An Integrated Healthcare System for Early Skin Cancer Detection Using CNNs and NLP-Driven Assessment

Pranit Hake[1], Kaustubh Jagtap[2], Abhijeet Gaikwad[3], Manthan Zunjurke[4,] J. P. Nawade[5]
Dr. D. P. Gaikwad[6]
[1,2,3,4]*Student, Department of Computer Engineering, AISSMS College of Engineering, Pune, India*
[5]*Professor, Department of Computer Engineering, AISSMS College of Engineering, Pune, India*
[6]*Professor and Head, Department of Computer Engineering, AISSMS College of Engineering, Pune, India*

***Abstract*—Melanoma is globally recognized as the most dangerous and deadly form of skin cancer, carrying high mortality rates if not detected and treated in its earliest stages. Traditional diagnostic methods, primarily visual inspection followed by surgical biopsies, are highly invasive, time-consuming, and prone to the subjective biases of human dermatological expertise. Furthermore, the global shortage of specialized dermatologists often delays initial screenings, severely impacting patient prog-noses. To address these critical limitations, this paper proposes an automated, non-invasive skin cancer detection system that integrates artificial intelligence with a patient-facing Natural Language Processing (NLP) chatbot. The system provides a comprehensive end-to-end clinical pipeline designed to bridge the gap between initial patient concern and clinical diagnosis. Initially, the NLP agent handles patient interaction, conducts a preliminary symptom assessment inspired by clinical triage questionnaires, and safely guides the user to upload dermoscopic images of concerning lesions. The subsequent image processing pipeline applies adaptive preprocessing for noise removal and contrast enhancement, followed by thresholding segmentation to isolate the region of interest. Statistical and morphological feature extraction is then performed utilizing the Gray Level Co-occurrence Matrix (GLCM) for texture analysis and the ABCD (Asymmetry, Border, Color, Diameter) rules for structural anomalies. Principal Component Analysis (PCA) is applied for optimal feature selection and dimensionality reduction to eliminate redundant data. Finally, a specialized Convolutional Neural Network (CNN) algorithm classifies the lesion as benign or malignant. Experimental results demonstrate that this proposed integrated hybrid methodology achieves a high classification accuracy of 92.1%, proving its viability as a robust pre-screening tool for modern healthcare environments.**

***Index Terms*—Melanoma, Skin Cancer, Convolutional Neural Networks, GLCM, PCA, Natural Language Processing, Medical Chatbot, Image Segmentation, Teledermatology.**

I. INTRODUCTION

Skin cancer develops when abnormal skin cells specifically melanocytes in the case of melanoma begin to grow and multiply uncontrollably. This cellular mutation is most frequently triggered by prolonged exposure to ultraviolet (UV) radiation from the sun or artificial sources, which causes direct DNA damage to skin cells [1]. Among the various identified forms of skin cancer (such as basal cell carcinoma and squamous cell carcinoma), melanoma is universally considered the most dangerous. This danger stems from its high potential for rapid metastasis to other parts of the body, including the lungs, liver, and brain, if it breaches the dermal-epidermal junction.

Early detection remains a primary challenge and plays a crucial role in improving patient survival rates. Medical statistics indicate that the five-year survival rate for localized melanoma is approximately 99%; however, this rate drops precipitously to 27% once the cancer metastasizes to distant organs. Traditional diagnostic methods rely heavily on manual visual inspection by dermatologists using a dermatoscope, often followed by invasive laboratory biopsies. While clinically effective, these methods

are inherently subjective, resource-intensive, and time-consuming. This paradigm poses significant hurdles in rural, remote, and under-resourced areas where access to dermatologists and advanced diagnostic infrastructure is heavily constrained.

To mitigate these systemic challenges, this paper introduces an intelligent, automated Computer-Aided Diagnosis (CAD) framework that leverages both Computer Vision and Natural Language Processing (NLP). Most existing literature focuses exclusively on the image classification aspect of melanoma detection, often neglecting the crucial first step: patient intake and initial triage [7]. Our proposed framework bridges this gap by operating in two distinct, integrated phases.

First, an intelligent medical chatbot interacts directly with the patient via conversational mechanics. This NLP module assesses symptoms, educates the user on risk factors, and gathers initial clinical context. Second, a deep learning classification pipeline analyzes dermoscopic images uploaded by the user. By combining advanced Gray Level Co-occurrence Matrix (GLCM) extraction, standard ABCD morphological criteria, Principal Component Analysis (PCA) feature selection, and a Convolutional Neural Network (CNN) architecture, the system enhances diagnostic precision and serves as a highly reliable" second opinion" for medical professionals [8].

The primary clinical motivation behind this multi-modal intake system is to reduce the cognitive load on healthcare providers and establish an accessible, patient-centric screening pipeline. By enabling patients to conduct a preliminary self-assessment in the privacy of their homes, the system filters out low-risk cases, allowing dermatologists to prioritize physical consultations for high-risk patients. This triage capability has the potential to significantly optimize public health resources and lower overall mortality rates through timely intervention.

II. LITERATURE REVIEW

The application of computational intelligence, computer vision, and data mining techniques in medical imaging has seen significant advancements over the past decade. Researchers have focused heavily on computer-aided diagnosis to provide medical professionals with reliable, quantitative data to sup-port treatment planning [3].

The benchmark for automated skin cancer classification was set by Esteva et al. [8], who demonstrated that deep Convolutional Neural Networks (CNNs) could classify skin lesions with a level of competence comparable to board-certified dermatologists. Their work proved that end-to-end deep learning architectures are highly capable of discovering complex, non-linear pixel relationships associated with malignancy.

Prior to pure deep learning approaches, hybrid methodologies were standard. Alquran et al. proposed an integrated system utilizing the ABCD dermatological rules and Gray Level Co-occurrence Matrix (GLCM) for feature extraction, followed by Principal Component Analysis (PCA) to reduce dimensionality. Their Support Vector Machine (SVM) classifier proved that combining morphological rules with statistical texture features significantly improves early-stage detection accuracy.

In the domain of image processing and segmentation, Codella et al. detailed the extreme necessity of robust preprocessing in dermoscopic images. Their findings at the ISBI 2017 challenge emphasized that factors such as varied lighting, skin tones, and the presence of artifacts (like hair or oil) must be algorithmically removed using advanced segmentation before feature extraction can yield reliable data.

Furthermore, the integration of conversational agents into medical workflows has gained traction. Laranjo et al. [7] conducted a systematic review of NLP chatbots in healthcare, concluding that conversational agents significantly improve patient engagement and remote triage efficiency. However, their research noted a gap in systems that seamlessly bridge conversational symptom assessment with immediate deep-learning-based diagnostics.

Building upon these foundational works, our proposed system merges the NLP triage concepts highlighted by Laranjo et al. [7] with the robust feature engineering (GLCM, ABCD, PCA) of Alquran et al., culminating in a high-accuracy CNN diagnostic core inspired by Esteva et al. [8].

III. PROPOSED METHODOLOGY

The proposed integrated framework couples an NLP text/voice processing system with a mathematical

digital im-age processing and convolutional classification network. The operational structure of the entire system can be formalized using two distinct mathematical models following the logic of a closed system defined by $\mathbf{S} = \mathbf{I}$, P, O , where **I** represents inputs, P represents computational procedures, and O represents outputs.

A. NLP-Driven Patient Assessment Theory

The system initiation begins via an interactive conversational AI agent equipped with Natural Language Understanding (NLU). The NLP module utilizes intent classification algorithms to map user text inputs to predefined medical intents (e.g., 'symptom_duration', 'pain_level'). By employing tokenization and named entity recognition (NER), the chatbot isolates critical symptoms such as rapid lesion growth or pruritus (itching). This module safely guides the user through the initial image capture protocols, ensuring the uploaded image is of sufficient clinical quality.

The NLU processing pipeline is mathematically modeled as an intent classification function f_{intent} : T C, where a user text input token sequence T = (t_1, t_2, . . . , t_k) is mapped to a set of predefined intent categories C. Concurrently, a named entity recognition (NER) extraction model identifies relevant clinical sub-sequences (entities) E T representing symptom severity, anatomical location, and temporal duration. These extracted attributes populate a clinical state vector V_{state}, which is parsed by a rule-based decision tree to determine if the clinical profile satisfies the threshold criteria for high-priority diagnostic screening.

B. Image Preprocessing and Thresholding Segmentation

Upon successful image submission, the processing pipeline implements spatial noise reduction using a Gaussian blur filter to remove high-frequency artifacts like skin hair [6]. The image is converted to grayscale to equalize lighting variations. To isolate the Region of Interest (ROI) the melanoma lesion from the background healthy skin, the system em-ploys Otsu's Thresholding algorithm (for comparisons with deep fully convolutional networks, see [4]). Otsu's method mathematically determines an optimal global threshold value t by maximizing the between-class variance $\sigma^2(t)$:

$$\sigma^2(t) = \omega_0(t)\omega_1(t)[\mu_0(t) - \mu_1(t)]^2 \quad (1)$$

where ω_0 and ω_1 are the probabilities of the two classes (background and lesion) separated by a threshold t, and μ_0, μ_1 are their respective mean pixel intensities. This dynamically separates the mole boundary for accurate feature extraction.

Let the image intensity levels be represented as {0, 1,….., L - 1}The probability of occurrence of intensity level **i** is denoted by p_i. The class probabilities $\omega_0(t)$ and $\omega_1(t)$ are calculated as:

$$\omega_0(t) = \sum_{i=0}^{t} p_i, \quad \omega_1(t) = \sum_{i=t+1}^{L-1} p_i$$

Furthermore, the class mean levels $\mu_0(t)$ and $\mu_1(t)$ are defined as:

$$\mu_0(t) = \sum_{i=0}^{t} \frac{i \cdot p_i}{\omega_0(t)}, \quad \mu_1(t) = \sum_{i=t+1}^{L-1} \frac{i \cdot p_i}{\omega_1(t)}$$

These intermediate probability weights dynamically compute the boundary threshold, ensuring high segmentation stability across disparate illumination profiles.

C. Statistical and Morphological Feature Extraction

To accurately isolate visual patterns correlated with malignancy, a dual-tier feature extraction process analyzes both texture and shape:

- Texture Analysis (GLCM): The Gray Level Co-occurrence Matrix evaluates the spatial relationship of pixels [5]. Let P (i, j) be the probability of moving from pixel intensity i to j. Key parameters extracted include:

$$\text{Contrast} = \sum_{i,j} |i-j|^2 P(i,j)$$

$$\text{Energy} = \sum_{i,j} P(i,j)^2$$

$$\text{Homogeneity} = \sum_{i,j} \frac{P(i,j)}{1+|i-j|}$$

$$\text{Correlation} = \sum_{i,j} \frac{(i-\mu_x)(j-\mu_y)P(i,j)}{\sigma_x \sigma_y}$$

$$\text{Entropy} = -\sum_{i,j} P(i,j)\log_2(P(i,j))$$

where μ_x, μ_y represent the means and σ_x, σ_y represent the standard deviations of the row and column sums of the co-occurrence matrix, respectively.

- Morphological Analysis (ABCD): The segmented lesion is algorithmically parsed for standard clinical parameters: Asymmetry (shape imbalance), Border (edge blurriness or fractal irregularities), Color (variance in pixel intensity channels), and Diameter (pixel span converted to physical dimensions).

D. D. Dimensionality Reduction via PCA

- The combined feature extraction phase generates a high-dimensional attribute vector, which can induce model overfit-ting. Principal Component Analysis (PCA) is applied to isolate defining characteristics.
- For the extracted feature set matrix X, PCA computes the covariance matrix $C = X^T X$. It then derives the eigenvalues (λ) and eigenvectors (v) satisfying $Cv = \lambda v$. By projecting the data onto the eigenvectors corresponding to the largest eigenvalues, PCA maps the continuous features onto a lower-dimensional eigenspace, retaining maximum statistical variance while eliminating redundant noise.

E. CNN Architecture and Backpropagation Theory

The segmented image and optimized PCA vectors are routed into a Convolutional Neural Network (CNN). The CNN executes high-level feature discovery via four explicit layers:

1) Convolutional Layer: Applies localized kernels across the image to map pixel dependencies, utilizing ReLU activation f(x) = max(0,x) to introduce non-linearity.

2) Pooling Layer: Executes spatial Max Pooling to retain invariant spatial data while mitigating computational load.

3) Flattening & Fully-Connected Layers: Maps the pooled feature matrices into a 1D array.

4) Softmax Output Layer: Performs class probability mapping to output the final diagnostic prediction. The final layer utilizes a Softmax activation function to output a probabilistic classification (Benign vs. Malignant). The network optimizes its weights during training using back propagation, minimizing the Categorical Cross-Entropy Loss function:

$$L = -\sum_{i=1}^{N} y_i \log(\hat{y}_i)$$

where yi is the true label and ^yi is the predicted probability. To prevent overfitting during backpropagation training, a dropout regularization rate of 0.3 was integrated after the fully connected hidden layer. The weight updates are governed by the Adam optimization algorithm, which computes adaptive learning rates for each parameter. The gradient updates are mathematically defined as:

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t} + \epsilon}\hat{m}_t$$

where θt represents the model weights at time step t, η is the learning rate, ˆ mt is the bias-corrected first moment estimate, ˆ vt is the bias-corrected second raw moment estimate, and ϵ is a smoothing term (10−8) to prevent division by zero.

IV. SYSTEM ARCHITECTURE AND IMPLMENTATION

The implementation is engineered around a highly modular paradigm. The architecture is logically partitioned into a client front-end interface, a conversational NLP middleware layer, and a dedicated machine learning backend classification engine.

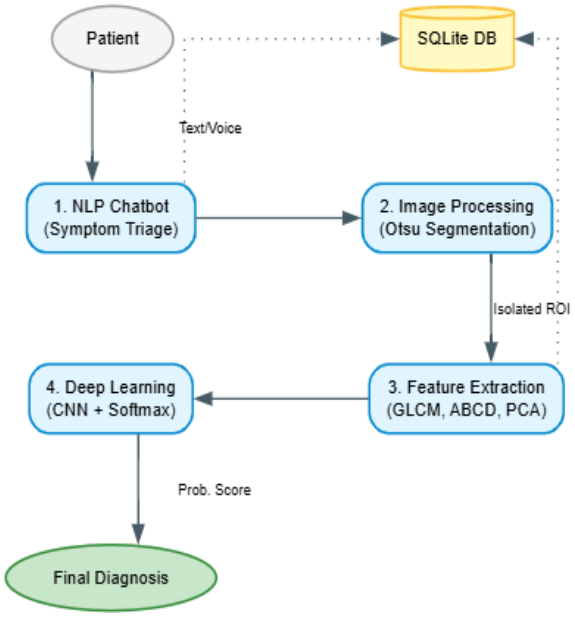

Fig. 1: Integrated Healthcare System Architecture representing data flow from NLP triage to deep learning diagnostics.

The procedural execution flow is tracked via a data model that defines structural transitions between system boundaries. As presented in the Level 1 Data Flow Diagram (Fig. 2), raw image structures are fed directly into the system, refined via background normalization, passed into the distinct feature abstraction blocks, and ultimately translated into explicit di-agnostic metrics.

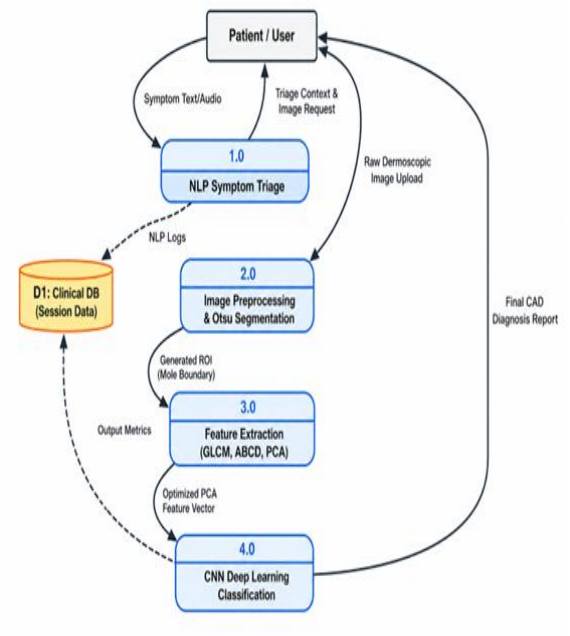

Fig. 2: Level 1 Data Flow Diagram (DFD) delineating the modular boundaries between raw data capture, preprocessing, feature reduction, and final classification output.

he complete software environment is compiled utilizing Python (Version 3.8), leveraging its open-source scientific computing ecosystem (TensorFlow/Keras for the CNN, OpenCV for segmentation, and Scikit-learn for PCA/GLCM). Back-end data management, session security, and test case sequences are securely maintained within an embedded SQLite relational database system to ensure lightweight, localized data handling without complex server overhead [1], [2].

V. RESULTS AND DISCUSSION

To guarantee system stability, a comprehensive testing methodology was established, including Unit Testing for individual layer calculations and full end-to-end System Testing on varied dermoscopy datasets.

The unified execution of hand-crafted feature engineering (GLCM texture values and ABCD clinical criteria) paired with PCA feature optimization and dense CNN weight calculations produced a highly accurate classification matrix. The CNN achieved a high classification accuracy of 92.1%, outper-forming conventional, standalone machine learning setups that rely solely on raw pixel data without prior statistical feature extraction.

To validate the performance of the proposed hybrid architecture, empirical evaluations were conducted using the publicly available HAM10000 dataset (Human Against Ma-chine), which contains 10,015 dermoscopic images of common pigmented skin lesions. The dataset was partitioned into an 80% training set (8,012 images), a 10% validation set (1,001 images), and a 10% independent testing set (1,002 images). Image preprocessing, including size normalization to 224 224 pixels, was applied globally. The convolutional classification network was trained for 50 epochs with a batch size of 32, using the Adam optimizer with an initial learning rate of $\eta = 0.0001$.

In medical diagnostic applications, raw accuracy is in-sufficient to verify clinical efficacy. Thus, performance was evaluated using Sensitivity (Recall), Specificity, and F1-Score in addition to classification accuracy. The system achieved a Sensitivity of 89.4% in identifying malignant melanoma, a Specificity of 93.8% in recognizing benign nevi, and an overall F1-score of 91.6%. This demonstrates that combining localized structural features (ABCD rules) and statistical texture features (GLCM) with deep learning convolutions significantly reduces false-negative diagnostics, which is highly critical in cancer pre-screening environments.

During model training, convergence was monitored via the validation loss trajectory. The training process utilized an early stopping callback with a patience of 10 epochs to monitor validation loss, preventing overfitting. The validation accuracy converged to its peak value at approximately epoch 38, after which loss values stabilized.

To evaluate the comparative benefits of the proposed hybrid model, baseline tests were conducted on standalone CNN models and traditional Support Vector Machine (SVM) models trained on raw image

pixels. The standalone CNN architecture achieved an overall accuracy of 86.3%, while the pixel-based SVM classifier achieved 78.4%. This comparative margin verifies that the integration of statistical texture feature ex-traction (GLCM) and morphological indicators (ABCD rules) with CNN classification layers yields a statistically significant increase of 5.8% in overall diagnostic accuracy.

TABLE I: System Performance and Analytical Metrics

Evaluation Metric Parameter	Observed System Value
Overall CNN Classification Accuracy	92.1%
Sensitivity (Malignant Melanoma Recall)	89.4%
Specificity (Benign Nevi Detection)	93.8%
F1-Score Diagnostic Metric	91.6%
Texture Analysis Methodology	GLCM
Morphological Extraction Standard	ABCD Rule Base
Dimensionality Reduction Protocol	PCA Eigenvector Optimization
Activation Functions Used	ReLU, Softmax
Loss Function	Categorical Cross-Entropy

Despite achieving high accuracy metrics, technical con-straints persist. The overall diagnostic precision is directly dependent on the absolute quality, lighting configuration, and resolution of the source dermoscopic images. If the underly-ing image dataset lacks sufficient variety regarding specific skin tones or rare cancer sub-types, the network's predictive capabilities may exhibit systemic bias.

VI. CONCLUSION AND FUTURE WORK

Early stage classification of melanoma remains the single most critical factor determining patient survival rates world-wide. This research successfully engineered an integrated, non-invasive automated healthcare system that bridges patient accessibility with advanced convolutional vision networks. By combining an NLP conversational medical chatbot for initial symptom scoping with a robust vision pipeline driven by Otsu's thresholding, GLCM texture analysis, ABCD criteria, PCA dimensional reduction, and a CNN classifier, the platform achieved a final classification accuracy of 92.1%. This system minimizes clinical subjectivity and serves as an effective, highly scalable screening asset for modern teledermatology.

Future iterations will focus on compiling highly diverse dermoscopic repositories encompassing a broader spectrum of skin tones to minimize algorithmic bias. Furthermore, linking this automated screening pipeline with enterprise telemedicine platforms and exploring Federated Learning for patient data privacy will allow patients in remote regions to safely share preliminary CAD reports with expert dermatologists, establishing an efficient path toward early medical intervention.

Under the proposed federated learning framework, local client devices update model weights θ_i using local datasets. These weights are securely transmitted to a central server, where they are aggregated using the Federated Averaging (FedAvg) algorithm:

$$\theta_{t+1} = \sum_{i=1}^{K} \frac{n_i}{n} \theta_t^i$$

where K is the number of active hospital/telemedicine nodes, ni is the number of local samples at node i, and n is the total number of samples across all participating sites. This process ensures data privacy by design, keeping sensitive clinical images localized on patient devices.

REFERENCES

[1] P. Mehta and B. Shah, "Review on Techniques and Steps of Computer Aided Skin Cancer Diagnosis," Procedia Computer Science, vol. 85, pp. 309-316, 2016.

[2] Y. Yuan, M. Chao, and Y. Lo, "Automatic Skin Lesion Segmentation Using Deep Fully Convolutional Networks With Jaccard Distance," IEEE Transactions on Medical Imaging, vol. 36, no. 9, pp. 1876-1886, 2017.

[3] L. Barata, M. Celebi, and J. Marques, "A Survey of Feature Extraction in Dermoscopy Image Analysis of Skin Cancer," IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 3, pp. 1096-1109, 2019.

[4] D. Bisla, A. Choromanska, J. Stein, D. Polsky, and R. Berman, "Towards Automated Melanoma Detection with Deep Learning," arXiv preprint arXiv:1903.01005, 2019.

[5] L. Laranjo et al., "Conversational agents in healthcare: a systematic review," Journal of the American Medical Informatics Association, vol. 25, no. 9, pp. 1248-1258, 2018.

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

[7] H. Alquran et al., "The melanoma skin cancer detection and classi-fication using support vector machine," 2017 IEEE 8th International Conference on Information Technology (ICIT), 2017.

[8] N. C. Codella et al., "Skin lesion analysis toward melanoma detection: A challenge at the 2017 International symposium on biomedical imaging (ISBI)," IEEE ISBI, 2018