International Journal of Computer Trends and Technology
ISSN: 2231–2803 / https://doi.org/10.14445/22312803/IJCTT-V71I7P105
Volume 71 Issue 7, 29-43, July 2023

Original Article

Image Recognition from Deep Features using Implicit Split Logarithmic Loss-Based Optimized Random Forest

Abhinav Babbar[1], Pooja Gupta[2], Alok Gupta[3]

[1,3]*Department of Computer Science and Engineering, Optum Global Advantage, India.*
[2]*PGT Computer Science, Government of Haryana, India.*

[1]*Corresponding Author : abhinav.babbar@gmail.com*

Received: 25 May 2023 Revised: 29 June 2023 Accepted: 12 July 2023 Published: 27 July 2023

***Abstract** - Image recognition collects meaningful data extracted from an image and performs various machine-based visual tasks. It is used in numerous applications such as self-driving cars, guiding autonomous robots, accident avoidance systems, and labeling the content of images with mega tags. It is considered a sub-category of digital technology that deals with regularities in image data, recognizing patterns, and then classifying them into categories by interpreting image pixel patterns. The traditional methods involved in image recognition perform sequential procedures like collecting, analyzing, and categorizing images manually, leading to time consumption and misclassification. So, ML algorithms such as instance-based, clustering, DT (Decision Tree), and others evolved to predict the images. But these methods involved over-fitting problems and uncertain outcome values leading to decreased accuracy. Hence the proposed system aims to improve the accuracy of predicting the images with suitable DL (Deep Learning) and ML (Machine Learning) based algorithms. The feature extraction is accomplished by implementing CNN (Convolutional Neural Network) algorithm. As CNN requires a large amount of training time and limited effectiveness for sequential data, an ORF (Optimised Random Forest) is employed. ISLLF (Implicit Split Log Loss Function) is used with ORF to predict the system's behaviour. The ISLLF is implemented during the training process and is used for optimizing classification models. The effectiveness of the respective model is calculated through F1-score, accuracy, recall, precision, confusion matrix, and ROC-AUC score. The proposed model is further compared with existing models to determine the proposed method's efficiency.*

***Keywords** - Image recognition, Feature extraction, Deep Learning, Convolutional Neural Networks, Classification, Machine Learning, Random forest.*

1. Introduction

In the era of modern technology, a huge amount of image data is generated every day. Subsequently, recognition of images is essential for various factors like identification, detection, and classification of the image data. It can be attained by image recognition technology that processes the images faster and more precisely than manual image inspection. It identifies features or objects in the image data with the help of software functions. There are numerous image recognition applications such as health care, agriculture, robotics, e-commerce, digital marketing, etc. Image recognition is the fundamental equipment of these applications. The recognition of objects in images and uses that data to create choices as part of a greater system. It serves these systems and becomes conscious, basically allowing better decisions by providing insight into the approach. Various procedures for image recognition, together with ML [1-3] and DL (Deep Learning) [4] techniques, are widely used. DCNN (Deep Convolutional Neural Network) [5-7] techniques tend to serve a non-iterative learning approach to retrain neurons at fully connected layers resulting in better performance. Traditional image recognition technology at present cannot attain fast real-time high-accuracy performance.

Accordingly, numerous techniques in image recognition like AI (Artificial Intelligence), ML, and DL is used, where DL has revolutionized traditional image recognition techniques by attaining larger accuracy in errands like image classification, object detection, semantic segmentation, and SLAM (Simultaneous Localization and Mapping). Moreover, DL models are recently emerging as promising tools to achieve performance. Major improvements in DL and enrichments in device abilities, including image sensor resolution, power consumption computing power, memory capacity, and optics, have augmented the spread of vision-based applications with cost-effectiveness and upgraded performance. The accessibility of an enormous amount of video data in the system supports this issue. While computer vision algorithms tend to be domain-specific, as well as it also offers enhanced tractability, as CNN [8] techniques with backgrounds are retrained with conventional input data. Numerous models are used for image classification, such as data augmentation, CNN, embedding visualization of fashion MNIST dataset, and PCA (Principal Component

Analysis). Correspondingly, a reconstruction-based, semi-supervised DNN for image irregularity discovery is used. Here the network is trained on normal data to enhance a normality method by mapping the images in a low-dimension feature space method [9].

Though several approaches are implemented in image recognition, they produce some drawbacks such as accuracy, high computational cost, not being compatible with larger datasets, and others. Behalf of reducing these limitations, the proposed system employs DCNN in feature classification and ORF for image classification combined with ISLLF. As CNN produces better feature extraction ability, it is widely used in computer vision. With the increase in network architecture optimization, the performance of DCNN is boosted significantly. Further, CNN enhances kernels that map image characteristics to dimensionality, allowing improved memory management and maintaining relevant image information. Early studies mostly concentrate on three dimensions such as depth, width, and cardinality. But the proposed system depends on corners and edges for feature extraction. Computers use the CIFAR-10 dataset to recognize objects, as the images are labeled with one of 10 mutual exclusive classes.

Further classification process involves an RF algorithm that uses multiple decision trees. By using this method, RF outperforms other classification methods by achieving better prediction accuracy and generalization by avoiding over-fitting problems. It is also more robust to training samples and predicting noise in training datasets. Further, the classifier is evaluated using an implicit split log loss function for performance analysis of the proposed system. If the dataset is smaller, then classification based on CNN produces greater difficulty, as they learn to recognize each pattern with large examples for training samples. Hence RF is used for image classification since they use split use method and assists in producing better accuracy.

The main contributions of the proposed method are as follows:

- To perform feature extraction on the CIFAR-10 dataset using CNN.
- To implement classification of CNN-based feature extracted images, using ORF with ISLLF for evaluation of binary classification
- To calculate the efficacy of the respective system by comparison with conventional works based on precision, accuracy, F1-score recall and confusion matrix, and ROC-AUC score for optimized image recognition.

1.1. Paper Organisation

The paper is categorized based on effective feature extraction and image classification methods. Whereas the existing works are analysed on similar domains with different approaches, as shown in Section II. Further, Section III signifies the methodology executed in the proposed system. Sequentially the results and outcomes attained by the proposed method are shown in Section IV. Lastly, the conclusion and future work of the proposed model is determined in Section V.

2. Review of Literature

The review of existing methods based on different approaches for optimized feature extraction and image classification is discussed in this section. Image recognition is the ability of a machine vision to predict the objects represented in digital format. The recommended paper has used the DL algorithm for image classification. Here, the dataset utilized is Cityscapes. The technique results have stated that the system has produced better outcomes [32]. The suggested paper has deployed a CNN-based feature extraction of GBC (Ground Based Cloud) images. The dataset utilized in the approach is TSI-880 (Total Sky Imager). The data has been taken from the NREL (National Renewable Energy Laboratory) database. Experimental results of the recommended method have stated that the method could automatically extract features of GBC images that are utilized for forecasting DNI and has produced better accuracy [11]. The suggested approach has utilized 2D-3D CNN for the classification of images by the multi-branch feature fusion method.

The dataset developed in the suggested method is the HIS (Health Information System) dataset. This method has attained ideal classification with a smaller number of training data. The results of this technique have stated that the method has produced good results with better accuracy [12]. The mentioned approach has used CNN in extracting the depth features of various kinds of SAR gray images. It has been further classified with the ELM algorithm. The experiment has shown that the recommended approach has attained good results with better accuracy[13]. The suggested technique has employed a lightweight framework integrated with edge computing and mobile sensing. The dataset utilized in the suggested approach is the MNIST dataset. The feature extraction of the model has utilized CNN over the encrypted data. The recommended approach has intended for low latency in the network and low overhead in mobile devices. The result of the approach has attained better accuracy [14].

The intimated approach has employed the AIA (Automation Image Annotation) technique with CNN feature extraction to resolve a semantic gap between the original image and its semantic information. The dataset utilized in the suggested approach is Crel15k and laprtc12. The method splits the image in means of similarity into a semantic neighbor group. Thus results of the mentioned approach have attained good results with better accuracy [15, 16]. Similarly, the suggested approach has implemented CNN for the feature extraction of the image. Data sets are from two sources, such as ASH image with ALL-IDB. It has also explored other ML algorithms like SVM (Support Vector Machine), K-NN (K-Nearest Neighbour), Naïve Bayes, and DT (Decision Tree) for evaluating the performances. Therefore, the experiment results have attained enhanced results with 88.25% accuracy [17]. Correspondingly, the insisted approach has used DL based framework to identify images. It is attained by the combination of CNN with LSTM (Long Short-Term Memory). CNN [18] has been utilized for the feature extraction of the image, while LSTM is utilized to capture sequential information and identification of features in the input data. Thus results of the method have produced optimized results with 91% accuracy [19]. The recommended approach has implemented DCNN (Deep CNN) for feature extraction and image data identification with six stages of data augmentation techniques. The data set which has been utilized in this technique is the open dataset that consists of 39 diverse classes of plant leaves. Simultaneously, the method has been trained with diverse training epochs, dropouts, and batch sizes. It has been further evaluated with the other transfer learning approaches. Thus the result has provided good outcomes with 94.46% accuracy [20]. Mutually, the suggested approach has also utilized DCNN-based decision fusion, for the identification of coronavirus disease by the input image data. Here the dataset used is a set of CT images. In this, decision fusion has been utilized in the approach, by fusing the identification from the various individual models. So, the results of the experiments has shown that the approach has achieved 86% accuracy in identification results [33].

The suggested approach has represented 13 layers of CNN for identifying the fruit species with the given input image data. The dataset obtained in the method is the 3200 fruit image dataset. Nearly three kinds of data augmentation methods have been utilized, such as noise injection, image rotation, and gamma correction. Moreover, the maximum pooling and average pooling have been compared to produce efficient precision. As a result, the system has attained 94.94% accuracy [22]. Likewise, the suggested paper has compared three image classifications of IoT (Internet of Things) edge devices with several ML algorithms. Five kinds of the dataset have been used in this process, such as MNIST Digits, Fashion-MNIST, CIFAR-10, Chest X-Ray, and Faces in the Wild. It has discovered the similarity between the various image classification like image resolution, algorithm phase, and device hardware and algorithm type. Therefore, the results of the procedure experiment show that the RF has performed better than the other two algorithms [23]. The intimated paper has employed an object-oriented RF scheme for forest image classification. Here, the dataset used is the GF-2 remote sensing image data that has been taken from the Laoshan area of the Maoershan forest. Where the optimal feature space has been constructed based on spatial characteristics of the spectrum texture, terrain, multi-scale segmentation, and vegetation index. Thus the performance has been compared with the SVM classifier for evaluation. The experiment results have shown that the system has attained 83% accuracy with better results [24].

Similarly, the insisted method has used an RF algorithm to classify land cover image data. The dataset utilized in the approach is the sentinel SAR imagery dataset. Besides, the process of classification has been developed by Python and cloud computing for statistical analysis. As a result, the experiment's outcomes have stated that the method has produced good results with better precision [25]. Even a multi-scale segmentation algorithm for extracting large-scale characteristics data based on the hierarchical structure of the input image data has been used. It has been attained by the RF technique. Therefore, the result of the experiment determines that the recommended approach has produced better results with 86% accuracy [26].

Further, PCNN (Pulse Coupled Neural Network) and SORF algorithm for seabed sediment classification by the image data has been implemented [27]. The dataset used in the suggested paper is the image data of the sediments in the region of Jiaodong peninsula, China. Moreover, the experiment results have stated that it produced optimization with 95% accuracy. The recommended paper has deployed RF for the organization of forest varieties in the region of northwest Gabon. Despite this, this paper has aimed at the valuation of Sentinel-2 satellite images as well as RF classifier [28] to map forest cover and forest varieties in northwest Gabon.

Additionally, the motive of the system has to predict the impact of several spectral bands by the Sevtinel-2 satellite, DEM (Digital Elevation Model), and NDVI (Normalised Difference Vegetation) [29]. The classification has been achieved by using the RF classifier. The overall accuracy for the intimated system has been found to be 97.4%.

2.1. Problem Identification

The major issues found by the analysis of conventional studies are,

- From the study, it is inferred that the CNN model used in the system is time-consuming and not stable throughout the classification process of medical images [15].
- Overall accuracy decreased by using imperfect images. The system validated CNN based on only the convolutional layer and pooling layer but not on the fully-connected layer [20].
- The fully connected layers used in the method consume more training parameters. This system decreases the generalization ability of the system [11].

3. Proposed Methodology

Image recognition is a supervised learning, which recognizes objects and trains a model to identify them using labeled example photos. However, raw pixel data alone does not offer adequate representation and thus produces some variations with misclassification. Hence, the proposed system involves DL and ML algorithms in feature extraction and image classification. CIFAR-10 dataset is used, as it contains RGB (Red, Green, Blue) colored images with a resolution of 32x32 pixels of mechanical and live objects. There are a total of totally 10 classes in the dataset, such as 6 animals (birds, cats, deer, dog, frog, horse) and 4 inanimate classes (airplane, automobile, ship, truck).

Further, the dataset is loaded with 6,000 images per class, with 50,000 images in the training set and 10,000 images in the testing set. Next, the data pre-processing method is involved to remove adversarial noise from the provided dataset. This method helps to improve the quality of the image. Since DCNN is an efficient method for handling huge volumes of datasets, it is implemented for feature extraction of input images. Also, it helps reduce data dimension and produce less redundant datasets. It also helps in extracting target image features automatically and makes them clear visibility. The neural network includes sets of convolutional layers and pooling layers. Here the convolutional layer transforms the image by convolution process (purpose filter effect). This layer is represented as a series of digital filters. The pooling layer transforms the neighboring pixels into a single pixel and then decreases the image dimension. The output from the CNN is then fed to the training and testing data model. The first part is the training dataset, which is used to train the first interference detection of the CNN model.

Then the test model is used to evaluate the extraction performance. Then the trained dataset is fed to classifiers to classify extracted features. An ML algorithm called optimized RF is a technique implemented to classify the feature-extracted dataset. Generally, the RF algorithm is an ensembling learning method implemented to increase the efficiency of binary classifiers by avoiding misclassification and over-fitting problems. The ISLLF is a cross-entropy loss used with ORF to evaluate prediction quality in binary classification. It is employed to predict the accuracy of binary labels and reduce the variation between actual and anticipated outcomes. The CNN and RF work with the train models, but the test data samples are used in the prediction phase. As a result, the system's performance is predicted in terms of accuracy, precision, recall, F1-score, and confusion matrix.

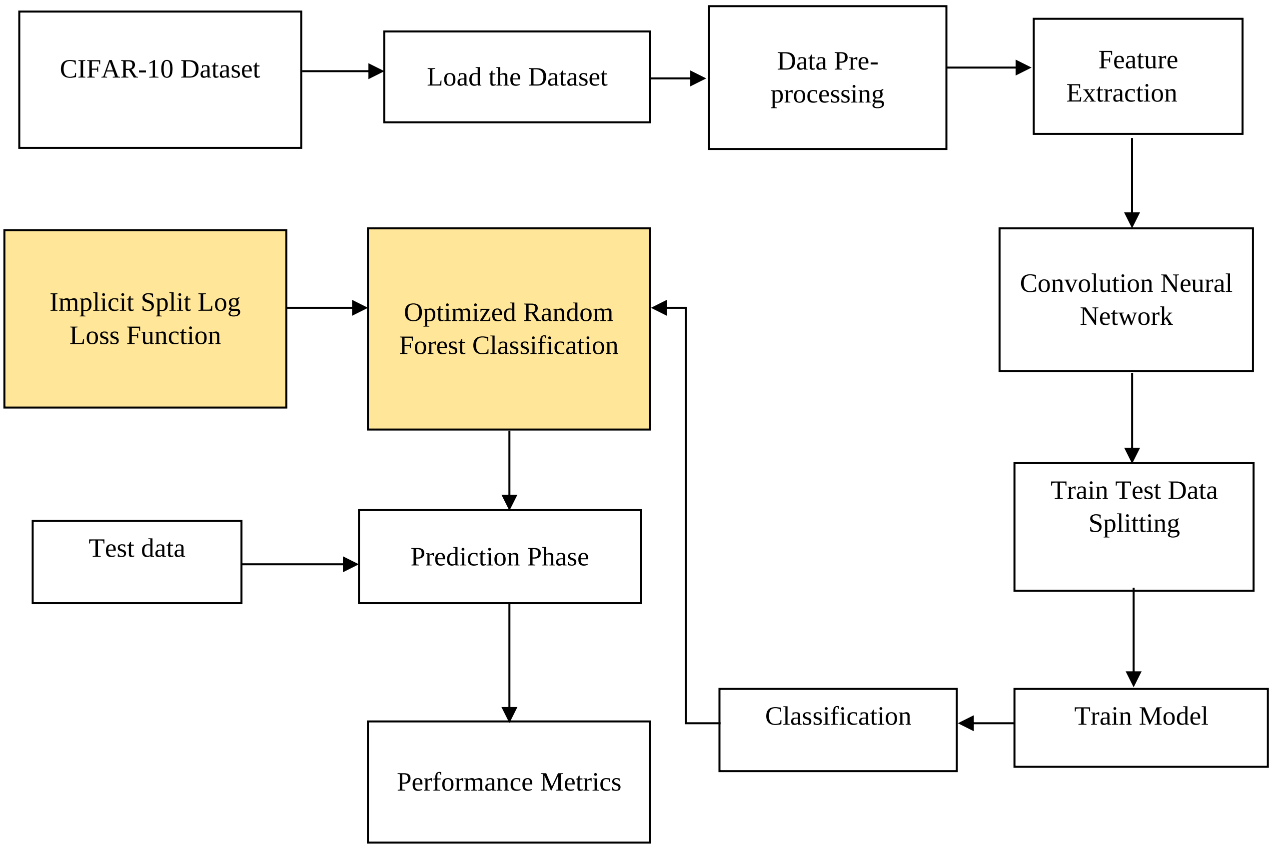

Fig. 1 Overall process involved in image recognition

3.1. Feature Extraction using CNN (Convolutional Neural Network)

The proposed system considers the CNN as they portray highly-layered structured NN combined with convolutional, pooling, and classification layers. Initially, the input images are pre-processed to typical normalization; afterwards, the data stream into a set of convolutional layers accompanied by pooling layers in which the redundancy reduction and feature extraction manifests. Then every attribute is grouped partly, and the subsequent attributes count for a part of the arrangement of the labeled class. Then the selected attributes are directed to a fully-connected layer, and this level delivers an approximation of arrangement. The steps involved in feature extraction using CNN are represented in algorithm I.

Algorithm I: Steps Involved in Feature Extraction Using CNN

Step 1: Load the CIFAR − 10 dataset
using a dataset loader such as cifar10. load_data().
Step 2: Pre − process the data by normalizing the
pixel values to be between 0 and 1.
Step 3: CNN model with the desired architecture,
such as multiple convolutional layers followed
by one or more dense layers.
input shape = (64,64,5), with 50,000 input images
Step 4: Compile the model using a suitable
optimizer = ADAM and loss
= CATEGORICAL CROSSENTROPHY function.
Step 5: Create a new model that takes the original inputs
but outputs the activations of the desired layer
(e. g. , the second − to − last layer).
Here, the con_{5blk} layer is taken to get the extracted features.
Step 6: Use the feature extractor to extract features from both
the training and test sets, using the predict() method.
Step 7: Store the extracted features in suitable variables
(e. g. , train_features and test_features).

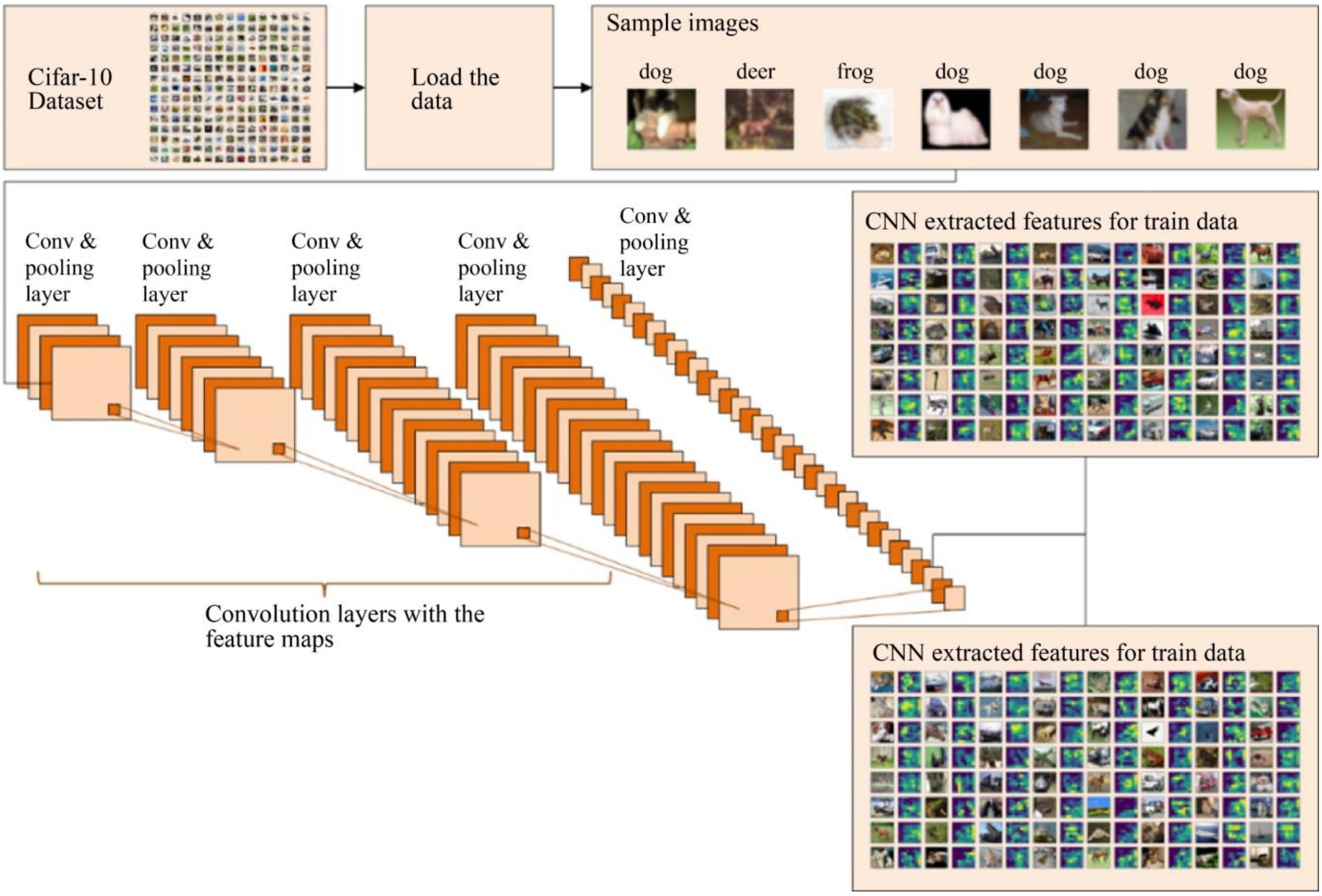

Fig. 2 CNN layer involved in feature extraction using CNN

Three different layers of CNN involved in feature extraction is represented in figure.2. The conventional layer is the core building block, and a large number of computation prevails. The input image consists of three measurements, such as corners and edges, with equivalent RGB consisting in the image. The kernel, a feature detector, is used to transfer across accessible areas of the image, evaluating whether the attribute exists. At every convolution process, the CNN employs a ReLU (Rectified Linear Unit) transformation to feature map, developing non-linearity to the system. The convolutional layer has the ability to follow the initial layer by another layer. So, the assembly of CNN turns hierarchical, as the further layers can view the pixels of prior layers within accessible fields.

Then the pooling layer is involved, decreasing the input parameters. It is related to the conventional layer but does not involve any weights. Instead, the kernel implies an aggregation function to the values, thus collecting the output array. Two methods involved in the pooling layer are maximum pooling and average pooling. The final stage is a fully connected layer where each node in the output layer connects directly to a node in the previous layers. This layer implements classification depending on features extracted by previous layers and their corresponding filters, thus producing a probability of 0 to 1.

3.2. Classification based on ISLLF-ORF (Implicit Split Log Loss Function-Optimised Random Forest)

Random forest is an ensemble technique using a classical decision tree classifier. It categorizes multiple decision trees obtained by randomization of features and observations. It is obtained by selecting each node of trees, k-variables randomly, as candidates for split with the minimum correlation between decision trees. This results in reduced variance and bias. Here, ORF is implemented in the classification of images for better precision and is capable of interacting with huge datasets. It consists of a set of tree-structured algorithms with identically distributed random values. It uses two generalized parameters, such as the interdependence of individual classifiers and exactitude. It also predicts the missing values by using the data handling method. K-classification trees in RF are to set weak classifiers as strong classifiers. The classification tree consists of several nodes, in which the root node indicates a training set, and each internal node indicates a weak classifier that segregates samples based on certain attributes.

The RF relies on many self-learning decision trees since instead of using a single tree; multiple trees can become robust and a strong classifier.

$$\mathrm{num_{tree}}\,,\ \max_{\mathrm{depth}}\mathrm{Tree}\,,\ \min_{\mathrm{samples}}\ \text{to split node}$$

The method to define an implicit split function is input features, and training points try to minimize the heterogeneity of two resulting subsets of data. Thus the ISLLF is given by,

$$\mathrm{L} = \mathrm{sum}(\mathrm{y_{true}} * \text{implicit logloss}\ (\mathrm{y_{pred}})) \qquad (1)$$

Where $\mathrm{y_{true}}$ is true class label (one − hot encoded),

$\mathrm{y_{pred}}$ is predicted class probability

(summing up to 1), and log is the natural logarithm.

During training a classification model, the split log loss should represent the predicted value as equal to the actual value.

Each tree in the random forest works on the basis of major layers of randomness.

- Random selection of a subset of the training data (with replacement) to create a bootstrap sample.
- Grow the tree recursively by selecting the best split among a random subset of features at each node based on the implicit split log loss function.
- Stop growing the tree when either the maximum depth is reached or the lesser amount of images needed to split a node is not met.

At each node, a random subset of CNN extracted features is selected and the feature that maximizes the information gain (based on the implicit split log loss) is chosen as the split criterion.

The information gain is given by:

$$\mathrm{IG} = \mathrm{L(parent)} - \left[\left(\frac{\mathrm{n_{left}}}{\mathrm{n_{total}}}\right) * \mathrm{L(left_{child})} + \left(\frac{\mathrm{n_{right}}}{\mathrm{n_{total}}}\right) * \mathrm{L(right_{child})}\right] \qquad (2)$$

Where L(parent) is the implicit split log loss of the parent note,

L(left $_{\text{child}}$) and L (right $_{\text{child}}$) are the implicit split log loss of the left and right child nodes,

n $_{\text{left}}$ and n $_{\text{right}}$ are the number of samples in the left and right child nodes,

n $_{\text{total}}$ is the total number of samples in the parent node.

In case of a log loss value closer to 0, the system behaves better, as it is the measure of uncertainty. So the value of log loss should be as low as possible. The ORF method used in image classification is validated based on ISLLF. During training, a classification model should represent the predicted value equal to the actual value. So, log loss evaluation is considered a better method for the loss function during optimizing and training classification models. The system's performance is predicted based on the probabilities of 0 to 1. Algorithm II represents how the ORF classifies the images and predicts the performance with ISLLF.

The following are the steps involved in image classification based on ISLLF-ORF.

Step 1: Predit the class probabilities for each test saple by averaging the predicted class probabilities of all trees in the random forest.

Step 2: Calculate the implicit split log loss on the test data using the predicted class probabilities and the true class labels.

Step 3: Optimize the hyperparameters of the random forest (the volume of trees, increased depth, a lesser amount of images needed to split a node) Using cross – validation to minize the implicit split log loss.

Step 4: Train the final random forest on the entire training data

Step 5: Calculate the efficacy of the final RF with test data using F1 – score, precision, accuracy, recall, and confusion matrix.

Algorithm II: ISLLF-ORF (Implicit Split Log Loss Function-Optimised Random Forest)

Input data: Extracted feature from dataset efd $= (a_1, b_1) \ldots \ldots . (a_n b_n)$,
features feat, number of Split trees in forest C
function Random forest(efd, feat)
Group extracted features with its multiple labels
calculate wei, the vector of weights that are assigned to the split of trees
randomforest classifier(rfc) $\rightarrow$ α
for i belongs to 1, … C do
efd(i) → A random samples from efd
$rfc_i \rightarrow Implicit_{RandomizedTreeLearn}(efd(i), feat)$
RFC—→ RFC ∪ of rfc_i
end for
return RFC
end **function**
Function $Implicit_{RandomizedTreeLearn}(efd, feat)$
while existing conditions are not met for every branch
do for every remaining tree, do
f → rand subset of feat
split on the best features in f
end for
end while
return the implicit split tree
end function

4. Results And Discussion

The results produced by implementing the proposed system are discussed in this section. Additionally, dataset descriptions, performance metrics, identified outcomes, and comparative analyses of conventional studies are presented.

4.1. Dataset

The study has considered the CIFAR-10 dataset, wherein the dataset encompasses 60,000 color images in 10 classes, such as 6 classes of animals and 4 classes of inanimate objects. The dataset contains 50,000 training images and 10,000 testing images.

4.1.1. Sample Dataset

The dataset consisting of 60,000 images of 10 classes, is taken as input and processed for image recognition. This type of dataset quickly adapts to the implementation of different algorithms. Hence, various images of animals and non-living things are used in the image recognition process.

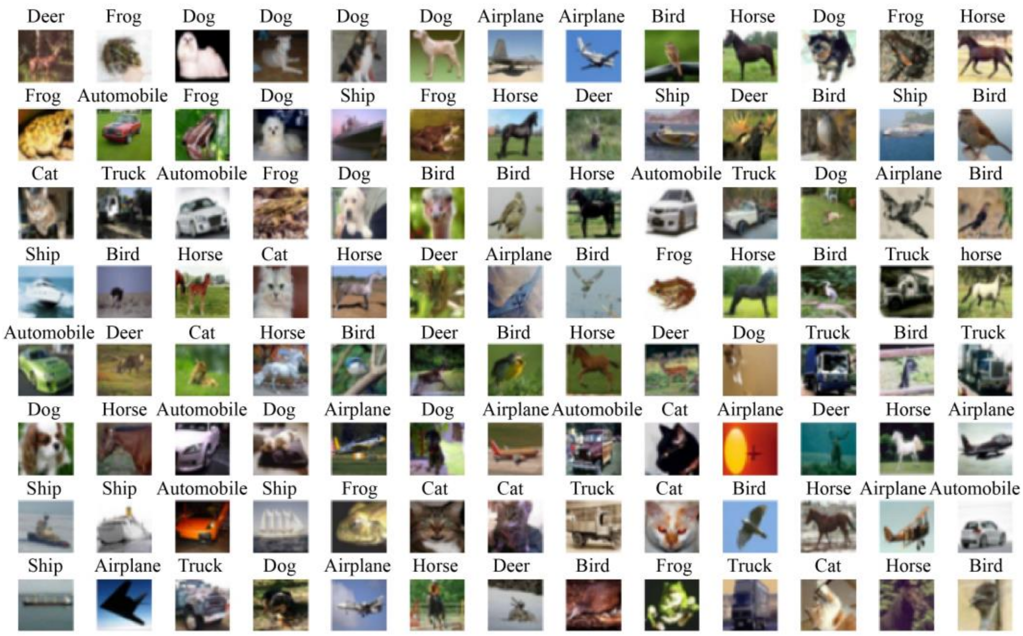

Fig. 3 Sample images of dataset

From the figure.3, it is inferred that using the CIFAR-10 dataset, the image recognition of all classes is performed. It is better to use a separate evaluation dataset. This is done by splitting into training and testing datasets. The learning behavior of the system is estimated in the training phase, and the system's performance is validated in the testing phase.

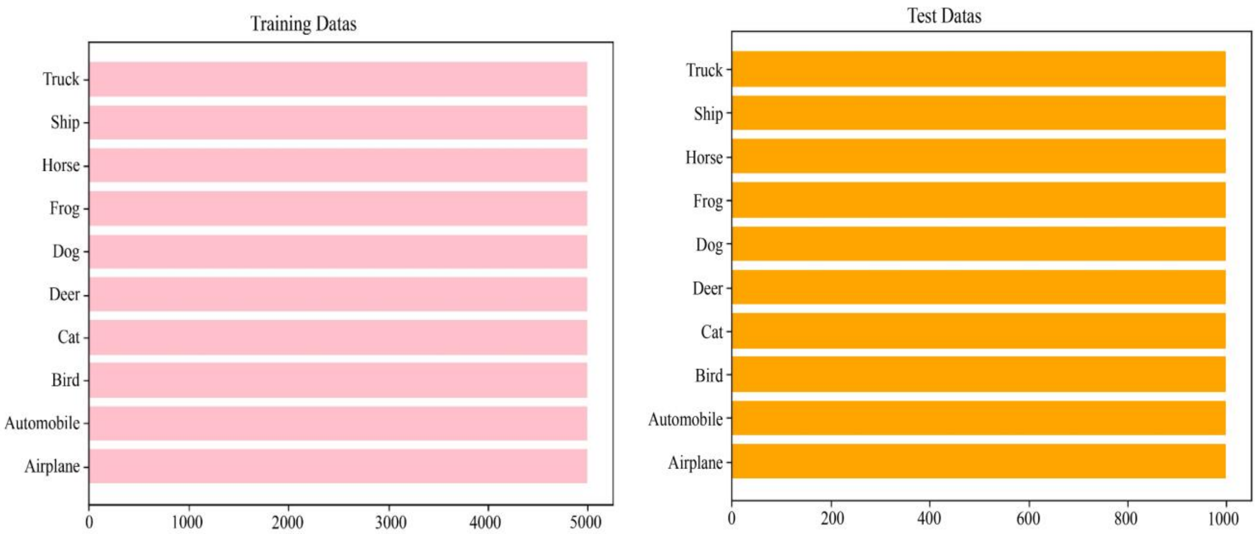

Fig. 4 Training dataset and testing dataset

Figure. 4 represents the training and testing datasets implemented in image recognition. Nearly 50,000 images in the training and 10,000 images in the testing phases are considered.

4.2. Performance Metrics

To estimate the performance of the proposed system, three major parameters are used precision, accuracy, F1 score, and recall.

4.2.1. Accuracy

The accuracy of the system is calculated by using equation (3), dividing the total number of correctly classified models by the total number of classified models,

$$\text{Accuracy} = \frac{\text{Tp+Tn}}{\text{Tp+Fp+Fn+Tn}} \qquad (3)$$

Where Tp, Tn, Fp, and Fn describe true positive, true negative, false positive, and false negative.

4.2.2. Precision

The precision of the system is represented as the capability of the classifier to predict the positive class correctly and is given by the equation (4),

$$\text{Precision} = \frac{\text{Tp}}{\text{Tp+Fp}} \qquad (4)$$

4.2.3. Recall

It is given by the equation (5) of dividing the number of times the classifier identified the negative class by all the times the negative class is obtained,

$$\text{Recall} = \frac{\text{Tp}}{\text{Tp+Fn}} \qquad (5)$$

4.2.4. F1 Score

It is a metric that depends on precision and recall and is given by the equation (6),

$$\text{F1 Score} = 2 + \frac{\text{Precision}*\text{Recall}}{\text{Precision+Recall}} \qquad (6)$$

4.2.5. ROC Curve

To visualize the efficacy of the respective system for binary classification, ROC is used. It is the representation of producing continuous variables of specificity and sensitivity. The ROC curve is an evaluation metric in which the curve is plotted between TP (True Positive) and FP (False Positive) at various threshold levels. The performance is thus demonstrated by using a confusion matrix.

4.2.6. AUC Curve

The AUC is a curve used to calculate the two-dimensional area under the entire ROC curve. It is used for binary classification problems to evaluate the performance. The AUC curve measures the ability to distinguish between classes and is used as a summary of the ROC curve.

4.3. Experimental Results

The results produced by the system denote the accuracy and behavior of the method implemented. Several sample images of different classes are used, and their outcomes are predicted.

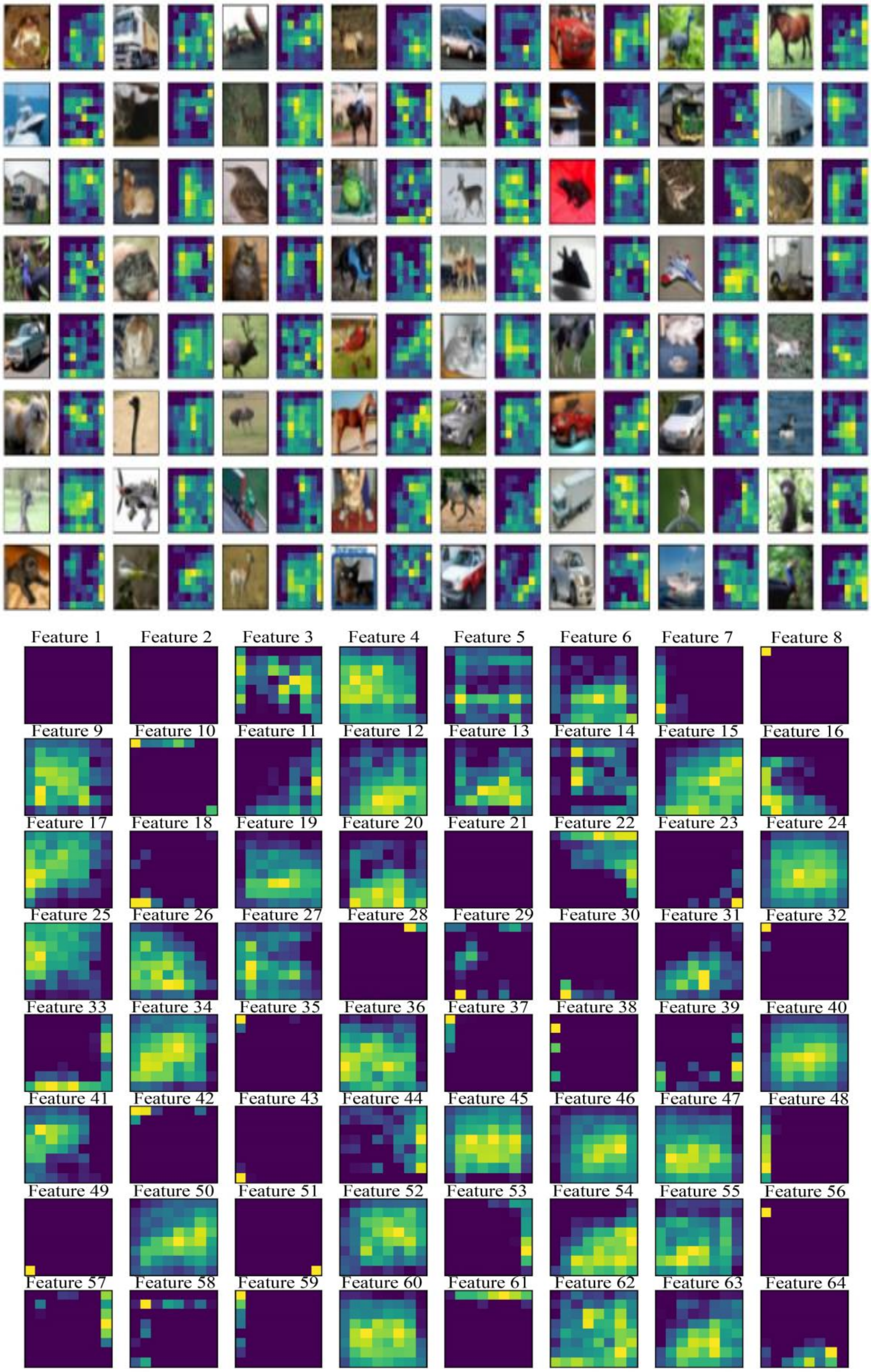

Fig. 5 Feature maps of different images extracted during training phase

From figure.5, the images used in the training phase corresponding to their extracted images are represented. Whereas the prediction phase uses the trained images and testing images to recognize the system's performance.

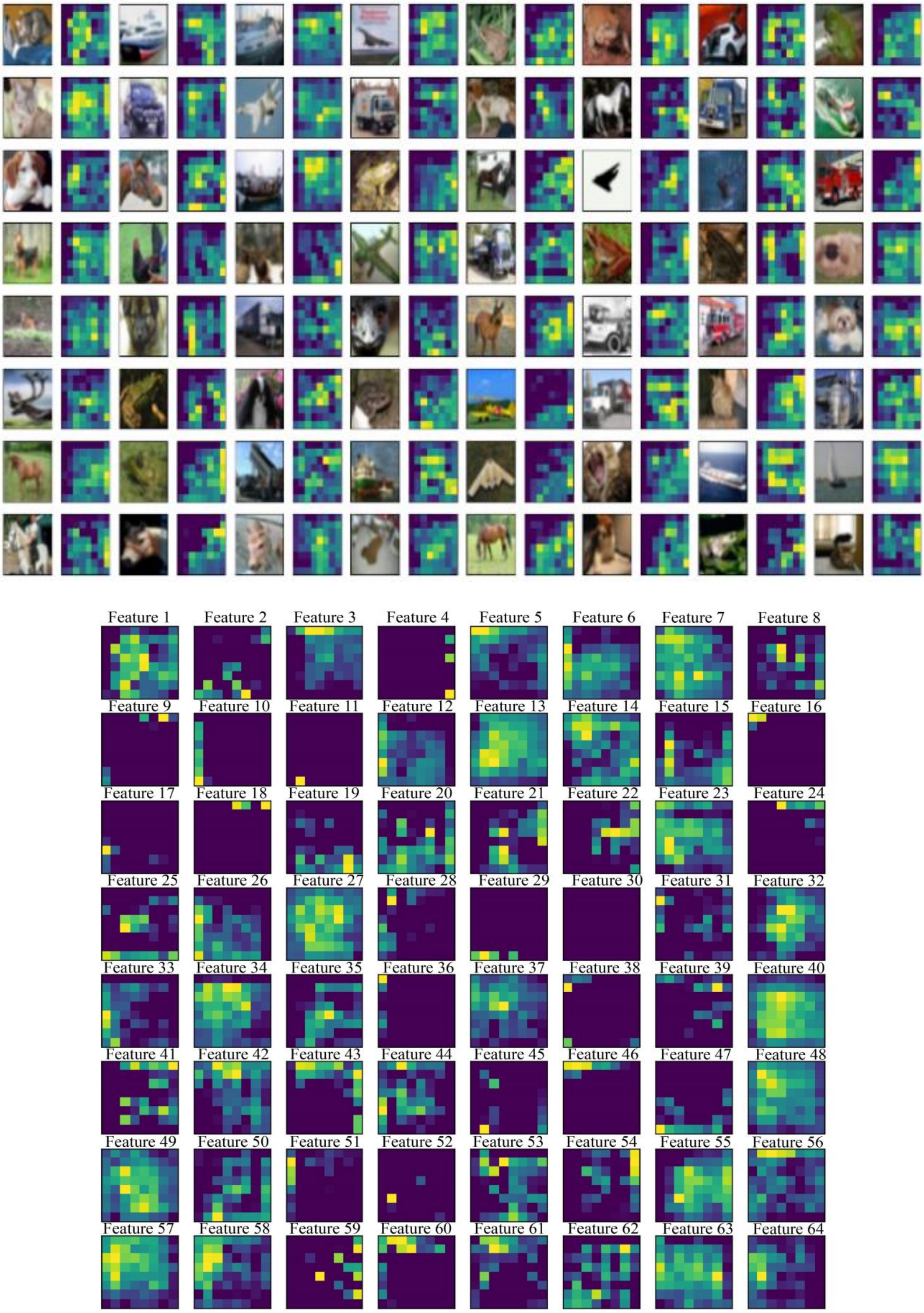

Fig. 6 Feature maps of different images extracted during testing phase

Figure. 6 denotes the images used in the testing phase and their corresponding extracted images. Accuracy is determined for predicting the efficiency of the proposed system. It is used to measure the statistical errors for the long-term validity of the system. The accuracy, precision, recall, and F1 score of the proposed system are found and represented in table.1.

Table 1. Accuracy prediction of proposed method

Accuracy	Precision	Recall	F1-score	ROC-AUC score
0.98	0.98	0.98	0.98	0.858185

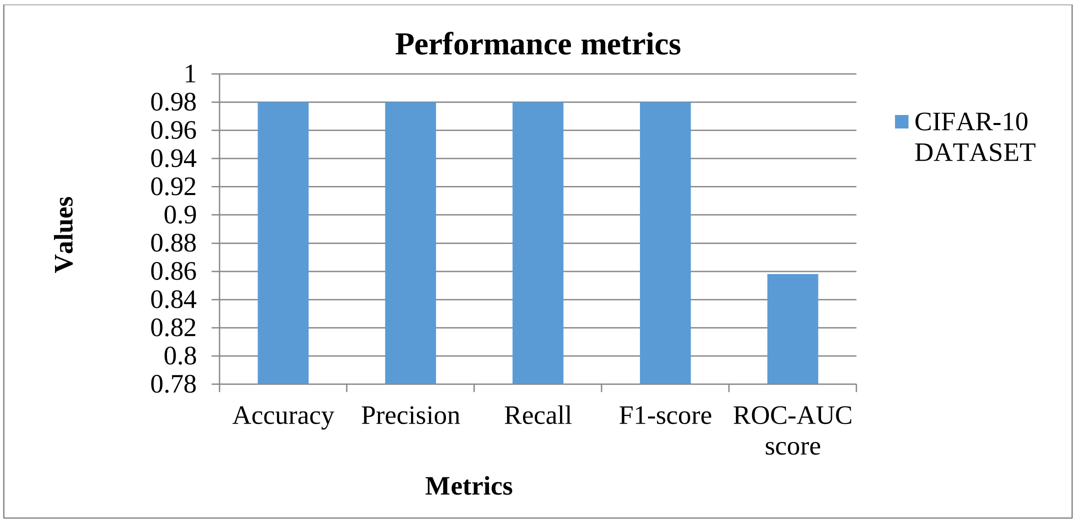

Fig. 7 Overall result analysis of the proposed system

It is found that the precision is 0.98, accuracy is 0.98, F1 score is 0.98, recall is 0.98, and ROC-AUC score is 0.858185. A graphical representation of performance metrics is shown in figure.7. Confusion matrix is performed to estimate the effectiveness of the classification methods for the given test dataset. The matrix is categorized into two parameters such as actual values and predicted values.

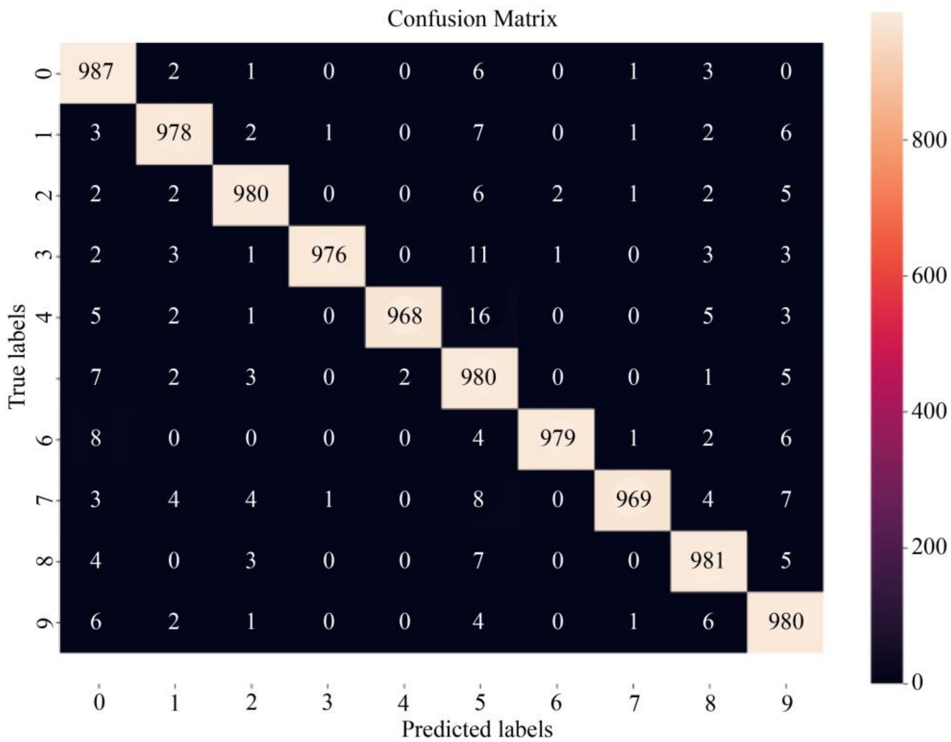

Fig. 8 Confusion matrix

The actual values are the true values for given interpretations, and the predicted values are those expected values of the system. From figure.8, it is inferred that the confusion matrix of the proposed system predicts maximum accuracy with minimum confusion, which occurs while distinguishing images between different classes. This implies that the proposed system performed with better accuracy.

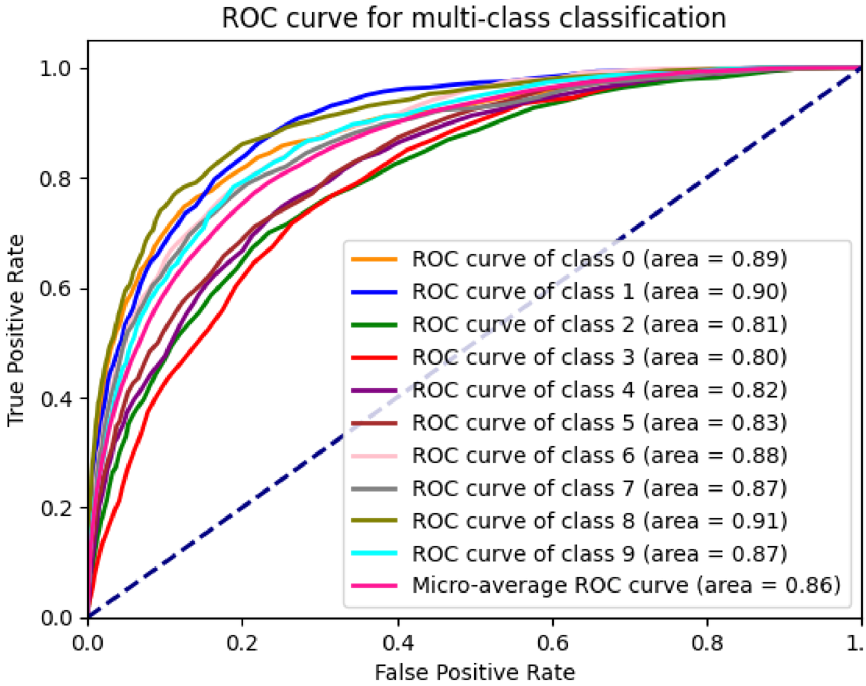

Fig. 9 ROC-AUC curve

If the predicted AUC is high, the system performance is also increased with maximum classification accuracy. In the proposed system, the values are found to be less than 1. Several numbers of true positives and true negatives than false positives and false negatives in the ROC-AUC curve produce a high chance of classification. Thus from the figure.9, it is inferred that the proposed system has produced a ROC-AUC score of 0.858185, indicating that the system produces better precision.

4.4. Comparative Analysis

To extract the underlying image features, pooling layers and learnable filters are employed to evade over-fitting by producing optimized precision during testing and evaluation. The training accuracy, validation accuracy, and testing accuracy of various existing CNN models are compared with the proposed system and represented in table.2.

Table. 2 Comparative analysis of different CNN models with proposed system [30]

	CNN Framework	Training Accuracy (%)	Validation Accuracy (%)	Testing Accuracy (%)
Existing methods	**NET III**	69.33	71.37	70.77
	NET III	74.59	75.31	74.73
	NET III	75.16	73.27	72.68
	NET III	77.04	78.39	78.29
	Existing Model	77.53	77.04	76.27
	Proposed model	98	98	98

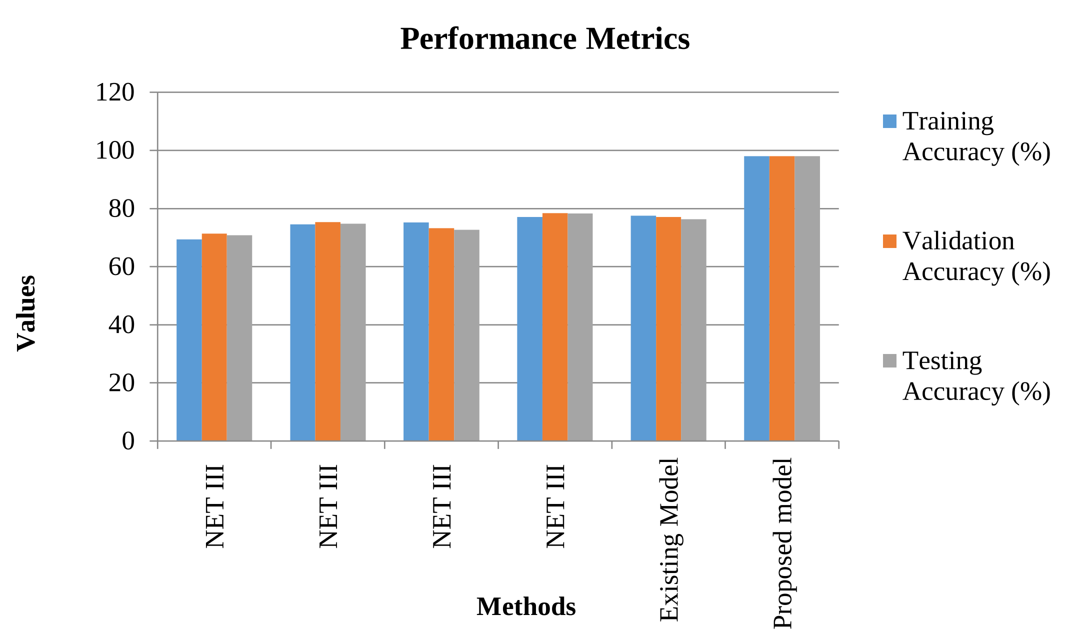

Fig. 10 Performance metrics of proposed system compared with existing methods [30]

Figure. 10 shows the graphical representation of performance metrics of existing methods with the proposed system. From the comparative analysis of different CNN models with the proposed system, it is found that the training accuracy, validation accuracy, and testing accuracy of the proposed system are 98%, 98%, and 98%. It is also observed that there is no over-fitting, and the accuracies predicted are equal.

In a network, skip connections provides benefits but create major issues when the dimensions between layers are not constant. In such a case, the zero padding technique matches the dimensions between layers. But this leads to the complexity of the network architecture resulting in increased computational costs and parameter numbers. So, the ORF algorithm is used in classification to overcome these limitations. The accuracy and loss predicted from different existing methods and the proposed system is shown in table.3.

Table. 3 Accuracy and loss of existing method and proposed method [31]

Model	Accuracy	Loss
ResNet50	0.8848	0.5358
InceptionV3	0.8006	0.9051
VGG16	0.7526	1.0064
Existing Model	0.92	0.4754
Proposed model	0.98	0.022

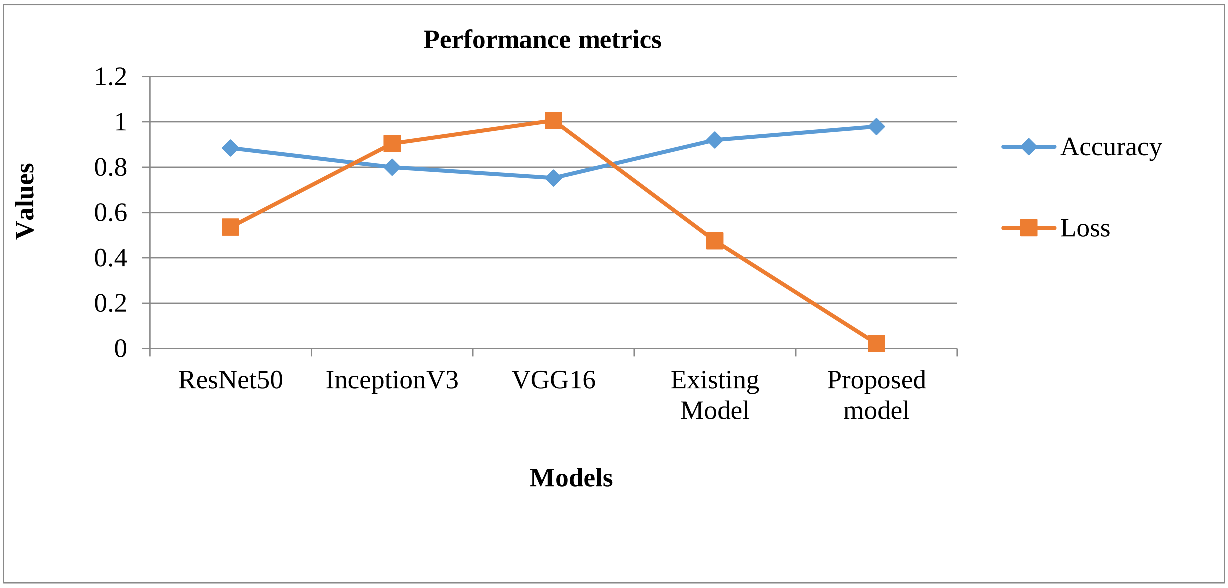

Fig. 11 Performance metrics of proposed system compared with existing methods [31]

Figure. 11 states that the respective method attained better precision and less computational loss compared to existing methods such as ResNet50, InceptionV3, and VGG16. But by applying ORF, the proposed system has produced an accuracy of 0.98 and a loss of 0.022, representing that the proposed system produced better results than the conventional methods. The CNN is implemented instead of other algorithms, as it portrays weight sharing feature that decreases the number of trainable network parameters during feature extraction. This leads to avoiding over-fitting and enhancing generalization. Large-scale network implementation is more easily applicable with CNN than other neural network algorithms. On the other view, RF, along with the implicit split log loss function, is employed in image classification as it performs well in noise immunity, is not prone to over-fitting issues, and is due to its fast training speed. Hence, CNN for feature extraction and ISLLF-ORF for image classification of feature extracted images is done in the proposed system and thus produces better results when compared with other conventional methods.

5. Conclusion

There are specifically different approaches used in image recognition, but it causes over-fitting problems. In this proposed system, CNN, along with random weights, were used in feature extraction with the CIFAR-10 dataset. The extracted features from CNN were fed to the ORF classifier for image classification. It helped classify the images, as they produced better capability of complex visualization.

The outcome was predicted by using ISLLF to evaluate the behavior of the proposed system. Thus by using these advantageous methods, the proposed system produced better results with 98% of precision. The performance was calculated based on parameters such as precision of 0.98, accuracy of 0.98, F1-score of 0.98, recall of 0.98, and ROC-AUC score of 0.858185.

Moreover, the results of the comparative analysis stated that the proposed model produced better outcomes than existing methods. The CIFAR-10 dataset used in the system can cause blurred and noisy images. When these images are implemented with CNN, it can compress the quality of the images, and thus misclassification might occur. Hence, as future work, high-quality and uncompressed images can be used for feature extraction and classification of various dataset images, using different ML and DL for enhanced image recognition.

References

[1] Lushan Cheng, Xu Zhang, and Jie Shen, "Road Surface Condition Classification Using Deep Learning," *Journal of Visual Communication and Image Representation*, vol. 64, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[2] Hao Fei et al., "Cotton Classification Method at the County Scale Based on Multi-Features and Random Forest Feature Selection Algorithm and Classifier," *Remote Sensing*, vol. 14, no. 4, p. 829, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[3] Mohammadreza Sheykhmousa et al., "Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 6308-6325, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[4] Xiaoxue Wang et al., "Land-Cover Classification of Coastal Wetlands Using the RF Algorithm for Worldview-2 and Landsat 8 Images," *Remote Sensing*, vol. 11, p. 1927, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[5] Safa Riyadh Waheed et al., "CNN Deep Learning-Based Image to Vector Depiction," *Multimedia Tools and Applications*, vol. 82, pp. 20283-20302, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[6] Ahmad Naeem et al., "Deep Learned Vectors' Formation Using Auto-Correlation, Scaling, and Derivations with CNN for Complex and Huge Image Retrieval," *Complex & Intelligent Systems*, vol. 9, pp. 1729-1751, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[7] Khadija Kanwal et al., "Deep Learning Using Symmetry, Fast Scores, Shape-Based Filtering and Spatial Mapping Integrated with Cnn for Large Scale Image Retrieval," *Symmetry*, vol. 12, no. 4, p. 612, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[8] El houssaine Hssayni, Nour-Eddine Joudar, and Mohamed Ettaouil, "KRR-CNN: Kernels Redundancy Reduction in Convolutional Neural Networks," *Neural Computing and Applications*, vol. 34, pp. 2443-2454, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[9] Pankaj Mishra, Claudio Piciarelli, and Gian Luca Foresti, "A Neural Network for Image Anomaly Detection with Deep Pyramidal Representations and Dynamic Routing," *International Journal of Neural Systems*, vol. 30, no. 10, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[10] R. Sakthi Prabha, and M. Vadivel, "Brain Tumor Stages Prediction using FMS-DLNN Classifier and Automatic RPO-RG Segmentation," *SSRG International Journal of Electrical and Electronics Engineering*, vol. 10, no. 2, pp. 110-121, 2023. [CrossRef] [Publisher Link]

[11] Xin Zhao et al., "3D-CNN-Based Feature Extraction of Ground-Based Cloud Images for Direct Normal Irradiance Prediction," *Solar Energy*, vol. 181, pp. 510-518, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[12] Zixian Ge et al., "Hyperspectral Image Classification Method Based on 2D–3D CNN and Multibranch Feature Fusion," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 5776-5788, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[13] Peng Wang, Xiaomin Zhang, and Yan Hao, "A Method Combining CNN and ELM for Feature Extraction and Classification of SAR Image," *Journal of Sensors*, vol. 2019, pp. 1-8, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[14] Kai Huang et al., "A Lightweight Privacy-Preserving CNN Feature Extraction Framework For Mobile Sensing," *IEEE Transactions on Dependable and Secure Computing*, vol. 18, no. 3, pp. 1441-1455, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[15] Yanchum Ma et al., "CNN-Feature Based Automatic Image Annotation Method," *Multimedia Tools and Applications*, vol. 78, pp. 3767-3780, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[16] Yuanyuan Qing et al., "Deep and Wide Feature Based Extreme Learning Machine for Image Classification," *Neurocomputing*, vol. 412, pp. 426-436, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[17] Nizar Ahmed et al., "Identification of Leukemia Subtypes from Microscopic Images Using Convolutional Neural Network," *Diagnostics*, vol. 9, p. 104, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[18] Mohd Anul Haq, "Planetscope Nanosatellites Image Classification Using Machine Learning," *Computer Systems Science and Engineering*, vol. 42, no. 3, pp. 1031-1046, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[19] Yongliang Qiao et al., "Individual Cattle Identification Using A Deep Learning Based Framework," *IFAC-PapersOnLine*, vol. 52, no. 30, pp. 318-323, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[20] G. Geetharamani and J. Arun Pandian, "Identification of Plant Leaf Diseases Using a Nine-Layer Deep Convolutional Neural Network," *Computers & Electrical Engineering*, vol. 76, pp. 323-338, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[21] Sonia Mittal, and Sanskruti Patel, "Age Invariant Face Recognition Techniques: A survey on the Recent Developments, Challenges and Potential Future Directions," *International Journal of Engineering Trends and Technology*, vol. 71, no. 5, pp. 435-460, 2023. [CrossRef] [Publisher Link]

[22] Yu-Dong Zhang et al., "Image-Based Fruit Category Classification by 13-Layer Deep Convolutional Neural Network and Data Augmentation," *Multimedia Tools and Applications*, vol. 78, pp. 3613-3632, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[23] Salma Abdel Magid, Francesco Petrini, and Behnam Dezfouli, "Image classification on IoT Edge Devices: Profiling and Modeling," *Cluster Computing*, vol. 23, pp. 1025-1043, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[24] Li Linhui, Jing Weipeng, and Wang Huihui, "Extracting the Forest Type from Remote Sensing Images by Random Forest," *IEEE Sensors Journal*, vol. 21, no. 16, pp. 17447-17454, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[25] Jose Manuel Monsalve-Tellez, Jorge Luis Torres-León, and Yeison Alberto Garcés-Gómez, "Evaluation of SAR and Optical Image Fusion Methods in Oil Palm Crop Cover Classification Using the Random Forest Algorithm," *Agriculture*, vol. 12, no. 7, p. 955, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[26] Xianyu Yu, Kaixiang Zhang, and Yanghui Zhang, "Land use Classification of Open-Pit Mine Based on Multi-Scale Segmentation and Random Forest Model," *Plos one*, vol. 17, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[27] Xue Ji, Bisheng Yang, and Qiuhua Tang, "Seabed Sediment Classification Using Multibeam Backscatter Data Based on the Selecting Optimal Random Forest Model," *Applied Acoustics*, vol. 167, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[28] Hu Gensheng et al., "Identification of Tea Leaf Diseases by Using an Improved Deep Convolutional Neural Network," *Sustainable Computing: Informatics and Systems*, vol. 24, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[29] Yun Yang et al., "Two-Stage Selective Ensemble of CNN via Deep Tree Training for Medical Image Classification," *IEEE Transactions on Cybernetics*, vol. 52, pp. 9194-9207, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[30] Ajala Sunday Adeyinka, "Convolutional Neural Network Implementation for Image Classification using CIFAR-10 Dataset," 2021. [Google Scholar] [Publisher Link]

[31] Ali Abdullah Yahya et al., "A Novel Image Classification Method Based on Residual Network, Inception, and Proposed Activation Function," *Sensors*, vol. 23, no. 6, p. 2976, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[32] Hironobu Fujiyoshi, Tsubasa Hirakawa, and Takayoshi Yamashita, "Deep Learning-Based Image Recognition for Autonomous Driving," *IATSS Research*, vol. 43, no. 4, pp. 244-252, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[33] Arnab Kumar Mishra et al., "Identifying COVID19 from Chest CT Images: A Deep Convolutional Neural Networks Based Approach," *Journal of Healthcare Engineering*, vol. 2020, 2020. [CrossRef] [Google Scholar] [Publisher Link]