Project Requirements

Project Title: Oral Cancer Detection Web App using AI Model Trained on Kaggle Multi-Cancer Dataset

Project Overview:

The project aims to develop a web application that utilizes an AI model trained on the Kaggle multi-cancer dataset to detect oral cancer cells in input images. The AI model will be trained for classification, distinguishing between oral cancer cells and cells from other types of cancers. The web app will provide a user-friendly interface for users to upload an image and receive predictions regarding the presence of oral cancer cells.

Project Components:

1**. Dataset Collection and Preparation:**

- Utilize the Kaggle multi-cancer dataset, specifically focusing on oral cancer-related images.

- Preprocess the images, ensuring uniform size, quality, and annotations relevant to oral cancer.

2. **AI Model Training:**

- Develop a deep learning model for image classification using convolutional neural networks (CNNs).

- Utilize the preprocessed dataset to train the model to differentiate between oral cancer cells and other cancer cells.

- Employ techniques to enhance model performance and reduce overfitting, such as data augmentation, transfer learning, and hyperparameter tuning.

3**. Model Evaluation:**

- Evaluate the trained model using appropriate metrics like accuracy, precision, recall, and F1-score to assess its performance.

4. **Model Deployment:**

- Choose a suitable platform or service for deploying the trained AI model. Options could include cloud platforms like AWS, Google Cloud, or Microsoft Azure.

- Prepare the model for deployment and integration within the web application.

5. **Web App Development:**

- Create a web application using a web development framework such as Flask, Django, or Node.js.

- Design an intuitive and user-friendly interface allowing users to upload an image for analysis.

- Integrate the trained model into the web app to provide real-time predictions based on the uploaded image.

6. **User Authentication and Data Privacy:**

- Implement a secure user authentication system to safeguard user data and ensure privacy.

- Comply with applicable data protection regulations to maintain the privacy and confidentiality of user information.

7. **Documentation:**

- Create comprehensive documentation detailing the project, including an overview, methodology, model architecture, deployment instructions, and how to use the web app.

- Provide a step-by-step guide on replicating the project and training the model.

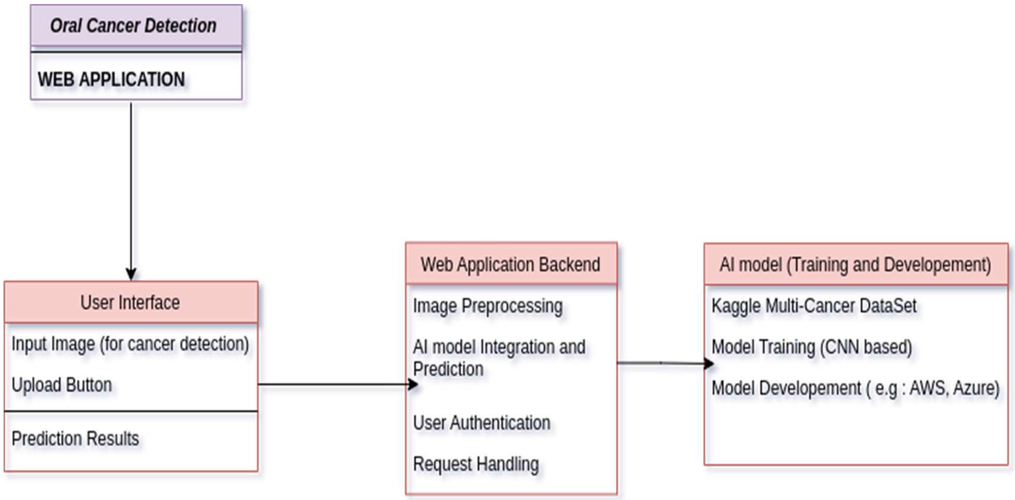
Expected Outcomes:

- A functioning web application that enables users to upload images of cells for oral cancer detection.
- An AI model capable of accurately classifying the uploaded images as either oral cancer cells or cells from other types of cancers.
- Documentation that allows for the replication of the project and further improvements.

This project aims to contribute to the early detection of oral cancer through an accessible and user-friendly web application, ultimately assisting in timely medical intervention and improved patient outcomes.

Context Diagram for Oral Cancer Detection Web App

The context diagram illustrates the interactions between the system and its external entities, detailing the inputs, outputs, and high-level components of the Oral Cancer Detection Web App.


Oral Cancer Detection
WEB APPLICATION
User Interface
Input Image (for cancer detection)
Upload Button
Prediction Results
Web Application Backend
Image Preprocessing
AI model Integration and Prediction
User Authentication
Request Handling
AI model (Training and Developement)
Kaggle Multi-Cancer DataSet
Model Training (CNN based)
Model Developement (e.g : AWS, Azure)

High-Level Components:

1. **User Interface:**

- **Inputs:**

  - Image of cells for oral cancer detection uploaded by the user.
- **Outputs**:
  - Displayed prediction results indicating the presence of oral cancer cells.

2. **Web Application Backend**:

- **Components**:
  - Image Preprocessing: Processes the uploaded image for input to the AI model.
  - AI Model Integration and Prediction: Integrates the trained AI model and makes predictions.
  - User Authentication: Ensures secure user access and data privacy.
  - Request Handling: Manages and handles user requests and interactions.

3. **AI Model and Model Training/Deployment:**

- **Components**:
  - Kaggle Multi-Cancer Dataset: Provides the training data for the AI model.
  - Model Training (CNN-based): Trains the AI model to classify oral cancer cells using the dataset.
  - Model Deployment: Deploys the trained model to make real-time predictions in the web application.

The system works by allowing users to upload images of cells via the web interface. The uploaded image is preprocessed and then passed to the AI model for prediction. The model determines whether the cells are oral cancer cells or not, and the prediction results are displayed to the user in the interface. The AI model's predictions are based on the training it received using the Kaggle Multi-Cancer Dataset, and the model is deployed and integrated into the web application for real-time usage. User authentication ensures secure usage of the application.

System Requirements:

Functional Requirements:

1. **User Authentication:**

- Steps:

1. User opens the web app.
2. User is prompted to log in or create an account.
3. User enters login credentials or creates a new account.
4. System validates the credentials and authenticates the user.
5. User gains access to the app's functionalities upon successful authentication.

2. **Upload Image for Detection:**

- Steps:

1. User navigates to the "Upload Image" section.
2. User selects an image from their device.
3. User confirms the image selection.
4. System processes the image for AI model input.

3. **Predict Oral Cancer:**

- Steps:

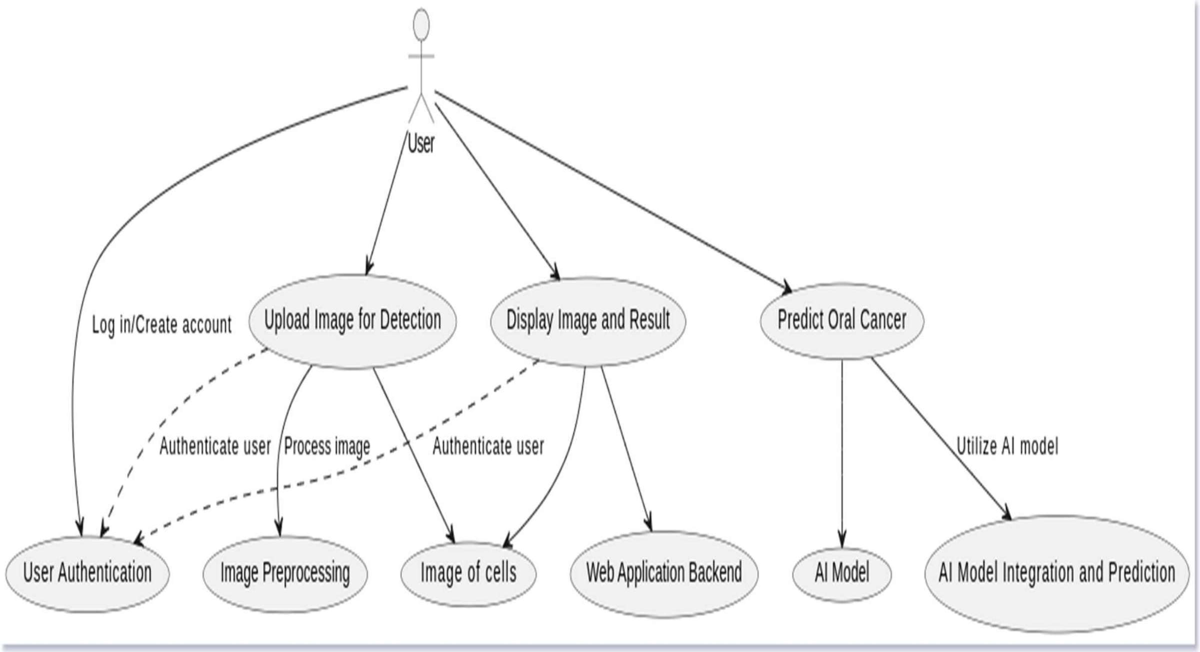
1. User initiates the prediction process.
2. System utilizes the AI model to predict oral cancer based on the preprocessed image.
3. System displays the prediction result (presence or absence of oral cancer cells).

4**. Display Image and Result**:

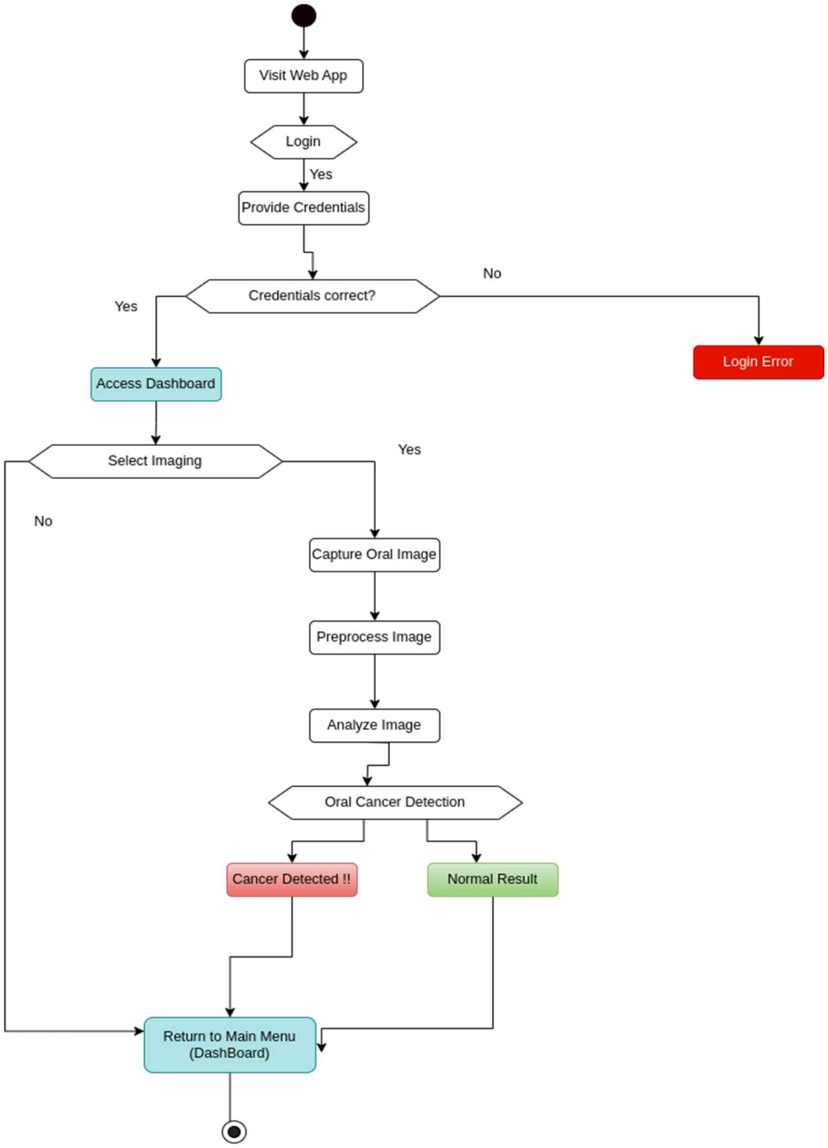
- Steps:

1. System displays the uploaded image to the user.
2. System displays the prediction result (presence or absence of oral cancer cells) alongside the image.

Use Case Diagram:



Activity Diagrams:



Non-Functional Requirements:

Usability Related Requirements:

1. **Responsive Design:**

   - The web application must provide a seamless experience across various devices, adapting to different screen sizes and resolutions.

2. **Intuitive User Interface:**

   - The user interface should be intuitive, requiring a minimal number of steps for users to achieve their objectives.

3. **Cross-Browser Compatibility:**

- The application should be compatible with major web browsers to ensure a consistent experience regardless of the user's browser choice.

Performance Related Requirements:

1. **Response Time:**

- The system should respond to user interactions within a maximum of 3 seconds, maintaining an efficient user experience.

2**. Image Processing Time:**

- Image preprocessing and model prediction should be completed within an acceptable timeframe to ensure timely results.

Scalability Related Requirements:

1. **Concurrent User Support:**

- The system should handle a good number of concurrent users while maintaining performance and response times.

2. **Scalability for Increased Data:**

- The system should scale efficiently as the dataset size or user base grows to maintain consistent performance.

Security Related Requirements:

1. **Data Encryption:**

- All data transmitted between the client and server should be encrypted using strong encryption algorithms and protocols.

2. **Authentication and Authorization:**

- Users should be required to authenticate before accessing the system, and appropriate authorization mechanisms should be in place.

Project Deliverables:

Software Development and Deployment Deliverables:

1. **Web Application:**

- The functional and visually appealing web application that allows users to upload images for oral cancer detection and view the prediction results.

2. **AI Model Integration:**

- The integrated AI model for oral cancer prediction within the web application.

3. **Model Training and Deployment:**

- Trained AI model for oral cancer detection and its deployment for real-time prediction.

4. **Database and Data Storage:**

- Database setup and data storage mechanisms required for the application to store relevant information securely.

5. **User Authentication System:**

- The secure user authentication system that enables users to create accounts, log in, and access the application securely.

Documentation and Manuals:

1. **Project Report:**

- A comprehensive report detailing the project's objectives, methodologies, architecture, implementation, evaluation, and future recommendations.

2. **User Guides:**

- Documentation providing instructions for users on how to navigate and utilize the web application effectively.

3. **Technical Documentation:**

- In-depth technical documentation covering system architecture, AI model specifications, and deployment instructions.

Training Materials:

1. **Model Training Instructions:**

- Instructions and materials used to train the AI model, including dataset information and model training configurations.