RESEARCH

Open Access



Autonomous detection of nail disorders using a hybrid capsule CNN: a novel deep learning approach for early diagnosis

Gunjan Shandilya¹, Sheifali Gupta¹, Salil Bharany¹, Ateeq Ur Rehman^{2*}, Upinder Kaur³, Hafizan Mat Som⁴ and Seada Hussen^{5*}

Abstract

Major underlying health issues can be indicated by even minor nail infections. Subungual Melanoma is one of the most severe kinds since it is identified at a much later stage than other conditions. The purpose of this research is to offer novel deep-learning algorithms that target the autonomous categorization of six forms of nail disorders by employing images: Blue Finger, Clubbing, Pitting, Onychogryphosis, Acral Lentiginous Melanoma, and Normal Nail or Healthy Nail Appearance. Based on this, we build an initial baseline CNN model, which is then further advanced by the introduction of the Hybrid Capsule CNN model by the reduction of space hierarchy deficiencies of the classic CNN model. All these models were trained and tested using the Nail Disease Detection dataset with intensive uses of techniques of data augmentation. The Hybrid Capsule CNN model, thus, provided superior classification accuracy compared to the others; the training accuracy was 99.40%, while the validation accuracy was 99.25%, whereas the hybrid model outperformed the Base CNN model with astounding precision, recall of 97.35% and 96.79%. The hybrid model additionally leverages the capsule network and dynamic routing, offering improved robustness against transformations as well as improving spatial properties. The current study consequently provides a very viable, economical, and accessible diagnostic tool, especially for places with a paucity of medical services. The proposed methodology provides tremendous capacity for early diagnosis and better outcomes for the patient in a healthcare scenario.

Clinical trial number Not applicable.

Keywords Convolutional neural network (CNN), Nail classification, Deep learning (D.L), Nail diseases, Hybrid model, Capsule network (CapsNet)

*Correspondence: Ateeq Ur Rehman 202411144@gachon.ac.kr Seada Hussen seada.hussen@aastu.edu.et ¹Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab, India ²School of Computing, Gachon University, Seongnam-si
 13120, Republic of Korea
 ³Department of Computer Science and Engineering, Lovely Professional University, Phagwara, Punjab 144411, India
 ⁴Computer and Information Sciences Department, Faculty of Science and Information Technology, Universiti Teknologi Petronas, Perak, Malaysia
 ⁵Department of Electrical Power, Adama Science and Technology University, Adama 1888, Ethiopia



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Introduction

The architectural complexity of the nail unit proves to be an important marker for the general health condition and very often represents alterations coinciding with most diseases. Architectural changes in the nails constitute important diagnostic information within a broad spectrum of diseases-from cancer and dermatological diseases to respiratory and cardiovascular diseases [1]. This study develops an intricate classification system for nail diseases based on the anatomical characteristics of the nail unit for the enhancement of accuracy in dermatological diagnosis. Detailed diagnosis of nail diseases such as onychogryphosis, cyanosis, clubbing, and koilonychia enhances the accuracy of dermatological examination and alerts the clinician to more generalized health issues including hypoxia or anemia due to an iron deficiency [2]. Besides, changes in nails may include manifestations like pitting in psoriasis or onycholysis in eczema: two diseases with a long duration.

The foundation for nail disease diagnosis has historically been a dermatologist's visual examination [3]. These evaluations were useful, but they are frequently unreliable because they heavily rely on the clinician's experience and expertise. Furthermore, in more severe situations, this subjectivity could result in a delayed diagnosis or even an incorrect one [4]. Acral lentiginous melanoma, for instance, is infamously difficult to identify at an early stage. A pigmented streak in the nail may be the first sign of this uncommon but fatal skin cancer. There could be catastrophic consequences if this illness is not diagnosed or is not recognized on time [5]. For more complicated or ambiguous situations, further diagnostic techniques including biopsy and dermoscopy were frequently used. These are crucial techniques, especially when the illness is first developing. They are not always conclusive, though, and they take a long time [6].

This study uses a nail disease image collection that includes tagged images from six different kinds of nail diseases to overcome these limitations. The suggested model has been trained and validated using the dataset, guaranteeing a comprehensive assessment of the system's performance. This labeled dataset gives a good description of ordinary diseases in nail, which helps accelerate the learning of a model and improves the accuracy of classification.

A hybrid strategy based on a CNN and CapsNet was adopted in this work. Although the CapsNets are very good at maintaining spatial hierarchies that ensure the model preserves crucial knowledge about the relationships between the various components within the image of the nail, the CNNs remain excellent for efficiency in extracting important features from photos [7]. This will hybridize the Capsule CNN model, which integrates the benefits of both architectures, such that minor differences are provided in the presentations of nail disease. The purpose of this proposed model is to solve the drawbacks of traditional diagnostic procedures by offering a more automated and accurate tool for identifying a range of nail diseases. The approach can greatly enhance clinical outcomes by providing faster, more accurate, and reliable nail disease identification. Additionally, it can lessen the need on subjective evaluations, improving the consistency and accessibility of dermatological diagnoses, especially in areas with limited resources where access to skilled clinicians might not always be guaranteed. This automated classification approach is a significant advancement in dermatology that will enable quicker and more accurate identification of nail diseases.

The major contributions of our study are stated below:

- In our paper, Shearing, 20-degree rotation, width/ height shifting, zooming, and horizontal flipping are applied as data augmentation strategies so that the network receives an enhanced dataset and does not overfit a particular set of input.
- This study proposed two models, a Base CNN model with four convolutional blocks for feature extraction, batch normalization for stable training, max-pooling for dimensionality reduction and Hybrid Capsule CNN to further improve understanding of spatial cues.
- The Hybrid Capsule CNN incorporates a primary capsule layer, dynamic routing to capture relationships between features, and a length output layer to estimate probabilities based on capsule vector magnitudes, ensuring robust spatial and hierarchical feature representation.
- Model-specific loss functions are utilized in this study, multi-class output using Sparse Categorical Cross-Entropy for the Base CNN and improved separation and accuracy using Margin Loss for the Hybrid Capsule CNN.

The structure of this study is summarized as follows: Sect. 1 presents the Introduction, Sect. 2 presents the Literature Review, Sect. 3 describes the Proposed Methodology, Sect. 4 provides the experimental results obtained for both models, and Sect. 5 discusses the Conclusion with the future scope of this study.

Literature review

The examination of fingernail color and texture as health indicators has filled a critical information gap in medicine, marking a major advancement in healthcare. This work opens a new way for early disease diagnosis by using nail traits as diagnostic tools. This study reveals intricate patterns and spatial linkages in nail photos using deep learning techniques, specifically CNNs, which enable automated and objective illness prediction. The authors of [8] have developed a mobile application that uses a deep learning model to identify disorders solely from uploaded nail photos. Additionally, a custom dataset containing three classes of nail photos selected and acquired from skin specialists has been constructed for this case. The VGG16 model yielded a 92% accuracy rate. The authors of [9] proposed a fine-tuned ResNet101V2 model to classify nail diseases into three classes. Their model attained an accuracy of 89%, precision of 90.8%, recall of 87.4% and f-1 score of 87.8%.

According to a study [10], CNN can be used to create an automated system for the identification of diseases related to the skin and nails. The created model is used in the "DermaDoc" web application to forecast these conditions and assist users by giving information about the lesion and, if available, temporary relief treatments. Psoriasis, eczema, guttate psoriasis, sebaceous cysts, paronychia, and yellow nail syndrome are diseases that are taken into consideration. With the use of several data augmentation methods and a transfer learning strategy on CNN, a greater accuracy of roughly 92.5% has been attained. By analyzing nail color, the framework of [11] can identify nine nail illnesses, such as Beau's lines, hyperpigmentation, onychomycosis, and others, in nails that are black, blue, red, white, and yellow. CNNs are used for feature extraction, and as there was no pre-existing dataset, a custom dataset of 18,025 photos from nine different illness classes was produced. The model was evaluated against many methods, including KNN, SVM, ANN, and others. With a kappa value of 84.8%, recall of 89.1%, precision of 89.8%, and accuracy of 88.05%, it demonstrated good performance in differentiating nail conditions. The research of [12] utilizes VGG16 and VGG19 models to detect nail disease in three classes. They employed 723 image datasets for the training and testing of the model. Their results concluded that VGG16 outperformed VGG19 in this classification task as VGG16 achieved 94% accuracy whereas VGG19 attained 92% accuracy. Whereas in another study conducted by [13] VGG16 attained an accuracy of 96%. The authors developed a customized dataset of 333 images and these images were divided into three classes.

A hybrid model for the classification of dermatological diseases that combines CNN and RF is presented in his study [14]. Using a data set of 15,000 photos that were divided into four categories—acne, hair loss, nail fungus, and skin allergy—the RF makes the classification judgment while CNN extracts features. 95.2% F1-score, 95.72% recall, 96.08% accuracy, and 95.12% precision were the accuracy findings. The model's effectiveness is further supported by a confusion matrix and performance table, and comparisons with more sophisticated methods demonstrate how competitively well it performs in the classification of dermatological illnesses. By combining CNNs with LSTMs, a novel method for increasing accuracy and efficacy in nail disease diagnosis was introduced by [15]. The CNN-LSTM model consistently shown exceptional accuracy when evaluated on a range of nail situations. The model achieved 94% accuracy for nail fungus, 92% for Beau's Lines, 93% for hangnails, and 91% for ingrown toenails, proving its effectiveness in identifying varied nail-related disorders. Deep Neural Networks (DNNs) developed on two skin image datasets, DermNet and ISIC Archive, are used in the study of [16] to accurately classify skin diseases. The model's DermNet results showed that it could predict 23 diseases with 80% accuracy and 98% AUC, and 622 sub-classes with 67% accuracy and 98% AUC. It categorized seven diseases on ISIC Archive with 99% AUC and 93% accuracy. The study demonstrates how accurately the model can diagnose skin conditions, almost matching human accuracy. This presents a possibility for large-scale, real-time diagnosis utilizing clinical pictures. The improved Vision Transformer for automated psoriasis detection is presented in the research of [17]. Following pre-processing, a CNN retrieves the entire image's features, which are then concatenated with every layer of the Vision Transformer encoder to preserve the entire image at every step. Preprocessed photos are split into patches at the same time and sent into the transformer using positional encoding. The suggested model yielded an F-Score of 96.5% and a classification accuracy of 97.7%. This study proposes a hybrid Capsule CNN model to detect and classify nail diseases. Table 1 gives the literature in tabular form, stating some recent advancements in the categorization field.

The literature review highlights crucial progress in the application of deep learning techniques in image classification for the detection of nail and skin diseases. Multiple models, such as VGG16, CNN-LSTM, and Vision Transformer, showed high accuracy across various studies. For instance, up to 96% classification accuracy was reached using VGG16 in [13] for nail disease classification. Hybrid approaches, such as CNN-LSTM, showed great accuracy for classifying nail fungus and Beau's Lines disease by attaining 94% and 92% accuracy for each class in [15]. Vision Transformer-based approaches presented in [17] attained an F-Score of 96.5% and 97.7% classification accuracy in detecting psoriasis in nails. This was made possible through a perfect combination of the results, hence due to custom datasets like those of 18,025 images of nail disease used in [11]. In addition, practical implementations in the form of the "DermaDoc" web platform [8], and [10] mobile applications show how the models could find their way into reality diagnostics, targeting diseases such as psoriasis, eczema, and fungal infections in nails [19–20]. Still, several challenges remain despite these developments. Studies often have limitations in

S.No	Disease Class Name	Number of	Sample Image 1	Sample Image 2
1.	Melanoma	753		
2.	Healthy	343		
3.	Onychogryphosis	689		
4.	Blue Finger	612		
5.	Clubbing	783	6	
6.	Pitting	655		6

 Table 1
 Showing the name of the disease class, number of class-wise images and sample images of each class

terms of dataset size and diversity, thereby limiting their generalization ability across wider populations.

Advancements in deep learning and hybrid techniques [21-22] have significantly enhanced diagnostic capabilities across various medical domains [23] as in CenterFormer a transformer-based framework for dental plaque segmentation, which underscores the effectiveness of advanced architectures in medical image analysis [24-25], as well as authors [26-28] proposed a hybrid denoising scheme the importance of integrating preprocessing and deep learning. These methodologies inspire the development of a Hybrid Capsule CNN [27-30] in this study to autonomously detect nail disorders [29]. Furthermore, insights from [31] as on complex data linking in medical records could guide future extensions for integrating diagnostic results with patient histories. This work aims to bridge the gap in dermatological diagnostics [32] by enabling early and accurate detection of nail disorders [33 and 34].

Furthermore, whereas some models rely on CNNs, their combinations, or variations thereof, capsule networks [27, 30, 32] still remain relatively little explored. Based on these gaps, this study introduces a Hybrid Capsule CNN model designed to classify nail diseases into six categories with improved spatial and hierarchical feature learning. This novel method builds upon earlier work and seeks to improve classification accuracy and robustness with a step toward automated dermatological diagnostics [33–34].

Proposed methodology

This section describes the proposed framework that has been used for the classification of nail diseases into six classes. Figure 1 is a descriptive figure of the suggested framework used to achieve the process of nail disease classification. From the figure, it can be noticed that data gathering is the very first step in the proposed process, which means a collection of images of nails with different conditions to prepare a dataset for testing and training purposes. For this study, the Nail Disease Detection dataset has been used to gather different nail disease images. These preprocessed images are resized to a resolution and, subsequently, undergo several data augmentation techniques like shearing, rotation by 20 degrees, shifting based on width and height, zooming, and horizontal flipping to diversify the dataset. The images are standardized so that they become suitably fit for the model and are then divided into three subsets: training, validation, and testing sets. The next step of the technique involves model development: a base CNN model and a Hybrid Capsule CNN model are defined. The Base CNN model is designed using four convolutional blocks. Each convolutional block is comprised of max-pooling layers for dimensionality reduction, batch normalization layers for training stability, and convolutional layers for feature extraction. After undergoing two convolutional blocks, the output is flattened and subsequently passed via a final layer of classification.

The Hybrid Capsule CNN model, on the other hand, incorporates a capsule network to enhance spatial comprehension while improving upon the base CNN architecture. After the convolutional layers, a primary capsule layer is introduced, converting the feature maps into capsules. A capsule layer that uses dynamic routing to ascertain the relationships between the features comes next. The final layer, the length output layer, estimates the probability of different nail ailments by calculating the size of each capsule.

The images are classified into six categories which comprise of nail diseases Clubbing, Melanoma, Onychogryphosis, Pitting, Blue Finger, and Healthy. The two models that is CNN and Hybrid Capsule CNN models are compared, with the best-performing model being highlighted. Before being tested on the test set, both models are trained and validated on the training and validation sets. A variety of measures are used to evaluate the models' performance, including confusion matrices, training and validation loss graphs, as well as classification metrics. The outcomes are shown visually to show the models' dependability and accuracy in classifying data.

To show the best classification performance, several critical hyperparameters have been carefully tuned for both models throughout all the steps in training. A total of thirty epochs was employed during the training of the models. Thirty-two images of nail diseases were dealt with in one epoch. Since the Adam optimizer performed at a very high level while handling big datasets, the careful learning rate of 0.0001 was assigned. This allowed for progressive learning gains without forcing sudden changes to the model's weights. A dropout rate of 0.45 was chosen to ensure that the models were not relying too heavily on any one pattern in the training data since some of the neurons would be randomly dropped every time training was going to pass through, thus preventing overfitting while training. The Base CNN model employed the Sparse Categorical Cross-entropy loss function because it effectively manages categorical output and is well-suited for multi-class classification challenges. The Margin loss function, which was created especially for capsule networks, was used for the Hybrid Capsule CNN model. This improved generalization and performance during the training phase by promoting better separation between the output predictions. These hyperparameter selections were crucial in preventing overfitting and lowering loss while allowing the models to learn intricate features, which resulted in an accurate classification of nail disorders.



Fig. 1 Proposed framework

Nail disease detection dataset

This study focuses on the nail disease classification. For this, the Nail Disease Detection dataset taken from Kaggle platform [18] has been used to train and test the model. This dataset comprises images aimed at identifying and classifying common nail diseases. The dataset includes 3835 images of nail conditions such as Onychogryphosis, Pitting, Melanoma, Blue Finger, Clubbing, and

Healthy nails. Table 1 displays the sample images from the dataset.

Data preprocessing

This section discusses the various methods employed for preprocessing the images according to the requirements of the model. The images are collected from different personnel, so the images can be of different sizes and some sort of noise can be also there in the images. Initially, the pictures are loaded and adjusted to a standard size of 128 by 128 pixels to guarantee that the model's input dimensions are consistent. Following resizing, the photos' pixel values are normalized by dividing each one by 255. This reduces the range of pixel intensity from 0 to 255 to 0 to 1. Because the normalization keeps the input values small, it aids in the model's faster convergence during training. The pre-processed images are then stored for later use in testing or training. Next, the images are augmented in terms of improving the diversity of class-wise images, this will resolve the issue of imbalance dataset. Figure 2. showcases the different transformations applied to the images during the implication of data augmentation.

While performing various image transformation methods, the system utilizes the following formulas:

a) **Rotation**: It increases the model's resistance to rotational fluctuations by rotating the image up to 20 degrees. This is performed as given by Eq. 1:

$$I'(a',b') = I(a.\cos(\theta) - b.\sin(\theta), a.\sin(\theta) + b.\cos(\theta))$$
(1)

where:

- θ states the rotation angle in radians.
- After rotation, the new pixel location is I'(a',b') rather than I(a, b), which represents the initial pixel location.
- b) **Width Shift**: This transformation contributes to the model's ability to generalize to small changes in object position by translating the image by 20% of its width horizontally. This is done with:

$$I'(a',b') = I(a + \Delta a, b)$$
(2)

where:

- Δa denotes the horizontal shift, and it is calculated as 20% of the image width.
- After moving horizontally, the new pixel location is I'(a',b') instead of I(a, b), which is the original pixel location.
- c) **Height Shift**: This method gives the model a 20% vertical translation, increasing its resistance to vertical displacements. It is given by:



Fig. 2 Different data augmentation transformations

$$I'(a',b') = I(a, b + \Delta b)$$
 (3)

where:

- The vertical shift, represented by Δy , is equal to 20% of the image height.
- (a, b) is the first-pixel location. After moving vertically, the new pixel location is I'(a',b').
- d) **Shearing**: To handle minor picture deformations, shear transformation introduces a "skew" effect by shifting each pixel in the image in a direction according to its distance from an axis. It is given as:

$$I'(a',b') = I(a+m.b,b)$$
(4)

where:

- The shear factor, m, is fixed at 0.2 (20% shear).
- Following the shearing transformation, the new pixel location is I'(x', y').
- e) Zooming: Arbitrary zooming in or out of the image is applied to aid in the model's generalization to various object sizes. It is calculated as given in Eq. 5:

$$I'(a',b') = I(s_a.a, s_b.b)$$
(5)

where:

- s_a, s_b are the zoom factors
- I'(a', b') is the zoomed-in or zoomed-out pixel location.
- f) Horizontal Flip: It rotates the image in a horizontal position which aids in making the model invariant to shifts in the left and right orientation and it can be given as stated in Eq. 6:

$$I'(a',b') = I(w-a,b)$$
(6)

Where:

- w is the width of the image.
- After flipping horizontally, the new pixel location is I'(a', b').

Through the process of data augmentation, a more diversified and diverse dataset is produced, which guarantees that the model is exposed to many viewpoints of the same image and enhances the model's ability to generalize effectively in testing [19]. After the application of data augmentation, the number of images increased to 7670 in the count. Figure 3 gives the details of the set-wise images after the data-splitting step. From the figure, it can be seen that there are 5369 train images, 1534 validation images and 767 test images.

Base CNN model

This section discusses the base CNN architecture. The base CNN architecture includes multiple essential elements intended for image classification. The model starts with the input layer and is fed images, usually the 128×128 size range images for consistency. The four convolutional blocks that make up the architecture's foundation each have a convolutional layer that uses filters to identify patterns or other features in the images. It is calculated as:

$$Parameters = (f_h \cdot f_w \cdot InputChannels + 1) \cdot Filters \quad (7)$$

The filters define the number of output feature maps, the input channels correspond to the original feature map's dimensions, and f_h and f_w represent the filter's height and breadth, respectively.

As the network gets deeper, these layers record both high-level and low-level characteristics. Following every convolution operation, a Batch Normalization Layer is



SPLITTING OF DATASET

Fig. 3 Set-wise description of dataset splitting

used to normalize the output. By keeping the mean and variance of the activations within a predetermined range, this normalization helps to stabilize and speed up the training process [20]. It is calculated as:

$$Parameters = 2 \cdot Output Channels \tag{8}$$

A Max-Pooling Layer follows, which keeps the most important characteristics while decreasing the spatial dimensions of the feature maps. This aids in downsampling the data and increases the model's computing efficiency. Its output shape is determined by:

$$H_{out} = \frac{H_{in} - Pool \ Size + 2 \cdot Padding}{Stride} + 1 \quad (9)$$

$$W_{out} = \frac{W_{in} - Pool \ Size + 2 \cdot Padding}{Stride} + 1 \quad (10)$$

After the sequence of operations of convolution, normalization, and pooling, the output goes to the flatten layer that takes the 2D feature maps and makes them a 1D vector to be passed into the fully connected layers. Once the flattened feature vector is processed, Dense Layer 1 puts together the extracted features to discover more intricate associations. Following this dense layer, a Dropout Layer is introduced to prevent overfitting [21]. During training, it randomly disables a portion of the neurons to improve the model's generalization ability. With each neuron representing a potential class, the Dense Layer 2, the output layer, generates the final classification. A Softmax

Table 2 Base CNN model summary

Layer type	Output shape	Parameters					
Input	(128, 128, 3)	0					
Conv2D	(None, 128, 128, 32)	896					
Batch_Norm	(None, 128, 128, 32)	128					
Max-Pool2D	(None, 64, 64, 32)	0					
Conv2d_1	(None, 64, 64, 64)	18,496					
Batch_Norm_1	(None, 64, 64, 64)	256					
Max-Pool2D_1	(None, 32, 32, 64)	0					
Conv2d_2	(None, 32, 32, 128)	73,856					
Batch_Norm_2	(None, 32, 32, 128)	512					
Max-Pool2D_2	(None, 16, 16, 128)	0					
Conv2d_3	(None, 16, 16, 256)	295,168					
Batch_Norm_3	(None, 16, 16, 256)	1024					
Max-Pool2D_3	(None, 8, 8, 256)	0					
Flatten	(None, 16384)	0					
Dense	(None, 512)	8,388,736					
Dropout	(None, 512)	0					
Dense_1	(None, 6)	3078					
Total Parameters: 8,787,112							
Trainable Parameters: 8,78	Trainable Parameters: 8,787,112						
Non-Trainable Parameters	:: 0						

activation function is applied to this layer for multi-class classification to transform the raw outputs into probabilities for each class.

$$P(y_i) = \frac{\exp(z_i)}{\sum_{j=1}^{n} \exp(z_i)}$$
(11)

where Z_i is the raw output (logit) for class i. The following formula is used to determine the Dense Layer 2 parameters:

$$Parameters = (Input Units + 1) \cdot Output Units$$
 (12)

This architecture is appropriate for complex image classification tasks as it integrates batch normalization and dropout for regularization, convolutional layers for feature extraction, and fully connected layers for classification. Table 2 presents the summary of the base CNN model. The table displays the quantities of trainable and non-trainable attributes.

Capsule network

A neural network architecture called a Capsule Network (CapsNet) was created to get around some of the drawbacks of conventional CNNs, most notably its incapacity to represent hierarchical spatial relationships between features. Scalar activations, the foundation of feature extraction in CNNs, identify the existence of a feature but obscure important details about its spatial characteristics [22–23]. On the other hand, capsules—collectives of neurons that produce vectors, are introduced by capsnets. These vectors express an entity's pose, position, orientation, and other instantiation factors in addition to just showing the entity's existence (such as an object part). The orientation of the capsule vector encodes the entity's attributes, while its length indicates the likelihood that the entity exists. The vector is subjected to a squashing function, which maintains orientation while compressing length to a value between 0 and 1, to guarantee that capsule outputs are bounded between 0 and 1. Capsule Networks are extremely useful for medical image analysis because they encode the presence and spatial properties, such as orientation and size, of features, thereby being insensitive to variations such as rotation or scaling. They can model part-whole relationships quite effectively and capture anatomical hierarchies that are crucial in medical data. Therefore, it is ideal to use these networks to handle the complexity and variability of medical imaging modalities such as X-rays, MRIs, and CT scans. Figure 4 shows the CapsNet Architecture.

In the Primary Capsule Layer, the initial significant capsule layer, the feature maps generated by convolutional layers are organized into capsules. For example, multiple feature maps can form a capsule, each contributing a



Fig. 4 Capsule network architecture

dimension to the vector. The solitary neuron of a CNN cannot encode as much information about a detected feature as a capsule vector, which aids the network in understanding the relationships between object pieces. Additionally, CapsNets [24] employs a novel routing strategy called dynamic routing by agreement. This strategy ensures that outputs from lower-level capsules, which are responsible for detecting fundamental features, only reach higher-level capsules, which are responsible for detecting more sophisticated features, if there is a strong consensus regarding the outcome. Every lower-level capsule employs a transformation matrix that discovers relationships between them to forecast a higher-level capsule. Routing coefficients denote the significance of each capsule's prediction in determining the eventual output of the higher-level capsule. The result is a prediction vector. These coefficients are dynamically modified by an iterative procedure that assesses the agreement between capsules and a softmax function. The routing coefficient is raised, strengthening the link between those capsules, if a forecast matches the output of the higherlevel capsule.

CapsNets are designed with a margin loss function for the image categorization task [27]. This loss is different as traditionally CNNs utilize the softmax or cross-entropy loss algorithms, This encourages a capsule's output to have a length close to 1 if it detects the correct class and close to 0 otherwise. The margin loss is provided with higher and lower criteria to ensure that the correct class capsules produce long vectors while the erroneous class capsules produce short ones. The network can also be assisted in capturing fine-grained properties by employing a reconstruction loss. By rebuilding the input image using the output of the capsules, the technique aims to minimize the mean squared error, a measure of the difference between the original and rebuilt images. This leads to the encouragement of complex, discriminative features in the capsules. In conclusion, some of the main issues with CNNs are resolved by Capsule Networks, which preserve spatial hierarchies through the use of vector-based representations and dynamic routing. As a result, the network can handle changes like rotations and perspective shifts more effectively. Because CapsNets can learn more informative feature representations and are more resilient to affine transformations than CNNs, which may lose spatial information through pooling operations, they outperform CNNs in applications that demand detailed recognition.

Hybrid capsule CNN model

This section discusses the design of a hybrid Capsule CNN model. The Hybrid Capsule CNN Model blends traditional CNNs with Capsule Networks to integrate the benefits of both architectures [30]. The Hybrid Capsule CNN model incorporates a capsule network to enhance spatial comprehension while improving upon the base CNN architecture [32]. After the convolutional layers, a primary capsule layer is introduced, converting the feature maps into capsules. A capsule layer that uses dynamic routing to ascertain the relationships between the features comes next. The final layer, the length output layer, estimates the probability of different nail ailments by calculating the size of each capsule. The proposed model addresses key limitations of traditional methods by using dynamic routing and capsule vectors. Dynamic routing avoids the information loss caused by pooling in CNNs, ensuring that features are adaptively preserved and hierarchically structured. Capsule vectors, unlike scalar activations, encode both the presence and spatial attributes (e.g., pose, rotation) of features, enabling the model to capture part-whole relationships and spatial transformations effectively. These features improve the model's ability to handle variability and structural complexity in data, overcoming the limitations of traditional CNNs.

The Hybrid CNN-CapsNet Model integrates the advantages of feature extraction in CNN with spatial relationship capture in CapsNets. The architecture starts with an input layer feeding the image into four hierarchical convolutional blocks, which consist of convolutional layers for the detection of patterns, batch normalization for the stabilization of training by normalizing the feature maps, and max-pooling layers, reducing the spatial dimensions but keeping the crucial features. This sequence allows the model to learn progressively abstract and complex features. Further generalization and prevention of overfitting during training are achieved by using a dropout layer after the last convolution block, which randomly deactivates neurons to minimize reliance on specific features. The convolutional blocks form the foundation for feature extraction, which is critical for moving into the Capsule Network part of the architecture.

Now after feature extraction, this output is passed to Primary Capsule Layer, so the process now is no longer a traditional CNN processing but has shifted to the CapsNet framework. Here the feature maps are grouped into capsules that are vectors, which may represent not only the existence of a feature but also spatial attributes such as the position, orientation, and size. This representation overcomes one of the main drawbacks of CNNs, which are usually prone to losing spatial hierarchies due to max-pooling. The model then uses dynamic routing to make sure that capsules interact effectively and routes the outputs to higher-level capsules based on agreement, keeping meaningful spatial relationships. The Higher-Level Capsule Layer encapsulates abstract and complex feature relationships, allowing the network to represent objects or parts with spatial and hierarchical precision. Lastly, the Length Layer calculates the norm of capsule vectors, which actually reflects class probabilities, and therefore provides an exact classification. The hybrid approach closes the gap between CNN's feature extraction efficiency and spatial awareness abilities of CapsNets, and hence, is highly suitable for applications where strong object-part relationships are involved, like in medical image analysis or complicated object detection. The outputs from the capsule layer are vectorized by employing a squashing function as given by Eq. 7:

$$v_i = \frac{\|x_i^2\|}{1 + \|x_i^2\|} \cdot \frac{x_i}{\|x_i\|}$$
(13)

where:

the input to the capsule is represented by the symbol x_i . The squashing function makes sure that the output vector, v_i , is normalized to have a length between 0 and 1.

A capsule vector aids the network in understanding the interactions between the image parts, which a single neuron in a CNN cannot encode as much information about a detected feature. Furthermore, dynamic routing by agreement is a novel routing strategy used by CapsNets that ensures that outputs from lower-level capsules, or those in charge of detecting basic features, only reach higher-level capsules, or those in charge of detecting more complex features, if there is a strong consensus regarding the outcome.

$$c_{ij} = softmax\left(b_{ij}\right) \tag{14}$$

Here, c_{ii} is the coupling coefficient between capsule i in the current layer and capsule j in the subsequent. The network can ascertain how much of the output of capsule i contributes to capsule j. Each lower-level capsule forecasts a higher-level capsule by utilizing a transformation matrix to find correlations between them. Routing coefficients show how much importance each capsule's forecast ought to have in determining the higher-level capsule's final output. The outcome is a vector for prediction. An iterative process measuring the agreement between capsules and a SoftMax function dynamically modifies these coefficients. If the prediction matches the output of the higher-level capsule, the routing coefficient is increased, strengthening the connection between those capsules. The Higher-Level Capsule Layer receives the output from this layer after that. The probability that the input corresponds to the class that each capsule in the layer represents is indicated by the length of the capsule vector in this layer. Every capsule produces a vector as its output, and the length of the vector indicates the likelihood that the class was identified.

Ultimately, the categorization result is provided by the Length (result) stage. The probability of the discovered class is correlated with the length of each capsule vector. To guarantee that the vector lengths of the proper class capsule and the incorrect class capsules are almost equal, this model employs a Margin Loss function. For image classification tasks, the hybrid model is particularly useful because it combines the strong routing mechanism and spatial awareness of CapsNets with the feature extraction capability of CNNs. The margin loss is defined as follows:

$$L_{i} = T_{j} \cdot \max(0, m^{+} - ||v_{j}||)^{2} + \lambda - (1 - T_{j})$$

$$\cdot \max(0, ||v_{j}|| - m^{-})^{2}$$
(15)

Here, the output capsule vector is represented by v_j , the margins (hyperparameters) are represented by m^+ and m^- . The ground truth is represented by T_j .



Fig. 5 Base CNN model accuracy and loss evaluation (a) Training and validation loss (b) Training and validation accuracy

Results and discussion

This section of the article presents a graphical examination of the model's performance during the training, validation, and testing stages. The Base CNN model's performance evaluation is covered in Sect. 4.1, while the Hybrid Capsule CNN model's performance evaluation is covered in Sect. 4.2. After that, both models are compared in Sect. 4.3, and the model that performs the best is chosen to classify nail illnesses into six different groups. Last but not least, Sect. 4.4 presents the categorization outcomes attained during the best-performing model's testing phase.

Accuracy and loss representation for base CNN model

This section displays various performance measures achieved during the training and testing phase of the base CNN model.

Figure 5 showcases the graphical view of the accuracy and loss development models The two graphs are the performance of the model throughout thirty epochs showing accuracy and loss for both the training set and validation set. From the loss graph, it is apparent that the model is learning efficiently and minimizing mistakes since the training loss declines smoothly to 0.02 by the end of the thirty-first epoch from the very high level at the start of 0.85. There are similar trends in the validation loss, which begins from 0.44 and decreases down to 0.11 while showing in detail how well the model generalizes to unseen data without overfitting.



Fig. 6 Confusion matrix for base CNN model

From Fig. 5(b) it can be analyzed that the training accuracy of the accuracy graph rises from 61.04 to 99.4%, showing great improvement from the training set in terms of prediction accuracy. Then, the validation accuracy grew stronger from a higher beginning position of 85.41% up to 97.75%. Both of the figures grow progressively, which means the model is consistent and well-trained on both datasets.

Figure 6 shows the confusion matrix that was obtained when the base CNN Model was tested. The confusion matrix provides an analysis of how well the Base CNN model performs as it compares expected and actual class

Page 13 of 19

labels, citing nail diseases into six categories. Off-diagonal numbers are those having errors, while the diagonal elements show correct predictions. The model got 124 pictures correct for Class 0 while getting 1 image wrong for Class 1, 1 image wrong for Class 2, and 2 images wrong for Class 5. Although this results in some tiny error margins regarding the classification of some of the pictures into neighboring classes, it is still quite a good performance for Class 0. Since only one picture was confused with Class 0 while another one was with Class 5, Class 1 got 126 correct pictures and showed practically a perfect performance for that class. Class 2 produced 125 correct predictions, but of course, with uncertainly over 1 image wrongly classified as Class 1, 1 as Class 3 and 1 as Class 5. It hence indicates that even though the model performs very efficiently, there are times that it will misinterpret small changes of illness trends. Though there were some mistakes in the light of the 122 correctly categorized Class 3 images, the overall classification was still quite strong. Two were incorrectly classified as Class 0, and one as Class 5. Three of the 125 images in Class 4 were falsely identified as Class 1, showing huge overlapping features of these two classes. The rest of the 125 images were classified correctly. It correctly classified 128 images, although in some confusion; one was classified into Class 0 and the other in Class 3. Summarizing, the confusion matrix shows that it does a good job of differentiating between the six groups since most of the predictions are on the diagonal, thus representing correct classes.

Table 3 gives the performance parameters achieved during the evaluation of the base CNN model. This table gives a deeper view of how well this model works for most categories. The classes with high precision demonstrate the accurate predictions made by the model. The precision varies from 97.04% for Class 5 to 98.43% for Class 0. True positive forecasting proportions in comparison with all positive predictions vary from 92.61% for Class 0 to 95.19% for Class 3, which indicates a result of recall to be quite lower but impressive, measuring the skill of the model in identifying all the real positives. That is, in most images, the model correctly identifies the class, but very few images exist where the model classifies the true label incorrectly.

Table 3 Performance measures for base CNN model

Class label	Precision %	Recall %	F-1 score %	Accuracy %
0	98.43	92.61	95.47	97.75
1	97.71	94.92	96.41	
2	97.14	94.13	95.60	
3	97.38	95.19	96.31	
4	97.28	93.88	95.60	
5	97.04	93.43	95.22	

The model does not deviate much, actually steadily performing on all six classes to demonstrate its power. Indeed, as seen above, the F1-score, or harmonic mean of precision and recall, balances these two measures with values ranging from 95.22% (Class 5) to 96.41% (Class 1). The model again yields consistently high accuracy with a measure of the percentage of accurate predictions for every class at 97.75%, hence suggesting that the model is in general effective for the accurate classification of nail diseases. This thorough analysis shows that the model could achieve almost perfect classifying performance based on the fact that it correctly and reliably distinguishes the six different types of disease. Although there were a few small misclassifications highlighted by the confusion matrix, the model did a fairly good job overall. By the conclusion of 30 epochs, the model had learned a great deal, as evidenced by the training and validation metrics. The model is a trustworthy tool for disease classification because the class-wise performance measures demonstrate that it is accurate and efficient across all six nail disease categories.

Accuracy and loss representation for hybrid CNN model

From the performance metrics presented in the two tables, the behavior of the model while training and validating could be gathered comprehensively with its capability to classify classes. Figure 7 gives a graphical view of the hybrid Capsule CNN model accuracy and loss development. From Fig. 7(a) the model's loss can be seen that T.L decreases gradually from a high initial value to a much better fit to training data and, after some minor oscillations, maintains approximately zero good convergence. The V.L is approximately the same, in terms of trend; it is decreasing and settling which is good regarding generalization to unseen data. The proximity of the training and the validation loss lines indicates that the model is not overfitting.

Further evidence of the model's effectiveness is provided by the accuracy graphs of Fig. 7(b) for training and validation. The T.A steadily increases as the procedure goes on, reaching almost 100% at the end. The V.A, which consistently follows the training accuracy and stabilizes at 99%, shows how resilient the model is and how well it can generalize to new data. From the figure, it can be seen that during the 16th epoch, the T.A drops but then it further increases to an optimum value.

The confusion matrix obtained during the hybrid Capsule CNN Model's testing is shown in Fig. 8. This indicates that the model performs well in categorization for each of the six classes. The confusion matrix shows that the Hybrid CNN-CapsNet model performs well overall, with most predictions correctly matching the actual classes, as seen in the high numbers along the diagonal. However, there are a few misclassifications, such as three



Fig. 7 Graphical presentation of the hybrid capsule CNN model (a) Model training and validation loss (b) Model training and validation accuracy



Fig. 8 Confusion matrix for hybrid capsule CNN model

 Table 4
 Performance parameters for the hybrid capsule CNN model

Class Label	Precision %	Recall %	F-1 Score %	Accuracy %
0	97.79	96.37	97.07	99.25
1	97.08	97.36	97.21	
2	96.50	97.44	96.96	
3	97.42	96.57	96.99	
4	97.79	96.37	97.18	
5	97.55	96.68	97.11	

samples from Class 3 being incorrectly predicted as Class 0 and one sample from Class 1 also classified as Class 0. These mistakes can occur due to overlapping features between certain classes, making them harder to distinguish. The matrix shows that the model works admirably overall, with very few small classification errors.

Table 4 illustrates the model's performance in accuracy, precision, recall, and F1-score on six different classes. Class 0 performs highly in the prediction of this class with very few misclassifications with a precision of 97.79%, recall of 96.37%, and an F1-score of 97.07%. Class 1 achieves 97.08% precision, 97.36% recall, and 97.21% F1-score which confirms the great efficiency of the model in producing correct classifications of data belonging to this class. Class 2 performances are uniform throughout this category with an F1-score at 96.96% based on a precision of 96.50% and a recall of 97.44%. The classification performance of the model was quite balanced for Class 3 as well with precision at 97.42%, recall at 96.57%, and F1-score at 96.99%. Class 4 performs similarly with an accuracy of 97.79%, recall at 96.37% and an F1-score of 97.18%, Class 4 performs similarly. Lastly, Class 5 reveals the resilience of the model with all the six classes crossed with an accuracy of 97.55%, a precision is 96.68%, and an F1-score of 97.11%.



Performance Comparison Chart for both the Models



Table 5 Compariso	Table 5 Comparison of both the models							
Parameters	Base CNN Model	Hybrid Capsule CNN Model						
Feature Representation	• Uses scalar outputs where each neuron represents the presence of a feature.	• Capsules produce vector outputs that represent both feature presence and spatial properties.						
Dynamic Routing	CNN layers have fixed connections without dynamic communication.	• Utilizes dynamic routing by agreement, allowing flexible com- munication between capsules based on feature agreement.						
Handling Spatial Hierarchies	 It has difficulty in capturing part-whole relationships in complex structures. 	 Effectively captures part-whole relationships, improving its ability to handle spatial hierarchies and transformations. 						
Robustness to Transformations	• Less robust to variations like rotation, scaling, or translation.	• More resistant to such transformations due to capsule vectors and dynamic routing.						
Training Complexity	 Simpler and faster to train with lower computational requirements. 	 More complex and computationally demanding due to its capsule architecture and routing process. 						
Overfitting and Generalization	 Prone to overfitting, often requiring dropout and other regularization methods. 	 Better generalization and less prone to overfitting, thanks to richer feature representations and routing mechanisms. 						
Common Application Areas	• Extensively used for image classification, object detection, and related tasks.	• Suitable for tasks requiring strong spatial understanding, such as detailed image recognition and segmentation.						
Interpretability	• Limited interpretability, as neurons only capture whether a feature is present or not.	 Provides higher interpretability by capturing both feature presence and spatial characteristics within capsules. 						

. . *c* .

The slight reduction in the Precision index of the hybrid Capsule CNN model for Class Label 0 and 1, compared to the base CNN model, the reason for this can be the feature overlap within the dataset. If the features of Class Label 0 and 1 significantly overlap with other classes, the dynamic routing mechanism might result in a marginal reduction in precision while aiming to preserve spatial relationships for all classes. Overall, all these classes have accuracy levels over 96% and provide substantial evidence that the model is useful for low error rates of its capability in classifying six different nail disease groups. These results establish that the model has a great generalizing capacity and depends on reliability at an overall validation accuracy of 99.25%.

Comparison of both the model's performance

This section discusses the comparison of the performance of both Base CNN and Hybrid Capsule CNN models. Figure 9 gives the comparison chart for the performances of both models. The figure depicts the overall performance of the models in terms of precision, recall, f-1 score and accuracy.

As may be observed from Fig. 9, the Hybrid Capsule CNN model has outperformed the base CNN model in terms of all the performance measures. From this comparison chart, consequently, it can be said that the Hybrid Capsule CNN model is well suited for the classification of nail diseases into six classes.

Table 5 gives a comparison table for Base CNN and Hybrid Capsule CNN Model. From this table, it can be easily analyzed how the hybrid Capsule CNN model is better than base CNN model.

Visualization of the classification results

This section showcases the classification results achieved during the testing of the best-performing model. Figure 10 displays the classification results labeled with actual class and predicted class label.



Fig. 10 Visualization of classification results

 Table 6
 Performance analysis of proposed model on different optimizers

•	necuracy	Precision	Recall	F-1 Score
Adam	99.25 ± 0.50	97.35 ± 0.40	96.79 ± 0.45	97.07 ± 0.38
Adagrad	95.67 ± 0.62	94.23 ± 0.58	95.01 ± 0.55	94.62 ± 0.60
SGD	97.90 ± 0.47	97.65 ± 0.42	96.78 ± 0.45	97.21 ± 0.41
RMSprop	98.56 ± 0.53	97.03 ± 0.48	97.79 ± 0.43	97.41 ± 0.46
	Adam Adagrad SGD RMSprop	Adam 99.25±0.50 Adagrad 95.67±0.62 SGD 97.90±0.47 RMSprop 98.56±0.53	Adam 99.25±0.50 97.35±0.40 Adagrad 95.67±0.62 94.23±0.58 SGD 97.90±0.47 97.65±0.42 RMSprop 98.56±0.53 97.03±0.48	Adam 99.25±0.50 97.35±0.40 96.79±0.45 Adagrad 95.67±0.62 94.23±0.58 95.01±0.55 SGD 97.90±0.47 97.65±0.42 96.78±0.45 RMSprop 98.56±0.53 97.03±0.48 97.79±0.43

Analysis of proposed model with different optimizers

The standard deviation results provided in Table 6 were from several runs of the Hybrid Capsule CNN segmentation model, each with a different optimizer. In each case, the model was trained and tested on several independent experiments or folds, and then its performance metrics, that is, accuracy, precision, recall, and F1-score, were computed for each run. The standard deviation of each metric was then calculated to capture the variability in the performance across different runs. This allows us to evaluate consistency and reliability in the results yielded by the model under each optimizer. Table 6 describes the proposed model's performance achieved on different optimizers.

From Table 6 it can be analyzed that with an accuracy of 99.25%, precision, recall, and F1-score values of 97.35%, 96.79%, and 97.07%, respectively, and extremely

low SD values, Adam is the best performer due to its consistent and dependable performance. With the highest recall (97.79%), good precision (97.03%), and accuracy of 98.56%, RMSprop comes in second with an F1-score of 97.41%, demonstrating consistent results as well. With a 97.90% accuracy rate and good metrics (97.65% precision, 96.78% recall, and 97.21% F1-score) with little variance, SGD also performs well. Despite its effectiveness, Adagrad performs marginally worse than the others, with slightly higher SD values and accuracy, precision, and recall of 95.67%, 94.23%, and 95.01%, respectively, as well as an F1-score of 94.62%. Considering all analysis, Adam is the best optimizer, offering the most accuracy and consistent performance; RMSprop and SGD are other excellent options.

Comparison of proposed model with transfer learning models

The performance of various models has been shown in Table 7, in which the Proposed Hybrid Capsule CNN, is compared using several metrics. Along with the standard deviations (SD), which show how consistent the findings are over several runs, the table displays the performance of several segmentation models based on accuracy, precision, recall, and F1-score. DenseNet121 obtains an

Tab	le 7	Com	parative a	anal	vsis of	pro	posed	mode	el wit	h ot	her	transf	fer	learning	mod	el	S
			1		/		1										

S.No	T.L Model	Accuracy	Precision	Recall	F-1 Score
1.	DenseNet121	87.91±1.25	90.82±1.10	86.85 ± 1.35	89.19±1.20
2.	ResNet50	92.89 ± 0.85	90.34 ± 0.95	91.08 ± 0.90	90.39 ± 0.80
3.	VGG19	89.37 ± 1.10	85.67 ± 1.20	86.09 ± 1.25	85.02 ± 1.15
4.	Proposed Hybrid Capsule CNN	99.25 ± 0.50	97.35 ± 0.40	96.79 ± 0.45	97.07 ± 0.38

Table 8 Comparison with current state of art

Reference	Technique/ Model used	Number of Images/ Dataset	Number of Classes	Performance Measures			
				Accuracy %	Other Parameters		
[7]/2022	Deep Learning	120/Patient Data	2	82	Recall 72.7% Specificity 72.9% AUC 0.755		
[<mark>18</mark>]/2023	VGG16	723/Nail Disease Dataset	3	94	-		
[<mark>19</mark>]/2022	Ensemble of CNNs	185/Patient Data	2	95	Precision – 94%		
[<mark>20</mark>]/2022	UNet	723/Nail Disease Dataset	2	86.49	-		
[21]/2024	CNN MobileNetV2	1159/Nail Disease image dataset	8	97	-		
Proposed Hybrid Capsule CNN Model		3835/Nail Disease Detection Dataset	6	99.25	Precision – 97.35% Recall – 96.79% E-1 Score – 97.07%		



Fig. 11 Graphical presentation of comparison analysis

F1-score of 89.19% with an accuracy of 87.91%, precision of 90.82%, and recall of 86.85%. There is moderate stability in its SD values. With a 92.89% accuracy rate and balanced precision (90.34%) and recall (91.08%), ResNet50 outperforms, resulting in an F1-score of 90.39%. Its lower SD values suggest more dependable outcomes. VGG19 performs marginally worse, with considerable consistency demonstrated by its SD values and accuracy of 89.37%, precision of 85.67%, recall of 86.09%, and F1-score of 85.02%. Notably, the Proposed Hybrid Capsule CNN achieves an exceptional accuracy of 99.25%, precision of 97.35%, recall of 96.79%, and F1-score of 97.07%. Very low standard deviations (SD) for example, ± 0.50 for accuracy indicate remarkable consistency and dependability. Considering all these performance parameters, the Proposed Hybrid Capsule CNN is the greatest option for the segmentation assignment since it performs the best across all parameters and has exceptional consistency.

Comparison with the current state of art

To identify nail diseases into six different categories Onychogryphosis, Pitting, Melanoma, Blue Finger, Clubbing, and Healthy nails this section compares the performance of the suggested Hybrid Capsule CNN Model with particular recent studies. The proposed study and the actual State of Academic Knowledge are contrasted in Table 8. According to the table, the proposed Hybrid Capsule CNN Model fared better than rival models like VGG16, Ensemble Model, UNet, and others. With an accuracy of 99.25%, total precision of 99.13%, recall of 99.13, and f-1 score of 99.12%, it was the most successful. These remarkable outcomes show how the proposed framework may more accurately and efficiently classify photos of nail diseases. The closest rival to the suggested model is CNN-MobileNetV2. Figure 11 shows the comparison analysis graphically. From the figure, it can be seen that the proposed hybrid Capsule CNN model has outperformed other models efficiently in classifying the corn leaf diseases.

In comparison, methods like VGG16, UNet, and other ensemble-based CNNs [17-21] perform well but fall short in overall accuracy and specificity for multi-class classification tasks. The Hybrid Capsule CNN's strong performance can be attributed to its advanced architecture, which likely combines the advantages of capsule networks in capturing spatial hierarchies with the robustness of CNNs in feature extraction. Given its impressive performance on the Nail Disease Detection Dataset, the model has potential applications in other areas of dermatology. It could be extended to detect and classify various skin conditions, including eczema, psoriasis, or melanoma, by training on diverse datasets. Its high precision and recall make it suitable for tasks where accurate detection and minimal false negatives are critical. However, further studies with larger, diverse datasets and different dermatological applications would be needed to validate its generalizability.

Conclusion and future work

In conclusion, this research began with the development of a Base CNN model for nail disease classification and progressed to the creation of a more advanced Hybrid Capsule CNN model to improve classification performance. The integration of capsule networks into the Hybrid model significantly enhanced its ability to capture spatial hierarchies and handle transformations, leading to better overall classification outcomes. The Nail Disease Detection dataset has been employed to conduct the training and testing of both models. With an accuracy of 99.25%, the Hybrid Capsule CNN model provides a more accurate, robust, and dependable solution for automated nail disease classification then Base CNN model with 97.75 accuracy. Its potential applications extend to medical diagnostics and healthcare automation, where accurate disease detection is critical for effective treatment.

The proposed Hybrid Capsule CNN shows great potential for real-world medical applications, but its clinical adoption would require thorough validation on diverse, real-world datasets that reflect variations in patient demographics, imaging conditions, and disease stages. To adapt it for clinical use, the model could be integrated into hospital workflows or telemedicine platforms as a decision-support tool, with user-friendly interfaces enabling easy input and interpretation of diagnostic results. With further fine-tuning, the model could also be extended to detect a broader range of dermatological conditions, transforming it into a versatile clinical asset for improving diagnostic accuracy and patient care.

Acknowledgements

Not Applicable.

Author contributions

Gunjan Shandilya: Conceptualization; Data curation; Formal analysis; Methodology; Writing - original draft; Software. Sheifali Gupta: Investigation; Methodology; Writing - original draft; Writing - review & editing. Salil Bharany: Writing, Reviewing and Editing; Project administration; Investigation; Methodology. Ateeq Ur Rehman: Writing - review & editing; Methodology; Conceptualization. Upinder Kaur: Writing, Reviewing and Editing; Project administration; Investigation; Methodology. Hafizan Mat Som: Writing, Reviewing and Editing; Project administration; Investigation. Validation. Seada Hussen: Writing - review & editing; Software; Resources; Methodology.

Funding

No funding is available for this research.

Data availability

The dataset used in this study is publically available on the Kaggle repository at: https://www.kaggle.com/datasets/nikhilgurav21/nail-disease-detection-d ataset.

Declarations

Ethics approval and consent to participate

This article does not contain any studies with human participants performed by any of the authors.

Consent for publication

Not Applicable.

Competing interests

The authors declare no competing interests.

Received: 11 October 2024 / Accepted: 24 December 2024 Published online: 30 December 2024

References

- Jansen P, et al. Deep learning assisted diagnosis of onychomycosis on wholeslide images. J Fungi (Basel). 2022;8(9):912.
- Aishwarya A, Goel, Nijhawan R. A deep learning approach for classification of onychomycosis nail disease. in Lecture notes in Electrical Engineering. Cham: Springer International Publishing; 2020. pp. 1112–8.
- Karunarathne HHM, Senarath GLCCS, Pathirana KPTP, Samarawickrama HM, Walgampaya N. Nail abnormalities detection and prediction system, in 2023 5th International Conference on Advancements in Computing (ICAC), 2023, pp. 394–399.https://doi.org/10.1088/1742–6596/1201/1/012052
- Di Biasi L, Auriemma Citarella A, De Marco F, Risi M, Tortora G, Piotto S. Exploration of genetic algorithms and CNN for melanoma classification. in Communications in Computer and Information Science. Cham: Springer Nature Switzerland; 2022. pp. 135–8.
- Lilly KK, Koshnick RL, Grill JP, Khalil ZM, Nelson DB, Warshaw EM. Cost-effectiveness of diagnostic tests for toenail onychomycosis: a repeated-measure, single-blinded, cross-sectional evaluation of 7 diagnostic tests. J Am Acad Dermatol. 2006;55(4):620–6.
- Fan Z, Liu Y, Ye Y, Liao Y. Functional probes for the diagnosis and treatment of infectious diseases. Aggregate. 2024;e620. https://doi.org/10.1002/agt2.620.
- Kim YJ, Han SS, Yang HJ, Chang SE. Prospective, comparative evaluation of a deep neural network and dermoscopy in the diagnosis of onychomycosis. PLoS ONE. 2020;15(6):e0234334. https://doi.org/10.1371/journal.pone.023433 4.
- Mehra M, D'Costa S, D'Mello R, George J, Kalbande DR. Leveraging Deep Learning for Nail Disease Diagnostic, 2021 4th Biennial International Conference on Nascent Technologies in Engineering (ICNTE), NaviMumbai, India, 2021, pp. 1–5. https://doi.org/10.1109/ICNTE51185.2021.9487709
- Yamaç SA, Kuyucuoğlu O, Köseoğlu ŞB, Ulukaya S. Deep Learning Based Classification of Human Nail Diseases Using Color Nail Images, 2022 45th International Conference on Telecommunications and Signal Processing

(TSP), Prague, Czech Republic, 2022, pp. 196–199. https://doi.org/10.1109/TS P55681.2022.9851300

- H MB, Krishnan DAAJ. and K. S D, Automated Detection of skin and nail disorders using Convolutional Neural Networks, 2021 5th International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2021, pp. 1309–1316. https://doi.org/10.1109/ICOEI51242.2021.9452959
- Marulkar S, Narain B. Nail Disease Prediction using a Deep Learning Integrated Framework, 2023 3rd International Conference on Intelligent Technologies (CONIT), Hubli, India, 2023, pp. 1–6. https://doi.org/10.1109/CO NIT59222.2023.10205721
- 12. Coşar Soğukkuyu DY, Ata O. Classification of melanonychia, Beau's lines, and nail clubbing based on nail images and transfer learning techniques, PeerJ Comput. Sci., vol. 9, no. e1533, p. e1533, 2023.
- Nijhawan R, Verma R, Ayushi S, Bhushan R, Dua, Mittal A. An integrated deep learning framework approach for nail disease identification, in 2017 13th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), 2017, pp. 197–202.
- Kumar A, Tiwari KK. Disease Classification in Dermatology: A CNN-RF Hybrid Approach, 2024 3rd International Conference on Artificial Intelligence For Internet of Things (AlIoT), Vellore, India, 2024, pp. 1–5. https://doi.org/10.1109 /AlIoT58432.2024.10574607
- Roy PS, Kukreja V, Nisha Chandran S, Choudhary A. Emperical Analysis of Nail Diseases through Using Hybrid Algorithms of LSTM and CNN, 2024 IEEE International Conference on Computing, Power and Communication Technologies (IC2PCT), Greater Noida, India, 2024, pp. 54–59. https://doi.org/10.1109/I C2PCT60090.2024.10486393
- Pattnayak P, Patnaik S, Sameer S, Rout S, Patra SS. Utilizing Deep Neural Networks for Enhanced Diagnosis of Dermatological Conditions, 2024 International Conference on Inventive Computation Technologies (ICICT), Lalitpur, Nepal, 2024, pp. 260–265. https://doi.org/10.1109/ICICT60155.2024.10544747
- 17. Vishwakarma G, Nandanwar AK, Thakur GS. Optimized vision transformer encoder with cnn for automatic psoriasis disease detection. Multimed Tools Appl. 2023;83:59597–616. https://doi.org/10.1007/s11042-023-16871-z.
- 18. Gurav N. Nail disease detection dataset, Kaggle, 2021. [Online]. Available: http s://www.kaggle.com/datasets/nikhilgurav21/nail-disease-detection-dataset
- Regin R, Gautham Reddy G, Sundar Kumar C, Jaideep CVN. Nail disease detection and classification using deep learning. Cent Asian J. Jun. 2022;3. https://doi.org/10.17605/cajmns.v3i3.828.
- Coşar Soğukkuyu DY, Ata O. Classification of melanonychia, Beau's lines, and nail clubbing based on nail images and transfer learning techniques, PeerJ Comput. Sci., vol. 9, no. e1533, p. e1533, 2023. https://doi.org/10.7717/peerj-c s.1533
- 21. Daniel CR III, Elewski BE. The diagnosis of nail fungus infection revisited. Arch Dermatol. 2000;136(9):1162–4.
- Di Biasi L, De Marco F, Auriemma Citarella A, Barra P, Piotto SP, Tortora G. Hybrid approach for the design of CNNs using genetic algorithms for melanoma classification. Lecture Notes in Computer Science. Cham: Springer Nature Switzerland; 2023. pp. 514–28.
- 23. Winkler JK et al. Mar., Melanoma recognition by a deep learning convolutional neural network—Performance in different melanoma subtypes and

localisations, European Journal of Cancer, vol. 127. Elsevier BV, pp. 21–29, 2020. https://doi.org/10.1016/j.ejca.2019.11.020

- Wang Y, Xu Y, Song J, Liu X, Liu S, Yang N. Tumor Cell-Targeting and Tumor Microenvironment–Responsive nanoplatforms for the Multimodal Imagingguided Photodynamic/Photothermal/Chemodynamic treatment of Cervical Cancer. Int J Nanomed. 2024;19:5837–58. https://doi.org/10.2147/IJN.546604 2.
- Bing P, Liu W, Zhai Z, Li J, Guo Z, Xiang Y, Zhu L. A novel approach for denoising electrocardiogram signals to detect cardiovascular diseases using an efficient hybrid scheme. Front Cardiovasc Med. 2024;11:1277123. https://doi. org/10.3389/fcvm.2024.1277123.
- Roy VK, Thakur V, Nijhawan R. Vision transformer framework approach for yellow nail syndrome disease identification. in Lecture notes in networks and systems. Singapore: Springer Nature Singapore; 2022. pp. 413–25.
- Kyriakou A et al. Fungal Infections and Nail Psoriasis: An Update, Journal of Fungi, vol. 8, no. 2. MDPI AG, p. 154, Feb. 03, 2022. https://doi.org/10.3390/jof 8020154
- Song W, Wang X, Guo Y, Li S, Xia B, Hao A. CenterFormer: a Novel Cluster Center enhanced Transformer for Unconstrained Dental Plaque Segmentation. IEEE Trans Multimedia. 2024;26:10965–78. https://doi.org/10.1109/TMM. 2024.3428349.
- Hussain M, Fiza M, Khalil A, et al. Transfer learning-based quantized deep learning models for nail melanoma classification. Neural Comput Applic. 2023;35:22163–78. https://doi.org/10.1007/s00521-023-08925-y.
- Maithresh A, Nikhil V, Saipuneeth H, Reddy GVS. Exploring the superiority of CapsNet over CNN for early detection of lung cancer A comparative analysis, in 2023 International Conference on Inventive Computation Technologies (ICICT), 2023, pp. 322–329. https://doi.org/10.1109/ICICT57646.2023.1013427
- Li Q, You T, Chen J, Zhang Y, Du C. LI-EMRSQL: linking information enhanced Text2SQL parsing on Complex Electronic Medical Records. IEEE Trans Reliab. 2024;73(2):1280–90. https://doi.org/10.1109/TR.2023.3336330.
- Kruthika KR, Rajeswari, Maheshappa HD. CBIR system using Capsule networks and 3D CNN for Alzheimer's disease diagnosis. Inf Med Unlocked. 2019;14:59–68.
- Li Z, Koban KC, Schenck TL, Giunta RE, Li Q, Sun Y. Artificial Intelligence in Dermatology Image Analysis: current developments and future trends. J Clin Med. 2022;11:6826. https://doi.org/10.3390/jcm11226826. 22. MDPI AG.
- Haenssle HA, Fink C, Schneiderbauer R, Toberer F, Buhl T, Blum A, Kalloo A, Hassen ABH, Thomas L, Enk A, et al. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. Ann Oncol. 2018;29(8):1836–42.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.