International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 6 Issue 2, January-February 2022 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Deep Learning Applications and Image Processing

Ahmet Özcan[1], Mahmut Ünver[2], Atilla Ergüzen[3]

[1]Instructor, Department of Computer Engineering, Kırıkkale University, Turkey
[2]Assistant Professor, Department of Vocational High School, Kırıkkale University, Turkey
[3]Associate Professor, Department of Computer Engineering, Kırıkkale University, Turkey

ABSTRACT

With the rapid development of digital technologies, the analysis and processing of data has become an important problem. In particular, classification, clustering and processing of complex and multi-structured data required the development of new algorithms. In this process, Deep Learning solutions for solving Big Data problems are emerging. Deep Learning can be described as an advanced variant of artificial neural networks. Deep Learning algorithms are commonly used in healthcare, facial and voice recognition, defense, security and autonomous vehicles. Image processing is one of the most common applications of Deep Learning. Deep Learning software is commonly used to capture and process images by removing the errors. Image processing methods are used in many fields such as medicine, radiology, military industry, face recognition, security systems, transportation, astronomy and photography. In this study, current Deep Learning algorithms are investigated and their relationship with commonly used software in the field of image processing is determined.

KEYWORDS: *Deep Learning, Image Processing, CNN, Neural Networks, Yolo*

How to cite this paper: Ahmet Özcan | Mahmut Ünver | Atilla Ergüzen "Deep Learning Applications and Image Processing" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-6 | Issue-2, February 2022, pp.1-5, URL: www.ijtsrd.com/papers/ijtsrd49142.pdf

IJTSRD49142

I. INTRODUCTION

In recent years, the use of Deep Learning mechanisms and techniques in image processing applications has become popular. The availability of powerful computing environments on the Internet [1]; and the fact that cloud providers offer ready-made machine learning and artificial intelligence laboratories to researchers [2] have accelerated these efforts.

The concept of machine learning has been used since the 1990s. But it has not evolved for many years. The reason is that the computers are not powerful enough, the data sets are small, the wrong installation and the use of wrong activation functions [3]. The increase in computer power, fast access to internet; and increasing technological capabilities in cloud environments have increased the interest in machine learning and deep learning approaches in recent years.

Deep learning is a machine learning method that predicts the results of a given dataset and its structure consists of more than one artificial neural network. It can be termed as a subfield of machine learning. Researchers use Deep Learning software to analyze complex and large data sets and process image, text, and audio data more accurately and quickly. Behaviors that can be perceived by humans, such as smart home devices perceiving commands, autonomous vehicles distinguishing pedestrians, recognizing fresh and spoiled food, are among the topics of Deep Learning and successful results are obtained.

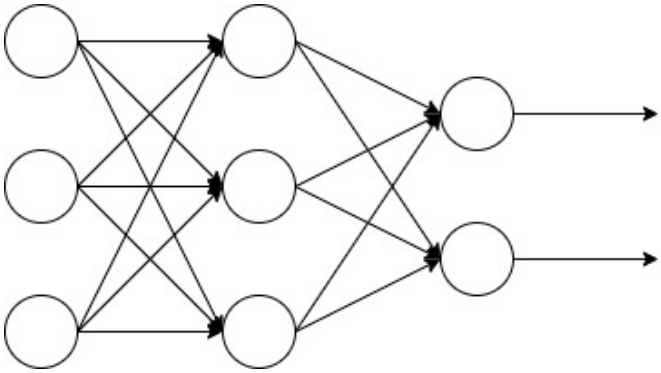

Figure 1 Neural network structure

Nowadays, there are many different neural network architectures designed for different purposes. The simple architecture of a neural network is shown in Figure 1. The names of these architectures are defined

by the type of layers used and the way the layers are interconnected.

The deep learning algorithm Convolutional Neural Network (CNN) is still widely used for object recognition [4, 5]. CNN has shown successful results in image processing applications such as image segmentation and classification [6]. As an example, consider the image of a bird in Figure 2. To determine whether it is really a bird or some other object, send the pixels of the image as arrays to the input layer of the neural network (multilayer networks for object classification). Hidden layers implement various computational methods and perform feature extraction by making changes. Convolutional layer performs the process of feature extraction from the image sent to the system. As the last layer, there is a fully connected layer that defines the object in the image.

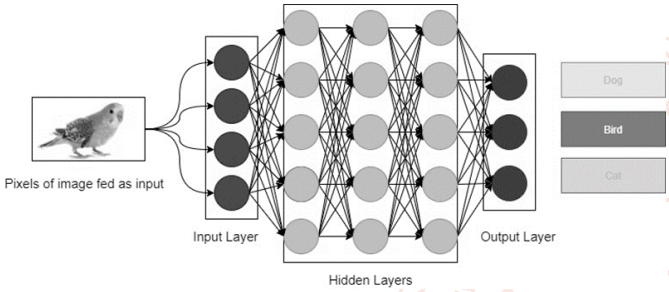

Figure 2 Convolutional Neural Network to identify the image of a bird

Convolutional Neural Networks are inspired by biology. CNN consists of three basic layers: Convolutional Layer, Pooling Layer and Fully-connected Layer. Any number of Convolutional and Pooling layers can be applied sequentially. Then the fully-connected layer is used. If the classification problem is to be solved with multiple labels, the softmax layer is used as the last layer. In the fully connected layer, the three dimensional input is reduced to one dimension [7]. In the literature, there are many deep learning models that use Convolutional Neural Network (CNN). The most commonly used CNN models include LeNet, AlexNet, ZFNet, VGGNet, GoogLeNet and ResNet.

LeNet uses a 5-layer CNN. It was developed in 1998 to classify handwriting on bank checks. This model uses mean pooling method to reduce the size [8]. AlexNet consists of 25 layers. It uses max pooling and softmax activation function. Due to the pooling layer, it is similar to LeNet [9]. ZFNet has a similar structure to AlexNet. In this model, the filter sizes in the first layer were changed and the object detection was improved. Moreover, a new technique called Deconvolution Network was developed in this model and the success rate was increased [10]. VGGNet is a CNN model supported by a graphics processing unit (GPU). In this model, the double or triple convolutional layers are followed by the pooling layer. Although the number of layers is high in this model, the data size decreases from input to output [11]. GoogLeNet is a CNN model with 144 layers. This model follows a different sequential approach using parallel network segments [12]. ResNet consists of 152 CNN layers. It has a deeper structure compared to other CNN architectures [13].

II. DEEP LEARNING APPLICATIONS

There are many software/libraries developed for Deep Learning. In this part of the article, you will find information about the most commonly used software and its features.

II.I. THEANO

It was developed by the Université de Montréal in Python. Of course, it uses the NumPy library of the Python language. It has GPU support. With GPU support, it can perform operations 140 times faster than CPU. It can also perform mathematical calculations using KERAS and BLOCKS applications. It can effectively define mathematical expressions, including multi-dimensional arrays, enabling optimization and evaluation. It can run on a variety of operating systems. It can be pre-trained by Lasagne's Model Zoo. It is capable of running on multiple nodes. It has extensive unit testing and self-validation features.

II.II. TENSORFLOW

Written in C++ and Python by the Google Brain team. There is support for Linux, Windows and Mac operating systems OS. It has GPU support. It is a free and open source software library. It focuses on neural network training and inference. It can perform its computations on TPUs and GPUs. Today, it supports Python as well as many other languages such as C++, Java, C#, Javascript, and R. It is widely used. The reason for its popularity is that TensorFlow libraries are prepared for different platforms. It consists of a large library prepared for mobile apps, IoT apps, web apps, and artificial intelligence apps. Developers use TensorFlow for artificial intelligence algorithms that are used on the most popular mobile devices

II.III. CAFFE

It was developed by the Berkeley Vision and Learning Center in the C++ software language. It has support for multiple platforms. It has GPU support. It encourages development with an extensible code structure. It provides good solutions for research experiments and industrial applications. It supports deep learning algorithms for image classification and image segmentation. With an Nvidia K40 GPU, it can process over 60 million images per day. Caffe's accessible fast ConvNet implementation. Caffe

software supports CNN, RCNN, LSTM and fully connected neural networks.

III. IMAGE PROCESSING APPLICATIONS

This section will describe two commonly used image processing applications.

III.I. YOLO

It is an algorithm that has been widely used in recent years for computer recognition of objects. Its most important feature is real-time object recognition. The general average accuracy (mAP) values used in object recognition are widely used because they are better than others. In Yolo algorithm, a single neural network is applied to the image. The network divides the image into regions. It estimates the bounding boxes and probabilities for each region and the bounding boxes are weighted by the probabilities determined. Yolo looks at the entire image once during the test. Unlike the R-CNN algorithm, Yolo makes predictions with a single network evaluation. This way of working makes it 1000 times faster than the R-CNN algorithm and 100 times faster than the Fast R-CNN algorithm [14].

III.II. SSD

SSD algorithm is used for real-time object detection. SSD speeds up the process as the region does not know the bid network. However, it uses optimization methods such as multiscaling and standard boxes to improve the accuracy drop. With these improvements, SSD tries to match the accuracy of Faster R-CNN on low resolution images and further increase the speed. The SSD algorithm consists of 2 stages. In the first stage, feature maps are created. In the second stage, convolutional filters are applied to detect objects.

IV. RELATED WORKS

There are many articles in the literature about the LeNet algorithm used for recognizing objects and scripts. Recognition of Arabic letters and numbers using LeNet-5 [15], recognition of gasses for electronic noses [16], recognition of road signs using the optimized LeNet5 algorithm [17], recognition of Covid-19 disease from CT images using LeNet-5 architecture [18] has been observed.

While examining the studies using AlexNet Deep Learning model, high performance was observed in vegetable classification [19], a review of AlexNet, AlexNet and VGG-16 for ear detection [20], a proposed AlexNet architecture for diabetic retinopathy image classification [21], AlexNet and VGG architecture using non-foldable layers. There are studies such as comparison [22]. The ZFNet deep learning model can be considered as a version of the AlexNet algorithm. It has been used in areas such as detection of defects in corn plants [23], detection of fungal diseases in plants [24], classification of fruits sold in retail stores [25], and diagnosis of Covid 19 disease using Deep Learning [26]. The VGGNet algorithm has been used in painting style classification [27], chicken disease identification [28], motion detection [20], fish species classification [29], single shot multiple object detection [30]. GoogLeNet algorithm has been used to test the reliability of CNN on GPUs [31], for face recognition and classification [32], for artificial intelligence based classification of clothing [33], for Covid-19 disease detection [34].

Theano software has been used by researchers for machine learning applications in finance [35] and for water vapor removal applications in TDS data [36]. Tensor Flow is the most widely used library for image processing software. It has been used for static number crunching [37], X-ray image classification [38], and machine learning in business and finance [39]. The Yolo image processing algorithm is used to solve many image processing problems, such as one-step object detection, detection of thermal objects in harsh weather conditions, accurate detection of apple blossoms in natural environments, detection of medical masks in the fight against Covid-19, and detection of hazards on sidewalks.

V. CONCLUSION

Deep learning algorithms have been used for many years. The acceleration of computer processors has made deep learning algorithms popular again. The method, which initially started with artificial neural networks, produces effective results as the number of layers is increased and the manipulations within the layers are differentiated. A foldable neural network model that can be created for one problem may not be a suitable solution for another problem. Deep learning models are commonly used to solve image processing problems. Libraries created for image processing make problem solving very easy. Libraries created for image processing are constantly being updated and given new capabilities. Problems that could not be solved with the first versions of the libraries are solved faster with the new versions

REFERENCES

[1] V. Monga, Y. Li, and Y. C. Eldar, “Algorithm Unrolling: Interpretable, Efficient Deep Learning for Signal and Image Processing,” IEEE Signal Process. Mag., vol. 38, no. 2, pp. 18–44, 2021, doi: 10.1109/MSP.2020.3016905.

[2] Microsoft, “Azure Machine Learning,” 2021. https://azure.microsoft.com/tr-tr/services/machine-learning/#product-overview (accessed Oct. 05, 2021).

[3] Beyaz.net, "Derin Öğrenme (Deep Learning) Nedir?," beyaz.net, 2019. https://www.beyaz.net/tr/yazilim/makaleler/derin_ogrenme_deep_learning_nedir.html (accessed Nov. 13, 2021).

[4] G. E. Hinton, "A practical guide to training restricted Boltzmann machines," in Neural networks: Tricks of the trade, Springer, 2012, pp. 599–619.

[5] A. Angelova, A. Krizhevsky, and V. Vanhoucke, "Pedestrian detection with a large-field-of-view deep network," in 2015 IEEE international conference on robotics and automation (ICRA), 2015, pp. 704–711.

[6] A. Kaya, V., Tuncer, S., Baran, "Derin öğrenme yöntemleri kullanilarak nesne tanima," Int. Sci. Technol. Conf. Lefkoşa, Kıbrıs (Kktc), 2 - 04 Eylül 2020, pp. 277–287, 2020.

[7] Akcayol and M. Ali, "Derin Öğrenme-Deep Learning (Convolutional neural networks)," w3.gazi.edu.tr, 2016. https://w3.gazi.edu.tr/~akcayol/files/DL_L6_CNNs.pdf (accessed Dec. 11, 2021).

[8] Y. Le Cun et al., "Handwritten digit recognition: Applications of neural network chips and automatic learning," IEEE Commun. Mag., vol. 27, no. 11, pp. 41–46, 1989.

[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," Adv. Neural Inf. Process. Syst., vol. 25, pp. 1097–1105, 2012.

[10] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in European conference on computer vision, 2014, pp. 818–833.

[11] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv Prepr. arXiv1409.1556, 2014.

[12] M. A. KIZRAK and B. BOLAT, "Derin öğrenme ile kalabalık analizi üzerine detaylı bir araştırma," Bilişim Teknol. Derg., vol. 11, no. 3, pp. 263–286, 2018.

[13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.

[14] J. Redmon and A. Farhadi, "YOLO v.3," Tech Rep., pp. 1–6, 2018, [Online]. Available: https://pjreddie.com/media/files/papers/YOLOv3.pdf.

[15] A. El-Sawy, E.-B. Hazem, and M. Loey, "CNN for handwritten arabic digits recognition based on LeNet-5," in International conference on advanced intelligent systems and informatics, 2016, pp. 566–575.

[16] G. Wei, G. Li, J. Zhao, and A. He, "Development of a LeNet-5 gas identification CNN structure for electronic noses," Sensors, vol. 19, no. 1, p. 217, 2019.

[17] C. Zhang, X. Yue, R. Wang, N. Li, and Y. Ding, "Study on traffic sign recognition by optimized Lenet-5 algorithm," Int. J. Pattern Recognit. Artif. Intell., vol. 34, no. 01, p. 2055003, 2020.

[18] M. R. Islam and A. Matin, "Detection of COVID 19 from CT Image by The Novel LeNet-5 CNN Architecture," in 2020 23rd International Conference on Computer and Information Technology (ICCIT), 2020, pp. 1–5.

[19] L. Zhu, Z. Li, C. Li, J. Wu, and J. Yue, "High performance vegetable classification from images based on alexnet deep learning model," Int. J. Agric. Biol. Eng., vol. 11, no. 4, pp. 217–223, 2018.

[20] C. Alippi, S. Disabato, and M. Roveri, "Moving convolutional neural networks to embedded systems: the alexnet and VGG-16 case," in 2018 17th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN), 2018, pp. 212–223.

[21] T. Shanthi and R. S. Sabeenian, "Modified Alexnet architecture for classification of diabetic retinopathy images," Comput. Electr. Eng., vol. 76, pp. 56–64, 2019.

[22] W. Yu, K. Yang, Y. Bai, T. Xiao, H. Yao, and Y. Rui, "Visualizing and comparing AlexNet and VGG using deconvolutional layers," 2016.

[23] S. Bansal and A. Kumar, "Deep Learning for Maize Crop Deficiency Detection," in Modeling, Simulation and Optimization: Proceedings of CoMSO 2020, 2021, pp. 473–483.

[24] K. Bhargavi and B. S. Babu, "Application of Convoluted Neural Network and Its Architectures for Fungal Plant Disease Detection," in Artificial Intelligence and IoT-

Based Technologies for Sustainable Farming and Smart Agriculture, IGI Global, 2021, pp. 314–324.

[25] J. L. Rojas-Aranda, J. I. Nunez-Varela, J. C. Cuevas-Tello, and G. Rangel-Ramirez, "Fruit Classification for Retail Stores Using Deep Learning," in Mexican Conference on Pattern Recognition, 2020, pp. 3–13.

[26] H. Catal Reis, "COVID-19 Diagnosis with Deep Learning," Ing. e Investig., vol. 42, no. 1, 2022.

[27] Z. Yang, "Classification of picture art style based on VGGNET," in Journal of Physics: Conference Series, 2021, vol. 1774, no. 1, p. 12043.

[28] L.-D. Quach, N. Pham-Quoc, D. C. Tran, and M. F. Hassan, "Identification of Chicken Diseases Using VGGNet and ResNet Models," in International Conference on Industrial Networks and Intelligent Systems, 2020, pp. 259–269.

[29] E. Prasetyo, N. Suciati, and C. Fatichah, "Multi-level residual network VGGNet for fish species classification," J. King Saud Univ. Inf. Sci., 2021.

[30] K. Sathish, S. Ramasubbareddy, and K. Govinda, "Detection and localization of multiple objects using VGGNet and single shot detection," in Emerging Research in Data Engineering Systems and Computer Communications, Springer, 2020, pp. 427–439.

[31] Y. Ibrahim, H. Wang, and K. Adam, "Analyzing the Reliability of Convolutional Neural Networks on GPUs: GoogLeNet as a Case Study," in 2020 International Conference on Computing and Information Technology (ICCIT-1441), 2020, pp. 1–6.

[32] R. Anand, T. Shanthi, M. S. Nithish, and S. Lakshman, "Face recognition and classification using GoogleNET architecture," in Soft Computing for Problem Solving, Springer, 2020, pp. 261–269.

[33] S.-K. Noh, "Classification of Clothing Using Googlenet Deep Learning and IoT based on Artificial Intelligence," Smart Media J., vol. 9, no. 3, pp. 41–45, 2020.

[34] X. Yu, S.-H. Wang, X. Zhang, and Y.-D. Zhang, "Detection of COVID-19 by GoogLeNet-COD," in International Conference on Intelligent Computing, 2020, pp. 499–509.

[35] T. C. Nokeri, "Investment Portfolio and Risk Analysis," in Implementing Machine Learning for Finance, Springer, 2021, pp. 167–178.

[36] C. Thanapirom, C. J. Yi, N. Cota, R. Jintamethasawat, W. Kusolthossakul, and K. Prasertsuk, "An ADAM Optimization for Water-Vapor Effect Removal in THz-TDS Data," in 2020 45th International Conference on Infrared, Millimeter, and Terahertz Waves (IRMMW-THz), 2020, pp. 1–2.

[37] S. Lagouvardos, J. Dolby, N. Grech, A. Antoniadis, and Y. Smaragdakis, "Static analysis of shape in TensorFlow programs," 2020.

[38] D. Boinon et al., "Patient satisfaction with a rapid diagnosis of suspicious breast lesions: Association with distress and anxiety," Breast J., vol. 24, no. 2, pp. 154–160, 2018, doi: 10.1111/tbj.12856.

[39] I. Hull, "Machine Learning for Economics and Finance in TensorFlow 2," Apress, Berkeley, CA, 2021.