

Project Title: MelanoAI: Intelligent Skin Cancer Screening

Objective:

Build an Artificial Intelligence model to accurately classify skin lesions as benign or malignant, and provide a user-friendly interface for interaction.

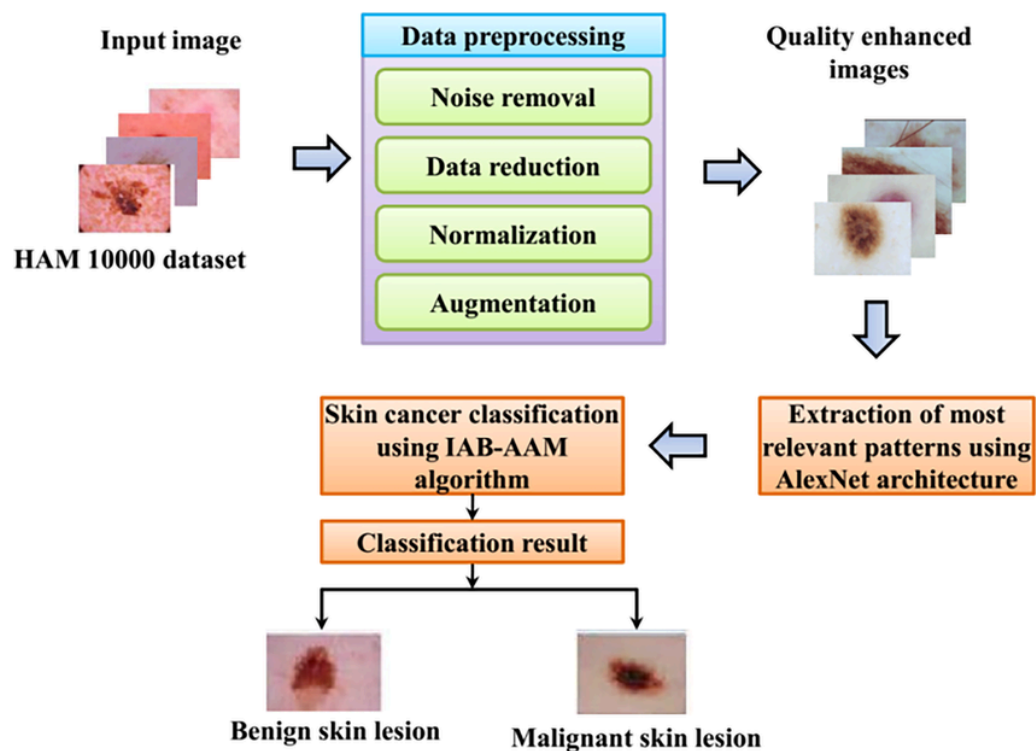
Dataset:

- **Dataset:** A dataset containing images of skin lesions and their corresponding labels (benign or malignant).

Project Workflow:

1. Data Collection and Exploration
2. Exploratory Data Analysis (EDA)
3. Data Preprocessing
4. Model Building
5. Model Evaluation
6. Deployment and Documentation

Architecture Diagram



Components

User Interface

- **Image Upload:** Allows users to upload diagnostic images.
- **Prediction Display:** Shows the malignancy probability prediction.

Data Processing Pipeline

- **Data Preprocessing:** The uploaded images undergo preprocessing steps such as resizing and normalization to prepare them for classification.
- **Model Training and Optimization:** The preprocessed images are used to train and optimize the AI model for image recognition.

Storage and Model Deployment

- **Storage:** The trained model and any necessary data are stored in a database or file storage system.

Detailed Plan:

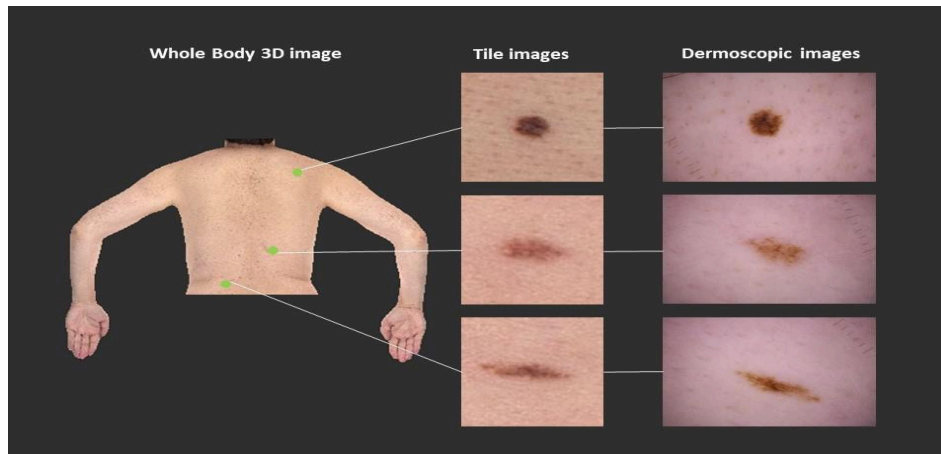
Milestone 1: Weeks 1-2

Data Collection and Exploration

- **Understand the Problem Statement**
 - Define the classification task for skin cancer detection.
 - Understand the features and labels available in the dataset.
- **Collect Data**
 - Download and load the dataset.
 - Familiarize with the structure and content of the dataset.
- **Initial Data Exploration**
 - Examine the dataset for initial insights.
 - Identify the types of features (image data, metadata).

Exploratory Data Analysis (EDA)

- **Image Analysis**
 - Analyze the images to understand their quality and characteristics.
 - Visualize the distribution of benign vs. malignant lesions.
- **Metadata Analysis**
 - Explore any additional metadata provided.
 - Identify any patterns or correlations with the labels.
- **Visualization**
 - Use image visualization tools to examine sample images.
 - Summarize insights from the EDA.



Milestone 2: Weeks 3-4

Data Preprocessing

- **Image Preprocessing**
 - Resize images to a consistent size.
 - Normalize pixel values.
- **Data Augmentation**
 - Apply techniques such as rotation, flipping, and zooming to augment the dataset.
- **Handling Missing Values**
 - Identify and treat any missing values in the metadata.
- **Feature Engineering**
 - Extract relevant features from the images (if applicable).

UI Development

- Begin UI development focusing on image upload functionality and basic layout.
- Develop components for image upload and display, considering user experience and design.

Milestone 3: Weeks 5-6

Model Building

- **Split Data**
 - Split the dataset into training and testing sets.
- **Model Selection**
 - Explore various deep learning architectures (e.g., CNN, ResNet, EfficientNet).
- **Model Training**
 - Train multiple models on the training data.
 - Use cross-validation to evaluate models.
- **Hyperparameter Tuning**
 - Use techniques like grid search or random search to tune hyperparameters for the best-performing models.

Model Evaluation

- **Evaluate Models**
 - Use metrics such as accuracy, precision, recall, F1-score, and ROC-AUC.
- **Compare Models**
 - Compare the performance of different models.
 - Select the best model based on evaluation metrics.
- **Final Model Training**
 - Retrain the best model on the entire training set.

Milestone 4: Weeks 7-8

Deployment and Documentation

- **Model Deployment**
 - Deploy the final model using a web framework like Flask or Django.
 - Create an API for the skin cancer detection model.
- **Presentation and Documentation**
 - Prepare a presentation detailing the problem statement, data collection, preprocessing steps, model building, and evaluation.
 - Document the project comprehensively, including code, methodologies, and findings.
- **GitHub Submission**
 - Upload the final code, documentation, and presentation to GitHub.

Milestone Evaluation:

Milestone 1: Week 2

- Understanding of the problem statement and dataset.
- Initial data exploration and insights.
- Completed EDA with visualizations.

Milestone 2: Week 4

- Completed data preprocessing.
- Data augmentation and handling of missing values.

Milestone 3: Week 6

- Model building and initial training results.
- Selection of the best model based on evaluation metrics.

Milestone 4: Week 8

- Deployment of the model.

- Comprehensive project documentation and final presentation.