Uncertainty-Aware Domain Adaptation for Vitiligo Segmentation in Clinical Photographs

Wentao Jiang[1] WENTAO.JIANG@RICE.EDU
Vamsi Varra[2] * VAMSIVARRA@GMAIL.COM
Caitlin Perez-Stable[3] CAITLIN.PEREZ-STABLE@BCM.EDU
Harrison Zhu[3] HARRISON.ZHU@BCM.EDU
Meredith Apicella[3] MEREDITH.APICELLA@BCM.EDU
Nicole Nyamongo[3] NICOLE.NYAMONGO@BCM.EDU
[1] *Electrical and Computer Engineering, Rice University, Houston, TX, United States*
[2] *U.S. Dermatology Partners, Houston, TX, United States*
[3] *School of Medicine, Baylor College of Medicine, Houston, TX, United States*

Editors: Under Review for MIDL 2026

Abstract

Accurately quantifying vitiligo extent in routine clinical photographs is crucial for longitudinal monitoring of treatment response. We propose a trustworthy, frequency-aware segmentation framework built on three synergistic pillars: (1) a domain-adaptive pretraining strategy on the ISIC 2019 dermoscopy dataset; (2) an architectural refinement via a ConvNeXt V2-based encoder enhanced with a novel High-Frequency Spectral Gating (HFSC) module and stem-skip connections to capture subtle textures; and (3) a clinical trust mechanism employing K-fold ensemble and Test-Time Augmentation (TTA) to generate pixel-wise uncertainty maps. Extensive validation on an expert-annotated clinical cohort demonstrates superior performance, achieving a Dice score of 85.05% and significantly reducing boundary error (95% Hausdorff Distance improved from 44.79 px to 29.95 px), consistently outperforming strong CNN (ResNet-50 and UNetPP) and Transformer (MiT-B5) baselines. Notably, our framework demonstrates high reliability with zero catastrophic failures and provides interpretable entropy maps to identify ambiguous regions for clinician review. This work establishes a robust and practical tool for objective vitiligo assessment. Our approach suggests that the proposed framework is a practical and reliable component for vitiligo management in routine clinical practice.

Keywords: Vitiligo Segmentation, Skin lesion segmentation, Clinical Photography, Frequency-domain Analysis, Uncertainty Quantification, Domain Adaptation, Trustworthy AI.

1. Introduction

Vitiligo is a chronic autoimmune disorder characterized by melanocyte destruction, leading to distinct depigmented patches. Affecting 0.5–2% of the global population, it imposes a substantial psychosocial burden, particularly on individuals with darker skin tones or facial involvement (Bergqvist and Ezzedine, 2020; Picardo et al., 2022). In clinical practice, treatment efficacy is primarily tracked through semi-quantitative indices such as the Vitiligo Area Scoring Index (VASI), which rely on visual estimation of lesion extent (Hamzavi et al., 2004; Bibeau et al., 2024). However, such manual assessments are inherently subjective and

* Corresponding author: vamsivarra@gmail.com

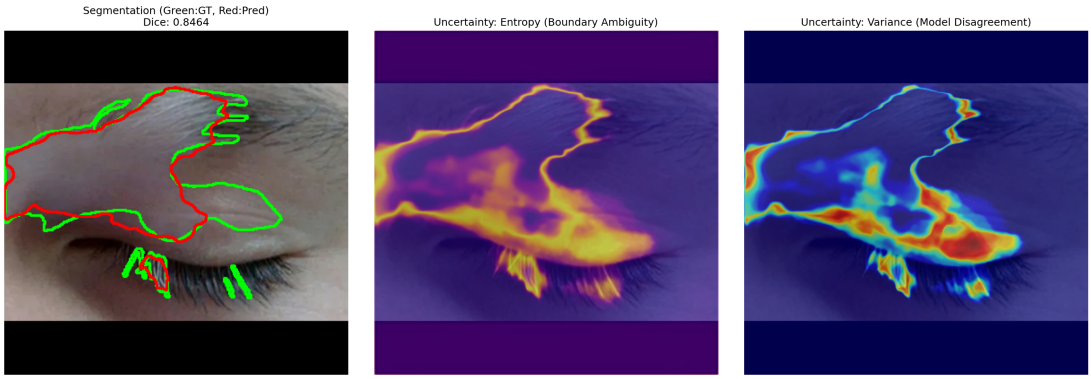


Figure 1: **Visual Abstract.** Overview of the proposed framework addressing boundary ambiguity and trustworthy assessment in vitiligo segmentation.

exhibit significant inter-observer variability, impeding standardized longitudinal monitoring (Nugroho et al., 2013). This underscores a well-recognized need for automated, objective measurement tools to support vitiligo management (Mazzetto et al., 2025).

Automated segmentation of vitiligo lesions from clinical photographs presents a promising solution. Early approaches relied on semi-automatic techniques such as region growing and interactive thresholding (Nugroho et al., 2013; Neri et al., 2020), while recent work has demonstrated that convolutional neural networks (CNNs) can achieve superior accuracy (Low et al., 2020; Owida et al., 2025; Fan and Wang, 2024; Biswas et al., 2025). Despite this progress, most models are trained on limited, single-center cohorts acquired under controlled conditions. Their performance often degrades in real-world settings due to heterogeneous lighting, complex backgrounds (e.g., clothing, medical equipment), and diverse skin tones. This is a manifestation of the pervasive domain-shift challenge in medical imaging (Fogelberg et al., 2023; Gilani et al., 2024), highlighting the need for strategies that explicitly enhance model generalization beyond curated datasets.

In parallel, many recent methods for dermoscopic skin lesion analysis have introduced design ideas that are directly useful for our task. State-of-the-art methods have shown that explicitly modeling high-frequency spectral information is crucial for resolving ambiguous boundaries (Hu et al., 2025; Tang et al., 2024; Thachankattil and Sujith, 2025), and that leveraging large-scale pre-training or foundation models is key to overcoming data scarcity (Zhou et al., 2026; Kirillov et al., 2023). However, directly translating these dermoscopy-centric advances to vitiligo photography remains non-trivial for three reasons. First, the imaging context is fundamentally different: vitiligo lesions are often multifocal, partially occluded, and embedded in cluttered, color-inhomogeneous clinical backgrounds, unlike centered lesions on roughly homogeneous fields in dermoscopy (Fogelberg et al., 2023). Second, most vitiligo-specific models operate purely in the spatial domain. They suffer from an inherent **"spectral bias,"** prioritizing low-frequency shapes while neglecting the high-frequency cues that define subtle depigmentation transitions and boundary fuzziness (Hu et al., 2025). Third, nearly all existing methods function as deterministic "black boxes" that provide no estimate of predictive confidence, leaving clinicians unable to distinguish reliable segmentations from potential failures—a major barrier to trustworthy deployment in

settings where out-of-distribution cases (e.g., extreme phototypes or low-quality smartphone images) are inevitable (Tsaneva-Atanasova et al., 2025; Li et al., 2025).

To address these gaps, we propose a trustworthy, frequency-aware segmentation framework specifically engineered for vitiligo assessment from clinical photographs. Our design is structured around three synergistic pillars: (i) architectural refinement for fine-grained boundary recovery, (ii) a data-efficient training strategy that leverages dermatology-specific priors to mitigate domain shift, and (iii) an uncertainty-aware inference pipeline that provides pixel-wise confidence maps for clinical quality control.

Our main contributions are as follows:

- **Frequency-enhanced architectural refinement.** We introduce a Frequency-aware Dual-branch Network built upon a ConvNeXt V2 encoder, equipped with a novel High-Frequency Spectral Gating (HFSC) module and stem-skip connections. By explicitly modeling high-frequency spectral components and propagating early texture cues, the network is able to delineate indistinct lesion rims and subtle depigmentation transitions.

- **Anatomy-guided, data-efficient training paradigm.** We design a robust training regimen that combines domain-adaptive pre-training on the large-scale ISIC 2019 dataset with an ROI-constrained dual-task loss. By utilizing dynamic skin masks to suppress background interference, this strategy functions as an implicit **anatomy-guided hard negative mining** mechanism, forcing the model to focus optimization on the challenging decision boundary between lesion and normal skin.

- **Trustworthy uncertainty quantification.** Moving beyond deterministic prediction, we implement a test-time augmentation (TTA) and K-fold ensemble pipeline to generate pixel-wise uncertainty maps. This system exposes regions of low predictive confidence (capturing both epistemic and aleatoric uncertainty), providing clinicians with an interpretable "risk map" for quality assurance, aligning with emerging frameworks for trustworthy clinical AI (Tsaneva-Atanasova et al., 2025; Wang et al., 2019).

Extensive validation on an expert-annotated clinical cohort demonstrates that our framework achieves superior segmentation accuracy (Dice score of 85.05%) and boundary precision (HD95 of 29.95 px) compared with both heavyweight CNN and Transformer baselines. Moreover, the resulting uncertainty maps highlight challenging regions and low-confidence cases, supporting more reliable and interpretable quantitative assessment of vitiligo in routine clinical practice.

2. Methodology

2.1. Overview

The proposed framework, illustrated in Fig. 2, represents a frequency-aware, uncertainty-quantified segmentation system tailored for clinical vitiligo assessment. The system is founded on a **Domain-Adapted Encoder** (Fig. 2, Panel A), which is strategically initialized with large-scale dermatological priors from the ISIC 2019 dataset. This pre-training serves as the critical backbone, enabling robust feature extraction despite the scarcity

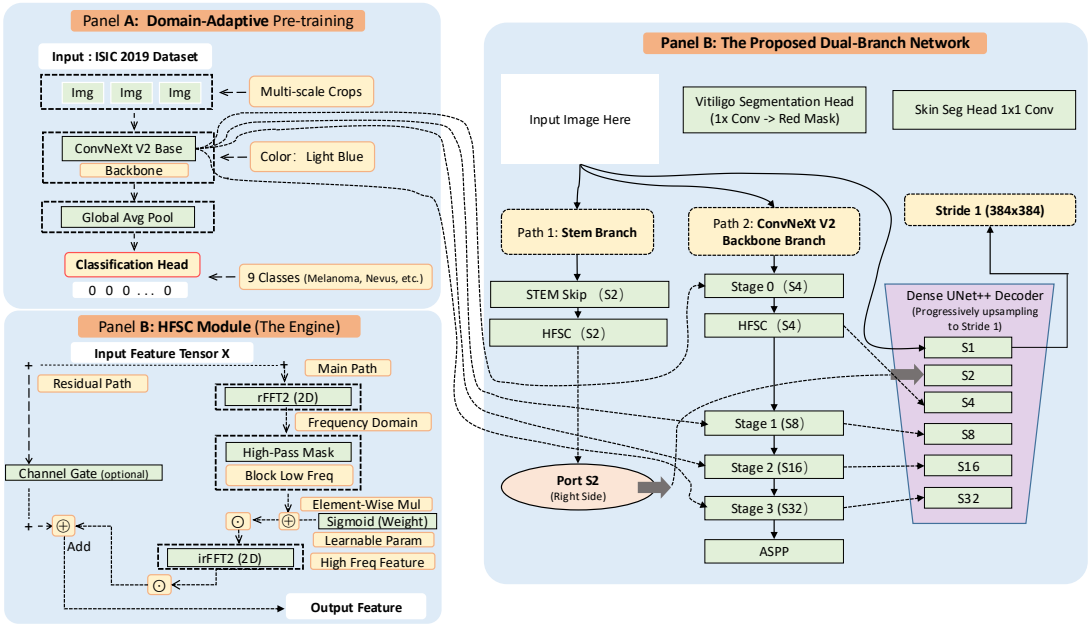


Figure 2: **Overview of the proposed Trustworthy Frequency-aware Segmentation Framework.** The system is grounded on three synergistic pillars: **(A) Domain-Adaptive Pre-training (Top-Left):** The ConvNeXt V2 encoder is initialized with ISIC 2019 dermatological priors to mitigate domain shift. **(B) High-Frequency Spectral Gating (Bottom-Left):** A novel HFSC module explicitly filters spectral components via FFT to recover fine-grained boundary harmonics lost in downsampling. **(C) Main Architecture (Right):** The encoder features (enhanced by HFSC at shallow stages) are aggregated via a dense U-Net++ decoder. The network is optimized using an *Anatomy-Guided Dual-Task Strategy*, where an auxiliary skin head constrains the learning of the vitiligo head. During inference, a TTA-based ensemble generates pixel-wise **Uncertainty Maps** for clinical quality control.

of annotated vitiligo data. Complementing this, a **High-Frequency Spectral Gating (HFSC)** module (Fig. 2, Panel B) is introduced to explicitly recover fine-grained texture cues lost during downsampling. These multi-scale features are aggregated via a dense **U-Net++ Decoder** (Right, Panel C) and optimized using an **Anatomy-Guided Dual-Task Strategy**. Finally, a test-time augmentation (TTA) ensemble is employed to ensure trustworthy, uncertainty-aware inference.

2.2. Frequency-aware Dual-Branch Architecture

2.2.1. Domain-Adapted Encoder: Bridging the Data Gap

Training deep segmentation networks on limited clinical cohorts often leads to overfitting. Furthermore, standard ImageNet pre-training introduces a domain discrepancy, as object-centric features differ fundamentally from the amorphous patterns of skin lesions. To overcome this, we employ a **ConvNeXt V2** backbone explicitly pre-trained on the massive

ISIC 2019 dermoscopy dataset. By transferring these rich dermatological representations, we establish a strong initialization "anchor." To address rapid spatial downsampling, we further introduce a **Stem-Skip connection**. This path preserves the full-resolution spatial topology from the input stem (Stride 2), preventing the loss of fine-grained boundary cues before they reach deep semantic layers.

2.2.2. High-Frequency Spectral Gating (HFSC)

Standard convolutional operators are inherently limited by their local receptive fields and behave as low-pass filters, creating a "spectral bias" that progressively smooths out high-frequency details. This is particularly detrimental for vitiligo segmentation, where the distinction between lesion and skin relies on sharp boundary gradients. To overcome this, we introduce the **High-Frequency Spectral Gating (HFSC)** module. Unlike spatial attention (e.g., CBAM), HFSC leverages the **global receptive field** of the Fourier domain. The module operates in three rigorous steps:

Frequency Transformation. We first transform the spatial features into the spectral domain using the 2D Real Fast Fourier Transform (rFFT). This converts local pixel intensities into global frequency components:

$$\mathcal{F}(X)(u,v) = \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} X(h,w) e^{-j2\pi(\frac{uh}{H}+\frac{vw}{W})} \tag{1}$$

where (u, v) represent the frequency coordinates.

Learnable Spectral Gating. To isolate lesion boundaries, we design a high-pass filtering mechanism. We construct a static high-pass mask M_{high} to attenuate low-frequency components. Crucially, we introduce a **learnable channel-wise gating parameter** $\mathcal{W}_{gate}$ to dynamically modulate the intensity of the high-frequency injection:

$$\tilde{X}_{freq} = \mathcal{F}(X) \odot M_{high} \odot \sigma(\mathcal{W}_{gate}) \tag{2}$$

where $\odot$ denotes element-wise multiplication. This allows the network to selectively amplify informative texture cues while suppressing spectral noise.

Dual-Domain Fusion. The modulated spectrum is transformed back to the spatial domain via Inverse FFT (IFFT). These recovered high-frequency details are then injected into the spatial stream via a residual connection. Crucially, we employ a **Channel Attention** mechanism $A_{ch}(\cdot)$ to adaptively re-weight the high-frequency features before fusion:

$$X_{out} = X + \text{IFFT}(\tilde{X}_{freq}) \odot A_{ch}(X) \tag{3}$$

where $\odot$ denotes element-wise multiplication with channel-wise broadcasting. This design creates a **dual-domain representation**, leveraging the channel attention to selectively emphasize boundary cues that are most relevant to the segmentation task while suppressing noise.

2.3. Multi-scale Decoder and Aggregation

The extracted multi-scale features are aggregated via a decoder to generate the final segmentation map. Following [Zhou et al.], we employ the **U-Net++ architecture**, characterized by dense nested skip connections, to recover fine-grained details. *(Note: For visual clarity, the schematic representation in Fig. 2 simplifies these dense internal connections to highlight the overall multi-scale information flow.)*

Specifically, we integrate an Atrous Spatial Pyramid Pooling (ASPP) module at the bottleneck (Stride 32) to capture multi-scale contextual information. The decoder outputs two parallel prediction maps: a primary *Vitiligo Map* and an auxiliary *Skin Map*, which are used to compute the ROI-constrained loss.

2.4. ROI-Constrained Dual-Task Optimization Strategy

Standard segmentation losses treat all pixels equally, allowing background artifacts to dominate gradients. To address this, we propose an **ROI-Constrained Optimization Strategy**. We leverage the auxiliary Skin Head to generate a dynamic region-of-interest (ROI) mask, forcing the Vitiligo Head to learn exclusively within valid skin areas.

2.4.1. Formulation

Let $\hat{y}_{vit}$ and y_{vit} denote the predicted probability and ground truth for vitiligo, respectively, and $M_{skin} \in \{0, 1\}$ denote the ground truth skin mask. The standard Focal Loss is defined as $\mathcal{L}_{focal}$. We formulate the **Masked Focal Loss** as:

$$\mathcal{L}_{masked\_focal} = \frac{\sum_{i,j} M_{skin}^{(i,j)} \cdot \mathcal{L}_{focal}(\hat{y}_{vit}^{(i,j)}, y_{vit}^{(i,j)})}{\sum_{i,j} M_{skin}^{(i,j)} + \epsilon} \tag{4}$$

where ϵ is a smoothing term. Similarly, the **Masked Dice Loss** is formulated to maximize the overlap only within the skin region:

$$\mathcal{L}_{masked\_dice} = 1 - \frac{2 \sum_{i,j} (M_{skin}^{(i,j)} \cdot \hat{y}_{vit}^{(i,j)} \cdot y_{vit}^{(i,j)}) + \epsilon}{\sum_{i,j} (M_{skin}^{(i,j)} \cdot \hat{y}_{vit}^{(i,j)} + M_{skin}^{(i,j)} \cdot y_{vit}^{(i,j)}) + \epsilon} \tag{5}$$

To prevent hallucinations in non-skin areas, we introduce a background suppression term $\mathcal{L}_{bg}$, which imposes a binary cross-entropy penalty on the vitiligo head for all pixels outside the skin mask ($M_{bg} = 1 - M_{skin}$):

$$\mathcal{L}_{bg} = \frac{\sum_{i,j} M_{bg}^{(i,j)} \cdot \text{BCE}(\hat{y}_{vit}^{(i,j)}, 0)}{\sum_{i,j} M_{bg}^{(i,j)} + \epsilon} \tag{6}$$

The final objective function is a weighted sum:

$$\mathcal{L}_{total} = \lambda_1 \mathcal{L}_{masked\_focal} + \lambda_2 \mathcal{L}_{masked\_dice} + \lambda_3 \mathcal{L}_{bg} + \lambda_4 \mathcal{L}_{skin\_aux} \tag{7}$$

2.4.2. Theoretical Interpretation: Anatomy-Guided Hard Negative Mining

This strategy implicitly functions as an **anatomy-guided hard negative mining** mechanism. In clinical images, pixels can be categorized into "Easy Negatives" (background artifacts) and "Hard Negatives" (normal skin). Training on the full image allows easy negatives to dominate the gradient since they constitute the majority of pixels. Our dynamic skin mask M_{skin} effectively zeros out the loss contribution from these easy negatives. This compels the network to focus its optimization capacity entirely on distinguishing vitiligo from normal skin (the challenging decision boundary), thereby significantly enhancing discriminative power in ambiguous transition zones.

2.5. Trustworthy Inference Pipeline

To transition from a deterministic "black box" to a trustworthy clinical tool, we implement an uncertainty estimation pipeline.

K-Fold Ensemble. We aggregate predictions from $K = 5$ models trained on different data subsets to capture epistemic uncertainty (model knowledge gaps).

Test-Time Augmentation (TTA). For each input, we generate $N = 8$ variations (rotations and flips) to capture aleatoric uncertainty (data noise).

The final prediction is the mean of these $K \times N$ probabilistic outputs. We quantify pixel-wise uncertainty using **Predictive Entropy**:

$$H(x) = -\sum_{c \in \{0,1\}} p(y = c|x) \log p(y = c|x) \tag{8}$$

High entropy values indicate regions of low confidence, serving as a visual alert for clinicians to perform manual review.

3. Experiments and Results

3.1. Datasets and Experimental Setup

To evaluate the efficacy and robustness of our framework, we curated a comprehensive experimental setup involving two distinct imaging domains: standard clinical photography (macro) and dermoscopy (micro).

Primary Clinical Cohort (In-Distribution). Our primary dataset comprises RGB clinical photographs of vitiligo, curated from the repository established by Zhang et al. (Zhang et al., 2021). Unlike dermoscopic images, these samples represent "in-the-wild" clinical settings, characterized by multifocal lesions, complex backgrounds (e.g., clothing), and varying lighting conditions. The dataset (total 978 files) was partitioned using a strict patient-level split to prevent data leakage:

- **Development Set (839 images):** Used for 5-fold cross-validation. Crucially, this set includes **198 negative background patches** (regions with no lesions) derived from healthy skin areas. Including these "hard negatives" during training forces the model to discriminate between normal skin textures and actual vitiligo, rather than simply segmenting any skin-colored area.

- **Hold-out Test Set (118 images):** Comprising 15% of the patients. We explicitly removed the artificial negative patches from this set to simulate a realistic clinical diagnostic scenario where the focus is on active lesion assessment.

External Dermoscopic Cohort (Out-of-Distribution). To rigorously stress-test the model's generalization capability, we employed an external dataset consisting of dermoscopic vitiligo images, originally curated by Biswas et al. (Biswas et al., 2025). This dataset introduces a severe **domain shift**: unlike clinical photos, dermoscopic images feature magnified textures, uniform illumination, and gel artifacts. Evaluating on this cohort without fine-tuning allows us to verify whether our architecture learns transferable dermatological features rather than overfitting to the specific visual style of clinical cameras.

3.2. Implementation Details

Data Augmentation. To mitigate overfitting and simulate clinical acquisition variability, we employed a robust augmentation pipeline uniformly across all models. Strategies included: (i) **geometric transformations** (random flips, 90°/180° rotations, elastic deformation); (ii) **photometric jittering** (brightness, contrast, HSV shifts) to handle lighting inconsistencies; and (iii) **dynamic multi-scale training** (0.75×–1.25×) to ensure robustness to varying camera distances.

Training Configuration. The framework was implemented in PyTorch and trained on a single NVIDIA A100 GPU. We employed the AdamW optimizer with a weight decay of $1e-4$. The learning rate was scheduled using a cosine annealing strategy, starting from $3e-4$ (encoder) and $1e-3$ (decoder) after a 3-epoch linear warmup. The loss weights were empirically set to $\lambda_1 = 1.0$ (Focal), $\lambda_2 = 1.0$ (Dice), $\lambda_3 = 0.1$ (Background), and $\lambda_4 = 0.5$ (Skin Aux).

3.3. Comparative Analysis with SOTA Methods

We compared our framework against SOTA CNN (ResNet50-UNet, UNet++ Vanilla) and Transformer-based (MiT-B2/B5) segmentation models. As shown in Table 1, our method achieves superior performance across all metrics.

Quantitative Superiority. Our framework achieves a Dice score of **85.05%** and a Recall of **86.70%**, consistently outperforming the heavy-weight Transformer baseline (MiT-B5, Dice 84.07%). Most notably, we achieved a significant reduction in boundary error, lowering the **95% Hausdorff Distance (HD95)** from 44.79 px (ResNet50) and 30.90 px (MiT-B5) to **29.95 px**. This validates that our frequency-aware design effectively resolves the boundary ambiguity that challenges standard architectures. Furthermore, our method demonstrated high reliability with a **0.0% failure rate** on the test set.

3.4. Ablation Study: Dissecting the Three Pillars

To validate the effectiveness of our proposed components, we conducted a step-by-step ablation study (Table 2).

Table 1: **Comparison with State-of-the-Art Methods.** Best results are highlighted in **bold**.

Method	Dice (↑)	IoU (↑)	Recall (↑)	HD95 (px, ↓)
ResNet50-UNet	80.24	70.00	82.83	44.79
UNet++ Vanilla	81.70	72.06	81.64	39.68
MiT-B2-UNet (Transformer)	83.63	74.71	**86.77**	30.46
MiT-B5-UNet (Transformer)	84.07	75.01	86.76	30.90
Ours (ConvNeXtV2++ HFSC)	**85.05**	**75.98**	86.70	**29.95**

1. Impact of Domain Adaptation (M1 vs. M2). Training from scratch (M1) yielded suboptimal performance (Dice 72.50%). Incorporating ISIC 2019 pre-training (M2) led to a massive **11.5% improvement in Dice** and a dramatic reduction in HD95 (49.7 px → 30.9 px). This confirms that domain knowledge transfer is the *fundamental prerequisite* for this task, serving as the "anchor" for effective feature learning.

2. Failure of Standard Spatial Attention (M2 vs. M4). Interestingly, directly applying generic spatial-channel attention (CBAM, M4) resulted in performance degradation (Dice dropped to 83.09%). This suggests that generic attention modules may overfit to high-frequency noise (e.g., hairs, artifacts) in dermatological images rather than focusing on lesion semantics.

3. Effectiveness of Frequency Gating (M4 vs. M5). In contrast, our proposed **HFSC module (M5)** successfully reversed this trend, boosting the Dice to **84.72%**. Unlike CBAM, HFSC explicitly filters spectral components, allowing the model to enhance structural boundaries without being distracted by noise. M5 achieves the highest Recall among all variants.

4. Structure vs. Detail (M3 vs. M5). While adding ASPP (M3) improved context modeling, combining it with HFSC (M5) yielded the best overall performance. Although the raw HD95 of M5 (30.76 px) is comparable to M3 due to the model's high sensitivity to fine details, the visual quality demonstrates superior boundary continuity.

Table 2: **Ablation Study.** M1: Random Init. M2: ISIC Init. M3: +ASPP. M4: +CBAM. M5: +HFSC (Ours).

ID	Pre-train	Context	Attention	Dice	Recall	HD95
M1	Random	-	-	72.50	72.12	49.70
M2	**ISIC**	-	-	84.01	85.51	30.96
M3	ISIC	ASPP	-	84.56	**85.89**	**29.46**
M4	ISIC	-	CBAM	83.09	84.97	33.58
M5	**ISIC**	**ASPP**	**HFSC**	**84.72**	**85.89**	30.76

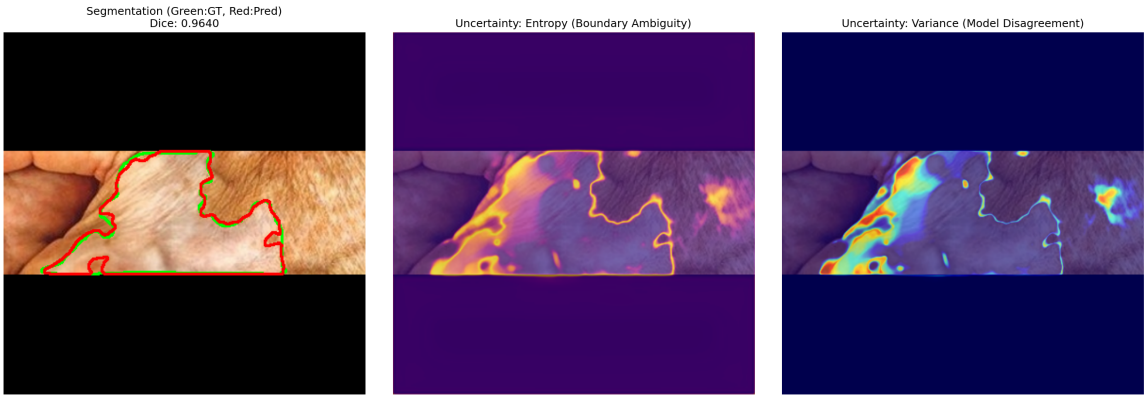


Figure 3: **Trustworthy Inference.** Left: Segmentation result (Green: GT, Red: Prediction, Dice: 0.8464). Middle: Predictive Entropy Map highlighting boundary ambiguity. Right: Variance Map showing model disagreement.

3.5. Qualitative Analysis and Generalization

Uncertainty-Aware Assessment. Fig. 3 visualizes the model's predictive confidence. The **Entropy Map** (Middle) clearly highlights the "boundary ambiguity" zones with warm colors, strictly correlating with the fuzzy transition regions of the lesion. The **Variance Map** (Right) captures the disagreement between ensemble models, effectively flagging potential failure modes. This provides clinicians with an interpretable "risk map" for quality assurance.

Generalization to Dermoscopy (Zero-shot). To verify robustness, we evaluated the model on the external dermoscopic dataset without any fine-tuning. As shown in Fig. 4, despite the severe domain shift (magnified texture, gel artifacts), our model accurately delineates the lesions. This confirms that the **Frequency-aware Dual-Branch** architecture has learned transferable dermatological semantics rather than overfitting to the clinical photography domain.

4. Conclusion

In conclusion, we presented a trustworthy, frequency-aware framework that effectively resolves the challenges of boundary ambiguity and domain shift via spectral gating and anatomy-guided learning, establishing a robust and interpretable tool for objective vitiligo assessment in routine clinical practice.

Data Availability Statement

The Primary Clinical Cohort is publicly available via Zhang et al. (Zhang et al., 2021). Regarding the External Dermoscopic Cohort (Biswas et al., 2025), while the original repository access may be intermittent, the specific subset used for validation has been archived by the authors. **The raw data, along with the precise data split indices and eval-**

uation protocols, are available from the corresponding author upon reasonable request.

References

Christina Bergqvist and Khaled Ezzedine. Vitiligo: A Review. *Dermatology (Basel, Switzerland)*, 236(6):571–592, 2020. ISSN 1421-9832. doi: 10.1159/000506103.

Kristen Bibeau, Kathleen Butler, Mingyue Wang, Konstantina Skaltsa, and Iltefat H. Hamzavi. Psychometric Evaluation of the Facial and Total Vitiligo Area Scoring Index Instruments in the TRuE-V Phase 3 Studies. *Dermatology and Therapy*, 14(8):2223–2234, August 2024. ISSN 2193-8210. doi: 10.1007/s13555-024-01223-y.

Rahul Biswas, Md. Faruk Abdullah Al Sohan, Rifat Al Mamun Rudro, and Shanaj Parvin. VGG19-ResSE: an optimized hybrid model for accurate segmentation and classification of vitiligo lesions. *Neural Computing and Applications*, 37(32):27227–27248, November 2025. ISSN 1433-3058. doi: 10.1007/s00521-025-11670-z. URL https://doi.org/10.1007/s00521-025-11670-z.

Zhuangzhuang Fan and Changqing Wang. Deep Learning-Based Segmentation of Lesions from Wide-Field Vitiligo Images. *International Journal of Pattern Recognition and Artificial Intelligence*, 38(10):2457007, August 2024. ISSN 0218-0014. doi: 10.1142/S0218001424570076. URL https://www.worldscientific.com/doi/10.1142/S0218001424570076. Publisher: World Scientific Publishing Co.

Katharina Fogelberg, Sireesha Chamarthi, Roman C. Maron, Julia Niebling, and Titus J. Brinker. Domain shifts in dermoscopic skin cancer datasets: Evaluation of essential limitations for clinical translation. *New Biotechnology*, 76:106–117, September 2023. ISSN 1871-6784. doi: 10.1016/j.nbt.2023.04.006. URL https://www.sciencedirect.com/science/article/pii/S1871678423000213.

Syed Qasim Gilani, Muhammad Umair, Maryam Naqvi, Oge Marques, and Hee-Cheol Kim. Adversarial Training Based Domain Adaptation of Skin Cancer Images. *Life*, 14(8):1009, August 2024. ISSN 2075-1729. doi: 10.3390/life14081009. URL https://pmc.ncbi.nlm.nih.gov/articles/PMC11355601/.

Iltefat Hamzavi, Hem Jain, David McLean, Jerry Shapiro, Haishan Zeng, and Harvey Lui. Parametric modeling of narrowband UV-B phototherapy for vitiligo using a novel quantitative tool: the Vitiligo Area Scoring Index. *Archives of Dermatology*, 140(6):677–683, June 2004. ISSN 0003-987X. doi: 10.1001/archderm.140.6.677.

Gefeng Hu, Wen Zhu, Xinyi Liao, and Qingbo Li. WA-NET: enhanced boundary-aware segmentation of skin lesions via frequency-spatial feature fusion and attention-guided edge refinement. *Scientific Reports*, 15(1):41598, November 2025. ISSN 2045-2322. doi: 10.1038/s41598-025-25583-3. URL https://www.nature.com/articles/s41598-025-25583-3. Publisher: Nature Publishing Group.

Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C. Berg, Wan-Yen Lo, Piotr Dollár, and Ross Girshick. Segment anything, 2023. URL https://arxiv.org/abs/2304.02643.

Shiman Li, Mingzhi Yuan, Xiaokun Dai, and Chenxi Zhang. Evaluation of uncertainty estimation methods in medical image segmentation: Exploring the usage of uncertainty in clinical deployment. *Computerized Medical Imaging and Graphics: The Official Journal of the Computerized Medical Imaging Society*, 124:102574, September 2025. ISSN 1879-0771. doi: 10.1016/j.compmedimag.2025.102574.

Makena Low, Victor Huang, and Priyanka Raina. Automating Vitiligo Skin Lesion Segmentation Using Convolutional Neural Networks. In *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, pages 1–4, April 2020. doi: 10.1109/ISBI45749.2020.9098682. URL https://ieeexplore.ieee.org/document/9098682. ISSN: 1945-8452.

Roberto Mazzetto, Alvise Sernicola, Jacopo Tartaglia, Christian Ciolfi, and Mauro Alaibac. Potential of automated image analysis for the measurement of vitiligo lesions. *Frontiers in Medicine*, 12, August 2025. ISSN 2296-858X. doi: 10.3389/fmed.2025.1623408. URL https://www.frontiersin.org/journals/medicine/articles/10.3389/fmed.2025.1623408/full. Publisher: Frontiers.

Paolo Neri, Michela Fiaschi, and Giovanni Menchini. Semi-Automatic Tool for Vitiligo Detection and Analysis. *Journal of Imaging*, 6(3):14, March 2020. ISSN 2313-433X. doi: 10.3390/jimaging6030014. URL https://www.mdpi.com/2313-433X/6/3/14. Publisher: Multidisciplinary Digital Publishing Institute.

Hermawan Nugroho, M. Hani Ahmad Fadzil, Norashikin Shamsudin, and S. H. Hussein. Computerised image analysis of vitiligo lesion: evaluation using manually defined lesion areas. *Skin research and technology: official journal of International Society for Bioengineering and the Skin (ISBS) [and] International Society for Digital Imaging of Skin (ISDIS) [and] International Society for Skin Imaging (ISSI)*, 19(1):e72–77, February 2013. ISSN 1600-0846. doi: 10.1111/j.1600-0846.2011.00610.x.

Hamza Abu Owida, Ibrahim Abd El-Fattah, Suhaila Abuowaida, Nawaf Alshdaifat, Hamza A. Mashagba, Azlan B. Abd Aziz, Alaa Alzoubi, Samia Larguech, and Samir Salem Al-Bawri. A deep learning-based dual-branch framework for automated skin lesion segmentation and classification via dermoscopic Images. *Scientific Reports*, 15(1):37823, October 2025. ISSN 2045-2322. doi: 10.1038/s41598-025-21783-z. URL https://www.nature.com/articles/s41598-025-21783-z. Publisher: Nature Publishing Group.

M. Picardo, R. H. Huggins, H. Jones, R. Marino, M. Ogunsola, and J. Seneschal. The humanistic burden of vitiligo: a systematic literature review of quality-of-life outcomes. *Journal of the European Academy of Dermatology and Venereology: JEADV*, 36(9):1507–1523, September 2022. ISSN 1468-3083. doi: 10.1111/jdv.18129.

Shu Tang, Haiheng Ran, Shuli Yang, Zhaoxia Wang, Wei Li, Haorong Li, and Zihao Meng. A frequency selection network for medical image segmentation. *Heliyon*, 10(16):e35698,

August 2024. ISSN 2405-8440. doi: 10.1016/j.heliyon.2024.e35698. URL https://www.sciencedirect.com/science/article/pii/S240584402411729X.

Anjali Thachankattil and Abhishek Sujith. SKIN-ORBIT: A bio-mimetic oscillatory resonance-based inference topology for universal skin lesion segmentation. *Alexandria Engineering Journal*, 128:1177–1202, September 2025. ISSN 1110-0168. doi: 10.1016/j.aej.2025.08.012. URL https://www.sciencedirect.com/science/article/pii/S111001682500883X.

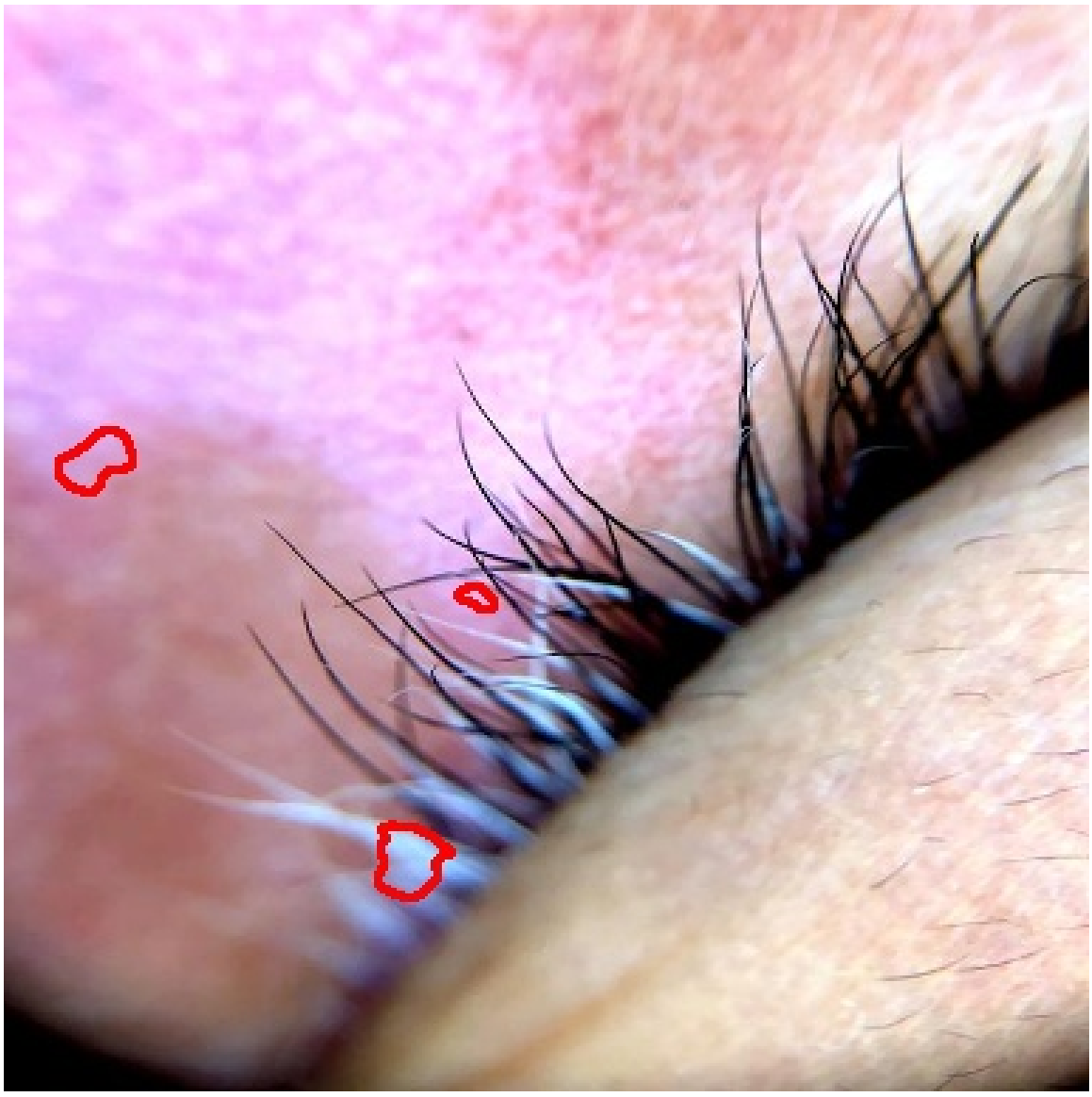
Krasimira Tsaneva-Atanasova, Giulia Pederzanil, and Marianna Laviola. Decoding uncertainty for clinical decision-making. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, 383(2292):20240207, 2025. ISSN 1364-503X. doi: 10.1098/rsta.2024.0207. URL https://pmc.ncbi.nlm.nih.gov/articles/PMC11904615/.

Guotai Wang, Wenqi Li, Michael Aertsen, Jan Deprest, Sébastien Ourselin, and Tom Vercauteren. Aleatoric uncertainty estimation with test-time augmentation for medical image segmentation with convolutional neural networks. *Neurocomputing*, 338:34–45, April 2019. ISSN 0925-2312. doi: 10.1016/j.neucom.2019.01.103. URL https://www.sciencedirect.com/science/article/pii/S0925231219301961.

Li Zhang, Suraj Mishra, Tianyu Zhang, Yue Zhang, Duo Zhang, Yalin Lv, Mingsong Lv, Nan Guan, Xiaobo Sharon Hu, Danny Ziyi Chen, and Xiuping Han. Design and Assessment of Convolutional Neural Network Based Methods for Vitiligo Diagnosis. *Frontiers in Medicine*, 8, October 2021. ISSN 2296-858X. doi: 10.3389/fmed.2021.754202. URL https://www.frontiersin.org/journals/medicine/articles/10.3389/fmed.2021.754202/full. Publisher: Frontiers.

Yifeng Zhou, Huiling Gong, Ruiyun Qiu, Shaofeng Wei, Zhixun Li, and Wei Zhang. MAFF-Net: SAM-powered Mixed Multi-scale Perception Adaptation and frequency-guided feature fusion for robust skin lesion segmentation. *Biomedical Signal Processing and Control*, 113:109222, March 2026. ISSN 17468094. doi: 10.1016/j.bspc.2025.109222. URL https://linkinghub.elsevier.com/retrieve/pii/S1746809425017331.

0.48
(a)



0.48
(b)