



# Classification bootstrap resnet50 for potato diseases datasets

Hasanal Fachri Satia Simbolon

Sistem dan Teknologi Informasi, Institut Teknologi Sawit Indonesia, Medan, Indonesia

## ARTICLE INFO

### Article history:

Received Dec 21, 2023

Revised Jan 11, 2024

Accepted Jan 12, 2024

### Keywords:

Accuracy;  
Bootstrap;  
Classification;  
Potato diseases;  
ResNet50.

## ABSTRACT

Potato diseases can have negative impacts on crop production, food security, income generation, trade competitiveness, consumer health, environmental sustainability, etc. Therefore, it is important to prevent, monitor, diagnose, treat, and control potato diseases using various methods such as cultural practices (e.g., crop rotation, intercropping), biological control (e.g., beneficial insects), chemical control (e.g., fungicides), physical control (e.g., pruning), etc. Classification Bootstrap ResNet50 is a topic related to image classification using a deep learning model called ResNet50. ResNet50 is a convolutional neural network that has 50 layers and can classify images into 1000 object categories, such as animals, plants, vehicles, etc. ResNet50 is based on the concept of residual learning, which means that it uses skip connections to avoid the problem of vanishing gradients and improve the performance of the network. We used classification bootstrap and ResNet50 to analyze a dataset of potato images with different diseases. We compared the performance of ResNet50 with other models, such as Inception-v3 and Xception, and used bootstrap to evaluate the accuracy and variability of the models. We also used data augmentation and dropout techniques to optimize the models and prevent overfitting. We define some functions to run training by doing Early Stop, train\_steps, val\_steps, predictions. In the research I conducted, I got the highest accuracy results, reaching 100% accuracy by testing epochs with a total of 0-41 epochs.

*This is an open access article under the [CC BY-NC](https://creativecommons.org/licenses/by-nc/4.0/) license.*



### Corresponding Author:

Hasanal Fachri Satia Simbolon,  
Sistem dan Teknologi Informasi,  
Institut Teknologi Sawit Indonesia,  
Jalan Willem Iskandar, Medan Estate, Kota Medan, North Sumatera, 20226, Indonesia.  
Email: [hasanal.fachri12@gmail.com](mailto:hasanal.fachri12@gmail.com)

## 1. INTRODUCTION

Classification is a process of categorizing data or objects into predefined classes or categories based on their features or attributes. It is a type of supervised learning technique where an algorithm is trained on a labeled dataset to predict the class or category of new, unseen data. The main objective of classification is to build a model that can accurately assign a label or category to a new observation based on its features. (jasim, 2020).

There are different types of classification problems, such as binary classification, multiclass classification, and multi-label classification. Binary classification involves predicting whether an input belongs to one of two classes, such as spam or not spam.

Multiclass classification involves predicting which class an input belongs to among several classes, such as flower species. Multi-label classification involves predicting which classes an input belongs to among multiple labels, such as movie genres. (Krizhevsky, 2012).

There are also different types of classifiers that can be used for classification problems, such as linear classifiers, non-linear classifiers, and ensemble classifiers. Linear classifiers create a linear decision boundary between classes, such as logistic regression and support vector machines. Non-linear classifiers create a non-linear decision boundary between classes, such as k-nearest neighbors and naive Bayes. Ensemble classifiers combine multiple base classifiers to improve the overall performance, such as random forests and neural networks.

ResNet-50 is a convolutional neural network that is 50 layers deep. It is a variant of the ResNet model, which stands for Residual Network. ResNet models use shortcut connections to avoid the problem of vanishing or exploding gradients when training very deep networks. ResNet-50 can classify images into 1000 object categories, such as keyboard, mouse, pencil, and many animals. (Li, 2023).

Classification Bootstrap ResNet50 is a topic related to image classification using a deep learning model called ResNet50. ResNet50 is a convolutional neural network that has 50 layers and can classify images into 1000 object categories, such as animals, plants, vehicles, etc. ResNet50 is based on the concept of residual learning, which means that it uses skip connections to avoid the problem of vanishing gradients and improve the performance of the network. Bootstrap is a technique that involves sampling data from different distributions and combining them to create new data. Bootstrap can be used for data augmentation, which means that it can increase the size and diversity of the training data by applying transformations such as rotation, cropping, flipping, etc. Bootstrap can also be used for transfer learning, which means that it can leverage the knowledge learned from one task or domain to another task or domain. (Mukti, 2019).

## 2. RESEARCH METHOD

Bootstrap is a way of estimating the uncertainty or variability of a statistic or a model based on resampling from the original data. It is useful when the underlying distribution of the data is unknown or when the sample size is small. The bootstrap method can be applied to various problems, such as hypothesis testing, confidence interval construction, model selection, and performance evaluation.

The basic steps of the bootstrap method are:

- Choose a statistic or a model that you want to estimate or evaluate.
- Choose a number of bootstrap samples (B) and a sample size (n).
- For each bootstrap sample (b = 1, 2, ..., B):
- Draw n observations randomly with replacement from the original data.
- Calculate the statistic or fit the model on the bootstrap sample and record the result.
- Analyze the distribution of the bootstrap results and derive the desired estimates or evaluations.

There are different types of bootstrap methods, such as the standard bootstrap, the bias-corrected bootstrap, the percentile bootstrap, the bootstrap-t, the parametric bootstrap, the smooth bootstrap, and the wild bootstrap. Each type has its own advantages and disadvantages depending on the situation and the goal.

Some of the benefits of the bootstrap method are:

- It is easy to implement and understand.
- It does not require any assumptions about the data distribution or the population parameters.
- It can handle complex statistics or models that do not have analytical formulas or solutions.
- It can provide accurate and robust estimates or evaluations in many cases.

Some of the limitations of the bootstrap method are:

- It can be computationally intensive and time-consuming, especially for large datasets or complex models.
- It can be sensitive to outliers or extreme values in the data.
- It can be unreliable or misleading when the data is not representative of the population or when the sample size is too small.
- It can have difficulties with discrete or dependent data or with complex structures or dependencies in the data.

ResNet50 is a way of designing and training a deep convolutional neural network for image recognition and classification tasks. ResNet50 is a variant of ResNet, which stands for Residual Network. ResNet is a network architecture that uses shortcut connections to avoid the problem of vanishing or exploding gradients when training very deep networks. ResNet50 has 50 layers, including 48 convolutional layers, one max pooling layer, and one average pooling layer. ResNet50 uses a bottleneck design for the building blocks, which reduces the number of parameters and matrix multiplications. ResNet50 can classify images into 1000 object categories, such as keyboard, mouse, pencil, and many animals.

Some of the steps of the research method about ResNet50 are:

- Choose a dataset of images that you want to classify, such as ImageNet, CIFAR-10, or your own custom dataset.
- Choose a pre-trained ResNet50 model or train your own ResNet50 model from scratch using a deep learning framework, such as TensorFlow, PyTorch, or Keras.
- Fine-tune the ResNet50 model on your dataset by adjusting the hyperparameters, such as learning rate, batch size, number of epochs, etc. You can also modify the network structure, such as adding or removing layers, changing the activation functions, etc.
- Evaluate the performance of the ResNet50 model on your dataset by using metrics, such as accuracy, precision, recall, F1-score, etc. You can also compare the ResNet50 model with other models, such as VGG, Inception, or AlexNet.
- Analyze the results and draw conclusions about the ResNet50 model and its applications.

Potato disease is a term that refers to any disorder or condition that affects the growth, development, or quality of potato plants or tubers. Potato diseases can be caused by various agents, such as bacteria, fungi, viruses, nematodes, insects, and weeds. Some of the common potato diseases are:

- Bacterial wilt: This is a bacterial disease that infects the vascular system of potato plants and causes them to wilt and die. The symptoms include yellowing and browning of leaves, stunting and wilting of stems, and black or brown rings on tubers. The disease can spread through infected seeds, soil, water, tools, or animals.
- Septoria leaf spot: This is a fungal disease that causes small to irregular spots with a grey center and dark margin on the leaves of potato plants. The spots usually start on the lower leaves and move upwards. The disease can reduce the yield and quality of potatoes by affecting photosynthesis and nutrient uptake.
- Late blight: This is a fungal disease that causes blistering and scalding of leaves on potato plants. The affected leaves turn brown or black in color and eventually rot and dry out. The disease can also cause spots on the underside of leaves covered with a white powder-like substance. The stems and tubers can also be infected by the fungus.
- Early blight: This is another fungal disease that causes dark brown to black lesions on the leaves of potato plants. The lesions are circular or irregular in shape and may have yellow halos around them. The disease can also cause wilting and dieback of plants.
- Common scab: This is a fungal disease that causes small to large pustules on the surface of potato tubers. The pustules are filled with white spores that can infect new tubers when they are harvested or stored. The disease can reduce the marketability and shelf life of potatoes.

- Black scurf/canker: This is a fungal disease that causes dark brown to black lesions on the stems and roots of potato plants. The lesions may have sunken centers with yellow halos around them. The disease can also cause wilting and death of plants.
- Viral diseases: These are viral diseases that infect different parts of potato plants or tubers. Some examples are potato virus X (PVX), which causes mosaic patterns on leaves; potato virus Y (PVY), which causes yellowing between veins; potato spindle tuber viroid (PSTVd), which causes abnormal growths called spindle tubers; black leg (BLV), which causes necrosis in vascular tissues; soft rot (SR), which causes softening in tubers; pink rot (PR), which causes discoloration in tubers; black heart-disorder (BHD), which causes heart-shaped deformities in tubers; etc.

### 3. RESULTS AND DISCUSSIONS

In this research, tests were carried out on potatoes with lots of potato image data with various characteristics as follows:

Total Classes = 7

- Black Scurf: 58 images
- Blackleg: 60 images
- Common Scab: 62 images
- Dry Rot: 60 images
- Healthy Potatoes: 80 images
- Miscellaneous: 74 images
- Pink Rot: 57 images

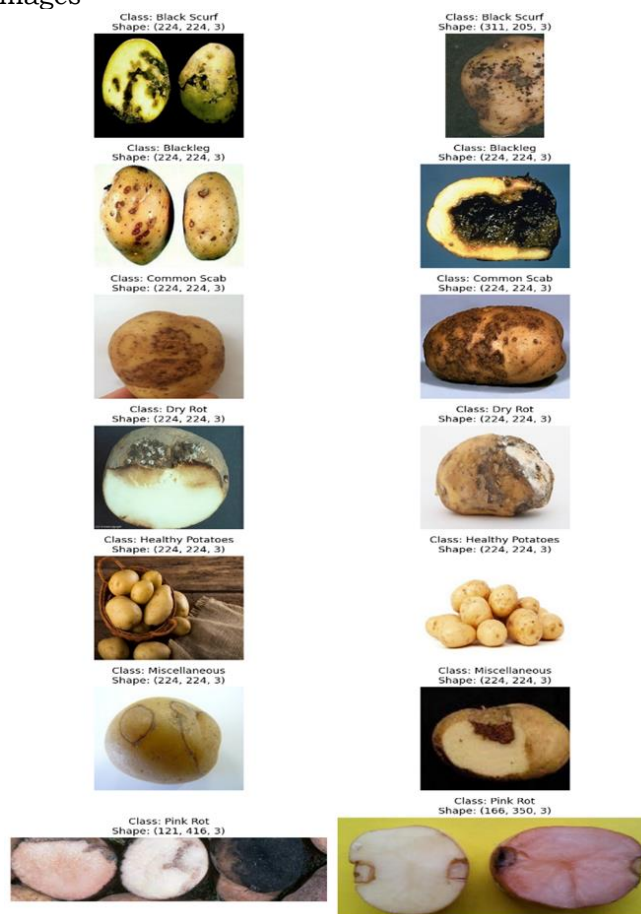


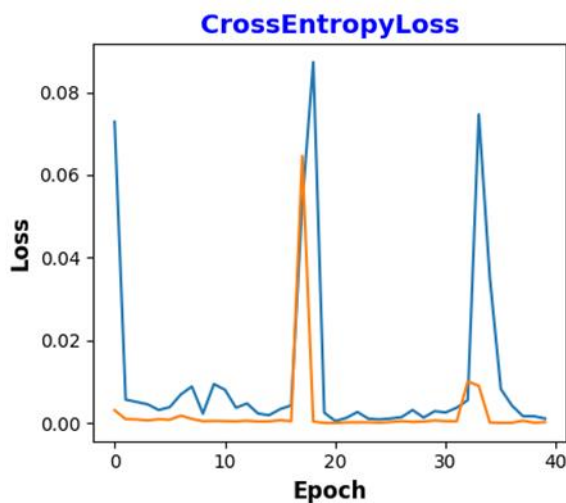
Figure 1 Potato Disease Image

In the research carried out regarding potato diseases which has been shown previously, the author will test the potato disease data with several steps, namely EarlyStopping which is a stop at the epoch which is carried out in order to limit the process so that there is no overfitting of the data, then train\_step is carried out with a value of 15% which is the testing stage in order to obtain accuracy values and cross entropy values. Next, carry out the val\_step test with a test value of 15% of the data in order to get validation results for accuracy and loss and next is a prediction of the resulting confusion matrix value.

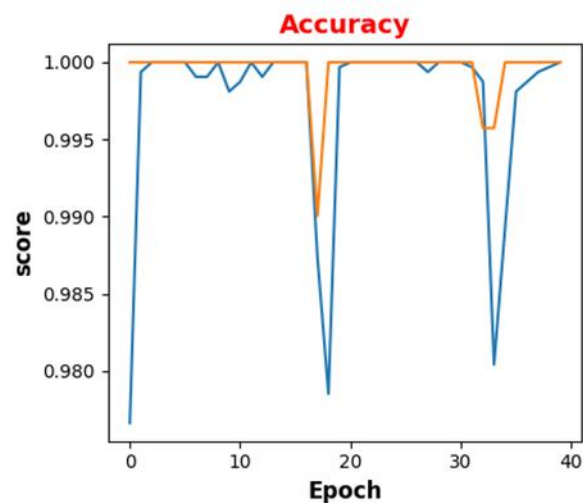
Now we take 70% of our new dataframe for training, 15% validation and 15% for testing.

```
Epoch: 1 | Train Loss: 0.0728 | Train Accuracy: 0.9766 | Val Loss: 0.0032 | Val Accuracy: 1.0000
Epoch: 2 | Train Loss: 0.0057 | Train Accuracy: 0.9994 | Val Loss: 0.0010 | Val Accuracy: 1.0000
Epoch: 3 | Train Loss: 0.0052 | Train Accuracy: 1.0000 | Val Loss: 0.0009 | Val Accuracy: 1.0000
Epoch: 4 | Train Loss: 0.0046 | Train Accuracy: 1.0000 | Val Loss: 0.0007 | Val Accuracy: 1.0000
Epoch: 5 | Train Loss: 0.0032 | Train Accuracy: 1.0000 | Val Loss: 0.0010 | Val Accuracy: 1.0000
Epoch: 6 | Train Loss: 0.0039 | Train Accuracy: 1.0000 | Val Loss: 0.0009 | Val Accuracy: 1.0000
Epoch: 7 | Train Loss: 0.0070 | Train Accuracy: 0.9991 | Val Loss: 0.0019 | Val Accuracy: 1.0000
...
Epoch: 38 | Train Loss: 0.0017 | Train Accuracy: 0.9994 | Val Loss: 0.0006 | Val Accuracy: 1.0000
Epoch: 39 | Train Loss: 0.0017 | Train Accuracy: 0.9997 | Val Loss: 0.0001 | Val Accuracy: 1.0000
Epoch: 40 | Train Loss: 0.0011 | Train Accuracy: 1.0000 | Val Loss: 0.0003 | Val Accuracy: 1.0000
Epoch: 41 | Train Loss: 0.0017 | Train Accuracy: 1.0000 | Val Loss: 0.0001 | Val Accuracy: 1.0000
```

We define a function to visualize the evolution of metrics and loss during training. The crossentropy graph decreases as the epoch value gets higher and the accuracy graph displays high accuracy results at epochs between 17 and 18. The following is a graphic display based on the epoch value in the training data based on crossentropy loss and accuracy:



Figures 2. Cross Entropy Loss



Figures 3. Accuracy

The confusion matrix test graph below shows the true predicted results produced on potato disease data, which shows that healthy potatoes have a greater value because the existing data shows that the potatoes studied have healthy potato elements. The following is a confusion matrix graph on potato disease data:

**Confusion Matrix Test**

True	Black Scurf	90	0	0	0	0	0	
	Blackleg	0	75	0	0	0	0	
	Common Scab	0	0	107	0	0	0	
	Dry Rot	0	0	0	96	0	0	
	Healthy Potatoes	0	0	0	0	116	0	
	Miscellaneous	0	0	0	0	0	105	
	Pink Rot	0	0	0	0	0	88	
		Black Scurf	Blackleg	Common Scab	Dry Rot	Healthy Potatoes	Miscellaneous	Pink Rot
		Predicted						

Figures 4. Confusion matrix

#### 4. CONCLUSION

Classification is a process of categorizing data or objects into predefined classes or categories based on their features or attributes. It is a type of supervised learning technique where an algorithm is trained on a labeled dataset to predict the class or category of new, unseen data. Classification can be used for various applications, such as plant disease recognition and diagnosis. Bootstrap is a way of estimating the uncertainty or variability of a statistic or a model based on resampling from the original data. It is useful when the underlying distribution of the data is unknown or when the sample size is small. Bootstrap can be applied to various problems, such as hypothesis testing, confidence interval construction, model selection, and performance evaluation. ResNet50 is a convolutional neural network that is 50 layers deep. It is a variant of ResNet, which stands for Residual Network. ResNet models use shortcut connections to avoid the problem of vanishing or exploding gradients when training very deep networks. Potato disease is a term that refers to any disorder or condition that affects the growth, development, or quality of potato plants or tubers. Potato diseases can be caused by various agents, such as bacteria, fungi, viruses, nematodes, insects, and weeds. Some of the common potato diseases are bacterial wilt, septoria leaf spot, late blight, early blight, common scab, and black scurf/canker. In this study, we used classification bootstrap and ResNet50 to analyze a dataset of potato images with different diseases. We compared the performance of ResNet50 with other models, such as Inception-v3 and Xception, and used bootstrap to evaluate the accuracy and variability of the models. We also used data augmentation and dropout techniques to optimize the models and prevent overfitting. Our results showed that ResNet50 achieved the highest accuracy and lowest error rate for detecting potato diseases, followed by Inception-v3 and Xception. We also found that bootstrap provided a reliable estimate of the uncertainty and variability of the models. We concluded that ResNet50 is a suitable model for potato disease classification and that bootstrap is a useful method for model evaluation. Our suggestion is that researchers can continue by combining other methods to obtain This research produces types of potatoes that have disease by looking at the characteristics and characteristics based on the pictures that have been made so that consumers cannot consume potatoes with the potato characteristics that have been mentioned, so it can help farmers to better sort out types of potatoes that cannot be used. consumed by humans

because of disease. Our suggestion is that researchers can continue by combining other methods to obtain higher accuracy and use palm disease data to develop in further research.

#### REFERENCES

- Al Gaashani MS, Shang F, Abd El-Latif AA. (2023). Ensemble learning of lightweight deep convolutional neural networks for crop disease image detection. *J Circuits Syst Comput* 32(05):2350086. <https://doi.org/10.1142/S021812662350086X>
- Arshad, M.S., Rehman, U.A., Fraz, M.M. (2021). Plant disease identification using transfer learning. In: 2021 International Conference on Digital Futures and Transformative Technologies (ICoDT2), pp. 1–5. IEEE.
- Bhakta I, Phadikar S, Majumder K et al. (2023). A novel plant disease prediction model based on thermal images using modified deep convolutional neural network. *Precision Agric* 24(1):23–39
- Chen, J. Zhang, D. Nanekharan, Y. A. (2020). Identifying plant diseases using deep transfer learning and enhanced lightweight network. <https://doi.org/10.1007/s11042-020-09669-w>
- Dasgupta, S. R. Rakshit, S. Mondal, D. Kole, D. K. (2020). Detection of diseases in potato leaves using transfer learning. *Computational Intelligence in Pattern Recognition* Springer Singapore
- Fauzan Muhammad., Aniati Murni Arimurthy., Dina Chahyati. (2023). Transfer Learning dengan Metode Fine Tuning pada Model Network VGG16 dan ResNet50. *Indonesian Journal of Computer Science*.
- Islam, F., Hoq, M.N., Rahman, C.M. (2019). Application of transfer learning to detect potato disease from leaf image. In: IEEE International Conference on Robotics, Automation, Artificial-Intelligence and Internet-of-Things (RAAICON), pp. 127–130. IEEE.
- Jasim, M.A., Al-Tuwaijari, J.M. (2020). Plant leaf diseases detection and classification using image processing and deep learning techniques. In: 2020 International Conference on Computer Science and Software Engineering (CSASE), pp. 259–265. IEEE.
- Jin W, Yane Li, Hailin F, Ren Lijin Du, Jian XW. (2020). Common pests image recognition based on deep convolutional neural network. *Comput Electron Agric* 179(1):105834. <https://doi.org/10.1016/j.compag.2020.105834>
- Kashyap CK, Rashmi M, Chandan C, Kangkana B. (2021). Automated recognition of optical image based potato leaf blight diseases using deep learning - ScienceDirect. *Physiol Mol Plant Pathol* 117:101781. <https://doi.org/10.1016/J.PMPP.2021.101781>
- Nashrullah, Faiz., Wibowo, Suryo Adhi., Budiman, Gelar. (2020). Investigasi Parameter Epoch Pada Arsitektur ResNet50 Untuk Klasifikasi Pornografi. *Journal of Computer, Electronic, and Telecommunication* ISSN 2723-4371, E-ISSN 2723-5912 [doi.org/10.52435/complete.v1i1.51](https://doi.org/10.52435/complete.v1i1.51)
- Li, Weiming, Siqi Yu, Runhuang Yang, Yixing Tian, Tianyu Zhu, Haotian Liu, Danyang Jiao, Feng Zhang, Xiangtong Liu, Lixin Tao, and et al. (2023). Machine Learning Model of ResNet50-Ensemble Voting for Malignant-Benign Small Pulmonary Nodule Classification on Computed Tomography Images. *Cancers* 15, no. 22: 5417. <https://doi.org/10.3390/cancers15225417>
- Liu, Y., Chen, X., Wang, Z., Wang, Z.J., Ward, R.K., Wang, X. (2018). Deep learning for pixel-level image fusion: recent advances and future prospects. *Information Fusion* 42, 158–173.
- Liyuan Wang., Yulong Chen., Xiaoye Wang., Ruixing Wang., Hao Chen., Yin Hai Zhu. (2023). Research on Remote Sensing Image Classification Based on Transfer Learning and Data Augmentation. *International Conference on Knowledge Science, Engineering and Management KSEM 2023: Knowledge Science, Engineering and Management* pp 99–111
- Lv, Z., Hu, Y., Zhong, H., Wu, J., Li, B., Zhao, H. (2018). Parallel K-means clustering of remote sensing images based on mapReduce. In: Wang, F.L., Gong, Z., Luo, X., Lei, J. (eds.) *WISM. LNCS*, vol. 6318, pp. 162–170. Springer, Heidelberg. [https://doi.org/10.1007/978-3-642-16515-3\\_21](https://doi.org/10.1007/978-3-642-16515-3_21)
- Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., Johnson, B.A. (2019). Deep learning in remote sensing applications: a meta-analysis and review. *ISPRS J. Photogramm. Remote. Sens.* 152, 166–177.
- Mora-Romero, G.A., Félix-Gastélum, R., Bomberger, R.A. et al. (2022). Common potato disease symptoms: ambiguity of symptom-based identification of causal pathogens and value of on-site molecular diagnostics. *J Gen Plant Pathol* 88, 89–104. <https://doi.org/10.1007/s10327-021-01045-2>

- Mukti, I.Z., Biswas, D. (2019). Transfer learning based plant diseases detection using ResNet50. In: 4th International Conference on Electrical Information and Communication Technology (EICT), pp. 1–6. IEEE.
- R. Ferjaoui, M. A. Cherni, F. Abidi and A. Zidi. (2022). Deep Residual Learning based on ResNet50 for COVID-19 Recognition in Lung CT Images. 8th International Conference on Control, Decision and Information Technologies (CoDIT), Istanbul, Turkey, pp. 407-412, doi: 10.1109/CoDIT55151.2022.9804094.
- R. Kumar, D. Singh, A. Chug and A. P. Singh. (2022). Evaluation of Deep learning based Resnet-50 for Plant Disease Classification with Stability Analysis. 6th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, pp. 1280-1287, doi: 10.1109/ICICCS53718.2022.9788207.
- Sanjay A. Agrawal., Vaibhav D., Rewaskar., Rucha A., Agrawal., Swapnil S. Chaudhari., Yogendra Patil., Nidhee S. Agrawal. (2023). Advancements in Nsfw Content Detection: A Comprehensive Review of Resnet-50 Based Approaches. International Journal Of intelligent Systems and Applications In Engineering.
- Shabbir, A., et al. (2021). Satellite and scene image classification based on transfer learning and fine tuning of resnet50. Math. Probl. Eng. 2021, 1–18.
- Suherman, Endang., Rahman, Ben., Hindarto, Djarot., Santoso, Handri. (2023). Implementation of ResNet-50 on End-to-End Object Detection (DETR) on Objects. Sinkron : Jurnal dan Penelitian Teknik Informatika Volume 8, Number 2.
- Tiwari, D., Ashish, M., Gangwar, N., Sharma, A., Patel, S., Bhardwaj, S. (2020). Potato leaf diseases detection using deep learning. In:4th International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 461–466. IEEE.
- Too, E. C. Yujian, L. Njuki, S. Yingchun, L. (2019). A comparative study of fine-tuning deep learning models for plant disease identification. Comput. Electron. Agric.
- Umamaheswari S; Vishal N R; Pragadesh N R; Lavanya S. (2023). Performance Analysis of ResNet50 Architecture based Pest Detection System. 9th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2023, pp. 578-583, doi: 10.1109/ICACCS57279.2023.10112802.