[1,*]A. Josephine Anitha

[2]Dr. D. Gladis

Artificial Gorilla Troops Optimization with Ensemble Recurrent Neural Network for Remote Sensing Image Classification

Journal of Electrical Systems

Abstract: - Remote sensing images (RSIs) generally consist of the classification and analysis of largescale spatial data attained from aerial platforms, satellites, or other remote sensors. It plays an important part in extracting relevant data from RS information, and facilitating applications like disaster assessment, land cover mapping, environmental monitoring, and urban planning. Recurrent neural networks (RNN), a kind of neural network structure, have been effectively used for multiple tasks, namely image classification, time series analysis, and natural language processing and are more suitable for processing sequential data. On the other hand, Ensemble learning includes independently training various models and combining their predictions for making final decisions. Ensemble method aims to accomplish maximum accuracy, better robustness, and improved generalization by leveraging the diversity and complementary strengths of various models. This article introduces an Artificial Gorilla Troops Optimization Algorithm with Ensemble Recurrent Neural Network for Remote Sensing Image Classification (AGTOA-ERNN) model for RSIs. To accomplish this, the AGTOA-ERNN model categorizes the RSIs into different classes. In the presented AGTOA-ERNN technique, dynamic histogram equalization (DHE) technique can be applied to enhance the contrast level of the image. For feature extraction purposes, the AGTOA-ERNN technique uses EfficientNetB7 model can be exploited. Besides, ensemble RNN model comprising long short term memory (LSTM) and gated recurrent unit (GRU) techniques can be utilized for the classification process. Finally, the hyperparameter tuning of the ensemble RNN models takes place using the AGTOA model, which helps in accomplishing enhanced classification performance. To illustrate the betterment of the AGTOA-ERNN technique, a widespread simulation analysis takes place on Eurosat dataset. The comprehensive comparison result shows the upgraded achievement of the AGTOA-ERNN technique over other present techniques under different metrices.

Keywords: Remote sensing images; Images classification; Computer vision; Ensemble learning; Recurrent neural network

I. Introduction

RS is a way of gathering information from the objects or any scent with aids to reflection and emission of electromagnetic radiation from a distance without physical body contact with the objects [1]. These are made by satellites or aircrafts. This creates the process faster when forming it feasible to collect images from risky inaccessible locations covered to the wider field [2]. There are many application areas of RSI comprising the prediction of earthquakes, observation of air quality, urban planning and management of lands, etc. The manual identification and detection of objects and images from RS satellite imagery are difficult and highly expensive [3]. There have been many systems made for detecting the objects and classification of images from RSI. In recent years, there have considerable efforts to integrate machine learning (ML) algorithms for the detection and classification of images in the object. Despite the developments in RS approaches and based on the object image analysis tools, the classification efficiency of complex images remains insufficient [4]. The fundamental reason for this is insufficient because of the strong variations in the spatial and spectral features of the images, which make more difficulties to the classification of heterogeneous lands covered by the classes [5].

The latest research has proved that convolutional neural network (CNN) used deep learning (DL) approaches [6]. Due to its outstanding performance and accuracy, it is generally used in RSI classification. Despite the training DL methods from scratches, the additional complexities maybe lead to over-fit [7]. Deep transfer learning (DTL) algorithm emerged as a novel learning model to overcome these challenges. With the usage of TL model, it can be feasible to drastically decrease the requirements for training time and training data in the targeted field [8]. The framework does not necessitate training from the ground up in the targeted field, and the training and trial sets doesn't necessarily be essential to be a self-reliant or uniform distribution [9]. The initial objective of TL is to increase the performances of designated learners in particular fields by transporting the acquired comprehension to different but relevant source fields [10]. TL's various types of application chances have been created; it is a notable and attractive subject in the domain of ML techniques.

[1] *Corresponding author: Research Scholar, PG and Research Department of Computer Science, Presidency College (Autonomous), University of Madras, Chennai 600 005

[2] Principal, Bharathi Women's College, Chennai 600 108

This article introduces an Artificial Gorilla Troops Optimization Algorithm with Ensemble Recurrent Neural Network for Remote Sensing Image Classification (AGTOA-ERNN) model for RSIs. In the presented AGTOA-ERNN model, dynamic histogram equalization (DHE) technique can be applied to enhance the contrast level of the image. For feature extraction purposes, the AGTOA-ERNN technique uses EfficientNetB7 model can be exploited. Besides, ensemble RNN model comprising long short term memory (LSTM) and gated recurrent unit (GRU) techniques can be utilized for the classification process. Finally, the hyperparameter tuning of the ensemble RNN models takes place employing the AGTOA, which helps in accomplishing enhanced classification performance. To illustrate the betterment of the AGTOA-ERNN technique, a widespread simulation analysis takes place on Eurosat datasets.

II. Literature Review

Boulila et al. [11] devised a newly dispersed DL-based algorithm for the classification of big RSI. Especially, a Distributed CNN was introduced to handle RSI classification (RS-DCNN). At first, the training dataset was prepared for RS-DCNN. Next, a pixel-based CNN architecture across the big data clustering is implemented by the BigDL model, which ensures data-parallel training on the framework of Apache Spark. The authors in [12], exploit the BigEarthNet Sentinel-2 multi-spectral datasets for benchmarking DL algorithms for the multiclass, multilabel LULC image classification problems. This benchmark consists of typical CNN model along with the non-convolutional techniques namely Vision Transformer and Multi-Layer Perceptron. The study was conducted on the EfficientNet and wide residual network (WRN) framework, and leverage inference rate, classifying performance and training duration. An architecture based on dense connection convolutional network and global context spatial attention (GCSA) for the feature extraction of multiscale global scenes, named GCSANet was introduced by Chen et al. [13]. Several sample spacing is used to optimize the smoothness in the data space neighbourhood, and the mixup function is utilized for enhancing the spatial mixed data of RSIs. The features of multiscale surface object were extracted, and the densely connected backbone networks can reinforce the internal dense connection.

Li et al. [14] examine a DL based classifying algorithm for RSI, especially for high spatial resolution RS (HSRRS) images with multi-scene class and more modifications. Especially, to develop the classification method in urban regions, first we considered the four different deep neural network (DNN) models: 1) capsule network (CapsNet); 2) CNN; 3) identical model with the dissimilar training round based on CapsNet (SMDTR-CapsNet); and 4) similar model with training round based on CNN (SMDTR-CNN). The authors in [15], performed cropland segmentation with two different classes of labels mainly found in RS dataset that is source of "weak supervision": (1) image-level labels and (2) labels encompassed single geotagged points. We have shown that (1) a UNet image classifiers transported to segmentation outperforms pixel-level approaches like support vector machine (SVM), logistic regression (LR), and random forest (RF) and (2) UNet trained on the single labelled pixel per image. Li et al. [16] introduced a collaborated boosting framework (CBF) for combining knowledge-based ontology reasoning (OR) and data-driven DL modules iteratively. This study adopted the DSSN model, take the incorporation of the initial images, and inferred channel as DSSN input. Furthermore, the OR model is made up of extra-and-intra-taxonomy reasoning. Especially, based on the domain knowledge, the intra-taxonomy reasoning directly correct the misclassification of the DL algorithm.

For the scene classification of RSIs under the few shot scenarios, the authors in [17], devised a deep SSL architecture named RS-FewShotSSL, when we only have a few (lesser than 20) labelled scenes per class. In such cases, classical DL solutions that finetune CNN model, pretrained on the ImageNet datasets, fails intensely. In this work, a DL algorithm is pretrained from the ground up during the pretext tasks employing a massive amount of unlabelled scenes. Next, the model is finetuned on the labelled scenes during the so-called downstream or the main tasks. Saha et al. [18] introduced multi-target DA where single classifier is learned for unlabelled target domain. To construct multi-target classifiers, it might be advantageous to efficiently aggregate features from the labelled sources and unlabelled target domain. By doing this, based on the graph neural network, we exploit coaching that can able to leverage unlabelled datasets. Then, a sequential adaptation strategy was used that firstly adapted to the target domain, which makes it easier for adapting to the closest target domains.

III. The Proposed Model

In this study, a new AGTOA-ERNN method was introduced for RSIs classification. The AGTOA-ERNN technique exploits the DL models to categorize the RSIs into different classes. To achieve this, the presented

AGTOA-ERNN method encompasses various sub-procedures like DHE based pre-processing, EfficientNet-B7 feature extractor, ensemble RNN based classification, and AGTOA based hyperparameter tuning. Fig. 1 depicts the entire flow of AGTOA-ERNN technique.

3.1. Image Pre-processing

In a preliminary stage, the DHE technique is used to preprocess the RSIs. DHE is an extended version of the normal histogram equalization method utilized in image processing [19]. It is intended to improve the image contrast by redistributing the intensity value in a way that increases the total contrast while retaining the local details.

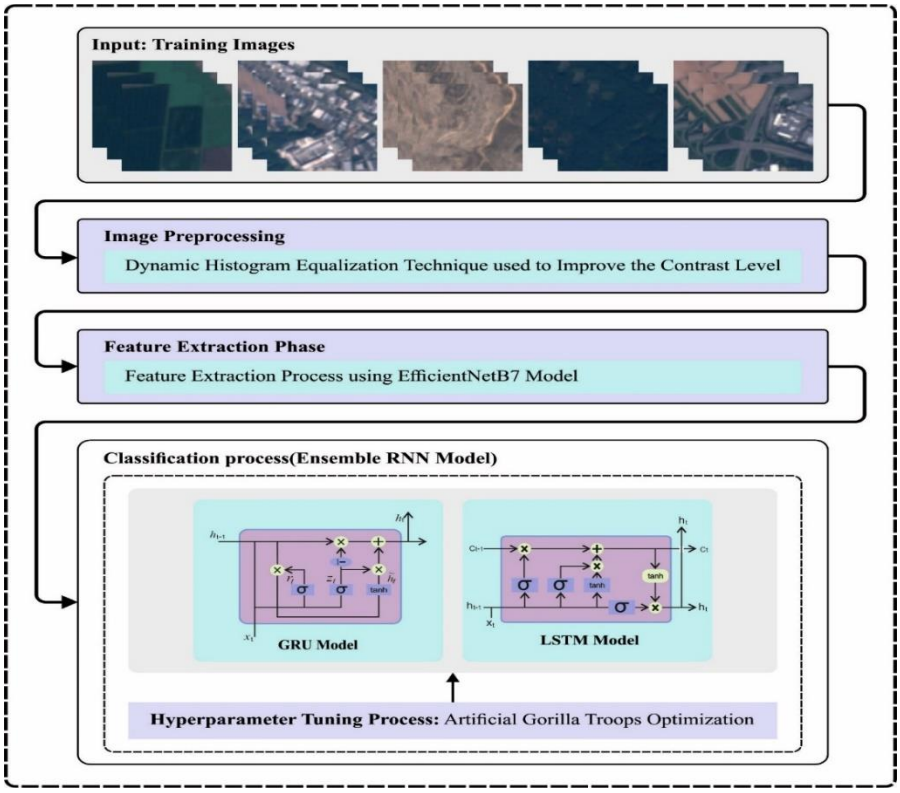

Fig. 1. Overall flow of AGTOA-ERNN approach

The normal histogram equalization works overall image, using the cumulative distribution function (CDF) of the image histogram to map the value of original intensity to newest value that is evenly spread out through the available dynamic ranges. This technique functions well once the image has a normal distribution of intensity value. But in certain cases, the normal histogram equalization causes undesirable effects, namely intensifying noise or excessively improving large areas of uniform intensity that might lead to unnatural-looking image. DHE's aim is to overcome this limitation by using histogram equalization locally on small region of the image, instead of globally overall image. The procedure of DHE was given below:

- Divide the image into overlapping or non-overlapping regions.
- Calculate the histogram of all the regions.
- Compute the cumulative distribution function (CDF) for all the regions.
- Map the original intensity value of all the regions to new value using the region-specific CDF.
- The processed region was combined to form the last enhanced image.

DHE could better adapt to local variation in intensity by using histogram equalization on small regions, which leads to more enhancement of the image. It helps to disclose details that could be hidden in brighter or darker regions of the image and enhance the total image visibility.

3.2. Feature Extraction

EfficientNet-B7 model is used for generating a relevant set of feature vectors in this study. DL model based on EfficientNetB7 that is called innovative architecture in convolution network was constructed [20]. EfficientNet-B7 demonstrates a special feature in applying a scaling approach in a compound co-efficient to equally measure

the parameters together with width, depth, and resolution. Firstly, we use convolutional network for learning feature maps and then the input images are classified. All the convolution layers are followed by the ReLU activation layer. Activation defines whether a neuron should be activated or not by measuring a weighted sum. It is used to present nonlinearity in the neuron output. Then, max-pooling is used for reducing dimensionality, downsampling the input images, and preparing it to process. The pretrained approach was utilized on the ImageNet datasets and the last layer of all the models. Then, dropout, batch normalization, and FC layers were exploited. The dropout layer avoids overfitting. Also, initialization of weight kernels with orthogonal weight is implemented. This avoids the layers' activation output from disappearing or exploding in the forward pass through CNN. Finally, an FC layer with 3 neurons was supplemented to characterize the class score (viz., output layer).

3.3. Ensemble RNN based Classification

In this work, ensemble RNN approach is employed for the RSIs classification. Schmidhuber et al in 1997 first proposed LSTM as a kind of RNN and add a gating mechanism based on the RNN comprising three gating mechanisms: input, output, and forget gates [21]. The input gate will be deciding whether to input the data to the memory cells, the output gate outputs the data, and the forget gate to forget the data in the memory cells. Concisely, LSTM may decrease the memory problem by forgetting the data.

The transition equation of LSTM is given below:

$$i_t = \sigma(\bar{i}_t) = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \quad (1)$$

$$f_t = \sigma(\tilde{f}_t) = \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + b_f\right) \quad (2)$$

$$g_t = \tanh(\tilde{g}_t) = \tanh\left(W_{xg}x_t + W_{hg}h_{t-1} + b_g\right) \quad (3)$$

$$o_t = \sigma(\tilde{o}_t) = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \quad (4)$$

$$c_t = c_{t-1} \otimes f_t + g_t \otimes i_t \quad (5)$$

$$m_t = \tanh(c_t) \quad (6)$$

$$h_t = o_f \otimes m_t \quad (7)$$

$$y_t = \mathcal{W}_{yh}h_t + b_y \quad (8)$$

Where h_t indicates the hidden state, i_t refers to the input gate, c_t represent the long memory cell and 0_t stands for the output gate, and f_t shows the forget gate.

GRU decreases the gate mechanism of LSTM and only has reset and update gates. The transition equation of GRU was shown below:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad (9)$$

$$r_t = \sigma\left(W_r \cdot \left[h_{t-1}, x_f\right]\right) \quad (10)$$

$$\tilde{h}_t = \tanh\left(W \cdot \left[r_t \times h_{t-1}, x_f\right]\right) \quad (11)$$

$$h_t = (1 - z_t) \times h_{t-1} + z_t \times \tilde{h}_t \quad (12)$$

Where h_t and h_{t-1}indicates the output and the output of last time. z_t and r_t denotes the update and reset gates, correspondingly. Compared to LSTM, the gating mechanism of GRU is easier and simpler to be trained with similar performance. Therefore, the advantage has led it to replace LSTM. The layers of LSTM and GRU exploit $tanh$ activation function for the resultant layer. The last output is generated by the linear activation function of the densely connected layer.

3.4. Hyperparameter Tuning using AGTOA

The AGTOA is utilized in this work for hyperparameter tuning process. The AGTOA is based on the common behaviors of the gorillas' living intelligence and life span [22]. The two fundamental stages of the MHs (intensification and diversification) are modelled by the creature's behaviors and their social life in nature. Consequently, the social life comprises migration, group formation, protecting new-born gorillas from predators, attracting female gorillas', and feeding them. But the mentor of each group modelled is a single gorilla termed Silverback gorillas, owing to its silver hair all over the body, responsible for handling the respective groups. The communal behaviors including protecting group members, migration, the search for food, and fighting between other group members, are attained with another important gorilla called black back gorilla, next to the silverback

gorillas. In this work, troops imply the gorilla group. Based on gorillas' individual and social interactions, a mathematical structure is established and gives meaningful insight for exploitation and exploration. Besides, five different identifiers were used; among them, three (potential to find better places within the search space, relocation towards the new space, and behaviours of other gorillas' in the same group) are employed in the exploration stage. On the other hand, the remaining operators are applied during the exploitation stage that enhances the quality search technique of the AGTOA.

The three conditional parameters are performed according to the relocation of gorillas in the disquisition stage, approach towards other individuals, and better food location search (optimized global site selection). The position vector for the best gorillas' (silverback gorilla) or the targeted gorilla is modelled, as given in Eq. (1), using three probabilities.

$$GY(t+1) = \begin{cases} (upperbound - lowerbound) \times rnd_1 + lowerbound, & rnd < \mathrm{s} \\ (rnd_2 - K) + Y_r(t) + R \times H & \mathrm{rnd} \geq 0.5 \\ Y(i) - R \times \left(R \times \left(Y(r) - GY_r(r)\right) + rnd_1 \times \left(Y(r) - GY_r(r)\right)\right) & rnd < 0.5 \end{cases} \quad (13)$$

Where rnd_1, rnd_2, rnd_3, and rnd_4: random integers that differ from zero to one, s indicates the parameter that needs to allocate the value within [0,-1] at the beginning, which decides the gorillas' relocating probabilities to newest location, $y_r(r)$ and $GY_r(r)$: randomly chosen individual from the troop and for modification in all the stages, correspondingly. Fig. 2 shows the structure of AGTOA.

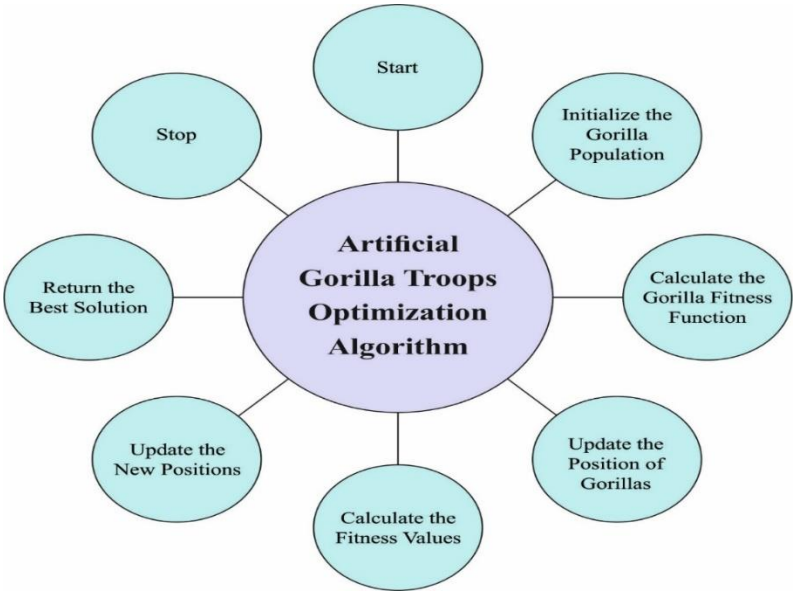

Fig. 2. Flowchart of AGTOA

Furthermore, K, and H were modelled by the following expression:

$$K = E \times \left(1 - \frac{currentiteration}{tota1countsofiterations}\right) \quad (14)$$

$$E = \cos\left(2 \times rnd_4\right) + 1 \quad (15)$$

$$R = K \times l \quad (16)$$

Where the current and maximal iterations are taken for the implementation of the algorithm, and $E$is derived from Eq. (15). rnd_4 denotes the random integer within [0,1]. As well, leadership factor l was introduced in Eq. (16) to realize its dominance and scale the leadership of the silverback gorillas. A random value with the range of $-K$ to K, H can be modelled as follows.

$$H = A \times Y(r) \quad (17)$$

$$A = [-K, K] \quad (18)$$

During the exploitation stage, two major factors, to act according to the rivalry and silverback for the female individuals, are taken for the implementation of the process. This aspect was eventually selected in accordance with the group integration and silverback. This can be based on K operator. When the K value is greater than or equivalent to $1N$ parameters, the group follows the silverback gorillas; otherwise (for less value), they have inter-rivalry for the female gorillas.

$$GY(t+1) = R \times MG \times (Y(t) - Y_{SG}) + Y(t) \tag{19}$$

In Eq. (19), MG was modelled as follows:

$$MG = \left(\left| \frac{1}{N} \sum_{i=1}^{N} G\, Y_i(r) \right|^{g} \right)^{\frac{1}{g}} \tag{20}$$

$$g = 2^{R} \tag{21}$$

In Eq. (20), Y_{SG} represent the location of silverback gorilla that suggest the best solution, N is the overall amount of creatures, and $Y_i(r)$ denotes the gorilla location at the specific iteration. The g value with respect to the R can be represented by Eq. (21). Rivalry for the young female gorilla is given as follows:

$$GY(i) = Y_{SG} - (Y_{SG} \times IF - Y(t) \times IF) \times CV \tag{22}$$

$$IF = 2 \times rnd_5 - 1 \tag{23}$$

$$CV = \beta \times E \tag{24}$$

$$E = \begin{cases} N_1, & rnd \geq 0.5 \\ N_2, & rnd < 0.5 \end{cases} \tag{25}$$

In Eq. (22) IF and CV correspondingly represent the impact force and coefficient vector that can be modelled as shown in Eqs. (23) and (24), rnd_5 portrays the random value within [0,1], and the parameters E and β represent the determination of solution that is shown in Eq. (25). Furthermore, the efficiency of AGTOA was enhanced by improving the chaotic maps, effective oppositional based learning algorithms, or hybridizing with other MHs.

The process of selecting the fitness function (FF) is a major aspect in the AGTOA approach. Solution encoding is implemented to measure the solution candidate goodness. At this point, the accuracy value is the crucial condition implemented for designing a FF.

$$Fitness = \max(P) \tag{26}$$

$$P = \frac{TP}{TP + FP} \tag{27}$$

Where TP and FP portrays the true and the false positive values.

IV. Results and Discussion

This article validates the performance of the AGTOA-ERNN approach on the EuroSat Dataset [23], comprising 10 classes (AnnualCrop, Herbaceous Vegatation, Forest, Industrial, Highway, Pasture, Residential, Permanent Crop, Sea Lake, and River). Fig. 3 demonstrates the ample images. Fig. 4 exhibits the sample of original, pre-processing, and feature maps images.

Fig. 3. Sample Images

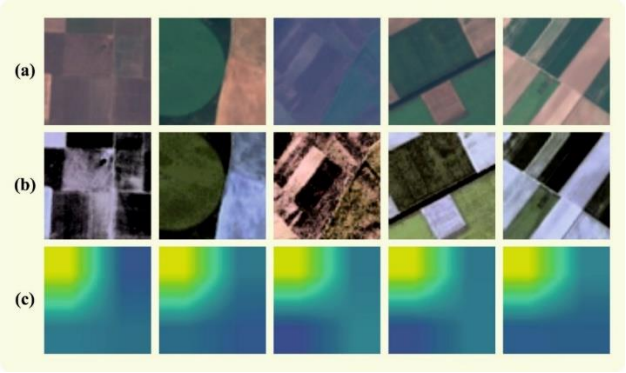

Fig. 4. Images (a) Original (b) Pre-Processed (c) Feature Maps

Fig. 5 illustrates the classifier outputs of the AGTOA-ERNN approach on without optimization. Fig. 5a portrays the confusion matrix provided by the AGTOA-ERNN approach. The figure also portrayed that the AGTOA-ERNN technique has precisely recognized and classified the complete 10 classes. Likewise, Fig. 5b portrays the PR investigation of the AGTOA-ERNN model. The figure portrayed that the AGTOA-ERNN technique attained highest PR achivement under 10 classes. Lastly, Fig. 5c portrays the ROC study of the AGTOA-ERNN algorithm. The figure showed that the AGTOA-ERNN algorithm has given an output in greater values with high ROC under 10 classes.

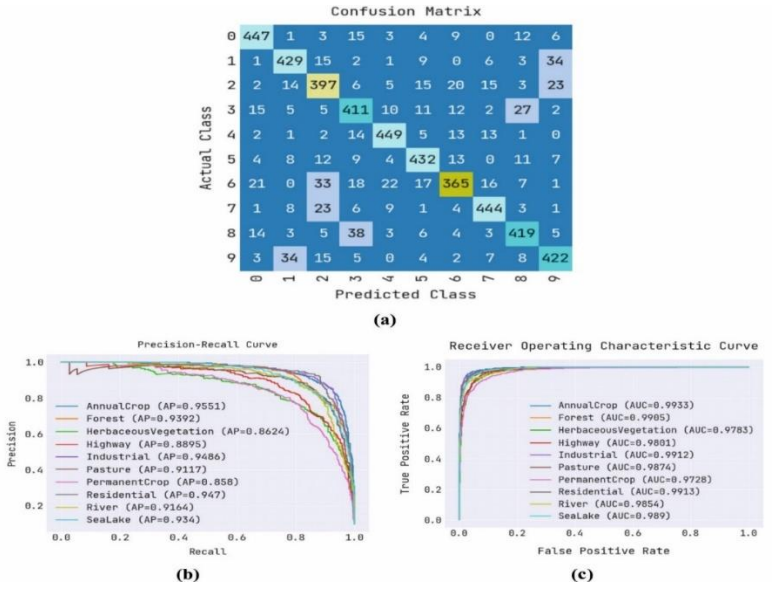

Fig. 5. Performance of Without Optimization (a) Confusion Matrix (b-c) PR and ROC Curve

Fig. 6 showing the TR_accu_y and VL_accu_y of the AGTOA-ERNN approach without optimization. The TL_accu_y is defined by assessing the AGTOA-ERNN approach on TR dataset, while the VL_accu_y is assessed by computing discrete test data achievement. The output also exhibited that TR_accu_y and VL_accu_y grows with an epoch upsurge. Consequently, the achievement of the AGTOA-ERNN model gets improve on the TR/TS dataset as epochs surges.

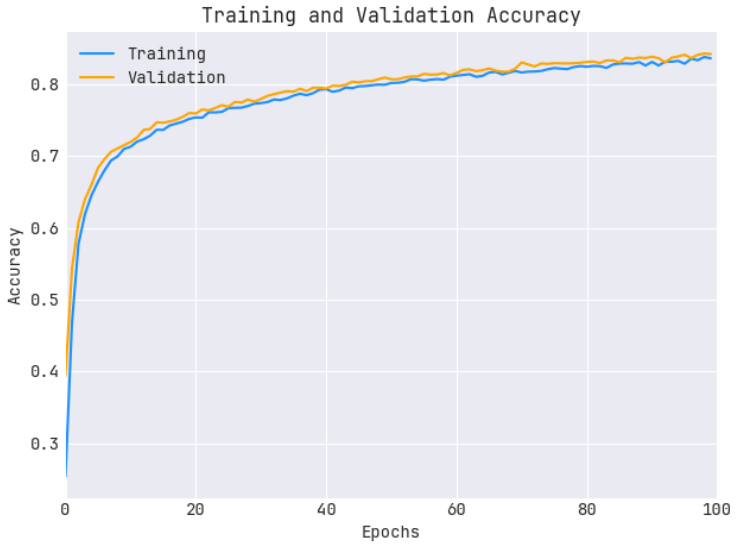

Fig. 6. $Accu_y$ outcome of AGTOA-ERNN system without optimization

In Fig. 7, the TR_loss and VR_loss outputs of the AGTOA-ERNN technique without optimization is shown. The TR_loss describes the error amid the anticipated and original value achievement on the TR data. The VR_loss represent the achievement measure of the AGTOA-ERNN system on a discrete data of validation. The outputs portrayed that the TR_loss and VR_loss tends to decrease with rising epochs. It also illustrated that the greater achievement of the AGTOA-ERNN system and its ability to generate precise classification. The lessened TR_loss and VR_loss values demonstrate the maximum achievement of the AGTOA-ERNN technique in comprehending patterns and relationships.

Fig. 7. Loss outcome of AGTOA-ERNN system without optimization

Fig. 8 indicated the classifier outputs of the AGTOA-ERNN system with optimization. Fig. 8a indicated the confusion matrix presented by the AGTOA-ERNN system. The figure indicated that the AGTOA-ERNN system precisely recognized and classified the complete 10 classes. Likewise, Fig. 8b illustrates the PR investigation of the AGTOA-ERNN method. The figure described that the AGTOA-ERNN method has attained highest PR achievement under 10 classes. Lastly, Fig. 8c demonstrates the ROC analysis of the AGTOA-ERNN approach.

The figure revealed that the AGTOA-ERNN approach has given an output in greater outputs with highest values of ROC under 10 classes.

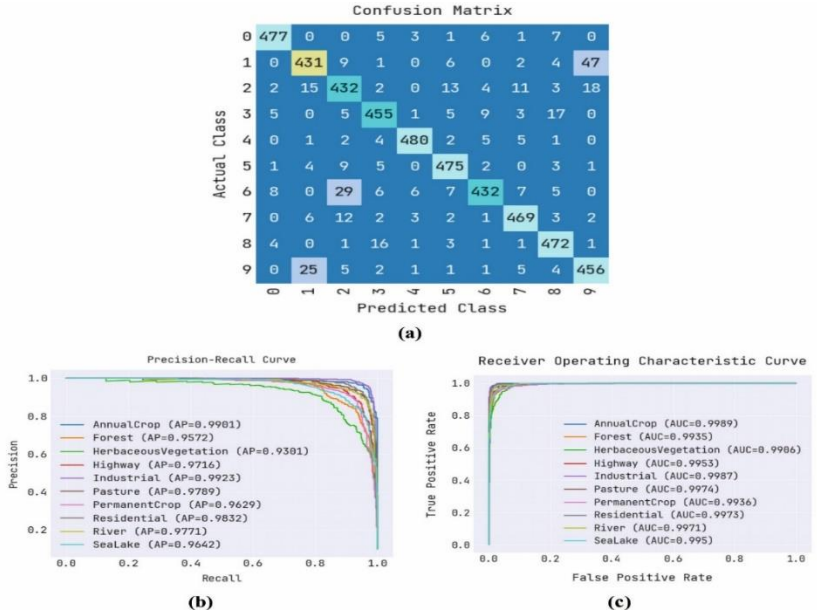

Fig. 8. Performance of With Optimization (a) Confusion Matrix (b-c) PR and ROC Curve

Fig. 9 showing the TR_accu_y and VL_accu_y of the AGTOA-ERNN technique with optimization. The TL_accu_y is defined by the assessment of the AGTOA-ERNN technique on TR dataset, while the VL_accu_y is evaluated by assessing the achievement on a discrete test data. The outputs also portray that TR_accu_y and VL_accu_y surges with an epoch growth. Finally, the achievement of the AGTOA-ERNN method gets improve on the TR/TS dataset with an epoch upsurge.

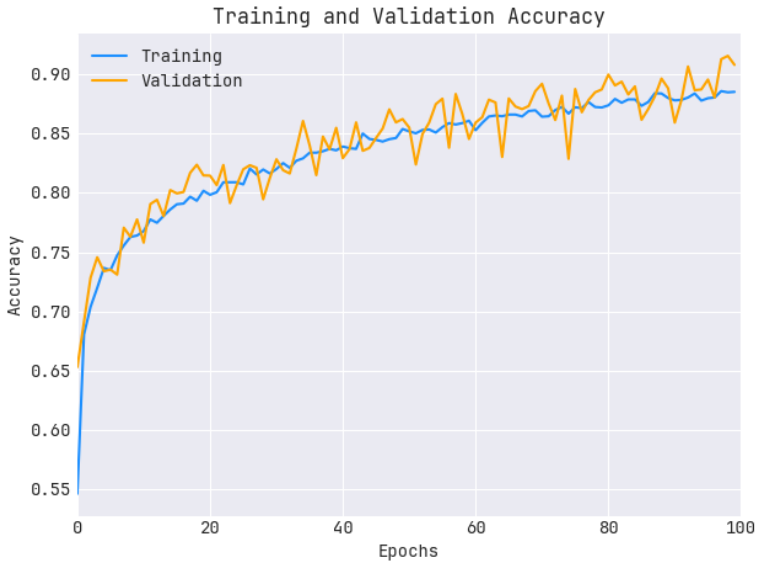

Fig. 9. $Accu_y$ outcome of AGTOA-ERNN system with optimization

In Fig. 10, the TR_loss and VR_loss outputs of the AGTOA-ERNN method with optimization is illustrated. The TR_loss illustrates the error amid the anticipated and original TR data values. The VR_loss illustrates the achievement measure of the AGTOA-ERNN method on discrete validation data. The outputs also portrayed that the TR_loss and VR_loss lessens with growing epochs. It also illustrated the improved achievement of the

AGTOA-ERNN method and its capacity in generating precise classification. The lessened TR_loss and VR_loss values portray the improved achievement of the AGTOA-ERNN method in comprehending patterns and relationships.

Fig. 10. Loss outcome of AGTOA-ERNN system with optimization

Table 1 and Fig. 11 reports the classifying outputs of the AGTOA-ERNN technique and without optimization of ensemble model. The outputs highlighted that the AGTOA-ERNN technique appropriately detected the classes with the inclusion of optimization. Without optimization, the AGTOA-ERNN technique attains $accu_y$ of 84.30%,$prec_n$ of 84.30%, $sens_y$ of 84.30%, $spec_y$ of 98.26%, $F1_{score}$ of 84.27%, and AUC_{score} of 98.59%. Also, with optimization, the AGTOA-ERNN technique attains $accu_y$ of 91.58%,$prec_n$ of 91.62%, $sens_y$ of 91.58%, $spec_y$ of 99.06%, $F1_{score}$ of 91.57%, and AUC_{score} of 99.57%.

Table 1 Classification outcome of AGTOA-ERNN approach without and with optimization

Metrics	Without Optimization	With Optimization
Accuracy	84.30	91.58
Precision	84.30	91.62
Sensitivity	84.30	91.58
Specificity	98.26	99.06
F1-Score	84.27	91.57
AUC Score	98.59	99.57

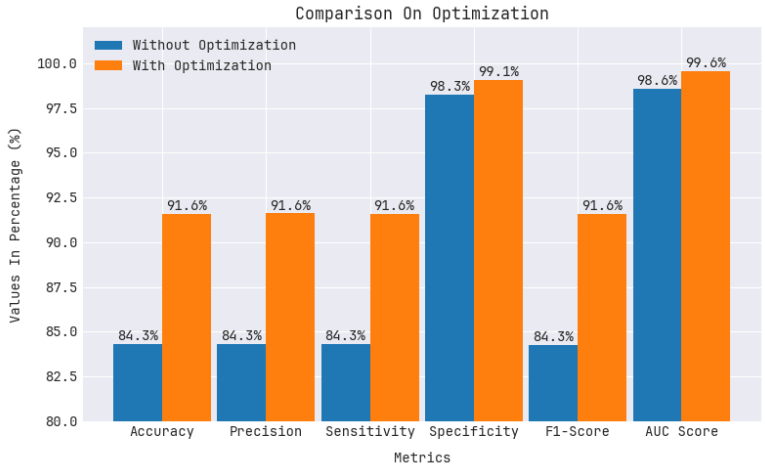

Fig. 11. Classification outcome of AGTOA-ERNN approach without and with optimization

At last, an extensive comparison study of the AGTOA-ERNN technique is clearly exhibited in Table 2 and Fig. 12 [24-26]. The resulted values inferred that the random initialization model attains least $accu_y$ of 76.76%. Along with that, the PSO-LSTM and U-Net methods have resulted in slightly boosted $accu_y$ of 88.20% and 86.40% respectively. Meanwhile, the IMO-LSTM and HGO-LSTM methods have attained slightly closer $accu_y$ values of 90.09% and 90% respectively. But the AGTOA-ERNN method confirmed its supremacy with maximum $accu_y$ of 91.58%. These outputs concluded the enhanced achievement of the AGTOA-ERNN method.

Table 2 $Accu_y$ outcome of AGTOA-ERNN method with other recent approaches

Methods	**Accuracy (%)**
AGTOA-ERNN	91.58
Random Initialization	76.76
PSO-LSTM	88.20
IMO-LSTM	90.09
HGO-LSTM	90.00
U-Net	86.40
RefineNet	92.80

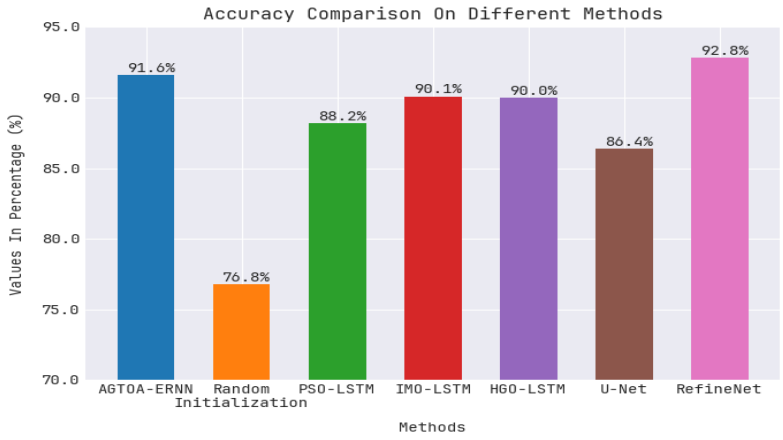

Fig. 12. $Accu_y$ outcome of AGTOA-ERNN approach with other recent methods

V. Conclusion

In this study, a new AGTOA-ERNN technique was presented for classifying the RSIs. The AGTOA-ERNN technique exploits the DL models to categorize the RSIs into different classes. To achieve this, the presented AGTOA-ERNN technique comprises several subprocesses such as DHE based pre-processing, ensemble RNN based classification, and AGTOA-based hyperparameter tuning. Also, the hyperparameter tuning of the ensemble RNN models occurs using the AGTOA, which helps in accomplishing enhanced classification performance. To illustrate the betterment of the AGTOA-ERNN technique, a widespread simulation analysis takes place on Eurosat dataset. The comparison outputs stated the improved accomplishment of the AGTOA-ERNN method over other present methods under diverse evaluation metrics. In the future, the achievement of AGTOA-ERNN method can be enhanced by the procedure of feature fusion.

Declarations

Competing Interests and Funding: The authors did not receive support from any organization

for the submitted work.

Conflict of Interest: The authors have expressed no conflict of interest.

References

[1] Li, Y., Chen, W., Zhang, Y., Tao, C., Xiao, R. and Tan, Y., 2020. Accurate cloud detection in high-resolution remote sensing imagery by weakly supervised deep learning. *Remote Sensing of Environment*, *250*, p.112045.

[2] Wu, X., Hong, D. and Chanussot, J., 2021. Convolutional neural networks for multimodal remote sensing data classification. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, pp.1-10.

[3] Martins, V.S., Kaleita, A.L., Gelder, B.K., da Silveira, H.L. and Abe, C.A., 2020. Exploring multiscale object-based convolutional neural network (multi-OCNN) for remote sensing image classification at high spatial resolution. *ISPRS Journal of Photogrammetry and Remote Sensing*, *168*, pp.56-73.

[4] Maxwell, A.E., Warner, T.A. and Guillén, L.A., 2021. Accuracy assessment in convolutional neural network-based deep learning remote sensing studies—Part 2: Recommendations and best practices. *Remote Sensing*, *13*(13), p.2591.

[5] Li, W., Chen, H., Liu, Q., Liu, H., Wang, Y. and Gui, G., 2022. Attention mechanism and depthwise separable convolution aided 3DCNN for hyperspectral remote sensing image classification. *Remote Sensing*, *14*(9), p.2215.

[6] Li, Y., Zhang, Y. and Zhu, Z., 2020. Error-tolerant deep learning for remote sensing image scene classification. *IEEE transactions on cybernetics*, *51*(4), pp.1756-1768.

[7] Dong, L., Du, H., Mao, F., Han, N., Li, X., Zhou, G., Zheng, J., Zhang, M., Xing, L. and Liu, T., 2019. Very high resolution remote sensing imagery classification using a fusion of random forest and deep learning technique—Subtropical area for example. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *13*, pp.113-128.

[8] Yang, X., Yang, X., Zhang, C. and Wang, J., 2023. SAR Image Classification Using Markov Random Fields with Deep Learning. *Remote Sensing*, *15*(3), p.617.

[9] Cao, X., Yao, J., Xu, Z. and Meng, D., 2020. Hyperspectral image classification with convolutional neural network and active learning. *IEEE Transactions on Geoscience and Remote Sensing*, *58*(7), pp.4604-4616.

[10] Ghadi, Y.Y., Rafique, A.A., Al Shloul, T., Alsuhibany, S.A., Jalal, A. and Park, J., 2022. Robust object categorization and Scene classification over remote sensing images via features fusion and fully convolutional network. *Remote Sensing*, *14*(7), p.1550.

[11] Boulila, W., Sellami, M., Driss, M., Al-Sarem, M., Safaei, M. and Ghaleb, F.A., 2021. RS-DCNN: A novel distributed convolutional-neural-networks based-approach for big remote-sensing image classification. *Computers and Electronics in Agriculture*, *182*, p.106014.

[12] Papoutsis, I., Bountos, N.I., Zavras, A., Michail, D. and Tryfonopoulos, C., 2023. Benchmarking and scaling of deep learning models for land cover image classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *195*, pp.250-268.

[13] Chen, W., Ouyang, S., Tong, W., Li, X., Zheng, X. and Wang, L., 2022. GCSANet: A global context spatial attention deep learning network for remote sensing scene classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *15*, pp.1150-1162.

[14] Li, W., Liu, H., Wang, Y., Li, Z., Jia, Y. and Gui, G., 2019. Deep learning-based classification methods for remote sensing images in urban built-up areas. *Ieee Access*, *7*, pp.36274-36284.

[15] Wang, S., Chen, W., Xie, S.M., Azzari, G. and Lobell, D.B., 2020. Weakly supervised deep learning for segmentation of remote sensing imagery. *Remote Sensing*, *12*(2), p.207.

[16] Li, Y., Ouyang, S. and Zhang, Y., 2022. Combining deep learning and ontology reasoning for remote sensing image semantic segmentation. *Knowledge-Based Systems*, *243*, p.108469.

[17] Alosaimi, N., Alhichri, H., Bazi, Y., Ben Youssef, B. and Alajlan, N., 2023. Self-supervised learning for remote sensing scene classification under the few shot scenario. *Scientific Reports*, *13*(1), p.433.

[18] Saha, S., Zhao, S. and Zhu, X.X., 2022. Multitarget domain adaptation for remote sensing classification using graph neural network. *IEEE Geoscience and Remote Sensing Letters*, *19*, pp.1-5.

[19] Rao, B.S., 2020. Dynamic histogram equalization for contrast enhancement for digital images. *Applied Soft Computing*, *89*, p.106114.

[20] Moussaid, A., Zrira, N., Benmiloud, I., Farahat, Z., Karmoun, Y., Benzidia, Y., Mouline, S., El Abdi, B., Bourkadi, J.E. and Ngote, N., 2023, February. On the Implementation of a Post-Pandemic Deep Learning Algorithm Based on a Hybrid CT-Scan/X-Ray Images Classification Applied to Pneumonia Categories. In *Healthcare* (Vol. 11, No. 5, p. 662). MDPI.

[21] Qi, C., Ren, J. and Su, J., 2023. GRU Neural Network Based on CEEMDAN–Wavelet for Stock Price Prediction. *Applied Sciences*, *13*(12), p.7104.

[22] Gürses, D., Mehta, P., Patel, V., Sait, S.M. and Yildiz, A.R., 2022. Artificial gorilla troops algorithm for the optimization of a fine plate heat exchanger. *Materials Testing*, *64*(9), pp.1325-1331.

[23] Patrick Helber, Benjamin Bischke, Andreas Dengel. 2018. Introducing EuroSAT: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification. IEEE International Geoscience and Remote Sensing Symposium, 2018Zhang, C., Chen, Y., Yang, X., Gao, S., Li, F., Kong, A., Zu, D. and Sun, L., 2020. Improved remote sensing image classification based on multi-scale feature fusion. *Remote Sensing*, *12*(2), p.213.

[24] Stateczny, A.; Bolugallu, S.M.; Divakarachari, P.B.; Ganesan, K.; Muthu, J.R. Multiplicative Long Short-Term Memory with Improved Mayfly Optimization for LULC Classification. Remote Sens. 2022, 14, 4837. https://doi.org/10.3390/rs14194837

[25] Naushad, R.; Kaur, T.; Ghaderpour, E. Deep Transfer Learning for Land Use and Land Cover Classification: A Comparative Study. Sensors 2021, 21, 8083. https://doi.org/10.3390/s21238083