

A Constructive Auto-Encoder Based Convolutional Neural Network Model for Skin Lesions Prediction Using Compression Model

Arpita Roy¹, Surabhi Saxena², Mohammed Alisha³, Jagjit Singh Dhattewal⁴, Tarak Hussain⁵

¹Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India

²Department of Computer Science, CHRIST University, Bengaluru, India

³Department of Artificial Intelligence and Machine Learning, Aditya University, Surampalem, Kakinada District, Andhra Pradesh, India

⁴School of Computer Science & Artificial Intelligence SR University, Ananthasagar, Hasanparthy (M), Warangal, Telangana 506371, India.

⁵Department of Artificial Intelligence and Data Science, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India

¹arpitaroy@kluniversity.in, ²saxenasurabhi1987@gmail.com, ³mohammedalisha@gmail.com, ⁴jagjits247@gmail.com,

⁵tariqsheekh2000@gmail.com

Corresponding author: Tarak Hussain

Department of Artificial Intelligence and Data Science, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India

ARTICLE INFO

ABSTRACT

Received: 24 Dec 2024

Revised: 12 Feb 2025

Accepted: 26 Feb 2025

In medical imaging, skin lesion prediction and classification is highly crucial while predicting skin malignancy. Various prevailing deep learning-based CAD diagnosis approaches show poor performance. It is incredibly challenging to diagnose skin lesions with complex features like artefacts, boundary analysis, low contrast images with poor foreground and background images, and constraint training datasets. Also, it relies on the appropriate tuning of millions of parameters that causes poor generalization, overfitting, and massive consumption of computing resources. This research concentrates on modelling an efficient framework that performs some preliminary processes like pre-processing, segmentation and classification of skin lesions for automated prediction of skin lesions. The anticipated framework is composed of five stages: 1) pre-processing with adaptive median filtering; 2) segmentation with enhanced active contour-based approach; 3) ROI region compression with hybrid Lampel Ziv Wolch; 4) Non-ROI region compression with Fractal Image Compression (FIC) and 5) Classification with auto-encoder CNN. The proposed auto-encoder-based CNN model is designed with sub-networks connected in series, which is more efficient than conventional approaches. The simulation is done in the MATLAB 2020a environment. The performance of the anticipated model is compared with some standard methods like CNN and CNN-PSO. The predicted model shows a better trade-off than the prevailing approaches.

Keywords: Skin lesion, auto-encoder CNN, region of interest, deep learning, classification

1. INTRODUCTION

The skin lesions are diagnosed by obtaining the images using a computer that involves extraction of the lesion boundary that gets differentiated from the adjacent healthy skin that is the segmentation of skin lesion till a few years ago [1]. This process continued to feature calculation depending on rules like the ABCD and CASH rules. It is generated with the help of dermatologists depending on the attained segmentation. Eventually, these features are used to train the traditional machine learning models such as random decision forests and support vector machines to suggest the diagnosis [2]. The deep learning technology is used to directly predict diagnosis from the images, followed by bypassing the segmentation available in the commonplace. It is visible in the other imaging modalities since skin lesion segmentation is an intermediary process in the dermatological analysis [3]. Instead of prioritizing to produce the prediction accuracy of the clinical task ultimately, such as clinical management, the system projects the same trend where this model depreciates in prediction to diagnose [4]. No work is published in predicting disease management directly. At the same time, the diagnosis is based on deep learning under the dermatological criteria from the images attaining the achievement level of medical professionals. Also, when an automated prediction model is used to decide the diagnosis, the dermatologist or general physician needs to decide on the management of the

disease that would be the treatment plan or another course of action [5]. However, diagnosing underlying skin criteria accurately from an image alone is impossible in many circumstances. It is illustrated with an example of the evaluation of 'majority decision' attained from about a hundred dermatologists in a recent study to classify melanoma, resulting in 71.8% sensitivity in terms of ground truth in diagnosis [6].

Considering the lesion's visual presentation is enigmatic rather than diagnosing the criteria. Then the accurate work can be to perform a biopsy to get the needed information. Machine learning-based techniques are used to classify underlying skin criteria and directly utilize the predicted skin criteria to decide disease management [7]. This decision may not be differentiated by various decision management in a single class [7]. The decision management (for instance, a follow-up visit needs to be scheduled to observe the progression of skin lesion) can be essential to approve the diagnosis (While the inadequate information is available within the image) [8]. After that, it needs to proceed. It illustrated with an instance the decision on clinical management for a naïve with no unusual features requiring extra tasks. Instead of the nevi having unique features, the clinical follow-up or excision is opted by the dermatologist based on the severity of distinctive features [9]. Hence, it helps investigate the performance of an automatic prediction system for skin disease management based on artificial intelligence [10]. This system can propose clinical management decisions considered the second choice or propose to patients directly in under-served communities [11]. However, the prediction of decisions management believes as the simple calculation issue to mention instead of predicting the diagnosis and concluding the direction.

The non-invasive skin imaging technique is Dermoscopy that helps represent the magnified and illuminated skin lesion image to enhance the marks clarity. The deeper skin lever has a visual impact that can be increased when the reflection of the skin surface is eliminated [12]. Moreover, the melanoma dermoscopy images that can be recognized automatically are complex because of some features. Primarily, the skin lesions segmentation is challenging and considered difficult work due to the intra-class lesion differentiation like size, texture, shape, colour, and location. Secondly, the similarities among non-melanoma and melanoma lesions are high. Lastly, the conditions of the surrounding circumstance such as veins, hair, colour calibration charts, and the ruler marks are considered. These problematic issues are overcome by considering many trials. At the earlier stage, scientists managed to differentiate the melanoma lesions and non-melanoma lesions with the help of low-level features hand-crafted [13]. Few scientists introduced algorithms to choose features of proper hand-crafted.

On the other hand, high visual similarity affects the features and intra-class, and artefacts of dermoscopy images have high variations that turn to poor outcomes. A different set of scientists tried segmentation methods to dispose of unwanted features and backgrounds. In addition, the segmentation steps and the classification depend on the low-level factors and low discrimination abilities that lead to poor outcomes [14]. The author proposed the set of high-level intuitive features (HLIF) to explain the lesion border volume asymmetrically. The small HLIF group for low-level features are utilized to obtain a more semantic meaning of the feature set. The intuitive rationale decision is provided for classification purpose allowed by the proposed system. The classification rate of 87.38% is achieved using the mentioned system. The traditional NN and SVM classify skin cancer.

The segmentation and enhancement process is carried out with the help of pre-processing steps where LBP and the dissimilarity on the texture analysis of inverse probabilities are used. The rate of accuracy is 71.4% is achieved by the system. The computer-aided diagnosis system is anticipated by author et al. [15], where the features are taken out with the help of the ABCD rule. For the classification, SVM is utilized. The rate of accuracy for the system was 75.1%. The texture and colour descriptors are used to acquire the lesion's region for the classification. The system achieved the rate of classification, about 81%. The input images have noises are eliminated with the help of mentioned filtering techniques that are proposed. Then the ABCD rule is applied to acquire the features of skin lesions. The model achieves 79% of accuracy. However, these models fail to provide superior prediction accuracy with fewer samples and do not concentrate on the compression problem. Henceforth, this research focuses on modelling an efficient framework for effectual prediction using DL concepts and providing superior outcomes. The significant contributions are listed below;

1) The dataset for predicting skin lesions are acquired from the online resources and provided for some preliminary steps like 1) pre-processing with adaptive median filtering; 2) segmentation with enhanced active contour-based

approach; 3) ROI region compression with hybrid LZW; 4) Non-ROI region compression with Fractal Image Compression (FIC) and 4) Classification with auto-encoder.

2) The hybrid LZW and FIC extract the ROI of the provided image, and the idea is constructed and reconstructed with the anticipated model.

3) Finally, classification is done with an auto-encoder based CNN model, and the comparison is made among the various existing approaches.

The work is structured as: section 2 provides a comprehensive analysis of various prevailing approaches and discusses the pros and cons of those models. In section 3, the methodology is elaborated well to offer extensive knowledge of segmentation and classification. The numerical outcomes of the anticipated model are provided in section 4, followed by the research summary in section 5.

2. RELATED WORKS

Most studies' implementation of machine learning in dermatology focused on categorizing skin lesions for various diseases. It includes non-melanoma skin cancer (NMSC), atopic dermatitis, melanoma, onychomycosis and psoriasis rosacea. The research focused primarily on CNN to recognize and classify the image. The feature extraction is mainly carried out with the help of pre-trained CNN that is AlexNet, and those features were classified using more simple machine learning algorithms like support vector machines [16] and KNN. Most CNNs can extricate features and perform the image classification from E2E learning.

Some common invasive cancer is melanoma in fifth place, and the incidence is kept increasing across the globe [18]. The vast majority of mortalities related to skin cancer are identified with the help of melanomas. Skin cancer is shielded with the entire body skin investigation visually. The NHIS provides data specifies the screening rates are slow significantly (13% in women and 16% in men) [19]. The primary deep learning algorithm concentrated on melanoma classification to mention the low screening rates and enhance the evaluation successively. The infectious melanoma is classified from one of the primary milestone researches with considered accuracy proposed by Esteva et al. in 2017 [20]. The CNN, namely Google Inception V3, is used by this traditional research which pre-trained before on 1.28 million general objects images. The authors train the algorithm using transfer learning with 129,450 clinical and dermoscopic images. The primary classifier exhibits commensurate accuracy in classifying seborrheic keratosis from keratinocyte carcinoma. Benign nevi compared to infectious melanoma when two dermatologists attained 66% and 65.56% accuracy. At the same time, the CNN achieved an accuracy of 72.1% correspondingly. The overall AUC is 91% acquired by CNN, identical to the average predictions of the output of 21 dermatologists. More research showed the commensurate accuracy, sensitivity, AUROC, and specificity to practising dermatologists [21] or certified dermatologists since the leveraged transfer learning for the classification of lesions into multiple skin cancer classes and evaluates the probability of infectious region. It is essential to consider the average accuracy to diagnosis by dermatologists related to specificity and sensitivity while determining the machine learning models to screen in general. The critical note describes the authors' systematic review of predicted research about the melanomas' diagnostic accuracy. The sensitivity was 81% to 100% for dermatologists and was 42% to 100% [40] for primary care physicians (PCPs). One of the studies that reported specificity would be 98% for PCPs when no other studies have reported the specificity in the review for dermatologists. The dermatologists have specificity, and PCPs varies from 70% - 89% and 70% - 87% [22]. The dermatologists have sensitivity and PCPs ranging from 82% - 100% and from 70% - 88% for the referral or biopsy accuracy. Henceforth, the accuracy of referral or biopsy and the same diagnostic accuracy is demonstrated by the machine learning models can be proved usefully to screen melanoma in general [23].

CNN's can show the upper-level specificity and sensitivity in melanoma classification are established by the Brinker and colleagues studies, contrasted to the certified dermatologists and the under training dermatologists [24]. These authors utilized the transfer learning over pre-trained ResNet50 CNN and trained it with 4204 melanoma and nevi proved images of biopsy (1:1) in one of the studies. One hundred forty-four dermatologists (92 junior dermatologists and 52 board-certified dermatologists) are used to determine the 804 dermoscopic images, proven based on biopsy for melanoma compared to nevi. The greater specificity and sensitivity are attained by trained CNN contrasted with junior dermatologists and board-certified dermatologists. Even though these findings are promised, the same overall

dataset is utilized for validation and training [25]. The algorithm is also not validated externally. Henceforth, it is unclear if these outcomes are evident to other data sets. The authors have explained the limitations and suggested future research which fine-tuned the CNN with less sample of modern images previous to use these new data sets.

When dermatologists work along with machine learning techniques, the mean accuracy can be grown to about 1.36%. Nonetheless, the forthcoming validation in predicted clinical trials needs to be performed than dermatologists before claiming the performance superiority of algorithm in real-world settings. Bissoto et al. [26] investigated with the latest study, which combines the artificial intelligence and the human achieves the better classification of images than to dermatologists alone or CNN [27] classification only when few researchers marked the algorithms of artificial intelligence over dermatologists [28]. It is essential to find a better option to enhance the chances of combining artificial intelligence and humans when the outcomes are not significantly statistical [29] – [30].

3. METHODOLOGY

The methodology section includes five different phases: 1) pre-processing with adaptive median filtering; 2) segmentation with enhanced active contour-based approach; 3) ROI region compression with hybrid LZW; 4) Non-ROI region compression with Fractal Image Compression (FIC) and 4) Classification with auto-encoder. The outcomes acquired from the above phases are validated using MATLAB 2020a simulation environment, and the results are compared with various existing approaches like standard CNN and CNN with PSO.

a. Dataset

PH2 datasets are constructed with the collaboration of the dermatology service and some research universities. The images (dermoscopic) are attained from the Tuebingen mole analyzer system with 20x magnification. It is made of 8-bit RGB color images with 768*560 pixel resolution and the database comprises 200 dermoscopic images with 80 atypical nevi, 80 nevi and 40 melanomas. The images are skin type I/II based on the classification scale. Thus, the skin color varies from white to cream-white color. The images are chosen with more concentration with superior resolution, quality, and dermoscopic features [31]. The images are measured by the expert dermatologists with the following parameters: 1) manual skin lesions segmentation; 2) histological and clinical diagnosis; 3) dermoscopic criteria (blue-white veil, regression area, streaks, pigment network, asymmetry, colors and dots/globules). Dermatologists carry manual segmentation and image annotation with the customized tool for retrieving DerMAT images.

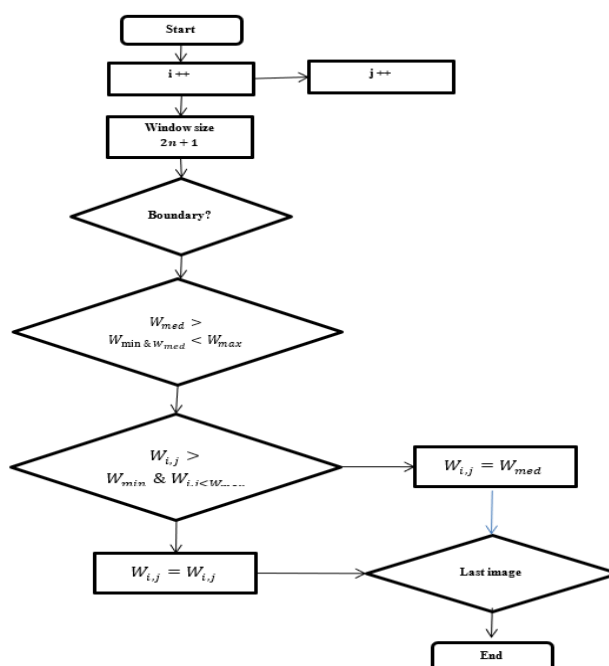


Fig 1 Flow diagram of adaptive median filtering

b. Adaptive median filtering

It is hypothetically a grey image with 8 bits; when the pixels comes under salt and pepper and ranges from 0 or 255. The PDF is mathematically expressed as in Eq. (1):

$$P(x) = \begin{cases} Pa, x & = a \\ Pb, x & = b \\ 0 & Other \end{cases} \quad (1)$$

From the above Eq. (1), a is 0 and b is 255, and these are considered white and black noise points; Pa and Pb specify probability related to ' a ' and ' b '. The preliminary median filter sets the ordering points with filtering window and uses intermediate values for sequential sort to replace specific window points. It is mathematically expressed as in Eq. (2):

$$I(i, j) = Median(n(k)) \quad (2)$$

Based on these formulas, K specifies the number of pixels over the window, n specifies the grey value sequence sort, i and j specifies the horizontal and vertical pixel coordinate. In the conventional methods, when the noise probability $P(x)$ is higher than 25%, the filtering process is declined sharply and it is due to the fact that median filter is not discriminate towards the noise points. The window is invalid with more than 50% noise point, and the median point is substituted with noise point that causes filter loss. It provides a condition to differentiate the substituted median filter whether or not the noise point is practical to eliminate the filter problem during failure. Various steps must be followed: 1) set an outer loop layer to predict the maximal value among the windows set as W_{max}, W_{min} and W_{med} . Then, set $A1 = W_{med} - W_{min}$, $A2 = W_{med} - W_{max}$, if $A1 > 0$ and $A2 < 0$, then move to inner loop and set as B . The window size is based on $W + 1$ to perform constant expansion. In the successive layers B , $B1 = W_{i,j} - W_{min}$, $B2 = W_{i,j} - W_{max}$. If $B1 > 0$ and $B2 < 0$ unchanged current pixel values; else, $W_{i,j} = W_{med}$. The flow diagram of the adaptive median filter is shown in Fig 1.

c. Enhanced active contour-based segmentation

Based on the non-local operator, this work is enhanced with the color image segmentation for color texture image segmentation. It is expressed as in Eq. (3):

$$\min_{u, \emptyset \in \{0,1\}} \{E(u_{ij}, \emptyset) = \sum_{i=1}^2 \sum_{j=1}^3 \alpha_i \int_{\Omega} ((u_{ij} - f_j)^2 + \beta_i |\nabla u_{ij}|^2 + \lambda_i |\nabla_{NL} u_{ij}|^2) X_i(\emptyset) dx + \sum_{i=1}^1 \gamma \int_{\Omega} |\nabla_{\emptyset_i}| dx \} \quad (3)$$

Here, $f(x) = (f_1(x), f_2(x), f_3(x))$; $\Omega \in R^3$ specifies color texture image and $f_i(x)$ specifies color image channel, $(u_{ij} - f_j)^2$ specifies fidelity term with the processed image which is similar to original appearance. The initial smooth term $|\nabla u_{ij}|^2$ defines the active contour is influenced by the image gradients. The non-local term represents the active contour model, controlled by the slices of similar texture. $\alpha_i, \beta_i, \gamma_i, \lambda$ are penalty parameters that manages the item contribution. It is considered as the piecewise smooth segmentation and non-local segmentation process. When λ_i specifies zero, it becomes the multi-channel active contour model. When β_i represents zero, the model becomes the non-local mean approach. The above Eq. (3) is alternated with the minimization method. It is expressed as in Eq. (4):

$$\min_{u_{ij}} \{E(u_{ij}) = \sum_{i=1}^2 \sum_{j=1}^3 \alpha_i \int_{\Omega} ((u_{ij} - f_j)^2 + \beta_i |\nabla u_{ij}|^2 + \lambda_i |\nabla_{NL} u_{ij}|^2) X_i(\emptyset) dx \} \quad (4)$$

Here, Euler-Lagrange equation is provided to generate the energy function. It is expressed as in Eq. (5):

$$u_{ij} = \frac{f(x)x_i(\emptyset) + \beta_i^* C + \lambda_i \int_{\Omega} (X_i(\emptyset(y)) + X_i(\emptyset(x))) u_{i,j}(y) w(x, y) dy}{X_i(\emptyset) + 4 * \beta_i + \lambda_i \int_{\Omega} (X_i(\emptyset(y)) + X_i(\emptyset(x))) w(x, y) dy} \quad (5)$$

$$b^{k+1} = b^k + v^k, \quad b^0 = v^0 = 0; \quad \text{Here, } b = \text{Bregman iteration parameter} \quad (6)$$

$$\begin{cases} \nabla \cdot (v - \nabla \emptyset - b^{k+1}) + \sum_{i=1}^2 Q_i(u) \frac{\partial X_i(\emptyset)}{\partial(\emptyset)} & \text{in } \Omega \\ (v - \nabla \emptyset - b^{k+1}) \cdot \vec{n} & \text{on } \Omega \end{cases} \quad (7)$$

$$v^{k+1} = \text{Max} \left(|\nabla \emptyset^{k+1} + b^{k+1}| - \frac{\gamma}{\theta}, 0 \right) \frac{\nabla \emptyset^{k+1} + b^{k+1}}{|\nabla \emptyset^{k+1} + b^{k+1}|} \quad (8)$$

The algorithm for the anticipated segmentation model is provided below:

Algorithm 1: Enhanced active contour-based segmentation

1. **Initialize** $k = 0$; $\vec{b}^0 = \vec{v}^0 = 0$; $u_{ij}^0 = f_j^0$; $j = 1, 2, 3$; $i = 1, 2$; no. of iterations;
 2. **For** all iterations = 0 to N , do
 3. Update every weight from $w(x, y) = \exp \left\{ -\frac{G_{\sigma}^* (|f(x+) - f(y+)|)^2}{h^2} \right\}$; $// f(x) \rightarrow \text{grey image intensity}; x \in \Omega, y \in \Omega, G_{\sigma} \rightarrow \text{Gaussian kernel function}, \sigma \rightarrow \text{SD of Gaussian Kernel}; h \rightarrow \text{threshold of patch similarities}, w: \Omega * \Omega \rightarrow R$ specifies similarity index among patch centred points x & y .
 4. Evaluate u_{ij} from Eq. (5);
 5. Compute $\emptyset$ from Eq. (6);
 6. Compute v^{k+1} from Eq. (8);
 7. $k = k + 1$;
 8. **end for**
-

d. ROI region compression using hybrid enhanced LZW

Lebel-Ziv-Welch (LZW) coding is a dictionary-based technique with several image pixels are identified using a single index. The provided dictionary is composed of character sequences chosen from the image. An index is allocated to every character sequence over a maximum 4096 characters. Also, it is known as greedy algorithm where the dictionary is updated in alliteration for the available strings, i.e. character sequence. The following are the Pseudocode of LZW coding:

Pseudocode 1: LZW

1. Initialize the dictionary for all input strings with length 1;
 2. Predict the longest string for all the current input symbols based on the dictionary usage;
 3. Emit dictionary index for all the strings and remove the input;
 4. Go to successive symbols and do the same for all input symbols;
 5. Proceed with the process;
-

The anticipated model is tested over the skin lesions, and the samples are acquired from the online repository where the images possess various dimensions. All the available images are tested with 8-bit image depth. The ROI region extraction separates the background and foreground images using the proposed model, and the bounding box operation is applied over the entire image volume. The coordinate values of all slices are recorded, and the global coordinate values are predicted without refusing the active image slice portion.

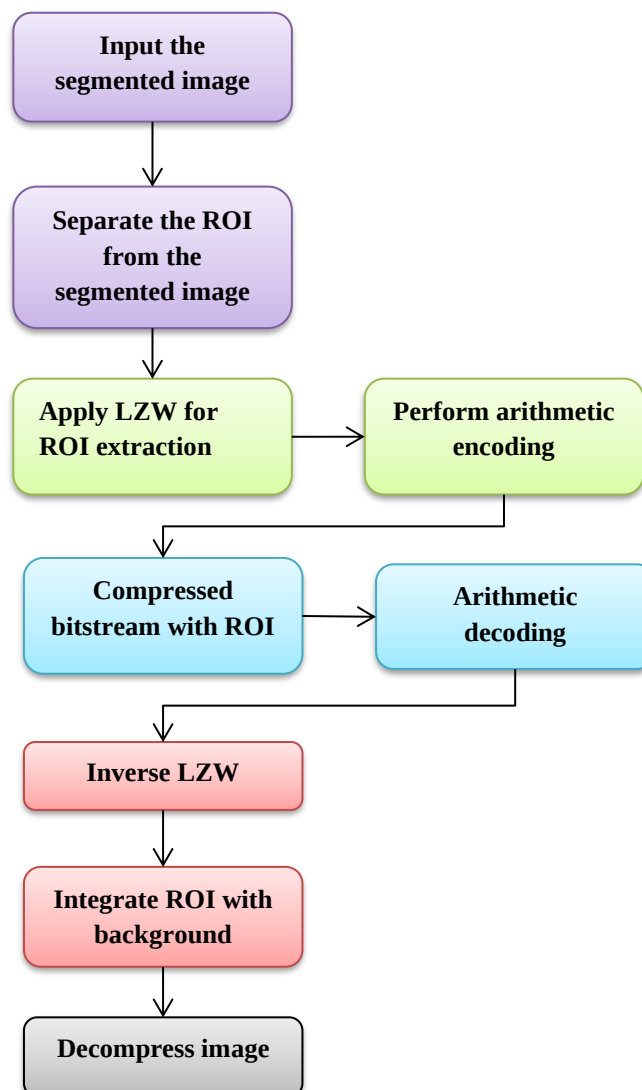


Fig 2 ROI region compression using hybrid LZW

The anticipated model eliminates the foremost blank slices and the end slices with essential information. It is termed null image slices. The bounding box volume is designed with the global coordinates of morphological operation. Moreover, all the encoder and decoder process is applied for ROI extraction (See Fig 2). After the decoding process, the entire image is acquired by integrating the ROI to the background with the encoder coordinates. After ROI extraction, LZW coding is applied to the extracted ROI using arithmetic encoding. The outcomes are attained from the LZW coding and fed to the encoder, resulting in the compressed bit-stream of the ROI. Indeed of the compressed bitstream, the bounding volume coordinate is required for reconstructing the actual image volume sent to the decoder. Here, decoding is done with the compressed bitstreams. The inverse LZW follows the inverse decoding operation with arithmetic decoding. The inverse function is applied over the compressed bitstream to decompress ROI. It is provided for further process, known as the reconstruction process. After completing compression and decompression, ROI is extracted and integrated with the background. The bounding box coordinates include the actual or original size of the original image volume. Therefore, the compressed image is decompressed using the decoding process.

e. Non-ROI region compression using Fractal image compression (FIC)

Here, the image is composed of essential information to make further processing. The image is compressed and partitioned into the series of $N * N$ pixel sub-blocks and does not overlap and deal with the complete image. It is known as a range block, and the encoded image is partitioned into domain blocks with $2N * 2N$ and D coinciding with one another. Before completing the encoding process, D is averaged by the successive neighbourhood pixels, and the size is diminished as R blocks. The average sample is subjected to various kinds of equidistance transformation and constitutes a codebook. It is essential to predict the matching block over the codebook Ω . The block is approximated with the luminance transform for matching the block, $R = s \cdot D + o \cdot 1$, where one is known as the unit matrix $N * N$, s, o is represented by the contrast and brightness with the adjustment factors of D blocks. The algorithm for the anticipated model is shown below:

Algorithm 2: Non-ROI region compression using Fractal image compression (FIC)

Input: Image size $M * M$;

Output: Fractal encoded file $\{t, s, o, dx, dy\}$;

1. Divide the block by partitioning the image F and partition into R block whose $N * N$ size; //does not overlap with one another;
2. Intercepting window $2N * 2N$ is moved in the vertical/horizontal image directions using step size δ , and intercepted block is constituted with domain block.
3. Perform equidistant and average sampling transformation of all blocks to form codebook;
4. Predict the matching block D_j^i that fulfils $d(R_i W_i(D_j^i)) = \min \left\| R_i - (s_i \cdot (t_k(D_j^i)) + o_i) \right\|^2$; // D_j^i specifies domain block after average sampling, $t_k \in \{t_1, \dots, t_8\}$ is eight isometric transforms, s_i and o_i are contrast and luminance factor for all gradation transformation.
5. Compute s and o using Eq. (9) & Eq. (10);

$$s = \frac{n \sum_{p=1}^n d_p r_p - \sum_{p=1}^n d_p \sum_{p=1}^n r_p}{n \sum_{p=1}^n d_p^2 - (\sum_{p=1}^n d_p)^2}; \quad (9)$$

$$o = \frac{1}{n} \left[\sum_{p=1}^n r_p - s \sum_{p=1}^n d_p \right]; \quad (10)$$

f. Auto-encoder CNN algorithm

It is the simplest form of CNN with single hidden fully connected layers, i.e. input/output layers, as in Fig 2. The number of nodes over the input layer is similar to the output layers (See Fig 3). The provided auto-encoder model generates a newer representation for the input data via a mapping pair $y \rightarrow z \rightarrow y'$. The initial one is the encoder map $z = f(y)$, and the second one is the decoder map $y' = g(z)$. The dimensionality of the data is reduced using the encoder model. The encoder transforms the input data dimensionality to a smaller dimension. Similarly, the decoder intends to reconstruct the input data dimensionally reduced to the original. The auto-encoder is generally forced to prioritize the input factors during the training process. The encoder for CNN maps the input point to code through the sigmoid activation function as in Eq. (11):

$$z = f(y) = \sigma(Wy^T + b) \quad (11)$$

Here, ' W ' is a weighted matrix; ' b ' is a bias vector. ' z ' is the latent representation of the image where the sigmoid function transforms the input values to the activation function, which is generally nearer to 0 or 1, respectively. The decoder needs to map the activation function for the reconstruction of y' to the simple dimensional space as in Eq. (12):

$$y' = g(z) = \sigma(W'z + \bar{b})^T \quad (12)$$

The decoder's weighted matrix is the encoder weighted matrix transpose, $W' = W^T$. The encoder training represents the prediction of optimal values of W , b and $\bar{b}$ that reduces the cost function. The stacked encoder is constructed with layers to the encoder and encoder side. Here, the encoder is adopted for CNN followed by convolution layers for extracting spatial features from the input image. The extracted features are processed via CNN for learning dependencies over the signers.

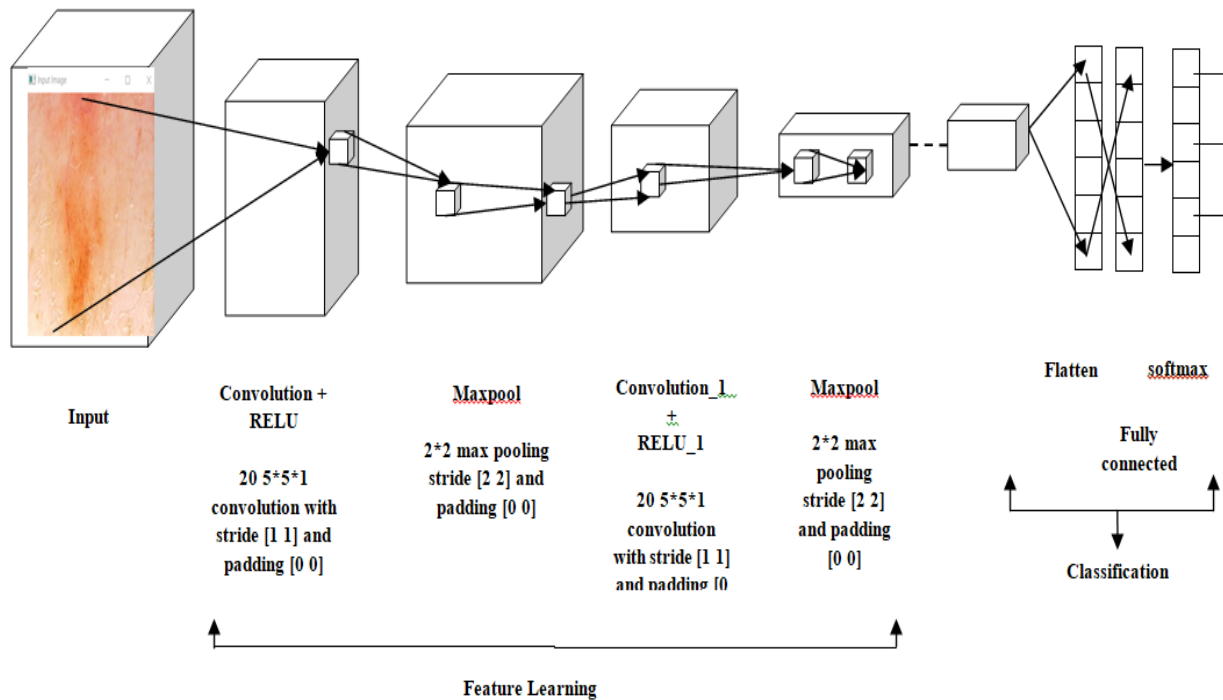


Fig 3 Auto-encoder CNN based architectural model

4. Numerical results and discussion

The experimentation is done in a Python simulation environment in windows OS, 64 GB memory. The complete experiment is based on an open-source deep learning framework. Some standard metrics are considered to evaluate the performance of the anticipated model at diverse stages. They are discussed below:

a. Recall (sensitivity): The proportion of positives (actual) that are identified appropriately.

$$Sensitivity = \frac{TP}{TP + FN} \quad (13)$$

b. Precision: It is the proportion of predicted positive observations (correctly) to the total indicated positive observations. It is expressed in Eq. (14):

$$Precision = \frac{TP}{TP + FN} \quad (14)$$

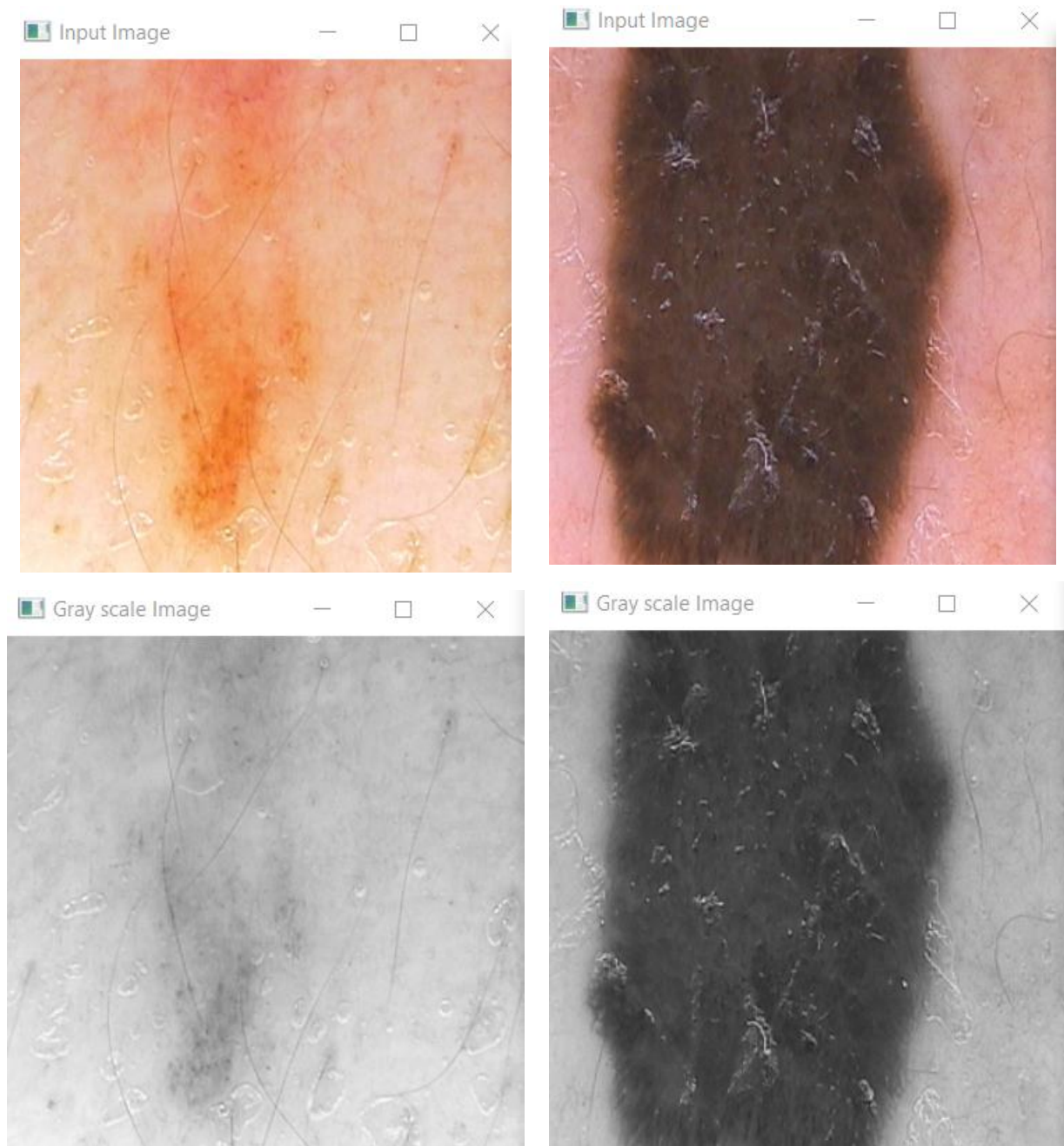
c. F1-score: It is the weighted average of recall and precision. It is expressed as in Eq. (15):

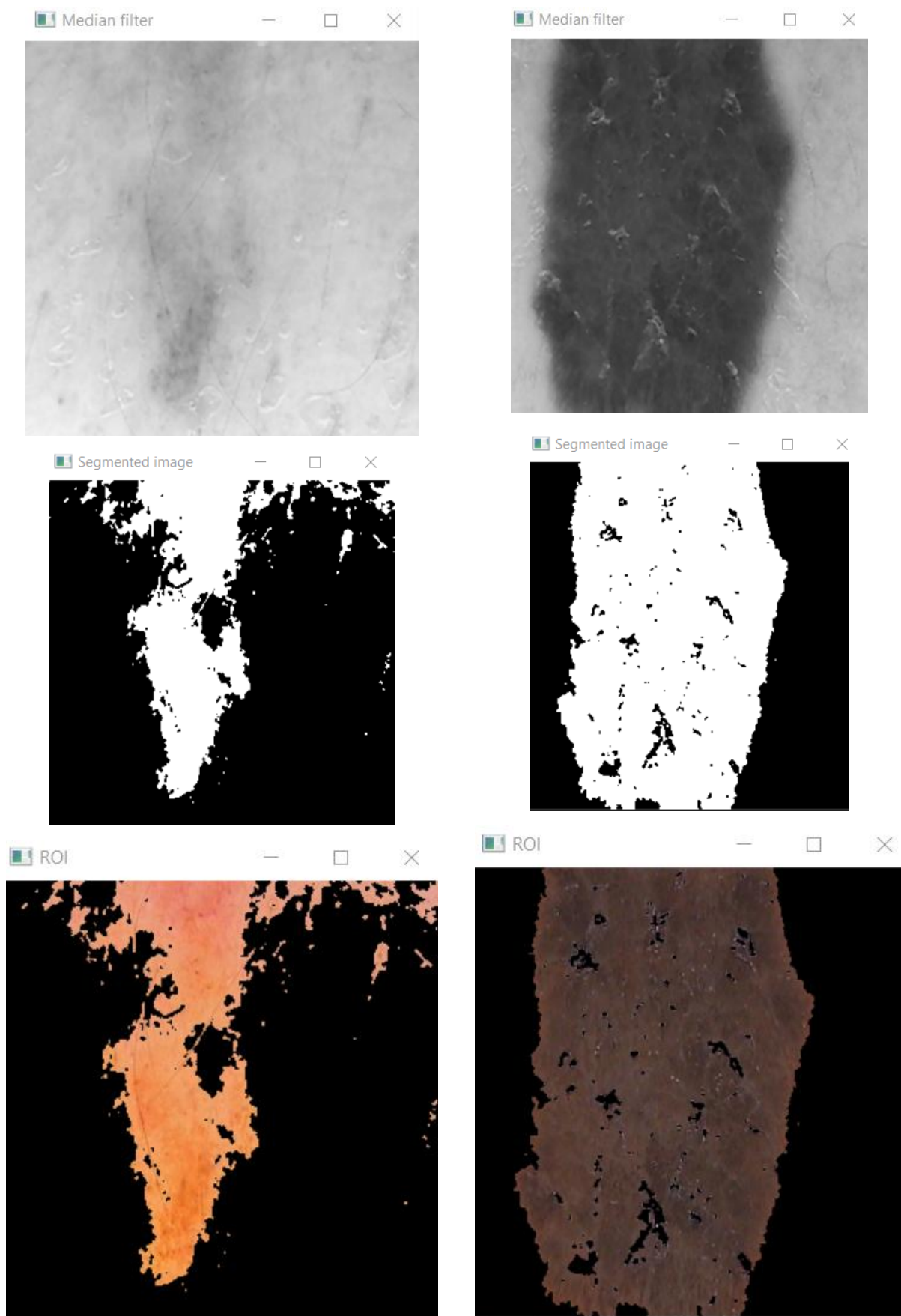
$$F1 - score = \frac{2 * (recall * precision)}{recall + precision} \quad (15)$$

d. Accuracy: It is the proportion of predicted observations (appropriately) to the total observations. It is expressed as in Eq. (16):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (16)$$

The segmentation outcomes are analyzed during the classification time by categorizing the samples as well-segmented or not. The results of the segmented images are appropriately classified as melanoma or non-melanoma. Fig 4 depicts the sample processed image using the compression approach. Here, FP is total false-positive outcomes, FN is total false-negative outcomes, TP is total true-positive, and TN is total true-negative outcomes.





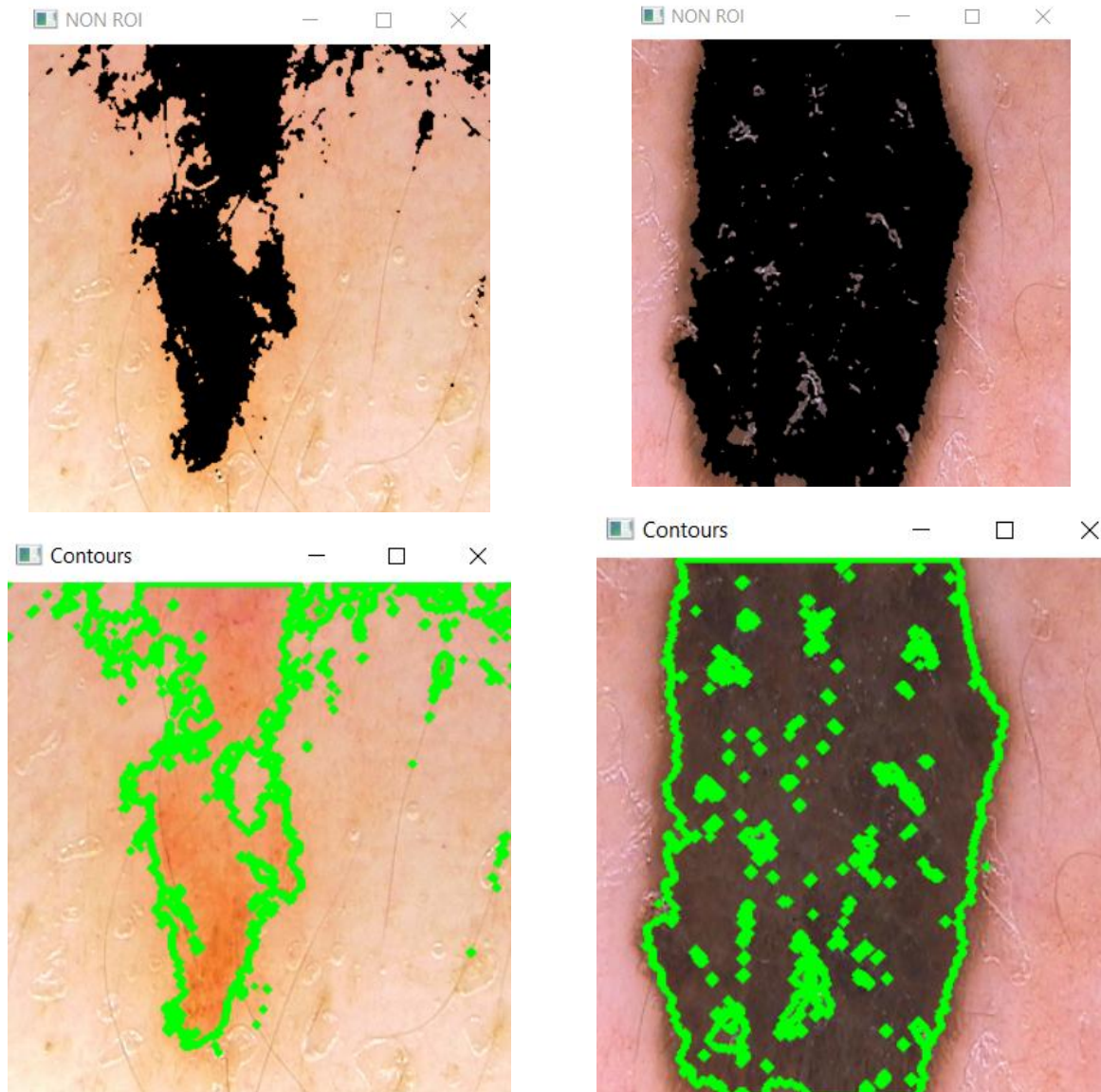


Fig 4 Sample processed image

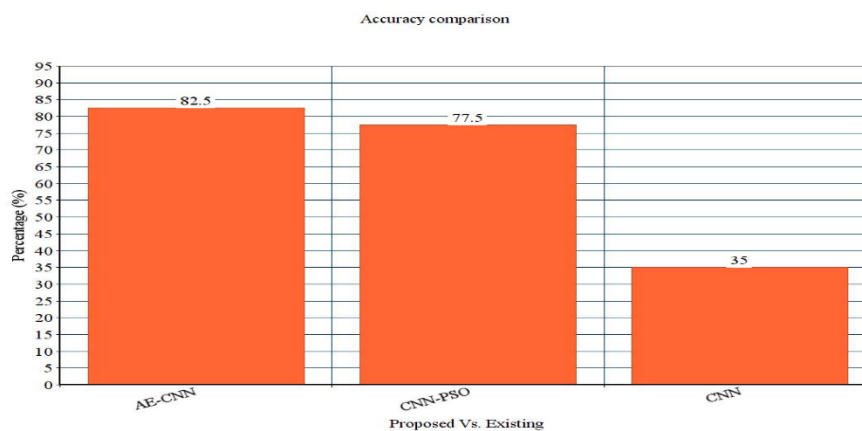


Fig 5 Accuracy comparison

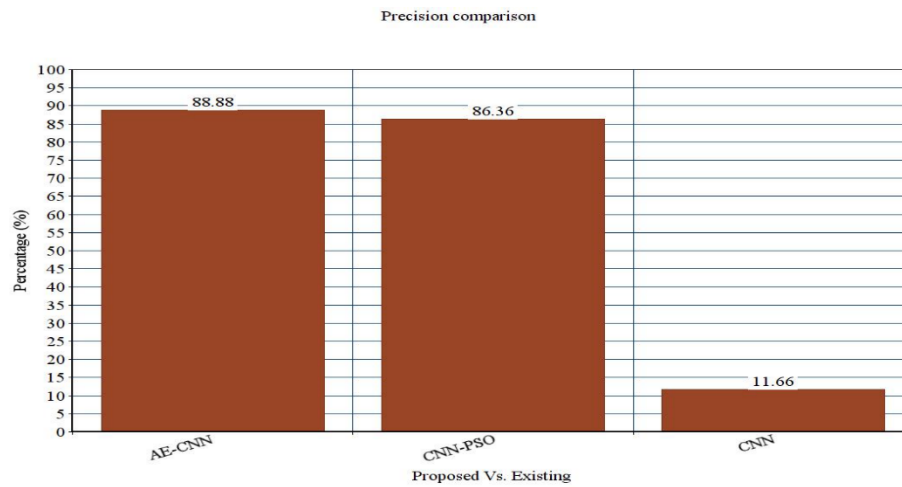


Fig 6 Precision comparison

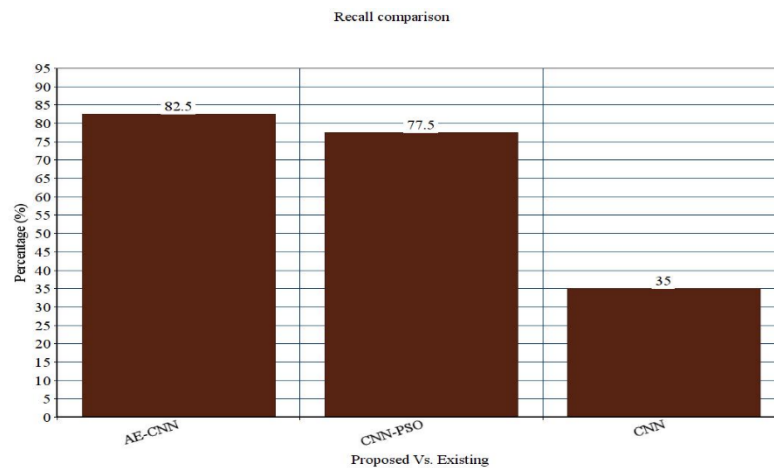


Fig 7 Recall comparison

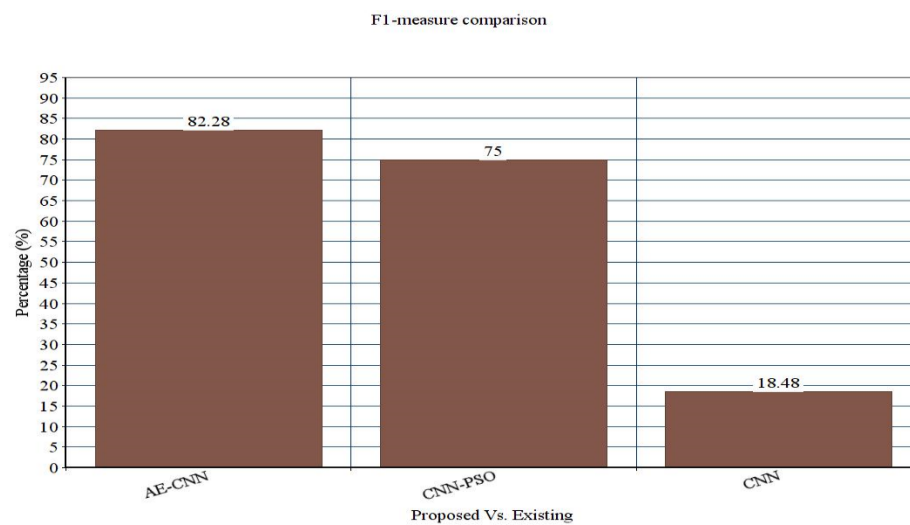


Fig 8 F1-measure comparison

Fig 5 depicts the comparison of the accuracy of proposed vs existing. The comparison is made among AE-CNN, CNN-PSO and standard CNN approaches. The accuracy of the anticipated AE-CNN is 82.5% which is 5% and 47.5% higher than CNN-PSO, and CNN approaches. The precision of the anticipated AE-CNN is 88.88% which is 2.52% and 77.22% higher than CNN-PSO and CNN approaches (See Fig 6). The recall of the anticipated AE-CNN is 82.5% which is 5% and 47.5% higher than CNN-PSO, and CNN approaches as in Fig 7. The F1-measure of the anticipated AE-CNN model is 82.28% which is 7.28% and 63.8% higher than other approaches (See Fig 8). Based on the analysis, it is proven that the anticipated model is superior to different approaches.

5. CONCLUSION

Here, a novel Auto-encoder based Convolutional Neural Networks (AE-CNN) model is exhibited for predicting skin lesions using the PH2 dataset. The anticipated model categorizes various lesions based on the provided pre-processed, segmented and compressed ROI images. The model works effectually even in the case of imbalanced images among the classes. The performance metrics of the anticipated AE-CNN model is accuracy, precision, recall and F-measure and provide the numerical outcome of 82.5%, 88.8%, 82.5% and 82.87%, respectively. This research shows that the model performance increased with the increased amount of images over the classes to deal with the imbalanced image problem among the classes. It is observed that the weights of the provided architectural models are fine-tuned, and the performance metrics are drastically superior than fine-tuning layers. The model predict the unknown images with the auto-encoder model with multi-class representation. The simulation is done with MATLAB 2020a, and the performance is compared with various existing approaches like standard CNN and hybrid CNN-PSO. The primary research constraint is the analysis of compression techniques that consumes enormous time and lead to computational complexity. However, in the future, eliminating these compression techniques and the study with various other approaches provides some promising results.

REFERENCES

- [1] Masood and A. A. Al-Jumaily, "Computer-aided diagnostic support system for skin cancer: A review of techniques and algorithms," *Int. J. Biomed. Imag.*, vol. 2013, Oct. 2013, Art. No. 323268.
- [2] National Cancer Institute. (2018). Skin Cancer (Including Melanoma)- Patient Version. Accessed: Dec. 22, 2018. [Online]. Available: <https://www.cancer.gov/types/skin>
- [3] National Cancer Institute. (2017). Cancer Statistics. Accessed: Dec. 22, 2018. [Online]. Available: <https://www.cancer.gov/aboutcancer/understanding/statistic>
- [4] Ridell and H. Spett, "Training set size for skin cancer classification using Google's inception v3," PhD dissertation, School Comput. Sci. Commun., KTH Roy. Inst. Technol., Stockholm, Sweden, 2017.
- [5] Mehta and B. Shah, "Review on techniques and steps of computer-aided skin cancer diagnosis," *Procedia Comput. Sci.*, vol. 85, pp. 309–316, Jan. 2016.
- [6] Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, "Deep learning for visual understanding: A review," *Neurocomputing*, vol. 187, pp. 27–48, Apr. 2016.
- [7] Pathan, K. G. Prabhu, and P. C. Siddalingaswamy, "Techniques and algorithms for computer-aided diagnosis of pigmented skin lesions— A review," *Biomed. Signal Process. Control*, vol. 39, pp. 237–262, Jan. 2018.
- [8] Kanimozhi and A. Murthi, "Computer-aided melanoma skin cancer detection using artificial neural network classifier," *Singaporean J. Sci. Res. J. Sel. Areas Microelectron.*, vol. 8, no. 2, pp. 35–42, 2016.
- [9] Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.
- [10] Dubai, S. Bhatt, C. Joglekar, and S. Patil, "Skin cancer detection and classification," in *Proc. 6th Int. Conf. Elect. Eng. Inform. (ICEEI)*, Nov. 2017, pp. 1–6.
- [11] Rehman, S. H. Khan, S. M. D. Rizvi, Z. Abbas, and A. Zafar, "Classification of a skin lesion by interference of segmentation and convolution neural network," in *Proc. 2nd Int. Conf. Eng. Innov. (ICEI)*, Jul. 2018, pp. 81–84.
- [12] Masood, A. Al-Jumaily, and K. Anam, "self-supervised learning model for skin cancer diagnosis," in *Proc. 7th Int. IEEE/EMBS Conf. Neural Eng. (NER)*, Apr. 2015, pp. 1012–1015.
- [13] Nasr-Esfahani et al., "Melanoma detection by analysis of clinical images using convolutional neural network," in *Proc. 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2016, pp. 1373–1376.

- [14] Yoshida, M. E. Celebi, G. Schaefer, and H. Iyatomi, "Simple and effective pre-processing for automated melanoma discrimination based on cytological findings," *Proc. IEEE Int. Conf. Big Data*, Dec. 2016, pp. 3439–3442.
- [15] Harangi, "Skin lesion classification with ensembles of deep convolutional neural networks," *J. Biomed. Inform.*, vol. 86, pp. 25–32, Oct. 2018.
- [16] Feigelson, J. D. Powers, M. Kumar, N. M. Carroll, A. Pathy, and D. P. Ritzwoller, "Melanoma incidence, recurrence, and mortality in an integrated healthcare system: A retrospective cohort study," *Cancer Med.*, vol. 8, no. 9, pp. 4508–4516, Aug. 2019.
- [17] Schadendorf, A. C. J. van Akkooi, C. Berking, K. G. Griewank, R. Gutzmer, A. Hauschild, A. Stang, A. Roesch, and S. Ugurel, "Melanoma," *Lancet*, vol. 392, no. 10151, pp. 971–984, Sep. 2018.
- [18] Conforti, R. Giuffrida, R. Vezzoni, F. S. S. Resende, N. Meo, and I. Zalaudek, "Dermoscopy and the experienced clinicians," *Int. J. Dermatology*, vol. 59, no. 1, pp. 16–22, Jan. 2020.
- [19] Lucas, S. Yazar, A. R. Young, M. Norval, F. R. de Gruijl, Y. Takizawa, L. E. Rhodes, C. A. Sinclair, and R. E. Neale, "Human health concerning exposure to solar ultraviolet radiation under changing stratospheric ozone and climate," *Photochem. Photobiology. Sci.*, vol. 18, no. 3, pp. 641–680, Mar. 2019.
- [20] Wang, C.-Y. Lee, Z. Tu, and S. Lazebnik, "Training deeper convolutional networks with deep supervision," 2015, arXiv:1505.02496. [Online]. Available: <http://arxiv.org/abs/1505.02496>
- [21] Milton, "Automated skin lesion classification using an ensemble of deep neural networks in ISIC 2018: Skin lesion analysis towards melanoma detection challenge," 2019, arXiv:1901.10802. [Online]. Available: <http://arxiv.org/abs/1901.10802>
- [22] Rashid, M. A. Tanveer, and H. Aqeel Khan, "Skin lesion classification using GAN based data augmentation," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 916–919
- [23] Baur, S. Albarqouni, and N. Navab, "MelanoGANs: High-resolution skin lesion synthesis with GANs," 2018, arXiv:1804.04338. [Online]. Available: <http://arxiv.org/abs/1804.04338>
- [24] Ahmad, M. Usama, C.-M. Huang, K. Hwang, M. S. Hossain, and G. Muhammad, "Discriminative feature learning for skin disease classification using deep convolutional neural network," *IEEE Access*, vol. 8, pp. 39025–39033, 2020.
- [25] Carcagni, M. Leo, A. Cuna, P. L. Mazzeo, P. Spagnolo, G. Celeste, and C. Distance, "Classification of skin lesions by combining multilevel learnings in a DenseNet architecture," in *Proc. Int. Conf. Image Anal. Process. Trento, Italy: Springer*, 2019, pp. 335–344.
- [26] Bissoto, E. Valle, and S. Avila, "The six fronts of the generative adversarial networks," 2019, arXiv:1910.13076. [Online]. Available: <http://arxiv.org/abs/1910.13076>
- [27] Zhang, C. Lu, X. Li, H.-J. Kim, and J. Wang, "A fully convolutional network based on DenseNet for remote sensing scene classification," *Math. Biosci. Eng.*, vol. 16, no. 5, pp. 3345–3367, 2019.
- [28] Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen, "Improved techniques for training GANs," in *Proc. Adv. Neural Inf. Process. Syst.*, 2016, pp. 2234–2242.
- [29] Denton, S. Chintala, and R. Fergus, "Deep generative image models using a Laplacian pyramid of adversarial networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 1486–1494.
- [30] Wang, M. Jiang, C. Qian, S. Yang, C. Li, H. Zhang, X. Wang, and X. Tang, "Residual attention network for image classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 3156–3164.
- [30] Hussain, T., Urlamma, D., Vericharla, R. *et al.* Augmenting traffic flow efficiency using multi-agent systems (MAS). *Int. j. inf. tecnol.* (2025). <https://doi.org/10.1007/s41870-025-02452-w>
- [31] Wang, M. Jiang, C. Qian, S. Yang, C. Li, H. Zhang, X. Wang, and X. Tang, "Residual attention network for image classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 3156–3164.
- [32] Chellam, V. V., Veeraiah, V., Khanna, A., Sheikh, T. H., Pramanik, S., & Dhabliya, D. (2023). A machine vision-based approach for tuberculosis identification in chest X-rays images of patients. In *International Conference on Innovative Computing and Communication* (pp. 23–32). Springer Nature Singapore. https://doi.org/10.1007/978-981-19-3840-8_3.
- [33] Khanna, A., Selvaraj, P., Gupta, D., Sheikh, T. H., Pareek, P. K., & Shankar, V. (2021). Internet of things and deep learning enabled healthcare disease diagnosis using biomedical electrocardiogram signals. *Expert Systems*, 39(3), e12864. <https://doi.org/10.1111/exsy.12864>.

- [34] Dang, N., Saraf, V., Khanna, A., Gupta, D., & Sheikh, T. H. (2020). Malaria detection on Giemsa-stained blood smears using deep learning and feature extraction. In A. Khanna, D. Gupta, S. Bhattacharyya, V. Snasel, J. Platos, & A. Hassanien (Eds.), *International Conference on Innovative Computing and Communications* (Vol. 1087, pp. 829–841). Springer. https://doi.org/10.1007/978-981-15-1286-5_70
- [35] Mohana, T. V., Buradkar, M. U., Alaskar, K., Sheikh, T. H., & Kumbhkar, M. (2023). Deep learning-based regulation of healthcare efficiency and medical services. In *AI and IoT-based Intelligent Health Care & Sanitation* (pp. 176–190).