Hyperspectral Image Reconstruction in Remote Sensing: LaplaceGAN Synthesis Coupled with VGG-UNet Classification

Shikha Sain [1,*], Dr. Monika Saxena [2]

[1] *Department of Computer Science, Banasthali Vidyapith, Jaipur ,Rajasthan 304022*

Abstract

Enhancing and reconstructing environmental images involve refining visual data to improve quality and reconstructing scenes. In remote sensing, this aids in accurate analysis, contributing to advanced understanding and decision-making. This study focuses on advancing hyperspectral image analysis in remote sensing through the design of a deep learning-based model aimed at enhancing and reconstructing environmental images. An integral aspect involves introducing a novel approach using LaplaceGAN to generate synthetic images with high fidelity, building upon real images as a foundational basis. Furthermore, the study proposes the implementation of a specialized VGG-UNet architecture tailored for the classification of hyperspectral images, specifically addressing the nuances of remote sensing data. To assess the model's efficacy, a comparative analysis is conducted, pitting the performance of VGG-UNet against alternative methods such as Res-UNet and Faster R-CNN in the context of remote sensing image classification. This research aims to contribute to the field by designing a deep learning model that not only analyzes hyperspectral images comprehensively but also enhances and reconstructs environmental images, thereby advancing the most recent methods for better comprehension and judgement in a range of remote sensing applications.

Keywords

Hyperspectral images, Remote sensing, LaplaceGAN, VGG-UNet [1]

Symposium on Computing & Intelligent Systems (SCI), *May 10, 2024, New Delhi, INDIA*

* Corresponding author.
† These authors contributed equally.
id4shikha93@gmail.com (Shikha Sain); Muskan.saxena@gmail.com (Dr. Monika Saxena)

1. Introduction

A new era in image processing has been brought about by the development of deep learning. In the realm of environmental science, where the acquisition of high-quality imagery is crucial for accurate analysis and decision-making, Deep learning model applications are now the main focus of research and development. Environmental images, ranging from satellite captures of Earth's surface to ground-level snapshots of ecosystems, are often beset with challenges that compromise their utility[1]. Factors such as atmospheric interference, low resolution, and noise frequently degrade the visual quality of these images, impeding their efficacy in various applications such as environmental monitoring, land cover classification, and climate change analysis.

Neural networks with numerous layers are used in deep learning, a type of machine learning, to automatically learn and extract characteristics from incoming data. Convolutional neural networks (CNNs) are a particular kind of neural networks intended for tasks involving images. Have emerged as powerful tools for deciphering complex patterns within environmental images. Enhanced images contribute to more accurate land cover classification, facilitate object detection, and improve anomaly detection in environmental monitoring [2]. In applications such as disaster management, climate monitoring, and ecological preservation, the availability of high-quality images is instrumental in informed decision-making.

However, deploying deep learning models for environmental image enhancement is not without its challenges. Limited availability of diverse and comprehensive datasets, interpretability of deep learning models, and computational demands pose significant hurdles [3]. This research paper seeks to delve into the transformative deep learning's promise for improving and rebuilding environmental imagery. By examining the most recent developments, difficulties, and possible uses, this study seeks to add to the growing conversation on how deep learning may influence environmental image processing in the future. Researchers and scientists are now investigating the use of environmental remote sensing, also known as imaging spectroscopy, to find and identify minerals, terrestrial plants, artefacts, and backdrops, we collect hyper spectral data.

The following are the research's primary goals:

1. To provide deep learning frameworks for the thorough examination and categorization of hyperspectral photos acquired by remote sensing.
2. To introduce an innovative approach based on LaplaceGAN for generating synthetic images with high fidelity, utilizing real images as a basis.
3. To propose the implementation of a VGG-UNet architecture specifically created specifically for the purpose of classifying hyperspectral photos obtained using remote sensing technology.
4. Conduct a comparative analysis, pitting the performance of VGG-UNet against alternative methods such as Res-UNet and Faster R-CNN, to assess their efficacy in the context of remote sensing image classification.

2. Literature Review

[4] investigated that deep learning has shown to be one of the most effective machine learning strategies for a range of inference problems in recent years. [5] investigated the impact of image motion artifacts on cardiac magnetic resonance (MR) segmentation and evaluated many methods for concurrently correcting aberration and segmenting heart chambers. [6] reviewed the state of the art in terms of research on using deep learning to several picture fusion scenarios, including multi-modal, sharpening, and image fusion from digital photography.[7]. examined in order to provide a comprehensive summary of recent advancements in deep learning-based image super-resolution.

[8] investigated the current approaches to data augmentation, encouraging advancements, and meta-level choices for data augmentation implementation will all be covered in this study. [9] outlined a deep learning architecture of U-net-based mapping paradigm for urban villages. The study area is located in Guangzhou City, China. For investigation, a building boundary vector file and an eight pan-sharpened band, 0.5-meter spatial resolution Worldview satellite picture were employed. [10] proposed the use of deep learning based on augmented convolutional neural networks (CNNs) for the real-time detection of apple leaf disease. The apple leaf disease dataset (ALDD) is first created in this study by using techniques for picture interpretation and data augmentation. The dataset consists of complex photos captured in actual field situations as well as images captured in laboratories.

Reference	**Technology**	**Major Inclusion**	**Parameters**
[11]	CNN	Utilizing Partial Convolutions for Image In painting for Irregular Holes	Spectral Dimension
[12]	Decision Tree Classifier	Using data from Landsat-7 ETM+, a decision tree classifier may be used to categorize land uses in an agricultural region	Classifier Six Land Use Classes
[13]	Target-Adapted CNN	Utilizing time series segmentation to detect shifts in vegetation patterns	DBEST
[14]	GIS Tool	GIS, a possible tool for enabling the creation and use of thematic data to assess the potentiality of groundwater	Thematic Info

[15]	Image Processing	Image fusion can be attempted in the frequency domain	Frequency Domain

3. Methodology

3.1. Dataset Description

In this study, we collected hyper spectral images of remote sensing from below website. https://www.kaggle.com/datasets/sciencelabwork/hyperspectral-image-sensing dataset-ground-truth. The process of hyperspectral remote sensing involves the collection and examination of a broad spectrum of electromagnetic wavelengths.

3.2. Data Pre-processing

Data preprocessing involves the application of techniques to handle various aspects of the data:

- **Data Cleaning:** The dataset is inspected for artifacts, noise, and inconsistencies. The goal of this procedure is to find and fix any problems that might compromise the accuracy and dependability of the data. Common data cleaning techniques include removing outliers, addressing missing values, and handling corrupted data points.
- **Data Normalization**: Normalization is performed to ensure that all features in the dataset have a consistent scale. This step is crucial for training machine learning models because it stops certain characteristics from controlling the learning process because of variances in their strength.

3.3. LaplaceGAN

The LaplaceGAN (LapGAN) serve a crucial purpose in environmental image enhancement and reconstruction by leveraging their generative capabilities to enhance the photos' clarity, authenticity, and interpretive power. As shown in Fig. 1, the generator and discriminator are the two main parts of the Laplace Generative Adversarial Network (LapGAN).

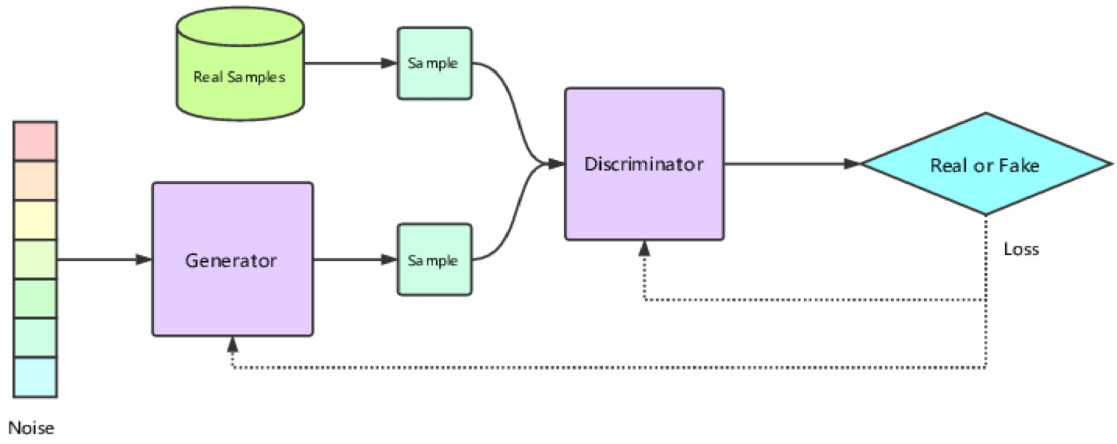

Figure 1 The structure of LaplaceGAN

Within the LapGAN framework, the generator is designed to intake both noise and labels as input. In this setup, the discriminator not only provides the probability of authenticity (real or fake) but also specifies the category of the input sample. The intricacies of the model are elucidated as follows:

Generator

In the Laplace GAN, the generator processes noise vectors and load pattern labels encoded in one-hot format to produce synthetic load profiles. The architecture involves a fully-connected layer for initial mapping, followed by two transposed convolution layers with a 2 × 2 stride and a 4 × 4 kernel for up-sampling. To stabilize training and ensure smooth gradient flow, batch normalization layers are applied after each computational step. The final transposed convolution layer, crucial for shaping generated loads within the same interval as normalized real loads, uses a Sigmoid activation function.

Discriminator

The discriminator is used to assess the authenticity and quality of the model-generated pictures of synthetic environments. Its main objective is to differentiate between produced and actual pictures. Ensuring the produced images exhibit biologically plausible structures. Engaging in adversarial training with the generator, the discriminator continually adapts to maintain its ability to differentiate between real and synthetic data. By providing feedback to the generator, the discriminator guides the improvement of synthetic images, fostering a dynamic interplay that refines both components. Ultimately, the discriminator acts as a critical quality control mechanism, contributing to the production of high-quality, biologically relevant environmental images through the LapGAN training process.

The detailed results of the conducted

3.4. Build Model

VGG-UNET

A deep learning architecture called the Proposed VGG-UNet combines the advantages of the VGG16 and U-Net models [16] While the U-Net model is an architecture ideally suited for image segmentation, the VGG16 model is a potent feature extractor as shown in figure.

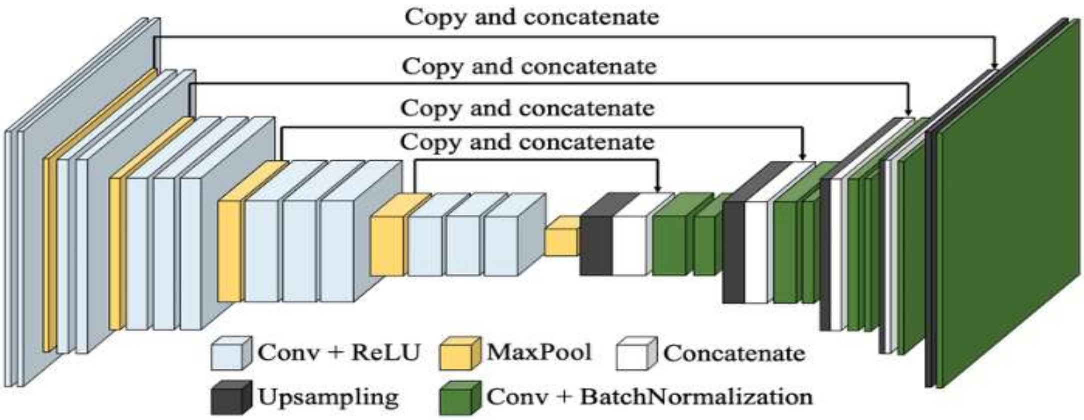

Figure 2 Architecture of VGG-UNET

The encoder and the decoder are the two primary components of the VGG-UNet architecture. While the decoder creates the segmentation mask, the task of extracting features from the input image falls to the encoder. The VGG16 model serves as the basis for the encoder in the proposed VGG-UNet. The max pooling layer comes after each of the 16 convolutional layers that make up the VGG16 model. Five blocks with varying numbers of filters each comprise the convolutional layers. Every block has filters that are 3 by 3 in size. The U-Net model serves as the foundation for the proposed VGG-UNet decoder.

3.5. Test-Train Split

The training set and the test set are the two primary subsets of the dataset. The test set is used to evaluate how effectively the trained model generalizes to fresh, untested data, while the training set is used to train the model. A validation set is often constructed in order to avoid overfitting and fine-tune the model during training.

Train the network:

Using the train data, the proposed VGG-UNet model and Res-UNet, Faster R-CNN are trained. The performance of the proposed model is assessed and contrasted with these two additional Res-UNet, Faster R-CNN architectures.

3.6. Performance Metrics

Subsequently, the training set is employed for the purpose of training various models, including VGG-UNet model and Res-UNet, Faster R-CNN. The efficacy of the proposed algorithm will be evaluated using the performance measures listed below. The accuracy, sensitivity, precision, and F1-score of the confusion matrix are used to evaluate a technique's efficacy.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (2)$$

$$\Pr ecision = \frac{TP}{TP + FP} \quad (3)$$

$$F1 - score = 2 * \frac{\Pr ecisison * \operatorname{Re} call}{\Pr ecision + \operatorname{Re} call} \quad (4)$$

4. RESULTS

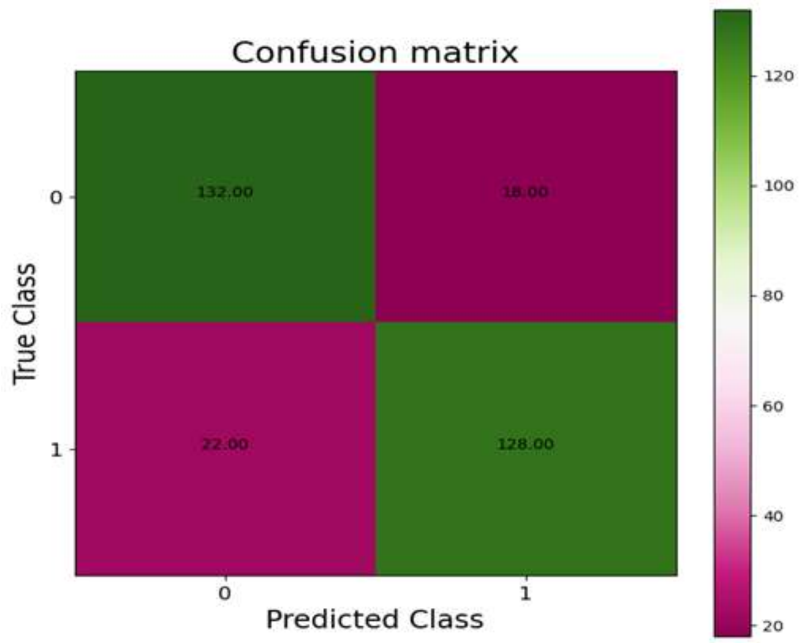

Figure 3 VGG-UNet

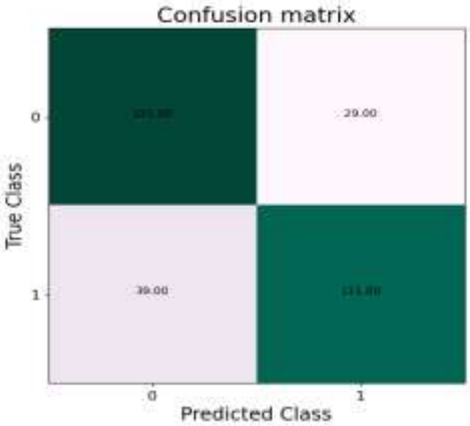

Figure 4 Res-UNet

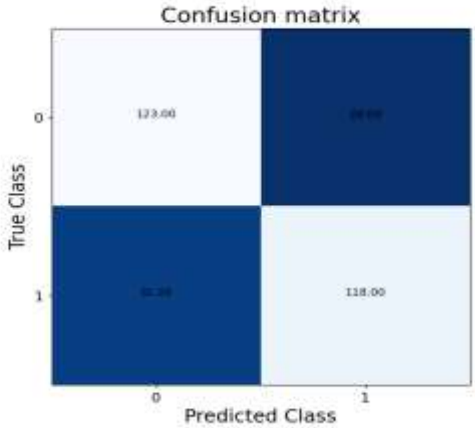

Figure 5 Faster R-CNN

The confusion matrix displays how well the suggested VGG-UNet classifier in contrast to established models like Res-UNet and Faster R-CNN for hyperspectral image classification in remote sensing. The proposed model demonstrated proficiency in identifying 128 hyperspectral and 112 non-hyperspectral images. However, it faced challenges in accurately classifying non-hyperspectral images, resulting in 18 false positives and a 53% recall rate for that class. Conversely, it exhibited an 85% recall for true hyperspectral images but misclassified 22 as non-hyperspectral. Notably, the focus should be on refining the model to minimize misclassifications of non-hyperspectral images. Despite these challenges, the proposed architecture outperformed existing models by achieving higher correct classifications and fewer misclassifications, showcasing its potential for improved hyperspectral image classification in remote sensing applications.

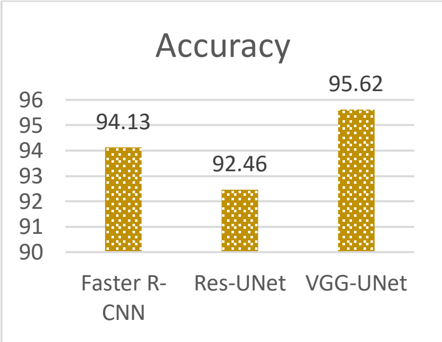

Figure 6 comparison of accuracy

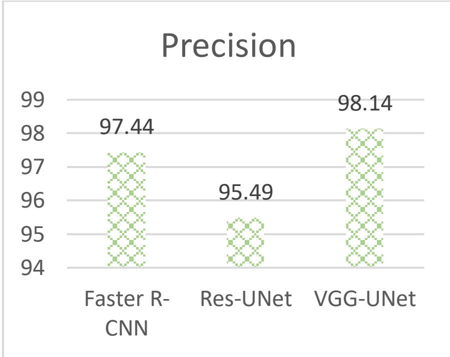

Figure 7 comparison of precision

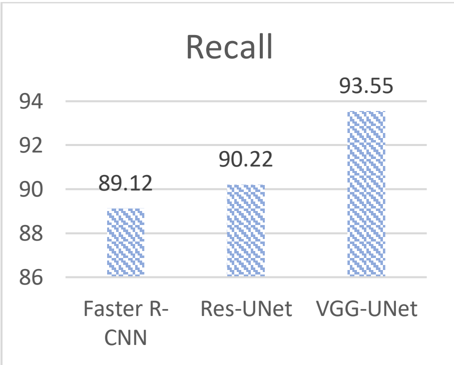

Figure 8 comparison of recall

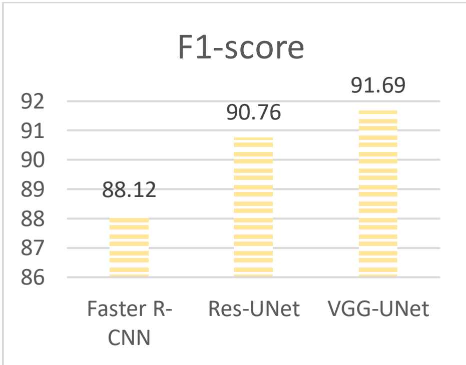

Figure 9 comparison of F1-score

The graphical comparisons highlight the superior performance of the proposed VGG-UNet classifier against established models (Res-UNet and Faster R-CNN) in hyperspectral image classification for remote sensing. The VGG-UNet consistently outperforms with the highest accuracy (95.62%), precision (98.14%), recall (93.55%), and F1-score (91.69%). These findings highlight the VGG-UNet model's remarkable ability to categorizes hyperspectral pictures reliably, highlighting its usefulness in remote sensing applications. The thorough assessment based on several metrics confirms that the VGG-UNet is a solid and trustworthy option for developing hyperspectral image analysis in the remote sensing domain.

5. Conclusion.

Finally, the assessment of the suggested VGG-UNet classifier against established models, Res-UNet and Faster R-CNN, for hyperspectral image classification in remote sensing has yielded valuable insights. The study proposes the implementation of a specialized VGG-UNet architecture tailored for the classification of hyperspectral images, specifically addressing the nuances of remote sensing data. To assess the model's efficacy, a comparative analysis is conducted, pitting the performance of VGG-UNet against alternative methods such as Res-UNet and Faster R-CNN in the context of remote sensing image classification. Complementing these findings, the F1-score, recall, accuracy, and precision comparisons depicted in the graphs consistently showcased the superior performance of the VGG-UNet model. With a notably high accuracy of 95.62%, precision at 98.14%, recall at 93.55%, with a remarkable 91.69% F1-score, the VGG-UNet outperformed Res-UNet and Faster R-CNN. These metrics collectively underscore the VGG-UNet's efficacy in hyperspectral image classification, reinforcing its potential for enhanced remote sensing applications. Despite facing challenges, the proposed architecture demonstrated its superiority through higher correct classifications and fewer misclassifications, emphasizing its promise for improved hyperspectral image analysis. With a foundation for model refinement and breakthroughs in accurate hyperspectral image classification for environmental monitoring and decision-making processes. the studies thorough evaluation and comparative analyses advance deep learning techniques in remote sensing.

References

[1] Z. Guan, X. Miao, Y. Mu, Q. Sun, Q. Ye, and D. Gao, "Forest Fire Segmentation from Aerial Imagery Data Using an Improved Instance Segmentation Model," *Remote Sens.*, vol. 14, no. 13, 2022, doi: 10.3390/rs14133159.

[2] X. Han, L. Xue, F. Shao, and Y. Xu, "A power spectrum maps estimation algorithm based on generative adversarial networks for underlay cognitive radio networks," *Sensors (Switzerland)*, vol. 20, no. 1, 2020, doi: 10.3390/s20010311.

[3] N. Sharma and M. Hefeeda, "Hyperspectral reconstruction from RGB images for vein visualization," *MMSys 2020 - Proc. 2020 Multimed. Syst. Conf.*, pp. 77–87, 2020, doi: 10.1145/3339825.3391861.

[4] K. De Haan, Y. Rivenson, Y. Wu, and A. Ozcan, "Deep-Learning-Based Image Reconstruction and Enhancement in Optical Microscopy," *Proc. IEEE*, vol. 108, no. 1, pp. 30–50, 2020, doi: 10.1109/JPROC.2019.2949575.

[5] I. Oksuz *et al.*, "Deep Learning-Based Detection and Correction of Cardiac MR Motion Artefacts during Reconstruction for High-Quality Segmentation," *IEEE Trans. Med. Imaging*, vol. 39, no. 12, pp. 4001–4010, 2020, doi: 10.1109/TMI.2020.3008930.

[6] H. Zhang, H. Xu, X. Tian, J. Jiang, and J. Ma, "Image fusion meets deep learning: A survey and perspective," *Inf. Fusion*, vol. 76, no. August, pp. 323–336, 2021, doi: 10.1016/j.inffus.2021.06.008.

[7] Z. Wang, J. Chen, and S. C. H. Hoi, "Deep Learning for Image Super-Resolution: A Survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 10, pp. 3365–3387, 2021,

doi: 10.1109/TPAMI.2020.2982166.
[8] C. Shorten and T. M. Khoshgoftaar, “A survey on Image Data Augmentation for Deep Learning,” *J. Big Data*, vol. 6, no. 1, 2019, doi: 10.1186/s40537-019-0197-0.
[9] Z. Pan, J. Xu, Y. Guo, Y. Hu, and G. Wang, “Deep learning segmentation and classification for urban village using a worldview satellite image based on U-net,” *Remote Sens.*, vol. 12, no. 10, pp. 1–17, 2020, doi: 10.3390/rs12101574.
[10] P. Jiang, Y. Chen, B. Liu, D. He, and C. Liang, “Real-Time Detection of Apple Leaf Diseases Using Deep Learning Approach Based on Improved Convolutional Neural Networks,” *IEEE Access*, vol. 7, pp. 59069–59080, 2019, doi: 10.1109/ACCESS.2019.2914929.
[11] S. Nie, L. Gu, Y. Zheng, A. Lam, N. Ono, and I. Sato, “Deeply Learned Filter Response Functions for Hyperspectral Reconstruction,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 4767–4776, 2018, doi: 10.1109/CVPR.2018.00501.
[12] M. Q. Huang, J. Ninić, and Q. B. Zhang, “BIM, machine learning and computer vision techniques in underground construction: Current status and future perspectives,” *Tunn. Undergr. Sp. Technol.*, vol. 108, pp. 1–78, 2021, doi: 10.1016/j.tust.2020.103677.
[13] J. Zhao *et al.*, “Deep learning in hyperspectral image reconstruction from single rgb images—a case study on tomato quality parameters,” *Remote Sens.*, vol. 12, no. 19, pp. 1–14, 2020, doi: 10.3390/rs12193258.
[14] C. Engineering and S. Attila, “hyperspectral image classification using deep learning,” 2021.
[15] A. J. X. Guo and F. Zhu, “Improving deep hyperspectral image classification performance with spectral unmixing,” *Signal Processing*, vol. 183, pp. 1–24, 2021, doi: 10.1016/j.sigpro.2020.107949.
[16] S. Krishnamoorthy, Y. Zhang, S. Kadry, and W. Yu, “Framework to Segment and Evaluate Multiple Sclerosis Lesion in MRI Slices Using VGG-UNet,” *Comput. Intell. Neurosci.*, vol. 2022, 2022, doi: 10.1155/2022/4928096.