2024 8th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT) | 979-8-3503-5442-3/24/$31.00 ©2024 IEEE | DOI: 10.1109/ISMSIT63511.2024.10757203

Classification Flower Images Based On Deep Learning And Machine Learning

1st Asia Kamal Mustfa
dept. Computer Science
Computer Science and Mathematics College
Tikrit University
Tikrit, Iraq
asia.k.mustfa@st.tu.edu.iq

2nd Salwa Khalid Abdulateef
dept. Computer Science
Computer Science and Mathematics College
Tikrit University
Tikrit, Iraq
Khalid.salwa@tu.edu.iq

3rd Qabas Amer Hameed
dept. Computer Science
Computer Science and Mathematics College
Tikrit University
Tikrit, Iraq
qabas.a.hameed@tu.edu.iq

4th Mohamed Aktham Ahmed
Department of computer science
College of Computer Science and Mathematics
Tikrit University
Tikrit,,Iraq
mohamed.aktham@tu.edu.iq

5th Zaydoon Al-qaysi
Department of computer science
College of Computer Science and Mathematics
Tikrit University
Tikrit,,Iraq
zaidoonalqaissi@gmail.com

***Abstract*—Image classification has grown increasingly popular due to the growing significance of machine learning and deep learning. Flower images may sometimes exhibit resemblances in terms of hue, form, and visual characteristics. The problem lies in the classification of flowers. This work employs a hybrid approach that integrates deep learning and machine learning techniques to classify 17 discrete flower species. In order to do this, we utilised the ResNet_50, PCA, and SVM architecture to classify several species from the "Oxford-17" dataset. With this goal in mind, we have made efforts to improve our model in order to get more accuracy compared to similar methods. Prior to inputting our images into our pretrained model, we resized them, and subsequently fine-tuned the model. The dataset was partitioned into two distinct sets: a training set and a testing set. We attained a precision rate of 95.58% while utilising the "Oxford-17" dataset. Our approach outperformed previous machine learning and deep learning-based methods on this dataset.**

Keywords— Machine Learning, Deep Learning, Flower Classification, ResNet_50, PCA, SVM.

I. Introduction

The study of flowers and the skill of recognising them have captivated humanity for countless years [1]. Flowers, a type of botanical organism, play a crucial role in maintaining ecological balance. Flowers have the ability to provide sustenance to a wide range of organisms, including nearly all insect species on Earth. Additionally, flowers are utilised in several beneficial ways for humans, particularly in the field of pharmaceuticals. Identifying flowers continues to be a difficult task for the majority of individuals. The user's text is "[2]." The main factor for this is the existence of several flowers that share similar colour, shape, and appearance. In addition, images of different flowers commonly have comparable elements in their surroundings, such as leaves, grass, and other similar objects. There are roughly 250,000 identified species of flowering plants, which are classified into about 350 families[5]. Identifying and categorising these items requires substantial time and exertion [6]. Manual categorization is a laborious process that can consume a significant amount of time and may result in mistakes as time goes on [7]. Therefore, it is crucial to design a computer-aided method that can classify flowers quickly and accurately [8]. The study of flower categorization holds significant importance in the field of botany. Conventional flower categorization systems struggle to accurately account for the influence of a bloom's background. This leads to an inadequate categorization effect. The development of massive data and rapid advancements in Internet technology have led to the increased use of deep learning in the field of picture categorization research [9]. The flower categorization was performed using machine learning and deep learning methodologies. Machine learning is an artificial intelligence technique employed to detect and analyse patterns within datasets [10]. Deep learning is a subset of machine learning [11] that allows computers to autonomously extract data [12].

This study presents a hybrid method can classify a different type of flowers. This approach is based on Residual Neural Network (ResNet_50) to Feature Extraction (FE), Principal Component Analysis (PCA) to Feature Selection (FS), and Support Vector Machine (SVM)for Classification.

The paper is structured as follows: Section 2 presents a comprehensive summary of the current methodologies and approaches that have been reviewed in the literature. Section 3 offers an elaborate elucidation of the suggested methodology, framework, and measurements utilized to attain the intended results. Section 4 analyzes the outcomes and assesses the proposed methodology. Section 5 encompasses the final remarks of the report and outlines the potential areas for future work.

II. Literature Review

Many academics are attempting to propose various answers to the challenge of flower images classification modeling using the utilization of machine learning and image processing techniques to attain precise and automated classification. In recent times, scholars have started employing deep learning algorithms to yield more precise outcomes. This section showcases the most pertinent deep learning-driven flower images classification modeling utilizing deep learning algorithms, which include:

In 2019, the researcher utilised Support Vector Machines (SVM) for classification and applied Speeded Up Robust Features (SURF) and Local Binary Patterns (LBP) for feature extraction. This study seeks to assess the efficacy of feature descriptors in flower classification using the Local Binary Pattern (LBP) and Speeded-Up Robust Features (SURF). The Oxford_17 dataset outperformed other datasets and classifiers, with an accuracy of 87.2%[13].

In accordance with the earlier reference [14], the authors presented the Convolutional Neural Network (CNN) model for Utilising a reduced dataset for training purposes leads to enhanced precision in the model's outcomes. An inquiry into the indigenous floral species of Bangladesh is commenced following a suggestion. Flowers including Chapa, Kadam, Kath Golap, Shapla, Rongon, Radhachura, and Rojonigondha are found in the area. We have demonstrated that a Convolutional Neural Network (CNN) architecture can achieve an accuracy rate of 85% in classification. Isha Patel and her colleagues [15] constructed a deep convolutional neural network utilising the Faster R-CNN framework and the NAS-FPN architecture. Accurately identifying, locating, and categorising flower photographs in databases with several categories may be a difficult task, especially when working in the agricultural industry. The research employs many dataset classifications. Dataset 1 contains a total of 8189 photographs, which depict 102 distinct varieties of flowers.When evaluated using the 102-flower class dataset, the NAS-FPN model combined with the Faster R-CNN approach got an impressive mean average accuracy (mAP) score of 87.6% [15].

The study employed a Random Forest algorithm to classify flower photos. The study's classification difficulty is presented by the Oxford 102 Flowers dataset, which exhibits significant diversity within each class. The suggested method's efficacy is assessed using stratified k-fold cross-validation tests, resulting in an accuracy of 88.74%. The text is referenced by the number 16. A 2021 study by Rongxin Lv et al. introduced a technique that integrates saliency detection with the VGG-16 convolutional neural network. The conventional convolutional neural networks and other methodologies for flower classification include inherent limitations, and the objective of our methodology is to rectify such shortcomings. The classification results have been unsatisfactory since the efforts to reduce the impact of the flowery backdrop have been ineffective. Experiments conducted on the Oxford flower-102 dataset [9] demonstrate that the algorithm achieves a 91.9% accuracy rate on these datasets. The study examines the effectiveness of transfer learning in the task of categorising pictures by utilising pre-trained deep learning models such as Alex Net [17]. Image classification, a crucial domain in deep learning, is the primary subject of the research piece, where pretrained models are employed. The data collection is organised into five distinct categories: dandelion, sunflower, tulip, rose, and chamomile. The Alex Net model attained an accuracy of 86.28% after being trained on these datasets. In 2021, Shuai Cao et al. introduced Visual Attentional-driven DCNNs (VA-DCNNs), a technique for guiding attention. The classification of flowers is a complex undertaking that requires very sophisticated image recognition techniques, and researchers focus their efforts on addressing these issues. These issues include a lack of sufficient training data, similarities across different classes, and inadequate accuracy in categorising certain varieties of flowers. Researchers have demonstrated that VA-DCNNs may obtain accuracies of up to 85.7% using the publically available Flowers 17 dataset [18]. Ari Peryanto and his colleagues in 2021. It was dependent on this paradigm. Convolutional neural networks (CNNs) are unrivalled in the field of image processing. Manual classification is prone to human fallibility and subjectivity, which might result in inconsistent results. Scaling up the manual method is difficult, and there is no guarantee that the classification results will be accurate or consistent. A collection was created by compiling 1200 random photographs using Google Image Search. The Rose category has 400 data units, the Tulip category contains 400 flowers, and the Aster category also contains 400 flowers. The CNN model attained an accuracy of 91.6% in categorising flower photos. CNN had a precision rating of 91.6% in the categorization of flower photographs. CNN demonstrated a recall rate of 91.6% in accurately categorising photographs of flowers. The F1 Score for identifying floral images using CNN was 91.6% [19].The primary goal of this research is to develop a software that can accurately identify various types of flowers from photographs utilising pre-trained models such as ResNet, MobileNet, DenseNet, and Inception. Two datasets were employed for training the models: the Oxford 17 dataset, which consists of seventeen distinct flower variations, and the flower dataset, which includes five different flower types. Subsequently, the models were evaluated to determine their performance. The success of deep learning methods is contingent upon the number of classes in the dataset and the use of the Adam optimizer. Here is the level of precision exhibited by these models: Attained accuracy rates of 93%, 94%, 91%, and 92% on the Oxford_17 dataset [20]. In 2023, the Associates of Rini Nuraini were established. This research tries to identify sunflower photographs using an approach that extracts first-order features. This technique is employed to derive statistical measures such as entropy, variance, skewness, kurtosis, and mean. Subsequently, these attributes are entered into the Multiclass SVM identification programme. The dataset consists of seven sunflower cultivars: Red Sun, Fiesta Del Sol, Sunny Smile, Teddy Bear, Early Russian, Cherry Rose, and Velvet Queen. The model achieved an average accuracy rate of 79%. The user's input is "[21]".

III. Proposed Methodology

This section introduces the classification model for flower images as explained in Figure 1.

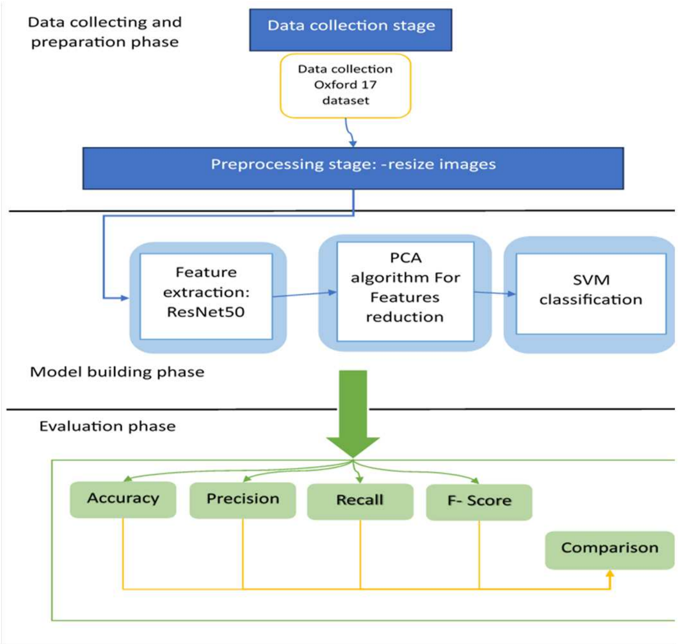

Fig. 1. Study research methodology.

A. Dataset

The study included "Oxford-17"dataset. The Oxford-17 dataset includes1,360 images for each of 17 flower class,80 image in each class, with varying poses, sizes, and perspectives [18]. (Fig 1). It was created by Nilsback and Zisserman, a flower species seen in England [20]. This dataset consists the following flowers (cowslip, tulip, tigerlily, crocus, bluebell, lilyvalley, snowdrop, windflower, sunflower, pansy, iris, fritillary, dandelion, daisy, daffodil, colts' foot, and buttercup) [22].

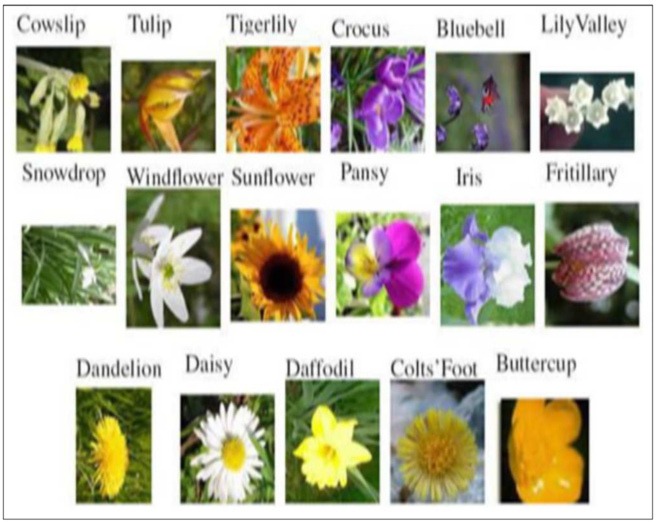

Fig.2. "Oxford-17" dataset sample images

The images of the flowers used in this study were obtained from the publicly accessible Kaggle database.

B. Pre-processing

Pre-processing is one of the key steps that requires precise attention, as data quality has a direct impact on the efficiency and precision of the intended results. During this phase, images are improved for processing and resized to 224x224. This process involves extracting flowers from a given image. Figure 3 illustrates an instance of an image extracted from the dataset prior to and subsequent to the preprocessing stage.

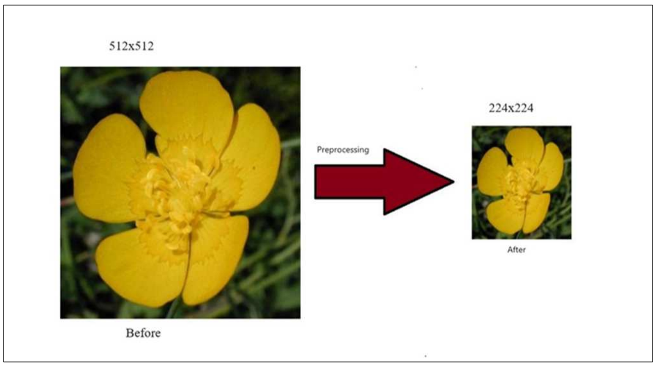

Fig. 3. Preprocessing Stage.

C. Model

The recommended models consist of three components: Residual Neural Network (ResNet_50) for extracting features, Principal Component Analysis (PCA) for selecting features, and Support Vector Machine (SVM) for classifying. The models have been built in Colab, a comprehensive software environment.

ResNet50 is a convolutional neural network consisting of 50 layers that was created using the ImageNet dataset. ResNet architectures incorporate shortcut connections, which distinguish them from conventional ESAs. Shortcut connections do not provide additional parameters or augment computational complexity [20]. This work used the ResNet_50 model to extract features from floral photos obtained from the Oxford_17 dataset. The initial number of features was 100352.

Principal Component Analysis (PCA) is a method employed to reduce the number of variables in certain datasets, hence decreasing their dimensionality. Improves understanding while retaining a substantial amount of the original information. This is accomplished by generating novel variables that are not influenced by one another. To solve the eigenvalue/eigenvectors problem [23], it is important to identify the newly added variables, also referred to as the major components. This work use Principal Component Analysis (PCA) using the ResNet_50 model on the Oxford_17 dataset to identify and select relevant features. The number of features after doing Principal Component Analysis (PCA) is 572.

Support Vector Machines (SVM), sometimes referred to as Support Vector Networks, are models and methods used in supervised learning to analyse data and identify patterns for the purpose of classification and regression analysis [24].

SVMs can categorize linear and nonlinear data. SVM classifies training data by translating it into multidimensional space and creating hyperplanes in higher dimensions. SVM is an optimum hyperplane-based mathematical learning system [25]. In this study we use SVM after ResNet_50 and PCA on Oxford_17 dataset for classify 17 types of flower images.

D. Model Evaluation

To determine the model's performance, an accuracy test was done to measure the level of correctness exhibited by the model. This test evaluates the classifier's capability to produce accurate diagnoses. The accuracy equation is represented by equation (1). Furthermore, we employed Precision (PREC), F_Score, and Recall (Rec) as evaluation metrics to gauge the effectiveness of our proposed model, as these metrics are widely used and effective in assessing model performance. The following equations provide an explanation for these measurements [26-28]:

$$ACC = \frac{TP + TN}{TP + TN + FN + FP} \quad (1)$$

$$PREC = \frac{TP}{TP + FP} \quad (2)$$

$$F\ Score = \frac{2 \times precision \times recall}{precision \times recall} \quad (3)$$

$$Rec = \frac{TP}{TP + FN} \quad (4)$$

TP represents the count of true positive results (which are the predicted positive outcomes).

FN stands for the count of false negatives, which are positive outcomes that go against the predictions.

FP represents the number of false positives. (The result is bad, as expected.)

TN is the number of instances where a prediction. (which was anticipated to be positive, actually turned out to be negative).

IV. Results and Discussion

The collected findings demonstrate the feasibility of the provided concept, which is founded on ResNet_50, PCA, and SVM models, its efficiency in classification based on_ flower images. Pre-learning models were trained using "Oxford-17" dataset.70% of the images using as training and 30% as test data.

The model was trained using the "Oxford_17" dataset, resulting in an accuracy of 95.58%. The accuracy was determined to be 96%, the recall was 95%, and the F1-score was 95. The most significant findings are shown in Table I.

TABLE I. MODEL

Model	Dataset	Accuracy	Precision	Recall	F1-Score
Special Model (ResNet_50, PCA, SVM)	Oxford-17	95.58%	96%	95%	95 %

The proposed model is compared to previous models that utilise Mobilenet_v2, Resnet152v2, Inception_v3, Quadratic SVM, and VA-ResNet50, all of which employ the same database. The findings indicate that our model outperforms these models in terms of accuracy. Tables II provide a summary of the results obtained by comparing our model with selected previous models on the Oxford_17 dataset.

TABLE II. COMPARED TO PREVIOUS STUDIES

Paper	Model	Dataset	No. of images	No. of classes	Accuracy
[20]	Mobilenet_v2,	Oxford_17	1360	17	93%
[20]	Resnet152v2	Oxford_17	1360	17	92%
[20]	Inception_v3	Oxford_17	1360	17	91%
[13]	Quadratic SVM	Oxford_17	640	8	87.2%
[18]	VA-ResNet50	Oxford_17	1360	17	85.7%
Our proposed model	ResNet50+PCA+SVM	Oxford_17	1360	17	95.58%

The proposed model is compared to different datasets utilising the same paradigm, such as Flower Recognition. The datasets used include the CNN Keras dataset [17], the Google flowers dataset [19], and the Oxford_102 dataset [15]. Figure 4 provides a summary of the findings obtained by comparing our proposed model with different datasets, using the same model as indicated. Figure 5 displays the confusion matrix representing the test data.

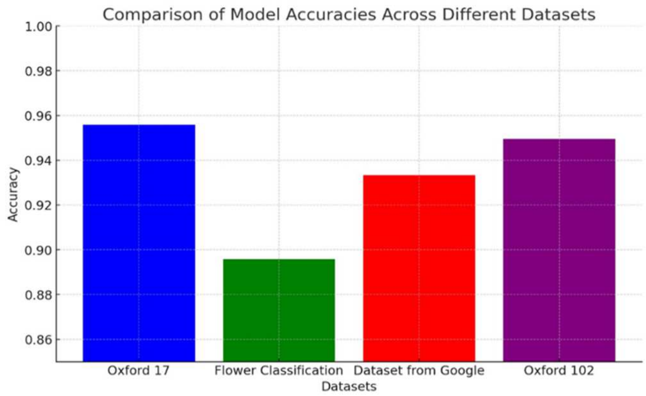

Fig. 4.Comparison model accuracies across different datasets.

Figure 5 displays the confusion matrix for the test data, which is a methodical way of mapping predictions to the original classes of the data. It is important to note that confusion matrices are only useful in supervised learning frameworks where the output distribution is known. Figure 5 accurately predicts and incorrectly predicts various flowers in the Oxford_17 dataset. For example, it correctly predicts 19 Snowdrop flowers and incorrectly predicts 1 Snowdrop flower. Similarly, it correctly predicts 18 Tigerlily flowers and incorrectly predicts 0 Tigerlily flowers. The same pattern continues for other flowers in the Oxford_17 dataset. Furthermore, Figure 6 showcases the evaluation of our proposed model on different datasets used in this study.

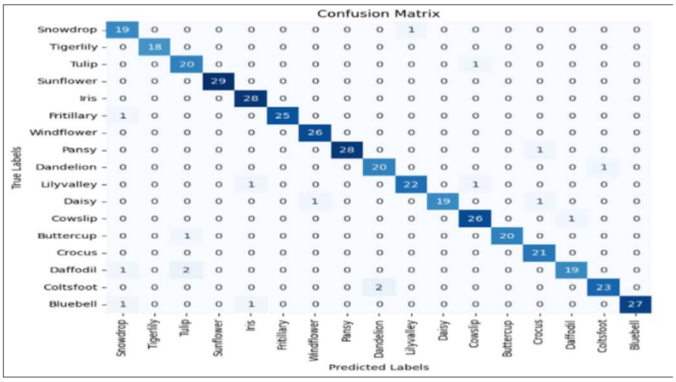

Fig.5. Confusion matrix on the test data

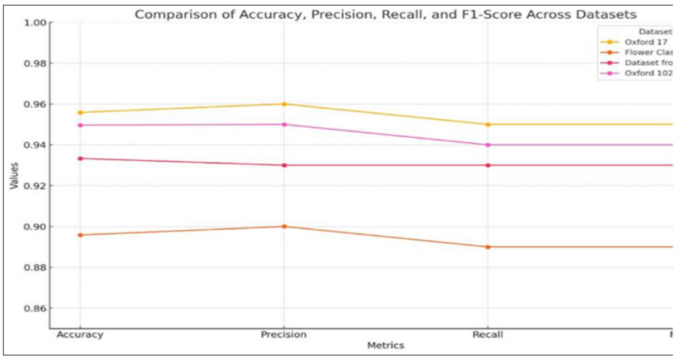

Fig. 6. Comparison of model evaluation across all datasets.

V. Conclusion

This study presents a flower categorization model that combines machine learning and deep learning techniques. The model was trained on the Oxford-17flower dataset using ResNet_50, PCA, and SVM. The classification results demonstrated that the suggested strategy was able to get a satisfactory level of accuracy. For our upcoming study, we want to employ an alternative hybrid approach that combines machine learning and deep learning techniques to classify floral images. In addition, we may apply the suggested approach to additional categorization tasks, hence increasing its use.

References

[1] Rabindra, C. S. P.: A Review on Flower Image Recognition. International journal of computer sciece and engineering 7(12), 206-216 (2029).

[2] Toğaçar, M., Ergen, B., and Cömert, Z. J. M. Classification of flower species by using features extracted from the intersection of feature selection methods in convolutional neural network models, 158, p. 107703, (2020).

[3] Mete, B. R. and Ensari, T. Flower Classification with Deep CNN and Machine Learning Algorithms. 2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), pp.1-5: IEEE Xplore, Turkey (2019).

[4] Hai, P. B., Tuan, D. A. and Nam, H. N. Performance evaluation of the multiclass classification of flowers on diverse datasets. In 2023 3rd International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), pp. 1-6: IEEE (2023).

[5] Hiary, H., Saadeh, H., and Yaqub, M. J. I. C. V., Flower classification using deep convolutional neural networks, vol. 12, no. 6, pp. 855-862, (2018).

[6] Chen, C., Yan, Q., M. Li, and Tong J., Classification of blurred flowers using convolutional neural networks, in Proceedings of the 2019 3rd International Conference on Deep Learning Technologies, pp. 71-74, (2019).

[7] ABDULATEEF, Salwa Khalid; AHMED, Saadaldeen Rashid Ahmed; SALMAN, Mohanad Dawood. A novel food image segmentation based on homogeneity test of K-means clustering. In: IOP Conference Series: Materials Science and Engineering. IOP Publishing, 2020. p. 032059.

[8] Alipour, N., Tarkhaneh, O., Awrangjeb, M., and Tian, H., Flower image classification using deep convolutional neural network, in 2021 7th International conference on web research (ICWR), 2021, pp. 1-4: IEEE, (2021).

[9] Zuo, R. Lv, Z. Li, J., and Liu, J., Flower classification and recognition based on significance test and transfer learning, in 2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE), 2021, pp. 649-652: IEEE, (2021).

[10] Jain., and Yadav, A., Analysis of performance of machine learning algorithms in detection of flowers, in 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), pp. 706-709: IEEE, (2021).

[11] Shinde, P. P., and Shah, S., A review of machine learning and deep learning applications, in 2018 Fourth international conference on computing communication control and automation (ICCUBEA), pp. 1-6: IEEE, (2018).

[12] Chauhan, N. K., and Singh, K., A review on conventional machine learning vs deep learning, in 2018 International conference on computing, power and communication technologies (GUCON, pp. 347-352: IEEE), (2018).

[13] Dhar, P. J. I. J. o. I., Graphics and Processing, S., A new flower classification system using LBP and SURF features, vol. 11, no. 5, pp. 13-20, (2019).

[14] Islam, S., Foysal, M. F. A. and Jahan, N., A computer vision approach to classify local flower using convolutional neural network, in 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 1200-1204: IEEE, (2020).

[15] Patel, I., Patel, S. J. I. J. o. S. and Research, T., An optimized deep learning model for flower classification using nas-fpn and faster r-cnn, vol. 9, no. 03, pp. 5308-5318, (2020).

[16] Siregar, F. R., and Al Maki, W. F., Hybrid method for flower classification in high intra-class variation, in 2020 3rd International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), pp. 73-78: IEEE, (2020).

[17] Cengıl, E., and Çinar, A., Multiple classification of flower images using transfer learning, in 2019 International Artificial Intelligence and Data Processing Symposium (IDAP, pp. 1-6: IEEE), (2019).

[18] Cao, S., Song, B. J. M. B., and Engineering, Visual attentional-driven deep learning method for flower recognition, vol. 18, no. 3, pp. 1981-1991, (2021).

[19] Peryanto, A., Yudhana, A., and Umar, R. J. K. I. J. I. K. d. I., Convolutional neural network and support vector machine in classification of flower images, vol. 8, no. 1, pp. 1-7, (2022).

[20] Duman, B., Süzen, A. A. J. I. J. o. A. N., and Applications, A Study on Deep Learning Based Classification of Flower Images, vol. 14, no. 2, pp. 5385-5389, (2022).

[21] Nuraini, R., Destriana, R., Nurnaningsih, D., Daniarti, Y., and Alexander, A. D. J. J. R., Sunflower Image Classification Using Multiclass Support Vector Machine Based on Histogram Characteristics, vol. 7, no. 1, pp. 146-152, (2023).

[22] Cıbuk, M., Budak, U., Guo, Y., Ince, M. C., and Sengur, A. J. M., Efficient deep features selections and classification for flower species recognition, vol. 137, pp. 7-13, (2019).

[23] Hasan, B. M. S., Abdulazeez, A. M. J. J. o. S. C., and Mining, D., A review of principal component analysis algorithm for dimensionality reduction, vol. 2, no. 1, pp. 20-30, (2021).

[24] Pisner, D. A., and Schnyer, D. M., Support vector machine, in Machine learning: Elsevier, 101-121, (2020).

[25] Chandra, M. A., and Bedi, S. J. I. J. o. I. T., Survey on SVM and their application in image classification, vol. 13, no. 5, pp. 1-11, (2021).

[26] Vujović, Ž. J. I. J. o. A. C. S., and Applications, Classification model evaluation metrics, vol. 12, no. 6, pp. 599-606, (2021).

[27] Abdulateef, S. K., Ismael, A. N., Salman, M. D., Feature weighting for parkinson's identification using single hidden layer neural network, International Journal of Computing, vol. 22, no. 2, pp. 225-230, doi: 10.47839/ijc.22.2.3092, (2022)