DERMA AI

Abhishek
Yash Raj
Aman Kudiyal
Parveen Kumar
Group 69
ML-mid sem project

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI

Motivation

Skin health is an essential component of overall well-being, yet diagnosing skin conditions remains a challenge. Conditions such as acne, eczema, or life-threatening diseases like melanoma require early detection to minimize complications. Traditional diagnostic methods rely heavily on manual visual inspections by dermatologists, which can be subjective, time-consuming, and susceptible to errors.

With advancements in machine learning, there is a growing need for automated systems to improve the accuracy, speed, and accessibility of skin condition diagnosis. These systems are especially critical in regions with limited access to specialized dermatological care, where early intervention can significantly enhance patient outcomes.

Derma AI seeks to address this gap by leveraging state-of-the-art deep learning techniques, particularly **Convolutional Neural Networks (CNNs)**. The project aims to create a scalable solution capable of identifying diverse skin conditions with precision. By reducing diagnostic errors and enabling faster, more reliable assessments, **Derma AI** holds the potential to transform dermatological care, improving patient outcomes and operational efficiency on a global scale.

Literature review

1. Multimodal Skin Lesion Classification Using Deep Learning

- **Yap et al. (2018)** introduced a multimodal approach that combines **dermatoscopic and macroscopic images** with patient metadata for classifying skin lesions into five disease categories.
- Utilizing **ResNet-50** for feature extraction, the study employed late fusion to integrate image features with metadata.

Key Findings:

- Combining image types and metadata improved performance for melanoma detection (**AUC = 0.866**) compared to using dermatoscopic images alone (**AUC = 0.831**).
- Dermatoscopic images consistently outperformed macroscopic images, and metadata provided marginal gains in complex multi class tasks (**mAP = 0.729**).

Implication: Highlights the potential of **multimodal data integration** to enhance classification accuracy while emphasizing the strength of dermatoscopic images for precise lesion detection.

Literature review

2. Artificial Intelligence in Dermatology: Advancements and Challenges in Skin of Color (SOC)

- This paper explores the underrepresentation of **Skin of Color (SOC)** in dermatology datasets, which impacts the diagnostic accuracy of AI models.
- Traditional scales like the Fitzpatrick Skin Phototype (FST) are criticized for insufficiently representing darker skin tones. The **Monk Skin Tone (MST) scale** is proposed as a more inclusive alternative.
- Challenges Identified:
  - **Dataset Bias:** AI tools like VisualDx and MelaFind underperform in diagnosing SOC due to biased datasets.
  - **Image Quality:** Variations in lighting and overexposure affect model accuracy, especially for darker skin tones.

Proposed Solutions:

- Standardizing image capture protocols with tools like **CLEAR Derm**.
- Incorporating **diverse skin tone datasets** to mitigate bias and improve equitable dermatological care.

Dataset description

Dataset Overview

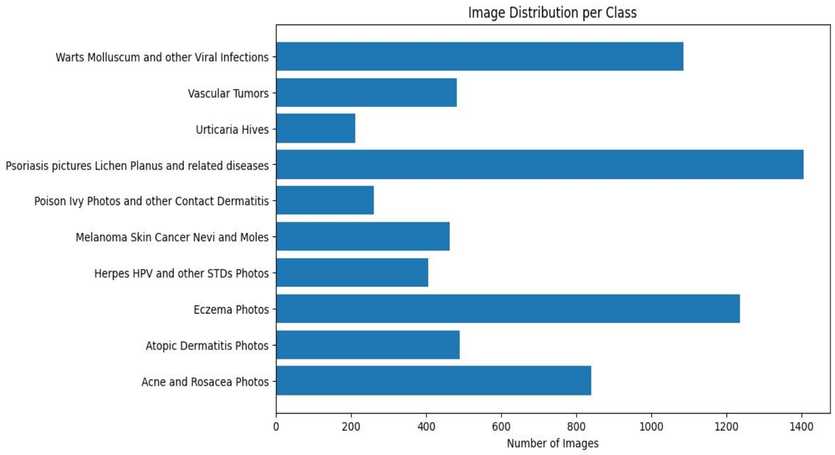
- **Skin Conditions:** The dataset comprises images from **10 distinct skin condition classes**, including:
  - Common conditions: **Eczema**, **Acne**, **Psoriasis**
  - Severe conditions: **Melanoma**, **Skin Cancer**
- **Image Count:**
  - **Training Set:** 6,877 images
  - **Validation Set:** 1,743 images
- **Source:** Data sourced from reputable dermatology repositories like the **ISIC archive**, with variability in skin tones, lighting, and image quality.

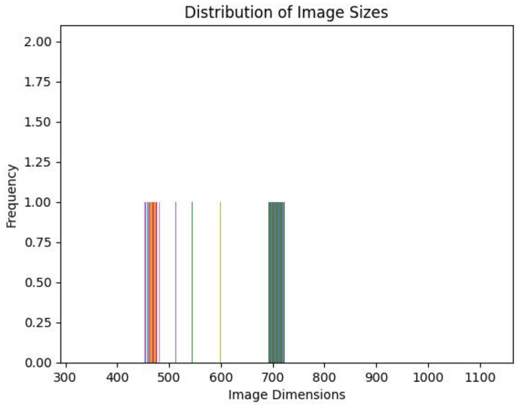
Class Distribution

- **Class Imbalance:**
  - Significant disparity in the number of images across classes. Some conditions like **Eczema** are overrepresented, while rare conditions like **Melanoma** are underrepresented.
  - **Data Augmentation** and **resampling techniques** were applied to ensure balanced representation during training.

Dataset description

- Significant variation exists in the number of images per class.
- The dataset consists of ten classes of skin conditions, with the following number of images per class:

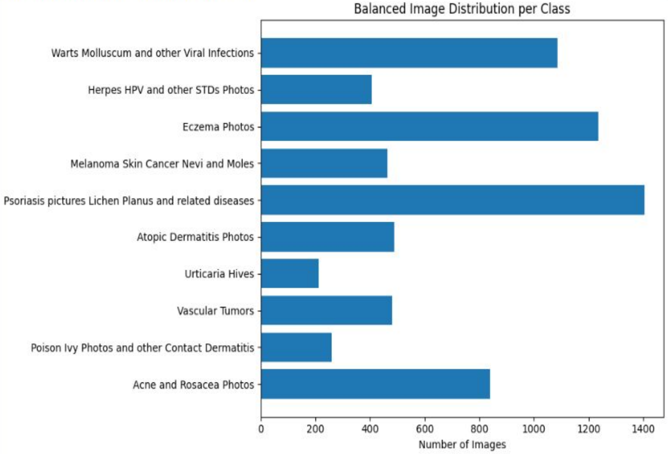




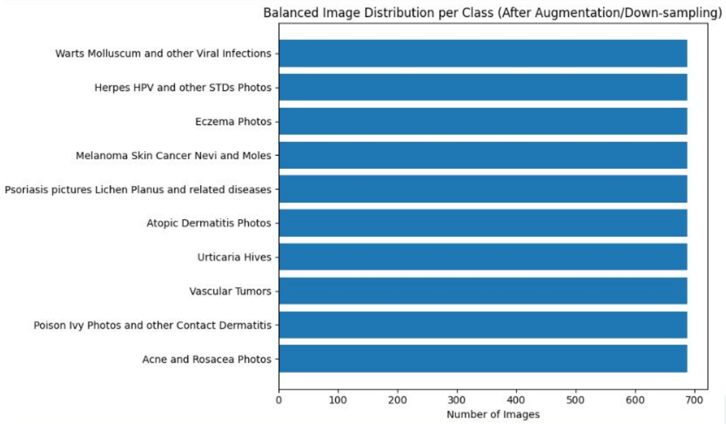
- The total training set contains 6,877 images.

Dataset description

Unbalanced Dataset



Balanced Dataset



Dataset description

Preprocessing

- **Image Resizing:** All images resized to **224x224 pixels** for consistency across models.
- **Normalization:** Pixel values scaled to the range [0, 1] to improve model stability.
- **Brightness & Contrast Adjustment:** Enhanced image quality for better generalization under varied lighting conditions.
- **Data Augmentation:** Applied transformations such as **rotation**, **zooming**, **flipping**, and **cropping** to increase variability, especially for underrepresented classes.

Visual Distribution

- **Class-wise Visualization:** Illustrations of the class distribution and the impact of preprocessing techniques on balancing the dataset.

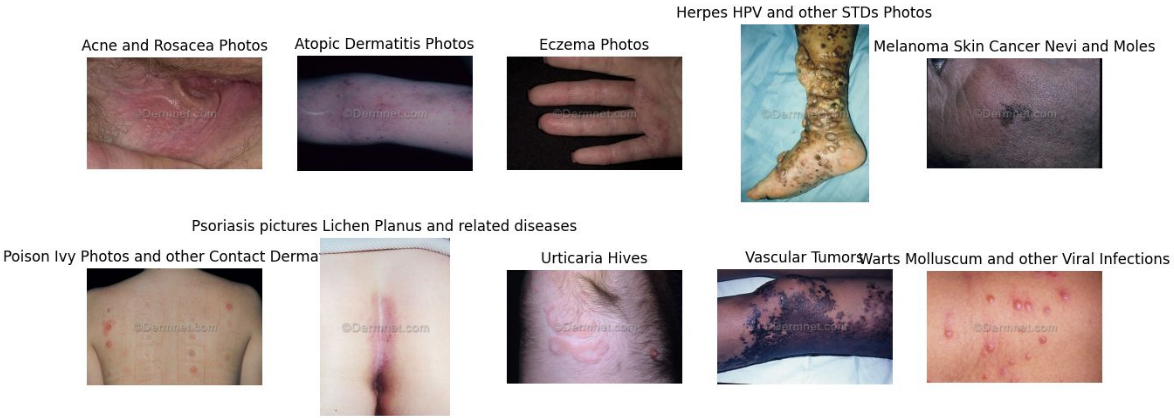
Dermnet

ISIC Challenge

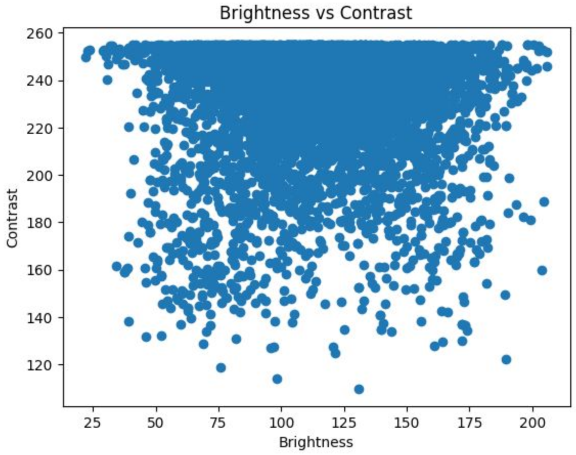
Skin Disease Classification Dataset - Mendeley Data

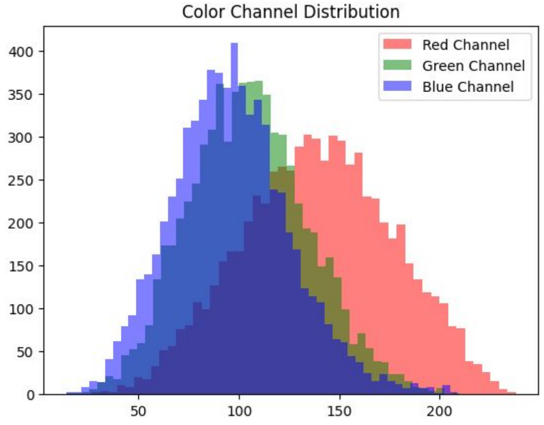
Dataset Description

- The validation set contains 1,743 images.
- There is considerable class imbalance, with some classes significantly underrepresented.
- The class imbalance could lead to biased model performance, favoring the more represented classes.
- This imbalance may result in poorer generalization for underrepresented conditions.
- Addressing the imbalance using data augmentation or weighted loss functions could improve model generalization across all classes.



Dataset Description





Color Distribution:

- Strong red tones (peak at 250-300)
- Dominant blue presence (peak around 400)
- Green channel moderately present
- Clear separation between red and blue channels

The image appears to have good color separation, balanced exposure, and strong contrast, suggesting quality image capture with distinct color elements.

Methodology

Approach Overview :

The methodology for **Derma AI** includes two primary techniques for skin condition classification:

1. The methodology for Derma AI includes two primary techniques for skin condition classification:
2. Convolutional Neural Networks **(CNNs) and its variations**
3. HOG Feature Extraction with Random Forest Classifier
4. Pre-Trained Models for Transfer Learning Data Augmentation & Preprocessing
5. Evaluation Metrics

Methodology

1. Convolutional Neural Network (CNN) Approach

- **Architecture:**
  - Sequential CNN model with layers:
    - **Convolutional Layers:** 32, 64, 128 filters, 3x3 kernel sizes, **ReLU activation**.
    - **MaxPooling Layers:** 2x2 pooling applied after each convolution.
    - **Fully Connected Layer:** Flattened output with 128 units.
    - **Output Layer:** Softmax activation for **multi-class classification**.
- **Training:**
  - **Optimizer:** Adam
  - **Loss Function:** Categorical Cross-Entropy
  - **Epochs:** 20, with **early stopping** to prevent overfitting.

Methodology

2. HOG Feature Extraction + Random Forest Classifier Approach

- **Feature Extraction:**
  - **Histogram of Oriented Gradients (HOG)** used:
    - 9 orientations, 8x8 pixels per cell, 2x2 cells per block.
    - Combined HOG features with **color histograms** to create feature vectors.

- **Classifier:**
  - **Random Forest Classifier** with 100 estimators.
  - **LabelEncoder** used for encoding disease categories.

- **Data Split:** 80% training, 20% testing.

- **Evaluation: Accuracy**, **Precision**, **Recall**, **F1-score** metrics.

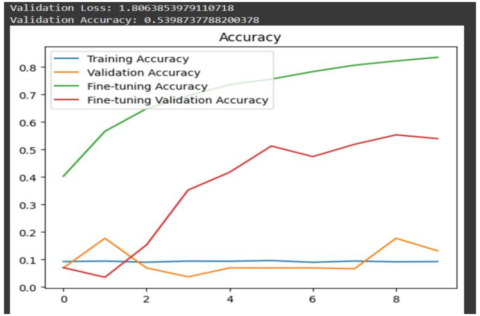
Methodology

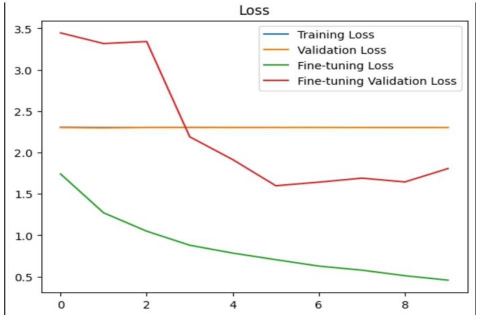
3. Pre-Trained Models for Transfer Learning :

- **Models Evaluated:**
  - **ResNet-50**: Best-performing model with **training accuracy of 89.59%** and **validation accuracy of 68.38%**.
  - **VGG16**: Validation accuracy of **38.64%**, moderate performance.
  - **EfficientNetB0**: Validation accuracy of **53.98%**.
  - **MobileNetV2**: Validation accuracy of **16.01%**.
  - **MobileNetV3Small**: Low validation accuracy of **7.06%**.
  - **NASNetMobile**: Poor performance with validation accuracy of **15.72%**.

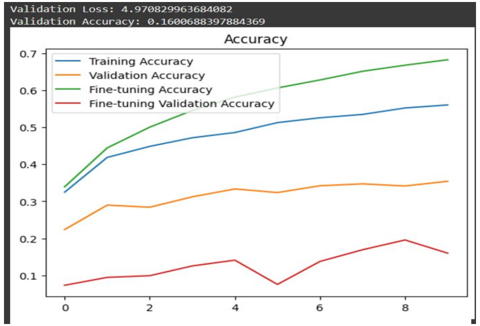
Results

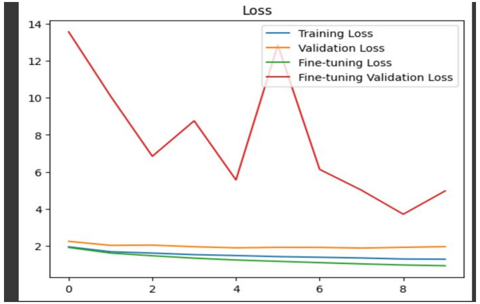
EfficientNetB0



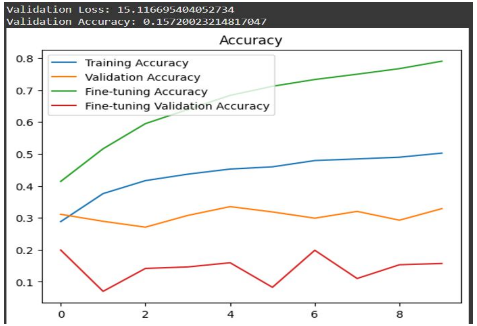


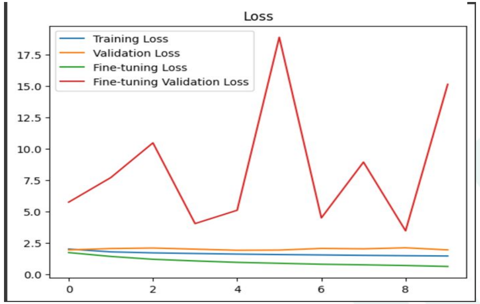
MobileNetV2





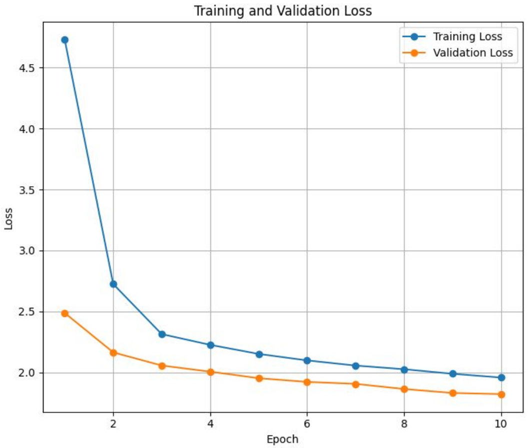
NASNetMobile





Results

VGG16

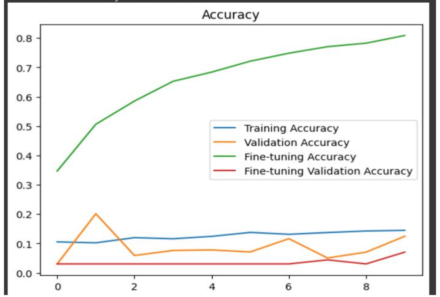


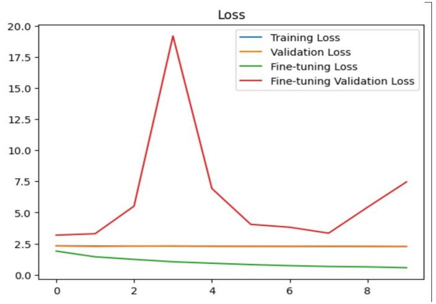

```
Epoch 4: val_loss improved from 2.05729 to 2.00680, saving model to best_vgg16_model.keras
198/198 ━━━━━━━━━━━━━━━━━━━━ 959s 5s/step - accuracy: 0.2139 - loss: 2.2428 - val_accuracy: 0.3112 - val_loss: 2.0068 - learning_rate: 1.0000e-04
Epoch 5/10
198/198 ━━━━━━━━━━━━━━━━━━━━ 0s 4s/step - accuracy: 0.2448 - loss: 2.1341Epoch 5: loss = 2.1523, accuracy = 0.2393, val_loss = 1.9534, val_accuracy = 0.3321

Epoch 5: val_loss improved from 2.00680 to 1.95340, saving model to best_vgg16_model.keras
198/198 ━━━━━━━━━━━━━━━━━━━━ 931s 5s/step - accuracy: 0.2448 - loss: 2.1342 - val_accuracy: 0.3321 - val_loss: 1.9534 - learning_rate: 1.0000e-04
...

Epoch 10: val_loss improved from 1.83219 to 1.82293, saving model to best_vgg16_model.keras
198/198 ━━━━━━━━━━━━━━━━━━━━ 940s 5s/step - accuracy: 0.3054 - loss: 1.9805 - val_accuracy: 0.3864 - val_loss: 1.8229 - learning_rate: 1.0000e-04
Restoring model weights from the end of the best epoch: 10.
```

Results

MobileNetV3Small

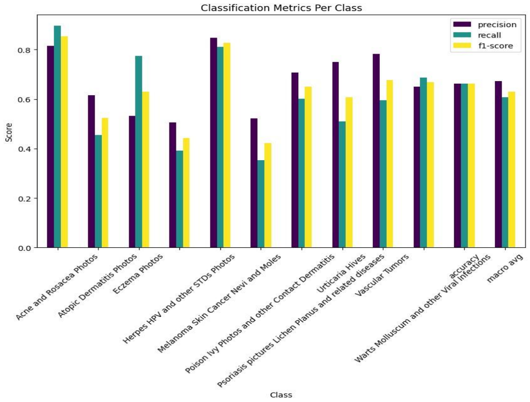
```
Validation Loss: 7.453986644744873
Validation Accuracy: 0.07056798785924911
```

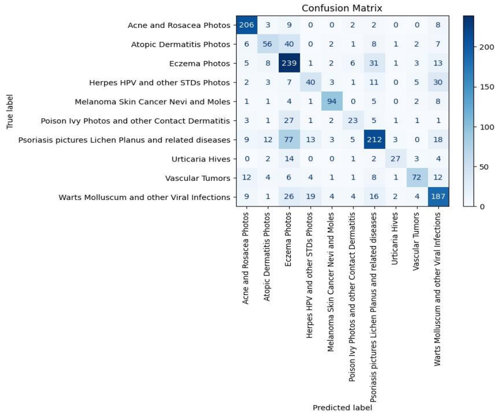




ResNet

```
WARNING:absl:Compiled the loaded model, but the compiled metrics have yet to be built. `model.compile_metrics` will be empty until you train or evaluate
Model loaded successfully!
Found 1743 images belonging to 10 classes.
C:\Users\99710\AppData\Local\Packages\PythonSoftwareFoundation.Python.3.9_qbz5n2kfra8p0\LocalCache\local-packages\Python39\site-packages\keras\src\trair
  self._warn_if_super_not_called()
55/55 ━━━━━━━━━━━━━━━━━━━━ 109s 2s/step - accuracy: 0.7273 - loss: 1.1370
Test Loss: 1.4602503776550293
Test Accuracy: 0.6632243394851685
55/55 ━━━━━━━━━━━━━━━━━━━━ 106s 2s/step
Predicted Classes:  [0 1 0 ... 9 9 9]
Actual Classes:  [0 0 0 ... 9 9 9]
```





Results/Analysis/conclusion

Results

- **Best Model: ResNet-50**
  - **Training Accuracy:** 89.59%
  - **Validation Accuracy:** 68.38%
  - ResNet-50 outperformed other models, proving highly effective for complex dermatological image classification.
- **Other Models Performance:**
  - **SVM:** Accuracy: 33.87%, ROC AUC: 0.7483
  - **KNN:** Accuracy: 19.70%, ROC AUC: 0.6016
  - **MobileNetV2 and MobileNetV3Small:** Low performance with validation accuracy below 20%, highlighting challenges with lightweight models.

Class-wise Evaluation:

- **High Precision:** For classes like **Eczema** and **Psoriasis**, with accurate predictions.
- **Challenges:** Misclassifications in **Herpes HPV**, **Vascular Tumors**, and **Melanoma**, indicating difficulties in distinguishing these categories due to overlapping features and class imbalance.

Results/Analysis/conclusion

Analysis

- **Confusion Matrix:**
  - Diagonal entries dominate, showing true positive predictions for most classes.
  - Some confusion in rare categories like **Vascular Tumors** and **Melanoma**, likely due to limited data and feature overlap.
- **Key Metrics:**
  - **Precision & Recall:**
    - High precision for balanced classes but lower recall for underrepresented ones like **Herpes HPV** and **Vascular Tumors**.
  - **F1-Score:**
    - Balanced classes like **Eczema** showed strong F1-scores, but rare classes like **Herpes HPV** had lower scores, reflecting precision-recall imbalance.

Model Challenges:

Class Imbalance:

- Underrepresented classes performed poorly due to limited training data.
- **Overfitting & Underfitting:**
  - **ResNet-50** showed signs of overfitting, with high training accuracy but a gap in validation accuracy.

Timeline

Phase	Proposed Dates	Planned Activities	Actual Completion
Proposal Submission	27 Aug 2024	Submit project proposal	**Completed on time**
Exploring Project Scope	3 Sep - 10 Sep 2024	Literature review, dataset collection, and initial preprocessing	**Completed on time**
Data Preprocessing & Analysis	10 Sep - 1 Oct 2024	Dataset exploration, cleaning, augmentation, balancing, and splitting	**Completed by 5 Oct 2024**
Model Training & Development	1 Oct - 5 Nov 2024	Model selection, training, and fine-tuning (CNN, ResNet-50, SVM, etc.)	**Completed on time**
Final Model Evaluation	6 Nov - 20 Nov 2024	Model performance evaluation, confusion matrix analysis, tuning	**Completed on time**
Report and Presentation	21 Nov - 29 Nov 2024	Final report writing, presentation creation, results analysis	**On track**

Individual team members' contributions

Team Members and Contributions

1. **Aman Kudiyal :**
    - **Traditional Machine Learning Models:**
        - Implemented and evaluated **Support Vector Machine (SVM)** and **K-Nearest Neighbors (KNN)**, achieving accuracy of 33.87% and 19.70%, respectively.
        - Trained and evaluated the **Basic CNN model** with training accuracy of 26.77% and validation accuracy of 30.89%.
    - **Contributions to Report and Documentation:**
        - Prepared and documented methodologies for model training and results analysis.

Individual team members' contributions

Abhishek :

- **Deep Learning Model Development:**
  - Implemented and fine-tuned **ResNet-50**, achieving **89.59% training accuracy** and **68.38% validation accuracy**.
- **Confusion Matrix Analysis:**
  - Performed in-depth analysis of the confusion matrix, highlighting misclassifications and challenges with underrepresented classes like **Herpes HPV** and **Vascular Tumors**.
- **Recommendations for Model Improvement:**
  - Suggested data augmentation, class balancing, and hyperparameter tuning to address misclassifications.

Individual team members' contributions

Parveen Kumar

- **Experiments with Lightweight Models:**
    - Evaluated and trained models such as **MobileNetV2**, **EfficientNetB0**, and **NASNetMobile**, providing performance insights on their suitability for skin condition classification.
    - Applied **hyperparameter tuning** to **MobileNetV3Small**, although performance remained suboptimal.

- **Data Augmentation and Preprocessing:**
    - Contributed significantly to **data augmentation** strategies (rotation, zoom, flipping) to improve performance on underrepresented classes.
    - Assisted in enhancing image quality for better model performance.

Individual team members' contributions

Yash Raj Ojha :

- **Visualization and Analysis:**
    - Created **data visualizations** such as class distributions, model comparisons, and training curves, which helped in better interpreting model performance.

- **Report Consolidation and Presentation Preparation:**
    - Worked on integrating team members' contributions into a cohesive **final report**.
    - Designed the **presentation slides** and ensured alignment with project goals, focusing on clarity and key insights.

- **Collaborative Analysis:**
    - Contributed to interpreting the model evaluation results and recommended improvements for future iterations of the model.

Thank You