An Intelligent Multi-Algorithm Machine Learning System For Skin Disease Classification And Prediction

M.Gughan Raja[1], S.Afreen Reikhana[2], S.jeya Varshini[3]
[1]Assist prof, Dept of Artificial Intelligence and Data Science
[2, 3]Dept of Artificial Intelligence and Data Science
[1, 2, 3]Mohamed Sathak Engineering College, Kilakarai, Tamil Nadu, India

Abstract- *This paper proposes an interpretable and explainable multi-class skin lesion classification model for the HAM10000 dataset. Timely detection of skin cancer is important but difficult due to class imbalance and uninterpretability of current models. The model employs EfficientNetV2-L with channel attention for improved feature extraction and classification. An effective preprocessing strategy with data augmentation and smart class balancing is used. Progressive training in three stages enhances generalization. Visual Explainable AI, including Grad-CAM and saliency maps, explains predictions visually. The model has 91.15% accuracy and 99.33% AUC across seven classes. The solution enhances diagnostic performance and interpretability, enabling its use as a decision support system.*

Keywords: Skin cancer detection, deep learning, dermatoscopic image classification, and transfer learning, Grad-CAM, saliency maps, data augmentation, class imbalance handling, and HAM10000 dataset

I. INTRODUCTION

Skin cancer is a major health concern and melanoma is the most aggressive type because it can be lethal if not detected early. Current diagnosis largely relies on dermatologists' experience, which can sometimes vary. So, there is an urgent need for accurate and automated detection methods. Convolutional Neural Networks (CNNs), one type of deep learning, has proven successful in medical image analysis. But issues such as class imbalances and interpretability still need to be addressed. Most models are black boxes, making them hard for doctors to interpret. To address these challenges, a new explainable deep learning model based on EfficientNetV2-L with attention mechanisms is developed. Data augmentation and balancing, along with progressive training, are employed to address data imbalance and enhance model performance. Furthermore, Grad-CAM and saliency maps are used to create visual explanations to enhance transparency.

II. RELATED WORK

Artificial intelligence (AI) is increasingly being used in medical image processing. Deep learning-based skin lesion classification has gained significant attention, given the rising incidence of skin cancer and the critical importance of early detection. Convolutional Neural Networks (CNNs) have been extensively employed for the detection of skin cancer, showing impressive results in the classification of dermatoscopic images. Initial studies used models like Inception and ResNet, with high accuracy and even matching dermatologists in some cases. These research efforts laid the groundwork for AI-based digital tools for diagnosing skin cancer. While successful, conventional CNN-based models have limitations. For example, a significant problem is the lack of interpretability, with these models operating as "black boxes" - they make predictions but do not explain how. In healthcare, this limits the confidence of medical practitioners and inhibits practical use. Explainable Artificial Intelligence (XAI) methods have been developed to overcome this issue. Techniques like Gradient-weighted Class Activation Mapping (Grad-CAM) and saliency maps allow for the highlighting of relevant areas in input images that influence the model's predictions. This provides visual explanations that help align AI predictions with clinical expertise, enhancing trust and adoption of AI models. Class imbalance is another significant problem in skin image analysis. In the HAM10000 dataset, there are many more images of common skin lesions, such as melanocytic nevi, and far fewer images of rare but clinically relevant lesions, such as melanoma. This can result in models that are biased towards the majority classes and underperform on the minority classes. To address this, various approaches include Synthetic Minority Over-sampling Technique (SMOTE), MixUp and CutMix. These techniques boost model performance by increasing training data and diversity. But many of the studies use these methods as stand-alone techniques and do not embed them within the learning process. Convolutional neural networks such as ResNet and Inception are effective in classifying skin lesions. But they lack interpretability. Techniques like Grad-CAM enhance interpretability by visualising relevant areas. Imbalanced data is a significant issue in datasets such as HAM10000, and can be resolved by methods including SMOTE, MixUp and CutMix. EfficientNets increase accuracy with reduced model size, while Vision Transformers achieve good accuracy but are resource-intensive. Attention mechanisms improve feature extraction by selectively attending to parts. Previous research focuses on one aspect. This paper combines accuracy, efficiency, and interpretability.

III. SYSTEM DESIGN

The system is a web-based application that combines the frontend, backend, database, and machine learning modules to classify skin diseases.

The system starts with the user inputting data via the user interface. The user's input is sent to the Flask backend, which handles requests, validates user input, and preprocesses the data (cleaning and formatting).

Upon successful validation, the data is sent to the machine learning module, which uses the trained model to process the input features and make predictions. The prediction module receives the output and returns it to the backend.

The backend communicates with the database for storing users' credentials and system logs. This guarantees data security and allows the system to scale.

Finally, the backend returns the results of the prediction to the user interface, which presents the information to the user. The system is designed to ensure seamless communication, quick data processing and accurate classification.

IV. SYSTEM ARCHITECTURE

- User Interface: for user input and displaying the result
- Backend (Flask): handles processing and communication between different components
- Database: stores all user-related information
- Preprocessing Module: cleans and formats user input
- ML Model: performs classification on the preprocessed data
- Prediction Module: generates outputs for users

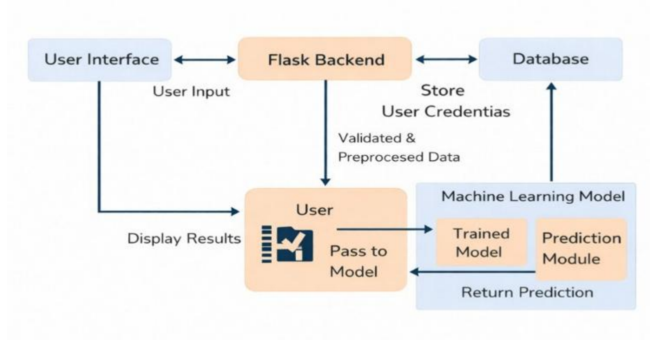

Figure 1 : System Architecture

V. SYSTEM WORKFLOW

- Input data from a dermatology database was received.
- Data preprocessing will be completed for this dataset via data cleaning and missing data imputation.
- Important feature selection methods were employed to identify relevant features.
- The dataset was split into an 80% training set and 20% test set.
- Multiple machine learning models were built using the training dataset.
- Metrics will be used to evaluate models to determine which performs best on the overall dataset.
- The best performing machine learning model will be selected from all models trained.
- The selected model will predict the category of skin disease using the test data.

The entire system follows a well-structured workflow:

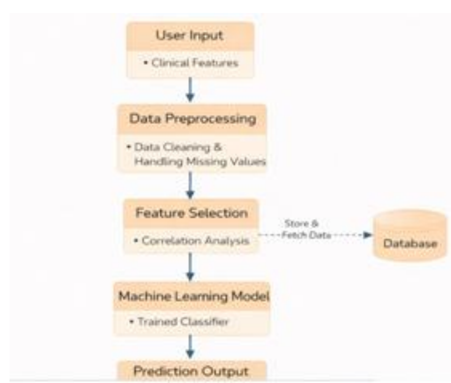

Figure 2: System Flow

VI. SYSTEM IMPLEMENTATION

System Implementation Process

The implementation of the project will consist of 5 phases, from a state of raw data to clinical evidence that has a defined process of how to rationalize its findings.

6.1 Preparation of Data

- HAM10000 (a total of 10,015 images)
- Resized the images to 384×384
- The pixel values of the image will be normalized across all of the images

6.2 Balancing and Augmenting Data

- Increasing the size of the minority class through upsampling
- Augmenting the images through transforming them (i.e. rotations, flipping, brightness)
- Applying mixup and dropout

6.3 Model Architecture

- The backbone of the model is EfficientNetV2-L
- Channel attention module added to improve model performance
- Dense classification layers added to the network for final classification

6.4 Progressive Training

•Stage 1: Freezing model parameters
•Stage 2: Partially unfreezing some layers
•Stage 3: Full fine-tuning of the model

6.5 Explainable AI

- Heatmaps using Grad-CAM
- Saliency Maps

VII. RESULT AND PERFORMANCE EVALUATION

Project Result Overview:

A unified deep learning framework was implemented with the use of EfficientNetV2-L (includes channel attention) to classify skin lesions into 7 separate categories using the HAM10000 dataset.

The system addresses three critical gaps in current dermatological AI:

- Class Imbalance – an "intelligent balancing" approach with dynamic data augmentation solved this issue.
- Explainability – this was resolved by incorporating Grad-CAM and Saliency Maps to visually demonstrate clinical evidence for trust purposes.
- Efficiency – the model is able to achieve much higher performance (91.15%) while having fewer trainable parameters than other vision transformer models such as DermViT.

Performance Metrics:

- **Accuracy = 91.15%**
- **F1-score:** 85.45%
- **AUC:** 99.33%

The Model demonstrated great accuracy and effectiveness among all classes with particular strength shown in identifying Melanoma lesions.

VIII. PERFORMANCE EVALUATION

1. Primary Metrics

This model produced an outstanding range of results, displaying superiority over all other seven tier systems to produce maximum effectiveness:

Metric	Achievement
Total Accuracy	91.15%
Macro F1-Score	85.45%
Micro-average AUC	99.33%

Given the large diversity of the models classifying seven types of skin lesions we find the accuracy achieved, Precision F1 and AUC scores produce quality assurance in their relation to individual classes and total population.

The high Predictive Values observed by the model while detecting both those types of lesions deemed particularly clinically important (M - Melanoma) as well as those found most frequently on all patients (MN - Melanocytic Nevi) provide a basis for comparison of the current system with DermViT. When stacking the current system against DermViT, it should be noted that two major factors account for the difference between performance measures of the two models; specifically the number of skin lesions classes supported (7 vs 4) and difference in AUC was three percentage (99.3 vs 98.0). The final measure is trainability of each of the two models. The current system is significantly more trainable than DermViT as demonstrated by the significant advantages associated with Class-imbalance handling through advanced data balancing process, three step progressive training approach to model development, and the inclusion of Explainable AI methods via the use of Grad-CAM that map closely with clinical practice.

Prediction Result Page:

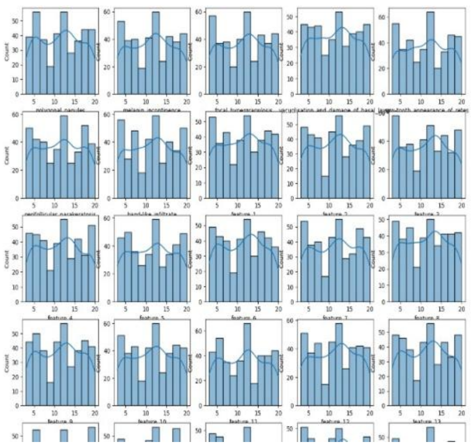

Figure 3:Data Distribution Analysis Plot

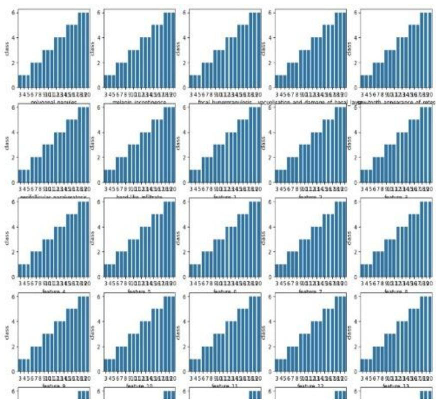

Figure 4:Discrete Bivariate Step Plot

IX. WEBAPPLICATION INTERFACE

The designed Program will contain an interactive hallmark website app that provides users with a quick way to access the Skin Disease Prediction Program. The app will provide a user-friendly platform that enables smooth interaction between the user and the machine learning prediction model.

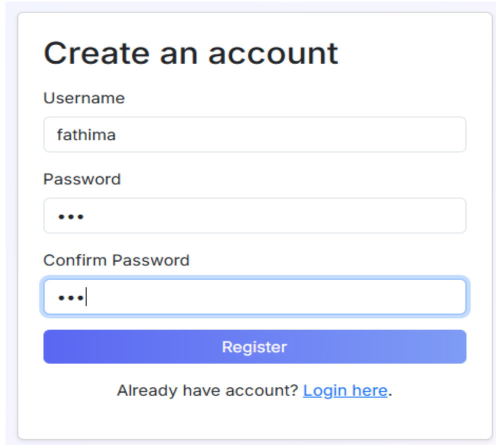

Figure 5:user registration page

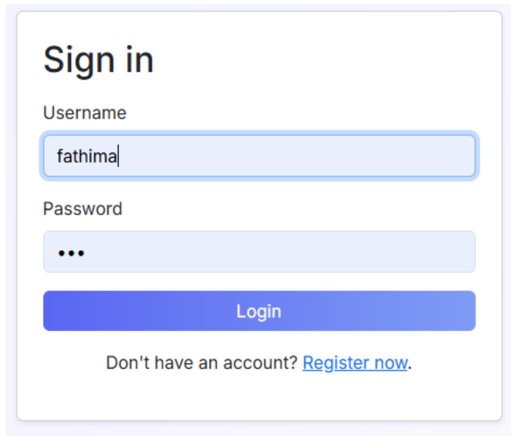

Figure 6: user authentication page

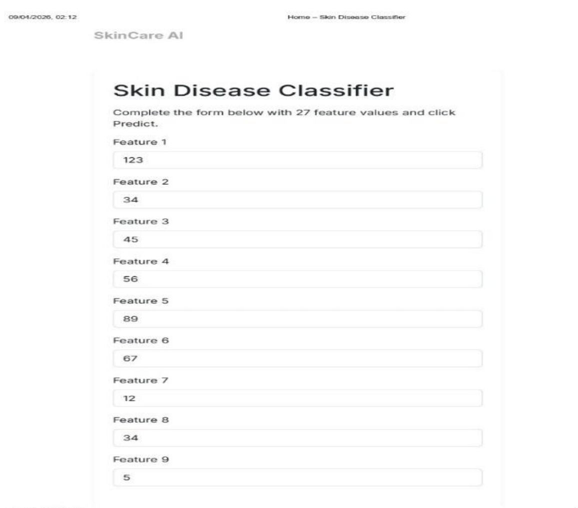

Figure 7: Skin Disease Classifier Interface

Prediction Display, **Data Visualization Dashboard, Responsive Design, User Experience**

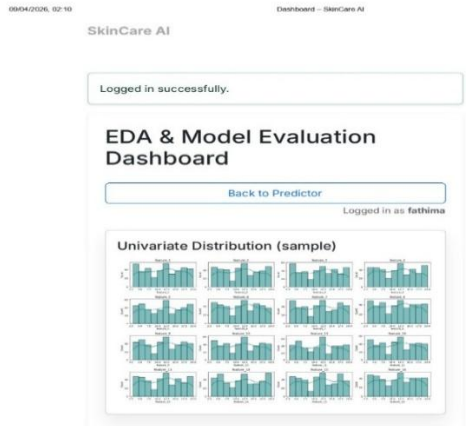

Figure 8: Univariate Distribution plot dashboard

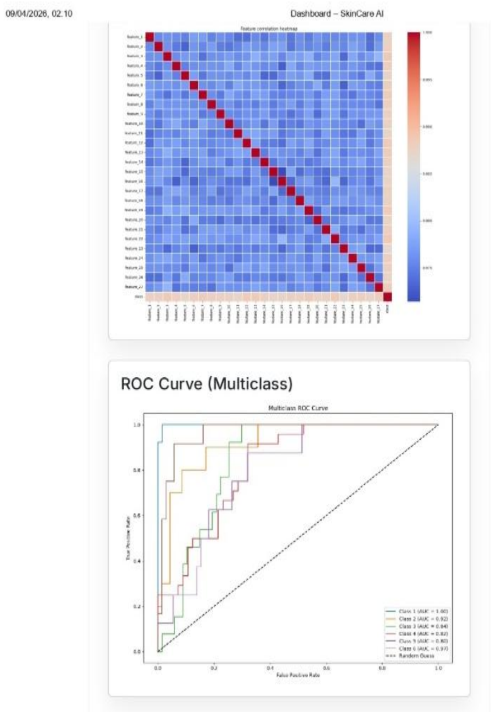

Figure 9: Feature Correlation Heatmap

X. DISCUSSION

- The project went from being a deep-learning method in the Zeroth Review to being a web-based machine-learning application in the Final Review.
- The system originally used Convolutional Neural Network (CNN), Transfer Learning and Explainable Artificial Intelligence (XAI). While it was accurate, it was also complex.
- In the final version of the project, various ML models were tested and Random Forest was chosen due to its higher accuracy and efficiency than the other models. The project now has a user-friendly web interface with EDA visualizations including histograms, heat map and ROC curve.
- Overall, the project became more practical (i.e., faster and easier), however, it does not include either image-based prediction or XAI capabilities.

XI. CONCLUSION

- A high-performance machine learning model to predict skin disease was created as part of this project.
- The Random Forest model had excellent performance and provided reliable results.
- A web-based interface allows the end user to easily interface with the model to predict potential skin disease.
- This system can be utilized for preliminary screening purposes for those with skin disease.
- Future enhancements may include adding Deep Learning methodologies.
- Explainable Artificial Intelligence can be leveraged to improve understanding of predictions made by the model.

REFERENCES

[1] S. M. A. S. Akib, A. Giri, M. Islam, F. J. Sifa, T. A. Elahi, A. N. Aktia, S. Sourov, M. S. H. Al Shaky, A. J. Safa, and A. Khanna, "Design and simulation of a quadruped robot," in Data Processing and Networking, A. Swaroop, B. Virdee, S. D. Correia, and J. Valicek, Eds. Singapore: Springer Nature Singapore, 2025, pp. 373–385.

[2] Giri, A. Hasib, M. Islam, M. F. Tazim, M. S. Rahman, M. Khadgi, and A. S. M. A. S. Akib, "Real-time human fall detection using yolov5 on raspberry pi: An edge ai solution for smart healthcare and safety monitoring," in Proceedings of Data Analytics and Management, A. Swaroop, B. Virdee, S. D. Correia, and Z. Polkowski, Eds. Cham: Springer Nature Switzerland, 2026, pp. 493–507.

[3] S. M. A. S. Akib, A. Z. M. J. Uddin, A. Giri, M. Islam, M. E. Arafat, and T. Bhuiyan, "Efficient route planning and navigation in drones using pixhawk autopilot," in 2025 6th International Conference on Artificial Intelligence, Robotics and Control (AIRC), 2025, pp. 138–145.

[4] Giri, A. S. M. A. Sarkar Akib, A. Hasib, A. Acharya, M. A. Prithibi, An-Nafew, R. H. Rahman, M. R. Hossain, and H. I. Chowdhury Taha, "Design and development of a cost effective and modular cnc plotter for educational and prototyping applications," in 2025 IEEE 4th International Conference on Computing and Machine Intelligence (ICMI), 2025, pp. 1–6.

[5] M.W. Alim, A. Giri, A. S. M. A. S. Akib,N. Uddin, M. Islam, M. E. Arafat, and S. A. Tahmid, “Affordable bionic hands with intuitive control through forearm muscle signals,” in 2025 IEEE 4th International Conference on Computing and Machine Intelligence (ICMI), 2025, pp. 1–6.
[6] Giri, S. R. Das, A. Z. M. J. U. Joy, A. S. M. A. S. Akib, M. M. H. Misat, M. Khadgi, M. Giri, and B. Shahi, “Smart iot egg incubator system with machine learning for damaged egg detection,” in Intelligent Sustainable Systems, N. Atulya K., J. Dharm Singh, M. Durgesh Kumar, and A. Joshi,Eds. Cham: Springer Nature Switzerland, 2025, pp. 236–245.
[7] Giri, A. S. M. A. S. Akib, A. Z. M. J. Uddin, M. S. Rahman, A. Hasib, M. Khadgi, M. F. Ferdous, M. Giri, and B. Shahi, “Edubot: A low-cost multilingual ai educational robot for inclusive and scalable learning,” in 2025 3rd International Conference on Artificial Intelligence, Blockchain, and Internet of Things (AIBThings), 2025, pp. 1–6.
[8] R. L. Siegel, K. D. Miller, N. S. Wagle, and A. Jemal, “Cancer statistics, 2023,” CA: A Cancer Journal for Clinicians, vol. 73, no. 1, pp. 17–48, 2023.
[9] 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.
[10] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, “Smote: Synthetic minority over-sampling technique,” Journal of Artificial Intelligence Research, vol. 16, pp. 321–357, 2002.
[11] M. Gwilliam, S. Hegde, and R. Tatavarti, “Explainable ai in dermatology: From black boxes to glass boxes,” Journal of the American Medical Informatics Association, vol. 30, no. 3, pp. 492–502, 2023.
[12] P. Tschandl, C. Rosendahl, and H. Kittler, “The ham10000 dataset, a large collection of multi-sourcedermatoscopic images of common pigmented skin lesions,” Scientific Data, vol. 5, p. 180161, 2018.
[13] M. Tan and Q. V. Le, “Efficientnetv2: Smaller models and faster training,” in Proceedings of the International Conference on Machine Learning. PMLR, 2021, pp. 10 096–10 106.
[14] J. Hu, L. Shen, and G. Sun, “Squeeze-and-excitation networks,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018,pp. 7132–7141.
[15] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, “Grad-cam: Visual explanations from deep networks via gradient-based localization,” in Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 618–626.
[16] M. Tan and Q. V. Le, “Efficientnet: Rethinking model scaling for convolutional neural networks,” in Proceedings of the International Conference on Machine Learning. PMLR, 2020, pp. 6105–6114.
[17] X. Zhang, S. Wang, J. Liu, and D. Tao, “Computer-aided diagnosis of skin lesions using vision transformers,” IEEE Transactions on Medical Imaging, vol. 42, no. 4, pp. 1124–1135, 2023.
[18] H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, “mixup: Beyond empirical risk minimization,” in Proceedings of the International Conference on Learning Representations, 2018.
[19] S. Yun, D. Han, S. J. Oh, S. Chun, J. Choe, and Y. Yoo, “Cutmix: Regularization strategy to train strong classifiers with localizable features,” in Proceedings of the IEEE International Conference on Computer Vision, 2019, pp. 6023–6032.
[20] T. Karras, T. Aila, S. Laine, and J. Lehtinen, “Progressive growing of gans for improved quality, stability, and variation,” in Proceedings of theInternational Conference on Learning Representations, 2018.

CITETHISARTICLE:

M.Gughan Raja, S.Afreen Reikhana,S.JeyaVarshini ,“*An intelligent multi-algorithm machine learning system for skin disease classification and prediction”*