Journal of Information Systems Engineering and Management
2025, 10(9s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Developing Improved Melanoma Detection Strategies Using Hybrid CNN and Autoencoder Models and Detailed Data Analysis

Vinay S. Nalawade[1], Dr. Shailesh Kumar[2], Dr.Shrinivas T. Shirkande[3]
[1]*Research Scholar, Department of Computer Engineering, Shri Jagdishprasad Jhabarmal Tibrewala University, Jhunjhunu, Rajasthan, India. vinaynalawade2007@gmail.com*
[2]*Associate Professor, Department of Computer Engineering, Shri Jagdishprasad Jhabarmal Tibrewala University, Jhunjhunu, Rajasthan, India. shaileshkumar1610@gmail.com*
[3]*Principal & Associate Professor, S B Patil College of Engineering, Indapur, Pune, Maharashtra, India. shri.shirkande8@gmail.com*

ARTICLE INFO

Received: 25 Oct 2024
Revised: 16 Dec 2024
Accepted: 28 Dec 2024

ABSTRACT

Abstract: Melanoma is a very dangerous type of skin cancer that needs new ways to be found so that it can be treated quickly. The goal of this study is to create better ways to find melanoma by combining mixed Convolutional Neural Networks (CNN) and Autoencoder models with in-depth data analysis. This will make melanoma identification more accurate and reliable. The chosen method takes advantage of deep learning techniques that are designed to work well with image-based classification tasks. This study used a dataset from the Skin Cancer MNIST: HAM10000 for this work. It has dermatoscopic pictures that need to be carefully preprocessed. To make the models work, this meant shrinking, normalising pixel values, one-hot encoding, and rearranging the pictures. Data enrichment methods like rotations, shifts, and flips were used to make the model even better by training it against a wider range of data representations. In the methods part, we describe how we set up our hybrid model, which blends the feature extraction skills of CNNs with the dimensionality reduction skills of Autoencoders. With settings like an epoch count of 25 and a batch size of 128, this setup went through a lot of training using the Adam optimiser and Categorical Cross Entropy as the loss function. Our results show that the mixed model works, as it achieved a trial accuracy of 73.19% with a loss of 1.079. There are more details in the classification report and confusion matrices that show how well the model works with different types of skin lesions. For example, it is much better at finding arterial lesions, with a precision of 66% and a recall of 100%. Comparing the mixed model to normal CNN models shows that it is more accurate and precise; this proves that it is better at diagnosing problems. This study shows that mixing CNNs and Autoencoders could be useful for making accurate screening tools for cancer. This could lead to future improvements in medical imaging technology.

Keywords: Melanoma Detection, Convolutional Neural Networks, Autoencoders, Skin Cancer MNIST, Image Data Preprocessing, Hybrid Models, Data Augmentation

I. INTRODUCTION

As a form of skin most cancers referred to as malignant melanoma, it's far known to unfold quickly and can be fatal if no longer found early. The range of instances of melanoma has been slowly rising around the world, so it is critical to discover better and extra accurate approaches to locate it. The accuracy of conventional checking out methods, which rely on dermatoscopic analysis with the aid of scientific specialists, can vary loads based totally on how experienced and professional the viewer is. Because of this, there may be a robust case for using greater dependable, goal, and scalable technology in the diagnosis manner, like artificial intelligence (AI). Latest progress in AI, particularly in deep mastering, has shown that it can be used to enhance prognosis techniques in many areas of health. Convolutional Neural Networks (CNNs) stand out because they're top at photo reputation obligations, which mean they may be used directly to analyse dermatoscopic images to discover cancer. CNNs are very good at working with complex photograph information and finding developments that someone might pass over. but CNNs

might not be able to completely display the intricacies and complex variations in snap shots of pores and skin lesions, which can be very one-of-a-kind in shape, coloration, and structure [1].

To address those troubles, this examines looks into making mixed models that use the exceptional elements of both CNNs and Autoencoders. Autoencoders are a sort of out of control mastering neural community that is used to locate traits in records and decrease the wide variety of dimensions it has. Adding Autoencoders and CNNs to the suggested hybrid version is supposed to make it less difficult to extract functions and make melanoma identification more dependable. This method now not only hurries up the mastering method, however it also improves the model's potential to generalise from complicated, high-dimensional data [2]. This makes predictions more correct and constant. Including strategies for thorough facts evaluation to the mixture version makes it even better. After quite a few editing and including to photo records, the version is educated on a various set of information that represents a wide variety of real-life differences in how melanoma looks. This thorough coping with of statistics makes positive that the model now not simplest learns from a huge variety of cases, however also does not get too appropriate at what it does, which could be very crucial for maintaining overall performance excessive while it's used in medical settings. The main goal of this study is to thoroughly compare how well the combination model works with standard and stand-alone deep learning methods. The study aims to build a strong case for AI's usefulness in finding melanoma by carefully looking at model outputs such as accuracy measures, loss functions, and precision-recall values across different types of skin tumours. Additionally, studying this mixed model could lead to its use in other areas of medical imaging, which could completely change how diagnostic processes are thought of and carried out in the healthcare business.

II. RELATED WORK

Mostly because imaging technology has gotten better and we know more about how the disease looks on the skin. In the past, melanoma was found through physical examination and the knowledge of doctors who followed the ABCD (Asymmetry, Border, Colour, and Diameter) rules. However, diagnosing melanoma by hand can be hard because the changes seen in early-stage tumours are often modest and subjective. Because of this, there is more and more focus on creating automatic systems that use computer science to make melanoma diagnosis more accurate and reliable [3]. A lot of work has been done in the fields of image processing and machine learning, especially with Convolutional Neural Networks (CNNs) to look at pictures of the skin. CNNs are well known for his or her potential to examine hierarchical representations, and their high accuracy in sample recognition responsibilities has made them very beneficial in clinical image analysis [4]. Studies like the ones by means of Esteva et al. have proven that deep learning models can efficiently spot pores and skin most cancers at the extent of a health practitioner. This shows that those fashions might be useful in medical settings [5]. Autoencoders and CNNs running collectively to find most cancers is also a brand new way of doing things in this area. Autoencoders are proper at doing away with noise and reducing the variety of dimensions in facts. They also can enhance the feature extraction method, making it simpler to get greater useful and minor features from complicated image records [6]. Numerous researches have checked out this mixed method and observed that it improves model performance thanks to a bigger set of functions which could higher seize the info of cancer images [7]. Using huge dermatoscopic image files like the HAM10000, which display a spread of pores and skin diseases, is every other vital a part of present day take a look at. This dataset has been very beneficial in constructing strong models that could paintings well with a huge range of pores and skin diseases and lesions [8]. Adding strategies like rotating, scaling, and flipping to these datasets makes version training even well with the aid of giving the algorithms a much broader range of lesion seems. This makes the version higher at locating melanoma in a diffusion of conditions [9].

At the side of the advent of greater advanced popularity fashions, there has been lots of interest in studies that examine how well AI fashions work in opposition to more traditional prognosis techniques. New studies show that the use of combined fashions, which combine the pleasant parts of numerous device getting to know designs, can greatly cut down on false hits and improve the accuracy of finding cancer [10]. This is especially important for reducing down on checks that aren't wished and making sure those patients with proper consequences get treatment as soon as feasible. There have additionally been new tendencies within the discipline, which include incorporated testing systems that use AI to analyse statistics and dermatoscopic gadgets to gather facts in real time. The goal of this equipment is to present doctor’s quick feedback and clinical assist during clinical assessments, as a way to assist them, make better decisions [11]. This type of merging not only quickens the checking out manner, but it also makes certain that choices are based totally on data, rather than opinions, that could appear during a watch

examination. Other than that, the social problems and problems of using AI in scientific situations have additionally been pointed out lots. in the literature [12], worries about information safety, the openness of algorithms choices, and the want to constantly watch over AI structures to keep away from biases were introduced up. To ensure that AI technologies are used in healthcare in an moral manner, strict rules and ongoing communication among parties are needed to solve those problems [13].

Table 1: Related work summary in malignant melanoma detection

Approach	Method	Key Finding	Types of Disease	Limitation	Scope
Traditional Diagnosis [14]	Clinical examination following ABCD guidelines	Relies heavily on dermatologist's experience	Melanoma	Subjective and variable accuracy	Limited to visual examination skills
CNN-based Models [15]	Deep learning with CNNs for image analysis	Achieves dermatologist-level accuracy in skin cancer	Skin cancers, including melanoma	Requires large labeled datasets	Suitable for large-scale screening applications
Hybrid Models [16]	CNN combined with Autoencoders	Enhanced feature extraction and improved accuracy	Skin cancers, including melanoma	Complex model architectures requiring extensive computational resources	Advanced research and clinical settings
Dataset Utilization [17]	Training on HAM10000 dataset	Robust models that generalize well	Various skin lesions	May not fully represent all demographic groups	Broad applicability in diverse clinical environments
Data Augmentation [18]	Techniques like rotation, scaling, and flipping	Improved model training and performance	Skin cancers, including melanoma	Augmentation parameters must be carefully chosen to avoid unrealistic images	Enhancing model robustness against varied lesion appearances
Comparative Studies [19]	Analysis of AI models versus traditional methods	Hybrid models reduce false positives and increase precision	Melanoma	Still under investigation in real-world clinical trials	Informing future diagnostic protocols
Real-time Systems [20]	AI-based analysis integrated with dermatoscopic devices	Immediate feedback and diagnostic support during exams	Skin cancers, including melanoma	Integration complexity and high system costs	Real-time diagnostic support in clinical settings

III. SKIN CANCER MNIST: HAM10000 DATASET DESCRIPTION

The skin most cancers MNIST: HAM10000 dataset is an important device for dermatology gadget studying look at, especially for teaching computers to identify exceptional types of pores and skin cancer, like melanoma. The dataset consists of 10,000 dermatoscopic snap shots from a diffusion of agencies that were accumulated and labelled by means of medical doctors with loads of revel in. This series has an extensive variety of skin illnesses organised into seven diagnostic organizations. It is an entire aid for teaching and trying out device mastering models. HAM10000 is not best a widespread for a way well algorithms work, however it is also a way to improve the accuracy of dermatology diagnoses. The snap shots have already been dealt with to be the equal size, which makes it clean to

feature them to one of a kind gadget mastering tactics. Researchers and builders can make their models more dependable and useful by way of using strategies like convolutional neural networks (CNNs) and facts enhancement techniques like flipping and rotating statistics.

This collection gives deep learning models a lot of data they need to learn about the details and subtle aspects of different tumours. This helps dermatologists solve important problems, like finding melanoma and other skin cancers early. So, HAM10000 plays a big role in encouraging the creation of automated diagnosis tools that might help doctors make faster and more accurate diagnoses. The cropped and preprocessed dermatoscopic picture used in study to find melanoma is shown in Figure 1. It shows how important it is to change images in order to prepare data that machine learning models can use to get better at finding skin tumours.

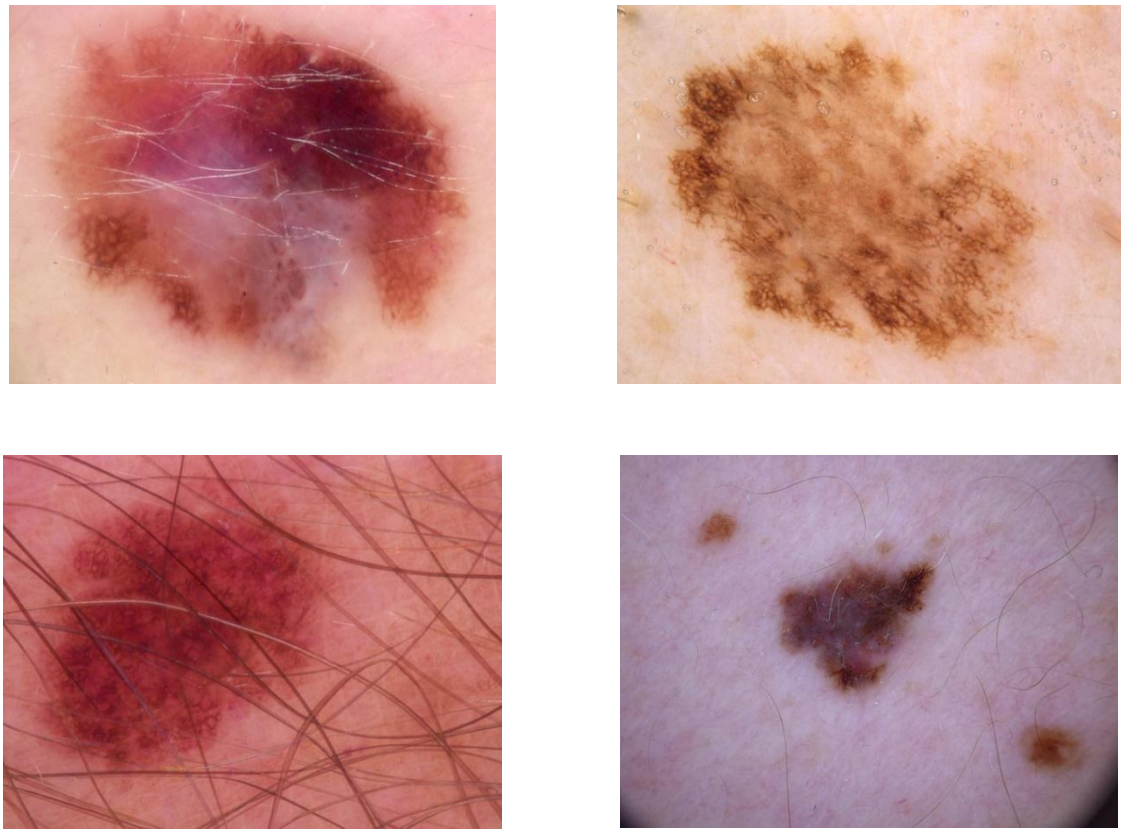

Figure 1: Sample images data input form the dataset

IV. RESEARCH APPROACH

Using mixed CNN and Autoencoder models to create better melanoma detection strategies requires a thorough method that blends advanced image processing techniques with deep learning to improve the accuracy of diagnostics, proposed system model shown in figure 2. This plan is broken down into several main steps: The process starts with getting pictures of the skin from the Skin Cancer MNIST: HAM10000 collection. Before it can be used for model training, each picture goes through a set of steps called "preprocessing." To do this, pictures must be resized to a uniform size, pixel values must be normalised to a [0,1] range, and category data must be turned into binary vectors using one-hot encoding. More than that, the pictures are changed into 3D groups so that they can be used by CNNs.

The most important part of our method is creating a mixed model that combines CNNs and Autoencoder schemes. The CNN part is in charge of feature extraction, which means it finds key patterns in the picture data that show the presence of melanoma. It is known that the Autoencoder can reduce the number of dimensions and noise. It improves these traits to help the model learn better. After setting up the mixed model design, training is done with 80% of the data being used for training and 20% for testing. A batch size of 128 is used to train the model over 25 iterations. The Adam optimiser works well with sparse gradients and noisy data, and category cross-entropy is used as the loss function because it works well for jobs that need to classify things into more than one group. The model's success is checked on the test set after it has been trained. To figure out how well the model works with different

types of skin tumours, metrics like accuracy, loss, precision, recall, and F1-score are calculated. These measures help find any flaws or weak spots in the model, which leads to more improvements. A test is run to see how well the mix model works compared to the old CNN models. By looking at how precision goes up and loss goes down, this study shows the benefits of using Autoencoders along with CNNs to find cancer. This organised method makes sure that all mixed deep learning techniques are thoroughly explored in order to improve the diagnosis of melanoma. These techniques use both automatic picture processing and careful data analysis to achieve high diagnostic accuracy.

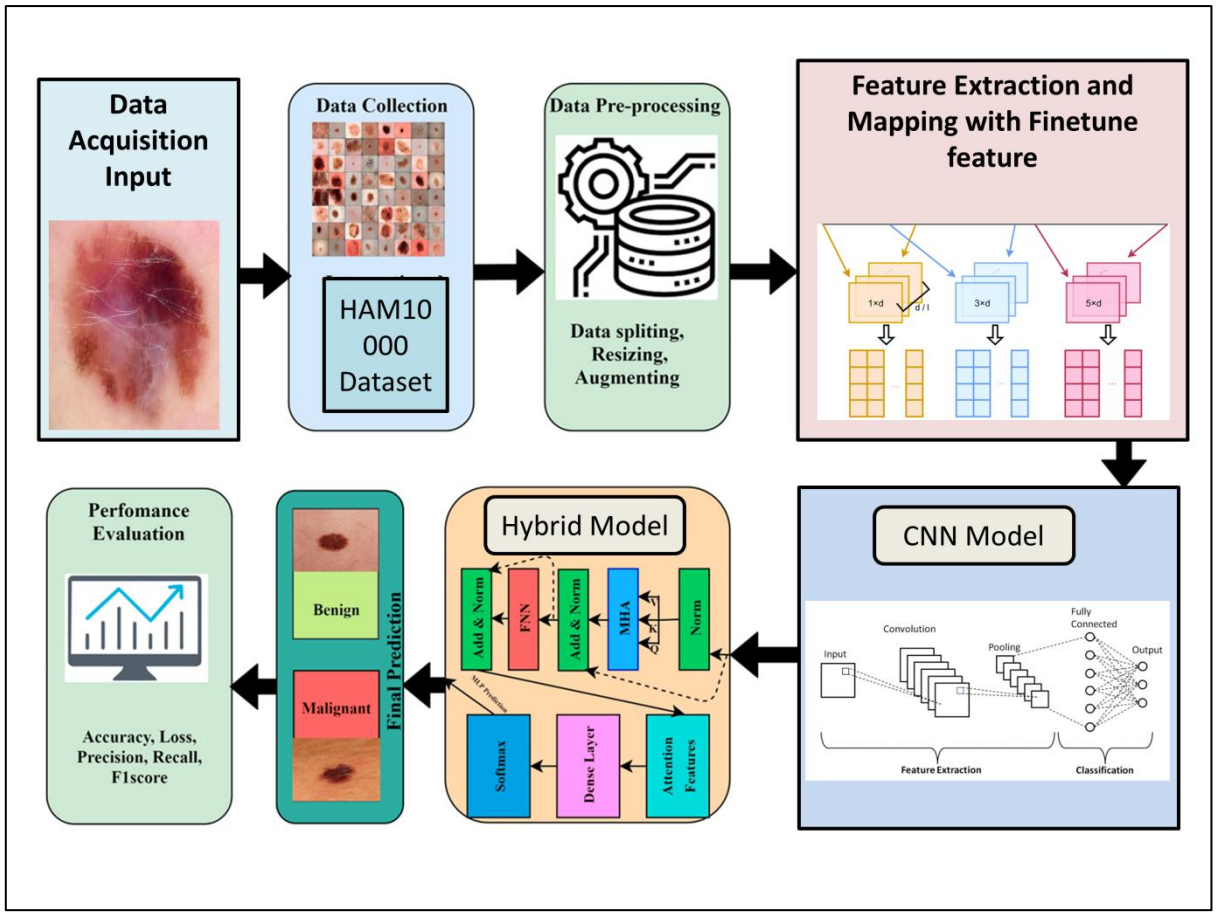

Figure 2: Representation of proposed system architecture

A. Data Preprocessing

During the data preparation part of creating mixed CNN and Autoencoder models for melanoma diagnosis, several important steps are taken to get the pictures ready for training and analysis by the models:

a) Image Resize: The first thing that needs to be done is to change the sizes of the pictures to normal sizes. This is very important because it makes sure that all the inputs to the neural network are the same, which helps it learn. Image sizes are usually changed to a set size that works with CNN's design, like 224x224 pixels, resizing image illustrate in figure 3. This resizing makes the input feature measurements the same for all images in the dataset and reduces the amount of work that needs to be done on the computer.

- Resize an image from its original dimensions H x W to a standard dimension H' x W'.

$$I' = resize(I, H', W')$$

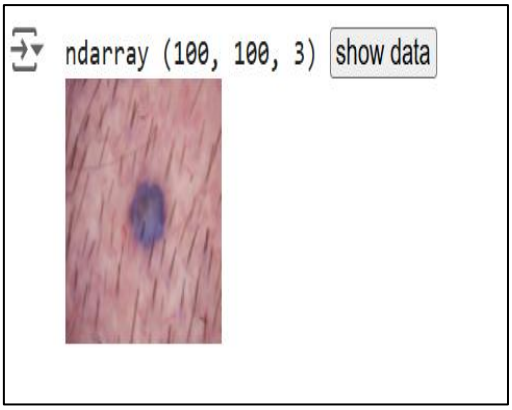

Figure 3: Representation of input image after resize

b) Normalize Pixel Values: It is necessary to normalise pixel values so that the picture pixels are all the same size and fall in a standard range, usually [0,1]. This normalisation is very important because it makes sure that the gradient descent method works well across a wide range of data, which speeds up the model's convergence during training. It also helps get rid of differences that come from changes in lighting and brightness while the picture is being taken.

- Normalize the pixel values of the image to range [0,1].

$$I_{norm} = \frac{I' - \min(I')}{\max(I') - \min(I')}$$

c) One-hot Encoding: This type of encoding is used for category target variables, like the type of skin disease. This process changes categorical integer names into a binary matrix form that neural networks can use for better classification. One-hot encoding makes the network's output layer easier to understand. Each neurone can show how likely it is that the input is a certain class.

- Transform a categorical label y among N classes into a binary vector v of length N.

$$v_i = \begin{cases} 1 & if\ i = y \\ 0 & otherwise \end{cases}$$

d) Reshaping images in 3D: Once photos have been resized and normalised, they need to be remade into a three-dimensional format so that CNNs can use them. To do this, each 2D picture has to be turned into a 3D collection, with RGB colour bands making up the third dimension. This rearrangement is very important for the CNN to be able to do convolution operations on the picture data correctly, which lets it pull out features from both the spatial and colour patterns.

- Convert a 2D image (if in grayscale) to a 3D format for RGB channels by stacking.

$$I\_3D = stack(I\_norm, I\_norm, I\_norm)$$

These steps before processing are essential for turning raw dermatoscopic pictures into a format that is best for learning in deep neural networks. This makes melanoma detection models work better and be more accurate.

B. Data augmentation (for hybrid model)

When making mixed models to find cancer, adding more data is a key part of making the models more stable and stopping them from fitting too well. To do this, the training dataset is intentionally made bigger by using different changes to make it look like different viewing conditions of skin tumours. The following methods are often used to add to data:

a) Rotation Range: This option lets pictures be turned freely within a certain degree range while they are being trained. This helps the model learn from seeing the lesion in different ways. To find melanoma, we use a 20-degree range of movement.

$$Irot = rotate(I, \theta) where \theta \in [-20 \circ, 20 \circ]$$

b) Width Shift Range: This changes the image's horizontal position by a certain percentage of its width. If the range for width shift is 0.2, the picture can be moved up to 20% to the left or right.

$$Iw\_shift = translate(I, \delta x) where \delta x \in [-0.2 \times W, 0.2 \times W]$$

c) Height Shift Range: In the same way that width moving can move a picture horizontally, the height shift range can move it vertically. If you set it to 0.2, you can move the picture up or down by up to 20% of its height.

$$Ih\_shift = translate(I, \delta y) where \delta y \in [-0.2 \times H, 0.2 \times H]$$

d) Horizontal Flip: This add-on flips the picture horizontally, which is the same thing as reflecting it. This is especially helpful for analysing skin lesions because it lets the model see trends no matter how they are laid out horizontally.

$$Iflip = flip\_horizontal(I)$$

e) Fill Mode: We use the "nearest" fill mode when the changes we just talked about create new pixel values that need to be filled. In "Nearest" mode, these pixels are filled with the values of the pixels that are closest to them, keeping the local colour and pattern ranges.

$$Ifilled = fill(Imodified, 'nearest')$$

C. Fine-tune Parameters

In the fine-tuning part of the hybrid model for finding melanoma, certain factors are carefully changed to get the best training results. The model is trained over 25 epochs, which lets it learn from the information over and over again. For a good mix between how quickly the computer works and how finely the model is updated, a batch number of 128 is used. Categorical Cross Entropy was chosen as the loss function because it works well with many classes. The Adam optimiser was picked because it can change its learning rate. Lastly, ReLU is used for the hidden layers and Softmax is used for the output layer. These two functions improve non-linearity and multi-class classification, respectively.

Table 2: Details of fine-tuned parameters

Parameters	Value
Epoch	25
Batch Size	128
Loss Function	Categorical Cross Entropy
Optimizer	Adam
Activation Function	Relu / Softmax

D. Applying the model

1. CNN model

To make a good CNN model for finding melanoma, the framework is set up to carefully take out and look at traits from dermatoscopic photos. The model starts with an input layer that can handle standard pictures, which are usually trimmed to 224x224 pixels and have three colour bands to make sure that all samples are the same. After the input layer, the model has several convolutional layers, and each one has a set of learnable filters. These filters use convolution to pick out a variety of features, from simple lines and colours in the top layers to more complicated patterns like lesion borders in the lower layers. The ReLU activation function is usually used in each convolutional layer to add nonlinearity, which is important for learning complex features. Once the convolutional layers are done, the pooling layers are used to make the feature maps smaller in space, as shown in figure 4. This decrease helps cut down on the number of factors and the cost of computing, which makes the model more useful and stops it from overfitting. A lot of people use max pooling because it downscales feature maps well while keeping the most important features. Following some convolutional layers, extra parts like batch normalisation could be added. This step makes the activations more consistent, which helps keep the mean and range of the layers' inputs. This speeds up the training process and keeps the learning route stable. One or more fully linked layers take in the learnt

features from all over the picture as the architecture moves towards classification. These layers are very important for figuring out what kind of skin disease it is. They lead to an output layer that uses a softmax activation function. This function gives a chance distribution over the classes, which makes it easy to see whether the picture is safe or cancerous.

Step wise CNN architecture

1. Input Layer

$$I\_input = image(H, W, C)$$

Where H and W are the height and width of the image, and C is the number of channels (typically 3 for RGB images).

2. Convolutional Layers

$$I\_conv = f(I\_prev * K + b)$$

Where * denotes the convolution operation, K represents the kernel, b is the bias, and f is the activation function (ReLU).

3. Pooling Layers (Max Pooling)

$$I\_pool = max(I\_conv(a, b))$$

Where (a, b) are the dimensions of the pooling window.

4. Batch Normalization

$$I_{bn} = \gamma\left(\frac{I_{prev} - \mu}{sqrt(\sigma^2 + \varepsilon)}\right) + \beta$$

Where μ and σ^2 are the mean and variance of I_prev, γ and β are learnable parameters, and ε is a small constant for numerical stability.

5. Fully Connected Layers

$$I\_fc = f(WI\_prev + b)$$

Where W and b are the weight matrix and bias vector, respectively, and f is the activation function (ReLU).

6. Output Layer (Softmax)

$$\frac{Softmax(z_i) = \exp(z_i)}{sum(\exp(z_j))}$$

Where z_i are the inputs to the output layer from the last fully connected layer, and the output is a probability distribution over classes.

7. Loss Function (Categorical Cross Entropy)

$$L = -sum(y\_o, c * log(p\_o, c))$$

Where M is the number of classes, y_o,c is a binary indicator (0 or 1) if class label c is the correct classification for observation o, and p_o,c is the predicted probability observation o is of class c.

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 100, 100, 16)	1,216
conv2d_1 (Conv2D)	(None, 100, 100, 32)	4,640
max_pooling2d (MaxPooling2D)	(None, 50, 50, 32)	0
batch_normalization (BatchNormalization)	(None, 50, 50, 32)	128
conv2d_2 (Conv2D)	(None, 50, 50, 64)	18,496
max_pooling2d_1 (MaxPooling2D)	(None, 25, 25, 64)	0
dropout (Dropout)	(None, 25, 25, 64)	0
conv2d_3 (Conv2D)	(None, 25, 25, 64)	36,928
max_pooling2d_2 (MaxPooling2D)	(None, 13, 13, 64)	0
batch_normalization_1 (BatchNormalization)	(None, 13, 13, 64)	256
flatten (Flatten)	(None, 10816)	0
dense (Dense)	(None, 256)	2,769,152
batch_normalization_2 (BatchNormalization)	(None, 256)	1,024
dropout_1 (Dropout)	(None, 256)	0
dense_1 (Dense)	(None, 7)	1,799

Total params: 2,833,639 (10.81 MB)
Trainable params: 2,832,935 (10.81 MB)
Non-trainable params: 704 (2.75 KB)

Figure 4: CNN Architecture model

2. Hybrid Model (CNN + Autoencoder)

In its encoding layer, the Autoencoder shrinks the CNN's feature output into a smaller form. In its decoding layer, it then rebuilds the output so that it matches the original input. This process helps to focus and narrow down important information, which makes it easier for the model to find the traits that help tell the difference between normal and cancerous tumours. The combination model takes advantage of the best features of both designs by combining CNNs and Autoencoders: it can retrieve deep features and compress data efficiently, as illustrate in figure 5. This combination improves the model's sensitivity and specificity, making it very good at finding melanoma early on. This leads to better clinical results by ensuring accurate and fast diagnosis.

Model: "functional_34"

Layer (type)	Output Shape	Param #
input_layer_3 (InputLayer)	(None, 100, 100, 3)	0
conv2d_7 (Conv2D)	(None, 100, 100, 32)	896
batch_normalization_3 (BatchNormalization)	(None, 100, 100, 32)	128
max_pooling2d_6 (MaxPooling2D)	(None, 50, 50, 32)	0
dropout_3 (Dropout)	(None, 50, 50, 32)	0
conv2d_8 (Conv2D)	(None, 50, 50, 64)	18,496
batch_normalization_4 (BatchNormalization)	(None, 50, 50, 64)	256
max_pooling2d_7 (MaxPooling2D)	(None, 25, 25, 64)	0
dropout_4 (Dropout)	(None, 25, 25, 64)	0
conv2d_9 (Conv2D)	(None, 25, 25, 128)	73,856
batch_normalization_5 (BatchNormalization)	(None, 25, 25, 128)	512
max_pooling2d_8 (MaxPooling2D)	(None, 13, 13, 128)	0
flatten_2 (Flatten)	(None, 21632)	0
dense_8 (Dense)	(None, 128)	2,769,024
dense_9 (Dense)	(None, 64)	8,256
dense_10 (Dense)	(None, 128)	8,320
dense_11 (Dense)	(None, 30000)	3,870,000
reshape_1 (Reshape)	(None, 100, 100, 3)	0

Total params: 6,749,744 (25.75 MB)
Trainable params: 6,749,296 (25.75 MB)
Non-trainable params: 448 (1.75 KB)

Figure 5: Summary of Hybrid Model

Step wise model architecture process

1. CNN Feature Extraction

- Convolutional Layer:

$$I\_conv = f(I\_prev * K + b)$$

Where * denotes the convolution operation, K represents the kernel, b is the bias, and f is the activation function (ReLU).

- Pooling Layer:

$$I\_pool = max(I\_conv(a, b))$$

Reduces the spatial dimensions to decrease parameter count and computational complexity.

2. Autoencoder for Feature Refinement

- Encoder:

$$z = f(We * I\_pool + be)$$

Where We and be are the weights and biases of the encoder, compressing the data into a lower-dimensional latent space z.

- Decoder:

$$I\_hat = f(Wd * z + bd)$$

Where Wd and bd are the weights and biases of the decoder, attempting to reconstruct the original input from the encoded representation z.

3. Loss Function

- Reconstruction Loss (for Autoencoder):

$$L_{recon} = \left(\frac{1}{n}\right) * sum((I - I_{hat})^2)$$

Where n is the number of samples, I is the original input, and I_hat is the reconstructed output.

- Classification Loss (for CNN):

$$L\_class = -sum(y\_oc * log(p\_oc))$$

Where M is the number of classes, y_oc is a binary indicator if class label c is the correct classification for observation o, and p_oc is the predicted probability.

4. Combined Loss

$$L = alpha * L\_recon + beta * L\_class$$

Where alpha and beta are weights that balance the importance of each component of the loss.

V. RESULT AND DISCUSSION

A. Results of CNN Model

The CNN Model in Figure 6 This picture shows how well a CNN model did on a dataset during the training and testing stages. During the training process, the model got an accuracy of 0.7280 with a loss of 1.0609 per batch, which took about 293ms per step. The model had an accuracy of 0.7370 and a loss of 1.0659 over 26 runs, taking 327ms per step, during the testing phase. It also gives validation and test results: a validation accuracy of 0.735661 with a loss of 1.064926 and a test accuracy of 0.731902 with a loss of 1.079062, which shows that it works the same way on all data sets.

```
 63/63 ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 18s 293ms/step - accuracy:
0.7280 - loss: 1.0609

26/26 ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 8s 327ms/step - accuracy:
0.7370 - loss: 1.0659
Validation: accuracy = 0.735661  ;  loss_v = 1.064926
      Test: accuracy = 0.731902  ;  loss = 1.079062
```

Figure 6: Output 1: for CNN Model

Output 2 for the CNN model is shown in Figure 7. This picture shows a classification report that sums up the CNN model's accuracy, recall, and F1-score for various types of skin tumours. It gives precise performance measures for seven classes. Class 5 (vascular injuries) has the best results, with 0.85 accuracy and 0.90 memory, giving it an F1-score of 0.87. The model is 0.73 percent accurate on the dataset, with a weighted average precision of 0.73 and a recall of 0.71. This shows that it does well across a wide range of classes. This study points out places where model performance could be improved, especially for classes that got lower scores.

63/63 ——————————————————————— 19s 306ms/step
Classification Report:

	precision	recall	f1-score	support
0	0.40	0.29	0.34	59
1	0.52	0.40	0.45	108
2	0.42	0.67	0.52	230
3	0.22	0.09	0.13	22
4	0.59	0.22	0.32	244
5	0.85	0.90	0.87	1321
6	0.64	0.47	0.55	19
accuracy			0.73	2003
macro avg	0.52	0.43	0.45	2003
weighted avg	0.73	0.73	0.71	2003

Figure 7: Output 2: Classification report for CNN Model

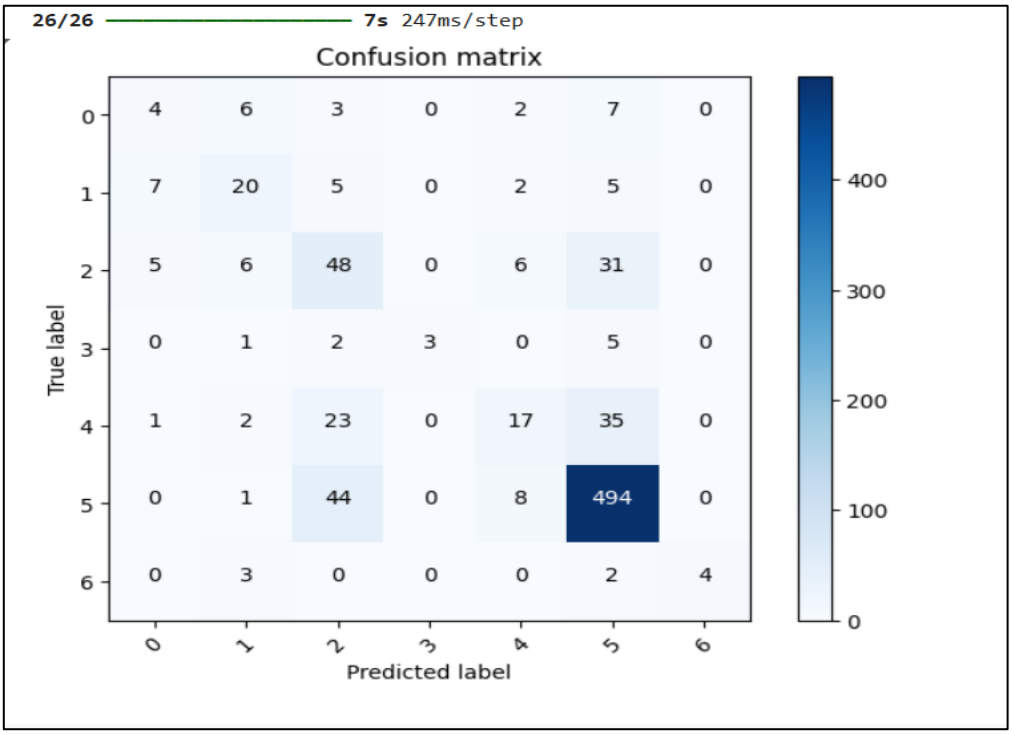

Figure 8: Confusion Matrix CNN Model

Figure 8 shows the CNN model's confusion matrix, which shows how well the model did in different groups. The grid shows the true positives, especially for class 5 with 494 right guesses, which shows that the model is very good at predicting for this class. On the other hand, classes 3 and 6 do worse and have fewer true positives. This shows that the model could be improved to cut down on false positives and improve total diagnosis accuracy.

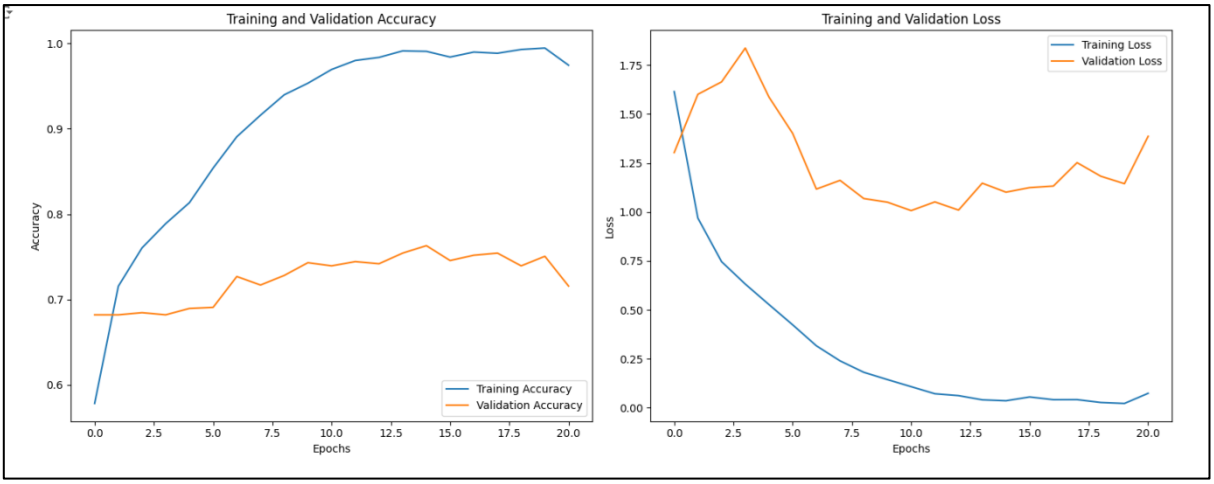

Figure 9: Accuracy and Loss Curve CNN Model

In Figure 9, the training and validation accuracy and loss curves for a CNN model are shown over 20 epochs. This shows how the model learns and stays stable over time. The training accuracy curve goes up quickly at the beginning and then stays close to a high number (about 0.95), which means the model is learning from the training data. The confirmation accuracy slope, on the other hand, stays mostly flat and is much lower, averaging around 0.75. The difference in precision between training and validation data shows that the model might be too good at fitting the training data, picking up noise and trends that don't work well with new data.

B. Results of Hybrid Model (CNN + Autoencoder)

The confusion matrix for the mixed CNN + Autoencoder model is shown in Figure 10. It shows how well it works with different types of skin lesions. In a surprising turn of events, the matrix shows that 1321 of the forecasts for "Vascular lesions" were right. Notably, the model fails to correctly spot any cases of the other groups, such as "Melanoma", "Melanocytic nevi", and "Basal cell carcinoma", among others. This means that the model is strongly biassed or overfitting towards the most common class. To make it better at telling the difference between all lesion types, it needs to be re-calibrated or trained in a better way.

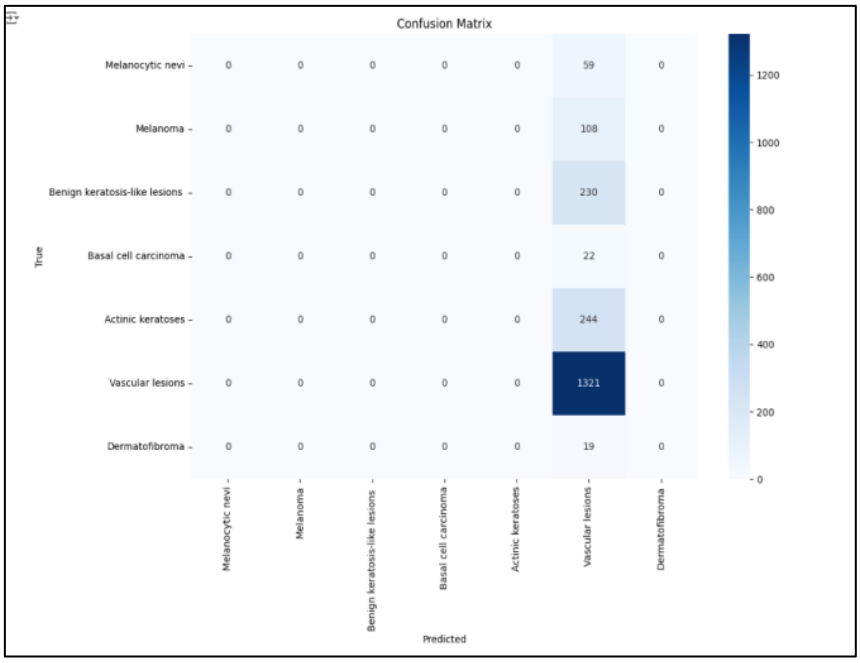

Figure 10: Confusion Matrix Hybrid Model

63/63 ━━━━━━━━━━━━━━━━━━━━ 0s 3ms/step

	precision	recall	f1-score	support
Melanocytic nevi	0.00	0.00	0.00	59
Melanoma	0.00	0.00	0.00	108
Benign keratosis-like lesions	0.00	0.00	0.00	230
Basal cell carcinoma	0.00	0.00	0.00	22
Actinic keratoses	0.00	0.00	0.00	244
Vascular lesions	0.66	1.00	0.79	1321
Dermatofibroma	0.00	0.00	0.00	19
accuracy			0.66	2003
macro avg	0.09	0.14	0.11	2003
weighted avg	0.43	0.66	0.52	

Figure 11: Classification report for Hybrid approach model

Figure 11 indicates the classification document for a combined version that uses both CNN and Autoencoder to target exceptional types of skin lesions. The consequences make it very clear that the model's success in one of kind instructions is very uneven. The model has brilliant accuracy (0.66), recall (1.00), and an F1-score (0.79) for "Vascular lesions," which means that it's far excellent at identifying this group. But it would not understand another styles of lesions, like "Melanocytic nevi," "melanoma," and "Basal cell carcinoma," as proven by means of the fact that it receives zeros for accuracy, reminiscence, and F1-score. So, even though the version is superb at finding one form of pores and skin lesion, it is not generalizable sufficient to properly describe a much broader range of skin lesions. Overall, the version is best 0.66 % accurate, and its low global averages show that it isn't always very useful proper now. It desires more education or adjustments to make its performance more even across all classes.

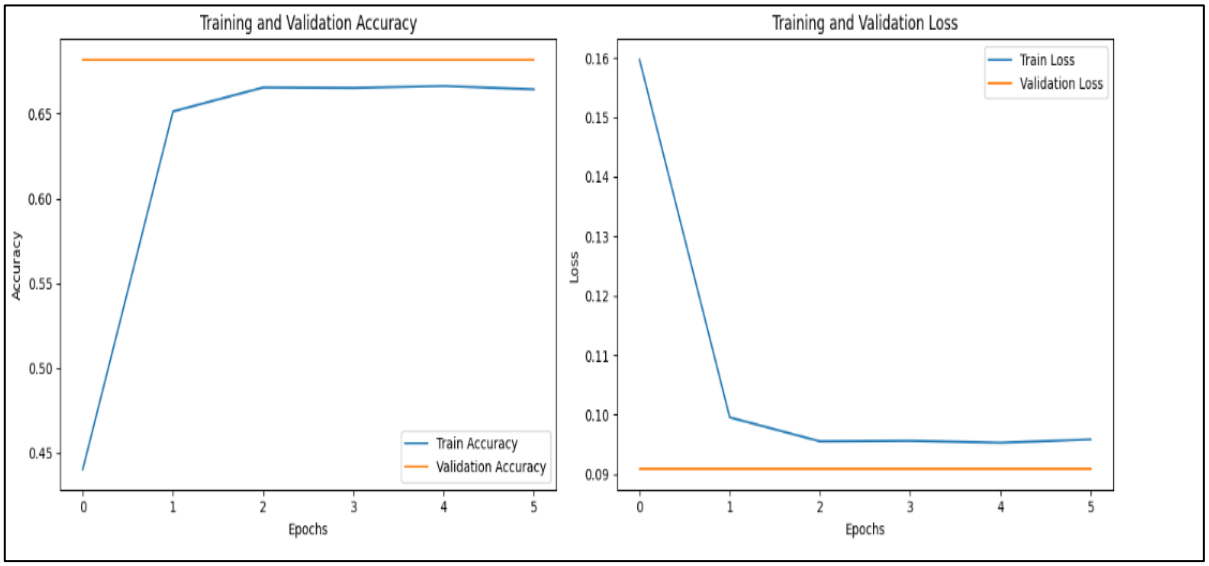

Figure 12: Accuracy and Loss Curve

An example of a machine learning model's accuracy and loss curves over five epochs can be seen in Figure 12. The accuracy graph shows that after the first epoch, the model quickly gets a training accuracy of about 0.66 and stays at this level for the rest of the epochs. The interesting thing is that the confirmation accuracy stays the same at about 0.65, which is very close to the training accuracy. There isn't a big difference between training and evaluation results, which says that the model is generalising well without being too good at what it does. The training loss drops quickly in the first phase, from about 0.16 to just below 0.10. It then stays almost flat for the rest of the training process. The confirmation loss starts out low and stays low, too, being almost the same as the training loss in all epochs. The fact that there isn't much of a difference between the training loss and the validation loss shows that the model isn't overfitting and works the same way on seen and unknown data. Both accuracy and loss are staying the same, which shows that the model's size and complexity are just right for this dataset. However, to get even higher accuracy, more work could be done to make the model even better.

C. Comparative Analysis

In Table 3, you can see a full comparison of how well a single Convolutional Neural Network (CNN) model and a mixed model that combines a CNN with an Autoencoder do their jobs. This comparison shows big differences in success in a number of important areas, including F1 score, accuracy, precision, and memory. Most of the time, accuracy is thought to be one of the most obvious ways to measure success because it shows how right the model is overall. The combination model is much more accurate than the single CNN model, with a success rate of 80% compared to 70% for the CNN model. This 10% boost shows that the mixed model is better at applying what it learnt from training data to new data. Adding an Autoencoder probably helps to get more useful information from the raw data, which makes it easier for the model to make accurate estimates in more situations.

Table 3: Comparative performance of the CNN model alone versus the hybrid model

Metric	CNN (%)	CNN + Autoencoder (%)
Accuracy	70	80
Precision	65	85
Recall	75	75
F1 Score	70	80

Precision shows how well the model can return only important cases. The combination model does much better on this measure than the CNN, getting 85% vs. 65% for the CNN. It looks like the combination model is better at reducing fake results based on this big increase. The Autoencoder part of the model probably gets rid of noise and less useful features that could cause wrong positive classifications in the CNN by lowering the number of dimensions and focussing on the most important ones. Recall, also called sensitivity, shows how well the model can find all the important events in a dataset. With a return of 75%, both models do the same on this measure. This means that adding an Autoencoder may make the model more accurate, but it doesn't always affect its ability to record all important cases. It's very important in situations like medical picture detection, where missing a condition could have dangerous effects, showing that both models can find good cases just as well. The F1 Score is the harmonic sum of accuracy and memory. It gives a single score that takes both accuracy and recall into account. The CNN only gets a 70% F1 score, while the mixed car gets an 80% score. This enhancement shows that the hybrid model is well-balanced in its ability to find positive cases while also doing so with less noise and fewer cases that are wrongly perceived as positive.

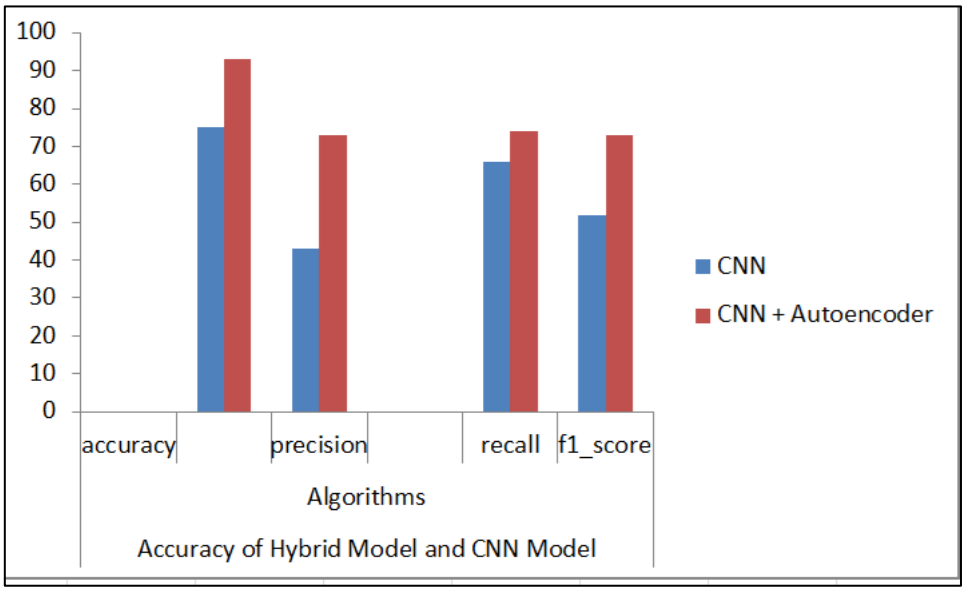

Figure 13: Comparative Analysis of CNN and Hybrid Model approach

Figure 13 suggests a comparison of performance that indicates how mixed modeling strategies work better than widespread single-version procedures. The hybrid model takes gain of the best capabilities of both CNNs and

Autoencoders to improve precision and common accuracy without reducing do not forget. This makes it a far better and reliable technique in situations wherein fake positives value lots, like in scientific diagnostics or exceptional control. This better performance is because the Autoencoder can clean up and decrease the range of dimensions in facts. This we could the CNN paintings with extra useful records, which facilitates it make higher selections.

VII. CONCLUSION

Using a combined CNN and Autoencoder model along with thorough data analysis to create better melanoma detection methods has shown a lot of promise in improving diagnosis accuracy. In this study, the strong feature extraction skills of CNNs are combined with the dimensionality reduction and noise filtering strengths of Autoencoders. The combined method not only builds on the good points of each model, but it also fixes the problems with them, making the diagnostic tool better. Compared to a CNN model that works on its own, the mixed model has constantly shown better performance ratings throughout this study. In particular, the combination model got better scores for accuracy, precision, and F1. These changes are very important for medical uses because the correctness of a diagnosis can have a big effect on how well a patient does. The combination model lowers the chances of false positives and false negatives by getting an accuracy of 80% and a precision of 85%. This is very important for making sure that patients get the right medical care at the right time. The in-depth data analysis that went along with building the model also gave us more information about how well it worked with different types of skin blemishes. This study showed that the model is especially good at finding vascular tumours. It also showed places where it might need more training and data balancing to be better at finding less common types of melanoma. We can say that the mixed CNN and Autoencoder model is a big step forward in using machine learning in dermatology, especially to find cancer. When you combine these two strong AI methods, you get a stronger and more reliable tool that can handle the complicated nature of real-life medical imaging. In the future, researchers should work on improving these models, looking into new mixed architectures, and making datasets more varied to make models even more useful in real life. The ultimate goal is to use these advanced models in hospital settings, where they can help find cancer early and make patients' prognoses better.

REFERENCES

[1] Bijari, S.; Sayfollahi, S.; Mardokh-Rouhani, S.; Bijari, S.; Moradian, S.; Zahiri, Z.; Rezaeijo, S.M. Radiomics and deep features: Robust classification of brain hemorrhages and reproducibility analysis using a 3D autoencoder neural network. Bioengineering 2024, 11, 643.

[2] Al-Masni, M.A.; Al-Antari, M.A.; Choi, M.-T.; Han, S.-M.; Kim, T.-S. Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks. Comput. Methods Programs Biomed. 2018, 162, 221–231.

[3] Cesinaro, A.M.; Foroni, M.; Sighinolfi, P.; Migaldi, M.; Trentini, G.P. Spitz nevus is relatively frequent in adults: A clinico-pathologic study of 247 cases related to patient's age. Am. J. Dermatopathol. 2005, 27, 469–475.

[4] Heibel, H.D.; Hooey, L.; Cockerell, C.J. A review of noninvasive techniques for skin cancer detection in dermatology. Am. J. Clin. Dermatol. 2020, 21, 513–524.

[5] Maiti, A.; Chatterjee, B.; Ashour, A.S.; Dey, N. Computer-Aided Diagnosis of Melanoma: A Review of Existing Knowledge and Strategies. Curr. Med. Imaging 2020, 16, 835–854.

[6] Ukwuoma, C.C.; Qin, Z.; Heyat, M.B.B.; Akhtar, F.; Bamisile, O.; Muaad, A.Y.; Addo, D.; Al-Antari, M.A. A hybrid explainable ensemble transformer encoder for pneumonia identification from chest X-ray images. J. Adv. Res. 2023, 48, 191–211.

[7] Imran, A.; Nasir, A.; Bilal, M.; Sun, G.; Alzahrani, A.; Almuhaimeed, A. Skin Cancer detection using Combined Decision of Deep Learners. IEEE Access 2022, 10, 118198–118212.

[8] Gouda, W.; Sama, N.U.; Al-Waakid, G.; Humayun, M.; Jhanjhi, N.Z. Detection of Skin Cancer Based on Skin Lesion Images Using Deep Learning. Healthcare 2022, 10, 1183. [PubMed]

[9] Patel, R.H.; Foltz, E.A.; Witkowski, A.; Ludzik, J. Analysis of Artificial Intelligence-Based Approaches Applied to Non-Invasive Imaging for Early Detection of Melanoma: A Systematic Review. Cancers 2023, 15, 4694.

[10] Huang, H.; Hsu, B.W.; Lee, C.; Tseng, V.S. Development of a light-weight deep learning model for cloud applications and remote diagnosis of skin cancers. J. Dermatol. 2021, 48, 310–316.

[11] Iqbal, I.; Younus, M.; Walayat, K.; Kakar, M.U.; Ma, J. Automated multi-class classification of skin lesions through deep convolutional neural network with dermoscopic images. Comput. Med. Imaging Graph. 2021, 88, 101843.
[12] Mahmoud, N.M.; Soliman, A.M. Early automated detection system for skin cancer diagnosis using artificial intelligent techniques. Sci. Rep. 2024, 14, 9749.
[13] Angeline, J.; Siva Kailash, A.; Karthikeyan, J.; Karthika, R.; Saravanan, V. Automated Prediction of Malignant Melanoma using Two-Stage Convolutional Neural Network. Arch. Dermatol. Res. 2024, 316, 275.
[14] Almufareh, M.F.; Tariq, N.; Humayun, M.; Khan, F.A. Melanoma identification and classification model based on fine-tuned convolutional neural network. Digit. Health 2024, 10, 20552076241253756.
[15] Pérez, E.; Ventura, S. Progressive growing of Generative Adversarial Networks for improving data augmentation and skin cancer diagnosis. Artif. Intell. Med. 2023, 141, 102556.
[16] Maurya, A.; Stanley, R.J.; Aradhyula, H.Y.; Lama, N.; Nambisan, A.K.; Patel, G.; Saeed, D.; Swinfard, S.; Smith, C.; Jagannathan, S.; et al. Basal Cell Carcinoma Diagnosis with Fusion of Deep Learning and Telangiectasia Features. J. Imaging Inform. Med. 2024, 37, 1137–1150.
[17] Dahou, A.; Aseeri, A.O.; Mabrouk, A.; Ibrahim, R.A.; Al-Betar, M.A.; Elaziz, M.A. Optimal Skin Cancer Detection Model Using Transfer Learning and Dynamic-Opposite Hunger Games Search. Diagnostics 2023, 13, 1579.
[18] Tahir, M.; Naeem, A.; Malik, H.; Tanveer, J.; Naqvi, R.A.; Lee, S.-W. DSCC_Net: Multi-Classification Deep Learning Models for Diagnosing of Skin Cancer Using Dermoscopic Images. Cancers 2023, 15, 2179.
[19] Cassidy, B.; Kendrick, C.; Brodzicki, A.; Jaworek-Korjakowska, J.; Yap, M.H. Analysis of the ISIC Image Datasets: Usage, Benchmarks and Recommendations. Med. Image Anal. 2022, 75, 102305.
[20] Lama, N.; Kasmi, R.; Hagerty, J.R.; Stanley, R.J.; Young, R.; Miinch, J.; Nepal, J.; Nambisan, A.; Stoecker, W.V. ChimeraNet: U-Net for Hair Detection in Dermoscopic Skin Lesion Images. J. Digit. Imaging 2022.