

CV A - Hinton's  
Prophet

# Self-Driving Car

Portofolio Project 3



## OBJECTIVE

“Mengembangkan sistem cerdas berbasis teknologi AI yang mampu melakukan pelacakan objek untuk kebutuhan smart city”

## Business Understanding

Proyek ini bertujuan untuk mengatasi masalah kemacetan dan kecelakaan di kota-kota besar di seluruh dunia dengan mengembangkan teknologi **Self-Driving Car berbasis AI**. Dengan kemampuan AI untuk melakukan segmentasi objek dan pemetaan otomatis dari citra kamera, diharapkan sistem ini dapat:

1. **Mengurangi angka kecelakaan** yang disebabkan oleh kelalaian pengemudi.
2. **Mengurangi kemacetan** dengan memberikan kemampuan bagi kendaraan untuk mengenali lingkungan sekitarnya dan melakukan navigasi secara efisien.

Proyek ini akan menghasilkan model yang mampu mengenali berbagai objek di jalan seperti kendaraan lain, pejalan kaki, rambu lalu lintas, dan lain-lain melalui **segmentation models**.



## Domain Understanding

Proyek ini beroperasi dalam domain **Computer Vision** dan **Autonomous Driving (Self-Driving Cars)** dengan fokus utama pada **Object Segmentation**. Teknologi ini memerlukan:

1. **Dataset:** Menggunakan **Cityscapes Dataset** yang berisi gambar-gambar dari lingkungan perkotaan dari perspektif pengambilan kamera di mobil.
2. **Model yang digunakan:**
  - FCN8s (Fully Convolutional Network 8s)
  - U-Net
3. **Evaluasi:** Dilakukan untuk menentukan model mana yang paling efektif dalam melakukan segmentasi objek pada dataset.
4. **Publikasi:** Hasil dari proyek ini akan dipublikasikan di **GitHub**.



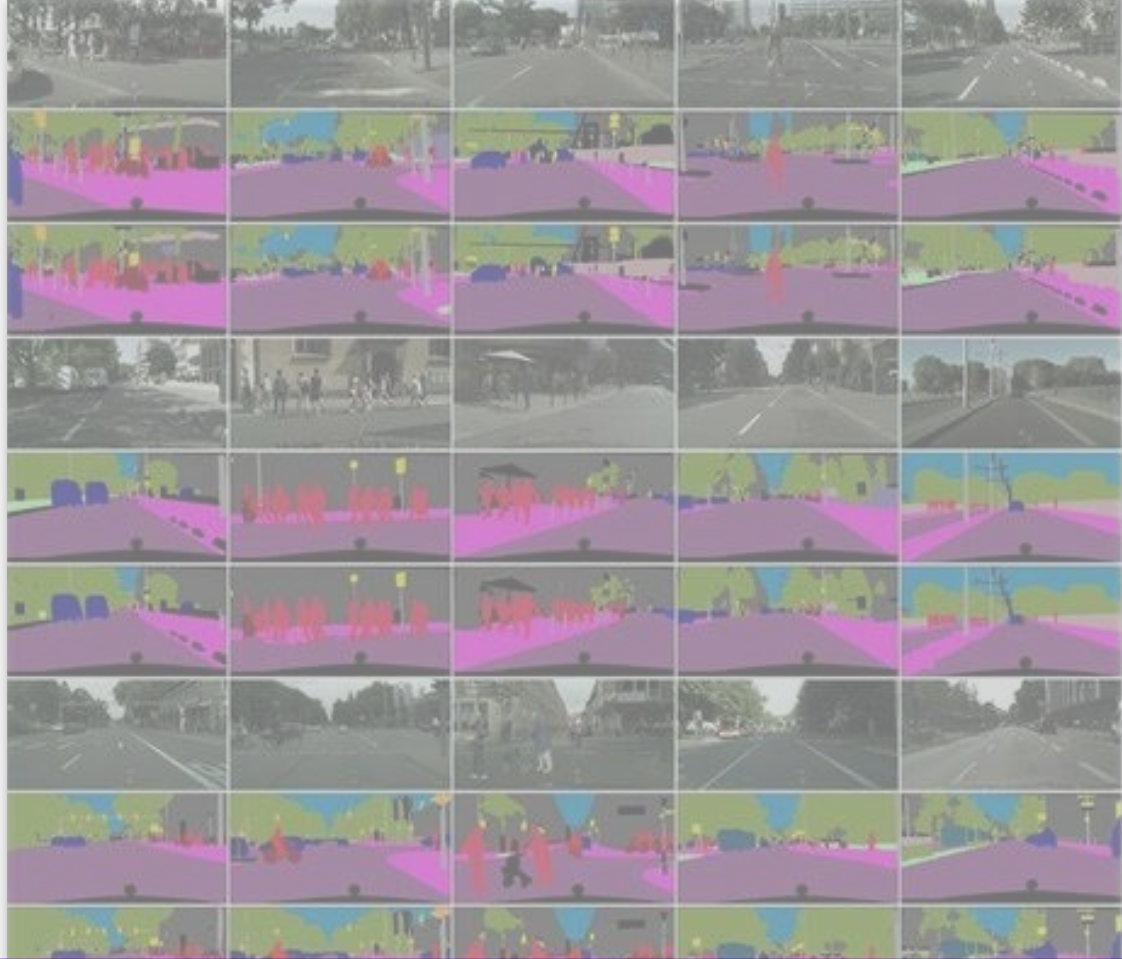
## Tools



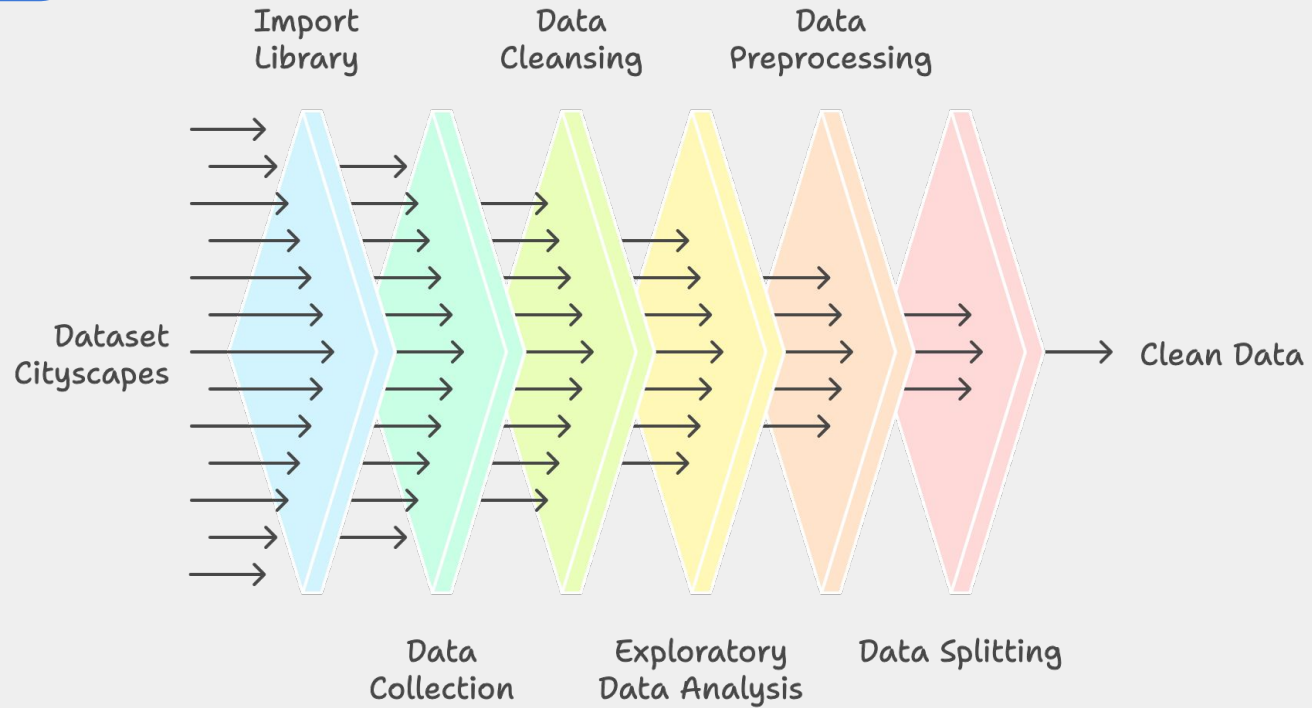


## Data Understanding

- Menggunakan dataset Cityscapes, dataset yang banyak digunakan dalam penelitian visi komputer, khususnya untuk tugas-tugas yang terkait dengan deteksi objek, segmentasi. Dataset ini dirancang untuk mendukung pengembangan teknologi kendaraan otonom.
- Ukuran dan Variasi: Dataset Cityscapes terdiri dari ribuan gambar yang diambil dari berbagai kota di Eropa, dengan anotasi mendetail untuk 30 kelas objek, seperti jalan, kendaraan, pejalan kaki, dan bangunan.
- Dalam dataset ini, terdapat 3.000 gambar untuk pelatihan (training) dan 750 gambar untuk pengujian (testing).



## Data Preparation



## Data Collection

```
From: https://drive.google.com/uc?id=10wNCh9NMkYEX8PtJCGUjUxWbgoXqnh
To: /content/annotations_prepped_train-20241111T025913Z-001.zip
100% [██████████] 2.30M/2.30M [00:00<00:00, 124MB/s]
Downloading...

From: https://drive.google.com/uc?id=14vLTnRa3no-7NxA_tSNsePLjdD3JE09
To: /content/annotations_prepped_train-20241111T025903Z-001.zip
100% [██████████] 566k/566k [00:00<00:00, 75.3MB/s]
Downloading...

From (original): https://drive.google.com/uc?id=10wNCh9NMkYEX8PtJCGUjUxWbgoXqnh
From (redirected): https://drive.google.com/uc?id=10wNCh9NMkYEX8PtJCGUjUxWbgoXqnh&confirm=t&uiid=ab302baa-1388-4c47-bd64-764a6b7d7d57
To: /content/images_prepped_test-20241111T025016Z-001.zip
100% [██████████] 28.6M/28.6M [00:00<00:00, 72.6MB/s]
Downloading...

From (original): https://drive.google.com/uc?id=1dnltnJdNTVMF80Jp70xW2vM0SPPhlp2R4S6&confirm=t&uiid=d993c1ee-a081-4d34-8970-3d61aed67d94
From (redirected): https://drive.google.com/uc?id=1dnltnJdNTVMF80Jp70xW2vM0SPPhlp2R4S6&confirm=t&uiid=d993c1ee-a081-4d34-8970-3d61aed67d94
To: /content/images_prepped_train-20241111T024959Z-001.zip
100% [██████████] 103M/103M [00:00<00:00, 108MB/s]
```

## Data Preprocessing

```
data_transforms = A.Compose([
    A.Resize(256,256),
    A.Normalize(),
    ToTensorV2()
])
```

Data  
Splitting =  
80:20

## Data Cleansing

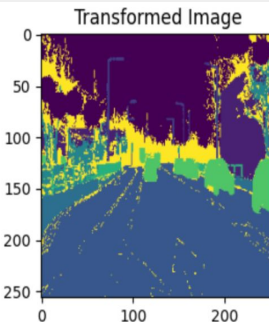
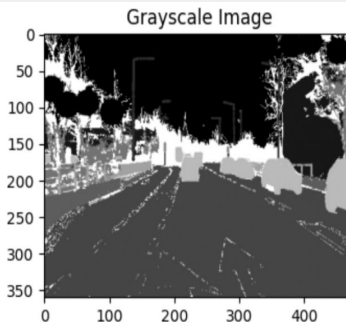
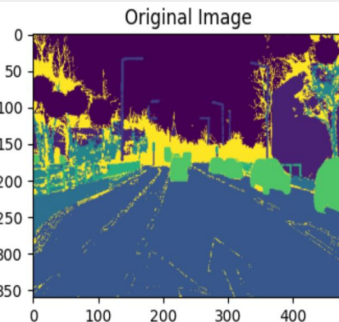
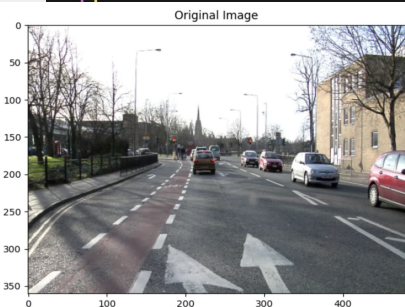
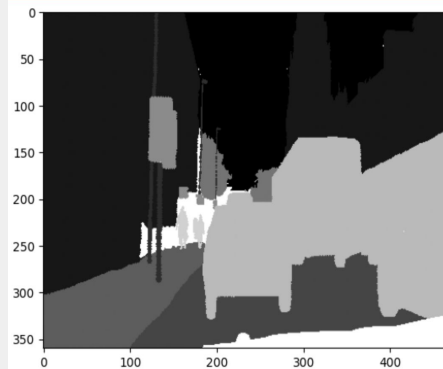
```
PATH = "/content/data/images_prepped_train"
files_list = os.listdir(PATH)

for file_name in files_list:
    if "(" in file_name and ")" in file_name:
        file_path = os.path.join(PATH, file_name)
        os.remove(file_path)
        print(f'Deleted: {file_name}')
```

```
Deleted: 0006R0_f03840(1).png
Deleted: 0016E5_07860(1).png
Deleted: 0016E5_07710(2).png
Deleted: 0016E5_07650(1).png
Deleted: 0016E5_07800(1).png
Deleted: 0016E5_07740(1).png
Deleted: 0016E5_07770(1).png
Deleted: 0016E5_07830(2).png
Deleted: 0016E5_07920(1).png
Deleted: 0016E5_07920(2).png
```

## EDA

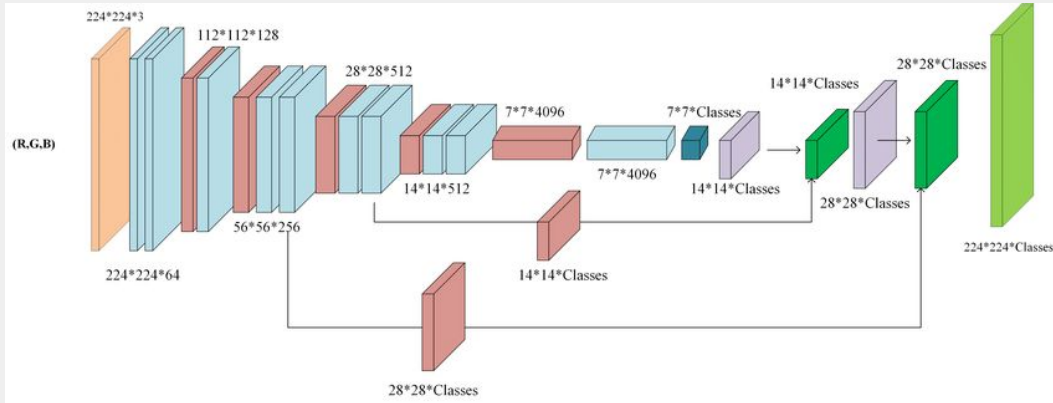
The number of train dataset is 367  
The number of test dataset is 101



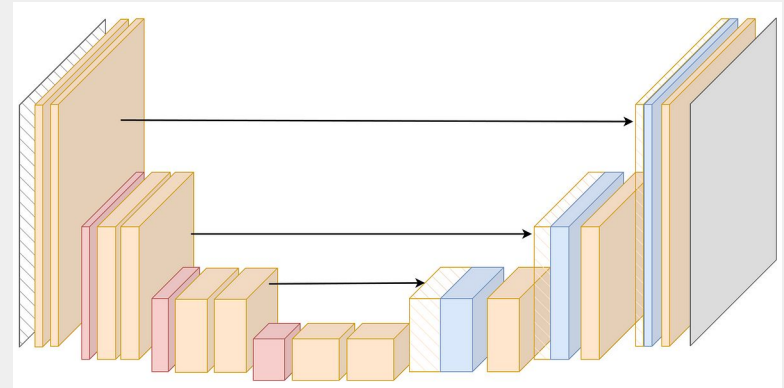


## Architecture & Modelling

FCN8s

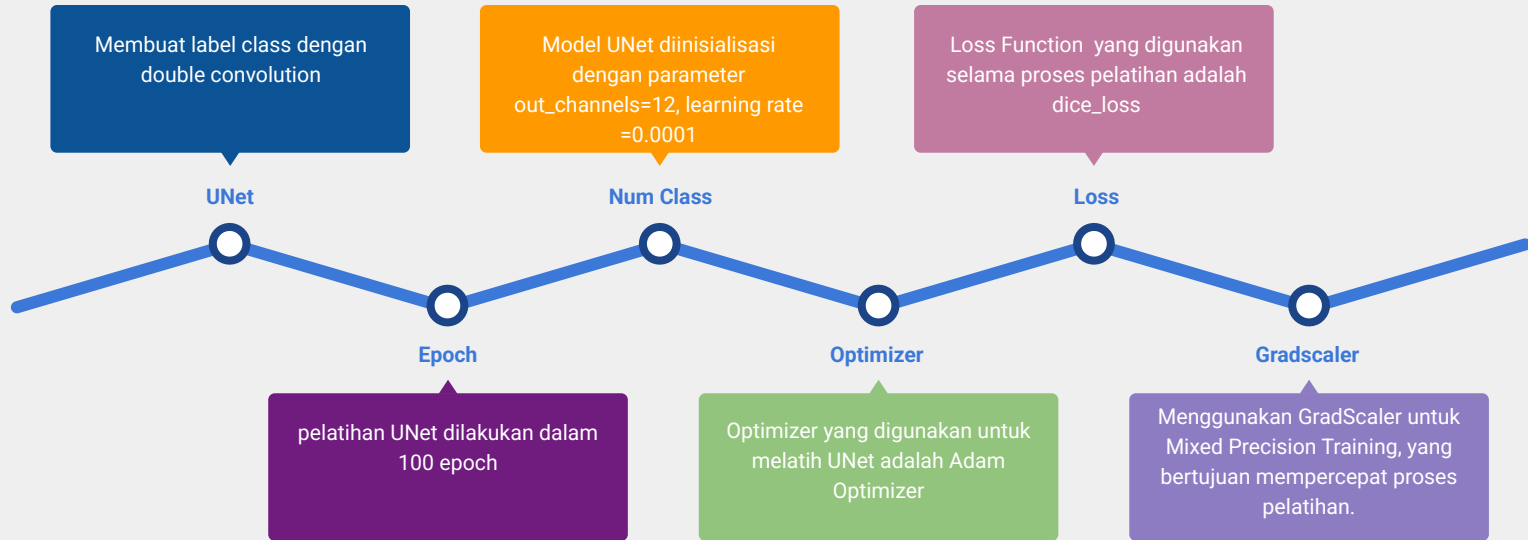


UNet

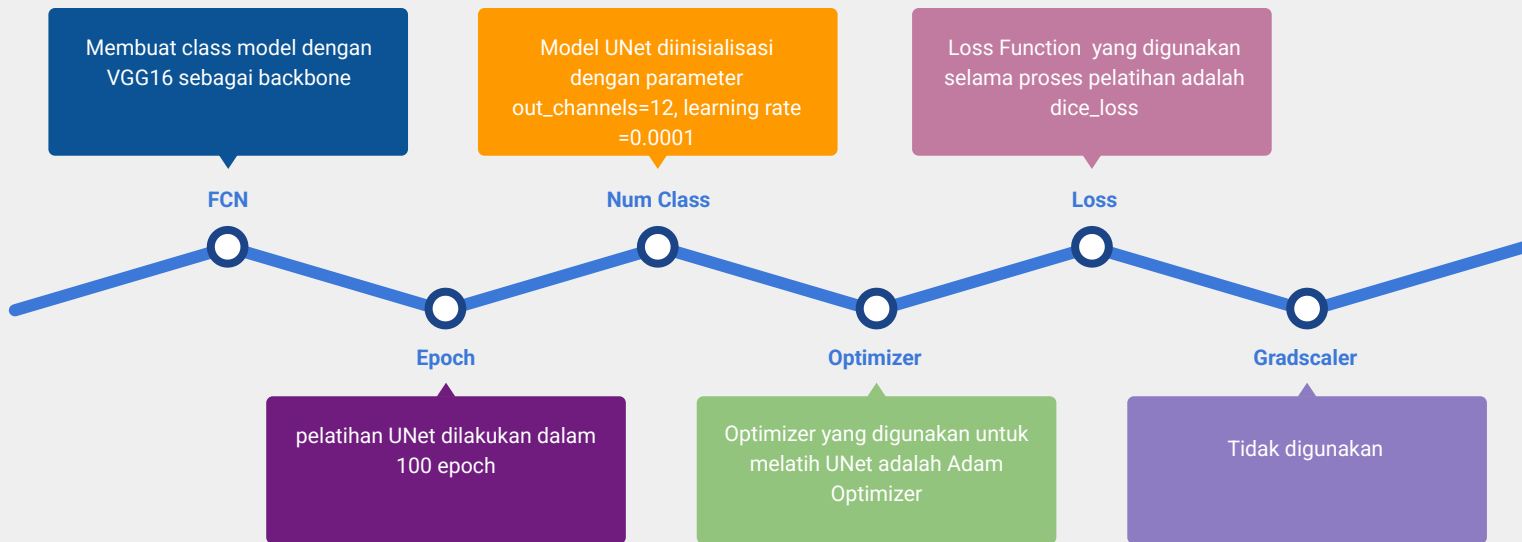


**FCN8s** menggunakan arsitektur *Fully Convolutional* dengan *skip connections* yang hanya menambahkan informasi dari lapisan sebelumnya, menyebabkan hilangnya detail spasial penting. Sebaliknya, **UNet** menerapkan arsitektur *encoder-decoder* dengan *skip connections* yang menggabungkan fitur dari *encoder* ke *decoder* melalui *concatenation*. Pendekatan ini memungkinkan **UNet** mempertahankan lebih banyak informasi spasial dan menghasilkan segmentasi yang lebih akurat. **FCN8s** lebih efisien secara komputasi, sementara **UNet** lebih kompleks namun lebih unggul dalam presisi.

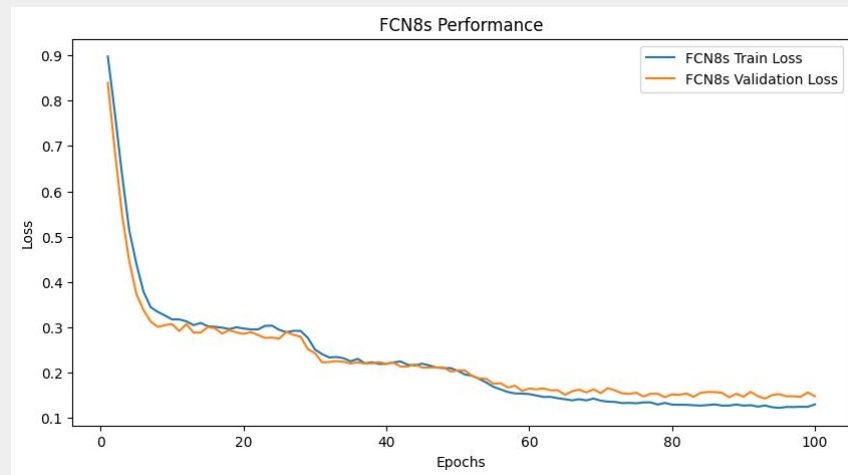
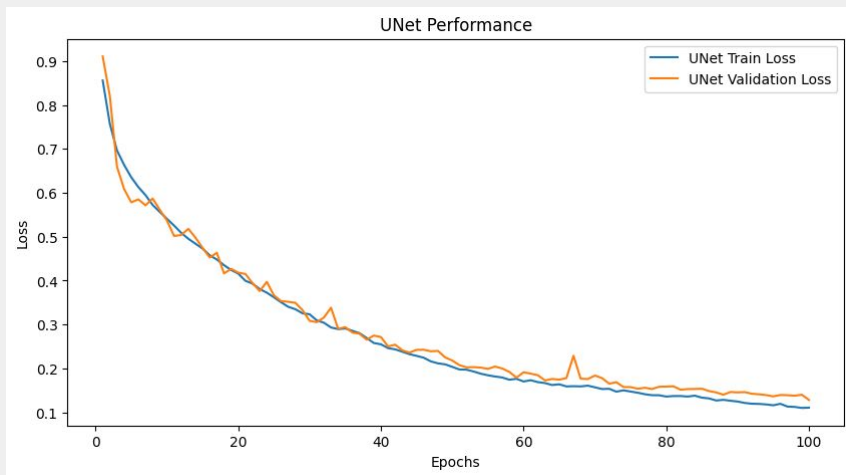
# UNet



## FCN8s



## Model Performance Comparison



## Evaluation and Comparison

### Loss Comparison

Model	Training Loss	Val Loss	Test Loss
UNet	0.11061112284660339	0.12824018796284994	0.15502086281776428
FCN	0.1301909029483795	0.14829587936401367	0.1457064151763916



## Evaluation and Comparison

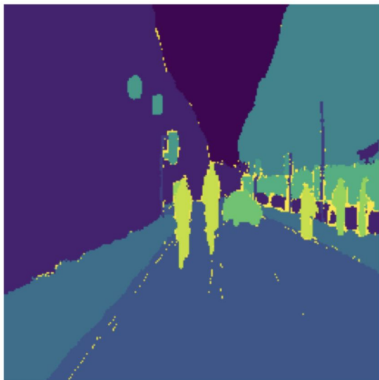
### Dice Coeff Comparison

Model	Training	Val	Test
UNet	0.059680	0.060127	0.051266
FCN	0.062916	0.063093	0.054208

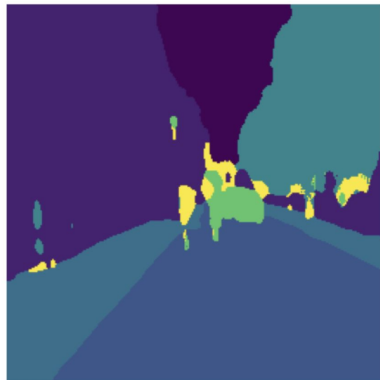
## Model Performance on TestSet

Ground Truth and Prediction Comparison of Image No 1

Ground Truth



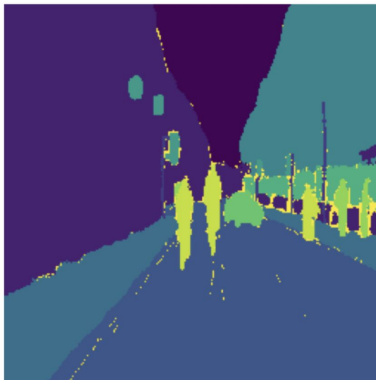
Prediction



UNet

Ground Truth and Prediction Comparison of Image No 1

Ground Truth



Prediction



FCN

## Conclusion

- FCN bekerja lebih baik dari UNet jika dilihat dari nilai loss dan koefisien dice. Tapi, visualisasi hasil dari UNet lebih bagus dibandingkan FCN8s. Artinya FCN lebih baik dalam membedakan objek, tapi kurang akurat dalam membuat mask.
- So, kalau tujuannya untuk membedakan objek, FCN adalah pilihan yang lebih cocok. Tapi kalau fokusnya lebih ke membuat mask, maka UNet lebih direkomendasikan

*"I think people need to understand that deep learning is making a lot of things, behind-the-scenes, much better."* - **Geoffrey Hinton**