

Kuş ve Drone Resimlerini Deep Learning Kullanarak İkili Sınıflandırma

Kadir Kemal Temel
170114021

Mekatronik Mühendisliği, Bursa Teknik Üniversitesi

June 4, 2025

Özet:

Veri Madenciliğine Giriş dersi projesi olarak Data in Brief dergisinde yayımlanan "YOLO-based segmented dataset for drone vs. bird detection for deep and machine learning algorithms" isimli makalenin veri seti deep learning teknikleri ve modelleri kullanılarak ikili sınıflandırma yapılmak için seçilmiştir. Transfer learning, sıfırdan eğitilen modeller (training from scratch) ve özgün bir mimari ile yapılan sınıflandırma testleri gerçekleştirilmiş ve performansları karşılaştırılmıştır.

Anahtar Kelimeler: Deep Learning, Transfer Learning, Classification, Veri Madenciliği

1 Giriş

Veri Madenciliği, klasik istatistik ve Makine Öğrenimi gibi bilgisayar bilimi algoritmalarının birleşiminden ortaya çıkan ve amacı büyük miktarda veriden yeni ve faydalı bilgi çıkarmak olan bir disiplindir [1]. Veri madenciliği kullanılacak verilere göre farklı yöntemler sunmaktadır. Bunlar:

- Sınıflandırma (Classification):** Eğitim veri setlerinden sınıf tahminleri yapan bir gözetimli öğrenme (supervised learning) tekniğidir. Bayesian ağları (bayesian networks), karar ağaçları (decision trees), destek vektör makineleri (support vector machines), sinir ağları (neural networks) ve derin öğrenme (deep learning) örnek olarak verilebilir.
- Kümeleme (Clustering):** Sınıflandırmanın tersi olarak gözetimsiz öğrenme tekniği ile bilinmeyen sınıfların olduğu bir veri setini, veriler arasındaki benzerlikleri kullanarak çözer.
- İlişkilendirme Araştırması (Research of Associations):** Veri setindeki objeler arasındaki anlamlı ilişkileri keşfeder.
- Tahmin (Prediction):** Verilerin geleceği hakkında makul tahminlere yol açabilecek veri kalıplarının keşfedilmesinden oluşur.

Bu proje kapsamında sınıflandırma yöntemi için derin öğrenme tekniği kullanılmıştır. Derin öğrenme, verilerdeki üst düzey çıkarımları modellemek için algoritmalar kullanan bir makine öğrenimi alt alanıdır. Genellikle gizli katmanlara sahip yapay sinir ağlarından oluşur [2].

Proje sonunda oluşturulan ve kullanılan modeller kesinlik (precision), geri çağırma (recall), F1 skoru, karışıklık (confusion) matrisi, kayıplar ve doğrulukları karşılaştırılmıştır.

- Kesinlik:** Doğru tahmin edilen pozitif verilerin tahmin edilen pozitif verilere oranıdır.
- Recall:** Duyarlılık olarak da bilinen bu terim doğru tahmin edilen pozitif verilerin gerçek sınıftaki tüm verilere oranıdır.
- F1 Skoru:** Kesinlik ve recall değerlerinin ağırlıklı ortalamasıdır. Bu puan hem yanlış pozitifleri hem de yanlış negatifleri kullanılarak hesaplanır.

- Confusion Matrisi:** Doğru pozitif, doğru negatif, yanlış pozitif ve yanlış negatif sayılarını gösteren tablo biçimli bir sonuç üretir.
- Kayıp Değeri:** Hata oranı.
- Doğruluk:** Eğitimde kullanılan modelin test verisindeki doğruluk oranı.

2 Yöntem

2.1 Eğitilmiş Modeller

Transfer öğrenimi (transfer learning), belirli bir veri koleksiyonu için geliştirilen önceden eğitilmiş modellerin, farklı veri setleri kullanılarak yeni bir model oluşturmak olarak adlandırılmaktadır [3]. Proje uygulama aşamasında önceden eğitilip farklı alanlarda kullanılan modeller hem transfer öğrenimi, hem de sıfırdan öğrenim ile test edilip sonuçları gözlemlenmiştir. Kullanılan görüntü sınıflandırma modelleri ise şunlardır:

- ResNet18:** 18 gizli katmanla oluşturulmuş, bazı katmanlar arası kısa yolları bulunan bir sinir ağı modeli [4].
- ResNet50:** ResNet34 modelinin her ikili katmanlarının üçer katmanlı hale getirilmesiyle oluşturulmuş sinir ağı modelidir [4].
- EfficientNet-B0:** 18 gizli katmanlı, yüksek doğruluk ve düşük hesaplama maliyeti arasında dengeli bir performans sağlamak amacıyla geliştirilmiş bir modeldir [5].
- Mobilenet-V2:** Mobil ve gömülü cihazlar için geliştirilmiş, hafif ve verimli mimariye sahip bir modeldir [6].
- Vgg-16:** 3x3 filtrelerden oluşan basit fakat güçlü olabilmesi amacıyla geliştirilmiş ve genellikle transfer öğrenimi için kullanılan bir modeldir [7].

2.2 Oluşturulan Model

Proje kapsamında oluşturulan model, iki sınıf arasında (kuş ve drone) görüntü tabanlı sınıflandırma yapmak için üretilmiştir. PyTorch kütüphaneleri yardımıyla oluşturulmuş, derin öğrenmeye dayalı bir **Convolutional Yapay Ağ (CNN)**'dir.

Giriş olarak 3 renk kanallı 224x224 boyutunda resim al-

maktadır. Çıkışında ise 2 boyutlu tensör, softmax uygulanarak iki sınıfa ayrılır. 5 gizli katmana sahip olan modelde konvolüsyon yapılırken 3x3 boyutlu bir filtre uygulanmaktadır. Ayrıca ezber yapmasına karşın %50 oranda dropout uygulanmıştır.

3 Sonuç

Kullanılan farklı modellerin eğitimi 16 batch ve 10 epoch değerleriyle yapılmıştır.

3.1 Eğitim Aşaması

Modeller eğitildikten sonra doğruluk ölçüleri elde edilip karşılaştırılmıştır. Figür 1 incelendiğinde en yüksek eğitim kaybının bile %2.5 değerinin altında olduğu görülmekte. 12 modelden ise 7'sinin neredeyse kayıpsız bir eğitim süreci geçirmiş olduğu tercih edilemeyecek bir durumdur. Aynı modellerin doğrulama aşamasında gösterdiği performans ise Figür 2 'de gösterilmiştir.

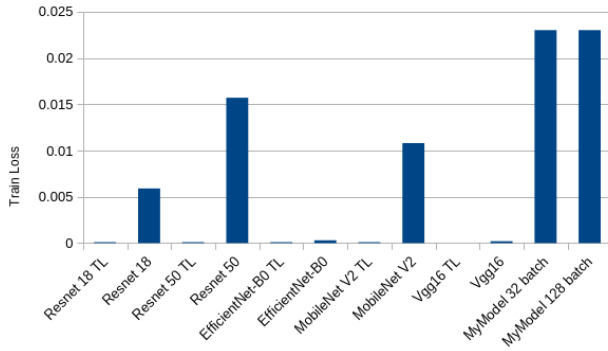


Figure 1. Eğitim Kaybı Grafiği

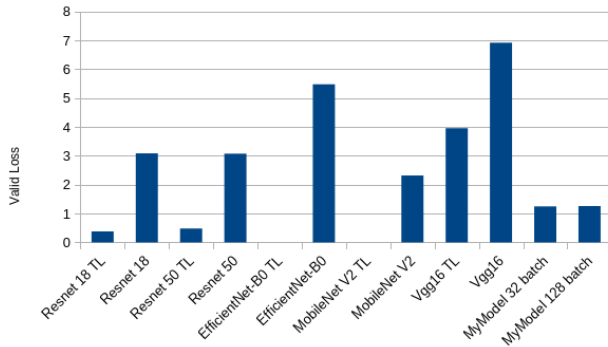


Figure 2. Eğitim Validasyon Grafiği

Eğitim kayıplarının çok düşük olduğu bir senaryoda validasyon kayıpları incelendiğinde sadece Efficientnet-B0 modeli ve MobileNet V2 modelinin eğitim kayıplarıyla tutarlılık gösterdiği görülmektedir. Kesin bir çıkarım yapabilmek Figür 3'de görülen eğitim doğruluğu grafiği incelenmelidir.

Grafik incelendiğinde eğitim doğrulukları en az %50 oranında başarı sağladığı görülmekte. Bu verilerle şimdiden çoğu modelin öğrenmek yerine ezber yaptığı söylenebilir.

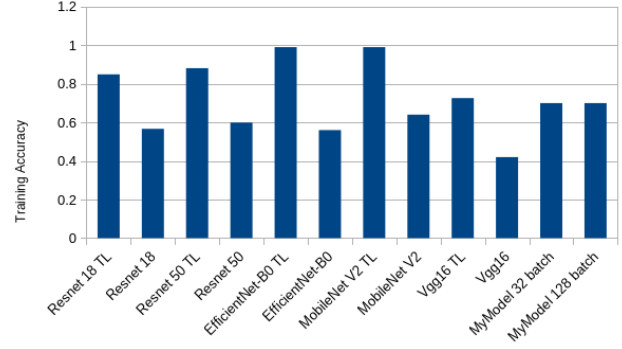


Figure 3. Eğitim Doğruluğu Grafiği

3.2 Test

Modeller eğitim aşamasından test aşamasına geçtiğinde başarılı bir şekilde eğitilen modeller göz önüne çıkmaktadır. Figür 4 incelendiğinde test aşamasındaki kayıp oranları sadece 4 model için kabuledilebilir durumdadır. Figür 5'te ise kayıpları az olan Resnet 18 transfer öğrenim, Resnet 50 transfer öğrenim, EfficientNet-B0 transfer öğrenim ve MobileNet V2 transfer öğrenimli modellerin başarılı eğitim sürecinden geçtiği gözlemlenmektedir. Tarafımdan oluşturulmuş modelin ise doğruluğu %70 oranında kaldığı ve 128 batch büyüklüğü ile yapılan eğitim kaybının daha düşük olduğu ortaya çıkmıştır.

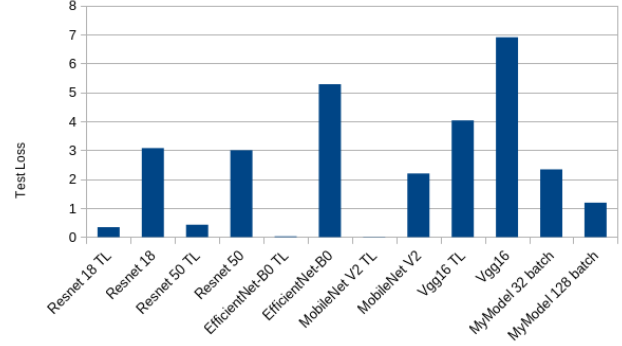


Figure 4. Test Kaybı Grafiği

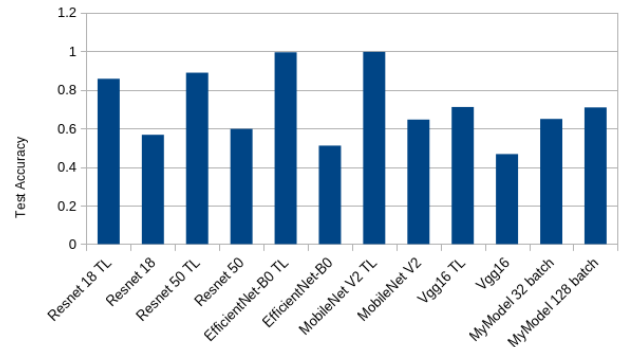


Figure 5. Test Doğruluğu Grafiği

Modellerin Figür 6'da recall, Figür 7'de kesinlik ve Figür 8'de F1 skorları incelendiğinde en başarılı olan modellerin daha önceden tespit edilmiş modeller olduğu değişmemekte ve özellikle EfficientNet-B0 ve MobileNetV2 modellerinin transfer öğrenim yapılmış halleri %100 doğruluğa çok yakın bir performans gösterdiği ve overfitting ihtimali bulunduğu anlaşılmıştır. Tarafımdan oluşturulmuş modelin ortalama bir performans gösterdiği ve Resnet modellerinin kullanım açısından daha güvenli olduğu ortaya çıkmıştır.

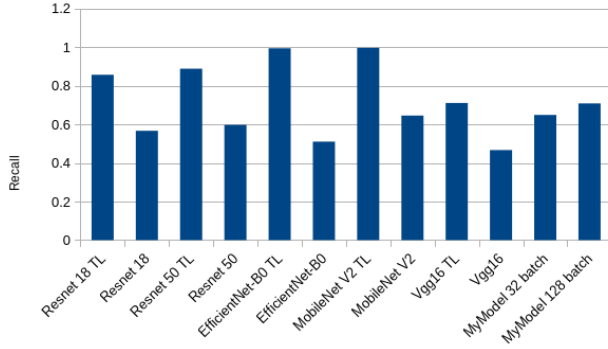


Figure 6. Recall Grafiği

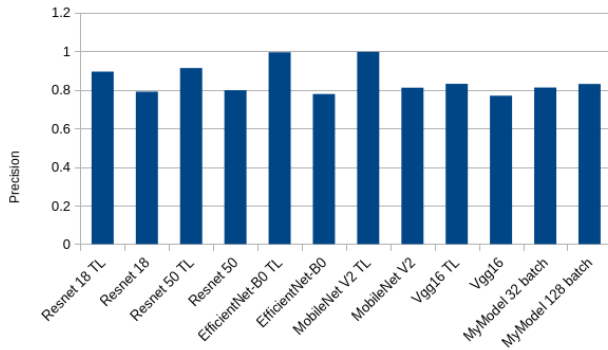


Figure 7. Kesinlik Ölçüsü Grafiği

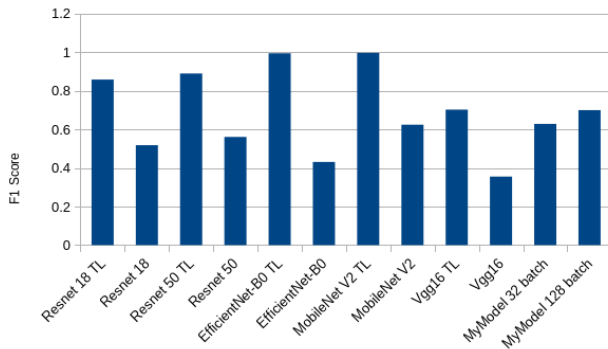


Figure 8. F1 Skoru Grafiği

Son olarak modellerin confusion matrisleri incelendiğinde, Resnet 18 TL modeli hiçbir kuşu drone olarak tanımazken 889 resimlik bir test verisinde 127 drone görselini kuş olarak

tanımıştır. Resnet 18 modelinde ise bu sayı 385'e yükselmiştir.

Resnet 50 modellerini incelediğimizde yine kuşlar drone olarak tanınmamış fakat normal eğitilen model 358 dronu kuş olarak görürken transfer öğrenim yapılan modelde bu sayı 99'dur.

EfficientNet B0 modelinin test sonucunda 435 dronu kuş olarak etiketlediği görülürken transfer öğrenimli modelinde bu sayı yalnızca 5'dir.

MobileNet V2 modeli diğer modellerle ortalama şekilde bir sonuç verip 315 dronu kuş olarak tanıırken transfer öğrenimi gerçekleştirildiğinde EfficientNet B0 modelinden daha iyi performans verip sadece 3 drone resmini kuş olarak etiketlemiştir.

Vgg16 modelini incelediğimizde 474 drone resmini kuş olarak etiketlemişken transfer öğrenim sonucu bu sayı 257 gibi ortalama bir sayıya düşmüştür.

Tarafımdan oluşturulan model 128 batch sayısı ile eğitim gördüğünde 259 drone, 64 batch ile 305, 32 batch ile ise 312 yanlış etiketleme yapmıştır.

Modellerin hiç birisi kuşları yanlış etiketlemezken bütün modellerde drone fotoğrafları yanlış etiketlenmektedir.

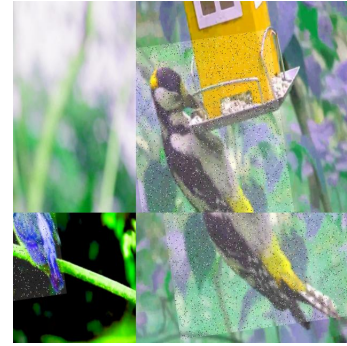


Figure 9. Eğitim İçin Hazırlanmış Kuş Resmi

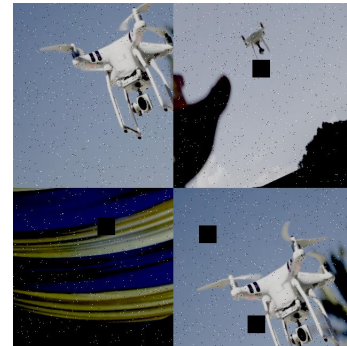


Figure 10. Eğitim İçin Hazırlanmış Drone Resmi

Eğitim veri seti incelendiğinde Figür 9 ve Figür 10 da görüleceği üzere YOLO adı verilen bir nesne tanıma algoritması için renk düzenlemesi, gürültü eklemesi, nesne kırılması gibi etkiler yapılmış olduğu görülmektedir. Bu eğitimlerin sonucundan da anlaşılabacağı üzere eğitim seti sınıflama yapan derin öğrenme modelleri için bu haliyle uygun değildir. YOLO tarzı bir algoritmayla ve segmentasyon

yardımıyla modellerin daha yüksek doğruluk payı olacağı aşıkardır.

Acknowledgements

This research received support during the XXX course, instructed by Professor XXX, PhD at the School of XXX, XXX.

References

- [1] C. Zucco, “Data mining in bioinformatics,” pp. 399–408, 2025.
- [2] H. A. Issad, R. Aoudjit, and J. J. Rodrigues, “A comprehensive review of data mining techniques in smart agriculture,” *Engineering in Agriculture, Environment and Food*, vol. 12, no. 4, pp. 511–525, 2019.
- [3] J. Gupta, S. Pathak, and G. Kumar, “Deep learning (cnn) and transfer learning: A review,” *Journal of Physics: Conference Series*, vol. 2273, p. 012029, may 2022.
- [4] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” 2015.
- [5] M. Tan and Q. V. Le, “Efficientnet: Rethinking model scaling for convolutional neural networks,” 2020.
- [6] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, “Mobilenetv2: Inverted residuals and linear bottlenecks,” 2019.
- [7] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” 2015.