HBO-DEVIT: VISION TRANSFORMER BASED ATTENTION-GUIDED EVOLUTIONARY ARCHITECTURE FOR SHIP-ICEBERG CATEGORISATION IN ARCTIC SAR IMAGES ESTIMATION

ANUVAB SEN[1], SUBHABRATA ROY[1], UDAYON SEN[1], CHHANDAK MALLICK [2], AND SUJITH SAI[3]

[1]IIEST SHIBPUR, [2]JADAVPUR UNIVERSITY AND [3]NIT ROURKELA

ABSTRACT

Remote offshore regions face significant risks from drifting icebergs. Traditional instruments like Unmanned Aerial Vehicles (UAV) are impractical due to harsh weather conditions. Utilizing Synthetic Aperture Radar (SAR) imagery for iceberg monitoring is a viable solution, derived from dual-polarized Sentinel-1. This paper proposes a novel Hybrid Bayesian Optimized Differential Evolution based Vision Transformer (HBO-DEViT) model for efficiently detecting and categorising ships and icebergs in SAR images. The developed model showed promising results for iceberg and vessel detection, achieving near 99.1% accuracy on the Statoil/C-CORE dataset. This research has a significant impact on navigational safety and possesses broad applicability to address threats in the maritime domain.

INTRODUCTION

For maritime safety and efficiency, detecting ships and icebergs is crucial, which helps to prevent collisions, optimize routes, and improve situational awareness in ice-prone waters. Synthetic Aperture Radar (SAR) serves as an active remote sensing technique, utilizing radar signals to generate high-resolution earth surface images. Its capability to penetrate cloud cover and provide continuous monitoring of marine regions proves indispensable for detecting ships and icebergs even in complex and harsh weather conditions. Till date, the SAR ship detection field can be roughly divided into two development stages: traditional methods and deep learning-based approaches. Traditional methods mainly rely on manually extracted features, but these methods are susceptible to environmental influences. With the gradual maturity of Deep Learning and computer vision, its application fields are increasingly becoming wider [1].

MOTIVATION

The motivation of the work:

- Deep learning based ship detection methods utilizing SAR images have shown outstanding performance and have surpassed traditional methods in performance and universality. Vision Transformers (ViTs) have made remarkable strides in computer vision tasks, particularly in image classification and segmentation. ViTs excel in capturing global contextual information within images. This can be extremely crucial for tasks in complex scenes such as maritime landscapes, where distant relationships between elements are vital for accurate classification [2].

The contributions of the proposed work:

- This paper presents an innovative Hybrid Bayesian Optimized Differential Evolution based Vision Transformer (HBO-DEViT) model, designed to enhance ship-iceberg detection from SAR images [3].

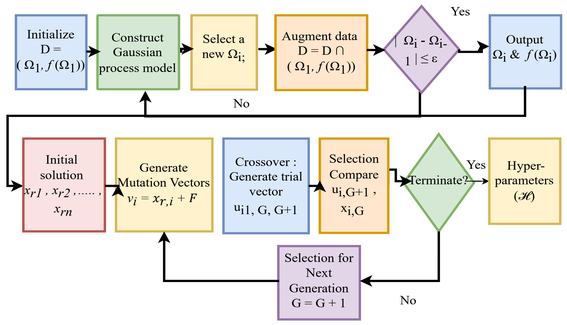


Figure 1: The working mechanism of the HBO-DE algorithm based hyper-parameters selection approach for the arctic SAR images estimation task.

PROPOSED WORK

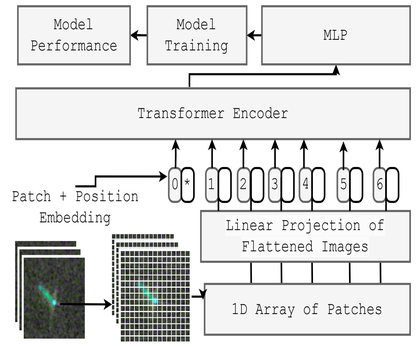


Figure 2: ViT Architecture

The output layer adopts softmax activation :

$$\text{softmax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{n} e^{z_j}} \quad \forall i$$

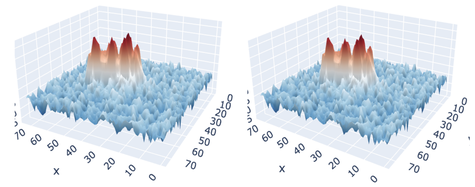
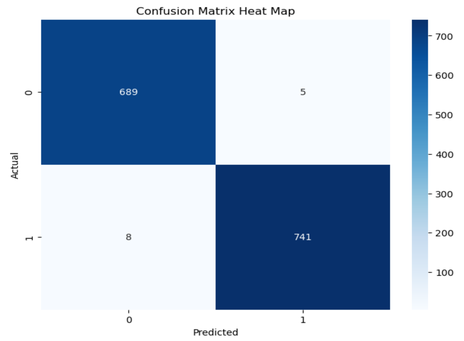


Figure 3: 3-D Surface Plot for Ship: HH, HV1 bands

RESULTS AND CONCLUSION



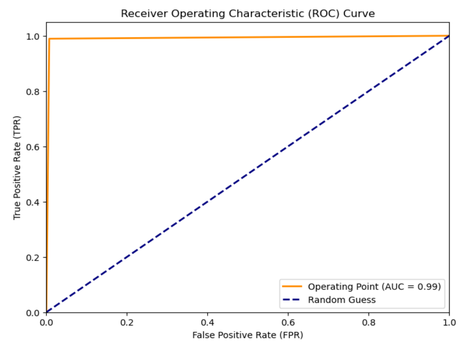


Table 1: REPRESENTATIVE BEST HYPER-PARAMETERS OBTAINED FROM HBO-DE ALGORITHM USED TO TRAIN THE PROPOSED MODEL

$\mathcal{B}$	$\mathcal{E}$	Patch Size	$\mathcal{D}$	η
24	987	5×5	0.1	0.06324

Table 2: COMPARISON OF ACCURACY AND LOG LOSS VALUES WITH EXISTING MODELS

Models	Accuracy (%)	Log Loss
Chen et al. [?]	95.0	0.08
Taylor et al. [?]	95.4	0.22
Maoguo Gong et [?]	88.0	0.28
ViT	94.2	0.27
HBO-DEViT	99.1	0.07

Experimental results showcase that our model stands as a SOTA benchmark in remote sensing, achieving an impressive accuracy of 99.1%, while maintaining low log loss and desirable precision, recall, and F1-score values.

REFERENCES

[1] Jimin Yu, Guangyu Zhou, Shangbo Zhou, and Maowei Qin. A fast and lightweight detection network for multi-scale sar ship detection under complex backgrounds. *Remote Sensing*, 14(1), 2022.

[2] Anuvab Sen, Vedica Gupta, and Chi Tang. Differential evolution algorithm based hyperparameter selection of gated recurrent unit for electrical load forecasting, 2023.

[3] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Weissenborn, et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.