Recep Tayyip Erdoğan Üniversitesi Fen ve Mühendislik Bilimleri Dergisi, 5(1) 141-149 2024
Recep Tayyip Erdogan University Journal of Science and Engineering, 5(1) 141-149 2024
ISSN: 2687-2315 / E-ISSN: 2757-7686

An Investigation of Deep Learning Object Recognition on Dangerous Dog Breeds

İclal ÇETİN TAŞ [1*]

[1] Electrical-Electronics Engineering Department, Ostim Technical University, Ankara, 06374, Türkiye

Sorumlu Yazar/Corresponding Author
E-mail: iclal.cetintas@ostimteknik.edu.tr

Araştırma Makalesi/Research Article
Geliş Tarihi/Received: 20.07.2023
Kabul Tarihi/Accepted:24.05.2024

ABSTRACT

Many deep learning approaches have been developed to solve artificial intelligence problems with deep learning architectures. Due to its robust feature extraction and learning capabilities, it is frequently preferred in object recognition processes. Detection of dogs, one of the most preferred pets today, is essential for different purposes. It is preferred in analyses made based on gender. In this article, deep learning methods and deep learning and segmentation methods are used together to detect the dog in a data set consisting of 3 different dangerous dog breeds. This demonstrates the effectiveness of integrating segmentation techniques with deep learning architectures in improving object recognition performance. The findings suggest that such combined methods can be highly beneficial for precise and reliable detection in various applications involving animal recognition and classification. In the results obtained, it was seen that the accuracy rate increased to 88.33% with the tissue segmentation method used before NasNetLarge. The study highlights the potential of advanced deep learning techniques in achieving high accuracy rates in complex object detection tasks.

Keywords: ResNet, NasNetLarge, Inceptionresnetv2, Texture segmentation, Object recognition

Derin Öğrenme Nesne Tanımasının Tehlikeli Köpek Irkları Üzerinde İncelenmesi

ÖZ

Derin öğrenme mimarileri ile yapay zekâ problemlerinin çözümü için birçok derin öğrenme yaklaşımı geliştirilmiştir. Güçlü özellik çıkarma ve öğrenme yetenekleri sebebiylede nesen tanıma işlemlerinde sıkça tervih edilmektedir. Günümüzde en cok tercih edilen evcil hayvanların başında gelen köpeklerin tespiti farklı amaçlarla önem arz etmektedir. Cins bazında yapılan analizlerde tercih edilmektedir. Bu makalede, derin öğrenme yöntemleri ile 3 farklı tehlikeli köpek ırkından oluşan bir veri setinde köpeğin tespiti için derin öğrenme ve segmentasyon yöntemi birlikte kullanılmıştır. Bu, nesne tanıma performansını iyileştirmede segmentasyon tekniklerini derin öğrenme mimarileriyle entegre etmenin etkinliğini göstermektedir. Bulgular, bu tür kombine yöntemlerin, hayvan tanıma ve sınıflandırmayı içeren çeşitli uygulamalarda kesin ve güvenilir tespit için oldukça faydalı olabileceğini göstermektedir. Elde edilen sonuçlarda neural architecture search network (NasNetLarge) öncesinde kullanılan doku bölütleme yöntemiyle doğruluk oranını %88,33'e çıkardığı görülmüştür. Karmaşık nesne algılama görevlerinde yüksek doğruluk oranlarına ulaşmada gelişmiş derin öğrenme tekniklerinin potansiyelini vurguluyor.

Anahtar Kelimeler: ResNet, NasNetLarge, Inceptionresnetv2, Doku segmentasyonu, Nesne tanıma

Cite as;
Çetin Taş, İ. (2024). An investigation of deep learning object recognition on dangerous dog breeds. *Recep Tayyip Erdogan University Journal of Science and Engineering,* 5(1), 141-149. Doi: 10.53501/rteufemud.1330367

1. Introduction

Dogs are one of man's best friends. The pet population has been increasing worldwide in recent years. Figure 1 shows the increase between 2021 and 2022 in Europe. However, negative situations may occur when a suitable process is not followed for the animals that are owned or purchased.

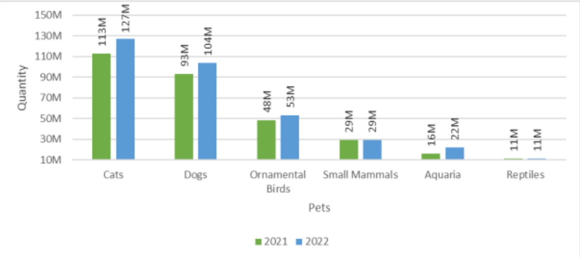

Figure 1: Pet population rate of increase in Europe from 2021 to 2022 (URL-1)

Some dog species can pose a danger to the environment from time to time. Although it is forbidden to keep these species as pets in some countries, this is not the case in every country. For this reason, in some cases, it is important to make early or instant detections in order to take precautionary measures. Object-oriented methodologies are at the forefront of the methods that allow detection. Object detection has been an important topic in the field of image processing and computer vision, which has become the focus of many researchers in recent years. Object detection has many sub-procedures, such as image classification, visual tracking, semantic segmentation, and so on. Object detection is done in two different ways, one is traditional methods and the other is an object detection algorithm based on deep learning. Traditional object detection algorithms include procedures such as feature extraction, and window-based operations, which are very slow and low performance compared to others. Deep learning-based approaches that eliminate these situations at a high rate have gained popularity (Wang et al., 2020).

Dog breed detection plays a crucial role in various applications, ranging from animal welfare to smart surveillance systems. It involves the identification and classification of different dog breeds based on their visual characteristics. Traditional methods of dog breed recognition relied heavily on human expertise, making them subjective and time-consuming. However, recent advancements in deep learning, a subfield of machine learning, have revolutionized this task by enabling automated and accurate identification of dog breeds.

Deep learning techniques, particularly Convolutional Neural Networks (CNNs), have emerged as a powerful tool for image recognition tasks, including dog breed detection. CNNs are inspired by the visual cortex of the human brain and excel at extracting features from images. They consist of multiple layers of interconnected nodes that process images in a hierarchical manner, gradually learning to recognize patterns and complex structures.

Fine grain recognition has been applied to the Stanford dog's dataset (URL-2). This method is difficult due to large intra-class variations, such as poses versus subtle local differences between classes. They emphasized in their work that the key to addressing this problem is to localize the distinctive parts to extract the pose-invariant properties. As a result of the study, 75.50% success was achieved by using Googlenet (Sermanet et al., 2014).

In another study using convolutional neural networks for the same dataset, Simon et al. associates achieved 68.61% success by using Alexnet (Simon and Rodner, 2015).

The other study achieved 94.5% success in detecting dogs and their breeds by using Yolo v3 on the data set created with data obtained from 8 different dog breeds. Mean average precision was used for performance evaluations (Wang et al., 2020).

Liu et al. (2016) presented Fully Convolutional Attention Networks (FCANs) for detailed recognition. With the proposed model with fully convolutional architecture, it has been tried to create a much faster way than previous reinforcement learning-based visual attention models during both training and testing. They proposed a method based on finding the optimal decision policy for generating actions from observations characterized by the parameters of the FCANs to maximize the expected gains at all-time steps. They have achieved 93.1% accuracy by asserting extensive experimental procedures on four different fine-grained criteria (Finetune baseline, Random regions, Center regions, Attention regions)

Raduly et al. (2018) as one of the multi-class classifications, tried to find a solution to an image recognition problem, which determines the breed of a dog in a particular image, by using convolutional neural networks. Two different networks were trained and evaluated in the Stanford Dogs dataset (URL-2). As a result of the analyzes made using NasNet-A mobile and Inception Residual Networks-v2 (Inception-ResNet-v2 networks), 90.69% accuracy rate of the Inception-ResNet-v2 method was obtained in solving the problem. They also developed an application and made it available for use.

A similar application has been made by proposing a CNN-based method to identify and breed dogs in potentially complex images. Results yielded approximately 85% accuracy in breed classification for 50 dog classes and 64% accuracy for the other 120 less common dog breeds. An iOS application is used to support image classification algorithms (Sinnott et al., 2019).

In this study, after texture segmentation of the images of 3 dangerous breeds in the Stanford dog dataset (URL-2), object detection was performed using 3 different deep learning architectures. The results were compared. Materials and methods will be examined in the flow of the study after the introduction. In the third chapter, the details of the experimental procedure and the analysis of the data obtained as a result will be made. In the last part, the evaluation of the results of the study and the predictions for future studies will be discussed.

2. Materials and Methods

In the study, the steps in Figure 2 were followed on the data. Before the object recognition process with deep learning, the accuracy of deep learning architectures in object recognition was analyzed by highlighting the objects on the image more than other components by using segmentation methods.

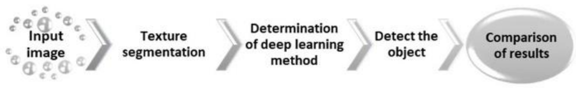

Figure 2: Flowchart of this study

2.1 Materials

The image dataset was obtained from Kaggle which is name Stanford dog image for object detection (URL-2). The analysis was carried out on the data set consisting of 120 different dog species. Sample images are shown in Figure 3. The dataset contains 12.000 images, approximately 100 per breed. Within the scope of the study, approximately 300 images of American Staffordshire terrier, boxer and doberman dogs were used from the data.

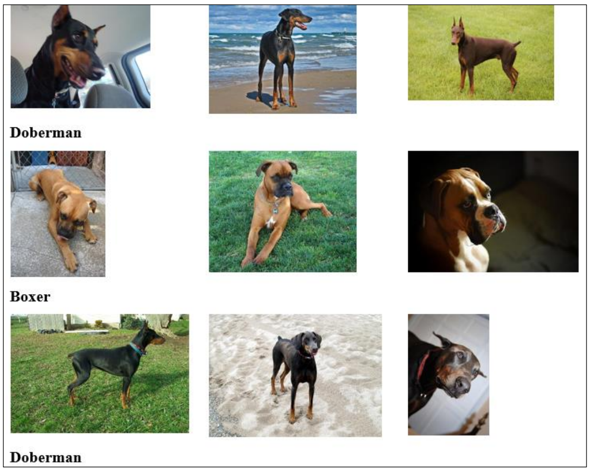

Figure 3: Sample images

2.2 Methods

Deep learning has been investigated with many different applications because of its strong learning capabilities and easier solutions to more complex problems. In the studies, it is seen that pre-processing procedures are applied in datasets. In some cases, post-processing methods are used as well as pre-processing to facilitate solving the classification, detection, or segmentation problem. The integrity and result oriented that pre- process and post-process methods provide in deep learning applications is a very interesting area (Salvi et al., 2021)

2.2.1 Texture Segmentation by Using Gabor Filter and K-means Clustering

Texture segmentation is the process of dividing an image into regions or segments based on the similarity of the texture patterns within each region. This is a challenging task because texture patterns can vary greatly in terms of their frequency, orientation, and scale.

Gabor filters are linear filters that are designed to mimic the receptive field properties of simple cells in the primary visual cortex of the human brain. These filters can capture both the frequency and orientation information present in an image. The filter responses are obtained by convolving an image with a set of Gabor filters of different scales and orientations.

Gabor filters are typically defined in the spatial domain but it can also be defined in the frequency domain. They are characterized by two key parameters: the spatial frequency and the orientation. The spatial frequency represents the number of oscillations (or cycles) of the filter along a given direction, and the orientation represents the angle of the filter concerning the horizontal axis.

The Gabor filter can be mathematically defined as the product of a Gaussian function and a complex sinusoidal function in Equation1:

$$\mathrm{G}(\mathrm{x},\mathrm{y}) = \exp\left(-\frac{\mathrm{x}^2+\mathrm{y}^2}{2\sigma^2}\right) \times \exp(\mathrm{j}2\pi\mathrm{f}\mathrm{x}) \tag{1}$$

where x and y represent the spatial coordinates, σ is the standard deviation of the Gaussian function, f is the spatial frequency of the sinusoidal function, and j is the imaginary unit.

The real part of the Gabor filter response represents the magnitude of the response at a particular spatial frequency and orientation, and the imaginary part represents the phase. By using multiple Gabor filters of varying spatial frequencies and orientations, we can obtain a representation of the texture properties of an image. Gabor filters are commonly used in applications such as texture analysis, object recognition, and face recognition. They are particularly useful for capturing texture properties that are difficult to represent using traditional filters, such as those found in natural scenes.

Clustering methods are used to group data showing similar characteristics in a data set. Clustering can simply be summarized as "the work of organizing objects into certain groups according to their similarity". In clustering methods, when there are no output values, only the input values are grouped. these methods are named as unsupervised learning methods. k-means algorithm is an unsupervised learning and clustering algorithm. The k value in k-means determines the number of clusters and should take this value as a parameter. They are called k-means because they form k unique clusters, and the center of each cluster is the mean of the values in the cluster. The algorithm puts statistically similar records in the same group. An element is allowed to belong to only one set. The cluster center is the value that represents the cluster. In this algorithm, the 'k' parameter specifies how many clusters the data will be divided into. Although there are several analysis methods for the selection of this parameter, the best is to run the algorithm at different k values and get the one that works best for us. Because different numbers of groups can expose different features (Liu and Yu, 2009; Dhanachandra et al., 2015; Zheng et al., 2018; Iong and Chen, 2021; Sahin et al., 2022;). The formulated form of the algorithm is shown in Equation 2.

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_i^{(j)} - c_j \right\|^2 \tag{2}$$

J: objective function, *k*: number of clusters, *n*: number of cases, x_i: case *i*, c_j: centroid for cluster *j*

$\left\| x_i^{(j)} - c_j \right\|^2$ can be defined as the distance function. One of the most important cases is distance calculation for k-means algorithm.

Texture segmentation by using Gabor filter and k-means clustering methods steps can be summarized as shown below.

- To perform texture segmentation using Gabor filter and K-means clustering, we can follow these steps.
- Apply a set of Gabor filters of different scales and orientations to the input image.
- Compute the filter responses for each pixel in the image.
- Flatten the filter responses into a feature vector for each pixel.
- Use K-means clustering to cluster the feature vectors into K clusters, where K is a user-defined parameter.
- Assign each pixel in the image to the nearest cluster centroid.
- Generate a segmented image by replacing each pixel in a cluster with the centroid of that cluster.
- Post-process the segmented image if necessary

This technique can be used for a variety of applications such as object recognition, image retrieval, and medical image analysis. However, it is important to note that the performance of this technique is highly dependent on the choice of parameters such as the number of Gabor filters, the number of clusters, and the initialization of the K-means algorithm.

2.2.2 CNN Architecture

Convolutional neural networks are a type of multilayer perceptron. Its design was inspired by the visual center of animals. In addition to image processing, it is frequently used in fields such as sound processing, natural language processing and biomedicine, and applications that produce successful solutions are made especially in the image field. Basic CNN architecture is shown in Figure 4.

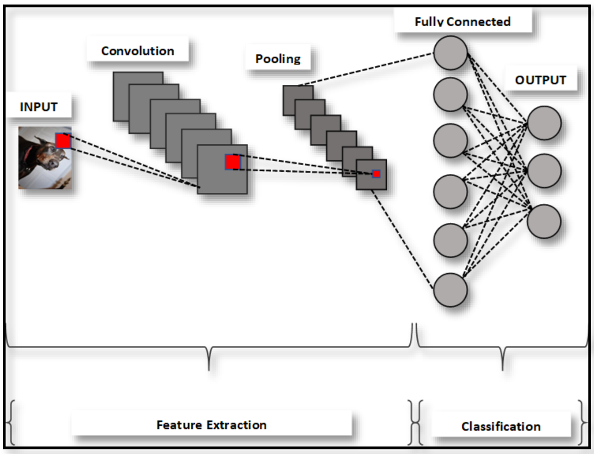

Figure 4: CNN architecture

The AlexNet architecture is a structure that has managed to significantly increase the “ImageNet” classification accuracy. There are 5 convolutional layers and 3 fully connected layers in its design. AlexNet uses Rectified Linear Unit (ReLu) as the characteristic feature of activation CNN structures in non-linear parts. One reason why ReLu is preferred over the sigmoid function used in neural network architectures is that it is faster when training the model. Another reason is that the sigmoid function causes a vanishing gradient problem. As the values of the sigmoid function get very large or small, its derivative approaches zero. It is called a vanishing gradient. Therefore, it becomes difficult to update the weights in the model. In cases where a very deep model is considered, the first layers may have difficulty learning because of this problem (Gao et al., 2020; Chen et al., 2022). Another problem solved with AlexNet architecture is to prevent overfitting by using dropout. To explain most shortly, overfitting means memorizing the given train data.

Residual Networks (ResNet) deep learning model has a deep convolutional neural network architecture. It contains 50 layers in structure, and the architecture of ResNet-50 starts with a convolution of 7 × 7 cores and 64 different cores, all with 2 steps in size, so we get one layer. It contains different convolutional layers with different kernels (Sai Bharadwaj Reddy and Sujitha Juliet, 2019 Rajpal et al., 2021; Murcia-Gómez et al., 2022).

The main contribution of the technology is the creation of a new search scope known as NasNet search scope. The optimal convolution layer is usually called cells, and these cells can be found in a small dataset and can be applied to a larger dataset. By stacking these network components further, the network can handle more complex and larger datasets. This network convolution design is named NasNet architecture (Ilhan et al., 2022; Alfarhood et al., 2023). The extended architecture in question is called NasNetlarge.

Inception-ResNet-V2 is a combination of Inception and ResNet architectures with improved recognition and classification performance. It is a created model. Resnet architecture is about going deep, Inception is about going wide. Therefore, with the Inception-ResNet-V2 architecture, we can achieve the optimum result in going both deep and wide. Inception-ResNet-V2 is a convolutional neural network algorithm based on the Inception architecture family and containing residual connections. This also allows for a significant simplification of the starting blocks. This structure allows optimization of the residual layer by changing the size of the initial convolution operation to 1×1. It also has the advantage of transferring the previous activation value to the output, even in conditions where learning stops (Ferreira et al., 2018; Özgüret al., 2021; Bozkurt, 2023).

In this study, a segmentation process is applied using K-means clustering and Gabor filtering methods. In the second step, it was used to predict the most probable breed with pre-trained deep learning models. This prediction is obtained by passing the image through the trained mesh and the output is a probability distribution over different dog breed. The breed with the highest probability is considered the predicted dog breeds. In this way, the process has been followed to detect dangerous dog breeds.

3. Result and Discussion

There are factors other than the target in the images in the dataset used in this study. For this reason, it is thought that making the target object more specific will contribute to the process. Texture Segmentation by using Gabor Filter and K-means Clustering has been applied to make the target object more visible. The results of this strengthening and preprocessing step are shown in Figure 5.

Figure 5: a) Gabor filter output b) K-Means Clustering result c) Texture segmentation results

After segmentation, object recognition was performed for each sample with 4 different pre-trained network architectures. The same process was applied to all images of each dangerous dog breeds. One of the result images is shown in Figure 6.

Figure 6: Object recognition result for a sample image by using NasNetlarge architecture

This accuracy is %89.2 for this image. This process was repeated for 4 different architectures. For each architecture, the average values of all results belonging to 3 different types were obtained. These values obtained for performance comparison are shown in Table 1.

Table 1: Obtained result accuracies from 4 different pre-trained networks

Accuracy (%)	Alex net	Res net18	NasNet large	Inception resnetv2
Dangerous dog breed dataset	73.68%	80.70%	88.33%	84.63%

When the results are examined, it is seen that the best result is obtained with the NasNetlarge architecture used after texture segmentation. The NasNetlarge input size is 331×331. Since Image is net-based, it can detect 1000 different objects. NasNet networks are scalable CNN architecture and consist of simple blocks such as separable convolution and jointing improved by the reinforcement learning method. NasNet-based architectures are created by repeating these blocks according to network capacity.

Deep learning-based dangerous dog breed systems have found practical applications in various domains, such as veterinary medicine, pet adoption platforms, and even automated dog show judging. The ability to accurately classify dog breeds using deep learning models not only enhances the efficiency of these applications but also contributes to our understanding of canine genetics and behavior.

In conclusion, segmentation methods have emerged as a vital component in deep learning-based dog breed detection systems. They improve the localization of relevant regions, enhance the model's robustness to variations and occlusions, facilitate breed-specific feature extraction, enhance interpretability, and contribute to dataset augmentation and annotation. By incorporating segmentation techniques into the breed detection pipeline, researchers and practitioners can achieve higher accuracy rates, more reliable predictions, and a deeper understanding of the intricate visual characteristics that define different dog breeds.

4. Conclusion

In this paper, we delve into the dangerous dog breed detection using deep learning. We explore the architecture and workings of CNN models, discuss various datasets and training techniques, and present state-of-the-art approaches and their performance. Furthermore, we examine the challenges and limitations of this technology and highlight potential future directions for research and development. By understanding the advancements and possibilities of dog breed detection using deep learning, we can unlock new opportunities for leveraging computer vision in the world of canines.

Existing object detection is usually based on the distinction of different objects. In this study, besides basic object detection, subspecies detection was made. A model based on segmentation and pre-trained deep network architectures is proposed so that the said analysis can be performed with higher accuracy. Experimental results showed that the accuracy rate of the analyzes made using only the deep learning model with the proposed method was exceeded. It shows that it has solved the problem of dog detection and classification. It is foreseen to be used in systems that include security, remote sensing systems and analysis on animals. In future studies, different methods for the detection of hybrid breeds will be differentiated or the development of hybrid models will have a positive reflection on the accuracy rates.

Author contribution

All studies for the article, such as material procurement, data collection, data processing, literature review, writing, and critical review, were carried out by İclal ÇETİN TAŞ.

Funding declaration

This research has not received a specific grant from any commercial or non-commercial funding agency.

Conflict of Interest Statement

The author(s) declare that they have no conflict of interest.

Ethical standards

No Ethics Committee Approval is required for this study

References

Alfarhood, S., Alrayeh, A., Safran, M., Alfarhood, M., and Che, D. (2023). Image-based Arabian camel breed classification using transfer learning on CNNs. *Applied Sciences 2023*, 13(14), 8192. https://doi.org/10.3390/APP13148192

Bozkurt, F. (2023). Skin lesion classification on dermatoscopic images using effective data augmentation and pre-trained deep learning approach. *Multimedia Tools and Applications*, 82(12), 18985–19003. https://doi.org/10.1007/S11042-022-14095-1/TABLES/9

Chen, H.C., Widodo, A.M., Wisnujati, A., Rahaman, M., Lin, J.C.W., Chen, L. and Weng, C.E. (2022). AlexNet convolutional neural network for disease detection and classification of tomato leaf. *Electronics*, 11(6), 951. https://doi.org/10.3390/ELECTRONICS11060951

Dhanachandra, N., Manglem, K. and Chanu, Y.J. (2015). Image segmentation using K -means clustering algorithm and subtractive clustering algorithm. *Procedia Computer Science*, 54, 764–771. https://doi.org/10.1016/J.PROCS.2015.06.090

Ferreira, C.A., Melo, T., Sousa, P., Meyer, M.I., Shakibapour, E., Costa, P. and Campilho, A. (2018). Classification of Breast Cancer Histology Images Through Transfer Learning Using a Pre-trained Inception Resnet V2. **In**: *Image Analysis and Recognition. ICIAR 2018. Lecture Notes in Computer Science, vol 10882* (Eds.Campilho, A., Karray, F., ter Haar Romeny, B), Springer, Cham. https://doi.org/10.1007/978-3-319-93000-8_86

Gao, Z., Shao, Y., Xuan, G., Wang, Y., Liu, Y. and Han, X. (2020). Real-time hyperspectral imaging for the in-field estimation of strawberry ripeness with deep learning. *Artificial Intelligence in Agriculture*, 4, 31–38. https://doi.org/10.1016/J.AIIA.2020.04.003

Ilhan, I., Bali, E., and Karakose, M. (2022). An ımproved deepfake detection approach with NasNetLarge CNN. *2022 International Conference on Data Analytics for Business and Industry, ICDABI 2022*, 598–602. https://doi.org/10.1109/ICDABI56818.2022.10041558

Iong, J. and Chen, Z. (2021). Automatic vehicle license plate detection using K-Means clustering algorithm and CNN. *Journal of Electrical Engineering and Automation (EEA)*. https://doi.org/10.36548/jeea.2021.1.002

Liu, D. and Yu, J. (2009). Otsu method and K-means. *Proceedings - 2009 9th International Conference on Hybrid Intelligent Systems, HIS 2009*, 1(2), 344–349. https://doi.org/10.1109/HIS.2009.74

Liu, X., Xia, T., Wang, J., Yang, Y., Zhou, F., Lin, Y., and Research, B. (2016). Fully convolutional attention networks for fine-grained recognition. *arXiv*,1603.06765. https://doi.org/10.48550/arXiv.1603.06765

Özgür, A., Bozkurt Keser, S.N. and Nur, S. (2021). Meme kanseri tümörlerinin derin öğrenme algoritmaları ile sınıflandırılması. *Turkish Journal of Nature and Science*, 10(2), 212–222. https://doi.org/10.46810/TDFD.957618

Murcia-Gómez, D., Rojas-Valenzuela, I. and Valenzuela, O. (2022). Impact of ımage preprocessing methods and deep learning models for classifying histopathological breast cancer ımages. *Applied Sciences*, 12(22), 11375. https://doi.org/10.3390/APP122211375

Ráduly, Z., Sulyok, C., Vadászi, Z. and Zölde, A. (2018). Dog breed ıdentification using deep learning. *SISY 2018 - IEEE 16th International Symposium on Intelligent Systems and Informatics, Proceedings*, 271–275. https://doi.org/10.1109/SISY.2018.8524715

Rajpal, S., Lakhyani, N., Singh, A.K., Kohli, R., and Kumar, N. (2021). Using handpicked features in conjunction with ResNet-50 for improved detection of COVID-19 from chest X-ray images. *Chaos, Solitons and Fractals*, 145, 110749. https://doi.org/10.1016/J.CHAOS.2021.110749

Sahin, M.E., Tawalbeh, L. and Muheidat, F. (2022). The security concerns on cyber-physical systems and potential risks analysis using machine learning. *Procedia Computer Science*, 201(C), 527–534. https://doi.org/10.1016/J.PROCS.2022.03.068

Sai Bharadwaj Reddy, A. and Sujitha Juliet, D. (2019). Transfer learning with RESNET-50 for malaria cell-image classification. *Proceedings of the 2019 IEEE International Conference on Communication and Signal Processing, ICCSP 2019*, 945–949. https://doi.org/10.1109/ICCSP.2019.8697909

Salvi, M., Acharya, U.R., Molinari, F. and Meiburger, K.M. (2021). The impact of pre- and post-image processing techniques on deep learning frameworks: A comprehensive review for digital pathology image analysis. *Computers in Biology and Medicine*, 128, 104129. https://doi.org/10.1016/J.COMPBIOMED.2020.104129

Sermanet, P., Frome, A. and Real, E. (2014). Attention for fine-grained categorization. *arXiv*, 1412.7054v3. https://doi.org/10.48550/arXiv.1412.7054

Simon, M. and Rodner, E. (2015). Neural activation constellations: unsupervised part model discovery with convolutional networks. *arXiv*, 1504.08289. https://doi.org/10.48550/arXiv.1504.08289

Sinnott, R.O., Wu, F. and Chen, W. (2019). A Mobile application for dog breed detection and recognition based on deep learning. *Proceedings - 5th IEEE/ACM International Conference on Big Data Computing, Applications and Technologies, BDCAT 2018*, 87–96. https://doi.org/10.1109/BDCAT.2018.00019

URL-1, (2023). https://europeanpetfood.org/_/news/new-fediaf-facts-figures-highlights-the-growth-of-european-pet-ownership/, July 19, 2023.

URL-2, (2023). https://www.kaggle.com/datasets/jessicali9530/stanford-dogs-dataset, May 12, 2023

Wang, C., Wang, J., Du, Q. and Yang, X. (2020). Dog breed classification based on deep learning. *Proceedings - 2020 13th International Symposium on Computational Intelligence and Design, ISCID 2020*, 209–212. https://doi.org/10.1109/ISCID51228.2020.00053

Zheng, X., Lei, Q., Yao, R., Gong, Y. and Yin, Q. (2018). Image segmentation based on adaptive K-means algorithm. *Eurasip Journal on Image and Video Processing*, 2018(1), 1–10. https://doi.org/10.1186/S13640-018-0309-3/FIGURES/14