

# Image Segmentation

 by AI Cloud Solutions

# Problem Statements

In this project, the objective is to build a semantic segmentation model using the U-Net architecture to classify each pixel in images captured from the CARLA self-driving simulator. The dataset is sourced from the Lyft-Udacity Perception Challenge and contains RGB images paired with segmentation masks. To simplify the workflow and reduce training time, only the dataA subset is used. All images are resized and normalized, and the masks are reduced to single-channel class maps using argmax.

The model is implemented in TensorFlow/Keras and trained using standard callbacks like early stopping and learning rate reduction. Once trained, it's evaluated using metrics such as accuracy, IoU, F1 score, precision, recall, and specificity. Visual results are also generated to compare the model's predictions with the actual masks. The output includes the trained model file, performance plots, and sample prediction visuals

## Model Architecture – U-Net

The U-Net model is constructed using custom functions:

### Encoder (Downsampling)

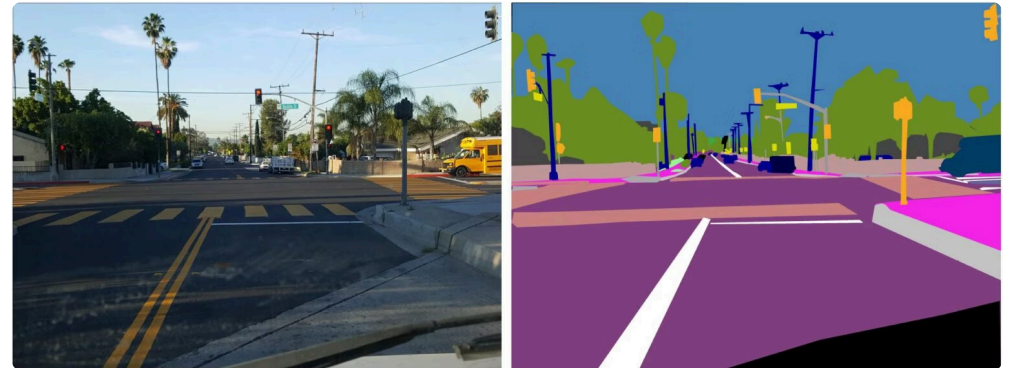
- Two Conv2D → BatchNorm → ReLU layers
- MaxPooling (except at bottleneck)
- Skip connections are retained for decoder

### Decoder (Upsampling)

- Conv2DTranspose for upsampling
- Concatenation with corresponding encoder skip connection
- Two Conv2D → BatchNorm → ReLU layers

### Output

- Final 1x1 Conv2D layer with sigmoid activation to predict 13 classes.

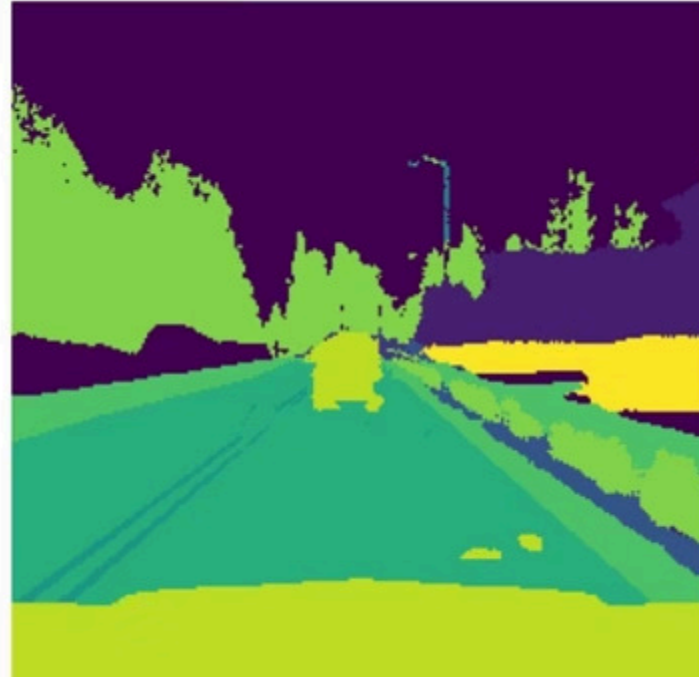
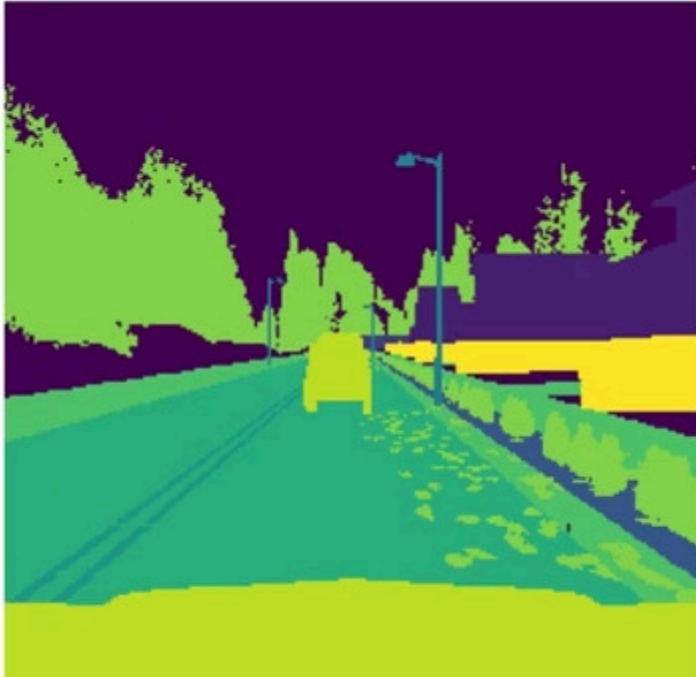


# Evaluation Metrics

After training, true and predicted masks are generated for all datasets:

- Metrics computed per class:
  - **Precision, Recall, Specificity**
  - **IoU, TDR, F1-Score**
- Evaluation is shown for:
  - Train Set
  - Validation Set
  - Test Set

All metrics are presented in tabular format using pandas DataFrames.



# Visualisation

# Prediction Visualization

A utility function `show_predictions()` displays:

- Input image
- Ground truth mask
- Predicted mask (post-processing with argmax)

This helps validate model performance visually.

# Dataset & Assumptions

- Dataset: [Lyft Udacity Challenge – CARLA Simulated Images](#)
- Only `dataA/` is used for experimentation to reduce training time.
- Image-mask pairs are loaded from Google Drive and preprocessed before training.

# Summary

This notebook delivers a full segmentation pipeline:

- Modular U-Net implementation
- Realistic dataset subset
- Structured metrics
- Clear evaluation and visualization tools

It is built to serve both educational clarity and real-world deployment readiness.

# Resources

[Github URLG](#)

ChatGPT Problem Solving Prompts