

Skin Cancer Classification Using 5-Channel I- RGB-U Images Generated from RGB Images and a Simple BandCross-Transformer

Racharla Aravind¹ and Dr. Bitla Prabhakar²

¹ PG Scholar, Department of ECE,

²Professor, Department of ECE,

^{1,2}JNTUH University College of Engineering Jagtial, Jagtial, TS, India

¹aravindracharla360@gmail.com, ²bprabhakar2008@gmail.com

Abstract: Early and reliable skin cancer screening remains an important challenge in medical image analysis because benign and malignant lesions often exhibit similar visible appearances. This paper presents a skin cancer classification framework that extends a standard RGB lesion image into a five-channel I-*I*-RGB-U representation and classifies the lesion using a Simple BandCross-Transformer (SimpleBCT). The proposed system first applies MST++ spectral reconstruction to generate two additional bands—near-infrared (710–800 nm) and ultraviolet (300–390 nm)—from the original RGB image. These generated bands are fused with the original R, G, and B bands to form a five-channel input that provides richer spectral and structural information than RGB alone. A transformer-based classifier then performs band-aware lesion recognition by learning cross-band interactions and global spatial context. The supplied implementation classifies seven lesion categories from the HAM10000 diagnostic taxonomy and provides a graphical user interface that displays the predicted class and its medical description. Experimental results show a test accuracy of 0.7518 with the highest performance for melanocytic nevi (precision 0.8545, recall 0.9165, F1-score 0.8844). The uploaded screenshots demonstrate correct sample predictions for nevus, benign keratosis-like lesions, basal cell carcinoma, and actinic keratosis / intraepithelial carcinoma. The study indicates that multispectral augmentation combined with transformer-based cross-band learning provides an effective and interpretable approach for non-invasive skin lesion screening.

Keywords: skin cancer classification, I-*I*-RGB-U image, MST++, hyperspectral reconstruction, SimpleBCT, band cross transformer, dermoscopy, HAM10000, medical image analysis.

I. INTRODUCTION

Skin cancer is among the most common forms of cancer and includes both melanoma and non-melanoma lesions. Early diagnosis is crucial because treatment outcome improves significantly when the lesion is recognized before local invasion or metastasis. In clinical practice, dermatologists evaluate lesions visually using features such as asymmetry, border irregularity, color variation, diameter, and evolution. However, manual assessment remains subjective and can be difficult when benign and malignant lesions share similar visual patterns [1], [2].

Computer-aided diagnosis has therefore become an active research area. Earlier systems relied on handcrafted features and conventional classifiers, whereas recent methods increasingly use deep neural networks to learn features directly from the lesion image. Public datasets such as HAM10000 and the ISIC challenge collections have accelerated this progress by enabling standardized benchmarking of classification algorithms [2]–[4].

Despite the success of deep learning, most existing skin cancer classifiers still operate on ordinary RGB images. RGB images represent visible-light appearance well but cannot directly capture non-visible wavelength responses that may



reveal useful tissue differences. The uploaded project addresses this limitation by generating an extended five-channel I-RGB-U image from a single RGB input. Two additional bands—near-infrared and ultraviolet—are computationally reconstructed using MST++, and the resulting five-channel representation is classified by a Simple BandCross-Transformer that learns both spectral interactions and global lesion context.

II. RELATED WORK

Deep learning for skin lesion classification gained strong momentum after dermatologist-level image classification results were reported using convolutional neural networks [1]. The HAM10000 dataset and the ISIC challenge series later became standard resources for the evaluation of automated skin lesion analysis systems [2]–[4].

Spectral reconstruction and hyperspectral-style processing have recently emerged as promising directions because skin tissue interacts differently across wavelengths. MST++ was introduced as an efficient transformer-based spectral reconstruction model that progressively estimates spectral information from RGB observations [5]. This idea is attractive for medical imaging because it can provide richer diagnostic cues without requiring specialized hyperspectral hardware.

Transformer-based recognition models have also been adopted for skin lesion analysis because self-attention can capture long-range dependencies such as asymmetry, global shape, and distributed pigmentation patterns better than purely local convolution operators [6], [8], [9]. In addition, explainable AI is becoming important in dermatology because clinical trust increases when the model can indicate what features or bands influenced its decision [9], [10]. These developments motivate the proposed combination of MST++-based five-channel image generation and band-aware transformer classification.

III. METHODOLOGY

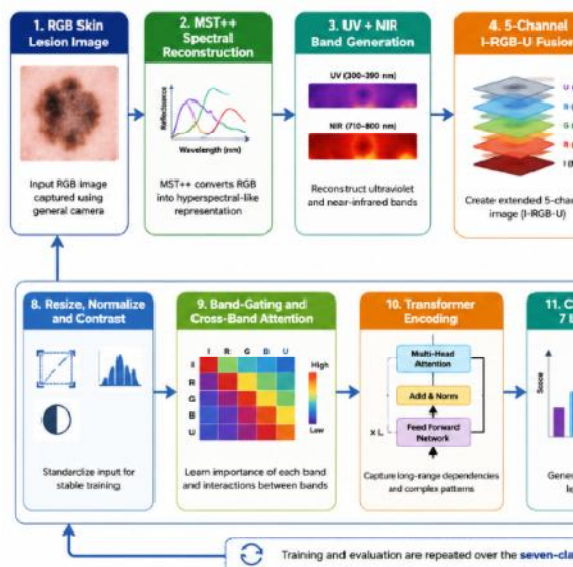


Figure 1. Proposed workflow for five-channel skin lesion classification using MST++ and SimpleBCT.

3.1 Overall Processing Flow

The proposed system follows five major stages: RGB lesion acquisition, spectral reconstruction, five-channel fusion, preprocessing and feature learning, and final seven-class classification. A standard dermoscopic or clinical skin lesion image is first supplied to the system. MST++ reconstructs two additional spectral channels, namely ultraviolet (UV) and near-infrared (NIR). These two bands are merged with the original RGB bands to form a five-channel I-RGB-U



image. The fused image is then normalized, resized, and fed to the SimpleBCT classifier, which generates class probabilities for the lesion categories.

3.2 RGB Skin Lesion Representation

Let the input skin lesion image be denoted by $X_RGB(x,y)$, where each pixel contains three visible color components. The RGB input may be written as:

$$X_RGB(x,y) = [R(x,y), G(x,y), B(x,y)]$$

This representation describes visible color and texture patterns, but it does not explicitly encode ultraviolet or near-infrared responses. Since some lesion properties may become more distinguishable outside the visible range, the system extends the RGB representation before classification.

3.3 MST++ Spectral Reconstruction

The spectral enhancement stage uses MST++ to transform the RGB input into a hyperspectral-like representation. The model learns a mapping from the visible image space to an extended spectral space. In simplified form, the reconstruction can be represented as:

$$H_hat = F_MST++(X_RGB)$$

where $F_MST++(.)$ denotes the learned reconstruction function and H_hat is the estimated multi-band spectral output. From this reconstructed spectral information, the project selects two diagnostically relevant bands: one ultraviolet band in the range 300–390 nm and one near-infrared band in the range 710–800 nm.

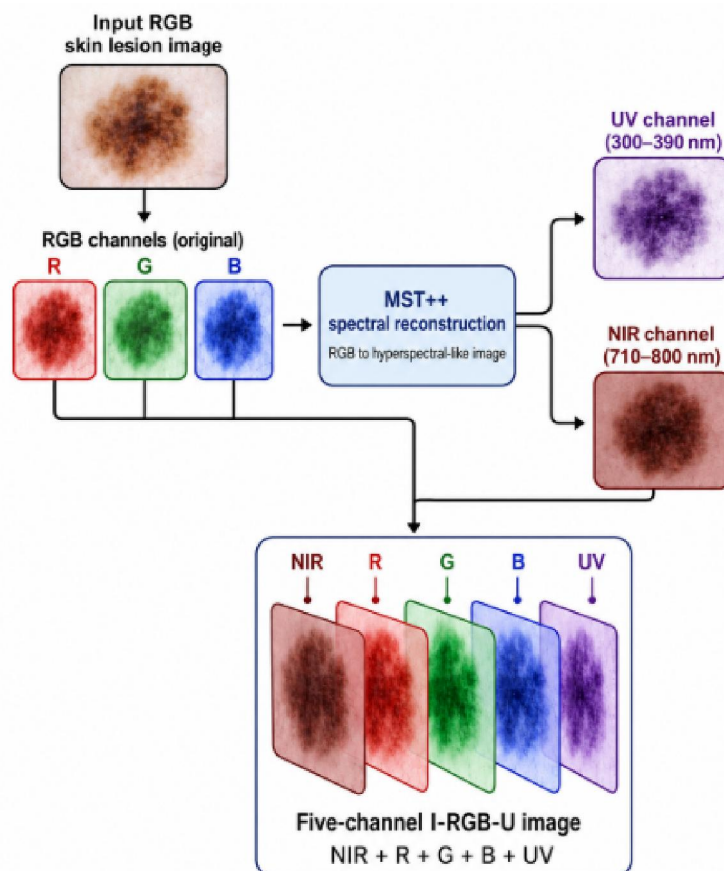


Figure 2. I-RGB-U image formation from an RGB lesion image using MST++ reconstruction of UV and NIR bands.



3.4 Five-Channel I-RGB-U Fusion

After spectral reconstruction, the original RGB channels are combined with the generated near-infrared and ultraviolet channels to form the five-channel lesion representation:

$$X_{IRGBU}(x,y) = [I(x,y), R(x,y), G(x,y), B(x,y), U(x,y)]$$

Here $I(x,y)$ represents the reconstructed near-infrared information and $U(x,y)$ represents the reconstructed ultraviolet information. The resulting I-RGB-U image provides both visible and non-visible cues. This allows the classifier to learn not only superficial pigmentation and texture, but also wavelength-dependent tissue responses that may support lesion discrimination.

3.5 Preprocessing and Data Preparation

Before classification, each five-channel image is resized to the network input size and intensity normalized to a stable numeric range. Data augmentation is used to improve robustness to imaging variation and class imbalance. Typical augmentation operations include horizontal flipping, rotation, and mild contrast variation. These steps help the model generalize across different lesion orientations, illumination conditions, and appearance variations present in the dataset.

3.6 SimpleBCT Classifier

The classification stage uses a simplified BandCross-Transformer (SimpleBCT). Unlike a standard CNN, this model is designed to learn relationships both within each spectral band and across different bands. The network contains four main mechanisms: patch embedding, band-gating, cross-band attention, and classification by a fully connected head. First, the five-channel image is divided into patches and projected into an embedding space. If P_i denotes the i -th image patch, the patch embedding may be written as:

$$z_i = W_e P_i + b_e$$

where W_e and b_e are learnable embedding parameters. Positional encodings are added so that spatial order is retained.

Band-gating adaptively weights the contribution of each channel. If the channel feature vector is denoted by b , the gated band response can be written as:

$$g = \text{sigma}(W_g b + b_g)$$

$$b' = g \odot b$$

where $\text{sigma}(\cdot)$ is the sigmoid function, g is the learned gate, and $\odot$ denotes element-wise multiplication. This mechanism helps the model emphasize the most informative spectral bands for the current lesion.

Cross-band attention then learns how the five channels interact. Given query Q , key K , and value V matrices, scaled dot-product attention is computed as:

$$\text{Attention}(Q,K,V) = \text{softmax}((QK^T)/\text{sqrt}(d_k)) V$$

This operation enables the model to learn dependencies between ultraviolet, visible, and infrared cues. The transformer encoder stacks these attention operations with feed-forward layers and normalization to build a global spectral-spatial representation.

Finally, the classification head produces class scores for the seven lesion categories. If z is the final feature vector, the softmax probability for class c is:

$$P(c | z) = \exp(z_c) / \sum_{k=1}^7 \exp(z_k)$$

The predicted class is the one with the highest probability. The training objective is categorical cross-entropy:

$$L = - \sum_{c=1}^7 y_c \log(P(c | z))$$

where y_c is the one-hot label of class c . In the user interface, the final class label and its medical description are displayed directly after inference.

3.7 Diagnostic Classes

The experimental system follows the seven-class lesion taxonomy commonly used with HAM10000: akiec (actinic keratoses / intraepithelial carcinoma), bcc (basal cell carcinoma), bkl (benign keratosis-like lesions), df (dermatofibroma), mel (melanoma), nv (melanocytic nevi), and vasc (vascular lesions). This multi-class setting is more



clinically meaningful than a simple binary benign-versus-malignant decision because it distinguishes several important lesion subtypes.

IV. RESULTS AND DISCUSSION

The uploaded screenshots provide both qualitative and quantitative evidence of the system's performance. Four screenshots show example GUI predictions for different lesion types, while one screenshot reports the overall test accuracy and detailed class-wise precision, recall, and F1-scores. Together, these outputs confirm that the trained model is functioning as a seven-class lesion classifier.

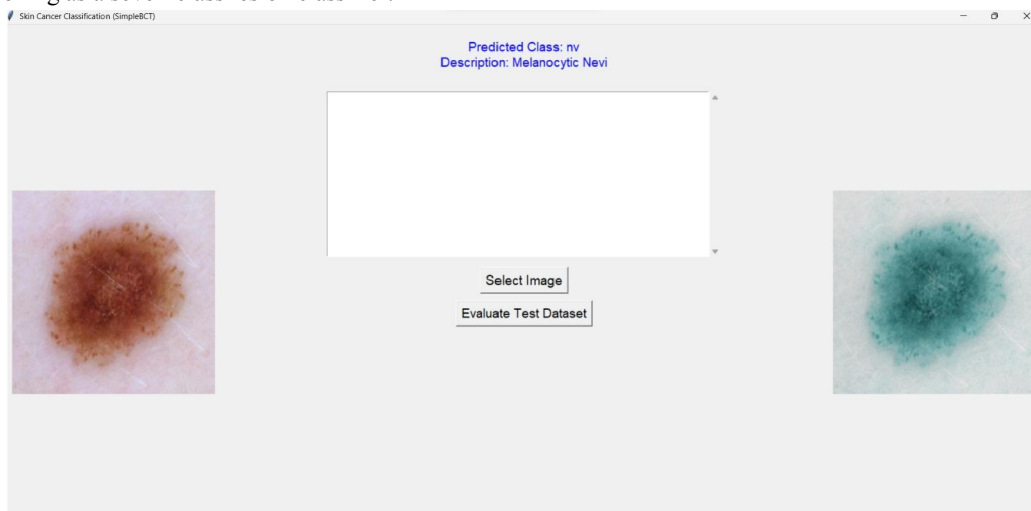


Figure 3. GUI output for a sample lesion classified as nv (Melanocytic Nevus).

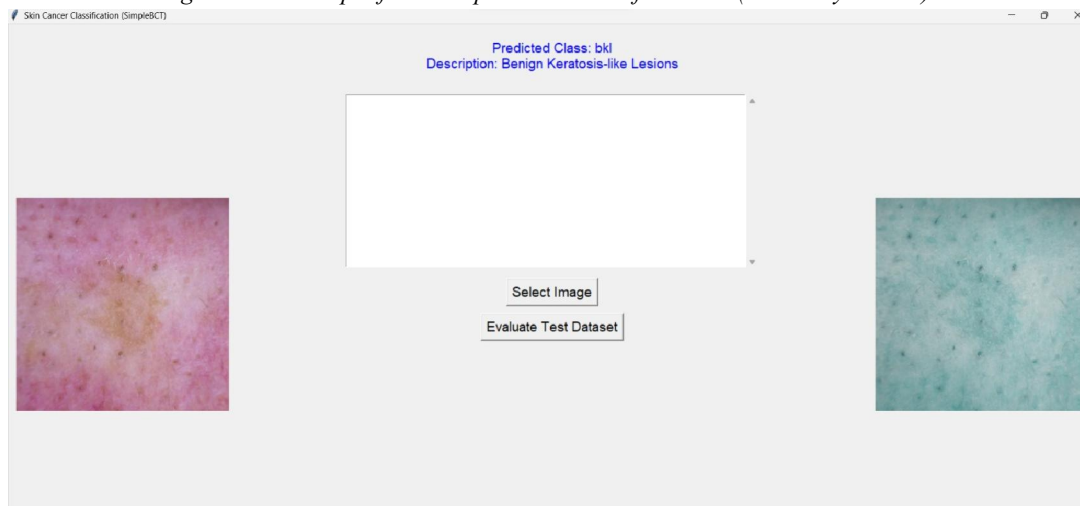


Figure 4. GUI output for a sample lesion classified as bkl (Benign Keratosis-like Lesions).



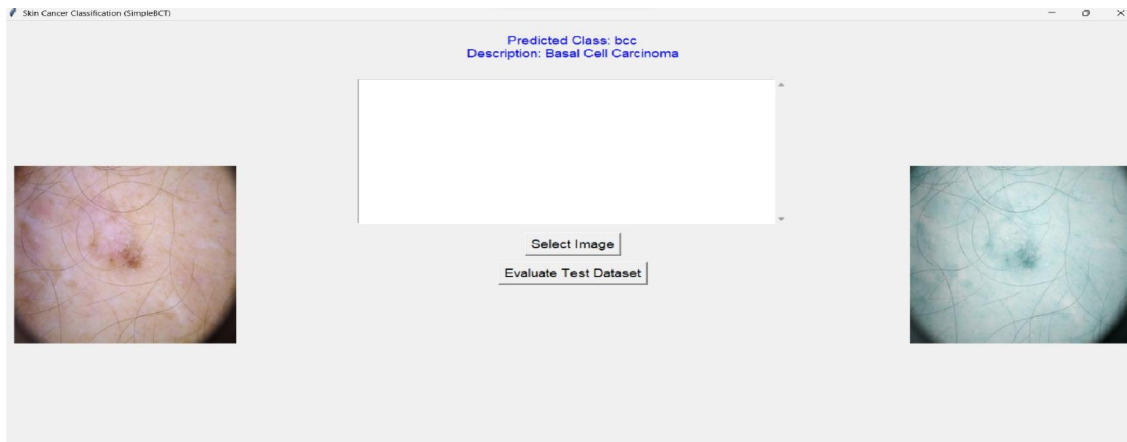


Figure 5. GUI output for a sample lesion classified as bcc (Basal Cell Carcinoma).

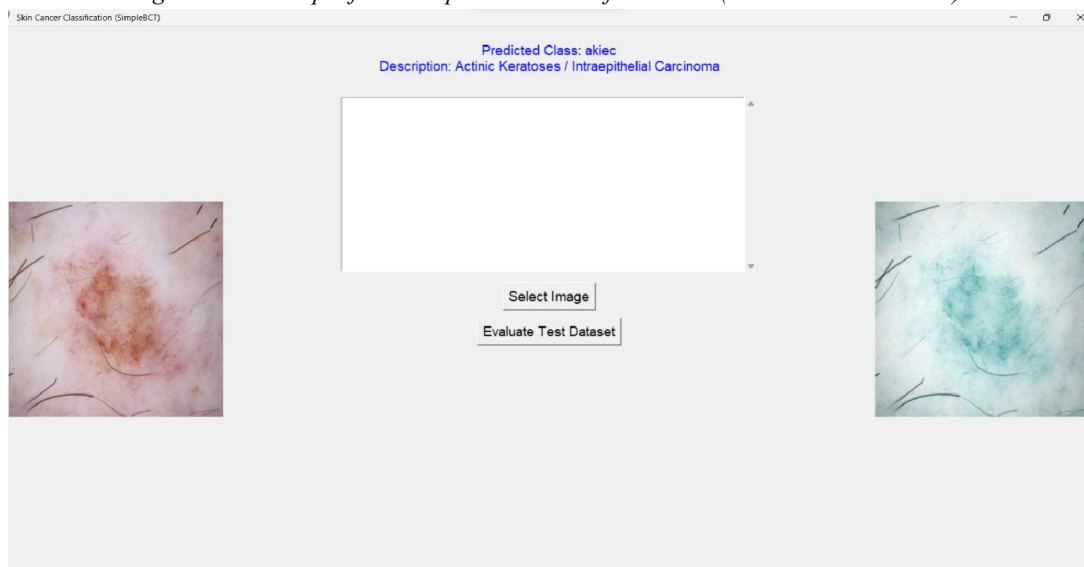


Figure 6. GUI output for a sample lesion classified as akiec (Actinic Keratoses / Intraepithelial Carcinoma).



```
Test Accuracy: 0.7518
Classification report:

```

	precision	recall	f1-score	support
akiec	0.4545	0.4082	0.4301	49
bcc	0.6667	0.5195	0.5839	77
bkl	0.4620	0.4788	0.4702	165
df	0.2222	0.1176	0.1538	17
mel	0.4560	0.3413	0.3904	167
nv	0.8545	0.9165	0.8844	1006
vasc	0.6667	0.4545	0.5405	22
accuracy			0.7518	1503
macro avg	0.5404	0.4623	0.4934	1503
weighted avg	0.7346	0.7518	0.7406	1503

```

Saved confusion matrix to: results\confusion_matrix.png

```

Figure 7. Console output showing overall test accuracy and the classification report.

4.1 Overall Accuracy

The reported test accuracy is 0.7518, which corresponds to 75.18% correct classification over 1,503 test images. The weighted-average F1-score is 0.7406, showing that the model achieves useful overall performance despite the difficulty of the seven-class dermatology task and the class imbalance of the dataset.

4.2 Class-wise Performance

The best performing class is melanocytic nevi (nv), which achieves precision 0.8545, recall 0.9165, and F1-score 0.8844 over 1,006 samples. This strong result is expected because nevus images are more numerous in the dataset and often exhibit relatively stable visual patterns. Basal cell carcinoma also shows good results with an F1-score of 0.5839. Benign keratosis-like lesions and actinic keratosis categories obtain moderate performance. The weakest performance occurs for dermatofibroma, which has only 17 samples in the test report, suggesting that limited sample availability affects generalization.

Class	Precision	Recall	F1-score	Support
akiec	0.4545	0.4082	0.4301	49
bcc	0.6667	0.5195	0.5839	77
bkl	0.4620	0.4788	0.4702	165
df	0.2222	0.1176	0.1538	17
mel	0.4560	0.3413	0.3904	167
nv	0.8545	0.9165	0.8844	1006
vasc	0.6667	0.4545	0.5405	22

4.3 Interpretation of Sample Predictions

The GUI predictions are clinically meaningful and align with the corresponding lesion descriptions. The screenshots show successful predictions for nevus, benign keratosis-like lesions, basal cell carcinoma, and actinic keratosis / intraepithelial carcinoma. In each case, the interface displays the abbreviated class identifier and its expanded medical description, which improves usability for academic demonstration or decision-support scenarios.

4.4 Discussion

The results indicate that the proposed I-RGB-U plus SimpleBCT framework is able to distinguish common skin lesion classes with reasonable performance. The strong result for the nevus class suggests that the spectral-spatial representation successfully captures dominant lesion characteristics. Lower performance for minority classes such as dermatofibroma and melanoma points to a need for better class balancing, data augmentation, or specialized cost-



sensitive training. The absence of the confusion matrix image prevents a more detailed error analysis, but the class-wise report already indicates that the main challenges are class imbalance and similar appearance among diagnostically related lesions.

V. CONCLUSION

This paper presented a complete skin cancer classification framework based on five-channel I-RGB-U images generated from standard RGB lesion images. The method uses MST++ to reconstruct ultraviolet and near-infrared bands and then classifies the fused five-channel image using a Simple BandCross-Transformer that learns cross-band and global spatial relationships. The uploaded results demonstrate practical seven-class lesion classification with a test accuracy of 75.18% and successful GUI predictions for several representative lesion types. The study confirms that spectral augmentation from RGB images can be a useful alternative to expensive hyperspectral imaging hardware. Future work can improve the system by including more balanced datasets, attention-based explainability maps, confusion matrix analysis, calibration across different acquisition devices, and comparison against standard CNN and transformer baselines.

REFERENCES

- [1] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, 2017, doi: 10.1038/nature21056.
- [2] P. Tschandl, C. Rosendahl, and H. Kittler, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Scientific Data*, vol. 5, Art. no. 180161, 2018, doi: 10.1038/sdata.2018.161.
- [3] N. C. F. Codella et al., "Skin Lesion Analysis Toward Melanoma Detection: A Challenge at the 2017 International Symposium on Biomedical Imaging, Hosted by the International Skin Imaging Collaboration," in *Proc. IEEE 15th Int. Symp. Biomedical Imaging (ISBI)*, 2018, pp. 168–172.
- [4] N. C. F. Codella et al., "Skin Lesion Analysis Toward Melanoma Detection 2018: A Challenge Hosted by the International Skin Imaging Collaboration," *arXiv:1902.03368*, 2019.
- [5] Y. Cai, J. Lin, Z. Lin, H. Wang, Y. Zhang, H. Pfister, R. Timofte, and L. Van Gool, "MST++: Multi-stage spectral-wise transformer for efficient spectral reconstruction," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2022, pp. 745–755.
- [6] A. Dosovitskiy et al., "An Image is Worth 16×16 Words: Transformers for Image Recognition at Scale," in *Proc. International Conference on Learning Representations (ICLR)*, 2021.
- [7] H.-Y. Huang, C.-H. Hsiao, C.-W. Liu, and C.-Y. Wang, "Classification of Skin Cancer Using Novel Hyperspectral Imaging Engineering via YOLOv5," *Journal of Clinical Medicine*, vol. 12, no. 3, Art. no. 1134, 2023.
- [8] G. M. S. Himel, M. M. Rahman, M. R. Islam, and M. A. Hossain, "Skin Cancer Segmentation and Classification Using Vision Transformer," *Diagnostics*, vol. 14, no. 4, Art. no. 362, 2024.
- [9] G. H. Dagnaw, B. K. Belete, and T. G. Debelee, "Skin cancer classification using vision transformers and explainable artificial intelligence," *Journal of Medical Artificial Intelligence*, vol. 7, 2024.
- [10] S. Ali et al., "Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence," *Information Fusion*, vol. 99, Art. no. 101805, 2023.

