Human Skin Disease Prediction

Using Machine Learning Models:
SVM, KNN, and CNN

By: [Your Name]

Abstract

- This study employs ML models to predict skin diseases using Kaggle-sourced data.
- Three models—SVM, KNN, and CNN—are used to classify diseases like eczema and psoriasis.
- CNN showed superior performance. The study aims to aid dermatologists with accurate, early detection.

Introduction

- • Eczema & Psoriasis are chronic skin diseases.
- • Early detection is vital for better treatment.
- • ML techniques can improve diagnostic accuracy.
- • Challenge: Lack of large annotated datasets.
- • Aim: Build models to detect multiple diseases effectively.

Literature Survey

- • Esteva et al. (2017): CNNs for skin cancer diagnosis.
- • Karamizadeh et al. (2020): SVM/KNN for eczema detection.
- • Wang et al. (2018): CNN for psoriasis classification.
- • Ali et al. (2021): CNN outperforms SVM/KNN.
- • Hekler et al. (2019): Dataset limitations and model explainability issues.

Dataset

- • Source: Kaggle
- • Classes: Healthy, Eczema, Psoriasis
- • Severity levels: No disease, Mild, Moderate, Moderate-Severe, Severe, Very Severe
- • Preprocessing: Augmentation, resizing, normalization

Methodology Overview

- • Models used: SVM, KNN, CNN
- • Data pre-processing: Augmentation, scaling, normalization
- • Feature extraction: Manual (SVM/KNN) vs Automatic (CNN)
- • Model Training: Transfer learning with VGG19, AlexNet, MobileNetV2

CNN Model Architecture

- • Input: 224x224 color images
- • Conv Layers: 32 → 64 → 128 filters (ReLU)
- • MaxPooling: 2x2 after each conv layer
- • Dense Layers: 256 → Dropout → 128 → Softmax
- • Output: 3 classes (Healthy, Eczema, Psoriasis)

SVM Model Summary

- • Works on feature vectors (color, texture)
- • Uses kernels: Linear, RBF
- • Effective for smaller datasets
- • Simple and interpretable
- • Requires tuning hyperparameters

KNN Model Summary

- • Distance-based classification
- • Feature extraction: color histograms, texture, edges
- • Sensitive to feature scaling
- • Best K chosen via cross-validation

Model Training & Compilation

- • Platform: Google Colab
- • Optimizer: Adam (lr=0.00001)
- • Loss: Categorical Cross-Entropy
- • Metrics: Accuracy
- • Trained for 50 epochs using transfer learning

Results: Accuracy Comparison

- SVM: Train - 63.32%, Test - 80%
- KNN: Train - 53%, Test - 60%
- CNN: Train - 65%, Test - 88%
- CNN performed best overall

Conclusion & Future Work

- • CNN showed the best performance (88% accuracy)
- • SVM effective but limited by class imbalance
- • KNN struggled with minority class detection
- • Future: Dataset balancing, hybrid models, mobile integration

References

- • Esteva et al. (2017), Nature
- • Karamizadeh et al. (2020), Materials Today
- • Wang et al. (2018), CNN for Psoriasis
- • Ali et al. (2021), Comparative ML
- • Hekler et al. (2019), Dataset Challenges