Haploid Diploid Maize Seeds Classification Using Residual Network

Wahyudi Setiawan[1](✉) and Yoga Dwitya Pramudita[2]

[1] Department of Information System, University of Trunojoyo Madura, Bangkalan, Indonesia
wsetiawan@trunojoyo.ac.id
[2] Department of Informatics, University of Trunojoyo Madura, Bangkalan, Indonesia
yoga@trunojoyo.ac.id

Abstract. Maize seed breeding is an important basis for getting better production. Maize seeds consist of two types: diploid and haploid. Haploid seed can accelerate maize breeding results in just two to three generations. In contrast to diploid (normal) which requires up to eight generations. In this article, we discuss about the classification of haploid-diploid seeds. The dataset uses rovile public data with a total number of 3,000 images. Training data consists of 2,400 images, the rest is testing data. The data consists of 1,230 haploid and 1,770 diploids. The experiment contained of preprocessing, feature extraction, and classification. Preprocessing using closing morphology. While feature extraction and classification using ResNet50. As a comparison, this study also used VGG16, and MobileNet. The parameters during the training process use epoch 50, batch-size 10, learning rate 0.0001, and Root Means Square Propagation Optimization. The experimental results showed accuracy using ResNet50, VGG16, and MobileNet at 98.16%, 97.83%, and 97.83%, respectively.

Keywords: Maize Leaves Diseases · Image Classification · Convolutional Neural Network · Residual Network

1 Introduction

Maize is the natural resource commodities that are the mainstay of agriculture in Indonesia besides rice. The maize needs in Indonesia has increased every year. This is due to the large demand for maize in line with the growth of the population. Also, the need for raw materials for the processed food industry, both as food and animal feed ingredients. Based on the coordinating ministry for economic affairs in 2022, maize acts as a staple food with domestic needs of 15 million tons [1]. The average productivity of maize nationally in 2020 reached 54.74 quintal/hectare according to data from the Badan Pusat Statistic (BPS) [2]. Therefore, new innovations and strategies in farming are needed to improve and meet the daily needs. One way that can be used as an option to overcome this is by hybridizing [3]. The technique of hybridization is the process of cross-pollination between different elders of genetic composition [4]. This technique shortens the breeding period, increases breeding efficiency, results in better and predictable hybrid maize

I. H. Agustin (Ed.): ICONNSMAL 2022, AISR 177, pp. 49–59, 2023.
https://doi.org/10.2991/978-94-6463-174-6_6

seed varieties [5]. One of the important processes is the selection of haploid and diploid seeds. The facts say that selecting seeds is still carried out manually, in such a way it can reduce the success rate, time and energy unefficient. For this reason, it is necessary to develop an automatic selection that will save time, effort and increase the success of selection [6].

A. Haploid and Diploid

Haploid and diploid are actually the name for the chromosomes in maize seeds. A haploid chromosome (n) is a cell that has only one set of chromosomes, while a diploid chromosome (2n) is a homologous chromosome that is paired or has two chromosomes in a set. The body of chromosomes or autosomes are diploid in nature, while genital chromosomes or gonosomes have haploid. The haploid (n) has 10 chromosomes on (n), while the diploid (2n) has 20 chromosomes and has 30 chromosomes inside the endosperm (triploid) cell. In general, maize has 10 pairs of chromosomes.

Cells with haploid (n) will divide meiosis two times, while cells with diploid chromosomes will divide to multiply mitosis. Cleavage on the haploid chromosome (n) does not produce an identic sapling with its parent due to crossbreeding in the process. Meanwhile, the division of diploid chromosomes (2n) will produce genes that are identic to their parents. The function of the haploid (n) is used for sexual reproduction and genetic diversity in living beings. While the function of the diploid (2n) is used for the growth and development of living things. Haploid seeds are engineered that can accelerate maize breeding results in just two to three generations. In contrast to diploid (normal) which requires up to eight generations. The difference between haploid and diploid views is shown in Fig. 1 [7].

B. Related Research

In 2018, a study was conducted on the identification of haploid maize seeds by Wang et al. [8]. The study used Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) methods. The accuracy result obtained is 93%.

Altuntaş et al., 2018 [6], have conducted research on the classification of haploid and diploid maize seeds. His research uses the Support Vector Machine (SVM). The accuracy result for haploid is 94.25%, diploid is 77.91% and the total overall accuracy is 81.36%

Veeramani, et al., 2018 [5], conducted a study to haploid maize seeds classification. It was explained that the study used Deep Convolutional Networks (Deep Sort) which is a further deepening of the Convolutional Neural Network (CNN), Support Vector Machine (SVM), random forest classifier from decision trees, and logistic regression. The accuracy results obtained were Deep Sort (96.8%) SVM (87.6%), Random Forest (84.5%), Logistic regression (77.5%).

Fig. 1. Haploid (left) and diploid (right) seeds.

In 2020, Dönmez [7] conducted a study on the classification of haploid and diploid maize seeds. The study used the k-Nearest Neighbors (k-NN), Support Vector Machine (SVM), Decision Tree using pre-trained models from CNN such as AlexNet, GoogLeNet, ResNet-18, ResNet-50, and VGG-16. The largest accuracy results achieved using the ResNet50, along with the results k-NN (87.5%), SVM (91.4%), and Decision Tree (82.5%).

Previous research used a variety of machine learning and deep learning methods, especially CNN. The results of the study have not obtained optimal performance. The best accuracy using deep sort yields 96.8%. In this study we aim to improve the accuracy of performance system. The dataset comes from the research of Altuntas et al. The study resulted in an average accuracy of 81.36% [6]. The novelties of this research include:

- Preprocessing consists of masking using closing morphology which aims to remove the background from the foreground of the image.
- The classification uses a comparison of CNN's architecture, Resnet50. For comparison the experiment also used the sequential CNN VGG16, and the CNN mobile and embedded system type CNN Mobilenet.

2 Research Method

This section describes the steps taken to classify diploid haploid seed images. The research consists of five steps, including:

- Image input 224 × 224
- Preprocessing all images with masking morphology closing
- The data is divided into two parts, training and testing with a ratio of 80:20.
- Training data processed using CNN to produce the best model.
- The best model is used to classify testing data for subsequent recognition. Data includes haploid or diploid. The system proposed is shown in Fig. 2.

A. Input Dataset

The data used as input in this research is a dataset from the Robotic Vision and Learning Research Group. The data can be downloaded on the https://www.rovile.org/datasets/haploid-and-diploid-maize-seeds-dataset/. The data totals 3,000 images consisting of 2 classes, haploid 1,230 and diploid 1,770. Image data is 224 × 224 pixels in jpg format. The image of the rovile is shown in Fig. 3.

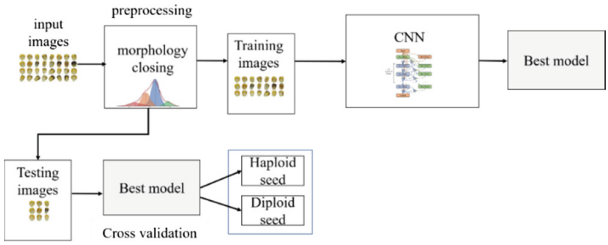

Fig. 2. Haploid Diploid Classification System Proposed

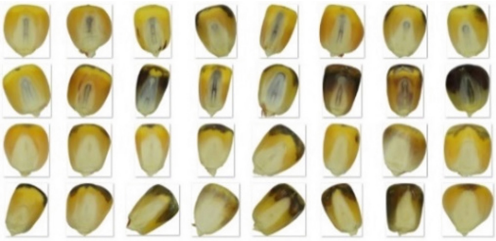

Fig. 3. Haploid Diploid Images (taken from rovile dataset).

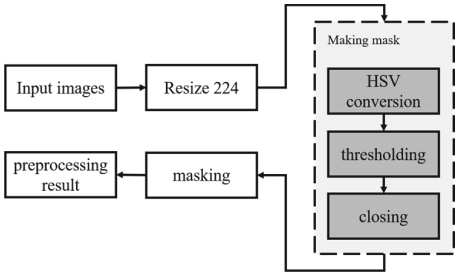

Fig. 4. Preprocessing steps

The difference between the two types of seeds is that there are color embryos in diploid, while in haploids they are colorless.

B. Preprocessing Using Masking Morphology Closing

Preprocessing in this study aims for separating the foreground from the background. Only foreground (object) maize seeds are used for the next process. Preprocessing uses masking between the original image and the image resulting from the morphology closing. The stages of preprocessing are shown in Fig. 4.

1) *HSV Conversion*

RGB colors use a scale of 0 to 255. RGB color combinations determine the color intensity of each pixel. The human eye is sensitive to RGB colors. 24-bit image file format with 8-bit RGB components each. RGB color normalization is determined in Eq. 1–3.

$$r = \frac{R}{R+G+B} \tag{1}$$

$$g = \frac{G}{R+G+B} \tag{2}$$

$$b = \frac{B}{R+G+B} \tag{3}$$

r, g, b = red, green, blue color intensity.

while Hue Saturation Value (HSV) is a color model derived from the results of non-linear transformations of primary colors. The HSV equation is shown in Eqs. 4, 5 and 6 [8].

$$H = \arccos \frac{1/2(2R - G - B)}{\sqrt{\left((R - G)^2 - (R - B) - (G - B)\right)}} \tag{4}$$

$$S = \frac{max(R, G.B) - \min(R, G, B)}{max(R, G.B)} \tag{5}$$

$$V = max(R, G.B) \tag{6}$$

2) *Thresholding*

Threshold aims to segmentation between foreground and background. Threshold uses two values, lower and upper. The lower value is set [12,70,70] while the upper value is set [106, 255, 230]. The algorithm for double threshold is as follows [9]:

- Select two RGB threshold values, lower (T1) and upper (T2).
- Partition the image into three regions, R1 for pixels below T1, R2 for pixels between T1 and T2, and R3 for pixels above T2.
- The pixels on R2 set the value to 0 or white, for R1 and R3 setting the value to 1 or black.

3) *Morphology Closing*

The purpose of morphological operations is to make improvements to the image pixels thresholding. The study used closing morphology consisting of dilation and erosion operations. Dilation is the process of merging the background points (0) into part of the object (1), based on the structuring of the S element used. Meanwhile, Erosion is the process of removing the dots of the object (1) into part of the background (0), based on the structuring of the S element used. Dilation and erosion are shown in Eqs. 7 and 8.

$$D(A,\ S) = A \oplus B \tag{7}$$

$$D(A,\ S) = A \ominus B \tag{8}$$

D = Dilation; E = Erosion; A = citra; S = structuring element.

4) *Masking*

Masking performs substruction between the original image and the image resulting from morphology closing. To tidy up the threshold results, a closing morphological operation was performed on the mask image.

After creating the mask, it is merged with the RGB image. Provided that, if the mask image value is not = 0 (value = 255), then the original image is displayed, if it is equal to 0 the image will still display a black background. The masking results are shown in Fig. 5.

C. *Split Data Training and Testing*

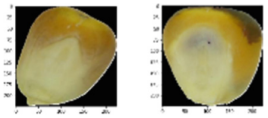

Fig. 5. Preprocessed image

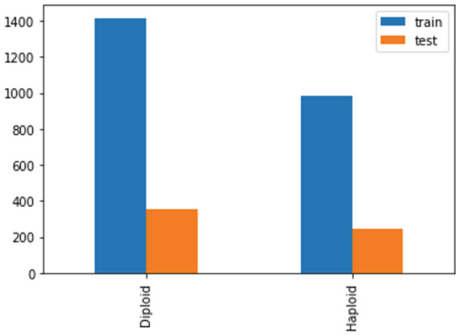

Fig. 6. Number of diploid vs haploid images

Split training and testing data with a percentage of 80 compared to 20. Split data is shown in Fig. 6.

D. *Residual Network*

Since 2012 there has been a new method in computer vision with the emergence of CNN deep learning. In fact, this method has existed since LeCun introduced LeNet in 1998, but since then there has been no further development. Since ImageNet Large Scale Visual Recognition Challenge (ILSVRC) conducted an image classification and detection competition in 2010, all research groups have been competed to be the best. In 2012 alexnet became the winner on the ILSRC beating conventional machine learning methods. Since then it has continued to rapidly evolve new CNN models [10, 11].

However, there is a tendency for researchers to add deeper layers (add layers) to solve complex problems and to improve accuracy in classification and detection. This certainly makes it difficult to do training which causes saturation and decreased accuracy at the same time. Therefore, ResNet made improvements to this problem.

Resnet was introduced by Kaiming He et al. in the paper "Deep Residual Learning for Image Recognition". ResNet won ILSVRC 2015 with an error rate of 3.57% for classification and detection. It also featured COCO detection 2012 [12–14].

This very deep network training problem has been addressed with the introduction of ResNet or residual networks and this Resnet is created from Residual Blocks. There is a direct connection that skips several layers in between. This connection is called a "skip connection". This is the core of the residual block. Since the connection passes this, the output of the layer is not the same now. Without using this skip connection, the input is multiplied by the weight of the layer followed by adding bias [13]. In this study, Resnet-50 was used. The ResNet50 architecture is shown in Fig. 7 [15–17].

E. *Confusion Matrix*

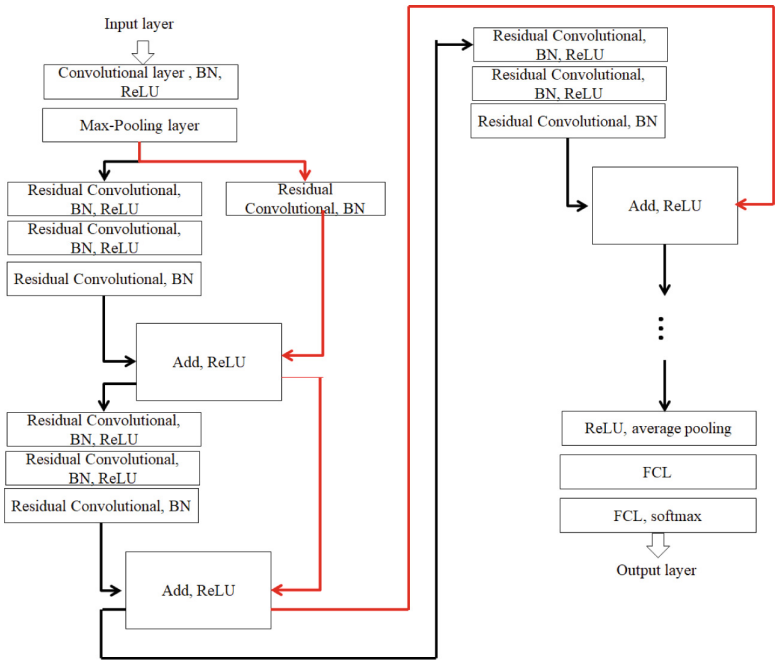

Fig. 7. ResNet50 Architecture. Description: BN = Batch Normalization, FCL = Fully Connected Network, the red connecting line is a *skip connection.* There are 16 skip connections on resNet50

In the evaluation step, confusion matrix is used as a solution to calculate the predicted accuracy value resulting from the training and validation process. The selection of matrix confusion is used to calculate accuracy, precision, and recall. The values based on the information provided in the identification process including, False Positives (FP), False Negatives (FN), True Positives (TP), and True Negatives (TN). A binary confusion matrix is shown in Fig. 8 [18].

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{9}$$

$$Precision = \frac{TP}{TP + FP} \tag{10}$$

$$Recall = \frac{TP}{TP + FN} \tag{11}$$

3 Result and Discussion

Testing was conducted on hardware with specifications of processor Intel (R) Core (TM) i5-9300H CPU@ 2.40 GHz (8 CPUs), ~ 2.4 GHz, Memory 8 GB RAM 1TB HDD, GPU NVIDIA GEFORCE GTX 1050 (2 GB). While the software needs are Python library openCV, numpy, and matplotlib. The components or modules needed include scikit-learn 0.24.1, and keras 2.7.1

Total Observations (n)		Predicted label	
		diploid	haploid
Actual label	diploid	True Positive (TP)	False Positive (FP)
	haploid	False Negative (FN)	True Negative (TN)

Fig. 8. Confusion matrix

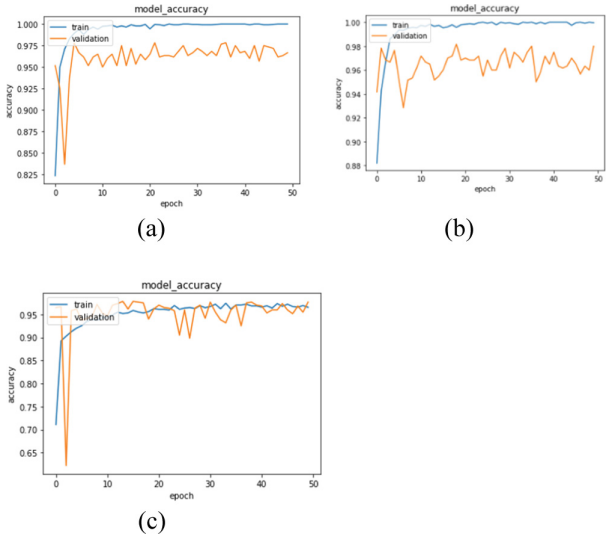

Fig. 9. Graph the test results using (a) ResNet50, (b) VGG16, and (c) mobile net

The experiment was conducted using three different CNN architectures, CNN with skip connection ResNet50, CNN sequential VGG16, and CNN depthwise separable mobilenet. The parameters during the training process use epoch 50, batch-size 10, learning rate 0.0001, and Root Means Square Propagation Optimization. The results of the experiment are shown in Fig. 9 (a), (b), and (c). The Figure is a training and validation graph. The results showed the accuracy of ResNet50, VGG16, and MobileNet which were 98.16%, 97.83%, and 97.83% respectively. In addition, Figs. 10 (a), (b), (c) are the matrix of the three CNN architectures. Meanwhile, Table 1 shows the accuracy, precision, and recall performance of each of the CNN architectures tested.

The test results showed that ResNet50 produced better accuracy compared to others. ResNet50 is a CNN architecture that has 48 convolutional layers, one max-pooling layer and one average pooling layer. Resnet solved the vanishing gradient problem on CNN sequential. These problems can cause a continuous decrease in recognition performance. For this reason, the concept of "skip connection" is used on ResNet. Skip Connection is a direct connection through several layers of the model [19].

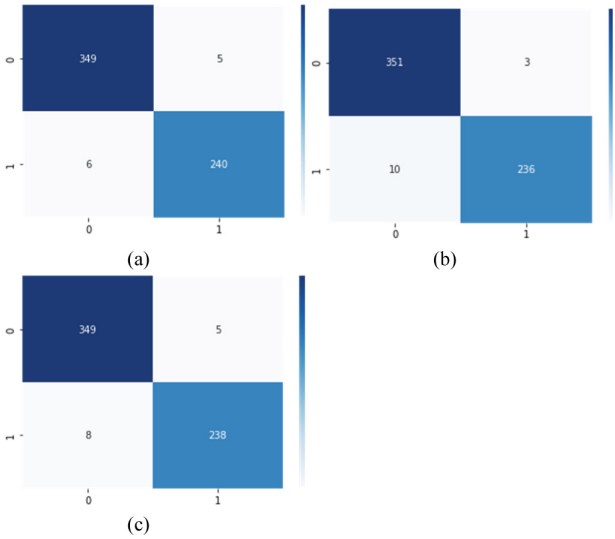

Fig. 10. Confusion matrix using (a) ResNet50, (b) VGG16, and (c) mobile net.

Table 1. Performance Result

CNN	Accuracy (%)	Precision (%)	Recall (%)
ResNet50	98.16	98.59	98.31
VGG16	97.83	99.15	97.23
MobileNet	97.83	98.59	97.76

The second experiment used VGG16. This architecture is a successor of alexnet. It is one of the sequential CNN architectures. VGG16 consists of 13 convolution layers, and three Fully Connected Layers. In addition, there are five max-pooling layers. The kernel filter on VGG uses only 3×3. VGG is from the Visual Gometry Group at Oxford University. This architecture is known to be reliable for vision recognition, it's just that the parameters owned are too large to more than 143 million. It requires a large amount of resources to process it [20].

While the third CNN architecture for the experiment is MobileNet. This architecture is a type of CNN created for the benefit of mobile and embedded vision applications. Embedded vision is an application made on hardware for control and automation purposes. Mobilenet uses depthwise separable convolution to build lightweight CNN on mobile and embedded devices [21].

4 Conclusion

Classification of haploid and diploid maize seeds has been carried out. The research phase consists of preprocessing, feature extraction, and classification using CNN. Preprocessing using masking of the original image with the image of the closing morphology. Furthermore, the feature extraction and classification process use CNN ResNet50. For comparison used CNN Sequential VGG16 and CNN mobilenet architectures. The results showed the best performance on ResNet50 with accuracy, precision, and recall of 98.16%, 98.59%, and 98.31%, respectively. To optimize performance, tuning parameters can be added in the next study.

Acknowledgment. This research is one of the outputs of penelitian kolaborasi nasional, Penelitian Mandiri, University of Trunojoyo Madura No 227/UN46.4.1/PT01.03/2022.

References

1. S. Moegiarso, "Pemerintah Dorong Peningkatan Produksi Jagung Nasional, Melalui Intensifikasi dan Ekstensifikasi , Khususnya Perluasan Lahan Baru , Untuk Memenuhi Kebutuhan Nasional dan Ekspor," Kementerian Koordinator Bidang Perekonomian Republik Indonesia, 2022. https://www.ekon.go.id/ publikasi/detail/ 4403/pemerintahdorong-peningkatan-produksi-jagung-nasional-melalui-intensifikasi-dan- ekstensifikasi-khususnya-perluasan-lahan-baru-untuk-memenuhi- kebutuhan-nasional-dan-ekspor (accessed Nov. 25, 2022).
2. K. Astuti, R. O. Prasetyo, and N. . Khasanah, Analysis of Maize and Soybean Productivity in Indonesia 2020 (the Result of Crop Cutting Survey), vol. 05100.2103, no. 5203029. 2021.
3. W. Liao, X. Wang, D. An, and Y. Wei, "Hyperspectral imaging technology and transfer learning utilized in haploid maize seeds identification," 2019 Int. Conf. High Perform. Big Data Intell. Syst. HPBD IS 2019, pp. 157–162, 2019, doi: https://doi.org/10.1109/HPBDIS.2019.8735457.
4. R. Yunianti, S. Sujiprihati, and Syukur. Muhammad, "Teknik Persilangan Buatan," Bogor, 2010. [Online]. Available: https://repository.ipb.ac.id/jspui/bitstream/123456789/60268/1/PRO2010_RYU.pdf.
5. B. Veeramani, J. W. Raymond, and P. Chanda, "DeepSort: Deep convolutional networks for sorting haploid maize seeds," BMC Bioinformatics, vol. 19, no. Suppl 9, pp. 1–9, 2018, doi: https://doi.org/10.1186/s12859-018-2267-2.
6. Y. Altuntaş, A. F. Kocamaz, R. Cengiz, and M. Esmeray, "Haploid ve Diploid Mısır Tohumlarının Görüntü İşleme Teknikleri ve Destek Vektör Makineleri Kullanılarak Sınıflandırılması," 2018 26th Signal Process. Commun. Appl. Conf., pp. 1–4, 2018.
7. J. Taylor, C. P. Chiou, and L. J. Bond, "A methodology for sorting haploid and diploid corn seed using terahertz time domain spectroscopy and machine learning," AIP Conf. Proc., vol. 2102, no. May, 2019, doi: https://doi.org/10.1063/1.5099809.
8. P. Ganesan and V. Rajini, "Assessment of satellite image segmentation in RGB and HSV color space using image quality measures," 2014 Int. Conf. Adv. Electr. Eng. ICAEE 2014, 2014, doi: https://doi.org/10.1109/ICAEE.2014.6838441.
9. H. Ye and S. Yan, "Double Threshold Image Segmentation algorithm based on adaptive filtering," in IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), 2017, pp. 1008–1011.

10. J. Gu et al., "Recent advances in convolutional neural networks," Pattern Recognit., vol. 77, pp. 354–377, 2018, doi: https://doi.org/10.1016/j.patcog.2017.10.013.
11. G. Huang and K. Q. Weinberger, "Densely Connected Convolutional Networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 4700–4708, [Online]. Available: https://arxiv.org/pdf/1608.06993.pdf.
12. K. He, X. Zhang, S. Ren, and J. Sun, "Delving Deep into Rectifiers : Surpassing Human-Level Performance on ImageNet Classification," 2015.
13. K. He, X. ZHang, S. Ren, and J. Sun, "Deep residual learning for image steganalysis," Multimed. Tools Appl., pp. 1–17, 2015, doi: https://doi.org/10.1007/s11042-017-4440-4.
14. M. F. Haque, H. Y. Lim, and D. S. Kang, "Object detection based on VGG with ResNet network," ICEIC 2019 - Int. Conf. Electron. Information, Commun., no. January 2019, 2019, doi: https://doi.org/10.23919/ELINFOCOM.2019.8706476.
15. G. G. A. Celano, "A ResNet-50-based Convolutional Neural Network Model for Language ID Identification from Speech Recordings," SIGTYP 2021 - 3rd Work. Res. Comput. Typology Multiling. NLP, Proc. Work., pp. 136–144, 2021, doi: https://doi.org/10.18653/v1/2021.sigtyp-1.13.
16. H. Mujtaba, "Introduction to Resnet or Residual Network," 2020. [Online]. Available: https://www.mygreatlearning.com /blog/resnet/.
17. A. Deshpande, V. V. Estrela, and P. Patavardhan, "The DCT-CNN-ResNet50 architecture to classify brain tumors with super-resolution, convolutional neural network, and the ResNet50," Neurosci. Informatics, vol. 1, no. 4, p. 100013, 2021, doi: https://doi.org/10.1016/j.neuri.2021.100013.
18. I. Markoulidakis, G. Kopsiaftis, I. Rallis, and I. Georgoulas, "Multi-Class Confusion Matrix Reduction method and its application on Net Promoter Score classification problem," ACM Int. Conf. Proceeding Ser., pp. 412–419, 2021, doi: https://doi.org/10.1145/3453892.3461323.
19. K. He and J. Sun, "Deep Residual Learning for Image Recognition," IEEE Xplore, pp. 1–9, 2015.
20. K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," in ICLR, 2015, pp. 1–14, doi: https://doi.org/10.1016/j.infsof.2008.09.005.
21. A. G. Howard et al., "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," 2017, [Online]. Available: http://arxiv.org/abs/1704.04861.