Optimal Deep Learning-Based Image Classification for IoT-Enabled UAVs in Remote Sensing Applications

Sanjar Mirzaliev[1,*], Samandarboy Sulaymanov[1]

[1]Tashkent State University of Economics, Tashkent, Uzbekistan

Emails: s.mirzaliev@tsue.uz, s. Sulaymanov@tsue.uz

Abstract

Unmanned Aerial Vehicles (UAVs), together with Internet of Things (IoT) technology, have emerged as robust tools for remote sensing (RS) and data collection in different sectors, including environmental monitoring, agriculture, and disaster management. The incorporation of data from UAVs with IoT sensors on the ground can provide a holistic view of the environment, improving the quality of input for image classification. Deep learning (DL) models-based image classification is a key component of IoT-assisted UAVs, transforming them from data collection tools into intelligent decision-making platforms. Especially, Convolutional Neural Networks (CNNs) can automatically recognize objects, patterns, and anomalies in images captured by UAVs. Therefore, the study presents an automated image classification with the Tyrannosaurus optimization algorithm using deep learning (AIR-TROADL) method on the IoT-aided UAV network. The AIR-TROADL technique aims to examine the UAV images for the identification and classification of images into distinct categories. In the projected AIR-TROADL method, an enhanced ShuffleNet model is exploited for feature extraction. Besides, the hyperparameter tuning of enhanced ShuffleNet model can be performed by using TROA, which in turn boosts the classification performance. Finally, the classification of images takes place using the attention-based gated recurrent unit (AGRU) model. A series of simulations have been conducted to exhibit the promising outcome of the AIR-TROADL technique. The comparative outcomes highlighted that the AIR-TROADL method reaches high efficiency over its recent approaches in terms of distinct measures.

Keywords: Unmanned aerial vehicles; Image classification; Remote sensing; Deep learning; Computer vision

1. Introduction

Unmanned Aerial Vehicles (UAVs)/drones recently have gained great interest because it has a Remote Sensing (RS) platform for several practical applications like satellite imagery processing, precision agriculture, traffic monitoring, and search and rescue [1]. In recent developments in technology, the integration of camera sensors provides the chance for UAV applications to analyze, monitor, and detect passive and active attacks as well as hazards at event acts [2]. For example, flooding threats, landslide-prone areas road collisions, and fire spots in forested areas. Additionally, their small-size UAVs provide quick deployment and can be in the loop of task dangerous verdicts to manage the obtainable resources and then enhance mitigation, prevention, and risk assessment. Based on the time, timely and precise achievement of a landform and its change collection has become very significant [3]. This kind of data is highly advantageous to make sense of the effect of human behaviour on natural sources through dissimilar times by monitoring the earth at a 3D scale. The improvement of RS methodology in spectral, temporal resolution, and spatial RS data and the excellent advancements of data and communication methods in integration, transmission, management capabilities, and data storage are significantly altering the method we see the Earth [4].

DOI: https://doi.org/10.54216/IJAACI.060101
Received: July 24, 2023 Revised: September 18, 2023 Accepted: February 02, 2024

Machine learning (ML) is one of the primary methods for land cover classification by using single RS images. Moreover, single RS images produce only rapid spectral data of the earth's surface that provides restricted features obtainable for classification and then leads to poor classification outcomes, especially for dissimilar vegetation and yield types [5]. In addition, the recognition results from single images are influenced by weather, season, and other factors which make the technique not relevant to a complete study on land cover changes [6]. The traditional approaches employed manifold features such as background, color, textual, shape, edge, foreground subtraction, and removal to organize diverse objects. These kinds of techniques gain extraordinary results for different remote surveillance applications [7]. Normally, it requires a high range of information and that is not capable of identifying various objects whose features do not exist in the training datasets. With the high evolution of enhanced DP, many convolution neural network (CNN) based techniques are utilized by investigators to divide dissimilar regions of interest and objects into aerial data sets [8]. These refined models suggestively improved the general and simplification correctness of the segmentation process for different kinds of applications.

Recently, DL techniques have increasingly concerned more attention in the RS field [9]. The latest technology artificial intelligence (AI) has proved that a data-driven deep neural network (DNN) aids in recognizing basic successive needs from multitemporal RS observations. Recurrent neural networks (RNN), CNN, self-attention systems, and their alternatives are more popular DL frameworks for processing time-series satellite data [10]. The learning feature of DNN decreases the job of human feature engineering as well as enhances the performance and simplification of the technique.

This study presents an automated image classification with the Tyrannosaurus optimization algorithm using deep learning (AIR-TROADL) method on IoT-aided UAV networks. The AIR-TROADL technique aims to inspect the UAV images for the identification and classification of images into distinct categories. In the presented AIR-TROADL method, an enhanced ShuffleNet model is exploited for feature extraction. Besides, the hyperparameter tuning of the enhanced ShuffleNet model is performed by using TROA, which in turn boosts the classification performance. Finally, the classification of images takes place using the attention-based gated recurrent unit (AGRU) model. A series of simulations have been carried out to exhibit the superior performance of the AIR-TROADL approach.

2. Related Works

Selvam [11] introduced an Earthworm Optimizer with Deep-TL Enabled Aerial Image Classification (EWODTL-AIC) technique in IoT-based UAV network. The proposed method primarily utilizes the AlexNet framework as an extraction feature to generate optimum feature vectors. Secondarily, the hyperparameter values of the AlexNet framework have been determined by the deployment of the earthworm optimizer (EWO) method. Lastly, the XGBoost technique was utilized in classifying aerial images. Mitra et al. [12] developed a public BC-based access control operation to monitor wildlife. Subsequently, the authors assume this presented security model for the BC part for evaluating the overall computational time required to improve a wide-ranging number of blocks in a BC and then various counts of transactions per block. In [13], the authors presented an innovative method to determine an internal and 3D region by evaluating the signal power. The scientific representation depends on a path loss algorithm and DT using the ML method. The authors designed 2D and 3D frameworks for collecting highly accurate data in the nodes.

In [14], the authors offered comprehensive issues and concerns about the application of BC in a 5G-based IoT-assisted IoD platform. The authors developed and analyzed a novel BC-assisted secure model for data management between IoD communication objects. This introduced method can have the capability to block numerous possible attacks, which are needed in IoT-assisted IoD platforms. In [15], the authors considered the 3D assignment and resource allocation of numerous UAV-mounted BS from the IoT but, the stable activities for the confined channel resources, UAV-BSs, and signal intervention have been considered. In this studied method, the overall broadcast power of IoT devices was decreased, subjected to a SINR threshold for all devices. In [16], the authors developed a DL technique AgriSegNet for automatically identifying the agricultural land anomalies employing multiscale attention semantic segmentation of UAV attained images. This developed system was beneficial in monitoring agricultural land and crops to raise the effectiveness of precision agricultural methods.

In [17], the UAV-based Integrated Farm Management (UAV-IFM) was developed to enhance smart livestock farming. Reliable and protective monitoring of livestock in farming to fork has constructed potential by a sensor that can be extensive effects for identifying and comprising disease occurrences and avoiding the subsequent economic losses and food-relevant health illnesses. The study aims to increase the evaluation method therefore, smart livestock agriculture leads to highly extensive acceptance and provides growth-promoting support to

DOI: https://doi.org/10.54216/IJAACI.060101
Received: July 24, 2023 Revised: September 18, 2023 Accepted: February 02, 2024

agriculturalists. Peng et al. [18] examined a UAV-enabled mobile edge computing (MEC) network. The age of information (AoI) has been combined as a significant effectiveness metric to satisfy a new need for IoT application. Regarding the dimensional explosion, the DRL model could be employed, and a double DQL network (DDQN) method has been developed for understanding the intelligent and additional experience path development of the UAV.

3. The Proposed Model

In this study, we have established an automated image classification method, named AIR-TROADL method on the IoT-aided UAV networks. The AIR-TROADL method aims to examine the UAV images for the identification and classification of images into distinct categories. In the projected AIR-TROADL approach, three major stages are involved namely enhanced ShuffleNet model, TROA-based hyperparameter tuning, and AGRU-based classification. Fig. 1 defines the working flow of the AIR-TROADL model.

3.1. Structure of Enhanced ShuffleNet Model

An enhanced ShuffleNet model can be used in the projected AIR-TROADL method. ShuffleNet has a related concept with ResNet, MobileNet, and Xception [19]. The channel attention feature and depthwise convolution (DWConv) are used to increase the residual module of ResNet, enhance the operational efficacy of the model, and ensure network accuracy. Unlike classical residual structure, which integrates the feature of the deep and non-deep networks attained by various Conv layers, the inverted residual structure splits the input feature mapping into $X1$ and $X2$, $X2$ through DWConv and twice 1x1 Conv+batch normalization+activation function, $X1$ and $X2$. Assume that the overall amount of channels is $G \times n$, and the input is split into G groups. First, the channels are divided into 2D (G, n), later transposed 2D into (n, G), and lastly reshaped into 1D $G \times n$.

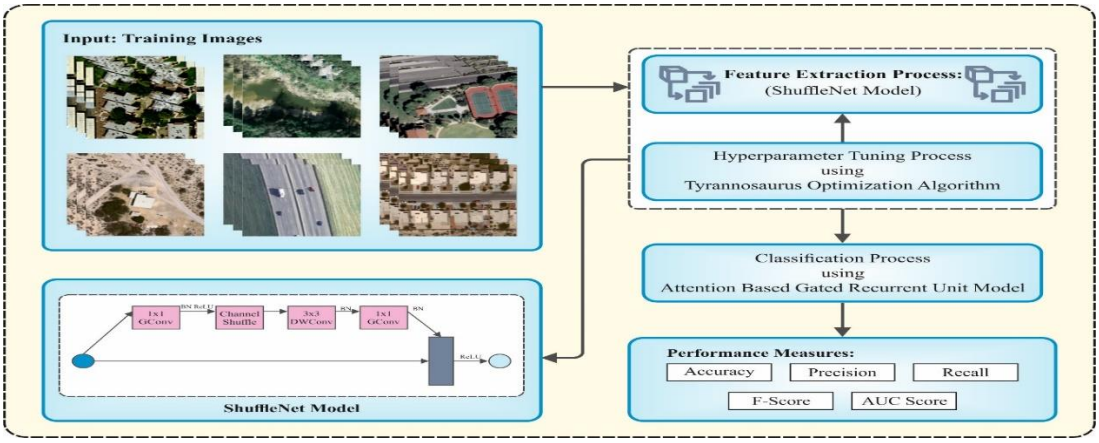

Fig. 1. Overall procedure of AIR-TROADL algorithm

The Squeeze-and-Excitation Network (SE) and Selective Kernel Network (SK) are added to make more accurate classification. At first, a feature map U with dimension of $H \times W$ and entire quantity of channels C is compressed into feature vectors $(1, 1, C)$ through the global pooled F_{sq}:

$$Z_c = F_{sq}(U_c) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{w} U_c\,(I, j). \tag{1}$$

For further adding non-linear conditions, the linear mapping and activation functions are added to the feature vector that best fit the composite relation between channel attentions. Lastly, the channel attention features calculated are multiplied by the feature mapping to attain the channel output. Feature extraction of channel proportion can be performed on the DWConv output once the SE channel attention has been added to the DWConv.

SK is similar to SE. The only difference is that SENet carries out channel attention, whereas SKNet carries out on Conv kernels. SKNet exploits Conv check feature mappings of varying dimensions in the network for

DOI: https://doi.org/10.54216/IJAACI.060101
Received: July 24, 2023 Revised: September 18, 2023 Accepted: February 02, 2024

extracting the features of dissimilar scales and after feature fusion of different scales, the channel attention is extracted. Also, SK is used in DWConv to compare with SE.

Firstly, the input feature mapping is calculated by the DWConv ($conv\_1$) with size of 3x3 convolution kernel, and DWConv with size of 3x3 Conv kernel and dilation aspect of 2 in dissimilar scales; later the pooling layer has been calculated as SE channel feature, and the 2 resultant feature mappings are added for global pooling; consequently, the outcome of 2-channel attention features can be multiplied with $conv\_1$ and $conv\_2$ from the channel to attain 2 feature mappings of varied channel features at dissimilar scales, later the 2 feature mappings can added to attain output features of the SK attention.

3.2. Design of AGRU based Classification

For image classification, the AGRU model can be utilized. The problems of gradient vanishing or disappearance are easy to take place in the computation of backpropagation once the layers of the RNN are comparatively deep [20]. Thus, RNN failed to capture the long-term dependency of sequence data. Hochreiter and Schmidhuber proposed the LSTM based on RNN that captures the long-term dependence on sequences. GRU network is based on LSTM, introduced by Cho et al. Consequently, GRU is used for learning long-term dependency features. GRU contains updated gate z_t and reset gate r_t. Through a set of equations, the GRU computes the hidden state expressed as follows:

$$z_t = \sigma\left(W^{(z)} \cdot [h_{t-1}, x_t]\right)$$

$$r_t = \sigma\left(W^{(r)} \cdot [h_{t-1}, x_t]\right)$$

$$\tilde{h}_t = \tanh\left(W \cdot [r_t * h_{t-1}, x_t]\right)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t. \quad (2)$$

Where update gate z_t decides what amount of data from the prior time is replaced into the existing state, and reset gate r_t used decides what amount of data from the prior instant is disregarded The $\tilde{h}_t$ candidate activation is calculated by the reset gate r_t (defines which information must be maintained). The activation function is σ. The entry-wise multiply operation is represented as $*$. Lastly, the actual activation of GRU, which is linear approximation between the $\tilde{h}_t$ and h_{t-1} candidate and prior activations . The output of the attention model is expressed as follows:

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \quad (3)$$

Where $= XW_U, U \in (Q, K, V)$, and X denotes the input, K is the transpose of matrix K, W_U indicates the learnable matrix and d_k is the dimension of keys. $softmax$ is an activation function described as $(x_i) = \left(e^{x_i} / \sum_j e^{x_j}\right)$, where x_i denotes the i^{th}dimension of N-dimension vector $x(i, j = 1,2, \ldots, N)$. Fig. 2 depicts the architecture of AGRU.

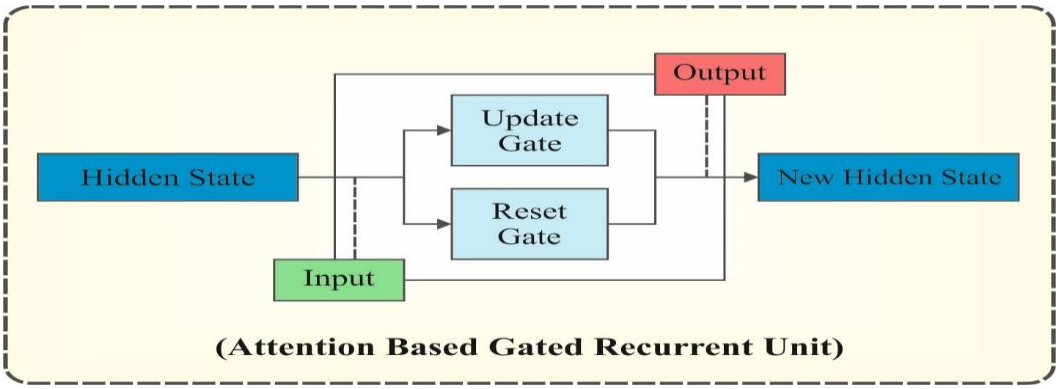

Fig. 2. Architecture of AGRU

DOI: https://doi.org/10.54216/IJAACI.060101
Received: July 24, 2023 Revised: September 18, 2023 Accepted: February 02, 2024

3.3. Process involved in TROA-based Hyperparameter Tuning

Lastly, the TROA is utilized for the optimal hyperparameter selection of the AGRU algorithm. Tyrannosaurus rex also known as T. rex or T-Rex, is the largest carnivore creature in the world and best-represented theropods [21]. They particularly live in western North America. T. rex is an apex predator such as Wolf, Lion, and so on. On the other hand, some experiments have proved that T. rex is a scavenger and apex predators. Among all dinosaurs of its type, it is well-known for its hunting behaviours since it has a stronger bite force.

The three major steps of TROA are given below:

1) Initialization

As a population-based approach, TROA produces the prey randomly in the search range. Assume x as the position or location of prey, and it can be stochastically produced in the search range from up and low boundaries as follows.

$$X_i = rand(np, dim) * (ub - lb) + lb \quad (4)$$

In Eq. (4), $X_i = [x_1, x_2, \cdots x_n]$ represents the target location and $i = 1,2 \cdots n$, where, n represents the size, the dimension of the search range is dim, $and\ ub$ and lb are the upper and lower limits. np is the No. of population,

The black circle characterizes the T. rex location and the Green circles signify the prey location.

2) Hunting and chasing

Similar to the apex predator, T. rex hunts the target prey. Once the prey is spotted, it hunts the target. Sometimes the prey may run away, or it protects itself from hunting. The T. rex hunting includes catching and chasing the target, conversely, the T. rex hunts, it hunts at random manner.

$$X_{new} = \begin{Bmatrix} X_{new} & if\ rand() < Er \\ Random & else \end{Bmatrix} \quad (5)$$

In Eq. (5), the Estimate of getting the prey scattered is Er, once the T. rex begin its hunting, the prey begins to shatter and hunt the target by changing the place as follows.

$$X_{new} = x + rand() * sr * (tpos * tr - target * pr) \quad (6)$$

In Eq. (6), the rate of success of hunting ranges in the interval $[0.1,1]$ is sr. Once the rate of success is 0, then the prey has run away and hunting becomes unsuccessful, accordingly, the prey position should be modified. Target is the least possible prey location for the T. rex location. T. rex chasing rate is tr. Consider that T. rex's average running rate of 30 miles/h and walking speed of 6.7 miles/h. The prey's running speed is pr which ranges from zero to one, but the running speed of the prey must be lesser than the T. rex rate.

3) Selection

The selection technique relies on the prey position, viz., existing and the prior locations of the target prey. Once the T. rex has failed, the prey position becomes 0, if the prey escapes or defends itself from hunting.

$$X_i^{k+1} = \begin{cases} update\ the\ target\ position\ if\ f(X) < f(X_{new}) \\ target\ is\ zero \qquad otherwise \end{cases} \quad (7)$$

In Eq. (7), the FF for the random initial location of the prey is $f(X)$, and the FF for the updated location of prey is$f(X_{new})$.

Algorithm 1: Pseudocode of TROA
Begin
Randomly initialize the prey position based on Eq. (4).
Compute the fitness according to Eq. (4).
Find the nearby location of prey and makes it the target for T. rex.
Begin T. rex hunting according to Eqs. (5) and (6).
Compute the fitness for the updated location of prey.
If $f(X) < f(X_{new})$then upgrade the prey location and target
If the condition is not satisfied, then the target is equivalent to 0
End

DOI: https://doi.org/10.54216/IJAACI.060101
Received: July 24, 2023 Revised: September 18, 2023 Accepted: February 02, 2024

The fitness selection is a major factor in the TROA method. Encoding solution is employed to evaluate the goodness of solution candidate. Here, the accuracy values are the most important condition intended to design an FF.

$$Fitness = \max(P) \tag{8}$$

$$P = \frac{TP}{TP + FP} \tag{9}$$

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

4. Results and Discussion

The simulation analysis of the AIR-TROADL model is tested on the UCM dataset [22]. The database contains a set of 100 images with 21 classes as given in Table 1. The database encompasses images with the same size of 256x256 pixels.

Table 1: Details on database

Labels	Classes	No. of Instances
C0	Agricultural	100
C1	Airplane	100
C2	Baseballdiamond	100
C3	Beach	100
C4	Buildings	100
C5	Chaparral	100
C6	Denseresidential	100
C7	Forest	100
C8	Freeway	100
C9	Golgcourse	100
C10	Harbor	100
C11	Intersection	100
C12	Mediumresidential	100
C13	Mobilehomepart	100
C14	Overpass	100
C15	Parkinglot	100
C16	River	100
C17	Runway	100
C18	Sparseresidential	100
C19	Storagetanks	100
C20	Tenniscourt	100
Total No. of Instances		**2100**

The experimental outcomes of the AIR-TROADL methodology with the test database is demonstrated in Fig. 3. The confusion matrices provided by the AIR-TROADL method at 70:30 of the TR Phase (TRPH) /TS Phase (TSPH) is shown in Figs. 3a-3b. The outcomes denoted that the AIR-TROADL algorithm is accurately identified and categorized for each 21 class. Furthermore, the PR examination of the AIR-TROADL methodology is revealed in Fig. 3c. The simulation value specified that the AIR-TROADL technique obtains superior PR performance with each 21 class. Besides, the ROC analysis of the AIR-TROADL technique is shown in Fig. 3d. The outcome defined that the AIR-TROADL model leads to effectual outcomes with enriched ROC values on different 21 classes.

DOI: https://doi.org/10.54216/IJAACI.060101
Received: July 24, 2023 Revised: September 18, 2023 Accepted: February 02, 2024

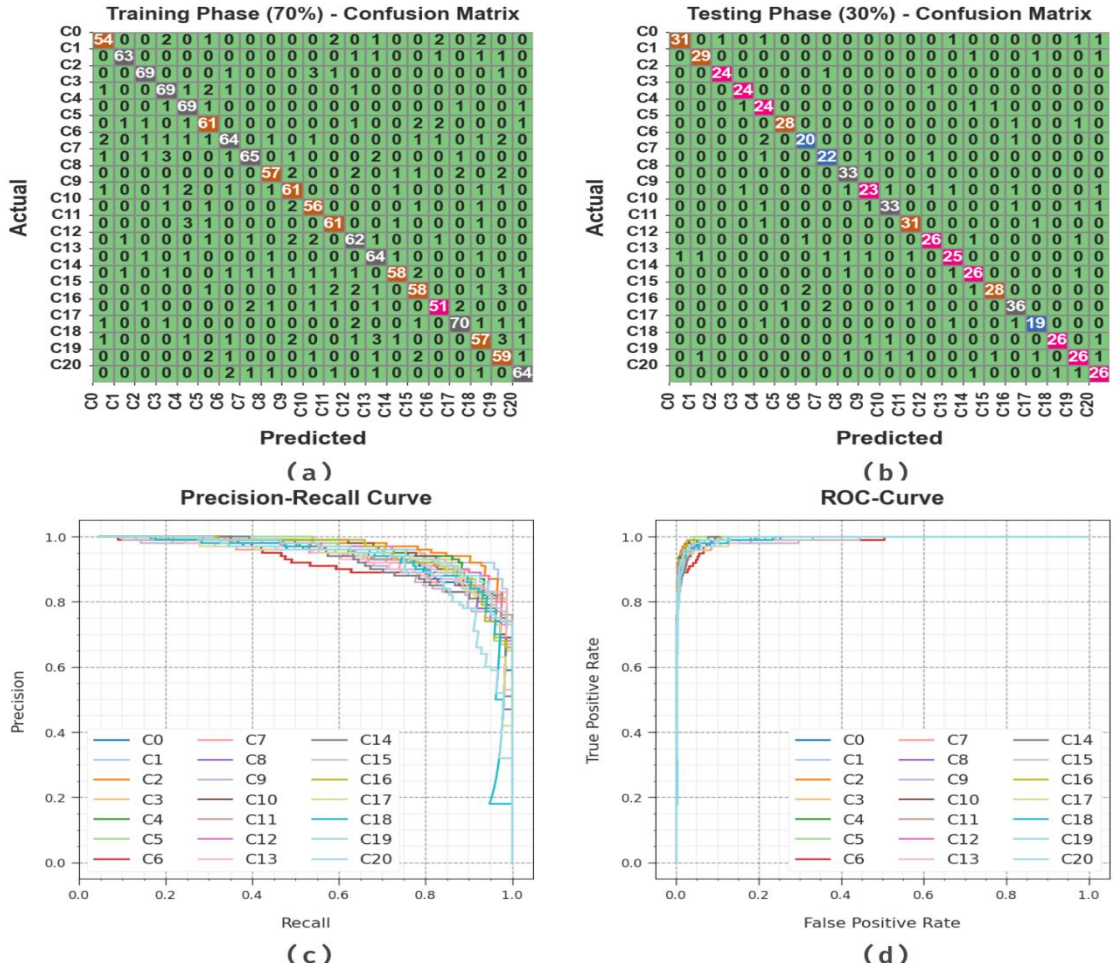

Fig. 3. Classifier performance (a-b) Confusion matrices, (c) PR curve, and (d) ROC

A detailed classification analysis of the AIR-TROADL technique is demonstrated under various classes is shown in Table 2 and Fig. 4. The experimental value infer that the AIR-TROADL model obtains effective results with all 21 classes. With 70% of the TRPH, the AIR-TROADL model attains average $accu_y$, $prec_n$, $reca_l$, F_{score}, and AUC_{score} values of 98.85%, 87.96%, 87.86, 87.84%, and 93.63%, correspondingly. Moreover, based on 30% of the TSPH, the AIR-TROADL method obtains average $accu_y$, $prec_n$, $reca_l$, F_{score}, and AUC_{score} values of 98.94%, 88.93%, 88.96%, 88.85% and 94.20, correspondingly.

Table 2: Classifier outcome of AIR-TROADL technique with 70:30 of TRPH/TSPH

Class Labels	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}	AUC_{score}
TRPH (70%)					
C0	98.84	88.52	84.38	86.40	91.94
C1	99.46	95.45	92.65	94.03	96.22
C2	99.25	93.24	92.00	92.62	95.82
C3	98.84	86.25	92.00	89.03	95.61
C4	99.05	87.34	94.52	90.79	96.90
C5	98.64	84.72	87.14	85.92	93.18
C6	98.50	87.67	83.12	85.33	91.24
C7	98.98	92.86	86.67	89.66	93.15
C8	98.91	90.48	85.07	87.69	92.32
C9	98.71	85.92	87.14	86.52	93.21
C10	98.91	84.85	90.32	87.50	94.81
C11	99.05	88.41	91.04	89.71	95.24

DOI: https://doi.org/10.54216/IJAACI.060101
Received: July 24, 2023 Revised: September 18, 2023 Accepted: February 02, 2024

C12	98.64	84.93	87.32	86.11	93.27
C13	98.84	84.21	92.75	88.28	95.95
C14	98.78	92.06	81.69	86.57	90.67
C15	98.57	85.29	84.06	84.67	91.67
C16	98.91	87.93	85.00	86.44	92.25
C17	99.05	92.11	89.74	90.91	94.66
C18	98.30	83.82	80.28	82.01	89.75
C19	98.44	79.73	88.06	83.69	93.50
C20	99.12	91.43	90.14	90.78	94.86
Average	**98.85**	**87.96**	**87.86**	**87.84**	**93.63**
TSPH (30%)					
C0	99.05	96.88	86.11	91.18	92.97
C1	99.21	93.55	90.62	92.06	95.15
C2	99.68	96.00	96.00	96.00	97.92
C3	99.52	92.31	96.00	94.12	97.83
C4	98.41	77.42	88.89	82.76	93.86
C5	99.52	96.55	93.33	94.92	96.58
C6	99.05	86.96	86.96	86.96	93.23
C7	98.89	84.62	88.00	86.27	93.67
C8	99.37	89.19	100.00	94.29	99.66
C9	98.25	85.19	76.67	80.70	88.00
C10	98.73	91.67	86.84	89.19	93.17
C11	99.21	91.18	93.94	92.54	96.72
C12	99.05	89.66	89.66	89.66	94.58
C13	98.41	86.21	80.65	83.33	89.99
C14	98.73	83.87	89.66	86.67	94.41
C15	99.21	93.33	90.32	91.80	94.99
C16	98.57	87.80	90.00	88.89	94.58
C17	99.05	86.36	86.36	86.36	92.94
C18	99.37	96.30	89.66	92.86	94.74
C19	97.94	81.25	78.79	80.00	88.89
C20	98.57	81.25	89.66	85.25	94.33
Average	**98.94**	**88.93**	**88.96**	**88.85**	**94.20**

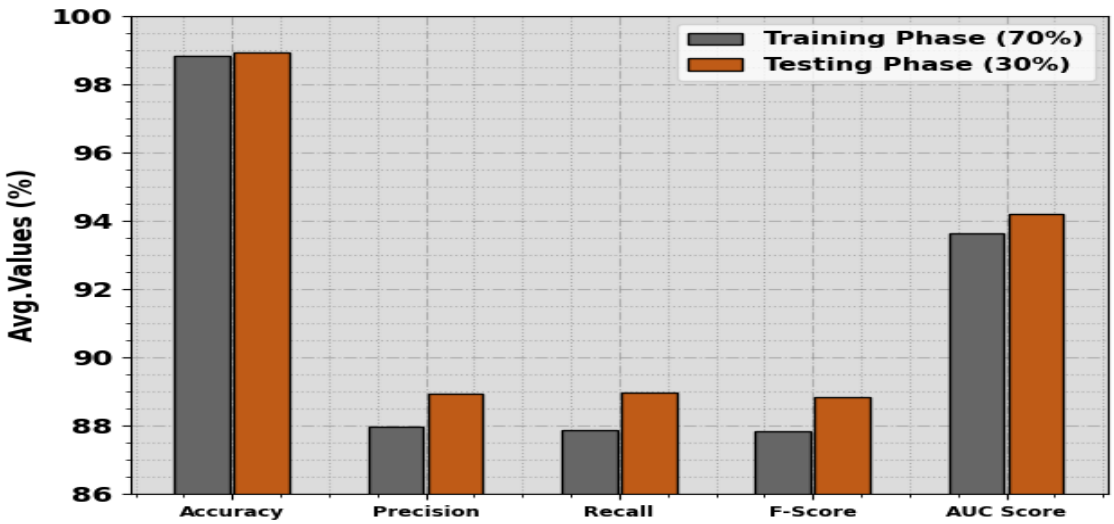

Fig. 4. Average of AIR-TROADL methodology with 70:30 of TRPH/TSPH

DOI: https://doi.org/10.54216/IJAACI.060101
Received: July 24, 2023 Revised: September 18, 2023 Accepted: February 02, 2024

The TRA and TES $accu_y$ curves are plotted to define the performance of the AIR-TROADL algorithm, as given in Fig. 5. The TRA and TES $accu_y$ curves exhibit the outcomes of the AIR-TROADL model over various epochs. The experimental value offers considerable details about the learning task and generalization capabilities of the AIR-TROADL model. With increasing epochs, the TRA and TES $accu_y$ curves achieve improvement. It is remarked that the AIR-TROADL method obtains improved TES accuracy that can be proficient in identifying the patterns in the TRA and TES datasets.

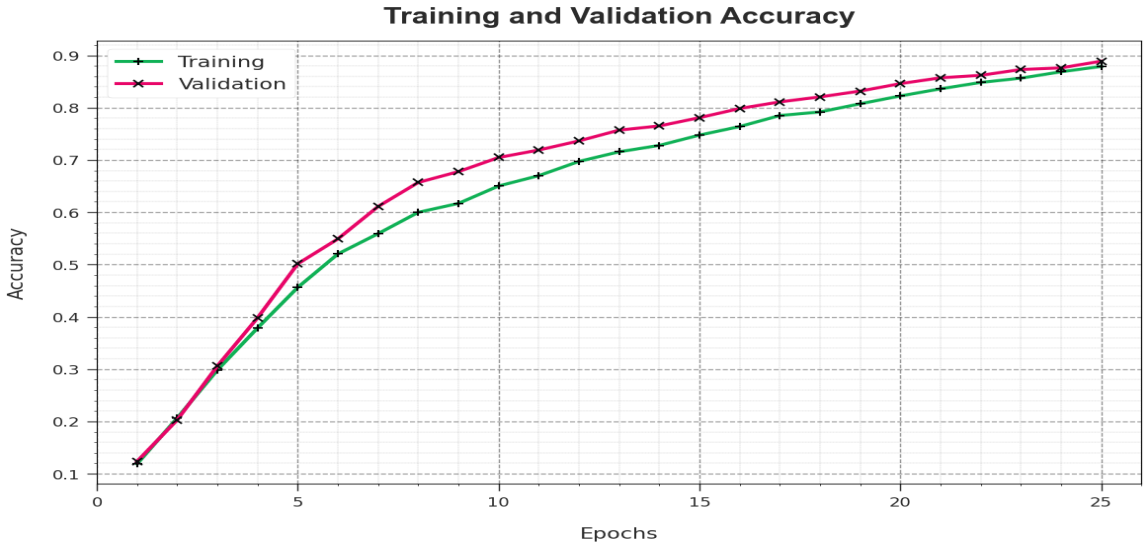

Fig. 5. $Accu_y$ Curve of AIR-TROADL algorithm with 70:30 of TRPH/TSPH

The TRA and TES loss values of the AIR-TROADL method over epochs. The TRA loss indicates the model loss gets decreased over epochs is demonstrated in Fig. 6. The loss values become minimized as the model adjusts the weight to decrease the predictive error with the TRA and TES data. The loss curves show that what extent the model fits the training dataset. The TRA and TES loss is gradually diminished and demonstrates that the AIR-TROADL model learns the patterns revealed in the TRA and TES set. The AIR-TROADL approach changes the parameters to diminish the variance between the actual and predicted training label

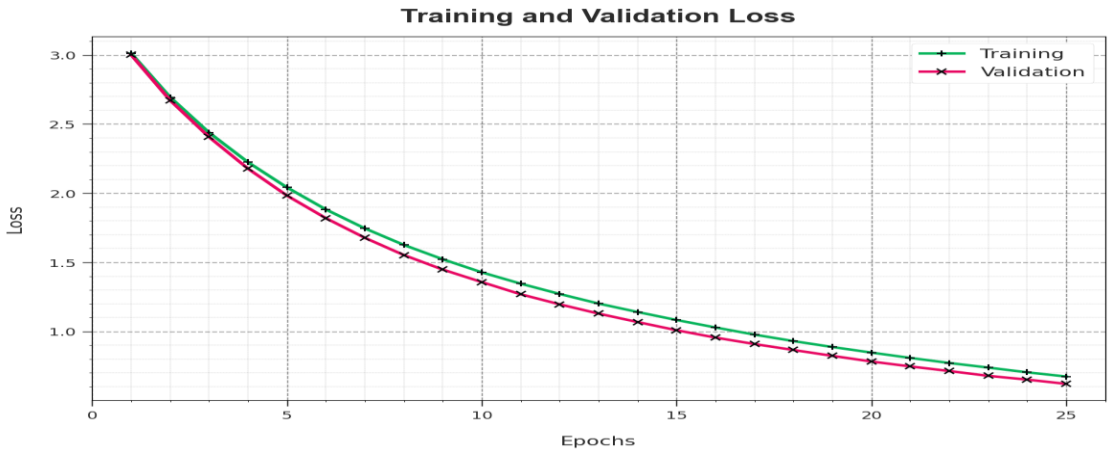

Fig. 6. Loss curve of AIR-TROADL method with 70:30 of TRPH/TSPH

Table 3 and Fig. 7 highlight the comparison study of the AIR-TROADL method with existing techniques [23]. The simulated value indicates that the CAGNet-BiLSTM technique obtains worse outcomes while the VGGNet, CAVGG-BiLSTM, and GNet-RBF models attain slightly improved performance. Meanwhile, the RNet-RBF and DLIRV2-AIC models attain considerably closer outcomes. However, the AIR-TROADL technique offers better performance with a maximum $accu_y$ of 98.94%, F_{score} of 88.93%, $prec_n$ of 88.96%, and $reca_l$ of 88.85%.

DOI: https://doi.org/10.54216/IJAACI.060101
Received: July 24, 2023 Revised: September 18, 2023 Accepted: February 02, 2024

Table 3: Comparative outcome of AIR-TROADL method with existing DL techniques

Model	$Accu_y$	F_{score}	$Prec_n$	$Reca_l$
VGGNet Algorithm	94.40	78.50	79.10	82.30
CAVGG-Bi LSTM	94.50	79.80	79.30	84.00
GNet-RBF Model	94.10	81.50	80.00	86.80
CAGNet-Bi LSTM	91.40	81.80	79.90	87.10
RNet-50 Model	97.90	79.70	80.90	82.00
RNet-RBF Model	98.30	80.60	79.90	84.60
DLIRV2-AIC	98.57	86.30	87.40	87.90
AIR-TRAVEL	98.94	88.93	88.96	88.85

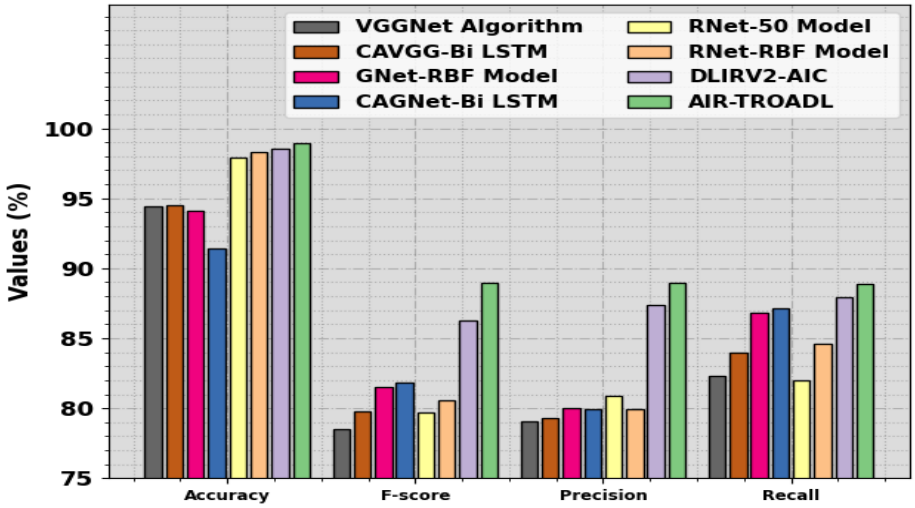

Fig. 7. Comparative outcome of AIR-TROADL model with existing DL techniques

In Table 4 and Fig. 8, the computation time (CT) analysis of the AIR-TROADL algorithm is compared with recent DL approaches. The outcomes indicated that the AIR-TROADL technique offers a reduced CT of 0.53s. On the other hand, VGGNet, CAVGG-Bi LSTM, GNet-RBF, CAGNet-Bi LSTM, RNet-50, RNet-RBF, and DLIRV2-AIC models obtain increased CT values of 1.58s, 1.45s, 2.03s, 1.48s, 2.02s, 1.68s, and 1.43s respectively. These simulated values show an exceptional performance of the AIR-TROADL technique.

Table 4: CT outcome of AIR-TROADL method with existing DL techniques

Model	Computational Time (sec)
VGGNet Algorithm	1.58
CAVGG-Bi LSTM	1.45
GNet-RBF Model	2.03
CAGNet-Bi LSTM	1.48
RNet-50 Model	2.02
RNet-RBF Model	1.68
DLIRV2-AIC	1.43
AIR-TROADL	0.53

DOI: https://doi.org/10.54216/IJAACI.060101
Received: July 24, 2023 Revised: September 18, 2023 Accepted: February 02, 2024

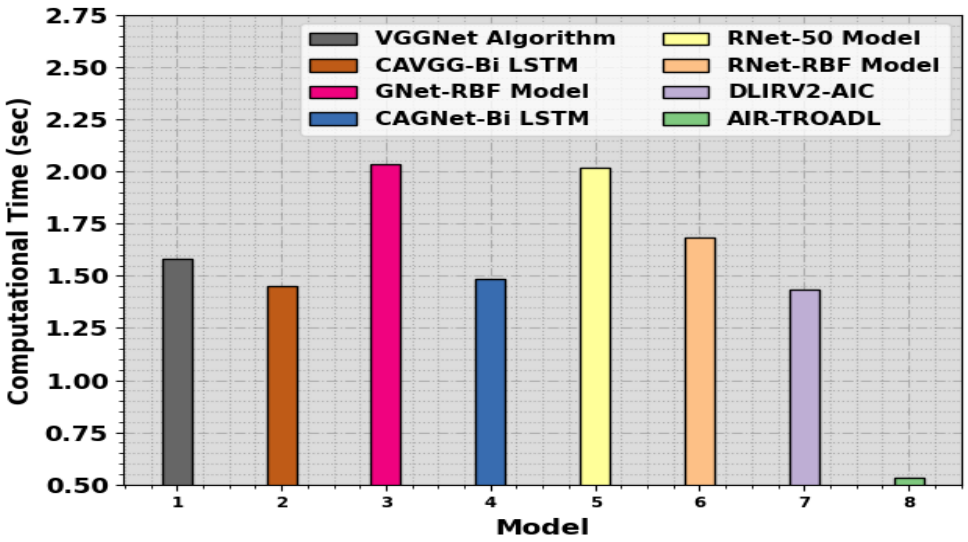

Fig. 8. CT outcome of AIR-TROADL algorithm with recent DL systems

5. Conclusion

In this study, we have established an automatic image classification algorithm, called the AIR-TROADL method on the IoT-aided UAV network. The AIR-TROADL method aims to examine the UAV images for the identification and classification of images into distinct categories. In the projected AIR-TROADL method, three major phases are involved namely enhanced ShuffleNet model, TROA-based hyperparameter tuning, and AGRU-based classification. Besides, the hyperparameter tuning of the enhanced ShuffleNet model is performed by the use of TROA, which in turn boosts the classification performance. Finally, the classification of images takes place using the AGRU approach. A series of simulations can be carried out to exhibit the superior performance of the AIR-TROADL technique. The comparative outcomes emphasized that the AIR-TROADL method reaches high efficiency over its recent approaches in terms of distinct measures.

References

[1] G. Cheng, J. Han, P. Zhou, and L. Guo, ‘‘Multi-class geospatial object detection and geographic image classification based on collection of part detectors,’’ ISPRS J. Photogramm. Remote Sens., vol. 98, pp. 119–132, Dec. 2014.

[2] A. M. Cheriyadat, ‘‘Unsupervised feature learning for aerial scene classification,’’ IEEE Trans. Geosci. Remote Sens., vol. 52, no. 1, pp. 439–451, Jan. 2014.

[3] F. Zhang, B. Du, and L. Zhang, ‘‘Saliency-guided unsupervised feature learning for scene classification,’’ IEEE Trans. Geosci. Remote Sens., vol. 53, no. 4, pp. 2175–2184, Apr. 2015.

[4] L. Ye, L. Wang, Y. Sun, R. Zhu, and Y. Wei, ‘‘Aerial scene classification via an ensemble extreme learning machine classifier based on discriminative hybrid convolutional neural networks features,’’ Int. J. Remote Sens., vol. 40, no. 7, pp. 2759–2783, Apr. 2019.

[5] L.-J. Zhao, P. Tang, and L.-Z. Huo, ‘‘Land-use scene classification using a concentric circle-structured multiscale Bag-of-Visual-Words model,’’ IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 7, no. 12, pp. 4620–4631, Dec. 2014.

[6] A. Coates, A. Ng, and H. Lee, ‘‘An analysis of single-layer networks in unsupervised feature learning,’’ in Proc. 14th Int. Conf. Artif. Intell. Statist., 2011, pp. 215–223.

[7] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, ‘‘Gradient-based learning applied to document recognition,’’ Proc. IEEE, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

[8] A. Darwish, A. E. Hassanien, and S. Das, ‘‘A survey of swarm and evolutionary computing approaches for deep learning,’’ Artif. Intell. Rev., vol. 53, no. 3, pp. 1767–1812, Mar. 2020.

[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, ‘‘Imagenet classification with deep convolutional neural networks,’’ in Proc. Adv. Neural Inf. Process. Syst. (NIPS), 2012, pp. 1097–1105.

DOI: https://doi.org/10.54216/IJAACI.060101
Received: July 24, 2023 Revised: September 18, 2023 Accepted: February 02, 2024

[10] F. Al-Turjman, H. Zahmatkesh, and L. Mostarda, "Quantifying uncertainty in the Internet of medical things and big-data services using intelligence and deep learning," IEEE Access, vol. 7, pp. 115749–115759, 2019.
[11] SELVAM, R.P., 2022. Earthworm optimization with deep transfer learning enabled aerial image classification model in IoT-enabled UAV networks. *Fusion: Practice and Applications*, *7*(1), pp.41-1.
[12] Mitra, A., Bera, B. and Das, A.K., 2021, May. Design and testbed experiments of public blockchain-based security framework for IoT-enabled drone-assisted wildlife monitoring. In *IEEE INFOCOM 2021-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)* (pp. 1-6). IEEE.
[13] Khan, I.U., Alturki, R., Alyamani, H.J., Ikram, M.A., Aziz, M.A., Hoang, V.T. and Cheema, T.A., 2021. RSSI-controlled long-range communication in secured IoT-enabled unmanned aerial vehicles. *Mobile information systems*, *2021*, pp.1-11.
[14] Bera, B., Saha, S., Das, A.K., Kumar, N., Lorenz, P. and Alazab, M., 2020. Blockchain-envisioned secure data delivery and collection scheme for 5g-based IoT-enabled internet of drones environment. *IEEE Transactions on Vehicular Technology*, *69*(8), pp.9097-9111.
[15] Liu, Y., Liu, K., Han, J., Zhu, L., Xiao, Z. and Xia, X.G., 2020. Resource allocation and 3-D placement for UAV-enabled energy-efficient IoT communications. *IEEE Internet of Things Journal*, *8*(3), pp.1322-1333.
[16] Anand, T., Sinha, S., Mandal, M., Chamola, V. and Yu, F.R., 2021. AgriSegNet: Deep aerial semantic segmentation framework for IoT-assisted precision agriculture. *IEEE Sensors Journal*, *21*(16), pp.17581-17590.
[17] Mishra, S., 2023. Internet of Things enabled deep learning methods using unmanned aerial vehicles enabled integrated farm management. *Heliyon*, *9*(8).
[18] Peng, Y., Liu, Y., Li, D. and Zhang, H., 2022. Deep reinforcement learning based freshness-aware path planning for UAV-assisted edge computing networks with device mobility. *Remote Sensing*, *14*(16), p.4016.
[19] Fang, S., Yang, J., Wang, M., Liu, C. and Liu, S., 2022. An improved image classification method for cervical precancerous lesions based on ShuffleNet. *Computational Intelligence and Neuroscience*, *2022*.
[20] Xiong, L., Zhang, L., Huang, X., Yang, X., Huang, W., Zeng, H. and Tang, H., 2021. DCAST: a spatiotemporal model with DenseNet and GRU based on attention mechanism. *Mathematical Problems in Engineering*, *2021*, pp.1-12.
[21] Sahu, V.S.D.M., Samal, P. and Panigrahi, C.K., 2023. Tyrannosaurus optimization algorithm: A new nature-inspired meta-heuristic algorithm for solving optimal control problems. *e-Prime-Advances in Electrical Engineering, Electronics and Energy*, *5*, p.100243.
[22] http://weegee.vision.ucmerced.edu/datasets/landuse.html
[23] Minu, M.S. and Canessane, R.A., 2022. Deep learning-based aerial image classification model using inception with residual network and multilayer perceptron. *Microprocessors and Microsystems*, *95*, p.104652.

DOI: https://doi.org/10.54216/IJAACI.060101
Received: July 24, 2023 Revised: September 18, 2023 Accepted: February 02, 2024