ISSN 2581 – 3781

AJCSE
ASIAN JOURNAL OF COMPUTER SCIENCE ENGINEERING

Available Online at www.ajcse.info
Asian Journal of Computer Science Engineering 2025;10(4):1-12

RESEARCH ARTICLE

RRCNN-EM: A Deep Learning Framework for Hyperspectral Road Surface Segmentation to Enhance Automotive Safety

Priya Boopalan[1]*, Senthilkumar Srinivasan[2]

[1]ZF Active Safety and Electronics, Michigan, United States, [2]Apple Inc., California, United States.

Received: 28-10-2025; Revised: 22-11-2025; Accepted: 24-11-2025

ABSTRACT

Recent advances in hyperspectral imaging technology have created new opportunities for enhancing roadway safety through automated surface-condition monitoring. Traditional road inspection techniques, relying on RGB or infrared imagery, often fail to distinguish visually similar materials such as dry pavement, oil spills, water films, or surface cracks – conditions that significantly affect vehicle traction and traffic safety. To address these limitations, this research proposes a recurrent residual convolutional neural network with ensemble modeling (RRCNN-EM) for hyperspectral road-surface segmentation and hazard classification. The proposed framework integrates a Weiner filter for denoising hyperspectral data, followed by RRCNN-based segmentation to preserve spatial dependencies and handle high-dimensional spectral information. In the feature extraction stage, local gradient tangent pattern, gray-level co-occurrence matrix, and spectral indices are employed to enhance the representation of material textures and surface variations. The extracted features are classified using deep learning (DL)-based ensemble models, including deep belief network, quantum neural network, and restricted Boltzmann machine, to improve accuracy and generalization. Experimental analysis on hyperspectral datasets representing asphalt, concrete, water, and oil-contaminated surfaces demonstrates that the proposed RRCNN-EM achieves superior performance compared with conventional convolutional neural networks and U-Net architectures in terms of Accuracy, Precision, Recall, F1-Score, Jaccard Index, and Dice Coefficient. The findings highlight the potential of hyperspectral DL for proactive transportation safety management, enabling early detection of hazardous roadway conditions and supporting data-driven infrastructure maintenance planning.

Key words: Deep belief network, hyperspectral image, quantum neural network, recurrent residual convolutional neural network, remote sensing, restricted Boltzmann machine

INTRODUCTION

In recent decades, hyperspectral imaging (HSI) technology has witnessed significant growth in diverse domains, including environmental monitoring, agriculture, material identification, medical diagnostics, and food quality assessment. More recently, HSI has shown increasing potential in transportation, particularly for enhancing roadway safety, supporting autonomous driving, and monitoring traffic infrastructure.[1] Initially developed in remote sensing research, primarily by NASA for Earth observation, HSI captures rich spectral and spatial information, enabling detailed material and surface characterization beyond what conventional RGB or infrared imaging can provide.[2]

The capability of HSI to identify road surface anomalies – such as cracks, potholes, water accumulation, oil spills, and ice patches – has direct implications for vehicle safety. Detecting these conditions in real-time is critical for maintaining vehicle traction, reducing accident risk, and informing driver assistance systems under adverse weather conditions.[3] By providing high-resolution spectral data, HSI can support predictive maintenance and proactive hazard detection, thereby enabling safer driving environments.

Despite its potential, HSI analysis for transportation faces several challenges, including high-dimensional data, limited labeled datasets, class imbalance, and computational complexity.[4]

Address for correspondence:
Priya Boopalan
E-mail: priyaboopalan@yahoo.com

Traditional computer vision methods often struggle to provide the real-time accuracy needed for autonomous or connected vehicle applications. Recent advances in deep learning (DL) and machine learning (ML) have improved the automated identification of surface conditions. Convolutional neural networks (CNNs), recurrent neural networks (RNNs), and ensemble models have been applied to segment road surfaces and detect anomalies with high precision.[5-7] Feature extraction methods, such as gray-level co-occurrence matrix (GLCM), local gradient tangent patterns (LGTP), and spectral indices, have enhanced the detection of texture, moisture, and material composition, which are essential for assessing traction and safety.

Importantly, HSI-based systems can be integrated with vehicle sensors and driver assistance platforms to segment road scenes into drivable and non-drivable areas, detect lane markings, and identify hazards such as debris or water patches.[8,9] Such integration supports near real-time safety interventions, including adaptive cruise control, lane-keeping assistance, and automated emergency braking. While most studies have used satellite-based datasets like Sentinel-2, the adoption of road-level HSI or autonomous vehicle sensor datasets would improve relevance to on-road safety scenarios and system integration.

This paper explores the application of HSI for transportation safety by proposing a recurrent residual convolutional neural network (RRCNN) for road surface segmentation, coupled with an ensemble classification framework to improve detection robustness. The contributions of this research include:

- Developing an HSI-based approach for accurate road surface segmentation
- Linking road surface anomaly detection with vehicle safety-relevant outputs to support proactive hazard management
- Implementing an ensemble classification strategy to enhance reliability under diverse environmental conditions.[10]

This research work is structured as follows: Section 2 explains the literature review and Section 3 demonstrates the overall process and proposed description. Section 4 indicates the results of the proposed method, and Section 5 briefs the overall conclusion of the paper.

LITERATURE REVIEW

Several approaches are presented in the literature for coupling spatial dimension and spectral information in HSI.

Huang *et al*.[11] presented a spatial-spectral transformer network (SS-TMNet) with multi-scale convolution to extract information from spatial-spectral regions. This SS-TMNet comprises two key modules: A spatial-spectral attention module (SSAM) and a multi-scale 3D convolution projection module (MSCP). For extracting spectral-spatial features, MSCP utilized multi-scale 3D convolutions. The spectral and spatial features were extracted using the SSAM module, which consists of three branches: Spectral attention, height, and width. Finally, SS-TMNet was evaluated on three HSI datasets: Houston2013, Pavia University, and Indian Pines. This method was compared with progressive transformer-grounded and CNN-based approaches and demonstrated superior accuracy on three datasets. However, self-supervised learning needs to be developed to enhance classification performance.

Ullah *et al*.[12] introduced deep smooth wavelet CNN shots for HSI classification. The deep hyperspectral shots architecture comprises five functional components: Preprocessing, the core or base model, snapshot generation, further snapshot optimization, and ensemble assortment, culminating in the output portion. Spectral features were extracted using a deep smooth wavelet CNN and subsequently combined with 2-D CNN to extract spatial features. Cyclic Annealing Schedule (CAS) was employed to converge on different local minima, and the base model was saved. This method achieved higher accuracies on all three datasets: Indian Pines, Pavia University, and Salinas. However, a lightweight ensemble must be constructed to maintain classification performance while minimizing computational complexity.

Chhapariya *et al*.[13] developed a Deep Spectral-Spatial Residual Attention Network (DSSpRAN) for HSI classification. The method comprises two sections: Feature extraction and feature classification. Initially, spectral and spatial features were extracted using the Spectral Residual Attention Network (SRAN) and SpRAN, two core modules. These extracted features were then stacked in the classification module. The classification block

included a Batch Normalization (BN), Rectified Linear Unit (ReLU), convolutional layer, and dropout layers to learn low-level features. Finally, a fully connected layer was utilized to learn high-level features for classification and prediction. This method achieved improved classification accuracy with limited training samples, but further optimization is required to reduce computation time and enhance algorithm efficiency.

Wang *et al*.[14] established a multi-scale spatial feature extraction residual network (MFERN) for HSI classification. The network consisted of two modules: A spectral segmentation residual module (SSRM) and a multi-scale spatial feature extraction (MSFE) module. First, spectral bands of the input image were divided into several sub-bands by MSFE, and each sub-band was preprocessed with different receptive fields. Subsequently, dense block-based group convolution was combined with the MSFE module by SSRM to synthesize extracted feature maps. This module extracted feature information and achieved higher classification accuracy, although the complexity of this method was not fully optimized.

Zhang and Wang[15] devised a spatial proximity feature selection technique for HIS classification employing a residual spatial-spectral attention network. This method comprises three modules: Residual spatial attention, spatial proximity feature selection, and residual spectral attention module. Initially, the residual spatial attention module was utilized to select critical spatial information, assigning weights to various features. Subsequently, for spectral bands, a residual spectral attention module was deployed to emphasize valuable bands. Finally, features were effectively filtered by utilizing the spatial proximity feature selection module. This method demonstrated superior performance in terms of efficiency and accuracy, although with increased time consumption due to the iterative process.

Ge *et al*.[16] presented a pyramidal multiscale spectral-spatial convolutional network for HIS classification utilizing polarized self-attention. This network comprises three stages: Channel-wise feature extraction, classification network, and spatial-wise feature extraction. These networks were employed to extract spectral and spatial features and generate classification results. Subsequently, spatial and spectral features of HIS were extracted through polarized attention and pyramidal convolutional block integration. To converge the network, residual and one-shot aggregation were employed, and a classification was utilized to obtain the classification results. This network outperformed and produced spatially smoother classification maps. However, convolutional networks with PSA networks must be combined to apply to other HIS datasets.

Wang *et al*.[17] presented a novel Transformer Network CNN-enhanced Cross-Attention system for HIS classification. The system utilized a module for shallow feature extraction, employing various fusion approaches to integrate multi-scale spatial-spectral features. Subsequently, shallow features were converted into tokens by a module with different quantities, and a CNN with a transformer module was employed to enhance cross-attention. Finally, a classifier head was utilized, taking input pixels and classification labels as outputs. This network demonstrated superior classification performance, although a more lightweight operation must be designed to provide efficient classification models.

Li *et al*.[18] presented a crop classification application utilizing DL to enhance crop classification accuracy. Initially, samples were selected, and for classification experiments, various time-series images were constructed. Subsequently, multiple DL models were employed, including a benchmark model known as random forest (RF). Finally, all classification results were validated, compared, and analyzed. The results indicated that the integration of multi-spectral or spatial information with temporal features improved classification accuracy. However, the RNN and CNN need to be combined with a convolutional long short-term memory architecture to further improve image crop classification.

Mao *et al*.[19] developed a Dual-1 DCNN algorithm utilizing the CNN model for crop type detection. Initially, data from the European Space Agency (ESA) Sentinels Scientific data hub website, specifically S1A and S2 data, were downloaded. Subsequently, S1A data underwent preprocessing within the Sentinel application platform, encompassing stages such as radiometric calibration, speckle filtering, orthorectification, and re-projection. Similarly, S2 data underwent preprocessing, involving stages such as atmosphere calibration, cloud masking, resampling, and gap filling. Subsequently, the Dual-1 CNN model was

employed to integrate the S1A and S2 time-series data by constructing two Dual-1 CNN modules. Finally, the Dual-1 DCNN model was compared with ML methods, including RF and support vector machine. This algorithm enhanced the timeliness of crop identification, particularly for yam and cotton. However, the timeliness of crop mapping still requires further improvement.

Mohammadi *et al*.[20] developed a 3D fully convolutional network to supervise intermediate layers by employing two middle supervision methods: Supervised contrastive loss middle supervision and cross-entropy loss middle supervision. This technique considers features belonging to similar classes while separating features from different classes. Furthermore, feature clustering and discrimination were enhanced throughout the network to improve its performance, and the effectiveness of intersection over union and F1 loss functions was verified. Results indicated that these methods were suitable for crop-type mapping. However, loss functions and the capability of middle supervision should be focused more on discrimination in crops present in agricultural landscapes.

Khan *et al*.[21] introduced a robust LSTM-based DL model for the early identification of crop types. Initially, a ground-based survey was conducted to record the planted crops and coordinates in the respective fields. Subsequently, time-series images of mapped fields were acquired from the Sentinel-2 satellite. A DL-based LSTM was then utilized for the mapping of crops at an early stage. Finally, the LSTM layer was followed by a linear SoftMax layer and was trained using an Adam optimizer and a cross-entropy loss. This model achieved higher accuracy and performed well on small-holder farms, but the detection time needs to be minimized, and the model should be extended to include more crops by incorporating additional satellite data.

In a study conducted by Kaur *et al*.,[22] an artificial intelligence-based framework was developed to estimate crop yield using an optical satellite dataset and image fusion of microwave data. The proposed framework methodology involved pre-processing both the optical and microwave input datasets. Subsequently, images were fused using the Nearest-Neighbor-based image fusion algorithm. Following image fusion, the post-classification change detection technique was employed, which comprised two steps: Classification and change detection. Finally, after generating change maps and classifications, the soil moisture active and passive dataset was utilized for validating the classified and changed maps. The suggested framework demonstrated improved results and higher accuracy, but it was noted that the omission and commission errors in map production were significant.

PROPOSED METHODOLOGY

To support roadway safety and vehicle hazard detection, the input Sentinel-2 HSI undergoes preprocessing using the Wiener filter to reduce noise while preserving critical spectral information. The preprocessed data are then segmented using the RRCNN, which identifies road surface regions, lane markings, and potential hazard zones relevant for vehicular operations. Features are extracted from the segmented images using GLCM, vegetation index (VI)-based, and LGTP methods, capturing texture, material composition, and subtle anomalies that could affect traction and driving safety. A hybrid classification framework combining quantum neural network (QNN), deep belief network (DBN), and restricted Boltzmann machine (RBM) is employed to classify road surface conditions into 16 distinct categories. The classifier leverages both the extracted feature set and the segmentation output to improve accuracy and robustness under diverse environmental conditions. Figure 1 presents the overall workflow of the proposed system, illustrating how HSI can generate actionable information for real-time automotive safety applications.

Dataset Description

Sentinel-2 is a multispectral imaging program developed under the ESA Copernicus Program,

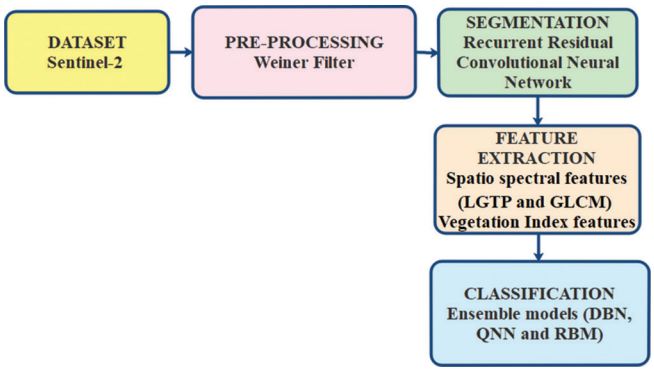

Figure 1: Overall flow diagram of the proposed system

originally aimed at environmental and security monitoring. The two satellite missions, Sentinel-2A (2015) and Sentinel-2B (2017), are equipped with advanced multispectral instrument sensors capable of capturing high-resolution optical imagery of land and water surfaces. With 13 spectral bands, Sentinel-2 provides detailed spectral information that can be leveraged to assess surface conditions, detect anomalies, and monitor changes in land cover and vegetation health.[23] Although designed for large-scale environmental monitoring, Sentinel-2 imagery offers valuable insights for transportation safety research, such as detecting road surface anomalies, water accumulation, and other conditions that may impact vehicle traction and accident risk. The program's frequent revisit cycle and wide-area coverage enable near real-time monitoring of road conditions across large regions.

Preprocessing

In the context of road surface analysis, HSI often contains noise due to sensor limitations, atmospheric effects, or environmental interference. To enhance image quality, the Wiener filter is employed, which reduces noise based on statistical calculations from each pixel's local neighborhood. The filter adaptively applies stronger smoothing in low-variance regions and preserves details in high-variance areas, making it particularly effective for maintaining critical features such as cracks, potholes, or water patches on road surfaces. Wiener filtering, one of the earliest and widely adopted techniques for suppressing additive random noise, minimizes the mean squared error between the reconstructed and original images, assuming that the noise is a stationary random process independent of pixel location.[24] By applying this preprocessing step, the subsequent segmentation and feature extraction processes operate on cleaner, more reliable data, enhancing the accuracy of road safety assessments.

Segmentation

Recently developed three DL models were employed to evaluate various benchmarks. The RCNN and its variants demonstrated superior performance in object recognition tasks.[25] Concurrently, the impact of increased recurrent residual in residual networks' convolutional operations was illustrated. Recurrent convolutional layer (RCL) operations are conducted across discrete time stages, as related to RCNN.[26] In this context, x_l represents an input sample of the residual RCNN (RRCNN) block in the l^{th} layer, and a pixel is assigned at (i,j) on the k^{th} feature map. Furthermore, let us assume that the network output $O_{ijk}^{l}(t)$ is in time step t. The output is provided in Equation (1):

$$O_{ijk}^{l}(t) = (w_k^f)^T * x_l^{f(i,j)}(t) + \left(w_k^r\right)^T * x_l^{r(i,j)}(t-1) + b_k \quad (1)$$

Here, the inputs are $x_l^{f(i,j)}(t)$ and $x_l^{r(i,j)}(t-1)$ for l^{th} RCL and also to standard convolutional layers. Weights of standard convolutional layers are w_k^f and w_k^r values, and b_k is represented as bias and RCL of k^{th} feature map. The ReLU activation function f considers the output of RCL and is expressed in Equation (2):

$$F(x_l, w_l) = f\left(O_{ijk}^{l}(t)\right) = \max(0, O_{ijk}^{l}(t)) \quad (2)$$

The outputs of l^{th} layer are represented as $F(x_l, w_l)$ for the RCNN unit. The upsampling also downsampling layers utilize the output of $F(x_l, w_l)$ in convolutional decoding as well as encoding units. RCNN final outputs are forwarded through a residual unit in the circumstance of R2U-Net, given in [Figure 2d]. Let us consider that the RRCNN-block outcome is represented as well as x_{l+1} and calculated as given in Equation (3):

$$x_{l+1} = x_l + F(x_l, w_l) \quad (3)$$

In this context, input samples of the RRCNN-block are denoted as x_l. For instance, during upsampling or downsampling layers, x_(l+1) sample serves as input to the convolutional units of R2U-Net. However, within the RRCNN block, the dimensions and quantity of feature maps remain consistent for the residual units depicted in [Figure 2d].

The presented DL models are considered stacked convolutional unit building blocks, as illustrated in [Figure 2b and d]. In this work, four distinct constructions were evaluated. First, feature concatenation and forward convolution layers were employed with U-Net to replicate and crop methods in the primary version of U-Net. Figure 2a provides a basic convolutional unit model description.

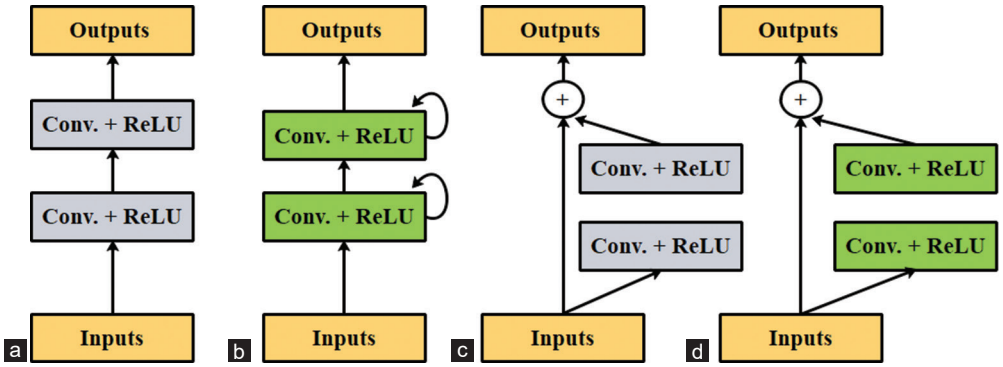

Figure 2: Different types of convolutional and recurrent convolutional units (a) Forward convolutional units, (b) Recurrent convolutional block, (c) Residual convolutional unit, and (d) Recurrent residual convolutional units

Second, U-Net with residual connectivity and forward convolutional layers was utilized, referred to as Residual U-Net (ResU-Net), as shown in [Figure 2c]. U-Net along with forward RCL was designated as RU-Net and considered the third architecture, as depicted in [Figure 2c]. Finally, the recurrent convolution layers along U-Net with connectivity in residuality were named R2U-Net, which is the last architecture and is illustrated in [Figure 2d]. In this context, t = 2 (0 ~ 2) represents a recurring convolution operation, involving only one single convolution layer and two RCL. In practice, feature map concatenation was applied to both RUNet and R2U-Net models.

Despite the potential for feature accumulation, this model demonstrates advantages during both training and testing phases. This feature accumulation enhances the representation of features across various time steps. Consequently, feature accumulation facilitates the extraction of low-level features for diverse modalities and segmentation tasks in medical imaging. Figure 2 presents the variants of RRCNN models.

Feature Extraction

Feature extraction enhances accuracy in learned approaches by representing inputted data features. In addition, minimization of data dimensionality is accompanied by the termination of replicated data. Here, consideration of two feature kinds is accomplished:

Spatiospectral features

LGTP features: It defines a prolonged form of traditional LBP. This contains three-valued codes in the classification of texture.[27] LGTP uses a three-level thresholding system: values within ±T of the center pixel are assigned 0, values above +T are assigned 1, and values below −T are assigned −1. This produces a 3-valued local texture pattern as described in Equation (4).

$$Lgtp' = \begin{cases} 1; I_b^{Seg} \geq i_c + T \\ 0; \left| I_b^{Seg} - i_c \right| < T \\ -1; I_b^{Seg} \leq i_c - T \end{cases} \quad (4)$$

Equation (4) has LTP code which organized at binary LBP codes. LTP code is here to widely present the noise as well as gray level conversion, where T is set as the threshold.

GLCM features: GLCM deals with measuring spatial links between pixels. GLCM features are Energy, Entropy, Variance, Correlation, Sum Average, Contrast, Homogeneity, Difference Variance, Sum Variance, Sum Entropy, Matthews correlation coefficient (MCC), Difference Entropy, IMC 1, and IMC 2.

VI features

VI is referred as a single value computed through the transformation of multiple spectral band observations. The VI features are subjugated for improving the vegetation availability, green features, and therefore guides for distinguishing from other objects in the image. This feature improves contrast among soil as well as vegetation which is given in Equation (5). The index value is between −1 and 1. The common range for green vegetation is 0.2 to 0.8.

$$NDVI = \frac{I_{b1}^{Seg} - I_{b2}^{Seg}}{I_{b1}^{Seg} + I_{b2}^{Seg}} \quad (5)$$

HSI Classification

As specified, a hybrid model with a grouping of DBN, QNN as well as RBN is helpful for HSI

classification, considering two inputs: I_b^{Seg} and *Fs*. Initially, the inputs are exposed to DBN, QNN, and RBN, respectively. After that, classification productions are averaged when compared to the ensemble for the determination of the last classification result.

DBN

As a more computationally efficient kind of feedforward neural networks, DBNs find extensive use. To build a DBN network, a trained neural network's slow learning needs the greatest training data. Using a minimal amount of labeled data at the fine-tuning step reduces the likelihood of overfitting. Other DL systems that employ DBNs to initialize their weights are RNN and CNN. DBN achieves better quality than further classifiers as a result. For output (o), probability $p(\tau)$ is represented in Equations (6) and (7), in which the temperature is described by T.

$$p(\tau) = \frac{1}{1 + e^{\frac{-\tau}{T}}} \tag{6}$$

$$o = \{ \begin{matrix} 0 & with\ probabilityof\ 1 - p(\tau) \\ 1 & with\ probability\ ofp(\tau) \end{matrix} \} \tag{7}$$

Equation depicts linked energy along the neuron formation state s. Bias as well as weight between neurons are indirect by $wt_{l,m}$ and ψ_l, and s_l indicates binary neuron state. The study contains Leaky Relu which is helpful for the activation function given in Equation (8).

$$E(\vec{V}, \vec{K}) = -\sum_{l<m} s_l s_m wt_{l,m} - \sum_l \psi_l s_l \tag{8}$$

The energy function Z is represented by the Equation, which also suggests a visible unit, a hidden unit, and a partition function; still, the traditional loss function is susceptible to outliers. Thus, the limits are set between 0 and 1 and a new loss function is described in Equations (9) and (10).

$$Loss = \frac{1}{N} \sum_i \sum_{c=1}^{U} y_{ic} log K_{ic} \tag{9}$$

$$Loss = -[\prod_i \prod \prod_{c=1}^{U} sigmoid(y_{ic}) logsigmoid(K_{ic})]^{\frac{1}{N}} \tag{10}$$

QNN

The results obtained from I_b^{Seg} and *Fs* are used to train the QNN. The three primary levels of the QNN are inputs, hidden units, and the output layer. The sigmoid function's total multilevel positions are shown as *Ns*. The quantum period for the quantum level is shown as $\emptyset^T$. Mathematically, the sigmoid function is given by Equation (11).

$$out = \frac{1}{N_s} \sum_{r=1}^{N_s} (\frac{1}{1 + exp(-out \pm \phi^r)}) s \tag{11}$$

QNN output referred to as *ORE* 1; DBN output referred to as *SCORE* 2, uncovered to weighted geometric mean revealed in Equation (12). Where, weight $w = w_1, w_2, \ldots\ldots, w_n s$, $w = w_1, w_2, \ldots\ldots w_n$.

$$C^{\text{out}} = (\prod_{i=1}^{n} SCORE_i)^{1/\sum_{i=1}^{n} w_i} \tag{12}$$

In this instance, all 16 classes are supplied as input. As data volumes rise, the batch process goes through steps of training and testing. This patch breaks up the dataset into many smaller parts, which are then delivered to the model individually. As a result, the classification method eventually produces 16 categories (output).

RBM

RBM is a probabilistic energy-based model (EBM) with two architecture layers: a visible layer containing binary stochastic units $v \in \{0,1\}$ and a hidden layer containing binary stochastic units $h \in \{0,1\}^{nh}$. RBMs were originally designed for modeling binary visible variables, but later extended to handle continuous and non-binary data as well. Initially, RBMs were helpful for binary visible variables, generalized to continuous variables, and non-binary. The energy function of the RBM (binary units) was signified as Equations (13) and (14).

$$E(v, h; \theta) = -\sum_{i,j} v_i W_{ij} h_j - \sum_i v_i a_i - \sum_j h_j b_j \tag{13}$$

Where $\theta = \{W, a, b\}$ is parametrized, and sum indices i, j run from 1 to n_v, n_h. The probability distribution in an energy function of RBM model follows the Boltzmann distribution.

$$q(v, h; \theta) = \frac{exp(-E(v, h; \theta))}{Z(\theta)} \tag{14}$$

Where $Z(\theta) = \sum_{v'} \sum_{h'} exp(-E(v', h'; \theta))$ is

normalization constant also called partition function, it was intractable, even for moderate values n_v and n_h in practice.

It was an essential RBM property along binary hidden units, which marginalization over hidden vectors and accomplished systematically. It allowed for describing probability distribution over visible units in the concept of free energy $F(v; \theta)$, as in Equations (15) and (16).

$$q(v;\theta)=\frac{exp\left(-F(v;\theta)\right)}{Z(\theta)},\ Z(\theta)=\sum_{v'} exp\left(-F\left(v';\theta\right)\right) \quad (15)$$

$$F(v;\theta)=-\sum_i v_i a_i-\sum_j In\left(1+exp\left(\sum_i v_i W_{ij}+b_j\right)\right) \quad (16)$$

Where partition function Z was calculated by totaling it in each noticeable configuration in Equation (3).

Although RBMs were initially presented as generative models, which was helpful in classification. RBMs are also seen as a feature extraction method, which was commonly trained to extract features, followed by nourishing to classifier network. In addition, RBMs were used as standalone discriminative models with convolutional deep BM for the classification of electroencephalograms in initial Alzheimer's disease diagnosis, as well as two-layer deep BM demonstrated to classify facial languages in images of thermal infrared.

Model's classes were denoted with a one-hot encoded vector $y=\left(1_{y=k}\right)_{k=1}^{C}$ and considered for RBM as extra noticeable units. Variables were linked to hidden units through matrix U and biases c. Thus, the energy of this model in Equation (17) was like Equation. (13).

$$E(v, h; \theta) = -\sum_{i,j} v_i W_{ij} h_j - \sum_i v_i a_i - \sum_j h_j b_j - \sum_k y_k c_k - \sum_{k,j} y_k U_{kj} h_j \quad (17)$$

Further, it leads to the expression's modification of some conditional probabilities for GS as follows in Equations (18) and (19).

$$p\left(y_k=1\mid h\right)=\frac{exp\left(c_k+\sum_j U_{kj}h_j\right)}{\sum_{k'} exp\left(c_{k'}+\sum_j U_{k'j}h_j\right)} \quad (18)$$

$$p\left(h_j=1\mid yk,\ v\right)=\sigma\left(\sum_i W_{ij}v_i+b_j+U_{kj}\right) \quad (19)$$

Still, a subsequent method has been more fitted in classification which was presented in Equation (20).

$$p\left(y_k=1\mid v\right)=\frac{exp\left(c_k\right)\prod_{j=1}^{n_h}\left(1+exp\left(\sum_i W_{ij}v_i+b_j+U_{kj}\right)\right)}{\sum_{k'} exp\left(c_{k'}\right)\prod_{j=1}^{n_h}\left(1+exp\left(\sum_i W_{ij}v_i+b_j+U_{k'j}\right)\right)} \quad (20)$$

Where the probability of every class of y_k was assumed as an instance v and calculated for the training model, then it maximized the probability of a true class in an image. Therefore, the discriminative objective function was signified in Equation (21).

$$\theta^* = \arg\min -\log\ p(y|v;\theta) \quad (21)$$

Next, gradients of the above objective function compared with parameters were calculated as demonstrated in Equation (22).

$$\frac{\partial_p\left(y_k\mid v;\theta\right)}{\partial\theta}=\sum_j \sigma\left(o_{kj}(v)\right)\frac{\partial_{o_{kj}(v)}}{\partial\theta}-\sum_{j,m}\sigma\left(o_{kj}(v)\right)p\left(y_m\mid v\right)\frac{\partial_{o_{kj}(v)}}{\partial\theta} \quad (22)$$

Where $o_{kj}(v) = b_j + U_{kj} + \sum_i W_{ij} v_i$. Therefore, Equation (22) was used for updating the parameters with SGD when training the models for the classification problem.

Standard metrics were used to evaluate and assess the RBMs classifier; specifically, classification accuracy was one among the most popular metrics that assesses the ratio among total number of instances classified as well as the number of well-classified instances. However, the matrix does not demonstrate when classes were unbalanced. Consequently, the input of each class was weighted with the inverse of its frequency of arrival; this process was known as the balanced version of accuracy.[28,29]

Nonetheless, the Log-loss function is used to set lower probabilities for the class of the given image shown in Equation (23).

$$L_{log}\left(Y,\ P\right)=-\frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{K} y_{i,k}\log p_{i,k} \quad (23)$$

Where Y was a vector of N one-hot signified K classes, so $y_{i,k} = 1$, if instance i resembles class k

as well as 0, and P is the vector of probabilities assigned for every instance in the model.

RESULTS

The proposed HSI segmentation and classification model is implemented on the Sentinel 2 dataset using the MATLAB 2023a software. For assessing the suggested approach, the following metrics are utilized: Accuracy, precision, sensitivity, specificity, false-positive rate (FPR), false-negative rate (FNR), negative predictive value, MCC, and F-measure. Equations for performance measures are articulated in Equations (24-32).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (24)$$

$$Precision = \frac{TP}{TP + FP} \quad (25)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (26)$$

$$Specificity = \frac{TN}{TN + FP} \quad (27)$$

$$FPR = \frac{FP}{FP + TN} \quad (28)$$

$$FNR = \frac{FN}{FN + TP} \quad (29)$$

$$NPV = \frac{TN}{TN + FN} \quad (30)$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (31)$$

$$F - measure = 2 \times \frac{prescision \times recall}{precision + recall} \quad (32)$$

Where FP means False Positive, FN signifies False Negative, TP symbolizes True Positive, and TN indicates True Negative.

Performance Analysis

Table 1 provides the performance analysis of classification models. Table 1 clearly shows that the proposed RRCNN-EM model achieved better performance than existing CNN, DBN, QNN, and RBM in terms of accuracy, precision, specificity, and sensitivity.

Statistical Analysis

Here, the statistical analysis of several classification models such as CNN, DBN, QNN, and RBM is accomplished for the HSI Segmentation and Classification model. Table 2 depicts the statistical study over U-Net with CNN, DBN, QNN, and RBM. Now, the comparison is made between the performance of U-Net in the proposed BOISO and other models, namely CNN, DBN, QNN, and RBM. The stochastic nature of metaheuristic algorithms requires running the algorithm multiple times for better and more accurate results. Here, the EM model attains a high fitness of 0.92 at the mean condition, CNN, DBN, QNN as well as RBM gain moderately fewer NN fitness of 0.88, 0.91, 0.89, and 0.89, respectively. In case of the maximal situation, the EM acquired higher value by 0.94 which is higher than that of CNN, DBN, QNN, and RBM. Higher HIS classification is attained when HC is collectively working with diverse features as well as RRCNN segmentation.

Classifier Comparison

Table 3 represents an analysis of classifiers in DNN, DBN, QNN, RBM, and CNN. In Table 3, the advanced model achieves higher accurateness having lesser false rates.

Table 1: Performance analysis of classification models

Methods	Accuracy	Precision	Specificity	Sensitivity
CNN	0.91	0.78	0.82	0.83
DBN	0.92	0.87	0.85	0.91
QNN	0.89	0.84	0.88	0.84
RBM	0.80	0.82	0.85	0.89
Proposed RRCNN-EM	0.94	0.91	0.95	0.93

CNN: Convolutional neural networks, DBN: Deep belief network, QNN: Quantum neural network, RBM: Restricted Boltzmann machine, RRCNN-EM: Recurrent residual convolutional neural network with ensemble modeling

Table 2: Statistical analysis for various classification methods

Methods	Standard deviation	Median	Mean	Minimum	Maximal
CNN	0.07	0.88	0.88	0.89	0.93
DBN	0.09	0.90	0.91	0.89	0.93
QNN	0.1	0.91	0.89	0.89	0.92
RBM	0.07	0.91	0.89	0.89	0.93
EM	0.03	0.92	0.92	0.9	0.94

CNN: Convolutional neural networks, DBN: Deep belief network, QNN: Quantum neural network, RBM: Restricted Boltzmann machine, EM: Ensemble modeling

Specifically, precision of the ensemble model is 0.94, which is higher than DNN with 0.91, DBN with 0.83, QNN with 0.85, RBM with 0.81, and CNN with 0.86. The FNR, Sensitivity, MCC, and F-measure of the ensemble model are 0.11, 0.92, 0.92, and 0.92.

Ablation Study for HSI Classification

Table 4 determines the ablation study of the proposed work along RRCNN-RBM, RRCNN-DBN, and RRCNN-QNN also extracted features. From Table 4, the proposed RRCNN-EM method achieved a higher precision of 0.92, while the other combinations such as RRCNN-RBM are 0.84, RRCNN-QNN attains a precision of 0.86, and RRCNN-DBN attains a precision of 0.85. Likewise, the FNR of model effort with RRCNN-QNN, RRCNN-DBN, and RRCNN-EM methods is 0.40, 0.37, and 0.11, respectively. Higher accuracy on HIS classification is attained by collectively working with diverse features as well as optimized RRCNN segmentation.

Table 3: HSI classification has diversified classifiers' evaluation

Parameters	DNN	DBN	QNN	RBM	CNN	Ensemble models
Sensitivity	0.88	0.84	0.89	0.83	0.82	0.92
F-measure	0.87	0.83	0.88	0.88	0.83	0.92
FPR	0.21	0.19	0.28	0.20	0.20	0.13
Precision	0.91	0.83	0.85	0.81	0.86	0.94
NPV	0.90	0.89	0.83	0.88	0.91	0.92
Specificity	0.89	0.91	0.82	0.90	0.83	0.93
MCC	0.83	0.84	0.82	0.83	0.82	0.92
Accuracy	0.84	0.86	0.82	0.85	0.81	0.94
FNR	0.12	0.29	0.42	0.25	0.20	0.11

HSI: Hyperspectral imaging, FPR: False-positive rate, FNR: False-negative rate, NPV: Negative predictive value, MCC: Matthews correlation coefficient

Segmentation Analysis of RRCNN and Conventional Methods

Table 5 defines the segmentation accuracy of the proposed RRCNN-EM technique with respect to Jaccard as well as dice coefficients. The dice as well as Jaccard values are higher for the advanced model. It is contradicted by Table 6.

Time Analysis

Table 5 demonstrates the analysis of time in the proposed RRCNN-EM technique in seconds. The time utilized by the RRCNN-EM algorithm for running will be less compared to CNN, DBN, QNN as well as RBM schemes.

Comparative Analysis

Table 7 provides the comparative analysis of classification models which depicts that the proposed RRCNN-EM model achieved superior performance than existing Multiple DL,[18] Dual 1DCNN,[19] and LSTM-Softmax layer.[21] Table 7 clearly shows that the proposed RRCNN-EM has obtained higher accuracy of 0.94 which is better than the existing Multiple DL,[18] Dual 1DCNN,[19] and LSTM-Softmax layer[21] which have 0.93, 0.86, and 0.93, respectively.

DISCUSSION

This research employs a new RRCNN-EM for HSI segmentation and classification using the Sentinel-2 dataset. The Weiner filter is first applied to the input HSI pre-processing. After segmenting the image with an RRCNN model, the LGTP, VI, and GLCM features are extracted. In addition,

Table 4: Ablation study for HSI classification

Parameters	Extracted features	RRCNN-RBM	RRCNN-QNN	RRCNN-DBN	RRCNN-EM
Accuracy	0.86	0.89	0.87	0.88	0.94
Precision	0.88	0.84	0.86	0.85	0.92
MCC	0.86	0.85	0.85	0.84	0.92
FPR	0.27	0.28	0.26	0.26	0.11
F-measure	0.89	0.83	0.90	0.82	0.92
Specificity	0.79	0.89	0.80	0.87	0.92
FNR	0.34	0.40	0.37	0.41	0.11
Sensitivity	0.88	0.81	0.89	0.80	0.92
NPV	0.85	0.82	0.86	0.81	0.92

HSI: Hyperspectral imaging, NPV: Negative predictive value, MCC: Matthews correlation coefficient, FPR: False-positive rate, FNR: False-negative rate

Table5: Time analysis of proposed RRCNN-EM over classification models

Methods	Time in seconds
CNN	2847.893
DBN	2903.673
QNN	2687.939
RBM	2890.678
RRCNN-EM	2593.426

CNN: Convolutional neural networks, DBN: Deep belief network, QNN: Quantum neural network, RBM: Restricted Boltzmann machine, RRCNN-EM: Recurrent residual convolutional neural network with ensemble modeling

Table 6: Segmentation analysis of RRCNN as well as conventional methods

Methods	Jaccard coefficient	Dice coefficient
CNN	0.83	0.82
DBN	0.81	0.84
QNN	0.83	0.86
RBM	0.84	0.85
RRCNN-EM	0.89	0.89

CNN: Convolutional neural networks, DBN: Deep belief network, QNN: Quantum neural network, RBM: Restricted Boltzmann machine, RRCNN-EM: Recurrent residual convolutional neural network with ensemble modeling

Table 7: Comparative analysis of the proposed method

Models	Accuracy
Multiple DL[23]	0.93
Dual-1DCNN[24]	0.86
LSTM-Softmax layer[26]	0.93
Proposed RRCNN-EM	0.94

DL: Deep learning, RRCNN-EM: Recurrent residual convolutional neural network with ensemble modeling

classification is carried out by an ensemble model, which includes methods such as DBN, QNN, and RBM. As a result, the overall classification outcome is estimated by the weighted geometric mean of both models. From the overall analysis, the sensitivity of the CNN, DBN, QNN, and SO models is minimal, i.e., 0.85, 0.84, 0.89, and 0.87, respectively, while EM shows a sensitivity of 0.939. The chosen procedure in 90% LP achieves its lowest FNR of 0.082768. The proposed scheme has an FPR of 0.13, while CNN, DBN, QNN, and RBM have the FPRs of 0.21, 0.17, 0.22, and 0.25, respectively. In the ablation study, the RRCNN-EM achieves a precision of 0.92; the RRCNN-RBM achieves a precision of 0.84, the RRCNN-QNN achieves a precision of 0.86, the RRCNN-DBN achieves a precision of 0.85, and the extracted features achieve a precision of 0.88, respectively.

CONCLUSION

This study presents an RRCNN framework for HSI segmentation, specifically targeting road surface analysis to support automotive safety applications. Preprocessing with the Wiener filter effectively mitigates noise in HSI, ensuring that critical features such as cracks, potholes, water accumulation, and ice patches are preserved. The RRCNN model segments the processed images, maintaining high-dimensional spatial and spectral information essential for detecting road hazards. Subsequent feature extraction using GLCM, Vegetation Indices (VI), and LGTP enhances the reliability and precision of hazard detection. An ensemble of DL classifiers, including RBM, QNN, and DBN, further improves classification accuracy. Simulation results demonstrate the framework's strong performance, achieving an overall accuracy of 0.94, with dice coefficient and Jaccard index values of 0.89, outperforming existing methods.

While this study employs Sentinel-2 satellite imagery, we acknowledge that vehicle-mounted or road-level hyperspectral sensors could provide higher spatial resolution and context more directly relevant to automotive safety. The proposed methodology is designed to be adaptable to such datasets, enabling near real-time hazard detection in dynamic driving conditions. Integrating this approach into driver-assistance systems, predictive maintenance alerts, or traffic management platforms could directly enhance vehicle safety and accident prevention by providing actionable information on road conditions.

Future work will focus on validating the framework with vehicle-level hyperspectral or autonomous vehicle sensor datasets, integrating additional segmentation models such as U-Net variants, and optimizing the system for near real-time deployment to strengthen its applicability in real-world automotive safety systems.

REFERENCES

1. Zade N, Gupte A, Gupta P, Detalle N, Mannion A, Voyle R. Spectral Feature Extraction and Ensemble Learning for Multiclass aircraft Damage Identification. Amsterdam: Elsevier; 2025.
2. Shah IA, Li J, Glavin M, Jones E, Ward E, Deegan B. Hyperspectral Imaging-Based Perception in Autonomous Driving Scenarios: Benchmarking Baseline Semantic Segmentation Models. [arXiv Preprint]; 2024.
3. Gutiérrez-Zaballa J, Basterretxea K, Echanobe J, Martínez MV, Martínez-Corral U, Carballeira OM, *et al*. On-chip hyperspectral image segmentation with

fully convolutional networks for autonomous driving. J Comput Appl Math 2023;405:113-24.
4. Sattar S, Li S, Chapman M. Developing a near real-time road surface anomaly detection approach for road surface monitoring. Measurement 2021;185:109990.
5. Valme D, Gallindos J, Liyanage D. Road condition estimation using deep learning with hyperspectral images: Detection of water and snow. Proc Eston Acad Sci 2024;73:77-91.
6. Choi JG, Kong CW, Kim G, Lim S. Car crash detection using ensemble deep learning and multimodal data from dashboard cameras. Expe Syst Applic 2021;165:113-23.
7. Paoletti ME, Haut JM, Plaza J, Plaza A. Deep learning classifiers for hyperspectral imaging: A review. ISPRS J Photogram Remote Sens 2019;158:279-317.
8. Shah IA, Li J, Brophy T, Glavin M, Jones E, Ward E, Deegan B. Multi-Scale Spectral Attention Module-based Hyperspectral Image Segmentation for Autonomous Driving [arXiv preprint]; 2025.
9. Zhang Y, Ding R, Shi H, Liu J, Yu Q, Cao G, *et al*. Ensemble network-based distillation for hyperspectral image classification with label noise. Remote Sens 2024;16:4247.
10. Sattar S, Li S, Chapman M. Developing a near real-time road surface anomaly detection system using hyperspectral imaging. Autom Constr 2021;130:103-15.
11. Huang X, Zhou Y, Yang X, Zhu X, Wang K. Ss-tmnet: Spatial-spectral transformer network with multi-scale convolution for hyperspectral image classification. Remote Sens 2023;15:1206.
12. Ullah F, Long Y, Ullah I, Khan RU, Khan S, Khan K, *et al*. Deep hyperspectral shots: Deep snap smooth wavelet convolutional neural network shots ensemble for hyperspectral image classification. IEEE J Sel Top Appl Earth Obs Remote Sens 2023;17:14-34.
13. Chhapariya K, Buddhiraju KM, Kumar A. A deep spectral-spatial residual attention network for hyperspectral image classification. IEEE J Sel Top Appl Earth Obs Remote Sens 2024;17:15393-406.
14. Wang J, Ren J, Peng Y, Shi M. Spectral segmentation multi-scale feature extraction residual networks for hyperspectral image classification. Remote Sens 2023;15:4219.
15. Zhang X, Wang Z. Spatial proximity feature selection with residual spatial-spectral attention network for hyperspectral image classification. IEEE Access 2023;11:23268-81.
16. Ge H, Wang L, Liu M, Zhao X, Zhu Y, Pan H, *et al*. Pyramidal multiscale convolutional network with polarized self-attention for pixel-wise hyperspectral image classification. IEEE Trans Geosci Remote Sens 2023;61:1-18.
17. Wang X, Sun L, Lu C, Li B. A novel transformer network with a CNN-enhanced cross-attention mechanism for hyperspectral image classification. Remote Sens 2024;16:1180.
18. Li Q, Tian J, Tian Q. Deep learning application for crop classification via multi-temporal remote sensing images. Agriculture 2023;13:906.
19. Mao M, Zhao H, Tang G, Ren J. In-season crop type detection by combing Sentinel-1A and Sentinel-2 imagery based on the CNN model. Agronomy 2023;13:1723.
20. Mohammadi S, Belgiu M, Stein A. Improvement in crop mapping from satellite image time series by effectively supervising deep neural networks. ISPRS J Photogramm Remote Sens 2023;198:272-83.
21. Khan HR, Gillani Z, Jamal MH, Athar A, Chaudhry MT, Chao H, *et al*. Early identification of crop type for smallholder farming systems using deep learning on time-series sentinel-2 imagery. Sensors 2023;23:1779.
22. Kaur R, Tiwari RK, Maini R, Singh S. A framework for crop yield estimation and change detection using image fusion of microwave and optical satellite dataset. Quaternary 2023;6:28.
23. Lu Q, Xie Y, Wei L, Wei Z, Tian S, Liu H, *et al*. Extended attribute profiles for precise crop classification in UAV-borne hyperspectral imagery. IEEE Geosci Remote Sens Lett 2024;21:1-5.
24. Reyana A, Kautish S, Karthik PS, Al-Baltah IA, Jasser MB, Mohamed AW. Accelerating crop yield: Multisensor data fusion and machine learning for agriculture text classification. IEEE Access 2023;11:20795-805.
25. Himali PV, Raja RA. Automatic identification of tree species from sentinel-2a images using band combinations and deep learning. IEEE Geosci Remote Sens Lett 2024;21:1-5.
26. Kılıçarslan S, Kılıçarslan S. A comparative study of bread wheat varieties identification on feature extraction feature selection and machine learning algorithms. Eur Food Res Technol 2024;250:135-49.
27. Ya'nan Z, Weiwei Z, Li F, Jianwei G, Yuehong C, Xin Z, *et al*. Hierarchical classification for improving parcel-scale crop mapping using time-series Sentinel-1 data. J Environ Manage 2024;369:122251.
28. Zanella MA, Martins RN, Da Silva FM, Carvalho LC, De Carvalho Alves M, Rosas JT. Coffee yield prediction using high-resolution satellite imagery and crop nutritional status in Southeast Brazil. Remote Sens Appl Soc Environ 2024;33:101092.
29. Gupta P, Gupta P, Singh SK, Thakur B, Kumar M. Ensemble methods-based comparative study of Landsat 8 operational land imager (OLI) and sentinel 2 multi-spectral images (MSI) for smart farming crop classification. Multimedia Tools Appl 2024;84:18025-68.