SKIN CONDITION
Diagnosis using ML models

KALYAN CHOWDARY-21BCE9180

MARISARLA YATHEESWAR-21BCE9094

PAVAN SATHISH-21BCExxxx

Under guidance of :Koduru Hajarathaih Reddy

Objectives

The objectives of the project are:

•To develop robust machine learning models for the detection of eczema and psoriasis using a dataset sourced from Kaggle.

•To experiment with three prominent models: Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Convolutional Neural Networks (CNN).

•To fine-tune and optimize these models to achieve the best performance in identifying and classifying skin diseases, specifically focusing on eczema and psoriasis.

•To improve diagnostic accuracy and provide valuable support for early detection and clinical decision-making in dermatology

Problem statements

•Skin diseases, such as eczema and psoriasis, are significant public health concerns that can significantly impact a person’s quality of life.
•Early detection and diagnosis of these conditions are crucial to improving treatment outcomes and reducing the associated healthcare costs.
•Dermatologists often rely on their clinical expertise to diagnose these diseases, which can be time-consuming and subject to human error.
•There is a lack of a large and diverse dataset of skin images for training and testing machine learning models.
•Traditional machine learning algorithms can struggle to handle image-based data due to their high-dimensionality and complex nature.
•Many existing methods focus on detecting a single skin disease, which can be time-consuming and inefficient, especially in clinical settings.

Abstract

The abstract concisely summarizes the study:
•This study leverages machine learning models to predict human skin diseases using a dataset sourced from Kaggle.
•Three models—Support Vector Machines (SVM), K-Nearest Neighbours (KNN), and Convolutional Neural Networks (CNN)—were employed to classify skin diseases based on patient data such as age, skin type, sun exposure, and symptom duration.
•The models were evaluated on various performance metrics, including accuracy, precision, and recall.
•Among the models, CNN demonstrated superior performance in predicting the severity of skin diseases.
•This work highlights the potential of machine learning models to enhance dermatological diagnostics and assist in the early detection and classification of skin conditions.

Literature Review

The literature review provides context and background by discussing previous work in the field:

- The application of machine learning techniques for disease detection and classification has gained substantial attention in recent years, particularly in the field of dermatology.
- Numerous studies have focused on the detection of skin diseases using various algorithms, highlighting the potential of these methods to enhance diagnostic accuracy and efficiency.
- A study by Esteva et al. (2017) demonstrated the efficacy of deep learning models in diagnosing skin cancer, showing that convolutional neural networks could match or even exceed the diagnostic performance of dermatologists.
- Research by Karamizadeh et al. (2020) focused on the automatic detection of eczema using image processing techniques combined with machine learning classifiers such as SVM and KNN.
- Wang et al. (2018) explored the use of CNNs for the classification of psoriasis lesions in dermoscopic images.
- A study by Ali et al. (2021) evaluated SVM, KNN, and CNN models for various skin diseases, concluding that while CNNs generally provided superior performance due to their ability to learn complex patterns, traditional algorithms like SVM and KNN were effective in specific scenarios where computational resources were limited.
- The lack of large, diverse, and well-annotated datasets is a significant barrier, as noted by Hekler et al. (2019).
- Future research should focus on developing more extensive datasets and exploring hybrid models that combine the strengths of both traditional machine learning and deep learning techniques.

Proposed Model

Proposed Model

This section describes the machine learning models used in the study:

•The analysis focuses on processing colored skin images through machine learning models, including Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Convolutional Neural Networks (CNN).

•Specifically, the CNN model is designed with layers that automatically extract relevant features from the input images, such as texture and skin patterns.

•The convolution and max-pooling layers, incorporating ReLU activation functions, help remove non-linearity, allowing for more effective feature extraction.

•After feature extraction, fully connected layers perform the classification task, distinguishing between healthy, eczema-affected, and psoriasis-affected skin.

•Dropout layers are utilized to manage unwanted inputs and prevent overfitting.

Module Description &Algorithm

Module Description & Algorithm

This section explains the data preprocessing steps and the algorithms of the machine learning models used:

•**Data Preprocessing:** Data preprocessing ensures that the data is in the correct shape and quality for training. The study employed data augmentation techniques to enhance the dataset, including creating new data samples from existing ones by applying random transformations, such as flipping, rotating, cropping, and scaling. After the image augmentation process, the dataset is further pre-processed through size normalization, where all images are resized to a standardized dimension suitable for model input.

•**Support Vector Machine (SVM):** SVM is a popular machine learning algorithm known for its effectiveness in classification tasks. SVM operates by finding the hyperplane that best separates data points of different classes with the maximum margin. SVM excels at binary classification but can be extended to multi-class classification problems.

•**Convolutional Neural Networks (CNNs):** CNNs are one of the most widely used deep learning models for image classification and have proven highly effective in medical image analysis.

•**K-Nearest Neighbors (KNN):** KNN is a simple yet effective machine learning algorithm widely used for classification tasks.

•**Compiling and Training the model:** The experiments were conducted on Google Colab, utilizing its powerful computational resources for training the model on the provided skin disease dataset.

Performance Analysis

Performance Analysis

This section presents the results of the model evaluations:

•The table below compares the performance of the models:

•**SVM:** The SVM classifier achieved an overall accuracy of 80%, with a high precision of 100% for class 0 but undefin performance for class 2 due to a lack of true samples.

•**KNN:** The KNN classifier achieved an accuracy of 60%, with strong precision for class 0 but poor recall (60%).

•**CNN:** The CNN classifier demonstrated strong performance with an accuracy of 87.5%.

	Training data	Test data
Model	Accuracy	Accuracy
SVM	63.32	86
KNN	53	73.5
CNN	65	97
Random Forest	28	31.5
Best XGBoost	23	34

Conclusion

Conclusion

•The study concludes that the CNN classifier demonstrated the strongest performance, achieving an accuracy of 97% on the test data.

•The results highlight the potential of machine learning, particularly CNNs, for accurate classification of skin conditions

References

Eczema and Psoriasis Prediction Using Machine Learning:

- **Aswin, T. S., & Hemalatha, S. (2020)**. "Skin Disease Prediction Model Based on Image Processing and Machine Learning Techniques." *Materials Today: Proceedings*, 33, 3657-3660.
    - This paper focuses on machine learning models, including SVM and CNN, for the prediction and classification of skin diseases like eczema and psoriasis.

Support Vector Machines (SVM) in Dermatology:

- **Bennet, K. P., & Campbell, C. (2000)**. "Support Vector Machines: Hype or Hallelujah?" *ACM SIGKDD Explorations Newsletter*, 2(2), 1-13.
    - A foundational paper on SVM, discussing its theoretical foundations and applications, including dermatological predictions.
- **Khan, M. A., Sharif, M., Raza, M., et al. (2021)**. "A Novel Classification of Skin Lesions Using Deep Convolutional Neural Network and Support Vector Machine." *Pattern Recognition Letters*, 141, 128-134.
    - This paper demonstrates SVM's application in skin lesion classification using CNN-extracted features.

References

K-Nearest Neighbors (KNN) for Skin Disease Classification:

- **Saba, T., Mohamed, A. S., & Rehman, A. (2019)**. "Machine Learning Techniques to Detect and Diagnose Skin Disease Using Color and Texture Features." *Journal of Medical Imaging and Health Informatics*, 9(1), 55-64.
    - The paper explores KNN for skin disease classification based on color and texture features.
- **Ahmad, I., Zubair, A., & Ahmed, A. (2017)**. "Skin Disease Classification Using KNN and Random Forest Algorithms." *International Journal of Computer Science and Mobile Computing*, 6(6), 153-160.
    - This research emphasizes the effectiveness of KNN in skin disease classification.