



T.C.

BİLECİK ŞEYH EDEBALI ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

BİLGİSAYAR MÜHENDİSLİĞİ TEZLİ YÜKSEK LİSANS PROGRAMI

**DENGESİZ CİLT LEZYONU VERİ SETLERİNDE GAN TABANLI VERİ ARTIRIMI
İLE DERİN ÖĞRENME SINIFLANDIRMA PERFORMANSININ İYİLEŞTİRİLMESİ**

YÜKSEK LİSANS TEZİ

SERKAN ÖZDEMİR

TEZ DANIŞMANI

DR. ÖĞR. ÜYESİ SALİM CEYHAN

BİLECİK, 2026

10812619

T.C.
BİLECİK ŞEYH EDEBALI ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
BİLGİSAYAR MÜHENDİSLİĞİ TEZLİ YÜKSEK LİSANS PROGRAMI

**DENGESİZ CİLT LEZYONU VERİ SETLERİNDE GAN TABANLI VERİ ARTIRIMI
İLE DERİN ÖĞRENME SINIFLANDIRMA PERFORMANSININ İYİLEŞTİRİLMESİ**

YÜKSEK LİSANS TEZİ

SERKAN ÖZDEMİR

TEZ DANIŞMANI
DR. ÖĞR. ÜYESİ SALİM CEYHAN

BİLECİK, 2026

10812619

BEYAN

“Dengesiz Cilt Lezyonu Veri Setlerinde GAN Tabanlı Veri Artırımı ile Derin Öğrenme Sınıflandırma Performansının İyileştirilmesi” başlıklı yüksek lisans tezi hazırlık ve yazım aşamasında bilimsel araştırma ve etik kurallarına uyduğumu, Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Faaliyetlerinde Üretken Yapay Zekâ Kullanımına Dair Etik Rehberine uygun olarak tez/dönem projemi hazırladığımı, başkalarının eserlerinden yararlandığım bölümlerde bilimsel etik kurallarına uygun olarak atıfta bulunduğumu, kullandığım verilerde herhangi bir tahrifat yapmadığımı, çalışmamın herhangi bir kısmının başka bir tez/dönem projesi olarak sunulmadığını, aksinin tespit edilmesi durumunda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

SERKAN ÖZDEMİR

.././20..

İmza:

ÖN SÖZ

Bu tez çalışmasının gerçekleştirilmesinde bana rehberlik yapan ve her bir safhasında kıymetli tecrübesini ve bilgisini benden esirgemeyen yüksek lisans tez danışmanım Sayın Dr. Öğr. Üyesi Salim CEYHAN'a sonsuz teşekkürlerimi iletir ve minnettarlığımı belirtmek isterim.

Bu aşamada bana her zaman destek olan Bilecik Şeyh Edebali Üniversitesi ailesine ve özellikle Mühendislik Fakültesi'ndeki kıymetli hocalarıma teşekkür ederim.

Tez savunma sürecinde şahsıma yol gösteren ve bu çalışmanın nihai haline gelmesinde büyük önemi olan kıymetli geri bildirimlerini benimle paylaşan Sayın jüri üyelerine en içten teşekkürlerimi ve saygılarımı sunarım.

Ayrıca akademik gelişimime katkıda bulunan ve bu süreci daha anlamlı kılan tüm değerli hocalarıma desteği, bu çalışmanın ortaya çıkmasında büyük rol oynamıştır.

Son olarak lisans eğitim döneminden itibaren benim her zaman en büyük destekçilerimden olan çok sevdiğim arkadaşım Sayın Doç. Dr. Yılmaz KAYA'ya ve hayatımın her aşamasında olduğu gibi yüksek lisans çalışmalarım sırasında da beni sabır ve içtenlikle destekleyen sevgili abim Serdar ÖZDEMİR'e en içten teşekkürlerimi sunarım.

ÖZET

Bu çalışmada, dengesiz cilt lezyonu veri setlerinin derin öğrenme modelleri üzerindeki etkisi incelenmiştir. Özellikle nadir fakat klinik açıdan kritik öneme sahip malign lezyonların doğru şekilde tespit edilmesi hedeflenmiştir. Bu doğrultuda, veri setindeki sınıf dengesizliği problemi gidermek amacıyla Derin Evrişimli Üretici Karşıt Ağlar (DCGAN) ve klasik veri artırma tekniklerini bir arada kullanan hibrit bir veri dengeleme yaklaşımı uygulanmıştır. Çalışmada HAM10000 veri setindeki dermatolojik görüntüler; dermato fibroma, vasküler lezyonlar ve diğer yaygın sınıfları içerecek şekilde düzenlenmiş ve hem orijinal dengesiz veri seti hem de yalnızca eğitim kümesine sentetik verilerin eklendiği dengelenmiş veri seti üzerinde analiz edilmiştir.

Söz konusu veriler derin öğrenme tabanlı sınıflandırma modelleri kullanılarak işlenmiş ve DenseNet121, ResNet50, VGG16, EfficientNetB0 ve ConvNeXt mimarileri ile karşılaştırmalı deneyler gerçekleştirilmiştir. Model eğitim sürecinde veri sızıntısını önlemek amacıyla sentetik veriler yalnızca eğitim setine dahil edilmiş, test seti ise gerçek verilerden oluşturulmuştur. Modellerin performansı doğruluk, Macro-Kesinlik, Macro-Duyarlılık ve Macro-F1 skoru metrikleriyle analiz edilmiştir. Elde edilen sonuçlar, dengesiz veri seti ile eğitilen modellerin yüksek doğruluk değerlerine ulaşmasına rağmen “Doğruluk Paradoksu” nedeniyle azınlık sınıflarını tespit etmekte yetersiz kaldığını göstermiştir. Önerilen hibrit dengeleme yöntemi ile eğitilen modellerde ise, klinik açıdan kritik azınlık sınıflarına ait duyarlılık (recall) değerlerinde belirgin iyileşmeler elde edilmiş; özellikle melanom sınıfında yanlış negatiflerin (kaçırılan malign vakaların) sayısı önemli ölçüde azalmıştır. Bu duyarlılık kazanımı, kesinlik ve genel doğrulukta sınırlı bir düşüşle birlikte geldiğinden, Macro-F1 skorundaki değişim modelden modele farklılık göstermiştir. Ayrıca, sınıf bazında elde edilen iyileşmelerin, üretilen sentetik görüntülerin kalitesiyle (FID skorlarıyla) doğrudan ilişkili olduğu gözlemlenmiştir.

Sonuç olarak, bu çalışmada önerilen GAN destekli hibrit veri artırma yaklaşımı, dengesiz veri setleri ile çalışan derin öğrenme modellerinde nadir ve klinik açıdan kritik lezyonların tespit duyarlılığını artırarak yanlış negatifleri azaltmış ve böylece klinik karar destek sistemlerinin güvenilirliğine anlamlı bir katkı sağlamıştır. Gelecekte yapılacak çalışmalarda, daha yüksek kaliteli sentetik görüntü üreten gelişmiş üretici modellerin ve sınıf dengesizliğine duyarlı kayıp fonksiyonlarının kullanılması, model performansını daha da artırabilir.

Anahtar Kelimeler: Cilt lezyonu sınıflandırması, sınıf dengesizliği, DCGAN, sentetik veri artırma, HAM10000, evrişimli sinir ağıları.

ABSTRACT

In this study, the impact of imbalanced skin lesion datasets on deep learning models is investigated, with a particular focus on improving the detection of rare but clinically critical malignant lesions. To address the class imbalance problem, a hybrid data balancing approach combining Deep Convolutional Generative Adversarial Networks (DCGAN) and conventional data augmentation techniques is proposed. Within the scope of the study, dermatological images including classes such as dermato fibroma, vascular lesions, and other common categories are utilized. Analyses are conducted on both the original imbalanced dataset and a balanced dataset created by adding synthetic images exclusively to the training set.

The data are processed using deep learning-based classification models, and comparative experiments are performed using DenseNet121, ResNet50, VGG16, EfficientNetB0, and ConvNeXt architectures. During the training process, synthetic data are incorporated only into the training set to prevent data leakage, while the test set consists solely of real data. Model performance is evaluated using accuracy, Macro-Precision, Macro-Recall, and Macro-F1 score. The results reveal that models trained on the imbalanced dataset achieve high accuracy; however, they suffer from the “Accuracy Paradox,” remaining inadequate at detecting minority classes. In contrast, models trained with the proposed hybrid balancing approach achieve substantial improvements in the recall (sensitivity) of clinically critical minority classes; in particular, the number of false negatives (missed malignant cases) in the melanoma class is markedly reduced. Since this sensitivity gain is accompanied by a limited decrease in precision and overall accuracy, the change in the Macro-F1 score varies from model to model. Moreover, the per-class improvements are observed to be directly correlated with the quality of the generated synthetic images (their FID scores).

In conclusion, the GAN-supported hybrid data augmentation strategy proposed in this study increases the detection sensitivity of rare and clinically critical lesions in deep learning models trained on imbalanced datasets, thereby reducing false negatives and providing a meaningful contribution to the reliability of clinical decision support systems. Future work may focus on employing advanced generative models that produce higher-quality synthetic images and class-imbalance-aware loss functions to further improve model performance.

Keywords: Skin lesion classification, class imbalance, DCGAN, synthetic data augmentation, HAM10000, convolutional neural networks.

İÇİNDEKİLER

	Sayfa
ÖN SÖZ	i
ÖZET	ii
ABSTRACT	iii
İÇİNDEKİLER	iv
TABLolar LİSTESİ	vi
ŞEKİLLER LİSTESİ	vii
KISALTMALAR VE SİMGELER LİSTESİ	viii
1. GİRİŞ	1
1.1. Literatür	3
1.1.1. Cilt lezyonu sınıflandırmasında derin öğrenme çalışmaları	3
1.1.2. Sınıf dengesizliğine yönelik algoritmik ve veri düzeyli yaklaşımlar	3
1.1.3. Çok modlu ve öznitelik tabanlı yaklaşımlar	4
1.1.4. HAM10000 veri setinde sınıf dengesizliği çalışmaları	5
1.1.5. Geleneksel veri artırma yöntemleri	6
1.1.6. GAN tabanlı sentetik veri üretimi	7
1.1.7. Güncel derin öğrenme mimarileri (ConvNeXt, DenseNet, EfficientNet)	8
1.1.8. Literatür değerlendirmesi ve araştırma boşluğu	9
1.2. Tezin amacı	9
1.3. Hipotez	9
2. KURAMSAL ÇERÇEVE	11
2.1. Yapay zekâ kavramı	11
2.2. Makine öğrenmesi	12
2.3. Derin öğrenme	12
2.4. Yapay sinir ağları	12
2.5. Performans değerlendirme yöntemleri	21
3. MATERYAL VE METOT	24
3.1. Veri seti analizi	24
3.1.1. Veri ön işleme ve bölümlleme	24

3.1.2.	Önerilen hibrit veri dengeleme stratejisi	25
3.1.3.	Rastgele alt örnekleme	26
3.2.	Uygulama değerlendirme	27
3.2.1.	Kullanılan model (GAN tabanlı sentetik veri üretimi – DCGAN)	27
3.2.2.	Klasik veri artırma ile boşluk doldurma	29
3.2.3.	Sınıflandırma modelleri	29
4.	BULGULAR VE TARTIŞMA	32
4.1.	Model performansının değerlendirilmesi	32
4.1.1.	Eğitim ve doğrulama sonuçları	33
4.1.2.	Dengesiz veri seti ile elde edilen bulgular	33
4.1.3.	GAN tabanlı dengelenmiş veri seti ile elde edilen bulgular	34
4.2.	Yöntemlerin karşılaştırılması	36
4.3.	Modelin iyileştirilmesi için öneriler	38
5.	SONUÇLAR VE ÖNERİLER	41
	KAYNAKÇA	43

TABLÖLAR LİSTESİ

	Sayfa
Tablo 3.1. Eğitim setinin sınıf bazında dengeleme kompozisyonu.	26
Tablo 4.1. Eğitim hiper-parametreleri ve donanım özellikleri.	32
Tablo 4.2. Dengeleme öncesi ve sonrası melanom (mel) tespit performansı.	35
Tablo 4.3. Sınıf bazında ortalama duyarlılık değerleri (beş modelin ortalaması).	35
Tablo 4.4. Dengesiz ve dengeli veri setlerinde model performansları.	38

ŞEKİLLER LİSTESİ

	Sayfa
Şekil 2.1.	Zeki sistemler (Russell ve Norvig, 1995). 11
Şekil 2.2.	Sinir hücresi biyolojik gösterimi. 13
Şekil 2.3.	Sinir hücresi matematiksel gösterimi. 14
Şekil 2.4.	Tam bağlaşımlı yapay sinir örneği (Lucas vd., 2018). 14
Şekil 2.5.	Tek katmanlı algılayıcı (Arı ve Berberler, 2017). 15
Şekil 2.6.	Çok katmanlı algılayıcı (Arı ve Berberler, 2017). 15
Şekil 2.7.	CNN mimarisi (Kızrak ve Bolat, 2018). 16
Şekil 2.8.	ReLU ve LeakyReLU aktivasyon fonksiyonlarının karşılaştırması. . . . 16
Şekil 2.9.	GAN mimarisi (Goodfellow vd., 2014). 18
Şekil 2.10.	DCGAN mimarisi (Radford vd., 2015). 18
Şekil 2.11.	LSUN sahne modellemesi için kullanılan DCGAN üretici mimarisi (Radford vd., 2015). 19
Şekil 2.12.	DenseNet121 mimarisi – Yoğun bağlantı yapısı (Huang vd., 2017). . . 20
Şekil 2.13.	ResNet50 mimarisi (He vd., 2016). 20
Şekil 2.14.	VGG16 mimarisi – Blok yapısı (Simonyan ve Zisserman, 2015). 21
Şekil 2.15.	EfficientNetB0 mimarisi (Tan ve Le, 2019). 21
Şekil 2.16.	ConvNeXt blok yapısı (V1 ve V2) (Liu vd., 2022). 22
Şekil 3.1.	Önerilen üç aşamalı hibrit veri dengeleme stratejisinin akış şeması. . . 27
Şekil 3.2.	Orijinal ve GAN üretimi melanom örnekleri. 28
Şekil 3.3.	Sınıflara göre gerçek ve DCGAN sentetik görüntü sayıları (parantez içinde ilgili FID skoru; FID’de düşük değer daha yüksek kaliteyi gösterir). 31
Şekil 4.1.	İncelenen beş mimarinin dengelenmiş veri seti üzerindeki eğitim perfor- mansı: (a) doğrulama doğruluğu, (b) doğrulama kaybı. 33
Şekil 4.2.	DenseNet121 modelinin dengesiz ve dengeli veri setlerindeki karmaşıklık matrisleri. 37
Şekil 4.3.	ResNet50 modelinin dengesiz ve dengeli veri setlerindeki karmaşıklık matrisleri. 37

KISALTMALAR VE SİMGELER LİSTESİ

AI :	Artificial Intelligence (Yapay Zekâ)
CNN :	Convolutional Neural Network (Evrışimli Sinir Ağı)
DCGAN :	Derin Evrışimli Üretici Karşıt Ağlar (Deep Convolutional Generative Adversarial Networks)
FID :	Fréchet Başlangıç Mesafesi (Fréchet Inception Distance)
GAN :	Generative Adversarial Networks (Üretici Karşıt Ağlar)
LeakyReLU :	Sızıntılı Doğrultulmuş Doğrusal Birim (Leaky Rectified Linear Unit)
MLP :	Çok Katmanlı Algılayıcılar (Multi-Layer Perceptrons)
ReLU :	Doğrultulmuş Doğrusal Birim (Rectified Linear Unit)
SLP :	Tek Katmanlı Algılayıcılar (Single-Layer Perceptron)
STGAN :	Kendi Kendini Aktaran GAN (self-Transfer GAN)
ViT :	Vision Transformer (Vizyon Dönüştürücü)
σ :	Sigmoid Aktivasyon Fonksiyonu

1. GİRİŞ

Cilt kanseri, dünya genelinde görülme sıklığı hızla artan ve halk sağlığı açısından önemli bir tehdit oluşturan hastalıklar arasında yer almaktadır. Özellikle son yıllarda artan ultraviyole maruziyeti, çevresel faktörler ve yaşam tarzı değişiklikleri, cilt kanseri vakalarında belirgin bir yükselişe neden olmuştur. Bu hastalık grubunda erken teşhis ve doğru tedavi planlaması, hastalığın seyrini doğrudan etkileyen en kritik faktörler arasında bulunmaktadır. Özellikle melanom gibi malign karakterli cilt lezyonları, erken evrede tespit edildiğinde yüksek oranda tedavi edilebilirken, geç tanı durumunda metastaz riski artmakta ve bu durum hayatta kalma oranlarını ciddi şekilde düşürmektedir. Nitekim GLOBOCAN 2022 verilerine göre dünya genelinde yaklaşık 331.722 yeni melanom vakası bildirilmiş ve hastalık 58.667 ölüme yol açmıştır (Bray vd., 2024). Melanom, tüm cilt kanserleri içinde görece küçük bir paya sahip olmasına karşın cilt kanserine bağlı ölümlerin büyük çoğunluğundan sorumludur (American Cancer Society, n.d.-a). Bu ölümcül seyri belirleyen en önemli etken tanı evresidir: lokalize (erken) evrede yakalanan melanomda beş yıllık görece sağkalım oranı %99'un üzerindeyken, hastalık uzak organ metastazı yaptığında bu oran yaklaşık %35'e kadar gerilemektedir (American Cancer Society, n.d.-b). Bu keskin farklılık, melanomun erken ve doğru tespitinin hastanın yaşam beklentisini doğrudan belirleyen kritik bir faktör olduğunu ortaya koymaktadır. Bu nedenle, erken tanı süreçlerini destekleyecek güvenilir ve hassas yöntemlerin geliştirilmesi büyük önem taşımaktadır.

Dermatoloji alanında yaygın olarak kullanılan dermoskopi, cilt yüzeyinin büyütülerek ve detaylı şekilde incelenmesini sağlayan non-invaziv bir görüntüleme tekniğidir. Dermoskopik görüntüler, lezyonların renk, yapı ve sınır özellikleri gibi mikroskobik detaylarını ortaya koyarak dermatologların daha doğru teşhis koymasına yardımcı olmaktadır. Bu yöntem, klinik uygulamalarda standart bir araç haline gelmiş olsa da dermoskopik görüntülerin yorumlanması uzmanlık gerektiren ve deneyime bağlı olarak değişkenlik gösterebilen bir süreçtir. Bu durum, otomatik ve nesnel karar destek sistemlerine olan ihtiyacı artırmaktadır. Son yıllarda yapay zekâ ve özellikle derin öğrenme alanındaki gelişmeler, tıbbi görüntü analizi süreçlerinde önemli ilerlemeler sağlamıştır. Evrişimsel Sinir Ağları (CNN) tabanlı modeller, görüntü sınıflandırma ve özellik çıkarımı konularında yüksek başarı performansları sergileyerek dermatolojik teşhis süreçlerinde etkin bir şekilde kullanılmaya başlanmıştır. Bu modeller, büyük veri setleri üzerinde eğitildiklerinde insan uzmanlara yakın, hatta bazı durumlarda daha yüksek doğruluk oranlarına ulaşabilmektedir. Ancak bu başarı, büyük ölçüde kullanılan veri setlerinin kalitesine, çeşitliliğine ve dengeli dağılımına bağlıdır.

Dermatolojik görüntü veri setlerinin önemli bir kısmı, doğası gereği sınıf dengesizliği problemi içermektedir. Örneğin, iyi huylu lezyonlar veri setlerinde genellikle fazla sayıda temsil edilirken, melanom, dermato fibroma ve vasküler lezyonlar gibi nadir ancak klinik açıdan kritik sınıflar oldukça sınırlı sayıda örnek içermektedir. Bu dengesizlik, derin öğrenme modellerinin

çoğunluk sınıflarına karşı yanlışlık geliştirmesine neden olmakta ve modelin genel doğruluğu yüksek görünse bile azınlık sınıflarını doğru şekilde tespit edememesine yol açmaktadır. Literatürde “Doğruluk Paradoksu” olarak bilinen bu durum, özellikle tıbbi uygulamalarda ciddi riskler doğurmaktadır. Çünkü yanlış negatif sonuçlar, yani mevcut bir hastalığın model tarafından tespit edilememesi, doğrudan hasta güvenliğini tehdit etmektedir.

Sınıf dengesizliği problemini çözmek amacıyla geleneksel veri artırma yöntemleri yaygın olarak kullanılmaktadır. Döndürme, ölçekleme, kırpma ve yansıtma gibi teknikler, mevcut görüntüler üzerinden yeni örnekler üretmekte ve veri setinin boyutunu artırmaktadır. Ancak bu yöntemler, yalnızca geometrik çeşitlilik sağlamakta ve lezyonların gerçek morfolojik varyasyonlarını yeterince temsil edememektedir. Bu nedenle, daha gerçekçi ve çeşitli veri üretimi sağlayabilecek yöntemlere ihtiyaç duyulmaktadır.

Bu noktada, Üretici Karşıt Ağlar (GAN) (Goodfellow vd., 2014) önemli bir çözüm olarak öne çıkmaktadır. GAN tabanlı modeller, gerçek veri dağılımını öğrenerek, eğitim veri setinde bulunmayan ancak istatistiksel olarak benzer ve klinik açıdan anlamlı yeni görüntüler üretebilmektedir. Bu özellikleri sayesinde GAN’lar, özellikle azınlık sınıflarının veri setinde daha iyi temsil edilmesini sağlayarak sınıf dengesizliği probleminin giderilmesine katkı sunmaktadır. Bu tez çalışmasında, dengesiz cilt lezyonu veri setlerinde sınıflandırma performansını artırmak amacıyla GAN tabanlı hibrit bir veri dengeleme stratejisi önerilmektedir. Önerilen yaklaşımda, geleneksel veri artırma yöntemleri ile GAN tabanlı sentetik veri üretimi birlikte kullanılarak daha dengeli ve çeşitli bir eğitim veri seti oluşturulmuştur. Çalışma kapsamında DenseNet121, ResNet50, VGG16, ConvNeXt ve EfficientNetB0 gibi farklı derin öğrenme mimarileri kullanılarak kapsamlı bir analiz gerçekleştirilmiştir.

Araştırma iki aşamalı olarak yürütülmüştür. İlk aşamada, söz konusu modeller orijinal dengesiz veri seti üzerinde eğitilmiş ve elde edilen sonuçlar değerlendirilmiştir. Bu aşamada modellerin yüksek doğruluk değerlerine ulaşmasına rağmen azınlık sınıflarını tespit etmede yetersiz kaldığı gözlemlenmiştir. İkinci aşamada ise, önerilen hibrit veri dengeleme yöntemi uygulanarak azınlık sınıfları sentetik verilerle zenginleştirilmiş ve modeller yeniden eğitilmiştir. Elde edilen sonuçlar, dengelenmiş veri seti ile eğitilen modellerin özellikle azınlık sınıflarına ait tespit duyarlılığını artırdığını ve klinik açıdan kritik yanlış negatifleri azalttığını ortaya koymuştur. Bu çalışmanın temel katkıları şu şekilde özetlenebilir: İlk olarak sınıf dengesizliğinin derin öğrenme modelleri üzerindeki etkisi “Doğruluk Paradoksu” bağlamında kapsamlı bir şekilde analiz edilmiştir. İkinci olarak GAN tabanlı sentetik veri üretiminin nadir görülen cilt lezyonlarının tespit duyarlılığı üzerindeki katkısı, sınıf bazında ve karmaşıklık matrisleri üzerinden nicel olarak ortaya konulmuştur. Üçüncü olarak, sentetik veri kalitesi (FID) ile sınıflandırma başarımı arasındaki ilişki incelenmiştir. Son olarak farklı derin öğrenme mimarilerinin dengelenmiş veri setleri üzerindeki performansları karşılaştırılarak, bu modellerin genelleme kabiliyetleri değerlendirilmiştir.

Sonuç olarak bu tez çalışması, cilt kanseri teşhisinde kullanılan derin öğrenme tabanlı

sistemlerin güvenilirliğini artırmaya yönelik önemli bir yaklaşım sunmaktadır. Önerilen hibrit veri dengeleme yöntemi, özellikle sınırlı ve dengesiz veri setleri ile çalışılan tıbbi uygulamalarda daha doğru ve dengeli sonuçlar elde edilmesine katkı sağlamaktadır. Gelecekte yapılacak çalışmaların, daha büyük veri setleri, farklı GAN mimarileri ve ileri optimizasyon teknikleri ile bu yaklaşımı geliştirmesi beklenmektedir.

1.1. Literatür

Bu bölümde, cilt lezyonu sınıflandırması, sınıf dengesizliği ve veri artırma teknikleri üzerine literatürde yer alan güncel çalışmalar temalarına göre gruplandırılarak sunulmaktadır.

1.1.1. Cilt lezyonu sınıflandırmasında derin öğrenme çalışmaları

Bu doğrultuda, Ergün ve Kılıç (2021), veri artırma teknikleri ile destekledikleri çalışmalarında ResNet-34 mimarisini kullanarak %87,5 doğruluk oranına ulaşmışlardır. Benzer şekilde, Aruk ve Toprak (2025) tarafından gerçekleştirilen ve yirmi bir farklı CNN modelinin karşılaştırıldığı kapsamlı çalışmada, EfficientNet-B0 modelinin %84 doğruluk ve %78,59 F1 skoru ile en dengeli performansı sergilediği rapor edilmiştir. Arora vd. (2022) ise, 14 farklı derin öğrenme modelini değerlendirdikleri çalışmalarında DenseNet201 mimarisinin %82,5 doğruluk oranı ile diğer modellere kıyasla daha başarılı sonuçlar verdiğini belirtmiştir. Bunun yanı sıra, model performansını artırmak amacıyla topluluk öğrenme (ensemble learning) yöntemlerini kullanan çalışmalar da literatürde yer almaktadır. Örneğin, Chatterjee vd. (2024), InceptionV3 ve Xception modellerini birleştirerek %98 test doğruluğu elde etmiş; Selvaraj vd. (2024) ise, ağırlıklı topluluk modeli ile %97 makro-ROC-AUC değerine ulaşmıştır.

Ancak, literatürde yüksek doğruluk oranları raporlayan bu çalışmaların önemli bir kısmı, veri setlerindeki sınıf dengesizliği problemini ve bu durumun azınlık sınıflar üzerindeki etkisini yeterince detaylı şekilde ele almamıştır. Özellikle çoğunluk sınıflarının veri setinde baskın olması, modellerin bu sınıflara yönelik bir yanlılık geliştirmesine neden olmakta ve nadir görülen ancak klinik açıdan kritik öneme sahip lezyonların tespitinde performans düşüşüne yol açmaktadır. Bu durum, genel doğruluk yüksek görünse bile modelin gerçek klinik kullanımda güvenilirliğini sorgulanabilir hale getirmektedir.

1.1.2. Sınıf dengesizliğine yönelik algoritmik ve veri düzeyli yaklaşımlar

Sınıf dengesizliği problemini çözmeye yönelik yaklaşımlar genel olarak iki ana grupta incelenmektedir. Bunlardan ilki, modelin öğrenme sürecini doğrudan etkileyen algoritmik düzenlemelerdir. Bu kapsamda, Gayatri ve Aarthy (2024), sınıflandırılması zor örneklerle daha fazla ağırlık veren Focal Loss fonksiyonunun model performansını artırmadaki etkinliğini ortaya koymuştur. Benzer şekilde, Nguyen vd. (2022), derin öğrenme modellerine Soft-Attention mekanizması entegre ederek %90 doğruluk oranına ulaşmışlardır. Yao vd. (2021) ise, modifiye edilmiş veri artırma teknikleri ile kümülatif öğrenme stratejilerini bir arada kullanarak, sınırlı

veri setlerinde dahi yüksek performans elde edilebileceğini göstermiştir. Bu yöntemler, modelin öğrenme kapasitesini artırarak belirli ölçüde iyileşme sağlasa da veri setinin çeşitliliğini artırma konusunda sınırlı kalmaktadır.

Sınıf dengesizliğini gidermeye yönelik bir diğer yaklaşım ise, veri seviyesinde gerçekleştirilen iyileştirmelerdir. Geleneksel veri artırma yöntemleri (döndürme, yansıtma, ölçekleme vb.), mevcut verilerden yeni örnekler üretmekte ancak bu yöntemler yalnızca geometrik dönüşümlerle sınırlı kalmakta ve lezyonların morfolojik çeşitliliğini yeterince yansıtamamaktadır. Bu sınırlılık, araştırmacıları daha gelişmiş sentetik veri üretim tekniklerine yöneltmiştir. Bu bağlamda Üretici Karşıt Ağlar (GAN), gerçek veri dağılımını öğrenerek daha gerçekçi ve çeşitli görüntüler üretebilme kapasitesi sayesinde önemli bir alternatif olarak öne çıkmaktadır. Nitekim Su vd. (2024), Self-Transfer GAN (STGAN) olarak adlandırdıkları iki aşamalı bir mimari kullanarak yüksek çözünürlüklü sentetik cilt lezyonu görüntüleri üretmiş ve bu verilerin sınıflandırma performansı üzerinde anlamlı iyileşmeler sağladığını göstermiştir. Özellikle kesinlik, duyarlılık ve F1 skoru gibi metriklerde belirgin artışlar elde edilmesi, GAN tabanlı yaklaşımların etkinliğini desteklemektedir.

1.1.3. Çok modlu ve öznitelik tabanlı yaklaşımlar

Yap vd. (2018) deri lezyonu teşhisi için dermatoskopik ve makroskopik görüntüler ile hasta bilgilerini birleştiren çok modlu bir yöntem önermişlerdir. Model hem ikili hem çok sınıflı görevlerde tek görüntü kullanan yöntemden daha iyi performans göstermiştir. Ayrıca dermatoskopik görüntülerin daha yüksek doğruluk sağladığı gösterilmiştir. Çalışma 2917 vakalık veri setiyle yapılmıştır. Melanom, tehlikeli bir cilt kanseri türüdür ve erken teşhis büyük önem taşır. Biyopsi gibi geleneksel yöntemler zaman alıcı ve hastalar için rahatsız edicidir. Deif ve Hammam (2020) dermoskopik görüntüler kullanılarak otomatik teşhis yapabilen derin öğrenme tabanlı bir yöntem önermişler. 10.135 görüntü üzerinde VGG16, Inception_v3 ve ResNet-50 modelleri karşılaştırılmıştır. Kıl (hair) temizleme işleminin doğruluğu artırdığı görülmüştür. En yüksek başarıyı %96,9 ile VGG16 modeli elde etmiş, onu Inception_v3 (%94,2) ve ResNet-50 (%91,8) izlemiştir. Mahbod vd. (2019) cilt kanseri teşhisi için otomatik bir sınıflandırma yöntemi önermişlerdir. Yöntem, AlexNet, VGG16 ve ResNet-18 gibi önceden eğitilmiş CNN modellerinden elde edilen derin özellikleri kullanır ve bu özellikler SVM ile sınıflandırılır. Son aşamada farklı model çıktıları birleştirilir. ISIC 2017 veri seti üzerinde yapılan testlerde, melanom için %83,83 ve seboreik keratoz için %97,55 AUC başarımı elde edilmiştir. Benyahia vd. (2022), cilt lezyonlarının sınıflandırılmasında öznitelik çıkarım yöntemleri incelemişlerdir. 17 farklı önceden eğitilmiş CNN modeli öznitelik çıkarıcı olarak, 24 makine öğrenmesi algoritması ise sınıflandırıcı olarak test edilmiştir. Sonuçlara göre, ISIC 2019 veri setinde DenseNet201 + Fine KNN veya Cubic SVM en iyi performansı (%92,34 ve %91,71 doğruluk) göstermiştir. PH2 veri setinde ise önerilen yöntem %99 doğruluk ile diğer yaklaşımlardan daha başarılı olmuştur. Da Silva vd. (2026) yüksek frekanslı ultrason görüntüleri kullanılarak cilt lezyonlarının

sınıflandırılması için derin öğrenme modelleri değerlendirmişlerdir. B-mod ve Doppler verileri ayrı ayrı ve birlikte (Unity ve Cascade mimarileriyle) kullanılmıştır. En iyi sonuç Doppler modeliyle elde edilmiştir (%95 doğruluk, AUC 0,98). Unity modeli de yüksek performans göstermiştir (%90,5 doğruluk), Cascade ise daha düşük doğruluk ancak kötü huylu lezyonlarda daha yüksek güven sunmuştur. Ibrahim vd. (2026), cilt kanserinin erken ve doğru teşhisi için dermoskopik görüntüler ile hasta verilerini birleştiren çok modlu bir yöntem önermişlerdir. Görüntüler CLAHE ile iyileştirilmiş ve veri artırma uygulanmıştır. Ardından çok ölçekli bir sinir ağı ve dikkat mekanizması ile önemli bölgeler vurgulanmıştır. Elde edilen görüntü özellikleri hasta bilgileriyle farklı füzyon yöntemleri kullanılarak birleştirilmiştir. HAM10000 veri setinde yapılan deneylerde en iyi sonuç, eleman bazlı çarpım yöntemiyle elde edilmiştir (%94,02 doğruluk, %99,31 AUC).

1.1.4. HAM10000 veri setinde sınıf dengesizliği çalışmaları

Setiawan ve Soewito (2025), HAM10000 veri setindeki sınıf dengesizliği problemini çözmek için CRCDKD adlı yeni bir derin öğrenme yöntemi önermişlerdir. Model, “mean-teacher” yaklaşımı, kontrastif öğrenme ve bilgi damıtma tekniklerini birleştirerek az temsil edilen sınıflarda performansı artırmayı hedeflemiştir. Önerilen yöntem, dengeli çok sınıflı doğruluk %84,45 ve genel doğruluk %89,41 elde etmiştir. Ayrıca AUC %98,41 gibi yüksek bir performans göstermiş ve özellikle azınlık sınıflarda önceki yöntemlere göre daha iyi sonuçlar vermiştir. Khan vd. (2025), dermoskopik görüntülerden melanom ve diğer cilt kanserlerini otomatik olarak tespit etmek için derin öğrenme tabanlı bir yöntem önermişlerdir. Modelde CNN ve DenseNet yapıları, batch normalization, max-pooling ve ReLU aktivasyon fonksiyonları kullanılmıştır. Veri dengesizliği ve aşırı öğrenme problemleri veri artırma ve uygun optimizasyon yöntemleriyle azaltılmıştır. Önerilen model ISIC-2019 veri setinde %95,70, HAM10000 veri setinde ise %93,24 doğruluk elde etmiştir. Ayrıca çapraz doğrulamada %94,8 başarı göstererek mevcut yöntemlerden daha iyi performans sağlamıştır. Haque vd. (2025) HAM10000 veri seti kullanılarak 7 farklı cilt kanseri türünü sınıflandırmak için CNN tabanlı bir model geliştirmişlerdir. Veri ön işleme olarak yeniden boyutlandırma, normalizasyon ve veri artırma uygulanarak sınıf dengesizliği azaltılmıştır. Önerilen model Adam optimizasyonu ile eğitilmiş ve %98,17 eğitim, %99,60 doğrulama doğruluğuna ulaşmıştır. Ancak az temsil edilen sınıflarda hatalar görülmüştür. Genel olarak, ResNet ve EfficientNet gibi yöntemlerle rekabet edebilir performans gösterirken daha düşük hesaplama maliyeti sağlamıştır. Gupta vd. (2026) HAM10000 veri setindeki sınıf dengesizliği problemini çözmek için derin üretici modeller kullanılarak veri artırma yapmışlardır. DCGAN, VAE ve DDPM yöntemleri ile özellikle az örneğe sahip sınıflar çoğaltılmış ve daha dengeli bir veri seti oluşturulmuştur. Oluşturulan veri seti ResNet50 tabanlı bir sınıflandırıcı ile test edilmiştir. En iyi sonuç DDPM yöntemiyle elde edilmiştir (%88 doğruluk ve 0,88 macro F1). Ayrıca nadir sınıflarda belirgin iyileşme sağlanmış, örneğin vasküler lezyon sınıfında F1 skoru 0,99’a ulaşmıştır. Bununla birlikte, söz konusu çalışmada üretici modellerin yalnızca beş epoch boyunca eğitilmesi, üretilen sentetik görüntülerin FID veya Inception Score gibi nicel kalite

ölçütleriyle değerlendirilmemesi ve test kümesinin gerçek/sentetik bileşiminin açıkça belirtilmemesi nedeniyle, özellikle nadir sınıflarda raporlanan çok yüksek F1 skorlarının (ör. 0,99) temkinli yorumlanması gerekmektedir. Bu durum, sentetik veriyle yapılan artırmada değerlendirme protokolünün (sentetik örneklerin yalnızca eğitim kümesinde kullanılması, test kümesinin ise tümüyle gerçek görüntülerden oluşması) titizlikle tasarlanmasının önemini ortaya koymaktadır. Abini ve Anjali (2025), HAM10000 veri seti kullanılarak cilt lezyonlarını sınıflandırmak için CNN'lere alternatif olarak Capsule Network tabanlı CapSkinNet modeli önermişlerdir. Bu yapı, uzamsal bilgiyi daha iyi koruyarak benzer görünümlü lezyonlar arasındaki ayrımı artırmayı hedefler. Model %96,8 doğruluk elde etmiş, sınıflar için yüksek precision, recall ve F1-skorları sağlamıştır (yaklaşık %93,4–%96,5 aralığında). Özellikle melanocytic nevi ve actinic keratoses sınıflarında yüksek performans göstermiştir. Sonuçlar, CapSkinNet'in klasik CNN'lere göre daha başarılı ve klinik kullanım için umut vadeden bir yöntem olduğunu göstermektedir. Kumari vd. (2026) cilt kanseri sınıflandırması için farklı veri setleri üzerinde çalışan (cross-dataset) bir CNN tabanlı yöntem önermişlerdir. HAM10000 veri seti ile eğitim yapılmış, ISIC 2019 ve HAM10000'un bir kısmı ile test gerçekleştirilmiştir. Veri artırma ve transfer öğrenme teknikleri kullanılmıştır. Başlangıçta VGG16 modelinde aşırı öğrenme (overfitting) problemi görülmüş ve doğruluk düşmüştür. Yapılan iyileştirmeler sonrası model HAM10000'da %84,2, ISIC 2019'da ise %76,5 doğruluk elde etmiştir. Çalışma, veri çeşitliliği ve doğru model optimizasyonunun performans için önemli olduğunu vurgulamaktadır.

1.1.5. Geleneksel veri artırma yöntemleri

Geleneksel veri artırma yöntemleri, özellikle sınırlı ve dengesiz veri setlerinde model performansını artırmak için uzun süredir kullanılan temel yaklaşımlardan biridir. Bu yöntemler, orijinal veri üzerinde çeşitli dönüşümler uygulayarak yapay olarak yeni örnekler üretmeyi amaçlar ve böylece modelin genelleme kabiliyetini geliştirmeye yardımcı olur. Görüntü tabanlı problemlerde en yaygın kullanılan geleneksel teknikler arasında döndürme, yatay ve dikey çevirme, ölçekleme, kırpma, kaydırma ve parlaklık/kontrast değişiklikleri yer alır. Bu işlemler, veri setinin çeşitliliğini artırarak modelin farklı açılardan ve koşullardan gelen görüntülere karşı daha dayanıklı olmasını sağlar. Perez vd. (2018), çalışmalarında melanom sınıflandırması için veri artırma yöntemlerinin etkisi incelenmiştir. Inception-v4, ResNet ve DenseNet modelleri kullanılarak 13 farklı veri artırma senaryosu test edilmiştir. Bu senaryolar arasında klasik geometrik ve renk dönüşümlerinin yanı sıra elastik dönüşümler, random erasing ve farklı lezyon görüntülerinin karıştırıldığı yeni bir yöntem de yer almıştır. Sonuçlar, veri artırmanın hem eğitim hem de test aşamasında performansı önemli ölçüde artırdığını göstermiştir. En iyi yöntem, harici veri kullanılmadan ISIC 2017 görevinde 0,882 AUC elde ederek, ek veri ile eğitilen en iyi yarışma sonucunu (0,874) geride bırakmıştır. Bribesh ve Tareq (2025) çalışmalarında cilt lezyonu sınıflandırmasında veri artırma yöntemlerinin etkisi HAM10000 veri seti üzerinde incelenmiştir. Dört farklı artırma yaklaşımı karşılaştırılmıştır: renk uzayı dönüşümleri, gürültü ekleme, random erasing ve geometrik dönüşümler. Sonuçlar, veri artırma türünün modele göre

değişen performans etkisi olduğunu göstermiştir. Özellikle doygunluk ve renk tabanlı dönüşümler en iyi sonuçları vermiştir. En yüksek performans GoogLeNet modeliyle %82,73 doğruluk ve 0,8205 F1 skoru olarak elde edilmiştir. Chowdhury (2025), cilt kanseri sınıflandırması için derin öğrenme ve makine öğrenmesini birleştiren hibrit bir yaklaşım önermişlerdir. Veri ön işleme aşamasında gürültü giderme, kenar temizleme, kıl temizleme, kontrast artırma ve veri artırma uygulanmıştır. HRNet, VGG16/19 ve ResNet50/101 gibi modellerden özellik çıkarımı yapılmış, ardından bu özellikler dikkat mekanizmaları (CAB ve SENet) ile iyileştirilmiştir. Çok ölçekli özellik birleştirme (MSFF) sonrası sınıflandırma için XGBoost, Random Forest ve diğer algoritmalar kullanılmıştır. En iyi performans HRNet + XGBoost kombinasyonu ile elde edilmiştir. Blas Garcia vd. (2026), çalışmalarında 19 farklı cilt hastalığının klinik RGB görüntülerden otomatik sınıflandırılması için derin öğrenme tabanlı bir yöntem geliştirilmiştir. Veri dengesizliğini azaltmak ve model performansını artırmak için kontrast odaklı ön işleme, gelişmiş veri artırma ve iki aşamalı fine-tuning süreci kullanılmıştır. ResNet-50, VGG-16, DenseNet-121 ve EfficientNet-B3 modelleri karşılaştırılmıştır. En iyi sonuç %89,63 doğruluk ile ResNet-50 modelinde elde edilmiştir. EfficientNet-B3 ise %88,28 doğrulukla yakın performans göstermiştir. Çalışma, uygun eğitim stratejileri ve veri dengeleme teknikleri ile daha basit CNN modellerinin bile güçlü sonuçlar verebileceğini ortaya koymuştur.

1.1.6. GAN tabanlı sentetik veri üretimi

Rahman vd. (2024) GAN kullanılarak %100 sentetik veri üretiminin NIDS modellerinin eğitimi için potansiyeli incelemişlerdir. Bu yaklaşım, gerçek dünya verilerine olan bağımlılığı önemli ölçüde azaltarak daha esnek ve etik açıdan uygun bir model geliştirme süreci sunmaktadır. UNSW-NB15 veri seti için %90 doğruluk, %91 kesinlik, %90 duyarlılık ve %89 F1 skoru elde edilmiştir. NSL-KDD veri setinde %84 doğruluk, %85 kesinlik, %84 duyarlılık ve %84 F1 skoru elde edilmiştir. BoT-IoT veri setinde ise tüm metriklerde %100 başarı sağlanmıştır. Vijayakumari ve Mathew (2025) cilt kanseri veri setlerinde sınıf dengesizliğini gidermek için dikkat mekanizması içeren EA-GAN modelini önermektedir. ISIC 2020 veri setinde yapılan deneylerde, yüksek kaliteli sentetik görüntüler üretilmiş ve EfficientNet-B3 ile %94,2 doğruluk elde edilmiştir. Yöntem, özellikle azınlık sınıfların tespitinde önemli iyileşme sağlayarak tıbbi görüntüleme de etkili bir veri artırma yaklaşımı sunmaktadır. Nirmala ve Premaladha (2023) cilt kanseri sınıflandırmasında GAN ile üretilen sentetik görüntülerin CNN (geliştirilmiş VGG-16) ile kullanımını incelemişlerdir. GAN tabanlı veri artırma ile %96,33 doğruluk elde edilmiştir. Yöntem, geleneksel veri artırmaya göre daha başarılı sonuçlar sunarak veri eksikliği sorununa etkili bir çözüm önermektedir. Saranya vd. (2024) melanom tespiti için GAN tabanlı sentetik veri üretimi kullanarak sınıflandırma performansını artırmayı önermişlerdir. Model, ROC ve AUC gibi metriklerle değerlendirilmiş ve sentetik verinin veri çeşitliliğini artırarak başarıyı iyileştirdiği gösterilmiştir. Ayrıca sisteme entegre edilen dermatolog öneri modülü, hastaları uzmanlık ve konuma göre uygun doktorlarla eşleştirerek pratik bir klinik destek sunmaktadır. Shaik vd. (2026) sınırlı veri problemi ve sınıf dengesizliğini çözmek için GAN ile sentetik cilt

lezyonu görüntüleri üretmeyi önermektedir. ISIC veri seti üzerinde yapılan deneylerde, veri artırma yaklaşımı model performansını artırmış ve yanlış sınıflandırmaları azaltmıştır. Sonuçlar, GAN tabanlı veri artırmanın cilt kanseri teşhisinde etkili bir yöntem olduğunu göstermektedir. Tariq vd. (2026), GAN ve difüzyon modelleri ile üretilen sentetik cilt kanseri görüntülerini gerçek görüntülerden ayırt etmek için derin öğrenme tabanlı bir yöntem önermektedir. DenseNet169 modeli %98,50 test doğruluğu ile en iyi performansı göstermiştir. Sonuçlar, tıbbi görüntü adli bilişiminde yapay görüntülerin tespiti için etkili bir yaklaşım sunmaktadır. Lin vd. (2026) melanom görüntü üretimi için DCGAN, StyleGAN2 ve StyleGAN3 modellerini karşılaştıran kapsamlı bir benchmark sunmuşlardır. StyleGAN2 en iyi performansı göstermiş ve sentetik veri eklenmesi ile melanom tespit AUC değeri 0,925'ten 0,945'e yükselmiştir. Sonuçlar, GAN tabanlı sentetik görüntülerin sınıf dengesizliğini azaltmada ve tanısal doğruluğu artırmada etkili olduğunu göstermektedir.

1.1.7. Güncel derin öğrenme mimarileri (ConvNeXt, DenseNet, EfficientNet)

Sajol vd. (2024) HAM10000 veri seti üzerinde farklı CNN, Transformer ve hibrit derin öğrenme modellerini karşılaştırmışlardır. ConvNeXt V2 modeli %93,2 doğruluk ve düşük çıkarım süresi ile en iyi performansı göstermiştir. Sonuçlar, derin öğrenme tabanlı yöntemlerin cilt kanseri tespitinde etkili olduğunu ortaya koymaktadır. Naeem vd. (2025) ConvNeXt ve ViT mimarilerini birleştiren MedFusionNet modelini önererek cilt kanseri sınıflandırmasını geliştirmektedir. HAM10000 ve ISIC-2019 veri setlerinde sırasıyla %98,80 ve %97,90 doğruluk elde edilmiştir. Sonuçlar, hibrit CNN-Transformer modellerinin cilt kanseri teşhisinde yüksek performans sunduğunu göstermektedir. Faris vd. (2026) HAM10000 veri seti üzerinde ConvNeXt, EfficientNetV2 ve MobileNetV3 tabanlı üçlü ensemble model önermektedir. Sınıf dengesizliği için oversampling yerine dengeli örnekleme kullanılmış ve model %89,72 doğruluk elde etmiştir. Sonuçlar, ensemble ve veri artırma tekniklerinin cilt kanseri sınıflandırmasını iyileştirdiğini göstermektedir. Singh vd. (2024) HAM10000, ISIC 2019 ve ISIC 2020 veri setlerini birleştirerek 68.472 görüntü içeren ortak bir veri seti oluşturmuşlardır. DenseNet201, EfficientNetV2M ve ConvNeXtBase modelleri transfer öğrenme ile eğitilmiş ve ensemble yöntemi kullanılmıştır. Model %96,5 (çok sınıflı) ve %98,7 (ikili) doğruluk elde ederek yüksek performans göstermiştir. Lilhore vd. (2025) ConvNeXt, EfficientNetV2 ve Swin Transformer tabanlı SkinEHDLF modelini önermişlerdir. ISIC 2024 veri setinde %98,76 doğruluk ve %99,8 AUROC elde edilmiştir. Model, mevcut yöntemlere göre daha düşük yanlış pozitif oranı ve daha yüksek doğruluk sağlayarak cilt kanseri sınıflandırmasında üstün performans göstermektedir. Wei vd. (2023) model füzyonu ve attention mekanizması içeren bir CNN tabanlı cilt hastalığı sınıflandırma modeli önermişlerdir. HAM10000 veri setinde %95,29 doğruluk ve %89,99 F1 skoru elde edilmiştir. Sonuçlar, model füzyonunun ve veri artırma tekniklerinin sınıflandırma performansını artırdığını göstermektedir.

1.1.8. Literatür değerlendirmesi ve araştırma boşluğu

Literatürdeki mevcut çalışmalar genel olarak değerlendirildiğinde, cilt lezyonu sınıflandırmasında derin öğrenme modellerinin yüksek doğruluk oranları sergilediği görülmektedir. Ancak bu çalışmaların önemli bir kısmı, veri setlerindeki sınıf dengesizliği problemini ve bu durumun nadir görülen sınıflar üzerindeki etkisini yeterince detaylı ele almamıştır. Özellikle çoğunluk sınıflarının baskın olması, modellerin bu sınıflara yönelik yanlılık geliştirmesine ve genel doğruluk yüksek görünse de klinik açıdan kritik öneme sahip lezyonların tespitinde performans düşüşüne yol açmaktadır; bu durum literatürde “Doğruluk Paradoksu” olarak tanımlanmaktadır. Ayrıca, geleneksel veri artırma yöntemleri (döndürme, ölçekleme vb.) yalnızca geometrik çeşitlilik sağlamakta, lezyonların gerçek morfolojik varyasyonlarını temsil etmekte yetersiz kalmaktadır. Modern derin öğrenme mimarilerinin, özellikle ConvNeXt gibi güncel modellerin, GAN ile üretilen sentetik verilere olan adaptasyon yeteneklerinin literatürde yeterince incelenmediği de dikkat çeken bir diğer eksikliktir.

1.2. Tezin amacı

Tezin amacı, dengesiz cilt lezyonu veri setlerinde (özellikle HAM10000) derin öğrenme modellerinin çoğunluk sınıfına karşı geliştirdiği yanlılığı gidermek ve nadir fakat klinik açıdan kritik öneme sahip malign ve azınlık sınıfların (melanom, dermato fibroma, vasküler lezyonlar vb.) tespit duyarlılığını artırarak yanlış negatifleri azaltmaktır. Bu doğrultuda, Derin Evrişimli Üretici Karşıt Ağlar (DCGAN) ile klasik veri artırma tekniklerini birleştiren hibrit bir veri dengeleme stratejisi önerilmiştir. DenseNet121, ResNet50, VGG16, EfficientNetB0 ve ConvNeXt mimarileri; orijinal dengesiz veri seti ile yalnızca eğitim kümesine sentetik görüntüler eklenecek dengelenmiş veri seti üzerinde karşılaştırılmıştır. Böylece GAN tabanlı hibrit yaklaşımın “Doğruluk Paradoksu”nu aşmadaki ve özellikle azınlık sınıflarının tespit duyarlılığını (recall) artırarak klinik açıdan kritik lezyonların gözden kaçırılmasını azaltmadaki etkinliği sistematik olarak incelenmiştir.

1.3. Hipotez

Bu tez çalışmasında, dengesiz cilt lezyonu veri setlerinde GAN tabanlı veri artırımının, derin öğrenme modellerinin nadir ve klinik açıdan kritik azınlık sınıflarını tespit etme duyarlılığını artırabileceği ve bu sınıflardaki yanlış negatifleri (kaçırılan malign vakaları) azaltabileceği hipotezi savunulmaktadır. Bu kapsamda, DCGAN mimarisi ile klasik veri artırma tekniklerini birleştiren hibrit bir veri dengeleme stratejisi geliştirilmiş, DenseNet121, ResNet50, VGG16, EfficientNetB0 ve ConvNeXt modelleri hem orijinal dengesiz veri seti hem de yalnızca eğitim kümesinde sentetik görüntülerin eklendiği dengelenmiş veri seti üzerinde eğitilerek hipotezin doğruluğu test edilmiştir. Performans değerlendirmesinde doğruluk, Macro-Kesinlik, Macro-Duyarlılık ve Macro-F1 skorları birlikte kullanılmış; değerlendirmede, genel doğruluğun yanıltıcı olabildiği sınıf dengesizliği koşullarında özellikle azınlık sınıflarındaki duyarlılık ve yanlış

negatif oranları esas alınmıştır.

Çalışmanın giriş bölümünde, cilt kanseri, dermoskopi ve derin öğrenme uygulamalarına ilişkin genel bir çerçeve sunulmuş, ayrıca literatürde cilt lezyonu sınıflandırması, sınıf dengesizliği ve veri artırma teknikleri üzerine yapılan çalışmalar özetlenmiştir. Bu bölümde ayrıca çalışmanın amacı, önemi ve hipotezi ele alınmıştır. İkinci bölümde yapay zekâ, makine öğrenmesi, derin öğrenme, yapay sinir ağları ve derin öğrenme mimarileri ile performans değerlendirme yöntemlerini kapsayan kuramsal çerçeve sunulmuştur. Üçüncü bölümde çalışmada kullanılan HAM10000 veri seti, veri ön işleme ve bölümlleme adımları, önerilen hibrit veri dengeleme stratejisi (rastgele alt örnekleme, DCGAN tabanlı sentetik veri üretimi ve klasik veri artırma) ile deneysel kurulum süreci detaylandırılmıştır. Dördüncü bölümde, dengesiz ve dengeli veri setleri üzerinde elde edilen deneysel sonuçlar karşılaştırmalı olarak sunulmuş, “Doğruluk Paradoksu”, sınıf bazlı performanslar ve karmaşıklık matrisleri analiz edilmiştir. Son olarak beşinci bölümde çalışmanın sonuçları özetlenmiş, GAN tabanlı hibrit yaklaşımın katkıları ve sınırlılıkları tartışılmış ve gelecekte yapılabilecek çalışmalar için öneriler sunulmuştur.

2. KURAMSAL ÇERÇEVE

2.1. Yapay zekâ kavramı

Günümüzde yaşamın pek çok alanında etkisini giderek daha fazla hissettiğimiz yapay zekâ kavramının kökenleri aslında oldukça eskiye dayanmaktadır. Bilgiyi işleyerek tekrar eden görevleri daha hızlı ve verimli biçimde yerine getirmek amacıyla geliştirilen bilgisayarların hayatımıza girmesi, insanın işleri otomatikleştirme isteğini artırmış ve bu da yapay zekâ fikrinin önem kazanmasına yol açmıştır. Yapay zekânın temellerinin, felsefe, matematik, psikoloji, bilgisayar mühendisliği ve dilbilim gibi farklı disiplinlerde uzun zaman önce atıldığını söylemek mümkündür. “Makineler düşünebilir mi?” sorusunu ilk kez ortaya atan Alan Mathison Turing, İkinci Dünya Savaşı sırasında geliştirdiği elektromekanik araçlar sayesinde hem bilgisayar bilimine hem de yapay zekâ alanına önemli katkılar sağlamıştır. Bilgisayarlar, temel olarak insan beyninden esinlenerek tasarlanmış sistemlerdir. Bu sistemlerde veri, giriş birimleri aracılığıyla alınır, gerekli işlemlerden geçirilir ve sonunda istenen biçimde çıktı olarak sunulur. Bu süreçte yapay zekâ, verilerin işlenme biçimini geliştirmeyi ve insan benzeri düşünme ya da davranışlar sergileyebilen sistemler oluşturmayı hedeflemiştir.

Russell ve Norvig, bu doğrultuda zeki sistemler sınıflandırması şeklinde bir tasnife giderek, yapay zekâyı dört ana kategoriye ayırmaktadır. Bu sınıflandırma Şekil 2.1’de gösterilmiştir (Russell ve Norvig, 1995).

İnsan düşüncesiyle ilişkili olduğumuz faaliyetlerin otomasyonu, karar verme, problem çözme, öğrenme gibi faaliyetler... (Bellman, 1978) “Düşünen bilgisayarlar yapma konusundaki heyecan verici yeni çabalar” (Haugeland, 1985)		“Hesaplama modelleri kullanarak zihinsel yeteneklerin incelenmesi” (Charniak ve McDermott, 1985) “Algılama, mantık ve hareket etmeyi mümkün kılan hesaplamaların incelenmesi” (Winston, 1992)	
İnsan gibi düşünen sistemler		Akılcı düşünen sistemler	
İnsan gibi davranan sistemler		Akılcı davranan sistemler	
“İnsanların daha iyi olduğu şeyleri bilgisayarların nasıl yapabileceğine dair çalışma” (Rich ve Knight, 1991) “Zeki insanlar tarafından gerçekleştirilen işlevleri yerine getiren makineler yaratma sanatı” (Kurzweil, 1990)		“Akıllı davranışın otomasyonu ile ilgili olan bilgisayar bilimi kolu” (Luger ve Stubblefield, 1993) “Hesaplamalı süreçler açısından akıllı davranışları açıklamak ve taklit etmek isteyen bir çalışma alanı” (Schalkoff, 1990)	

Şekil 2.1. Zeki sistemler (Russell ve Norvig, 1995).

2.2. Makine öğrenmesi

Makine öğrenme süreci, verilerin işlenip değerlendirilmesinde başlayıp öncesinde karşılaştırılıp analizden geçirilen veriler vasıtasıyla karar verme safhasına kadar devam etmektedir. Makine öğrenmesi denetim kriteri baz alınarak 3 farklı kategoride sınıflandırılmıştır.

a) Denetimli Öğrenme (Supervised Learning): Söz konusu bu öğrenme yönteminde etiketlenmiş giriş ve çıkış verilerinin fonksiyonundan araştırmacı yardımı ile öğrenme işlevi yerine getirilir. Bu yaklaşım kullanılarak, güvenilir veri modelleri oluşturulur. Fakat bu durum büyük veri setleri için aynı zamanda yüksek maliyetli etiketleme zorluğunu da yanında getirir (Saba Raoof ve Durai, 2022). Denetimli öğrenme metotlarının uygulandığı algoritmalara örnek vermek gerekirse Destek Vektör Makineleri (Support Vector Machines) ve Rastgele Orman (Random Forest) verilebilir. Ayrıca Regresyon analizi (Regression Analysis) ve Karar Ağaçları (Decision Trees) da denetimli öğrenmenin uygulandığı diğer algoritma çeşitlerindendir.

b) Denetimsiz Öğrenme (Unsupervised Learning): Bu yöntemde model izlemesine ya da veri etiketlenmesi işlemlerine gereksinim duyulmaz. Kendi veri setinden veri bilgisi ve desenin keşfedilmeye odaklanıldığı bu yöntemde genellikle kümeleme işlemleri uygulanır. Etiket bilgisinin bulunmadığı bu modelde çıktı verilerinin doğruluk oranlarının düşük seviyede olması bir dezavantaj olsa da diğer modellere oranla daha az karmaşıklık barındırdığı için uygulanabilirliği yüksek seviyededir. Denetimsiz öğrenme metotlarının uygulandığı algoritmalara ilk olarak Temel Bileşenler Analizi (Principal Component Analysis) olmak üzere k-Ortalamalar Kümelemesi (k-Means Clustering) ve Kendi Kendine Organize Haritalar (Self Organizing Maps) algoritmaları örnek verilebilir.

c) Pekiştirmeli Öğrenme (Reinforcement Learning): Önceden etiketlenmiş veri gerektirmeyen bu öğrenme modelinde ajan, çevreyle etkileşimi sırasında deneme-yanılma ve ödül-ceza kavramları uygulanarak en yüksek ödülü alma hedefiyle çalıştırılır. Q-Learning, Asynchronous Advantage Actor-Critic (A3C), Policy Gradient Methods ve Actor-Critic Methods algoritmaları pekiştirmeli öğrenmeye örnek olarak verilebilir.

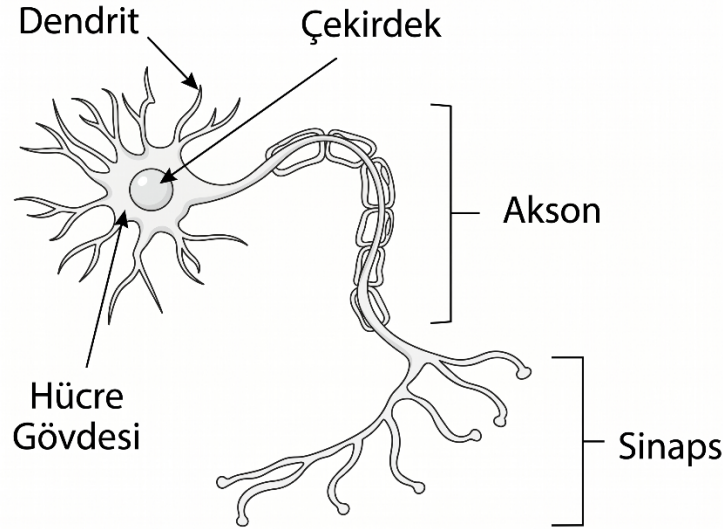
2.3. Derin öğrenme

Derin öğrenmede verilerin insan beyni tarafından ele alınıp işlenmesinden esinlenilmiştir. Derin öğrenme öncesinde, araştırmacı tarafından belirlenmesi gereken bir kural gereksinimi bulunmamaktadır. Ancak veriler arasında anlamlı bağlantılar kurulması için oldukça fazla veriye gereksinim vardır. Bu nedenle derin öğrenmede, önceki katmanların çıktısından elde edilen verilerin kullanıldığı çeşitli algoritmik sinir ağı katmanları oluşturulur (Alzubaidi vd., 2021).

2.4. Yapay sinir ağları

Biyolojik sinir sistemlerinin modellenmesi, sinir ağlarının temelini oluşturmuştur. Nöron, beynin en küçük hesaplama birimidir. Gelen sinyalleri iletmekle görevli akson nöronun çıkışını

oluşturur. Nöronlar arasında gelen sinyalleri toplayıp hücre girişini oluşturan yapı dendritlerdir. Tüm nöronlar arasında sağlıklı bir haberleşmenin kurulabilmesi için dendrit ve aksonlar birbirlerine sinaps ile bağlanmışlardır. Şekil 2.2’de yer alan biyolojik sinir hücresinin çiziminde akson, dendrit, çekirdek gibi yapılar yer almaktadır.



Şekil 2.2. Sinir hücresi biyolojik gösterimi.

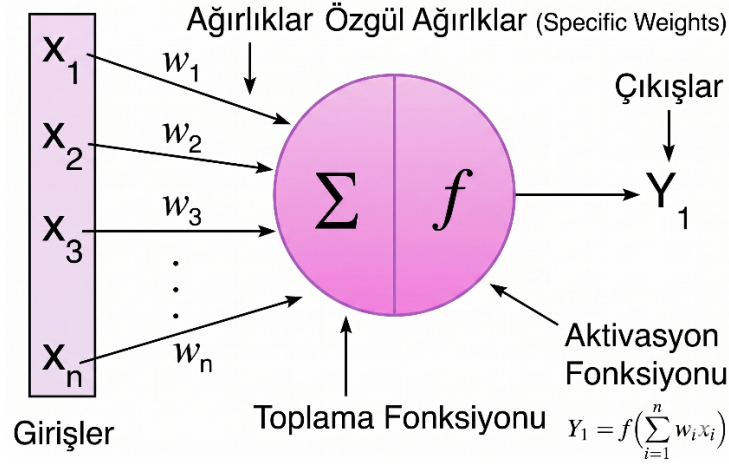
Yapay Sinir Ağının Modeli: Yapay sinir ağlarının en küçük birimi olan perseptron (algılayıcı), lineer bir fonksiyonla gösterilir. 1957 yılında Frank Rosenblatt tarafından algılayıcı tanımlanmıştır. Şekil 2.3’de gösterildiği gibi, bir algılayıcının çıktısı, girişlerin ağırlıklı toplamına bir aktivasyon fonksiyonu (f) uygulanmasıyla elde edilir:

$$Y_1 = f\left(\sum_{i=1}^n w_i x_i\right) \quad (F.1)$$

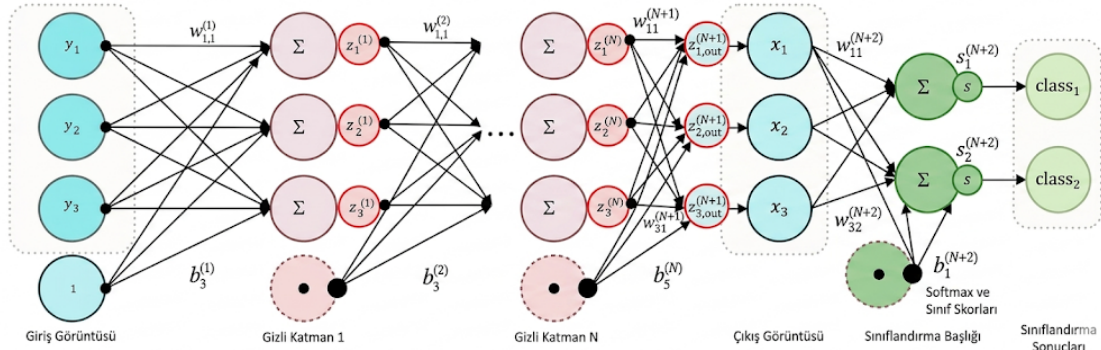
Burada x_i giriş değerlerini, w_i bu girişlere karşılık gelen ağırlıkları ve f aktivasyon fonksiyonunu göstermektedir.

Algılayıcılardan oluşan katmanlar yığını, yapay sinir ağları olarak bilinir. Her modül, verileri bir sonraki modüle aktarmak için farklı bir gösterime dönüştürür. Bu modüller, literatürde katmanlar olarak adlandırılır ve her bir katman çok sayıda birim veya nöron içerir (Lucas vd., 2018). Eğitim sırasında öğrenilen w ağırlık (weight) katsayısı ve b sapma (bias) katsayısı, Şekil 2.4’te gösterilen N gizli katmandan oluşan tam bağlaşımlı yapay sinir ağında yer almaktadır. Giriş birimine verilen girdilerin ağırlıklı toplamı, her katmanın çıktısıdır. Aynı katmandaki nöronlar arasında bağlantı yoktur ve diğer katmanlardaki nöronlar arasında ağırlık ve sapma katsayısı öğrenilir. Ardışık iki katman arasındaki nöronlar, birbirlerini çeşitli aktivasyon işlevleriyle etkiler. Modelin öğrenme düzeyi, bu fonksiyonların değerlerine göre belirlenir.

Tek Katmanlı Algılayıcılar (SLP) Modeli: Doğrusal sorunları çözmek amacıyla oluş-



Şekil 2.3. Sinir hücresi matematiksel gösterimi.



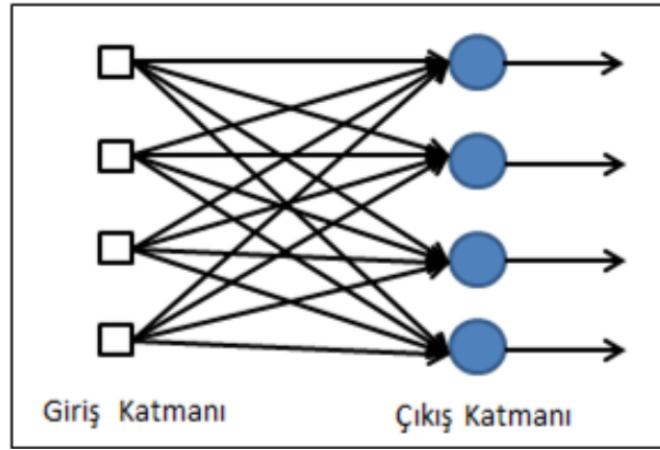
Şekil 2.4. Tam bağışlımlı yapay sinir örneđi (Lucas vd., 2018).

turulan tek katmanlı YSA'da girdi katmanı ve çıktı katmanı kullanılmaktadır. Söz konusu bu katmanlar bir tek nörondan oluşabileceđi gibi çok sayıda nöronu da içerebilir. Şekil 2.5'te basit düzeyde oluşturulan tek katmanlı algılayıcı gösterilmektedir (Arı ve Berberler, 2017).

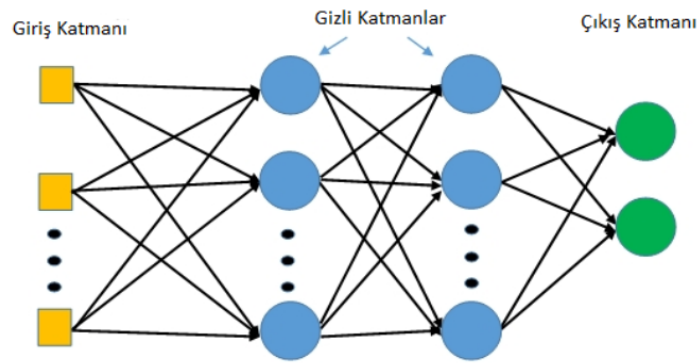
Çok Katmanlı Algılayıcılar (MLP) Modeli: Bir diđer yapay sinir ağı modeli olan çok katmanlı algılayıcı (MLP) ise, ayrıştırma gerçekleştirebilen modeldir. İleri beslemeli olan bu modelde, giriş ve çıkış katmanı arasında olmak üzere çok sayıda gizli katman yer almaktadır.

Söz konusu çok katmanlı algılayıcı modelin çalışma prensibini iki aşamada ele alabiliriz. İlk aşamada ağı uygulanan girişler, giriş katmanından çıkış katmanına doğru ileri yönde iletilir. Diđer aşamada ise, hataların geri yayılması algoritması uygulanır. Çıkış katmanına ulaşan gerçek değerler ile ağı tarafından tahmin edilen çıkış verileri arasındaki fark, algoritma tarafından işlenir ve ağı yeniden döndürölür.

Çok katmanlı algılayıcı modelinin doğrusal olmayan problemler üzerinde üstün bir çözme yeteneđi bulunmaktadır. Yüksek sayıda girdi ile yüksek sayıda çıktı verisi üretebilen bu model özellikle doğal dil işleme, görüntü tanıma gibi farklı alanlarda oldukça tercih edilen bir yapay sinir ağı modelidir. Hesaplama maliyetinin yüksek potansiyeli ve çok katmanlı yapısı modelin negatif yönleri arasında yer almaktadır. Çok sayıda parametre içermesi nedeniyle bu



Şekil 2.5. Tek katmanlı algılayıcı (Arı ve Berberler, 2017).

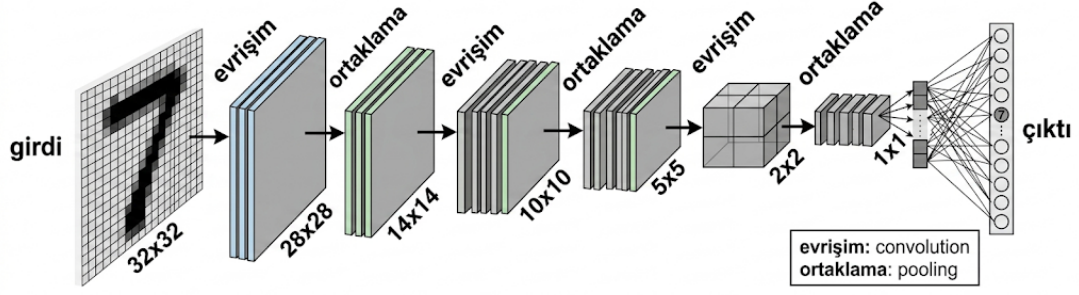


Şekil 2.6. Çok katmanlı algılayıcı (Arı ve Berberler, 2017).

parametrelerin uygulama öncesinde belirlenmesi gerekmektedir. Bu modelin eğitimi esnasında zaman zaman uyum problemleri yaşanabilir.

Evrişimli Sinir Ağları (CNN) Modeli: Derin ileri beslemeli bir sinir ağı olan CNN, genellikle görüntü işleme alanında nesnelerin tanımlanması, analiz edilmesi ve kategorizasyon gibi farklı alanlarda kullanılmaktadır. Gride benzer bir yapıya sahip olan bu sinir ağı modelinde matris çarpımı yerine katmanlar arasında evrişim ya da konvolüsyon olarak tanımlanan matematiksel bir işlem uygulanır. İlk katmanlarında daha basit ve temel niteliklerden meydana gelmektedir. Ancak katman sayısı ilerledikçe, derinleşen katmanlarda daha kompleks öznelikler elde edilir. Söz konusu bu kompleks öznelikler sayesinde görüntüler daha doğru ve net bir şekilde tasvir edilir (Kızrak ve Bolat, 2018). CNN, geleneksel sinir ağlarının aksine, sayısal veriler yerine matrisler kullanır. Bu modelde insan görsel sisteminin modellenmesi örnek alınmaktadır.

Aktivasyon Fonksiyonları: Yapay sinir ağlarında doğrusal olmayan öğrenmeyi mümkün kılan aktivasyon fonksiyonları, bir nöronun çıktısının hesaplanmasında belirleyici rol oynar. Evrişimli ağlarda en yaygın kullanılan fonksiyonlardan biri olan ReLU (Rectified Linear Unit), negatif girdileri sıfıra eşitlerken pozitif girdileri değiştirmeden geçirir (Eşitlik (F.2)). Bu davranış hesaplama açısından verimli olsa da, sürekli negatif girdi alan nöronların gradyanının sıfırlanarak öğrenmeyi durdurması anlamına gelen “ölü nöron” (dying ReLU) problemine yol açabilir.

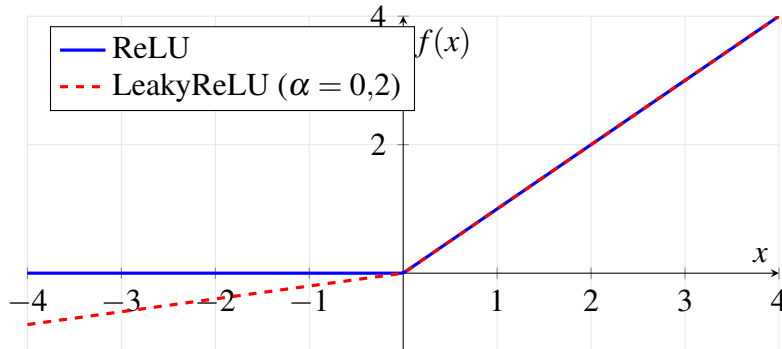


Şekil 2.7. CNN mimarisi (Kızrak ve Bolat, 2018).

LeakyReLU ise negatif bölgede küçük bir α eğimi tanımlayarak (Eşitlik (F.3)) bu sorunu hafifletir ve gradyan akışının sürdürülmesini sağlar. Bu nedenle, özellikle GAN mimarilerinin ayırt edici ağlarında LeakyReLU tercih edilmektedir. İki fonksiyonun karşılaştırması Şekil 2.8’de gösterilmiştir.

$$\text{ReLU}(x) = \max(0, x) \quad (\text{F.2})$$

$$\text{LeakyReLU}(x) = \begin{cases} x, & x \geq 0 \\ \alpha x, & x < 0 \end{cases} \quad (\text{F.3})$$



Şekil 2.8. ReLU ve LeakyReLU aktivasyon fonksiyonlarının karşılaştırması.

GAN (Generative Adversarial Networks) Mimari Yapısı: Generative Adversarial Networks (GAN), Ian Goodfellow ve ekibi tarafından 2014 yılında tanıtılan bir derin öğrenme mimarisidir (Goodfellow vd., 2014). İki temel bileşenden oluşur:

- **Üreteç (Generator – G):** $N(0, 1)$ veya $U(-1, 1)$ gibi bilinen bir dağılımdan örneklenen latent gürültü vektörünü (z) girdi olarak alır ve bunu gerçek verilere çok benzeyen sentetik örneklerle dönüştürür.

- **Ayırt Edici (Discriminator – D):** Gelen verinin gerçek mi yoksa üretici tarafından üretilmiş sahte mi olduğunu ayırt etmeye çalışır.

Bu iki ağ, adversarial (rekabetçi) bir şekilde eğitilir. Üreteç, ayırt ediciyi kandırmaya çalışırken; ayırt edici ise giderek daha iyi ayırt etme yeteneği kazanır. Bu süreç, ortak bir değer fonksiyonu üzerinde zıt yönlerde optimize edilen bir minimaks (minimax) oyunu olarak tanımlanır:

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log (1 - D(G(z)))] \quad (\text{F.4})$$

Burada $p_{\text{data}}(x)$ gerçek veri dağılımını, $p_z(z)$ latent gürültü dağılımını, $D(x)$ ayırt edicinin x örneğini gerçek olarak değerlendirme olasılığını ve $G(z)$ üreticinin z gürültüsünden ürettiği sentetik örneği göstermektedir. Ayırt edici D , gerçek örnekler için $\log D(x)$ terimini ve sahte örnekler için $\log(1 - D(G(z)))$ terimini birlikte maksimize etmeye çalışır. Sabit bir üreteç için optimum ayırt edici Eşitlik (F.5) ile verilir:

$$D^*(x) = \frac{p_{\text{data}}(x)}{p_{\text{data}}(x) + p_g(x)} \quad (\text{F.5})$$

Burada $p_g(x)$ üreteç tarafından üretilen örneklerin dağılımıdır. Üreteç ise ayırt ediciyi yanıltmak amacıyla $\log(1 - D(G(z)))$ terimini minimize etmeyi hedefler. Ancak eğitimin erken aşamalarında bu ifade zayıf gradyan ürettiğinden, pratikte üreteç Eşitlik (F.6)'deki doygun olmayan amaç fonksiyonunu maksimize edecek şekilde eğitilir:

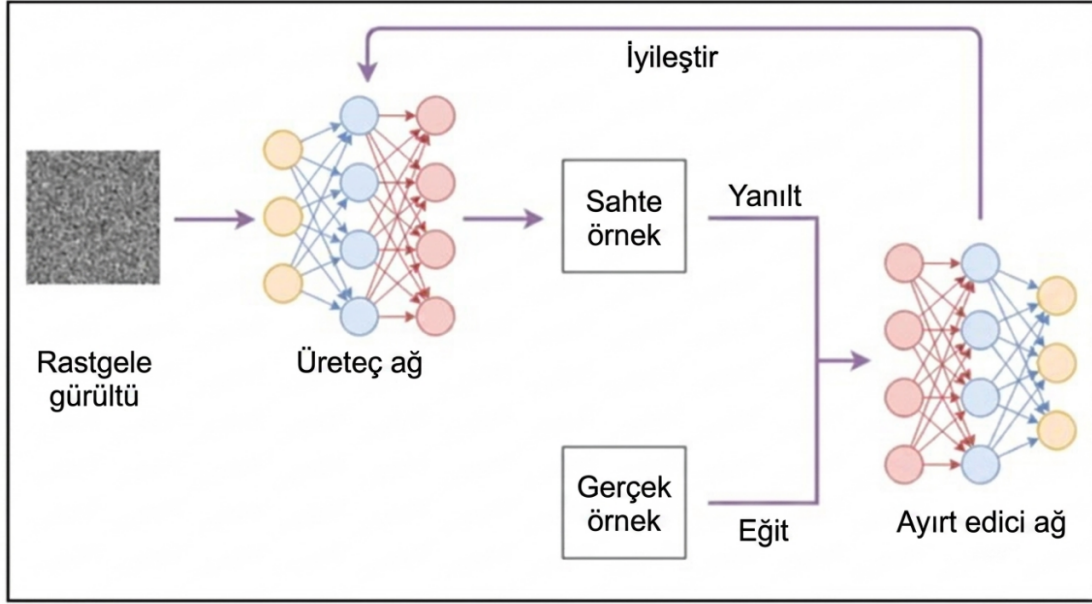
$$\max_G \mathbb{E}_{z \sim p_z(z)} [\log D(G(z))] \quad (\text{F.6})$$

Eğitim, p_g dağılımı gerçek veri dağılımı p_{data} 'ya yakınsayana kadar sürer; bu noktada ayırt edicinin çıktısı ideal olarak $D(x) = 0,5$ değerine ulaşır.

DCGAN (Deep Convolutional GAN) Mimari Özellikleri: DCGAN, klasik GAN'ın görüntü üretimi için geliştirilmiş ve daha stabil bir versiyonudur (Radford vd., 2015). Temel farkı, tamamen evrişimli katmanlar kullanmasıdır. DCGAN, Eşitlik (F.4)'te verilen minimaks amaç fonksiyonunu aynen benimser; farkı, üreteç ve ayırt edicinin derin evrişimli ağlar olarak tasarlanmasıdır. Model, kararlı bir eğitim için genellikle Adam optimizasyon algoritması ile ($\beta_1 = 0,5$ ve öğrenme oranı 0,0002) eğitilir. Radford vd. (2015), model kararsızlığını önlemek amacıyla Batch Normalization'ın üreticinin çıkış katmanı ile ayırt edicinin giriş katmanı hariç tüm katmanlara uygulanmasını önermektedir.

1. Üreteç (Generator) Mimarisi

- Giriş: 100 boyutlu rastgele gürültü vektörü (z)
- İlk adım: Project and Reshape (tam bağlantılı katman ile küçük bir tensöre dönüştürülür,

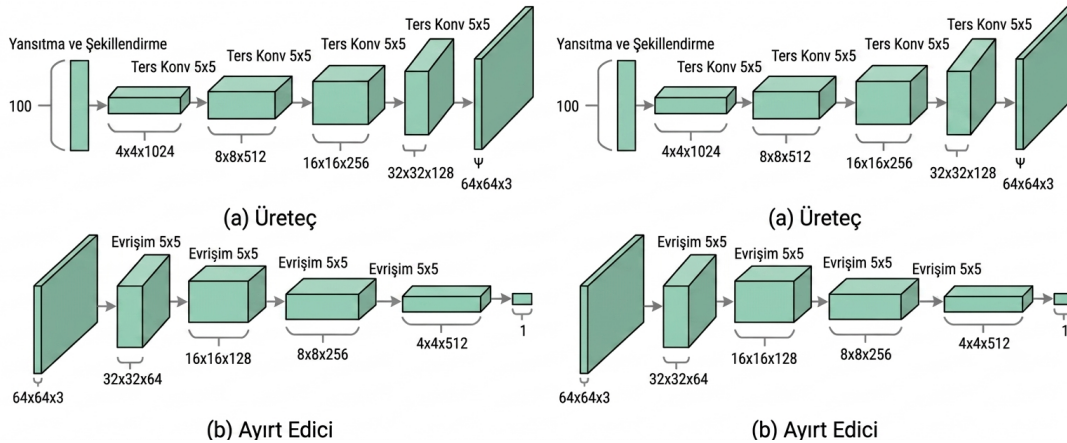


Şekil 2.9. GAN mimarisi (Goodfellow vd., 2014).

örneğin $4 \times 4 \times 1024$)

- Ardından Transposed Convolution (veya Fractionally-strided Convolution) katmanları ile görüntü boyutu adım adım büyütülür ($4 \times 4 \rightarrow 8 \times 8 \rightarrow 16 \times 16 \rightarrow 32 \times 32 \rightarrow 64 \times 64$)
- Çıkış katmanı hariç tüm katmanlarda Batch Normalization ve ReLU aktivasyon fonksiyonu uygulanır; çıkış katmanında ise Tanh aktivasyon fonksiyonu kullanılır.

Çıktı: $64 \times 64 \times 3$ (RGB görüntü) veya istenen boyutta sentetik görüntüdür.

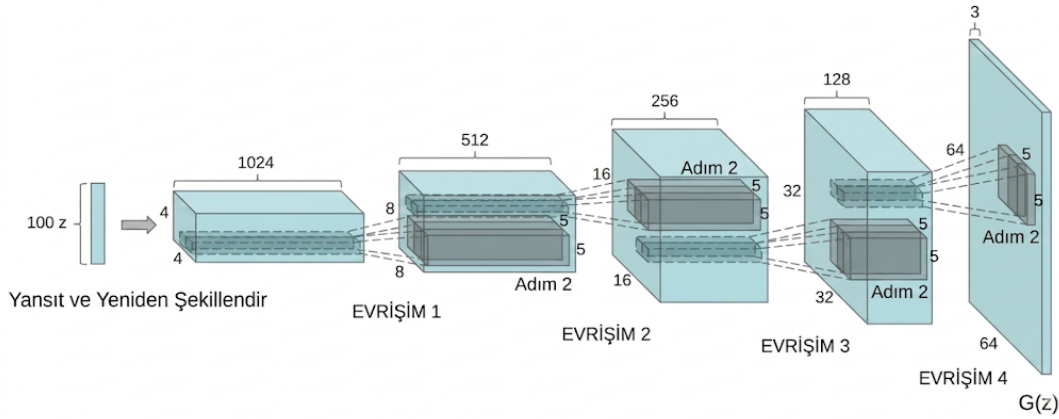


Şekil 2.10. DCGAN mimarisi (Radford vd., 2015).

Şekil 2.10'de DCGAN Üreteci (a) ve Ayırt Edici (b) mimarisi detaylı olarak gösterilmektedir. Üreteç, 100 boyutlu gürültüyü alıp adım adım ters evrişim (deconv) katmanlarıyla 64×64 piksel görüntüye dönüştürür.

2. Ayırt Edici (Discriminator) Mimarisi

- Giriş: $64 \times 64 \times 3$ gerçek veya sahte görüntü
- Convolution katmanları ile özellik çıkarımı yapılır (boyut küçültülür).
- Giriş katmanı hariç tüm katmanlarda Batch Normalization uygulanır; tüm katmanlarda LeakyReLU aktivasyon fonksiyonu (genellikle slope = 0,2) kullanılır.
- Pooling yerine Strided Convolution kullanılır (bilgi kaybını önler).
- Son katman: Sigmoid aktivasyonu ile “gerçek” (1) veya “sahte” (0) olasılığı verir.

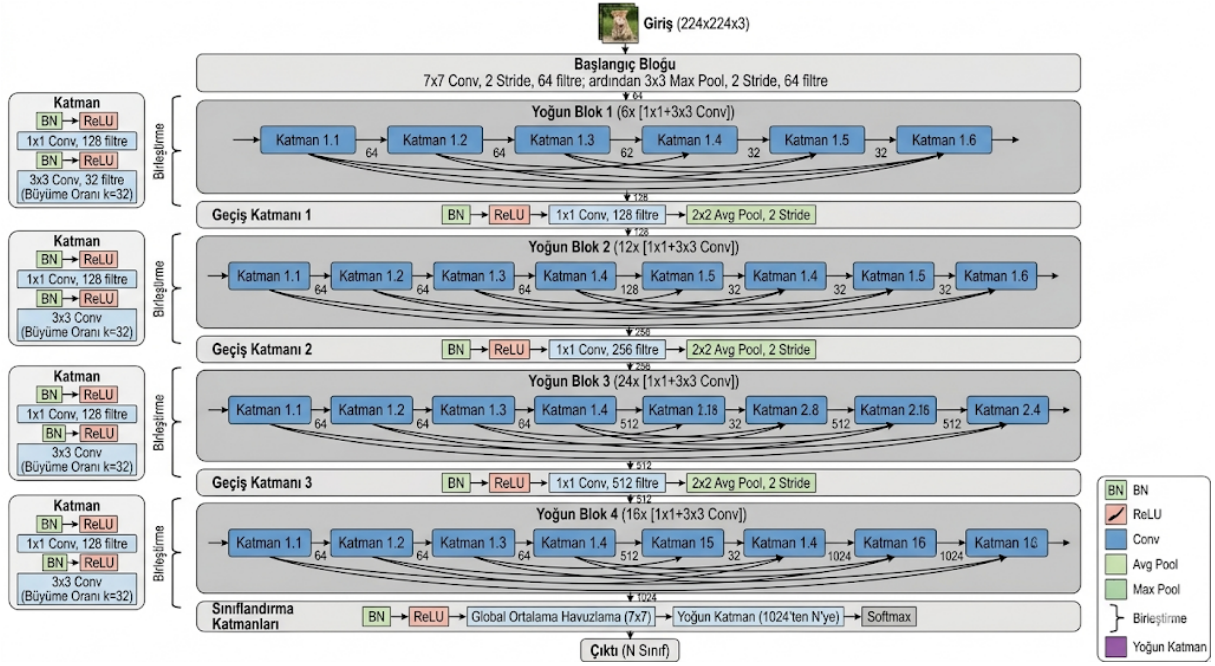


Şekil 2.11. LSUN sahne modellemesi için kullanılan DCGAN üretici mimarisi (Radford vd., 2015).

Şekil 2.11, orijinal DCGAN makalesinde (Radford vd., 2015) verilen ve LSUN veri seti üzerinde sahne görüntüleri (bedroom, church vb.) üretmek için tasarlanmış üretici mimarisini detaylı olarak göstermektedir. Şekil, DCGAN üreticinin katman detaylarını (stride = 2, filtre sayıları: 1024, 512, 256, 128 vb.) daha teknik olarak sunmaktadır.

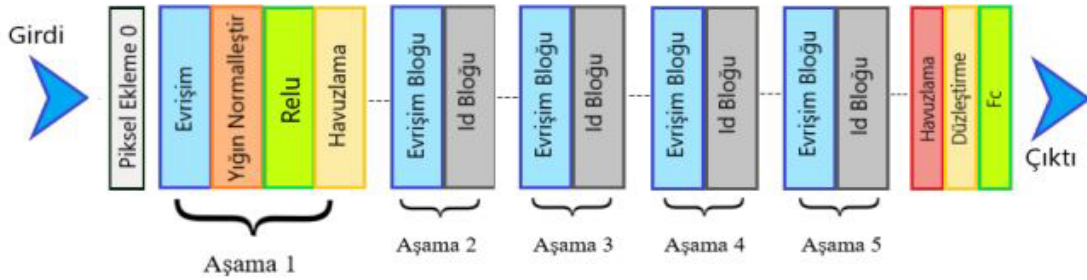
DenseNet121: DenseNet121, 2017 yılında önerilen bir CNN mimarisidir. Temel yeniliği yoğun bağlantı (dense connectivity) yapısıdır. Her katman, kendisinden önceki tüm katmanların çıktılarını (öznitelik haritalarını) girdi olarak alır. Bu sayede öznitelikler tekrar tekrar kullanılır, parametre sayısı azalır, gradyan akışı iyileşir ve vanishing gradient (kaybolan gradyan) sorunu büyük ölçüde çözülür. Mimari yapısı; başlangıçta 7×7 evrişim ve maxpooling, ardından katman sayıları 6, 12, 24 ve 16 olan 4 adet Dense Block, her Dense Block içinde Bottleneck katmanlar (1×1 Conv + 3×3 Conv) ve Transition Layer’lar arasında kanal sıkıştırma ile average pooling içerir. Toplamda 121 katman ve yaklaşık 8 milyon parametreden oluşur (Huang vd., 2017).

ResNet50: Derin ağlarda gradyan kaybını önleyen artık blok yapısına sahip ResNet50, 2015 yılında Kaiming He ve ekibi tarafından geliştirilen çığır açan bir mimaridir. Derin ağlarda yaşanan vanishing gradient sorununu çözmek için residual (artık) öğrenme kavramını getirmiştir. Ağ, hedef fonksiyonu doğrudan öğrenmek yerine, girdiye “kalan farkı” (residual) öğrenir. Mimari; Identity Block ve Bottleneck Block ($1 \times 1 \rightarrow 3 \times 3 \rightarrow 1 \times 1$ Conv) olmak üzere residual bloklardan oluşur; skip (shortcut) bağlantıları ile gradyan doğrudan akar ve ağ toplam 50 katman içerir.



Şekil 2.12. DenseNet121 mimarisi – Yoğun bağlantı yapısı (Huang vd., 2017).

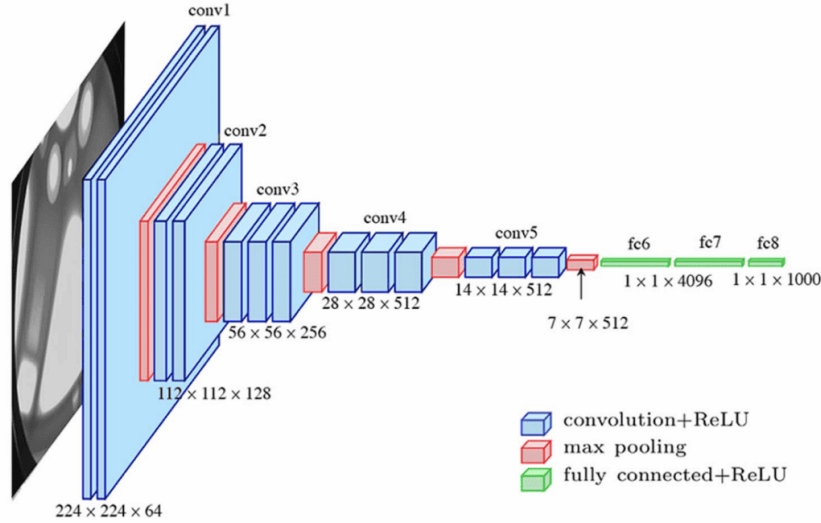
ResNet50, derinlik arttıkça performansın düşmemesini sağlar ve transfer öğrenmede en çok kullanılan omurga (backbone) ağlardan biridir (He vd., 2016).



Şekil 2.13. ResNet50 mimarisi (He vd., 2016).

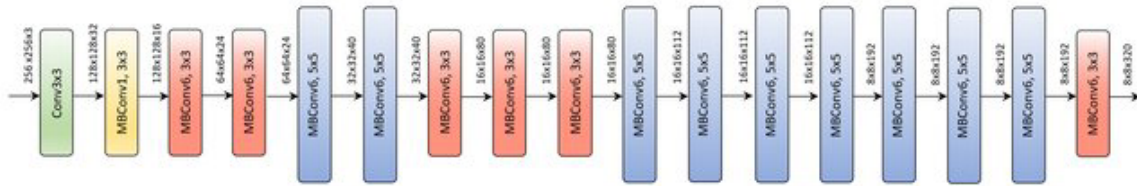
VGG16: VGG16, 2014 yılında Simonyan ve Zisserman tarafından önerilen klasik ve sade bir CNN mimarisidir (Simonyan ve Zisserman, 2015). “Derinlik önemlidir” prensibiyle küçük (3×3) filtreler kullanarak çok sayıda katman oluşturur ve güçlü bir öznetelik hiyerarşisi yakalar. Mimari; her blokta artan filtre sayısına ($64 \rightarrow 128 \rightarrow 256 \rightarrow 512 \rightarrow 512$) sahip 5 adet blok, her blokta birden fazla 3×3 evrişim ve 2×2 maxpooling içerir. Toplamda 13 evrişim katmanı ve 3 tam bağlantılı katman olmak üzere 16 katmandan oluşur; yüksek parametre sayısına (yaklaşık 138 milyon) rağmen basit ve tekrarlanabilir bir yapı sunar.

EfficientNetB0: EfficientNetB0, 2019 yılında geliştirilen, compound scaling yöntemiyle öne çıkan çok verimli bir mimaridir. Derinlik, genişlik ve çözünürlüğü aynı anda dengeli şekilde ölçeklendirilerek yüksek doğruluk ile düşük hesaplama maliyeti sağlar. Temel bloğu MBConv (Mobile Inverted Bottleneck Convolution) ve Squeeze-and-Excitation dikkat mekanizmasından oluşur; compound scaling ile derinlik, genişlik ve çözünürlük birlikte ölçeklenir. Ailenin en



Şekil 2.14. VGG16 mimarisi – Blok yapısı (Simonyan ve Zisserman, 2015).

küçük versiyonu olan EfficientNetB0 yaklaşık 5,3 milyon parametre içerir ve düşük FLOPs değeriyle kaynak verimliliği yüksektir (Tan ve Le, 2019).



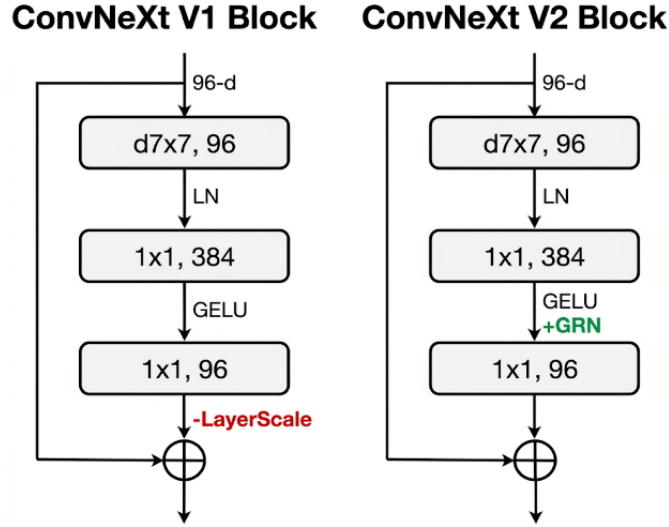
Şekil 2.15. EfficientNetB0 mimarisi (Tan ve Le, 2019).

ConvNeXt (Tiny): ConvNeXt Tiny, 2022 yılında Meta AI tarafından tanıtılan modern bir CNN mimarisidir. Vision Transformer’ların güçlü yönlerini (büyük çekirdek, Layer Norm, inverted bottleneck vb.) klasik CNN’e uyarlayarak Transformer seviyesinde performans elde ederken CNN’in verimliliğini korur. Mimari; patch stem (4×4 konvolüsyon), ResNet benzeri 4 aşamalı yapı, 7×7 depthwise convolution ile inverted bottleneck bloklarından oluşur. Batch Normalization yerine Layer Normalization, ReLU yerine GELU kullanılır; Layer Scale ve daha az aktivasyon gibi modern iyileştirmeler içerir (Liu vd., 2022).

2.5. Performans değerlendirme yöntemleri

Geliştirilen modellerin sınıflandırma performansı; Doğruluk (Accuracy), Makro-Kesinlik (Macro-Precision), Makro-Duyarlılık (Macro-Recall) ve Makro-F1 Skor metrikleri kullanılarak değerlendirilmiştir. Sınıf dengesizliğinin bulunduğu veri setlerinde, çoğunluk sınıfının etkisini dengelemek ve her bir sınıfın başarımını eşit ağırlıkla temsil etmek amacıyla “Makro Ortalama” yöntemi tercih edilmiştir.

Toplam C adet sınıfın bulunduğu bir veri setinde, her bir i sınıfı için Doğru Pozitif (TP_i), Yanlış Pozitif (FP_i) ve Yanlış Negatif (FN_i) değerleri hesaplanmıştır. Kullanılan metriklerin



Şekil 2.16. ConvNeXt blok yapısı (V1 ve V2) (Liu vd., 2022).

matematiksel ifadeleri aşağıdaki eşitliklerde sunulmuştur.

Doğruluk: Modelin tüm sınıflar üzerindeki genel doğru tahmin oranını, yani doğru sınıflandırılan örnek sayısının toplam örnek sayısına (N) oranını ifade eder:

$$\text{Doğruluk} = \frac{1}{N} \sum_{i=1}^C TP_i = \frac{\sum_{i=1}^C TP_i}{\sum_{i=1}^C (TP_i + FP_i)} \quad (\text{F.7})$$

Makro-Kesinlik: Her bir sınıf için hesaplanan kesinlik değerlerinin aritmetik ortalamasıdır. Modelin pozitif tahminlerinin ne kadarının doğru olduğunu gösterir:

$$\text{Makro-Kesinlik} = \frac{1}{C} \sum_{i=1}^C \frac{TP_i}{TP_i + FP_i} \quad (\text{F.8})$$

Makro-Duyarlılık: Her bir sınıf için hesaplanan duyarlılık değerlerinin aritmetik ortalamasıdır. Modelin ilgili sınıfa ait gerçek örnekleri ne oranda tespit edebildiğini ölçer:

$$\text{Makro-Duyarlılık} = \frac{1}{C} \sum_{i=1}^C \frac{TP_i}{TP_i + FN_i} \quad (\text{F.9})$$

Makro-F1 Skor: Kesinlik ve duyarlılık metriklerinin harmonik ortalaması olan F1-skorunun, tüm sınıflar için ortalamasını ifade eder. Sınıf dengesizliği durumunda modelin genel

başarısını gösteren en güvenilir metrik olarak kabul edilir:

$$\text{Makro-F1} = \frac{1}{C} \sum_{i=1}^C \frac{2 \times \text{Kesinlik}_i \times \text{Duyarlılık}_i}{\text{Kesinlik}_i + \text{Duyarlılık}_i} \quad (\text{F.10})$$

3. MATERYAL VE METOT

3.1. Veri seti analizi

Tez kapsamında, pigmentli deri lezyonlarının sınıflandırılmasına yönelik olarak literatürde geniş kabul görmüş olan HAM10000 (“Human Against Machine with 10000 training images”) veri setinden yararlanılmıştır. Söz konusu veri seti, farklı coğrafi popülasyonlardan ve çeşitli görüntüleme teknikleri kullanılarak elde edilmiş toplam 10.015 dermoskopik görüntüyü içermektedir. Bu yönüyle veri seti hem çeşitlilik hem de klinik temsil gücü açısından güçlü bir kaynak sunmaktadır.

HAM10000 veri seti, yedi farklı tanı sınıfını kapsamaktadır. Bu sınıflar sırasıyla; melano-sitik nevüs (nv), melanom (mel), benign keratoz benzeri lezyonlar (bkl), bazal hücreli karsinom (bcc), aktinik keratoz (akiec), vasküler lezyonlar (vasc) ve dermato fibroma (df) olarak tanımlanmaktadır. Veri setinde yer alan örneklerin önemli bir kısmı güvenilir tanı yöntemleri ile doğrulanmıştır. Özellikle tanılarının %50’sinden fazlası histopatolojik incelemelerle kesinleştirilmiş olup, geri kalan veriler ise uzman hekimlerin konsensüsü veya in-vivo konfokal mikroskopi gibi ileri görüntüleme teknikleri aracılığıyla doğrulanmıştır. Bu durum, veri setinin doğruluk ve güvenilirlik düzeyini artırmaktadır (Tschandl vd., 2018).

Model geliştirme sürecinde veri seti; eğitim, doğrulama ve test olmak üzere üç ayrı alt kümeye bölünmüştür. Eğitim veri kümesi, modelin temel öğrenme sürecini gerçekleştirmesi amacıyla kullanılmıştır. Doğrulama veri kümesi ise hiperparametre ayarlamaları, model optimizasyonu ve erken durdurma (early stopping) gibi süreçlerde performansın izlenmesi için değerlendirilmiştir. Test veri kümesi ise modelin daha önce görmediği veriler üzerindeki genelleme yeteneğini tarafsız bir şekilde ölçmek amacıyla ayrılmıştır. Bu kapsamda veri setinin %70’i eğitim, %15’i doğrulama ve %15’i test verisi olarak kullanılmıştır. Bu yapılandırma sayesinde modelin hem öğrenme süreci etkin bir şekilde yönetilmiş hem de elde edilen sonuçların güvenilirliği ve genellenebilirliği artırılmıştır.

3.1.1. Veri ön işleme ve bölümlendirme

Modelin genelleme performansını tarafsız ve güvenilir bir biçimde değerlendirebilmek, aynı zamanda eğitim verisinin test aşamasına herhangi bir şekilde sızmasını önlemek amacıyla bu çalışmada “Önce Ayır, Sonra İşle” yaklaşımı benimsenmiştir. Bu yaklaşım, veri ön işleme adımlarının veri bölme işleminden sonra gerçekleştirilmesini esas alarak, veri sızıntısı (data leakage) riskini minimize etmeyi hedeflemektedir.

İlk aşamada, veri seti sınıflar arasındaki dağılım dengesi korunacak şekilde üç alt kümeye ayrılmıştır. Bu doğrultuda verilerin %70’i eğitim kümesi, %15’i doğrulama kümesi ve kalan %15’i test kümesi olarak belirlenmiştir. Bu bölme işlemi sırasında stratified (katmanlı) örnek-

leme yöntemi kullanılarak her bir sınıfın tüm alt kümelerde benzer oranlarda temsil edilmesi sağlanmıştır. Böylece modelin farklı veri parçaları üzerinde dengeli ve tutarlı bir performans sergilemesi amaçlanmıştır.

İkinci aşamada, görüntü ön işleme adımları uygulanmıştır. Tüm dermoskopik görüntüler, kullanılan derin öğrenme mimarilerinin giriş boyut gereksinimlerini karşılamak amacıyla 224×224 piksel çözünürlüğe yeniden boyutlandırılmıştır. Bunun yanı sıra, modelin daha hızlı ve kararlı bir şekilde öğrenmesini sağlamak amacıyla görüntülere kanal bazlı normalizasyon işlemi uygulanmıştır. Bu kapsamda, yaygın olarak kullanılan ImageNet veri setine ait ortalama ($\mu = [0,485, 0,456, 0,406]$) ve standart sapma ($\sigma = [0,229, 0,224, 0,225]$) değerleri esas alınmıştır. Bu normalizasyon işlemi, piksel değerlerinin standart bir dağılıma çekilmesine yardımcı olarak model eğitim sürecinde daha stabil ve verimli bir optimizasyon sağlamaktadır. Bu iki aşamalı süreç sayesinde hem veri bütünlüğü korunmuş hem de modelin gerçek dünya verileri üzerindeki performansını daha doğru yansıtacak bir değerlendirme altyapısı oluşturulmuştur.

3.1.2. Önerilen hibrit veri dengeleme stratejisi

Sınıf dağılımının dengesiz olduğu veri setlerinde, özellikle dermato fibroma (df) ve vasküler lezyonlar (vasc) gibi azınlık sınıfların model tarafından yeterince öğrenilememesi önemli bir sorun teşkil etmektedir. Bu sorunun önüne geçebilmek ve modelin tüm sınıflar üzerinde dengeli bir performans sergilemesini sağlamak amacıyla, bu çalışmada yalnızca eğitim veri kümesi üzerinde uygulanan üç aşamalı hibrit bir dengeleme stratejisi benimsenmiştir. Bu yaklaşımda, doğrulama ve test veri kümelerine kesinlikle müdahale edilmemiş, böylece model değerlendirme sürecinin objektifliği korunmuştur.

Uygulanan dengeleme stratejisinin temel amacı, eğitim veri setinde yer alan her bir sınıfın örnek sayısını eşitleyerek $N = 1000$ seviyesine ulaştırmaktır. Bu doğrultuda üç bileşenden oluşan bir yaklaşım izlenmiştir. İlk olarak, aşırı temsil edilen çoğunluk sınıfı (melanositik nevüs, nv) için kontrollü örnek azaltma (undersampling) uygulanarak örnek sayısı 1.000'e indirilmiş; böylece modelin baskın sınıfa aşırı uyum (overfitting) göstermesi engellenmiştir.

İkinci ve temel bileşen olarak, az temsil edilen her azınlık sınıfı için DCGAN ile sentetik dermoskopik görüntüler üretilmiştir. Bu adımda, her sınıf için mevcut gerçek örnek sayısını 1.000'e tamamlayacak miktarda sentetik görüntü üretilmiş; yani sınıf başına (1000 – gerçek örnek sayısı) kadar sentetik görüntü oluşturulmuştur. Dolayısıyla azınlık sınıflarının dengelenmesinde asıl katkı sentetik veri üretiminden gelmektedir.

Üçüncü ve tamamlayıcı bileşen olarak, sentetik üretim sonrasında 1.000 örneğe ulaşamayan sınıflardaki küçük artık boşluk, gerçek görüntülerin klasik veri artırma (data augmentation) teknikleriyle çoğaltılmasıyla tamamlanmıştır. Uygulanan dönüşümler döndürme, yatay/dikey yansıtma ve renk/parlaklık değişimlerini içermektedir. Bu adım yalnızca birkaç görüntülük eksiklikleri gidermeye yönelik bir tamamlama işlevi gördüğünden, Tablo 3.1'de görüleceği üzere

klasik veri artırmanın toplam katkısı oldukça sınırlı (sınıf başına 0–11 görüntü) kalmıştır. Bu üç bileşenli yaklaşım sayesinde eğitim setinde her sınıfın tam olarak 1.000 örnekle temsil edildiği dengeli bir dağılım elde edilmiştir.

Uygulanan hibrit dengeleme stratejisinin her sınıf için oluşturduğu nihai kompozisyon Tablo 3.1’de ayrıntılı olarak verilmiştir. Tabloda görüldüğü üzere, azınlık sınıflarının 1.000 örneğe tamamlanmasındaki baskın katkı DCGAN ile üretilen sentetik görüntülerden gelmekte; klasik veri artırma ise yalnızca kalan küçük boşluğu doldurmaktadır. Çoğunluk sınıfı olan nv ise rastgele alt örnekleme ile 1.000 örneğe indirilmiştir. Böylece eğitim setindeki toplam örnek sayısı 7.000 olacak şekilde tam bir sınıf dengesi sağlanmıştır.

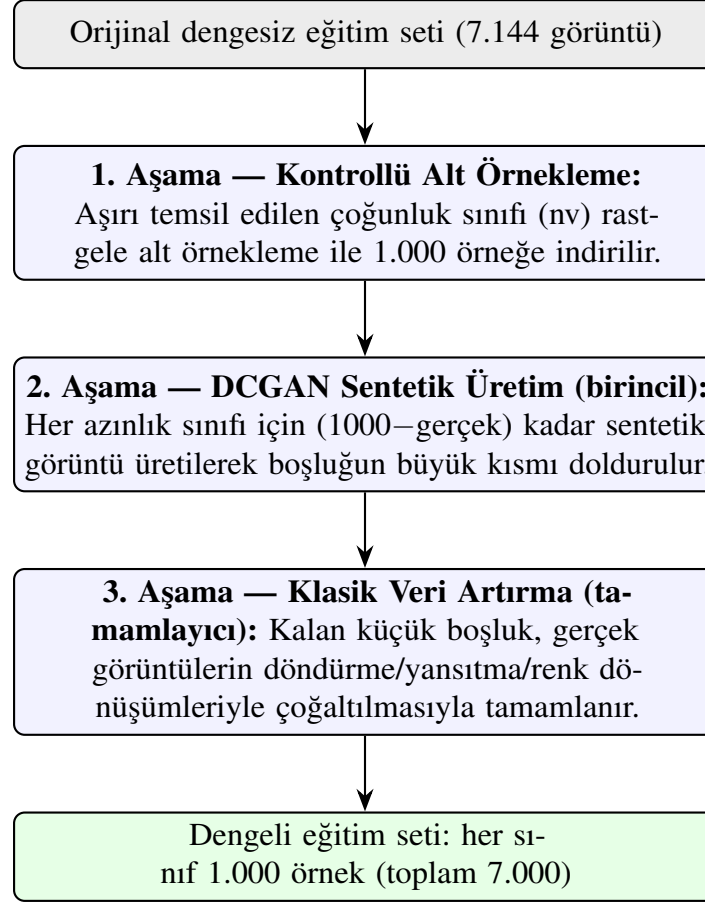
Tablo 3.1. Eğitim setinin sınıf bazında dengeleme kompozisyonu.

Sınıf	Orijinal	Klasik Artırma	DCGAN Sentetik	Dengeli	Yöntem
nv	4.784	–	–	1.000	Alt Örnekleme
mel	794	0	206	1.000	GAN
bkl	782	2	216	1.000	GAN + Klasik
bcc	366	11	623	1.000	GAN + Klasik
akiec	234	0	766	1.000	GAN
vasc	102	1	897	1.000	GAN + Klasik
df	82	4	914	1.000	GAN + Klasik
Toplam	7.144	18	3.622	7.000	–

Önerilen üç aşamalı hibrit veri dengeleme stratejisinin genel işleyişi Şekil 3.1’te akış şeması olarak özetlenmiştir.

3.1.3. Rastgele alt örnekleme

Veri setinde belirgin biçimde yüksek temsiliyete sahip olan ve bu nedenle modelin öğrenme sürecinde yanlılık geliştirmesine yol açabilecek çoğunluk sınıfları, özellikle melanositik nevüs (nv), kontrollü bir şekilde azaltılmıştır. Bu kapsamda, söz konusu sınıfa ait örnekler rastgele örnekleme (random under sampling) yöntemi kullanılarak 1.000 adet ile sınırlandırılmıştır. Gerçekleştirilen bu işlem, veri setindeki sınıf dengesizliğini azaltarak modelin yalnızca baskın sınıfa yönelmesini engellemeyi hedeflemektedir. Aksi durumda model, çoğunluk sınıfına ait görüntüleri daha kolay öğrenerek diğer sınıfları ihmal edebilir ve bu da genelleme performansının düşmesine neden olabilir. Rastgele seçim yöntemi sayesinde veri çeşitliliği mümkün olduğunca korunurken, aynı zamanda örnek sayısı dengelenmiştir. Bu yaklaşım, modelin tüm sınıflara daha eşit ağırlık vermesini teşvik ederek sınıflandırma performansının dengeli bir şekilde artmasına katkı sağlamıştır. Böylece özellikle azınlık sınıflar üzerindeki duyarlılık (sensitivity) ve genel doğruluk (accuracy) metriklerinde iyileşme hedeflenmiştir.



Şekil 3.1. Önerilen üç aşamalı hibrit veri dengeleme stratejisinin akış şeması.

3.2. Uygulama değerlendirme

3.2.1. Kullanılan model (GAN tabanlı sentetik veri üretimi – DCGAN)

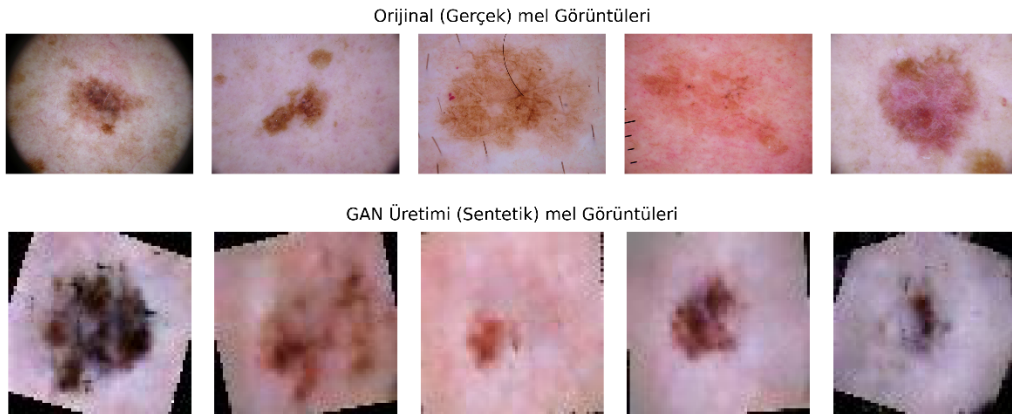
Veri setinde yeterli sayıda temsil edilmeyen azınlık sınıflarının (melanom (mel), bazal hücreli karsinom (bcc) ve aktinik keratoz (akiec)) model tarafından daha etkin öğrenilebilmesini sağlamak amacıyla, bu çalışmada Derin Evrişimli Üretici Karşıt Ağlar (DCGAN) tabanlı bir sentetik veri üretim yaklaşımı kullanılmıştır. Bu yöntem, sınırlı sayıdaki gerçek örneklerin ötesine geçerek veri çeşitliliğini artırmayı ve modelin genelleme yeteneğini güçlendirmeyi hedeflemektedir.

Önerilen DCGAN mimarisi iki temel bileşenden oluşmaktadır: üretici (generator) ve ayırt edici (discriminator) ağlar. Üretici ağ, 100 boyutlu rastgele bir latent gürültü vektörünü giriş olarak almakta ve bu vektörü aşamalı olarak işleyerek sentetik dermoskopik görüntüler üretmektedir. Bu süreçte, özellikle ters evrişim (ConvTranspose2d) katmanları kullanılarak düşük boyutlu temsil uzayından daha yüksek boyutlu görüntü uzayına geçiş sağlanmıştır. Katmanlar boyunca gerçekleştirilen yukarı örnekleme (upsampling) işlemleri sonucunda, 64×64 piksel çözünürlüğünde yeni görüntüler elde edilmiştir. Üretici ağın amacı, gerçek görüntülere mümkün olduğunca benzer ve ayırt edilmesi zor örnekler üretmektir.

Ayırt edici ağ ise hem gerçek hem de üretici tarafından oluşturulan sentetik görüntüleri giriş olarak alarak bunları doğru şekilde sınıflandırmayı öğrenmektedir. Bu ağ, evrişim katmanları ve doğrusal olmayan aktivasyon fonksiyonları (özellikle LeakyReLU) kullanılarak tasarlanmıştır. LeakyReLU tercih edilmesinin temel nedeni, klasik ReLU fonksiyonlarında görülebilen “ölü nöron” problemini azaltmak ve gradyan akışını daha sağlıklı bir şekilde sürdürmektir (bkz. Şekil 2.8). Ayırt edici ağın temel görevi, gerçek ve sahte veriler arasındaki farkı öğrenerek üretici ağı daha gerçekçi örnekler üretmeye zorlamaktır.

Modelin eğitim sürecinde karşılaşılabilecek kararsızlık problemlerini minimize etmek ve özellikle üretici ağın sınırlı sayıda örneğe odaklanarak çeşitliliği kaybetmesi anlamına gelen mod çökmesi riskini azaltmak amacıyla çeşitli iyileştirme teknikleri uygulanmıştır. Bu kapsamda, eğitim stabilitesini artırmak için Tek Yönlü Etiket Yumuşatma (One-Sided Label Smoothing) yöntemi kullanılmıştır. Bu teknikte, ayırt edici ağın gerçek veri için kullandığı etiket değeri 1,0 yerine 0,9 olarak belirlenmiş, böylece modelin aşırı güven geliştirmesi engellenmiştir. Bu yaklaşım, aynı zamanda gradyan sönmesi problemini azaltarak üretici ağın daha verimli bir şekilde güncellenmesine katkı sağlamıştır. Buna ek olarak, eğitim sürecinde giriş verilerine anlık veri artırma uygulanmıştır. Bu sayede modelin belirli örnekleri ezberlemesi önlenmiş, daha genelleştirilebilir ve çeşitli özellikleri öğrenmesi teşvik edilmiştir.

Önerilen DCGAN mimarisi kullanılarak elde edilen sentetik dermoskopik görüntüler incelendiğinde, modelin lezyonlara ait morfolojik ve yapısal özellikleri başarılı bir şekilde öğrendiği gözlemlenmiştir. Özellikle renk dağılımı, doku yapısı ve lezyon sınırları gibi kritik görsel özelliklerin, üretilen görüntülerde gerçeğe oldukça yakın biçimde temsil edildiği görülmüştür. Elde edilen sentetik görüntüler ile orijinal gerçek görüntüler arasındaki görsel karşılaştırmalar Şekil 3.2’de sunulmuş olup, iki veri türü arasındaki benzerlikler açık bir şekilde ortaya konmuştur.



Şekil 3.2. Orijinal ve GAN ürettiği melanom örnekleri.

Sentetik görüntülerin çeşitliliği Fréchet Inception Distance (FID) metriği ile değerlendirilmiş ve sınıflar bazında 349,23 ile 436,42 arasında değişen skorlar elde edilmiştir. Bu değerlerin görece yüksek çıkması, GAN modelinin hesaplama verimliliği adına 64×64 piksel çözünürlüğünde üretim yapması ve FID hesaplamasında kullanılan InceptionV3 ağının giriş boyutuna

yapılan ölçeklendirmenin neden olduğu bulanıklıktan kaynaklanmaktadır.

3.2.2. Klasik veri artırma ile boşluk doldurma

GAN tabanlı veri üretim sürecinin bazı sınıflar için yeterli düzeyde yakınsama sağlamaması ya da mevcut örnek sayısının son derece sınırlı olması nedeniyle, özellikle dermatofibroma (df) ve vasküler lezyonlar (vasc) sınıflarında alternatif bir veri çoğaltma stratejisine başvurulmuştur. Bu kapsamda, söz konusu sınıflarda veri çeşitliliğini artırmak amacıyla klasik veri artırma (data augmentation) teknikleri uygulanmıştır. Uygulanan veri artırma yöntemleri arasında görüntülerin belirli açılarla döndürülmesi, yatay ve dikey eksenlerde yansıtılması (flip işlemleri) ve parlaklık seviyelerinin değiştirilmesi gibi temel fakat etkili dönüşümler yer almaktadır. Bu işlemler sayesinde mevcut sınırlı sayıdaki görüntüden farklı varyasyonlar elde edilerek modelin daha zengin ve çeşitli bir veri dağılımı üzerinden öğrenmesi sağlanmıştır. Aynı zamanda bu yaklaşım, modelin belirli örneklerle aşırı uyum göstermesini (overfitting) engellemeye yardımcı olmuştur.

Tez kapsamında benimsenen hibrit dengeleme stratejisi kapsamında, hem GAN tabanlı sentetik veri üretimi hem de klasik veri artırma teknikleri birlikte kullanılarak eğitim veri setindeki tüm sınıfların örnek sayısı dengelenmiştir. Böylece her bir sınıfın eşit temsil edildiği bir eğitim kümesi oluşturulmuştur. Buna karşın, doğrulama ve test veri kümeleri üzerinde herhangi bir dengeleme ya da veri artırma işlemi uygulanmamış, bu kümeler orijinal halleriyle (dengesiz ve gerçek veri dağılımını yansıtacak şekilde) korunmuştur. Bu yaklaşım, model performansının gerçek dünya koşullarına daha yakın ve tarafsız bir biçimde değerlendirilmesini sağlamaktadır. Eğitim setinin dengeleme öncesi ve sonrası sınıf bazlı dağılımları ise Tablo 3.1’de ayrıntılı olarak verilmiştir.

3.2.3. Sınıflandırma modelleri

Dengelenmiş veri seti üzerinde, farklı mimari yaklaşımları temsil eden ve literatürde yaygın olarak kullanılan beş farklı derin öğrenme modeli eğitilmiştir. Bu modeller hem klasik CNN tasarımlarını hem de daha modern ve optimize edilmiş mimari yaklaşımları kapsayacak şekilde seçilmiştir. Böylece, farklı mimari tasarım felsefelerinin sınıflandırma performansı üzerindeki etkileri karşılaştırmalı olarak analiz edilebilmiştir.

İlk olarak yoğun bağlantı yapısıyla dikkat çeken DenseNet121 modeli kullanılmıştır. Bu mimari, her katmanın kendisinden önceki tüm katmanlardan doğrudan bağlantılar almasını sağlayarak özellik tekrar kullanımını mümkün kılmakta ve parametre verimliliğini artırmaktadır. İkinci olarak derin ağlarda sıkça karşılaşılan gradyan sönmesi problemini çözmek amacıyla geliştirilen ResNet50 modeli tercih edilmiştir. Bu mimaride yer alan artık (residual) bloklar, kimlik eşleme (identity mapping) sayesinde bilginin katmanlar arasında doğrudan aktarılmasına olanak tanımakta ve çok derin ağların daha stabil bir şekilde eğitilmesini sağlamaktadır. Üçüncü model olarak VGG16 kullanılmıştır. VGG16, sade ve düzenli yapısıyla bilinen, ardışık evrişim

katmanlarından oluşan klasik bir CNN mimarisidir. Küçük boyutlu filtreler (3×3) kullanılarak derinleştirilen bu yapı, görsel özelliklerin hiyerarşik olarak öğrenilmesine olanak tanımaktadır. Dördüncü olarak parametre verimliliği ve ölçeklenebilirlik açısından öne çıkan EfficientNetB0 modeli tercih edilmiştir. Bu mimari, derinlik, genişlik ve çözünürlük boyutlarını dengeli bir şekilde ölçeklendiren bileşik ölçekleme (compound scaling) yöntemi sayesinde daha az hesaplama maliyeti ile yüksek performans elde etmeyi hedeflemektedir. Son olarak modern mimari yaklaşımları temsil eden ConvNeXt (Tiny) modeli kullanılmıştır. Bu model, Vision Transformer (ViT) mimarisinin tasarım prensiplerini geleneksel CNN yapısına entegre ederek daha güncel ve güçlü bir yapı sunmaktadır.

Tüm bu modellerde transfer öğrenme (transfer learning) yaklaşımı benimsenmiştir. Bu kapsamda, modeller ImageNet veri seti üzerinde önceden eğitilmiş ağırlıklar ile başlatılmış ve böylece düşük seviyeli görsel özelliklerin yeniden öğrenilmesi yerine mevcut bilgiden faydalanılması sağlanmıştır. Model mimarilerinin son katmanları ise problem özelinde yeniden düzenlenerek, yedi sınıflı sınıflandırma görevine uygun hale getirilmiştir. Bu doğrultuda, son tam bağlı (fully connected) katmanlar kaldırılarak yerine yedi çıkış nöronuna sahip yeni bir sınıflandırma katmanı eklenmiştir.

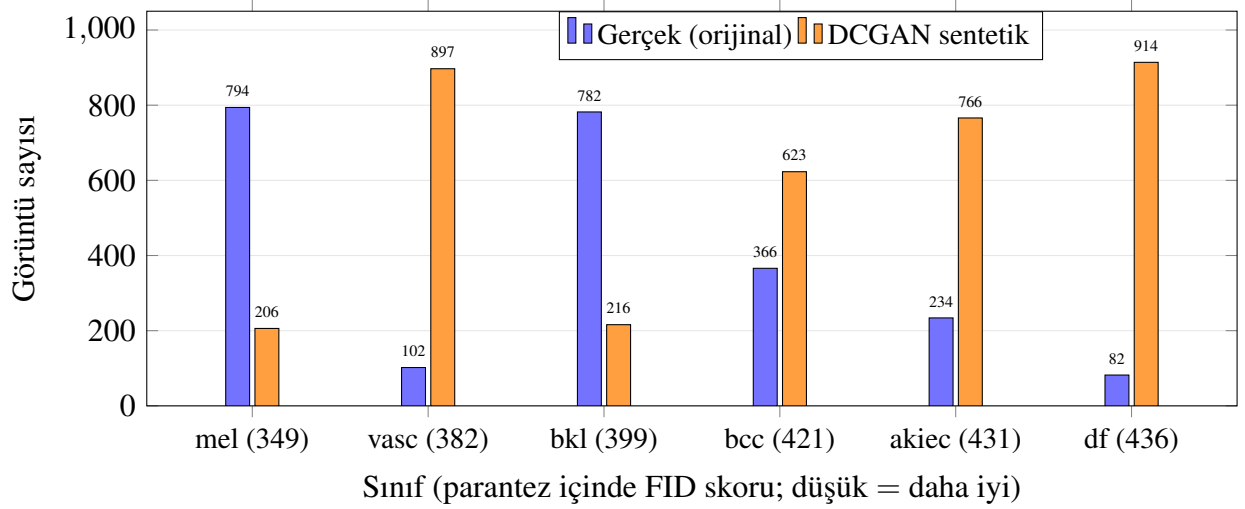
Tez kapsamında kullanılan veri seti, sınıf dağılımı korunacak şekilde stratified veri bölme yöntemi ile eğitim (%70), doğrulama (%15) ve test (%15) kümelerine ayrılmıştır. Veri bölme işlemi sırasında her bir sınıfın tüm alt kümelerde dengeli temsil edilmesi sağlanmış ve deneylerin tekrarlanabilirliği için tüm işlemlerde sabit bir rastgelelik değeri (random seed = 42) kullanılmıştır. Veri sızıntısını önlemek amacıyla her bir görüntünün yalnızca tek bir veri setinde yer almasına dikkat edilmiştir. Bununla birlikte, veri setinde tekrar eden örnekler ve hasta bazlı ayırım için ek kontroller uygulanmamış olup, bu durum medikal veri setleri açısından potansiyel bir sınırlılık olarak değerlendirilmektedir.

Çalışmada sentetik veri üretimi için DCGAN mimarisi kullanılmıştır. Modelde giriş gürültü vektörü boyutu 100 olarak belirlenmiş ve üretilen görüntüler 64×64 piksel çözünürlüğünde oluşturulmuştur. Eğitim süreci 80 epoch olarak planlanmış, ancak aşırı öğrenmeyi önlemek amacıyla erken durdurma mekanizması uygulanmıştır. Bu kapsamda patience değeri 10 ve minimum iyileşme eşiği 0,001 olarak belirlenmiştir. Model eğitimi sırasında batch size değeri 16 olarak seçilmiş, öğrenme oranı 0,0002 olarak ayarlanmış ve optimizasyon işlemi Adam optimizier kullanılarak gerçekleştirilmiştir. Adam optimizier için beta parametreleri (0,5, 0,999) olarak belirlenmiştir. Veri artırma aşamasında klasik tekniklerden yararlanılmış olup, görüntülere %50 olasılıkla yatay çevirme ve ± 10 derece rastgele döndürme işlemleri uygulanmıştır.

Üretilen sentetik görüntülerin kalitesi Fréchet Inception Distance (FID) metriği ile değerlendirilmiştir. FID hesaplamasında gerçek ve DCGAN tarafından üretilen görüntüler ayrı olarak değerlendirilmiş; klasik veri artırma ile elde edilen görüntüler, gerçek dağılımın deterministik dönüşümleri olduğundan hesaplamaya dahil edilmemiştir. Özellik çıkarımı için InceptionV3 modelinin pool_3 katmanından elde edilen 2048 boyutlu öznetelik vektörleri kullanılmıştır.

FID, sentetik görüntü dağılımı ile gerçek görüntü dağılımı arasındaki uzaklığı ölçer; **düşük değer daha yüksek kaliteyi** gösterir ve 0 mükemmel örtüşmeye karşılık gelir. Doğal görüntü kıyaslamalarında genellikle 50'nin altındaki değerler iyi kabul edilir. Ancak dermoskopik görüntüler, FID'nin dayandığı InceptionV3 ağıının eğitildiği doğal görüntü alanının dışında kaldığından ve üretim hesaplama verimliliği için 64×64 çözünürlükte yapıldığından, bu çalışmadaki mutlak FID değerleri yüksek çıkmıştır. Bu nedenle FID burada mutlak bir kalite ölçütü olarak değil, sınıflar arası *görelî* kalite karşılaştırması için kullanılmıştır.

Her sınıf için gerçek ve DCGAN sentetik görüntü sayıları, ilgili FID skorlarıyla birlikte Şekil 3.3'de sunulmuştur.



Şekil 3.3. Sınıflara göre gerçek ve DCGAN sentetik görüntü sayıları (parantez içinde ilgili FID skoru; FID’de düşük değer daha yüksek kaliteyi gösterir).

Şekil 3.3, sentetik veri kalitesi (FID) ile o sınıftaki gerçek örnek sayısı arasındaki güçlü ilişkiyi ortaya koymaktadır. En fazla gerçek örneğe sahip melanom (794 gerçek) sınıfında en düşük (en iyi) FID değeri (349) elde edilirken; yalnızca 82 gerçek örneğe sahip dermato fibroma sınıfında eğitim verisinin büyük bölümü sentetik görüntülerden oluşmakta ve FID en yüksek (436, en düşük kalite) seviyeye ulaşmaktadır. Bu eğilim, çok az gerçek örnekle eğitilen bir üreticinin daha düşük kaliteli sentetik görüntüler ürettiğini ve Bulgular bölümünde ortaya konan “düşük FID → yüksek duyarlılık kazanımı” ilişkisini doğrulamaktadır. Üretilen sentetik veriler sınıf dengesizliğini azaltmak amacıyla yalnızca eğitim setine dahil edilmiş; model performans değerlendirmesi ise tamamıyla gerçek verilerden oluşan test seti üzerinden gerçekleştirilmiştir.

4. BULGULAR VE TARTIŞMA

Bu bölüm, tahminlerin sonuçlarını incelemek ve modelin performansının bir analizini sunmak için tasarlanmıştır. Eğitim, PyTorch kütüphanesi kullanılarak NVIDIA RTX 4070 Ti GPU üzerinde gerçekleştirilmiştir. Optimizasyon algoritması olarak Adam ($lr = 1e-4$) tercih edilmiş ve kayıp fonksiyonu olarak Çapraz Entropi kullanılmıştır. Modeller 30 epoch boyunca eğitilmiş, aşırı öğrenmeyi önlemek için en iyi doğrulama başarısına sahip ağırlıklar kaydedilmiştir. Eğitim sürecinde kullanılan donanım özellikleri ve tüm hiperparametre ayarları Tablo 4.1’de detaylandırılmıştır.

Tablo 4.1. Eğitim hiper-parametreleri ve donanım özellikleri.

Parametre	Değer
Girdi Boyutu	224×224 px
Batch size	32
Optimizasyon	Adam ($lr = 1e-4$)
Epoch	30 (Erken Durdurma)
Loss	Cross-Entropy
Ağırlıklar	ImageNet (Pre-trained)
Donanım	NVIDIA RTX 4070 Ti
Kütüphane	PyTorch

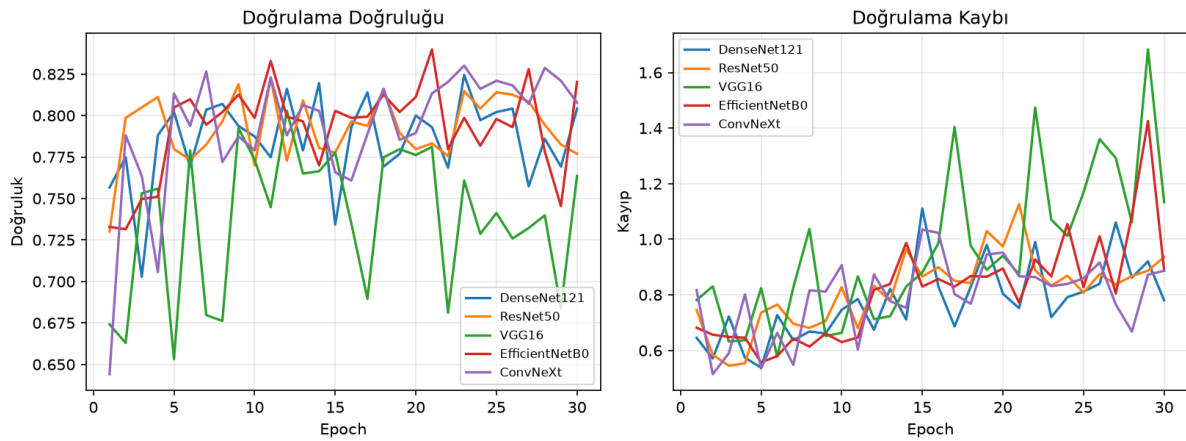
4.1. Model performansının değerlendirilmesi

Bu bölümde, önerilen hibrit veri dengeleme stratejisinin sınıflandırma başarımı üzerindeki etkisi, beş farklı derin öğrenme mimarisi (DenseNet121, ResNet50, VGG16, EfficientNetB0 ve ConvNeXt) kullanılarak karşılaştırmalı biçimde değerlendirilmektedir. Değerlendirmede, sınıf dengesizliği koşullarında tek başına yanıltıcı olabilen genel doğruluğun yanı sıra, azınlık sınıflarının başarımını daha güvenilir biçimde yansıtan Makro-Kesinlik, Makro-Duyarlılık ve Makro-F1 skorları ile sınıf bazlı duyarlılık (recall) değerleri ve karmaşıklık (confusion) matrisleri birlikte ele alınmaktadır. Bu çok yönlü yaklaşım, modellerin yalnızca ortalama performansını değil, klinik açıdan kritik nadir lezyonları doğru tespit edebilme yeteneğini de görünür kılmayı amaçlamaktadır. İzleyen alt bölümlerde sırasıyla modellerin eğitim ve doğrulama davranışları incelenmekte; ardından dengesiz veri seti üzerinde ortaya çıkan “Doğruluk Paradoksu” ortaya konmakta ve son olarak GAN tabanlı dengelenmiş veri seti ile elde edilen bulgular, özellikle azınlık sınıflarının tespit duyarlılığındaki değişim ve klinik açıdan kritik yanlış negatiflerdeki azalma bağlamında karşılaştırmalı olarak sunulmaktadır.

4.1.1. Eğitim ve doğrulama sonuçları

Bu bölümde sunulan tüm performans ölçütleri ve karmaşıklık matrisleri, modellerin eğitim ve doğrulama süreçlerinde kesinlikle karşılaşmadığı, veri setinin %15'lik kısmını oluşturan bağımsız test kümesi üzerinde elde edilmiştir. Bu yaklaşım sayesinde, modellerin değerlendirilmesi sırasında herhangi bir veri sızıntısı riski ortadan kaldırılmış ve elde edilen sonuçların güvenilirliği artırılmıştır.

Önerilen yaklaşımın performansı, farklı mimari tasarım prensiplerini temsil eden beş ayrı derin öğrenme modeli (DenseNet121, ResNet50, VGG16, EfficientNetB0 ve ConvNeXt Tiny) üzerinden karşılaştırmalı olarak değerlendirilmiştir. Deneysel çalışmalar iki aşamalı bir yapı içerisinde gerçekleştirilmiştir. İlk aşamada, herhangi bir dengeleme işlemi uygulanmamış olan orijinal dengesiz veri seti kullanılarak modellerin temel performansları belirlenmiştir. İkinci aşamada ise önerilen GAN tabanlı hibrit veri dengeleme yöntemi uygulanarak dengelenmiş veri seti oluşturulmuş ve aynı modeller bu yeni veri kümesi üzerinde yeniden eğitilmiştir. Bu iki aşamalı deneysel tasarım sayesinde, veri dengeleme stratejisinin model performansına olan etkisi izole edilerek analiz edilmiştir. İncelenen beş mimarinin dengelenmiş veri seti üzerindeki doğrulama doğruluğu ve doğrulama kaybı eğrileri Şekil 4.1'de karşılaştırmalı olarak sunulmuştur. Eğrilerden görüldüğü üzere, DenseNet121 ve ConvNeXt gibi modern mimariler daha hızlı ve kararlı bir yakınsama sergilerken, tüm modellerin eğitim süreci erken durdurma ile aşırı öğrenmeye karşı korunmuştur.



Şekil 4.1. İncelenen beş mimarinin dengelenmiş veri seti üzerindeki eğitim performansı: (a) doğrulama doğruluğu, (b) doğrulama kaybı.

4.1.2. Dengesiz veri seti ile elde edilen bulgular

Çalışmanın ilk aşamasında, herhangi bir veri artırma ya da dengeleme stratejisi uygulanmaksızın, orijinal HAM10000 veri seti kullanılarak modeller eğitilmiş ve elde edilen sonuçlar detaylı biçimde analiz edilmiştir. Bu aşamanın temel amacı, veri setinin doğal yapısında bulunan sınıf dengesizliğinin model performansı üzerindeki etkisini doğrudan gözlemlemektir. Özellikle

melanositik nevüs (nv) sınıfının veri setinde baskın bir çoğunluğa sahip olması, modellerin öğrenme davranışını belirgin şekilde etkilemiştir. Dengesiz veri seti ile gerçekleştirilen deneyler sonucunda modellerin genel doğruluk (accuracy) değerlerinin %81 ile %84 aralığında değiştiği görülmüştür. İlk bakışta bu oranlar yüksek bir performansa işaret ediyor gibi görünse de bu durum literatürde “Doğruluk Paradoksu” (accuracy paradox) olarak bilinen yanıltıcı bir probleme karşılık gelmektedir.

Bu paradoks, modelin yüksek doğruluk değerine sahip olmasına rağmen, gerçekte sınıflar arasında dengeli ve anlamlı bir öğrenme gerçekleştirememesini ifade etmektedir. İncelenen sonuçlar, modellerin veri setinde baskın olan nv sınıfına ait örüntüleri öğrenmeye aşırı derecede odaklandığını, buna karşılık melanom (mel), bazal hücreli karsinom (bcc) ve diğer azınlık sınıflara ait örnekleri ayırt etmede yetersiz kaldığını göstermektedir. Nitekim makro düzeyde hesaplanan F1 skorları, genel doğruluğun oldukça altında (yaklaşık 0,60–0,69 aralığında) kalmıştır. Bu tablo, yalnızca genel doğruluk metriğine odaklanmak yerine, sınıf bazlı performans ölçütlerinin (duyarlılık, kesinlik, F1-skoru vb.) de dikkate alınması gerektiğini açıkça ortaya koymaktadır.

Dengesiz veri seti üzerinde gerçekleştirilen deneylerde VGG16 modeli, diğer mimarilerle karşılaştırıldığında en düşük genel performansı sergileyen yaklaşım olmuştur. Elde edilen sonuçlara göre VGG16’nın Makro-F1 skoru 0,60 seviyesinde kalmış olup, bu durum modelin sınıflar arasında dengeli bir öğrenme gerçekleştiremediğini göstermektedir. Özellikle klinik açıdan kritik öneme sahip olan melanom (mel) sınıfında gözlenen performans düşüklüğü dikkat çekicidir. Bu sınıfa ait duyarlılık (recall) değerinin VGG16’da yaklaşık %28, ResNet50’de %35 ve DenseNet121’de %37 seviyesinde kalması, modellerin melanom vakalarının önemli bir bölümünü tespit edemediğini ortaya koymaktadır. Bu durum, tıbbi görüntüleme temelli karar destek sistemleri açısından ciddi bir risk oluşturmaktadır. Benzer şekilde, dermatofibroma (df) ve vasküler lezyonlar (vasc) gibi nadir sınıflarda da modellerin büyük çoğunluğu anlamlı bir öğrenme gerçekleştirememiştir. Daha güncel ve güçlü bir mimari olan ConvNeXt (Tiny) modeli dahi bu problemten bağımsız kalamamıştır. Bu bulgular, yalnızca mimari seçimin değil, veri dağılımının da model performansı üzerinde belirleyici bir rol oynadığını göstermekte ve özellikle nadir sınıfların doğru şekilde öğrenilebilmesi için veri dengeleme yöntemlerinin gerekliliğini ortaya koymaktadır.

4.1.3. GAN tabanlı dengelenmiş veri seti ile elde edilen bulgular

Önerilen GAN tabanlı hibrit dengeleme yaklaşımıyla eğitilen modeller incelendiğinde, azınlık sınıflarının veri içerisindeki temsil gücünün artırılmasının, özellikle bu sınıflara ait **duyarlılık (recall)** değerleri üzerinde belirgin ve olumlu bir etki yarattığı görülmektedir. Sınıf dengesizliğinin bulunduğu tıbbi görüntüleme problemlerinde en kritik metrik, klinik açıdan riskli sınıfların kaçırılma oranını (yanlış negatif) doğrudan yansıtan duyarlılıktır.

Bu bağlamda en dikkat çekici kazanım, klinik açıdan en kritik sınıf olan melanom (mel)

tespitinde elde edilmiştir. Dengeleme sonrası melanom duyarlılığı DenseNet121’de 0,37’den 0,57’ye, ResNet50’de 0,35’ten 0,65’e ve VGG16’da 0,28’den 0,47’ye yükselmiştir. Bu iyileşme, karmaşıklık matrislerinde en tehlikeli hata türü olan “melanomun iyi huylu nevüs (nv) olarak sınıflandırılması” (kaçırılan melanom) sayısındaki azalma ile doğrudan ilişkilidir. Örneğin ResNet50 modelinde iyi huylu olarak yanlış sınıflandırılan melanom sayısı 80’den 36’ya, DenseNet121’de ise 72’den 41’e düşmüştür. Söz konusu duyarlılık kazanımları Tablo 4.2’de özetlenmiştir.

Tablo 4.2. Dengeleme öncesi ve sonrası melanom (mel) tespit performansı.

Model	Duyarlılık (Dengesiz)	Duyarlılık (Dengeli)	Kaçan melanom (mel→nv)
DenseNet121	0,371	0,566	72 → 41
ResNet50	0,346	0,648	80 → 36
VGG16	0,277	0,465	77 → 67
EfficientNetB0	0,503	0,522	59 → 52
ConvNeXt	0,491	0,465	56 → 53

Benzer bir iyileşme vasküler lezyon (vasc) sınıfında da gözlemlenmiştir; bu sınıfa ait duyarlılık değeri incelenen beş modelin tamamında artmıştır (örneğin ConvNeXt’te 0,65’ten 0,95’e, EfficientNetB0’da 0,80’den 0,90’a).

Dengeleme stratejisinin yalnızca melanom sınıfıyla sınırlı olmayan genel etkisini ortaya koymak amacıyla, tüm sınıflar için beş model üzerinden hesaplanan ortalama duyarlılık değerleri Tablo 4.3’de sunulmuştur. Tabloda görüldüğü üzere, duyarlılık kazanımları sınıflar arasında türdeş değildir ve büyük ölçüde üretilen sentetik görüntülerin kalitesiyle (Şekil 3.3’deki FID skorlarıyla) örtüşmektedir. En düşük FID değerine, yani en yüksek sentetik kaliteye sahip melanom (FID 349) ve vasküler lezyon (FID 382) sınıflarında en yüksek duyarlılık artışları elde edilmiştir (+0,136 ve +0,140). Buna karşılık, en yüksek FID değerine sahip dermato fibroma (FID 436) ve aktinik keratoz (FID 431) sınıflarında kayda değer bir iyileşme sağlanamamıştır. Çoğunluk sınıfı olan melanositik nevüs (nv) sınıfında ise beklendiği üzere sınırlı bir düşüş gözlemlenmiştir. Bu bulgu, GAN tabanlı dengelemenin etkinliğinin, üretilen sentetik verinin gerçekçilik düzeyine doğrudan bağlı olduğunu göstermektedir.

Tablo 4.3. Sınıf bazında ortalama duyarlılık değerleri (beş modelin ortalaması).

Sınıf	Duyarlılık (Dengesiz)	Duyarlılık (Dengeli)	Değişim
nv (çoğunluk)	0,956	0,917	−0,039
mel	0,397	0,533	+0,136
bkl	0,654	0,651	−0,003
bcc	0,708	0,768	+0,060
akiec	0,526	0,509	−0,017
vasc	0,710	0,850	+0,140
df	0,400	0,388	−0,012

Bu sonuçlar, GAN tabanlı sentetik veri üretiminin ve klasik veri artırmanın, yeterli kali-

tede sentetik örnek üretilebilen az temsil edilen sınıfların model tarafından daha iyi tanınmasına katkı sağladığını göstermektedir.

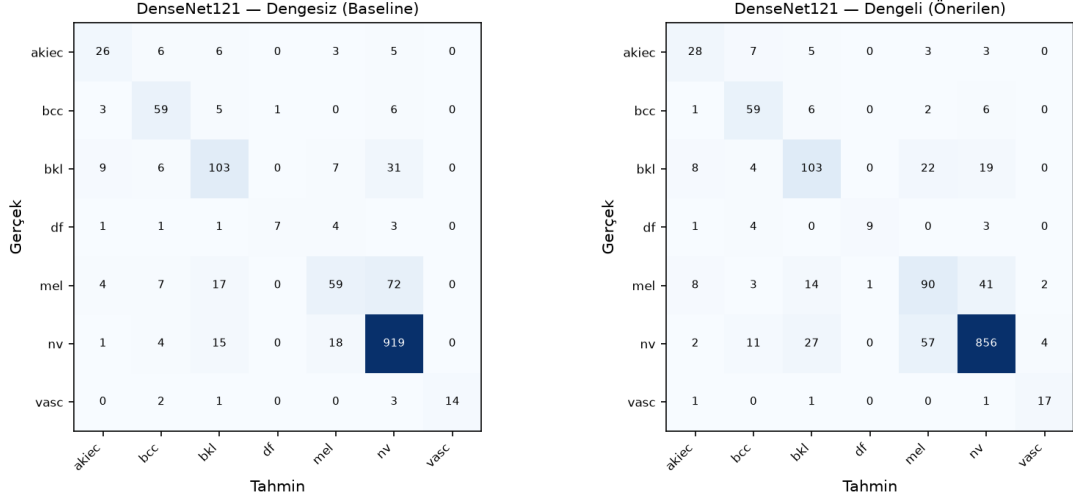
Öte yandan, bu duyarlılık kazanımı bir ödünleşim (trade-off) ile birlikte gelmektedir. Azınlık sınıflarına yönelik duyarlılığın artması, bu sınıflarda daha fazla yanlış pozitif üretilmesine ve dolayısıyla kesinlik (precision) ile genel doğruluk değerlerinde sınırlı bir düşüşe yol açmaktadır. Kesinlik ve duyarlılığın harmonik ortalaması olan ve her sınıfa eşit ağırlık veren Makro-F1 skoru bu iki eğilimi birlikte yansıttığından, dengeleme sonrası Makro-F1'deki değişim modelden modele farklılık göstermiştir: duyarlılık kazanımının kesinlik kaybını aştığı DenseNet121, ResNet50 ve EfficientNetB0 modellerinde Makro-F1 skoru artmış (ResNet50'de 0,64'ten 0,70'e belirgin bir artış), bu dengenin tersine döndüğü ConvNeXt ve VGG16 modellerinde ise sınırlı bir düşüş gözlemlenmiştir. Genel doğruluktaki bu sınırlı düşüş ise bir performans kaybı olarak değil, modelin yalnızca baskın sınıfı tahmin ederek “kolay başarı” elde etmek yerine, daha zor ve az temsil edilen sınıfları da ayırt etmeye yönelmesinin doğal bir sonucu olarak değerlendirilmektedir. Klinik tarama bağlamında, yanlış negatiflerin (kaçırılan malign vakaların) azaltılması genel doğruluktaki küçük bir düşüşten daha önemli kabul edildiğinden, elde edilen bu ödünleşim istenen yöndedir.

Önerilen yaklaşımın katkısını daha somut biçimde ortaya koymak amacıyla, en yüksek genel başarıyı gösteren DenseNet121 ile dengeleme stratejisinden en fazla fayda sağlayan ResNet50 modellerinin dengesiz ve dengeli veri setleri üzerindeki karmaşıklık matrisleri Şekil 4.2 ve Şekil 4.3'de sunulmuştur. Matrislerin köşegeni doğru sınıflandırmaları, köşegen dışı hücreler ise yanlış sınıflandırmaları göstermektedir. Her iki modelde de en kritik gelişme, gerçek melanom (mel) satırında gözlemlenmektedir: dengesiz durumda melanom örneklerinin büyük bir kısmı iyi huylu nevüs (nv) sütununa düşerken (DenseNet121'de 72, ResNet50'de 80 örnek), dengeleme sonrası bu yanlış negatifler belirgin biçimde azalmış (sırasıyla 41 ve 36'ya gerilemiş) ve köşegendeki doğru melanom tespitleri artmıştır (DenseNet121'de 59'dan 90'a, ResNet50'de 55'ten 103'e). Bu görsel analiz, önerilen hibrit dengeleme yönteminin temel katkısının, klinik açıdan en tehlikeli hata türü olan malign lezyonların gözden kaçırılmasını azaltmak olduğunu açıkça ortaya koymaktadır.

4.2. Yöntemlerin karşılaştırılması

Dengesiz ve dengeli veri setleri üzerinde Orijinal (Dengesiz) ve Önerilen (Dengeli) yaklaşımların performans karşılaştırması Tablo 4.4'de ayrıntılı olarak sunulmaktadır. Tablodaki tüm değerler, tamamıyla gerçek verilerden oluşan bağımsız test kümesi üzerinde elde edilmiştir.

Tablo 4.4 genel olarak değerlendirildiğinde, önerilen dengeleme stratejisinin en tutarlı katkısının **Makro-Duyarlılık** metriğinde ortaya çıktığı görülmektedir. Bu metrik, incelenen beş modelin tamamında dengeleme sonrası artmıştır; bu da modellerin azınlık sınıflara ait gerçek örnekleri daha yüksek oranda tespit edebildiğini göstermektedir. Buna karşılık, genel doğruluk ve kesinlik değerlerinde modele göre değişen sınırlı düşüşler gözlemlenmiş, bu durum Makro-F1



Şekil 4.2. DenseNet121 modelinin dengesiz ve dengeli veri setlerindeki karmaşıklık matrisleri.



Şekil 4.3. ResNet50 modelinin dengesiz ve dengeli veri setlerindeki karmaşıklık matrisleri.

skorundaki değişimin modele göre farklılaşmasına yol açmıştır.

Model bazlı karşılaştırmada, DenseNet121 ve ConvNeXt mimarileri her iki senaryoda da en yüksek genel başarıyı sergileyen modeller olmuştur. ResNet50, dengeleme stratejisinden en fazla fayda sağlayan model olarak öne çıkmış; hem Makro-Duyarlılık hem de Makro-F1 skorunda en belirgin iyileşmeyi göstermiştir. VGG16 ise daha eski bir mimariye sahip olması ve karmaşık veri dağılımlarını öğrenme kapasitesinin daha sınırlı olması nedeniyle, dengeli veri setine rağmen makro metriklerde en düşük performans sergilemiştir. Sonuç olarak, tablo bulguları veri dengesizliğinin model performansını yanıltıcı biçimde yüksek gösterebildiğini, buna karşılık dengeli veri kullanımıyla özellikle klinik açıdan kritik sınıflarda daha yüksek duyarlılık ve daha gerçekçi bir performans değerlendirmesi elde edildiğini ortaya koymaktadır.

Tablo 4.4. Dengesiz ve dengeli veri setlerinde model performansları.

Model	Veri Seti	Doğruluk	Makro Kesinlik	Makro Duyarlılık	Makro F1-Skor
DenseNet121	Dengesiz	0,8307	0,7698	0,6380	0,6820
	Dengeli	0,8132	0,7114	0,7009	0,6977
ConvNeXt	Dengesiz	0,8397	0,7707	0,6484	0,6886
	Dengeli	0,8160	0,6780	0,6839	0,6698
EfficientNetB0	Dengesiz	0,8300	0,7277	0,6419	0,6710
	Dengeli	0,8230	0,7390	0,6634	0,6861
ResNet50	Dengesiz	0,8139	0,6937	0,6120	0,6367
	Dengeli	0,8265	0,7498	0,6860	0,7014
VGG16	Dengesiz	0,8062	0,6689	0,5679	0,6027
	Dengeli	0,8090	0,7527	0,5633	0,5906

4.3. Modelin iyileştirilmesi için öneriler

Tez kapsamında elde edilen bulgular doğrultusunda, önerilen GAN tabanlı hibrit dengeleme yaklaşımının özellikle azınlık sınıfların tespit duyarlılığı üzerinde önemli katkılar sağladığı görülmüştür. Bununla birlikte, elde edilen sonuçlar model performansının daha da geliştirilebileceğini göstermekte olup, aşağıda gelecekteki çalışmalar için çeşitli iyileştirme önerileri sunulmuştur.

İlk olarak kullanılan GAN mimarisinin üretim kalitesinin artırılması önemli bir geliştirme alanı olarak öne çıkmaktadır. Çalışmada kullanılan DCGAN yapısı temel ve hesaplama açısından verimli olmakla birlikte, daha gelişmiş GAN türevleri (örneğin StyleGAN, CycleGAN veya Diffusion tabanlı modeller) kullanılarak daha yüksek çözünürlüklü ve morfolojik açıdan daha gerçekçi sentetik görüntüler üretilebilir. Bu sayede özellikle sınıflar arası ince farkların daha iyi öğrenilmesi ve kesinlik değerlerindeki düşüşün telafi edilmesi mümkün olabilir.

İkinci olarak, sentetik veri üretimi ile klasik veri artırma arasındaki katkı dengesinin yeniden gözden geçirilmesi önemli bir iyileştirme alanıdır. Bu çalışmada dengeleme büyük ölçüde DCGAN sentezine dayandırılmış; klasik veri artırma yalnızca kalan küçük boşluğu dolduran tamamlayıcı bir rol üstlenmiştir. Bunun bir sonucu olarak, en az sayıda gerçek örneğe sahip sınıfların (örneğin df ve akiec) eğitim verisinin yaklaşık %90'ı, yalnızca birkaç yüz gerçek görüntüden eğitilen bir GAN tarafından üretilen sentetik örneklerden oluşmuştur. Bulgular bölümünde de gösterildiği üzere, bu sınıflarda üretilen sentetik görüntülerin kalitesi (FID) en düşük seviyede kalmış ve duyarlılıkta kayda değer bir iyileşme sağlanamamıştır. Bu durum, çok sınırlı veriyle eğitilen bir üreticinin ürettiği sentetik verilere aşırı bağımlılığın bir sınırlılık oluşturduğunu göstermektedir. Gelecekteki çalışmalarda, her sınıfın mevcut gerçek örnek sayısına göre uyarlanan dengeli bir strateji; yani yeterli gerçek örneğe sahip sınıflarda GAN sentezine, çok az örnekli sınıflarda ise gerçek morfolojiyi koruyan klasik veri artırmaya daha fazla ağırlık verilmesi,

dengeleme sürecinin güvenilirliğini artırabilir. Bunun somut bir uygulaması olarak, sınıf başına üretilen sentetik görüntü sayısı sabit ve sınırlı bir değerle (örneğin yaklaşık 200 görüntü) tutulup 1.000'e kadar kalan boşluğun klasik veri artırma ile tamamlanması önerilebilir. Bu yaklaşım, en az gerçek örneğe sahip sınıflarda eğitim verisinin düşük kaliteli sentetik görüntülerle domine olmasını önleyerek dağılımın gerçek örneklerle daha yakın kalmasını sağlayabilir; ancak klasik veri artırmanın da yeni morfolojik çeşitlilik üretmediği göz önünde bulundurularak, bu tasarımın etkinliğinin deneysel olarak doğrulanması gerekmektedir.

Üçüncü olarak, sınıf dengesizliği ile mücadelede yalnızca veri seviyesinde değil, algoritma seviyesinde de iyileştirmeler yapılabilir. Örneğin, Focal Loss, Class-Balanced Loss veya ağırlıklı çapraz entropi gibi kayıp fonksiyonlarının kullanılması, modelin zor örneklerle daha fazla odaklanmasını sağlayarak azınlık sınıflardaki performansı daha da artırabilir. Bu tür yaklaşımlar, önerilen veri dengeleme yöntemi ile birlikte hibrit şekilde kullanılabilir.

Bu algoritma seviyesindeki önerinin etkinliği, tez kapsamında kontrollü ve veri sızıntısına kapalı bir ön deneyle sınanmıştır. Söz konusu deneyde, herhangi bir sentetik veri kullanılmaksızın yalnızca gerçek HAM10000 görüntüleri üzerinde lezyon bazlı katmanlı 5-katlı çapraz doğrulama (*StratifiedGroupKFold*) uygulanmış; her katta test kümesi tümüyle gerçek görüntülerden oluşturularak veri sızıntısı önlenmiştir. ResNet50 mimarisi üç ayrı koşulda eğitilmiştir: (i) standart çapraz entropi (temel durum), (ii) sınıf-dengeli ağırlıklı çapraz entropi (kayıp seviyesinde dengeleme) ve (iii) sınıf-dengeli yeniden örnekleme ile klasik veri artırma (veri seviyesinde dengeleme). Sonuçlar, her iki dengeleme yönteminin de temel duruma kıyasla makro-duyarlılığı artırdığını; en yüksek kazancın, makro-duyarlılığı 0,693 ($\pm 0,037$) değerinden 0,735'e ($\pm 0,025$) yükselten ağırlıklı kayıp yönteminde elde edildiğini göstermiştir. Bu yöntem özellikle en az örneğe sahip sınıfların duyarlılığını artırmış (dermato fibroma için 0,522'den 0,643'e, vasküler lezyon için 0,810'dan 0,894'e, akiec için 0,593'ten 0,664'e); buna karşılık çoğunluk sınıfı (nv) duyarlılığında ve genel doğrulukta beklenen sınırlı düşüş gözlenmiştir (sırasıyla 0,946'dan 0,908'e ve 0,842'den 0,828'e). Yeniden örnekleme yöntemi ise daha kararlı (güven aralıkları daha dar) ve genel doğruluğu neredeyse hiç düşürmeyen sonuçlar vermiş, en belirgin katkıyı orta büyüklükteki melanom sınıfında sağlamıştır (0,526'dan 0,626'ya); ancak en az gerçek örneğe sahip sınıflarda (df, vasc) kayda değer bir iyileşme sağlayamamıştır. Bunun nedeni, yeniden örneklemenin bu sınıflarda yalnızca aynı sınırlı gerçek görüntüleri tekrarlaması ve yeni morfolojik çeşitlilik üretememesidir. Elde edilen sınıf-bazlı kazançların bir kısmı, azınlık sınıflarındaki küçük test kümesi boyutundan kaynaklanan geniş güven aralıklarının içinde kaldığından, tek başına kesin bir istatistiksel üstünlük olarak yorumlanmamalıdır; buna karşılık makro düzeyde gözlenen tutarlı iyileşme, sınıf-dengeli kayıp fonksiyonlarının güçlü ve uygulaması kolay bir temel oluşturduğunu göstermektedir. Bu bulgu, önerilen GAN tabanlı hibrit yaklaşımın; sentetik veri üretmeden, sızıntısız ve düzgün çapraz doğrulama ile elde edilen bu basit dengeleme temeliyle karşılaştırmalı olarak değerlendirilmesi gerektiğine işaret etmektedir.

Dördüncü olarak model mimarilerinin daha ileri düzey versiyonlarının veya topluluk

yöntemlerinin kullanılması performansı artırabilir. Özellikle DenseNet ve ConvNeXt gibi başarılı modellerin birleşimi değerlendirilebilir. Bunun yanı sıra, dikkat mekanizmalarının modele entegre edilmesi, lezyon bölgelerinin daha etkin bir şekilde öğrenilmesini sağlayabilir. Veri tarafında ise, daha büyük ve çeşitli veri setlerinin kullanılması veya farklı veri setleri ile çapraz veri eğitimi yapılması modelin genelleme yeteneğini artıracaktır. Ayrıca klinik veriler (hasta yaşı, cinsiyet, lezyon lokasyonu gibi) ile görüntü verilerinin birlikte kullanıldığı çok modlu yaklaşımlar, daha gerçekçi ve klinik açıdan anlamlı sonuçlar üretebilir.

Son olarak, model performansının yalnızca metrikler üzerinden değil, klinik uygulanabilirlik açısından da değerlendirilmesi önemlidir. Yanlış negatif oranlarının minimize edilmesi, özellikle melanom gibi hayati risk taşıyan sınıflar için kritik bir gerekliliktir. Bu nedenle gelecekteki çalışmalarda model çıktılarının açıklanabilirliği artırılarak, klinisyenlerin karar süreçlerine güvenilir destek sağlanması hedeflenmelidir. Özellikle hiperparametre optimizasyonu ve öğrenme oranı gibi parametrelerin optimize edilmesi de modelin daha iyi sonuçlar vermesini sağlayabilir (Bergstra ve Bengio, 2012).

5. SONUÇLAR VE ÖNERİLER

Bu çalışma kapsamında, pigmente cilt lezyonlarının otomatik sınıflandırılmasında derin öğrenme modellerinin performansını olumsuz etkileyen temel problemlerden biri olan sınıf dengesizliği sorunu kapsamlı bir şekilde ele alınmış ve bu soruna çözüm olarak GAN tabanlı hibrit bir veri dengeleme stratejisi önerilmiştir. Önerilen yaklaşım, DCGAN tabanlı sentetik veri üretimi ile klasik veri artırma tekniklerini bir arada kullanarak hem veri çeşitliliğini artırmayı hem de modellerin az temsil edilen sınıfları daha etkin öğrenmesini amaçlamıştır. Gerçekleştirilen deneysel analizler, dengesiz veri setleri ile elde edilen yüksek doğruluk değerlerinin her zaman güvenilir bir performans göstergesi olmadığını açıkça ortaya koymuştur. Nitekim modellerin büyük çoğunluğu, veri setinde baskın olan sınıflara yönelerek yüksek doğruluk üretmiş, ancak klinik açıdan kritik öneme sahip azınlık sınıflarda yetersiz kalmıştır. Bu durum, literatürde “doğruluk paradoksu” olarak tanımlanan yanıltıcı başarı problemini doğrulamaktadır.

Önerilen hibrit dengeleme yöntemi uygulandığında ise, azınlık sınıflara ait duyarlılık (recall) değerlerinde belirgin iyileşmeler gözlemlenmiştir. Özellikle klinik açıdan en kritik sınıf olan melanom tespitinde, dengeleme sonrası duyarlılık değerleri birçok modelde önemli ölçüde artmış ve iyi huylu lezyon olarak yanlış sınıflandırılan (kaçırılan) melanom sayısı azalmıştır. Benzer bir iyileşme vasküler lezyon sınıfında da incelenen tüm modellerde elde edilmiştir. Bununla birlikte, elde edilen kazanımların tüm azınlık sınıflarında türdeş olmadığı; en yüksek duyarlılık artışlarının, en düşük FID değerine (en yüksek sentetik kaliteye) sahip melanom ve vasküler lezyon sınıflarında ortaya çıktığı görülmüştür. Buna karşılık, çok sınırlı sayıda gerçek örneğe sahip olması nedeniyle üretilen sentetik görüntülerin kalitesi görece düşük kalan dermatofibroma ve aktinik keratoz sınıflarında kayda değer bir iyileşme sağlanamamıştır. Bu bulgu, GAN tabanlı dengelemenin etkinliğinin, üretilen sentetik verinin gerçekçilik düzeyine doğrudan bağlı olduğunu ve sentetik veri üretiminin yalnızca veri sayısını artırmakla kalmayıp lezyonların morfolojik çeşitliliğini de yeterince zenginleştirmesi gerektiğini göstermektedir.

Bu duyarlılık kazanımı, kesinlik ve genel doğrulukta sınırlı bir düşüşü de beraberinde getirmiştir; dolayısıyla her sınıfa eşit ağırlık veren Makro-F1 skorundaki değişim modele göre farklılık göstermiştir. Ancak genel doğruluktaki bu sınırlı düşüş bir zayıflık olarak değil, modelin kolay (çoğunluk) sınıflara aşırı odaklanmak yerine zor ve az temsil edilen örnekleri ayırt etmeye yönelmesinin bir göstergesi olarak değerlendirilmelidir. Klinik tarama uygulamalarında yanlış negatiflerin (kaçırılan malign vakaların) azaltılması öncelikli bir hedef olduğundan, elde edilen bu ödünleşim istenen yöndedir ve önerilen yaklaşımın klinik açıdan anlamlı bir katkı sunduğunu ortaya koymaktadır.

Tez kapsamında elde edilen sonuçlar, model performansının daha da iyileştirilebileceğine işaret etmekte ve gelecekte yapılacak araştırmalar için çeşitli geliştirme alanları sunmaktadır. Öncelikle, kullanılan DCGAN mimarisinin StyleGAN veya difüzyon tabanlı modeller gibi daha

gelişmiş türevleri ile değiştirilmesi, sentetik veri kalitesini artırarak kesinlik değerlerindeki düşüşü telafi edebilir. Bunun yanı sıra, sınıf dengesizliğine duyarlı kayıp fonksiyonlarının (odak kaybı veya sınıf ağırlıklı kayıplar) kullanılması ve dikkat mekanizmalarının modele entegre edilmesi, azınlık sınıflardaki performansı daha da artırabilir. Nitekim tez kapsamında gerçekleştirilen sızıntısız çapraz doğrulamalı bir ön deney, sınıf-dengeli ağırlıklı kayıp fonksiyonunun herhangi bir sentetik veri kullanılmaksızın dahi makro-duyarlılığı tutarlı biçimde artırdığını göstererek bu yönün umut verici olduğunu ortaya koymuştur. Farklı model mimarilerinin birlikte kullanıldığı topluluk yaklaşımları ile klinik verilerin görüntü verilerine eklendiği çok modlu yaklaşımlar da tanısal doğruluğu artıracak önemli araştırma yönleridir. Ayrıca, model çıktılarının açıklanabilirliğinin artırılması, klinik uygulamalarda güvenilir karar desteği sağlanması açısından büyük önem taşımaktadır.

Son olarak, önerilen yöntemin farklı veri setleri ve farklı tıbbi görüntüleme alanları üzerinde test edilmesi, yaklaşımın genellenebilirliğini ortaya koyacaktır. Daha büyük ve çeşitli veri setleri ile yapılacak çalışmalar, modelin gerçek dünya senaryolarındaki performansını daha sağlıklı değerlendirme imkânı sunacaktır. Bu doğrultuda, önerilen yaklaşımın geliştirilmesi ve farklı alanlara uyarlanması, bilgisayar destekli teşhis sistemlerinin güvenilirliğini artırmada önemli katkılar sağlayacaktır.

KAYNAKÇA

- Abini, M. A., & Anjali, S. (2025). CapSkinNet: A Customized Capsule Network for Robust Multi-class Skin Lesion Classification on HAM10000 Dataset. *Congress on Intelligent Systems*, 170–188.
- Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., & Farhan, L. (2021). Review of Deep Learning: Concepts, CNN Architectures, Challenges, Applications, Future Directions. *Journal of Big Data*, 8(1), 53.
- American Cancer Society. (n.d.-a). *Key Statistics for Melanoma Skin Cancer* [Eriřim tarihi: 14 Temmuz 2026]. Retrieved July 14, 2026, from <https://www.cancer.org/cancer/types/melanoma-skin-cancer/about/key-statistics.html>
- American Cancer Society. (n.d.-b). *Survival Rates for Melanoma Skin Cancer* [SEER verileri, 2015–2021 tanılı vakalar; Eriřim tarihi: 14 Temmuz 2026]. Retrieved July 14, 2026, from <https://www.cancer.org/cancer/types/melanoma-skin-cancer/detection-diagnosis-staging/survival-rates-for-melanoma-skin-cancer-by-stage.html>
- Arı, A., & Berberler, M. E. (2017). Yapay Sinir Ağları ile Tahmin ve Sınıflandırma Problemlerinin Çözümü için Arayüz Tasarımı. *Acta Infologica*, 1(2), 55–73.
- Arora, G., Dubey, A. K., Jaffery, Z. A., & Rocha, A. (2022). Comparative Study of Fourteen Deep Learning Networks for Multi Skin Lesion Classification (MSLC) on Unbalanced Data. *Neural Computing and Applications*, 35(11), 7989–8015.
- Aruk, İ., & Toprak, A. N. (2025). Cilt Kanseri Tanısı için Farklı Evriřimsel Sinir Ağı Modellerinin Karşılaştırılması. *Journal of the Institute of Science and Technology*, 15(1), 25–38.
- Benyahia, S., Meftah, B., & Lézoray, O. (2022). Multi-features Extraction Based on Deep Learning for Skin Lesion Classification. *Tissue and Cell*, 74, 101701.
- Bergstra, J., & Bengio, Y. (2012). Random Search for Hyper-parameter Optimization. *Journal of Machine Learning Research*, 13(2).
- Blas Garcia, J., Flores Garcia, S., Daza Vergaray, A., Laiza Infantes, L., & Cier Paredes, S. (2026). Enhancing CNN-Based Multiclass Skin Disease Diagnosis from Clinical RGB Images Using CASE and C-BAFT Image Processing Techniques [Available at SSRN 6367882].
- Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R. L., Soerjomataram, I., & Jemal, A. (2024). Global Cancer Statistics 2022: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*, 74(3), 229–263. <https://doi.org/10.3322/caac.21834>
- Bribesh, R., & Tareq, W. Z. (2025). Evaluating Augmentation Strategies for CNN-Based Skin Lesion Classification on HAM10000. *International Conference on Advanced Engineering, Technology and Applications*, 1–16.
- Chatterjee, S., Gil, J.-M., & Byun, Y.-C. (2024). Early Detection of Multiclass Skin Lesions Using Transfer Learning-Based IncepX-Ensemble Model. *IEEE Access*, 12, 113677–113693.
- Chowdhury, S. H. (2025). HRNet-XGB: A Comprehensive Hybrid Framework for Automated and Interpretable Skin Cancer Detection with Multi-Scale Feature Fusion [SSRN].
- Da Silva, I. R. V., Jardim, A. G., de Souza Faés, G. R., Alva, T. A. P., Becker, C. D. L., & Botelho, V. R. (2026). Deep Learning-Based Skin Lesion Classification: A CNN Approach on High-Frequency Ultrasound Imaging. *Journal of Ultrasound in Medicine*, 45(5), 925–936.
- Deif, M. A., & Hammam, R. E. (2020). Skin Lesions Classification Based on Deep Learning Approach. *Journal of Clinical Engineering*, 45(3), 155–161.

- Ergün, E., & Kılıç, K. (2021). Derin Öğrenme ile Artırılmış Görüntü Seti Üzerinden Cilt Kanseri Tespiti. *Black Sea Journal of Engineering and Science*, 4(4), 192–200.
- Faris, M., Khalid, F., Shahid, M., Haque, Z. U., Husnain, S. I., & Ali, M. F. (2026). An Ensemble Deep Learning Framework for Automated Multi Class Skin Lesion Classification Using ConvNeXt-Tiny, EfficientNetV2-S, and MobileNetV3. *VFAST Transactions on Software Engineering*, 14(1), 60–72.
- Gayatri, E., & Aarth, S. L. (2024). Reduction of Overfitting on the Highly Imbalanced ISIC-2019 Skin Dataset Using Deep Learning Frameworks. *Journal of X-Ray Science and Technology*, 32(1), 53–68.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative Adversarial Nets. *Advances in Neural Information Processing Systems*, 27, 2672–2680.
- Gupta, M., Khanna, P., & Agrawal, A. K. (2026). Balancing the HAM10000 Dataset Using Deep Generative Models: Evaluating GAN, VAE and DDPM for Skin Lesion Classification. *2026 2nd International Conference on Cognitive Computing in Engineering, Communications, Sciences and Biomedical Health Informatics (IC3ECSBHI)*, 814–819.
- Haque, M., Khatun, K., Rohman, G., Rakib, A. A., & Billah, M. (2025). A Deep Learning Framework for Multi-class Skin Cancer Diagnosis: Addressing Class Imbalance and Performance Optimization Through Convolutional Neural Network Architectures. *European Journal of Computer Science and Information Technology*, 13(1), 63–71.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778.
- Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely Connected Convolutional Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4700–4708.
- Ibrahimi, M. A., Abbad, K., & Ben Abbou, R. (2026). A Novel Multimodal Deep Learning Architecture Integrating a Multi-scale Gated Residual Block and Soft Attention for Skin Lesion Classification. *Multimedia Tools and Applications*, 85(1), 40.
- Khan, A. R., Mujahid, M., Alamri, F. S., Saba, T., & Ayesha, N. (2025). Early-Stage Melanoma Cancer Diagnosis Framework for Imbalanced Data from Dermoscopic Images. *Microscopy Research and Technique*, 88(3), 797–809.
- Kızrak, M. A., & Bolat, B. (2018). Derin Öğrenme ile Kalabalık Analizi Üzerine Detaylı Bir Araştırma. *Bilişim Teknolojileri Dergisi*, 11(3), 263–286.
- Kumari, B., Chawla, P., & Arora, Y. (2026). Cross-Dataset Dermatology Framework: HAM10000 Training with ISIC Testing. *2026 International Conference on Computing, Sciences and Communications (ICCSC)*, 1–6.
- Lilhore, U. K., Sharma, Y. K., Simaiya, S., Alroobaea, R., Baqasah, A. M., Alsafyani, M., & Alhazmi, A. (2025). SkinEHDLF: A Hybrid Deep Learning Approach for Accurate Skin Cancer Classification in Complex Systems. *Scientific Reports*, 15(1), 14913.
- Lin, P. Y., Shen, Y., Mathew, N., Hu, R., Huang, S., Queen, C. M., & Zouridakis, G. (2026). Synthetic Melanoma Image Generation and Evaluation Using Generative Adversarial Networks. *Bioengineering*, 13(2), 245.
- Liu, Z., Mao, H., Wu, C.-Y., Feichtenhofer, C., Darrell, T., & Xie, S. (2022). A ConvNet for the 2020s. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 11976–11986.

- Lucas, A., Iliadis, M., Molina, R., & Katsaggelos, A. K. (2018). Using Deep Neural Networks for Inverse Problems in Imaging: Beyond Analytical Methods. *IEEE Signal Processing Magazine*, 35(1), 20–36.
- Mahbod, A., Schaefer, G., Wang, C., Ecker, R., & Ellinger, I. (2019). Skin Lesion Classification Using Hybrid Deep Neural Networks. *ICASSP 2019 – 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1229–1233.
- Naeem, M. A., Yang, S., Saleem, M. A., Javed, I., & Javeed, A. (2025). Automated Skin Cancer Detection Using MedFusionNet with Attention-Based Fusion of ConvNeXt and Vision Transformer. *Scientific Reports*.
- Nguyen, V. D., Bui, N. D., & Do, H. K. (2022). Skin Lesion Classification on Imbalanced Data Using Deep Learning with Soft Attention. *Sensors*, 22(19), 7530.
- Nirmala, V., & Premaladha, J. (2023). Synthetic Medical Image Augmentation: A GAN-Based Approach for Melanoma Skin Lesion Classification with Deep Learning. In *Deep Learning in Personalized Healthcare and Decision Support* (pp. 69–80). Academic Press.
- Perez, F., Vasconcelos, C., Avila, S., & Valle, E. (2018). Data Augmentation for Skin Lesion Analysis. *International Workshop on Computer-Assisted and Robotic Endoscopy*, 303–311.
- Radford, A., Metz, L., & Chintala, S. (2015). Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. *arXiv preprint arXiv:1511.06434*.
- Rahman, S., Pal, S., Mittal, S., Chawla, T., & Karmakar, C. (2024). SYN-GAN: A Robust Intrusion Detection System Using GAN-Based Synthetic Data for IoT Security. *Internet of Things*, 26, 101212.
- Russell, S. J., & Norvig, P. (1995). *Artificial Intelligence: A Modern Approach*. Prentice Hall.
- Saba Raoof, S., & Durai, M. S. (2022). A Comprehensive Review on Smart Health Care: Applications, Paradigms, and Challenges with Case Studies. *Contrast Media & Molecular Imaging*, 2022(1), 4822235.
- Sajol, M. S. I., Alvi, S. T., & Era, C. A. A. (2024). Performance Assessment of Advanced CNN and Transformer Architectures in Skin Cancer Detection. *2024 11th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, 1–8.
- Saranya, N., Rishikesh, R. R., & Gilbert Idayan, I. (2024). Analysis of GAN for Melanoma Skin Cancer Classification with Dermatologist Recommendation. *2024 International Conference on Recent Advances in Electrical, Electronics, Ubiquitous Communication, and Computational Intelligence (RAEEUCCI)*, 1–8.
- Selvaraj, K. M., Gnanagurusubbiah, S., Roy, R. R. R., & Balu, S. (2024). Enhancing Skin Lesion Classification with Advanced Deep Learning Ensemble Models: A Path Towards Accurate Medical Diagnostics. *Current Problems in Cancer*, 49, 101077.
- Setiawan, F., & Soewito, B. (2025). CRCDKD: A Novel Architecture for Medical Skin Cancer Classification on the Imbalanced HAM10000 Dataset. *Communications in Mathematical Biology and Neuroscience*.
- Shaik, A. R., Gangadhar, K., Sampath, K. S., Bharath, P. D. S. H., Tarun, M. S., & Arjun, M. (2026). Enhancing Skin Cancer Classification with GAN Based Data Augmentation. *2026 5th International Conference on Communication, Computing and Electronics Systems (ICCCES)*, 1753–1757.
- Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition [International Conference on Learning Representations (ICLR)]. *arXiv preprint arXiv:1409.1556*.

- Singh, P., Altaf Ahmed, M. D., Sharique, M. D., Gond, P., Verma, N., & Ranvijay. (2024). Skin Cancer Detection Using Ensembling of DenseNet, EfficientNet, and ConvNet. *International Conference on VLSI, Communication and Signal Processing*, 57–77.
- Su, Q., Hamed, H. N. A., Isa, M. A., Hao, X., & Dai, X. (2024). GAN-Based Data Augmentation Method for Imbalanced Multi-class Skin Lesion Classification. *IEEE Access*, 12, 16498–16513.
- Tan, M., & Le, Q. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *International Conference on Machine Learning*, 6105–6114.
- Tariq, O., Arshed, M. A., Kabir, M., Ijaz, K., Gherghina, Ş. C., & Batool, H. B. (2026). A Transfer-Learning Approach for Detection of Multiclass Synthetic Skin Cancer Images Generated by Deep Generative Models to Prevent Medical Insurance Fraud. *Mathematical and Computational Applications*, 31(1), 31.
- Tschandl, P., Rosendahl, C., & Kittler, H. (2018). The HAM10000 Dataset, a Large Collection of Multi-source Dermatoscopic Images of Common Pigmented Skin Lesions. *Scientific Data*, 5(1), 180161.
- Vijayakumari, G., & Mathew, O. C. (2025). Class-Specific Synthetic Data Augmentation Using EA-GAN for Enhanced Skin Cancer Detection. *2025 8th International Conference on Trends in Electronics and Informatics (ICOEI)*, 981–986.
- Wei, M., Wu, Q., Ji, H., Wang, J., Lyu, T., Liu, J., & Zhao, L. (2023). A Skin Disease Classification Model Based on DenseNet and ConvNeXt Fusion. *Electronics*, 12(2), 438.
- Yao, P., Shen, S., Xu, M., Liu, P., Zhang, F., Xing, J., Shao, P., Kaffenberger, B., & Xu, R. X. (2021). Single Model Deep Learning on Imbalanced Small Datasets for Skin Lesion Classification. *IEEE Transactions on Medical Imaging*, 41(5), 1242–1254.
- Yap, J., Yolland, W., & Tschandl, P. (2018). Multimodal Skin Lesion Classification Using Deep Learning. *Experimental Dermatology*, 27(11), 1261–1267.