2023 International Conference on Next-Generation Computing, IoT and Machine Learning (NCIM)
16-17 June 2023, Gazipur-1707, Bangladesh

An Ensemble-based Deep Learning Model for Multi-class Flower Recognition

Md Fozle Rabbi*, Sraboni Das Dola*, Noorjahan Binte Kader Taha*, Tamim Al Mahmud *, Muhammad Nazrul Islam †, Md. Abdur Razzaque ‡, and Faiz Al Faisal*
* Dept. of Computer Science and Engineering, Green University of Bangladesh, Dhaka, Bangladesh
†Dept. of Computer Science and Engineering, Military Institute of Science and Technology, Dhaka, Bangladesh
‡Dept. of Computer Science and Engineering, University of Dhaka, Dhaka-1000, Bangladesh
Email: faisal@cse.green.edu.bd

2023 International Conference on Next-Generation Computing, IoT and Machine Learning (NCIM) | 979-8-3503-1600-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/NCIM59001.2023.10212540

***Abstract*—Many species around the globe heavily depend on flowers, and even the use of flowers for medical purposes is enormous, which has been in practice since ancient times. In modern botany, agronomy, and species research, the identification and classification of flowers mainly depend on human intervention, which requires a lot of human resources, and the recognition accuracy rate is low. On the other hand, due to the vast number of flowers with different types, it is quite a challenge and increases the complexity of the accurate detection and classification process. Moreover, the traditional CNN model and other pre-trained models like- VGG16, VGG19, Xception, and MobileNet-V2 failed to ensure high validation accuracy. In this research, we have designed an ensemble-based deep learning model consisting of two pre-trained models DenseNet-201 and EfficientNet-B7, which we renamed "FlowerConvNet". This paper considers classifying five categories of flowers consisting of 4242 images, collected from the Kaggle flower recognition dataset. In this research study, the proposed deep learning model FlowerConvNet using transfer learning can obtain a competitive result with 95% validation accuracy.**

***Index Terms*—Convolutional Neural Network, deep learning, image processing, flower detection, ensemble learning, genetic algorithm.**

I. Introduction

Flowers are integral to human life as well as natural. They are valued for ornamental purposes in homes, in medicine, and as gifts for special occasions. In the medical sector, flowers offer a range of health benefits. Overall, flowers are an essential part of human life that provide practical and emotional benefits to people worldwide. It is estimated that there are somewhere between 250,000 to 400,000 flowering species on the planet, though only around 230,000 of these have been described [4]. Hundreds of thousands more flowering plants have been removed from official lists due to discoveries of duplicate species [21]. For humans, it's impossible to know all types of flowers. This research study works with 5 categorical flower species. As shown in Figure 1, these flowers are Dandelion, Rose, Daisy, Sunflower, and Tulip. Today in image processing, it is not easy to differentiate those flowers. The reason for this is that similarity in shape and color because of their diversity, such as grass and leaves, make it difficult to distinguish these species. Especially after a certain distance, it is not possible to recognize the species.

Convolutional Neural Network (CNN) is an Artificial Neural Network (ANN) that acts like the human brain. Kunihiko Fukushima introduces the concepts of CNN in 1980 inspired by biological neurons [27]. Nowadays, CNN is used for various purposes like prediction, classifications, and the like [23]–[25]. Though several studies have been conducted focusing on the identification and classification of flowers, a more effective approach could be proposed to assist botanists, scientists, or researchers to classify a flower with enhanced accuracy. Therefore, in this study, we proposed a deep learning method named FlowerConvNet using a transfer learning approach for better prediction. The proposed FlowerConvNet gives competitive results with 95% higher accuracy.

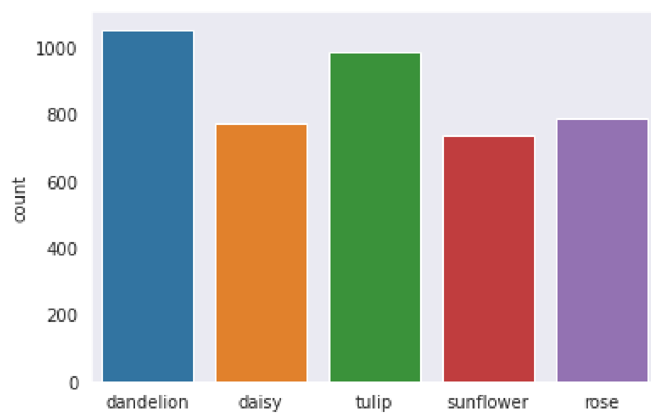

Fig. 1. Number of flowers with various types

II. Related Works

Nowadays, many researchers try to solve feature selection and classification problems. The difficulty is to find suitable descriptors of color, shape, and pattern and also for the classifier to have the capacity to select significant features [1]. The article [1] combined different features to enhance flower classification performance on a large dataset. Different features describe different aspects: the local shape/texture, the shape of the boundary, the overall spatial distribution of petals, and the color. A support vector machine (SVM) classifier was used for classification. The article [2] used the Gray Level Co-occurrence Matrix (GLCM) and a combination of Gabor for extracting the feature and then applying K-Nearest Neighbor as the classifier. This paper achieved 75% accuracy. A machine learning approach has been proposed using Weka software for plant disease detection [3] where they apply various machine learning techniques like the Zero-R algorithm, SVM [26], and Simple K-means.

Another research study [6] combined four models for extracting features, and for classification, they used Support Vector Machine (SVM). They also get the best accuracy in the ”Flower Recognition [7]” dataset. In our study, we also used the ”Flower Recognition” dataset and a similar type of model that used EfficientNet-B7 and DenseNet-201.

III. Methodology

Collecting data is a crucial part of the experimentation of the proposed technique. Among the two publicly available datasets: the Oxford Flower dataset [5] and Kaggle Flower Recognition [7] dataset, for this study we have considered the Flower Recognition [7] dataset. This dataset contains 4242 images of flowers. The dataset was divided into five classes Dandelion, Daisy, Tulip, Sunflower, and Rose (Figure 1). From these classes, each category such as Daisy has 764 images, Dandelion has 1052 images, Rose has 784 images, Sunflower has 733 images, and Tulip has 984 images.

A. Dataset Preprocessing

In the first step, we preprocessed our dataset to make it suitable for feeding into our model. In this study, two models were used, one which is DenseNet and the other is EfficientNet. Initially, we had reshaped our image data as $228 \times 228 \times 3$ height and width sizes. This study also tried to take different types of shapes. If we feed larger shapes of images, this model stops training or doesn't work to feed the neural network. After this, we normalized the images pixel using min-max normalizations. After normalizing those images, we created a training and validation set from the original dataset with an 80% and 20% ratio. Also, apply different types of augmentation, like- rescale, shear range, zoom range, horizontal flip, etc.

1) ***Reshaping****:* For our model, we converted all images into $228 \times 228 \times 3$ height and width sizes.

2) ***Normalization****:* Normalization helps to stabilize the gradient, allowing us to use a higher learning rate and converge more quickly [9].

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{1}$$

Here, x is the original data, x' is the normalized data, $x_{\max}, x_{\min}$ are respectively the maximum and minimum values of the original vector. This normalization technique is called scaling. This normalization scaling was applied to all the values in order to normalize between 0 to 1.

3) ***Augmentation****:* Image Augmentation is a technique that artificially increases the training dataset by creating new versions of existing images with some slight variations. The variations were made by applying a transformation to the images, rotating, scaling, flipping, and cropping. Figure 2 shows an example of a particular sample after introducing the data augmentation.

TABLE I
Range used for data augmentation

Augmentation Technique	Specified Value
Shear Range	.2
Zoom Range	.2
Horizontal Flip	True
Rescaling	1./255

Fig. 2. After Augmenting an Image.

B. Proposed FlowerConvNet Architecture

Our proposed method consists of three stages, as shown in Figure 3. Firstly, we preprocessed those images with different types of augmentations. Then, this study divides that dataset into training and validation sets. Then, the flower images were passed through as the inputs to the pre-trained CNN models. Finally, EfficientNet-B7 and DenseNet-201 models were considered, and hyperparameters were used for increasing the model accuracy. Two CNN models were trained on the flower dataset in our first experiment. In the feature extraction phase in DenseNet, Fully Connected (FC)-768 (FC-256, FC-128) layers were used. For EfficientNet-B7, we used FC-384 layers for improving the existing performances in both models. Here, we also used dropout layers, where 0.5 dropout layers considered for both models.

Since Flower images in the flower dataset are of different widths, heights, and sizes, resizing them to equal width and height is recommended to get optimum results. In this study, we resize the Flower images to either $224 \times 224 \times 3$ (or $299 \times 299 \times 3$) pixels since the input image dimensions of pre-trained CNN models are $224 \times 224 \times 3$ pixels except for the Xception, which requires the input image with a size of $299 \times 299 \times 3$ pixels. The study employs ensemble learning to train the dataset using significant pre-trained models such as

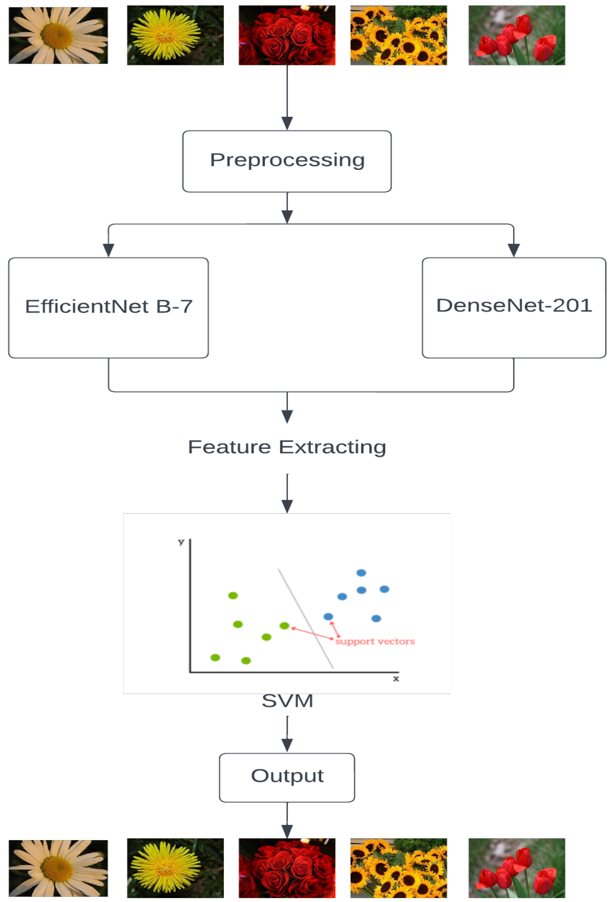

Fig. 3. FlowerConvNet Proposed Model

EfficientNet-B7 and DenseNet-201. Also, we have checked performance in different types of pre-trained models like ResNet-50, VGG-16, VGG-19, and Xception.

- **ResNet:** Residual Networks whose short name is ResNet. This model is used in Image Classification, Object Detection, and Image Segmentation [20]. This model introduced Residual Connections techniques, which resolved the vanishing gradient descent problem by producing Skip Connections techniques. Figure 4 shows the working process of a residual building block [22], where the formula f(x) + x represents the residual connections in ResNet architecture. F(x) is the residual mapping, learned from the training, and x is the input to the residual block.
- **EfficientNet-B7:** EfficientNet proposed a scaling method that uses scale-up CNNs in a more structured manner [10]. The scaling method involves finding the optimal combination of model depth, width, and resolution using a compound coefficient that controls the overall model size. EfficientNet model named with the scaling factor. Such as EfficientNet-B0, EfficientNet-B1, EfficientNet-B2, EfficientNet-B7, etc. Every model with successive models has a higher scaling factor and more parameters, and the result is more accurate.

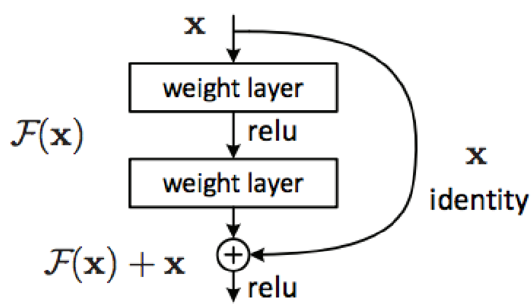

Fig. 4. Residual Skip Connection

- **DenseNet-201:** DenseNet also comes with several variations, such as DenseNet-121, DenseNet-169, and DenseNet-201 [12]. Where the number represents the number of layers in the network, in DenseNet, all the preceding layers were concatenated together with inputs to the current layers, which enables better gradient flow to the network. This pre-trained model reduces overfitting and increases the accuracy of training data.

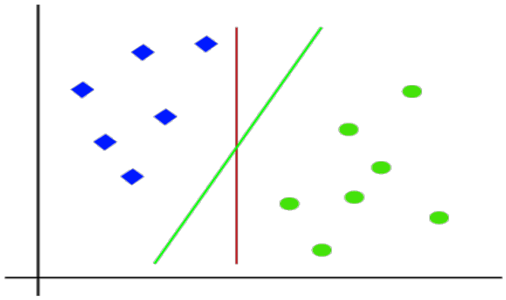

Fig. 5. Classification Technique using Support Vector Machine

C. Support Vector Machine (SVM) Classifier

The support vector machine (SVM) technique is widely used in regression and classification. The SVM method creates a boundary that divides different types of classes. In this manner, a hyperplane area is created. In the next step, the data attributes in the hyperplane were classified by SVM. The SVM method also prefers multiple classes. The visual design classified by the binary-class (two-class) SVM method is shown in Fig. 5. Support Vector Machine algorithm separates different types of data by using green boundary lines. The function of the line is $y = ax + b$. For two-dimensional linearly separable data separated by the boundary line. So, for two dimensional x with $x1$ and y with $x2$ and we get: $ax_1 - x_2 + b = 0$ and for, $x = (x_1, x_2)$ and $w = (a, -1)$, we get: $w \times x + b = 0$.

After creating the hyperplane, we considered the hyperplane to make predictions. We define the hypothesis function h as:

$$h(x_i) = \begin{cases} +1 & \text{if } w \cdot x + b \geq 0 \\ -1 & \text{if } w \cdot x + b < 0 \end{cases} \tag{2}$$

The point above or on the hyperplane was classified as class +1, and the point below the hyperplane will be classified as class -1 [11].

SVM can be classified into two classes. We can do multi-class classification with SVM use of a combination of two-class-based classification techniques, which is called one-versus another rest of class, and one vs versus of one class.

IV. Training Configuration

1) Epoch: We always need help with how many epochs to train to get good accuracy on our dataset. Too many epochs make the problem, which makes it overfitted in the trained dataset. We trained our dataset with 10-20 epochs for warm-up training. After using the EarlyStopping Keras function, our proposed approach did training in only 11 epochs. In this function, we also used the learning rate by a factor of 0.5 units.

2) Hyperparameters: In this study, we used various hyperparameters. Like learning rate, optimizers, dropout, batch-normalization, kernel initializer, that's it. Now describing those hyper-parameters:

- **Learning Rate**: The learning rate is a hyperparameter that controls how much to change the model in response to the estimated error each time the model weights were updated. It's a challenging task to choose the best learning rate for our model. For determining the learning rate, we used ReduceLROnPlateau. This is a Keras function, which helps us to determine when a metric has stopped improving. In these hyperparameters, we add a starting learning rate as lr = 1e-5 to lr = 1e-6 with a dropout rate of 0.2-0.7.
- **Optimizers**: Optimizer method are used to change the weights on the training times in neural network algorithms. In this study, we used Adam optimizers for this research analysis.

3) Adam Optimizers: In Convolutional Neural Networks (CNNs), an optimizer is an algorithm used to update the weights of the neural network during the training process. The optimizer is responsible for minimizing the loss function by adjusting the consequences of the neural network in the direction of the steepest descent of the loss function. There are various kinds of optimizers available now. Among all types of optimizers, Adam is the best optimizer nowadays. Most researcher now uses the Adam optimizer because that dynamically adjusts the learning rate based on the gradient of the loss function and the moments of the gradient descent.

4) Activation Function: In this study, we used two different activation functions. There are ReLu and Softmax. The ReLU (Rectified Linear Unit) is the most used activation function worldwide for its features. It is used in all convolutional neural networks or deep learning tasks. $f(x) = max(0, 1)$ Where x is the input and $f(x)$ is the output. The output of the function $f(x) = 0$ for the negative inputs and otherwise positive. SoftMax activations [18] functions were used for the output layers, while the Softmax activation function is considered mainly in image classifications and multiclass classification tasks.

5) Loss Function: The loss function calculates the distance between the current output generated by the algorithm and the expected output. It's a method to evaluate the model and how our algorithm works on the dataset. Loss function can be categorized into two groups. One is for classification tasks like discrete values, and the other is for continuous data in regression.

V. Experimental Results

In this study, the proposed model FlowerConvNet used two pre-trained models with transfer learning and ensemble learning approach. This proposed model uses two pre-trained models. The first model is Densenet 201 and another model is EfficientNet-B7. This study gives the best accuracy with a 95% validation score. This study also builds a custom neural network where this study adds four convolutional layers

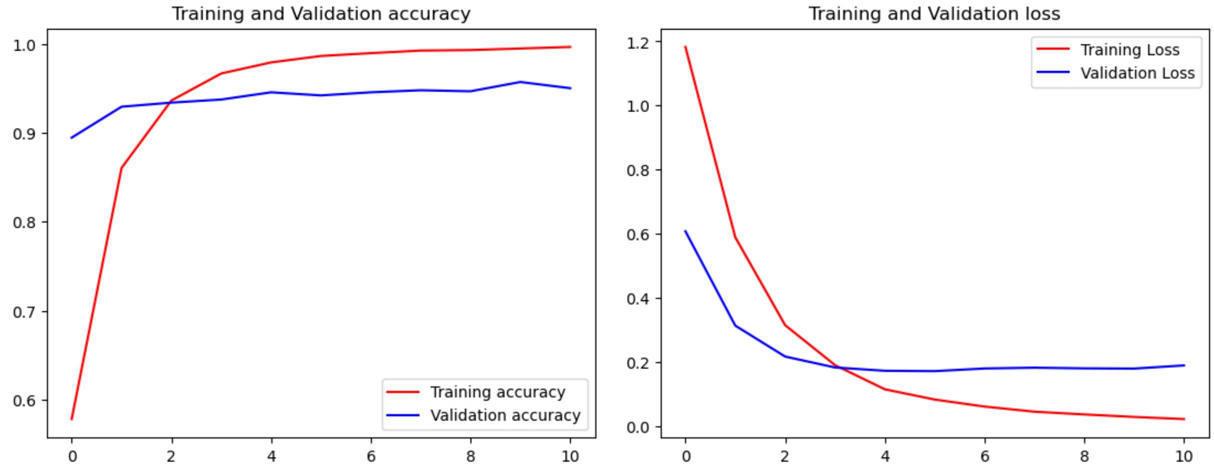

Fig. 6. Analysis of Model Accuracy and Model Loss on FlowerConvNet

respectively with 32, 64, 96, and 128 filters, as activation function ReLu and for pooling the information from the convolutional layer this study used MaxPool2D. After all, in the end, we used two dense layers. This custom neural network gives 84% accuracy. This study also trained our flower recognition dataset using other pre-trained models for better accuracy. MobileNet-V2 gives 89% accuracy, VGG-16 gives 84% accuracy, VGG-19 gives 89%, and ResNet-50 gives us 93% accuracy (see Table II).

TABLE II
OTHER PRETRAINED MODEL RESULTS

Model	Result (Validation accuracy)
VGG-16	84%
VGG-19	89%
Xception	85%
MobileNet-V2	89%
EfficientNet-B7	94%
Custom Neural Network	84%
Proposed Approach	**95%**

In this study, the metrics used to measure how ”accurate” our classifier is in predicting the class supplies and labels. For this measure, we used Accuracy, Recall, Precision, and F1-score. The formula for evaluating model performance we used in our study showing below:

$$Accuracy = \frac{(TP+TN)}{(TP+FP+TN+FN)} \tag{3}$$

$$Precision = \frac{TP}{(TP+FP)} \tag{4}$$

$$Recall = \frac{TP}{(TP+FN)} \tag{5}$$

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{6}$$

Here, True Positive (TP) is the number of correctly predicted positive instances, True Negative (TN) is the number of correctly predicted negative instances, False Positive (FP) is the number of incorrectly predicted positive instances and False Negative (FN) is the number of incorrectly predicted negative instances. Table III describes the analysis of precision, recall, and f1-score for each different class of flowers. The result analysis of the confusion matrix is shown in Fig. 7. In contrast, the model accuracy and model loss were showed in Fig. 6. Moreover, in Table IV, the authors have compared the accuracy analysis of other research publications with the same dataset.

TABLE III
CLASSIFICATION REPORT OF OUR FLOWERCONVNET MODEL

Flower Name	Precision	Recall	f1-score	Support
Dandelion	.98	.91	.94	144
Daisy	.98	.96	0.96	206
Sunflower	0.90	0.91	0.90	137
Tulip	0.97	0.97	0.97	172
Rose	0.91	0.95	0.93	204

TABLE IV
ACCURACY ANALYSIS WITH THE SAME DATASET

	Used Method	Acc. (%)
Baosu Guo, Jingwen Hu [14]	CNN and Genetic Algorithm	78%
Francois Luus, Naweed Khan, Ismail Akhalwaya [17]	Used CNN with dimension reduced technique	79%
Bo Chen et al. [15]	Custom CNN	85%
FlowerConvNet	**Our Proposed Approach**	95%

VI. CONCLUSION

The most significant contribution of this research is to introduce an ensemble-based deep learning model in order to classify the various classes of flowers like- Dandelion, Daisy, Tulip, Sunflower, and Rose. This research will definitely help the medical industry to a greater extent in order to identify the difference between multi-class flowers through image processing. This research considered two pre-trained models

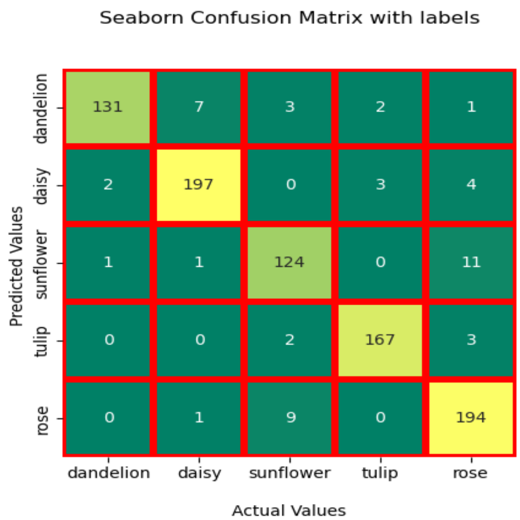

Fig. 7. Confusion matrix analysis on Kaggle dataset

DenseNet-201 and EfficientNet-B7 for model analysis, and in order to reduce the data processing overhead we used data augmentation techniques to expand our dataset and minimize overfitting. Regarding the result analysis on our model using various metrics; including accuracy, precision (showed 98% precision in percentage on Dandelion), recall (showed 97% recall in percentage on Tulip), and f1-score (showed 97% f1-score in percentage on Tulip). Our results ensured that our novel ensembled model outperformed others like- Baosu Guo et al. as having 78% [14], Francois Luus et al. as having 79% [17], Bo Chen et al. as having 85% of test data accuracy [15]. Finally, regarding the scope of future research, this novel ensemble model can be useful for plant detection and classification in terms of high accuracy and precision.

References

[1] Nilsback, Maria-Elena, and Andrew Zisserman. ”Automated flower classification over a large number of classes”, 2008 Sixth Indian Conference on Computer Vision, Graphics & Image Processing. IEEE, 2008.

[2] Guru, D. S., Y. H. Sharath, and S. Manjunath. ”Texture features and KNN in the classification of flower images.” IJCA, Special Issue on RTIPPR (1) (2010): 21-29.

[3] Ramya, R., et al. ”A review of different classification techniques in machine learning using WEKA for plant disease detection”, Int. Res. J. Eng. Technol.(IRJET) 5 (2018): 3818-3823.

[4] Nelson D. Types Of Flowers: List And Names - Science Trends. Science Trends - Explore More. Published November 29, 2018. Accessed May 29, 2023. https://sciencetrends.com/types-of-flowers-list-and-names/

[5] Visual Geometry Group - University of Oxford. Ox.ac.uk. Published 2023. Accessed May 29, 2023. https://www.robots.ox.ac.uk/vgg/data/flowers/

[6] Toğaçar, Mesut, Burhan Ergen, and Zafer Cömert. ”Classification of flower species by using features extracted from the intersection of feature selection methods in convolutional neural network models”, Measurement, Vol. 158, pp. 107703, 2020.

[7] A. Mamaev, Flowers Recognition, accessed on April 2023, [online] ”https://www.kaggle.com/datasets/alxmamaev/flowers-recognition”

[8] Ian Goodfellow, Yoshua Bengio, and Aaron Courville, Deep Learning, Accessed on April 2023, [online] ”http://www.deeplearningbook.org”.

[9] baeldung “Normalizing Inputs of Neural Networks Baeldung on Computer Science,” Baeldung on Computer Science, Accessed Dec., 2022,[online] ”https://www.baeldung.com/cs/normalizing-neural-network-inputs”

[10] Tan, Mingxing, and Quoc Le. ”Efficientnet: Rethinking model scaling for convolutional neural networks”, International conference on machine learning. PMLR, 2019.

[11] Awad, M., Khanna, R., ”Support Vector Machines for Classification”, In: Efficient Learning Machines. Apress, Berkeley, CA, 2015.

[12] He, Kaiming, et al. ”Deep residual learning for image recognition.” Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[13] Aditi, Sharia Wasika, Farjana Kabir, and Pintu Chandra Shill. ”Diagnosis of Diabetic Retinopathy Using Deep Learning Techniques.” 2021 5th International Conference on Electrical Information and Communication Technology (EICT). IEEE, 2021.

[14] Guo, Baosu, Jingwen Hu, Wenwen Wu, Qingjin Peng, and Fenghe Wu, ”The Tabu_Genetic Algorithm: A Novel Method for Hyper-Parameter Optimization of Learning Algorithms” Electronics 8, no. 5: 579, 2019.

[15] Bo Chen, Jun Liu, Jiong Liu, Jiacheng Sun, ”Flowers Classification vai Deep Learning”, http://noiselab.ucsd.edu/ECE228_2019/Reports/Report40.pdf

[16] Yohannes, Y., Udjulawa, D., & Ivan Sariyo, T. (2021). Klasifikasi Jenis Jamur Menggunakan SVM dengan Fitur HSV dan HOG. PETIR, 15(1), 113–120. https://doi.org/10.33322/petir.v15i1.1101

[17] Y. Ma, A. Fan, J. He, A. R. Nelakurthi and R. Maciejewski, ”A Visual Analytics Framework for Explaining and Diagnosing Transfer Learning Processes,” in IEEE Transactions on Visualization and Computer Graphics, vol. 27, no. 2, pp. 1385-1395, Feb. 2021, doi: 10.1109/TVCG.2020.3028888.

[18] M. Wang, S. Lu, D. Zhu, J. Lin and Z. Wang, ”A High-Speed and Low-Complexity Architecture for Softmax Function in Deep Learning,” 2018 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS), Chengdu, China, 2018, pp. 223-226, doi: 10.1109/APCCAS.2018.8605654.

[19] D. Kingma and J. Ba, “Adam: A method for stochastic optimization,” Proceedings of the 3rd International Conference on Learning Representations (ICLR 2015), 2015.

[20] C. Jianfei and Z. Changming, ”Research on Image Recognition Based on Improved ResNet,” 2020 IEEE 6th International Conference on Computer and Communications (ICCC), Chengdu, China, 2020, pp. 1422-1426, doi: 10.1109/ICCC51575.2020.9345181.

[21] Nelson, Daniel. “Types of Flowers: List and Names - Science Trends.” Science Trends - Explore More, 29 Nov. 2018, sciencetrends.com/types-of-flowers-list-and-names/. Accessed 12 Apr. 2023.

[22] Sivaram T. All You Need to Know About Skip Connections. Analytics Vidhya. Published August 24, 2021. Accessed May 25, 2023. https://www.analyticsvidhya.com/blog/2021/08/all-you-need-to-know-about-skip-connections/

[23] M. R. Sheikh, T. H. Masud, N. I. Khan and M. N. Islam, ”An Efficient Transfer Learning Model for Predicting Forged (Handwritten) Signature,” 2021 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2), Rajshahi, Bangladesh, 2021, pp. 1-4

[24] S. A. Suha and M. N. Islam, “An extended machine learning technique for polycystic ovary syndrome detection using ovary ultrasound image,” Sci. Rep., vol. 12, no. 1, Article no. 17123, 2022.

[25] M. N. Islam, T. T. Inan, S. Rafi, S. S. Akter, I. H. Sarker and A. K. M. N. Islam, ”A Systematic Review on the Use of AI and ML for Fighting the COVID-19 Pandemic,” in IEEE Transactions on Artificial Intelligence, vol. 1, no. 3, pp. 258-270, Dec. 2020,

[26] M. R. Hossain Bhuiyan, N. Saha Roy, M. K. Roy, M. Sultanul Islam Ovi, F. A. Faisal and M. Aminur Rahaman, ”Sedimentation Analysis on Padma River Using Machine Learning Techniques,” 2022 4th International Conference on Sustainable Technologies for Industry 4.0 (STI), Dhaka, Bangladesh, 2022, pp. 1-6, doi: 10.1109/STI56238.2022.10103351.

[27] The Brief History of Convolutional Neural Networks - Towards Data Science. Medium. Published February 24, 2021. Accessed May 29, 2023. https://towardsdatascience.com/the-brief-history-of-convolutional-neural-networks-45afa1046f7f