Project Report: Semantic Segmentation using Pascal VOC Dataset

1. Introduction
Semantic segmentation is a critical task in computer vision, involving the classification of each pixel in an image into predefined categories. This project initially aimed to utilize the Cityscapes dataset to develop a deep learning model for semantic segmentation of urban road scenes. However, due to difficulties in accessing the Cityscapes dataset, the Pascal VOC dataset was used as an alternative. This report outlines the challenges faced, the implemented solution, and possible alternatives.

2. Problem Statement
The original goal was to develop a model capable of segmenting road scenes into various classes such as vehicles, pedestrians, road signs, and buildings using the Cityscapes dataset. However, the inability to access the Cityscapes dataset due to login credential issues I shifted to the Pascal VOC dataset. The main challenge was to adapt the project to a different dataset while maintaining the objective of achieving high accuracy in segmentation tasks.

3. Solution Implemented
To address the challenge, the following steps were undertaken using the Pascal VOC dataset:

3.1 Dataset Preparation
- **Data Source**: Pascal VOC dataset, which includes images with annotations for 20 different object classes plus background making 21 classes.
- **Data Preprocessing**: Images and annotations were preprocessed to ensure compatibility with the deep learning framework used (PyTorch).

3.2 Model Architecture
Transforms:
- Resized images to 256x256 pixels.
- Normalized the images to match the input format expected by the pre-trained fnc_resnet50

3.3 Training and Validation
- **Dataset Split**: The dataset was divided into training (75%), validation (15%), and test sets (15%).
- **Hyperparameter Tuning**: The model was trained using the training set, and hyperparameters were tuned based on validation performance.

3.4 Evaluation
- **Metrics**: The model's performance was evaluated using pixel accuracy, and qualitative analysis through visual inspection of segmented images.

4. Challenges Faced
- **Dataset Access**: The primary challenge was the inability to access the Cityscapes dataset due to login credential issues. This necessitated a shift to the Pascal VOC dataset.

- **Adaptation**: Adapting the project plan and methodology to a different dataset with different characteristics and classes.
- **Implementation**: Ensuring the preprocessing, model design, and evaluation metrics were suitable for the Pascal VOC dataset.
- **Selecting suitable Model**: I tried various models and best one I could find was fnc_resnet50
- **Implementing Segment Anything Model (SAM):** At one Point I also tried to use [SAM](#) created by Meta but could not find appropriate way to include it in the project.

5. Possible Alternatives
- **Alternative Datasets**: Other datasets such as COCO (Common Objects in Context) or ADE20K could be considered for future work.
- **Model Variations**: Experimenting with different architectures, such as Fully Convolutional Networks (FCNs), SegNet or SAM to improve performance.
- **Advanced Techniques**: Incorporating advanced techniques such as Transfer Learning or using pre-trained models on similar tasks to enhance accuracy.

6. General Project Description
This project involves the development of a deep learning model for semantic segmentation using the Pascal VOC dataset. The aim is to accurately classify each pixel in an image into one of the 20 object classes provided by the dataset. The approach includes data preprocessing, CNN model design, data augmentation, training, and evaluation. Despite the initial challenge of accessing the Cityscapes dataset, the use of Pascal VOC allowed for successful segmentation model development and evaluation.

7. Results
Evaluation Metrics:
- **Pixel Accuracy**: 94.43%

Visual Results: Below are some screenshots showcasing the segmentation results:
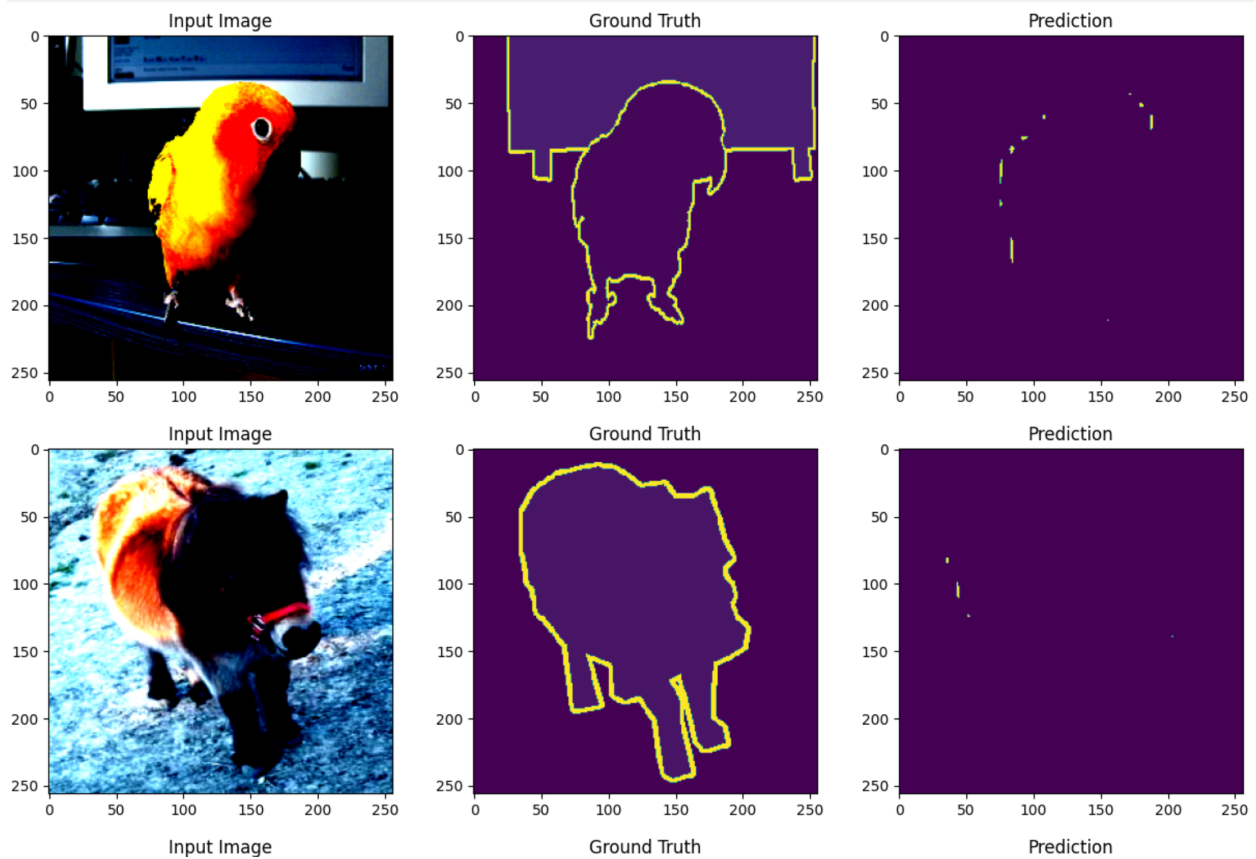
Something went very wrong with the prediction visualization, as I got higher frequency but the prediction visualization was not as good as expected.

8. Conclusion

Even though the project did not go as per planned, but the project successfully demonstrated the capability of using the Pascal VOC dataset for semantic segmentation tasks. The challenges encountered highlighted the importance of adaptability and the availability of alternative datasets. Future work could explore additional datasets and advanced techniques to further enhance segmentation accuracy and robustness.

9. References

- **Pascal VOC Dataset**: PASCAL Visual Object Classes
- **Cityscapes Dataset**: Cityscapes Dataset Overview