# Report on Attempt to Detect Features of Dog Faces in Zoomed-In Images

## Introduction

This project focuses on detecting dog faces in a set of provided images and identifying key facial landmarks. The process involves three main steps: data preparation, training a Single Shot Detector (SSD) model for face detection, and using Mediapipe's Face Mesh for landmark detection.

## Step 1: Data Preparation

The first step was to prepare the dataset. Images were collected from a specified directory and resized to 224x224 pixels, the required input size for the SSD model. Placeholder bounding box labels were created, assuming the entire image contained the dog face. This initial dataset preparation was crucial for the subsequent training and detection phases.

## Step 2: Training the SSD Model

A MobileNetV2-based SSD model was implemented for detecting dog faces. MobileNetV2 was chosen due to its efficiency and effectiveness in handling image classification tasks. The base model's output was processed through additional layers to predict bounding boxes around detected faces. The model was trained with the prepared images and labels, although a placeholder binary cross-entropy loss was used. This step aimed to fine-tune the model to recognize dog faces within the images.

## Step 3: Pose Estimation Using Mediapipe

Mediapipe's Face Mesh was utilized to detect facial landmarks on the identified dog faces. Face Mesh, typically used for human faces, was repurposed for this project to identify key points on dog faces. The input images were converted to RGB, and the face landmarks were extracted if any faces were detected.

## Detection and Visualization

The trained SSD model was employed to detect faces in new images. Detected bounding boxes were resized to fit the original image dimensions. Subsequently, Mediapipe Face Mesh was applied to the detected faces to identify landmarks. The results were visualized by overlaying bounding boxes and landmarks on the original images using Matplotlib.

## Challenges and Considerations

1. **Data Quality and Quantity**: The initial dataset had limited and possibly unrepresentative images of dog faces, impacting the model's ability to generalize well.
2. **Bounding Box Accuracy**: Placeholder bounding boxes might not accurately represent the dog's face, affecting the model's training quality.

3. **Landmark Detection**: Mediapipe Face Mesh is optimized for human faces, and its application to dog faces might not yield precise landmarks.
4. **Model Training**: The SSD model training used a placeholder implementation. A more sophisticated approach, such as incorporating a well-defined loss function and using a more extensive and annotated dataset, could improve detection performance.

**Conclusion**

This attempt provided a foundational understanding of the process involved in detecting dog faces and extracting facial landmarks. The integration of an SSD model with Mediapipe Face Mesh, despite its challenges, showcased the potential for such techniques in animal face detection. Future work should focus on enhancing the dataset quality, refining the bounding box annotations, and optimizing the model for better accuracy and robustness.

4o