Signal, Image and Video Processing
https://doi.org/10.1007/s11760-023-02859-7

ORIGINAL PAPER

HFRAS: design of a high-density feature representation model for effective augmentation of satellite images

Dipen Saini[1] · Rachit Garg[1] · Rahul Malik[2] · Deepak Prashar[1] · M. Faheem[3]

Received: 21 July 2023 / Revised: 15 October 2023 / Accepted: 20 October 2023

Abstract
Efficiently extracting features from satellite images is crucial for classification and post-processing activities. Many feature representation models have been created for this purpose. However, most of them either increase computational complexity or decrease classification efficiency. The proposed model in this paper initially collects a set of available satellite images and represents them via a hybrid of long short-term memory (LSTM) and gated recurrent unit (GRU) features. These features are processed via an iterative genetic algorithm, identifying optimal augmentation methods for the extracted feature sets. To analyse the efficiency of this optimization process, we model an iterative fitness function that assists in incrementally improving the classification process. The fitness function uses an accuracy & precision-based feedback mechanism, which helps in tuning the hyperparameters of the proposed LSTM & GRU feature extraction process. The suggested model used 100 k images, 60% allocated for training and 20% each designated for validation and testing purposes. The proposed model can increase classification precision by 16.1% and accuracy by 17.1% compared to conventional augmentation strategies. The model also showcased incremental accuracy enhancements for an increasing number of training image sets.

Keywords Satellite images · Classification · Augmentation · Long short-term memory · Gated recurrent unit

1 Introduction

1.1 Background

Training data regarding quality and quantity is crucial for the effectiveness of models and algorithms within the dynamic field of artificial intelligence and machine learning. Convolutional neural networks (CNNs) have gained significant traction as the most efficient method for image classification. Existing CNN models continue to have substantial flaws. It is common for datasets to lack training samples or to have an uneven class distribution [1]. Creating an extensive image collection takes time and resources. Data augmentation has become a potent strategy for improving machine learning models' resilience, generalizability, and overall efficacy [2].

Data augmentation can be used to meet the needs of the different types of training data and the amount of training data. In classification tasks, augmented data can also address the challenge of classes exhibiting excessive similarity or substantial disparity. Data augmentation is significant when a model is used to analyse the parts of an image. Let us say we want to pull out the details of a ship from a satellite image. A large amount of data is required because the ship's location, shape, and size constantly change, meaning the dataset needs to grow. To train a model for semantic segmentation, you must give it pairs of data, including the original image and a semantically labelled image. These two data pairs must be provided to train the model. As a direct result of this, we will simultaneously have to generate two identical images [3–5]. A unique progressive remote sensing ship image data augmentation approach was developed using ship simulation

✉ M. Faheem
muhammad.faheem@uwasa.fi

Dipen Saini
er.dipensaini@gmail.com

Rachit Garg
rachit.garg@lpu.co.in

Rahul Malik
rahulm@srmist.edu.in

Deepak Prashar
deepak.prashar@lpu.co.in

1 SCSE, Lovely Professional University, Punjab, India

2 CSE, SRM University, Chennai, India

3 School of Technology and Innovations, University of Vaasa, 65200 Vaasa, Finland

Published online: 11 November 2023

samples and an NST-based network. There are two steps to their procedure. A visible-light imaging simulation system is used to produce samples for the ship simulation first from images taken in the actual environment. The training dataset is made more diverse by this process. Second, researchers may recreate the simulated aesthetic in the real world utilizing some authentic images and a newly created NST-based network called Sim2RealNet [6]. Several ship targets were used to assess the suggested approach for classifying remote sensing images.

Conventional data augmentation methods, including those involving geometric transformations such as flipping, translation, and rotation, are used to generate augmented data. This enhanced dataset is subsequently utilized for training purposes, allowing the generation of an improved deep model [7]. A novel data augmentation technique, random image cropping and patching (RICAP), has been developed. This technique involves randomly cropping four images patched together to create a unique training image. In addition, the RICAP technique incorporates the class labels of the four images, leading to a notable benefit in using soft labels. An evaluation of RICAP using contemporary convolutional neural networks, including the shake-shake regularization model, is considered at the forefront of the field [1]. Due to global warming, forest fires have become a major cause of ecological harm. Because of its rapid updates and extensive coverage area, remote sensing (RS) is essential for monitoring forest fires. A major factor affecting classifier performance is the loss of significant image characteristics from the existing basic mixed sample data augmentation (MSDA) algorithms for smoke scene recognition. For MSDA, there is a brand-new technique called CAMMix. Using CAMMix, choose the area and mix the intensity throughout the significance map. A mixed mask that integrates class significance is produced by CAMMix using an auxiliary (AUX) classifier such that the distribution of the mixed sample closely resembles that of the original data [8].

A data augmentation technique, such as CutOut with iterative spatial–spectral training (ISST) [9–11], requires the input square area to be randomly masked before the training begins. Both the resilience of the convolutional neural network to errors and its overall performance might see an improvement as a result of this change. Generative adversarial networks have also been used for data augmentation in conventional RGB and satellite images. The method described is an unsupervised approach to data generation [12]. The generative model typically consists of the generator and the discriminator, which function like game components. The former's primary objective is to acquire the ability to produce visually authentic images and deceive the discriminator, distinguishing between genuine and artificially generated images. Several examples of generative adversarial networks (GANs) that have been utilized in the context of satellite imagery are DCGAN, CycleGAN, and SSSGAN. The progressive growth GAN technique generated high-resolution images [13]. Because these methods generate fresh samples, which are then rapidly modified before being stitched together at the image level, it might be challenging to determine the point at which one item concludes and another begins. Because borders play a significant part in semantic segmentation activities, the approaches described up to this point are inappropriate for upgrading samples for these pursuits. Because of its capacity to provide correct results, generative adversarial networks, often known as GAN, have been one of the most prominent unsupervised methods in recent years [14–16]. For instance, DCGAN and Marta GAN have been suggested to improve the image quality acquired by remote sensing. In contrast to the deep convolutional generative adversarial network (DCGAN), the Marta generative adversarial network (Marta GAN) can generate images with greater detail and resolution. Because of remote sensing images' inherent ambiguity and complexity, GAN-based augmentation algorithms have a tough time learning the target objects' distribution properties, ultimately leading to a low-quality augmentation effect. One of the many reasons why GAN-based augmentation algorithms have such a difficult time increasing image quality is because this is one of those reasons. For instance, the generated images could be of better quality and lack the majority of components generally agreed upon as essential. They can be augmented via UNet and its counterparts [17–19]. In addition, GAN-based augmentation algorithms cannot create matching semantic tag images, which are necessary for semantic segmentation and are typically annotated manually at a significant expense. This is because these images have to be annotated by hand. Because semantic segmentation is such an important endeavour [20–22], this is a considerable disadvantage. Therefore, it would be ideal to develop a technique for enhancing images collected by remote sensing by efficiently addressing the annotation complexity while minimizing the expense. Recently, a kind of convolutional neural network or CNN [23–25] was characterized as having the capacity to effectively satisfy the image translation challenge. This was accomplished via the use of neural networks. In this paper, we show how to use data augmentation as a pre-processing approach for training a deep CNN and empirically evaluate the efficacy of our data augmentation strategy for improving CNN representational power.

1.2 Motivation

As discussed in Sect. 1.1 above, a deep model's ability to describe things depends significantly on how different the training data is. However, the most advanced deep learning techniques in remote sensing mainly focus on making new multilayer representations. We have yet to look into how the

size and variety of the training dataset affect their performance. Deep learning cannot be used to its full potential in remote sensing because there is not enough training data. From what we have discussed, it is clear that researchers have come up with a wide range of feature representation models, most of which are more challenging to compute or have lower classification performance.

1.3 Contribution

This paper discusses these fundamental data limitations that make it hard to use deep learning's full power to classify images from remote sensing. We describe a way that suggests making a high-density feature representation model for efficiently augmenting satellite images to make remote sensing datasets bigger and varied and then using this dataset to train a deep CNN. The proposed model initially collects a small set of available satellite images and represents them via a hybrid of long short-term memory (LSTM) and gated recurrent unit (GRU) features. These features are processed via an iterative genetic algorithm (IGA), identifying optimal augmentation methods for the extracted feature sets. An iterative fitness function is modelled to analyse the efficiency of this optimization process, which assists in the incremental improvement of the classification process. The function uses an accuracy & precision-based feedback mechanism that helps in tuning the hyperparameters of the proposed LSTM & GRU feature extraction process.

In Sect. 4, the suggested model's accuracy, precision, and recall performances were evaluated and contrasted to those of conventional augmentation methods. In addition to providing suggestions for further enhancing the suggested augmentation model's performance in various use cases, this paper concludes with a few insightful observations.

2 Brief review of image augmentation models

A wide variety of deep learning-based techniques are proposed for the augmentation of images, and each of them varies in terms of their quantitative performance measures and qualitative characteristics. Deep CNNs have shown promising results when processing images, but their eloquence might be too exact. Existing datasets may be improved using techniques for data augmentation without introducing unintended bias. Modern CNN designs with more parameters make applying traditional data augmentation techniques useless. Table 1 presents additional discoveries about data augmentation. The information in Table 1 has been gathered from diverse research papers.

Table 1 Overview of different data augmentation techniques

Authors	Year	Model	Findings
Adedeji et al.	2022	DCGAN and WGAN-GP	The generalizability of deep classification models applied to satellite imagery can be enhanced by GANs [26]
Chen et al.	2022	IAug	Instance-level change augmentation (IAug) is a technique for using generative adversarial training to create bitemporal images that involve changes involving numerous structures. Building computer models are composited onto image backgrounds by IAug [27]
Nesteruk et al.	2022	XAug	Xtreme augment (XAug) is an automated technique for cataloguing and improving enormous collections of photographs [28]
Wang et al.	2022	SCNN	Siamese CNN (SCNN) technique that allows pairings of data from the same or different classes to be added to training datasets. New training samples may be produced by adding fresh data to an existing collection [29]
Chen et al.	2022	P-Nets	Prototypical networks (P-Nets) are a potent few-shot learning technique for identifying species in forests. Due to a lack of training data, overfitting still affects few-shot classifiers, which makes it challenging to train correct models [30]
Xie et al.	2022	CAMMix	Using CAMMix, choose the area and mix the intensity throughout the significance map. A mixed mask that integrates class significance is produced by CAMMix using an auxiliary (AUX) classifier such that the distribution of the mixed sample closely resembles that of the original data [8]
Shang et al.	2022	ITSA-MMP	ITSA starts by gradually accumulating data while using a spatial–spectral grouping technique (SRCS). High-quality additions to a sample pool are found using the box plot for representative sample selection (BPRSS) method (ASS)

Table 1 (continued)

Authors	Year	Model	Findings
			MMP projects the hyperspectral image into a lower dimensional subspace based on the ASS to enhance interclass separability and better comprehend the local structure of the data manifold. SVMs are then used to classify data that has been reduced via MMP [40]
Du et al.	2022	ENC DEC	An adversarial encoding network is used for physics-related data mining from synthetic and real-world datasets. Encoders (ENC) may extract valuable information from fictitious images via adversarial learning. To validate the encoding of traits, a classifier is utilized. A decoder (DEC) will recreate the encoded properties in the original image to avoid data loss [11]
Shen et al.	2021	Cut Mix	A cross-directional attention module is suggested to examine how images from before and after a disaster relate. The CutMix technique is used to deal with the problem of challenging classes [2]
Pérez-Hernández et al.	2021	Det DSCI	A two-level resolution-independent critical infrastructure detection (DetDSCI) methodology that employs a classification model to first identify the input image's spatial resolution before analysing it with the appropriate detector [42]
Luo et al.	2021	Deep Lab	This model is capable of picking up on subtle changes in subpixel-level remote sensing images. The generalization skills of the network are assessed using the satellite change detection datasets from Landsat 8, Google Earth, and Onera. The findings demonstrate a network accuracy of 95.1% and strong generalization abilities [43]
Kim et al.	2021	UNet	A data augmentation strategy for visual surveillance to enhance performance while maintaining their current network [31]
Xia et al.	2021	MDCF	Researchers develop a straightforward random erasing approach to improve cloud/snow identification [32]
Yamashita et al.	2021	RST	A data augmentation method that generates domain-independent visual representations using random style transfer (RST) from non-medical style sources like paintings may benefit computational pathology [33]
Kim et al.	2021	LA-CNN	A brand-new data augmentation technique called local augment (LA) is presented. LA alters the local bias attribute to produce distinctive augmented images and increase the network's augmentation effectiveness [21]
Nalepa et al.	2020	DNN	Combining the test-time technique with training-time strategies improves classification accuracy levels [44]
Zhu et al.	2020	SAGAN	First, a better self-attention generative adversarial network generates brand-new X-ray images of limited goods (SAGAN). Then, researchers demonstrate how to use a technique based on CycleGAN to transform familiar images into X-rays [34]
Pan et al.	2020	SSFA	A self-supervised feature augmentation (SSFA) network uses sampled photographs as inputs and produces features similar to those in upscaled images [35]
Zhang et al.	2020	BigAug	The findings demonstrate that BigAug outperforms 'shallower' stacked, BigAug's performance on an unknown domain exceeds 'deeper' stacked when trained on relatively small datasets, and BigAug outperforms conventional augmentation when trained on rather large datasets [36]
Takahashi et al.	2019	RI CAP	The random image cropping and patching (RICAP) technique involves the random cropping of four distinct images combined to form a new training image. In addition, the RICAP technique introduces a mixing of class labels among the four images, conferring a distinct advantage to the soft brands [1]
Mahdizadehaghdam et al.	2019	SPARSE GAN	An addition in GANs is done in which a third player, called a reconstructor, is added to maintain the desired diversity. The input–output relation is maintained, and an auto-encoding scheme is used. The output results are better than normal GAN results [37]

3 Design of the proposed High-density Feature Representation model for effective Augmentation of Satellite images

As per the analysis of existing feature representation models for augmenting satellite images, most have higher computational complexity or lower classification efficiency levels. The design of a high-density feature representation model for efficient augmentation of satellite images is discussed in this part to address these problems. As observed in Fig. 1, the proposed model initially collects a small set of available satellite images and represents them via a hybrid of short-term memory (LSTM) & gated recurrent unit (GRU) features. These features are processed via an iterative genetic algorithm (IGA), identifying optimal augmentation methods for the extracted feature sets. An iterative fitness function is modelled to analyse the efficiency of this optimization process, which assists in the incremental improvement of the classification process. The function uses an accuracy & precision-based feedback mechanism that helps in tuning the hyperparameters of the proposed LSTM and GRU feature extraction process.

At first, the proposed model pulls out many different sets of features from each image. These feature sets are extracted via a novel combination of long short-term memory

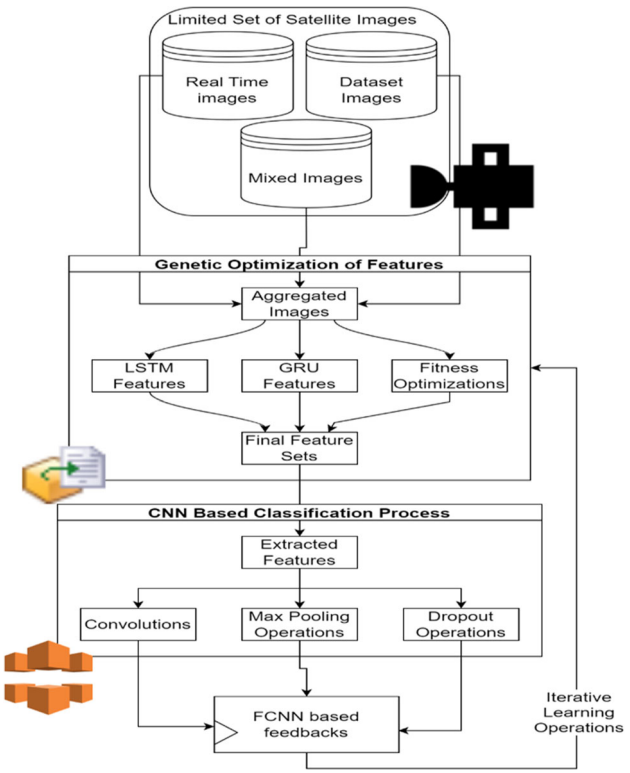

Fig. 1 Flow of the proposed feature representation process

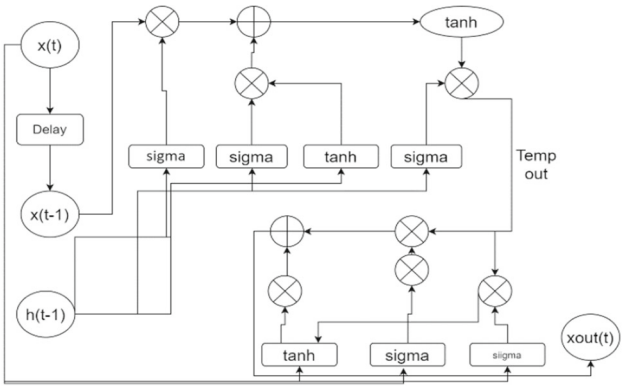

Fig. 2 Design of the LSTM & GRU-based feature extraction process

(LSTM) with gated recurrent unit (GRU)-based representation techniques. The reason for combining these techniques is due to their differential feature representation characteristics. The fused feature extraction model is depicted in Fig. 2, where different variance operations are combined with tangent operations to identify multimodal feature sets.

The model initially extracts initialization (i), temporal feature (f), and temporal output features via Eqs. 1, 2, and 3 as follows,

$$i = \text{var}\left(x_{in} * U^{i} + h_{t-1} * W^{i}\right) \tag{1}$$

$$f = \text{var}\left(x_{in} * U^{f} + h_{t-1} * W^{f}\right) \tag{2}$$

$$o = \text{var}\left(x_{in} * U^{o} + h_{t-1} * W^{o}\right) \tag{3}$$

where $U \& W$ represents variance constants for the LSTM & GRU processes, while h is a kernel matrix used to activate of these features [38, 39]. These features are combined to form a temporal convolutional feature set (C) via Eq. 4,

$$C_t' = \tanh\left(x_{in} * U^{g} + h_{t-1} * W^{g}\right) \tag{4}$$

All these features are used to generate the output feature matrix via Eq. 5,

$$T_{\text{out}} = \text{var}\left(f_t * x_{in}(t-1) + i * C_t'\right) \tag{5}$$

Based on this output feature matrix, a new kernel matrix is generated via Eq. 6,

$$h_{\text{out}} = \tanh(T_{\text{out}}) * o \tag{6}$$

These temporal output features are further processed via GRU- based operations. To perform these operations, an initial resistance (r) & impedance (z) metric is estimated via

Eqs. 7 & 8 as follows,

$$z = \text{var}(W_z * [h_{\text{out}} * T_{\text{out}}]) \tag{7}$$

$$r = \text{var}(W_r * [h_{\text{out}} * T_{\text{out}}]) \tag{8}$$

These metrics are augmented via Eqs. 9 & 10 to estimate updated kernel metric and output feature metrics as follows,

$$h'_t = \tanh(W * [r * h_{\text{out}} * T_{\text{out}}]) \tag{9}$$

$$x_{\text{out}} = (1 - z) * h'_t + z * h_{\text{out}} \tag{10}$$

These feature sets are capable of representing input images into multimodal sets. However, this feature extraction technique's efficiency must be validated to estimate efficient augmentation operations. To perform this task, an iterative genetic algorithm (IGA) is developed, which assists in evaluating high variance constants for the fused feature extraction process. This IGA model works as per the following process: To start the optimizer, set the following constants,

- Total iterations used for generation & configuration of solutions (N_i)
- Total solutions that will be generated & reconfigured (N_s)
- Rate at which the model will learn from other solutions (L_r)
- Initially, generate N_s solutions as per the following process,
- For each satellite image, generate rotated, zoomed, width shifted, height shifted, and scaled images via augmentation operations.
- Setup the values of $U\&W$ as per Eq. 11 & 12,

$$U = U(\text{Old}) \pm f * \text{STOCH}(L_r, 1) \tag{11}$$

$$W = W(\text{Old}) \pm f * \text{STOCH}(L_r, 1) \tag{12}$$

where $W(Old)\&U(Old)$represents old values for the LSTM & GRU constants, and $STOCH$ represents the production of number sets via a stochastic Markovian process.
- Using an iterative convolutional neural network (CNN), which is covered in the later sections of this text, classify satellite images based on these values by evaluating the LSTM and GRU features for each of the enhanced feature.
- After classification, estimate solution fitness as per Eq. 13,

$$f = \sum_{i=1}^{N_{\text{images}}} \frac{t_p}{t_p + t_n} + \frac{t_p + t_n}{t_p + t_n + f_p + f_n} + \frac{t_p + f_p}{t_p + t_n + f_p} \tag{13}$$

where t_p, t_n, $f_p\&f_n$ represents values of true positive rates, true negative rates, false positive rates, and false negative rates for the classification operations.
- Repeat this process for each solution, and then use Eq. 14 to figure out a solution fitness threshold.

$$f_{th} = \sum_{i=1}^{N_s} f_i * \frac{L_r}{N_s} \tag{14}$$

- Once these solutions are generated, check if $f > f_{th}$, and mark these solutions as 'not to be mutated', while marking all other solutions as 'to be mutated'.
- Scan all solutions for N_i iterations, and modify the solutions that are marked as 'to be mutated'.
- At each iteration, update the fitness and solutions fitness thresholds. The proposed algorithm is depicted in Table 2.

When all possible solutions have been found, pick the one with the highest fitness level and use its features to classify satellite images. This classification is done via a convolutional neural network (CNN), depicted in Fig. 3, wherein various convolutional, max pooling & drop out layers are connected to estimate augmented feature sets. The CNN processes LSTM & GRU features and classifies them into land-specific categories. The designed CNN model initially extracts convolutional feature sets from the LSTM & GRU feature sets via Eq. 16, which assists in extracting many features.

$$\text{Conv}_{\text{out}_{i,j}} = \sum_{a=-\frac{m}{2}}^{\frac{m}{2}} \sum_{b=-\frac{n}{2}}^{\frac{n}{2}} F_{\text{LSTM, GRU}}(i - a,\ j - b) * \text{ReLU}\left(\frac{m}{2} + a,\ \frac{n}{2} + b\right) \tag{15}$$

The window size for convolutional operations is represented by m, n, and a,b, which represent stride sizes, and ReLU represents a rectilinear unit model for activation of feature sets. The parameters are mentioned in Table 3 and design in Fig. 4.

The extracted features are given to a threshold engine, which assists in the estimation of the variance threshold via Eq. 16,

$$f_{\text{th}} = \left(\frac{1}{X_k} * \sum_{x \in X_k} x^{p_k}\right)^{1/p_k} \tag{16}$$

where $X\&p$ represents features' intensity and probability levels tuned by the CNN process.

Table 2 Algorithm of proposed methodology

Input:

Satellite images (dataset)

Parameters for LSTM & GRU feature extraction

Genetic algorithm parameters (e.g., population size, generations)

Hyperparameter tuning parameters (e.g., learning rate, batch size)

Output:

Optimized hyperparameters for LSTM & GRU feature extraction

Process:

1. Initialize a population of candidate augmentation methods using genetic algorithm parameters
2. Initialize the best_accuracy and best_precision variables to track the best performance achieved so far
3. Initialize a population of hyperparameters for LSTM & GRU feature extraction
4. Create an initial LSTM & GRU model with the initial hyperparameters
5. Evaluate the model's accuracy and precision on the dataset
6. Set the initial best_accuracy and best_precision values based on the evaluation
7. Start the genetic algorithm loop:

a. Evaluate the fitness of each candidate augmentation method using the current LSTM & GRU model

b. Select the top-performing augmentation methods based on the fitness

c. Generate a new population of augmentation methods through crossover and mutation

d. Repeat steps a–c for a specified number of generations

8. After genetic algorithm optimization, select the best-performing augmentation method
9. Iterate through hyperparameter tuning process:

a. Adjust hyperparameters (e.g., learning rate, batch size) of LSTM & GRU feature extraction

b. Create a new LSTM & GRU model with the adjusted hyperparameters

c. Evaluate the model's accuracy and precision on the dataset

d. Update best_accuracy and best_precision if better performance is achieved

e. Repeat steps a–c for a specified number of iterations or until convergence

10. Output the optimized hyperparameters for LSTM & GRU feature extraction

The max pooling layer removes all features with $f < f_{th}$, while passing others to consecutive layers. A fully connected neural network (FCNN)-based model is used to classify the characteristics collected at the final layer, aiding in estimating various image classes. This FCNN layer combines different weights (w) and biases with a SoftMax-based activation function (b) as per Eq. 17,

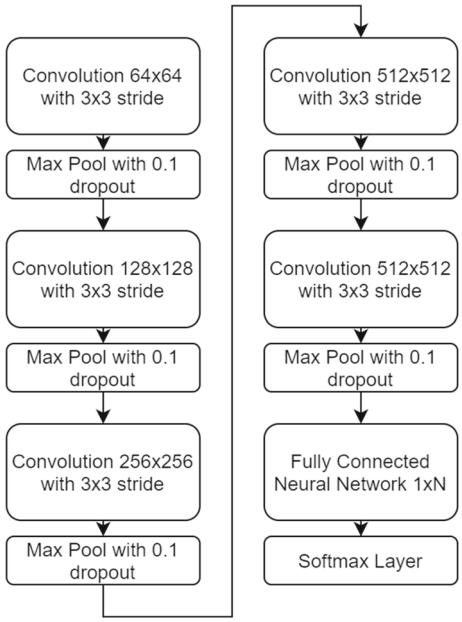

Fig. 3 Design of the CNN algorithm for augmented feature set classification

Table 3 Parameters used in the training model

S. No	Parameters	Value
1	Total layers	16
2	Window size	8,16,32,64,128,256
3	Kernel size	3,5,7,9,11,13,15
4	Activation function	SoftMax
5	Drop out	0.1

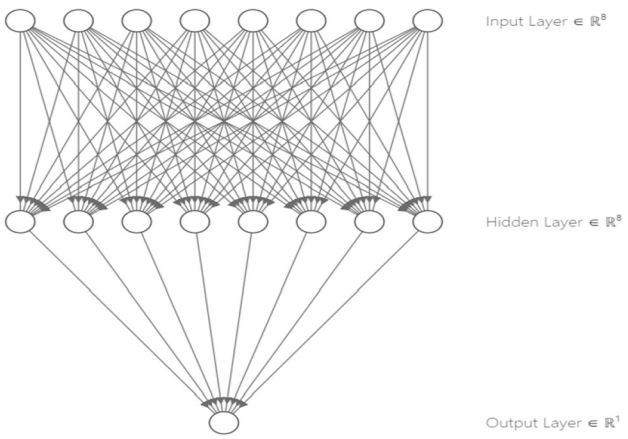

Fig. 4 Design of FCNN layers

$$c_{out} = \text{SoftMax}\left(\sum_{i=1}^{N_f} f_i * w_i + b\right) \tag{17}$$

where N_f represents a number of extracted features by the fused layers. The suggested model can categorize the images with high levels of efficiency since it uses CNN. The following section of this paper evaluates these efficiency levels and compares them to those of standard models.

4 Result analysis and comparison with standard augmentation techniques

The proposed model helps represent input images into multimodal feature sets by combining LSTM and GRU-based feature extraction algorithms. An effective iterative genetic algorithm (IGA) is trained using the collected features to help identify high-density augmentation operations and feature constants. As a result of these operations, the proposed model can improve the accuracy, precision, and recall of different satellite image classification applications. This model was verified on the following datasets to evaluate its performance:

- Copernicus image sets obtained from Kaggle.
- Sentinel image sets obtained from Google Earth Engine
- IEEE data port sets for different areas

These sets were aggregated to form a total of 100 k images, out of which 60% were used to train the model, while 20% each were used for validation & testing purposes. Based on this evaluation, the classification's accuracy (A_c) was compared with ISST [40], GAN [41], and UNet [17] with respect to total validation & test images (TVTI) for different applications. Results of these augmentations can be observed from Fig. 5a, b, and c, wherein different satellite images were used for the classification process.

The accuracy of this model is tabulated in Table 4 as follows,

Considering this evaluation and its visualization in Fig. 6, it can be seen that the proposed model can increase classification accuracy by 16.4% compared to ISST [40], 17.1% compared to GAN [41], and 13.6% compared to UNet [17], making it highly beneficial for a range of real-time classification applications. The reason for this enhancement is the use of accuracy during the optimization of fitness, which assists in estimating high-efficiency augmented feature sets. Table 5 shows the precision levels as follows:

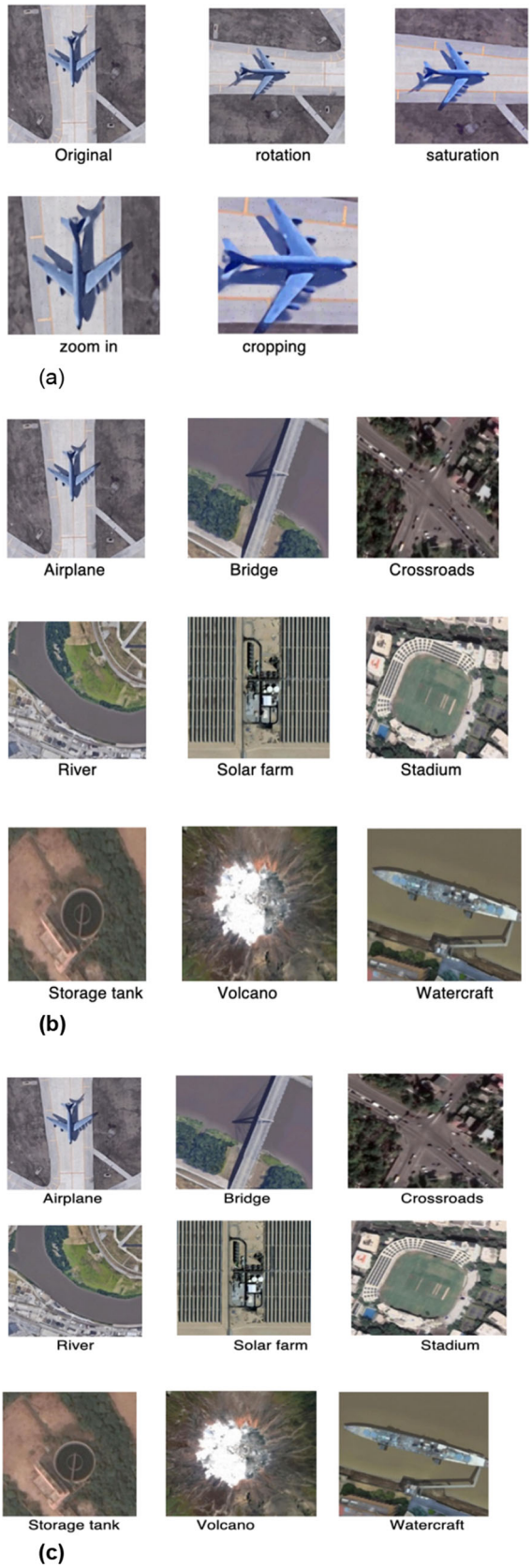

Fig. 5 **a** Use of different augmentation operations. **b** Classification of the augmented image sets. **c** Use of the augmentation for different application sets

Table 4 Accuracy obtained during the classification process

TVTI	A_c ISST [40]	A_c GAN [41]	A_c UNet [17]	A_c HFR AS
333	78.53	77.53	80.47	91.45
667	78.74	77.78	80.70	91.71
1000	78.95	78.04	80.92	91.96
1667	79.17	78.29	81.14	92.21
3333	79.39	78.54	81.36	92.46
8333	79.60	78.80	81.58	92.71
16,667	79.81	79.05	81.80	92.96
25,000	80.03	79.31	82.02	93.21
33,333	80.24	79.56	82.24	93.46
41,667	80.46	79.82	82.47	93.71
50,000	80.68	80.07	82.69	93.96
66,667	80.89	80.32	82.91	94.22
83,333	81.11	80.58	83.14	94.47
100,000	81.32	80.84	83.36	94.72

Table 5 Precision obtained during the classification process

TVTI	P_c ISST [40]	P_c GAN [41]	P_c UNet [17]	P_c HFR AS
333	75.78	76.37	78.42	88.02
667	75.98	76.62	78.64	88.26
1000	76.19	76.87	78.86	88.50
1667	76.40	77.11	79.07	88.74
3333	76.60	77.36	79.28	88.98
8333	76.81	77.60	79.48	89.21
16,667	77.02	77.86	79.70	89.46
25,000	77.23	78.11	79.91	89.70
33,333	77.44	78.35	80.12	89.94
41,667	77.65	78.60	80.34	90.18
50,000	77.86	78.85	80.55	90.42
66,667	78.07	79.10	80.77	90.66
83,333	78.28	79.35	80.98	90.90
100,000	78.49	79.59	81.19	91.14

Considering this evaluation and its visualization in Fig. 7, it can be seen that the proposed model can increase classification precision by 16.1% compared to ISST [40], 14.5% compared to GAN [41], and 12.2% compared to UNet [17], making it highly beneficial for a range of real-time classification applications. The reason for this precision enhancement is using LSTM & GRU during feature extraction, which assists in estimating high-efficiency augmented feature sets. Table 6 shows the recall levels as follows:

Considering this evaluation and its visualization in Fig. 8, it can be seen that the proposed model can increase classification recall by 38% compared to ISST [40], 34.9% compared to GAN [41], and 28.1% compared to UNet [17], making it highly beneficial for a range of real-time classification applications. This recall enhancement is due to the use of Iterative GA & LSTM with GRU during feature extraction, which assists in estimating high-efficiency augmented feature sets. These improvements allow the proposed model to identify classes in satellite images with high accuracy, precision, and recall. As a result, it applies to a wide range of real-time use cases.

Table 6 Recall obtained during the classification process

TVTI	R_c ISST [40]	R_c GAN [41]	R_c UNet [17]	R_c HFR AS
333	65.96	67.07	71.14	91.09
667	66.15	67.29	71.35	91.36
1000	66.34	67.51	71.54	91.61
1667	66.53	67.73	71.74	91.87
3333	66.72	67.95	71.93	92.12
8333	66.90	68.17	72.12	92.37
16,667	67.08	68.39	72.31	92.62
25,000	67.27	68.62	72.51	92.87
33,333	67.46	68.84	72.71	93.13
41,667	67.64	69.07	72.91	93.38
50,000	67.83	69.29	73.10	93.64
66,667	68.02	69.51	73.30	93.90
83,333	68.22	69.74	73.50	94.15
100,000	68.40	69.96	73.69	94.40

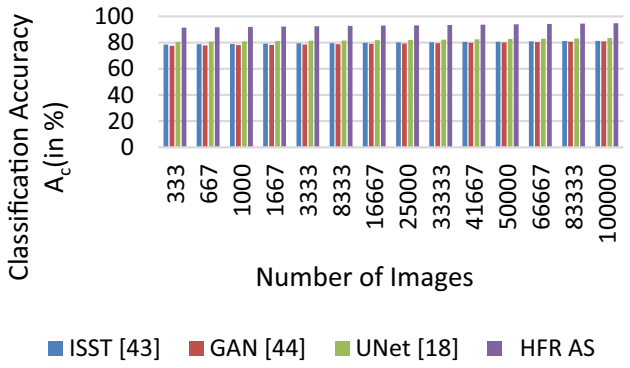

Fig. 6 Accuracy obtained during the classification process

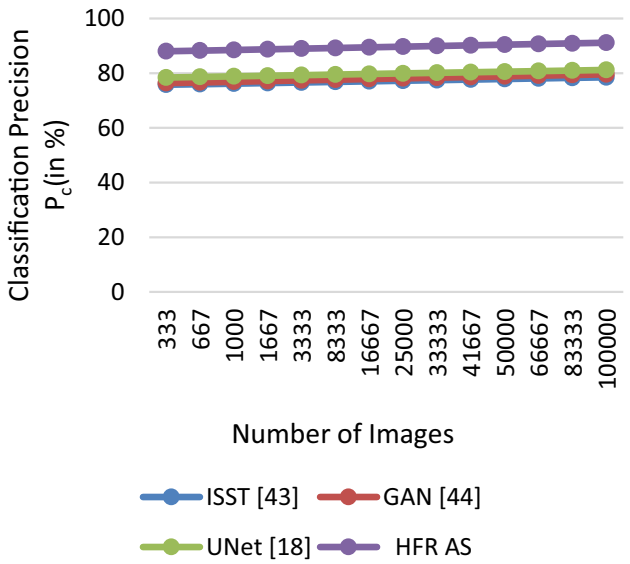

Fig. 7 Precision obtained during the classification process

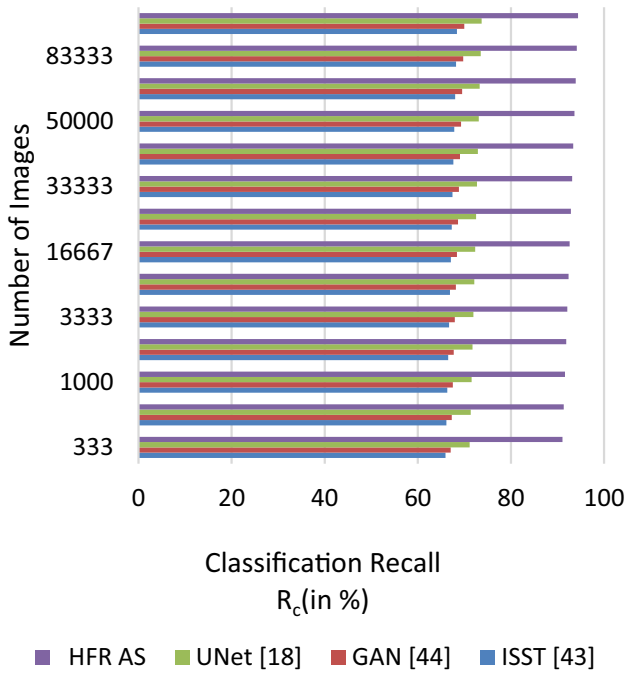

Fig. 8 Recall obtained during the classification process

5 Conclusion

According to our research, data augmentation is a significant method for preventing a model from becoming overly proficient and reducing the cost of labelling and cleansing the raw dataset. This study proposed a new model for improving the augmentation of satellite images that uses LSTM-based feature extraction with GRU-based feature extraction. First, this study combines LSTM-based feature extraction with GRU-based feature extraction, representing input images

as multimodal feature sets. The collected features train an efficient iterative genetic algorithm (IGA) that helps find high-density augmentation procedures and feature constants. These methods can improve the proposed model's accuracy, precision, and recall for various satellite image classification tasks. According to an evaluation of its accuracy, the suggested model may improve classification accuracy by 16.4% compared to ISST, 17.1% compared to GAN, and 13.6% compared to UNet, making it very beneficial for a range of real-time classification applications. The suggested model is helpful for a range of real-time classification scenarios since it may increase classification precision by 16.1% compared to ISST, 14.5% to GAN, and 12.2% to UNet. This improvement in accuracy is due to the use of LSTM and GRU during feature extraction, which helps estimate high-efficiency augmented feature sets. Estimates of recall levels show that the suggested model can improve classification recall by 38% compared to ISST, 34.9% to GAN, and 28.1% to UNet, making it very effective for a range of real-time classification applications.

As a future improvement, low-complexity and high-density feature extraction methods can be used together to improve the model. We can improve classification results using hybrid bioinspired models, autoencoders, Q-learning, or other deep learning methods.

Authors' contributions DS: Conceptualization; Methodology; Software; Visualization; Writing – original draft RG: Formal analysis; Resources; Writing – review & editing. RM: Data curation; Validation. DP: Investigation; Data curation; Supervision; Validation. MF: Data curation; Investigation; Software; Validation; Writing – review & editing.

Funding Open Access funding provided by University of Vaasa. The work of Muhammad Faheem was supported in part by the University of Vaasa, and in part by the Academy of Finland.

Data availability Data will be available on request.

Declarations

Competing interest The authors declare no competing interests.

References

1. Takahashi, R., Matsubara, T., Uehara, K.: Data augmentation using random image cropping and patching for deep CNNs. IEEE Trans. Circuits Syst. Video Technol. **30**(9), 2917–2931 (2020)
2. Shen, Y., Zhu, S., Yang, T., Chen, C., Pan, D., Chen, J., Xiao, L., Du, Q.: BDANet: multiscale convolutional neural network with cross-directional attention for building damage assessment from satellite images. IEEE Trans. Geosci. Remote Sens. **60**, 1–14 (2022)
3. Zhang, R., Lu, W., Wei, X., Zhu, J., Jiang, H., Liu, Z., Gao, J., Li, X., Yu, J., Yu, M., Yu, R.: A progressive generative adversarial method for structurally inadequate medical image data augmentation. IEEE J. Biomed. Health Informatics **26**(1), 7 (2022)
4. Ham, H.S., Lee, H.S., Chae, J.W., Cho, H.C., Cho, H.C.: Improvement of gastroscopy classification performance through image augmentation using a gradient-weighted class activation map. IEEE Access **10**, 99361–99369 (2022)
5. Li, Z., Zheng, C., Shu, H., Wu, S.: Dual-scale single image dehazing via neural augmentation. IEEE Trans. Image Process. **31**, 6213–6223 (2022)
6. Xiao, Q., Liu, B., Li, Z., Ni, W., Yang, Z., Li, L.: Progressive data augmentation method for remote sensing ship image classification based on imaging simulation system and neural style transfer. IEEE J. Selected Topics Appl. Earth Observat. Remote Sens. **14**, 9176–9186 (2021)
7. Yu, X., Wu, X., Luo, C., Ren, P.: Deep learning in remote sensing scene classification: a data augmentation enhanced convolutional neural network framework. GISci. Remote Sens. **54**(5), 741–758 (2017)
8. Xie, J., Yu, F., Wang, H., Zheng, H.: Class activation map-based data augmentation for satellite smoke scene detection. IEEE Geosci. Remote Sens. Lett. **19**, 1–5 (2022)
9. Hoang, P.M., Tuan, H.D., Son, T.T., Poor, H.V.: Qualitative HD image and video recovery via high-order tensor augmentation and completion. IEEE J. Selected Topics Signal Process. **15**(3), 688–701 (2021)
10. Huang, C., Zhao, J., Yu, Y., Zhang, H.: Comprehensive Sample augmentation by fully considering SSS imaging mechanism and environment for shipwreck detection under zero real samples. IEEE Trans. Geosci. Remote Sens. **60**, 1–14 (2022)
11. Du, S., Hong, J., Wang, Y., Xing, K., Qiu, T.: Physical-related feature extraction from simulated SAR image based on the adversarial encoding network for data augmentation. IEEE Geosci. Remote Sens. Lett. **19**, 1–5 (2022)
12. Abady, L., Cannas, E.D., Bestagini, P., Tondi, B., Tubaro, S., Barni, M.: An overview on the generation and detection of synthetic and manipulated satellite images. APSIPA Trans. Signal Inform. Process. **11**(1), 124 (2022)
13. Hao, X., Liu, L., Yang, R., Yin, L., Zhang, L., Li, X.: A review of data augmentation methods of remote sensing image target recognition. Remote Sens. **15**(3), 827 (2023)
14. Tiago, C., Gilbert, A., Beela, A.S., Aase, S.A., Snare, S.R., Sprem, J., McLeod, K.: A Data augmentation pipeline to generate synthetic labeled datasets of 3D echocardiography images using a GAN. IEEE Access **10**, 98803–98815 (2022)
15. Anaam, A., Bu-Omer, H.M., Gofuku, A.: Studying the applicability of generative adversarial networks on HEp-2 cell image augmentation. IEEE Access **9**, 98048–98059 (2021)
16. Kanwal, N., Perez-Bueno, F., Schmidt, A., Engan, K., Molina, R.: The devil is in the details: whole slide image acquisition and processing for artifacts detection, color variation, and data augmentation: a review. IEEE Access **10**, 58821–58844 (2022)
17. Zhang, J., Xing, M., Sun, G.C., Shi, X.: Vehicle trace detection in two-pass SAR coherent change detection images with spatial

feature enhanced UNET and adaptive augmentation. IEEE Trans. Geosci. Remote Sens. **60**, 1–15 (2022)
18. Hua, C.H., Kim, K., Huynh-The, T., You, J.I., Yu, S.Y., Le-Tien, T., Bae, S.H., Lee, S.: Convolutional network with twofold feature augmentation for diabetic retinopathy recognition from multi-modal images. IEEE J. Biomed. Health Inform. **25**(7), 2686–2697 (2021)
19. Miao, X., Zhang, Y., Zhang, J., Liang, X.: Hierarchical CNN classification of hyperspectral images based on 3-D attention soft augmentation. IEEE J. Selected Topics Appl. Earth Observat. Remote Sens. **15**, 4217–4233 (2022)
20. Qin, K., Ge, F., Zhao, Y., Zhu, L., Li, M., Shi, C., Li, D., Zhou, X.: Hapke data augmentation for deep learning-based hyperspectral data analysis with limited samples. IEEE Geosci. Remote Sens. Lett. **18**(5), 886–890 (2021)
21. Kim, Y., Uddin, A.F.M.S., Bae, S.H.: Local augment: utilizing local bias property of convolutional neural networks for data augmentation. IEEE Access **9**, 15191–15199 (2021)
22. Fan, R., Wang, H., Cai, P., Wu, J., Bocus, M.J., Qiao, L., Liu, M.: Learning collision-free space detection from stereo images: homography matrix brings better data augmentation. IEEE/ASME Trans. Mechatron. **27**(1), 225–233 (2022)
23. Hossain, M.T., Teng, S.W., Sohel, F., Lu, G.: Robust image classification using a low-pass activation function and DCT augmentation. IEEE Access **9**, 86460–86474 (2021)
24. Alzubaidi, L., Zhang, J., Humaidi, A.J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M.A., Al-Amidie, M., Farhan, L.: Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J. Big Data **8**(1), 1–74 (2021)
25. Monasterio-Exposito, L., Pizarro, D., Macias-Guarasa, J.: Label augmentation to improve generalization of deep learning semantic segmentation of laparoscopic images. IEEE Access **10**, 37345–37359 (2022)
26. Adedeji, O.: Image augmentation for Satellite Images. arXiv.org (2022)
27. Chen, H., Li, W., Shi, Z.: Adversarial instance augmentation for building change detection in remote sensing images. IEEE Trans. Geosci. Remote Sens. **60**, 1–16 (2022)
28. Nesteruk, S., Illarionova, S., Akhtyamov, T., Shadrin, D., Somov, A., Pukalchik, M., Oseledets, I.: Xtreme augment: getting more from your data through combination of image collection and image augmentation. IEEE Access **10**, 24010–24028 (2022)
29. Wang, W., Chen, Y., He, X., Li, Z.: Soft augmentation-based siamese CNN for hyperspectral image classification with limited training samples. IEEE Geosci. Remote Sens. Lett. **19**, 1–5 (2022)
30. Chen, L., Wei, Y., Yao, Z., Chen, E., Zhang, X.: Data augmentation in prototypical networks for forest tree species classification using airborne hyperspectral images. IEEE Trans. Geosci. Remote Sens. **60**, 1–16 (2022)
31. Kim, J.Y., Ha, J.E.: Spatio-temporal data augmentation for visual surveillance. IEEE Access **9**, 165014–165033 (2021)
32. Xia, M., Wang, Z., Han, F., Kang, Y.: Enhanced multi-dimensional and multi-grained cascade forest for cloud/snow recognition using multispectral satellite remote sensing imagery. IEEE Access **9**, 131072–131086 (2021)
33. Yamashita, R., Long, J., Banda, S., Shen, J., Rubin, D.L.: Learning domain-agnostic visual representation for computational pathology using medically-irrelevant style transfer augmentation. IEEE Trans. Med. Imaging **40**(12), 3945–3954 (2021)
34. Zhu, Y., Zhang, Y., Zhang, H., Yang, J., Zhao, Z.: Data augmentation of x-ray images in baggage inspection based on generative adversarial networks. IEEE Access **8**, 86536–86544 (2020)
35. Pan, X., Tang, F., Dong, W., Gu, Y., Song, Z., Meng, Y., Xu, P., Deussen, O., Xu, C.: Self-supervised feature augmentation for large image object detection. IEEE Trans. Image Process. **29**, 6745–6758 (2020)
36. Zhang, L., Wang, X., Yang, D., Sanford, T., Harmon, S., Turkbey, B., Wood, B.J., Roth, H., Myronenko, A., Xu, D., Xu, Z.: Generalizing deep learning for medical image segmentation to unseen domains via deep stacked transformation. IEEE Trans. Med. Imaging **39**(7), 2531–2540 (2020)
37. Mahdizadehaghdam, S. Krim, H.: Sparse Generative Adversarial Network. arXiv.org (2019)
38. Parsaeimehr, E., Fartash, M., Akbari Torkestani, J.: Improving feature extraction using a hybrid of CNN and LSTM for entity identification. Neural Process. Lett. **55**(5), 5979–5994 (2023)
39. Shihab, M.S.H., Aditya, S., Setu, J.H., Imtiaz-Ud-Din, K.M., Efat, M. I.A.: A hybrid GRU-CNN feature extraction technique for speaker identification. In 23rd International Conference on Computer and Information Technology (ICCIT) (2020)
40. Shang, X., Han, S., Song, M.: Iterative spatial-spectral training sample augmentation for effective hyperspectral image classification. IEEE Geosci. Remote Sens. Lett. **19**, 1–5 (2022)
41. Abady, L., Horváth, J., Tondi, B., Delp, E.J., Barni, M.: Manipulation and generation of synthetic satellite images using deep learning models. J. Appl. Remote Sens. **16**(04), 046504 (2022)
42. Perez-Hernandez, F., Rodriguez-Ortega, J., Benhammou, Y., Herrera, F., Tabik, S.: CI-dataset and DetDSCI methodology for detecting too small and too large critical infrastructures in satellite images: airports and electrical substations as case study. IEEE J. Selected Topics Appl. Earth Observat. Remote Sens. **14**, 12149–12162 (2021)
43. Luo, X., Li, X., Wu, Y., Hou, W., Wang, M., Jin, Y., Xu, W.: Research on change detection method of high-resolution remote sensing images based on subpixel convolution. IEEE J. Selected Topics Appl. Earth Observat. Remote Sens. **14**, 1447–1457 (2021)
44. Nalepa, J., Myller, M., Kawulok, M.: Training- and test-time data augmentation for hyperspectral image segmentation. IEEE Geosci. Remote Sens. Lett. **17**(2), 292–296 (2020)