RGConvNET:A Hybrid Hyperspectral Image Classification Model for Extracting Deep Features with Improved Accuracy

Vikash Ranjan
Department of Electronics and Telecommunication
IIIT Bhubaneswar
Bhubaneswar, India
c121015@iiit-bh.ac.in

Pradyut Kumar Biswal
Department of Electronics and Telecommunication
IIIT Bhubaneswar
Bhubaneswar, India
pradyut@iiit-bh.ac.in

***Abstract*— Hyperspectral imaging (HSI) has an important role in applications like urban planning, agriculture, surveillance, and remote sensing, where it helps identify various regions and materials. Although convolutional neural networks (CNNs) has significant attainment in feature extraction and classification, they face limitations in handling the dimensionality of HSI data and cannot efficiently extract local features under geometric transformations. In order to resolve this, we suggest a hybrid HSI classification model combining Recursive Gated Convolution (RG-Convolution) and U-Net. Three RG-Convolution blocks are used to capture spatial features, while U-Net handles spectral information. A dense network with global average pooling then maps the features to different classes, reducing dimensionality. The model, tested on benchmark datasets like Pavia University and Indian Pines, achieves superior performance compared to existing methods, with an overall accuracy of 98.56% and 99.84% for Indian Pines and Pavia University dataset, respectively. This outperforms current state-of-the-art algorithms.**

Keywords—HSI, Recursive Gated Convolution, U-NET, Dimensionality, global average pooling

I. INTRODUCTION

Hyperspectral images (HSIs) are valuable in applications such as food safety, precision agriculture, mining, military operations, and environmental monitoring, due to their rich spectral and spatial data, as seen in Fig. 1. In HSI classification each pixel is assigned by a class label. Generally, for the classification purpose various machine learning (ML) and deep learning (DL) algorithm are used. Some of the common ML methods, like Regression [1], KNN [3], Decision tree [1], SVM [4], kernel-based approaches [5], and regression models [6], have been applied, but they rely on time-consuming manual feature extraction. Recently, DL algorithms have achieved notable success in automatically extracting features. Examples include CNN , 2D CNN, and 3D CNN [7-9]. However, CNN models have certain disadvantages, such as handling the local minima or struggling with gradient descent, hindering learning and generalization. Pooling layers may also lead to the important data loss. While 3-D CNNs capture spatial details, they are computationally complex, making them less efficient for certain applications. In order to resolve these limitations, we suggest a novel approach utilizing Recursive Gated Convolutional Networks and U-Net for feature extraction and classification in hyperspectral imaging (HSI), combining the advantages of both methods to enhance the classification accuracy. The proposed model for classification follows a structured workflow:

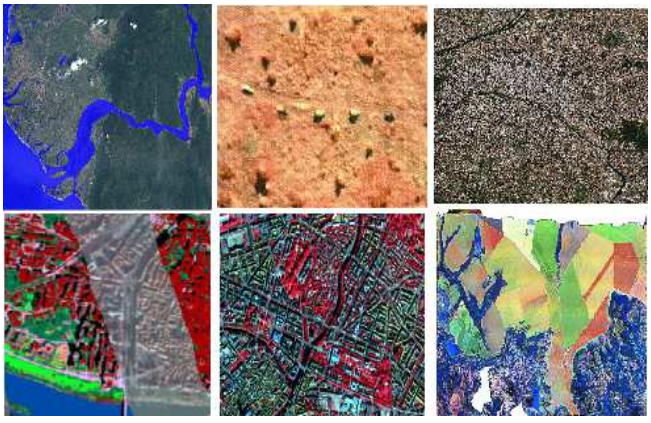

Fig. 1 Application of Hyperspectral Image a) Flood Monitoring b) Defense supervision or Tracking c) green space monitoring urban areas d) Spatial land use] e) Planning of the city f) Different varieties of corn shown by HSI .

• *Spatial feature abstraction using Recursive Gated Convolution*: Initially, a 3-layer Recursive Gated Convolution extracts spatial features from hyperspectral images, essential for accurate classification.

• *Spectral feature extraction using U-Net*: After spatial processing, the U-Net network is used to abstract spectral features. The final output from the third RG Convolution layer passes through U-Net's encoder and decoder, capturing key spectral properties for precise geospatial classification.

• *Classification Mapping*: Finally, 2D convolution and Global Average Pooling are applied, generating the final classification output for HSI mapping.

Our test findings uniquely demonstrate that the hybrid model which is suggested in this paper is more precise and efficient in classification of the hyper spectral images compared to other existing approaches. This demonstrates how the model may extract selective characteristics, which leads to a more precise classification.

We used the dataset provided by Pavia university and Indian pines for comparing the proposed model with the other existing ML and DL approaches in terms of various accuracy to demonstrate its superiority which in turn proved usefulness of the RGConvNet model for hyperspectral image classification.

This article is aligned as follows: Section 2 includes literature survey, 3rd section details the RGConvNET network architecture. 4th section discusses experimental setup, performance evaluation, data description, and result analysis. Section 5 highlights the originality of the approach and key findings. Finally, Section 6 includes the conclusion of the paper and future research scope.

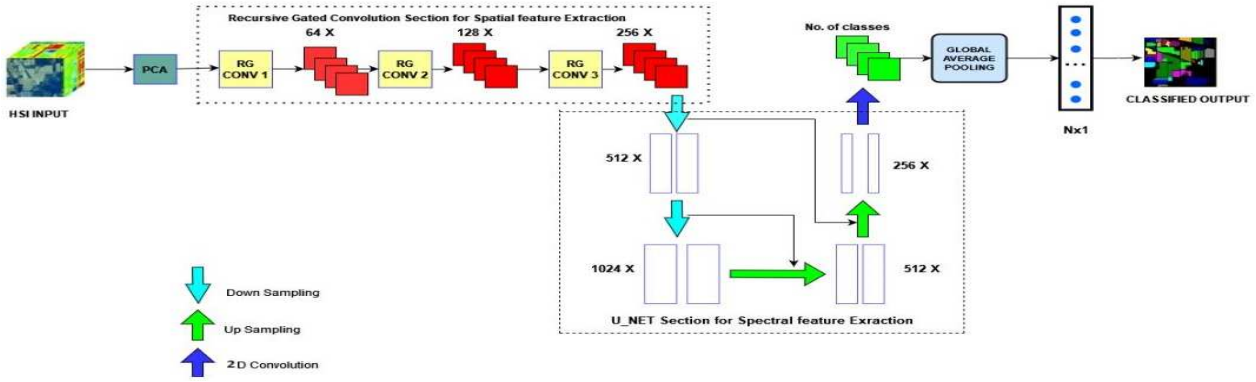

Fig. 2 Complete Architecture of the Proposed Model (RGConvNET)

II. Related Works

The deep learning methods [2][6] and neural translation encoder strategies like Gated Recurrent Units (GRU) [11] have significantly advanced HSI classification. This progress is supported by models such as HorNet [18], recurrent neural networks [11], 3D deep learning architectures [9], and cascaded recurrent neural networks , as well as deep learning combined with dimensionality reduction [19]. Architectures like UNET [12] have also been widely applied. However, a major limitation of CNNs is their inefficiency in processing curved and edge images, as they are primarily designed for square images. Innovations like AlexNet , which introduced multi-layer convolution and pooling with ReLU activation, and ResNet [18], which utilized skip connections for improved performance, have addressed some of these challenges. GoogleNet/Inception [26] further enhanced image classification with its 22-layer deep architecture. Song et al. [8] and Hang et al. [7] proposed models combining the concept of attention mechanism and multi-layer feature extraction for effective HSI classification.

III. RGConvNET Architecture

The architecture of the proposed model is shown in Fig. 2 and can be demonstrated by explaining the following section in detail:

A. Input Hperspectral data and PCA preprocessing:

Hyperspectral data $X \in \mathbb{R}^{H\times W\times D}$ consists of height H, width W, and spectral bands D, where D can be in the range of hundreds. A hyperspectral image is represented as a hypercube, where each pixel contains spectral information from narrow wavelength bands stored in 3D space. Along the x and y axes spatial data is distributed, while the spectral data is distributed along the z axis. The image is therefore represented in a 3D format (H, W, D). HSI image analysis starts with preprocessing the hypercube by Principal Component Analysis (PCA) [19] whose main function is to reduce the dimensionality to retain the most significant spectral information while reducing computational complexity. After performing PCA, the PCA transformed data will have 35 bands, which will have the major impact on the class. A fixed size sliding window of 5*5 is used for converting each pixel size to (5*5).

B. Recursive Gated Convolution (RGC) or g^{n}Convolution:

In order to achieve a high order spatial interaction, the recursive gated convolution is introduced. It is designed using standard convolutions, sigmoid functions, its linear projections, and element-wise multiplications. Recursive gated Convolution functions perform input-adaptive spatial mixing in a manner analogous to self-attention. The process begins with the PCA-reduced image $X_{PCA} \in \mathbb{R}^{H\times W\times 5}$ which is then passed through three layers of RGC units. Figure 3 illustrates one of these layer architectures, and the mathematical descriptions of all layers are provided below.

i. First Layer:

$$G^{(1)} = \sigma\left(w_g^{(1)} * X_{PCA} + b_g^{(1)}\right) \quad (1)$$

$$H_c^{(1)} = \left(w_c^{(1)} * X_{PCA} + b_c^{(1)}\right) \quad (2)$$

$$H^{(1)} = \left(G^{(1)} \odot H_c^{(1)}\right) \quad (3)$$

ii. Second Layer:

$$G^{(2)} = \sigma\left(w_g^{(2)} * H^{(1)} + b_g^{(2)}\right) \quad (4)$$

$$H_c^{(2)} = \left(w_c^{(2)} * H^{(1)} + b_c^{(2)}\right) \quad (5)$$

$$H^{(2)} = \left(G^{(2)} \odot H_c^{(2)}\right) \quad (6)$$

iii. Third Layer:

$$G^{(3)} = \sigma\left(w_g^{(3)} * H^{(2)} + b_g^{(3)}\right) \quad (7)$$

$$H_c^{(3)} = \left(w_c^{(3)} * H^{(2)} + b_c^{(3)}\right) \quad (8)$$

$$H^{(3)} = \left(G^{(3)} \odot H_c^{(3)}\right) \in \mathrm{R}^{H\times W\times 256} \quad (9)$$

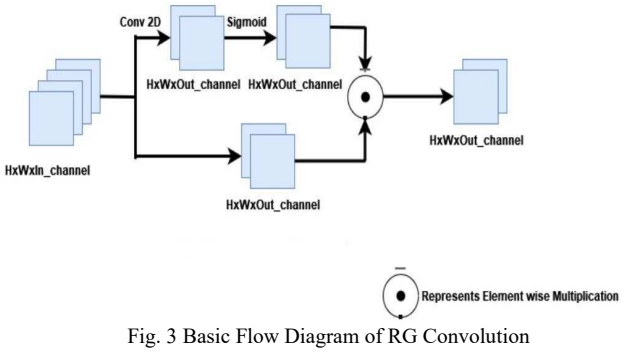

Fig. 3 Basic Flow Diagram of RG Convolution

In the above equations $W_g^{(t)}$ are the Weights (filters) for the gating mechanism at layer t. They determine how the input is filtered before passing through the sigmoid function. $b_g^{(t)}$ are the bias term associated with the gating filters at layer t. $W_c^{(t)}$ are convolutional weights for feature extraction at layer t. These filters convolve with the input to capture spatial features. $b_c^{(t)}$ bias term for the convolution operation at layer $H_c^{(t)}$Convolution output at layer t after applying the filters and bias. $H^{(t)}$Final output of the RGC layer after element-wise multiplication with the gating tensor. $\odot$ represents Element-wise multiplication, modulating the convolved features by their gated importance and σ represents the sigmoid function whose outputs values between 0 and 1, effectively controlling the flow of information based on the importance of the input features. The output of this step maps the feature with 256 channels.

C. U-Net Architecture for Spectral Feature Refinement

After achieving an efficient spatial interaction with the RG Conv, we then design the UNET as shown in Fig.2 to further enhance the model capacity by extracting the spectral features of the HSI. A general UNET architecture can be explained as it is shown in Fig 4. An encoding path is on the left of the design, while a decoding path is on the right. The encoder uses a conventional convolutional network topology, performing two successive 3x3 unpadded convolutions; after that, a 2x2 max-pooling operation is performed with a stride of 2 for downsampling and a rectified linear unit (ReLU). As downsampling progresses, the feature channels double at each step. During decoding, the feature map gets unsampled in each stage and applies a 2x2 up-convolution for halving feature channels. This is then concatenated with the appropriately cropped feature map from the encoding phase, and two 3x3 convolutions are performed, both using ReLU activation. Cropping ensures compensation for the border pixels that were lost in the convolution process.

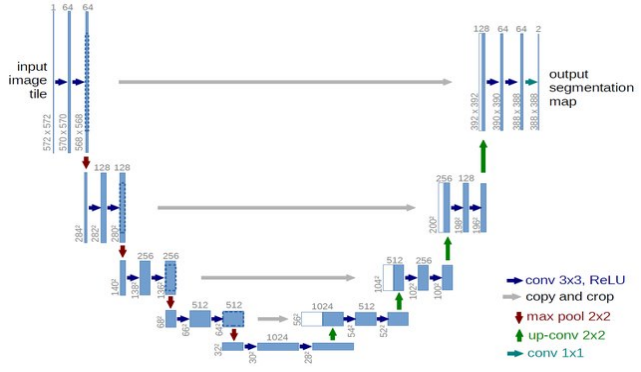

Fig. 4 Basic UNET Architecture

Using the idea from the general architecture of UNET from above Fig.4, we extended the output obtained from RG Convolution to this type of UNET structure where the input will have 256 feature channels, which will be increased to 512 and 1024 in the next two down-sampling respectively. the number of the feature channel will be once again bought to 256 after performing upsampling twice. A 2x2 convolution is performed at the output to get the requisite number of classes (16 in IP and 9 in PU) and finally, using global average pooling, output is obtained in the single dimension. The above process can be explained mathematically, as follows:

i. First Down Sampling (Encoder):

$$H_{down1} = ReLU\left(W_{down1} * H^{(3)} + b_{down1}\right) \quad (10)$$

Where the output of this stage is $H_{down1} \in \mathbb{R}^{H/2\times W/2\times 512}$

ii. Second Down Sampling (Encoder):

$$H_{down2} = ReLU\left(W_{down2} * H_{down1} + b_{down2}\right) \quad (11)$$

Where the output of this stages is $H_{down2} \in \mathbb{R}^{H/4\times W/4\times 1024}$

iii. First Up Sampling (Decoder):

During both the up-sampling stage three processes are performed Viz. Interpolation, concatenation and finally convolution to obtain the requisite output.

$$H_{up1} = \left(interpolate\left(H_{down2}\right) \oplus H_{down1}\right) \quad (12)$$

$$\in \mathbb{R}^{H/2\times W/2\times(1024+512)}$$

After this convolution is performed as

$$H_{up1} = ReLU\left(W_{up1} * H_{up1} + b_{up1}\right) \quad (13)$$

The output of this stage is

$$H_{up1} \in \mathbb{R}^{H/2\times W/2\times 512} \quad (14)$$

iv. Second Stage Up Sampling:

During this stage, the same operations are repeated to get the 16-channel output.

$$H_{up2} = \left(interpolate\left(H_{up1}\right) \oplus H^{(3)}\right) \in \mathbb{R}^{H\times W\times(512+256)} \quad (15)$$

After this convolution is performed as

$$H_{up2} = ReLU\left(W_{up2} * H_{up2} + b_{up2}\right) \quad (16)$$

And the output in this stage is

$$H_{up2} \in \mathbb{R}^{H\times W\times 16} \quad (17)$$

Finally, Average Global Pooling is performed to get the final output:

$$H_{GAP} = \frac{1}{H\times W}\sum_{i=1}^{H}\sum_{j=1}^{W} H_{up2}\,(i,j,k)\ for\ k = 1,2,3\ldots..16 \quad (18)$$

Where we will get $H_{GAP} \in \mathbb{R}^{16}$, a 1D vector representing the final feature as the output.

IV. Result analysis

A. Data Analysis

In order to assess the efficiency of the suggested model, two of the most popular datasets of HSI are utilized. Those are described as below:

1. To collect the Indian Pines (IP) dataset at a 20-meter ground sampling distance, the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) was used. With an image size of 145x145 pixels, it has 220 spectral bands, each having a spectral resolution of 10 nm. There are sixteen different classes in the dataset.

2. Reflective Optics System Imaging Spectrometer (ROSIS) is a special sensor that was utilised to acquire the Pavia University (PU) dataset. It is comprised of a 610x340 pixel picture with 103 spectral bands that span from 430 nm to 860 nm. Nine land cover classes make up the dataset's classification.

B. Experimental Setup

RGConvNET is compiled and executed using the Pytorch platform. Optimization is done with the help of Adam Optimiser. The learning rate is considered as 0.001. The experimental setup includes the following hardware and software specifications: CPU: Intel Core i5-1035G1 (base frequency 1.00 GHz, turbo boost up to 1.19 GHz), GPU: NVIDIA Quadro RTX 8000, RAM: 8 GB, and operating system: Windows 11. During training the dataset 50 epochs were used with early stopping to stop the model from overfitting. The patch size is considered as 5x5.

C. Feature Extraction

RG Convolution is a integration of two deep learning models consisting of convolutional and activation sections using a sigmoid function. RG Convolution is used for the spatial feature extraction. It comprises of three convolution layers, namely RGCONV 1, RGCONV 2, and RGCONV 3. Each of these has different dimensional sizes (64, 128, and 256 numbers).

The second part consists of UNET architecture which employed to extract spectral feature of HSI. Additionally, we utilize PCA during the preprocessing for dimensionality reduction over the HSI hypercube, at the input of the hyperspectral image. At the output, a 2D convolution layer used with Global average pooling for getting the final classification.

D. Experimental Analysis:

The efficiency of the suggested model is assessed using different types of Accuracy used for classification of HSI. The first one of them is Overall Accuracy (OA) that represents the total classification accuracy and is calculated by dividing total number of correctly classified samples (N_C) to the number of reference samples (N_A). In all the class levels, average classification accuracy is used and is denoted by AA. The last one is the Kappa coefficient (κ), which is obtained by dividing the observed accuracy and the expected classification accuracy.

$$OA = \frac{N_A}{N_C} \times 100\% \tag{19}$$

$$AA = \frac{1}{Cl} \sum_{m=1}^{Cl} \frac{N_C^m}{N_{All}^m} \times 100\% \tag{20}$$

$$\kappa = \frac{OA - P_e}{1 - P_e} \times 100\% \tag{21}$$

Tables 1 and 2 represent the comparison of accuracy of individual classes of Indian pines and Pavia University datasets for the suggested model and other existing approaches along with OA, AA, and κ. We haven't used a conventional model to evaluate our concept. Rather, Random Forest, SVM, MLP, RNN, GRU, CNN-1D, MorphCNN, and RLSBSA are used for comparison, which proclaimed to be the best models for the classification of the HSI. RGConvNET achieves the overall accuracy of 99.84%, average accuracy as 99.77% and 99.79% kappa accuracy for the Pavia University dataset. Similarly, for the IP dataset, the test accuracy achieved is 98.56%, 98.39%, and 98.35% as overall, average and kappa accuracy, respectively. The proposed model has a better test result for these two datasets than the state-of-art algorithms. The test results are visually compared as illustrated in the figure (Fig 5-6).

V. Novelty and summurization

In this research work, we propose an innovative method for hyperspectral image (HSI) classification by designing a hybrid model named RGConvNET. This model uniquely integrates two deep learning techniques—Recursive Gated Convolution and UNET—to enhance the overall performance of Hyperspectral image classification.

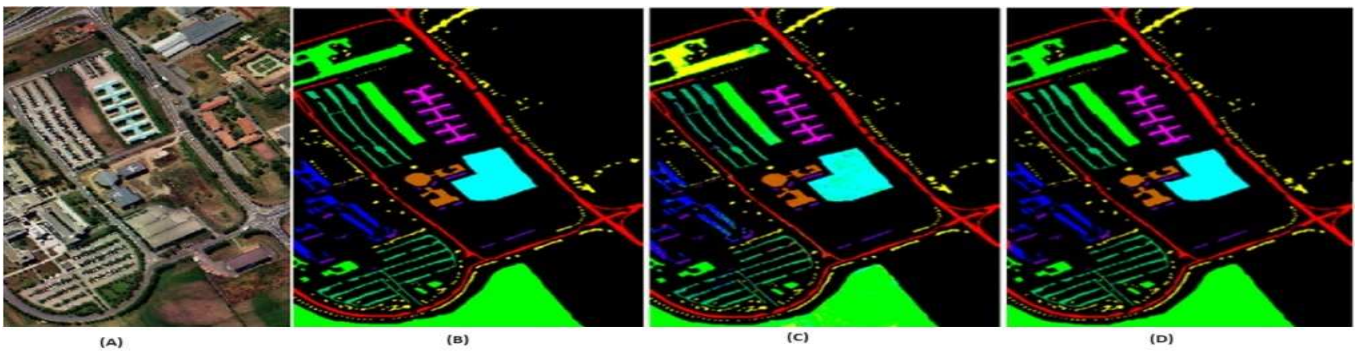

Fig. 5 Classification mapping (A) RGB image (B) UNET (C) Ground Truth (D) RGConvNET Model considering Pavia University dataset

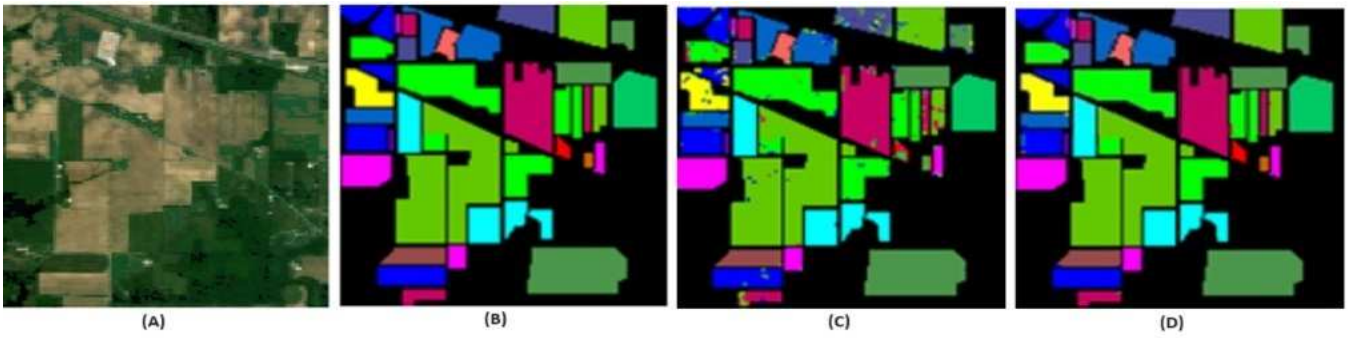

Fig. 6 Classification mapping (A) RGB image (B) UNET (C) Ground Truth (D) RGConvNET Model Model considering Indian Pines dataset

TABLE I. COMPARISION OF ALL CLASSES OF PAVIA UNIVERSITY DATASET WITH EXISTING IN TERMS OF ACCURACY

:

Class	SVM	RF	RNN	GRU	MLP	CNN 1D	UNET	RLSB S-A	Proposed Model
1.	89.98	82.23	84.53	83.08	77.25	87.18	94.52	99.8	99.96
2.	74.39	65.81	75.13	76.9	80.1	89.64	97.12	99.9	100
3.	38.42	66.72	68.37	65.17	54.79	71.1	85.08	98.6	99.88
4.	98.24	97.77	93.5	90.72	92.05	95.32	97	94.9	99.91
5.	95.98	99.37	99.37	99.23	99.51	99.48	99.25	99.9	100
6.	51.43	91.62	89.94	85.07	74.86	88.28	93.92	100	99.95
7.	80.63	87.36	87.2	82.94	90.17	86.77	84.98	98.4	99.62
8.	96,67	90.46	90.37	85.85	90.42	90.43	96.62	99.5	98.64
9.	94.92	93.71	98.44	94.52	93.51	97.33	97.05	99.3	100
OA	77.44	77.8	82.05	77.07	80.7	89.09	95.51	99.7	99.84
AA	80.18	86.12	87.43	83.83	83.63	89.5	93.95	99.4	99.77
κ	70.34	72.6	76.89	70.84	74.76	85.5	93.95	99.6	99.79

TABLE II. COMPARISION OF ALL CLASSES OF INDIAN PINES DATASET WITH EXISTING IN TERMS OF ACCURACY

Class	SVM	RF	RNN	GRU	MLP	CNN 1D	UNET	RLSB S-A	Proposed Model
1.	85.33	88	73.6	58.4	77.6	80.8	92.27	91.3	94.44
2.	55.11	80	81.45	75.5	81.1	79.38	84.05	97.8	96.84
3.	22.77	69.55	64.55	63.37	70.35	74.26	79.34	97.3	97.28
4.	13.13	48.48	47.07	29.49	53.33	31.92	52.14	95.4	96.80
5.	41.60	87.23	86.94	87.59	88.4	90.73	91.66	94.7	100
6.	94.06	96.33	95.93	95.31	96.38	96.39	95.74	98.7	100
7.	0	50	10.2	0	0	0	0	93.3	100
8.	91.33	100	99.8	99.52	99.12	99.84	100	100	100
9.	40	50	80	56	66	50	44.44	100	100
10.	26.83	76.54	75.35	71.13	78.53	81.83	80.77	95	99.22
11.	81.6	87.7	83.19	78.86	82.29	80.39	88.54	99	98.98
12.	28.95	77.3	78.58	71.91	83.19	84.75	88.46	97.6	97.05
13.	86.25	97.5	98	97	97.75	97.75	87.64	100	100
14.	91.07	91.38	92.92	90.28	92.88	93.32	98.82	99.1	99.60
15.	10.1	80.81	87.88	75.56	93.54	89.9	69.44	99.6	96.75
16.	71.96	87.27	88.64	94.09	96.82	82	84	98.6	97.29
OA	60.8	84.12	82.95	79.07	84.2	84	87.45	97.9	98.56
AA	52.47	79.91	77.66	71.16	78.49	76.76	77.33	96.9	98.39
κ	64.41	81.87	80.56	76.12	82.01	81.81	85.75	97.6	98.35

This technique enables our model to identify key features within hyperspectral images, enhancing the precision of classifications and marking a significant advancement in hyperspectral image analysis.

Spectral Feature Extraction with UNET: Our proposed model also incorporates the UNET architecture, a well-known deep learning framework, to extract spectral features from hyperspectral images. Although UNET is commonly used for various image classification tasks, its use in hyperspectral imaging remains relatively new. By performing this step our model will be able to extract high level spectral information. This model plays a pivotal role in ensuring accurate HSI classification, offering a novel perspective on hyperspectral image analysis through its unique adaptation.

Spectral and Spatial Features Fusion: One of the key innovations of our work is the combination of spatial features extracted by RG Convolution with the spectral features obtained from UNET. This dual feature extraction approach allows the model to capitalize on the strengths of both techniques, enhancing classification performance noticeably. The combined approach of feature extraction and fusion in our research is a special and significant advancement in the field of HSI classification. The model's

sensitivity to hyperparameters like learning rate, kernel size, and the depth of gated convolutions is evident. Optimal tuning ensures robust spatial and spectral feature extraction, avoiding overfitting or underfitting. In contrast to other existing models, the RGC layers dynamically integrate spatial features, while the U-Net architecture effectively amplifies spectral features. This fusion outperforms traditional models in accuracy and computational efficiency. Conclusively, our research advances the existing model in HSI classification while introducing alternative approaches and viewpoints that could be applied to other image analysis fields. The execution of this progressive concept places this research as a one of the most significant contributions to the field of hyperspectral image classification.

VI. Conclusion and Future Scope

This letter introduced a novel HSI classification model that combines RG Convolution and UNET, effectively extracting global spatial and interactive spectral features. The proposed hybrid model was trained, tested, and validated on two different HSI datasets, the outcomes highlight its superiority over the most advanced existing algorithms in the future. We intend to examine other potential combinations alongside advanced fusion strategies to optimize the Classification and segmentation in HSI.

References

[1] Shenming, Q., Xiang, L. & Zhihua, G. A new hyperspectral image classification method based on spatial-spectral features. Sci Rep **12**, 1541 (2022).

[2] X. X. Zhu et al., "Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources," in IEEE Geoscience and Remote Sensing Magazine, vol. 5, no. 4, pp. 8-36, Dec. 2017.

[3] Ying Li and Bo Cheng, "An improved k-nearest neighbor algorithm and its application to high resolution remote sensing image classification," 2009 17th International Conference on Geoinformatics, Fairfax, VA, 2009, pp. 1-4.

[4] Hong-Min Gao, Meng-Xi Xu, Ming-Gang Xu, Xin Wang and Feng-Chen Huang, "A method of hyperspectral image classification based on posterior probability SVM and MRF," 2013 International Conference on Machine Learning and Cybernetics, Tianjin, 2013, pp. 235-240.

[5] C. Shi, D. Liao, T. Zhang and L. Wang, "Hyperspectral Image Classification Based on Expansion Convolution Network," in IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-16, 2022.

[6] X. Cao, J. Yao, Z. Xu and D. Meng, "Hyperspectral Image Classification With Convolutional Neural Network and Active Learning," in IEEE Transactions on Geoscience and Remote Sensing, vol. 58, no. 7, pp. 4604-4616, July 2020.

[7] W. Song, S. Li, L. Fang, and T. Lu, "Hyperspectral image classification with deep feature fusion network," IEEE Trans. Geosci. Remote Sens., vol. 56, no. 6, pp. 3173–3184, Jun. 2018.

[8] R. Hang, Z. Li, Q. Liu, P. Ghamisi, and S. S. Bhattacharyya, "Hyper spectral image classification with attention-aided CNNs," IEEE Trans. Geosci. Remote Sens., vol. 59, no. 3, pp. 2281–2293, Mar. 2021

[9] S N. S. Diya, S. Z. Mishu, M. Sultana, M. A. Hossain and M. Al Mamun, "Hybrid 2D-3D convolution neural network for hyperspectral image classification," 2023 26th International Conference on Computer and Information Technology (ICCIT), Cox's Bazar, Bangladesh, 2023, pp. 1-6.

[10] K. Cho, B. V. Merrienboer, D. Bahdanau, and Y. Bengio, "On the ¨ properties of neural machine translation: Encoder-decoder approaches," arXiv preprint arXiv:1409.1259, 2014..

[11] Y. Rao, W. Zhao, Y. Tang, J. Zhou, S. N. Lim, and J. Lu, "HorNet: Efficient high-order spatial interactions with recursive gated convolutions," in Proc. Adv. Neural Inf. Process. Syst., vol. 35, 2022, pp. 10353–10366.

[12] S. Sivagami, P. Chitra, G. S. R. Kailash, and S. Muralidharan, "Unet architecture based dental panoramic image segmentation," in 2020 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET), 2020, pp. 187 191: IEEE.

[13] T. -T. -H. Le, J. Kim and H. Kim, "Classification performance using gated recurrent unit recurrent neural network on energy disaggregation," 2016 International Conference on Machine Learning and Cybernetics (ICMLC), Jeju, Korea (South), 2016, pp. 105-110.

[14] . Ge, G. Cao, X. Li and P. Fu, "Hyperspectral Image Classification Method Based on 2D–3D CNN and Multibranch Feature Fusion," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 13, pp. 5776-5788, 2020.

[15] S. Liu and W. Deng, "Very deep convolutional neural network based image classification using small training sample size," 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR), Kuala Lumpur, Malaysia, 2015, pp. 730-734.

[16] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770-778.

[17] K. Roy, A. Deria, C. Shah, J. M. Haut, Q. Du and A. Plaza, "Spectral–Spatial Morphological Attention Transformer for Hyperspectral Image Classification," in IEEE Transactions on Geoscience and Remote Sensing, vol. 61, pp. 1-15, 2023.

[18] K Ghotekar, R., Shaw, K., & Rout, M. (2024). Deep Feature segmentation model Driven by Hybrid Convolution Network for Hyper Spectral Image Classification. International Journal of Computing and Digital Systems, 16(1), 719-738..

[19] S. Feng and H. Wang, "Comparison of PCA and LDA Dimensionality Reduction Algorithms based on Wine Dataset," 2021 33rd Chinese Control and Decision Conference (CCDC), Kunming, China, 2021, pp.