BİNGÖL ÜNİVERSİTESİ
TEKNİK BİLİMLER DERGİSİ

Bingöl Üniversitesi Teknik Bilimler Dergisi
Bingol University Journal of Technical Science
e-ISSN: 2757-6884

2026 Sayı: 2, Sayfa: 1-22
Araştırma

2026 Issue: 2, Page: 1-22
Research

Doi: 10.5281/zenodo.22943937

Mineral Görüntülerinin Sınıflandırılmasında Derin Öğrenme Mimarilerinin Karşılaştırmalı Analizi ve Performans Optimizasyonu

Yusuf ÇELİK[1*], Anıl UTKU[2]

[1*] Munzur Üniversitesi, Bilgisayar Mühendisliği Bölümü, Tunceli/Turkiye.
ORCID No: 0000-0002-7859-7543, e-mail: celikyusuf@munzur.edu.tr
[2]Munzur Üniversitesi, Bilgisayar Mühendisliği Bölümü, Tunceli/Turkiye.
ORCID No: 0000-0002-7240-8713, e-mail: anilutku@munzur.edu.tr

(Alınış/Arrival: 17.04.2026, Kabul/Acceptance: 09.09.2026, Yayınlanma/Published: 25.09.2026)

Özet

Minerallerin doğru ve hızlı bir şekilde sınıflandırılması; jeoloji, madencilik ve endüstriyel uygulamalar açısından büyük önem taşımaktadır. Geleneksel yöntemler çoğunlukla uzman bilgi ve deneyimine dayalı olup zaman alıcı ve öznel sonuçlar üretebilmektedir. Bu çalışmada, mineral görüntülerinin otomatik sınıflandırılması amacıyla derin öğrenme tabanlı bir yaklaşım önerilmiş ve farklı model mimarilerinin performansları karşılaştırmalı olarak analiz edilmiştir. Çalışmada, 15 farklı mineral sınıfına ait toplam 39.354 görüntü içeren Mineral Photos veri seti kullanılmıştır. Deneysel süreçte, ResNet50, EfficientNet-B0, MobileNetV3-Large, DenseNet121 ve Vision Transformer (ViT-B/16) olmak üzere beş farklı önceden eğitilmiş model değerlendirilmiştir. Tüm modeller aynı eğitim protokolü altında 5-katlı çapraz doğrulama yöntemi ile test edilmiştir. İlk aşama sonuçlarına göre EfficientNet-B0 modeli %74.39 doğruluk ile en yüksek performansı elde ederken, ResNet50 (%71.37) ve DenseNet121 (%70.13) modelleri rekabetçi sonuçlar sunmuştur. Buna karşılık ViT-B/16 modeli %48.61 doğruluk ile diğer modellere kıyasla daha düşük performans göstermiştir. İkinci aşamada, EfficientNet tabanlı modeller üzerinde daha gelişmiş deneyler gerçekleştirilmiş ve EfficientNet-V2-S modelinin %83.14 ortalama doğruluk değeri ile daha yüksek ve daha kararlı bir performans sergilediği belirlenmiştir. Ayrıca model, kesinlik, duyarlılık ve F1-skoru metriklerinde de dengeli sonuçlar üretmiştir. Elde edilen bulgular, bu çalışmada kullanılan Mineral Photos veri seti ve uygulanan deneysel koşullar altında, değerlendirilen CNN tabanlı modellerin ViT-B/16 modeline kıyasla daha yüksek sınıflandırma performansı sergilediğini göstermektedir.

Anahtar Kelimeler: Mineral sınıflandırma, derin öğrenme, görüntü işleme, transfer öğrenme, konvolüsyonel sinir ağları.

Comparative Analysis of Deep Learning Architectures and Performance Optimization in Mineral Image Classification

Abstract

Accurate and rapid classification of minerals is of great importance in geology, mining, and industrial applications. Traditional methods are often based on expert knowledge and experience, and can produce time-consuming and subjective results. In this study, a deep learning-based approach is proposed for the automatic classification of mineral images, and the performance of different model architectures is analyzed comparatively. The Mineral Photos dataset, containing a total of 39,354 images belonging to 15 different mineral classes, was used in the study. In the experimental process, five different pre-trained models were evaluated: ResNet50, EfficientNet-B0, MobileNetV3-Large, DenseNet121, and Vision Transformer (ViT-

* Corresponding author/Sorumlu yazar

B/16). All models were tested using a 5-fold cross-validation method under the same training protocol. According to the first stage results, the EfficientNet-B0 model achieved the highest performance with %74.39 accuracy, while the ResNet50 (%71.37) and DenseNet121 (%70.13) models presented competitive results. In contrast, the ViT-B/16 model showed lower performance compared to other models with an accuracy of %48.61. In the second stage, more advanced experiments were conducted on EfficientNet-based models, and it was determined that the EfficientNet-V2-S model exhibited higher and more stable performance with an average accuracy of %83.14. Furthermore, the model produced balanced results in precision, recall, and F1-score metrics. The findings indicate that, under the Mineral Photos dataset used in this study and the applied experimental conditions, the evaluated CNN-based models exhibited higher classification performance compared to the ViT-B/16 model.

Keywords: Mineral classification, deep learning, image processing, transfer learning, convolutional neural networks.

1. GİRİŞ

Mineraller, jeoloji, madencilik, malzeme bilimi, endüstriyel üretim ve ekonomik kaynak yönetimi gibi çok sayıda alanda kritik öneme sahip doğal oluşumlardır [1]. Bir mineralin doğru şekilde tanımlanması, cevher arama, rezerv tespiti, kayaç analizi, endüstriyel hammadde ayrıştırma, mücevherat sınıflandırması ve laboratuvar tabanlı jeolojik incelemeler açısından temel bir gereklilik oluşturmaktadır [2]. Geleneksel mineral tanımlama süreçleri çoğunlukla uzman görüşüne, mikroskobik incelemeye, kristal yapı analizine, renk-parlaklık-doku değerlendirmesine ve bazı durumlarda kimyasal testlere dayanmaktadır [3]. Ancak bu yöntemler zaman alıcı, maliyetli ve uzman bağımlı olmalarının yanı sıra, özellikle görsel olarak benzer özellikler taşıyan mineral türlerinde öznel değerlendirmelere açık olabilmektedir. Son yıllarda yapay zekâ ve özellikle derin öğrenme yöntemlerinin görüntü analizi alanında sağladığı önemli gelişmeler, mineral sınıflandırma gibi uzmanlık gerektiren problemlerde otomatik ve yüksek doğruluklu çözümler geliştirilmesini mümkün hâle getirmiştir [4]. Derin öğrenme modelleri, ham görüntülerden doğrudan ayırt edici özellikler öğrenebilmekte ve klasik görüntü işleme tabanlı manuel öznitelik çıkarım süreçlerine olan ihtiyacı azaltmaktadır [5]. Özellikle konvolüsyonel sinir ağları (CNN) tabanlı mimariler, doku, kenar, renk geçişleri, yüzey yapısı ve ince morfolojik farklılıkların öğrenilmesinde oldukça başarılı sonuçlar vermektedir [6]. Buna ek olarak, son dönemde transformer tabanlı görsel mimariler de görüntü sınıflandırma alanında dikkat çekici bir araştırma yönü hâline gelmiştir [7].

Mineral görüntülerinin sınıflandırılması genel amaçlı nesne sınıflandırma problemlerine kıyasla daha zorlu bir yapıya sahiptir. Bunun temel nedenleri arasında sınıflar arası yüksek görsel benzerlik, aynı mineralin farklı renk ve biçim varyasyonlarıyla gözlenebilmesi, farklı ışıklandırma ve çekim koşulları, arka plan değişkenliği ve veri setlerinde görülen sınıf dengesizliği yer almaktadır [8]. Örneğin, bazı mineral türleri benzer kristal yapı, parlaklık veya yüzey dokusu sergileyebilirken, aynı mineral türü farklı örneklerde oldukça farklı görsel özellikler gösterebilmektedir. Bu durum, sınıflandırma problemini daha karmaşık hâle getirmekte ve güçlü temsil öğrenme yeteneğine sahip modellere ihtiyaç doğurmaktadır. Literatürde görüntü tabanlı sınıflandırma problemlerinde çok sayıda başarılı derin öğrenme uygulaması bulunmasına rağmen, mineral sınıflandırmaya odaklanan çalışmaların sayısı görece sınırlıdır. Mevcut çalışmaların önemli bir kısmı ya az sayıda sınıf içeren veri setleri üzerinde yürütülmüş, ya yalnızca tek bir model mimarisi ile değerlendirme yapmış, ya da sınırlı deneysel karşılaştırmalar sunmuştur. Ayrıca son yıllarda geliştirilen farklı mimari ailelerin, örneğin klasik derin CNN modelleri ile transformer tabanlı yaklaşımların, aynı mineral veri seti

üzerinde sistematik ve adil koşullar altında karşılaştırıldığı çalışma sayısı oldukça azdır. Bunun yanında, performans değerlendirmesinin yalnızca tek bir doğruluk metriği üzerinden verilmesi, modellerin gerçek genelleme yeteneklerinin ve kararlılıklarının yeterince analiz edilememesine yol açmaktadır.

Bu çalışmada, söz konusu boşluğu gidermek amacıyla Kaggle platformunda paylaşılan ve 15 farklı mineral sınıfına ait toplam 39.354 görüntü içeren Mineral Photos veri seti kullanılmıştır. Veri seti, Quartz, Calcite, Fluorite, Pyrite, Baryte, Pyromorphite, Beryl, Malachite, Wulfenite, Copper, Cerussite, Gypsum, Hematite, Azurite ve Smithsonite sınıflarını içermekte olup, sınıflar arasında belirgin bir örnek sayısı dengesizliği bulunmaktadır. Bununla birlikte veri seti, farklı mineral formlarını ve çeşitli görüntüleme koşullarını içermesi bakımından gerçek dünya senaryolarını yansıtan zengin bir görsel çeşitlilik sunmaktadır. Bu özellikler, veri setini derin öğrenme tabanlı sınıflandırma modellerinin karşılaştırmalı olarak değerlendirilmesi için uygun bir test ortamı hâline getirmektedir. Veri setinin 15 sınıf ve 39.354 görüntüden oluştuğu, ayrıca sınıf dağılımında belirgin dengesizlik bulunduğu çalışmanın veri seti tanımında açıkça belirtilmiştir.

Bu kapsamda çalışmada, farklı mimari karakteristiklere sahip beş adet önceden eğitilmiş derin öğrenme modeli değerlendirilmiştir: ResNet50, EfficientNet-B0, MobileNetV3-Large, DenseNet121 ve Vision Transformer (ViT-B/16). Bu modellerin seçilmesindeki amaç, artık bağlantılar kullanan derin CNN mimarileri, yoğun bağlantı yapısına sahip ağlar, hafif ve verimli mobil mimariler, bileşik ölçeklendirme yaklaşımı sunan modeller ve transformer tabanlı görsel öğrenme yaklaşımlarını aynı deneysel çerçevede incelemektir. Böylece yalnızca en yüksek doğruluk değerini veren modelin belirlenmesi değil, aynı zamanda farklı mimari yaklaşımların mineral sınıflandırma problemine ne ölçüde uyum sağlayabildiğinin ortaya konulması hedeflenmiştir.

Bu çalışma, literatüre birkaç açıdan özgün katkı sunmaktadır. İlk olarak, mineral görüntülerinin sınıflandırılmasına yönelik olarak farklı derin öğrenme mimarilerini aynı veri seti ve aynı deneysel koşullar altında karşılaştırmalı biçimde değerlendirmektedir. İkinci olarak, yalnızca temel CNN modelleri değil, aynı zamanda hafif mimariler ve transformer tabanlı bir model de analize dâhil edilerek mimari çeşitlilik sağlanmaktadır. Üçüncü olarak, çalışma iki aşamalı bir deneysel yapı sunmaktadır: ilk aşamada geniş model karşılaştırması yapılmakta, ikinci aşamada ise en umut verici model ailesi üzerinde daha derinlemesine optimizasyon ve performans analizi gerçekleştirilmektedir. Dördüncü olarak, model başarımı yalnızca doğruluk metriği ile değil kesinlik, duyarlılık, F1-skoru, doğrulama kaybı, kararlılık ve fold bazlı sonuçlar çerçevesinde değerlendirilmektedir. Beşinci olarak, sınıf dengesizliği ve sınıflar arası görsel benzerlik gibi gerçek dünya problemlerinin model başarımı üzerindeki etkileri tartışılarak, mineral sınıflandırma alanı için uygulanabilir çıkarımlar sunulmaktadır.

Çalışmanın literatüre temel katkıları aşağıdaki gibi özetlenebilir:

- 15 sınıf ve 39.354 görüntü içeren geniş ölçekli bir mineral veri seti üzerinde kapsamlı bir derin öğrenme karşılaştırması gerçekleştirilmiştir.
- ResNet50, EfficientNet-B0, MobileNetV3-Large, DenseNet121 ve ViT-B/16 modelleri aynı deneysel koşullar altında değerlendirilerek mimariler arası adil bir karşılaştırma sunulmuştur.
- İlk aşama sonuçlarına göre öne çıkan EfficientNet tabanlı modeller üzerinde gelişmiş veri artırma, sınıf ağırlıklandırma, AdamW optimizasyonu ve öğrenme oranı planlayıcısı gibi stratejilerle ikinci aşama deneyler yürütülmüştür.

- EfficientNet-V2-S modelinin %83.14 doğruluk ile daha yüksek ve daha kararlı performans sunduğu gösterilmiştir.
- Mineral Photos veri seti ve bu çalışmada uygulanan deneysel protokol kapsamında, değerlendirilen CNN tabanlı modellerin ViT-B/16 modeline kıyasla daha yüksek performans gösterdiği deneysel olarak ortaya konulmuştur.

2. LİTERATÜRDEKİ ÇALIŞMALAR

Son yıllarda derin öğrenme tabanlı görüntü analizi yöntemleri, jeoloji, mineraloji ve petrografi alanlarında giderek daha fazla kullanılmaya başlanmıştır. Özellikle kayaç, cevher, ince kesit ve hiperspektral mineral görüntülerinin otomatik olarak tanınması, uzman bağımlılığını azaltma, analiz süresini kısaltma ve daha tutarlı sonuçlar üretme potansiyeli nedeniyle dikkat çekmektedir. Bununla birlikte, güncel literatür incelendiğinde çalışmaların büyük bölümünün ince kesit görüntüleri, kayaç çekirdekleri, cevher örnekleri veya hiperspektral veriler üzerinde yoğunlaştığı, doğrudan çok sınıflı RGB mineral fotoğraflarını aynı deney düzeneğinde çoklu derin öğrenme mimarileriyle karşılaştıran çalışma sayısının daha sınırlı olduğu görülmektedir.
Chawshin ve ark. [9], bütün çekirdek CT görüntülerinden litoloji sınıflandırması yapmak için evrişimli sinir ağları kullanan bir iş akışı önermiştir. Çalışmada 2B CT dilimleri üzerinden eğitim yapılmış, istenmeyen artefaktlar ön işleme ile elenmiş ve sonrasında benzer litofasiyes sınıfları karışıklık matrisi yardımıyla birleştirilerek dört ana kayaç sınıfı elde edilmiştir. Yazarlar, CNN tabanlı yaklaşımın özellikle ince tabakalar ve sınırların belirlenmesinde manuel tanımlamaya tamamlayıcı bilgi sunduğunu vurgulamıştır. Bu çalışma, jeolojik görüntülerde derin öğrenmenin yalnızca sınıflandırma değil, aynı zamanda yorumlama hassasiyetini artırma potansiyelini göstermesi açısından önemlidir.

Xu ve ark. [10], mikroskobik kayaç görüntülerinden akıllı litoloji tanımlaması amacıyla yedi farklı CNN mimarisini karşılaştırmıştır: Xception, MobileNetV2, Inception-ResNetV2, InceptionV3, DenseNet121, ResNet101V2 ve ResNet-101. Transfer öğrenme ve veri artırma kullanılan çalışmada, Xception tabanlı modelin eğitim kümesinde %97.66, test kümesinde ise %98.65 doğruluk elde ettiği bildirilmiştir. Ayrıca yalnızca doğruluk değil, hız metriği de değerlendirilmiş. Böylece model seçiminin hem performans hem de uygulanabilirlik boyutunda ele alınması sağlanmıştır. Bu yönüyle çalışma, mimari karşılaştırmanın jeolojik görüntü sınıflandırmada ne kadar kritik olduğunu açık biçimde ortaya koymaktadır.
Zhou ve ark. [11], cevher görüntülerinin sınıflandırılması için geliştirilmiş bir CNN yaklaşımı önermiştir. Çalışmada yetersiz eğitim verisi problemini aşmak amacıyla transfer öğrenme ve veri artırma kullanılmış, ardından MobileNet tabanlı özellik çıkarımı ile SENet bileşeni birleştirilmiştir. ScienceDirect kaydına göre bu hibrit yaklaşım cevher sınıflandırmada yüksek performans üretmiş ve otomatik cevher tanıma için daha az eğitim verisiyle etkili bir çözüm sunduğunu göstermiştir. Makalenin özet sayfasında yaklaşık %96 düzeyinde sınıflandırma başarısı rapor edilmiştir. Bu çalışma, özellikle sınırlı veri koşullarında hafif mimariler ve dikkat mekanizmalarının birlikte kullanılmasının yararını göstermektedir.

Zhou ve ark. [12], kayaç ince kesit görüntülerinin sınıflandırılmasında adaptive pooling ve second-order pooling temelli bir CNN yaklaşımı önermiştir. Çalışma, ince kesit görüntülerinde klasik konvolüsyonel özniteliklerin yanı sıra daha yüksek mertebeden ilişkilerin de öğrenilmesini hedeflemiştir. Yayın kaydında bu yaklaşımın ince kesit tanıma doğruluğunu artırdığı ve yaklaşık %89.08 doğruluk rapor ettiği görülmektedir. Bu araştırma, jeolojik görüntülerde yalnızca derin ağ seçiminin değil, ağ içi havuzlama stratejilerinin de sonuçlar üzerinde belirgin etkisi olabileceğini göstermektedir.

Rizaldy ve ark. [13], mineral sınıflandırmasını iyileştirmek için tek bir veri türüne bağlı kalmak yerine hiperspektral görüntü, RGB görüntü ve 3B nokta bulutunu bir araya getiren çok akışlı bir sinir ağı önermiştir. Çalışmada hiperspektral akış üzerinde kanal dikkat modülü de kullanılmıştır. Sonuçlar, çok kipli birleşimin tek kipli kullanıma göre LWIR, SWIR ve VNIR veri kümelerinde sırasıyla %19.2, %4.4 ve %5.6 doğruluk artışı sağladığını göstermiştir. Bu çalışma, mineral sınıflandırmada veri füzyonunun özellikle spektral ve geometrik tamamlayıcılığı kullanarak önemli performans kazanımları sağlayabildiğini göstermesi bakımından dikkat çekicidir.

Li ve ark. [14], derin transfer öğrenmeye dayalı otomatik litoloji tanımlama sistemi geliştirmiştir. Araştırmacılar yaklaşık 150.000 kayaç görüntüsünden oluşan büyük bir veri çerçevesi kurmuş ve 12 yaygın mühendislik kayaç türü için uçtan uca CNN tabanlı bir tanıma modeli önermiştir. Springer kayıtlarına göre önerilen model, %95'in üzerinde doğruluk ve görüntü başına yaklaşık 0.08 saniye tanıma süresi elde etmiştir. Bu çalışma, büyük ölçekli veri kümeleriyle transfer öğrenmenin birlikte kullanımının saha uygulamaları için pratik ve hızlı çözümler üretilebileceğini göstermektedir.

Ma ve ark. [15], tek kutuplu (single polarized) ve çapraz kutuplu (orthogonal polarized) kayaç görüntülerinin birlikte kullanılmasına dayalı bir özellik füzyon modeli önermiştir. NFNet tabanlı mimari ile geliştirilen bu yaklaşımda iki farklı görüntü kipinden çıkarılan öznitelikler giriş ve özellik düzeyinde birleştirilmiştir. Test kümesinde %98.66 doğruluk elde edilmiştir. Bu performans, yalnızca tek kutuplu görüntü kullanımına göre %4.91 ve yalnızca çapraz kutuplu görüntü kullanımına göre %1.55 artış anlamına gelmiştir. Bu sonuç, ince kesit analizinde çoklu optik görünüm bilgisinin sınıflandırma başarısını anlamlı biçimde artırabildiğini göstermektedir.

Wang ve ark. [16] ise veri azlığı sorununa odaklanmış ve ince kesit kayaç görüntülerinin sınıflandırılması için few-shot learning yaklaşımı geliştirmiştir. Çalışmada PPL ve XPL görüntülerini birleştiren çapraz dikkatli özellik füzyonu (CAFF) ve prototip ağ tabanlı özellik seçimi modülü kullanılmıştır. ResNet50 ve Swin-Transformer omurgalarıyla yürütülen deneylerde, 5-way 5-shot görevinde önerilen ProtoNet-CAFF-FS modelinin ortalama doğruluğu %96.70 ve %99.16'ya kadar çıkmıştır. Bu çalışma, özellikle az örnekli jeolojik veri senaryolarında meta-öğrenme ve çok kaynaklı özellik birleştirmenin güçlü bir alternatif sunduğunu göstermektedir.

Bu güncel çalışmalar birlikte değerlendirildiğinde üç temel eğilim öne çıkmaktadır. Birincisi, transfer öğrenme ve önceden eğitilmiş derin ağların jeolojik görüntü sınıflandırmada yaygın ve etkili biçimde kullanılmasıdır. İkincisi, PPL/XPL, RGB/hiperspektral veya görsel/elementel veri gibi çok kipli birleşimlerin performansı belirgin şekilde artırabilmesidir. Üçüncüsü ise, yakın dönem çalışmaların önemli kısmının ince kesit, çekirdek, cevher veya hiperspektral veriler üzerine yoğunlaşması, buna karşılık büyük ölçekli RGB mineral fotoğrafı veri kümelerinde farklı CNN ve transformer tabanlı mimarilerin aynı deney protokolü altında sistematik biçimde karşılaştırıldığı araştırmaların görece az olmasıdır.

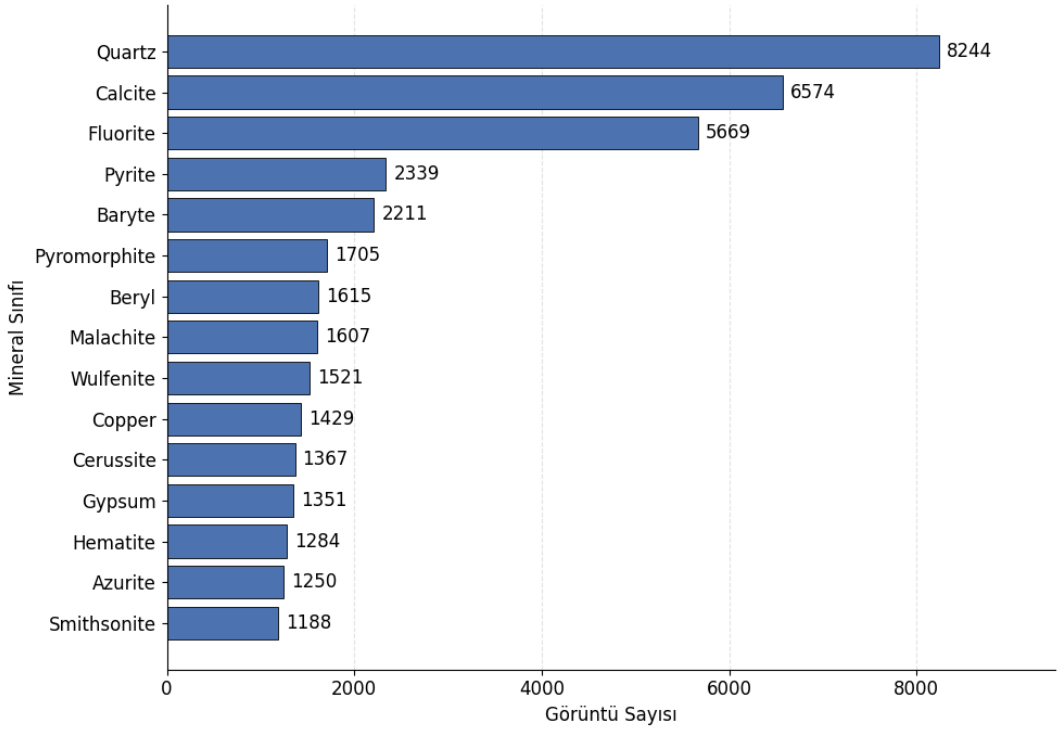
Bu nedenle mevcut çalışma, 15 sınıf ve 40 binden fazla görüntü içeren geniş ölçekli bir mineral fotoğraf veri kümesi üzerinde ResNet50, EfficientNet-B0, MobileNetV3-Large, DenseNet121 ve ViT-B/16 gibi farklı mimari aileleri aynı deneysel kurgu altında karşılaştırarak literatürdeki önemli bir boşluğu doldurmayı hedeflemektedir. Ayrıca çalışmanın ikinci aşamasında EfficientNet tabanlı modeller üzerinde daha ileri optimizasyon stratejileri uygulanması, literatürde sık görülen "tek aşamalı model raporlama" yaklaşımının ötesine geçmektedir.

3. MATERYAL VE METOT

Bu çalışmada, mineral görüntülerinin otomatik sınıflandırılmasına yönelik olarak derin öğrenme tabanlı bir yaklaşım benimsenmiştir. Önerilen yöntem, veri setinin hazırlanmasından başlayarak model seçimi, eğitim süreci ve performans değerlendirme aşamalarını kapsayan sistematik bir süreçten oluşmaktadır. İlk aşamada, farklı mineral türlerini içeren geniş kapsamlı bir görüntü veri seti kullanılarak veri ön işleme ve düzenleme işlemleri gerçekleştirilmiştir. Ardından, farklı mimari özelliklere sahip derin öğrenme modelleri belirlenmiş ve bu modeller transfer öğrenme yaklaşımı ile probleme uyarlanmıştır. Çalışmanın devamında, seçilen modeller aynı deneysel koşullar altında eğitilmiş ve karşılaştırılabilir sonuçlar elde edilebilmesi amacıyla standart bir eğitim protokolü uygulanmıştır. Model performanslarının güvenilir bir şekilde değerlendirilebilmesi için çapraz doğrulama yöntemi kullanılmış ve elde edilen sonuçlar farklı performans metrikleri ile analiz edilmiştir. Ayrıca, model başarımını artırmak amacıyla veri artırma teknikleri ve hiperparametre optimizasyonu gibi stratejiler de sürece dahil edilmiştir. Bu bölümde, çalışmada kullanılan veri seti, tercih edilen derin öğrenme modelleri ve performans değerlendirme metrikleri ayrıntılı olarak açıklanmakta, önerilen yöntemin tüm bileşenleri sistematik bir şekilde sunulmaktadır.

3.1. Veri Seti

Çalışmada kullanılan veri seti, Kaggle platformunda paylaşılan Mineral Photos veri setidir [17]. Veri setindeki sınıf dağılımı şekil 1'de sunulmuştur. Toplamda 15 sınıf ve 40354 görüntüden oluşan veri setinde sınıflar arasında belirgin bir dengesizlik bulunmaktadır. Quartz sınıfının diğer sınıflara göre daha yüksek sayıda örnek içerdiği görülmektedir. Buna karşın smithsonite sınıfının veri setinde daha düşük görsel sayısı ile temsil edildiği görülmektedir.



Şekil 1. Mineral Photos veri setinde yer alan sınıflara ait görüntü sayılarının dağılımı

Veri setinde yer alan görüntüler minerallerin farklı formlarını ve çeşitli ışıklandırma koşullarını içerecek şekilde oluşturulmuştur [17]. Bu durum, veri setinin görsel çeşitliliğini artırarak modelin farklı koşullara karşı daha genellenebilir özellikler öğrenmesine olanak sağlamaktadır. Veri setine ait her bir sınıfı temsil eden örnek görüntüler Şekil 2'de sunulmuştur.



Şekil 2. Mineral Photos veri setinde yer alan her bir sınıfı temsil eden örnek görüntüler

Veri setinde yer alan mineraller, renk, doku ve kristal yapı bakımından önemli farklılıklar göstermektedir. Bununla birlikte, bazı mineral sınıfları benzer görsel özelliklere sahip olduğundan sınıflar arasında ayrım yapılmasını zorlaştırabilmektedir. Bu görsel çeşitlilik, veri setinin sınıflandırma açısından zorluğunu artırmaktadır.

3.2. Veri Artırma

Derin öğrenme modellerinin eğitiminde, sınırlı veri miktarı ve görüntüler arasındaki yönelim farklılıklarının model performansını olumsuz etkilemesini önlemek amacıyla veri artırma tekniklerinden yararlanılmıştır. Veri artırma işlemleri yalnızca eğitim verileri üzerinde uygulanmış, doğrulama aşamasında ise herhangi bir veri artırma gerçekleştirilmemiştir. Böylece model performansının tarafsız bir şekilde değerlendirilmesi amaçlanmıştır.

Çalışmanın ilk deneysel aşamasında, farklı derin öğrenme mimarilerinin aynı koşullar altında karşılaştırılabilmesi amacıyla tüm görüntüler 224 × 224 piksel boyutuna yeniden ölçeklendirilmiş ve yalnızca rastgele yatay çevirme işlemi uygulanmıştır.

İkinci deneysel aşamada ise, yüksek performans gösteren EfficientNet tabanlı modeller üzerinde daha kapsamlı veri artırma stratejileri uygulanmıştır. Bu aşamada giriş görüntüleri 384 × 384 piksel boyutuna yeniden ölçeklendirilmiş, rastgele yatay ve dikey çevirme, ±30° rastgele döndürme ile parlaklık ve kontrast değişimlerini içeren renk değişimi işlemleri gerçekleştirilmiştir. Bu sayede eğitim verilerinin çeşitliliği artırılarak modellerin farklı yönelim, aydınlatma ve görünüm koşullarına karşı daha dayanıklı özellikler öğrenmesi hedeflenmiştir.

Doğrulama verileri üzerinde ise yalnızca yeniden boyutlandırma ve ImageNet normalizasyonu uygulanmıştır.

3.3. Sınıflandırma Modelleri

Bu çalışmada, mineral görüntülerinin sınıflandırılması amacıyla ResNet50, EfficientNet-B0, MobileNetV3-Large, DenseNet121 ve ViT-B/16 olmak üzere beş adet derin öğrenme tabanlı model kullanılmıştır. Seçilen modeller, literatürde yaygın olarak kullanılan ve farklı mimari yaklaşımları temsil eden güçlü sınıflandırma algoritmalarıdır. Bu sayede, CNN ile transformer tabanlı yaklaşımlar arasında kapsamlı bir karşılaştırma yapılması amaçlanmıştır.
Tüm modeller, transfer öğrenme yaklaşımı kapsamında ImageNet veri seti üzerinde önceden eğitilmiş ağırlıklarla başlatılmıştır. Önceden eğitilmiş modellerin kullanılması, özellikle sınırlı veri setlerinde modelin daha hızlı öğrenmesini sağlamakta ve genelleme performansını artırmaktadır. Bu doğrultuda, modellerin son sınıflandırma katmanları kaldırılarak, yerine veri setindeki 15 sınıfa uygun yeni tam bağlantılı (fully connected) katmanlar eklenmiştir.
ResNet50, derin sinir ağlarında karşılaşılan gradyan kaybolması problemini çözmek amacıyla önerilen artık bağlantı (residual connection) yapısını kullanmaktadır [18]. Bu modelde, katmanlar arasında doğrudan bağlantılar kurularak bilginin daha etkin şekilde aktarılması sağlanmaktadır [19]. 50 katmanlı derin bir mimariye sahip olan ResNet50, özellikle karmaşık görsel özelliklerin öğrenilmesinde başarılı sonuçlar vermektedir [20]. Mineral görüntülerindeki ince doku farklılıklarının yakalanmasında bu modelin derin yapısı önemli avantaj sağlamaktadır.

EfficientNet-B0, derinlik, genişlik ve çözünürlük parametrelerini dengeli bir şekilde ölçeklendiren bileşik ölçeklendirme (compound scaling) yöntemi ile geliştirilmiştir [21]. Bu yaklaşım sayesinde model, daha az parametre ile yüksek doğruluk elde edebilmektedir. Özellikle hesaplama maliyeti ile performans arasındaki dengeyi optimize etmesi nedeniyle EfficientNet-B0, görüntü sınıflandırma problemlerinde oldukça etkili bir model olarak öne çıkmaktadır [22]. Bu çalışmada da EfficientNet-B0 modelinin yüksek performans sergilemesi beklenmektedir.

MobileNetV3-Large, mobil ve gömülü sistemler için optimize edilmiş hafif bir derin öğrenme mimarisidir [23]. Bu model, derin ayrılabilir konvolüsyonlar (depthwise separable convolutions) ve squeeze-and-excitation blokları gibi teknikler kullanarak parametre sayısını azaltmakta ve hesaplama maliyetini düşürmektedir [24]. Aynı zamanda, otomatik mimari arama ile optimize edilmiştir. Bu özellikleri sayesinde MobileNetV3-Large, düşük donanım kaynakları ile çalışan sistemler için uygun bir çözüm sunarken rekabetçi performans elde edebilmektedir [25].

DenseNet121, her katmanın kendisinden önceki tüm katmanlarla doğrudan bağlantı kurduğu yoğun bağlantı (dense connectivity) yapısını kullanmaktadır [26]. Bu sayede özellik yeniden kullanımı artırılmakta ve modelin parametre verimliliği sağlanmaktadır. DenseNet mimarisi, özellikle sınırlı veri setlerinde aşırı öğrenmenin önlenmesine katkı sağlamaktadır [27]. Mineral sınıflandırma probleminde, benzer görsel özelliklere sahip sınıfların ayırt edilmesinde DenseNet121'in güçlü temsil yeteneği önemli avantaj sunmaktadır.

ViT-B/16, görüntü sınıflandırma problemlerinde konvolüsyonel yapılar yerine transformer tabanlı bir yaklaşım kullanmaktadır [28]. Bu modelde görüntüler, sabit boyutlu parçalara (patch) bölünmekte ve bu parçalar sıralı veri olarak işlenmektedir [29]. Self-attention mekanizması sayesinde model, görüntü üzerindeki uzun menzilli bağımlılıkları

öğrenebilmektedir [30]. Ancak transformer tabanlı modeller genellikle büyük veri setlerine ihtiyaç duymakta olup, sınırlı veri durumlarında performans düşüşü gözlenebilmektedir [31]. Bu çalışmada ViT modelinin performansı, CNN tabanlı modeller ile karşılaştırmalı olarak değerlendirilmiştir.

3.4. Deneysel Kurulum

Bu çalışmada deneyler iki aşamada gerçekleştirilmiştir. İlk aşamada, farklı derin öğrenme mimarilerinin mineral görüntülerini sınıflandırma performanslarının karşılaştırılması amaçlanmıştır. Bu kapsamda ResNet50, EfficientNet-B0, MobileNetV3-Large, DenseNet121 ve ViT-B/16 modelleri aynı deneysel koşullar altında eğitilmiş ve elde edilen sonuçlar karşılaştırılmıştır. Tüm modellerde önceden eğitilmiş ağırlıklar kullanılmış, son sınıflandırma katmanları veri setindeki sınıf sayısına uygun olarak yeniden düzenlenmiş ve modeller uçtan uca eğitilmiştir. Model performanslarının güvenilir biçimde değerlendirilebilmesi amacıyla beş katlı çapraz doğrulama yöntemi uygulanmıştır.

Veri bölme işlemi görüntü düzeyinde gerçekleştirilmiştir. Beş katlı çapraz doğrulama sürecinde her mineral sınıfının veri setindeki dağılımının katlarda yaklaşık olarak korunması sağlanmıştır. Veri setinde görüntülerin aynı fiziksel mineral örneğine veya aynı çekim kaynağına ait olup olmadığını belirten bir grup kimliği sunulmadığından, benzer kaynaklı görüntülerin farklı katlarda yer alıp almadığı ayrıca kontrol edilememiştir.

İlk aşamada elde edilen sonuçlar doğrultusunda en yüksek performansı gösteren EfficientNet tabanlı modeller ikinci aşamada daha ayrıntılı olarak incelenmiştir. Bu aşamada EfficientNet-B0 ve EfficientNet-V2-S modelleri daha yüksek giriş çözünürlüğü, sınıf ağırlıklandırması ve gelişmiş eğitim stratejileri kullanılarak yeniden eğitilmiş ve performansları karşılaştırılmıştır. Eğitim sürecinde doğrulama başarımı izlenmiş, en yüksek doğrulama başarımını sağlayan model ağırlıkları kaydedilmiş ve değerlendirmelerde bu ağırlıklar kullanılmıştır.

Her iki deneysel aşamada da modeller aynı veri bölme stratejisi kullanılarak eğitilmiş ve tüm performans değerlendirmeleri doğrulama sonuçlarının ortalaması alınarak gerçekleştirilmiştir. Böylece farklı model mimarilerinin ve uygulanan eğitim stratejilerinin sınıflandırma başarımına etkisi adil ve karşılaştırılabilir bir biçimde değerlendirilmiştir.

Tüm deneyler Google Colaboratory (Google Colab) bulut tabanlı çalışma ortamında gerçekleştirilmiştir. Model geliştirme ve eğitim süreçlerinde Python programlama dili ile derin öğrenme ve görüntü işleme kütüphanelerinden yararlanılmıştır. Deneylerin tekrarlanabilirliğini artırmak amacıyla kullanılan temel yazılım bileşenlerinin ve kütüphanelerin sürümleri kayıt altına alınmıştır. Hesaplama işlemleri, Google Colab tarafından sağlanan GPU hızlandırıcısı kullanılarak gerçekleştirilmiştir. Kullanılan çalışma ortamına ilişkin temel yazılım ve donanım özellikleri Tablo 1'de sunulmuştur.

Tablo 1. Deneylerde kullanılan yazılım ve donanım ortamı

Bileşen	Sürüm / Özellik
Çalışma ortamı	Google Colaboratory (GPU runtime)
İşletim sistemi	Linux (Google Colab çalışma ortamı)
Programlama dili	Python 3.12
Derin öğrenme çatısı	PyTorch 2.6.0
Görüntü/model kütüphanesi	Torchvision 0.21.0

Bileşen	Sürüm / Özellik
Makine öğrenmesi kütüphanesi	Scikit-learn 1.6.1
Sayısal hesaplama kütüphanesi	NumPy 2.0.2
CUDA	CUDA 12.4
GPU	NVIDIA Tesla T4
GPU belleği	15 GB

3.5. Hiperparametre Seçimi

Bu çalışmada gerçekleştirilen ilk deneysel aşamada, farklı derin öğrenme mimarilerinin performanslarının adil bir şekilde karşılaştırılabilmesi amacıyla tüm modeller aynı eğitim parametreleri kullanılarak eğitilmiştir. İlk deneyler sonucunda en yüksek sınıflandırma başarısını gösteren EfficientNet tabanlı modeller belirlenmiş ve ikinci deneysel aşamada bu modeller üzerinde daha ayrıntılı performans analizleri gerçekleştirilmiştir. Bu kapsamda, modelin öğrenme başarısını ve genelleme yeteneğini artırmaya yönelik giriş görüntü boyutu, eğitim süresi, optimizasyon algoritması, öğrenme oranı, veri artırma stratejileri ve düzenlileştirme yöntemleri yeniden düzenlenmiştir. Ayrıca deneylerin tekrar üretilebilirliğini sağlamak amacıyla Stratified K-Fold veri bölme yöntemi kullanılmış, veri karıştırma etkinleştirilmiş ve rastgelelik tohumu değeri 42 olarak belirlenmiştir. İkinci deneysel aşamada kullanılan hiperparametreler Tablo 2'de sunulmuştur.

Tablo 2. EfficientNet tabanlı modeller için kullanılan hiperparametreler

Parametre	Değer
Model	EfficientNet-B0, EfficientNet-V2-S
Giriş boyutu	384 × 384
Yığın büyüklüğü	32
Epok sayısı	40
Çapraz doğrulama	5 katlı Stratified K-Fold (random state = 42)
Optimizasyon algoritması	AdamW
Başlangıç öğrenme oranı	1×10^{-4}
Ağırlık azaltma katsayısı (Weight Decay)	1×10^{-2}
Kayıp fonksiyonu	Sınıf ağırlıklı çapraz entropi kaybı
Öğrenme oranı planlayıcısı	Cosine Annealing Warm Restarts
Veri artırma	Yatay ve dikey çevirme, rastgele döndürme, parlaklık ve kontrast değişimi
Erken durdurma	Patience = 5
Normalizasyon	ImageNet

3.6. Performans Değerlendirme Metrikleri

Bu çalışmada geliştirilen derin öğrenme tabanlı sınıflandırma modellerinin performanslarının kapsamlı bir şekilde değerlendirilmesi amacıyla birden fazla metrik kullanılmıştır. Çok sınıflı sınıflandırma problemlerinde tek bir metriğe bağlı kalmak, model performansını yeterince temsil edemeyebileceğinden, farklı açılardan değerlendirme yapılması hedeflenmiştir. Bu doğrultuda doğruluk (accuracy), kesinlik (precision), duyarlılık (recall) ve F1-skoru değerleri birlikte analiz edilmiştir.

Doğruluk metriği, modelin doğru sınıflandırdığı örneklerin toplam örnek sayısına oranını ifade etmektedir [32]. Çok sınıflı sınıflandırma problemlerinde en yaygın kullanılan performans ölçütlerinden biridir. Doğruluk metriği Eş.1 kullanılarak tanımlanmaktadır:

$$\text{Doğruluk} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \tag{1}$$

Burada TP (True Positive) doğru sınıflandırılan pozitif örnekleri, TN (True Negative) doğru sınıflandırılan negatif örnekleri, FP (False Positive) yanlış pozitif tahminleri ve FN (False Negative) ise yanlış negatif tahminleri ifade etmektedir. Ancak veri setindeki sınıf dengesizliği durumlarında doğruluk metriği tek başına yeterli bir değerlendirme sunmamaktadır.

Kesinlik metriği, modelin pozitif olarak tahmin ettiği örneklerin ne kadarının gerçekten doğru olduğunu ölçmektedir [33]. Özellikle yanlış pozitif sonuçların kritik olduğu durumlarda önemli bir metriktir. Kesinlik değeri Eş.2 kullanılarak hesaplanmaktadır:

$$\text{Kesinlik} = \frac{\text{TP}}{\text{TP+FP}} \tag{2}$$

Mineral sınıflandırma probleminde, özellikle birbirine görsel olarak benzeyen sınıflar arasında yanlış sınıflandırmaların değerlendirilmesinde bu metrik önemli bir rol oynamaktadır.
Duyarlılık metriği, gerçek pozitif örneklerin ne kadarının doğru şekilde tespit edildiğini ölçmektedir [34]. Bu metrik özellikle modelin belirli bir sınıfa ait örnekleri kaçırma oranını değerlendirmek açısından önemlidir. Duyarlılık değeri Eş.3 kullanılarak hesaplanmaktadır:

$$\text{Duyarlılık} = \frac{\text{TP}}{\text{TP+FN}} \tag{3}$$

Dengesiz veri setlerinde, düşük sayıda örneğe sahip sınıfların doğru şekilde tanımlanıp tanımlanmadığını analiz etmek için duyarlılık metriği kritik öneme sahiptir.

F1-skoru, kesinlik ve duyarlılık metriklerinin harmonik ortalaması olarak tanımlanmaktadır. Bu metrik, modelin hem yanlış pozitif hem de yanlış negatif hatalarını dengeli bir şekilde değerlendirmektedir [35]. F1-skoru Eş.4 kullanılarak hesaplanmaktadır:

$$\text{F}-\text{skoru} = \frac{\text{2.Kesinlik.Duyarlılık}}{\text{Kesinlik+Duyarlılık}} \tag{4}$$

F1-skoru, özellikle sınıf dengesizliğinin bulunduğu veri setlerinde model performansını daha doğru yansıtan bir metrik olarak öne çıkmaktadır.

Ayrıca, çapraz doğrulama katları arasında elde edilen performans farklılıklarının belirsizliğini ve sonuçların güvenilirliğini nicel olarak göstermek amacıyla ortalama doğruluk değerleri için %95 güven aralıkları hesaplanmıştır. Beş katlı çapraz doğrulama kullanıldığından, güven aralıklarının hesaplanmasında Student-t dağılımı esas alınmış ve $CI_{95\%} = \bar{x} \pm t_{0.975,n-1}(s/\sqrt{n})$ eşitliği kullanılmıştır. Burada $\bar{x}$ katlar üzerinden elde edilen ortalama doğruluğu, s standart sapmayı ve n=5 çapraz doğrulama kat sayısını ifade etmektedir. Bu analiz, modellerin yalnızca

ortalama performanslarının değil, aynı zamanda elde edilen sonuçların belirsizlik düzeyinin de değerlendirilmesine olanak sağlamaktadır.

4. DENEYSEL SONUÇLAR

Mineral Photos veri seti üzerinde gerçekleştirilen deneyler iki aşamada yürütülmüştür. İlk aşamada, farklı mimari özelliklere sahip ve görüntü sınıflandırma çalışmalarında yaygın olarak kullanılan ResNet50, EfficientNet-B0, MobileNetV3-Large, DenseNet121 ve ViT-B/16 modellerinin mineral görüntülerinin sınıflandırılmasındaki performansları karşılaştırılmıştır. Bu modeller; artık bağlantılar, yoğun bağlantılar, hafif mimariler, bileşik ölçeklendirme ve dikkat tabanlı transformer yaklaşımı gibi farklı mimari özellikleri temsil etmeleri nedeniyle seçilmiştir. Böylece farklı derin öğrenme yaklaşımlarının aynı veri kümesi üzerindeki davranışları karşılaştırmalı olarak değerlendirilmiştir.

İlk deneysel aşamada elde edilen ortalama doğruluk sonuçları Tablo 3'te sunulmuştur.

Tablo 3. Farklı derin öğrenme modellerinin mineral sınıflandırma performanslarının karşılaştırılması

Model	Doğruluk
ResNet50	0.7137
EfficientNet-B0	0.7439
MobileNetV3-Large	0.6962
DenseNet121	0.7013
ViT-B/16	0.4861

Tablo 3 incelendiğinde, EfficientNet-B0 modelinin 0.7439 doğruluk değeri ile en yüksek sınıflandırma başarısını elde ettiği görülmektedir. ResNet50 (0.7137) ve DenseNet121 (0.7013) modelleri benzer doğruluk değerleri elde ederken, MobileNetV3-Large modeli daha hafif bir mimariye sahip olmasına rağmen 0.6962 doğruluk oranı ile rekabetçi bir performans sergilemiştir. Buna karşılık ViT-B/16 modeli 0.4861 doğruluk değeri ile diğer modellere göre daha düşük başarı göstermiştir. Bu sonucun yalnızca veri seti büyüklüğünden değil, aynı zamanda kullanılan eğitim stratejisi, hiperparametre seçimi ve transformer tabanlı mimarilerin veri kümesinin özelliklerine uyumundan da etkilenmiş olabileceği değerlendirilmektedir.

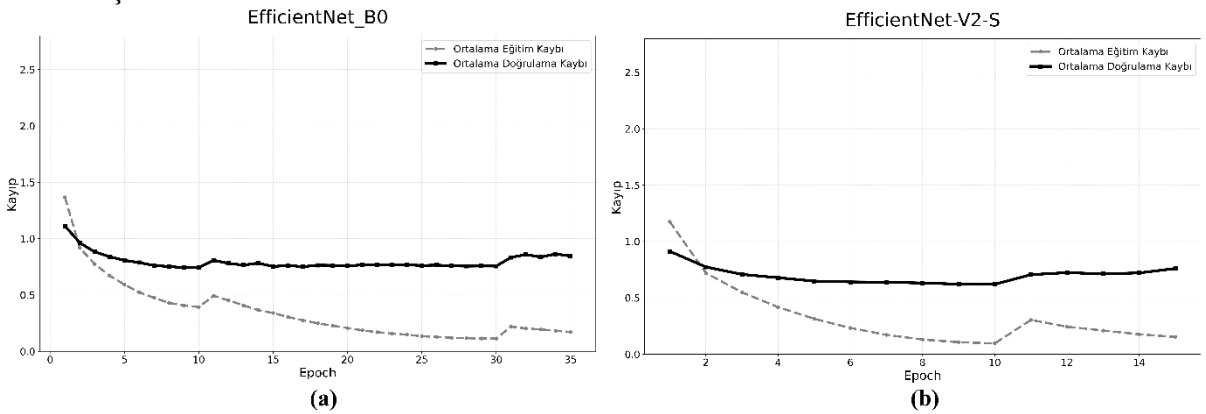
İlk deneysel aşamada elde edilen bulgular doğrultusunda, en yüksek doğruluk oranına ulaşan EfficientNet-B0 modeli temel alınarak ikinci deneysel aşama planlanmıştır. Bu aşamada EfficientNet-B0 modeli ile daha gelişmiş EfficientNet-V2-S mimarisi karşılaştırılmış ve geliştirilen eğitim stratejilerinin model performansına etkisi değerlendirilmiştir.

4.1. EfficientNet Tabanlı Modeller Üzerinde Gelişmiş Deneyler

Önceki deneylerde yüksek performans sergileyen EfficientNet tabanlı modeller üzerinde daha detaylı deneysel çalışmalar gerçekleştirilmiştir. Bu kapsamda EfficientNet-B0 ve EfficientNet-V2-S modelleri aynı deneysel kurulum altında karşılaştırılmıştır.

Her iki model için girdi görüntüleri 384×384 çözünürlüğe yeniden ölçeklendirilmiş ve veri setinin görsel özelliklerinden daha etkin faydalanılması amaçlanmıştır. Eğitim sürecinde yatay ve dikey çevirme, rastgele döndürme ve renk tabanlı artırma gibi veri artırma teknikleri uygulanmıştır. Doğrulama aşamasında ise yalnızca yeniden boyutlandırma ve normalizasyon işlemleri gerçekleştirilmiştir. Böylece veri sızıntısının önüne geçilmiştir.

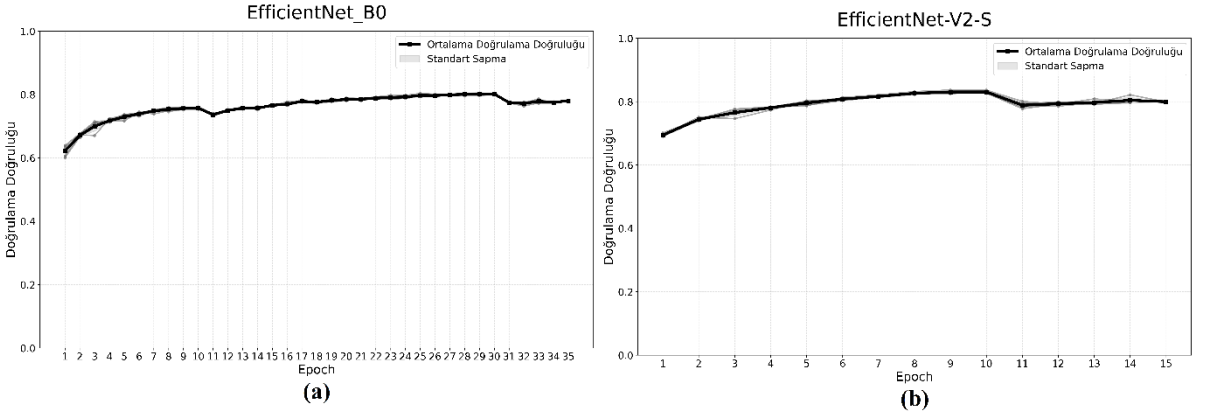
Modellerin eğitim ve doğrulama kayıp değerlerinin epoklara göre değişimi Şekil 3'te sunulmuştur.



Şekil 3. EfficientNet-B0 ve EfficientNet-V2-S modellerinin eğitim ve doğrulama kayıp değerlerinin epoklara göre değişimi

Şekil 3 incelendiğinde, her iki modelde de eğitim kaybının epok sayısı arttıkça düzenli olarak azaldığı görülmektedir. Eğitimler maksimum 40 epok olarak planlanmış olmasına rağmen, doğrulama kaybında iyileşmenin durması nedeniyle erken durdurma mekanizması devreye girmiş ve modeller farklı epok sayılarında eğitimi tamamlamıştır. EfficientNet-V2-S modelinin yaklaşık 15. epok, EfficientNet-B0 modelinin ise yaklaşık 35. epok civarında eğitimi sonlandırılmıştır. Bu durum, EfficientNet-V2-S modelinin daha kısa sürede yakınsamaya ulaştığını, EfficientNet-B0 modelinin ise benzer kararlılığa ulaşabilmek için daha uzun bir eğitim sürecine ihtiyaç duyduğunu göstermektedir

Modellerin doğrulama doğruluğu değerlerinin epoklara göre değişimi Şekil 4'te sunulmuştur.



Şekil 4. EfficientNet-B0 ve EfficientNet-V2-S modellerinin doğrulama doğruluğunun epoklara göre değişimi

Şekil 4 incelendiğinde, EfficientNet-V2-S modelinin daha yüksek doğrulama doğruluğuna ulaştığı ve eğitim süreci boyunca stabil bir performans sergilediği görülmektedir. Ayrıca EfficientNet-V2-S modelinin daha az sayıda epok içerisinde stabil performansa ulaştığı, buna karşın EfficientNet-B0 modelinin benzer performans seviyelerine ulaşmak için daha uzun bir eğitim süresine ihtiyaç duyduğu dikkat çekmektedir. Bu durum, EfficientNet-V2-S modelinin daha hızlı öğrenme ve daha verimli bir eğitim süreci sunduğunu göstermektedir.

EfficientNet-B0 ve EfficientNet-V2-S modellerinin genel performans sonuçları Tablo 4'te sunulmuştur.

Tablo 4. EfficientNet-B0 ve EfficientNet-V2-S modellerinin karşılaştırmalı performans sonuçları

Model	Doğruluk (ort ± ss)	%95 Güven Aralığı	Doğrulama kaybı (ort ± ss)	En iyi fold doğruluğu	Kesinlik	Duyarlılık	F1-skoru
EfficientNet-B0	0.7939 ± 0.0176	0.7720–0.8158	0.7545 ± 0.0405	0.8203	0.7935	0.8113	0.8010
EfficientNet-V2-S	**0.8314 ± 0.0033**	**0.8273–0.8355**	**0.6186 ± 0.0124**	**0.8367**	**0.84**	**0.84**	**0.84**

Tablo 4 incelendiğinde, EfficientNet-V2-S modelinin tüm performans ölçütlerinde EfficientNet-B0 modeline göre daha başarılı sonuçlar elde ettiği görülmektedir. Ortalama doğruluk değerinin %79,39'dan %83,14'e yükselmesi ve standart sapmanın ±0,0176'dan ±0,0033'e düşmesi, modelin yalnızca daha yüksek doğruluk üretmediğini, aynı zamanda farklı katlamalarda daha tutarlı sonuçlar verdiğini göstermektedir. Buna ek olarak, EfficientNet-B0 için %95 güven aralığı 0.7720–0.8158 olarak hesaplanırken, EfficientNet-V2-S için bu aralık 0.8273–0.8355 olarak elde edilmiştir. EfficientNet-V2-S modelinin daha dar güven aralığı, çapraz doğrulama katları boyunca performans değişkenliğinin daha düşük ve modelin tahmin performansının daha kararlı olduğunu desteklemektedir. Ayrıca iki modelin doğruluk için elde edilen %95 güven aralıklarının örtüşmemesi, gözlenen performans farkının yalnızca katlar arası rastlantısal değişkenlikle açıklanmasının düşük olabileceğine işaret etmektedir.

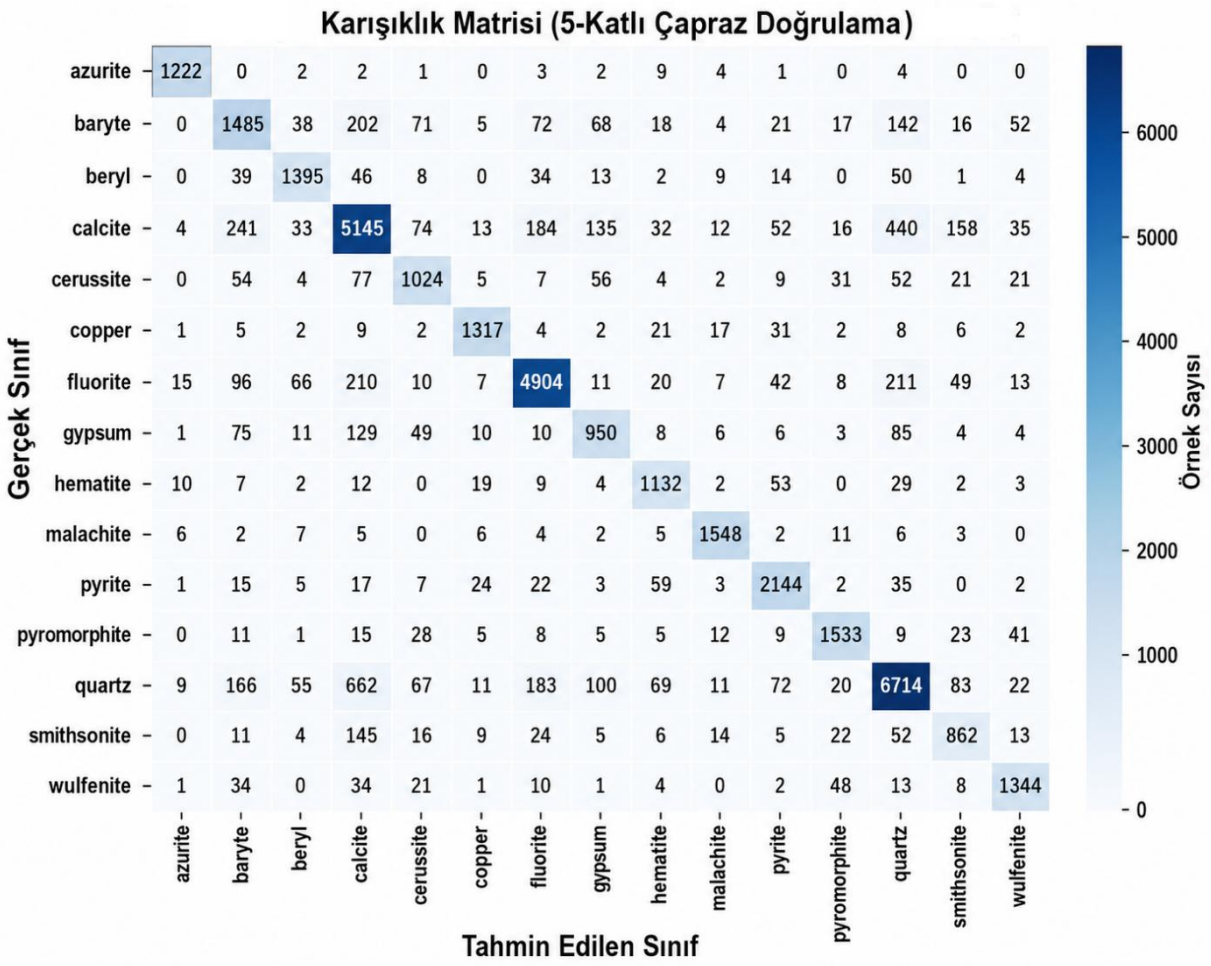
Daha yüksek performans gösteren EfficientNet-V2-S modelinin beş katlı çapraz doğrulama sürecinde elde edilen fold bazlı doğruluk ve doğrulama kaybı değerleri Tablo 5'te sunulmuştur.

Tablo 5. EfficientNet-V2-S modelinin 5-katlı çapraz doğrulama sonuçları

Çapraz Doğrulama Katı	**Doğruluk**	**Doğrulama Kaybı**
Fold 1	0.8275	0.6232
Fold 2	0.8333	0.6059
Fold 3	0.8301	0.6398
Fold 4	0.8294	0.6170
Fold 5	0.8367	0.6072
Ortalama ± SS	0.8314 ± 0.0033	0.6186 ± 0.0124

Tablo 5 incelendiğinde, EfficientNet-V2-S modelinin tüm çapraz doğrulama katlarında benzer performans sergilediği görülmektedir. Doğruluk değerlerinin 0.8275 ile 0.8367 arasında değişmesi ve standart sapmanın yalnızca ±0.0033 olması, modelin farklı veri bölünmelerinde kararlı ve tutarlı sonuçlar ürettiğini göstermektedir. Benzer şekilde doğrulama kaybındaki düşük değişim de eğitim sürecinin istikrarlı olduğunu desteklemektedir. Bu fold bazlı dağılımdan hesaplanan %95 güven aralığının 0.8273–0.8355 gibi dar bir aralıkta olması da EfficientNet-V2-S modelinin farklı veri bölünmelerine karşı düşük değişkenlik gösterdiğini ve elde edilen ortalama doğruluk değerinin güvenilirliğini desteklemektedir.

EfficientNet-V2-S modeli için, tüm fold'ların en iyi çıktıları birleştirilerek elde edilen karışıklık matrisi Şekil 5'te sunulmuştur.



Şekil 5. EfficientNet-V2-S modeli için 5-katlı çapraz doğrulama sürecinde elde edilen en iyi model çıktılarının birleştirilmesiyle oluşturulan karışıklık matrisi

Şekil 5'te verilen karışıklık matrisi incelendiğinde, tahminlerin büyük ölçüde ana köşegen üzerinde yoğunlaştığı ve modelin mineral sınıflarını yüksek doğrulukla sınıflandırdığı görülmektedir. Bununla birlikte, bazı sınıflar arasında belirgin karışıklıklar bulunmaktadır. Özellikle Calcite sınıfına ait görüntülerin en fazla Quartz, Baryte ve Fluorite sınıfları ile karıştırıldığı gözlenmiştir. Benzer şekilde, Quartz sınıfındaki hatalı tahminlerin önemli bir kısmının Calcite ve Fluorite sınıflarına yöneldiği görülmektedir. Buna karşılık Azurite, Malachite, Copper ve Wulfenite sınıflarında yanlış sınıflandırma sayısının oldukça düşük olması, modelin bu minerallere ait ayırt edici görsel özellikleri daha başarılı şekilde öğrendiğini göstermektedir. Calcite, Quartz, Baryte ve Fluorite sınıfları arasında gözlenen karışıklıkların, bu minerallerin benzer renk tonları, yüzey dokuları ve kristal yapıları nedeniyle ortak görsel özellikler paylaşmasından kaynaklandığı değerlendirilmektedir. Buna karşın, Azurite, Malachite, Copper ve Wulfenite minerallerinin daha belirgin renk ve doku özelliklerine sahip olması, modelin bu sınıfları daha yüksek doğrulukla ayırt edebilmesine katkı sağlamış olabilir.

Veri setindeki belirgin sınıf dengesizliği nedeniyle, genel doğruluk değerinin yanı sıra modelin her mineral sınıfındaki başarımının ayrı ayrı değerlendirilmesi önem taşımaktadır. Bu amaçla, EfficientNet-V2-S modelinin beş katlı çapraz doğrulama sürecinde elde edilen tahminleri birleştirilerek her sınıf için kesinlik, duyarlılık ve F1-skoru değerleri hesaplanmıştır. Sınıf bazlı performans sonuçları Tablo 6'da sunulmuştur.

Tablo 6. EfficientNet-V2-S modelinin sınıf bazlı performans sonuçları

Sınıf	Kesinlik	Duyarlılık	F1-skoru	Destek
Azurite	0.9622	0.9776	0.9698	1250
Baryte	0.6627	0.6716	0.6671	2211
Beryl	0.8585	0.8638	0.8611	1615
Calcite	0.7668	0.7826	0.7746	6574
Cerussite	0.7431	0.7491	0.7461	1367
Copper	0.9197	0.9216	0.9207	1429
Fluorite	0.8952	0.8651	0.8799	5669
Gypsum	0.6980	0.7032	0.7006	1351
Hematite	0.8121	0.8816	0.8454	1284
Malachite	0.9376	0.9633	0.9503	1607
Pyrite	0.8705	0.9166	0.8930	2339
Pyromorphite	0.8949	0.8970	0.8960	1705
Quartz	0.8553	0.8145	0.8344	8244
Smithsonite	0.6974	0.7256	0.7112	1188
Wulfenite	0.8638	0.8836	0.8736	1521

Tablo 6 incelendiğinde, EfficientNet-V2-S modelinin sınıflar arasında farklı performans düzeyleri sergilediği görülmektedir. En yüksek F1-skorları Azurite (0.9698), Malachite (0.9503) ve Copper (0.9207) sınıflarında elde edilmiştir. Buna karşılık Baryte (0.6671), Gypsum (0.7006) ve Smithsonite (0.7112) sınıflarında daha düşük F1-skorları gözlenmiştir. Özellikle Smithsonite sınıfının veri setinde en az sayıda örnekle temsil edilen sınıf olması, sınıf dengesizliğinin az temsil edilen sınıfların performansı üzerindeki etkisini göstermektedir. Bununla birlikte, yalnızca örnek sayısının performansı açıklamak için yeterli olmadığı da görülmektedir. Örneğin Baryte sınıfı Smithsonite'dan daha fazla örnek içermesine rağmen daha düşük F1-skoruna sahiptir. Bu durum, sınıf sıklığının yanı sıra mineraller arasındaki görsel benzerliğin de sınıflandırma başarımında önemli rol oynadığını göstermektedir. Özellikle Baryte, Calcite, Quartz ve Fluorite sınıfları arasında karışıklık matrisinde gözlenen hatalar bu bulguyu desteklemektedir.

Modelin sınıflandırma performansının görsel olarak daha ayrıntılı değerlendirilebilmesi amacıyla, doğru sınıflandırılan örnek görüntüler ile en sık karıştırılan sınıf çiftlerinden seçilen hatalı sınıflandırma örnekleri Şekil 6'da sunulmuştur. Her örnek için gerçek sınıf ve model tarafından tahmin edilen sınıf birlikte verilmiştir.



Şekil 6. EfficientNet-V2-S modeli tarafından doğru sınıflandırılan örnek görüntüler ile en sık karıştırılan sınıf çiftlerine ait hatalı sınıflandırma örnekleri. Her görüntü için gerçek sınıf ve model tarafından tahmin edilen sınıf gösterilmektedir

Şekil 6'da yer alan hatalı sınıflandırma örnekleri incelendiğinde, bazı görüntülerde karmaşık arka planlar, farklı çekim koşulları ve aydınlatma farklılıklarının bulunduğu görülmektedir. Bunun yanı sıra, bazı minerallerin benzer renk tonları, kristal morfolojileri ve yüzey dokuları da sınıflandırmayı güçleştirebilecek görsel özellikler arasında yer almaktadır. Bu tür farklılıkların modelin hata yaptığı örneklerde etkili olmuş olabileceği değerlendirilmektedir.

5. TARTIŞMA

Bu çalışmada, mineral görüntülerinin sınıflandırılması problemine yönelik olarak farklı derin öğrenme mimarilerinin performansları kapsamlı bir şekilde karşılaştırılmıştır. Elde edilen bulgular, model mimarisinin veri setinin yapısı ile olan uyumunun sınıflandırma performansı üzerinde belirleyici bir rol oynadığını açıkça ortaya koymaktadır.

Deneysel sonuçlar incelendiğinde, Mineral Photos veri seti ve bu çalışmada kullanılan eğitim protokolü altında, ResNet50, EfficientNet-B0, MobileNetV3-Large ve DenseNet121 modellerinin ViT-B/16 modeline kıyasla daha yüksek performans sergilediği görülmektedir. Özellikle EfficientNet-B0 modelinin %74.39 doğruluk oranı ile en yüksek performansa ulaşması, bu mimarinin mineral görüntülerindeki karmaşık dokusal ve yapısal özellikleri daha etkin bir şekilde temsil edebildiğini göstermektedir. EfficientNet'in bileşik ölçeklendirme

yaklaşımı sayesinde hem parametre verimliliği hem de temsil gücü açısından dengeli bir yapı sunması, bu çalışmada gözlenen yüksek performansın olası nedenlerinden biri olarak değerlendirilmektedir.

Benzer şekilde, ResNet50 ve DenseNet121 modellerinin de yüksek performans değerleri elde etmesi, derin CNN mimarilerinin özellik çıkarım kabiliyetlerinin bu problem için oldukça uygun olduğunu göstermektedir. ResNet mimarisindeki artık bağlantılar (residual connections), gradyan kaybolması problemini azaltarak daha derin ağların etkin şekilde eğitilmesini sağlarken DenseNet mimarisindeki yoğun bağlantı yapısı, özellik yeniden kullanımını artırarak daha zengin temsil öğrenimine olanak tanımaktadır.

Öte yandan, MobileNetV3-Large modelinin daha düşük parametre sayısına rağmen rekabetçi bir performans sergilemesi, hafif mimarilerin de uygun problem tanımı ve veri yapısı altında etkili olabileceğini göstermektedir. Bu durum, özellikle gerçek zamanlı uygulamalar ve düşük donanım kaynaklarına sahip sistemler için önemli bir avantaj sunmaktadır.

Çalışmada dikkat çeken en önemli bulgulardan biri, ViT-B/16 modelinin bu çalışmada daha düşük performans göstermesinde, veri setindeki sınıf dengesizliği, bazı sınıflardaki sınırlı örnek sayısı ve kullanılan eğitim stratejisinin etkili olmuş olabileceği değerlendirilmektedir. Transformer tabanlı modellerin veri gereksinimleri de bu sonuç üzerinde etkili olabilecek faktörlerden biri olabilir. Bununla birlikte, yalnızca tek bir transformer mimarisinin değerlendirilmiş olması nedeniyle bu bulgu tüm transformer tabanlı modeller için genellenmemelidir.

Veri setindeki sınıf dengesizliği de model performansları üzerinde önemli bir etkiye sahiptir. Özellikle Quartz gibi yüksek örnek sayısına sahip sınıfların model tarafından daha kolay öğrenildiği, buna karşın Smithsonite gibi düşük örnek sayısına sahip sınıflarda hata oranlarının arttığı gözlemlenmiştir. Bu durum, sınıf ağırlıklandırma ve veri artırma tekniklerinin önemini bir kez daha ortaya koymaktadır.

EfficientNet tabanlı modeller üzerinde gerçekleştirilen ileri düzey deneyler, model mimarisinin daha gelişmiş versiyonlarının performansı önemli ölçüde artırabildiğini göstermektedir. EfficientNet-V2-S modelinin %83.14 doğruluk oranı ile daha yüksek ve daha stabil bir performans sergilemesi, modern mimari iyileştirmelerin sınıflandırma başarısına doğrudan katkı sağladığını ortaya koymaktadır. Ayrıca düşük standart sapma değeri, modelin farklı veri alt kümeleri üzerinde tutarlı sonuçlar ürettiğini göstermektedir.

Eğitim süreci boyunca elde edilen kayıp ve doğruluk grafiklerinin analizi, EfficientNet-V2-S modelinin daha hızlı yakınsadığını ve daha stabil bir öğrenme süreci sergilediğini göstermektedir. Buna karşın EfficientNet-B0 modelinde doğrulama kaybındaki dalgalanmalar, modelin bazı veri alt kümelerinde aşırı öğrenme eğilimi gösterebildiğini ortaya koymaktadır. Bu durum, daha gelişmiş optimizasyon teknikleri ve düzenlileştirme (regularization) yöntemlerinin önemini vurgulamaktadır.

Şekil 5'te sunulan karışıklık matrisi analizi incelendiğinde, modelin bazı mineral sınıflarını ayırt etmekte zorlandığı görülmektedir. Özellikle görsel olarak benzer özellikler taşıyan sınıflar arasında (örneğin quartz–calcite ve quartz–fluorite) daha fazla hata yapıldığı dikkat çekmektedir. Bu durum, mineral görüntülerinin doğası gereği yüksek benzerlik içermesi ile ilişkilidir ve daha ayırt edici özellik çıkarım yöntemlerinin geliştirilmesi gerektiğini göstermektedir.

Genel olarak değerlendirildiğinde, bu çalışmanın bulguları, Mineral Photos veri seti ve kullanılan deneysel koşullar kapsamında, değerlendirilen CNN tabanlı modellerin güçlü sınıflandırma performansı sunduğunu göstermektedir. ViT-B/16 modelinin daha düşük performansı ise bu veri setinin yapısı, sınıf dengesizliği ve uygulanan eğitim koşullarıyla birlikte değerlendirilmelidir. Bu nedenle elde edilen sonuçlar, CNN ve transformer tabanlı mimariler arasında genel bir üstünlük ilişkisi olarak yorumlanmamalıdır. Gelecek çalışmalarda farklı transformer mimarileri, daha büyük ve dengeli veri setleri ve alternatif eğitim stratejileri kullanılarak daha kapsamlı karşılaştırmalar yapılabilir.

6. SONUÇLAR

Bu çalışmada, mineral görüntülerinin otomatik olarak sınıflandırılması amacıyla farklı derin öğrenme mimarileri kapsamlı bir şekilde analiz edilmiş ve karşılaştırılmıştır. Çalışma kapsamında beş farklı önceden eğitilmiş model kullanılarak, bu modellerin aynı veri seti ve eğitim koşulları altında performansları değerlendirilmiştir. Elde edilen sonuçlar, model mimarisinin veri yapısı ile uyumunun sınıflandırma başarısı üzerinde kritik bir etkiye sahip olduğunu ortaya koymuştur.

Deneysel bulgular, Mineral Photos veri seti ve bu çalışmada uygulanan eğitim koşulları altında, değerlendirilen CNN tabanlı modellerin ViT-B/16 modeline kıyasla daha yüksek performans elde ettiğini göstermiştir. Özellikle EfficientNet-B0 modeli ilk deneysel aşamada en yüksek doğruluk değerine ulaşmış, ResNet50 ve DenseNet121 modelleri de rekabetçi sonuçlar sergilemiştir. Buna karşılık ViT-B/16 modeli aynı deneysel koşullar altında daha düşük performans göstermiştir. Bununla birlikte, bu sonuç yalnızca bu çalışmada kullanılan veri seti, model seçimi ve eğitim protokolü kapsamında yorumlanmalıdır. Çalışmada yalnızca tek bir transformer tabanlı mimarinin değerlendirilmiş olması nedeniyle, elde edilen bulgular CNN ve transformer mimarileri arasında genel bir üstünlük ilişkisi olarak değerlendirilmemelidir.

Buna karşılık, ViT-B/16 modelinin bu çalışmada daha düşük performans göstermesi, kullanılan veri setinin büyüklüğü, sınıf dengesizliği ve uygulanan eğitim protokolü ile birlikte değerlendirilmelidir. Bu durum, model seçiminde yalnızca mimari yeniliklerin değil, aynı zamanda veri setinin büyüklüğü ve dağılımının da dikkate alınması gerektiğini ortaya koymaktadır.

EfficientNet tabanlı modeller üzerinde gerçekleştirilen ileri düzey deneyler, daha gelişmiş mimarilerin performansı önemli ölçüde artırabildiğini göstermiştir. EfficientNet-V2-S modelinin daha yüksek doğruluk ve daha düşük varyans ile daha stabil sonuçlar üretmesi, modern mimari iyileştirmelerin sınıflandırma performansına doğrudan katkı sağladığını göstermektedir. Bu bağlamda, yalnızca model seçiminin değil, aynı zamanda uygun eğitim stratejileri, veri artırma teknikleri ve hiperparametre optimizasyonunun da model başarısında belirleyici olduğu anlaşılmıştır.

Çalışma kapsamında elde edilen bulgular, mineral görüntülerinin sınıflandırılması gibi karmaşık görsel problemler için derin öğrenme tabanlı yaklaşımların etkili bir çözüm sunduğunu göstermektedir.

6.1. Çalışmanın Sınırlılıkları

Bu çalışmanın bazı sınırlılıkları bulunmaktadır. İlk olarak, deneysel değerlendirmeler yalnızca Mineral Photos veri seti üzerinde gerçekleştirilmiştir. Bu nedenle modelin farklı veri kümeleri

üzerindeki genellenebilirliği değerlendirilmemiştir. Ayrıca veri setindeki sınıflar arasındaki dengesiz örnek dağılımı ve bazı mineraller arasındaki görsel benzerlikler model performansını sınırlayan faktörler arasında yer almaktadır. Bunun yanı sıra, önerilen model bağımsız bir dış veri seti ve gerçek saha görüntüleri üzerinde değerlendirilmemiştir. Farklı veri kümeleri, bağımsız dış doğrulama ve gerçek saha görüntüleri üzerinde yapılacak değerlendirmeler, modelin genellenebilirliğinin daha kapsamlı biçimde ortaya konulmasına katkı sağlayabilir.

Çalışmada transformer tabanlı yaklaşımları temsilen yalnızca ViT-B/16 modeli kullanılmıştır; dolayısıyla CNN ve transformer mimari aileleri arasındaki performans karşılaştırmasına ilişkin bulgular diğer transformer modellerine doğrudan genellenmemelidir.

Ayrıca, model kararlarının daha iyi anlaşılabilmesi için açıklanabilir yapay zekâ (Explainable AI - XAI) yöntemlerinin entegrasyonu, özellikle kritik uygulamalar açısından önemli bir araştırma alanı olarak öne çıkmaktadır. Bunun yanı sıra, CNN ve transformer tabanlı yaklaşımların birleşiminden oluşan hibrit mimariler ile daha yüksek performans elde edilmesi mümkün olabilir.

KAYNAKLAR

[1] Tuğrul, A., & Yılmaz, M. (2023). Agrega Üretim Sahalarındaki Belirsizlik ve Risklerin Azaltılmasında Jeoloji ve Mühendislik Jeolojisi Araştırmaların Önemi. Jeoloji Mühendisliği Dergisi, 47(1), 87-102.

[2] Balaram, V. (2023). Advances in analytical techniques and applications in exploration, mining, extraction, and metallurgical studies of rare earth elements. Minerals, 13(8), 1031.

[3] Balaram, V., & Sawant, S. S. (2022). Indicator minerals, pathfinder elements, and portable analytical instruments in mineral exploration studies. Minerals, 12(4), 394.

[4] Long, T., Zhou, Z., Hancke, G., Bai, Y., & Gao, Q. (2022). A review of artificial intelligence technologies in mineral identification: classification and visualization. Journal of Sensor and Actuator Networks, 11(3), 50.

[5] Ling, Q., & Isa, N. A. M. (2023). Printed circuit board defect detection methods based on image processing, machine learning and deep learning: A survey. IEEE access, 11, 15921-15944.

[6] Taye, M. M. (2023). Theoretical understanding of convolutional neural network: Concepts, architectures, applications, future directions. Computation, 11(3), 52.

[7] Wang, Y., Deng, Y., Zheng, Y., Chattopadhyay, P., & Wang, L. (2025). Vision transformers for image classification: A comparative survey. Technologies, 13(1), 32.

[8] Tsangaratos, P., Ilia, I., Spanoudakis, N., Karageorgiou, G., & Perraki, M. (2025). Machine learning approaches for real-time mineral classification and educational applications. Applied Sciences, 15(4), 1871.

[9] Chawshin, K., Berg, C. F., Varagnolo, D., & Lopez, O. (2021). Lithology classification of whole core CT scans using convolutional neural networks. SN Applied Sciences, 3(6), 668.

[10] Xu, Z., Ma, W., Lin, P., & Hua, Y. (2022). Deep learning of rock microscopic images for intelligent lithology identification: Neural network comparison and selection. Journal of Rock Mechanics and Geotechnical Engineering, 14(4), 1140-1152.

[11] Zhou, W., Wang, H., & Wan, Z. (2022). Ore image classification based on improved CNN. Computers and Electrical Engineering, 99, 107819.

[12] Zhou, Z., Yuan, H., & Cai, X. (2023). Rock thin section image identification based on convolutional neural networks of adaptive and second-order pooling methods. Mathematics, 11(5), 1245.
[13] Rizaldy, A., Afifi, A. J., Ghamisi, P., & Gloaguen, R. (2024). Improving mineral classification using multimodal hyperspectral point cloud data and multi-stream neural network. Remote Sensing, 16(13), 2336.
[14] Li, S., Dong, Y., Zhang, Z., Lin, C., Liu, H., Wang, Y., ... & Zhang, G. (2024). Study on automatic lithology identification based on convolutional neural network and deep transfer learning. Discover Applied Sciences, 6(6), 322.
[15] Ma, W., Han, T., Xu, Z., & Lin, P. (2025). Feature fusion of single and orthogonal polarized rock images for intelligent lithology identification. AI in Civil Engineering, 4(1), 5.
[16] Wang, M., Guo, W., Yang, F., Yan, B., Xu, Y., Jiang, J., & Huang, J. (2025). Rock thin section image classification in low data scenarios using few-shot learning. Computers & Geosciences, 203, 105962.
[17] Geillon, F. Mineral Photos Veri Seti. Kaggle. Erişim adresi: https://www.kaggle.com/datasets/floriangeillon/mineral-photos
[18] Zhang, L., Bian, Y., Jiang, P., & Zhang, F. (2023). A transfer residual neural network based on ResNet-50 for detection of steel surface defects. Applied Sciences, 13(9), 5260.
[19] Wen, L., Li, X., & Gao, L. (2020). A transfer convolutional neural network for fault diagnosis based on ResNet-50. Neural Computing and Applications, 32(10), 6111-6124.
[20] Ali, R. R., Yaacob, N. M., Alqaryouti, M. H., Sadeq, A. E., Doheir, M., Iqtait, M., ... & Yaacob, S. S. (2025). Learning architecture for brain tumor classification based on deep convolutional neural network: Classic and ResNet50. Diagnostics, 15(5), 624.
[21] Padhi, J., Korada, L., Dash, A., Sethy, P. K., Behera, S. K., & Nanthaamornphong, A. (2024). Paddy leaf disease classification using EfficientNet B4 with compound scaling and swish activation: A deep learning approach. IEEE Access, 12, 126426-126437.
[22] Ali, H., Shifa, N., Benlamri, R., Farooque, A. A., & Yaqub, R. (2025). A fine tuned EfficientNet-B0 convolutional neural network for accurate and efficient classification of apple leaf diseases. Scientific Reports, 15(1), 25732.
[23] Zhang, J., Yang, X., Fu, X., Wang, B., & Li, H. (2025). LDL-MobileNetV3S: an enhanced lightweight MobileNetV3-small model for potato leaf disease diagnosis through multi-module fusion. Frontiers in Plant Science, 16, 1656731.
[24] Duan, Z., & Liu, H. (2025). Improving Network Structure for Efficient Classification Network Based on MobileNetV3. IEEE Access, 13, 191296-191308.
[25] Yehulu, G. T., Walle, Y. M., Haile, M. B., Yohannes, S. T., Tegegne, A. G., Tesfaye, D. G., & Baye, G. A. (2025). Mango fruit disease detection and classification using MobileNetV3_large model. Scientific African, e03061.
[26] Albelwi, S. A. (2022). Deep architecture based on DenseNet-121 model for weather image recognition. International Journal of Advanced Computer Science and Applications, 13(10).
[27] Prakash, N. N., Rajesh, V., Namakhwa, D. L., Pande, S. D., & Ahammad, S. H. (2023). A DenseNet CNN-based liver lesion prediction and classification for future medical diagnosis. Scientific African, 20, e01629.
[28] Hong, S., Wu, J., Zhu, L., & Chen, W. (2024). Brain tumor classification in VIT-B/16 based on relative position encoding and residual MLP. Plos one, 19(7), e0298102.

[29] Appati, J. K., Ziamah, B., Akrofi, H. A., & Dodoo, A. A. (2025). SARS detection in chest CT scan images using the bootstrapped ViT-B/16 model. Iran Journal of Computer Science, 8(2), 537-551.
[30] Kumar, P., Mathew, J., & Sanodiya, R. K. (2025). Optimized loss and self attention for enhanced domain adaptation in remote sensing image classification. Scientific Reports, 15(1), 40259.
[31] Ganesh, P., Chen, Y., Lou, X., Khan, M. A., Yang, Y., Sajjad, H., ... & Winslett, M. (2021). Compressing large-scale transformer-based models: A case study on bert. Transactions of the Association for Computational Linguistics, 9, 1061-1080.
[32] Vujović, Ž. (2021). Classification model evaluation metrics. International Journal of Advanced Computer Science and Applications, 12(6), 599-606.
[33] Cabot, J. H., & Ross, E. G. (2023). Evaluating prediction model performance. Surgery, 174(3), 723-726.
[34] Foody, G. M. (2023). Challenges in the real world use of classification accuracy metrics: From recall and precision to the Matthews correlation coefficient. Plos one, 18(10), e0291908.
[35] Obi, J. C. (2023). A comparative study of several classification metrics and their performances on data. World Journal of Advanced Engineering Technology and Sciences, 8(1), 308-314.