Machines and Algorithms
http://www.knovell.org/mna

Research Article

Enhancing Pixel-wise Segmentation with BZNet in Hyperspectral Imaging

Aiman khan[1*], Fatima Yousaf [1] and Sania Iqbal[2]

[1]Department of Computer Science, Bahauddin Zakariya University, Multan, Pakistan
[2]Department of Information Technology, Bahauddin Zakariya University, Multan, Pakistan
[*]Corresponding Author. Email: aimankhanji123@gmail.com
Received: 26 May 2022; Revised: 15 June 2022; Accepted: 03 August 2022; Published: 17 August 2022
AID: 001-02-000009

Abstract: Hyperspectral imaging is a technology used to capture images over a broad range of wavelengths, aiding in the identification of materials based on their distinct spectral signatures. This method possesses capabilities such as detecting various materials, penetrating obscurants like smoke, and providing detailed insight into material composition. The classification of hyperspectral data involves labeling pixels or spectrum ranges on a surface according to their reflective attributes. This task remains challenging due to the data's high dimensionality and the presence of mixed pixels, occurring when different objects share a pixel. Challenges arise from factors like environmental variations, lighting, and structural conditions, affecting classifier accuracy and generalization. Leveraging the power of deep learning, our novel BZNet (a customized SegNet model) introduces an innovative deep fully convolutional neural network architecture for pixel-wise semantic segmentation. The BZNet model enhances the decoder by incorporating skip connections to integrate hierarchical feature maps from the encoder, aligning them with high-resolution input feature maps for precise semantic classification. Extensive testing on various datasets, including combinations thereof, reveals that our BZNet model outperforms SegNet in terms of accuracy, memory usage, and computational speed. Key aspects of our research include the creation of a comprehensive training dataset, drawing from online resources such as the Indian Pines, Salinas, and Pavia University datasets. Multiple classification algorithms, including SegNet, Transnet, HyperNet, and ResNet50, were employed, coupled with various feature extraction methods like Gabor and Laplacian of Gaussian filters. Through rigorous experimentation, we identified the Unet algorithm with the conv2d filter extraction as the most effective. Notably, we achieved remarkable accuracy rates of 97.78% for Indian Pines, 96.67% for Salinas, and an impressive 99.22% for Pavia University classification. These findings underscore the efficacy of our proposed hyperspectral imaging classification system, which holds promise for a wide range of applications. The integration of deep learning techniques, careful dataset curation, and meticulous experimentation has yielded a robust and accurate solution for material identification using hyperspectral data.

Keywords: Hyperspectral Imaging, Deep Learning, Semantic Segmentation, Material Classification, BZNet Model.

1 Introduction

Hyperspectral imaging (HSI) has emerged as a pivotal technology in the realm of remote sensing, offering the ability to capture images spanning a broad spectrum of wavelengths [1]. This technique facilitates the identification of diverse materials and objects based on their distinctive spectral signatures [2]. The process of hyperspectral classification involves the assignment of labels to spatial and spectral features within data, aiming to comprehend their interconnectedness rather than segregating them [3]. In the context of remote sensing, HSI classification poses a significant challenge, and this study focuses on addressing it comprehensively.

The challenge of hyperspectral image (HSI) classification for remote sensing stands as the central focus of this study. HSI classification is a critical domain within data science and computer vision, tasked with the holistic categorization of both spatial and spectral features in data, without isolating them. This challenge is particularly relevant within remote sensing, where HSI classification presents unique complexities.

The motivation behind this research is twofold. Firstly, the scarcity of training data for remote sensing applications poses a significant hurdle [4]. Secondly, the complex task of simultaneously managing spatial and spectral features necessitates innovative approaches [5]. Despite ongoing efforts, the intricacies of hyperspectral classification persist, driving the need for comprehensive solutions. The primary objectives of this research are as follows:

- Dataset Combination: Develop a classification system capable of effectively utilizing combined datasets, contributing to improved accuracy and generalization.
- NPZ Slices and Dimensionality Reduction: Implement NPZ slices and dimensionality reduction techniques to enhance dataset manageability and facilitate a unified classifier system.
- Unified Classifier System: Create a unified classifier system tailored to both combined datasets (Indian Pines, Salinas, Pavia University) and individual datasets.
- Merged Synthetic Dataset: Train and test the proposed classifier on a merged synthetic dataset, evaluating its performance against shuffled combined datasets.
- Performance Evaluation: Thoroughly evaluate the performance of the proposed system using a diverse set of metrics, showcasing its effectiveness in hyperspectral image classification and semantic segmentation.

The research study offers several significant contributions:

- Combination of Multiple Datasets: The integration of multiple datasets enhances the classification model's robustness and generalization capabilities.
- BZNet Architecture: The creation of the BZNet architecture, an extension of the SegNet model, introduces a deep fully convolutional neural network framework for pixel-level semantic segmentation.
- Enhanced Classification with Skip Connections: Incorporating skip connections in BZNet's decoder enhances the fusion of hierarchical feature maps from the encoder, elevating semantic classification accuracy.
- Memory Management: Despite increased trainable parameters, memory constraints are effectively managed through strategic max-pool layer implementation.
- Optimized Pre-Processing: The study undertakes various pre-processing operations, converting hyperspectral datasets into NPZ format while addressing high-dimensional data challenges.

Through an exploration of hyperspectral imaging, classification challenges, the innovative BZNet model, research motivations, problem statements, objectives, and contributions, this study aims to advance the fields of remote sensing and computer vision.

2 Literature Review

Hyperspectral imaging has become a dynamic field that connects remote sensing and computer vision, showing incredible promise. Hyperspectral sensors can gather detailed information about colors and light across a wide range, making them useful for many things. They help us understand the environment and find minerals, among other applications. However, handling this complex data to classify and understand it accurately is tough. This challenge has led researchers to explore advanced methods, especially in machine learning and deep learning. This review explores how hyperspectral images are classified, looking at the difficulties of dealing with lots of information, mixed-up data, and how deep learning helps solve these issues.

2.1 Machine Learning Models

Remotely sensed hyperspectral (HS) images, due to their extensive coverage and multi-channel spectral data, serve as a valuable resource for researchers and scientists. Efforts have been directed towards harnessing the hidden information within these images, employing diverse machine learning techniques. Support Vector Machine (SVM) classifiers, for instance, utilize spectral data for band-based classification [6], where each band's spectral range forms the input feature matrix. Another approach, K-Nearest Neighbors (KNN) and its variations [7], solely exploit spectral attributes as pixel properties. Spectral data's significance in achieving reasonably accurate classification has been recognized in research. Integrating spectral and contextual spatial information has proven advantageous, especially given the similarity of neighboring pixels' spectral features within an HSI. Multispectral images were the primary focus of the earliest attempts at integrating spectral and spatial characteristics [8][9]. Later, Pessaries et al. unveiled a brand-new technique for integrating geographical information called Morphological Profiles (MPs) [10]. MPs employ morphological techniques on grayscale images using a fixed-shape structural component, yielding opening and closing operations (SE). This method's expansion to multi/hyperspectral images saw the introduction of Enhanced Morphological Profiles (EMP) [11] [12]. EMP incorporates MP on principal components obtained through Principal Component Analysis (PCA), effectively addressing dimensionality and band correlation challenges.

2.2 Deep Learning Models

In recent times, the subject of HS classification has garnered significant attention. This section delves into existing methodologies based on deep learning for this purpose. Deep learning techniques have gained traction due to their ability to hierarchically acquire data and learn distinctive and meaningful features. The automatic extraction and representation of these features have been made possible [13] owing to the superior information representation capacity of deep structures. In applications involving classification and target detection, the careful design of deep network topologies can substantially improve identification accuracy. Considering the extreme specificity and complexity of HS images, it is incredibly difficult to extract characteristics from HSI data [14]. To address these obstacles, the adoption of deep learning techniques has gained traction in hyperspectral (HS) feature extraction, classification, and target recognition. The initial forays into using deep learning for these tasks involved Stacked Autoencoders (SAE) and Deep Belief Networks (DBN). However, these methods require reshaping the 3-dimensional structure of HSI into 1-dimensional input vectors, leading to the loss of valuable spatial features. In contrast, Convolutional Neural Networks (CNN), as indicated by their criteria [15], have proven effective in processing unsupervised input HSI and achieving enhanced classification accuracy, unlike SAE and DBN [16]. In order to capture the spectral properties of HS, [17] presented a CNN built on a 5-layered network to address the problem of HS picture classification. They found that using high-dimensional data classifications increased the results of classification. Another method was a 2D-CNN approach, which was presented by Yue et al. [18]. By utilizing Principal Component Analysis (PCA), the dimensionality of the HS image was decreased while keeping the first three principal components. However, the standalone 2D-CNN model captures spatial information from neighboring pixels while disregarding spectral information.

Ghasrodashti et al. [19] devised a robust autoencoder with multiple layers and spatial enhancement capabilities, employing an unsupervised approach to extract a multitude of features from HS images. To extract spatial and spectral information, they used multi-scale functionality weight training, fuzzy patterns, and similarity angle map standards. In addition, Roy et al. [20] suggested a hybrid CNN-based model for HS picture classification that integrates spectral and spatial characteristics. They utilized a 2D-CNN for spectral feature extraction and a 3D-CNN for spatial data, resulting in the extraction of more abstract features. This model makes comprehensive use of both spatial and spectral information in the image.

However, it's worth noting that the CNN model has a large number of parameters and requires substantial training data and time.

A notable challenge in HS image classification is the limited training dataset, which makes the classification model susceptible to overfitting and results in poor generalization [21]. To address this, Ghasrodashti et al. [22] developed an enhanced sparse classification method for HS images based on a hidden Markov random field. By constructing a dictionary with minimal spectral-spatial correlation through sparse coding, they improved the objective function of the sparse classifier, yielding favorable results. The authors in [23] introduced a novel RPNet aimed at enhancing the computational efficiency of the model. This unique approach involves a convolutional kernel layer distinct from the conventional CNN, eliminating the need for training through random projection. This not only enhances feature extraction capabilities but also reduces computational load.

Nonetheless, there are certain limitations in this work. While spatial information within the HS image holds significant value, it tends to be overlooked and underutilized. As image processing techniques progress [24], spatial elements are becoming increasingly crucial in HS classification. Among the various methods for extracting spatial information, Gabor filters have garnered substantial attention due to their ability to provide distinctive features. Gabor filters serve as a valuable tool for unsupervised feature extraction, particularly in describing texture and spatial characteristics within HS images. As a result, the model is less dependent on training data and can extract texture and spatial information [25]. Numerous studies have shown that the use of Gabor filters in HSI categorization produces better outcomes.

For instance, Feng Xiao et al. [26] improved efficiency and accuracy for HS classification by combining a 3-dimensional Gabor filter with an SVM. Wang Liguo et al.'s [27] successful texture feature extraction from photos made use of spectral information decomposition through experimental techniques. Following dimensionality reduction through PCA, Chen et al. [28] devised a methodology wherein 2-dimensional Gabor features were derived from hyperspectral (HS) spectral data. These features were then input into a classification model based on 2D Convolutional Neural Networks (CNN). This innovative approach not only improved classification accuracy but also mitigated reliance on extensive training datasets and eliminated superfluous flattening procedures.

Originally designed for the purpose of segmenting medical images, the u-net architecture and its variations have found applications beyond medicine, extending to fields like hyperspectral (HS) imaging and agriculture. This framework stands out as a preferable choice compared to other segmentation methods due to its ability to generate high-quality segmented feature maps while working with a limited number of samples. This section delves into the existing literature to explore the strategic utilization of u-net topologies by researchers. The pioneering work in the realm of end-to-end and pixel-to-pixel classification through fully convolutional neural networks (FCNs) was introduced by Log et al. [29]. This innovative approach replaced traditional fully connected layers with up-sampling layers, and de-convolution was leveraged to maintain the original shape of input images. However, FCN-based models typically exhibit limitations. Challenges arise when dealing with smaller objects or grouping together comparable objects from different classes, resulting in difficulties for classification [30]. As a consequence, FCN-based networks often generate suboptimal segmentation maps. To address this issue, both the U-Net model and the Atrous Spatial Pyramid Pooling method (DeepLab) have demonstrated potential in tackling such intricacies [31]. Additionally, Conditional Random Fields (CRFs) have been incorporated into FCNs to improve segmentation accuracy. For instance, Yu et al. employed dilated convolutions to aggregate context, preserving spatial features [32]. To extract multi-scale features, Chen et al. used complex convolutions and spatial pooling [33]. Ghiasi et al. introduced Laplacian pyramid reconstruction and refinement to deal with low-feature maps [34]. In the pursuit of segmentation, encoder-decoder-based deep architectures [35] such as Seg-Net [36] have been adopted, along with CNN pooling techniques.

The prowess of u-net architecture shines prominently in the domain of hyperspectral imaging (HSI) when it comes to segmentation and classification tasks. U-net topologies incorporate skip connections and concatenation algorithms, where up-sampled outputs of equivalent dimensions to the contracting path are concatenated to facilitate effective segmentation. This architectural design facilitates feature extraction across multiple scales, allowing the incorporation of additional layers and demanding less training data. Notably, Redouane et al. adapted the u-net architecture for pixel-level classification of oceanic eddies, a task that typically demands substantial data for conventional CNNs. The extension to deep u-net has been harnessed for pixel-level sea-land classification. Despite its remarkable performance, the u-net architecture's application in the hyperspectral domain has garnered comparatively less attention, with recent research publications addressing this gap. A Multiscale Independent Component Analysis (MICA) for hyperspectral image categorization was developed by Ronald Kemker et al. [37]. Their approach focuses on learning lower-level features to detect various patterns like bars, edges, gradients, and textures within the dataset. They also proposed the Stacked Convolution Autoencoder Transfer Learning (SCATL) algorithm, aiming to learn deeper features for identifying components, pixel resolution and channel connections, and particles. The authors evaluated these methods on datasets from Indian Pines, Salinas Valley, and Pavia University. The Stacked Convolution Autoencoder (SCA) method yielded exceptional results, but it demanded more time for training and testing due to the utilization of convolutional neural networks in the classification process.

Wei Zhu et al. [38] presented a graph-based non-local total variation approach for classification. They introduced the primal-dual hybrid gradient approach to handle the variation in their method.

Sen Jia et al. [39] proposed a classification approach based on learning superpixels. They employed a multitask learning algorithm to address limited sample size issues. The process involves dividing a hyperspectral image into homogeneous superpixel chunks, applying a 2D Gabor filter to create a Gabor Cube, extracting superpixel features, reducing dimensionality using spatial-spectral Schrödinger Eigen maps, and finally using SVM for classification on datasets like Indian Pines, Pavia, and Salinas.

Yanni Dong et al. [40] described an ensemble learning approach for dimensionality reduction, which is followed by support vector machine classification. Their ensemble discriminative local metric learning (EDLML) method derives metrics from input samples and their neighbors to create a subspace for metric learning. This approach is adaptable to various datasets and not restricted to normally distributed data. It aims to transform features from the original space into a lower-dimensional space.

Han Zhai et al. [41] introduced a sparse subspace cluster analysis method for classification. Their approach assumes that all intensities within a subspace composed of specific pixels share the same value. In place of fuzzy c-means, random field clustering, and k-means, they concentrated on sparse subspace clustering. The authors used a l2 norm regularizer to provide spatial-spectral metadata to the subspace clustering and a four-neighbor approach to achieve spatial homogeneity among pixels.

Claudio Persil et al. [42] proposed a kernel-based feature selection (kFS) approach for classification. This method selects a subset of consistent and relevant parameters, simplifying data transfer between input and output domains. They employed a genetic algorithm for feature selection, which exhibited good accuracy with a small number of samples when dealing with non-parametric multi-objective functions.

Junshi Xia et al. [43] introduced an ensemble classifier approach called Clustering of Forests for hyperspectral image classification. They utilized a range of decision tree classifiers, including random forest and rotation random forest, to address the computational complexity associated with multispectral images. A novel ensemble method called Rotation Random Forest through Kernel Principal Component Analysis (RoRF-KPCA) was proposed, involving KPCA for dividing the feature space into subsets and combining the resultant feature sets for classification of HS data. This method aimed to capture high-order statistics in the classification process.

2.3 Discussion

The above discussed papers showcase a diverse array of innovative methodologies for hyperspectral image classification, utilizing various machine learning and deep learning techniques. These approaches highlight the ongoing efforts to leverage the rich spectral and spatial information contained within hyperspectral images for accurate and meaningful classification. The adoption of deep learning models, such as Convolutional Neural Networks (CNNs), Stacked Autoencoders (SAEs), and the U-Net architecture, underscores the potential of hierarchical feature extraction in handling the complexities of hyperspectral data. Furthermore, novel techniques like graph-based non-local total variation, ensemble learning with dimensionality reduction, and sparse subspace cluster analysis demonstrate a commitment to addressing challenges related to limited training datasets, dimensionality, and spatial-spectral integration. These advancements contribute to the growing understanding of hyperspectral image classification and pave the way for improved applications in fields ranging from remote sensing and agriculture to medical imaging. As this field continues to evolve, researchers are pushing the boundaries of machine learning and deep learning to unlock the full potential of hyperspectral data analysis.

3 Proposed Methodology

In the realm of hyperspectral (HS) imaging, which continues to gain substantial popularity, the demand for precise and efficient classification techniques has escalated. Conventional classification methods, including pixel-based approaches, often prove to be sluggish and imprecise when applied to HS data. Our novel HS data classification approach adopts a semantic methodology. By circumventing the intricate and time-consuming tasks associated with segmentation, our technique achieves accurate outcomes within a considerably shorter duration. After the initial image pre-processing steps, our suggested approach involves segmenting hyperspectral (HS) images into npz files. For the purpose of classifying individual sub-pixels within the HS image, we utilize a Unet model. The efficiency of our technique in producing precise results within a compressed timeframe positions it as an optimal solution for hyperspectral image classification. The subsequent section will elaborate on the process of structuring a training image set.

Similar to other classification frameworks, our proposed approach for addressing the current issue is comprised of five fundamental components. The initial pivotal stage involves capturing the image requiring digitization into numerical data. Subsequently, this acquired image is transformed into the designated format to align with our pre-processing classification system. The primary goal of this step is to convert the input image, whether multiscale or high-resolution, into three-dimensional data that can be swiftly processed for classification purposes. Hyperspectral (HS) images that have previously undergone treatment are segmented after the pre-processing stage.

3.1 Image Acquisition

The initial phase we introduce marks the inception of our proposed system. To transform the essential image data into manipulatable integers, our system initiates by procuring the requisite image data. These images are expected to possess a three-dimensional resolution. Upon acquisition within our proposed framework, the image is promptly directed to the pre-processing phase for subsequent enhancements.

3.2Pre-processing

This segment encompasses three principal steps:

- Resizing the Images
- Updating Label Values
- Normalization
- Implementing Factor Analysis (FA)
- Saving as npz Files

3.2.1 Resizing the Images

Uniform image dimensions are universally compatible with all deep learning networks. Nevertheless, the dataset employed for this investigation comprises images with varying sizes in terms of height and width. This challenge can be addressed through one of two methods. The first approach involves padding, entailing the addition of supplementary columns and rows to images, thereby standardizing their dimensions (e.g., zero padding, constant padding, etc.). However, this approach might lead to heightened computational expenses. The alternative technique, image resizing, mitigates computational burdens by proportionally adjusting image sizes to achieve uniformity. In our study, we adopted image resizing, which involves scaling input images to dimensions of 256x256 pixels, ensuring consistency across the dataset.

3.2.2 Updating Label Values

As different datasets were amalgamated, it became necessary to modify the dataset labels. This was imperative due to the fact that each dataset encompasses a distinct range of classes, spanning from 0 up to a specific threshold—such as 16 classes in the context of Indian pines. Consequently, the process of updating label numbers was meticulously executed to harmonize the consolidated dataset while ensuring label coherence.

3.2.3 Image Normalization

Deep neural network weights are commonly initialized using small random values, typically falling within the (0–1) range. Subsequently, these weights are adjusted based on the loss computed through an optimization technique. In contrast, the intensity range of input images, often spanning from 0 to 255, is significantly wider. Failing to normalize or rescale such images can result in challenges like the emergence of exploding gradients. This, in turn, can adversely affect the learning process. To mitigate these concerns, input images undergo rescaling before being fed into the network, employing a 0-1 normalization approach. This precautionary step helps avert potential issues, thereby contributing to the effective training of the network. The tags associated with this context include: MTI: Cardiomegaly, Manual: Cardiomegaly, Pulmonary Congestion, 58.

3.2.4 Implementing Factor Analysis (FA)

The amalgamated dataset contains a significant number of elevated frequency ranges, necessitating increased computational capacity. However, through the implementation of the Frequency Amplitude's High-Spectrum (FA's HS) technique, the dataset's dimensions were effectively reduced while retaining essential information.

3.2.5 Saving as npz Files

After completing the previous steps, it's recommended to reprocess or generate new npz files containing data and labels to address the resolved challenges effectively. An essential aspect of the proposed approach is the pre-processing stage. Its main goal is to convert the acquired image into a suitable format for subsequent steps. This entails processing the raw image into npz files for more effective computing, as well as duties like noise removal and normalization.

3.3 Proposed Architecture

HSI images are multi-dimensional data cubes or images with high dimensions (w x H x L), where w stands for the spatial width of the image, H for the image's height, and L for its spectral dimension. The 3D image cube's spectral information is provided by this spectral dimension. It's important to note that the HSI data's high spectral dimension results from the reflection of images obtained from limited spectrum bands. Due to the interconnected nature of these narrow bands, HSI data becomes complex for many deep learning techniques to handle.

To address this complexity, a well-known dimensionality reduction technique is employed on HSI data. Let's consider the variable $O \in R^{(w, H, L)}$ to describe the HSI image cube. Here, O represents the actual input HSI image, with w as width, H as height, and L as spectral data or depth. The label vector for each pixel of the original image "O" is denoted by $y = (y(1), y(2), y(3), ..., y(n)) \in R^{(nC)}$, where n represents the number of pixels and C denotes the categories of land covers.

The difficulty comes from managing mixed classes with a lot of interclass variation and similarity. In order to overcome this, factor analysis (FA) is used on the original image "O" to decrease redundancy in the HSI data. By using FA, the spectral dimension is decreased from L to D while maintaining the spatial information. The data cube for the HSI image is represented as P R (w x H x D) after Principal Component Analysis (PCA), where P is the low-dimensional input data, H is height, W is width, and D is the new spectral band dimension.

The proposed framework uses a 3D architecture to categorize pixels that are HS (Hyperspectral). The model incorporates encoding and decoding routes using a convolutional neural network (CNN) design. The encoding approach transforms HS images and employs convolutional, batch normalization, and pooling layers to extract spatial and spectral data. Convolutional layers are used in the decoding process to enlarge the image while preserving the same image size as the contracting portion through concatenation. After each convolutional block, the design doubles the number of channels at the encoder and cuts them in half at the decoder. The 3D model handles spatial and spectral information within a single phase, making use of 2D convolutional neural networks.

The convolutional operation in the model involves the convolution of input data with 3D kernels. The convolutional operation is produced by the dot product of the kernel and the input data. Spatial dimensions are covered by the kernel when it is run over the input data, and the activation function upholds the model's nonlinearity. The model's parameters and kernel weights are supervised trained using the Adam optimizer.

The decoder section of the model makes the most contribution, expanding features from the previous level to make them stand out more. Skip connections are employed to enhance predictions and reduce loss, resulting in clearer outcomes. The final part of the model involves fully connected layers with a softmax activation function for multiclass classification. Training utilizes batch normalization with a batch size of 30 and 200 epochs, using a learning rate of 0.001 for the experiments.

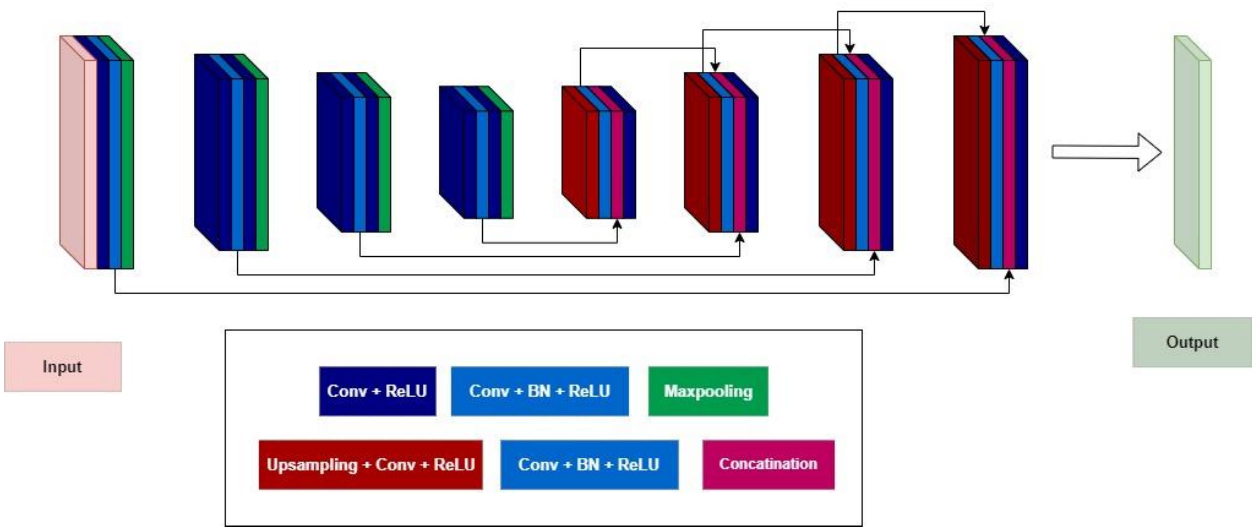

Figure 1: BZNet Architecture

4 Results and Evaluation

We assess the outcomes achieved through the application of our BZNet deep learning architecture to hyperspectral (HS) data. Presented below is a concise summary of the model employed in our study, along with a presentation of the experimental results:

4.1 System's Description

Table 1: Specifications of system

System Processor	3.30 GHz
GPU	NVIDIA GeForce RTX 2080 Ti
RAM	8.00 GB
SSD	1TB
Operating System	Windows 10 Pro
Tool	PyCharm Community Edition 2020.1.2 x64
Language	Python

4.2 Analytical Experiment

In our analytical experiment, we conducted multiple training and testing iterations to optimize parameters for desired outcomes. In the first test, we used the UNET final layer on 150 HS scenes. The second experiment involved training BZNet from scratch on preprocessed data. We trained for 200 epochs, monitoring training and validation loss. After saving weights, we predicted scenes using the trained model in the testing phase.

4.3 Training and Validation Losses

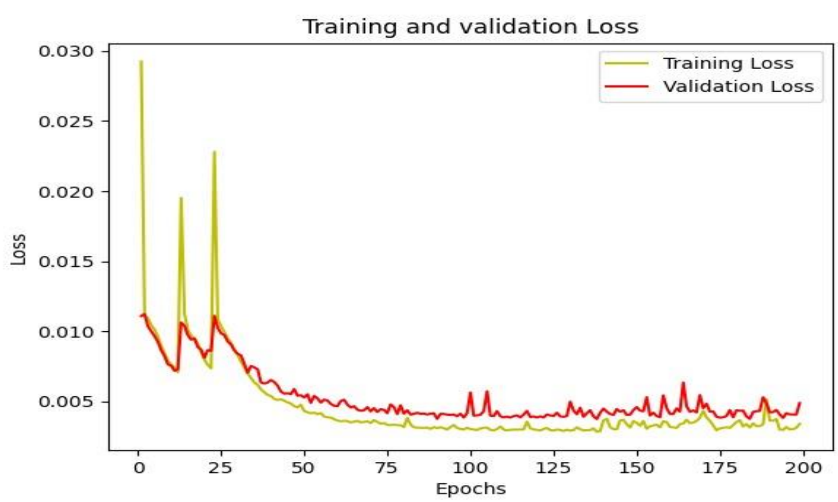

Figure 2: Losses of Training and Validation UNET

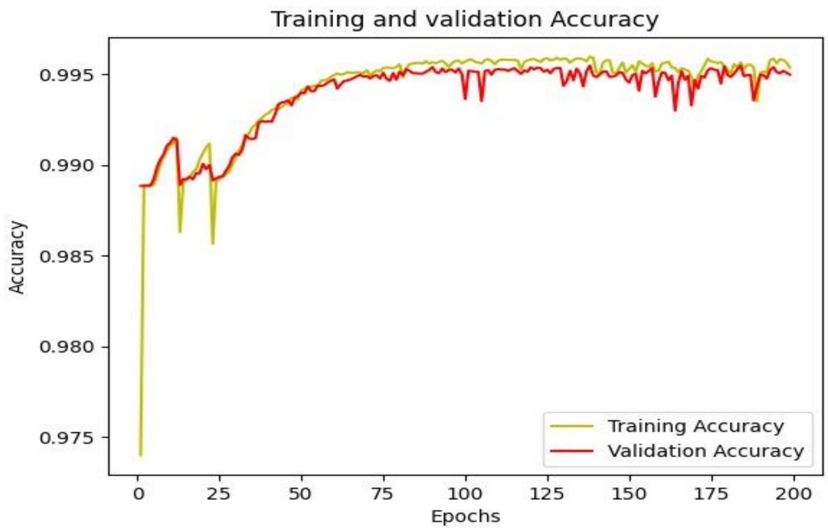

Figure 3: Accuracy of Training and Validation UNET

Figure 4: Training and Validation Accuracy of BZNet

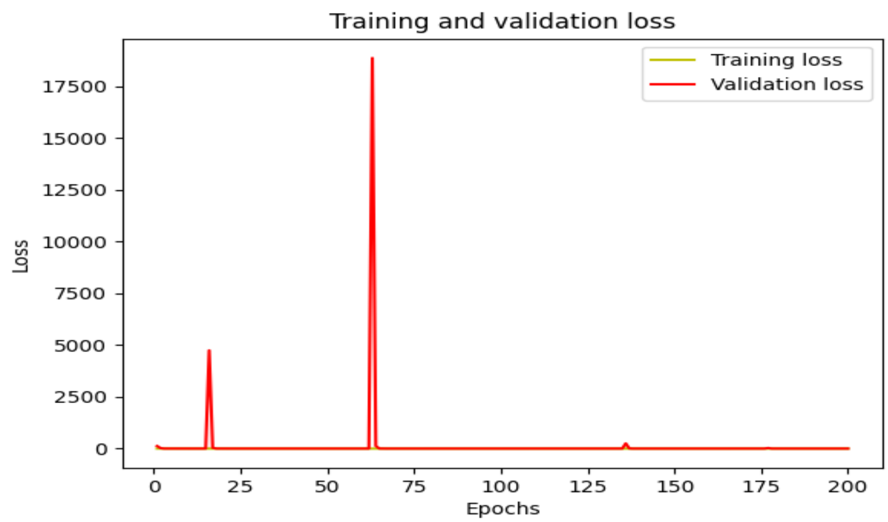

Figure 5: Training and Validation Losses on BZNet

The figures show two test images predictions from our suggested model next to the corresponding original reports.

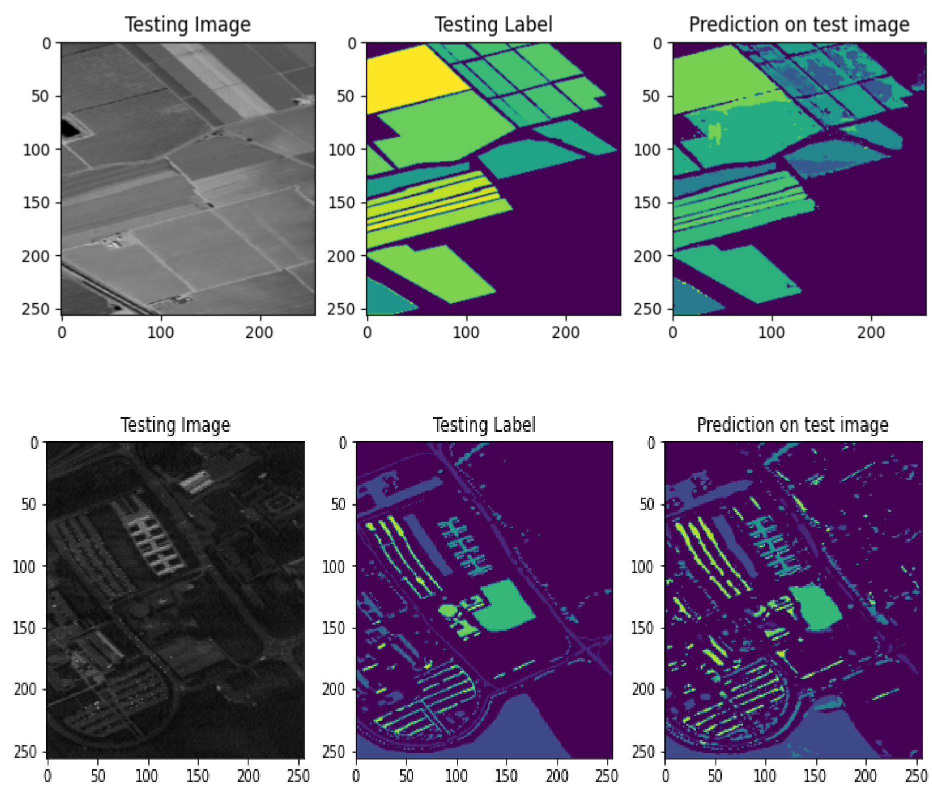

Figure 6: Predictions of Two Test Images on proposed model

4.4 Results

Table 2: Results evaluation on BZNet and UNET models

Dataset	Method	IOU	Value accuracy
HS images	BZNet	0.8888	0.987
	UNET	0.866	0.985

5 Conclusion and Future Direction

This research introduces a novel model, BZNet, designed for the semantic segmentation of hyperspectral images. Built upon an encoder-decoder architecture, BZNet combines elements from UNET and SegNet. In the encoder, BZNet inherits SegNet layers while generating feature maps akin to UNET's encoder. The decoder integrates fresh skip connections. To enhance performance and adaptability across memory sizes, dimension reduction through factor analysis (FA) is employed. This model serves both segmentation and classification purposes. FA is utilized for dimension reduction, but future exploration could involve techniques such as PCA or novel methods. Key findings include the model's adaptability to varying image sizes and datasets.

Future directions should address research gaps. The proposed segmentation system can be adapted for classification tasks. Additionally, alternatives like PCA can be explored for HS dimension reduction. Fine-tuning the network's architecture may optimize computational efficiency.

References

[1] Lodhi, V., Chakravarty, D., & Mitra, P. (2018). Hyperspectral imaging for earth observation: Platforms and instruments. Journal of the Indian Institute of Science, 98, 429-443.
[2] Stanco, F., Battiato, S., & Gallo, G. (Eds.). (2017). Digital imaging for cultural heritage preservation: Analysis, restoration, and reconstruction of ancient artworks. CRC press.
[3] Cao, X., Yao, J., Xu, Z., & Meng, D. (2020). Hyperspectral image classification with convolutional neural network and active learning. IEEE Transactions on Geoscience and Remote Sensing, 58(7), 4604-4616.
[4] Li, S., Song, W., Fang, L., Chen, Y., Ghamisi, P., & Benediktsson, J. A. (2019). Deep learning for hyperspectral image classification: An overview. IEEE Transactions on Geoscience and Remote Sensing, 57(9), 6690-6709.
[5] Zhu, X. X., Tuia, D., Mou, L., Xia, G. S., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep learning in remote sensing: A comprehensive review and list of resources. IEEE geoscience and remote sensing magazine, 5(4), 8-36.
[6] Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE transactions on pattern analysis and machine intelligence, 39(12), 2481-2495.
[7] M. A. U. Rehman, "A New Scale Invariant Optimized Chain Code for Nastaliq Character Representation," 2010 Second International Conference on Computer Modeling and Simulation, 2010, pp. 400-403, doi: 10.1109/ICCMS.2010.493.
[8] Abacha, A. B., et al. (2017). NLM at ImageCLEF 2017 caption task. Working Notes of CLEF 2017-Conference and Labs of the Evaluation Forum (CLEF 2017), Dublin, Ireland, September 11-14, 2017., CEUR Workshop Proceedings.
[9] Xia, J., et al. (2016). "Hyperspectral image classification with canonical correlation forests." 55(1): 421-

431.
[10] Xu, K., et al. (2015). Show, attend and tell: Neural image caption generation with visual attention. International conference on machine learning, PMLR.
[11] Xu, Y., et al. (2018). "Hyperspectral image classification via a random patches network." 142: 344-357.
[12] Yang, H., et al. (2011). "Unsupervised hyperspectral band selection using graphics processing units." 4(3): 660-668.
[13] Yin, J., et al. (2021). "Spatial-spectral network for hyperspectral image classification: A 3-D CNN and Bi-LSTM framework." 13(12): 2353.
[14] Roy, S. K., et al. (2020). "Attention-based adaptive spectral–spatial kernel ResNet for hyperspectral image classification." 59(9): 7831-7843.
[15] Simonyan, K. and A. J. a. p. a. Zisserman (2014). "Very deep convolutional networks for large-scale image recognition."
[16] Szegedy, C., et al. (2015). Going deeper with convolutions. Proceedings of the IEEE conference on computer vision and pattern recognition.
[17] Tarabalka, Y., et al. (2009). "Segmentation and classification of hyperspectral images using minimum spanning forest grown from automatically selected markers." 40(5): 1267-1279.
[18] Theiler, J., et al. (2011). "Sparse matrix transform for hyperspectral image processing." 5(3): 424-437.
[19] Tsochatzidis, L., et al. (2017). "Computer-aided diagnosis of mammographic masses based on a supervised content-based image retrieval approach." 71: 106- 117.
[20] Tu, B., et al. (2020). "Spectral–spatial hyperspectral classification via structuralkernel collaborative representation." 18(5): 861-865.
[21] Vinyals, O., et al. (2015). Show and tell: A neural image caption generator. Proceedings of the IEEE conference on computer vision and pattern recognition.
[22] Wang, L., et al. (2016). "Hyperspectral image classification by combining empirical mode decomposition with Gabor filtering." 37(2): 284-290.
[23] Xia, J., et al. (2014). Rotation-based ensemble classifiers for high-dimensional data. Fusion in Computer Vision, Springer: 135-160.
[24] Xia, J., et al. (2016). "Class-separation-based rotation forest for hyperspectral image classification." 13(4): 584-588.
[25] Zhang, H., et al. (2016). "Spectral–spatial sparse subspace clustering for hyperspectral remote sensing images." 54(6): 3672-3684.
[26] Zhu, W., et al. (2017). "Unsupervised classification in hyperspectral imagery with nonlocal total variation and primal-dual hybrid gradient algorithm." 55(5): 2786-2798
[27] Chen, Y., et al. (2017). "Hyperspectral images classification with Gabor filtering and convolutional neural network." 14(12): 2355-2359.
[28] Gamanya, R., et al. (2007). "An automated satellite image classification design using object-oriented segmentation algorithms: A move towards standardization." 32(2): 616-624.
[29] García Seco de Herrera, A., et al. (2018). Overview of the ImageCLEF 2018 caption prediction tasks. Working Notes of CLEF 2018-Conference and Labs of the Evaluation Forum (CLEF 2018), Avignon, France, September 10-14, 2018., CEUR Workshop Proceedings.
[30] Ghasrodashti, E. K., et al. (2018). "Sparse-based classification of hyperspectral images using extended hidden Markov random fields." 11(11): 4101-4112.
[31] Ghasrodashti, E. K. and N. J. S. P. I. C. Sharma (2021). "Hyperspectral image classification using an extended Auto-Encoder method." 92: 116111.
[32] Guo, B., et al. (2008). "Customizing kernel functions for SVM-based hyperspectral image classification." 17(4): 622-629.
[33] He, K., et al. (2016). Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition.
[34] He, L., et al. (2017). "Recent advances on spectral–spatial hyperspectral image classification: An overview and new guidelines." 56(3): 1579-1597.
[35] Hu, W., et al. (2015). "Deep convolutional neural networks for hyperspectral image classification." 2015.
[36] Huang, K.-K., et al. (2021). "Hyperspectral image classification via discriminant Gabor ensemble filter."
[37] Jia, S., et al. (2017). "Superpixel-based multitask learning framework for hyperspectral image classification." 55(5): 2575-2588.
[38] Jiang, J., et al. (2020). "Multilayer spectral-spatial graphs for label noisy robust hyperspectral image classification."

[39] Kemker, R., et al. (2017). "Self-taught feature learning for hyperspectral image classification." 55(5): 2693-2705.
[40] Li, J., et al. (2011). "Hyperspectral image segmentation using a new Bayesian approach with active learning." 49(10): 3947-3960.
[41] Liu, B., et al. (2017). "A semi-supervised convolutional neural network for hyperspectral image classification." 8(9): 839-848.
[42] Persello, C., et al. (2015). "Kernel-based domain-invariant feature selection in hyperspectral images for transfer learning." 54(5): 2615-2626.
[43] Plaza, J., et al. (2009). "Multi-channel morphological profiles for classification of hyperspectral images using support vector machines." 9(1): 196-218.