https://doi.org/10.56038/oprd.v2i1.252

Conference Article

Endemic Plant Classification Using Deep Neural Networks

Melih Öz[1*], Alper Özcan[2*]

[1] Akdeniz University, Turkey, Orcid ID: 0000-0003-2157-1788, melihoz@akdeniz.edu.tr
[2] Akdeniz University, Turkey, Orcid ID: 0000-0002-5999-1203, alperozcan@akdeniz.edu.tr
* Correspondence: melihoz@akdeniz.edu.tr

(First received January 02, 2023 and in final form March 12, 2023)

3rd International Conference on Access to Recent Advances in Engineering and Digitalization
March 15 - 17, 2023

Reference: Hybrid Beamforming for Multi User Massive MIMO Systems. Orclever Proceedings of Research and Development,2(1), 59-67

Abstract

Endemic plants are those that are native to a specific geographic region and are found nowhere else in the world. These plants are crucial for biodiversity, conservation, cultural significance, and economic value. Turkey hosts more than 4000 endemic plants. Therefore, this makes Turkey the richest in Europe. Preserving this habitat holds importance. This study aims to conceptualize a possible application that helps individuals to identify endemic species using camera-captured images. Thus, aiding the preservation of the habitat. In this study, 23 selected species of Turkey's endemic biodiversity are classified using Deep Neural Network built. In line with the objective of this study, a dataset containing 253 images is created to train the network. The dataset is available at: github.com/melihoz/endemicdataset

Keywords: Endemic, Plant, Classification, Deep Neural Networks

1. Introduction

Turkey is covered by Caucasus, Irano-Anatolian, and Mediterranean biodiversity hotspots [1]. Thus, with over 30% endemics out of nearly 12000 natural vascular plant taxa, Turkey is one of the most important biodiversity centers in the world [2].

Endemic species, being limited to a particular geographic region, are comparatively more susceptible to both human-made and natural threats, such as having a restricted range, existing in small populations, limited distribution, dwindling population numbers, excessive exploitation by humans, short reproductive cycles, specific habitat requirements, and the need for stable and consistent environments. The extent of these factors exhibited by an endemic species determines its vulnerability to extinction. Thus, the conservation and monitoring of endemic species are crucial at a global level [3]–[5].

Deep Neural Networks (DNNs) have become a popular tool in computer vision applications because of their ability to tackle complex tasks. They operate by utilizing input and output data and learning a customized mapping function between them through a process known as training [6]. This process involves updating the filter kernels for each input and output. Depending on the task, a DNN may require a large number of training data points, up to millions in some cases [7]. The introduction of the Imagenet dataset in 2010, which consists of 1,461,406 images belonging to 1000 object classes, significantly boosted the number of studies on DNNs [8].

Image classification refers to the process of assigning a label or category to an image by utilizing machine learning algorithms that can recognize and analyse the patterns and features present in the visual content of the image [9], [10]. The primary objective of image classification is to enable computers to interpret and comprehend visual information as humans do, allowing them to automatically analyse and classify a vast number of images accurately. The applications of image classification are diverse, ranging from identifying objects in photographs to detecting illnesses in medical images [11]–[14].

Within the scope of this study, 23 endemic species from Turkey are chosen, regarding available images found online to construct train test split [15]. Those species are shown in Table 1 and sample images are shown in Figure 1. Classifying endemic plants using images can help to build systems that display specific plants' risk of extinction. Thus, individuals can take preventive actions. Therefore, this study aims to instantiate a concept that can be used to protect endemic plants by preventing over-collection by humans. Classifying endemic plants using images can help to build systems that display specific plants' risk of extinction.

Figure 1: Sample Images from the Dataset

2. Materials and Methods

In this section, the materials and methods used to conduct experiments are introduced.

2.1. Dataset

Within the course of this study a dataset is created. The dataset contains 23 species with total of 253 images. The images are collected manually using google search engine and several web pages [16]–[18]. summary of the dataset is given at Table 1.

Table 1: Dataset Properties

Species	Class Id	Number Of Images
Acer cappadocicum	1	6
Astragalus pinetorum Boiss.	2	9
Aubrieta olympica Boiss.	3	8
Bornmuellera cappadocica	4	6
Campanula betulifolia C.Koch.	5	10
Centaurea appendicigera C.Koch.	6	12
Cirsium pseudopersonata Boiss.& Bal., subsp. ***pseudopersonata***	8	9
Delphinium Formosum Boiss. & Huet	9	19
Doronicum macrolepis Freyn. & Sint	10	9
Draba bruniifolia Stev. subsp. ***armeniaca*** Coode & Cullen	11	13
Euphrasia minima Jacq. ex DC. subsp. ***davisii*** Yeo	12	10
Heracleum plathytaenium Boiss.	13	10
Lilium ciliatum P.H.Davis	14	12
Melampyrum arvense L. var. ***elatius*** Boiss.	15	15

Muscari aucheri (Boiss.) Backer	16	14
Muscari bourgaei Backer	17	11
Papaver lateritium koch	18	18
Phlomis russeliana (Sims.) Bentham	19	10
Primula longipes Freyn & Sint.	20	14
Ranunculus dissectus Bieb. subsp. ***huetii*** (Boiss.) Davis	21	9
Rhododendron ponticum L. subsp. ***ponticum***	22	14
Sempervivum furseorum Murhead	23	7
Trifolium pannonicum Jacq. subsp. ***elongatum*** (Willd.) Zoh.	24	8

2.2. Image Processing

The dataset contains varying size images to feed them to the DNN, the images resized to 256x256 by using bilinear interpolation [19].

2.2.1. Image augmentation

To increase dataset size images are objected to some augmentation operations. Those operations are horizontal, vertical shift, shearing, and rotation. A sample of those operations is given in Figure 2.

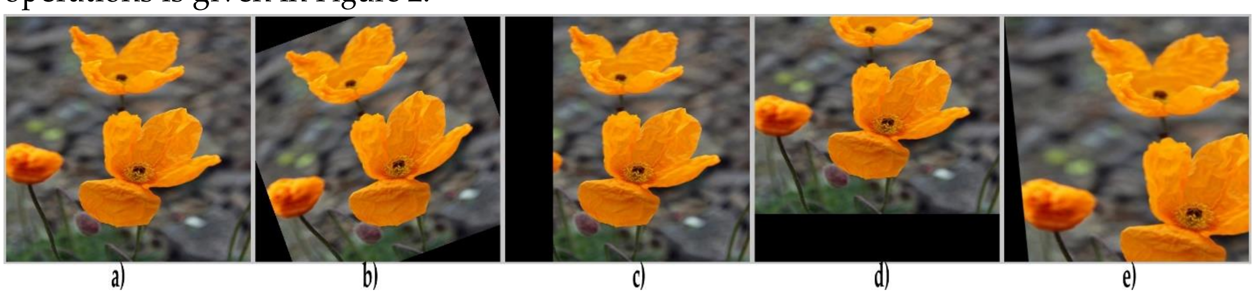

Figure 2 Augmentation Operations: a) Original Image b) Rotation Operation c) Vertical Shift d) Horizontal Shift e) Shearing.

2.3. Deep Neural Network

The classification operation done by using a DNN show in Figure 3. The network uses separable convolutions to extract more information without increasing computational power as regular convolutions. DNN creation, training, and testing done using Keras API [20].

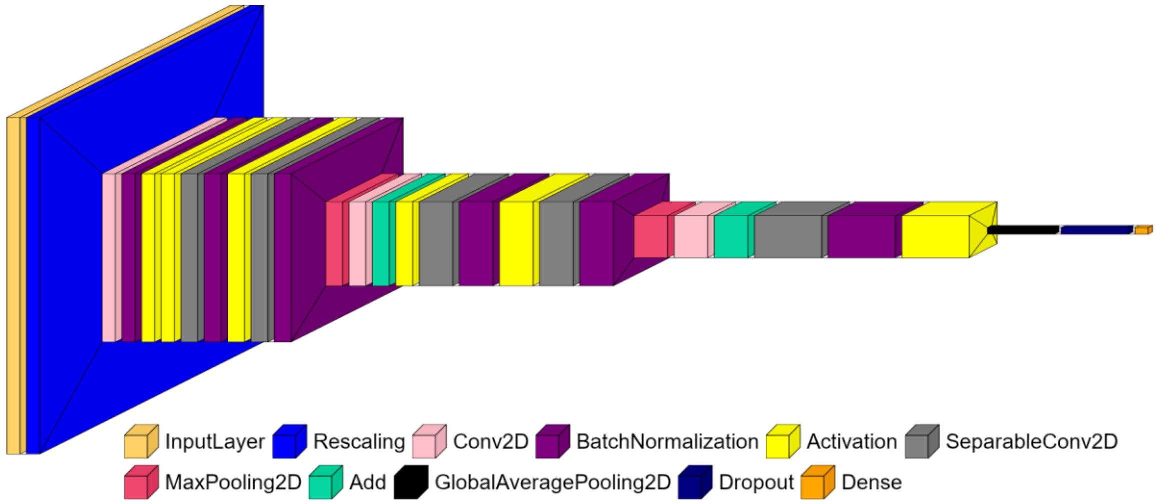

Figure 3 The DNN Designed for This Study

2.4. Evaluation

Segmentation performance evaluated using Accuracy, Precision, Recall, and F1-score metrics. Table 2 shows formulas of the metrics used. Accuracy measures the percentage of correctly classified images out of all the images in the dataset. Precision measures the proportion of correctly classified positive samples out of all the predicted positive samples. Recall measures the proportion of correctly classified positive samples out of all the actual positive samples. F1-score is the harmonic mean of precision and recall. It provides a balance between precision and recall [21].

Table 2: Evaluation Metrics Used and Their Formulas

Metric	Formulas
Accuracy	$\frac{True\ Positive\ (TP)\ +\ True\ Negative\ (TN)}{\text{TP + TN + False Positive (FP) + False Negative(FN)}}$
Precision	$\frac{TP}{TP + FP}$
Recall	$\frac{TP}{TP + FN}$
F1-score	$2 \times \frac{Precision\ x\ Recall}{Precision + Recall}$

3. Result

In this section experiments and their results are discussed. Experiments are conducted using Nvidia 1070ti graphics card and Keras API using Tensorflow 2.7.1 [22]. The dataset split into training and testing sets with a 5:1 ratio respectively and 5 folds constructed to train the network. Each fold is trained 10 times and the best performing network is chosen. The result for each split is shown in Table 3.

Table 3: Train and Test Performance of the Each Split

Set	Split	Accuracy	Precision	Recall	F1-score
Train	1	0.896	0.909	0.884	0.881
Train	2	0.91	0.926	0.9	0.905
Train	3	0.834	0.831	0.83	0.823
Train	4	0.863	0.861	0.858	0.855
Train	5	0.877	0.892	0.871	0.875
Test	1	0.732	0.685	0.674	0.662
Test	2	0.634	0.623	0.652	0.601
Test	3	0.683	0.652	0.659	0.612
Test	4	0.707	0.681	0.696	0.657
Test	5	0.756	0.75	0.761	0.711

Each split shows significant difference between train and test performance. However, that is expected due to the dataset size. There is no observation of direct correlation between training and testing performance of each split. Table 4 shows the average results.

Table 4: Average Results

Set	Accuracy	Precision	Recall	F1-score
Train	0.876	0.884	0.869	0.868

Test	0.702	0.678	0.688	0.6486

4. Discussion and Conclusion

This study shows promising results for the endemic plant segmentation task. However, the performance of the network can be increased. Further performance improvement is possible by increasing sample images per species and tweaking the network parameters. One of the biggest challenges for Endemic plant segmentation using DNNs is the lack of readily available datasets. Collaboration between individuals can help expand the dataset. Additionally, this effort can increase the number of species included. Implementing the DNN into a mobile application can help individuals to identify and protect endemic species. Another problem that needs to be addressed is some species look similar at the first glance. Figure 4 shows two species that are endemic and closely related. While there are visible differences, those differences can be minimized with image angle and lighting conditions. Moreover, those species can be classified by first identifying the main species then another specialized network can identify subspecies.

Figure 4: Species Left to Right Muscari Aucheri and Muscari Bourgaei

References

[1] Ç. H. Şekercioğlu *et al.*, "Turkey's globally important biodiversity in crisis," *Biological Conservation*, vol. 144, no. 12, pp. 2752–2769, Dec. 2011, doi: 10.1016/j.biocon.2011.06.025.

[2] C. Türe and H. Böcük, "Distribution patterns of threatened endemic plants in Turkey: A quantitative approach for conservation," *Journal for Nature Conservation*, vol. 18, no. 4, pp. 296–303, Dec. 2010, doi: 10.1016/j.jnc.2010.01.002.

[3] N. Coelho, S. Gonçalves, and A. Romano, "Endemic Plant Species Conservation: Biotechnological Approaches," *Plants*, vol. 9, no. 3, Art. no. 3, Mar. 2020, doi: 10.3390/plants9030345.

[4] B. Foggi, D. Viciani, R. M. Baldini, A. Carta, and T. Guidi, "Conservation assessment of the endemic plants of the Tuscan Archipelago, Italy," *Oryx*, vol. 49, no. 1, pp. 118–126, Jan. 2015, doi: 10.1017/S0030605313000288.

[5] K. IŞIK, "Rare and endemic species: why are they prone to extinction?," *Turkish Journal of Botany*, vol. 35, no. 4, pp. 411–417, Jan. 2011, doi: 10.3906/bot-1012-90.

[6] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

[7] F. Chollet, *Deep learning with Python*. Manning Publications Company, 2017.

[8] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*, 2009, pp. 248–255.

[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," 2012. [Online]. Available: http://code.google.com/p/cuda-convnet/

[10] I. Goodfellow *et al.*, "Generative Adversarial Nets," in *Advances in Neural Information Processing Systems 27*, Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2014, pp. 2672–2680. Accessed: May 21, 2019. [Online]. Available: http://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf

[11] C. Szegedy, A. Toshev, and D. Erhan, "Deep Neural Networks for Object Detection," in *Advances in Neural Information Processing Systems 26*, C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2013, pp. 2553–2561. Accessed: May 15, 2019. [Online]. Available: http://papers.nips.cc/paper/5207-deep-neural-networks-for-object-detection.pdf

[12] M. Yang, J. J. Yang, Q. Zhang, Y. Niu, and J. Li, "Classification of retinal image for automatic cataract detection," in *2013 IEEE 15th International Conference on e-Health Networking, Applications and Services, Healthcom 2013*, 2013, pp. 674–679.

[13] H. G. Akçay, B. Kabasakal, D. Aksu, N. Demir, M. Öz, and A. Erdoğan, "Automated Bird Counting with Deep Learning for Regional Bird Distribution Mapping," *Animals*, vol. 10, no. 7, Art. no. 7, Jul. 2020, doi: 10.3390/ani10071207.

[14] T. Danişman *et al.*, "PREDICTING THE LOCATION OF THE UTERINE CERVICAL OS FROM 2D IMAGES WITH CNN," *Mühendislik Bilimleri ve Tasarım Dergisi*, vol. 8, no. 5, Art. no. 5, Dec. 2020, doi: 10.21923/jesd.828457.

[15] Sefa AKBULUT, Zafer Cemal ÖZKAN, *ORMAN BOTANİĞİ DERS NOTLARI*.

[16] "Türkiyebitkileri.com - Anasayfa," Oct. 22, 2022. https://turkiyebitkileri.com/tr/ (accessed Mar. 07, 2023).

[17] "agaclar.net." http://www.agaclar.net/ (accessed Mar. 08, 2023).

[18] "Türkiye Bitkileri Listesi // bizimbitkiler.org.tr - Nezahat Gökyiğit Botanik Bahçesi - 2013." https://www.bizimbitkiler.org.tr/v2/index.php (accessed Mar. 07, 2023).

[19] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000.

[20] F. Chollet and others, *Keras*. GitHub, 2015. [Online]. Available: https://github.com/fchollet/keras
[21] A. A. Taha and A. Hanbury, “Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool,” *BMC Medical Imaging*, vol. 15, no. 1, p. 29, Aug. 2015, doi: 10.1186/s12880-015-0068-x.
[22] Martín Abadi *et al.*, “TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems.” 2015. [Online]. Available: https://www.tensorflow.org/