

ULUSLARARASI 3B YAZICI TEKNOLOJİLERİ  
VE DİJİTAL ENDÜSTRİ DERGİSİ

INTERNATIONAL JOURNAL OF 3D PRINTING  
TECHNOLOGIES AND DIGITAL INDUSTRY

ISSN:2602-3350 (Online)

URL: <https://dergipark.org.tr/ij3dptdi>

# DENGESİZ CİLT LEZYONU VERİ SETLERİNDE GAN TABANLI VERİ ARTIRIMI İLE DERİN ÖĞRENME SINIFLANDIRMA BAŞARIMININ İYİLEŞTİRİLMESİ

IMPROVING DEEP LEARNING CLASSIFICATION PERFORMANCE WITH GAN-BASED DATA AUGMENTATION IN IMBALANCED SKIN LESION DATASETS

**Yazarlar (Authors):** Serkan ÖZDEMİR<sup>id</sup>, Salim CEYHAN<sup>id</sup>

**Bu makaleye şu şekilde atıfta bulunabilirsiniz (To cite to this article):** Özdemir S., Ceyhan S., "Dengesiz Cilt Lezyonu Veri Setlerinde GAN Tabanlı Veri Artırımı ile Derin Öğrenme Sınıflandırma Başarımının İyileştirilmesi" *Int. J. of 3D Printing Tech. Dig. Ind.*, 10(1): 84-93, (2026).

DOI: 10.46519/ij3dptdi.1804803

Araştırma Makale/ Research Article

Erişim Linki: (To link to this article): <https://dergipark.org.tr/en/pub/ij3dptdi/archive>

# DENGESİZ CİLT LEZYONU VERİ SETLERİNDE GAN TABANLI VERİ ARTIRIMI İLE DERİN ÖĞRENME SINIFLANDIRMA BAŞARIMININ İYİLEŞTİRİLMESİ

Serkan ÖZDEMİR<sup>a</sup> , Salim CEYHAN<sup>b</sup> 

<sup>a</sup>\* Bilecek Şeyh Edebalı Üniversitesi, Lisansüstü Eğitim Enstitüsü, Bilgisayar Mühendisliği ABD, TÜRKİYE

<sup>b</sup> Bilecek Şeyh Edebalı Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, TÜRKİYE

\* Sorumlu Yazar: [serkanhan75@hotmail.com](mailto:serkanhan75@hotmail.com)

(Geliş/Received: 16.10.25; Düzeltme/Revised: 17.12.25; Kabul/Accepted: 13.01.26)

## ÖZ

Dengesiz cilt lezyonu veri setleri, derin öğrenme modellerinin çoğunluk sınıfına karşı yanlılık geliştirmesine neden olarak, nadir fakat hayati risk taşıyan malign lezyonların tespit başarısını düşürmektedir. Bu çalışma, sınıf dengesizliği problemini ele almak amacıyla, Derin Evrişimli Üretici Karşıt Ağlar (DCGAN) ve klasik veri artırma tekniklerini birleştiren hibrit bir veri dengeleme stratejisinin sınıflandırma performansına etkisini sistematik olarak incelemektedir. Çalışma kapsamında DenseNet121, ResNet50, VGG16, EfficientNetB0 ve ConvNeXt mimarileri; orijinal dengesiz veri seti ve veri sızıntısını önlemek amacıyla yalnızca eğitim kümesine sentetik görüntülerin eklendiği dengelenmiş veri seti üzerinde karşılaştırılmıştır. Deneysel sonuçlar, dengesiz veri setinde modellerin yüksek doğruluk (%89'a varan) göstermesine rağmen "Doğruluk Paradoksu" yaşandığını ve azınlık sınıflarının (dermatofibroma, vasküler lezyonlar) tespit edilemediğini ortaya koymuştur. Önerilen hibrit dengeleme yöntemiyle eğitilen modellerde ise, genel doğrulukta marjinal bir değişim gözlemlenirken, sınıflar arası dengeyi ifade eden Macro-F1 skorlarında ortalama %21 oranında belirgin bir artış sağlanmıştır. Dengelenmiş veri setinde en kararlı başarıyı DenseNet121 (Macro-F1=0.803; Doğruluk=0.868) ve ConvNeXt modelleri sergilerken, nadir sınıfların tespit başarısı dramatik şekilde yükselmiştir. Bu bulgular, GAN destekli hibrit veri artırımının, sınırlı veri ile çalışan klinik karar destek sistemlerinin güvenilirliğini ve genelleme yeteneğini artırmada etkili bir strateji olduğunu kanıtlamaktadır.

**Anahtar Kelimeler:** Cilt Kanseri, Dermoskopi, Derin Öğrenme, GAN Tabanlı Veri Artırımı, Sınıf Dengesizliği, HAM10000, Hibrit Dengeleme.

## IMPROVING DEEP LEARNING CLASSIFICATION PERFORMANCE WITH GAN-BASED DATA AUGMENTATION IN IMBALANCED SKIN LESION DATASETS

### ABSTRACT

Imbalanced skin lesion datasets cause deep learning models to develop a bias towards the majority class, thereby compromising the detection performance of rare but vital malignant lesions. This study systematically investigates the impact of a hybrid data balancing strategy, combining Deep Convolutional Generative Adversarial Networks (DCGAN) and classical data augmentation techniques, on classification performance to address the class imbalance problem. Within the scope of the study, DenseNet121, ResNet50, VGG16, EfficientNetB0, and ConvNeXt architectures were compared on the original imbalanced dataset and a balanced dataset where synthetic images were added exclusively to the training set to prevent data leakage. Experimental results revealed that although the models exhibited high accuracy (up to 89%) on the imbalanced dataset, an "Accuracy Paradox" occurred, and minority classes (dermatofibroma, vascular lesions) could not be detected. In models trained with the proposed hybrid balancing method, while a marginal change was observed in overall accuracy, a significant

average increase of 21% was achieved in Macro-F1 scores, which represent inter-class balance. On the balanced dataset, DenseNet121 (Macro-F1=0.803; Accuracy=0.868) and ConvNeXt models displayed the most stable performance, while the detection success of rare classes increased dramatically. These findings demonstrate that GAN-supported hybrid data augmentation is an effective strategy for enhancing the reliability and generalization capability of clinical decision support systems operating with limited data.

**Keywords:** Skin Cancer, Dermoscopy, Deep Learning, GAN-Based Data Augmentation, Class Imbalance, HAM10000, Hybrid Balancing.

## 1. GİRİŞ

Cilt kanseri, dünya genelinde insidansı en hızlı artan kanser türlerinden biri olup, erken teşhis ve tedavi süreçleri hastalığın prognozunu doğrudan etkilemektedir [1]. Özellikle melanoma gibi malign karakterli cilt lezyonlarının erken evrede tespiti, hayatta kalma oranlarını artırmada hayati bir rol oynamaktadır. Cilt yüzeyinin mikroskopik detaylarla incelenmesini sağlayan dermoskopi, non-invaziv bir görüntüleme tekniği olarak klinik teşhis süreçlerinin standart bir parçası haline gelmiştir [2]. Dünya Sağlık Örgütü verilerine göre melanom, cilt kanseri vakalarının küçük bir yüzdesini oluşturmasına rağmen, deri kanserine bağlı ölümlerin büyük çoğunluğundan sorumludur. Bu klinik tablo, dermoskopik görüntüleri analiz ederek dermatologlara yardımcı olabilecek, yüksek doğruluk ve hassasiyete sahip otomatik karar destek sistemlerinin geliştirilmesini zorunlu kılmaktadır [3].

Son yıllarda, derin öğrenme tabanlı yöntemlerin, özellikle Evrişimsel Sinir Ağları (CNN) mimarilerinin görüntü işleme alanındaki başarısı, dermatolojik teşhis süreçlerinde devrim niteliğinde ilerlemeler sağlamıştır [4]. Ancak, bu algoritmaların klinik ortamda güvenle kullanılabilmesi, eğitim verilerinin kalitesine ve çeşitliliğine sıkı sıkıya bağlıdır. Mevcut dermoskopik veri setleri (örneğin HAM10000), doğası gereği ciddi bir **sınıf dengesizliği** sorunu ile karşı karşıyadır. *Melanositik nevüs* gibi iyi huylu lezyonların aşırı temsiliyeti, *melanoma* veya *dermatofibroma* gibi nadir görülen ancak kritik öneme sahip sınıfların model tarafından yeterince öğrenilememesine neden olmaktadır. Bu durum, modellerin çoğunluk sınıfına karşı yanlılık geliştirmesine, genel doğruluk oranı yüksek görünse de malign vakaları kaçırma riskinin (yanlış negatif) artmasına yol açmaktadır.

Sınıf dengesizliğini gidermek amacıyla literatürde sıklıkla kullanılan döndürme, kırpma ve yansıtma gibi geleneksel veri artırma teknikleri, mevcut verinin geometrik varyasyonlarını üretmekte ancak lezyonların morfolojik çeşitliliğini artırmada yetersiz kalmaktadır. Bu kısıt, sentetik veri üretimi konusunda güçlü bir çözüm sunan **Üretici Karşıt Ağlar (GAN)** teknolojisinin kullanımını gündeme getirmiştir. GAN'lar, gerçek veri dağılımını öğrenerek, orijinal veri setinde bulunmayan ancak klinik olarak anlamlı ve gerçekçi yeni dermoskopik görüntüler üretebilme kapasitesine sahiptir.

Bu çalışma, dengesiz cilt lezyonu veri setlerinde sınıflandırma başarısını artırmak amacıyla GAN tabanlı hibrit bir veri dengeleme stratejisi önermektedir. Çalışma kapsamında, derin öğrenme tabanlı DenseNet121 [5], ResNet50 [6], VGG16 [7], ConvNeXt [8] ve EfficientNetB0 [9] mimarileri kullanılarak iki aşamalı bir analiz gerçekleştirilmiştir. İlk aşamada modeller orijinal dengesiz veri seti üzerinde eğitilmiş ve "Doğruluk Paradoksu" olarak adlandırılan yanıltıcı başarı oranları ortaya konulmuştur. İkinci aşamada ise, eğitim verisindeki azınlık sınıfları önerilen GAN mimarisi ile zenginleştirilerek dengelenmiş, bu sayede nadir lezyonların tespitindeki performans artışı nicel olarak kanıtlanmıştır. Çalışmanın temel katkıları şunlardır:

1. Sınıf dengesizliğinin derin öğrenme modelleri üzerindeki negatif etkisinin sistematik olarak analiz edilmesi.
2. GAN tabanlı sentetik veri üretiminin, nadir görülen cilt lezyonlarının (dermatofibroma, vasküler lezyonlar) tespitinde sağladığı kritik performans artışının gösterilmesi.
3. Farklı derin öğrenme mimarilerinin sentetik verilere adaptasyon yeteneklerinin karşılaştırılması olarak değerlendirilmesi.

Makalenin geri kalanı şu şekilde yapılandırılmıştır: **Bölüm 2**, konuyla ilgili literatürdeki güncel çalışmaları ve mevcut boşlukları incelemektedir. **Bölüm 3**, kullanılan veri seti, önerilen GAN mimarisi ve deneysel kurulumun metodolojik detaylarını sunmaktadır. **Bölüm 4**, dengesiz ve dengeli veri setleri üzerinde elde edilen deneysel bulguları karşılaştırmalı analizlerle sunmaktadır. Son olarak **Bölüm 5**, elde edilen sonuçların klinik önemini tartışmakta ve çalışmayı özetlemektedir.

## 2. İLGİLİ ÇALIŞMALAR

Cilt lezyonlarının dermoskopik görüntüler üzerinden otomatik tespiti ve sınıflandırılması, son yıllarda derin öğrenme tabanlı yöntemlerin, özellikle CNN ve transfer öğrenme tekniklerinin başarısıyla literatürde geniş yer bulmuştur. Bu bölümde, konuyla ilgili temel yaklaşımlar; standart CNN uygulamaları, sınıf dengesizliğine yönelik algoritmik çözümler ve üretici model tabanlı veri artırma çalışmaları olmak üzere üç ana başlık altında incelenmiştir.

Literatürdeki ilk grup çalışmalar, ön eğitimli derin öğrenme mimarilerinin HAM10000 ve ISIC gibi standart veri setleri üzerindeki başarısına odaklanmıştır. Ergün ve Kılıç [10], veri artırma teknikleriyle destekledikleri çalışmalarında ResNet-34 mimarisini kullanarak %87,5 oranında doğruluk elde etmiştir. Benzer şekilde Aruk ve Toprak [11], yirmi bir farklı CNN modelini kıyasladıkları kapsamlı analizlerinde, EfficientNet-B0 modelinin %84 doğruluk ve %78,59 F1-skoru ile en dengeli performansı sergilediğini raporlamıştır. Arora vd. [12] ise, DenseNet201 mimarisinin %82,5 doğruluk oranıyla 14 farklı model arasında öne çıktığını belirtmiştir. Performansı maksimize etmek amacıyla topluluk öğrenme yöntemlerine başvuran çalışmalar da mevcuttur. Örneğin, Chatterjee vd. [13] InceptionV3 ve Xception modellerini birleştirerek %98 test doğruluğuna ulaşırken, Selvaraj vd. [14] ağırlıklı topluluk modeliyle %97 makro-ROC-AUC skoru elde etmiştir.

Ancak, bu yüksek doğruluk oranlarına ulaşan çalışmaların birçoğu, veri setindeki sınıf dengesizliğinin azınlık sınıfları üzerindeki negatif etkisini ve modelin çoğunluk sınıfına olan eğilimini detaylı olarak irdelememiştir.

Sınıf dengesizliği sorununu aşmak için önerilen ikinci grup yaklaşım, modelin öğrenme sürecini veya kayıp fonksiyonlarını modifiye etmeye dayanmaktadır. Gayatri ve Aarthy [15], sınıflandırması zor olan örneklerle daha fazla ceza puanı vererek modelin odaklanmasını sağlayan “Focal Loss” fonksiyonunun etkinliğini göstermiştir. Nguyen vd. [16] ise, derin öğrenme modellerine “Soft-Attention” mekanizmasını entegre ederek %90 doğruluk başarısına ulaşmıştır. Yao vd. [17], modifiye edilmiş veri artırma ve kümülatif öğrenme stratejilerini birleştirerek küçük veri setlerinde dahi yüksek başarımlar elde edilebileceğini savunmuştur. Bu yöntemler algoritmik düzeyde iyileştirme sağlasa da, veri setinin fiziksel temsil gücünü artırmada sınırlı kalmaktadır.

Üçüncü ve güncel yaklaşım ise, azınlık sınıflarını sentetik verilerle zenginleştirmeyi hedefleyen veri seviyesindeki çözümlerdir. Geleneksel veri artırma tekniklerinin (döndürme, yansıtma vb.) sadece geometrik varyasyonlar üretmesi ve morfolojik çeşitlilik sağlayamaması, araştırmacıları Üretici Karşıt Ağlar (GAN) kullanımına yöneltmiştir. Bu bağlamda Su vd. [18], Self-Transfer GAN (STGAN) adını verdikleri iki aşamalı bir mimari ile yüksek çözünürlüklü sentetik lezyonlar üretmiş; bu verilerin sınıflandırma metriklerinde (kesinlik, duyarlılık ve F1-skoru) mevcut yöntemlere kıyasla belirgin iyileşmeler sağladığını ortaya koymuştur.

Literatür incelendiğinde, GAN tabanlı veri artırma çalışmalarının umut verici olduğu, ancak ConvNeXt gibi modern mimarilerin sentetik verilere adaptasyonunun ve “Doğruluk Paradoksu”nun (yüksek doğruluk ama düşük nadir sınıf başarısı) sistematik analizinin yeterince ele alınmadığı görülmektedir. Bu çalışma, literatürdeki bu boşluğu doldurmak amacıyla, GAN tabanlı hibrit dengeleme stratejisinin etkinliğini hem klasik (VGG, ResNet) hem de modern (EfficientNet, ConvNeXt) mimariler üzerinde karşılaştırmalı olarak değerlendirmeyi hedeflemektedir.

### 3. MATERYAL VE METOT

Bu bölümde, çalışmada kullanılan veri seti, önerilen hibrit veri dengeleme stratejisi, geliştirilen GAN mimarisi ve sınıflandırma deneylerinin gerçekleştirildiği teknik altyapı detaylandırılmıştır. Çalışmanın metodolojik akış şeması, veri sızıntısını önleyecek ve tekrarlanabilirliği sağlayacak şekilde kurgulanmıştır.

#### 3.1. Veri Seti (HAM10000)

Çalışmada, pigmentli cilt lezyonlarının sınıflandırılması amacıyla literatürde yaygın kabul gören HAM10000 (“Human Against Machine with 10000 training images”) veri seti kullanılmıştır [19]. Veri seti, farklı popülasyonlardan ve görüntüleme tekniklerinden elde edilen toplam 10.015 adet dermoskopik görüntüden oluşmaktadır. Veri setinde yer alan yedi tanı sınıfı ve dağılımları şöyledir: Melanositik nevüs (nv), Melanom (mel), Benign keratoz benzeri lezyonlar (bkl), Bazal hücreli karsinom (bcc), Aktinik keratoz (akiec), Vasküler lezyonlar (vasc) ve Dermatofibroma (df). Tanıların %50’sinden fazlası histopatolojik inceleme ile doğrulanmış olup, kalan vakalar uzman konsensüsü veya in-vivo konfokal mikroskopi ile teyit edilmiştir. Doğrulama ve %15’i test verisi olarak belirlenmiştir. Eğitim verisi modelin öğrenmesi için doğrulama verisi hiper parametre optimizasyonu ve erken durdurma gibi süreçler için, test verisi ise, modelin genelleme performansını bağımsız olarak değerlendirmek için kullanılmıştır.

#### 3.2. Veri Ön İşleme ve Bölümleme

Modelin genelleme yeteneğini objektif olarak değerlendirebilmek ve eğitim verisinin test sürecine sızmasını engellemek amacıyla “Önce Ayır, Sonra İşle” protokolü izlenmiştir.

**1. Veri Bölümleme:** Veri seti, sınıf oranlarını koruyacak şekilde %70 Eğitim, %15 Doğrulama ve %15 Test kümesi olarak üç parçaya ayrılmıştır.

**2. Görüntü İşleme:** Tüm görüntüler, derin öğrenme modellerinin giriş boyutuna uygun olarak  $224 \times 224$  piksel çözünürlüğüne yeniden boyutlandırılmış ve ImageNet standartlarına uygun olarak kanal bazlı normalizasyon ( $\mu = [0.485, 0.456, 0.406]$ ,  $\sigma = [0.229, 0.224, 0.225]$ ) işlemine tabi tutulmuştur.

#### 3.3. Önerilen Hibrit Veri Dengeleme Stratejisi

Dengesiz veri setlerinde azınlık sınıflarının (özellikle df ve vasc) öğrenilmesini sağlamak amacıyla, yalnızca eğitim kümesi üzerinde üç aşamalı bir hibrit dengeleme stratejisi uygulanmıştır. Hedef, eğitim setindeki her bir sınıfı eşit sayıda ( $N = 1000$ ) örneğe ulaştırmaktır.

##### 3.3.1. Rastgele alt örnekleme

Veri setinde aşırı temsiliyet gösteren ve modelin yanlılık geliştirmesine neden olabilecek çoğunluk sınıfları (nv), rastgele seçim yöntemiyle 1.000 örneğe indirilmiştir. Bu işlem, modelin tek bir sınıfa odaklanmasını engelleyerek sınıflandırma dengesini artırmıştır.

##### 3.3.2. GAN tabanlı sentetik veri üretimi (DCGAN)

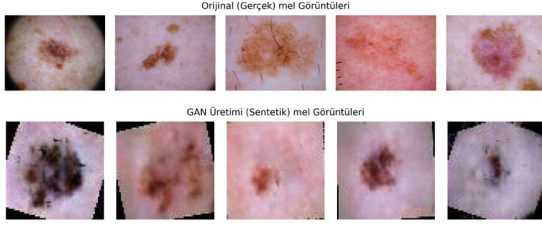
Örnek sayısı yetersiz olan azınlık sınıfları (mel, bcc, akiec) için Derin Evrişimli Üretici Karşıt Ağlar (DCGAN) mimarisi kullanılmıştır.

**Üretici (Generator):** 100 boyutlu latent gürültü vektöründen başlayarak, ters evrişim (ConvTranspose2d) katmanları aracılığıyla  $64 \times 64$  çözünürlüğünde sentetik dermoskopik görüntüler üretmiştir.

**Ayırt Edici (Discriminator):** Gerçek ve sentetik görüntüleri ayırmak üzere LeakyReLU aktivasyon fonksiyonuna sahip evrişim katmanlarından oluşturulmuştur.

**Eğitim Stabilitesi:** Mod çökmesiriskini azaltmak ve eğitimi stabilize etmek amacıyla Tek Yönlü Etiket Yumuşatma tekniği uygulanmıştır. Ayırt edici ağı aşırı güvenini kırmak ve gradyan sönmesini engellemek için gerçek etiket değerleri 1.0 yerine 0.9 olarak belirlenmiştir. Eğitim verilerine anlık veri artırma uygulanarak modelin ezberlemesi engellenmiştir.

Önerilen DCGAN mimarisi ile üretilen sentetik dermoskopik görüntülerin morfolojik yapısal özellikleri (renk dağılımı ve lezyon sınırları) başarılı bir şekilde öğrendiği gözlemlenmiştir. Eğitim setindeki azınlık sınıfları için üretilen sentetik örnekler ile orijinal gerçek görüntülerin görsel karşılaştırması Şekil 1’de sunulmuştur.



Şekil 1. Orijinal ve GAN üretimi melanoma 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 **64x64 piksel** çözünürlüğünde üretim yapması ve FID hesaplamasında kullanılan InceptionV3 ağının giriş boyutuna yapılan ölçeklendirmenin neden olduğu bulanıklıktan kaynaklanmaktadır.

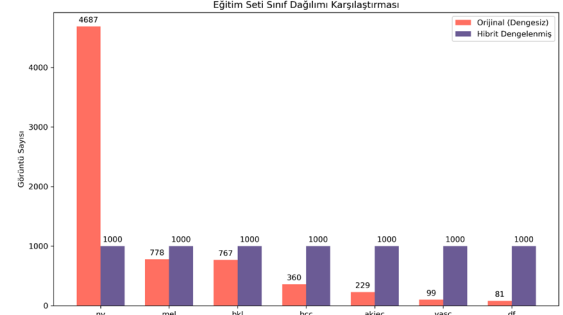
Ancak Bölüm 4'teki sınıflandırma başarılarındaki belirgin artış, skorlar yüksek olsa da üretilen verilerin modelin öğrenme kapasitesini artıracak yapısal niteliğe sahip olduğunu kanıtlamaktadır.

### 3.3.3. Klasik veri artırma ile boşluk doldurma

GAN eğitiminin yeterli yakınsamayı sağlayamadığı veya aşırı düşük örneklem sayısına sahip (df, vasc) sınıflarda, veri çeşitliliğini artırmak için klasik veri artırma teknikleri (döndürme, yatay/dikey çevirme, parlaklık değişimi) kullanılmıştır. Bu hibrit yaklaşım sayesinde, eğitim seti tamamen dengeli hale getirilirken, test ve doğrulama setlerinin orijinal (dengesiz ve gerçek) yapısı korunmuştur. Veri artırma işlemi öncesinde ve sonrasında eğitim setinin sınıf dağılımı Çizelge 1'de ve görsel karşılaştırması Şekil 2'de sunulmuştur.

Çizelge 1. Eğitim setinin dengeleme öncesi ve sonrası sınıf dağılımları

Sınıf	Orijinal (N)	Dengeli (N)	Yöntem
nv	4,687	1,000	Rastgele Silme
mel	778	1,000	GAN
bkl	767	1,000	GAN
bcc	360	1,000	GAN
akiec	229	1,000	GAN
vasc	99	1,000	GAN + Klasik
df	81	1,000	GAN + Klasik
<b>Toplam</b>	<b>7,001</b>	<b>7,000</b>	-



Şekil 2. Eğitim setinin dengeleme öncesi ve sonrası sınıf dağılımı

### 3.4. Sınıflandırma Modelleri

Dengelenmiş veri seti üzerinde, farklı mimari tasarım felsefelerini temsil eden beş derin öğrenme modeli eğitilmiştir:

- 1. DenseNet121:** Özellik tekrar kullanımını sağlayan yoğun bağlantılı mimari.
- 2. ResNet50:** Derin ağlarda gradyan kaybını önleyen artık blok yapısı.
- 3. VGG16:** Sade ve derin evrimsel katman yapısı.
- 4. EfficientNetB0:** Parametre verimliliği sağlayan ölçeklenebilir mimari.
- 5. ConvNeXt (Tiny):** Vision Transformer prensiplerini CNN yapısına uyarlayan modern mimari.

Tüm modellerde Transfer Öğrenme yöntemi uygulanmış; ImageNet ağırlıklarıyla başlatılan modellerin son tam bağlı katmanları 7 sınıf için yeniden yapılandırılmıştır.

### 3.5. Deneysel Kurulum

Eğitim işlemleri 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 Çizelge 2'de detaylandırılmıştır.

**Çizelge 2.** Eğitim hiperparametreleri 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

### 3.6. Performans Değerlendirme Kriterleri

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 matematiksel ifadeleri Eşitlik 1-4’te sunulmuştur:

**Doğruluk:** Modelin tüm sınıflar üzerindeki genel doğru tahmin oranını ifade eder.  $N$  toplam örnek sayısını göstermek üzere:

$$\frac{1}{N} \sum_{i=1}^C TP_i \quad (1)$$

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

$$\frac{1}{C} \sum_{i=1}^C \frac{TP_i}{TP_i+FP_i} \quad (2)$$

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

$$\frac{1}{C} \sum_{i=1}^C \frac{TP_i}{TP_i+FN_i} \quad (3)$$

**Makro-F1 Skor:** Kesinlik ve duyarlılık metriklerinin harmonik ortalaması olan F1-skorunun, tüm sınıflar için ortalamasını ifade eder. Sınıf dengesizliği durumunda modelin genel başarımını gösteren en güvenilir metrik olarak kabul edilir:

$$\frac{1}{C} \sum_{i=1}^C 2 \times \frac{Kesinlik_i \times Duyarlılık_i}{Kesinlik_i + Duyarlılık_i} \quad (4)$$

## 4. DENEYSEL BULGULAR

Bu bölümde sunulan tüm performans metrikleri ve karmaşıklık matrisleri, eğitim ve doğrulama süreçlerinde modellere gösterilmemiş olan ve veri setinin %15’ini oluşturan bağımsız test seti üzerinde elde edilmiştir. Böylece, önerilen yöntemin genelleme yeteneği herhangi bir veri sızıntısı riski olmaksızın değerlendirilmiştir.

Bu bölümde, pigmente cilt lezyonlarının sınıflandırılması amacıyla önerilen GAN tabanlı hibrit veri dengeleme yönteminin etkinliği, beş farklı derin öğrenme mimarisi (DenseNet121, ResNet50, VGG16, EfficientNetB0 ve ConvNeXt) üzerinden karşılaştırmalı olarak değerlendirilmiştir.

Deneyler, sınıf dengesizliğinin model performansı üzerindeki etkisini izole etmek amacıyla iki aşamada gerçekleştirilmiştir: (1) Orijinal dengesiz veri seti ile temel (baseline) performansın belirlenmesi ve (2) Önerilen yöntemle dengelenmiş veri seti ile performans artışının analizi.

### 4.1. Dengesiz Veri Seti ile Elde Edilen Bulgular (Baseline)

Çalışmanın ilk aşamasında, herhangi bir veri artırma veya dengeleme işlemi uygulanmadan orijinal HAM10000 veri seti üzerinde eğitilen modellerin sonuçları analiz edilmiştir. Bu aşama, veri setindeki doğal dengesizliğin (özellikle Melanositik Nevüs (nv) sınıfının baskınlığı) modellerde oluşturduğu yanlılığı ortaya koymaktadır.

Dengesiz veri seti üzerinde elde edilen sonuçlar incelendiğinde (Bkz. Çizelge 3), modellerin genel doğruluk değerlerinin %86 ile %89 arasında değiştiği görülmüştür. Ancak bu yüksek doğruluk oranları, “Doğruluk Paradoksu” olarak bilinen yanıltıcı bir durumdur. Modeller, veri setinin çoğunluğunu oluşturan nv sınıfını ezberleme eğilimi

göstermiş, buna karşın hayati öneme sahip azınlık sınıflarını öğrenememiştir.

**VGG16 Modeli:** Dengesiz veri setinde 0.53 Macro-F1 skoru ile en düşük performansı sergileyen mimari olmuştur. Özellikle Melanoma (mel) sınıfındaki Duyarlılık (Recall) değeri %57 seviyesinde kalmıştır; bu durum, kanserli vakaların neredeyse yarısının tespit edilemediği anlamına gelmektedir.

**Nadir Sınıf Başarısızlığı:** Tüm modeller, Dermatofibroma (df) ve Vasküler Lezyonlar (vasc) sınıflarında başarısız olmuştur. Örneğin, modern bir mimari olan ConvNeXt dahi dengesiz veri setinde df sınıfını öğrenememiş, bu sınıfta 0.00'a yakın bir F1 skoru üretmiştir.

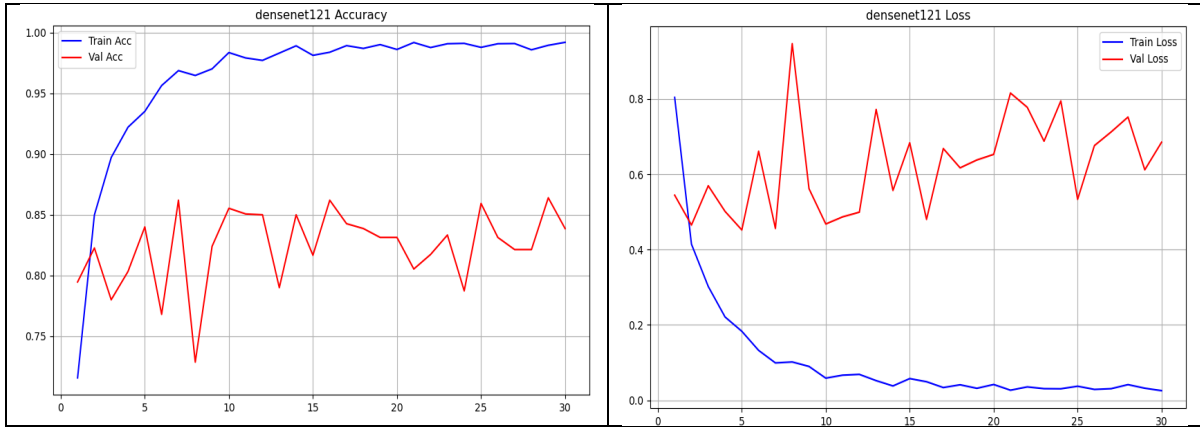
#### 4.2. GAN Tabanlı Dengelenmiş Veri Seti ile Elde Edilen Bulgular

Önerilen GAN tabanlı hibrit dengeleme stratejisi ile eğitilen modellerde, azınlık sınıflarının temsil gücünün artırılmasıyla birlikte performans metriklerinde istatistiksel olarak anlamlı iyileşmeler kaydedilmiştir.

Dengelenmiş veri seti sonuçlarına göre (Bkz. Çizelge 3), tüm mimarilerde Macro-F1 skorunda, dengesiz duruma kıyasla ortalama %21 oranında artış gözlemlenmiştir.

**DenseNet121 ve ConvNeXt Başarısı:** Önerilen yöntemle eğitilen DenseNet121 ve ConvNeXt modelleri, 0.80 seviyesinde Macro-F1 skoruna ulaşarak en başarılı modeller olmuşlardır. Özellikle ConvNeXt modelinin, dengesiz sette öğrenilemeyen Dermatofibroma (df) sınıfında 0.91 F1-skoruna (Kesinlik: 0.93, Duyarlılık: 0.88) ulaşması, sentetik verilerin morfolojik çeşitliliğe katkısını net bir şekilde ortaya koymaktadır.

**Kararlı Öğrenme:** Modellerin dengeli set üzerindeki genel doğruluk değerlerinin bir miktar düşmesi (%89'dan %86 bandına), beklenen ve olumlu bir durumdur. Bu, modelin artık sadece çoğunluk sınıfını tahmin ederek "kolay puan" toplamadığını, bunun yerine zor sınıfları ayırt etmeye çalıştığını göstermektedir.



Şekil 3. Densenet121 (dengeli) eğitim performansı: (a) Accuracy, (b) Loss

#### 4.3. Yöntemlerin Karşılaştırılması

Dengesiz ve dengeli veri setleri üzerinde Orijinal (Dengesiz) ve Önerilen (Dengeli) yöntemlerin performans karşılaştırması Çizelge 3'te detaylandırılmıştır. Tablo verileri, veri

dengeleme işleminin sadece azınlık sınıflarını iyileştirmekle kalmayıp, modelin genel sınıflandırma başarısını (Macro-F1) da optimize ettiğini göstermektedir.

**Çizelge 3.** Dengesiz ve dengeli veri setlerinde model performansları

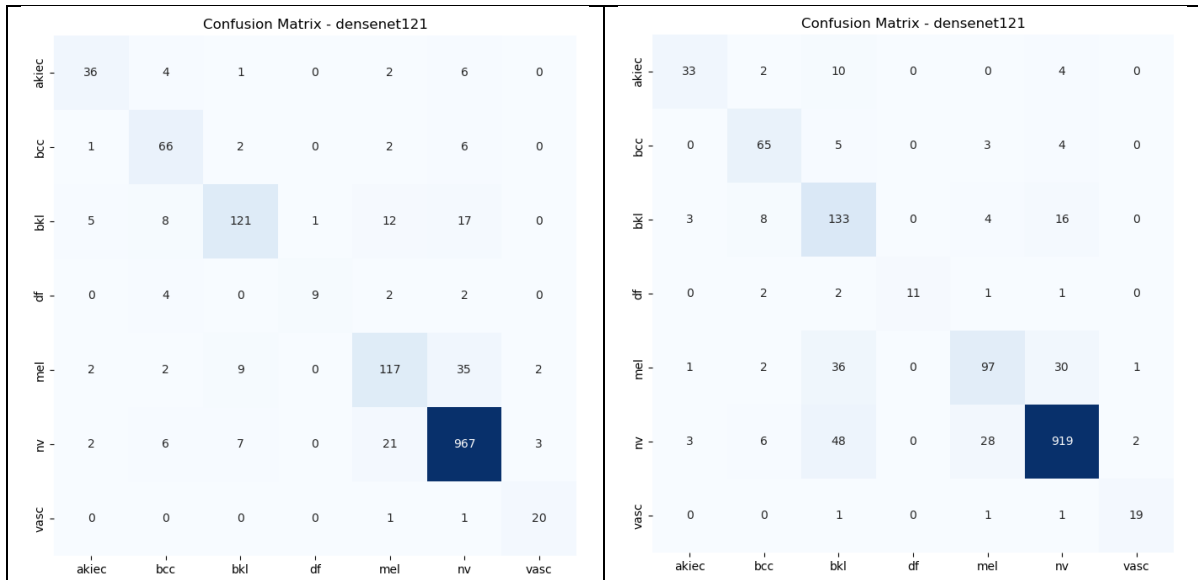
Model	Veri Seti	Doğruluk	Makro Kesinlik	Makro Duyarlılık	Makro F1-Skor
<b>DenseNet121</b>	Dengesiz	0.8875	0.8037	0.7818	0.5890
	<b>Dengeli</b>	<b>0.8682</b>	<b>0.8011</b>	<b>0.8131</b>	<b>0.8028</b>
<b>ConvNeXt</b>	Dengesiz	0.8928	0.8385	0.7788	0.5940
	<b>Dengeli</b>	<b>0.8675</b>	<b>0.8120</b>	<b>0.7969</b>	<b>0.8016</b>
<b>EfficientNetB0</b>	Dengesiz	0.8915	0.8199	0.7836	0.6051
	<b>Dengeli</b>	<b>0.8582</b>	<b>0.7960</b>	<b>0.7801</b>	<b>0.7847</b>
<b>ResNet50</b>	Dengesiz	0.8915	0.7936	0.7877	0.5719
	<b>Dengeli</b>	<b>0.8529</b>	<b>0.7831</b>	<b>0.7824</b>	<b>0.7766</b>
<b>VGG16</b>	Dengesiz	0.8642	0.7424	0.7270	0.5302
	<b>Dengeli</b>	<b>0.7983</b>	<b>0.7208</b>	<b>0.6720</b>	<b>0.6905</b>

#### 4.4. Karmaşıklık Matrisi Analizi

Modellerin sınıf bazlı hatalarını görselleştirmek amacıyla, en iyi F1 skoruna sahip DenseNet121 modelinin dengesiz ve dengeli durumlardaki karmaşıklık matrisleri karşılaştırılmıştır (Şekil 4).

Dengesiz veri setinde (Şekil 4-a), modelin özellikle Melanoma vakalarını sıklıkla iyi huylu Nevüs (nv) ile karıştırdığı (**35 örnek**) görülmüştür. Önerilen yöntemle eğitilen modelde ise (Şekil 4-b), modelin **Benign Keratoz (bkl)** sınıfındaki doğru tahmin sayısı **121'den 133'e** yükselerek belirgin bir iyileşme

sağlanmıştır. Ayrıca **Dermatofibroma (df)** sınıfındaki doğru tespit sayısı **9'dan 11'e** çıkmıştır. Bununla birlikte, dengeli eğitim setinde modelin Melanoma vakalarını Nevüs ile karıştırma oranı azalırken (**30 örnek**), morfolojik benzerlik nedeniyle Benign Keratoz (bkl) sınıfı ile karıştırma eğiliminin arttığı gözlemlenmiştir. Vasküler (vasc) lezyonlarda ise model, her iki senaryoda da yüksek başarı (Dengesiz: 20/22, Dengeli: 19/22) sergileyerek kararlılığını korumuştur. Bu sonuçlar, GAN tabanlı veri artırımının özellikle azınlık sınıflarındaki (df, bkl) öğrenme kapasitesini artırdığını doğrulamaktadır.



**Şekil 4.** Densenet121 karmaşıklık matrisleri: (a) Dengesiz, (b) Dengeli

#### 5. SONUÇ

Pigmente cilt lezyonlarının otomatik teşhisinde derin öğrenme modellerinin başarısını sınırlayan en temel faktörlerden biri olan sınıf dengesizliği problemi, bu çalışmada önerilen Üretici Karşıt Ağlar (GAN) tabanlı hibrit veri

dengeleme stratejisi ile ele alınmıştır. Çalışma kapsamında gerçekleştirilen kapsamlı deneyler, literatürde sıklıkla göz ardı edilen “Doğruluk Paradoksu” olgusunu somut verilerle ortaya koymuştur. Orijinal dengesiz veri seti üzerinde eğitilen DenseNet121, ResNet50, VGG16,

EfficientNetB0 ve ConvNeXt mimarilerinin, %89'a varan yüksek doğruluk oranlarına ulaşmalarına rağmen, bu başarının yanıltıcı olduğu ve modellerin eğitim setinin çoğunluğunu oluşturan melanositik nevüs sınıfını ezberleme eğilimi gösterdiği tespit edilmiştir. Özellikle VGG16 gibi geleneksel mimarilerin, dengesiz veri dağılımı altında hayati risk taşıyan melanoma vakalarının tespitinde %60'ın altında kalan duyarlılık oranları sergilemesi, veri dengeleme işlemi uygulanmayan modellerin klinik karar destek sistemlerinde kullanılmasının ciddi güvenlik riskleri barındırdığını kanıtlamaktadır.

Önerilen DCGAN ve klasik veri artırma tekniklerini içeren hibrit dengeleme yönteminin uygulanmasıyla, modellerin azınlık sınıflarını öğrenme kapasitesinde istatistiksel olarak anlamlı ve belirgin iyileşmeler kaydedilmiştir. Dengelenmiş veri seti ile eğitilen tüm mimarilerde, modelin genel başarısını ve sınıflar arası dengeyi ifade eden Macro-F1 skorlarında ortalama %21 oranında bir artış sağlanmıştır. Bu iyileşme, sentetik verilerin sadece sayısal bir dengeleme aracı olmadığını, aynı zamanda lezyonların morfolojik çeşitliliğini artırarak modelin karar sınırlarını daha hassas bir şekilde belirlemesine katkıda bulunduğunu göstermektedir. Özellikle modern mimarilerden ConvNeXt ve DenseNet121'in, dengesiz veri setinde neredeyse hiç tespit edilemeyen dermatofibroma ve vasküler lezyonlar gibi nadir sınıflarda %90'a varan başarı oranlarına ulaşması, önerilen yöntemin nadir görülen cilt hastalıklarının teşhisindeki etkinliğini doğrulamaktadır.

Sonuç olarak bu çalışmada sunulan GAN tabanlı yaklaşım, derin öğrenme modellerinin çoğunluk sınıfına karşı geliştirdiği yanlılığı başarılı bir şekilde minimize etmiş ve nadir görülen malign lezyonların tespit edilebilirliğini artırmıştır. Dengeli eğitim setlerinde gözlemlenen doğruluk değerindeki marjinal düşüş, modelin kolay tahmin stratejisini terk edip zorlu vakaları ayırt etmeye yönelik daha karmaşık öznitelikleri öğrenmesiyle açıklanmaktadır. Elde edilen bulgular, sınırlı ve dengesiz tıbbi görüntü veri setlerinde sentetik veri üretiminin, tanısal doğruluğu ve sistem güvenilirliğini artırmak için güçlü bir çözüm olduğunu ortaya koymaktadır. Gelecek çalışmalarda, önerilen yöntemin farklı görüntüleme modalitelerine uyarlanması ve

daha yüksek çözünürlüklü üretim mimarileri ile geliştirilmesi, bilgisayar destekli dermatolojik tanı alanına önemli katkılar sağlayacaktır.

## KAYNAKLAR

1. Kanca, E., Ayas, S., "Dengesiz veri kümelerinde topluluk yöntemlerine dayalı melanom sınıflandırılması", Karadeniz Fen Bilimleri Dergisi, Cilt 12, Sayı 2, Sayfa 1122-1137, 2022.
2. Özcan, Ö., Karaaltun, M., "Görüntü artırma tekniklerinin cilt kanseri türleri üzerinde evrimsel sınır ağları ile sınıflandırma başarılarının karşılaştırılması", Niğde Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi, Cilt 12, Sayı 4, Sayfa 1-1, 2025.
3. Sarı, A., Nizam, A., Aydın, M., "Implementing of transfer learning method in the diagnosis of skin diseases with convolutional neural networks", 7th International Conference on Computer Science and Engineering (UBMK), Sayfa 142-147, Diyarbakır, 2022.
4. Kaplan, A., Güldogan, E., Arslan, A.K., "Prediction of melanoma from dermoscopic images using deep learning-based artificial intelligence techniques", International Artificial Intelligence and Data Processing Symposium (IDAP), Sayfa 1-5, Malatya, 2019.
5. Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q., "Densely connected convolutional networks", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Sayfa 4700-4708, Honolulu, 2017.
6. He, K., Zhang, X., Ren, S., Sun, J., "Deep residual learning for image recognition", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Sayfa 770-778, Las Vegas, 2016.
7. Simonyan, K., Zisserman, A., "Very deep convolutional networks for large-scale image recognition", ArXiv Preprint, arXiv:1409.1556, <https://arxiv.org/abs/1409.1556>, May 22, 2025, 2014.
8. Liu, Z., Mao, H., Wu, C.Y., Feichtenhofer, C., Darrell, T., Xie, S.A., "Convnet for the 2020s", Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Sayfa 11976-11986, New Orleans, 2022.
9. Hoang, V.T., Jo, K.H., "Practical analysis on architecture of EfficientNet", 14th International Conference on Human System Interaction (HSI), Sayfa 1-4, Gdańsk, 2021.

10. Ergün, E., Kılıç, K., “Derin öğrenme ile artırılmış görüntü seti üzerinden cilt kanseri tespiti”, *Black Sea Journal of Engineering and Science*, Cilt. 4, Sayı 4, Sayfa 192-200, 2021.
11. Aruk, İ., Toprak, A.N., “Cilt kanseri tanısı için farklı evrimsel sinir ağı modellerinin karşılaştırılması”, *Journal of the Institute of Science and Technology*, Cilt 15, Sayı 1, Sayfa 25-38, 2025.
12. Arora, G., Dubey, A.K., Jaffery, Z.A., Rocha, A.A., “Comparative study of fourteen deep learning networks for multi skin lesion classification (MSLC) on unbalanced data”, *Neural Computing and Applications*, Cilt 35, Sayı 11, Sayfa 7989-8015, 2022.
13. Chatterjee, S., Gil, J.M., Byun, Y.C., “Early detection of multiclass skin lesions using transfer learning-based IncepX-ensemble model”, *IEEE Access*, Cilt 12, Sayfa 113677–113693, 2024.
14. Selvaraj, K.M., Gnanagurusubbiah, S., Roy, R.R.R., Balu, S., “Enhancing skin lesion classification with advanced deep learning ensemble models: A path towards accurate medical diagnostics”, *Current Problems in Cancer*, Cilt 49, Sayfa 101077, 2024.
15. Gayatri, E., Aarth, S.L., “Reduction of overfitting on the highly imbalanced ISIC-2019 skin dataset using deep learning frameworks”, *Journal of X-Ray Science and Technology*, Cilt 32, Sayı 1, Sayfa 53-68, 2024.
16. Nguyen, V.D., Bui, N.D., Do, H.K., “Skin lesion classification on imbalanced data using deep learning with soft attention”, *Sensors*, Cilt 22, Sayı 19, Sayfa 7530, 2022.
17. Yao, P., Shen, S., Xu, M., Liu, P., Zhang, F., Xing, J., Shao, P., Kaffenberger, B., Xu, R.X., “Single model deep learning on imbalanced small datasets for skin lesion classification”, *IEEE Transactions on Medical Imaging*, Cilt 41, Sayı 5, Sayfa 1242-1254, 2021.
18. Su, Q., Hamed, H.N.A., Isa, M.A., Hao, X., Dai, X.A., “GAN-based data augmentation method for imbalanced multi-class skin lesion classification”, *IEEE Access*, Cilt 12, Sayfa 16498-16513, 2024.
19. P. Tschandl, C. Rosendahl, and H. Kittler, “The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions”, *Scientific Data*, Cilt 5, Sayfa 180161, 2018.