Advanced Deep Learning Driven Geospatial Analysis for GLOF Risk Reduction: A Case Study from Pakistan's Northern Mountain Ranges

Nauman Ali Murad
Faculty of Computer Science
Ghulam Ishaq Khan Institute of Engineering Sciences and Technology
Topi, Pakistan
u2022479@giki.edu.pk

Abinta Mehmood Mir
Faculty of Computer Science
Ghulam Ishaq Khan Institute of Engineering Sciences and Technology
Topi, Pakistan
abinta.mehmood@giki.edu.pk

Nazia Shahzadi
Faculty of Computer Science
Ghulam Ishaq Khan Institute of Engineering Sciences and Technology
Topi, Pakistan
nazia.shahzadi@giki.edu.pk

***Abstract*— Glacial Lake Outburst Floods (GLOFs) pose a severe risk to populations in high-altitude areas, particularly in Pakistan's northern regions, where glacier melt has created 3,044 glacial lakes in Gilgit-Baltistan and Khyber Pakhtunkhwa [1]. The Chitral region is especially vulnerable, highlighting the need for robust monitoring and mitigation strategies. This paper focuses on using Deep Learning (DL) models for detecting and segmenting glacial lakes to mitigate GLOF risks. Using the Glacial Lakes Detection Dataset from High-Mountain Asia, covering 2080.12 km² and including around 30,121 glacial lakes, we selected 1,200 cloud-free true-color images with ground truth masks. Our study evaluates DL models like DeepLabV3+, U-Net, and YOLOv8 for lake segmentation and classification. Experimental results demonstrated model proficiency, with DeepLabV3+ (ResNet50 backbone and Dice loss function) achieving the highest IoU score of 77.2%. These findings highlight the potential of DL-based systems for enhanced GLOF monitoring and early warning.**

***Keywords*—Glacial Lake outburst floods (GLOF), remote sensing, geo-spatial data, climate change, semantic segmentation, satellite imagery, deep neural network**

I. INTRODUCTION

Glacial lakes in northern Pakistan serve as a crucial source of freshwater for the irrigation, drinking, and industrial needs of the local and downstream populations. These lakes are directly impacted by global warming, as evidenced by numerous prior studies [2]. The rising temperatures in the Chitral district have led to increased glacier melt, resulting in continuous damage to houses, lands, orchards, and forests. Glacial lake outbursts, or GLOFs, have emerged as a symbol of climate change in many mountainous regions across the world [3]. GLOFs are high-magnitude, low-frequency events that have profound effects on society and geomorphology.. Pakistan's Ministry of Climate Change has recorded 3,044 glacial lakes, which have increased significantly due to glacier retreat and rising global temperatures [1].

Recent DL advancements have enabled effective feature extraction and image segmentation. This research utilizes U-Net, DeepLabV3+, and YOLOv8-Seg [4] architectures to accurately segment small-scale objects like glacial lakes from satellite data, allowing efficient pixel-level identification and mapping.

II. LITERATURE REVIEW

GLOFs pose substantial risks to communities and infrastructure, necessitating their effective detection and monitoring. GLOFs are typically sudden and rapid events, leading to catastrophic flooding and significant geomorphic and socioeconomic impacts [2]. Remote sensing, combined with advanced DL models, has proven to be a valuable approach for the identification and prediction of GLOFs because of its capacity to precisely process large volumes of data from satellite imagery.

Previous research, such as the study by Basit et al., [5] and Siddique et al., [6] has utilized DL techniques, including Convolutional Neural Networks (CNN) and automated monitoring systems, to monitor glacial lake formations and enhance the accuracy and efficiency of GLOFs detection. These studies have laid the groundwork for the application of DL in glacial lake monitoring, demonstrating its potential in this domain.

In contrast, our study significantly advances this field by conducting extensive experimentation using state-of-the-art DL models, including DeepLabv3+, U-Net, and YOLOv8-Seg. These proposed models are renowned for their amazing performance in image segmentation and object detection tasks, making them well-suited for remote sensing applications. The study's comprehensive methodologies, involving rigorous training and evaluation processes [11], have resulted in improved accuracy and robustness in GLOFs detection.

The proposed models have exhibited great promise in accurately detecting and predicting GLOFs, outperforming earlier approaches. The DeepLabv3+ [8] model excels in semantic segmentation tasks, while U-Net [9] offers robust image segmentation capabilities. Additionally, YOLOv8-Seg [4], with its real-time object detection capabilities, further enhances the ability to promptly and accurately detect GLOFs. The extensive experimentation results underscore the effectiveness of these advanced models in the context of GLOFs monitoring, marking a significant advancement in this research area.

III. STUDY AREA AND DATASET

A. Glacial Lake Map

The provided glacial lake map shown below in Figure 1, delineates the area of study for our research. The map depicts both existing and projected glacial lakes situated in the High Mountain Asia region. The layout image has been published in paper called the Glacial Lake Evolution in High Mountain Asia [3].

B. Dataset Description

The dataset [7] used as a case study, comprises 1200 images and their corresponding ground truth masks, covering a 2080.12 km² area encompassing approximately 30,121 glacial lakes.

This dataset was generously provided by the Remote

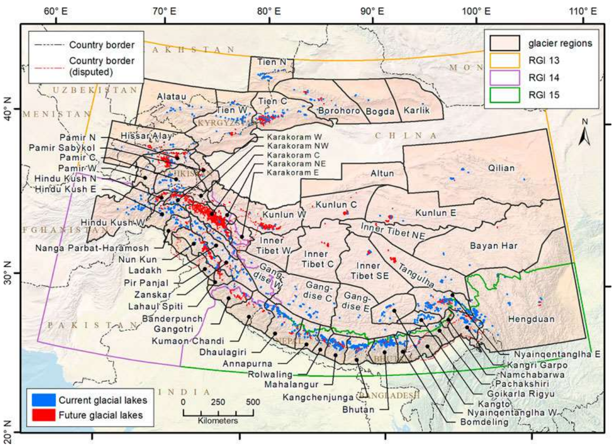

Figure 1. Glacial lake layout [3] for High Mountain Asia (HMA) region

Sensing and Spatial Analytics Lab at the Information Technology University, Pakistan [7]. To create this valuable resource, cloud-free, multispectral Sentinel-2 data was acquired from the Sentinel Hub EO Browser, ultimately retaining around 400 scenes for processing. From these scenes, 1200, 320x320 pixel crops containing glacial lakes were generated, preserving the original pixel resolution. Figure 2 shows the diversity of this dataset, showcasing true-color images with varying glacial lake radiometric signatures. Each image is classified using a binary system: 1 representing a lake and 0 indicating the absence of a lake. Our dataset [7] is divided into training and testing sets containing 1000 and 200 images, respectively. Figure 3 illustrates four examples from our training dataset, displaying both true-color images and their corresponding labels for ground truth masks.

IV. Methodology

We employed three distinct DL models with varying backbones and loss functions. DeepLabV3+ [8], an advanced CNN designed for image segmentation tasks, enhances the original DeepLab models by incorporating a more robust decoder module, which improves segmentation accuracy, particularly along object boundaries. The DeepLabV3+ model was paired with the ResNet50 backbone, utilizing pre-trained weights from ImageNet. We augmented the data using various techniques and utilized a batch size of 8. The optimizer we used was Adam, with a learning rate of 1e-3 and 0.001 for training. This approach exhibits efficient performance with minimal memory usage. We explored different functions for calculating loss, including binary cross-entropy, binary focal, and dice loss. The second model we utilized is the U-Net Encoder-Decoder segmentation network [9], paired with both ResNet50 and EfficientNetB0 backbones, also using ImageNet pre-trained weights as ImageNet has a vast collection of images that the models were pre-trained on.

All models were trained for 50 epochs. The DeepLabV3+ model was initially trained for 30 epochs, and then, after a brief pause, it was trained for an additional 20 epochs. We observed that the DeepLabV3+ model exhibits a temporary divergence during training, followed by enhanced performance compared to previous iterations. This behavior is consistently observed in the training of DeepLabV3+, and the reason for initially training it for 30 epochs and then an additional 20 epochs is to accommodate this pattern and achieve the best performance.

At last we implemented the YOLOv8-Seg [4] model which involves identifying and outlining individual objects in an image, providing a detailed understanding of spatial distribution.

Data was preprocessed, augmented and fed in to the DL models with varying batch sizes, learning rates and other hyperparatmers.

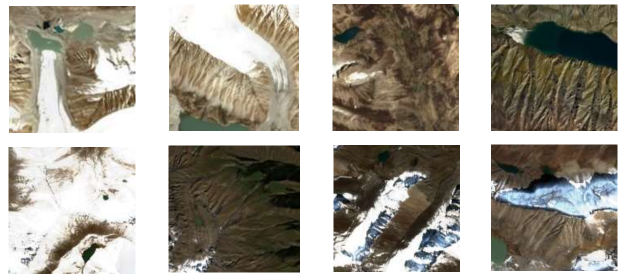

Figure. 2. Random images from the Sentinel-2 dataset displaying lakes of different colors and varying topographical features, including snow-covered mountains and glacial terrain.

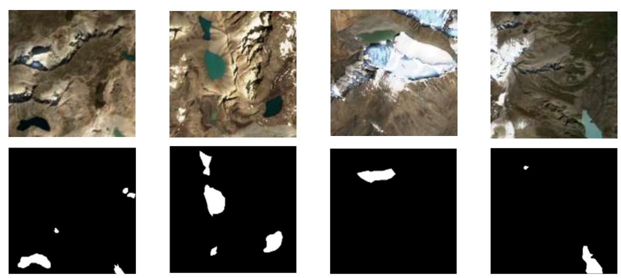

Figure. 3. Row 1 shows the Sentinel-2 RGB colored images of glacial lakes, Row 2 shows their respective ground truth masks.

A. *Proposed System Models*

Remote sensing image data encompass numerous intricate features, including glaciers, vegetation, exposed ground, and mountain shadows. This study exclusively focuses on glacial lake feature images, along with their corresponding masks. Several models have been created to selectively categorize and identify glacial lakes while extracting other valuable information for mapping purposes.

A schematic representation of the models utilized in our research is illustrated in Fig 4, 5, 6, and 7.

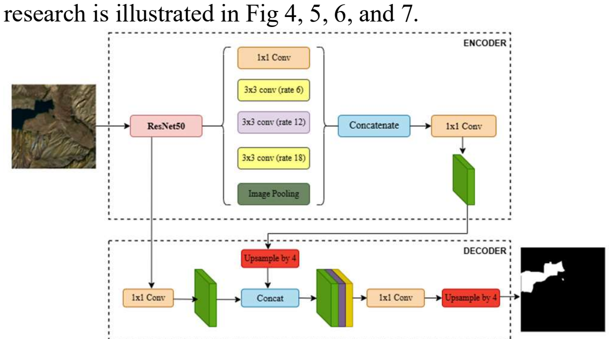

Figure 4. DeepLabV3+ with ResNet50 Backbone

The DeepLab series for semantic segmentation includes DeepLabV3+ [8], which utilizes Atrous Spatial Pyramid Pooling (ASPP) to enhance multi-scale context extraction. ASPP employs atrous convolutions at different rates to capture diverse features, allowing broader contextual aggregation without significantly increasing computation.

and segmenting them, but it did not achieve the highest performance among specialized segmentation models.

We used two main types of loss functions for our model to identify the error and gap: cross-entropy loss and dice loss, which were used in U-Net and DeepLabV3+ models, while

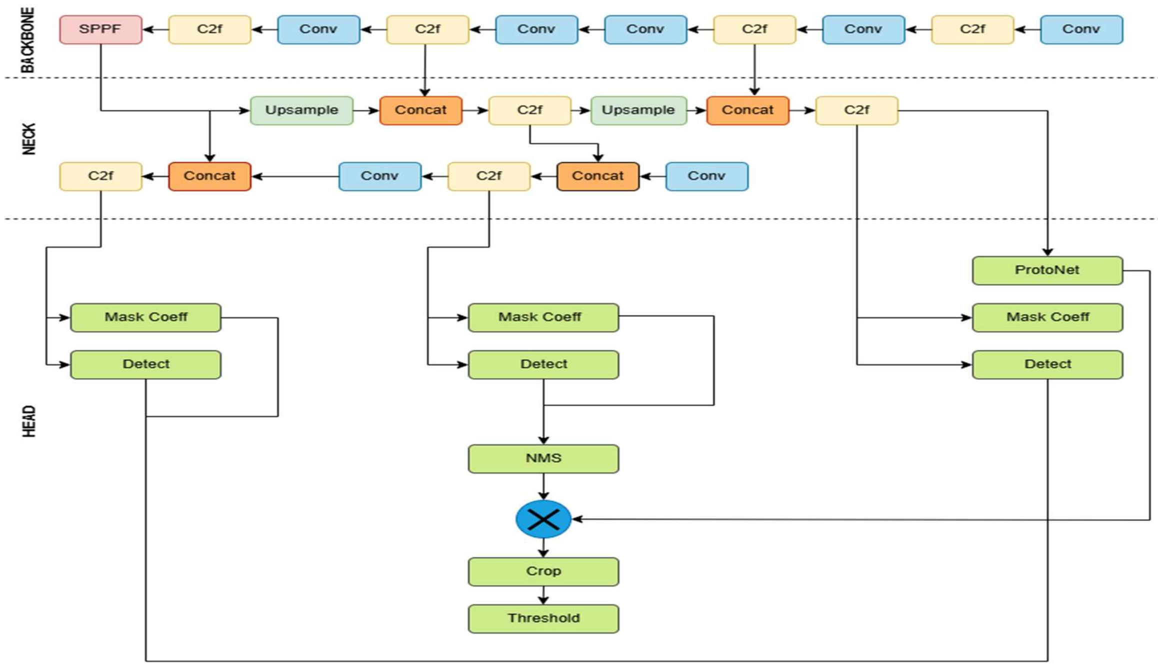

Figure. 5. The YOLOv8-Seg model shares similar fundamental architecture as the YOLOv8-Detect model, with an extra output module in the head that generates mask coefficients, as well as a supplementary Fully Convolutional Network layer (FCN) termed the Proto

The YOLOv8-Seg model is a significant advancement in the YOLO series for object segmentation, integrating a CNN with a segmentation head for simultaneous detection and precise delineation of object boundaries.

Trained on glacial lake images and ground truth masks, the YOLOv8-Seg model effectively segmented and classified the lakes with high confidence, accurately identifying their boundaries despite varying shapes, sizes, and environmental conditions.

The U-Net model is based on ecoder-decoder architecture for context capture and spatial reconstruction, with ResNet enhancing it through skip connections for better gradient flow and complex feature learning. In our study, the U-Net with a ResNet-50 backbone and pre-trained ImageNet weights performed well on glacial lake images, accurately classifying

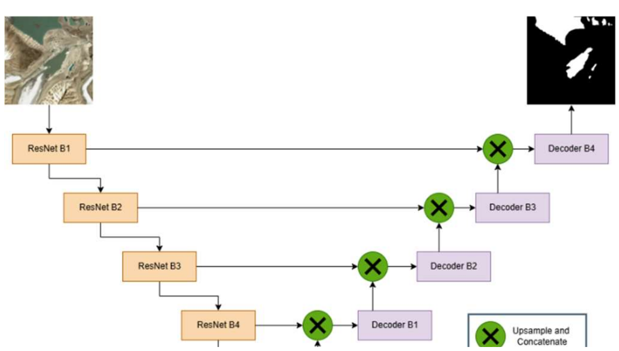

Figure 6. U-Net with ResNet50 Backbone

YOLOv8-Seg uses its own loss functions like box loss, class loss, and Dfl loss, in addition to cross-entropy loss.

The cross-entropy loss function (1) measures the discrepancy between the model's predicted and true probability distributions, assessing alignment with target labels in image segmentation tasks.

$$L = -\sum_{i=1}^{N} y_i \log(\hat{y}_i) \tag{1}$$

Dice Loss (2), derived from the Dice Similarity Coefficient, evaluates ML model performance in segmentation tasks by measuring the similarity between predicted and target segmentations. [7]

$$\mathcal{L}_{Dice} = 1 - \frac{2(y\overline{p}+1)}{(y+\overline{p}+1)} \tag{2}$$

B. *Evaluation Metrices*

The IoU score, in combination with precision, recall and F1-score, was our main evaluation metric.

Intersection over Union (IoU) measures object detection accuracy by assessing the overlap between predicted and actual bounding boxes or segmented regions.

$$IoU = \frac{|A \cap B|}{|A \cup B|} \tag{3}$$

Precision measures the quality of a model's positive predictions by dividing correct positives by total positive predictions, with higher values indicating better precision.

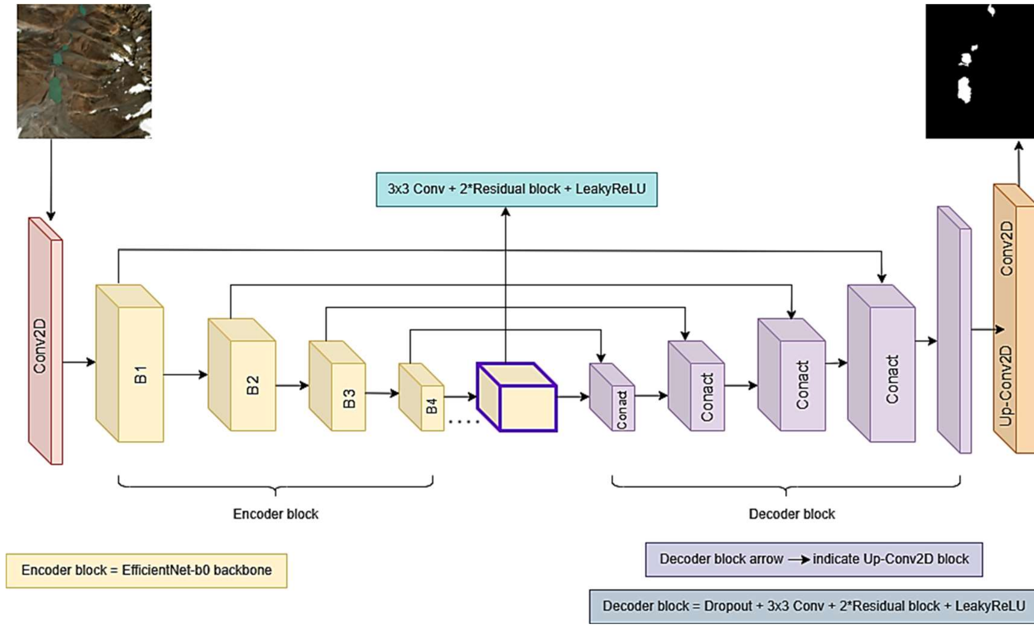

Figure 7. U-Net architecture with EfficientNet-B0 Backbone.

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

Recall evaluates a model's capacity to identify positive occurrences by dividing true positives by the total actual positives, including both true positives and false negatives.

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

F1-score (F1) rate is the harmonic mean of precision (4) value and recall rate (5):

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{6}$$

The Dice Coefficient (DL = 1 – Dice Coeff) is a quantitative formula that assesses the degree of similarity between two distinct sets, denoted as A and B. The coefficient value ranges from 0 to 1.

$$D = \frac{2|A \cap B|}{|A| + |B|} \tag{7}$$

TABLE I. EVALUATION METRICS TABLE

Model + Backbone	*Loss Function*	*Dice Coeff*	*Accuracy*	*Precision*	*Recall*
DeepLabV3+ ResNet50	Cross entropy	0.85	0.95	0.88	0.85
DeepLabV3+ ResNet50	Dice Loss	0.86	0.78	0.93	0.82
U-Net + EfficientNetB0	Cross entropy	0.76	0.99	0.90	0.69
U-Net	Cross entropy	0.56	0.99	0.69	0.72
U-Net + ResNet50	Cross entropy	0.74	0.99	0.88	0.68
YOLOv8	Multiple	0.73	0.89	0.90	0.84

V. RESULTS AND ANALYSIS

Our results include three types of images: a predicted segmentation mask, the predicted mask overlaid on the original image, and a GradCam visualization. GradCam produces a rough localization map that highlights the crucial regions in the image for the model's prediction of the target concept.

The DeepLabV3+ architecture exhibited significant performance variations based on the backbone and loss function used. With a ResNet50 backbone and cross-entropy loss, the model achieved an IoU of 75.5%. Using Dice Loss improved the Dice coefficient to 0.86 and increased the IoU to 79.2%. These findings suggest that the performance of the model can be considerably impacted by the choice of the loss function, with Dice Loss yielding better results. The results are shown in Table I.

The U-net [9] models, incorporating both ResNet50 and EfficientNetB0 backbones, aimed to balance gradient flow and computational efficiency. The U-Net with an EfficientNetB0 backbone and cross-entropy loss achieved an IoU of 69.1%. In contrast, the U-Net with a ResNet50 backbone and cross-entropy loss exhibited a lower IoU of 59%. The baseline U-Net without any advanced backbone yielded the lowest performance with a Dice coefficient of 0.56 and an IoU of 65%. These findings suggest that incorporating more sophisticated backbones like EfficientNetB0 can enhance the IoU score, although the improvements are contingent on the overall model architecture and loss function employed.

The YOLOv8-Seg [4] model, also demonstrated competitive performance. With multiple loss functions, YOLOv8 achieved an IoU of 72.9%. Despite not surpassing the top-performing configurations of DeepLabV3+ or U-Net, the YOLOv8 model exhibited substantial potential, particularly in terms of precision and recall.

Among the various models and configurations tested, the DeepLabV3+ with a ResNet50 backbone and Dice Loss function was found to be the top-performing model. This combination produced the highest IoU score of 79.2%, indicating its excellent ability to accurately segment glacial lakes in satellite images. The enhanced performance underscores the importance of selecting appropriate architectural components and loss functions to optimize model efficacy.

Since IoU was the primary evaluation metric for our models, we trained all of our models on a dataset of 1000 images and their corresponding segmentation masks [7]. After careful hyperparameter optimization and rigorous experimentation, our models demonstrated strong performance on the validation set.

A. Result table

IoU Score Calculation: Table II shows the IoU scores for the models evaluated in our study, using this metric to assess performance. The overlap between predicted bounding boxes and ground truth is measured by the IoU, making it crucial for image segmentation as well as object detection. Higher IoU scores indicate better model performance and more accurate alignment with true object boundaries.

High IoU scores are crucial for accurately outlining glacial lake boundaries in GLOFs detection, essential for effective monitoring and early warning systems. Comparing IoU scores across models helps identify the most reliable deep learning architectures for precise detection.

The DeepLabV3+ model with a ResNet50 backbone and Dice Loss function achieved the highest IoU score of 79.2%, demonstrating its superior performance in accurately predicting the spatial extent of glacial lakes. This model's combination of a powerful backbone and an effective loss function likely contributed to its high accuracy. In comparison, the same model with a cross-entropy loss function achieved a slightly lower IoU score of 75.5%, indicating that while still highly effective, the choice of loss function plays a essential role in performance optimization.

TABLE II. IOU RESULT TABLE

Model	***Backbone***	***Loss function***	***IoU%***
DeepLabV3+	ResNet50	Cross entropy	75.5%
DeepLabV3+	**ResNet50**	**Dice Loss**	**79.2%**
U-Net	EfficientNetB0	Cross entropy	69.1%
U-Net	None	Cross entropy	65%
U-Net	ResNet50	Cross entropy	59%
YOLOv8	None	Multiple	72.9%

The above table is used to compare the values of the IoU rate with each other to see which model produced the best results.

B. YOLOv8-Seg Results

The YOLOv8-Seg model was used to segment glacial lakes from the dataset, employing both anchor-based and anchor-free mechanisms to optimize multi-scale feature extraction for improved accuracy and efficiency.

Our models produced two output types: an inference function for classifying and segmenting glacial lakes on the validation set and analyzed validation images to generate corresponding masks. The predictions from both methods were equivalent, differing only in presentation as shown in Figure 9. The results of the inference function and the automatic prediction generation are shown in Figures 9 and 10, respectively.

C. Broader Implications and Societal Impact

The implications of this research extend beyond GLOF monitoring, as its methodologies can also be adapted to address other climate-induced hazards like landslides and flash floods in Pakistan's northern regions. Large-scale deployment must consider socio-economic and ethical challenges, particularly regarding false alarms that could disrupt vulnerable communities. Future research will aim to develop early warning systems that are socially equitable, ethically sound, and effective in mitigating risks without exacerbating vulnerabilities.

A central focus of this research is translating GLOF detections into actionable early warning systems. The proposed framework is designed to integrate with existing disaster risk management protocols in the region. Collaboration with local disaster management authorities will be crucial for ensuring that the model's predictive outputs effectively inform decision-making. By creating intuitive risk visualization interfaces and incorporating the model into automated alert systems, we aim to enable prompt responses and facilitate the deployment of mitigation strategies and evacuation procedures based on real-time risk assessments.

D. Limitations and future research directions

Although the proposed model shows promising results, scaling it in Pakistan’s northern regions presents challenges due to rugged terrain, limited infrastructure, and restricted data access. High-resolution data demands substantial computational power, and connectivity issues hinder real-time processing. Future work will focus on optimizing algorithms for low-resource settings to maintain accuracy.

In high-stakes GLOF-prone areas, false positives and negatives carry significant risks, potentially leading to unnecessary evacuations or missed warnings. To address this, future efforts will emphasize rigorous testing under varied conditions, alongside multi-sensor data fusion using hydrological, meteorological, and geological data to enhance prediction reliability and reduce erroneous alerts.

Future work can also focus on improving deep learning models and semantic segmentation techniques while dedicating efforts to climate change research and developing datasets to help prevent natural and human-caused disasters. Additionally, the future research will aim to optimize these models and loss functions for better segmentation performance in remote sensing applications, potentially integrating with future studies for impactful results.

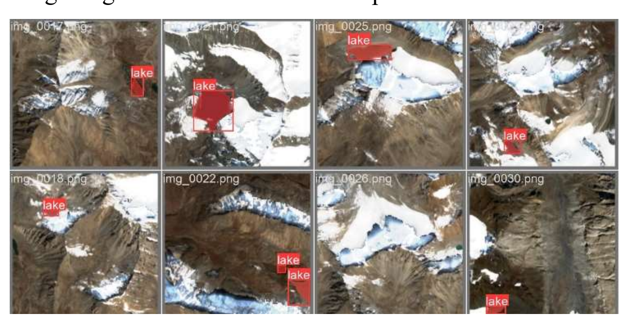

Figure. 8. The figure displays the ground truth for Sentinel-2 images, with the YOLOv8 model identifying glacial lakes. Each sub-image captures unique topographical features, including snow-covered mountains, rugged terrain, and multiple glacial lakes.

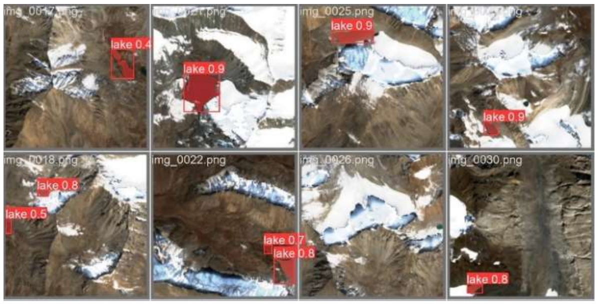

Figure. 9. The figure displays YOLOv8-seg model predictions, accurately identifying most glacial lakes with confidence scores. Red labels mark the lakes, and masks overlay them.

VI. CONCLUSION

Identifying and monitoring GLOFs is vital in addressing climate change challenges in high-altitude areas. Our study employs advanced deep learning methods to segment glacial lakes, achieving a 79.2% IoU score with the DeepLabv3+ model and ResNet50 backbone, surpassing previous studies.

The model's accuracy and adaptability are crucial for real-world applications in northern Pakistan, where GLOFs pose significant risks to life and infrastructure. Timely monitoring can provide alerts to mitigate these threats. This research highlights the importance of advanced architectures and data management in developing effective object detection systems, demonstrating substantial potential for societal impact.

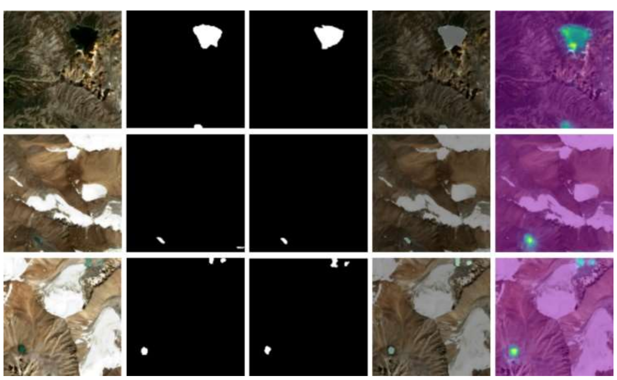

Figure. 10. The above figure shows Sentinel-2 RGB images from our dataset in Column 1, and the corresponding ground truth masks in Column 2. Column 3 showcases the model predictions, while Column 4 overlays these predictions on the original images. Column 5 displays Grad-CAM visualizations of the predictions. Specifically, the first row in Column 3 depicts the mask predicted by DeepLabV3+ResNet50 [8] with Dice Loss, the second row by DeepLabV3+ResNet50 with Cross-Entropy loss, the third row by U-Net [9] with EfficientNetB0 backbone and Cross-Entropy loss.

ACKNOWLEDGMENT

We sincerely thank the Remote Sensing and Spatial Analytics Lab at Information Technology University, Lahore, for the invaluable dataset that was crucial for our research on glacial lake classification and segmentation using deep learning techniques.

REFERENCES

[1] Scaling up of glacial lake outburst flood risk reduction in northern Pakistan | Climate Change adaptation. (2024, April 22). https://www.adaptation-undp.org/projects/scaling-glacial-lake-outburst-flood-risk-reduction-northern-pakistan

[2] Muneeb F, Baig SU, Khan JA, Khokhar MF. Inventory and GLOF Susceptibility of Glacial Lakes in Hunza River Basin, Western Karakorum. Remote Sensing. 2021; 13(9):1794. https://doi.org/10.3390/rs13091794

[3] S. K. Allen et al., "Potentially dangerous glacial lakes across the Tibetan Plateau revealed using a large-scale automated assessment approach", Sci. Bull., vol. 64, no. 7, pp. 435-445, 2019.

[4] Vadean, A., Achiche, S., & Aage, N. (2024). From Density to Geometry: YOLOv8 Instance Segmentation for Reverse Engineering of Optimized Structures. ArXiv. /abs/2404.18763.

[5] A. Basit, M. K. Bhatti, M. Ali, T. Fatima, B. Minchew and M. A. Siddique, "Deep Learning for Monitoring Glacial Lakes Formation using Sentinel 2 Multispectral Data," IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium, Kuala Lumpur, Malaysia, 2022, pp. 179-182.

[6] M. A. Siddique et al., "Towards Automated Monitoring Of Glacial Lakes In Hindu Kush And Himalayas Using Deep Learning," IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium, Pasadena, CA, USA, 2023, pp. 2165-2168, doi: 10.1109/IGARSS52108.2023.10281471.

[7] A. Basit, M. A. Siddique, July 4, 2022, "Glacial Lakes Detection Dataset", IEEE Dataport, doi: https://dx.doi.org/10.21227/th0c-8t68

[8] Chen, L., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. ArXiv. /abs/1802.02611

[9] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. ArXiv. /abs/1505.04597

[10] Sudre, C. H., Li, W., Vercauteren, T., Ourselin, S., & Cardoso, M. J. (2017). Generalised Dice overlap as a deep learning loss function for highly unbalanced segmentations. ArXiv. https://doi.org/10.1007/978-3-319-67558-9_28

[11] Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2016). Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. ArXiv. https://doi.org/10.1007/s11263-019-01228-7.

[12] J. M. Su, L. X. Yang, and W. P. Jing, "A U-net based semantic segmentation method for high resolution remote sensing image," Comput. Eng. Appl., vol. 55, no. 7, pp. 207–213, 2019

[13] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: Redesigning skip connections to exploit multiscale features in image segmentation," IEEE Trans. Med. Imag., vol. 39, no. 6, pp. 1856–1867, Jun. 2020

[14] Zhang M, Chen F, Zhao H, Wang J, Wang N. Recent Changes of Glacial Lakes in the High Mountain Asia and Its Potential Controlling Factors Analysis. Remote Sensing. 2021; 13(18):3757. https://doi.org/10.3390/rs13183757

[15] Furian W, Maussion F and Schneider C (2022) Projected 21st-Century Glacial Lake Evolution in High Mountain Asia. Front. Earth Sci. 10:821798. doi: 10.3389/feart.2022.821798

[16] Khan, Mohammad & Khan, Jawed & Ali, Zulfiqar & Ahmad, Imran & Ahmad, Muhammad. (2016). The challenge of climate change and policy response in Pakistan. Environmental Earth Sciences. 75. 10.1007/s12665-015-5127-7.

[17] Kaushik, Saurabh & Singh, Tejpal & Joshi, P.K. & Andreas, Dietz. (2022). Automated mapping of glacial lakes using multisource remote sensing data and deep convolutional neural network. International Journal of Applied Earth Observation and Geoinformation. 115. 103085. 10.1016/j.jag.2022.103085

[18] Hu S, Liu J, Kang Z. DeepLabV3+/Efficientnet Hybrid Network-Based Scene Area Judgment for the Mars Unmanned Vehicle System. Sensors. 2021; 21(23):8136. https://doi.org/10.3390/s21238136

[19] Cheng, Bowen & Girshick, Ross & Dollar, Piotr & Berg, Alexander & Kirillov, Alexander. (2021). Boundary IoU: Improving Object-Centric Image Segmentation Evaluation. 15329-15337. 10.1109/CVPR46437.2021.01508.

[20] Mingxing Tan and Quoc Le. Efficientnet: Rethinking model scaling for convolutional neural networks. In International conference on machine learning, pages 6105–6114. PMLR, 2019.