



An Automatic Fruit Image Classification System

Elham Mohammed Thabit A. Alsaadi^{a,*}, Meeras Salman Al-Shemarry^b,
Ahmed Sileh Gifal^c, Ghosoon k.munahy^d

^aDepartment of Information Technology, College of Computer Science and Information Technology, University of Kerbala, karbala, Iraq, elham.thabit@uokerbala.edu.iq

^bDepartment of Information Technology, College of Computer Science and Information Technology, University of Kerbala, karbala, Iraq, meeras.s@uokerbala.edu.iq

^cDepartment of Information Technology, College of Computer Science and Information Technology, University of Kerbala, karbala, Iraq, ahmed.alememe@uokerbala.edu.iq

^dDepartment of Information Technology, College of Computer Science and Information Technology, University of Kerbala, karbala, Iraq, ghosoon.k@uokerbala.edu.iq

ABSTRACT

Classifying fruits and vegetables is still challenging in daily production. Deep Convolutional Neural Networks (DCNNs) have made significant progress in solving prediction problems, such as object recognition, scene interpretation, and semantic segmentation, frequently outperforming humans in accuracy. In this study, we provide an effective fruit classification system in digital images utilizing deep learning techniques. By training the system on images of three different fruit categories: grape, citrus, and pomegranate, a deep learning strategy based on convolutional neural networks (CNNs) has been constructed to classify the item (fruit). We created an algorithm that automatically extracts and uses features from images in training.

The dataset used is 600, for training 80% were used, while the remaining images were used for testing. Based on our experiment, we discovered that 60x60 pixels is the ideal input image size, and 100 epochs is the perfect number. The accuracy of the test photos reached 97%, and the results are excellent. The findings demonstrate that the suggested methodology improves fruit classification ability overall.

Received: 17 / 01 / 2025

Accepted: 04 / 04 / 2025

Published: 30 / 06 / 2025

Keywords:

Deep Learning, DCNN, CNN



1. Introduction

Conventional techniques of classifying fruits have frequently depended on manual processes based on visual skills, which are laborious, time-consuming, and inconsistent [1], [2]. For fruit classification, external form appearance is the primary source. The fruit business has found computer machine vision and image processing techniques to be more and more helpful in recent years, especially for applications in quality inspection, size, color, and shape sorting [3], [4]. Several works in this field show that it is feasible to use machine vision systems to enhance product quality and eliminate the need for people to sort fruits by hand [5], [6]. A Convolutional Neural Network (CNN) is a class of

*Corresponding Author: Alsaadia, Elham

Email address: elham.thabit@uokerbala.edu.iq

deep neural networks, it is a multilayered neural network with a special architecture to detect complex features in data [7], [8]. Deep Learning permits computers to automatically elicit multiple levels of abstraction from raw data [9], [10]. It has played a distinguished role in solving various problems related to speech and image recognition, etc. [8]. This study proposed this technique to find an efficient solution for detecting and classifying fruits in still images. Many recent studies have been made to categorize and identify the images of objects (fruits): An effective deep learning framework for fruit classification was presented by (M. Shamim Hossain et al., 2019). More precisely, two distinct deep learning architectures serve as the foundation for the system. The first is a suggested light model of 6 layers of CNN, whereas the second is a refined Visual Geometry Group-16 pretrained deep learning model. This system has been tested on 2 color image datasets, one of which is publicly accessible. Clear fruit photographs make up the first dataset (dataset 1), whereas difficult-to-classify fruit images make up the second dataset (dataset 2). On dataset 1, the classification accuracy of 99.49% and 99.75% were attained for the first and second [11]. A. Mahdi et al. (2020) suggested a Convolutional Neural Network (CNN) classification model for orange photos. Deep learning CNN is used to classify oranges into five classes: good orange-grade-1, good orange-grade-2, immature orange, rotten orange, and damaged orange. The smartphone's camera was used to capture a total of 1000 orange photos. There are 200 photos in each class. The model has been validated using the K-Fold Cross-Valuation approach. The CNN hidden layer in this work has 256 nodes. According to the results, the ReLU activation function has a 96% accuracy rate, which is higher than the Tanh activation function's 93.8% accuracy rate [12]. (Khatun et al., 2020) proposed a convolutional neural network technique for fruit classification. Seven test samples were extracted from the 180 total; 20 images were used for stages of training and testing in order to gain the results. For creating and testing the suggested algorithm, anaconda software was used. For training and testing, a variety of fruits with varying backgrounds were selected. The accuracy rate of the suggested method was 98%. In this study, a CNN algorithm-based fruit categorization is investigated. Using the fruits-360 dataset, the accuracy and loss curves were produced for 5 different hidden layer combinations. This study uses a variety of computer vision-based techniques and algorithms for fruit classification and recognition. Improved CNN performance to achieve improved fruit classification [5]. This study aims to identify "what" objects (fruit) are inside an image for three types of fruits, which include: Grape, citrus, and pomegranate. It aims to develop an essential methodology for the automated classification of the three main groups of fruits (Grape, Citrus, and Pomegranate).

2. Work Challenges

In short, the challenges of this work are summarized as follows:

- The variations in the shapes and colors of various objects that fall under the same category.
- Lighting/Illumination Conditions: Differences in light intensity and brightness directions have a significant impact on the image's color.

3. Techniques of Deep Learning

One machine learning method that trains computers to perform human-like tasks is called deep learning [13]. This method uses visuals, text, or voice to teach a computer model to do categorization tasks [14], [15]. Neural network topologies with multiple layers and a sizable collection of labeled data are used to train models. The deep learning network for object classification is depicted in Fig. 1.

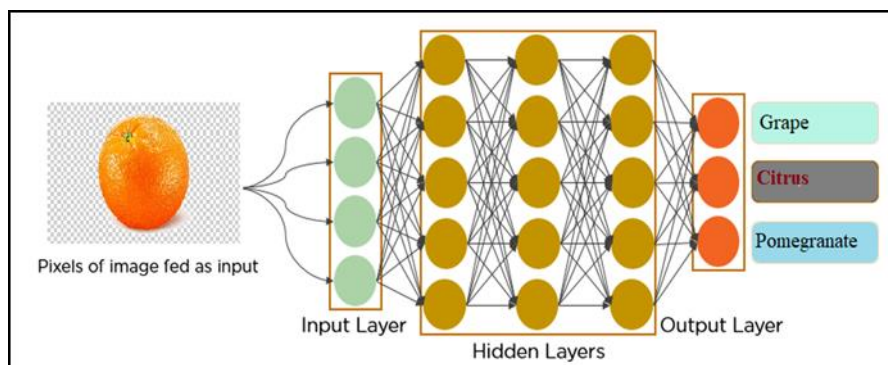


Fig. 1- Deep Learning Network for Image Classification [13]

3.1. Convolutional Neural Networks

CNNs are a type of deep neural network (DNN) that is most frequently used in deep learning for visual imagery analysis [19], [20]. It employs a unique method known as convolution. A mathematical operation on two functions that yields a third function that expresses how the shape of one is altered by the other is known as convolution [16], [17]. Multiple layers of artificial neurons make up convolutional neural networks. Typically, the first layer extracts fundamental properties like diagonal or horizontal edges. The subsequent layer receives this output and uses it to identify more intricate characteristics like corners or combinational edges [18], [19].

The classification layer generates a series of confidence scores (from 0 to 1) that indicate the likelihood that a picture belongs to a "class" based on the activation map of the last convolution layer [20].

The weighted outputs of the neurons in the preceding layer serve as the inputs to the neurons; if the layer is fully connected (FC), all of the outputs from the preceding layer are included. Weight regulates how much neuron output impacts the following neuron. Different sets of weighted outputs from earlier levels are used in each hidden layer. CNNs have shown promise in object detection and picture recognition. Convolution, ReLU, pooling, and the fully connected layer are the four fundamental parts of a neural network.

4. Proposed Method

In order to identify and categorize three different fruit varieties, grape, citrus, and pomegranate, we developed new algorithm in this paper. The algorithm is based on a robust DCNN-based model.

Information has been extracted from the overlapping of small regions obtained from the previous layers using three convolutional layers. The system prediction of the Grape class from three categories is shown in Fig. 4. The suggested algorithm for identifying and categorizing three different kinds of fruits is:

- The input comprises of 600 RGB images, of which 480 were used for stage of training and 120 for testing. A sample of the dataset is shown in Fig. 2.
- Identify fruits in a still image and categorize them into three groups (pomegranate, citrus, and grape).



Fig. 2- Sample Images of the Dataset

4.1. The Suggested Network Architecture

Max-pooling layers come after each of the three convolutional layers in the suggested model. This CNN's input was an RGB image, its first layer used 16 filters, its kernel size was 3, and its activation function was the Rectified Linear Unit (ReLU). The arrangement of the second layer is identical, with the exception of the 32 filters. Additionally, 64 filters were utilized in the third convolution layer. We use max_pooling with a pool_size of 2 and a stride of 1 to decrease the data after each layer. The two primary fruit classifications are detected using the ANN network design.

Each neuron's inputs are first multiplied by the weights, and the bias value is then added up. The outcomes are then sent through. Fig. 3 shows the proposed Architecture of CNN.

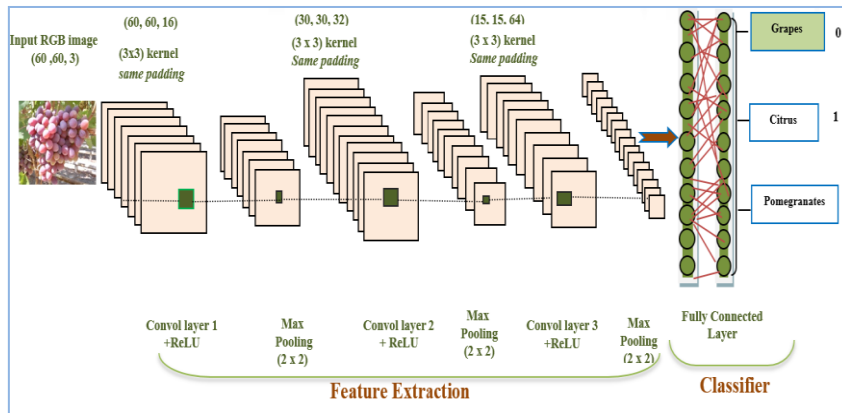


Fig. 3- Proposed Architecture of CNN

4.2. The Proposed Algorithm

The proposed algorithm (ConvNet) has been chosen, which it is distinct type of multilayer neural networks. It purposes to identify visual patterns from pixel images with minimal preprocessing for three categories of fruits. Figure 4 depicts the system predication of Pomegranates class from 3 categories.

ConvNet, a unique kind of multilayer neural network, has been selected for this suggested approach. It aims to extract visual patterns for three different fruit types from pixel images with pre-processing. The system prediction of the Pomegranates class from 3 categories is shown in Fig. 4.

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 60, 60, 16)	448
max_pooling2d_1 (MaxPooling2)	(None, 30, 30, 16)	0
conv2d_2 (Conv2D)	(None, 30, 30, 32)	4640
max_pooling2d_2 (MaxPooling2)	(None, 15, 15, 32)	0
conv2d_3 (Conv2D)	(None, 15, 15, 64)	18496
max_pooling2d_3 (MaxPooling2)	(None, 7, 7, 64)	0
dropout_1 (Dropout)	(None, 7, 7, 64)	0
flatten_1 (Flatten)	(None, 3136)	0
dense_1 (Dense)	(None, 500)	1568500
dropout_2 (Dropout)	(None, 500)	0
dense_2 (Dense)	(None, 3)	1503
Total params: 1,593,587		
Trainable params: 1,593,587		
Non-trainable params: 0		
predict accuracy: [0.9846042788356816, 0.6055045874293791]		
1/1 [=====] - 0s 120ms/step		
2		
pomegranates		

Fig. 4- System Prediction of Pomegranate Class

Training Algorithm

- Create Model of Convolutional Neural Network (CNN) which includes the following:
- Reading image with size 60 *60.
- Convoluting the image by using filters (16, 32, 64) and 3 max-pooling (pool size=2)
- Extracting feature maps, take the rectified feature as input to produce pooled feature map.
- Converting the entire result 2-D array from pooled feature map into a single long continuous linear vector by Flattened process.
- Feeding the Flattened matrix from the pooling layer as input to the fully connected layer to classify the image.
- Coding the outputs, the fruit is Grape when label =0, the fruit is Citrus when label =1, the fruit is Pomegranate when label =2.

Testing Algorithm

- Loading model.
- Reading colored image.
- Convoluting the image by using filters.
- Applying predicate process, according to the specific number of output.

5. Experimental Results

We used the following tests to assess our suggested approach:

5.1. As illustrated in Fig. 5, the system accurately predicts 480 distinct pomegranate images when they are used as input images.

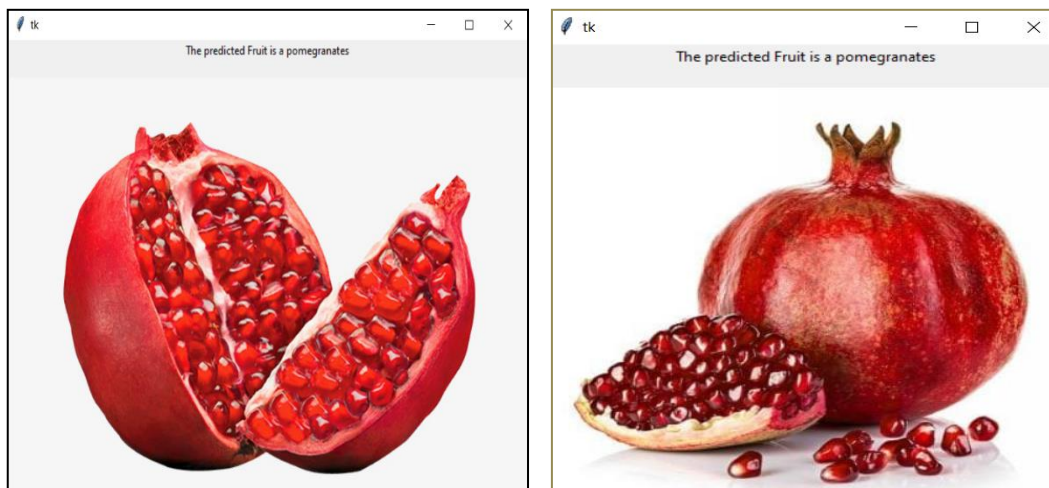


Fig. 5- Samples of Pomegranates Images

5.2. Citrus images were utilized as input; the network trained 480 images for various citrus varieties, and the tested images were effectively identified. Samples of citrus images are shown in Figure 6.

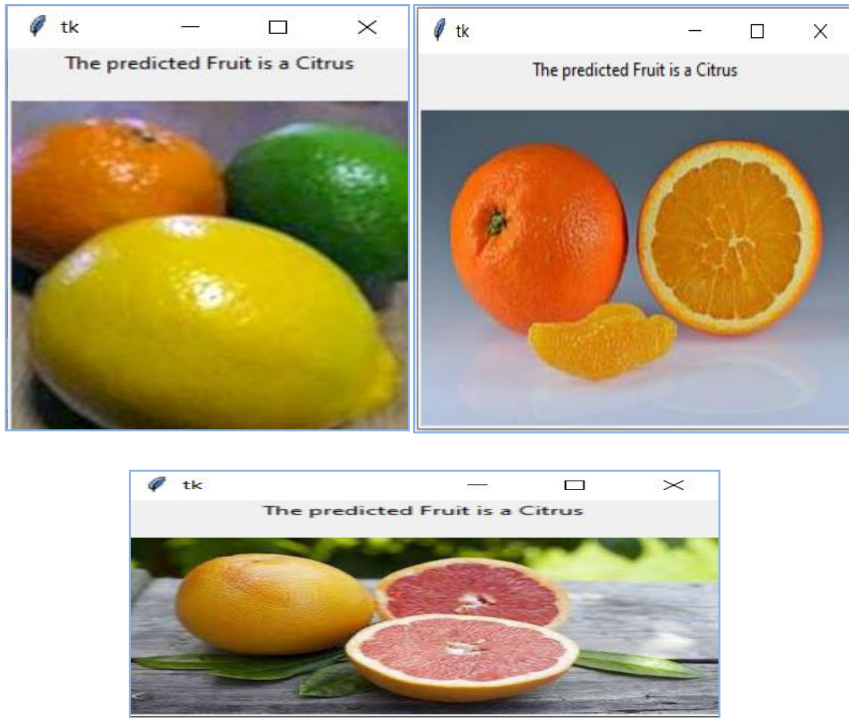


Fig. 6- Sample of Citrus Images

5.3. Images of grapes were utilized as input; the system trained 480 photographs of various grape varieties, and the tested images of grapes were effectively identified; an example of grape images is shown in Fig. 7.

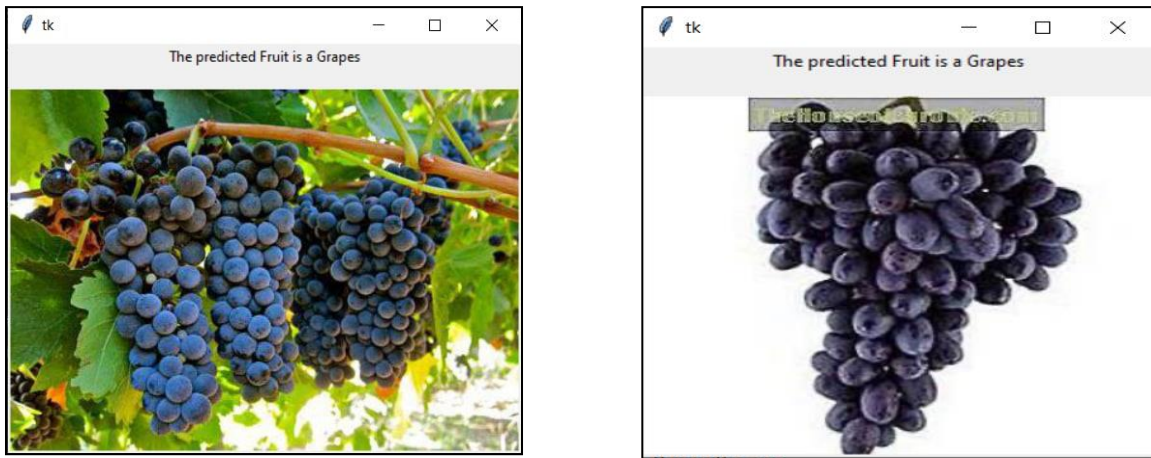


Fig. 7- Sample of Grapes Images

5.4. When the input image was dark or blurry, the system was able to classify them successfully as shown in Fig. 8.

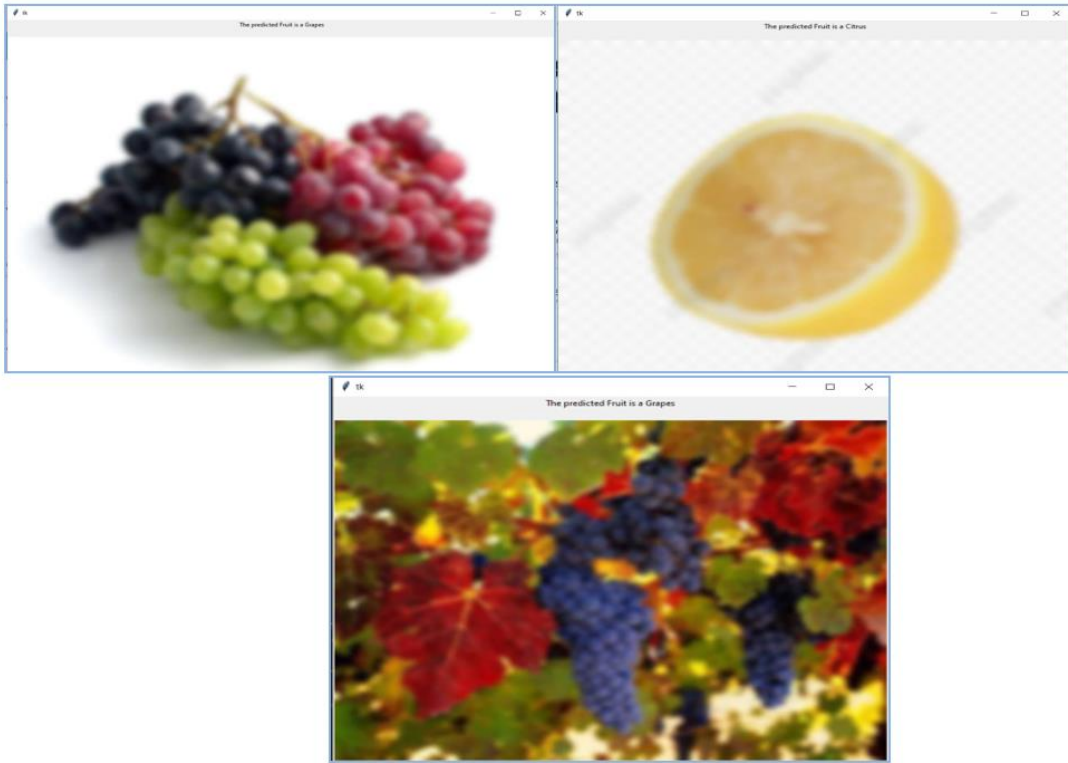


Fig. 8- Samples of Detected Blurry Images

5.5. The fifth test involves calculating the training duration. Based on Fig. 9, we can infer that increasing the image size and the number of epochs will both lengthen the training time.

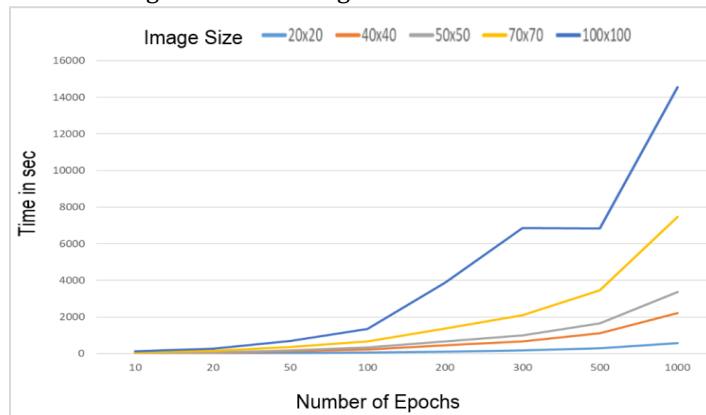


Fig.9- The Relation Among “Image Size, Number of Epochs, and Training Time”

6. Conclusion

To identify and categorize images into three classes, grape, citrus, and pomegranate, each of which contains multiple distinct fruits, we developed an artificial convolutional neural network using deep learning in this work. With a 97% accuracy rate and 100 iterations, our network produced excellent results. 480 images were used to train the system, while 120 images were used for testing. In this work, we built an artificial convolutional neural network to detect and classify images into three classes (Grape, citrus, and pomegranate), each class having several of different fruits. The suggested approach demonstrated a high fruit detection effectiveness even under an abnormal environment.

As far as we are aware, this effort identifies and categorizes fruits into the three fruit groups with all of their varieties, whereas the majority of other works concentrate on identifying particular fruit types in photos and occasionally classifying a small number of specific fruits (no more than ten different fruits).

7. Future Works

1. The number of fruit categories can be increased for classification.
2. Segment the fruit based on using CNN.
3. This work can be improved to work with video and road cameras to obtain a set of real photos.

References

- [1] 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.
- [2] P. Badre, S. Bandiwadekar, P. Chandanshive, A. Chaudhari, and M. S. Jadhav, "Automatically Identifying Animals Using Deep Learning," *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 6, no. 4, pp. 194–197, 2018.
- [3] A. Arora, A. Grover, R. Chugh, and S. S. Reka, "Real time multi object detection for blind using single shot multibox detector," *Wirel. Pers. Commun.*, vol. 107, no. 1, pp. 651–661, 2019.
- [4] F. Deng, X. Zhu, and J. Ren, "Object detection on panoramic images based on deep learning," in 2017 3rd International Conference on Control, Automation and Robotics (ICCAR), 2017, pp. 375–380.
- [5] M. Khatun, F. Ali, N. A. Turzo, J. Nine, and P. Sarker, "Fruits Classification using Convolutional Neural Network," *GRD Journals-Global Res. Dev. J. Eng.*, vol. 5, no. 8, 2020.
- [6] E. ALSAADI and N. K. El Abbadi, "Auto Animal Detection and Classification among (Fish, Reptiles and Amphibians Categories) Using Deep Learning," *J. Adv. Res. Dyn. Control Syst.*, vol. 11, no. 10-SPECIAL ISSUE, pp. 726–736, 2019.
- [7] B. R. Lee, "An image segmentation approach for fruit defect detection using k-means clustering and graph-based algorithm," *Vietnam J. Comput. Sci.*, vol. 2, no. 1, pp. 25–33, 2015.
- [8] E. M. T. A. Alsaadi and N. K. El Abbadi, "An automated mammals detection based on SSD-mobile net," in *Journal of Physics: Conference Series*, 2021, vol. 1879, no. 2, p. 22086.
- [9] K. T. Schütt, H. E. Saucedo, P.-J. Kindermans, A. Tkatchenko, and K.-R. Müller, "SchNet—A deep learning architecture for molecules and materials," *J. Chem. Phys.*, vol. 148, no. 24, p. 241722, 2018.
- [10] A. Azulay and Y. Weiss, "Why do deep convolutional networks generalize so poorly to small image transformations?," *arXiv Prepr. arXiv1805.12177*, 2018.
- [11] M. S. Hossain, M. Al-Hammadi, and G. Muhammad, "Automatic fruit classification using deep learning for industrial applications," *IEEE Trans. Ind. Informatics*, vol. 15, no. 2, pp. 1027–1034, 2018.
- [12] D. M. Asriny, S. Rani, and A. F. Hidayatullah, "Orange Fruit Images Classification using Convolutional Neural Networks," in *IOP Conference Series: Materials Science and Engineering*, 2020, vol. 803, no. 1, p. 12020.
- [13] H. Yousif, J. Yuan, R. Kays, and Z. He, "Animal Scanner: Software for classifying humans, animals, and empty frames in camera trap images," *Ecol. Evol.*, 2019.
- [14] N. Shone, T. N. Ngoc, V. D. Phai, and Q. Shi, "A deep learning approach to network intrusion detection," *IEEE Trans. Emerg. Top. Comput. Intell.*, vol. 2, no. 1, pp. 41–50, 2018.
- [15] A. Dingli and K. S. Fournier, "Financial time series forecasting—a deep learning approach," *Int. J. Mach. Learn. Comput.*, vol. 7, no. 5, pp. 118–122, 2017.
- [16] A. Rakhlin, "Convolutional neural networks for sentence classification," GitHub, 2016.
- [17] W. Rawat and Z. Wang, "Deep convolutional neural networks for image classification: A comprehensive review," *Neural Comput.*, vol. 29, no. 9, pp. 2352–2449, 2017.
- [18] N. Aloysius and M. Geetha, "A review on deep convolutional neural networks," in 2017 International Conference on Communication and Signal Processing (ICCSPP), 2017, pp. 588–592.
- [19] A. Ajit, K. Acharya, and A. Samanta, "A review of convolutional neural networks," in 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), 2020, pp. 1–5.
- [20] S. Seo, J. Huang, H. Yang, and Y. Liu, "Interpretable convolutional neural networks with dual local and global attention for review rating prediction," in Proceedings of the eleventh ACM conference on recommender systems, 2017, pp. 297–305.