

T.C.
GEBZE TEKNİK ÜNİVERSİTESİ
BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

POLEN GÖRÜNTÜLERİNİN TESPİTİ VE
SINIFLANDIRILMASI

FATİH KAAAN SALGIR

DANIŞMAN
PROF. DR. ERCHAN APTOULA

2022
GEBZE

2 Kullanılan Yöntem ve Elde Edilen Bulgular

2.1 Sınıflandırma

Yapılan ilk deneyler, modelin daha çok örneklenen sınıflara yanlı (biased) olduğunu göstermiştir. Bu problemini çözmek için sınıflardaki örnek sayısına göre ağırlıkları hesaplanmıştır. [1] Sınıf ağırlıkları, kayıp (loss) aralığını değiştireceği için, bu optimizasyon algoritmasına bağlı olarak eğitimin stabilitesini etkileyebilir. Deneylerde tercih edilen optimizasyon algoritması, ölçeklendirme değişikliğinden etkilenmeyecek şekilde seçilmiştir.

Farklı yöntemleri ve mimarileri kıyaslayabilmek için temel bir referans model geliştirilmiştir. Bu CNN modeli, her birinde bir pooling katmanına sahip 3 evrişim bloğundan ve ReLU fonksiyonu tarafından etkinleştirilen 128 birimli fully connected katmandan oluşur. Modelin eğitiminde ve sınanmasında, görüntüler ağı verilmeden önce, sabit en boy oranını koruyarak, 400'e 300 piksele yeniden ölçeklendirilmiştir. Model, 10^{-3} öğrenme oranına sahip Adam optimizasyon algoritması ile 200 epoch eğitilmiştir. Test sonuçlarına göre model yaklaşık %74 sınıflandırma doğruluğuna ulaşmıştır.

Öğrenme aktarımında (transfer learning) temel model olarak, ImageNet üzerinde önceden eğitilmiş olan MobileNet V2 kullanıldı. Evrişim bloğu dondurulup sınıflandırma bloğu eklendiğinde, temel modelde 2,3 milyon parametre ve dense katmanda 95 bin eğitilebilir parametre bulunmaktadır. 10^{-4} öğrenme oranı ve Adam optimizasyon algoritması ile 150 epoch eğitimden sonra model doğrulama (validation) kümesinde %72 başarıma ulaşmıştır. Performansı daha da artırmak ve ağırlıklara güncellemek için önceden eğitilmiş modelin ilk 56 katmanı çözülürken ve sınıflandırıcı ile birlikte 150 epoch daha eğitilmiştir. Bu adımında, aşırı öğrenmeyi (overfit) önlemek için öğrenme oranı 10 kat azaltılmış RMSprop optimizasyon algoritması kullanılmıştır. Modelin test kümesinde %90 sınıflandırma doğruluğuna ulaştığı gözlemlenmiştir.

2.2 Nesne Algılama

Nesne tespitinde amaç, tespit edilen nesne tanelerini çevreleyen sınırlayıcı kutular çizabilmektir. Böyle bir modeli sıfırdan eğitmek uzun zaman alacak ve çok büyük miktarda girdi verisi gerektirecektir. Bunun yerine deneyler, transfer öğrenme yöntemi ile Tensorflow Object Detection API kullanarak gerçekleştirilmiştir.

Türlere ve türlerin ait oldukları ailelere göre iki farklı nesne algılama görevi vardır. Tür düzeyinde nesne tespiti için COCO 2017 veri setinde önceden eğitilmiş model mimarileri kullanılmıştır:

- SSD ResNet50 V1 FPN 640x640 (RetinaNet50)
- Faster R-CNN ResNet50 V1 1024x1024

- EfficientDet D1 640x640

Aile düzeyi nesne tespitinde ise eğitimler EfficientDet D1 ve SSD ResNet50 kullanılarak yapılmıştır.

İlk model (SSD ResNet50 V1 FPN), referans sonuç elde etmek için sadece gerekli parametreler kullanılarak eğitime tabi tutulmuştur. Diğer modellerin konfigürasyonların başarımlar üzerindeki etkisini gösterebilmek için bu modelin sınıflandırma skorlarına da Tablo 1’de yer verilmiştir. Daha sonra bu modelin eğitimi parametreler güncellenerek tekrarlanmıştır.

Ön işlemede, aynalama, hue, kontrast, doygunluk ve parlaklık ayarlama yöntemleri kullanılmıştır. Aşırı öğrenmeyi daha iyi önlemek için, SSD modellerinde zaman içindeki öğrenme oranını kontrol eden cosine decay öğrenme oranı kullanılmıştır.

Referans model elde edilen sonuçlara göre, modelin nesnelerin konumlarının tespitinde iyi olduğu, ancak sınıflandırmada oldukça düşük performans verdiği gözlemlenmiştir. Daha iyi sınıflandırma elde etmek için sonraki modellerde sınıflandırma kaybının (loss) ağırlığı, konumlandırmaya kıyasla dereceli olarak arttırılmıştır.

EfficientDet D1 konfigürasyonunda, sınıflandırma loss fonksiyonu olarak focal loss kullanılmıştır. Bu loss fonksiyonu iyi sınıflandırılmış örneklerin ağırlığını düşürürerek, zor örneklerle odaklanır. RetinaNet’e göre, focal loss, veri kümesi dengesiz olduğunda sınıflandırma için oldukça yararlıdır. [2]

Model doğruluğunu iyileştirmek ve eğitim süresini azaltabilmek için, sınırlayıcı kutuların oluşturulması da konfigüre edilmiştir. Burada sınırlayıcı kutuların en boy oranları tanımlanabilir. Polen görüntüleri genellikle dairesel olduğu için sınırlayıcı kutuları kareye yakın olacaktır. Bu nedenle beklenen en boy oranı 1 olarak ayarlanmıştır.

Tablo 1’de, test kümesindeki sınıflandırma sonucu gösterilmiştir. Bu değerler sınıf dağılımlarının ağırlıklı ortalamaları baz alınarak hesaplanmıştır.

Tablo 1: Nesne algılama modellerinin sınıflandırma skorları

	Accuracy	Precision	Recall	F1 Score
SSD ResNet50 V1 FPN (Referans)	0.46	0.39	0.46	0.38
Faster R-CNN ResNet50	0.63	0.65	0.64	0.59
EfficientDet D1	0.82	0.84	0.82	0.80
SSD ResNet50 V1 FPN	0.86	0.87	0.86	0.85
SSD ResNet50 V1 FPN (Aile)	0.62	0.63	0.62	0.61
EfficientDet D1 (Aile)	0.88	0.87	0.88	0.87

3 Sonuç

Bu projede kapsamında polen görüntülerinin sınıflandırılması ve algılanması üzerine yapılan deneylerde, CNN'lerin eğitilmesi açısından sıfırdan öğrenme ve öğrenme aktarımı, çeşitli veri artırım teknikleri ile beraber denenmiştir. Sınıflandırmada sıfırdan eğitim yaklaşık %74 doğruluk sağlarken, MobileNetV2 ile öğrenim aktarımı %90 başarı oranına ulaşmıştır.

Nesne algılamada öğrenim aktarımı yöntemleri ile üç farklı mimari test edilmiştir. Bu CNN mimarilerini göz önüne alındığında tür düzeyi sınıflandırmada SSD ResNet50 ile %86 başarı oranı ile en iyi sonuç elde edilmiştir. Aile düzeyinde sınıflandırmada ise EfficientDet D1, 88% başarı oranı ile en iyi sonucu verdiği gözlenmiştir.

Kaynaklar

- [1] G. King and L. Zeng, “Logistic regression in rare events data,” *Political Analysis*, vol. 9, pp. 137–163, 2001.
- [2] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, “Focal loss for dense object detection,” in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2980–2988.
- [3] S. Battiato, A. Ortis, F. Trenta, L. Ascari, M. Politi, and C. Siniscalco, “Detection and classification of pollen grain microscope images,” in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2020, pp. 4220–4227. doi: 10.1109/CVPRW50498.2020.00498.
- [4] H. Menad, F. Ben-Naoum, and A. Amine, “Deep convolutional neural network for pollen grains classification,” in *JERI*, 2019.
- [5] A. R. d. Geus, C. A. Barcelos, M. A. Batista, and S. F. d. Silva, “Large-scale pollen recognition with deep learning,” in *2019 27th European Signal Processing Conference (EUSIPCO)*, 2019, pp. 1–5. doi: 10.23919/EUSIPCO.2019.8902735.
- [6] H. Yin, Y. Chen, J. Xiong, R. Xia, J. Xie, and K. Yang, “An improved local binary pattern method for pollen image classification and recognition,” *Computers Electrical Engineering*, vol. 90, p. 106983, 2021, issn: 0045-7906. doi: <https://doi.org/10.1016/j.compeleceng.2021.106983>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0045790621000161>.
- [7] A. B. Gonçalves, J. S. Souza, G. G. da Silva, *et al.*, “Polen23e: Image dataset for the brazilian savannah pollen types,” [Online]. Available: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0157044>.