüntüler gibi sınıflara sahip alanlar için düşük skorlar üretebilmektedir [26]. Buna ek olarak gerçek veri kümesi ve yapay veri kümesi arasındaki ilişki de dikkate alınmadığı için direkt olarak gerçek/yapay veriyi karşılaştırma imkanı sunmamaktadır.

2.1.2.3 Fréchet Başlangıç Mesafesi (Fréchet Inception Distance)

Yapay görüntülerin otomatik olarak kontrol edilmesine imkan tanıyan bir diğer skora yöntem olan Fréchet Başlangıç Mesafesi (FID) skoru da IS metriğiyle aynı yapıdan yararlanmaktadır. Hem gerçek hem de yapay görüntüler InceptionV3 sisteminden geçirilir ancak IS skorundan farklı olarak son çıktı katmanındaki sınıf olasılıklarının alınması yerine ağın çıktıya yakın ara katmanından görüntülerin özellik vektörleri alınır. Çıkarılan özellik vektörleri grubunun birbirlerine olan yakınlıkları Fréchet mesafesi kullanılarak ölçülür.

FID skorunun matematiksel hesaplaması Eşitlik 2.7’de verilmiştir. İlk olarak gerçek görüntü vektörlerinin ortalama değerleri μ , yapay görüntü vektörlerinin ortalama

değerleri μ_w olacak şekilde hesaplanır. Hesaplanan ortalama vektörlerin karesel farklarının toplamı alınır. Gerçek ve yapay görüntülerin özellik vektörlerinin eşdeğinti (covariance) matrislerini ifade eden Σ ve Σ_w değerlerinin toplamı ve çarpımlarının karekökleri ile hesaplama yapıldıktan sonra tr işlemi ile sonuç matrisinin çapraz boyuttaki elemanlarının toplamı alınarak hesaplama tamamlanır.

$$FID = ||\mu - \mu_w||_2^2 + \text{tr}(\Sigma + \Sigma_w - 2(\Sigma^{1/2}\Sigma_w\Sigma^{1/2})^{1/2}) \quad (2.7)$$

Aradaki mesafe ne kadar yakın olursa yapay görüntülerin gerçek görüntülere daha benzer bir dağılımda ve daha yüksek kalitede olduğu söylenebilir [27].

2.1.2.4 Öğrenilmiş Algısal Görüntü Yama Benzerliği (Learned Perceptual Image Patch Similarity)

Yapay olarak üretilen resimlerin kalitesinin yanı sıra bu görüntülerin çeşitliliğinin de ölçülmesi en yüksek performansa sahip GAN mimarisine karar vermede etkili bir adımdır. Bu amaç doğrultusunda Öğrenilmiş Algısal Görüntü Yama Benzerliği (LPIPS) metriği sistemin ürettiği yapay görüntüleri ve gerçek görüntüleri kullanarak GAN içerisindeki Üretici'nin kapasitesini gösteren bir benzerlik skoru üretir [28].

Bu benzerlik skorunun hesaplanması için hem gerçek görüntülerden hem de yapay görüntülerden rastgele seçilmiş veri ikilileri oluşturulur. Oluşturulan görüntü ikilileri ImageNet verisi üzerinde eğitilmiş bir yapay sinir ağına verilerek sınıflandırma katmanlarından önceki aktivasyon katmanlarının farkı alınır. Bütün ikililer arasındaki ortalama fark çeşitlilik skorunu verir. Yüksek bir skor Üretici'nin daha çeşitli görüntüler üretebilme kapasitesine sahip olduğunu belirtir.

2.2 Özellik Çıkarma (Feature Extraction)

Patolojik görüntülerin tespitinin yapılabilmesi amacıyla görüntülerin özelliklerinin çıkartılması gerekmektedir. Görüntülerden özellik çıkarımı için farklı yaklaşımlar kullanılabilir. Bunlardan bazıları aşağıdaki gibidir.

- Görüntü içerisindeki kenar, renk, doku bilgileri gibi karakteristik özellikler
- Alan bazı uzmanlık gerektiren patolojiyi işaret eden renk veya anatomik özellikler
- Ön eğitilmiş (pretrained) evrişimli sinir ağlarının (Convolutional Neural Network) (CNN) ara katmanlarından özellik vektörü çıkarımı

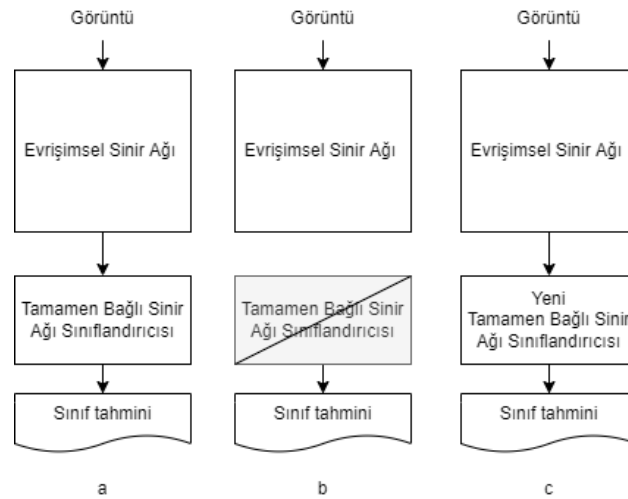
- Eğitildiği veri kümesini öğrenen otokodlayıcı sistemler

Derin öğrenme yaklaşımları ile eğitilen tanıma sistemlerinin başarısı renk, doku, kenar gibi öznitelikler çıkartılarak eğitilen klasik yaklaşımlardan daha yüksek başarımlar göstermektedir. Buna ek olarak bu özelliklerin patolojileri anlamlı bir şekilde ayırt edenlerini çıkartmak için görüntü ile ilgili alanda uzmanlık bilgisi gerekmektedir. Uzmanlar tarafından görüntülerden özelliklerin tek tek çıkartılması zaman alıcı bir iştir ve insan hatasına da açıktır.

2.2.1 Ön Eğitimli Evrişimli Sinir Ağı ile Öznitelik Çıkarımı

ImageNet (1.4 milyon etiketli görüntüye sahip 1000 sınıflı bir veri kümesi) gibi bir veri kümesinde ön eğitimli bir CNN yardımı ile özellik çıkarımı literatürde yaygın olarak kullanılmaktadır [29]. Büyük miktarda veri ile eğitilen bu modellerin sınıflandırma katmanından önceki ara katman çıktıları alınarak öznitelik vektörleri elde edilebilir. Bu öznitelik vektörleri farklı bir sistemde eğitmek üzere kullanılabilirler.

Ancak, eğitimi yapılan veri kümesi ile adapte edilecek problemin veri kümesinde farklar olduğu zaman alınan öznitelik vektörleri yeni veriyi tam olarak temsil etme kabiliyetine sahip olmayabilir. Yeni görüntülerin özelliklerinin anlamlı bir şekilde anlaşılabilmesi için sisteme yeni görüntülerin verilerek öğrenme aktarması (transfer learning) ve ince ayar (fine-tuning) yapılması gerekmektedir. Bu kapsamda sınıflandırıcı mimarisinde yapılması gereken değişiklikler Şekil 2.5'te verilmiştir.



Şekil 2.5 CNN transfer öğrenme şeması. (a) Ön Eğitimli CNN tabanlı sınıflandırıcı. (b) İlk olarak önceden eğitilmiş YSA'nın sınıflandırıcı katmanı evrişim katmanından kopartılır. (c) Yeni tamamen bağlı sınıflandırıcı katmanı sisteme eklenir.

Sistem yeni veri kümesine uyarlanırken ilk olarak evrişim katmanlarının ağırlıkları dondurularak sadece yeni eklenen sınıflandırma katmanı eğitilir. Bu aşamada hiç bir

katmanın ağırlıklarını dondurmada bütün sinir ağıнын eğitimi yapılırsa; yeni eklenen katmandan alınan yüksek hata değeri öznitelik çıkartan CNN bloğunun ağırlıklarını bozarak öğrenilen kaliteli filtrelerin bozulması riskini ortaya çıkartır. Bu kapsamda ilk olarak sınıflandırıcı katmanının özel olarak eğitilmesi, ardından bütün sistemin veya sistemin bir kısmının öğrenime açılması gerekir.

Öğrenme aktarılması tamamlandıktan sonra evrişim katmanlarının da yeni veri kümesine adapte olmasını sağlamak amacıyla ara katmanların ağırlıkları açılarak eğitime devam edilebilir. Önceden eğitilen evrişim katmanlarının çıkaracağı öznitelikleri bozmamak adına ince ayar eğitimi yapılırken sistemin öğrenme katsayısının küçük olması gerekmektedir.

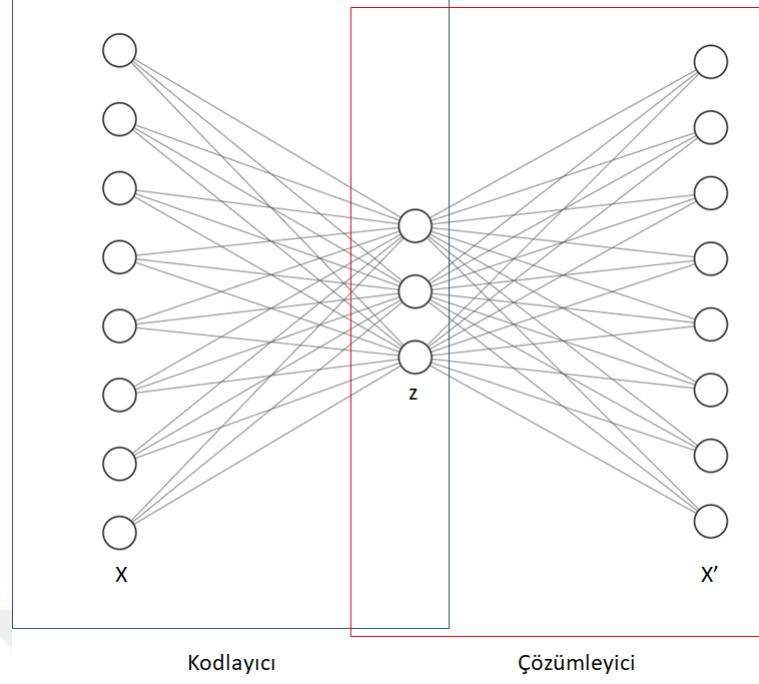
Sistemin eğitimi aşamasında yaptığı hataları ölçebilmesi ve dolayısıyla model ağırlıklarının güncellenebilmesi için uygun bir kayıp fonksiyonunun seçilmesi gerekmektedir. Normal ve patolojik görüntülerin ayrılabilmesi amacıyla eğitilecek bir sistemde ikili çapraz entropi (binary cross-entropy) kayıp fonksiyonu kullanılabilir. Eşitlik 2.8’de görülebilen bu maliyet fonksiyonunda çıktı için beklenen 0 veya 1 değerlerinin doğrulanarak log tabanlarındaki tahminlerine olan uzaklıkları hesaplanır. Bu formülde y_i 1 sınıfının olasılık değerini, $1 - y_i$ ise 0 sınıfının olasılık değerini ifade etmektedir.

$$\text{İkili Çapraz Entropi} = -(y_i * \log(\hat{y}_i) + (1 - y_i) * \log(1 - \hat{y}_i)) \quad (2.8)$$

Bu tür sistemler eğitildikleri veri kümesine bağımlı olduğu için etiket bilgisi bulunmayan ve tamamen farklı bir alandaki veri kümelerinden (günlük hayat görüntüleri ile eğitilmiş bir sistemden medikal görüntü özellikleri çıkartmak gibi) sağlıklı özellikler çıkartamayabilirler. Bu darboğazı aşmak için CNN ile karşılaştırılabilir özellik çıkartma kapasitesine sahip otokodlayıcı sistemlerden yararlanılabilir [30].

2.2.2 Otokodlayıcı ile Öznitelik Çıkarımı

Veri kümesinin büyük kısmında etiket bilgisi bulunmadığı zaman verinin ait olduğu uzayı anlamlandırarak öznitelik çıkarılmasına imkan sağlayan otokodlayıcı mimarisi kullanılabilir. Kodlayıcı (encoder) ve çözümleyici (decoder) adında iki ayrı yapay sinir ağından oluşan otokodlayıcı mimarisinin kodlayıcı parçası ile sisteme verilen bir görüntünün öznitelikleri çıkartılabilir. [31]. Örnek bir otokodlayıcı mimarisi Şekil 2.6’da görülebilir.



Şekil 2.6 Otokodlayıcı mimarisi

Sistemin eğitim aşamasında kullanılan yeniden oluşturma kayıp fonksiyonu Eşitlik 2.9'da verilmiştir. Kodlayıcının giriş katmanına (ϕ) verilen bilginin (X) boyutu bir darboğazda sıkıştırarak gizli uzayda temsil etmek üzere (z) azalır. Ardından çözümleyici ağı (σ) verilen gizli vektörden girdi tekrar oluşturulur ($\hat{X}$). Girdi X ile çıktı $\hat{X}$ arasındaki farkın karesi alınarak kayıp hesaplanır.

$$L_{recon}(X_i, \hat{X}_i) = \|X_i - \hat{X}_i\|^2 = \|x_i - \phi(\sigma(x_i))\|^2 \quad (2.9)$$

İki boyutlu görüntüler için evrimsel otokodlayıcılar kullanılır. Tamamen bağlı nöronlardan oluşan otokodlayıcılardan farklı olarak giriş, ara ve çıktı katmanlarda evrişim katmanlarından yararlanılmaktadır.

Sistemin eğitilebilmesi için kayıp fonksiyonu olarak yeniden oluşturma kayıp fonksiyonu kullanılır. Kayıp fonksiyonu girdi ve çıktı görüntülerinin alt bölümlerine (patch) ($i = 1, 2, \dots, p$) uygulanıp toplanır ve toplam parça sayısına bölünerek Eşitlik 2.10'da verilen maliyet fonksiyonu elde edilir. Eğitim belirlenen dönem (epoch) sayısı boyunca veya maliyet fonksiyonunun düşüşü durana kadar devam ettirilebilir.

$$J(\theta) = \frac{1}{p} \sum_{i=1}^p L[x_i, \hat{x}_i] \quad (2.10)$$

Otokodlayıcının nihai amacı aldığı girdiyi tekrar oluşturabileceği şekilde olabilecek en düşük boyutta kodlamaktır. Burada kodlamanın yapıldığı gizli vektör boyutu girdi boyutundan daha düşük olduğu için bilgiyi daha sıkışık formatta saklayabilmektedir. Görüntü gibi yüksek boyutta bilgi içeren veriler yerine bunların sıkıştırılmış temsillerinin iyi bir şekilde alınabilmesi için gizli vektör boyutunun çok düşük olmamasına dikkat edilmelidir.

2.3 Anormal Verilerin Tespiti

Veri kümesinde normalin dışında olan yani beklenmeyen durumların bulunduğu veriler anormal veri olarak adlandırılırlar. Sahte kredi kartı işlemleri, ağ sistemlerindeki sahte paketler, medikal görüntülerdeki patolojiler bu tür anormal verilere örnek olarak verilebilir.

Artan veri miktarından dolayı toplanan tüm veriyi etiketlemek ilgili alanda uzmanlık gerektirmekte ve etiketleme süreci insan gücü açısından çok maliyetli olmaktadır. Bu tür etiketsiz verilerde anormal noktaların tespiti için gözetimsiz yöntemlerin kullanılması gerekmektedir.

2.3.1 İzolasyon Tabanlı Karar Ağacı Ormanı ile Anormal Verilerin Tespit Edilmesi

Çok sayıda gözleme sahip olan ve bu gözlemlere ait olan özellik vektörlerinin de büyük olduğu veri kümelerinde gözetimli/gözetimsiz sınıflandırma işlemleri için hesaplama maliyeti düşük olan algoritmalara ihtiyaç vardır. Kümeleme tabanlı makine öğrenmesi algoritmaları test aşamasında bütün eğitim örneklerini işlediğinden dolayı hme zaman hem de yer karmaşıklığı yüksek olacak şekilde çalışmaktadır. Test edilen gözlemin eğitim için kullanılan gözlemler ile arasında yüksek maliyetli bir mesafe alma işlemi gerekmektedir.

Normal olmayan gözlemleri izole ederek gözetimsiz bir şekilde ayrıştırılmalarını sağlayan izolasyon tabanlı karar ağacı ormanı (Isolation Forest) (IForest) algoritması ile veri kümesindeki anormal gözlemlerin hızlı bir şekilde ve sistem hafızasında daha az yer tutarak tespit edilmesi sağlanmıştır [32].

IForest algoritmasında normal verilerin anlamlandırılması yerine anormal verilerin izole edilerek çıkartılması sağlanır. Verilerin izole edilmesi bir gözlemin veri kümesindeki diğer gözlemlerden ayrıştırılarak tek başına bırakılması olarak tanımlanabilir. IForest algoritmasında gözlemler arası mesafe hesaplaması veya gözlemlerin yoğunluk hesabı kullanılmaz. Bunun yerine veri kümesinden rastgele

seilen gzlemler ve zellikler aldıęı gzlemleri bařarılı bir řekilde izole edebilen izolasyon karar aęacına (iTree) verilirler.

Tam ikili aęa (full binary tree) yapısında olan iTree karar aęacının aldıęı zelliklere gre gzlemler tekilleřerek zelliklerine gre blnrler. Kullanılan aęacın zellięinden dolayı karar aęacının her dalında ya iki yaprak vardır ya da hi yapraęı yoktur. Karar aęacının eęitimi tamamlandıęında btn gzlemler izole olurlar. Aęacın eęitildięi veriler incelendięinde anormal veriler kke daha yakın olma eęilimi gstermektedir. Bundan dolayı alt seviyede yer alma eęilimi gsteren normal verilerden bařarılı bir řekilde ayrılmaktadırlar.

Birden fazla izolasyon karar aęacı ile topluluk ęrenmesi modeli kurularak IForest modeli oluřturulur. Eęitim tamamlandıęında aęa topluluęunda ortalama kk mesafesi dřk olan gzlemler anormal olarak iřaretlenir. Orman oluřturma iřleminin szde kodu Algoritma 2.1'de verilmiřtir.

Algoritma 2.1 İzolasyon Tabanlı Karar Aęacı Ormanı Algoritması

Girdiler:

Veri kmesi : X

Ormandaki aęa sayısı : t

Aęa bařına dřen alt rnek sayısı : ψ

ıktılar:

iTree ormanı

Algoritma:

- 1: **for each** $tree \in t$ **do**
 - 2: $X' \leftarrow ornekle(X, \psi)$ (Veriden gzlemleri geri koymadan rnekleme yap)
 - 3: $Forest \leftarrow Forest \cup iTree(X')$ (Aęacı oluřtur)
 - 4: **end for**
 - 5:
-

Tam ikili aęa olan iTree'lerin oluřturulması iřleminin zaęrılı (recursive) olarak alıřan szde kodu Algoritma 2.2'de verilmiřtir.

Algoritma 2.2 Tam İkili Yapıdaki iTree Ağacı Algoritması (iTree)

Girdiler:

Örneklenmiş veri kümesi : X'

Çıktılar:

iTree ağacı

Algoritma:

```
1: if  $X'$  bolunemiyorsa then
2:   return  $disDugum\{Boyut \leftarrow |X'|\}$ 
3: else
4:    $X'$  e ait  $Q$  ozellikleri icin
5:    $q \in Q$  olacak sekilde rastgele ozellikler sec
6:    $X'$  icerisindeki  $q$  ozelliginin en dusuk ve en yuksek noktaları arasindaki bir  $p$ 
      noktasini bolum yeri olarak sec
7:    $X_l \leftarrow filtrele(X', q < p)$ 
8:    $X_r \leftarrow filtrele(X', q \geq p)$ 
9:   return  $icDugum\{Sol \leftarrow iTree(X_l),$ 
10:                 $Sag \leftarrow iTree(X_r),$ 
11:                 $BolunenOzellik \leftarrow q,$ 
12:                 $BolumEsigi \leftarrow p\}$ 
13: end if
```

Eğitim tamamlandıktan sonra anormal olan verilerin tespit edilebilmesi için ağaçlar içerisinde kök düğümden kontrol edilecek gözlemin bağlı olduğu iç düğüme kadar giden yolun uzunluğu hesaplanır. Algoritma 2.3'te ağaç içerisinde bulunan gözlemin köke olan yol uzunluğunun özçağrılı olarak hesaplanması işleminin sözde kodu görülebilir.

Bir gözlemin anormallik skorunun çıkartılması için Eşitlik 2.11'de yer alan ve yol hesabında da kullanılan $c(\psi)$ fonksiyonu verilmiştir. Burada ağaç boyu önceden belirlenmiş olan, ağaç başına düşen gözlem sayısı parametresinin (ψ) harmoniği ($\ln(i) + EulerSabit$) alınır. Eşitlikte hesaplanan $c(\psi)$ değeri, yol uzunluğu olan $h(x)$ değerinin ortalamasına denk gelmektedir. Bir gözlemin anormallik skorunu hesaplarken $c(\psi)$ değeri $h(x)$ değerini normalize etmek için kullanılır. Eşitlik 2.12'de verilen hesaplama ile bütün ağaçlardan alınan ortalama yol uzunluğu kullanılarak anormallik skoru çıkartılır. Alınan sonuçta s değerinin 1'e yakın olması o gözlemin anormal olduğuna, 0.5'ten düşük olması da gözlemin normal olduğuna işaret eder.

Algoritma 2.3 Gözlemin Ağaç İçerisinde Köke Olan Yol Mesafesinin Bulunması Algoritması(YolUzunlugu)

Girdiler:Test gözlemi : x Bir iTree ağacı : T Yükseklik sınır : $hlim$ İlk oluşturulduğunda 0 olan anlık yol uzunluğu : e **Çıktılar:** x 'in köke olan yol uzunluğu**Algoritma:**

```
1: if  $T$  disDugumse veya  $e \geq hlim$  then
2:   return  $e + c(T.Boyut)$ 
3: end if
4:  $a \leftarrow T.BolunenOzellik$ 
5: if  $x_a < T.BolumEsigi$  then
6:   return  $YolUzunlugu(x, T.sol, hlim, e + 1)$ 
7: else  $\{x_a \geq T.BolumEsigi\}$ 
8:   return  $YolUzunlugu(x, T.sag, hlim, e + 1)$ 
9: end if
```

$$c(\psi) = \begin{cases} 2H(\psi - 1) - 2(\psi - 1)/\psi & \psi > 2 \text{ ise,} \\ 1 & \psi = 2 \text{ ise,} \\ 0 & \text{diğer koşullarda} \end{cases} \quad (2.11)$$

$$s(x, \psi) = 2^{-\frac{E(h(x))}{c(\psi)}} \quad (2.12)$$

Test zamanında daha hızlı çalışması ve sistem hafızasına bindirilen maliyetin daha düşük olması IForest algoritmasını gözetimsiz olarak anormal verilerin tespiti için ideal bir aday yapmaktadır.

2.3.2 Anormal Veri Tespiti Performans Analizi

Bütün veri kümesi içerisinde normal verilere kıyasla çok düşük sayıda yer alan anormal verilerin tespitinin performansını ölçmek için doğruluk puanını (accuracy score) kullanmak uygun değildir. Bunun yerine test verisi üzerinde yapılan tahminlerden oluşturulan karışıklık matrisi (confusion matrix) Tablo 2.1'de görülebilir.

Tablo 2.1 Hata Dizeyi

	Tahmin Edilen	
	Gerçek Pozitif (TP)	Yanlış Negatif (FN)
Gerçekleşen	Yanlış Pozitif (FP)	Doğru Negatif (TN)

Doğruluk puanı hesaplanırken Eşitlik 2.13'teki gibi olacak şekilde doğru tahmin edilen veri sayılarının toplamının veri sayısına oranı alınır. Örneğin veri kümesinde %99 oranında normal veri, %1 oranda anormal veri olduğu durumda bütün test verisine normal demek ile doğruluk puanı %99'a çıkacaktır. İlk bakışta sistemin performansı yüksek gözükse de anormal veriler yakalanamadığı için kullanılan sistem başarısızdır.

$$\text{Doğruluk Puanı} = \frac{TP + TN}{TP + FP + TN + FN} \quad (2.13)$$

Eğitilen sistemin anormal olarak tahmin ettiği verilerden gerçekten kaç tanesinin anormal olduğunun ölçülmesi için kesinlik (precision) puanı Eşitlik 2.14'te görüldüğü gibi pozitif tahmin edilen veri sayısının toplam pozitif tahmin sayısına oranlanması ile hesaplanır.

$$\text{Kesinlik} = \frac{TP}{TP + FP} \quad (2.14)$$

Diğer yandan sistemin anormal olarak tahmin etmesi gereken verilerin gerçekten kaç tanesini doğru tahmin edebildiğinin ölçülebilmesi için duyarlılık (recall) başarısı hesaplanmaktadır. Duyarlılık başarısı eşitlik 2.15'te görüldüğü gibi doğru tahmin edilen veri sayısının toplam anormal veri sayısına oranı alınarak hesaplanır.

$$\text{Duyarlılık} = \frac{TP}{TP + FN} \quad (2.15)$$

Kesinlik ve Duyarlılık değerinin Eşitlik 2.16'da görülen harmonik ortalaması alınarak F1 skoru hesaplanır. Sınıfları dengesiz veriye sahip olan veri kümelerinde yanlış model seçimi yapılmaması için tüm hataları içeren F1 skorunun dikkate alınması gerekmektedir.

$$F_1 = 2 * \frac{\text{Kesinlik} * \text{Duyarlılık}}{\text{Kesinlik} + \text{Duyarlılık}} \quad (2.16)$$

3

SİSTEM TASARIMI

Medikal görüntülerde hastalık veya patolojik bir durumun tespiti günümüzde alanında uzman doktorlar tarafından yapılmaktadır. Tıbbi uzmanlığa erişimin kısıtlı olduğu yerlerde doktorlara yardımcı olacak sistemlere ihtiyaç duyulmaktadır. Otomatik ve başarılı bir şekilde anormal durumların tespiti hem doktorun zaman kazanmasını sağlayacak hem de tıbbi hizmete erişimin kısıtlı olduğu durumlarda alanında uzman olmayan doktorlara yardımcı olacaktır.

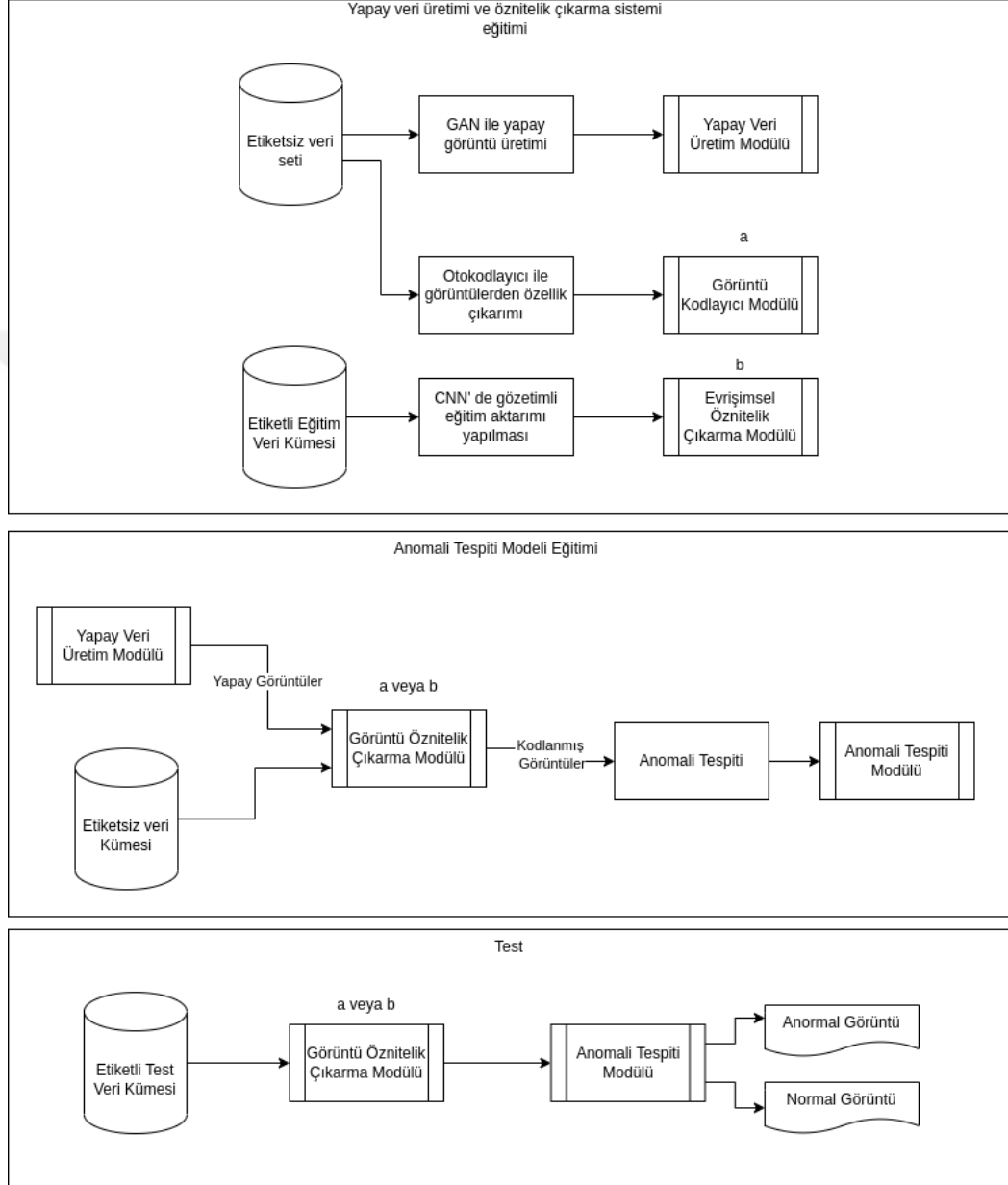
Anomali tespiti üzerine çalışan makine öğrenmesi yöntemleri yüksek miktarda veriye ihtiyaç duymaktadır. Kulak zarı görüntülerini toplamak zahmetli bir süreç olduğu için yapay olarak üretilmiş görüntülerin patoloji tespiti sistemlerinde kullanılarak nihai başarının artışına yardımcı olması hedeflenmektedir. Günümüzde yüksek gerçekçiliğe sahip görüntüler üretme kapasitesine sahip bir derin öğrenme mimarisi olan GAN'lar aracılığı ile yapay görüntülerin üretilmesi, yapay görüntülerin gerçek görüntülere benzer olmasını sağlayacaktır.

Bu çalışmada patoloji içeren kulak zarı görüntülerinin tespitini yapacak bir sistem tasarlanmış ve gerçekleştirilmiştir. İlk olarak GAN ile gerçek kulak zarı görüntülerinden yapay görüntüler üretebilecek bir model eğitilmiştir. Yapay görüntülerin kalitesi uzman doktorlar tarafından doğrulandıktan sonra veri arttırımı için kullanılmıştır.

Bir yandan da kulak zarı görüntülerinin özelliklerini çıkartabilecek kapasiteye sahip otokodlayıcı model eğitilerek bu modele girdi olarak verilen görüntülerin gizli vektörlerinin (latent vectors) alınabilmesi sağlanmıştır. Medikal görüntüler ile ilgili alan bilgisi gerekmeden yapay sinir ağından alınan özellik vektörü, özellik mühendisliği (feature engineering) ihtiyacını ortadan kaldırmaktadır.

Anormal kulak zarı görüntülerinin tespit edilebilmesi amacıyla Isolation Forest(IForest) öğrenme modeline otokodlayıcıdan çıkartılan görüntülerin gizli temsil vektörleri verilerek gözetimsiz eğitim yapılmıştır. Yapay olarak üretilmiş kulak zarı görüntüleri ile veri kümesinin boyutunun artırılmasının test başarısına etkisinin

ölçülebilmesi amacıyla ilk olarak sadece gerçek veri kümesi ile IForest modeli eğitilmiş ve sistem başarısı hesaplanmıştır. Sonrasında farklı miktarlarda üretilen yapay görüntüler gerçek görüntülerle birleştirilerek ayrı bir model eğitimi tamamlanmıştır ve bu modelin başarıları, ilk yapılan modelin başarısıyla kıyaslanmıştır. Tasarlanan sistem Şekil 3.1’de görülebilir.

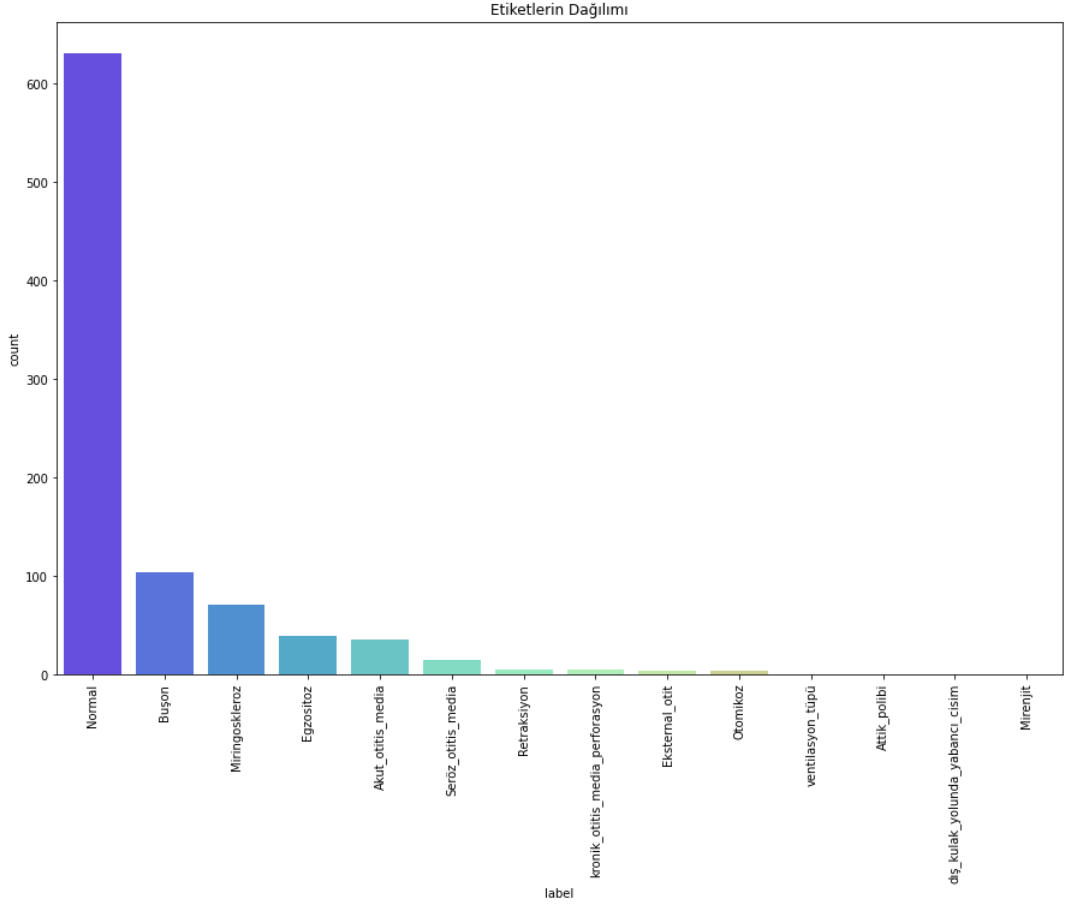


Şekil 3.1 Anormal kulak zarı görüntüsü tespiti sistemi tasarımı

3.1 Veri Kümesi

Bu çalışmada kullanılan veri kümesi bir KBB uzmanı tarafından otoskop cihazı yardımıyla alınan 1920x1080 çözünürlüğe sahip 10.088 adet kulak zarı görüntüsünden oluşmaktadır. Toplanan verilerde verilerin alındığı kişilerin cinsiyet,

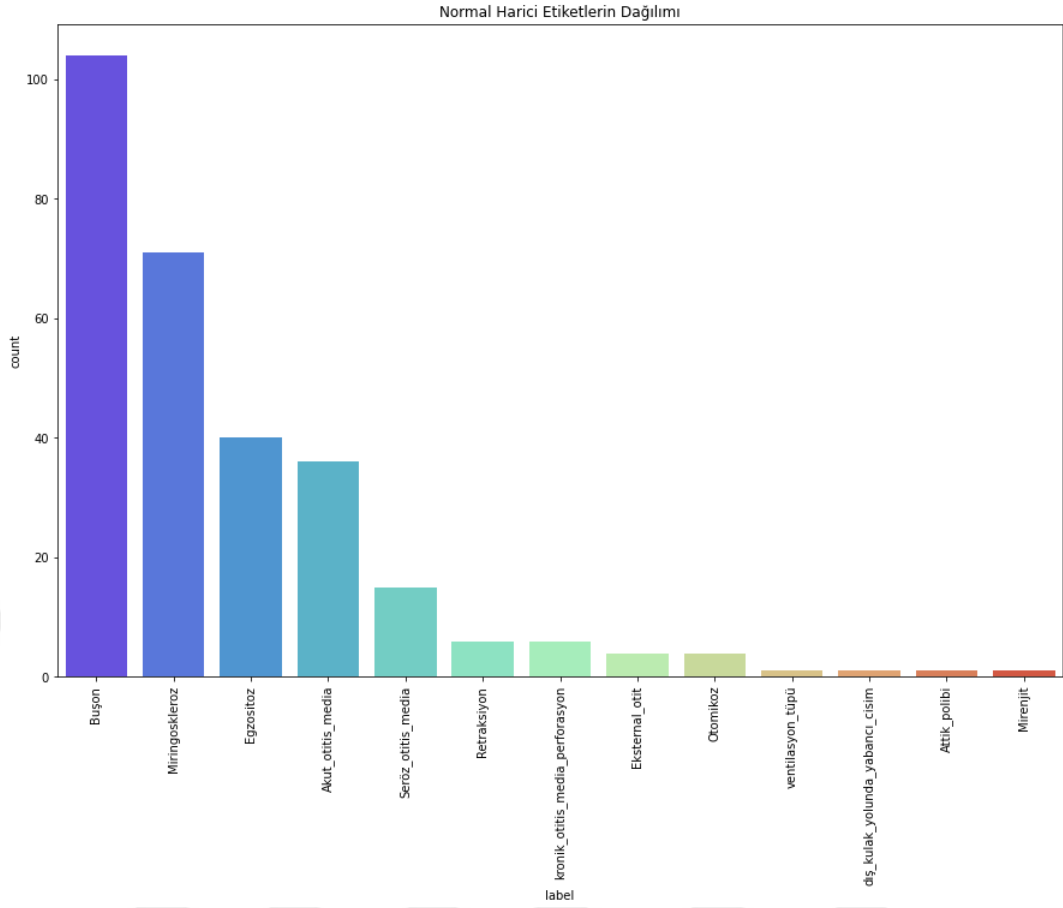
yaş gibi bilgileri dikkate alınmamıştır. Toplanan anonim veriler karıştırıldıktan sonra sistemin alt parçalarında kullanılmak amacıyla saklanmıştır. Bu verilerden rastgele seçilen 921 görüntü KBB uzmanı tarafından etiketlenmiştir. Toplanan etiketlerin sınıf dağılımları Şekil 3.2’de görülebilir. Patolojiye sahip verilerin dağılımı Şekil 3.3’te görülebilir.



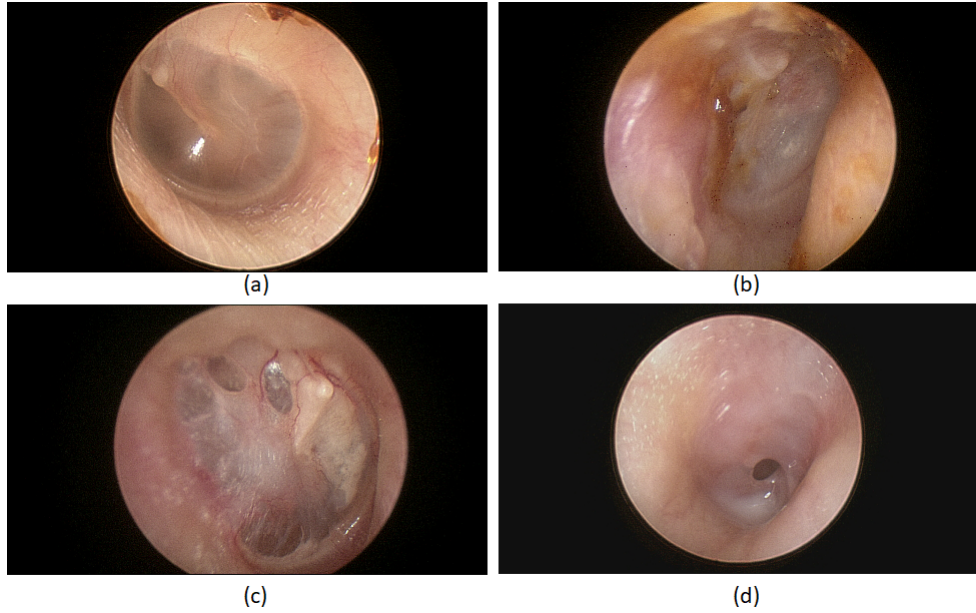
Şekil 3.2 Toplanan etiketlerin veri dağılımı

Sistemin verilen imgelerde sınıf yerine normal-anormal olarak sınıflandırma çıktısının alınabilmesi için *normal* haricindeki etiketler *anormal* olarak işaretlenmiştir. Gözetimli öğrenme aktarması yapılabilmesi için etiketlenen veri üzerinde 1 adet etiket olan görüntüler veri kümesinin dışına alınmıştır.

Veri kümesi eğitim ve test verisi olarak ikiye bölünmüştür. Bölünmenin ardından 86 adet anormal etiketli ve 190 adet normal etiketli görüntü içeren test verisinde sonuçlar incelenerek tez çalışması gerçekleştirilmiştir. Farklı etiketlere sahip görüntüler Şekil 3.4’te görülebilir.



Şekil 3.3 Toplanan patolojik etiketlerin veri dağılımı

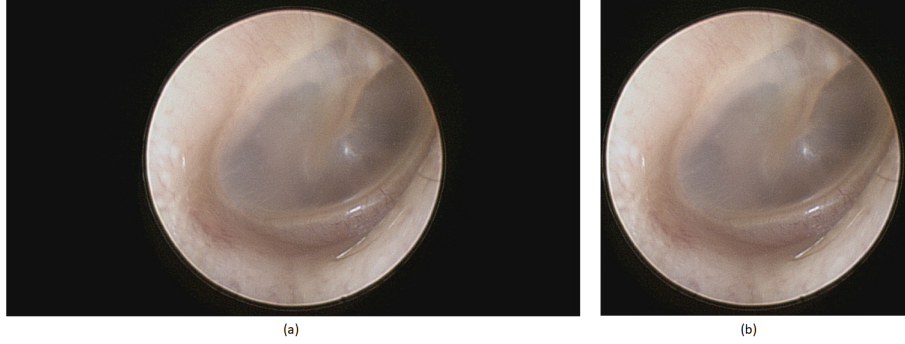


Şekil 3.4 Etiketlenen verilerden örnekler. (a): Normal, (b): Otitis Media, (c): Skleroz, (d): Zar Perforasyonu

3.2 Veri Ön İşleme

Veri kümesi incelendiğinde kulak zarı görüntülerinde önemli olan bölgenin resmin orta bölgesinde kaldığı ve çember etrafında kalan bölgenin siyah renkte olduğu

görülmüştür. Bu alanların eğitime verilmemesi amacıyla kulak zarı görüntüleri ortadan kesilecek şekilde otomatik olarak düzenlenmiş ve kaydedilmiştir. Alınan çıktılar Şekil 3.5'te görülebilir. Şekilde *a* görüntüsü orijinal görüntüyü, *b* görüntüsü hizalama işleminden sonra alınan çıktıyı göstermektedir.



Şekil 3.5 Kulak zarı görüntülerinin hizalanması

Evrişimli yapay sinir ağlarında kullanılan görüntülerin ön işlemlerden geçirilmesi bu yapay sinir ağlarının performansının artmasını sağlamaktadır [33]. Farklı görüntü piksellerinin standart aralıklarda olması verinin daha organize olarak tek tip olmasını ve dolayısıyla yapay sinir ağının aldığı görüntülerden daha anlamlı öznitelikler çıkartmasını sağlar.

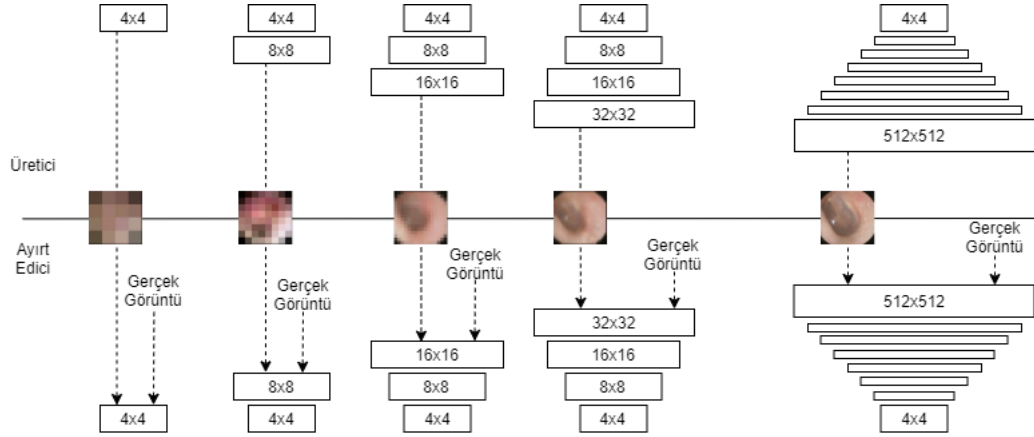
Bu kapsamda hem yapay veri üretimi sisteminin eğitimi hem de otokodlayıcı ile görüntü özelliklerinin çıkartılması için yapılan eğitim öncesinde kulak zarı görüntüleri standardizasyon işlemine sokulmuştur. Standardizasyon işlemi Eşitlik 3.1'de görülebilir. Verinin piksellerinin ortalama değeri (μ) ve standart sapması (Σ) ile görüntüler tek tip hale getirilmiştir.

$$X' = (X - \mu) / \Sigma \quad (3.1)$$

Veri kümesinde yer alan kulak zarı görüntü piksel değerlerinin ortalanması ve standardizasyonu işlemleri öncesinde kullanılacak yapay sinir ağının tasarımına göre veriler yeniden boyutlandırılmışlardır. Kullanılan yapay sinir ağları ve eğitimde kullanılan çözünürlükler Tablo 3.1'de görülebilir.

Tablo 3.1 Yapay sinir ağlarında kullanılan çözünürlükler ($n \times n$ ikilileri)

Model	4	8	16	32	64	128	256	512
WGAN-GP						✓		
PGGAN	✓	✓	✓	✓	✓	✓	✓	✓
StyleGAN	✓	✓	✓	✓	✓	✓	✓	
Otokodlayıcı						✓		



Şekil 3.6 Çalışmada kullanılan Kademeli Büyüyen GAN mimarisi ve örnek ara çıktıları

PGGAN ve StyleGAN ağları kademeli çözünürlükler ile eğitildikleri için ağın yapısındaki ilk çıktı katmanından alınması gereken (4 x 4) boyutu ile yeniden boyutlandırma yapılmıştır. Aynı şekilde ağların alt kademelerinde çözünürlük arttığı için girdi resmi yeniden uygun çözünürlüğe boyutlandırılmıştır.

3.3 Yapay Veri Üretimi

Veri kümesindeki sınıf dağılımlarının dengesiz olması veya toplamda yetersiz miktarda gözleme sahip olunması makine öğrenmesi temelli sistemlerde sınıf tahminleme başarısının düşük olmasına neden olmaktadır.

Sistemin eğitiminde kullanılan kulak zarı görüntülerinde sınıf bilgisinin olmamasından dolayı anormal görüntülerin tespiti için gözetimsiz yaklaşımlar kullanılmıştır ve aynı sebepten ötürü sınıflara göre veri kümesi üzerinde üst veya alt örnekleme işlemlerinin kullanılmasına engel olmuştur.

Yapay veri üretimi için literatürde yaygın olarak kullanılan çeşitli GAN mimarilerinin denenmesine ve bunların kalitesinin KBB uzmanları tarafından değerlendirilerek en gerçekçi görüntüleri üreten mimarin yapay veri üretimi aşamasında kullanılmasına karar verilmiştir. GAN mimarileri eğitildikleri verinin dağılımını öğrenmeyi hedeflediktedir. Bu sebeple üretilen görüntülerin de gerçek veriler ile benzer dağılımlar göstererek sonuç olarak sisteme daha çok veri beslenmesini sağlaması amaçlanmıştır.

Kullanılan veri kümesine en yakın görüntüleri üreten mimariyi bulmak amacıyla;

- Klasik GAN'dan farklı bir hata fonksiyonu kullanarak gerçek veri dağılımına daha yakın ve kaliteli örnekler üretmeyi amaçlayan WGAN-GP

- Çözünürlüğü kademeli olarak artırarak son çıktıda yüksek çözünürlüklü ve daha gerçekçi görüntüler üretmeyi amaçlayan PGGAN
- PGGAN'ın kalitesini iyileştirmeye yönelik üst değişkenlerde ve üretici yapısında değişiklikler yapılan StyleGAN

mimarilerinin kullanılmasına karar verilmiştir. Çalışmada kullanılan PGGAN mimarisinin yapısı ve ara çıktıları Şekil 3.6'da görülebilir. Eğitimi tamamlanan modellerden en gerçekçi kulak zarı görüntülerini üreten mimariyi bulmak amacıyla bir sonraki bölümde anlatılan model seçimi yaklaşımı uygulanmıştır. Eğitimin yapılan donanım kısıtlarından veya mimarinin limitlerinden dolayı her model için eğitim sonunda alınabilen maksimum çözünürlük Tablo 3.2'de görülebilir.

Tablo 3.2 Kullanılan Mimarilerin Çözünürlük Kapasitesi

Mimari	Çözünürlük
WGAN-GP	128 x 128
PGGAN	512 x 512
StyleGAN	256x256

3.4 Model Seçimi

İlk olarak toplanan verideki etiketsiz görüntüler kullanılarak seçilen GAN modeli eğitilmiş ve kullanıcının istediği miktarda veri üretebilecek modül elde edilmiştir. Farklı GAN mimarilerinden alınan çıktılarından en kaliteli olan mimarinin tespit edilebilmesi için Görsel Turing Testi yapılmıştır. Eğitilmiş olan GAN mimarilerinden alınan çıktılar orijinal veri kümesinden rastgele olarak seçilen görüntüler ile karıştırılmıştır ve standart olmaları için 256x256 çözünürlüğe alınarak uzmanlara görüntüleri gerçek veya yapay olarak oylayabileceği bir sistem hazırlanarak sunulmuştur.

Eğitim sonunda her modelden 500 adet yapay görüntü üretilmiş ve bu görüntüler arasından 30 tanesi rastgele seçilmiştir. Sonuç olarak elde edilen 120 adet yapay görüntü ile veri kümesinden alınan 30 gerçek görüntü birleştirilmiş ve bütün veriler karıştırılmıştır. Karıştırılan görüntüler bir anket sistemi vasıtasıyla kullanıcılara Görsel Turing testi olarak sunulmuştur. Uzmanlara anket yapısında sunulan testin çıktısı Şekil 3.7'de görülebilir. Yapılan çalışmada hazırlanan 150 adet görüntü iki KBB uzmanı tarafından gerçek/yapay olarak oylanmıştır.

Uzmanlar tarafından yapılan kontrollere ek olarak çıktı kalitesi ve çıktı çeşitliliği performanslarını değerlendirmek için Fréchet Başlangıç Mesafesi ve Öğrenilmiş Algısal Görüntü Yama Benzerliği skorları alınmıştır.

İmgelerde gerçek/yapay ayrımı

Aşağıdaki 150 imge, farklı yöntemlerle üretilen imgeler ile gerçek imgelerden oluşmaktadır. İmgeler karışık olarak sunulmaktadır. Yapılan çalışmada üretilen resimlerin gerçekçiliği ölçülecektir.

İmge gerçek mi yapay mı? *



☐ Gerçek

☐ Yapay

Şekil 3.7 Uzmanlara sunulan Görsel Turing testinin yapısı

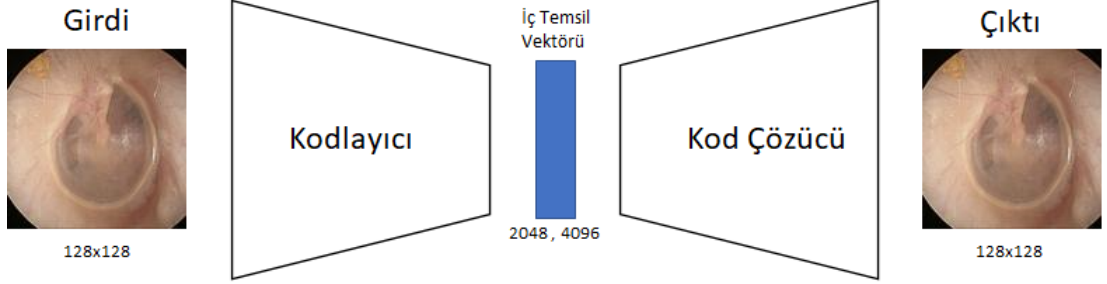
3.5 Özellik Çıkarma

Makine öğrenmesi tabanlı sistemlerde eğitime başlanmadan önce verinin karakteristiğini daha iyi yakalayabilmek amacıyla veriden özellik çıkarımı yapılmaktadır. Çıkarılan özellikler eğitim amacıyla makine öğrenmesi sistemlerine beslenmektedir. Tasarlanan sistemde kulak zarı görüntülerinin özelliklerinin çıkartılması ve performanslarının kıyaslanması amacıyla bir otokodlayıcı yapay sinir ağı eğitilmiştir ve ImageNet veri kümesi üzerinde eğitilen bir CNN üzerinde öğrenme aktarması uygulanmıştır.

Otokodlayıcı yapay sinir ağının kodlayıcı parçasından alınan sıkıştırılmış temsil vektörleri ve CNN ağının sınıflandırma katmanından önceki evrişim bloğu çıktısı patoloji tespiti sisteminde kullanılmıştır. Bu yaklaşımın seçilmesinin sebebi kulak zarı görüntüleri üzerinde elle yapılacak özellik çıkarma işleminin alan bilgisi gerektirmesi ve zaman alıcı bir yöntem olmasıdır.

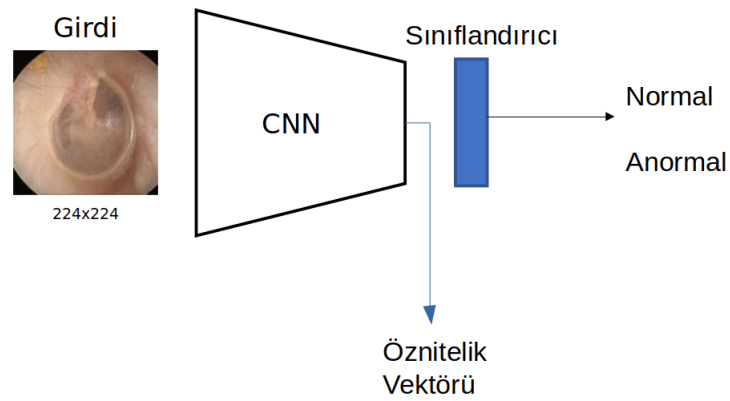
Otokodlayıcı eğilirken girdi görüntüsünü sıkıştırdıktan sonra en az kayıpla aynı görüntünün tekrar oluşturulması amaçlanmıştır. Böylece eğitim aşamasında sisteme verilen görüntülerden görüntülere ait gizli temsil uzayı öğrenilmiştir. Eğitimin ardından bu mimariye girdi olarak verilen görüntüler modelin ilk parçası olan kodlayıcıya verilerek temsil vektörleri kaydedilmiştir. Özellik çıkarımı için farklı

otokodlayıcı mimarileri kullanılabilir. Bu çalışmada kullanılan otokodlayıcı mimarisi Şekil 3.8’de görülebilir. Şekilde gösterilen gizli temsil vektörlerinin farklı boyutlardaki temsil kabiliyetlerini ölçmek amacıyla tamamen bağlı katman (fully connected layer) boyutu değiştirilerek farklı boyutlardaki vektörlerin patolojik görüntüleri tespitinde aldıkları başarı farkları incelenmiştir.



Şekil 3.8 Kullanılan otokodlayıcı yapısı

Öğrenme aktarımı ile eğitilen CNN’in sınıflandırıcı katmanı sistemden kopartılarak geriye kalan mimari öznetelik çıkarma sistemi olarak kullanılmıştır. Evrişim tabanlı sınıflandırıcı mimarisi olarak daha fazla parametre içeren modellere kıyasla benzer performansı daha az parametre ile yakalayabilen EfficientNet mimarisi tercih edilmiştir. Buna ek olarak EfficientNet ile yapılan öğrenme aktarımı çalışmalarında rakiplerine kıyasla daha yüksek performans alındığı gözlemlenmiştir [34]. Sistemin eğitimi aşamasında etiketlenmiş veri kullanılarak öğrenme aktarımı gerçekleştirilmiştir. Eğitim sonucunda alınan model ile Şekil 3.9’da görüldüğü şekilde öznetelik çıkarımı sağlanmıştır.



Şekil 3.9 CNN ile öznetelik çıkarımı

3.6 Patolojik Kulak Zarı Görüntüsü Tespiti

Tasarlanan sistemde patolojik kulak zarı görüntülerinin tespit edilebilmesi için izolasyon tabanlı karar ağacı ormanı algoritmasından yararlanılmıştır. Kullanılan kulak zarı görüntüleri veri kümesinde normal, orta kulak iltihabı, kulak zarı yırtıkları gibi sınıf bilgisi bulunmadığından dolayı gözetimsiz eğitim yapılan IForest algoritmasından yararlanılmıştır.

Veri kümesi üzerinde eğitilmiş olan yapay sinir ağlarından alınan öznitelik vektörleri IForest algoritmasının eğitiminde kullanılmıştır. GAN'lar yardımı ile üretilen yapay görüntüler aynı ağlardan geçirilerek öznitelik vektörleri elde edilmiştir. Elde edilen yapay görüntü vektörleri ek veri olarak patoloji tespiti yapan sistemi beslemiştir. Bu şekilde sisteme verilen ekstra verinin performansa olan etkisi ölçülmüştür. Orijinal veri kümesinin öznitelik vektörlerine yapay görüntülerden alınan öznitelik vektörleri farklı miktarlarda eklenmiş ve artan verinin sistem başarısına etkisi gözlenmiştir. Sistemin patolojik görüntü tespit performansını ölçmek için kesinlik, duyarlılık ve F1 skorlamaları kullanılmıştır.

4 DENEYSEL SONUÇLAR

Bu bölümde, deneylerin gerçekleştirildiği ortama dair teknik detaylar ve deney sonuçlarına yer verilmiştir. Deneyler kapsamında öncelikle farklı GAN mimarileri ile yapay kulak zarı görüntüleri üretilmiştir. Üretilen görüntüler veri setinden örneklenen gerçek görüntüler ile karıştırılıp Görsel Turing Testi şeklinde KBB uzmanlarının değerlendirmesine sunulmuştur. İnsan faktörü ile yapılan değerlendirmelere ek olarak yapay görüntülerin kalitesini değerlendirecek otomatik yöntemler kullanılmıştır.

Ardından etiketsiz kulak zarı görüntüleri verisi kullanılarak bir otokodlayıcı derin öğrenme ağı eğitimi yapılmıştır. Buna paralel olarak etiket bulunan eğitim verisi ile ön eğitilmiş bir CNN sınıflandırıcı modelinde öğrenme aktarması gerçekleştirilmiştir. Eğitilen otokodlayıcının kodlayıcı ağından görüntülerin gizli vektörleri ve CNN'e verilen etiketsiz görüntülerin öznelik çıkarımı katmanından alınan vektörler toplanmıştır.

Toplanan iki çeşit vektör izolasyon tabanlı karar ağacı (Isolation Forest) (IF) algoritmasına verilerek patolojik görüntü tespiti sisteminin eğitimi tamamlanmıştır. İki tip öznelik vektörünün IF algoritması ile aldığı başarımlar sınırlı sayıda etiketlenmiş veri ile ölçülmüştür. Test verisi hazırlanırken etiketlenen veri kümesindeki görüntüler sınıf dağılımı gözetilerek eğitim ve test verisi olarak bölünmüştür. Etiketlenen bütün verilerin gerçekleştirilen deneylerde test verisi olarak kullanılmamasının sebebi verilerin bir kısmının gözetimli öğrenme ile eğitilen CNN tabanlı sınıflandırma sisteminin eğitiminde kullanılmasıdır. Gözetimli öğrenme için ayrılan eğitim verisi izole edildiğinden bu verinin diğer deneylerde kullanılan test verisine sızması engellenmiştir.

Yapay görüntüler ile veri seti büyütmenin patoloji görüntü tespiti sistemi başarımına etkisini ölçmek için uzman doktorlar tarafından seçilen GAN yöntemi ile üretilen görüntüler kullanılmıştır. Farklı miktarlarda yapay görüntünün öznelik vektörleri eğitim verisine katılarak sistem başarımının değişimi deneylerle gözlenmiştir.

4.1 Deney Ortamı

Yapay kulak zarı görüntülerinin üretiminde kullanılan GAN'lar ve görüntü öznetelik vektörlerinin alınabilmesi adına eğitilen derin öğrenme ağları yüksek hesaplama gücüne ihtiyaç duymaktadır. Buna ek olarak yapay olarak artırılan verilerin de IF eğitiminde kullanılabilmesi için yüksek bellek kapasitesi ihtiyacı bulunmaktadır. Deneylerin gerçekleştirildiği donanım bilgileri Tablo 4.1'de verilmiştir.

Tablo 4.1 Çalışmada kullanılan sistemin özellikleri

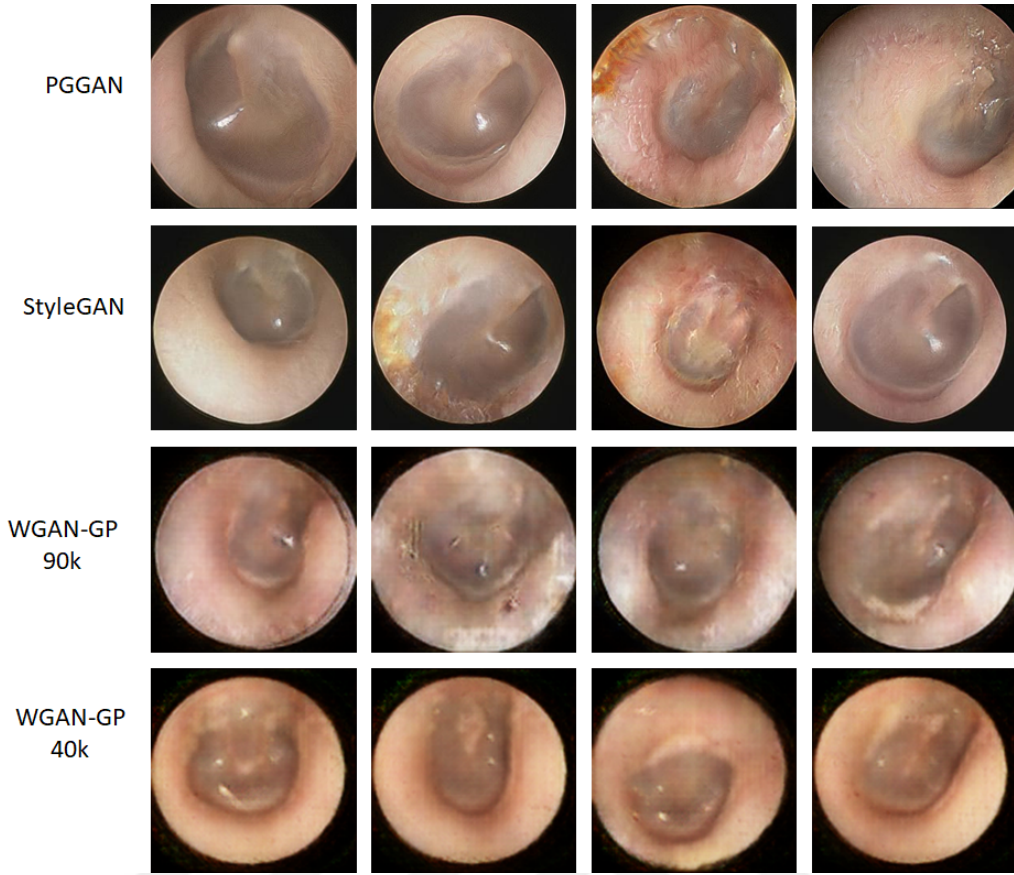
Sistem Özellikleri	
Çekirdek Sayısı	16
Hafıza (RAM)	64 GB
Disk	500 GB SSD
Grafik İşlemcisi (GPU)	NVIDIA RTX 2080 (x2)
İşletim Sistemi	Ubuntu 20.04

Yapılan çalışmada Pytorch[35], Tensorflow-Keras[36], scikit-learn[37] ve pyOD[38] makine öğrenmesi kütüphaneleri Python dili ile kullanılarak deneyler gerçekleştirilmiştir. Yapay veri üretimi modellerinde eğitimin tamamlanması bir hafta sürmektedir. Patoloji tespiti yapan modellerin eğitimi de model başına yaklaşık olarak beş dakika sürmektedir.

4.2 Görsel Turing Testi ile GAN Seçimi

Çalışma kapsamında kullanılan GAN yöntemlerinden ilki olan WGAN-GP mimarisi üzerinde yapılan deneylerde modelin eğitimi 120.000 iterasyon boyunca devam ettirilmiştir. Üreticiden alınan çıktıların kalitesi deneysel olarak izlenerek 40.000 ve 90.000 iterasyonlarında alınan çıktıların Görsel Turing Testinde kullanılmasına karar verilmiştir. StyleGAN modeli donanım kısıtlarından dolayı çözünürlük 256x256 olana kadar devam ettirilmiştir. PGGAN modeli, StyleGAN modelinden daha düşük parametre sayısına sahip olduğu için çözünürlük 512x512'ye ulaşana kadar devam ettirilmiştir. Eğitimi tamamlanan GAN'lardan alınan ve gözle kontrol edildiğinde kulak zarı yapısını koruyabilen yapay görüntüler Şekil 4.1'de verilmiştir.

Üretilen yapay görüntülerden kulak zarı yapısını koruyamayan yapay görüntüler de gözle incelenmiştir. Kötü üretim örneği olan bu yapay görüntüler Şekil 4.2'de verilmiştir. WGAN ağının başarısız örneklerinde görüntülerin renk uzaylarının dahi tam olarak oturmadağı görülmüştür. PGGAN ve StyleGAN ile üretilen yapay görüntülerin kenarlarındaki kahverengi tonunun baskın olması sistemlerin patolojik görüntüleri üretmede sorun yaşadığına işaret etmektedir.



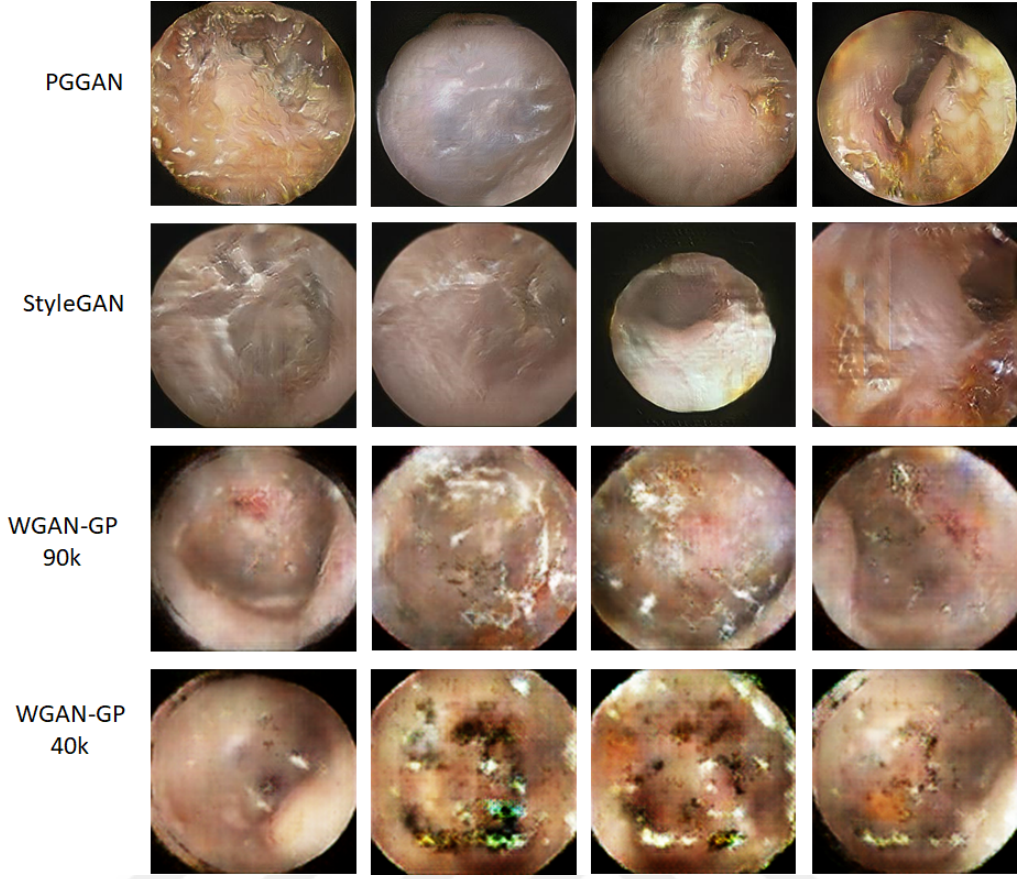
Şekil 4.1 Kulak zarı yapısının korunabildiği yapay çıktılar

KBB uzmanlarına sunulan Görsel Turing Testi (GTT) ile uzmanların yaptığı seçimler sonucunda en gerçekçi görüntü üretme kapasitesine sahip GAN mimarisi belirlenmiştir. Tablo 4.2’de yapılan Görsel Turing Testi çalışmasının sonuçları görülebilir.

Tablo 4.2 Görsel Turing testi sonuçları

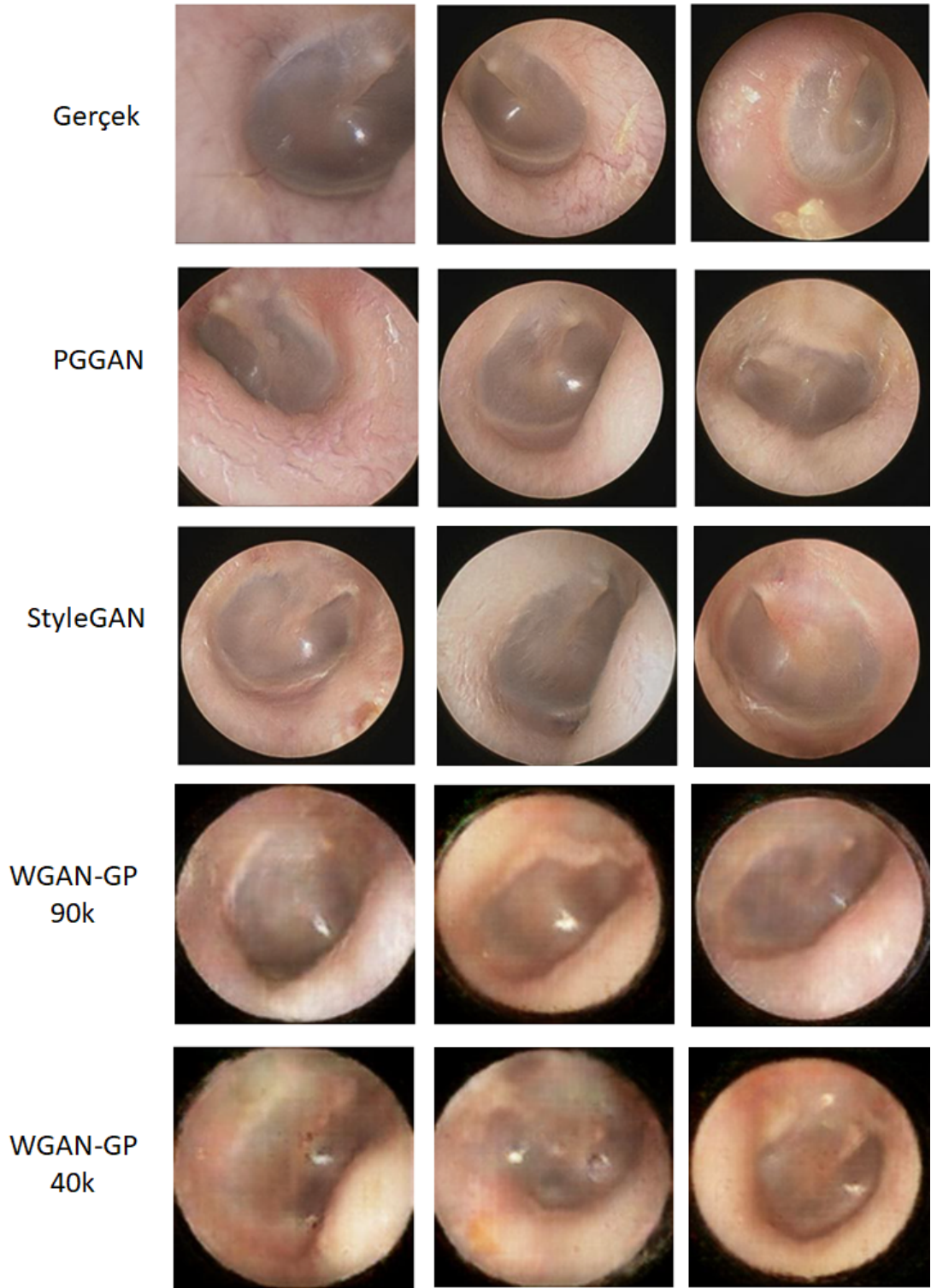
	Gerçek İmge	Yapay İmge	Gerçek Seçim Oranı(%)
Gerçek Veri	28	2	93
PGGAN	21	9	70
StyleGAN	18	12	60
WGAN-GP 40k	4	26	13
WGAN-GP 90k	3	27	10

Doktorların gerçek görüntüleri %94 oranda gerçek olarak işaretlemeleri ve yapay görüntüleri de en yüksek başarımda %70 oranda gerçek olarak sınıflandırmaları yapay üretilmiş görüntülerin gerçek veriye alternatif olarak kullanılabileceğini göstermektedir. Yapılan GTT’de oylanan gerçek kulak zarı görüntülerinden ve GAN mimarileri ile üretilen yapay kulak zarı görüntülerinden gerçek olarak seçilen görüntü örnekleri Şekil 4.3’te verilmiştir.

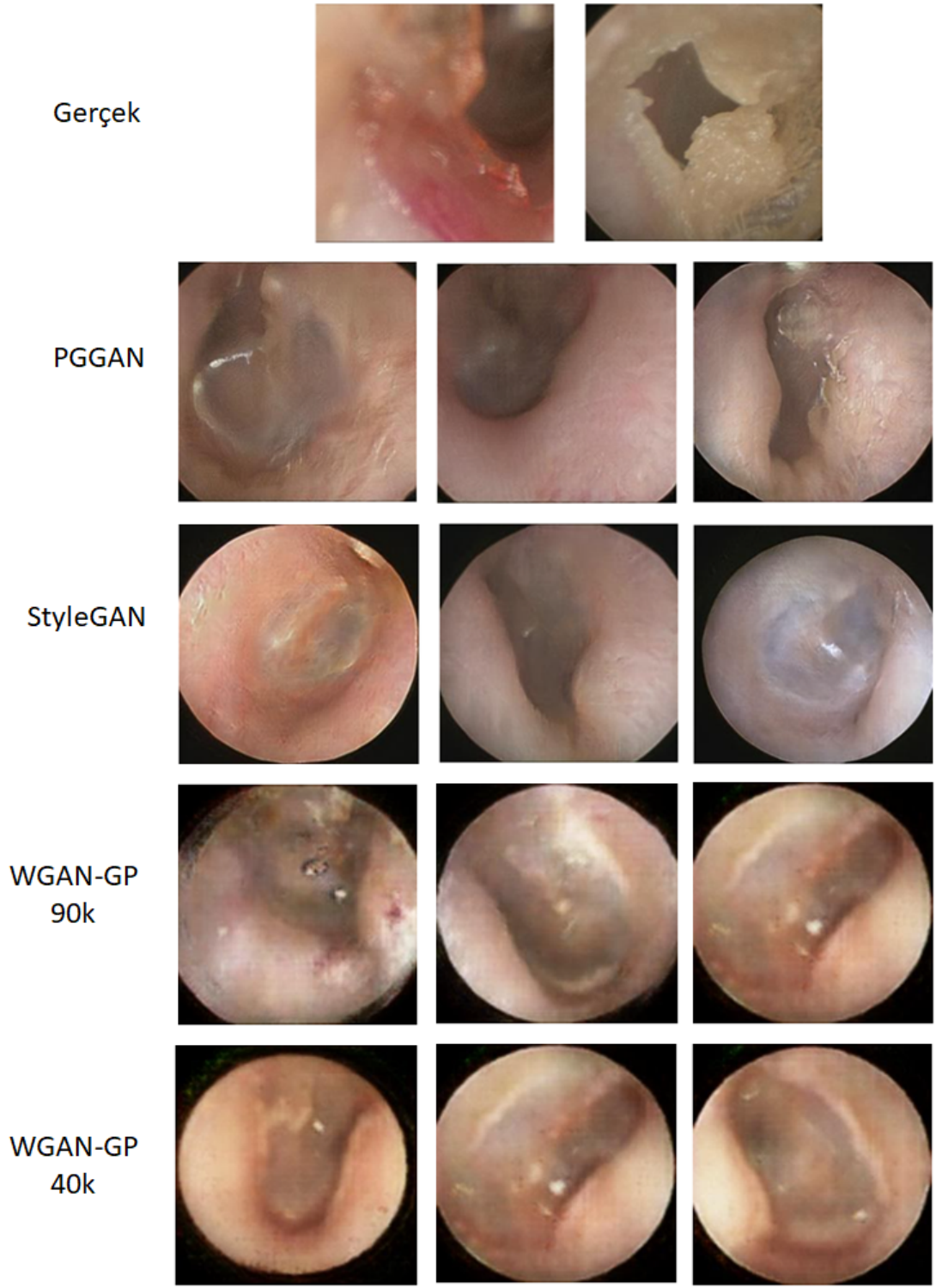


Şekil 4.2 Kulak zarı yapısının korunamadığı yapay çıktılar

Aynı şekilde doktorlar tarafından başarılı bir şekilde yapay olarak ayırt edilen görüntüler Şekil 4.4'te verilmiştir. Yapay olarak üretilen kulak zarı görüntülerinden alınan oylamalarda kulak zarı yapısını genel olarak koruyabilen görüntülerin gerçek olarak algılanabildiği anlaşılmaktadır. Nihai amacı eğitildikleri veri setinin gizli uzaydaki dağılımını öğrenmeye çalışmak olan GAN'larda PGGAN ve StyleGAN özelinde başarıya ulaşıldığı görülmüştür.



Şekil 4.3 Görsel Turing Testinde gerçek olarak oylanan görüntüler



Şekil 4.4 Görsel Turing Testinde yapay olarak oylanan görüntüler

4.3 Otomatize Yöntemler ile GAN Performansının Ölçülmesi

Eğitimi tamamlanan GAN'ların çıktı kalitesini ve yüksek varyasyonda çıktı üretebilme performanslarını ölçmek için FID ve LPIPS skorları alınmıştır. İlk olarak üreticilerden eğitim yapılan veri seti kadar yapay görüntü üretilerek eğitim veri seti ile beraber FID skorlaması sistemine verilmiştir. Alınan sonuçlar Tablo 4.3'te verilmiştir. Daha düşük skorun daha kaliteli çıktıya işaret ettiği bu performans ölçütünde GTT ile tutarlı sonuçlar alındığı görülmüştür. En yüksek kalitede görüntüleri PGGAN üretirken StyleGAN mimarisi de yakın bir performans göstermiştir.

Tablo 4.3 Eğitilen GAN'ların çıktı kalitesini belirten FID skorları

Model	Skor
WGAN-GP 40k	254.832
WGAN-GP 90K	169.466
PGGAN	40.428
StyleGAN	65.537

Eğitilen GAN'ların ürettiği görüntülerin çeşitliliğinin performansını ölçmek amacıyla üretilen yapay görüntüler Öğrenilmiş Algısal Görüntü Yama Benzerliği (LPIPS) modülüne verilmiştir. Farklı mimarilerden alınan sonuçlar Tablo 4.4'te verilmiştir. Daha yüksek skorun daha çok çeşitli yapay görüntü üretme kapasitesine işaret ettiği bu skorlama ölçütünde alınan sonuçların GTT ile tutarlı olduğu görülmüştür.

Tablo 4.4 Eğitilen GAN'ların çeşitli çıktı üretme kapasitesini belirten LPIPS skorları

Model	Ortalama Skor	Toplam Skor
WGAN-GP 40k	0.31624 +/- 0.00105	3189.91795
WGAN-GP 90K	0.33214 +/- 0.00125	3350.25514
PGGAN	0.52539 +/- 0.00098	5299.60464
StyleGAN	0.47891 +/- 0.00114	4830.80747

Gerçeklenen GTT ve koşulan otomatik skorlama ölçütleri doğrultusunda veri artırımı yöntemi olarak en çeşitli ve en yüksek kalitede görüntü üretme kapasitesine sahip mimarinin PGGAN olduğu tespit edilmiştir.

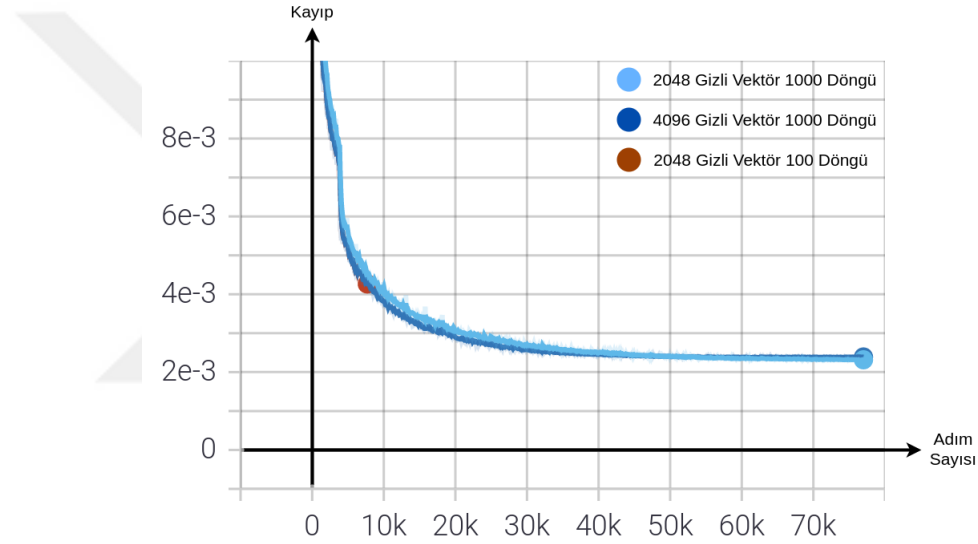
4.4 Öznitelik Çıkarım Modüllerinin Eğitimi

Patolojik görüntü tespiti sisteminde kullanılması için iki farklı öznitelik çıkarımı yaklaşımı denenmiştir. İlk olarak etiketsiz verilerin tamamı üzerinde evrişimli bir otokodlayıcı eğitimi gerçekleştirilmiş ve bu verilerden çıkartılan görüntü öznitelik vektörleri sonradan kullanılmak üzere kaydedilmiştir.

İkinci öznitelik vektörünün çıkartılabilmesi için ön eğitilmiş bir CNN sınıflandırıcı alınarak ikili sınıflandırma problemine uyarlanmıştır. Bu uyarlama kapsamında yapılan öğrenme aktarımı ile mimarinin kulak zarı görüntü vektörlerini de çıkarabilmesi sağlanmıştır.

4.4.1 Otokodlayıcı Eğitimi

Toplanan kulak zarı görüntülerinden özellik çıkarılabilmesi için bir evrişimli otokodlayıcı eğitilmiştir. Veri setindeki görüntüler 128x128 boyutuna düşürüldükten sonra eğitim gerçekleştirilmiştir. Farklı gizli katman boyutlarının patolojik kulak zarı görüntüsü tespitine etkisini gözlemlemek amacıyla otokodlayıcının darboğaz katmanının boyutu 2048 ve 4096 olacak şekilde iki farklı model eğitilmiştir.

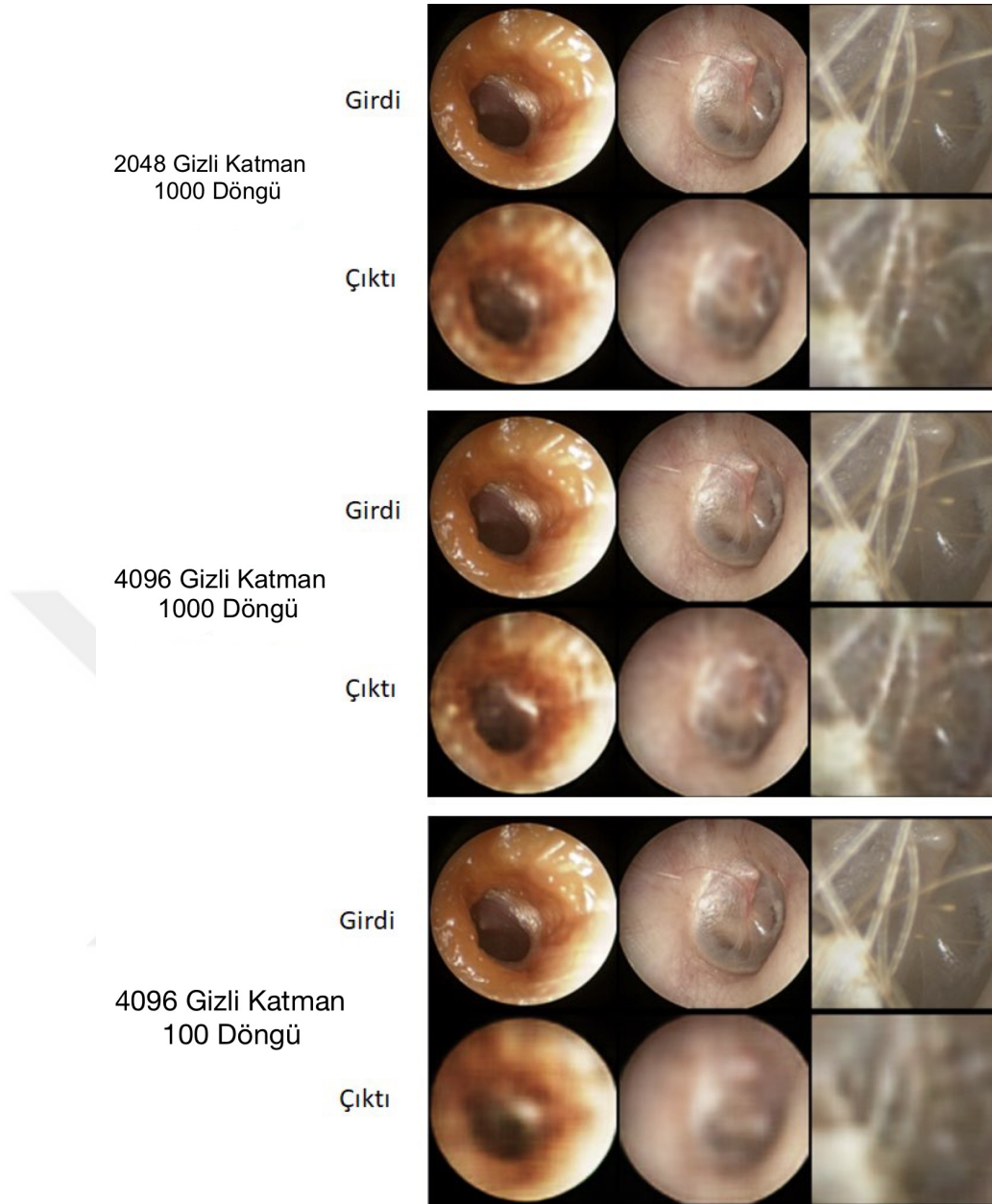


Şekil 4.5 Otokodlayıcının eğitim boyunca doğrulama verisinde aldığı kayıp grafiği

Eğitim aşamasında 10088 kulak zarı görüntüsü %98 eğitim %2 test verisi olacak şekilde rastgele bölünmüştür. Eğitim 1000 döngü boyunca devam ettirilmiş ve her döngüde (epoch) alınan doğrulama hatası kaydedilmiştir. Kaydedilen hatalar Şekil 4.5'te verilmiştir.

Modelin eğitimi boyunca 2048 ve 4096 gizli katman boyutunda benzer kayıp grafiğini izlediği gözlemlenmiştir. Eğitim erken durdurulduğunda patolojik görüntü sınıflandırma performansını görmek amacıyla 4096 gizli katman boyutundaki model 100 döngü ile bir kere daha eğitilmiş ve ağırlıkları kaydedilmiştir.

Modelin eğitimi boyunca girdi olarak aldığı kulak zarı görüntülerini tekrar yaratma performansı da incelenmiştir. Şekil 4.6'da eğitimi tamamlanan 3 model için girdi olarak verilen görüntü ve yeniden oluşturulan görüntü verilmiştir. Eğitimin sonunda



Şekil 4.6 Otokodlayıcının eğitim sonunda aldığı görüntüyü yeniden oluşturması

1000'er döngü boyunca eğitilen iki farklı gizli katman boyuna sahip modelin de girdi görüntüsünü ayırt edici bir şekilde oluşturabildiği görülmüştür.

Daha az döngü sayısı ile eğitilen modelde çıktı görüntülerinin girdi görüntülerine birebir benzememesinin sebebinin gizli uzayın boyutunun girdi görüntüsünden daha düşük olması sebebiyle kaybedilen bilgiden kaynaklı olduğu düşünülmektedir. Eğitim sonucunda elde edilen modeller görüntülerin öznitelik vektörlerinin çıkartılması için kullanılacaktır. Çıkartılan vektörler patolojik görüntü tespit sisteminde kullanılacaktır.

4.4.2 Ön Eğitimli CNN Mimarisinde Öğrenme Aktarımı

Toplanan ve etiketlenen kulak zarı görüntüleri kullanılarak bir CNN tabanlı ikili sınıflandırma modeli eğitilmiştir. Ön eğitimli model olarak EfficientNet ailesinden EfficientNet-B2 ,B3, B4, B5, B6 modelleri seçilmiştir. Birden fazla mimari üzerinde yapılanlar eğitimler sonucunda farklı derinliklerdeki ağlardan alınan öznitelik vektörlerinin patoloji tespiti başarımına etkisi gözlenmiştir. Eğitim bir uzman tarafından etiketlenen 645 adet görüntü kullanılarak gerçekleştirilmiştir.

Yapay sinir ağı hazırlanırken ilk olarak modelde tanımlı olan ve ImageNet verisindeki 1000 sınıfın tespiti için kullanılan en üst sınıflandırıcı katmanlar modelden kopartılmıştır. Bu katmanın yerine ikili sınıflandırmaya uygun yeni bir sınıflandırıcı katman (Dense Layer) eklenmiştir. Ara katmanlardan alınan çıktılar yeni eklenmiş olan sınıflandırıcı katmana bağlanmadan önce mutlak ortalama örnekleme (Global Average Pooling) işleminden geçirilerek boyutları düşürülmüştür. Sistemin eğitim verisini ezberlemesini engellemek ve modelin daha iyi genelleşebilmesi için Global Average Pooling katmanının ardına bir Seyreltme (Dropout) katmanı eklenmiştir. Son durumda eğitimi yapılacak sistemde EfficientNet-B2 kullanıldığında ortaya çıkan katman yapısı Tablo 4.5'teki gibidir.

Tablo 4.5 CNN Mimarisi

Katman Adı	Çıktı Şekli (Yığın Miktarı, En, Boy, Kanal)	Parametre Sayısı
Girdi Katmanı	(32, 224, 224, 3)	0
EfficientNet-B2 Tabanlı Öznitelik Çıkarım Katmanı	(32, 7, 7, 1048)	7768569
Mutlak Ortalama Örnekleme	(32, 1408)	0
Seyreltme	(32, 1408)	0
Sınıflandırıcı Katman	(32, 1)	1409

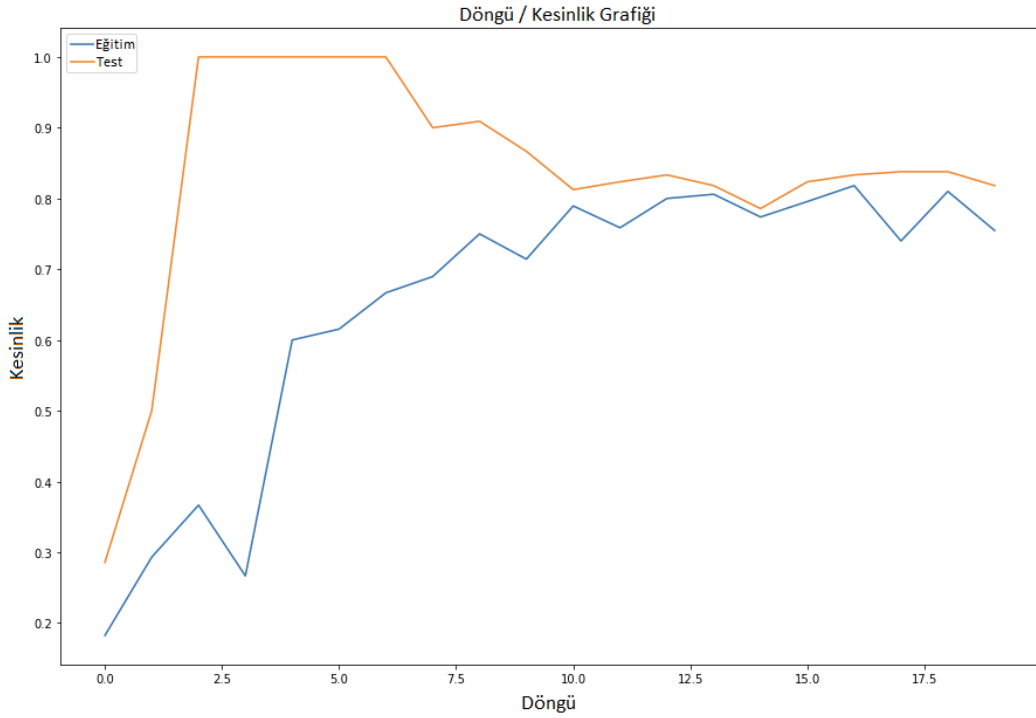
Farklı EfficientNet mimarileri kullanıldığında alınan öznitelik vektörlerinin boyutu Tablo 4.6'da görülebilir. Ağın derinliği arttıkça benzer oranda artan katman genişliği nedeni ile vektör boyutları da artmaktadır.

Tablo 4.6 Farklı EfficientNet mimarilerinden alınan öznitelik vektör boyutları

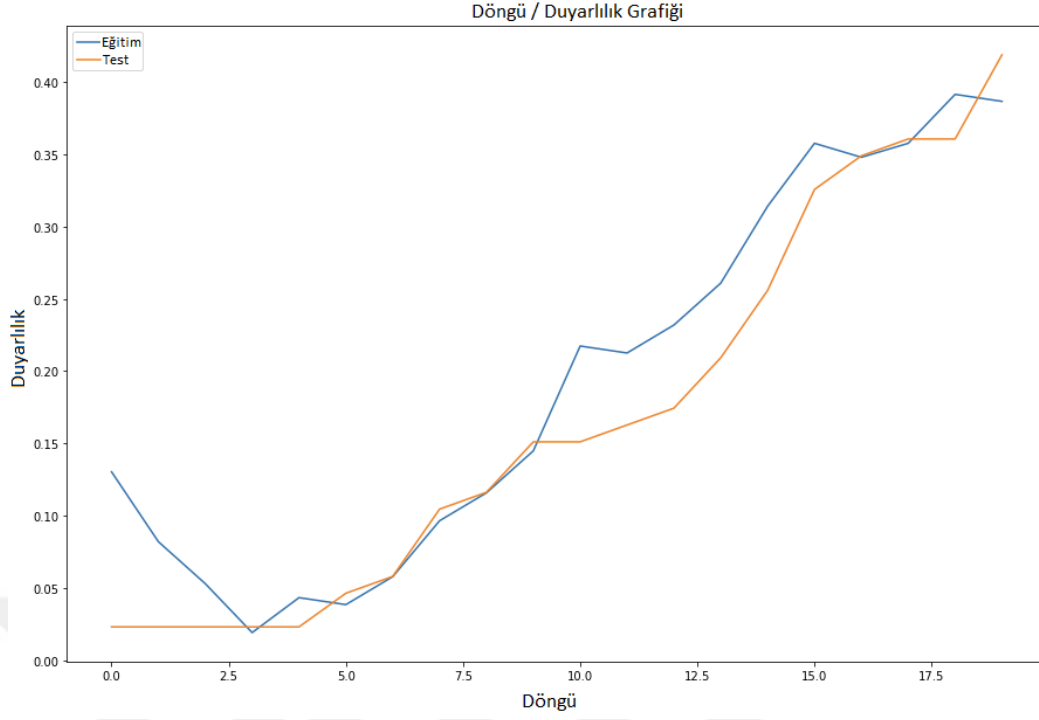
Mimari	Öznitelik Boyutu
EfficientNet-B2	1408
EfficientNet-B3	1536
EfficientNet-B4	1792
EfficientNet-B5	2048
EfficientNet-B5	2304

Koşulan ilk deneyde, evrişim bloklarının ağırlıkları dondurularak sadece son sınıflandırıcı katmanı eğitime açılan mimari 10^{-4} öğrenme katsayısı ile 100 döngü boyunca eğitilmiştir. Kayıp fonksiyonu olarak ikili çapraz entropi kullanılmıştır. Eğitimin optimizasyonu için Adam optimizasyonundan yararlanılmıştır. Sistemin test verisinde duyarlılık skorunun yetersiz olduğu görüldükten sonra ikinci deney koşulmuştur.

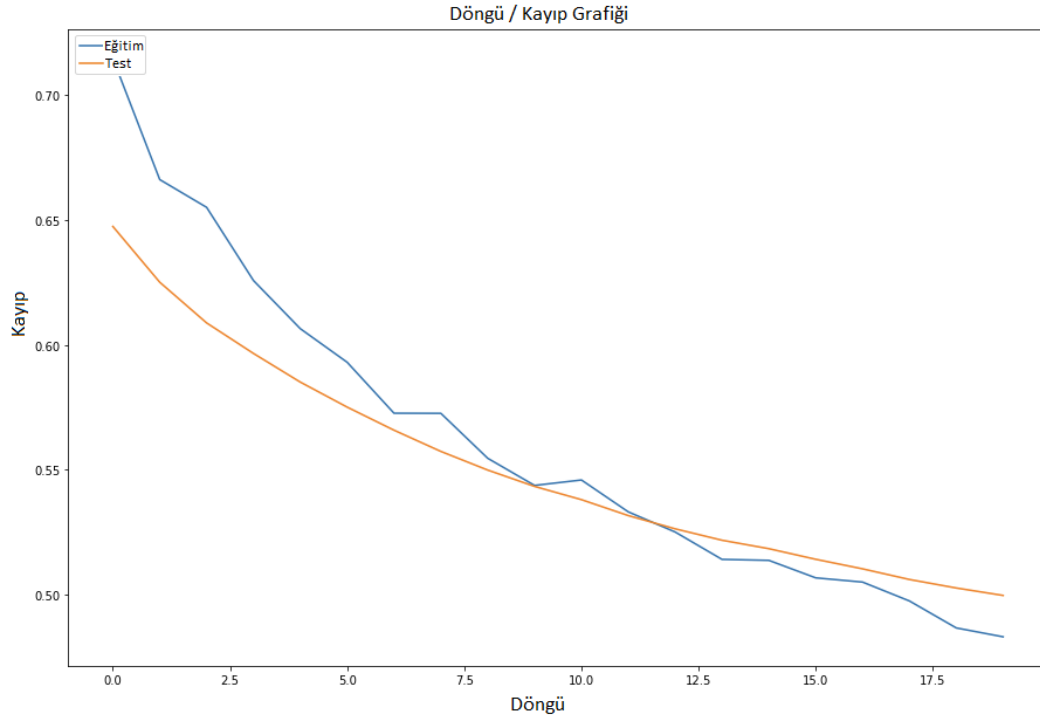
İkinci deney olarak gerçekleştirilen öğrenme aktarımı iki aşamalı gerçekleşmiştir. Her iki aşamada test verisinde en yüksek duyarlılık başarımları gösteren model ağırlıkları kaydedilmiştir. İlk olarak sistemin girdiden Global Average Pooling katmanına kadar olan kısmının ağırlıkları dondurulmuş ve 10^{-4} öğrenme katsayısı ile 20 döngü boyunca eğitilmiştir. Kayıp fonksiyonu olarak ikili çapraz entropi kullanılmıştır. Eğitimin optimizasyonu için Adam optimizasyonundan yararlanılmıştır. Eğitimi yapılan bütün EfficientNet çeşitleri için aynı yaklaşım kullanılmıştır. EfficientNet-B2 mimarisi kullanılarak gerçekleştirilen ilk eğitim aşamasında alınan kesinlik skoru, duyarlılık skoru ve kayıp değerleri Şekil 4.7, 4.8 ve 4.9’da görülebilir.



Şekil 4.7 EfficientNet-B2 ilk eğitim kesinlik grafiği



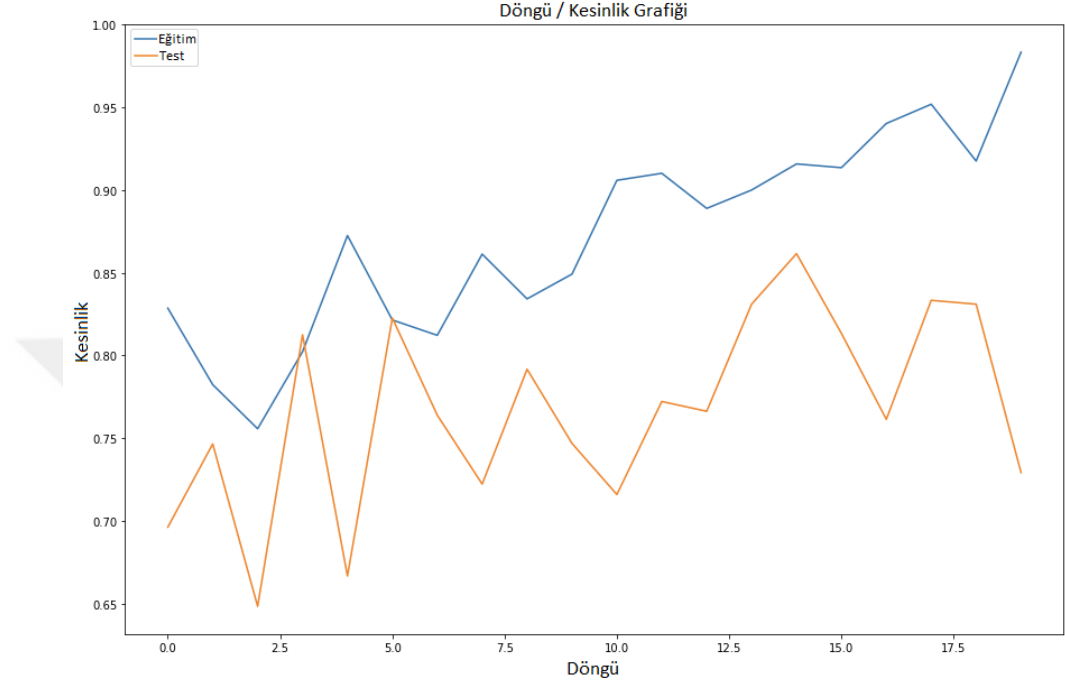
Şekil 4.8 EfficientNet-B2 ilk eğitim duyarlılık grafiği



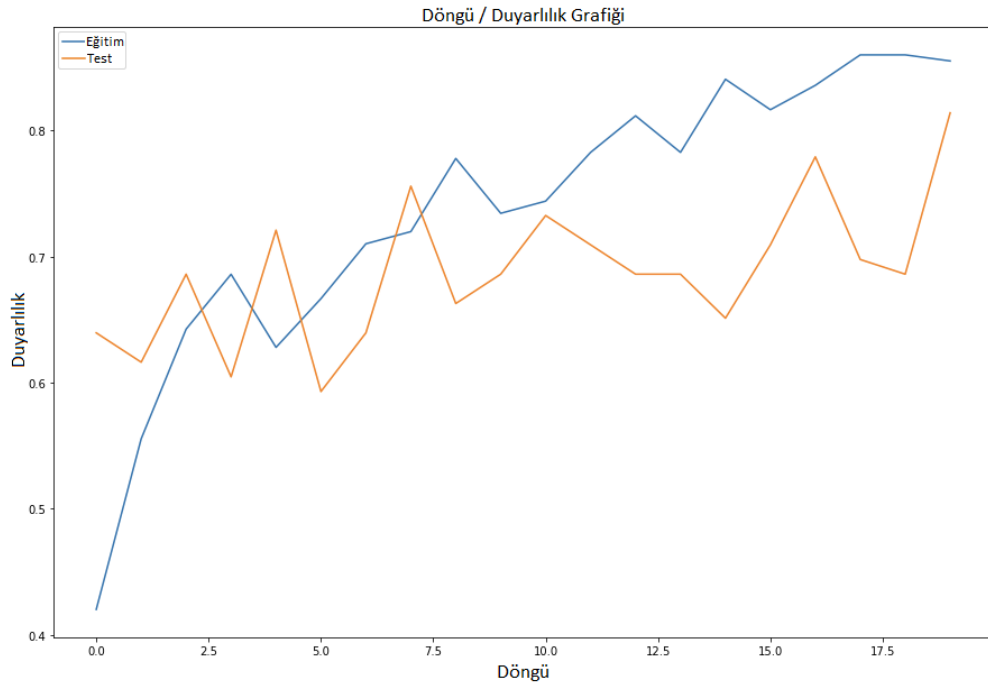
Şekil 4.9 EfficientNet-B2 ilk eğitim kayıp grafiği

İkinci aşamada sisteme ince ayar yapılması amacıyla modelin bütün katmanları eğitime açılarak öğrenme katsayısı 10^{-5} olacak şekilde değiştirilmiştir. Kayıp fonksiyonu ve optimizasyon yaklaşımı önceki adım ile aynı olacak şekilde eğitim gerçekleştirilmiştir. İlk aşamada en yüksek duyarlılığa sahip model alınarak eğitim 20

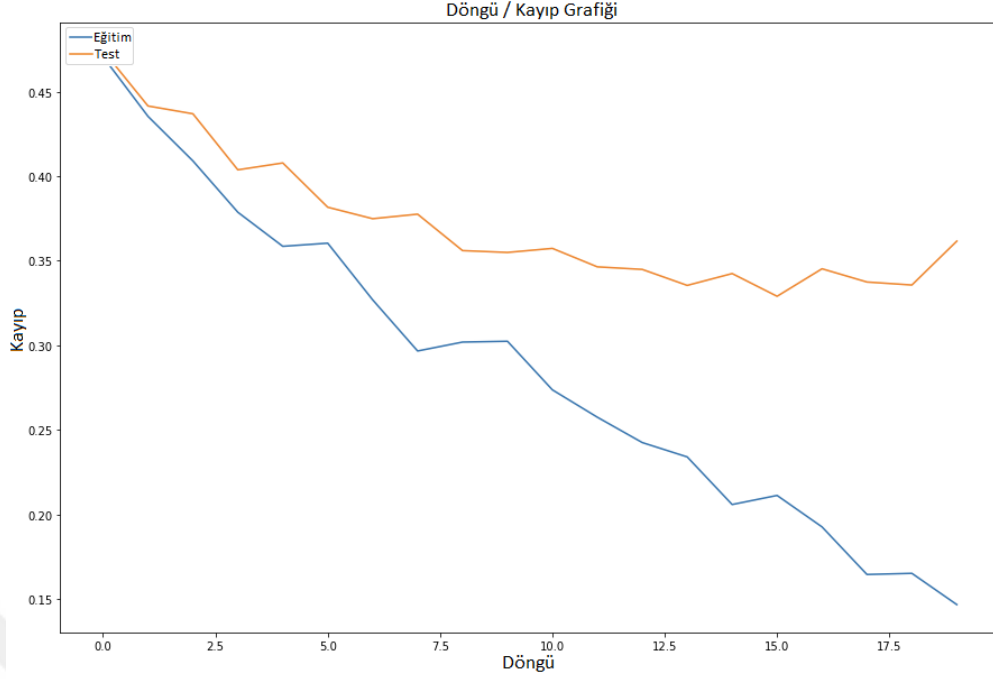
döngü boyunca sürdürülmüştür ve en yüksek duyarlılığa sahip model seçilmiştir. Eğitimi yapılan bütün EfficientNet çeşitleri için aynı yaklaşım kullanılmıştır. EfficientNet-B2 mimarisi kullanılarak gerçekleştirilen ikinci eğitim aşamasında alınan kesinlik skoru, duyarlılık skoru ve kayıp değerleri Şekil 4.10, 4.11 ve 4.12’de görülebilir.



Şekil 4.10 EfficientNet-B2 ikinci eğitim kesinlik grafiği



Şekil 4.11 EfficientNet-B2 ikinci eğitim duyarlılık grafiği



Şekil 4.12 EfficientNet-B2 ikinci eğitim kayıp grafiği

Sistemin eğitimi yapılırken ikinci eğitim aşamasında sistemin kesinlik başarımının düştüğü adımlarda duyarlılık başarımının arttığı gözlemlenmiştir. Böyle durumlarda sistem daha fazla yanlış alarm üretecektir ancak daha yüksek duyarlılık başarısı ile patolojik görüntülerin yanlış sınıflandırılmasının önüne geçilmiş olacaktır. Eğitimi tamamlanan EfficientNet-B2 tabanlı sistemin test verisi performansı Tablo 4.7'de görülebilir.

Tablo 4.7 EfficientNet-B2 tabanlı modelin farklı eğitim yaklaşımlarındaki test başarımları

Deney No	Kesinlik	Duyarlılık	F1
1	0.8275	0.2791	0.4173
2	0.7291	0.8139	0.7692

Daha düşük parametreye sahip mimaride koşulan deneylerin ardından ikinci deney yaklaşımı ile EfficientNet-B3, B4, B5, B6 mimarilerinin eğitimi gerçekleştirilmiştir. Yapılan deneyler sonunda her mimari için alınan test başarımları Tablo 4.8'de görülebilir.

Yapılan deneyler sonucunda gözetimli eğitim ile alınan sonuçların dengesiz veri dağılımına sahip kulak zarı görüntülerinde patolojik görüntülerin tespit edilebilmesi için yeterli bir başlangıç noktası sunduğu düşünülmektedir. Eğitimi tamamlanan ve farklı boyutlarda vektör üreten ağların öznetelik vektörlerinin başarıma etkisi gözetimsiz eğitimi yapılan izolasyon tabanlı karar ağacı ormanı ile test edilecektir.

Tablo 4.8 Eğitimi tamamlanan gözetimli sınıflandırıcı başarımları

Mimari	Kesinlik	Duyarlılık	F1
EfficientNet-B2	0.7291	0.8139	0.7692
EfficientNet-B3	0.722	0.848	0.78
EfficientNet-B4	0.708	0.848	0.772
EfficientNet-B5	0.683	0.802	0.732
EfficientNet-B6	0.755	0.825	0.788

4.5 İzolasyon Tabanlı Karar Ağacı Ormanı ile Patolojik Görüntülerin Tespiti

İzolasyon tabanlı karar ağacı (IForest) modelinde eğitim için kullanılacak öznitelik vektörleri eğitilen otokodlayıcı ve sınıflandırıcı modelden alınarak farklı deneyler olacak şekilde koşulmuştur. Farklı modellerden alınan farklı boyuttaki öznitelik vektörleri ile test performansı ölçülmüştür. Eğitim gerçekleşirken verilerin normal/anormal olarak kararlaştırılmasını sağlayan eşik (threshold) değerini belirleyen kirlilik oranı (contamination) 0.3 olarak belirlenmiştir. Seçilen değer etiketlenmiş kısıtlı sayıdaki verinin dağılımından yola çıkılarak hesaplanmıştır. Koşulan bütün deneylerde eğitim 5 kere koşulmuş ve en yüksek duyarlılık skoru veren model saklanmıştır.

İlk olarak IForest modeline ek veri sağlamadan sadece etiketsiz ve etiketli eğitim veri kümesinin vektörleri beslenerek alınan sonuçlar incelenmiştir. Alınan başarımlar Tablo 4.9’da görülebilir.

Tablo 4.9 Farklı modellerden çıkartılan öznitelik vektörleri kullanılarak alınan patolojik görüntü tespiti başarıları

Model	Gizli Vektör Boyutu	Döngü	Kesinlik	Duyarlılık	F1
Otokodlayıcı	2048	1000	0.356	0.488	0.412
Otokodlayıcı	4096	100	0.345	0.465	0.396
Otokodlayıcı	4096	1000	0.323	0.465	0.381
EfficientNet-B2	1408	X	0.576	0.616	0.596
EfficientNet-B3	1536	X	0.653	0.721	0.685
EfficientNet-B4	1792	X	0.704	0.802	0.750
EfficientNet-B5	2048	X	0.698	0.698	0.698
EfficientNet-B6	2304	X	0.611	0.674	0.641

Yapılan deney sonucunda gözetimli olarak eğitilen EfficientNet tabanlı sınıflandırıcı modellerden alınan öznitelik vektörlerinin gözetimsiz olarak eğitilen otokodlayıcıdan daha yüksek olduğu görülmüştür. CNN tabanlı sınıflandırıcı modellerinin gözetimli olarak eğitilirken patolojik görüntüleri daha iyi ayırt etmesini sağlayan öznitelikleri öğrenebildiği anlaşılmaktadır. Gözetimli olarak eğitilen farklı Efficient modellerinde

artan vektör boyutuna göre değişen Duyarlılık başarımının EfficientNet-B2 modeli haricinde yakın bir aralıkta seyrettiği görülmüştür.

Bu duruma ek olarak Bölüm 4.4.2’de alınan test performansının daha iyi olduğu gözlemlenmiştir. Bu durumun IForest modelinin patolojik görüntüleri modellemek için daha yüksek veriye ihtiyaç duyması olasılığından kaynaklanabileceği düşünülmektedir.

4.5.1 Yapay Görüntülerin Patolojik Görüntü Tespiti Başarımına Etkisi

Koşulan ilk deneylerde IForest modeli ile patolojik görüntü tespiti sistemindeki başarıların sınırlı veri kümesinde eğitilen gözetimli EfficientNet tabanlı sınıflandırıcılardan daha düşük olduğu görülmüştür. Var olan duyarlılık başarımına yaklaşmak veya bu başarıyı geçmek için IForest modeline ek veri beslenmesine karar verilmiştir.

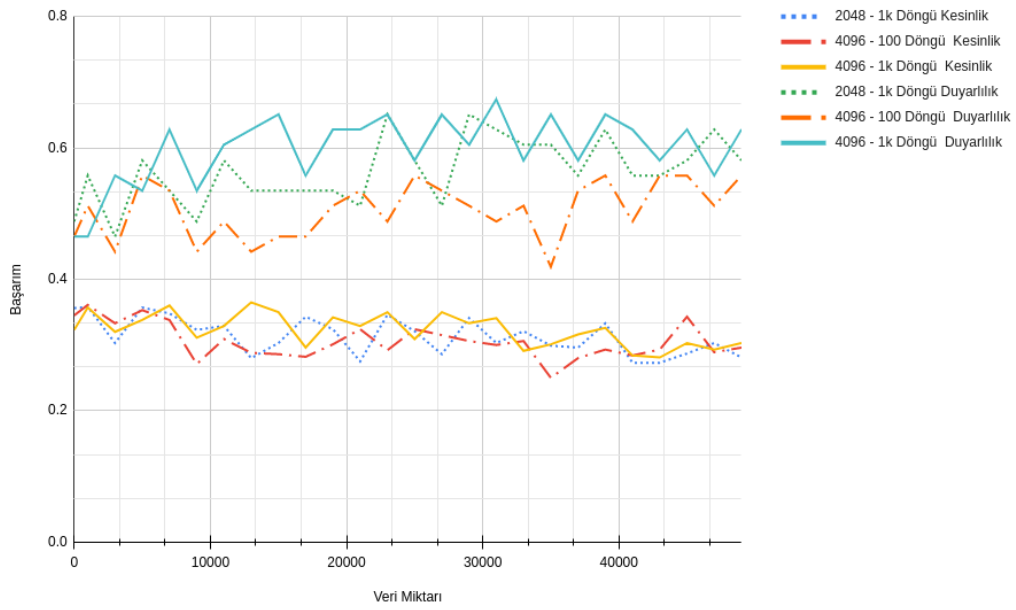
Doktorlar tarafından en gerçekçi bulunan PGGAN ağı ile farklı miktarlarda yapay görüntüler üretilmiş ve bu görüntülerin öznetelik vektörleri gerçek veriye eklenerek IForest modeline beslenmiştir. Eğitim yapılırken bütün seçenekler orijinal deneyle aynı tutulmuştur. Üretilen görüntü kümesinden daha düşük sayıdaki verilerin seçimi bu küme üzerinden rastgele yapılmıştır.



Şekil 4.13 Otokodlayıcı vektörü ile veri artırımı F1 skoru değişimi grafiği

İlk olarak performansı düşük olan otokodlayıcıya yapay görüntüler verilerek öznetelik

vektörleri alınmıştır. Toplamda üretilen 50000 görüntü orijinal eğitim verisine eklenerek deney gerçekleştirilmiştir. Farklı otokodlayıcılar ile gerçekleştirilen deney sonuçlarından alınan F1 skorları Şekil 4.13'te incelenebilir. Sistemde beslenen veri miktarı arttıkça F1 başarımının çok değişmediği görülmüştür. Buna ek olarak Şekil 4.14'te görülebileceği üzere sisteme eklenen veri miktarına göre kesinlik-duyarlılık değişimi de benzer davranışı göstermektedir. Fazla değişmeyen F1 skorunun aksine burada duyarlılık skorunun düşük miktarda da olsa arttığı, kesinlik skorunun ise düşüş eğilimine girdiği görülmüştür. Sonuç olarak sisteme veri eklendikçe patolojik görüntülerin daha doğru bir şekilde ayırt edilebildiği ancak daha fazla görüntünün de hatalı olacak şekilde patoloji sınıfı atandığı görülmüştür.



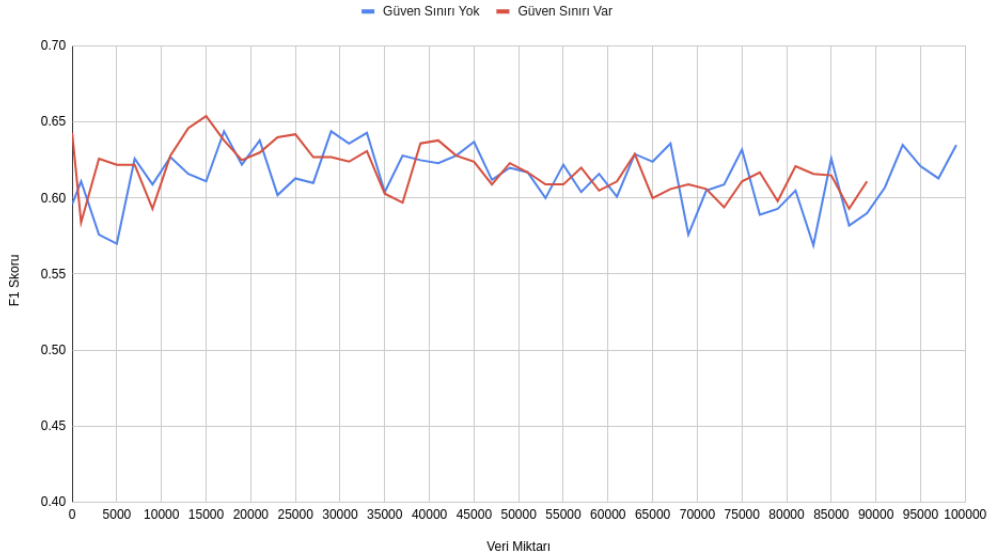
Şekil 4.14 Otokodlayıcı vektörü ile veri artırımı Kesinlik-Duyarlılık değişimi grafiği

Gerçeklenen ikinci deneyde ilk olarak 100.000 adet yapay kulak zarı görüntüsü PGGAN mimarisi ile üretilmiştir. Üretilen görüntüler EfficientNet tabanlı CNN'ler kullanılarak iki farklı yaklaşımda ele alınmıştır. Bu yaklaşımlar aşağıdaki gibidir:

- Üretilen yapay görüntünün direkt olarak IForest modeline verilmesi
- Üretilen yapay görüntünün CNN tabanlı sınıflandırıcıdan alınan sınıfa ait olma olasılığına göre elenerek IForest sistemine verilmesi

İlk senaryoda üretilen yapay görüntüler IForest sistemine her hangi bir ön kontrol uygulanmadan beslenmiştir. İkinci senaryoda ise öğrenme aktarması sırasında eğitilen model kullanılarak veri temizliği gerçekleştirilmiştir. Karar sınır değeri olarak tahmin olasılığı belirlenen 0.7 değerinden büyük görüntüler ve 0.3

EfficientNet-B2



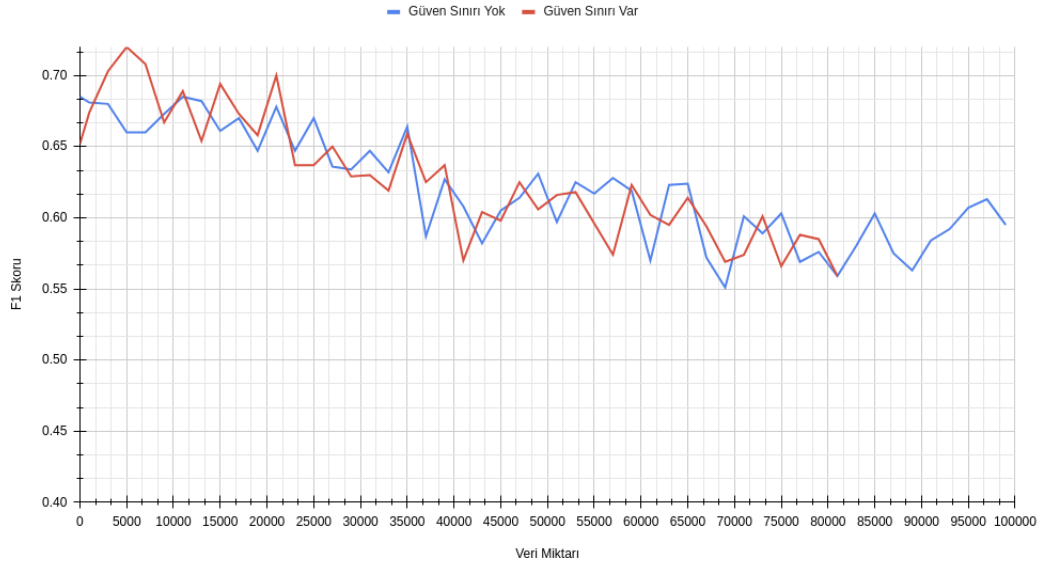
Şekil 4.15 EfficientNet-B2 tabanlı CNN vektörü ile veri artırımında F1 Skoru değişimi grafiği

değerinden küçük görüntüler seçilmiştir. Bu kontrol ile gözetimli eğitilen model kullanılarak oluşturulan IForest modelini potansiyel olarak yanıltabilecek görüntülerin temizlenmesi amaçlanmıştır.

Bu iki yaklaşım ile eğitim gerçekleştirildiğinde farklı EfficientNet mimarileri ile alınan F1 başarımları Şekil 4.15, 4.16, 4.17, 4.18, 4.19’da görülebilir. Her iki yöntemle alınan başarımlar incelendiğinde elde edilen en yüksek F1 skorunun, Bölüm 4.4.2’de elde edilen en yüksek F1 skorundan düşük olduğu görülmüştür.

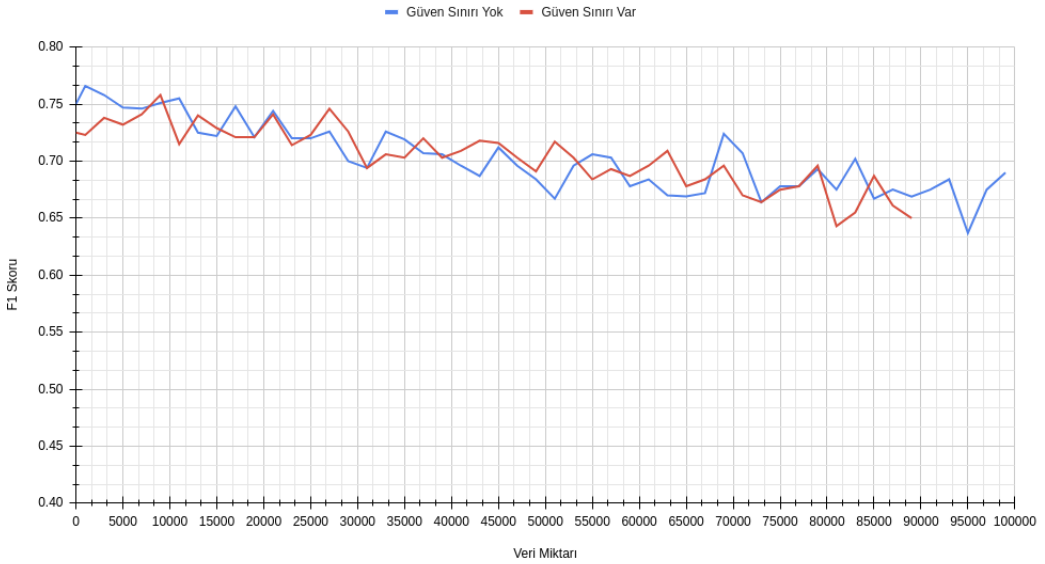
Şekil 4.20, 4.21, 4.22, 4.23, 4.24’te görülen kesinlik-duyarlılık değişimi incelendiğinde bu durumun düşme eğiliminde olan kesinlik başaımından kaynaklandığı anlaşılmıştır. Gerçeklenen deneylerde alınan en iyi kesinlik-duyarlılık başarımları EfficientNet-B3 mimarisi ile çıkartılan öznetelik vektörleri ile alınmıştır. Sisteme orijinal görüntülere ek olarak güven sınırından geçmiş 61.000 yapay görüntü eklendiğinde kesinlik ve duyarlılık skoru sırasıyla 0.435 ve 0.977 elde edilmiştir. Gözetimli eğitilen EfficientNet-B3 modeline kıyasla %29’lük bir kesinlik kaybı gözlemlenmiştir. Ancak düşen kesinlik başarımının aksine duyarlılık başarımı da %13 artarak sistemin patolojik görüntüleri daha başarılı olarak tespit etmesini sağlamıştır. EfficientNet-B6 tabanlı modelden alınan öznetelik vektörlerinde veri artırımı yapıldığında kesinlik ve duyarlılık skoru sırasıyla 0.554 ve 0.895 elde edilmiştir. Kesinlik başarımı gözetimli modele kıyasla %6 düşmüş olsa da duyarlılık başarımının %22 arttığı gözlemlenmiştir.

EfficientNet-B3



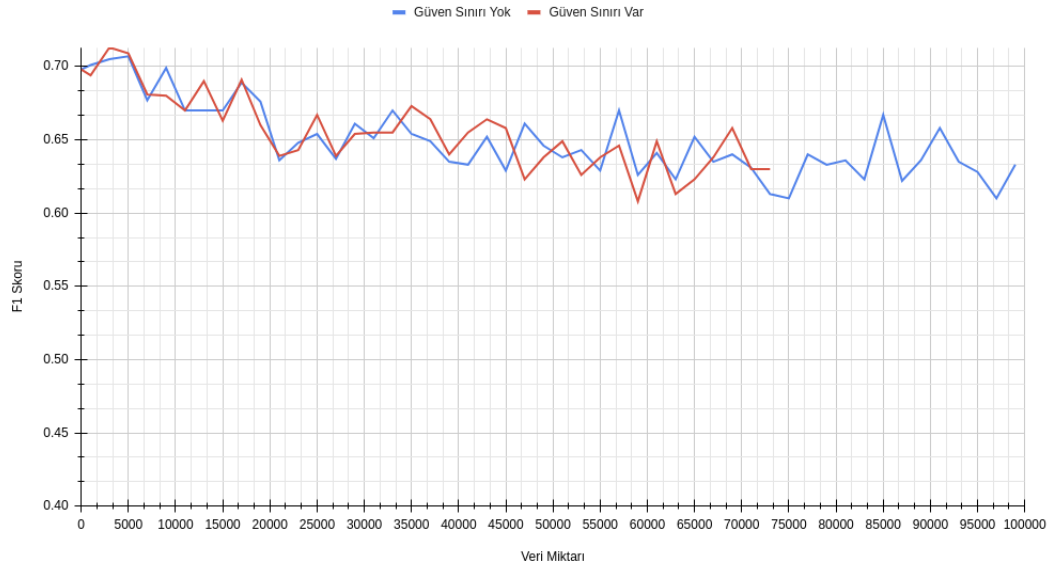
Şekil 4.16 EfficientNet-B3 tabanlı CNN vektörü ile veri artırımında F1 Skoru değişimi grafiği

EfficientNet-B4



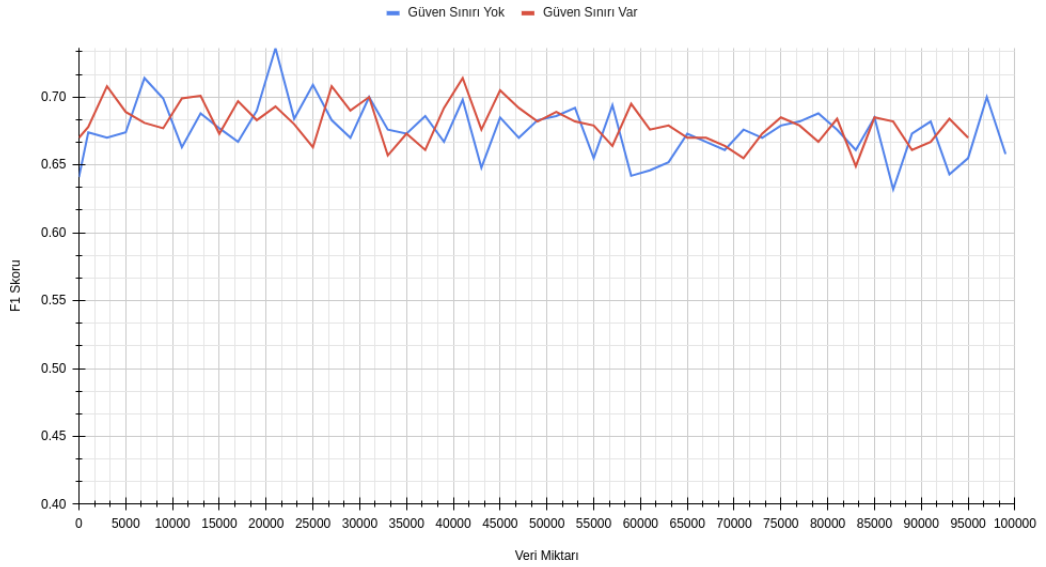
Şekil 4.17 EfficientNet-B4 tabanlı CNN vektörü ile veri artırımında F1 Skoru değişimi grafiği

EfficientNet-B5



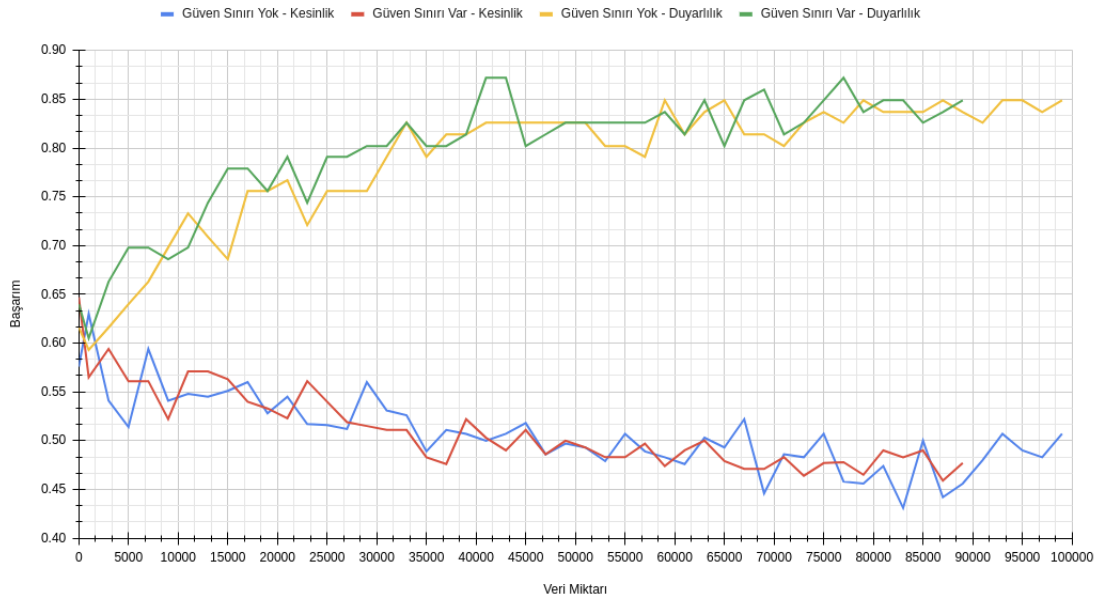
Şekil 4.18 EfficientNet-B5 tabanlı CNN vektörü ile veri artırımında F1 Skoru değişimi grafiği

EfficientNet-B6



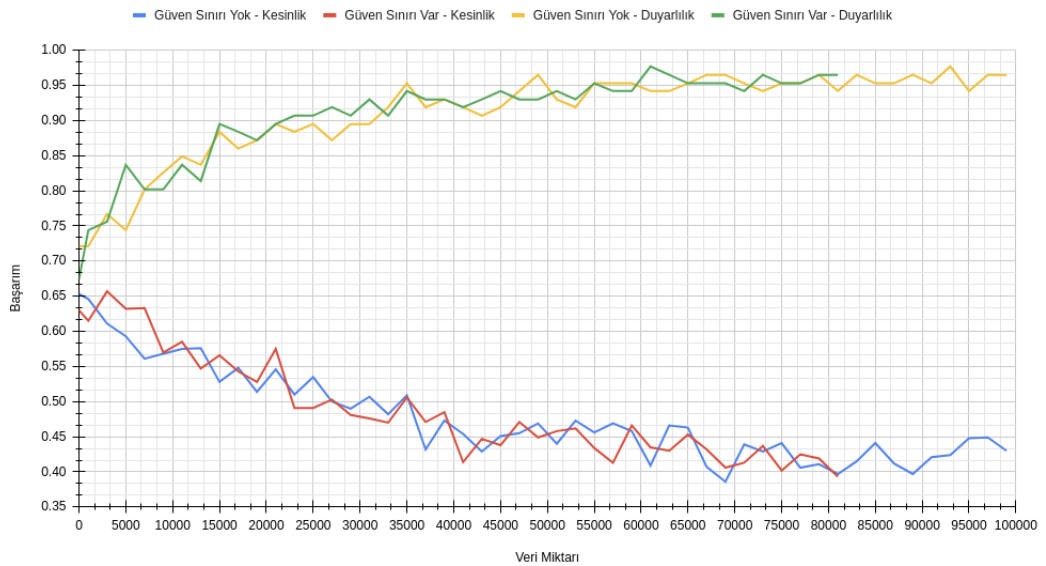
Şekil 4.19 EfficientNet-B6 tabanlı CNN vektörü ile veri artırımında F1 Skoru değişimi grafiği

EfficientNet-B2



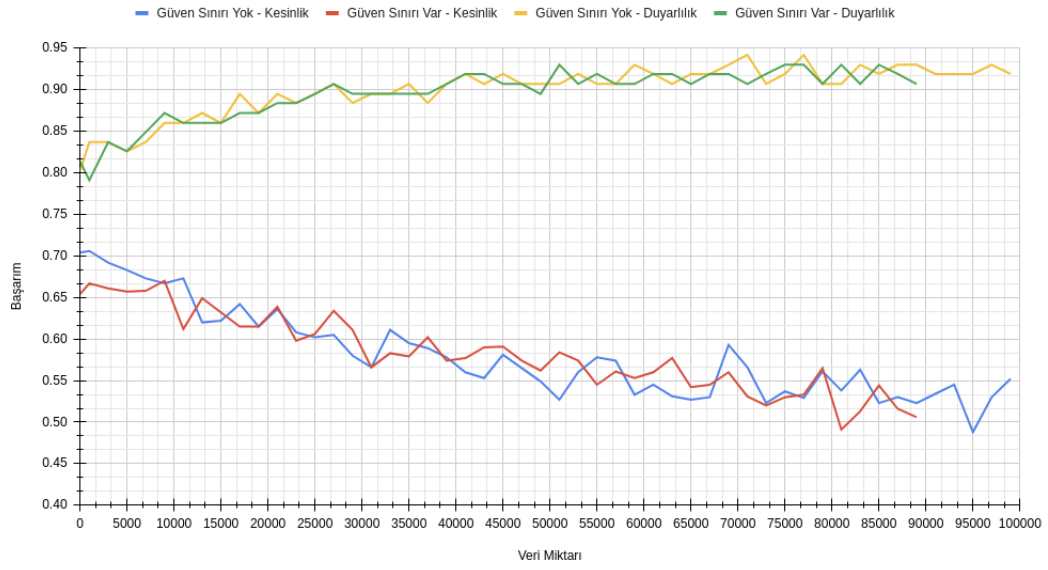
Şekil 4.20 EfficientNet-B2 tabanlı CNN vektörü ile veri artırımı Kesinlik-Duyarlılık değişimi grafiği

EfficientNet-B3



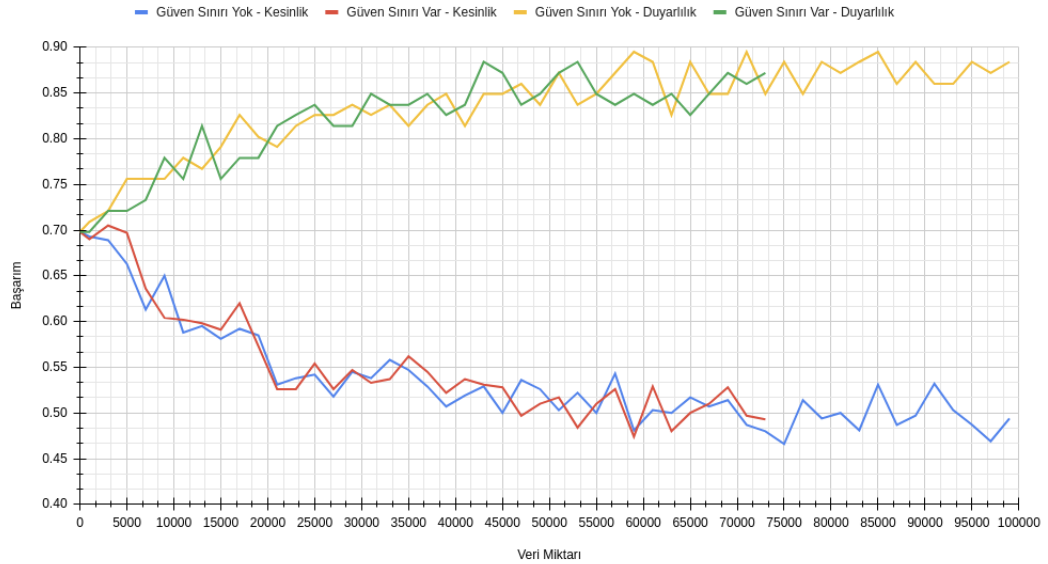
Şekil 4.21 EfficientNet-B3 tabanlı CNN vektörü ile veri artırımı Kesinlik-Duyarlılık değişimi grafiği

EfficientNet-B4

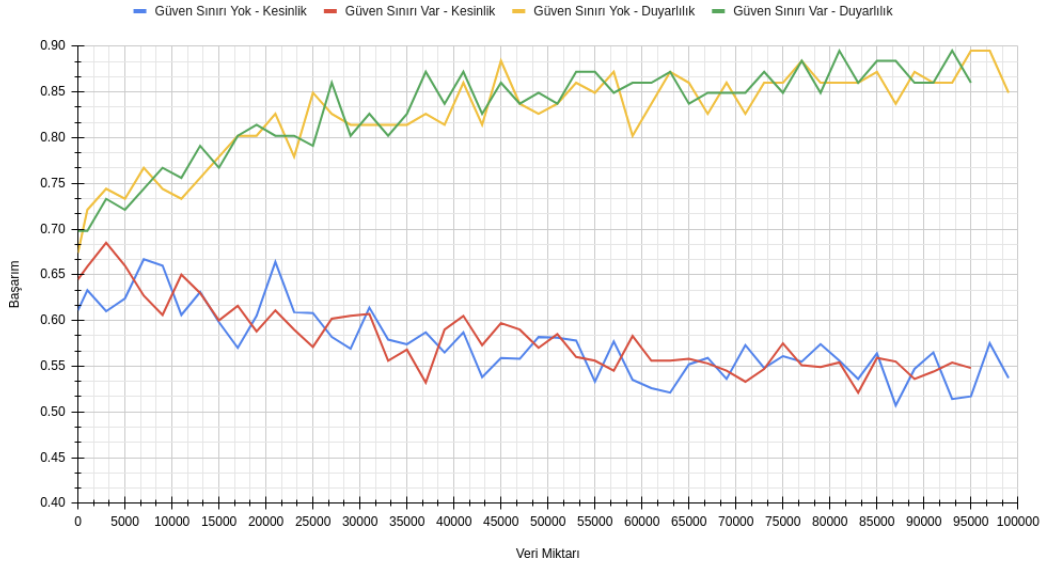


Şekil 4.22 EfficientNet-B4 tabanlı CNN vektörü ile veri artırımı Kesinlik-Duyarlılık değişimi grafiği

EfficientNet-B5



Şekil 4.23 EfficientNet-B5 tabanlı CNN vektörü ile veri artırımı Kesinlik-Duyarlılık değişimi grafiği



Şekil 4.24 EfficientNet-B6 tabanlı CNN vektörü ile veri artırımı Kesinlik-Duyarlılık değişimi grafiği

4.6 Deney Sonuçlarının Değerlendirilmesi

Bütün deney ve başarımlar değerlendirildiğinde sınırlı veriyle gözetimli olarak eğitilen CNN'den alınan öznetelik vektörlerinin otokodlayıcıya kıyasla daha başarılı bir şekilde veri artırımı için kullanılabileceği görülmüştür.

Yapay veri kullanarak veri kümesini büyütmenin sistem başarımına etkisini görmek amacıyla eğitilen IForest modeli CNN tabanlı sınıflandırıcının öznetelik çıkarma modülü ile birlikte kullanıldığında en yüksek duyarlılık skorunun elde edildiği görülmüştür. Alınan sonuç Tablo 4.10'da görülebilir.

Duyarlılık skoru temel alındığında en başarılı veri artırımı işlemi EfficientNet-B3 mimarisinden alınan öznetelik vektöründe elde edilmiştir. Orijinal veri kümesine ek olarak üretilen 61.000 görüntü ile duyarlılık başarısı 0.721'den 0.977'ye yükselmiştir. Buna ek olarak kesinlik başarımı 0.653'ten 0.435'e düşmüştür. F1 skoru da 0.685'ten 0.602'ye düşüş göstermiştir. Bu kapsamda patolojik görüntülerin tespiti her ne kadar yüksek başarımla yapılsa da bunun yanı sıra ortaya çıkan yanlış alarmlar da oldukça artmıştır. Patolojik görüntülerin doğru olarak tespit edilebilmesinin zorunlu olduğu bu senaryoda kesinlikteki başarımla düşüşü duyarlılıktaki yüksek artış sebebi ile göz ardı edilebilecek noktada değildir. EfficientNet-B6 mimarisinden alınan öznetelik vektörleri ile daha kesinlikteki başarımla düşüşüne kıyasla daha yüksek bir duyarlılık artışının elde edildiği görülmüştür.

Tablo 4.10 Veri artırımı ile alınan en yüksek duyarlılık başarımları

Mimari	Eklenen Veri Miktarı	Kesinlik	Duyarlılık	F1
EfficientNet-B2	0	0.576	0.616	0.596
	41000	0.503	0.872	0.638
EfficientNet-B3	0	0.653	0.721	0.685
	61000	0.435	0.977	0.602
EfficientNet-B4	0	0.704	0.802	0.75
	51000	0.584	0.93	0.717
EfficientNet-B5	0	0.698	0.698	0.698
	53000	0.484	0.884	0.626
EfficientNet-B6	0	0.611	0.674	0.641
	81000	0.554	0.895	0.684

Bütün öznitelikler kullanılarak veri artırımı yapıldığında sistemin duyarlılık başarımını artırmak adına daha fazla yanlış alarm üretecek şekilde kesinlik başarımını düşürdüğü gözlemlenmiştir. Bu kapsamda gözetimli sınıflandırma eğitimi aşamasında en yüksek kesinlik başarımına sahip EfficientNet-B6 modelinin iyi bir başlangıç noktası olduğu düşünülmektedir.

Son durumda IForest sisteminin duyarlılık başarımının artmış olması da önemli bir noktadır. Kesinliğin düşmesinden dolayı daha fazla yanlış alarm üretilse de hasta kişilerin yanlış sınıflandırılmasının engellenmesinin medikal alanda daha fazla önem arz ettiği düşünülmektedir. Yanlış alarmların tamamen göz ardı edildiği kapsamda EfficientNet-B6 mimarisi kullanılarak eğitilen gözetimli modele kıyasla aynı modeli öznitelik çıkarıcı olarak kullanan IForest ile alınan %22'lik duyarlılık artışı sistemin kullanılabilirliği açısından iyi bir işarettir.

5 SONUÇ VE ÖNERİLER

Bu çalışmada patolojik kulak zarı görüntülerini otomatik olarak tespit etmek amacıyla makine öğrenmesi ve derin öğrenme tabanlı bir sistem önerilmiştir. Sınırlı sayıda etiket bilgisi bulunan bir veri kümesi üzerinde deneyler koşulmuştur ve deneylerin sonuçları değerlendirilmiştir.

Önerilen sistemde veri artırımı yöntemi olarak kullanmak amacıyla yapay kulak zarı görüntülerinin üretimi yapılmıştır. Görüntülerin üretilmesi için farklı GAN mimarilerinden yararlanılmıştır. Üretilen görüntüler veri kümesinden örneklenen gerçek görüntüler ile karıştırılarak iki kulak burun boğaz uzmanına görsel turing testi olarak sunulmuştur. Sunulan görüntülerin uzmanlar tarafından gerçek/yapay olarak oylaması istenmiştir. Oylama sonucunda gerçek görüntülere %93 oranında gerçek oyu verilmiştir. En başarılı modelin ürettiği görüntülerin de %70 oranında gerçek olarak oylandığı gözlemlenmiştir. İnsanlar tarafından yapılan oylamaya ek olarak otomatik skorlama metrikleri ile yapay görüntülerin kalitesi ve çeşitliliği ölçülerek doktorlar tarafından en gerçekçi seçilen modelin 40.428 FID skoru ve 0.525 LPIPS skoru ile burada da en iyi başarımları aldığı gözlemlenmiştir. Alınan sonuçlardan yapay veri üretimi modelinin gerçek veri dağılımını başarılı bir şekilde öğrendiği görülmüştür.

Etiketlenen sınırlı veri kümesinde normal/patolojik görüntülerin dağılımının yüksek dengesizlik içerdiği görülmüştür. Bu kapsamda patoloji tespiti problemine gözetimsiz anomali tespiti problemi olarak yaklaşılmıştır. Gözetimsiz modelinin eğitiminde kullanılmak üzere iki farklı öznitelik çıkarımı yöntemi denenmiştir. Etiketsiz veri seti kullanılarak bir otokodlayıcı eğitimi gerçekleştirilmiştir ve girdi olarak aldığı görüntüleri başarılı olarak üretebildiği gözlemlenmiştir. Diğer bir yaklaşım olarak uzman bir doktor tarafından etiketlenen 645 adet görüntü ile Imagenet veri kümesinde eğitilmiş farklı EfficientNet modellerinde öğrenme aktarımı gerçekleştirilmiştir. Eğitilen modelin etiketli veri üzerinde patolojik sınıf için kesinlik değerinin 0.755, duyarlılık değerinin 0.825 ve F1 puanının 0.788 olduğu gözlemlenmiştir.

Gerçeklenen öznitelik çıkarımı modelleri kullanılarak etiketsiz verilerin ve en başarılı

GAN modeli ile üretilen görüntülerin vektörleri alınarak saklanmıştır. Yapay görüntü kullanmadan yapılan gözetimsiz IForest modeli eğitiminde 0.611 kesinlik, 0.674 duyarlılık ve 0.641 F1 skorları ile en dengeli modelin gözetimli eğitilen EfficientNet-B6 modeliyle üretilen özniteliklerden alındığı gözlemlenmiştir. Gözetimli sınıflandırma performansının altında alınan başarı puanlarını artırmak amacı ile farklı miktarlarda yapay görüntü eğitim verisine eklenerek deneyler yapılmıştır. Veri artırımı deneylerinde 9812 adet orijinal görüntüye eklenen 81000 yapay görüntü ile EfficientNet-B6 öznitelik çıkarımında kesinlik başarımının 0.611'den 0.554'e düştüğü ancak duyarlılık başarımının 0.674'ten 0.895'e yükseldiği gözlemlenmiştir. Bu bağlamda F1 puanının 0.641'den 0.684'e yükseldiği görülmüştür. Gözetimli sistem performansına kıyasla kesinlik başarımının düşük olduğu ancak sistemin patolojik görüntüleri daha yüksek başarımla ayırt edebildiği görülmüştür. Son durumda sistemden alınan yanlış alarmların arttığı ancak patolojik olan görüntülerin de kaçırılmadığı anlaşılmaktadır.

Yapay veri üretimi aşamasında yeterli etiket bilgisine sahip olunmaması sadece patolojik görüntüleri üretebilecek modellerin eğitilmesini engellemektedir. Gelecek çalışmalarda verideki etiket miktarının artırılarak sisteme direkt olarak normal veya patolojik sınıflı veriler besleyecek modeller geliştirilerek kesinlik başarımlarını iyileştirilmeye çalışılacaktır.

- [1] A. Esteva *et al.*, “Dermatologist-level classification of skin cancer with deep neural networks,” *Nature*, vol. 542, no. 7639, p. 115, 2017.
- [2] D. S. Kermany *et al.*, “Identifying medical diagnoses and treatable diseases by image-based deep learning,” *Cell*, vol. 172, no. 5, pp. 1122–1131, 2018.
- [3] D. W. Teele, J. O. Klein, B. Rosner, G. B. O. M. S. Group, “Epidemiology of otitis media during the first seven years of life in children in greater boston: A prospective, cohort study,” *Journal of infectious diseases*, vol. 160, no. 1, pp. 83–94, 1989.
- [4] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, F.E. Alsaadi, “A survey of deep neural network architectures and their applications,” *Neurocomputing*, vol. 234, pp. 11–26, 2017.
- [5] F. Piccialli, V. D. Somma, F. Giampaolo, S. Cuomo, G. Fortino, “A survey on deep learning in medicine: Why, how and when?” *Information Fusion*, vol. 66, pp. 111–137, 2021, ISSN: 1566-2535. DOI: <https://doi.org/10.1016/j.inffus.2020.09.006>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1566253520303651>.
- [6] C. Han *et al.*, “Gan-based synthetic brain mr image generation,” in *2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018)*, IEEE, 2018, pp. 734–738.
- [7] H. Salehinejad, E. Colak, T. Dowdell, J. Barfett, S. Valaee, “Synthesizing chest x-ray pathology for training deep convolutional neural networks,” *IEEE transactions on medical imaging*, vol. 38, no. 5, pp. 1197–1206, 2018.
- [8] H. Chen, P. Cao, “Deep learning based data augmentation and classification for limited medical data learning,” in *2019 IEEE International Conference on Power, Intelligent Computing and Systems (ICPICS)*, IEEE, 2019, pp. 300–303.
- [9] Z. Qin, Z. Liu, P. Zhu, Y. Xue, “A gan-based image synthesis method for skin lesion classification,” *Computer Methods and Programs in Biomedicine*, vol. 195, p. 105 568, 2020.
- [10] S. Motamed, P. Rogalla, F. Khalvati, “Data augmentation using generative adversarial networks (gans) for gan-based detection of pneumonia and covid-19 in chest x-ray images,” *Informatics in Medicine Unlocked*, vol. 27, p. 100 779, 2021.
- [11] S. K. Lim, Y. Loo, N.-T. Tran, N.-M. Cheung, G. Roig, Y. Elovici, “Doping: Generative data augmentation for unsupervised anomaly detection with gan,” in *2018 IEEE International Conference on Data Mining (ICDM)*, IEEE, 2018, pp. 1122–1127.

- [12] K. Ouardini *et al.*, “Towards practical unsupervised anomaly detection on retinal images,” in *Domain Adaptation and Representation Transfer and Medical Image Learning with Less Labels and Imperfect Data*, Springer, 2019, pp. 225–234.
- [13] X. Li, Y. Lu, C. Desrosiers, X. Liu, “Out-of-distribution detection for skin lesion images with deep isolation forest,” in *International Workshop on Machine Learning in Medical Imaging*, Springer, 2020, pp. 91–100.
- [14] N. Shvetsova, B. Bakker, I. Fedulova, H. Schulz, D. V. Dylov, “Anomaly detection in medical imaging with deep perceptual autoencoders,” *IEEE Access*, vol. 9, pp. 118 571–118 583, 2021.
- [15] C. Shorten, T. M. Khoshgoftaar, “A survey on image data augmentation for deep learning,” *Journal of Big Data*, vol. 6, no. 1, pp. 1–48, 2019.
- [16] I. Goodfellow *et al.*, “Generative adversarial nets,” *Advances in neural information processing systems*, vol. 27, 2014.
- [17] K. Cheng, R. Tahir, L. K. Eric, M. Li, “An analysis of generative adversarial networks and variants for image synthesis on mnist dataset,” *Multimedia Tools and Applications*, vol. 79, no. 19, pp. 13 725–13 752, 2020.
- [18] M. Arjovsky, S. Chintala, L. Bottou, “Wasserstein generative adversarial networks,” in *International conference on machine learning*, PMLR, 2017, pp. 214–223.
- [19] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, A. Courville, “Improved training of wasserstein gans,” *arXiv preprint arXiv:1704.00028*, 2017.
- [20] T. Karras, T. Aila, S. Laine, J. Lehtinen, “Progressive growing of gans for improved quality, stability, and variation,” *arXiv preprint arXiv:1710.10196*, 2017.
- [21] T. Karras, S. Laine, T. Aila, “A style-based generator architecture for generative adversarial networks,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 4401–4410.
- [22] A. Borji, “Pros and cons of gan evaluation measures,” *Computer Vision and Image Understanding*, vol. 179, pp. 41–65, 2019.
- [23] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, X. Chen, “Improved techniques for training gans,” *Advances in neural information processing systems*, vol. 29, pp. 2234–2242, 2016.
- [24] C. Han *et al.*, “Combining noise-to-image and image-to-image gans: Brain mr image augmentation for tumor detection,” *IEEE Access*, vol. 7, pp. 156 966–156 977, 2019.
- [25] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, “Imagenet: A large-scale hierarchical image database,” in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255. DOI: 10.1109/CVPR.2009.5206848.
- [26] S. Barratt, R. Sharma, *A note on the inception score*, 2018. arXiv: 1801.01973 [stat.ML].
- [27] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, S. Hochreiter, “Gans trained by a two time-scale update rule converge to a local nash equilibrium,” *Advances in neural information processing systems*, vol. 30, 2017.

- [28] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, O. Wang, “The unreasonable effectiveness of deep features as a perceptual metric,” in *CVPR*, 2018.
- [29] S. R. Dubey, “A decade survey of content based image retrieval using deep learning,” *CoRR*, vol. abs/2012.00641, 2020. arXiv: 2012 . 00641. [Online]. Available: <https://arxiv.org/abs/2012.00641>.
- [30] M. Chen, X. Shi, Y. Zhang, D. Wu, M. Guizani, “Deep features learning for medical image analysis with convolutional autoencoder neural network,” *IEEE Transactions on Big Data*, 2017.
- [31] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, P-A. Manzagol, L. Bottou, “Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion,” *Journal of machine learning research*, vol. 11, no. 12, 2010.
- [32] F. T. Liu, K. M. Ting, Z.-H. Zhou, “Isolation-based anomaly detection,” *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 6, no. 1, pp. 1–39, 2012.
- [33] K. K. Pal, K. S. Sudeep, “Preprocessing for image classification by convolutional neural networks,” in *2016 IEEE International Conference on Recent Trends in Electronics, Information Communication Technology (RTEICT)*, 2016, pp. 1778–1781. DOI: 10.1109/RTEICT.2016.7808140.
- [34] M. Tan, Q. V. Le, “Efficientnet: Rethinking model scaling for convolutional neural networks,” *CoRR*, vol. abs/1905.11946, 2019. arXiv: 1905 . 11946. [Online]. Available: <http://arxiv.org/abs/1905.11946>.
- [35] A. Paszke *et al.*, “Pytorch: An imperative style, high-performance deep learning library,” in *Advances in Neural Information Processing Systems 32*, H. Wallach, H. Larochelle, A. Beygelzimer, F. d’Alché-Buc, E. Fox, R. Garnett, Eds., Curran Associates, Inc., 2019, pp. 8024–8035. [Online]. Available: <http://papers.nips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>.
- [36] Martín Abadi *et al.*, *TensorFlow: Large-scale machine learning on heterogeneous systems*, Software available from tensorflow.org, 2015. [Online]. Available: <https://www.tensorflow.org/>.
- [37] F. Pedregosa *et al.*, “Scikit-learn: Machine learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [38] Y. Zhao, Z. Nasrullah, Z. Li, “Pyod: A python toolbox for scalable outlier detection,” *Journal of Machine Learning Research*, vol. 20, no. 96, pp. 1–7, 2019. [Online]. Available: <http://jmlr.org/papers/v20/19-011.html>.

Konferans Bildirisi

1. M. F. Eseoğlu, M. E. Karslıgil ve İ. Koçak, "Çekişmeli Üretici Ağlar ile Yapay Kulak Zarı Görüntülerinin Üretimi," 2021 29th Signal Processing and Communications Applications Conference (SIU), 2021, pp. 1-4, doi:10.1109/SIU53274.2021.9477848.

