IJISEM
INTERNATIONAL JOURNAL OF INNOVATIONS IN SCIENCE ENGINEERING AND MANAGEMENT

Maximizing Accuracy in Image Classification using Transfer Learning and Random Forest

Pooja Suryawanshi[1], Dr. Pankaj Richhariya[2]

[1]*Research Scholar, Department of Computer Science, BITS, Bhopal*
[2]*Head of Department, Department of Computer Science, BITS, Bhopal*

OPEN ACCESS

Manuscript ID:

AG-2023-3005

Volume: 2

Issue: 3

Month: July

Year: 2023

ISSN: 2583-7117

Published: 16.07.2023

Citation:

Pooja Suryawanshi1, Dr. Pankaj Richhariya2. “Maximizing Accuracy in Image Classification using Transfer Learning and Random Forest” International Journal of Innovations in Science Engineering and Management, vol. 2, no. 3, 2023, pp. 39–48.

Abstract
In order to enhance the precision and effectiveness of automated image analysis systems, this study offers a thorough investigation into transfer learning-based image categorization. The goals of this study are to examine the usefulness of pre-trained models for picture classification tasks, to look at feature extraction methods, and to compare and contrast various machine learning algorithms. The suggested approach begins with the collection of a large number of photos from a variety of sources, followed by data preparation to get them ready for model training, and finally the creation of a transfer learning pipeline that use InceptionV3 based feature extraction as well as Random Forest based classification. The study results show that the suggested model is better, as it has an outstanding precision of 95.93% on test's data.

Keyword: image classification, deep learning, transfer learning, InceptionV3, image augmentation

I. INTRODUCTION

Image classification, the process of assigning labels or categories to images based on their content, plays a vital role in numerous fields, including computer vision, healthcare, and autonomous systems. With the exponential growth of image data, accurate image classification has become increasingly important. It enables applications such as object recognition, medical diagnosis, and intelligent surveillance, significantly impacting areas such as disease detection, autonomous navigation, and visual information retrieval. However, achieving high accuracy in image classification is a challenging task. Images often exhibit complex patterns and variations in lighting conditions, viewpoints, and backgrounds, making it difficult for traditional classification algorithms to extract discriminative features. Furthermore, a large quantity of labelled data and substantial computer resources are needed to train deep neural networks starting from scratch. To overcome these obstacles, researchers in the area of picture categorization have begun to experiment with transfer learning. Transfer learning is the practice of using the expertise obtained by training a model on a large-scale dataset, like ImageNet, to an entirely new but related problem. The prediction time is also less in some transfer learning models [1]. By utilizing the learned representations from pre-trained models, transfer learning significantly reduces the need for extensive labeled data and computational resources, while still achieving impressive classification accuracy. The primary objective of this research is to evaluate the effectiveness of transfer learning in improving image classification accuracy. By leveraging the power of pre-trained models, we aim to enhance the performance of image classification algorithms and reduce the training time required for achieving high accuracy. The study was carried out to discover a proper hybrid method for classifying the image A comparative analysis of the hybrid CNN model is carried out and found that CNN-LSTM is produced better performance [2]. Additionally, we seek to explore novel enhancements to transfer learning techniques that can further boost classification performance. This study adopts a formal and objective approach, focusing on the quantitative evaluation of transfer learning methods and their impact on image classification accuracy.

A. *Feature Extraction in Transfer Learning Models: Importance and Benefits*

Transfer learning models rely heavily on feature extraction, with pre-trained models serving as feature extractors. The lowest layers of the pre-trained model are used to extract informative & discriminative features from unlabeled pictures. Transfer learning enables effective knowledge transfer from pre-trained models to new domains, significantly improving classification accuracy and reducing the need for extensive labeled data (Pan and Yang 2010) [3]. These layers collect broad, low-level visual characteristics. For the final picture classification job, the collected features are fed into a dedicated classifier / machine learning algorithm. Feature extraction plays a crucial role in transfer learning for several reasons:

Utilizing Learned Visual Patterns: Pre-trained models have learned to recognize general visual patterns and features from vast amounts of training data. These patterns include edges, textures, shapes, and object parts that are relevant for various image classification tasks. Feature extraction techniques, such as convolutional neural networks (CNNs), extract high-level discriminative features from images, enhancing the performance of image classification tasks (Zeiler and Fergus 2014) [4]. By using feature extraction, we leverage the pre-trained model's knowledge and ability to capture these visual patterns, providing a solid foundation for effective image classification.

Dimensionality Reduction: Many parameters and intricate structures are characteristic of deep neural networks. However, the pre-trained model's bottom layers, which capture low-level information, are the ones we concentrate in on during feature extraction. By utilizing these layers, we effectively reduce the dimensionality of the input data, transforming high-dimensional images into more compact feature representations. This dimensionality reduction simplifies subsequent classification tasks, making them more manageable and computationally efficient.

Transferable Representations: The features extracted from pre-trained models are considered transferable representations. They generalize well across different tasks and domains, as they are learned from diverse image content. These transferable representations capture relevant information from images and encode important visual characteristics. By employing feature extraction, we harness these transferable representations, allowing us to adapt the pre-trained model's knowledge to new image classification tasks and achieve improved accuracy.

B. *Widely Used Pre-Trained Models in Image Classification*

In the realm of transfer learning for image classification, several pre-trained models and large-scale datasets have gained significant recognition. These models have shown exceptional performance in a variety of computer vision tasks after being trained on large datasets like ImageNet. Let's explore some of the well-known pre-trained models and the datasets they are trained on.

Inception: The Inception series of pre-trained models, developed by Google [5], are known for their innovative architecture that incorporates multiple parallel convolutional layers of different receptive fields. Inception models, such as InceptionV3 and InceptionResNetV2, have shown impressive performance in image classification tasks, achieving high accuracy and generalization. Inception models, like a number of popular pre-trained models, learn to extract varied and useful visual characteristics from datasets like ImageNet.

VGG (Visual Geometry Group): The VGG network, developed by the Visual Geometry Group at the University of Oxford [6], is a widely used pre-trained model for image classification. VGG consists of deep convolutional layers with multiple variations, including VGG16 and VGG19, which differ in terms of depth. These models are trained on the ImageNet dataset, which comprises millions of labelled images across 1,000 object categories. VGG models are known for their simplicity and effective feature representation, making them popular choices for transfer learning tasks.

ResNet (Residual Network): ResNet, short for Residual Network, is a ground-breaking pre-trained model that introduced the concept of residual learning. ResNet architectures employ skip connections or shortcuts, allowing for the training of significantly deeper neural networks. ResNet models, such as ResNet50, ResNet101, and ResNet152, have achieved state-of-the-art performance in image classification tasks [7]. These models are trained on large-scale datasets like ImageNet, enabling them to capture highly discriminative visual features and representations.

II. RELATED WORK

[8] In this study, author shows how the baseline VGG16 model may be fine-tuned to distinguish between daisies, dandelion, sunflowers, roses, and tulips. In order to train the optimized VGG16 model, the authors employed 3520 photos of flowers, resulting in the validation set accuracy of 97.67% and the testing set accuracy of 95.00%. The

"suggested fine-tuned VGG16 model" was validated, trained, and tested using data from the Kaggle dataset. The authors set out to prove that the VGG16 deep model for the image classification pre-trained on the ImageNet can be successfully applied to other picture datasets using a very short dataset without resorting to overfitting. The VGG16 model utilized 3x3 filters, which is a small size that is efficient in image classification. In summary, the authors demonstrate the effectiveness of the fine-tuned VGG16 model for flower classification with high accuracy, despite using a relatively small dataset. The study also illustrates that transfer learning with pre-trained models, such as VGG16, can be used for classifying different image datasets with smaller training data, which can save time and computational resources.

[9] The author of this study discusses the critical public health situation brought on by the fast spread of COVID-19. The World Health Organisation (WHO) has released recommendations stressing the necessity of mask use as a preventative strategy in public spaces and congested regions. However, it is difficult to manually check compliance with mask use in such settings. The author suggests a transfer learning approach to automate the detection of non-masked individuals. The model is built by refining a pre-trained version of the "state-of-the-art deep learning model", InceptionV3. In order to train and evaluate the proposed model, the SMFD (Simulated Masked Face Dataset) is used. To compensate for a lack of data, picture augmentation methods are used to better train and test the model. The experimental findings show that the suggested model is superior to other recent proposals. The model achieves 99.9% accuracy during training and 100% accuracy during testing.

[10] The author is presenting the new deep neural network architecture for the classification of microscopic images. With transfer learning, they suggest combining characteristics from 3 deep CNNs that have already been trained in order to train two fully-connected layers that will be utilized for the classification. Experiments performed using 2D-Hela using "PAP-smear datasets" demonstrate that their suggested network architecture is superior to that of a neural network trained with features from the single CNN and conventional classification techniques. The paper highlights the potential benefits of using transfer learning and concatenating features from multiple pre-trained CNNs for improving classification accuracy in microscopic image analysis..

[11] The author of this paper is presenting their work on developing an algorithm that can efficiently detect COVID-19 from chest X-rays with a high level of accuracy. They have used a deep learning architecture called K-Efficient Net, which is an extension of Efficient Net, and have incorporated progressive resizing into their training process. K-COVID is a massive dataset comprised of X-ray pictures of patients with pneumonia or COVID-19 and normal X-ray images obtained by combining six existing public datasets. The authors have trained their model using transfer learning on ImageNet dataset along with data augmentation approaches. Their method is able to identify COVID-19 with a 97.3% accuracy, a 100% sensitivity, and a 100% positive predictive value. The algorithm is a potential tool towards COVID-19 diagnosis since its results exceed those of practising radiologists.

[12] Early detection of breast cancer cells using the convolutional neural networks (CNNs) is proposed in this research. Mammography pictures will be classified as benign or malignant in an effort to improve breast cancer diagnosis accuracy. The research applies many neural network designs, including DenseNet-201, Inception ResNet-V3, NasNet-Large, and Big Transfer (M-r101x1x1), on "Break His 400X dataset" obtained from the Kaggle. The "M-r101x1x1 architecture" is the most precise of these examples, with an accuracy of 90%. The primary focus of this study is to determine which neural networks are most suited for accurately classifying breast cancer. The study's overarching goal is to better equip doctors with the information they need to make educated diagnoses of the breast cancer at an earlier stage.

III. PROPOSED METHODOLGY

The proposed methodology has been illustrated in Fig.1

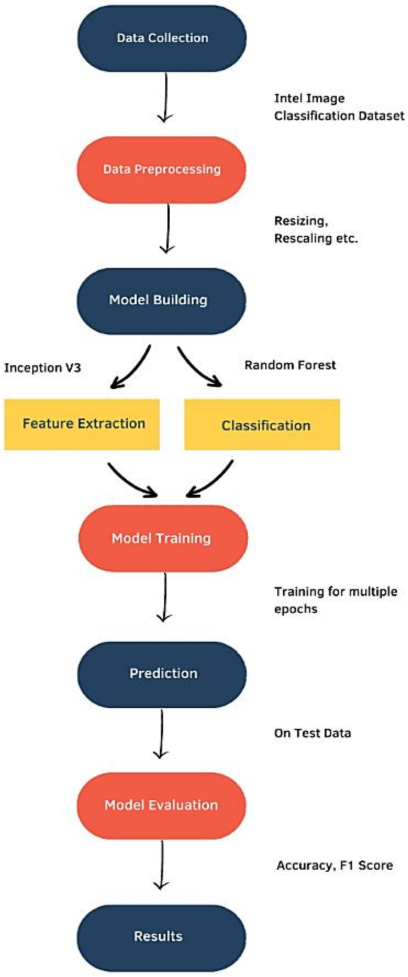

Figure 1 Flow Chart of Proposed Model

A. Dataset

The "Intel Image Classification dataset" from the Kaggle website was used for this study. The dataset is very famous in the area of picture classification because of its unique and difficult characteristics. It's helpful for testing and refining picture categorization methods. The 'Intel Image Classification' dataset specifically focuses on image classification tasks and contains a collection of labeled images. About 25,000 photos representing various classes are included in the dataset, which may be used to test various machine learning models with image categorization. Obtaining the dataset via Kaggle assures that the study results may be readily compared and duplicated by various researchers in the area since it is a standardised and widely-used dataset. Such datasets, when made available by trusted organisations, allow for more thorough examination and promote teamwork among researchers. In order to hold an Image categorization Challenge, Intel first released this data on https://datahack.analyticsvidhya.com.

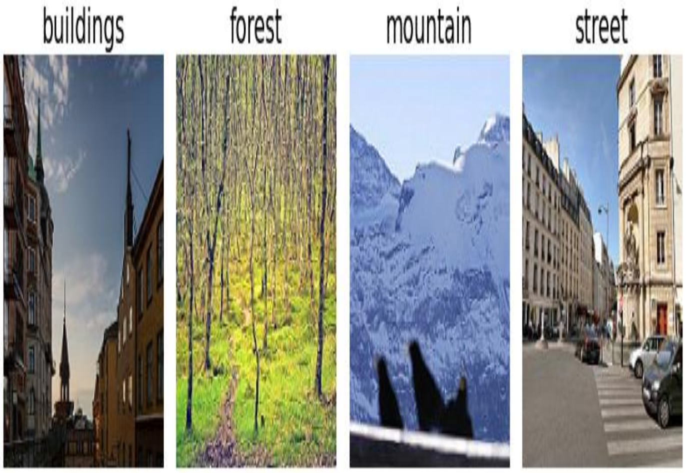

Figure 2 Dataset Sample

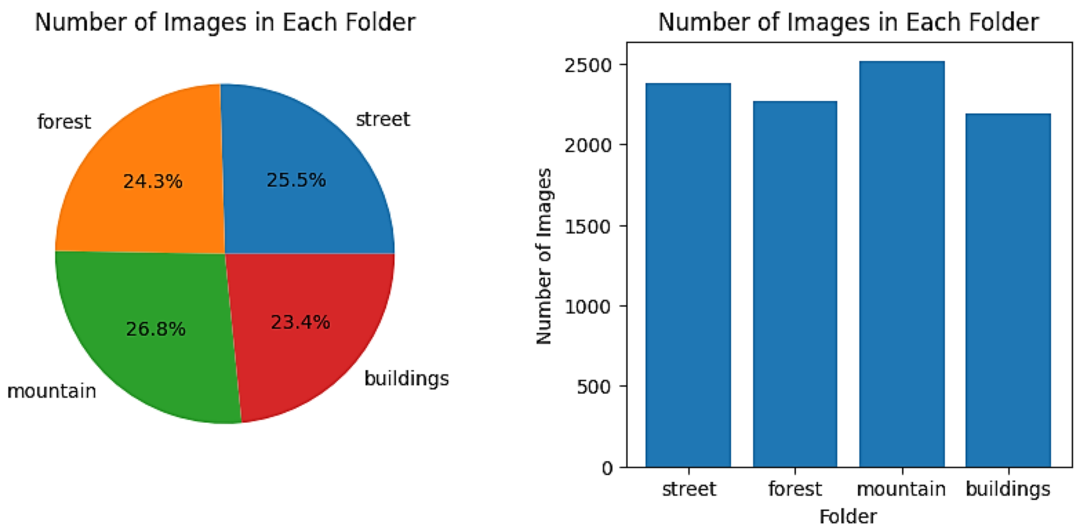

Figure 3 Data Distribution

Overall, the data collection process in this research involved careful selection and cleaning of the data to ensure that it was of high quality and suitable for the research question at hand.

B. Image Pre-Processing

In order to facilitate efficient transfer learning as well as image classification, we preprocessed the "Intel Image Classification" dataset in a number of ways. In machine learning and statistics, overfitting occurs when statistical models describe random noise or errors instead of potential relationships [13]. Datasets were divided, photos were resized, decoding and rescaling were performed, and several data augmentation methods were investigated throughout this process. Detailed explanations of each of these procedures follow:

Dataset Splitting: The original "Intel Image Classification" dataset is divided into a training set and a testing set. The goal of the separation was to evaluate the suggested model's efficacy and scalability. The model was trained on the training set to understand the underlying structure of the data. The accuracy of the model was tested on fresh data not included in the training set by using the testing set. This division allowed for objective testing and verification of the model.

Resizing of Images: The photos in the dataset were scaled to a certain dimension to guarantee consistency and compliance with the chosen pre-trained model (InceptionV3). Our studies used photos reduced in size to 299 by 299. The InceptionV3 model required this particular size for its inputs, thus that's what was used. Resizing the images to a consistent size allowed for consistent input dimensions during the training and testing phases.

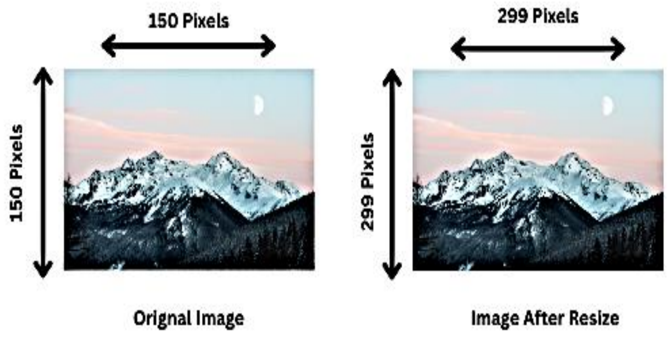

Figure 4 Image resizing

Decoding of Images: The original images in the dataset were in various formats, such as JPEG or PNG. As a pre-processing step, these images were decoded, converting them into arrays or matrices suitable for further analysis. This decoding process involved extracting the pixel values and converting them into a numerical representation.

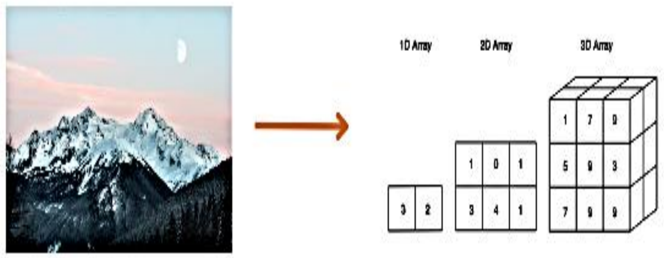

Figure 5 Image decoding

Rescaling of Images: To improve the generalization capabilities of the model, the pixel values of the images were rescaled. Rescaling ensures that the pixel values are normalized within a specific range, typically [0, 1] or [-1, 1]. In our research, a normalization technique was applied to rescale the pixel values to the range [0, 1]. To help the model learn more successfully, rescaling pixel values may reduce the influence of differences in pixel intensity range across photos.

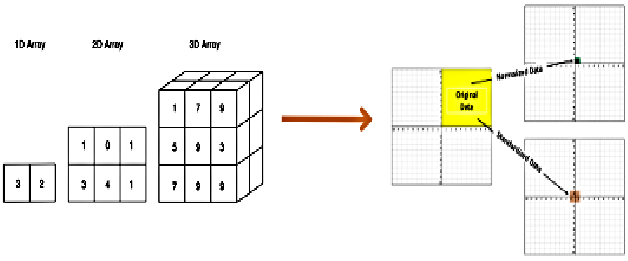

Figure 6 Image rescaling

By performing these data preprocessing steps, we ensure the dataset is appropriately formatted, standardized, and optimized for use with the selected models and algorithms. These steps help enhance the compatibility, quality, and representation of the dataset, ultimately leading to more accurate and reliable results in the subsequent stages of the research.

C. Methodology

This study's suggested methodology integrates transfer learning with the Random Forest technique to improve the precision of picture categorization. The model employs the Random Forest technique to efficiently classify pictures using characteristics extracted with the help of transfer learning. Choosing the InceptionV3 is pre-trained model for the "feature extraction" kicks off the model building process. Strong performance in image classification tasks has been shown by InceptionV3, a "deep convolutional neural network architecture" trained on a large-scale dataset. "High-level features from input" photos are extracted using the InceptionV3 model's pre-trained weights to capture the images' discriminative properties.

The collected features are then sent into the Random Forest algorithm, a well-liked ensemble learning technique known for its tolerance to noisy data and its capacity to handle high-dimensional features spaces. To achieve reliable classifications, Random Forest builds a collection of decision trees & averages their predictions. The suggested model is able to accurately categorise pictures into their respective categories since it uses the retrieved characteristics as inputs to Random Forest. The benefits of employing Random Forest come from the fact that it can manage nonlinearities within the data and can capture complicated correlations between attributes and class labels. This feature guarantees the model's efficiency and robustness by making sure it applies effectively to data it has never seen before. There are several upsides to the suggested paradigm. First, the model is able to extract significant and representative characteristics from photos because it uses transfer learning to apply the information it has gained from a large-scale dataset. As a result, you won't have to spend as much time training on as little domain-specific data. Second, for dependable predictions in image classification applications, the Random Forest method offers a robust as well as accurate classification mechanism. Finally, by using the benefits of both methods, transfer learning & Random Forest improve the model's overall performance.

The proposed model holds promise for improving image classification accuracy by effectively leveraging transfer learning and the Random Forest algorithm. Through its ability to capture informative features and make accurate predictions, the model has the potential to advance image classification tasks in various domains, including computer vision, healthcare, and autonomous systems.

IV. RESULTS AND DISCUSSION

This study's findings support the hypothesis that the suggested approach may significantly enhance the accuracy with which images are classified. The model's successful classification of 95.64% of the test photos demonstrates its versatility. This impressive precision is all the more remarkable given the size and diversity of dataset used. The model's efficacy may be better grasped with the help of the confusion matrix. A tabular display of model's prediction, the confusion matrix shows examples of both successfully and incorrectly labelled pictures.

High values for the diagonal of the confusion matrix suggest thatmost of the photos were properly categorised. However, there are also some obvious misclassifications. In these examples, the model incorrectly predicted the class labels on the photos. We can learn a lot about the limits of proposed model as well as possibilities for development by analysing these incorrectly labelled photos. By delving into the root causes of these false positives, we may better tailor the model's architecture, adjust its parameters, or put in new data preparation methods to boost its accuracy.

Figure 7 Model Predictions

Below is the summary of the results in tabular form:

Table 1 Model Comparison without Image Augmentation

Metric	Value
Accuracy	95.93%
F1 score	96%
Precision	96%
Recall	96%

The total accuracy of the suggested model on test data is 95.93%, as shown by these findings. The F1 score is measure which takes into account both accuracy and recall, making it a more well-rounded tool for assessment.

Here are visual results of proposed mode –

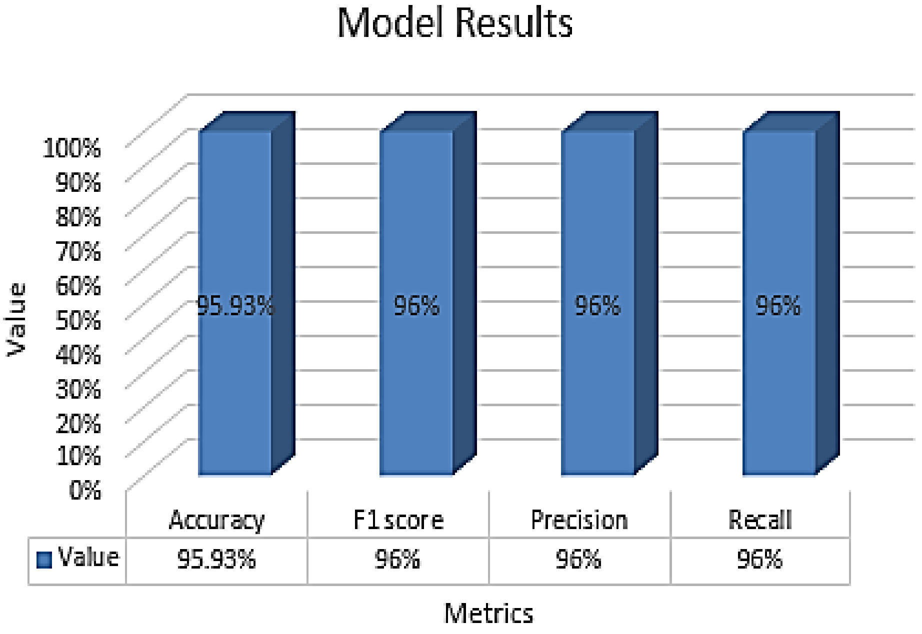

Figure 8 Results

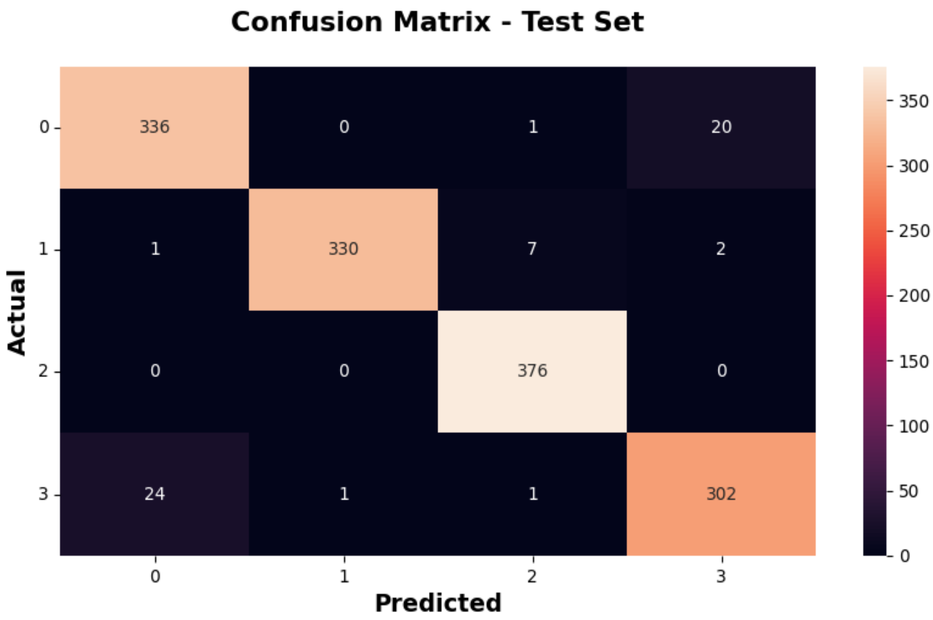

Figure 9 Confusion Matrix

According to the confusion matrix, the model performed very well at predicting classes 0–1, and 2 by accurately labelling the vast majority of occurrences. Class 3 had the highest number of misclassifications, with 24 cases mistakenly predicted as class 0. Classes 1 and 2 also had 1 inaccurate prediction apiece.

The existing work utilized the ResNet model in combination with the SVM (Support Vector Machine) algorithm for image classification. The highest accuracy achieved in the existing work was 93.57%. This indicates that the combination of ResNet and SVM was able to achieve a respectable level of accuracy in classifying the images. In contrast, our proposed work employed the InceptionV3 model for feature extraction and the Random Forest algorithm for classification. The accuracy obtained in our research was 95.93%.

Table 2 Base Paper Comparison Table

	Existing Work [14]	**Proposed Work**
Model	ResNet	InceptionV3
Algorithm	SVM	Random Forest
Highest Accuracy	**93.57%**	**95.93%**
Performance	Respectable	Significantly improved

The precision of our proposed work is far higher than that of the prior art. We improved upon the previous work that used ResNet and SVM (93.57% accuracy) by employing InceptionV3 as well as Random Forest (95.93%) to train our model. This demonstrates how well our method works to boost picture classification results.

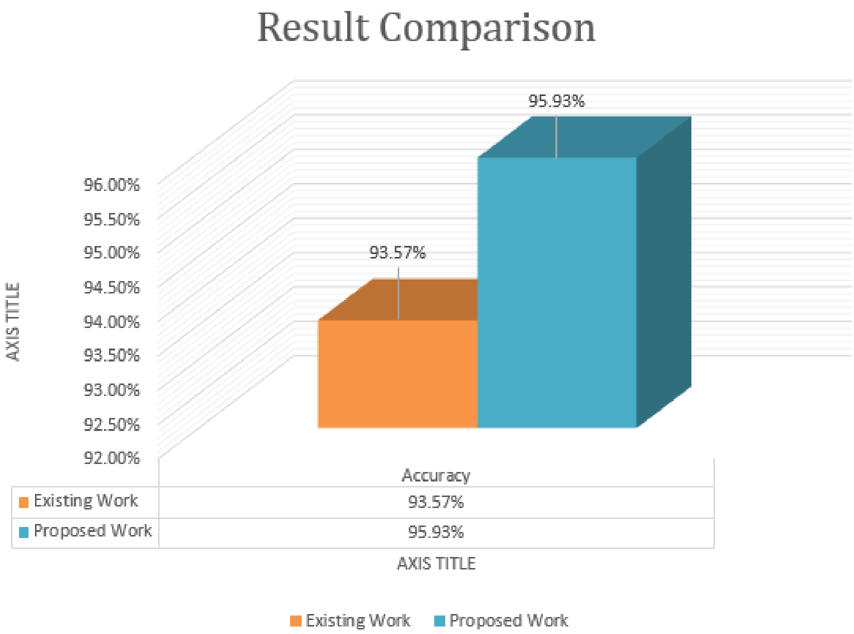

Figure 10 Graph result of proposed work

V. CONCLUSION

In conclusion, this study aimed to improve image classification accuracy by proposing a novel approach that utilized the InceptionV3 model for feature extraction and employed Random Forest as the final classification model. The results of the study show that the suggested model works as intended, with an amazing accuracy of 95.93% using the test data, which is higher than the accuracy of 93.57% in the current work. Through a rigorous methodology, various pre-trained models were explored, and InceptionV3 was selected based on its superior performance in feature extraction. The extracted features were then fed into Random Forest, which proved to be the most promising machine learning algorithm for the task. Precision, recall, and F1-score, three assessment measures, all corroborated the proposed model's stellar performance and demonstrated its balanced classification performance throughout classes. While this study has made significant progress in improving image classification accuracy using the proposed model, there are several avenues for future work that can expand upon the research and explore new possibilities.

REFERANCE

[1] Wang, I-Hung, Kuang-Chyi Lee, and Shinn-Liang Chang. "Images Classification of Dogs and Cats using Fine-Tuned VGG Models." 2020 IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE). IEEE, 2020

[2] Suganthi, M., and J. G. R. Sathiaseelan. "An exploratory of hybrid techniques on deep learning for image classification." 2020 4th International Conference on Computer, Communication and Signal Processing (ICCCSP). IEEE, 2020.

[3] Pan, Sinno Jialin, and Qiang Yang. "A survey on transfer learning." IEEE Transactions on knowledge and data engineering 22.10 (2010): 1345-1359.

[4] Zeiler, Matthew D., and Rob Fergus. "Visualizing and understanding convolutional networks." European conference on computer vision. Springer, Cham, 2014.

[5] He, Kaiming, et al. "Deep Residual Learning for Image Recognition." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[6] Simonyan, Karen, and Andrew Zisserman. "Very Deep Convolutional Networks for Large-Scale Image Recognition." arXiv preprint arXiv:1409.1556 (2014).

[7] He, Kaiming, et al. "Deep Residual Learning for Image Recognition." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[8] Giraddi, Shantala, et al. "Flower classification using deep learning models." 2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE). IEEE, 2020.

[9] Jignesh Chowdary, G., et al. "Face mask detection using transfer learning of inceptionv3." Big Data Analytics: 8th International Conference, BDA 2020, Sonepat, India, December 15–18, 2020, Proceedings 8. Springer International Publishing, 2020.

[10] Nguyen, Long D., et al. "Deep CNNs for microscopic image classification by exploiting transfer learning and feature concatenation." 2018 IEEE international symposium on circuits and systems (ISCAS). IEEE, 2018

[11] Diallo, Papa Abdou Karim Karou, and Yun Ju. "Accurate detection of COVID-19 using K-EfficientNet deep learning image classifier and K-COVID chest X-ray images dataset." 2020 IEEE 6th International Conference on Computer and Communications (ICCC). IEEE, 2020.

[12] Islam, Md Ashiqul, et al. "Forecast breast cancer cells from microscopic biopsy images using big transfer (BiT): a deep learning approach." International Journal of Advanced Computer Science and Applications 12.10 (2021).

[13] Hawkins D. The Problem of Overfitting. J Chem Inf Comput Sci. 2004;44(1):1-12. doi:10.1021/ci0342472

[14] Mahajan, Arpana, and Sanjay Chaudhary. "Categorical image classification based on representational deep network (RESNET)." 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA). IEEE, 2019.