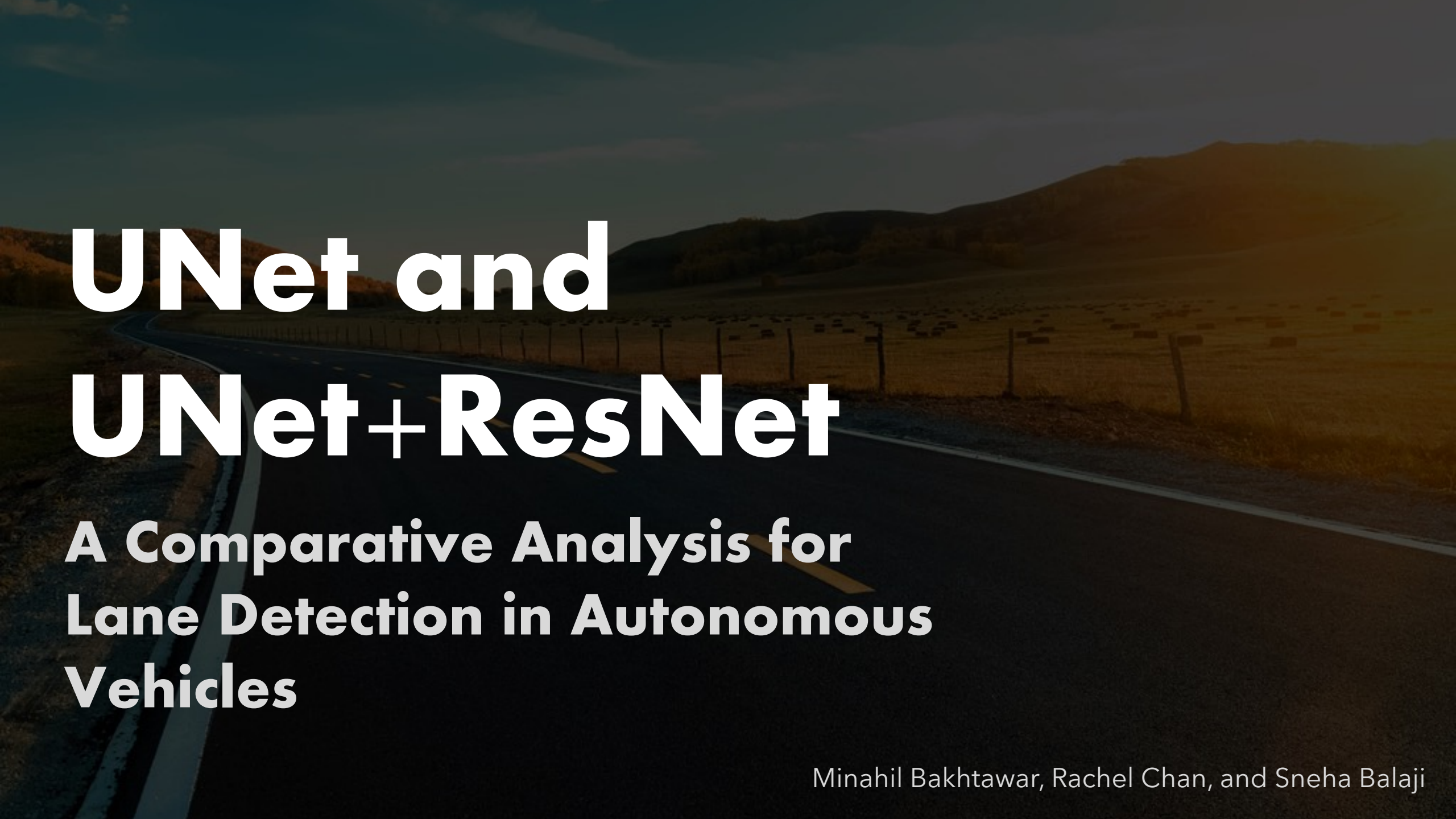


UNet and UNet+ResNet

A Comparative Analysis for Lane Detection in Autonomous Vehicles

Minahil Bakhtawar, Rachel Chan, and Sneha Balaji

The Lane Detection Problem

PURPOSE

Lane detection is very important for autonomous driving

CHALLENGES

Lighting, weather conditions, and road surface texture can cause challenges

MODELS

UNet is commonly used for lane detection

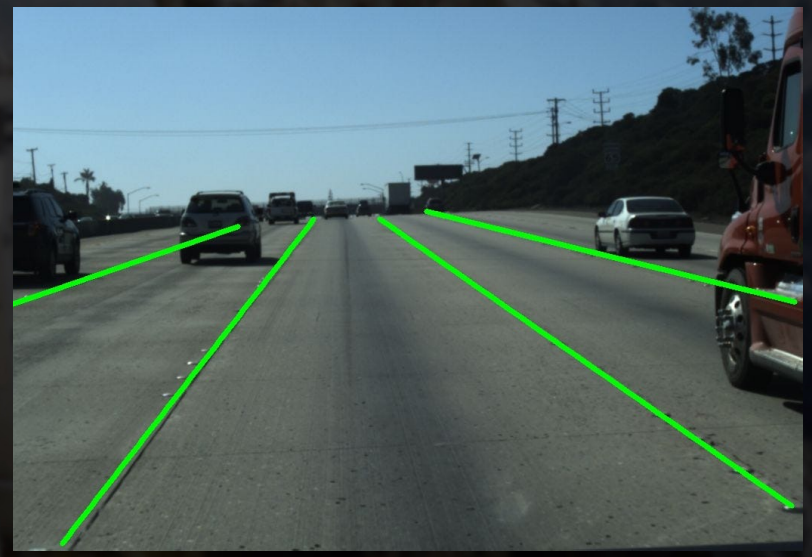
Other image segmentation tasks has used ResNet + UNet

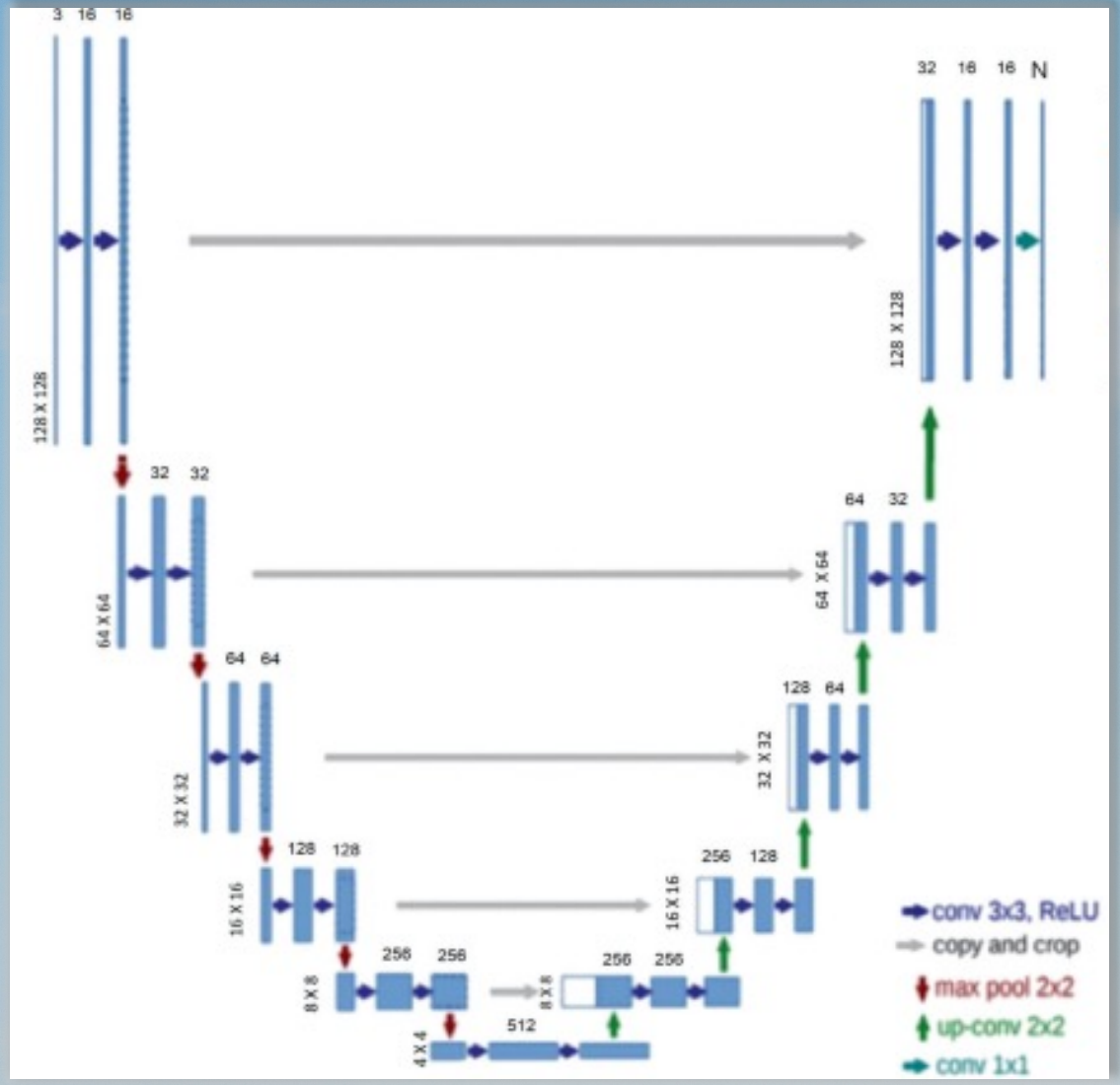
ResNet can:

- Improve global contextual understanding
- Reduce gradient diminishing problems

OUR APPROACH

Compared the performance of a UNet model against a UNet model with ResNet-50 encoder





UNet in Autonomous Driving

TASKS

UNet has been used for many tasks in autonomous driving like:

- Road Identification, road lane marking, surrounding obstacles
- Used to extract features from input images

VARIANTS

Modifications to UNet: CBAM, self-attention blocks, combinations of multiple UNet models

ISSUES

Issues in traditional UNet:

- Gradient diminishing
- Correlations between distant pixels in an image (global context)

UNet + ResNet in Image Segmentation

APPLICATIONS

UNet + ResNet combinations used in other applications like:

- Remote sensing, venous identification, aerial vehicle images, sea-land segmentation
- ResNet + UNet often outperforms just UNet in image segmentation tasks

BENEFITS

ResNet-50 is a powerful feature extractor

- Able to understand global context
- Mitigates gradient diminishing
- Skip connections allow for high-level detailed features

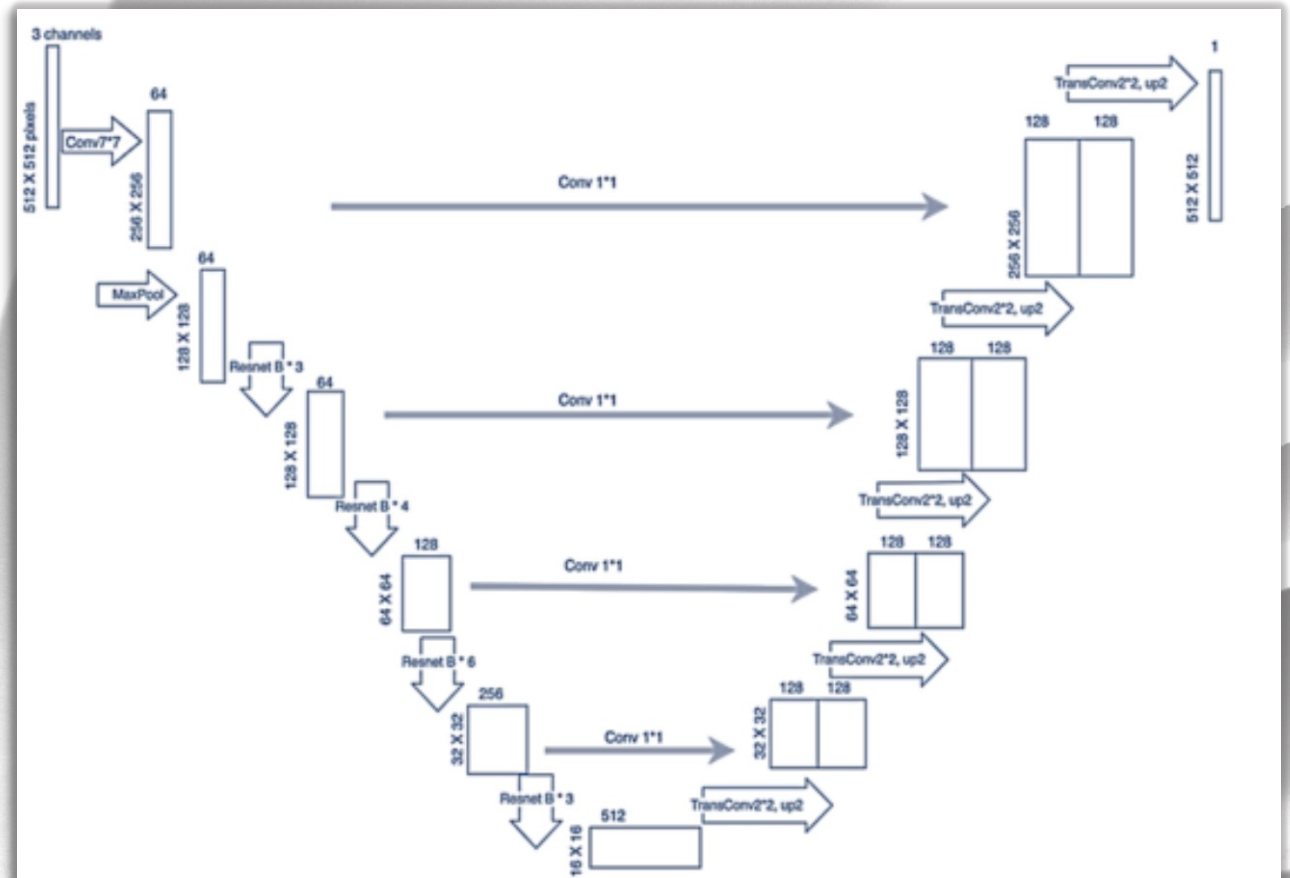


Fig. 3. UNet Architecture with ResNet Encoder [7]

Other Work in Image Segmentation

SegNet

Other works have used SegNet in automatic lane detection

UNet was found to have a lower MSE and average miss

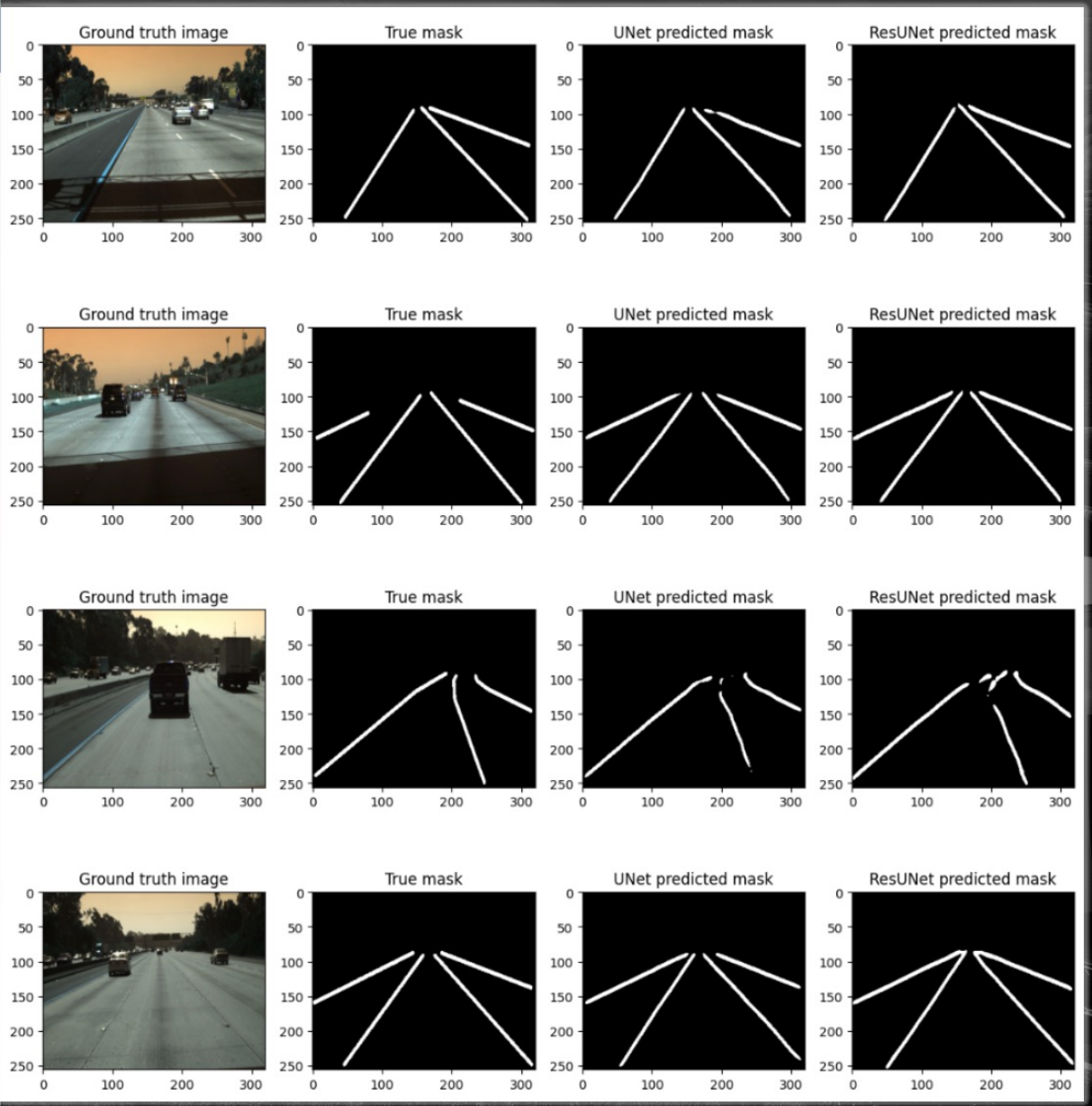
ResNet

ResNet has been explored for chest x-ray segmentation

Skip connections allow information to be retained

CNN

ResNet performed better than traditional CNN



Dataset

TUSIMPLE LANE DETECTION DATASET

Large collection of images from real-world driving scenarios

Diverse road, lighting, and weather conditions

ANNOTATIONS

Pixel-level annotations for lane markings

14,336 lane boundary annotations

CLASSES

7 different lane boundary classes

Pre-Processing

FRAME EXTRACTION

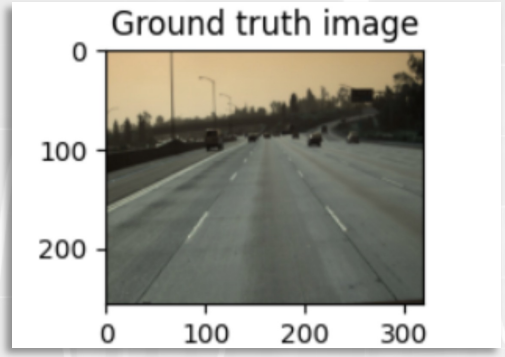
Last frame from dashcam clips are extracted as the raw ground truth image.

ANNOTATIONS

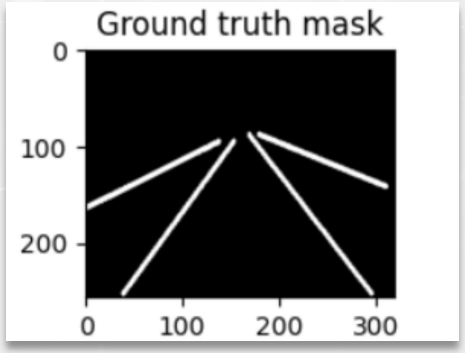
Lane annotations are applied to every pixel of the raw image file.

MASKS

Binary lane masks are generated using original classes, with lane type being ignored.




```
"lanes": [
      [-2, -2, -2, -2, 632, 625, 617, 609, 601, 594, 586, 578, 570, 563, 555,
547, 539, 532, 524, 516, 508, 501, 493, 485, 477, 469, 462, 454, 446, 438, 431,
423, 415, 407, 400, 392, 384, 376, 369, 361, 353, 345, 338, 330, 322, 314, 307,
299],
```

UNet Model

ENCODER

8 convolutional layers

Max pooling – noise suppression and feature invariance

BRIDGE

2 convolutional layers

Dropout – preventing overfitting, gaining model robustness

DECODER

8 Convolutional layers

Architectural symmetry

OUTPUT

1x1 convolution layer with **sigmoid** activation function

Res-UNet Model

ENCODER

ResNet50 – 5 stages, 3 convolutional layers per block

4 encoding blocks – input, stage 1, stage 2 block 3, stage 3 block 4

BRIDGE

ResNet50 – output from stage 4 block 6

DECODER

8 convolutional layers

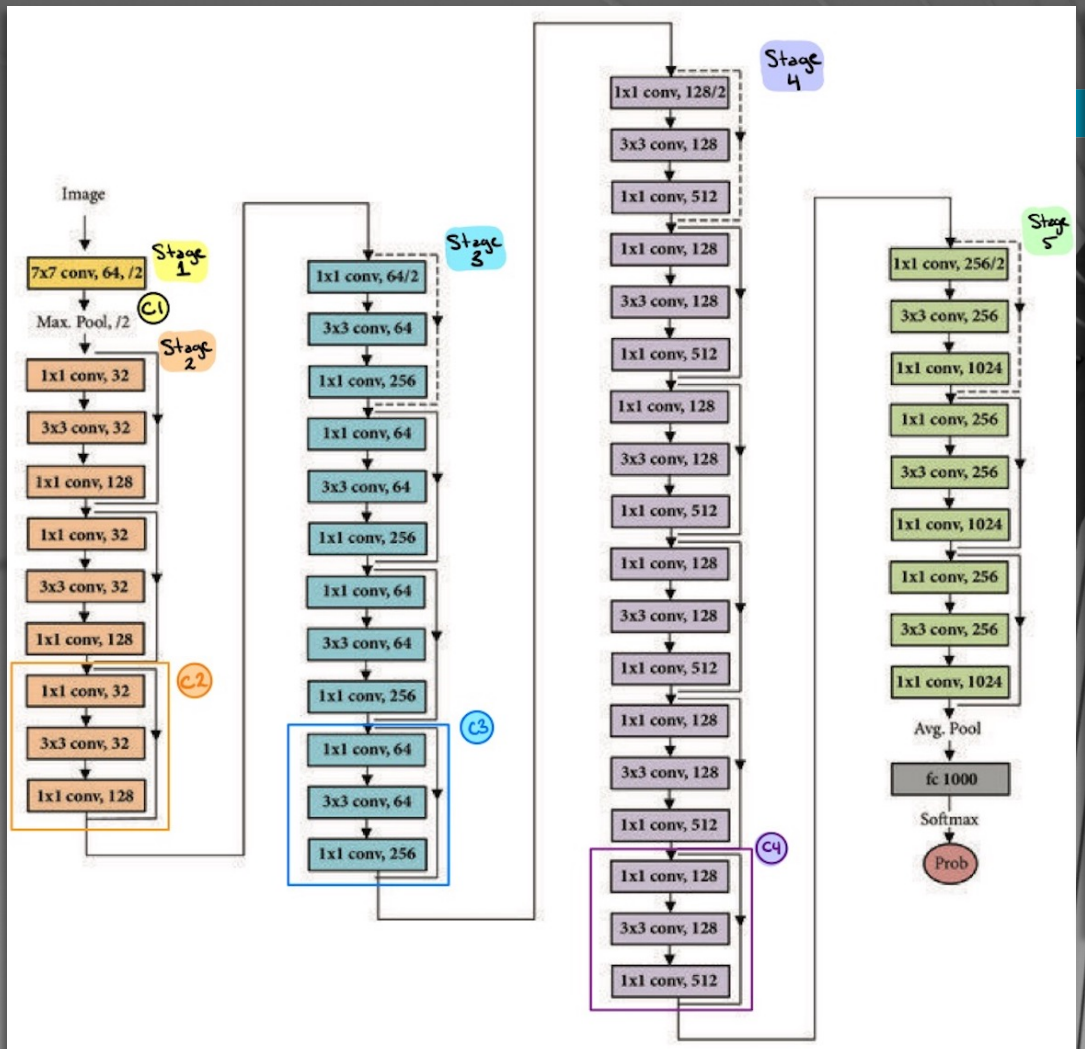
4 decoding blocks – same as UNet but using a Leaky ReLU activation instead to help prevent dying neuron problem

OUTPUT

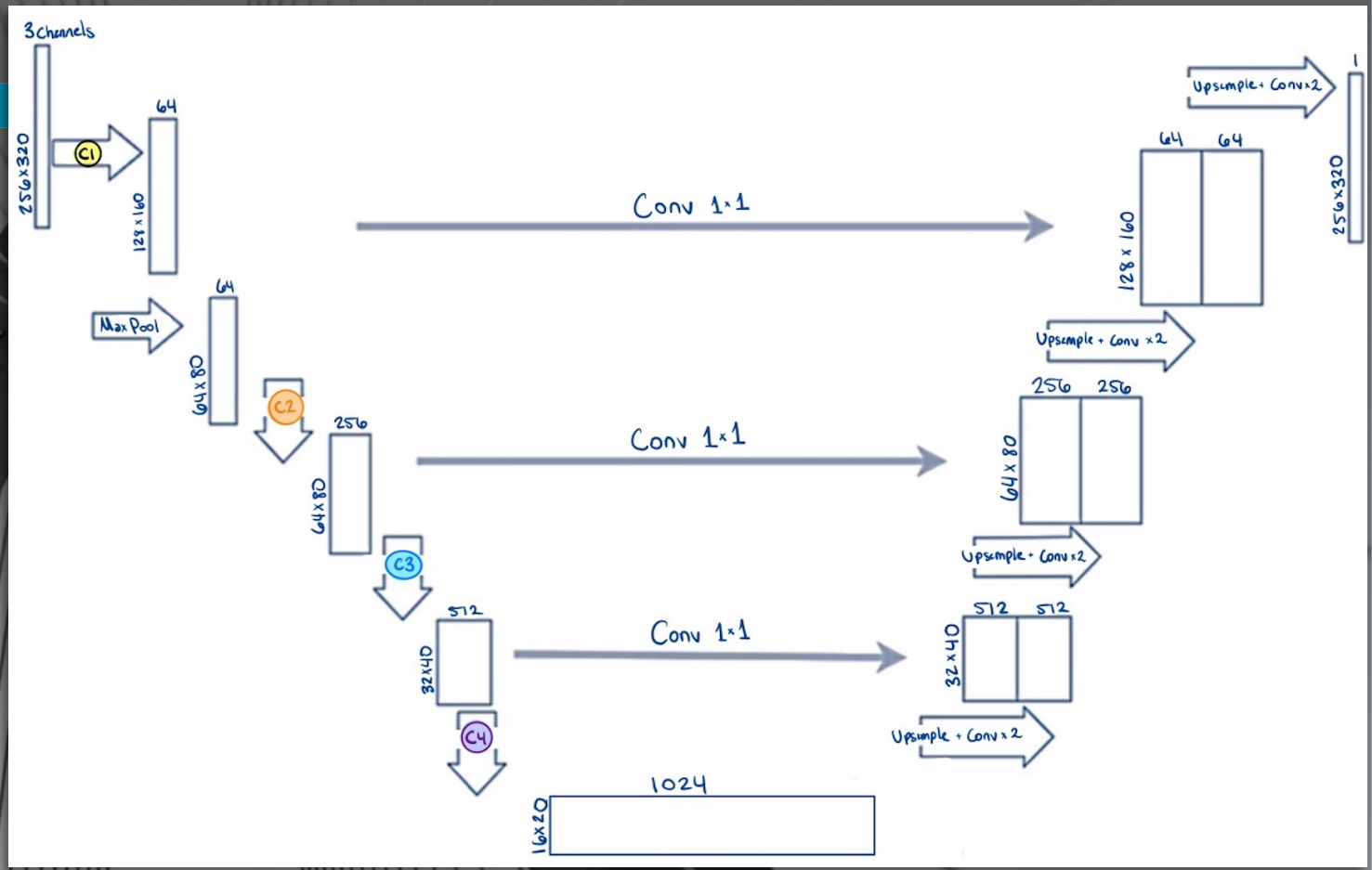
1x1 convolution layer with **sigmoid** activation function

Architecture Diagrams

ResNet-50



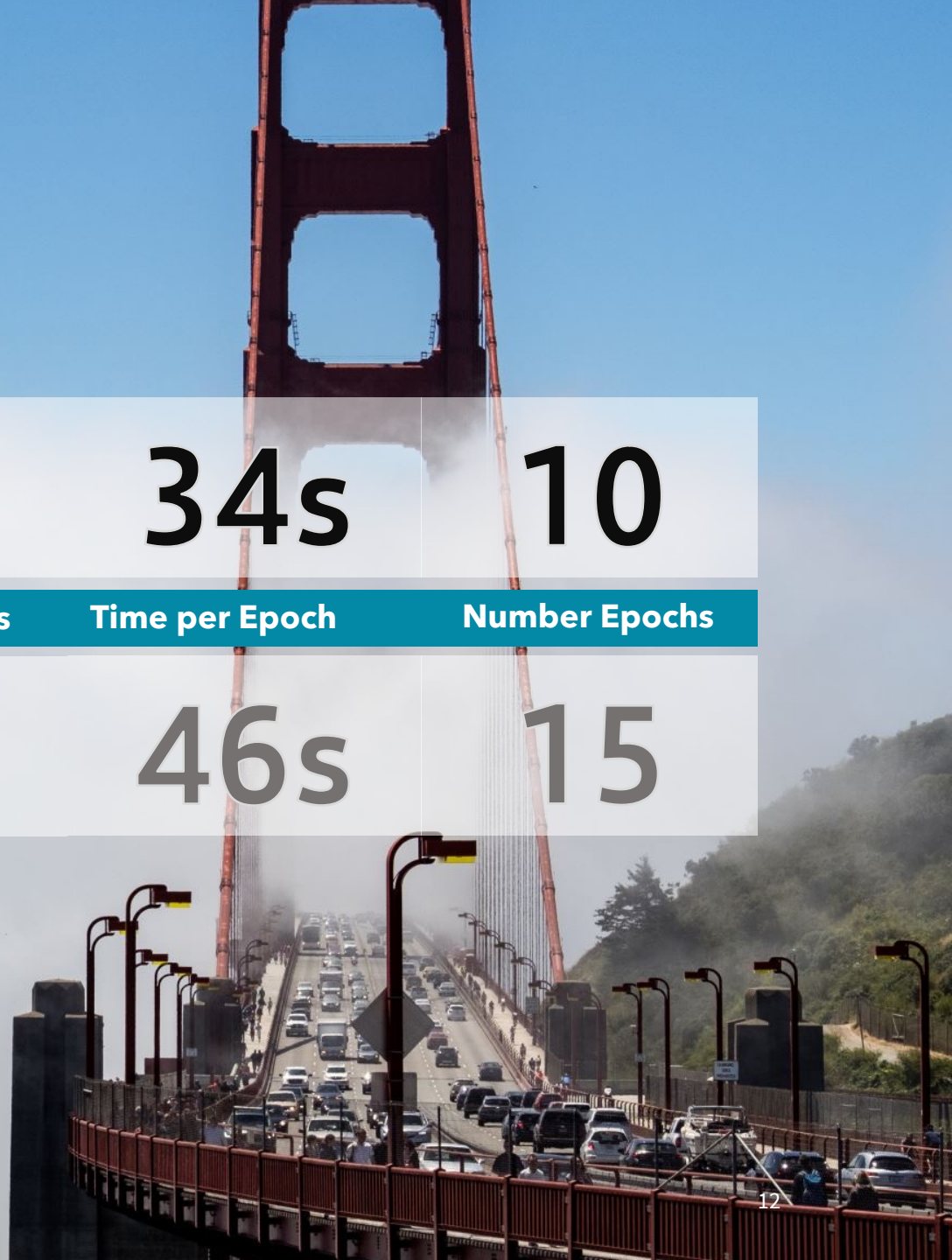
Res-UNet



PERFORMANCE RESULTS

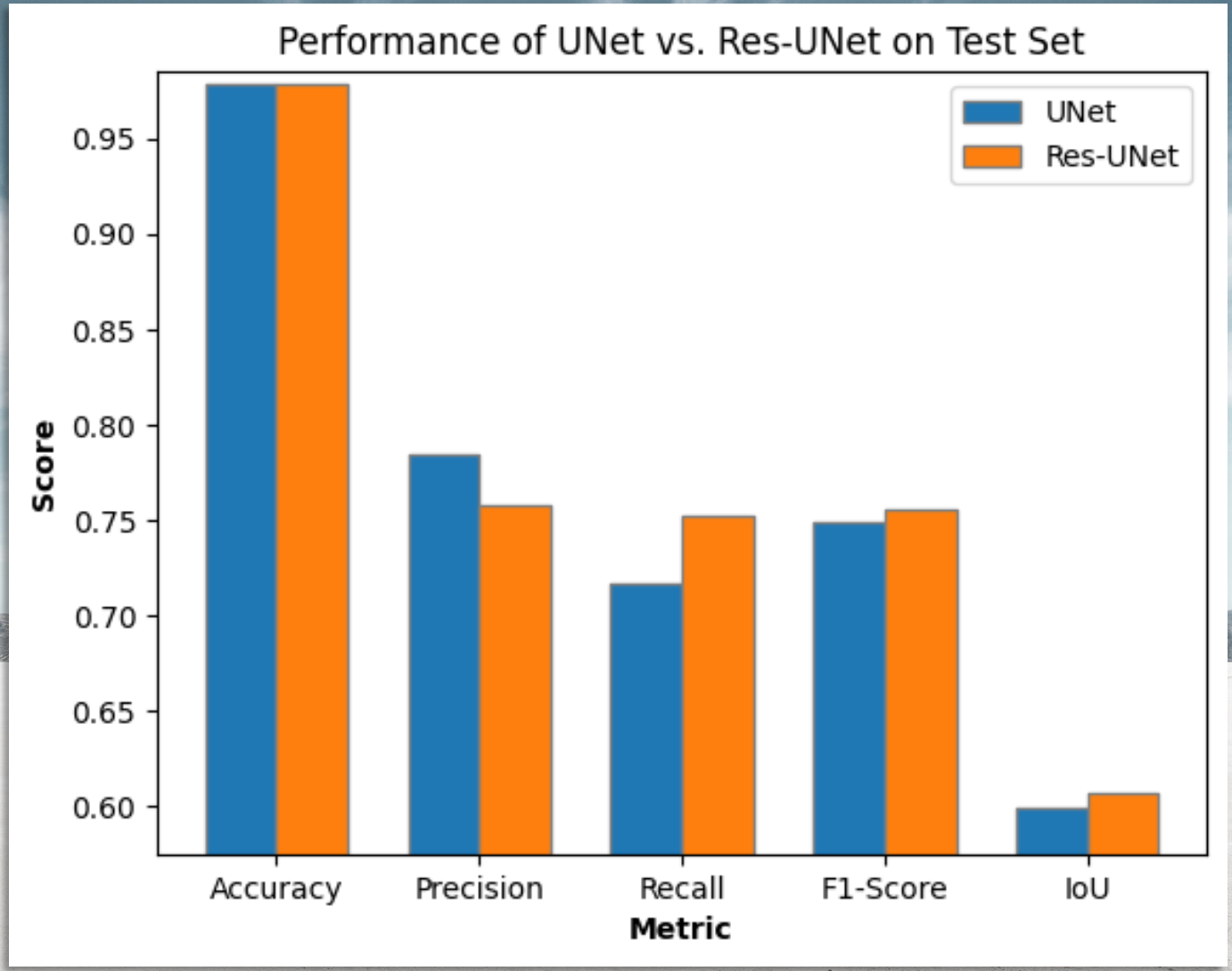
Training Cost

	Number of Parameters	Time per Epoch	Number Epochs
Res-UNet	21M	34s	10
UNet	31M	46s	15

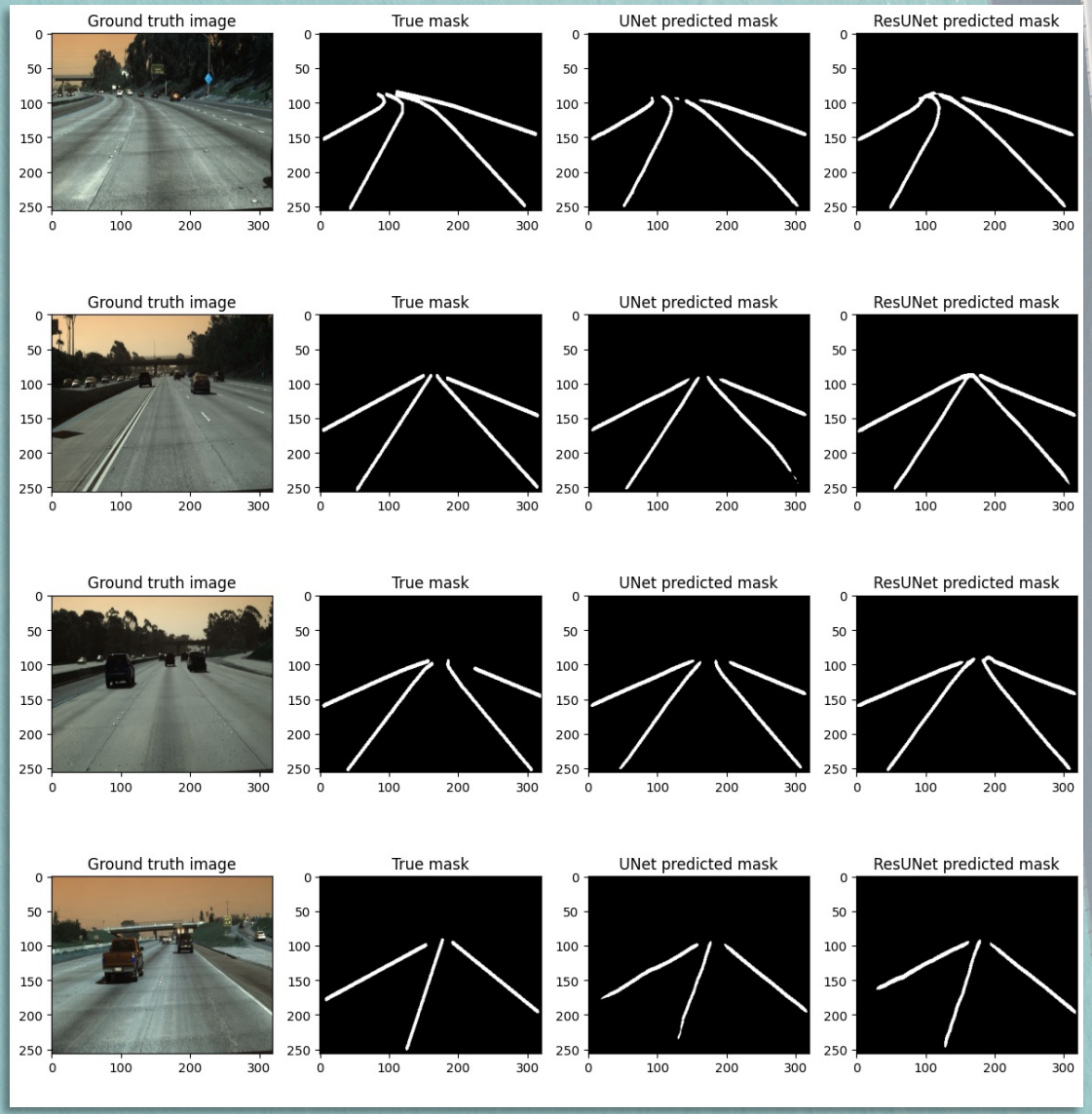


Performance Scores

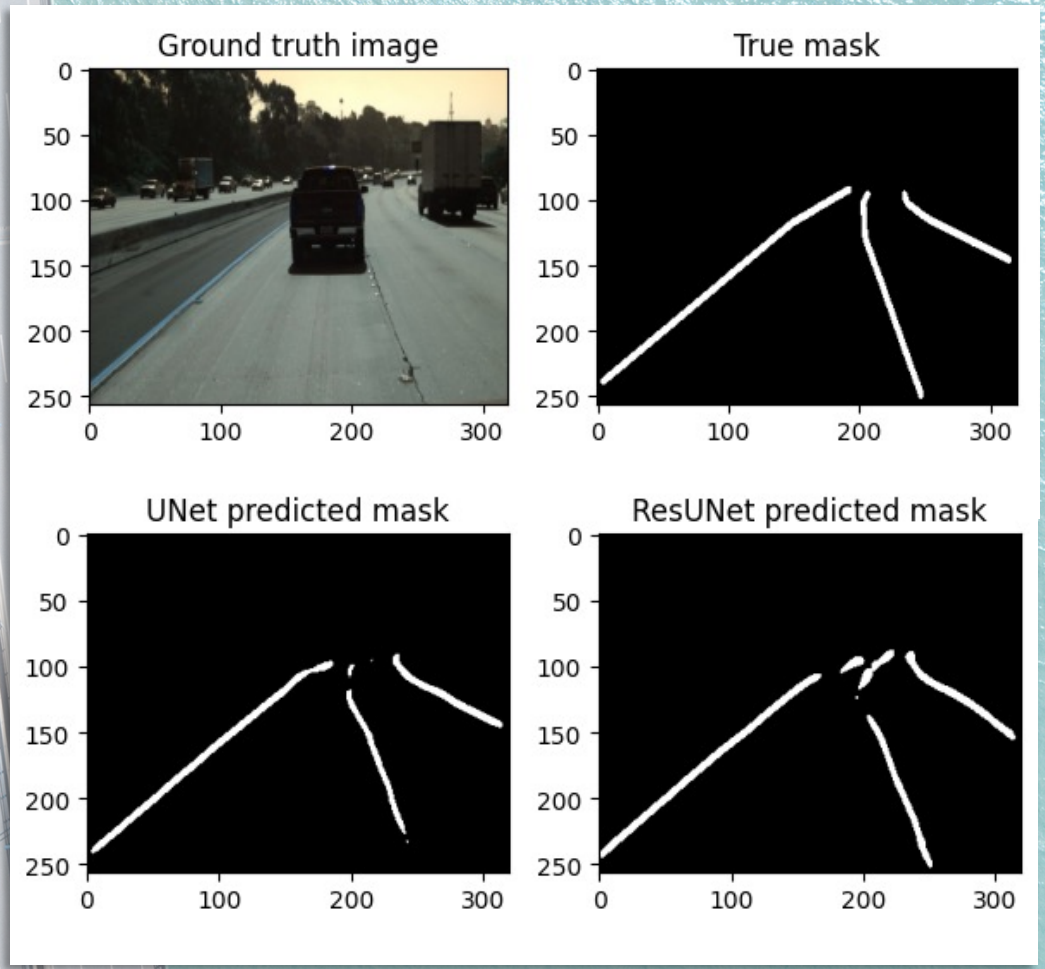
Model	Accuracy	Precision	Recall	F1-Score	IoU
UNet	***0.9787***	***0.7848***	0.7168	0.7493	0.5991
Res-UNet	0.9783	0.7582	***0.7524***	***0.7553***	***0.6068***



Predicted Masks



BLOCKED VIEW



Thank You!