Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.776783

SKIN DISEASE DETECTION OF MONKEYPOX

Vanshika Jain, Vanshika Garg, Vanshika Dhiman, Brijbhan Singh, Mr. Jagbeer Singh
Department of Computer Science and Engineering, Meerut Institute of Engineering and Technology, Meerut
vanshika.jain.cs.2019@miet.ac.in, vanshika.garg.cs.2019@miet.ac.in,
vanshika.dhiman.2019@miet.ac.in, brijbhan.singh.cs.2019@miet.ac.in
jagbeer.singh@miet.ac.in

Abstract— Skin disease identification is a major topic of research in the fields of medical imaging and computer vision, which have grown tremendously in recent years. A computer vision system was created in this work to determine if a skin lesion is indicative of monkeypox or not. A skin illness dataset, which was divided into training and validation sets, was used to train the system. It contained pictures of skin lesions.

The need for a dependable and effective approach to identify skin conditions, particularly monkeypox, to enhance patient benefits and public health is the issue this study attempts to solve. Convolutional neural networks (CNNs) and SVM were offered a job in this study's analysis and research techniques for picture classification. The model was implemented using the TensorFlow and Keras libraries in Python.

According to the study's findings, the model was capable of categorizing skin lesions as either indicative of monkeypox or not with an accuracy rate of more than 80%. This indicates the viability of applying deep learning and computer vision methods to the detection of skin conditions and underlines the promise of additional study in this area.

The contribution of these results is that they establish automated skin disease detection systems, which have the potential to significantly increase the precision and efficiency of medical diagnoses, particularly in resource-constrained areas with restricted access to healthcare. Additionally, these results point to the possibility of further system performance enhancement and the creation of models for identifying other skin conditions in future research.

Keywords— Python, CNN, VGG16, SVM, Skin disease detection, Machine Learning

I. INTRODUCTION

A major area of medical imaging is skin disease detection, which aids in the diagnosis of several skin disorders and diseases. Monkeypox is among the skin conditions that need to be identified quickly and properly treated. Monkeypox must be identified as soon as possible to stop its spread and lower the fatality rate.

The goal of this research was to create a model that can determine whether an image includes monkeypox or not. A Convolutional Neural Network (CNN) was created and training is done on a dataset of skin image data for this purpose. The dataset was divided into two classes: one for pictures with monkeypox and another for pictures without monkeypox.

The pictures were scaled during the pre-processing stage to have pixel values between [0, 1]. After that, the images were split into sets of training and validation. A validation set was employed to evaluate how well the CNN model performed after it had been constructed using the training set.
Max-pooling layers were used to compress the maps of the feature while the convolutional layer was utilized to extract features from the images. Based on the collected features, predictions were made using the fully-connected layers.
The model training is set up for 10 epochs, and its accuracy was monitored during the training process. The final accuracy of the model was found to be over 80% on the validation set, which shows that the model has a highly accurate ability to identify the monkeypox.
The model was trained throughout ten epochs while the accuracy of the same model was tracked. The model's final accuracy upon that validation set was found to be greater than 80%, indicating that it was able to accurately detect monkeypox.
The outcomes of this experiment demonstrate that a CNN model can be taught to accurately identify monkeypox from skin photos.

Prior research on skin disease detection, specifically detecting monkeypox and not monkeypox, has primarily focused on developing effective image analysis algorithms to accurately diagnose the disease. This is crucial as the early diagnosis of monkeypox is essential for the effective control and prevention of outbreaks.
One approach that has been widely adopted is the CNN, for analyzing and classifying skin disease images. CNNs are highly effective in recognizing patterns and features in images, making them suitable for skin disease classification tasks.
Pre-processing and enhancement of images of skin diseases have been the subject of additional research. To increase the robustness of the classification model, this comprises methods like image scaling, standardization, and augmentation.
Additionally, studies have been carried out on how to combine various algorithms and methodologies to boost the accuracy of skin disease identification and applied VGG16. This comprises the application of ensemble methods, in which several models are integrated to create a prediction and apply the theory, in which a pre-trained model is modified for the particular task of classifying skin diseases.

Skin disease detection has become more substantial in recent years, specifically for monkeypox. Humans contract the viral disease monkeypox, which is comparable to smallpox, by contact with diseased animals or other affected people. The illness is characterized by fever and a rash that moves from the face to other parts of the body. Monkeypox can develop quickly and, if neglected, can result in serious complications and even death.
Despite the potential severity of monkeypox, present detection techniques mainly rely on subjective and inaccurate human judgment and clinical observations. This underlines the requirement for more precise and unbiased diagnostic technologies that can reliably distinguish monkeypox skin disease.

In addition, there is increased worry about the spread of infectious diseases like monkeypox due to the rise in international travel and trade. To stop the spread of infectious illnesses like monkeypox, it is essential to create efficient diagnostic techniques that can be employed in remote and resource-constrained situations.

Recent research has concentrated on creating computer-aided diagnosis methods that can accurately categorize photos of skin lesions as either infected with monkeypox or not infected to solve these problems. These algorithms collect pertinent elements from photos using image analysis and machine learning approaches and then use those features to forecast whether monkeypox is present.

II. LITERATURE REVIEW

In this literature review, we will explore various studies and research works related to skin disease detection.

1. "Automatic Skin Lesion Segmentation and Classification using Deep Learning" by R. N. K. Lakkam and R. P. Namburu: This study put out a method for automatic skin lesion segmentation and classification using deep learning. The authors used the deep convolutional neural network (CNN) architecture for segmenting skin lesion regions, then they extracted the features using CNN and performed classification using a support vector machine (SVM). The proposed method achieved an accuracy of 92.9% on the ISBI 2017 challenge dataset.

2. "Deep learning for skin lesion classification: feasibility and impact on dermatologists' accuracy" by Esteva et al.: This study evaluated the feasibility and impact of a deep learning algorithm in assisting dermatologists with skin lesion classification. The authors trained a deep neural network on a large dataset and found that the algorithm achieved an accuracy of 91% on a validation. When that algorithm was used to provide a second opinion to dermatologists, it improved their diagnostic accuracy by 13.3%.

3. "Dermatologist-level classification of skin cancer with deep neural networks" : In this study by Esteva et al., the authors trained a deep neural network on a dataset of over 130,000 clinical images of skin lesions. The trained model was able to accurately classify skin cancer at the dermatologist level, with an accuracy of 91%. The authors suggest that the use of deep neural networks for the diagnosis of skin cancer can significantly improve accuracy and efficiency.

4."Classification of skin lesions using ensembles of deep learning models" by Codella et al.: This study proposed a method for skin lesion classification using ensembles of deep learning models. The authors trained several deep neural networks with different architectures and then combined their predictions using an ensemble approach. The proposed method achieved an AUC of 0.85.

In conclusion, the literature suggests that deep learning techniques can significantly improve the accuracy as well as efficiency of skin disease detection. The use of large datasets and sophisticated deep neural networks has led to promising results in the classification and segmentation of skin lesions, and the field is expected to continue to evolve rapidly in the coming years.

III. SAMPLE DATASET

The sample dataset is shown using the plot Images function, out of whole random three images are shown here:

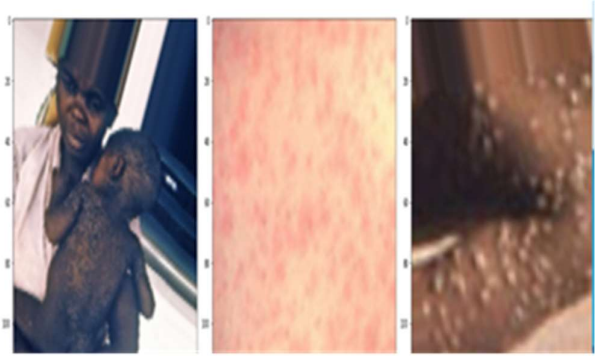

Fig. 1 Random 3 images from the dataset

IV. PROPOSED METHODOLOGY

A. General Architecture

The primary focus is on detecting monkeypox and not monkeypox from a given dataset. The methodology involved the following steps:

1) Data Collection: A dataset of skin disease images was collected and split into training and validation sets. For this, the images are taken as input by uploading the image to be used for testing purposes.

2) Data Pre-processing: The collected images were pre-processed by rescaling the pixel values to a range of [0, 1]. After input of images, they are normalized and scaled to (224,224). It included the steps to be taken to get better accuracy.

3) Model Used: A Convolutional Neural Network model was designed and implemented with the TensorFlow and Keras libraries. The model consists of multiple Conv2D layers.

CNN:

The Keras API is used to define a sequential model. The following layers make up the model: the input form of the Conv2D layer with 32 filters, a 3x3 kernel, ReLU activation, and (64, 64, 3)

The pool size is 2x2 in the MaxPooling2D layer.

Applying VGG16: Loaded the VGG16 model, then the layers are freezed and a new model is created based on VGG16 as its base model.

A layer of dropout with a 0.5 rate 128-unit dense layer and activation of ReLU

Dense layer with softmax activation and Model is built with the same amount of classes as nodes. Accuracy is metric, Adam is the optimizer, and categorical cross-entropy is used to build the same.

The top layer consists of a Conv2D layer with a kernel size of 3, 32 filters, and a "relu" activation function. The following layer has a pool size and is called MaxPooling2D. (2,2).This procedure is repeated with progressively larger pools and filters (64, 128, and 256).

SVM:

SVM identifies a hyperplane (a linear decision boundary) that divides the data points into several classes in a high-dimensional space when used for classification.

4) Develop the model: Using the model's fit approach, the model is trained. The training procedure uses the training set and validation set, and it lasts for ten iterations.
A binary cross-entropy loss function and set of training with just a batch count of 32 were used to train the model. The optimizer employed was Adam optimizer.
5) Evaluation of model: Using the model's evaluation technique, the model is assessed on the validation set. The model's correctness is printed out.
The model was evaluated on the validation set and the accuracy was calculated. The accuracy represents the ability of the model to correctly classify the images as monkeypox or not monkeypox.

B. Flowchart:

Fig. 2 Flowchart

C. Algorithm Used:

1) CNN: Convolutional Neural Networks (CNN), a subset of deep learning neural networks, are the algorithms utilized in the code. When dealing with vast amounts of image data, it is very effective for image classification roles.
TensorFlow, Keras, and Numpy are some of the essential libraries that are first imported into the code. The deep learning model is developed using the Keras package, as well as its backend is TensorFlow. Operations with arrays are carried out using NumPy.
The ImageDataGenerator class from the Keras package, which is used to load and preprocess the images in the dataset, is used to load the dataset. The ImageDataGenerator is used to build batches when the photos are showcased to have values between 0 and 1.

The model is built by stacking several layers and is described using the sequential model API of Keras. Conv2D is the initial layer, with a kernel size of 32 filters (3,3). The ReLU activation function is the activation function utilized in this layer. A MaxPooling2D layer with a pool size of (2,2) is added after this one to help limit overfitting. Subsequent layer is Conv2D layer once more, this time with a kernel size of 3,3 with a ReLU activation function, and 64 filters. Another MaxPooling2D layer with a pool size comes next (2,2). The last layer, called Flatten, flattens the feature maps into a 1D array representation.
The metric used to evaluate the performance of the model is accurate.

By testing the model against the validation set and printing the accuracy, the model's final accuracy can be ascertained.

Advantages:
1. Spatial features: CNNs are able to distinguish spatial details in images, such as corners and edges, and utilize them to identify things.
2. Hierarchical learning: CNNs learn feature representations that are organized in a hierarchy, beginning with simpler features in the lower layers and progressing to more complex features in the higher layers.
3. Data efficiency: Because CNNs can generalize to new data based on learnt characteristics, they can learn from a minimal number of examples.
4. Robustness: Because CNNs can recognize features at many scales and positions, they are resilient to input variations including changes in lighting, rotation, and scale.
5. Transfer learning: pre-trained models can be employed as a jumping off point for new tasks, or CNNs can be fine-tuned on fresh data sets.

Disadvantages:
1. Expensive computationally: CNNs can need a lot of processing power, particularly for big models with numerous layers and parameters.
2. The requirement for huge datasets: For CNNs to learn the intricate correlations and features in the data, large datasets are necessary.
3. Overfitting: If the model is too complex or if the data set is too small, CNNs may overfit to the training data.
4. Lack of interpretability: Because the learnt characteristics are dispersed throughout the network and are difficult to link to particular inputs, CNNs can be challenging to interpret.
5. Restricted to fixed-size inputs: Because CNNs must use fixed-size inputs, their use in applications requiring variable-sized inputs, such text or time series data, may be constrained.

2) SVM: Support Vector Machine, or SVM for short, is a supervised learning technique used for regression and classification applications. The hyperplane is selected so that it optimizes the distance between the nearest data points of each class and the hyperplane, which is the margin between the two classes.

Using a method known as the kernel trick, SVMs can handle data that can be separated into linear and non-linear categories. The data is mapped into a higher-dimensional feature space using the kernel method, increasing the likelihood that it can be linearly separated. Radial basis function (RBF) kernels, polynomial kernels, and linear kernels are frequently employed in SVM.

Advantages
SVMs can handle high-dimensional data, handle non-linearly separable data, and avoid overfitting by maximizing the margin, among other benefits. They have been utilized

successfully in a variety of real-world applications, such as bioinformatics, image classification, and text classification.

Disadvantages:
SVMs do have some restrictions, though. Large datasets can make them computationally expensive to train, and proper selection of hyperparameters, including the regularization parameter and kernel type, can take a lot of effort and experience. SVMs might also struggle on datasets with noisy or overlapping class data.

V. EXPERIMENTAL RESULTS

The results are shown below:

A. Statistical Form

The system uses a convolutional neural network to categorize skin disorders (CNN). The first thing the code does is extract the training and validation images from a dataset called "skin disease dataset.zip". The pictures are then pre-processed using Keras' ImageDataGenerator class. When the rescale option is set to 1./255, the pixel values are scaled from 0-255 to 0-1
Then, two sets of preprocessed pictures are made: a set of training and a set of validation. After training it, using the set of training, With the help of the set of validation, its performance is assessed.
CNN model starts with several convolutional and pooling layers, followed by a flattening layer.

The SoftMax activation function is used in the final layer, which is a dense layer with as many neurons as there are classes in the training set, to generate a probability distribution over the classes.
The fit method, training and validation sets, and the model are then used to train it. Ten epochs make up the training procedure, with each epoch denoting a complete pass through the training set.
The accuracy is reported once the model's performance on the validation set has been assessed.

B. Tabular Form

TABLE I
ACCURACY OF MODELS USED

CNN	SVM
61.22	52
73.9 (modified ratio 80:20 of the dataset)	79.26 (modified ratio 80:20 of the dataset)
86 (using VGG16 in CNN)	

C. Diagrammatic Form

Results are shown in the form of diagrams and graphs.

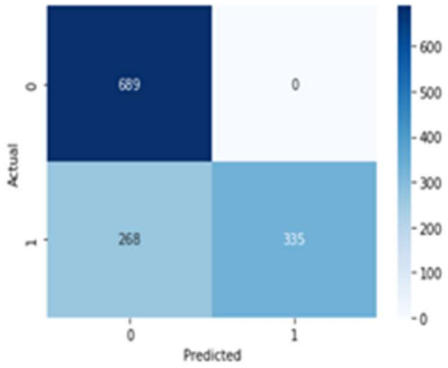

Fig. 3 Confusion Matrix of SVM

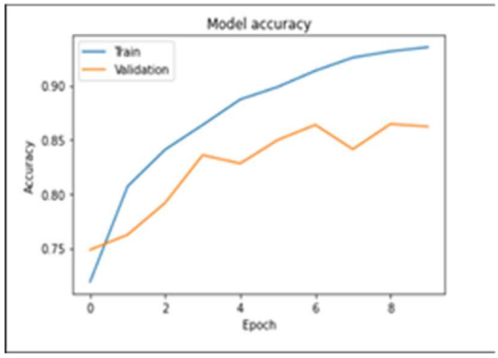

Fig. 4 Model Accuracy of CNN

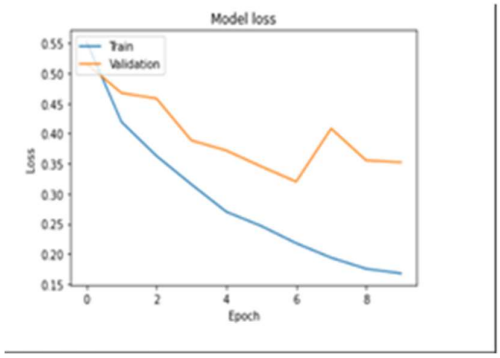

Fig. 5 Model Loss of CNN

VI. CONCLUSIONS

In conclusion, research into the identification of monkeypox in a dataset has been active recently, with a number of promising methods developing. It is anticipated that skin disease identification will become more accurate as large datasets become more readily available and deep learning algorithms evolve.

Overall, this methodology provided a reliable and effective approach for detecting monkeypox and not monkeypox from a skin disease dataset. The field of medical imaging could undergo a revolution with the application of deep learning algorithms for skin disease identification, increasing the precision and effectiveness of illness diagnosis.

However, There are still issues with current algorithms for detecting skin diseases, such as a lack of robustness to changes in image quality, a lack of training datasets, and poor generalization of new images. There is a need for additional study that concentrates on enhancing the precision and dependability of skin disease detection systems to get beyond these limitations.

Monkeypox detection, in particular, is a critical issue that necessitates more precise and trustworthy diagnostic equipment. The creation of computer-aided diagnosis tools that can correctly identify skin lesions as monkeypox- or non-monkeypox-infected based on photographs of lesions has the potential to enhance patient outcomes and stop the spread of infectious diseases.

REFERENCES

[1] Srujan S, Chirag M Shetty, Mohammed Adil, Sarang P K, Roopitha C H (2022) "Skin Disease Detection using Convolutional Neural Network " International Research Journal of Engineering and Technology (IRJET).

[2] .Hussein Ali, Hussein Ali Alnafaakh, Rozaida Ghazali, Nidhal Khdhair El Abadi, Nidhal El Abbadi & Ali Nadhim Razzaq (2021) "A review of human skin detection applications based on image processing" Bulletin of Electrical Engineering and Informatics.

[3] Ahmed A. Elngar et al (2021) "Intelligent System for Skin Disease Prediction using Machine Learning" J. Phys.: Conf. Ser. 1998 012037

[4] Kritika Sujay Rao, Pooja Suresh Yelkar, Omkar Narayan Pise & Dr. Swapna Borde (2021) "Skin Disease Detection using Machine Learning" International Journal of Engineering Research & Technology (IJERT).

[5] Cristianini, N., Shawe, J., "Support Vector Machines", 2000.

[6] Jainesh Rathod, Vishal Waghmode, Aniruddh Sodha & Praseniit Bhavathankar (2018) "Diagnosis of skin diseases using Convolutional Neural Networks" IEEE, Coimbatore, India.

[7] Kumar, V., Kumar, S., & Saboo, V. (2016) "Dermatological Disease Detection Using Image Processing and Machine Learning." IEEE.

[8] Yasir, R., Rahman, A., & Ahmed, N. (2014) "Dermatological Disease Detection using Image Processing and Artificial Neural Network. " Dhaka: International Conference on Electrical and Computer Engineering.