ConvNeXt-based Multi-Class Hydrocarbon Spill Classification in Hyperspectral Imagery

Milap Patel, Tathya Patel, Anuja Nair, *Member, IEEE*, Tarjni Vyas, Shivani Desai,
Sudeep Tanwar, *Senior Member, IEEE*
Department of Computer Science and Engineering, School of Technology, Nirma University, Ahmedabad, Gujarat, India
Emails: 22bce186@nirmauni.ac.in, 22bce352@nirmauni.ac.in, anuja.nair@nirmauni.ac.in,
tarjni.vyas@nirmauni.ac.in, shivani.desai@nirmauni.ac.in, sudeep.tanwar@nirmauni.ac.in

***Abstract*—This paper proposes a new approach of hydrocarbon spill detection using hyperspectral imaging (HSI) and fine-tuning ConvNeXt convolutional neural network (CNN). Hydrocarbon spill hyperspectral dataset (HSHD) containing 124 HSIs into four classes-cleans, gasoline, motor oil, and thinner is used in the training as well as testing phase. To overcome the computational complexity associated with the high spatial dimensions of HSIs ($1024 \times 1024 \times 20$), instead of resizing, each image is divided into 16 smaller patches of size $256 \times 256 \times 20$ to ensure that no critical spatial-spectral information is lost. The ConvNeXt model is adapted for 20 spectral channels and has its classification head modified for multi-class prediction. This patch-based approach, coupled with the model's spectral-spatial learning capabilities, allows for accurate classification with minimal misclassifications, as shown by the confusion matrix. The proposed framework underlines the efficacy of deep learning (DL) in hyperspectral data analysis, offering significant advantages for environmental monitoring and rapid hydrocarbon spill identification.**

Index Terms*—Hydrocarbon spill, ConvNeXt, CNN, spectral-spatial, hyperspectral imaging

I. INTRODUCTION

Oil spill and hydrocarbon spill is one of the significant environmental challenges, causing adverse and harmful effects on marine ecosystems and coastal communities. Such a release of hydrocarbons into water not only puts aquatic organisms at risk but impacts human activities related to these ecosystems. Oil spill detection, which in the past could only be based on visual analysis or restricted by remote sensing technologies, may be inefficient and time-consuming, as it might not even guarantee the exact nature of the oil involved. On the other hand, hyperspectral imaging (HSI) is advantageous since it captures more spectral information with high spatial coverage in various wavelength ranges; the data is thus useful in distinguishing and determining types of oils. This speeds up and becomes effective in terms of responding to oil spills [1].

HSI are rich information sources, which capture spectral data over hundreds of contiguous wavelength bands and go beyond the visible spectrum to the near-infrared and other regions. Traditional RGB images are a representation of only three bands-red, green, and blue-but in the case of HSIs, the spectral signature is highly detailed for each pixel in the spatial domain. This high spectral resolution allows the differentiation and identification of materials and substances through the patterns they give off in their unique reflectance characteristics, making HSIs extremely valuable in a number of applications, including environmental monitoring, mineral exploration, and chemical spill detection. HSI taps both the spectral and spatial information available and thus provides powerful means for accurate classification and analysis in complex real-world situations.

Many researchers have worked in the field of hydrocardon spill classification over the years. Sandhiya *et al.* [2] proposed an application of machine learning (ML) techniques for oil spill detection using satellite and drone imagery. They highlighted the significant role that the synthetic aperture radar (SAR) technology plays in identifying and monitoring oil pollution in marine environments. They trained ML methods like support vector machine (SVM), decision tree (DT), linear regression (LR) on labeled datasets comprising images of clean water and oil-contaminated water. Moving further, Sherif *et al.* [3] implemented a deep learning (DL) approach using an artificial neural network (ANN). The model was trained on processed satellite image data, employing techniques like gradient descent to minimize prediction errors. Bui *et al.* [4] presented another solution with data augmentation and attention mechanism. They used a tailored data augmentation strategy leveraging a conditional generative adversarial network (GAN), specifically the Pix2Pix framework was implemented to generate images that would mimic real oil spills to enhance the diversity of the training dataset. Then, a dual attention mechanism-based DL model was employed, integrating spatial and channel attention modules to boost oil spill classification accuracy. Yang *et al.* [5] studied high spectral resolution from HSI data and thermal infrared's sensitivity to temperature differences for identifying oil types. They focused on crude oil, emulsions, and refined products, collecting data using airborne HSI sensors and thermal cameras. The study employed SVM, RF, and convolutional neural network (CNN), for classification and found that combining modalities improved recognition accuracy.

Based on the research gaps found in state-of-the-art approaches, this paper proposes a new approach of hydrocarbon spill detection using HSI and fine-tuning ConvNeXt CNN. Hydrocarbon spill hyperspectral dataset (HSHD) containing 124 HSIs into four classes-cleans, gasoline, motor oil, and thinner is used in the training as well as testing phase. To overcome the computational complexity associated with the high spatial dimensions of HSIs ($1024 \times 1024 \times 20$), instead of resizing, each image is divided into 16 smaller patches of size $256 \times 256 \times 20$ to ensure that no critical spatial-spectral information is lost. The

ConvNeXt model is adapted for 20 spectral channels and has its classification head modified for multi-class prediction. This patch-based approach, coupled with the model's spectral-spatial learning capabilities, allows for accurate classification with minimal misclassifications, as shown by the confusion matrix. The proposed framework underlines the efficacy of DL in HSI data analysis, offering significant advantages for environmental monitoring and rapid hydrocarbon spill identification.

A. Motivation

1) Oil spills or hydrocarbon spills can severely damage marine ecosystems or agricultural lands sometimes. Early detection using HSI is important for quick response and minimal environmental damage.
2) Traditional detection methods do not have the accuracy and spectral resolution required in changing conditions. HSI provides superior accuracy in the identification and quantification of oil spills.
3) Recent developments in satellite and HSI sensor technology improve remote sensing ability, covering huge areas with oil spill monitoring very efficiently, so it finds usage in remedial actions.

B. Research Contributions

This paper has the following research contributions.

1) We utilise the HSHD dataset, which provides detailed HSI data to enhance the model's ability to accurately classify oil spills.
2) We propose a DL-based approach for detecting oil and hydrocarbon spills by fine-tuning various state-of-art CNN architectures, leveraging their strengths in feature extraction from HSI.
3) We demonstrate the model's predictive accuracy using various performance metrics, such as accuracy, precision, recall and F1-score, to evaluate its results .

C. Paper Organization

The paper is further organized as follows: Section II proposes the system model and the problem formulation, Section III explains the proposed framework, Section IV discusses the results, and Section V provides the conclusion and future work.

II. System Model and Problem Formulation

A. System Model

The system model of the proposed framework is illustrated in Fig. 1. The detection of oil spills using HSI data leverages advanced remote sensing and DL technologies. The system operates through a multi-step pipeline, starting with data acquisition via HSI sensors mounted on satellites and drones. Let $\mathcal{X} \subseteq \mathbb{R}^{H\times W\times C}$ represent the set of HSIs captured, where H and W denote the height and width of the images, respectively, and C is the number of spectral channels. Each image is segmented into smaller patches $\{P_i\}$, where $P_i \in \mathbb{R}^{h\times w\times C}$, with h and w representing the height and width of the patches, respectively. This segmentation facilitates localized analysis and reduces the computational overhead associated with processing large images. After data acquisition, the HSI data is transmitted to centralized data centers via secure communication channels for preprocessing and analysis. Let $\mathcal{X}' = \{P_i'\}$ denote the set of preprocessed patches. Each patch is then fed into the DL model for predictive analysis. The processed insights are subsequently communicated to relevant stakeholders, such as disaster management teams and local rescue operators, enabling timely remedial actions. The primary objective of the system is to classify each patch $\{P_i\}$ into its respective class $y_i \in \mathcal{Y}$. A DL-based approach is employed, fine-tuning various CNN architectures for feature extraction and classification. The models are optimized using the AdamW optimizer and CrossEntropyLoss function, with performance metrics such as accuracy, precision, recall, and F1-score used to evaluate their effectiveness.

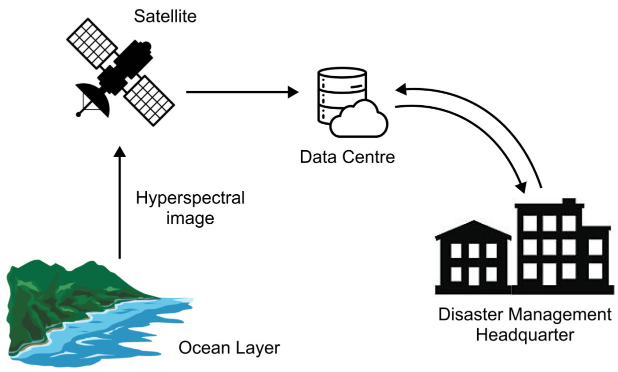

Fig. 1: System model.

B. Problem Formulation

Let the hyperspectral dataset be as follows:

$$\mathcal{D} = \{(P_i', y_i)\}_{i=1}^{N}, \tag{1}$$

where N is the total number of patches, $P_i' \in \mathbb{R}^{h\times w\times C}$ represents the i^{th} preprocessed patch of spatial dimensions $h\times w$ with C spectral channels, and $y_i \in \mathcal{Y}$ is its corresponding class label. The dataset comprises patches derived from high-resolution HSIs, where each patch inherits the label of its parent image, ensuring the preservation of spatial and spectral information. The objective is to train a DL model f_θ, parameterized by θ, that maps an input patch P_i' to its predicted label $\hat{y}_i$. Mathematically, this can be expressed as:

$$f_\theta(P_i') = \hat{y}_i, \quad \forall i \in \{1, 2, \ldots, N\}, \tag{2}$$

where $\hat{y}_i \in \mathcal{Y}$ is the predicted label for the i^{th} patch. The training process involves minimizing the CrossEntropyLoss $\mathcal{L}$, which measures the discrepancy between the true labels y_i and the predicted probabilities for each class. The loss function is defined as:

$$\mathcal{L}(\theta) = -\frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{K} 1(y_i = k)\log(p_\theta(y_i = k \mid P_i')) \tag{3}$$

where:

- K is the total number of classes,
- $p_\theta(y_i = k \mid P_i')$ represents the predicted probability for class k,
- $1(\cdot)$ is an indicator function that evaluates to 1 if $y_i = k$, otherwise 0.

The primary objective is to optimize the model parameters θ such that the loss function $\mathcal{L}(\theta)$ is minimized. This optimization problem can be formally expressed as:

$$\theta^* = \arg\min_{\theta} \mathcal{L}(\theta), \tag{4}$$

where θ^* represents the set of parameters that minimizes the average loss across all training samples.

By minimizing $\mathcal{L}(\theta)$, the model learns to accurately capture the spatial and spectral patterns within the HSI data, enabling robust classification across all classes. The patch-based approach ensures that the large spatial dimensions of the original images do not impose excessive computational burdens while preserving critical spatial-spectral details necessary for accurate classification. Additionally, the use of CrossEntropyLoss ensures the model effectively handles imbalances in class distributions, promoting stable and efficient learning.

III. The Proposed Framework

The proposed framework, as demonstrated in Fig. 2. It is a three-layered approach containing the data collection layer, intelligent layer, and application layer.

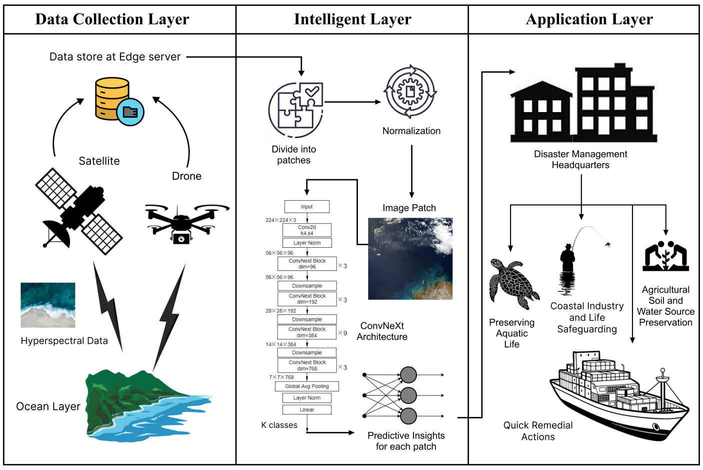

Fig. 2: Proposed framework.

A. Data Collection Layer

The data collection layer forms the base of the framework, using satellite and drone platforms equipped with HSI cameras to collect raw data in a number of electromagnetic spectra. HSI is fundamental in oil spill detection, since it provides detailed spectral information that differentiates oil from water, vegetation, and other substances due to its high spectral resolution. Recent satellite missions, such as ESA's Copernicus program with Sentinel satellites and the upcoming EnMAP, provide high-resolution HSI data over large ocean areas. The drones complement these satellites, being flexible and of higher spatial resolution, thus enabling fast, localized data acquisition with advanced HSI cameras. These are then transmitted to the centralized data center via satellite links or through 4G/5G networks, depending on the location of the platform. The infrastructure will provide real-time data ingestion, preprocessing, and storage of data, readying the data for further analysis by the intelligence layer for effective oil spill classification.

B. Intelligence Layer

1) Data Preprocessing: The HSHD [8] is a specific set of 124 HSIs to classify four different classes: clean samples (uncontaminated), and samples contaminated with gasoline, motor oil, or thinner. Every HSI has a resolution of $1024 \times 1024 \times 20$, where the first two dimensions are the spatial resolution, and the third dimension is the spectral channels. The dataset is stored in standard ENVI format, where each sample will have two accompanying files, such as the header metadata (.hdr) that provides wavelength information as well as the spatial resolution; the .dat file will carry the HSI data cube. To load images, Spectral Python (SPy) [9] was used: it enables transformation of ENVI files into appropriate numpy arrays which are more tractable for calculation.

Once loaded, the images undergo normalization along the spatial dimensions to ensure consistent scaling and facilitate convergence during model training. Due to the large spatial resolution of 1024×1024, training on full-sized images is computationally expensive and risks exceeding available hardware memory. However, directly reducing the spatial dimensions through downsampling can degrade performance by discarding critical spatial information, which is essential for accurate classification. To address this, each HSI is split into 16 smaller patches of $256 \times 256 \times 20$, with each patch inheriting the same label as the original image as shown in Fig. 3. This method does not lose any spatial information but increases the size of the training dataset, which helps mitigate overfitting by exposing the model to more variations during training. Lastly, these processed images are converted to tensors so that they can be used to train the models implemented in the PyTorch deeplearning framework . This preprocessing pipeline balances computational feasibility, meanwhile retaining rich spatial and spectral details required for optimal training of the ConvNeXt model.

2) Model Architecture and Training: The ConvNeXt model [10] was adapted and fine-tuned to classify HSIs into four classes: clean, gasoline, motor oil, and thinner. HSIs with 20 spectral channels were input to the model, necessitating modifications to the initial convolution layer to handle the 20-channel input. This was accomplished by replacing the stem cell with a 4×4 convolution with a stride of 4, which means it downsamples the input dimensions by a factor of 4×4 and projects the 20 channels into 96 feature maps. The stages in ConvNeXt consist of depthwise convolutions for spatial feature extraction, inverted bottleneck blocks with an expansion ratio of 4, and 1×1 convolutions for channel mixing. It represents

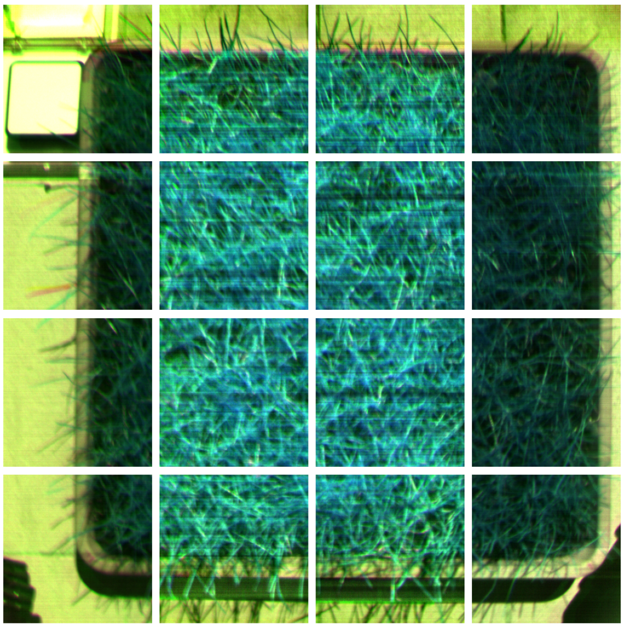

Fig. 3: A sample image of type clean (no contamination) from the dataset. The 11th, 8th and 4th band were selected as the red, green and blue components respectively, to display the HSI as RGB image. The image is divided into 16 patches.

the spatial feature extraction at every stage as follows:

$$Y_{i,h,w} = \sum_{u=1}^{k_h} \sum_{v=1}^{k_w} W_{i,u,v} X_{i,h+u,w+v} + b_i, \tag{5}$$

where k_h and k_w denote the kernel dimensions, W and b are the learnable weights and biases, and h, w index spatial dimensions. Depthwise convolutions separately process each input channel, reducing computational overhead.

ConvNeXt incorporates inverted bottleneck blocks, where the hidden dimension of the multi-layer perceptron (MLP) block is expanded by a factor of 4:

$$H' = \text{GELU}(W_1 H + b_1), \quad Y = W_2 H' + b_2, \tag{6}$$

where $W_1 \in \mathbb{R}^{d' \times d}$, $W_2 \in \mathbb{R}^{d \times d'}$, and d' is the expanded dimension. Layer normalization (LN) replaces batch normalization (BN) to stabilize training:

$$\text{LN}(X) = \frac{X - \mu}{\sqrt{\sigma^2 + \epsilon}} \cdot \gamma + \beta, \tag{7}$$

where μ and σ^2 are the mean and variance of the input features, and γ, β are learnable parameters. Gaussian error linear unit (GELU) activation further smooths nonlinear transformations:

$$\text{GELU}(x) = x \cdot \Phi(x), \quad \Phi(x) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{x}{\sqrt{2}}\right)\right]. \tag{8}$$

The classification head was adjusted to output logits for the four target classes via a fully connected layer, minimizing the cross-entropy loss:

$$\mathcal{L} = -\frac{1}{N} \sum_{i=1}^{N} \sum_{c=1}^{4} \mathbb{I}[y_i = c] \log p(c|x_i), \tag{9}$$

where $p(c|x_i)$ denotes the predicted probability for class c, and $\mathbb{I}[y_i = c]$ is an indicator function. The model was trained for 20 epochs using the AdamW optimizer with a weight decay of 10^{-2}. The learning rate was initialized at 10^{-4} and reduced by a factor of 0.1 via a ReduceLROnPlateau scheduler if the test loss did not improve for three consecutive epochs. Early stopping with a patience of 5 epochs was employed to prevent overfitting.

Algorithm 1 Oil Spill Detection using HSI and ConvNeXt

```
Input: Hyperspectral_Dataset
procedure OIL_SPILL_DETECTOR(Hyperspectral_Dataset, ConvNeXt)
Data preprocessing:
Create HyperspectralDataset class with following parameters:
patch_size = 512
image_size = 1024 × 1024 × 20 channels
for each Image ∈ Hyperspectral_Dataset do
    patches ← ExtractPatches(Image, patch_size)
    normalized_patches ← NormalizeSpectrum(patches)
end for
train_dataset, test_dataset ← Split(dataset, [0.7, 0.3])
dataloaders ← CreateDataLoader(batch_size = 16)
define ConvNeXt model:
model ← ConvNeXt_Small()
input_layer ← Conv2d(in_channels=20, out_channels=96)
output_layer ← Linear(out_features=4)
Training Configuration:
optimizer ← AdamW(lr=0.0001, weight_decay=0.01)
scheduler ← ReduceLROnPlateau(patience=3, factor=0.1)
loss_function ← CrossEntropyLoss()
for epoch in range(max_epochs) do
    Training Phase:
    for each batch ∈ train_dataloader do
        predictions ← model(batch)
        loss ← loss_function(predictions, labels)
        Backward propagation and optimize
    end for
    Evaluation Phase:
    for each batch ∈ test_dataloader do
        predictions ← model(batch)
        Calculate metrics (F1, Precision, Recall, AUROC)
    end for
    Update learning rate based on validation loss
    Check early stopping condition (patience = 5)
end for
Return trained model for oil spill classification
```

C. Application Layer

The application layer communicates processed data from the intelligence layer to disaster management authorities, translating HSI insights into actionable responses. Upon identifying oil spills with high precision, real-time information is communicated for the quick coordination of containment and remediation actions. Modern disaster management systems rely on real-time data feeds from environmental monitoring platforms to support dynamic decision-making. Advanced tools, such as geographic information systems (GIS) and decision support systems (DSS), help in visualizing the impacts of oil spills, forecasting environmental and economic damage. From this classification data, disaster management teams can initiate a unified response: deployment of containment booms, skimmers, and dispersants, mobilization of cleanup crews and notify affected stakeholders. This integrated approach, using HSI data and advanced tools, enhances the ability to mitigate environmental damage, protect ecosystems, and reduce economic losses.

IV. Results and Discussion

A. Experimental Setup

The proposed framework is implemented on Kaggle's code editor. For computational purposes, the Tesla T4 TPU is used. It is equipped with 2,560 CUDA cores and 16GB of GDDR6 memory, thus allowing for faster and more efficient processing. We implemented the model using Python version 3.10.14, along with essential libraries such as NumPy version 1.26.4 for numerical computations, Pandas version 2.2.3 for data manipulation and preprocessing, Matplotlib version 3.7.5 for data visualization, PyTorch version 2.4.0 for implementing the DL model, and Scikit-learn version 1.2.2 for additional ML functionalities.

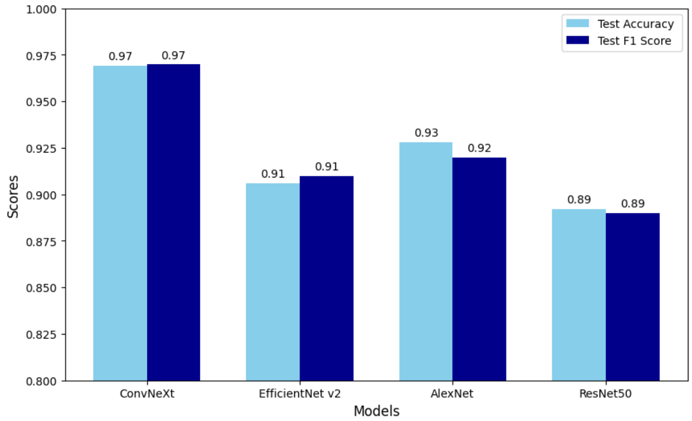

Fig. 4: Test time accuracy and F1-score comparison between different CNN models.

B. Performance Analysis and Discussion

Fig. 4 shows the comparative analysis of the four models demonstrating the best performance of ConvNeXt for the given task of classifying the oil spills from the dataset with a test accuracy of 96.93 % and an F1 score of 0.97, which leads to good generalization with balanced precision and recall, thus performing particularly well in identifying spill and non-spill regions correctly. AlexNet performed well, with a test accuracy of 92.82 % and an F1 score of 0.92, which indicates its capability as a simpler yet reliable architecture. EfficientNet v2 achieved moderate results, with a test accuracy of 90.6 % and an F1 score of 0.91, while ResNet50 showed the lowest performance among the models, with a test accuracy of 89.22% and an F1 score of 0.89. These results point to the superiority of ConvNeXt in the ability to tackle complex patterns in hydrocarbon spill detection, thereby suggesting suitability for deployment in real-world applications.

The superior performance of ConvNeXt can be attributed to its modern architecture that has the advancements from vision transformers (ViT) [11] with simplicity from CNNs. This hybrid design makes it possible for ConvNeXt to capture both global and local features, which are very important to differentiate between hydrocarbon spills and the background noise in high-resolution images. Moreover, the hierarchical feature extraction in the architecture of ConvNeXt is suitable for data with variable spill patterns and sizes. It should be noticed that all models were trained with the same batch size, which was 16 and patch size, which was 256, and other training parameters, as mentioned in III-B2. Thus, comparison should be fair. ConvNeXt's architectural improvements and effective feature representation would have been key factors to justify the superiority by this model.

Then, Fig. 5 describes the training and test performance of

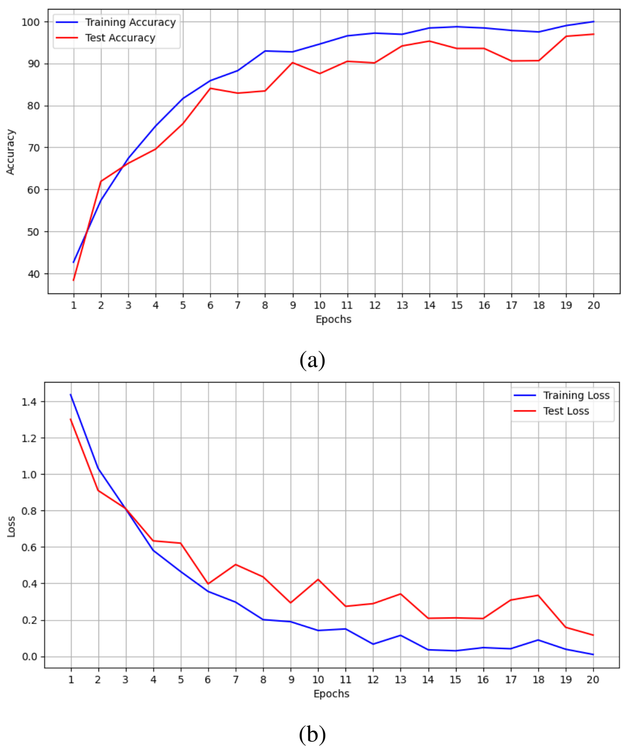

Fig. 5: Training vs test for (a) model accuracy and (b) loss over the epochs for oil spill detection.

the proposed ConvNext model on the given task in terms of accuracy and loss. The reduction in both train and test loss over the epochs (Fig. 5b) shows that the model was able to learn well from the dataset, and the overall training process is smooth. This can be credited to the normalized input images, and weight-regularized AdamW optimizer. The test loss is slightly higher than train loss, indicating that the model is slightly overfitting as the epochs increases. one can deduce the robustness of the model from the increasing accuracy of the model over the epochs (Fig. 5a).

Fig. 6 shows the confusion matrix for the classification

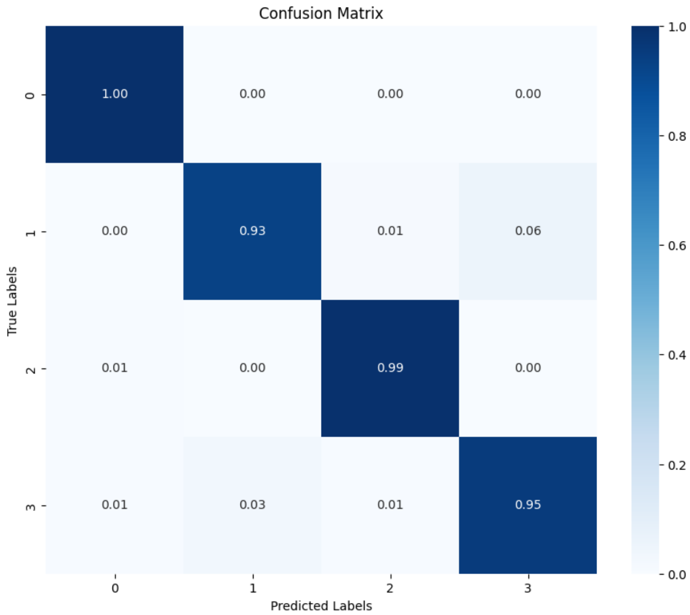

Fig. 6: Confusion matrix for ConvNeXt.

performance of the proposed fine-tuned ConvNeXt model on the HSHD. The model shows an overall high accuracy in all four classes: clean, gasoline, motor oil, and thinner, with strong diagonal values. In particular, the model shows a perfect classification rate for the clean class (class 0) with an accuracy of 100%. Moreover, class 2 with motor oil shows a very good correctness level of 99%, making almost all samples get well-classified. Class 1 with gasoline indicates a clear-cut classification with 93% accuracy, though there is a small portion of misclassification as thinner oil (class 3). Class 3 or thinner stands robustly at 95%, causing some minor misclassifications into the gasoline and motor oil classes, respectively.

As depcted from Fig. 6 the model is very efficient in distinguishing the clean class (class 0) from contaminated samples, which is crucial in practical scenarios for rapid detection of uncontaminated areas. Second, the high accuracy for motor oil (class 2) indicates that the spectral features of this class are distinct enough for the model to classify them with near-perfect precision. However, the slight misclassifications observed between gasoline (class 1) and thinner (class 3) suggest that these two hydrocarbons share some overlapping spectral characteristics, which might make them harder to differentiate. This overlap might be due to similarities in their chemical compositions.

V. Conclusion

In this work, we proposed a deep DL-based framework for hydrocarbon spill detection using HSI data. We leveraged the ConvNeXt CNN adapted to process 20 spectral channels, and classifying four classes.We divided large HSIs into smaller patches instead of resizing them. This approach ensured that critical spatial-spectral features were preserved and computational constraints could be efficiently managed. This patch-based method reduces the probability of information loss while enhancing model generalization performance across complex HSI datasets. The proposed approach attained high accuracy classification performance with four classes of hydrocarbon contamination, indicating high robustness and reliability. We showed the robustness of the ConvNext model by comparing it with other state-of-the-art CNN models like EfficientNet-V2, AlexNet, and Resnet50. Thus, the overall findings have highlighted the applicability of fine-tuned CNNs to HSI data analysis with many practical implications to environmental monitoring and disaster response systems. Future work could include integrating additional datasets, domain-specific augmentation techniques, and real-world deployment to enhance the model's applicability and performance.

References

[1] A. Bhargava, A. Sachdeva, K. Sharma, M. H. Alsharif, P. Uthansakul, and M. Uthansakul, "Hyperspectral imaging and its applications: A review," *Heliyon*, vol. 10, no. 12, p. e33208, 2024.

[2] S. S, R. R, S. M, and V. S, "Detection of oil spill events at sea using machine learning," in *2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA)*, pp. 53–58, 2023.

[3] K. Sherif, F. H. Rizk, A. M. Zaki, M. M. Eid, N. Khodadadi, A. Ibrahim, A. A. Abdelhamid, L. Abualigah, and E.-S. M. El-Kenawy, "Revolutionizing oil spill detection: A machine learning approach for satellite image classification," in *2024 International Telecommunications Conference (ITC-Egypt)*, pp. 245–250, 2024.

[4] N. A. Bui, Y. Oh, and I. Lee, "Oil spill detection and classification through deep learning and tailored data augmentation," *International Journal of Applied Earth Observation and Geoinformation*, vol. 129, p. 103845, 2024.

[5] J. Yang, Y. Hu, J. Zhang, Y. Ma, Z. Li, and Z. Jiang, "Identification of marine oil spill pollution using hyperspectral combined with thermal infrared remote sensing," *Frontiers in Marine Science*, vol. 10, 2023.

[6] J. M. Haut, S. Moreno-Alvarez, R. Pastor-Vargas, A. Perez-Garcia, and M. E. Paoletti, "Cloud-based analysis of large-scale hyperspectral imagery for oil spill detection," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, pp. 2461–2474, 2024.

[7] S. Jia, S. Jiang, Z. Lin, N. Li, M. Xu, and S. Yu, "A survey: Deep learning for hyperspectral image classification with few labeled samples," *Neurocomputing*, vol. 448, pp. 179–204, 2021.

[8] D. Rivas-Lalaleo and C. Hernandez, "Hydrocarbon spill hyperspectral dataset (hshd," 2024.

[9] Spectral Python Development Team, "Spectral python (spy)." Accessed: 2025-01-19.

[10] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A convnet for the 2020s," 2022.

[11] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, "An image is worth 16x16 words: Transformers for image recognition at scale," 2021.