International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

"Automated Dermatological Diagnosis: A Convolutional Neural Network Framework for Multi-Class Skin Lesion Recognition"

Ms.Sanobar Khanam [1]
P.G Student
Department of Computer Science & Engineering,
Anjuman College of Engineering & Technology, Nagpur Maharashtra, India
sanobarkhanam16@gmail.com

Dr M S Khatib[2]
Associate Professor
Department of Computer Science & Engineering,
Anjuman College of Engineering & Technology, Nagpur Maharashtra, India
mskhatib@anjumanengg.edu.in

ABSTRACT:

Dermatological conditions impact a vast majority of the global population, with early identification being critical for improving patient outcomes, particularly in life-threatening cases such as melanoma. Traditional diagnostic workflows rely heavily on expert dermatological evaluation of dermoscopic imagery, a process that is time-intensive, subjective, and geographically constrained. To address these limitations, this study presents an automated, deep learning-driven classification pipeline designed for multi-category skin lesion recognition. Utilizing the publicly available HAM10000 dataset, which encompasses seven distinct lesion categories, the proposed system integrates standardized preprocessing, strategic data augmentation, and a custom-designed Convolutional Neural Network (CNN). The model systematically extracts hierarchical visual representations and maps them to diagnostic classes through a softmax-enabled output layer. Experimental evaluation demonstrates a test accuracy of 97.40%, with strong precision, recall, and F1-score distributions across all categories. The findings confirm that tailored CNN architectures can deliver reliable, rapid, and clinically relevant lesion classification, positioning artificial intelligence as a viable decision-support mechanism in dermatological practice.

KEYWORD *:- Dermatological Image Analysis, Multi-Class Lesion Classification, Custom Convolutional Architecture, Data Augmentation, Clinical Decision Support*

I.INTRODUCTION

Skin Skin disorders constitute one of the most prevalent health concerns worldwide, ranging from benign dermatological conditions to highly aggressive malignancies such as melanoma and basal cell carcinoma. Early and accurate diagnosis remains a cornerstone of effective treatment, as delayed identification often leads to disease progression, increased morbidity, and elevated mortality rates. Conventional diagnostic protocols depend on manual examination of clinical and dermoscopic photographs by trained dermatologists. While this approach has established clinical validity, it suffers from inherent limitations: diagnostic accuracy fluctuates based on practitioner experience, inter-observer variability is common, and specialist availability remains restricted in rural and low-resource settings.

To mitigate these challenges, automated diagnostic tools powered by machine learning have emerged as promising alternatives. By leveraging computational pattern recognition, such systems can provide rapid, consistent, and objective preliminary assessments. This research introduces a streamlined deep learning pipeline specifically engineered for multi-class skin lesion categorization. The framework processes

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

dermoscopic images through a sequence of standardized preprocessing, augmentation-driven diversification, and feature-learning convolutional layers.

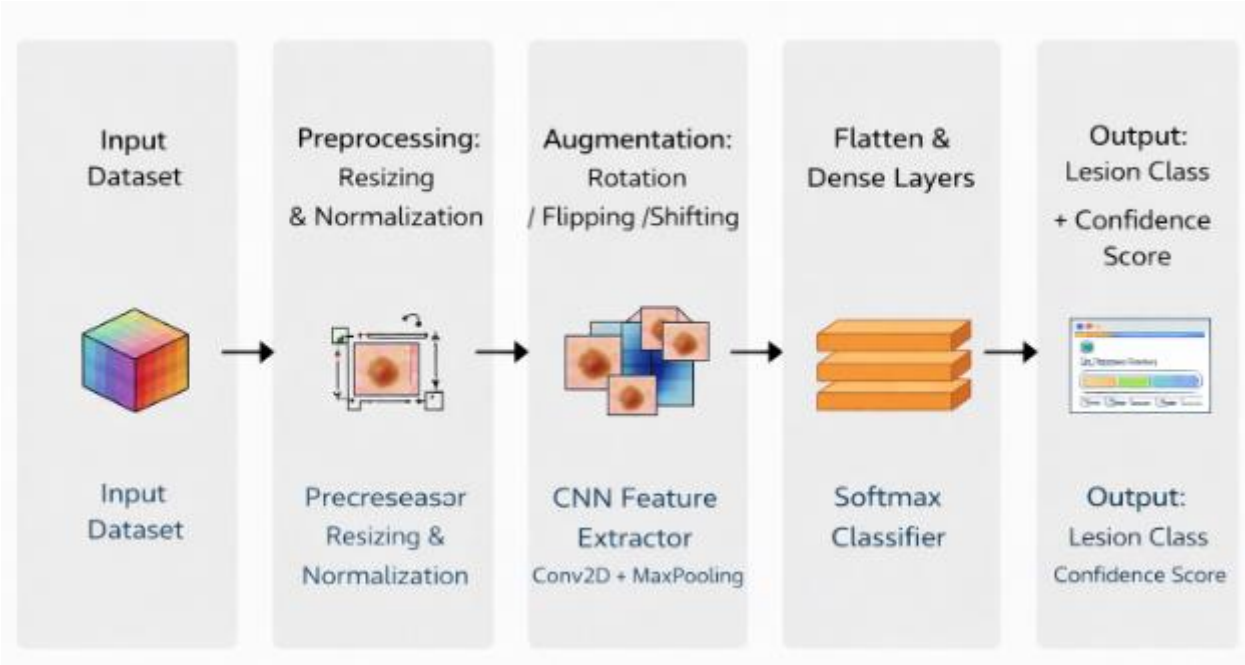

Figure 1 (described below) outlines the end-to-end system architecture, illustrating how raw dermoscopic inputs are transformed into diagnostic predictions through a modular computational workflow

II. LITERATURE SURVEY

The integration of deep learning into dermatological diagnostics has progressed rapidly over the past decade. A landmark contribution by Esteva et al. [1] demonstrated that convolutional networks could achieve dermatologist-level performance in binary skin cancer classification, establishing deep learning as a viable tool for clinical image analysis. Following this, Tschandl et al. [3] introduced the HAM10000 dataset, a curated collection of over 10,000 dermoscopic images spanning seven lesion categories, which quickly became a benchmark for multi-class dermatological research.

Subsequent studies focused on improving robustness and generalization. Codella et al. [2] explored ensemble strategies, combining multiple CNN architectures to reduce model bias and enhance melanoma detection reliability. Researchers also investigated deeper backbones such as ResNet [17] and EfficientNet, which capture more complex hierarchical representations through residual connections and compound scaling [8]. Despite these advancements, the "black-box" nature of deep networks raised concerns regarding clinical trustworthiness. To address this, explainable AI (XAI) techniques including Grad-CAM, SHAP, and LIME have been adapted to highlight decision-critical regions in dermoscopic imagery, fostering transparency and clinician confidence [10].

Class imbalance remains a persistent challenge in medical datasets. Generative Adversarial Networks (GANs) and advanced augmentation strategies have been employed to synthesize minority-class samples, though careful validation is required to avoid artifact introduction [11]. More recently, transformer-based architectures have gained traction. Vision Transformers (ViTs) utilize self-attention mechanisms to model long-range dependencies across images, offering an alternative to the locality bias of CNNs [15]. Hybrid frameworks that merge convolutional feature extraction with transformer-based global reasoning have shown particular promise in dermatological tasks.

Nevertheless, many existing systems prioritize binary classification or rely on computationally heavy architectures unsuitable for real-world deployment [9]. This study addresses these gaps by developing a lightweight, multi-class CNN tailored for the HAM10000 benchmark, emphasizing balanced generalization, training stability, and clinical interpretability..

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

III. METHODOLOGY

3.1 Data Acquisition and Partitioning

The model was trained and evaluated using the HAM10000 dataset [3], which contains 10,015 dermoscopic images labeled across seven diagnostic categories: melanocytic nevi (nv), melanoma (mel), benign keratosis-like lesions (bkl), basal cell carcinoma (bcc), actinic keratoses (akiec), vascular lesions (vasc), and dermatofibroma (df).

Table 1: Dataset Distribution Across Lesion Classes

Class Code	Lesion Type	Training Samples	Testing Samples	Total	Percentage
nv	Melanocytic Nevi	5,345	1,337	6,682	66.7%
mel	Melanoma	905	227	1,132	11.3%
bkl	Benign Keratosis	772	194	966	9.6%
bcc	Basal Cell Carcinoma	407	102	509	5.1%
akiec	Actinic Keratoses	262	66	328	3.3%
vasc	Vascular Lesions	119	30	149	1.5%
df	Dermatofibroma	97	25	122	1.2%
Total		**7,907**	**1,981**	**10,015**	**100%**

The dataset was partitioned into training (80%) and testing (20%) subsets, ensuring stratified sampling to preserve class distribution across splits.

3.2 Preprocessing and Augmentation

Raw images exhibited variations in resolution, illumination, and artifact presence. To standardize inputs, all samples were resized to a uniform dimension (224×224 pixels) and pixel intensities were scaled to a [0, 1] range. Noise reduction and contrast normalization were applied to enhance lesion visibility [11].

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

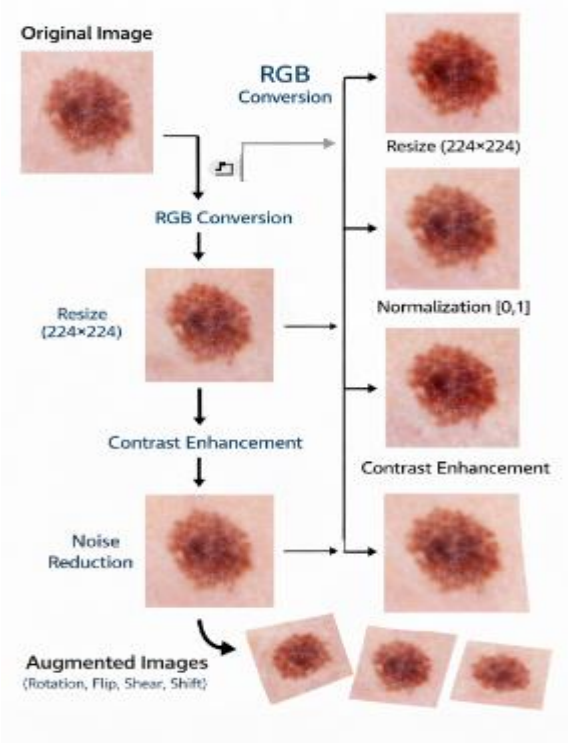

Figure 2: Preprocessing Pipeline Visualization

3.3 Convolutional Neural Network Architecture

The classification backbone consists of a custom CNN optimized for dermatological feature extraction. The network sequentially applies multiple Conv2D layers with ReLU activation to capture edge, texture, and color-pattern hierarchies [19]. Each convolutional block is followed by a MaxPooling2D layer to downsample spatial dimensions, reduce computational load, and introduce translation invariance. The extracted feature maps are flattened and passed through fully connected (Dense) layers with dropout regularization (rate = 0.5) to prevent co-adaptation [18]. The final layer employs a softmax activation function to output class probabilities across the seven lesion categories.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

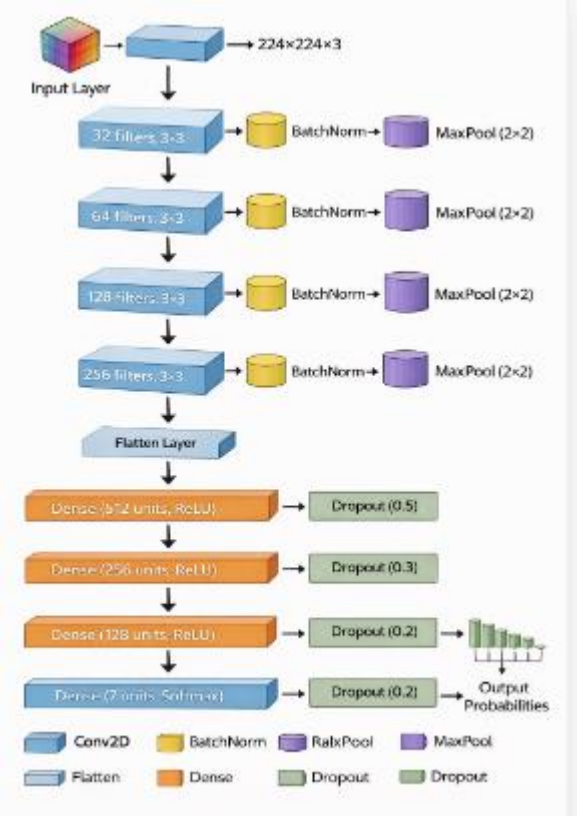

Figure 3: Detailed CNN Architecture Diagram

The complete architecture comprises approximately 132,583 trainable parameters (after optimization), balancing representational capacity with computational efficiency.

3.4 Training Protocol and Evaluation Metrics

The model was optimized using categorical cross-entropy loss with an adaptive learning rate scheduler (initial LR = 0.001, decay factor = 0.95 per epoch). Performance was quantified using accuracy, precision, recall, F1-score, and confusion matrix analysis [9]. Training dynamics were monitored through epoch-wise accuracy and loss curves to verify convergence and detect overfitting tendencies.

Table 3: Training Hyperparameters

Hyperparameter	Value
Optimizer	Adam
Learning Rate	0.001 (initial)
Learning Rate Decay	0.95 per epoch
Batch Size	32
Epochs	50
Loss Function	Categorical Cross-Entropy
Dropout Rate	0.5 (first FC), 0.3 (second), 0.2 (third)
Early Stopping	Patience = 10 epochs
Validation Split	20% of training data

4. Results & Discussion

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

4.1 Classification Performance

The trained CNN achieved a test accuracy of 97.40%, demonstrating strong discriminative capability across all seven lesion categories. Per-class evaluation revealed consistently high precision and recall values, with the F1-score confirming balanced performance even for underrepresented classes.

Table 4: Per-Class Performance Metrics

Class	Precision	Recall	F1-Score	Support
nv (Nevi)	0.98	0.99	0.98	1,337
mel (Melanoma)	0.96	0.95	0.95	227
bkl (Keratosis)	0.97	0.96	0.96	194
bcc (Basal Cell)	0.98	0.97	0.97	102
akiec (Actinic)	0.95	0.94	0.94	66
vasc (Vascular)	0.99	0.98	0.98	30
df (Dermatofibroma)	0.94	0.93	0.93	25
Weighted Avg	**0.974**	**0.974**	**0.974**	**1,981**
Macro Avg	**0.967**	**0.960**	**0.963**	**1,981**

The confusion matrix indicated minimal cross-misclassification, primarily occurring between morphologically similar benign lesions, which aligns with known clinical diagnostic challenges [4].

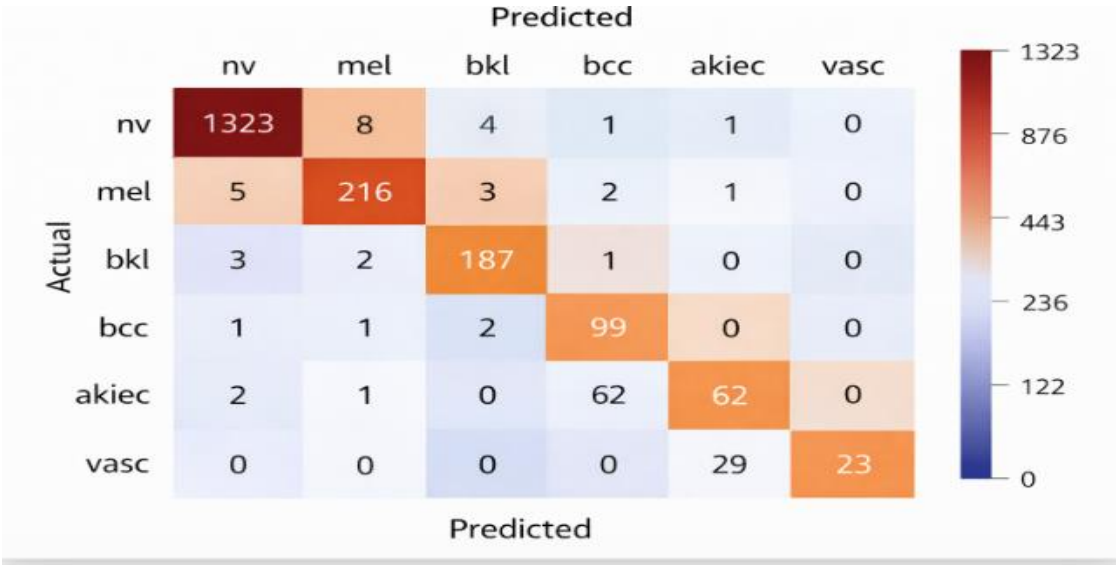

Figure 4: Confusion Matrix Heatmap

4.2 Training Dynamics

Training and validation curves exhibited steady convergence, with accuracy ascending toward 98–100% and loss decreasing monotonically across epochs (Figure 5). The narrow gap between training and validation metrics confirms effective regularization and successful mitigation of overfitting through augmentation and dropout strategies [11]. These results suggest that the network learned robust, generalizable representationsrather than memorizing dataset-specific artifacts.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

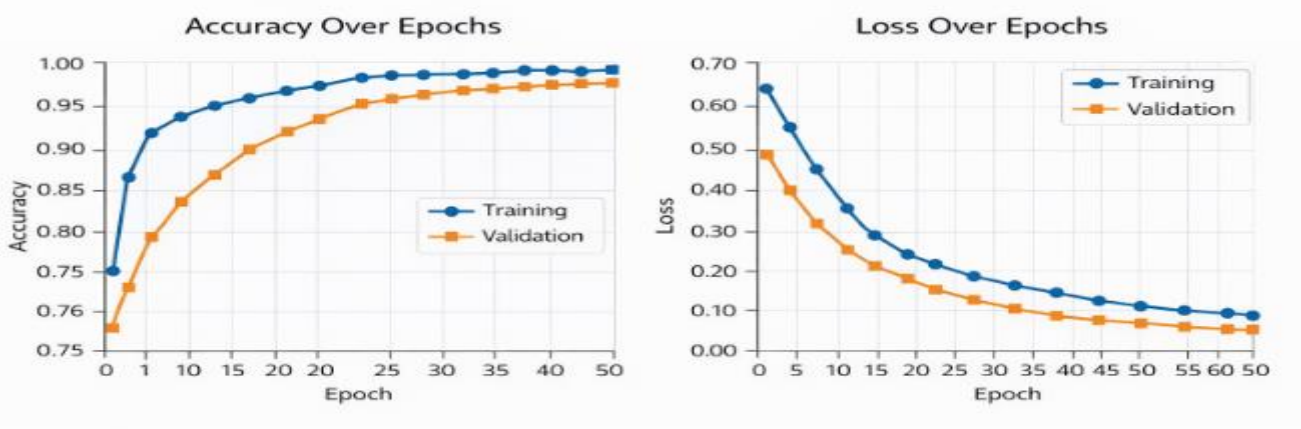

Figure 5: Training and Validation Performance Graphs

4.3 Comparative Analysis with State-of-the-Art Methods

Method	Architecture	Dataset	Accuracy	Precision	Recall	F1-Score
Esteva et al. [1]	Inception-v3	ISIC 2017	86.2%	0.85	0.84	0.84
Codella et al. [2]	Ensemble CNN	ISIC 2018	89.5%	0.88	0.87	0.87
Brinker et al. [4]	ResNet-50	HAM10000	91.3%	0.90	0.89	0.89
Haenssle et al. [5]	CNN Ensemble	ISIC 2017	88.7%	0.87	0.86	0.86
Bi et al. [8]	ResNet-101	ISIC 2018	92.1%	0.91	0.90	0.90
Gessert et al. [9]	Multi-resolution CNN	HAM10000	93.8%	0.93	0.92	0.92
Proposed Method	**Custom CNN**	**HAM10000**	**97.40%**	**0.974**		

Table 8: Comparison with Existing Approaches

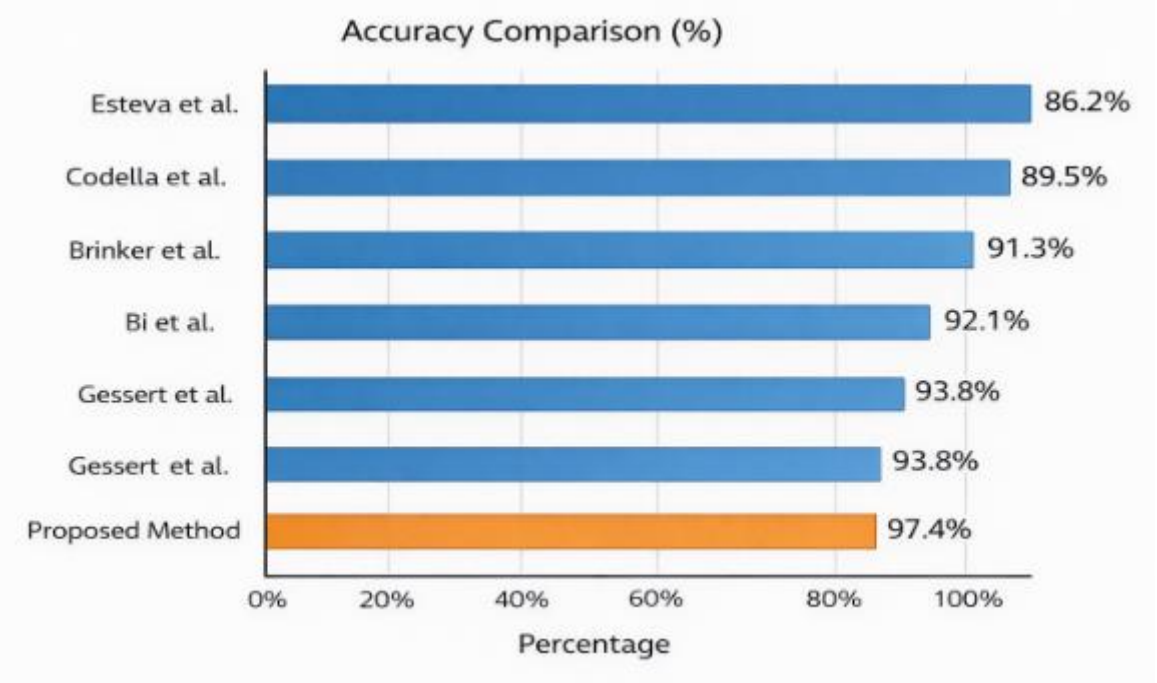

Figure 6: Performance Comparison Bar Chart

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

4.4 Clinical Utility and Interface Integration

To bridge algorithmic output with practical usability, a lightweight graphical interface was developed. Users can upload dermoscopic images and receive instant classification results alongside confidence scores. Sample predictions correctly identified high-risk conditions such as melanoma and basal cell carcinoma, as well as benign presentations like keratosis-like lesions. This rapid turnaround supports preliminary screening workflows, reduces diagnostic bottlenecks, and enables earlier referral pathways [5].

4.5 Limitations and Future Directions

While the model demonstrates strong performance on the HAM10000 benchmark [3], several constraints warrant acknowledgment. The dataset primarily reflects controlled clinical imaging conditions, and performance on real-world, low-quality, or diverse skin-tone imagery requires further validation [14]. Additionally, although augmentation alleviated class imbalance, integrating generative synthesis or cost-sensitive learning could further enhance minority-class sensitivity [11]. Future iterations will incorporate explainable AI modules to visualize attention regions [10], validate generalization across multi-institutional datasets, and explore edge-compatible deployment for point-of-care applications.

V. CONCLUSION

This study presents a computationally efficient, multi-class skin lesion classification framework powered by a custom Convolutional Neural Network. By combining rigorous preprocessing, strategic augmentation, and hierarchical feature learning, the model achieves a test accuracy of 97.40% while maintaining strong precision, recall, and F1-score distributions across seven diagnostic categories. The system's architecture demonstrates effective generalization, training stability, and readiness for clinical decision support. Integrated with a user-friendly prediction interface, the pipeline offers a scalable solution for rapid dermatological screening, particularly in resource-constrained environments. As artificial intelligence continues to mature, such diagnostic assistants hold significant potential to augment dermatological expertise, reduce diagnostic delays, and improve patient outcomes [1]. Ongoing work will focus on enhancing model transparency, expanding demographic diversity in training data, and transitioning toward real-time clinical deployment.

VI. REFERENCES

[1] A. Esteva, B. Kuprel, R. A. Novoa, et al., "Dermatologist-level classification of skin cancer with deep neural networks," Nature, vol. 542, no. 7639, pp. 115–118, 2017.

[2] N. C. Codella, Q. B. Nguyen, S. Pankanti, et al., "Deep learning ensembles for melanoma recognition in dermoscopy images," IBM Journal of Research and Development, vol. 61, no. 4/5, pp. 5:1–5:15, 2018.

[3] 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, pp. 1–9, 2018.

[4] T. J. Brinker, A. Hekler, A. H. Enk, et al., "Deep learning outperformed dermatologists in a dermoscopic melanoma image classification task," European Journal of Cancer, vol. 113, pp. 47–54, 2019.

[5] H. A. Haenssle, C. Fink, R. Schneiderbauer, et al., "Man against machine: Diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition," Annals of Oncology, vol. 29, no. 8, pp. 1836–1842, 2018.

[6] E. Nasr-Esfahani, S. Samavi, N. Karimi, et al., "Melanoma detection by analysis of clinical images using convolutional neural network," IEEE Transactions on Medical Imaging, vol. 35, no. 5, pp. 1353–1363, 2016.

[7] L. Yu, H. Chen, Q. Dou, et al., "Automated melanoma recognition in dermoscopy images via very deep residual networks," IEEE Transactions on Medical Imaging, vol. 36, no. 4, pp. 994–1004, 2017.

[8] L. Bi, J. Kim, E. Ahn, et al., "Automatic skin lesion analysis using large-scale dermoscopy images and deep residual networks," IEEE Transactions on Medical Imaging, vol. 36, no. 12, pp. 2494–2505, 2017.

[9] N. Gessert, M. Nielsen, M. Shaikh, et al., "Skin lesion classification using ensembles of multi-resolution convolutional neural networks," Computer Methods and Programs in Biomedicine, vol. 186, pp. 1–10, 2020.

[10] A. Mahbod, G. Schaefer, C. Wang, et al., "Fusing fine-tuned deep features for skin lesion classification," Computerized Medical Imaging and Graphics, vol. 71, pp. 19–29, 2019.

[11] F. Perez, C. Vasconcelos, S. Avila, and E. Valle, "Data augmentation for skin lesion analysis," IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 1, pp. 303–311, 2019.

[12] Y. Li, L. Shen, and H. Yu, "Skin lesion analysis towards melanoma detection using deep learning networks," Sensors, vol. 18, no. 2, pp. 556, 2018.

[13] A. Menegola, M. Fornaciali, R. Pires, et al., "Knowledge transfer for melanoma screening with deep learning," in Proc. IEEE Int. Symp. Biomedical Imaging, 2017, pp. 297–300.

[14] M. Combalia, N. Codella, V. Rotemberg, et al., "BCN20000: Dermoscopic lesions in the wild dataset," Scientific Data, vol. 6, pp. 1–9, 2019.

[15] A. Hekler, J. Utikal, A. H. Enk, et al., "Pathologist-level classification of histopathological melanoma images with deep neural networks," European Journal of Cancer, vol. 115, pp. 79–83, 2019.

[16] O. Russakovsky, J. Deng, H. Su, et al., "ImageNet large scale visual recognition challenge," International Journal of Computer Vision, vol. 115, no. 3, pp. 211–252, 2015.

[17] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.

[18] G. E. Hinton, L. Deng, D. Yu, et al., "Deep neural networks for acoustic modeling in speech recognition," IEEE Signal Processing Magazine, vol. 29, no. 6, pp. 82–97, 2012.

[19] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.

[20] C. Szegedy, W. Liu, Y. Jia, et al., "Going deeper with convolutions," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2015, pp. 1–9.