# Intelligent System for Real-Time Potholes Monitoring and Detection

Dev Dalmia [1], Reetu Jain[2][0000-0002-7199-2823] and Syed Abou Iltaf Hussain[3][0000-0002-5157-0607]

[1] Class of 2023, Bombay International School: Gilbert Building, 2nd Cross Ln, Babulnath, Khareghat Colony, Gamdevi, Mumbai, Maharashtra 400007, India.
devdalmia1@gmail.com
[2] Supervisor, On My Own Technology Private Limited, Mumbai, India
reetu.jain @onmyowntechnology.com
[3] Supervisor, On My Own Technology Private Limited, Mumbai, India
syed.hussai@onmyowntechnology.com

**Abstract.** Roads are the major connection between towns, cities, states, and countries. Due to poor construction, heavy pouring and very heavy weight vehicles, often potholes are formed on the asphalt layer. These potholes are also one of the major causes for road accidents. The main aim of this project is to build an efficient as well as effective system for detecting the potholes. In this regard, the image processing techniques are integrated with the convolutional neural network (CNN). Images of the potholes are taken from the video feed of the drone. The frames with pothole images are acquired and the features are extracted by the image processing which are then fed into the CNN model. The CNN model is trained and tested with 784 images which are labeled as normal and pothole. The model showed training and testing accuracy of 0.9902 and 0.9912 respectively. The strength of the proposed approach is its practical applicability and ability to detect potholes in real-time. In order to test the effectiveness of the prototype and the proposed approach, it is tested for 25 different instances for identifying the potholes in roads. Finally, the time-taken for detection of the potholes by the proposed approach is tested.

**Keywords:** Potholes detection, Image Processing, Convolutional Neural Network, Features extraction, Drone images.

## 1 Introduction

In this present fast paced world, roads play a very important role in connectivity. Roads are the major connection between towns, cities, states, and countries. In the year 2021 - 22, road construction per day in India is estimated to be about 28.64 km. In the last decade, the Indian road network has grown about 30% with a 300% increase in vehicle registration [1]. As per the Ministry of Road Transport and Highways, a total of 815140 major road accidents happened in India in the year 2019 and 2020 killing 282827 and injuring 799640 persons [2]. Of the total major road accidents occuring in India, around 60000 is due to poor road conditions claiming 2300 lives annually between 2016 to 2020. However just 1% of the lives are lost to accidents due to potholes.

Potholes are bowl shaped holes in the asphalt of the road caused mainly due to heavy pouring, poor road construction and weight of the overweight trucks and lorry [3]. The potholes if not repaired on time then it may lead to the reconstruction of the roads which may turn into a costly affair. Therefore detecting potholes is a critical component of maintenance and rehabilitation because of their serious effect on pavement quality and driver safety. In India potholes are detected by visual inspections which is a labor intensive and costly process [4]. Hence, in the present scenario many researchers have developed novel ideas on detecting the potholes.

### 1.1 Review of Contemporary Literature

In contemporary literature, real time detection of the potholes is primarily done by i. image-processing and ii. Convolutional Neural Network (CNN). In the article [5], proposed a method for automatic detection of potholes based on the surrounding shadows, elliptic shape, and grain surface texture information. In [6], the authors proposed three cascade classifiers to detect potholes based on the features of longitudinal-transversal cracks, and fatigue cracks based on local binary pattern. In the paper [7], the authors developed a supervised classifier based on the semantic texton forests (STFs) algorithm for detecting the potholes from the feed of the video camera. In the reference [8], the authors proposed a method that takes into account the canny filter and contour detection integrated with road color to detect the potholes. However the model failed to detect the potholes that are clogged due to waste deposition. In the article [9], the authors developed a method that could roughly estimate the potholes and act as a pre-processing step for other supervised classifiers. The model initially extracted the potholes from the RGB images by analyzing the difference in the color space and took the help of image segmentation, and spectral clustering technique to identify the potholes. In the article [10], the authors have developed a potholes detection system where the feed is taken from the videos. The disadvantage of the developed method is that it takes only one pothole at a time, hence it is very difficult to identify multiple potholes at a time.

With the advancement made in the field of technology, deep learning (DL) methods such as CNN are used for detecting the potholes. Some of the literature that employed CNN includes [11 - 14]. In the article [11], the authors developed a supervised DL technique named as CrackNet to classify the problem of road defect detection on smartphones. In [12], the authors provided a patch-based deep learning approach to achieve satisfying detection results on their own collected dataset. In [13], the study employed a DL technique to minimize the amount of the image annotation, thus reducing the cost of annotation by domain experts. In the research paper [14], the authors developed a new large-scale dataset with eight classes of road surface damage. They evaluated the single shot multibox detector with MobileNet and Inception V2 on their public dataset and achieved recall and precision greater than 75%. In the research paper [15] focussed on detecting the potholes in 2D vision with a neural network (NN) that showed precision of 95.2% and recall of 92%. Other papers are reviewed for compiling this research paper but limiting it to the most relevant papers.

### 1.2 Motivation and Novelties

From the literature reviewed for the study, it is observed that most of the research paper either used image processing or DL method for real time detection of the potholes. However the main problem in the techniques is the less accuracy and higher computational time in detecting the potholes for real time application. The present research bridges the gap by integrating image processing with CNN. The features of the potholes are extracted by image processing which are then exported to the CNN to classify the potholes from the manholes.

The remainder of this paper is organized accordingly. The section 2 of the paper discusses the preliminary concepts and the proposed methodology. The section 3 of the paper discusses the drone prototype which is used for detecting the potholes. Section 4 is the results obtained by applying the proposed method in the problem and finally section 5 concludes the paper.

## 2 Preliminaries and Methodology

In this section of the paper, the preliminary concept of the methods employed for real time detection of potholes is discussed.

### 2.1 2.1. IMAGE PROCESSING

Image processing is the process of extracting useful information from an image by performing different operations on it [15]. The different image processing steps which are used in this research paper are as follows:

**Frame acquisition**

In this step, the frame from the videos captured by the drone is extracted which are further processed for deriving features from it [16]. The GoPro camera captures images in the range of 24 fps to 240 fps. However the camera is setted for recording videos at 30 fps.

**Grayscale conversion**

In this step, the frame acquired from section 2.1.1 is converted into grayscale images. The grayscale images represent only the information of light intensity. Unlike the RGB images, the grayscale images typically display from the darkest black to the brightest white i.e. the image only consists of black, white, and gray colors [17]. The advantage of converting a RGB image to Grayscale image is that it reduces the computational complexity of the algorithm [18].

**Light intensity correction**

In this step of the present study, the intensity of the light is adjusted by filtering out the excessive light. This step is done in order for the efficient performance of the edge detection algorithm.

**Pixel intensity**

For feature extraction, pixels play a very important role. All the information of a picture is stored in the pixels of the image [19]. Therefore in this step the pixel intensity is adjusted so that maximum information could be obtained from the images.

**Edge detection**

Edge detection in image processing is the way of detecting the boundaries of objects in the images [20]. The process of edge detection works by detecting the discontinuities in brightness of the objects. This process is mostly used for data extraction. Canny edge detection [21] technique is used for detecting the edge of potholes in the project.

**Contour detection**

Contour detection can be defined as the process of joining all the continuous points along the boundary of the objects within an image that are having the same color and intensity [22]. The shape features can be extracted from the contour detection.

## 2.2 Convolutional Neural Network

In the present study, Convolutional Neural Network (CNN) is employed to classify the potholes from the images. In the past two decades CNN has been established as a state-of-the-art technology to classify images. The generic steps followed for identifying the impairment caused due to smoking are as follows:

**Dataset**

For the present study, the pothole images dataset collected from the kaggle database is used for training the CNN model. The shape, size and texture of the potholes are the features that are extracted from the images by following the steps from section 2.1.2 to 2.1.6. The features are then fed into the CNN model for efficient classification of the potholes. The dataset comprises 784 images of which 614 images were used for training the CNN and remaining 170 images were used for testing the model.
*the pothole image dataset is obtained from https://www.kaggle.com/datasets/atulyakumar98/pothole-detection-dataset?resource=download

### Preprocessing

In this step of the paper the images are standardized by converting them into 200*200 pixels Joint Photographic Experts Group (jpeg) images. As feeding non-standard pictures to develop CNN models may result in slow learning and harder training which may result in lower accuracy [23]. Above that, in order to prevent the images from overfitting some random noises are added to the images. Data augmenting in CNN model enhances the learning and predicting capability of the CNN model [24].

### Convolution operation

Convolution operation is the application of a filter to the input images resulting in a map of activations called a feature map that indicates the locations and strength of a detected feature in an image [25]. The convolution layer converts all the pixels in its receptive field into a single value by decreasing the size of the image and bringing all the necessary information together into a single pixel.

### Activation function

In CNN, the activation layer is a non-linearity layer that consists of an activation function whose input is the feature map that is generated in the convolution layer and output is the activation map. In the present study, the rectified linear unit (ReLU) is used as an activation function.

### Pooling operations

The pooling operation is done to reduce the number of dimensions in the feature maps which in turn the number of parameters to be learned is less and also the computation performed. The importance of this operation is accurate prediction by the model with variations in the position of the features in the input image. In the present study, the Max Pooling operation is employed.

### Layer stacking operations

Layer stacking operation is the repetition of the convolution, activation and pooling operation until the error computed between the input and the output is the least. In this paper, the error is computed by the cross-entropy method.

### Fully connected layer

The fully connected layer consists of neurons equal to the number of labels which are fully connected to neurons of the previous layer.

### Classification

In this step of the CNN model, the label of the input images are predicted. For the present study, CNN model classifies potholes from manholes and drainage holes on the road.

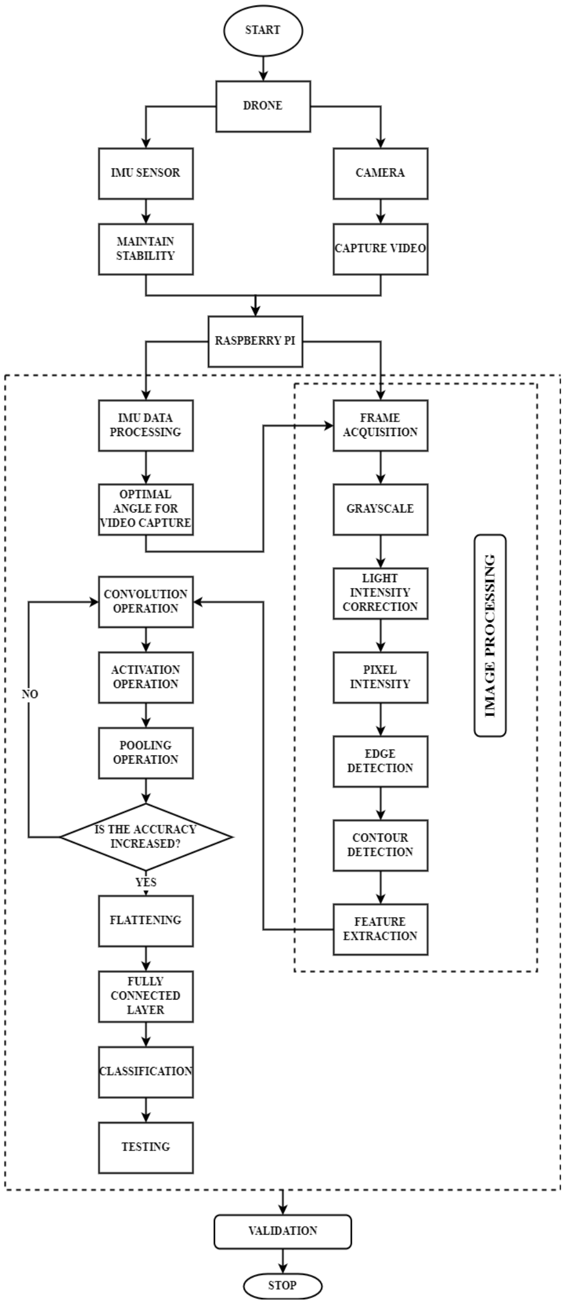


**Figure 1:** Flowchart for the proposed methodology

### 2.3 The Proposed Methodology

The comprehensive intention of the present research is to develop a potholes classification model by integrating the image processing and CNN algorithm. The CNN model is trained by the pothole images whose shape, size and textures are extracted by image processing. The flowchart for the proposed methodology is shown in figure 1 of the research paper.

## 3 The Prototype

The prototype of the drone is a first-person view (FPV) quadcopter. It is based on the Dragon HX5 frame which is 100% carbon fiber. The electronics used are the Mamba Flight Controller and Electronic Speed Controller Stack with a 120 Amp peak current. This is paired with 2450 kV Racerstar brushless motors. The drone utilizes the Futaba transmission protocol along with 5 W radio transmission for the FPV camera. Additionally, a GoPro is mounted on the underside to take pothole photos. The drone was flown in auto level mode in straight lines to take clear continuous video of the potholed road. The drone comprises IMU sensors which are used for maintaining stability and calculating the optimal angles for proper video capturing. Figure 2 shows a labeled picture of the prototype.

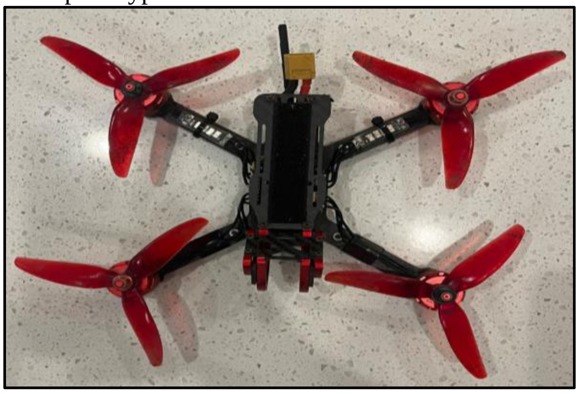

**Figure 2:** Labeled picture of the drone

## 4 RESULTS AND DISCUSSIONS

The proposed methodology for detecting the potholes is coded in Python 3.10 in a Window 10 system having RAM capacity of 8GB with i5, 1.6 GHz processor. The CNN is developed using the Tensorflow library which is run for 100 epochs with 20 steps per epoch.

### 4.1 Results

The training accuracy increased from 0.5863 in the 1st epoch to 0.9902 in the 100th epoch. However, the loss value as computed by the cross entropy method decreased from 0.6855 in the 1st epoch to 0.0344 in the 100th epoch. The graph for training accuracy and loss is shown in figure 3.

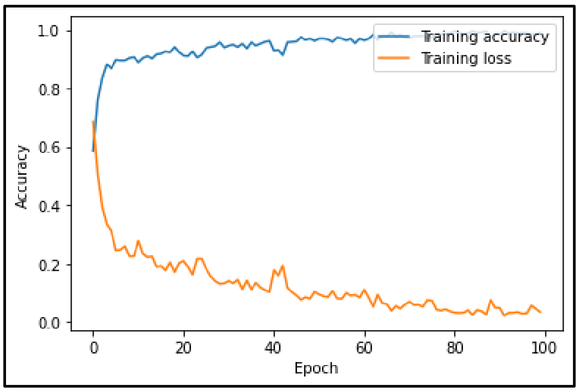


**Figure 3:** Training accuracy and loss graph

For validating the developed CNN model, the label for 170 images of the test dataset is predicted. The validation accuracy increased from 0.8073 in the 1st epoch to 0.9912 in the 100th epoch. On the other hand, the loss value decreased from 0.6061 in the 1st epoch to 0.1421 in the 100th epoch. The graph for validation accuracy and loss is shown in figure 4.

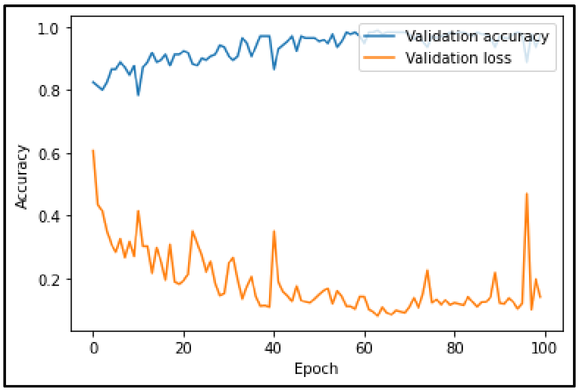


**Figure 4:** Testing accuracy and loss graph

Out of 170 images used for validating the CNN model, 88 images belong to the normal category whereas 82 images belong to the pothole category. Out of the 88 normal category images, 47 were rightly classified and the remaining 41 were falsely classified. Out of the 82 pothole category images, 48 were rightly classified and the remaining 34 were falsely classified. The confusion matrix for the CNN classification model is shown in figure 5.

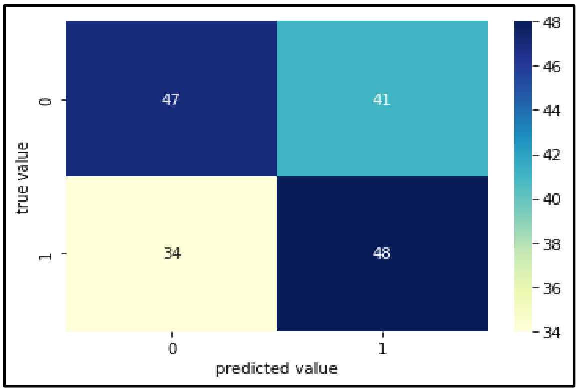


**Figure 5:** Confusion matrix

### 4.2 Validation of the Proposed Methodology

The strength of the proposed approach is its practical applicability and ability to detect potholes in real-time. In order to test the effectiveness of the prototype and the proposed approach, it is tested for 25 different instances for identifying the potholes in roads. The time taken for pothole identification is recorded and with the result obtained is shown in table 1.

**Table 1:** Summary table

| Instance no. | Hit/miss | Classified as | Object type | Time taken (s) |
|---|---|---|---|---|
| 1 | Hit | Normal | Clear road | 0.12 |
| 2 | Hit | Pothole | Pothole | 0.12 |
| 3 | Hit | Pothole | Mud puddle | 0.12 |
| 4 | Hit | Normal | Manhole | 0.11 |
| 5 | Hit | Pothole | Pothole | 0.12 |
| 6 | Miss | Normal | Manhole | 0.12 |
| 7 | Hit | Normal | Clear road | 0.12 |

| | | | | |
|---|---|---|---|---|
| 8 | Hit | Pothole | Pothole | 0.13 |
| 9 | Miss | Pothole | Manhole | 0.13 |
| 10 | Miss | Pothole | Manhole | 0.12 |
| 11 | Hit | Normal | Clear road | 0.12 |
| 12 | Hit | Normal | Clear road | 0.12 |
| 13 | Hit | Pothole | Pothole | 0.13 |
| 14 | Hit | Pothole | Pothole | 0.13 |
| 15 | Hit | Normal | Clear road | 0.13 |
| 16 | Hit | Pothole | Pothole | 0.12 |
| 17 | Hit | Pothole | Pothole | 0.12 |
| 18 | Miss | Normal | Pothole | 0.12 |
| 19 | Miss | Pothole | Manhole | 0.12 |
| 20 | Hit | Pothole | Pothole | 0.11 |
| 21 | Hit | Normal | Clear road | 0.12 |
| 22 | Hit | Pothole | Pothole | 0.12 |
| 23 | Miss | Pothole | Manhole | 0.11 |
| 24 | Miss | Pothole | Manhole | 0.13 |
| 25 | Hit | Pothole | Mud puddle | 0.13 |

## 4.3 Discussions

The points that are observed from table 1 is summarized as follows:

1. The proposed approach takes approximately 0.11 to 0.13 seconds in detecting potholes.
2. Out of the 25 instances, the proposed approach misclassified 7 instances.
3. The proposed approach failed to classify the manholes. This is due to the reason that the model is trained with no images of manholes.
4. 6 out of 7 instances, the manholes are falsely classified as potholes.

From the overall observation and discussion, it can be concluded that the proposed approach works well in detecting potholes at real-time. However due to no images trained for manholes, the proposed approach failed to identify it from the potholes.

## 5 Conclusion and Future Scope

The main aim of the present study is to develop an integrated model that is efficient as well as effective in detecting the potholes in real-time from the video feed obtained from the drones. In order to achieve the aim of the study, the proposed approach integrates image processing with the CNN. From the image processing the shape, size and texture of the images are extracted which are then fed into the CNN model to predict the classifications. The developed model showed training and testing accuracy of 0.9902 and 0.9912 respectively. For validating the CNN model it is tested for detecting the potholes for 25 instances. The model misclassified 7 instances of which 6 instances the model falsely classified manholes. This is due to the reason that the CNN model is not trained on the manholes and therefore the manholes are classified as potholes. In the future the proposed approach could be trained for detecting manholes from potholes.

## References


1. Why is India's traffic still among the worst in the world? | Travel - Hindustan Times retrieved on 12th September 2022.
2. Road Accidents in India 2020 retrieved from https://morth.nic.in/road-accident-in-india on 12th September 2022.
3. Wilhelmi, C. R., Danao, M. G. C., Gates, R. S., Zandonadi, R. S., & Hansen, A. C. (2014). Condition assessment of trucks used for soybean transport in Mato Grosso, Brazil. In 2014 Montreal, Quebec Canada July 13–July 16, 2014 (p. 1). American Society of Agricultural and Biological Engineers.
4. Kim, T., & Ryu, S. K. (2014). Review and analysis of pothole detection methods. Journal of Emerging Trends in Computing and Information Sciences, 5(8), 603-608.
5. Koch, C., & Brilakis, I. (2011). Pothole detection in asphalt pavement images. Advanced Engineering Informatics, 25(3), 507-515.
6. Tedeschi, A., & Benedetto, F. (2017). A real-time automatic pavement crack and pothole recognition system for mobile Android-based devices. Advanced Engineering Informatics, 32, 11-25.
7. Radopoulou, S. C., & Brilakis, I. (2016) Automated detection of multiple pavement defects. J Comput Civ Eng 31(2):04016057. https ://doi.org/10.1061/(asce)cp.1943-5487.00006 23
8. Nienaber, S., Booysen, M. J., & Kroon, R. S. (2015). Detecting potholes using simple image processing techniques and real-world footage.
9. Akagic, A., Buza, E., & Omanovic, S. (2017, May). Pothole detection: An efficient vision based method using rgb color space image segmentation. In 2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO) (pp. 1104-1109). IEEE.

10. Koch, C., Jog, G., & Brilakis, I. (2012). Towards automated pothole distress assessment using asphalt pavement video data. Journal of Computing in Civil Engineering, 27, 167.
11. Zhang, L., Yang, F., Zhang, Y. D., & Zhu, Y. J. (2016, September). Road crack detection using deep convolutional neural network. In 2016 IEEE international conference on image processing (ICIP) (pp. 3708-3712). IEEE.
12. Eisenbach, M., Stricker, R., Seichter, D., Amende, K., Debes, K., Sesselmann, M., ... & Gross, H. M. (2017, May). How to get pavement distress detection ready for deep learning? A systematic approach. In 2017 international joint conference on neural networks (IJCNN) (pp. 2039-2047). IEEE.
13. Feng, C., Liu, M. Y., Kao, C. C., & Lee, T. Y. (2017). Deep active learning for civil infrastructure defect detection and classification. Computing in civil engineering, 2017, 298-306.
14. Maeda, H., Sekimoto, Y., Seto, T., Kashiyama, T., & Omata, H. (2018). Road Damage Detection Using Deep Neural Networks with Images Captured Through a Smartphone, arvi preprint. arXiv preprint arxiv:1801.09454.
15. Balaji, S. R., & Karthikeyan, S. (2017, January). A survey on moving object tracking using image processing. In 2017 11th international conference on intelligent systems and control (ISCO) (pp. 469-474). IEEE.
16. Tan, L., Lv, X., Lian, X., & Wang, G. (2021). YOLOv4_Drone: UAV image target detection based on an improved YOLOv4 algorithm. Computers & Electrical Engineering, 93, 107261.
17. Liu, H. (2020). Robot systems for rail transit applications. Elsevier.
18. Kanan, C., & Cottrell, G. W. (2012). Color-to-grayscale: does the method matter in image recognition?. PloS one, 7(1), e29740.
19. Zheng, C., & Sun, D. W. (2008). Image segmentation techniques. In Computer vision technology for food quality evaluation (pp. 37-56). Academic Press.
20. Juneja, M., & Sandhu, P. S. (2009). Performance evaluation of edge detection techniques for images in spatial domain. International journal of computer theory and Engineering, 1(5), 614.
21. Rong, W., Li, Z., Zhang, W., & Sun, L. (2014, August). An improved CANNY edge detection algorithm. In 2014 IEEE international conference on mechatronics and automation (pp. 577-582). IEEE.
22. Gong, X. Y., Su, H., Xu, D., Zhang, Z. T., Shen, F., & Yang, H. B. (2018). An overview of contour detection approaches. International Journal of Automation and Computing, 15(6), 656-672.
23. Yu, J., Fan, Y., Yang, J., Xu, N., Wang, Z., Wang, X., & Huang, T. (2018). Wide activation for efficient and accurate image super-resolution. arXiv preprint arXiv:1808.08718.
24. Gu, S., Pednekar, M., & Slater, R. (2019). Improve image classification using data augmentation and neural networks. SMU Data Science Review, 2(2), 1.