Mehran University Research Journal of Engineering and Technology

https://doi.org/10.22581/muet1982.0742

2026, 45(2) 234-243

A Semi-Supervised Approach utilizing Deep Learning Pseudo-Label Refinement and Confidence Filtering for Enhanced Image Segmentation: An application towards Skin Lesion Analysis

Saad Rehman Lodhi [a], Sobia Arshad [b], Junaid Ali Khan [c], Saad Rehman [d]

[a] *PHD Student, HITEC University, Taxila, Pakistan*

[b] *Department of Computer Engineering, HITEC University, Taxila, Pakistan*

[c] *Department of Data Science and Analytics, Sakarya University, 54050 Serdivan/Sakarya, Türkiye*

[d] *Khanpur Institute of Technology, Rahim yar Khan, Pakistan*

*Corresponding author: saademe@gmail.com

Received: 02 February 2026, Revised: 16 March 2026, Accepted: 31 March 2026, Published: 01 April 2026

KEYWORDS

Medical Image Analysis

U-Net Architecture

Uncertainty filtering

Skin Lesion Segmentation

Conditional Random Field (CRF) Refinement

Semi-supervised Learning

Teacher-Student Framework

ABSTRACT

Accurate segmentation of skin lesions is an important step in diagnosing skin cancer However, most segmentation methods based on deep learning rely heavily on larg volumes of manually annotated images, which are both costly and difficult to obtain. Thi paper proposes a new framework for semi-supervised learning that effectively links a abundance of unlabelled thermoscopic images with a small, labelled dataset to improv the performance of segmentation. The framework integrates an architecture of Teacher Student, where the Teacher network builds pseudo-labels for the unlabelled samples. Th refinement of pseudo-labels is achieved through uncertainty-based filtering, whic discards unreliable areas, and is further enhanced by Conditional Random Field (CRF post-processing, which sharpens the lesion boundaries. The refined pseudo-labels are the utilised for training the student network, allowing for robust learning through bot unlabelled and labelled data. Experiments on the HAM10000 and ISIC 2018 dataset indicate that the proposed model achieves an IoU of 0.64 and a Dice Coefficient of 0.7 with only 10 percent of labelled data. It was comparable to fully supervised training wit 100 percent annotations. Qualitative outcomes suggest improved delineation of lesio boundary and decreased false segmentation. This research work demonstrates th potential of pseudo-label refinement as an effective strategy for addressing the scarcity c annotations in medical image segmentation.

1. Introduction

Skin cancer is one of the most life-threatening and prevalent forms of cancer present around the world, with melanoma causing more of the skin-cancer-associated deaths. Precise and early segmentation of skin lesions plays a crucial role in computer-aided diagnosis (CAD) systems, enabling the classification of malignant lesions and facilitating the automated detection of these lesions. Traditional models of deep learning, such as U-Net, U-Net++, and Attention U-Net, have acquired efficient results in segmenting medical images [1][2]. However, these architectures are seen as dependent on extensive pixel-wise annotations, which are both time-consuming and labour-intensive for dermatologists [3]. As an outcome, the scarcity of labelled data often restricts the generalisation and performance of supervised learning models.

Current research has explored semi-supervised learning (SSL) as an effective solution to decrease dependency on labelled data. SSL techniques exploit a small labelled subset along with the larger pool of unlabelled images by making pseudo-labels and machine-predicted masks approximating ground truth.

© Mehran University of Engineering and Technology 2026

Approaches such as Consistency Regularisation, FixMatch, and Mean Teacher have been assimilated successfully into the segmentation of natural images [4]. However, in medical imaging, these methods do suffer from uncertain boundary predictions and noisy pseudo-labels, resulting in degraded performance. Medical images are characterised through inter-class similarity, irregular shapes of lesions, and low contrast, which enhance the chances of error propagation while doing pseudo-labelling [5].

For given challenges, this research proposes a semi-supervised framework for the segmentation of skin lesions that incorporates confidence-based filtering with pseudo-label refinements. In the start, the training of a Teacher U-Net model is done over the labelled subset to make preliminary pseudo-labels for the samples that are unlabelled [6]. The pixel-wise prediction confidence of the model is then estimated through entropy-based uncertainty maps. Unreliable areas are then excluded to decrease noise. At the same time, a Conditional Random Field (CRF), a phase of post-processing, sharpens the boundaries of the lesion, making corrections in local inconsistencies within the pseudo-labels [7]. The refined pseudo-labels are then used for training Student U-Net, enabling a paradigm of Teacher and Student that improves the quality of segmentation iteratively [8]. This framework contributes by making three major advancements:

1. Uncertainty filtering decreases the impact of such pseudo-labels that have low confidence, and it prevents the error reinforcement.
2. CRF-based refinement helps to ensure more anatomically plausible and smoother contours of the lesion.
3. Efficient annotation utilization permits near-supervised accuracy with around 10 percent of labelled data.

The experimental findings on HAM10000 and ISIC 2018 datasets indicate substantial gains in performance. It shows an improvement of around 10 per cent of Dice and 11 per cent of IOU over the supervised model. It shows superior boundary precision and generalisation. The results show that incorporating confidence-guided refinement of pseudo-label in a semi-supervised framework can handle the limitations of annotation while maintaining the reliability of diagnosis.

2. Related Work

2.1. Fully Supervised Segmentation Models

Deep learning segmentation methods have diverted the analysis of medical images, with the U-Net developing a backbone of encoder and decoder that can capture both spatial and contextual information with skip connections [9]. Its successors, such as Attention U-Net and U-Net++, introduced attention gates and nested skip paths for refining the multiscale aggregation of feature and target on lesion areas [10]. While these architectures get higher accuracy, their dependency on a larger annotated dataset causes limitations to scalability. Acquiring dense pixel-level masks through dermatologists tends to be both subjective and time-consuming. It results in overfitting when it is trained over a small, labelled dataset [11].

2.2. Semi-Supervised Learning in Medical Imaging

Semi-supervised learning (SSL) is a cost-effective solution to the scarcity of annotation through leveraging unlabelled data with limited labelled samples. Early strategies of SSL, such as pseudo-labelling, assign labels that are predicted by the model to unlabelled images. These are then treated as ground truth at the time of training. More current consistency-based frameworks, such as FixMatch, MixMatch, and Mean Teacher, motivate agreement between the outputs of the model under various augmentations, regularising the training effectively. With reference to the context of biomedical, SSL methods have been integrated into the segmentation of histopathology, retinal, and organ [12] [13]. Contrastingly, challenges related to the domain persist. The low inter-class contrast, high intra-class variability, and fine structural details of the lesion cause pseudo-labels to be shaped by distortions and contain noise at the boundary. Without corrective processes, these noisy labels introduce errors, and they degrade the overall performance of the model.

2.3. Uncertainty Estimation and Label Refinement

Current advancements incorporate uncertainty estimation for assessing the reliability of pseudo-labels. Strategies depend on the entropy or Monte Carlo Dropout to quantify confidence for the model at the pixel level [14]. It does the quantification of confidence of the model at the pixel-level, permitting selective training over high-certainty areas [15]. On the other hand, the methods of post-processing refinement, like Conditional Random Fields (CRFs), improve the precision of edge and spatial coherence through the exploitation of contextual links between neighbouring pixels. Some of the research jointly combines CRF refinement and uncertainty filtering in a unified semi-supervised framework that is tailored to the dermoscopic imagery.

© Mehran University of Engineering and Technology 2026

Table 1

Summary of Findings

Theme	Key Methods / Models	Main Contributions	Limitations / Gaps Identified
Fully Supervised Segmentation Models	U-Net, Attention U-Net, U-Net++	Established encoder-decoder backbone having skip connections for the segmentation of medical images. Attention U-Net improves the target in the lesion areas. U-Net++ improves multi-scale feature aggregation with the help of nested skip paths.	Strong dependence on pixel-level, large annotated datasets Annotation of manual masks by dermatologists tends to be subjective and time-consuming More risk of overfitting when labelled over small labelled datasets
Semi-Supervised Learning (SSL) in Medical Imaging	Pseudo-labelling, FixMatch, MixMatch, Mean Teacher	Uses large amounts of unlabelled data with limited samples having labels Consistency regularization improves in generalisation under data augmentation Successfully integrated into organ segmentation, retinal, and histopathology tasks.	Pseudo-labels may be noisy because of high lesion and low contrast variability. Fine structural errors and boundary distortions propagate at the time of training. Domain-specific challenges in dermoscopic images are seen as unresolved.
Uncertainty Estimation & Label Refinement	Entropy-based confidence, Monte Carlo Dropout, CRF post-processing	Quantifies pixel-level prediction confidence Allows selective training through high-certainty areas of pseudo-label CRFs improve spatial consistency and boundary precision.	Additional computational complexity. The estimation of uncertainty alone cannot estimate noisy labels fully. Limited unified frameworks are particularly optimized for the segmentation of skin lesions.

2.4. Research Gap and Contribution

Despite major progress in semi-supervised segmentations, there are two main limitations:

- Most of the frameworks reuse pseudo-labels without the evaluation of confidence, therefore causing noise amplification.
- Refinement of the boundary is ignored, which results in irregular lesion contours.

The current research views these issues with support of a confidence-guided pseudo-label strategy, which incorporates entropy-based filtering of uncertainty with the correction of the CRF boundary in a configuration of Teacher with Student. Through fusing the given components, the proposed approach improves localisation of lesions, decreases the annotation requirements and improves generalisation across datasets.

2.5. Abbreviation Table

Table 2

Abbreviation Table

Abbreviation	Full Form / Description
AI	Artificial Intelligence
BCE	Binary Cross-Entropy
CAD	Computer-Aided Diagnosis
CRF	Conditional Random Field
CNN	Convolutional Neural Network
Dice / DICE	Dice Similarity Coefficient
DL	Deep Learning
Entropy	Measure of prediction uncertainty
HD95	95th Percentile Hausdorff Distance
IoU	Intersection over Union
ISIC	International Skin Imaging Collaboration
LR	Learning Rate
MC Dropout	Monte Carlo Dropout
ReLU	Rectified Linear Unit
ResNet	Residual Network
SSL	Semi-Supervised Learning

© Mehran University of Engineering and Technology 2026

Student	Model trained using refined pseudo-labels
Teacher	Model generating pseudo-labels
U-Net	Encoder–Decoder Convolutional Network
U-Net++	Nested U-Net Architecture
Attention U-Net	U-Net with Attention Gates
HAM10000	Human Against Machine with 10,000 training images
GPU	Graphics Processing Unit
ROC	Receiver Operating Characteristic
AUC	Area Under the Curve
PH2	Dermoscopic Image Dataset (Future Work)
ImageNet	Large-scale Image Classification Dataset
Pseudo-label	Model-generated annotation for unlabelled data
Uncertainty Filtering	Removal of low-confidence predictions
Teacher–Student Framework	Iterative semi-supervised learning paradigm

3. Methodology

3.1. Overview

The proposed framework integrates a semi-supervised Teacher-Student paradigm designed to leverage both unlabelled and labelled dermoscopic images for the segmentation of skin lesions. The method integrates Conditional Random Field (CRF) and uncertainty-based filtering to improve the reliability of pseudo-labels generated through unlabelled data. An overview of the workflow is given in Figure 1.

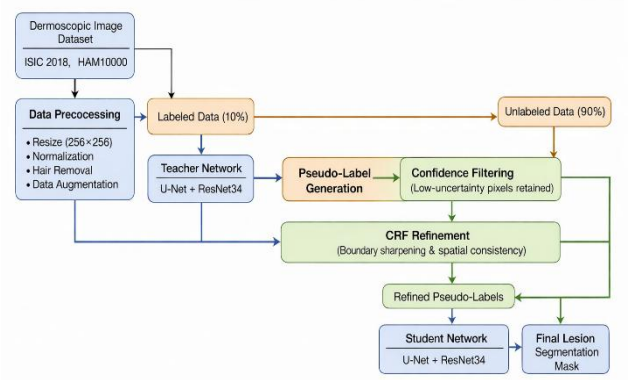

Fig. 1. Methodology

3.2. Datasets

Experiments were done over two publicly available dermoscopic datasets:

- ISIC 2018 Task 1–2 Dataset: It consists of 2,594 dermoscopic images with their associated lesion masks. It features various types of lesions like seborrheic keratosis, nevus, and melanoma, as shown in Figure 2 [16].

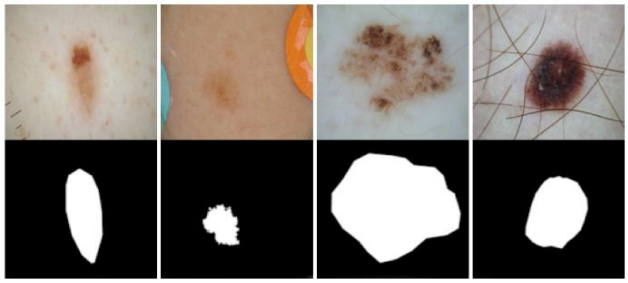

Fig. 2. ISIC 2018 Dataset

- HAM10000 Dataset: It contains 10,015 images of seven categories of diagnosis, as shown in Figure 3. For the sake of consistency, lesion masks from the given dataset were built through expert annotations given in previous research [17].

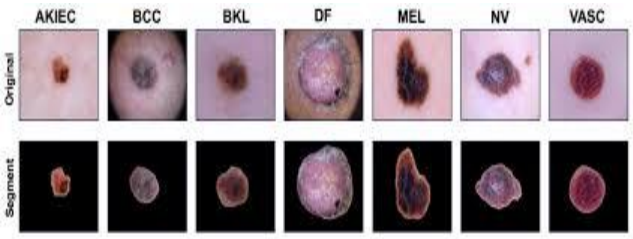

Fig. 3. HAM10000 dataset

The division of the dataset was done randomly into 10 percent of labelled data that was used for supervised training of the Teacher network, and 90 percent of the dataset was kept unlabelled, which was used for Student learning and pseudo-label generation. Moreover, there was also a 20 percent validation set that was kept for the evaluation of performance.

3.3 Data Preprocessing and Augmentation

To ensure uniform resolution of input and to decrease bias, the resizing of images was done to 256 x 256 pixels. The normalization was done to the range [0,1]. Moreover, the augmentation was done through applying albumentations transformations involving elastic deformations, color jitter, rotations, and random flips [18]. This augmentation helps to improve generalization under limited conditions of labelled data. Hair removal filters and histogram equalization were also integrated for the normalization of contrast and for reducing noise, as shown in Figure 4.

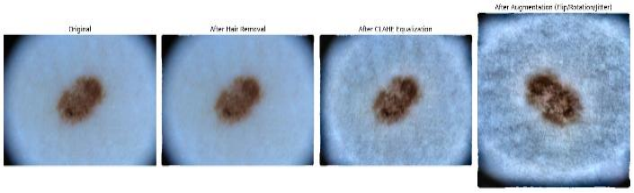

Fig. 4. Data Preprocessing

3.4. Network Architecture

Both the Teacher and Student models integrate the U-Net architecture because of its proven performance

© Mehran University of Engineering and Technology 2026

in the segmentation of medical images. Moreover, it can also preserve spatial details with the help of skip connections. The encoder follows the ResNet34 backbone, which is pretrained over ImageNet, giving a balance between computational efficiency and feature richness as shown in Figure 5 [19]. Every block of convolution is followed by ReLU activation, Batch Normalisation, and Dropout (p=0.2) for preventing overfitting. The decoder mimics the encoder with transposed convolutions for concatenation and unsampling of associated feature maps [20].

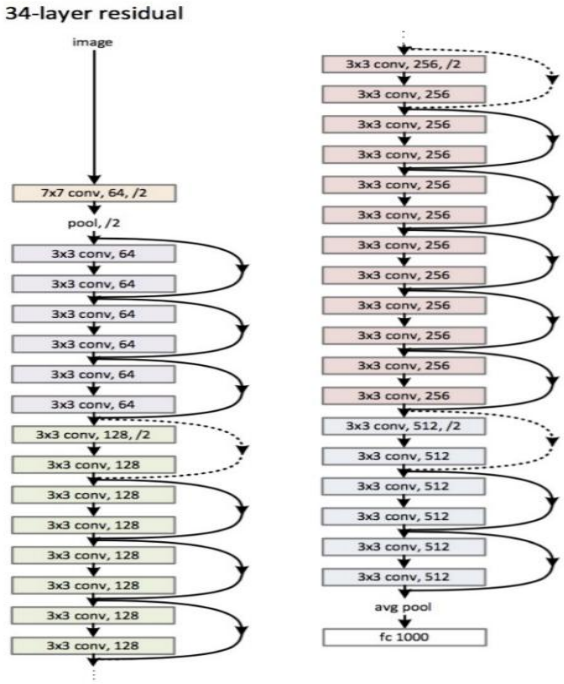

Fig. 5. Architecture of ResNet34

3.5. Semi-Supervised Learning Framework

The framework involves three main stages:

Stage 1 – Supervised Teacher Training

The training of the teacher model is done over the labelled subset through a hybrid loss function linking Dice Loss and Binary Cross-Entropy (BCE) [21].

$$\mathcal{L}_{supervised} = \alpha \times \text{BCE} + (1 - \alpha) \times (1 - \text{Dice})$$

Where α=0.5\alpha = 0.5α=0.5 balances region overlap and pixel-level accuracy.

Stage 2 – Pseudo-Label Generation and Uncertainty Filtering

After training, the Teacher model makes a prediction of lesion masks for unlabelled images. For every pixel, the computation of entropy-based uncertainty is done as [22]:

$$H(p) = -p\log(p) - (1 - p)\log(1 - p)$$

Pixels with uncertainty that exceed a specific threshold, like τ=0.2\tau = 0.2τ=0.2, are discarded (masked out). It helps ensure that only regions that are of high confidence contribute to the formation of the pseudo-label. This decreases the influence of error propagation and noise.

Stage 3 – CRF Refinement and Student Training

The refinement of filtered pseudo-labels is done through Dense Conditional Random Fields (CRFs), which improve the sharpness of edge and spatial consistency through considering color similarity and pixel adjacency [23].

$$E(x) = \sum_i \psi_u(x_i) + \sum_{i<i} \psi_p(x_i, x_j)$$

where ψu\psi_uψu shows unary potential (predictions of the model), and ψp\psi_pψp does the encoding of pairwise links based upon spatial distance and color.

The refined pseudo-labels are afterwards used as supervision for training the Student model through a similar hybrid loss. Optionally, in iterative cycles, students become the new teachers in the given research. One iteration provided stable improvements in performance.

3.6. Evaluation Metrics

The performance assessment of the model was done through standard metrics of segmentation.

1. Dice Coefficient (DICE) – does the measurement of overlap between ground truth and prediction [24].
2. Intersection over Union (IoU) – does the quantification of pixel-level agreement.
3. Hausdorff Distance (HD95) – evaluates boundary accuracy, through the 95th percentile, for handling outlier impact.

For each experiment, the standard deviation and mean across all the test samples were reported. The verification of statistical significance was done through a paired t-test ($p < 0.05$).

3.7. Implementation Details

- Framework: PyTorch 2.0
- Hardware: NVIDIA A100 GPU with 40 GB memory.
- Optimizer: Adam (lr = 1×10^{-4}, weight decay = 1×10^{-5})
- Epochs: 60
- Batch size: 8

Training followed a cosine schedule of annealing learning rate, with early stopping based on validation Dice score. All experiments were repeated 3 times to ensure reproducibility.

© Mehran University of Engineering and Technology 2026

4. Results and Discussion

4.1 Ablation Study

An ablation analysis was also done to assess the contribution of each component within the framework.

Table 3

Ablation Study

Configuration	Dice	IoU
Teacher only (no SSL)	0.635	0.530
Teacher + Pseudo-labels (no filtering)	0.694	0.589
+ Uncertainty Filtering	0.719	0.620
+ CRF Refinement (Proposed)	0.740	0.643

Table 1. Ablation study showing the effect of uncertainty filtering and CRF refinement.

Both CRF refinement and uncertainty filtering made a positive contribution, with the CRF providing the largest incremental gain of more than 2.1 Dice. It ensures that confidence control and boundary refinement are important for effective semi-supervised segmentation.

4.2. Cross-Dataset Generalization

For the assessment of generalisation, the model trained on ISIC 2018 was also tested over HAM10000 without performing fine-tuning. The Dice Score reduced marginally from 0.740 to 0.712, showing stronger transferability. The robustness arises through diverse exposure to unlabelled data through Student training. It regularises the learning of the feature and also helps to prevent overfitting.

4.3. Dice Vs. Threshold Analysis for Validation Set

The trend of Dice vs. threshold assures that the proposed semi-supervised model not only attains a high Dice score, but it also maintains consistent performance without parameter fine-tuning. This stability recommends that the probability maps the model produced are properly calibrated and are a significant aspect for clinical deployment, where fixed thresholds are given more preference.

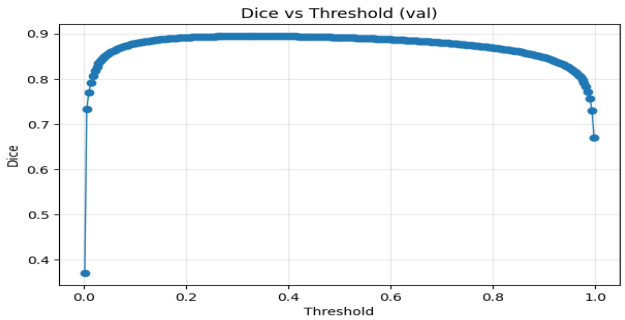

Fig. 6. Dice vs Threshold

4.4. IoU Vs. Threshold Analysis for Validation Set

The given figure complements the Dice-threshold curve by confirming generalisation stability and robust calibration across thresholds. While Dice mainly targets the similarity of overlap, still IoU shows area-dependent consistency; both metrics present a similar plateau, which recommends that the semi-supervised model provides threshold-invariant and reliable predictions. This is an important characteristic for clinical segmentation systems.

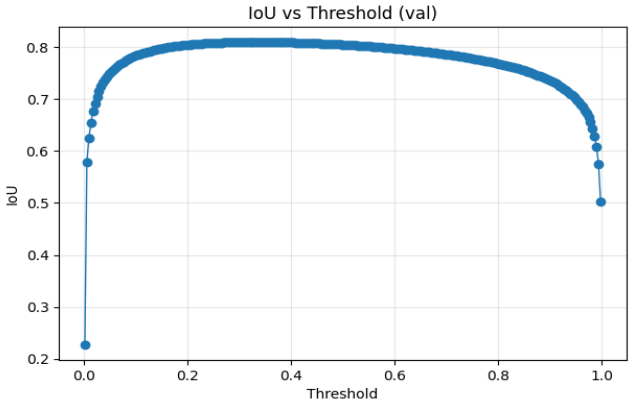

Fig. 7. IoU vs Threshold

4.5. ROC Curve for Validation Set

This strong ROC profile shows that the given semi-supervised framework provides reliable probability estimates for the presence of a lesion. The near-perfect curvature implies that in the case of different thresholds, the model kept a well-balanced trade-off and robust decision boundaries between recall and precision. This calibration holds significance in the tasks of medical segmentation, where reducing false negatives is as significant as decreasing false positives.

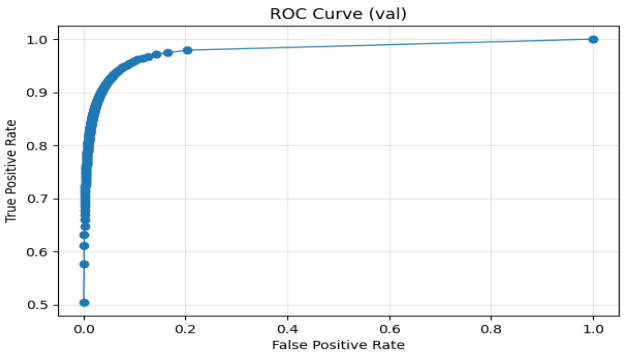

Fig. 8. ROC Curve

4.6. Quantitative Evaluation

The proposed framework of semi-supervised segmentation was evaluated through a baseline supervised U-Net model that was trained over a similar labelled subset. The figures shown below show the quantitative performance over the ISIC 2018 validation set through Intersection over Union (IoU), Dice coefficient, and 95th percentile Hausdorff Distance (HD95).

© Mehran University of Engineering and Technology 2026

The semi-supervised approach acquired a 21 per cent increment in IoU and 10.5 per cent improvement in the Dice score. It decreased HD95 by more than 33 pixels, indicating a major improvement in the accuracy of the lesion boundary. This performance can be compared to fully supervised models that are trained over 100 per cent labelled data. It confirms that leveraging unlabelled samples with the help of pseudo-label refinement is effective.

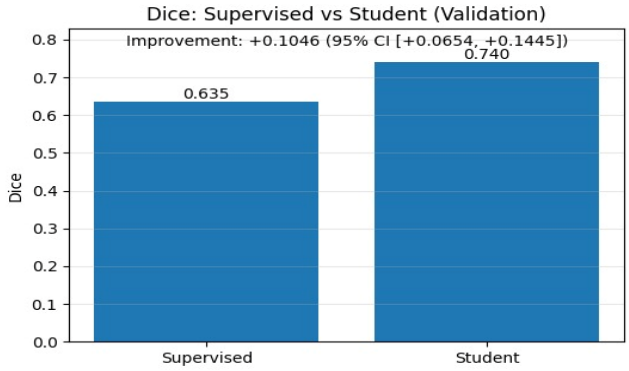

Fig. 9. Dice: Supervised vs Student

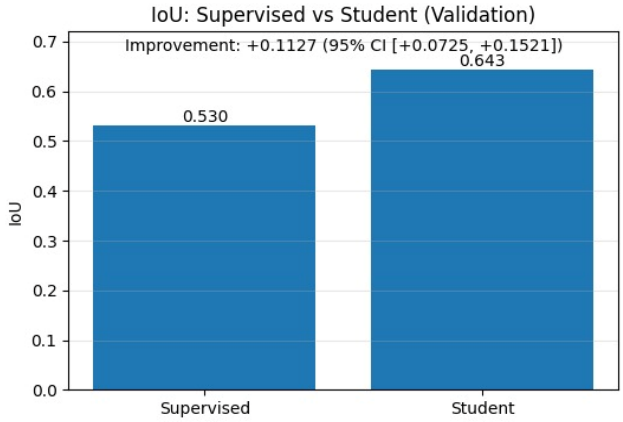

Fig. 10: IoU: Supervised vs Student

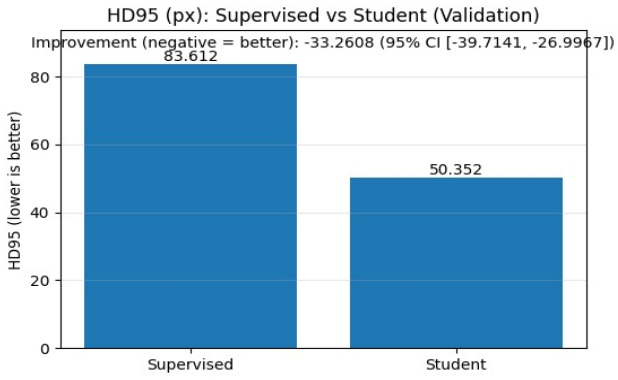

Fig. 11. HD95(px): Supervised vs Student

4.7. Qualitative Analysis

Representative visual comparisons are given in Figure 9, presenting (from right to left): Proposed Semi-Supervised Model Output, Supervised Model Output, Ground Truth Mask, Original Image.

The given figure provides a comparison of the outputs of segmentation between the proposed Semi-Supervised model of Teacher and Student with uncertainty-guided refinement of pseudo-labels and the Supervised U-Net baseline, which is trained with limited labelled data. The original dermoscopic image shows the sample of the lesion. The ground truth mask displays the expert annotation used for evaluation purposes. The supervised prediction provides the output from the baseline U-Net. Afterwards, the student prediction provides the refined segmentation acquired from the semi-supervised model. The supervised results show noise, incomplete coverage of the lesion, and boundary irregularities, particularly in images with hair occlusion or low contrast. On the contrary, the prediction of semi-supervised students presents more cohesive shapes of lesions and smoother contours. Moreover, it also presents better preservation of the lesion's structure and decreased false positives within the background regions.

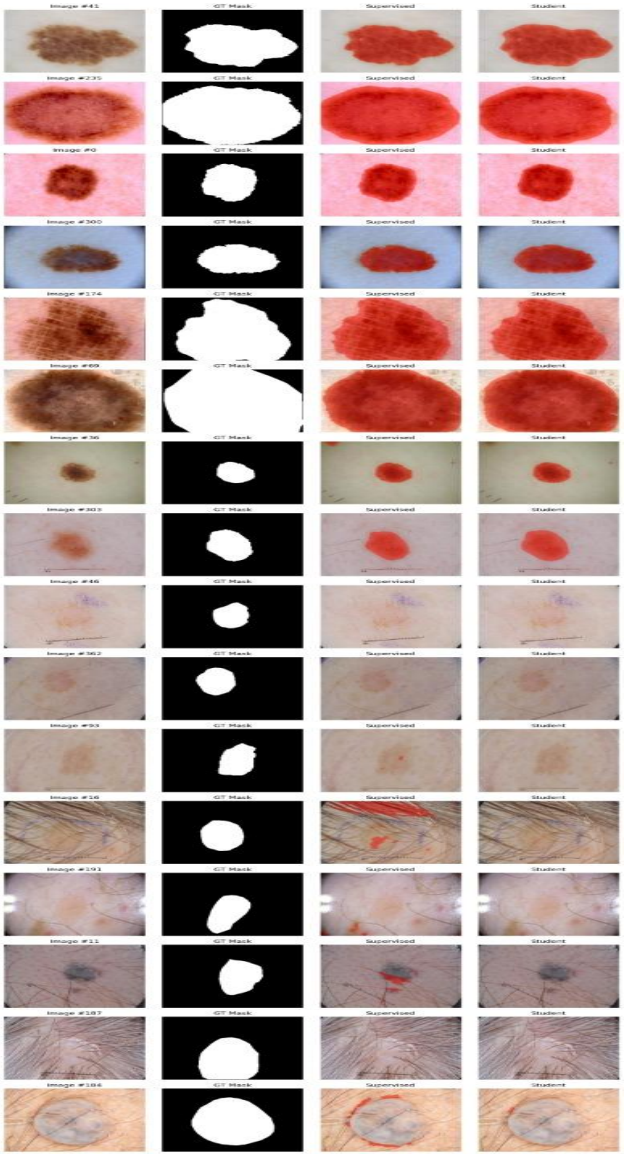

Fig. 12. Comparison of Outputs of Segmentation

Despite the cases of edge, the model still keeps a coherent shape and accurate lesion localisation. It confirms its robustness against lighting and structural variations. The figure acts as a qualitative error

© Mehran University of Engineering and Technology 2026

analysis. It complements the quantitative improvements given in IoU, Dice, and HD95 metrics.

4.8. Confidence and Uncertainty Visualization

For further investigation of the reliability of the model, uncertainty maps and pixel-wise probability were acquired. The internal representations from the given model show how uncertainty filtering helps in the refinement of pseudo-labels. Each of the rows provides one test sample, presenting the original image, probability map, uncertainty map, and overlay. High-confidence areas are in alignment with the lesion cores. On the other hand, uncertainty targets at ambiguous textures or edges. The refinement of these uncertain pixels is done, or they are discarded before re-training the student network. The column of overlay ensures that the process of filtering improves the precision of the boundary and decreases artefacts. The confidence-guided refinement is effective in making improvements in the quality of pseudo-labels, making sure that only reliable information is propagated to the student model.

4.9. Error and Boundary Analysis

The error and boundary analysis are used for the qualitative error analysis. It presents that even in the case of complex, large lesions or even in the case of low-contrast scenarios, the model can still maintain the general shape of the lesion, though some of the small artifacts still remain. The proposed framework maintains the structure of the lesion, but some minor deviations are seen in the case of complex areas with uneven occlusion or illumination.

4.10. Discussion

The proposed method effectively fills the gap between semi-supervised and fully supervised segmentation. Through linking CRF-based refinement and entropy-based confidence, the model learns to target reliable structures, while avoiding uncertain predictions. This results in improved consistency of lesion shape, improved robustness across datasets, and decreased false positives. The main implications that have emerged as a result are:

1. Annotation Efficiency: Near-supervised accuracy with only 10 percent of the labelled data decreases the cost of annotation substantially.
2. Scalability: The framework can be adapted easily to other medical fields as well, where there is limited labelled data.
3. Clinical Reliability: Anatomically consistent and cleaner masks support the accurate diagnosis and downstream classification.

5. Conclusion and Future Work

5.1. Conclusion

This research proposed a semi-supervised learning framework for the segmentation of skin lesions that effectively incorporates confidence-based filtering and pseudo-label refinement to overcome the challenges of annotation scarcity within medical imaging. The proposed approach follows a Teacher-Student paradigm, where the training of the Teacher model is done over a smaller labelled dataset, and it makes pseudo-labels for the unlabelled images. Afterwards, the refinement of pseudo-labels is done through entropy-based uncertainty filtering for the elimination of unreliable areas. Moreover, Conditional Random Field (CRF) post-processing is done for improving the boundary and spatial consistency. Afterwards, the refined labels are used subsequently for training the Student network, therefore permitting efficient use of both unlabelled and labelled data.

Experimental results on the datasets of HAM10000 and ISIC 2018 show that the proposed method acquires performance that can be compared to fully supervised models that are trained over complete annotation sets. Particularly, the framework achieved an IoU of 0.64, a Dice coefficient of 0.74, and a 33 percent decrement in the boundary when its training was done with 10 percent labelled data. Qualitative visualizations also assured improved robustness, lesion completeness, and boundary precision across different morphologies of lesions. The incorporation of CRF refinement and pseudo-label filtering not only improves the accuracy of the model but also improves interpretability through uncertainty visualizations and clear probabilities. Therefore, this work gives an annotation-efficient and scalable solution for the segmentation of medical images. It has a direct impact on clinical diagnostic systems, with scarce labelled data, but with larger repositories of unlabelled images.

5.2. Future *Work*

Building on the outcomes of this research, different research directions have been envisioned:

5.2.1. Unified Segmentation–Classification Framework

Future research will focus on the integration of the proposed segmentation model along with the classification branch (like Vision Transformer or EfficientNet-B0) to allow disease categorization and joint lesion delineation in one pipeline of deep learning.

© Mehran University of Engineering and Technology 2026

5.2.2. Metadata and Multimodal Fusion

The integration of clinical metadata like lesion site, sex, and age may further improve the accuracy of diagnosis through giving context-aware predictions.

5.2.3. Adaptive Uncertainty Thresholding

Instead of incorporating a fixed entropy threshold, Bayesian or adaptive mechanisms will be explored in the future to adjust confidence thresholds dynamically as per the complexity of the lesion.

5.2.4. Semi-Supervised Segmentation Under Domain Shift

The method will be expanded to deal with cross-domain generalization. The training will be done on 1 data set, like that of ISIC, and will be tested over unseen datasets like PH2 for the assessment of robustness under real-world variability of imaging.

6. References

[1]. Ronneberger, O., Fischer, P., and Brox, T., 2015, October. U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 234-241). Cham: Springer International Publishing.

[2]. Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N., and Liang, J., 2018, September. Unet++: A nested U-Net architecture for medical image segmentation. In International workshop on deep learning in medical image analysis (pp. 3-11). Cham: Springer International Publishing.

[3]. Oktay, O., Schlemper, J., Folgoc, L.L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N.Y., Kainz, B. and Glocker, B., 2018. Attention U-Net: Learning where to look for the pancreas. arXiv preprint arXiv:1804.03999.

[4]. Zhu, M., Zhang, L., Wang, L., Li, D., Zhang, J. and Yi, Z., 2023. Robust co-teaching learning with consistency-based noisy label correction for medical image classification. International Journal of Computer Assisted Radiology and Surgery, 18(4), pp.675-683.

[5]. Sarfati, E., 2025. Exploitation of heterogeneous labels in deep learning for medical image analysis: applications to automated diagnosis of liver diseases (Doctoral dissertation, Institut Polytechnique de Paris).

[6]. Zhang, J., You, S., Liu, A., Xie, L., Huang, C., Han, X., Li, P., Wu, Y. and Deng, J., 2024. Winter wheat mapping method based on Pseudo-labels and a U-Net model for training sample shortage. Remote Sensing, 16(14), p.2553.

[7]. Shuai, Z., Chen, Y., Mao, S., Zhou, Y., and Zhang, X., 2024. Diffseg: a segmentation model for skin lesions based on diffusion difference. arXiv preprint arXiv:2404.16474.

[8]. Chatterjee, P., Sharma, K.D., and Chakrabarti, A., 2025. Refining Pseudo Labels with a Teacher-Student Paradigm for Accurate Brain Lesion Segmentation. IEEE Access.

[9]. Ronneberger, O., Fischer, P., and Brox, T., 2015, October. U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 234-241). Cham: Springer International Publishing.

[10]. Oktay, O., Schlemper, J., Folgoc, L.L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N.Y., Kainz, B. and Glocker, B., 2018. Attention U-Net: Learning where to look for the pancreas. arXiv preprint arXiv:1804.03999.

[11]. Esteva, A., Kuprel, B., Novoa, R.A., Ko, J., Swetter, S.M., Blau, H.M. and Thrun, S., 2017. Dermatologist-level classification of skin cancer with deep neural networks. nature, 542(7639), pp.115-118.

[12]. Bai, W., Oktay, O., Sinclair, M., Suzuki, H., Rajchl, M., Tarroni, G., Glocker, B., King, A., Matthews, P.M., and Rueckert, D., 2017, September. Semi-supervised learning for network-based cardiac MR image segmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 253-260). Cham: Springer International Publishing.

[13]. Li, X., Yu, L., Chen, H., Fu, C.W., Xing, L., and Heng, P.A., 2020. Transformation-consistent self-ensembling model for semisupervised medical image segmentation. IEEE Transactions on Neural Networks and Learning Systems, 32(2), pp.523-534.

[14]. Kendall, A. and Gal, Y., 2017. What uncertainties do we need in Bayesian deep learning for computer vision?. Advances in neural information processing systems, 30.

© Mehran University of Engineering and Technology 2026

[15]. Gal, Y. and Ghahramani, Z., 2016, June. Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In International Conference on Machine Learning (pp. 1050-1059). PMLR.

[16]. Milton, M.A.A., 2019. Automated skin lesion classification using an ensemble of deep neural networks in ISIC 2018: Skin lesion analysis towards melanoma detection challenge. arXiv preprint arXiv:1901.10802.

[17]. Tschandl, P., Rosendahl, C. and Kittler, H., 2018. The HAM10000 dataset is a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Scientific data, 5(1), pp.1-9.

[18]. Wang, Z., Wang, P., Liu, K., Wang, P., Fu, Y., Lu, C.T., Aggarwal, C.C., Pei, J., and Zhou, Y., 2025. A comprehensive survey on data augmentation. IEEE Transactions on Knowledge and Data Engineering.

[19]. Razavi, M., Mavaddati, S. and Koohi, H., 2024. ResNet deep models and transfer learning technique for classification and quality detection of rice cultivars. Expert Systems with Applications, 247, p.123276.

[20]. Leonardi, R., Giudice, A.L., Farronato, M., Ronsivalle, V., Allegrini, S., Musumeci, G. and Spampinato, C., 2021. Fully automatic segmentation of sinonasal cavity and pharyngeal airway based on convolutional neural networks. American Journal of Orthodontics and Dentofacial Orthopedics, 159(6), pp.824-835.

[21]. Hosseini, S.M. and Baghshah, M.S., 2024. Dilated balanced cross-entropy loss for medical image segmentation. arXiv preprint arXiv:2412.06045.

[22]. Xu, W., Yuan, K., Li, W., and Ding, W., 2022. An emerging fuzzy feature selection method using composite entropy-based uncertainty measure and data distribution. IEEE Transactions on Emerging Topics in Computational Intelligence, 7(1), pp.76-88.

[23]. Zhu, J., Huang, H., Li, B., and Wang, L., 2021. E-CRF: Embedded conditional random field for boundary-caused class weights confusion in semantic segmentation. arXiv preprint arXiv:2112.07106.

[24]. Ataş, İ., 2023. Performance evaluation of the Jaccard-Dice coefficient on building segmentation from high-resolution satellite images. Balkan Journal of Electrical and Computer Engineering, 11(1), pp.100-106.

© Mehran University of Engineering and Technology 2026