ENPM673: Final Project

Rishie Raj

Uthappa Madettira

Group-11

UNIVERSITY OF MARYLAND FEARLESSLY FORWARD

Real-Time Semantic Segmentation

Content

- Background
- Network Architecture
- Results
- Metrics
- Challenges

Background

Previous Models

- Unet

  Encoder-Decoder structure with a bottleneck

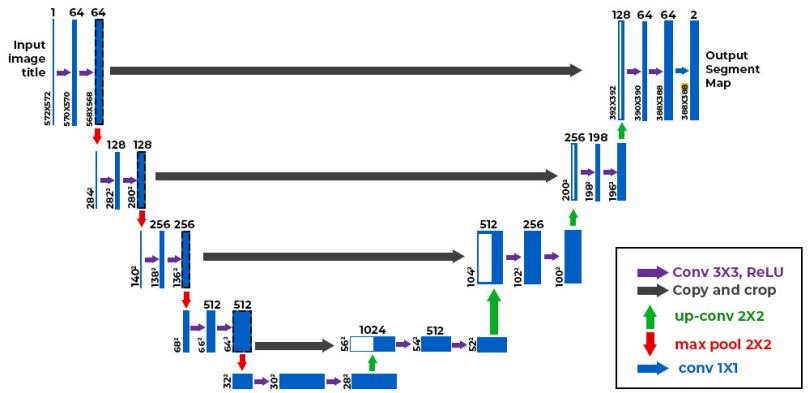
- SegNet

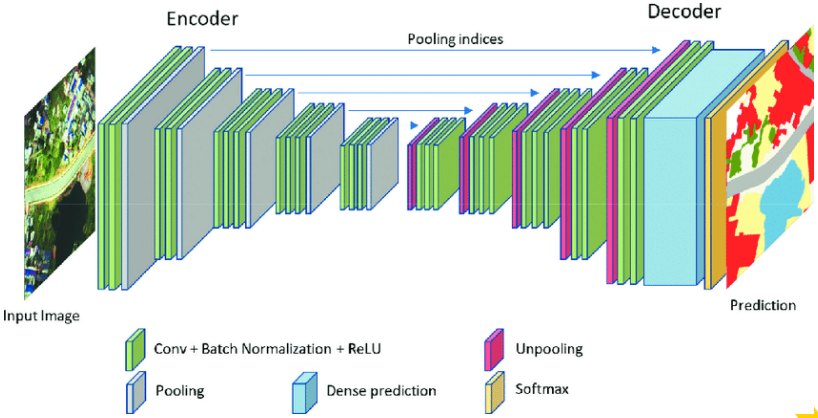
  Encoder-Decoder structure based on VGG16

- Issues

  Large models with a huge number of parameters

  Long inference time



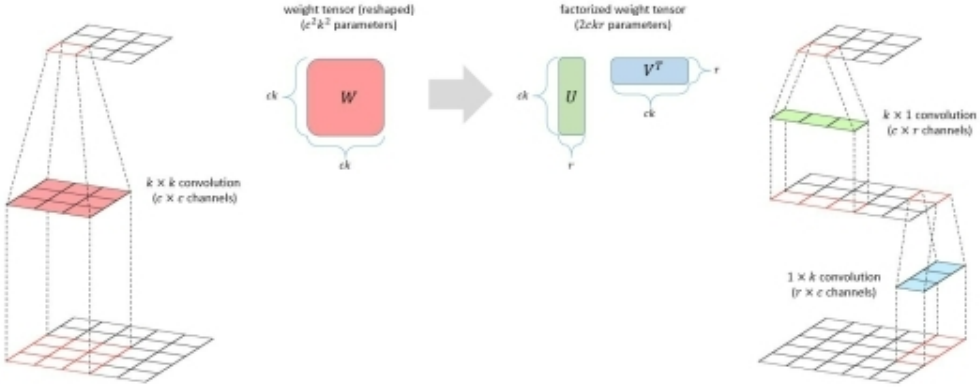


ENet: Efficient Net

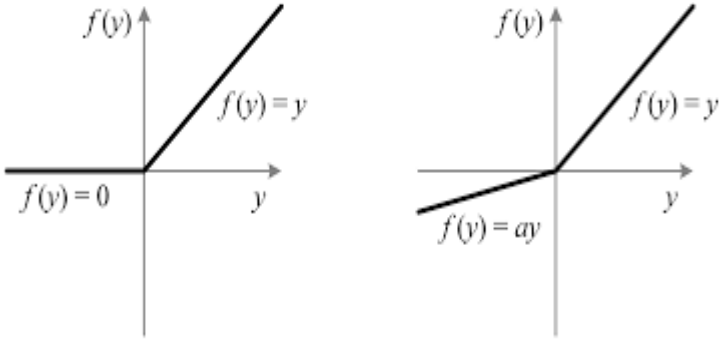
FCN architecture with Encoder/Decoder modules

Chaurasia et al. (2016)

ENet - Differences



- Uses filter factorization
- Uses 1 x 1 convolution operation
- Uses PReLU's

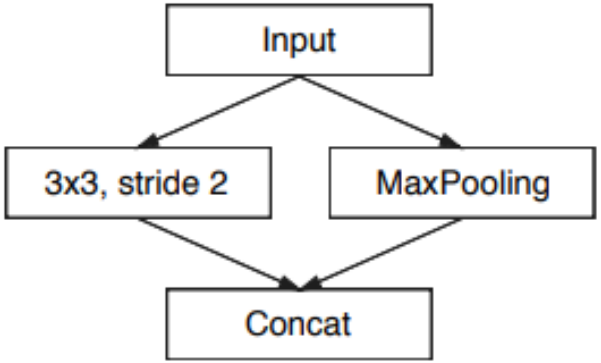


ENet - Architecture

- Initial stage

Early reduction of input size to save computation time

Concatenation of feature maps from 3x3 convolution and max-pooling

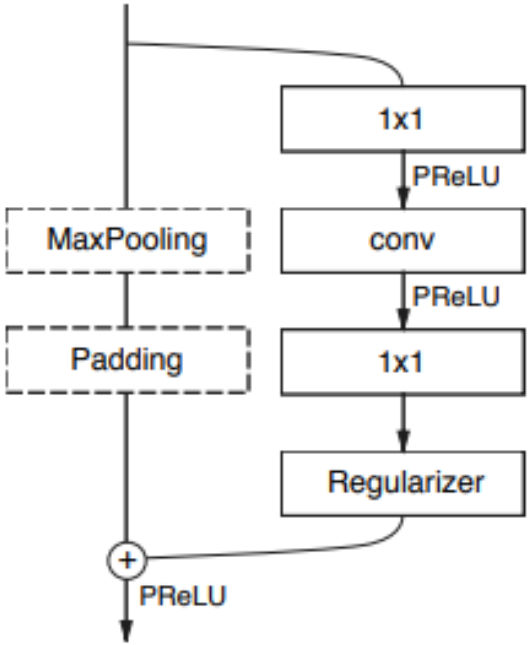


ENet - Architecture

- Encoder

Incorporates bottleneck blocks

Performs downsampling and captures the semantic content



ENet - Architecture

- Decoder

Incorporates bottleneck blocks

Performs upsampling to increase the resolution of feature maps

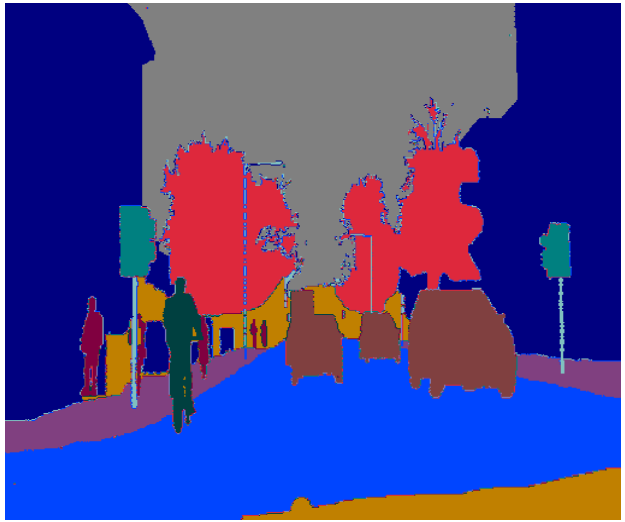
Full convolution layer at the end to produce segmented output

Results

Segmentation outputs

ENet – Segmentation output

Input image

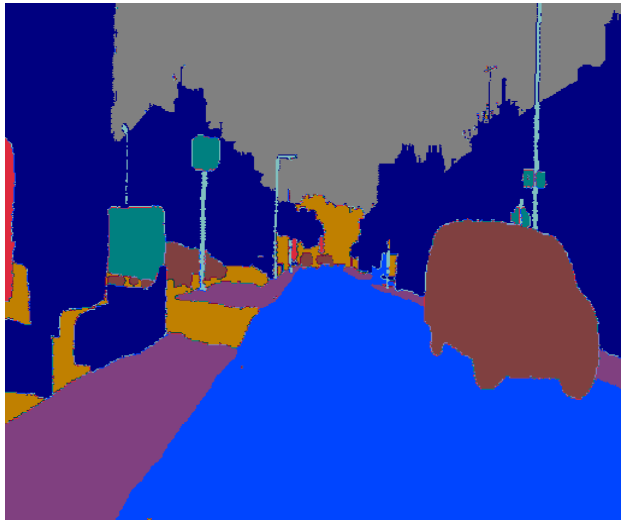


Ground truth

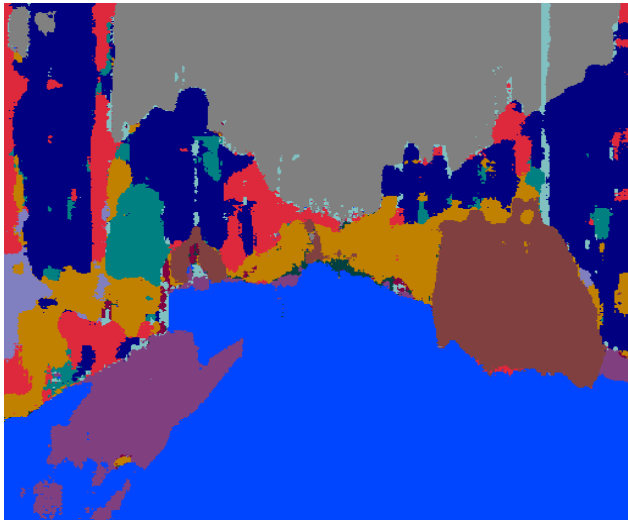
Output

ENet – Segmentation output

Input image



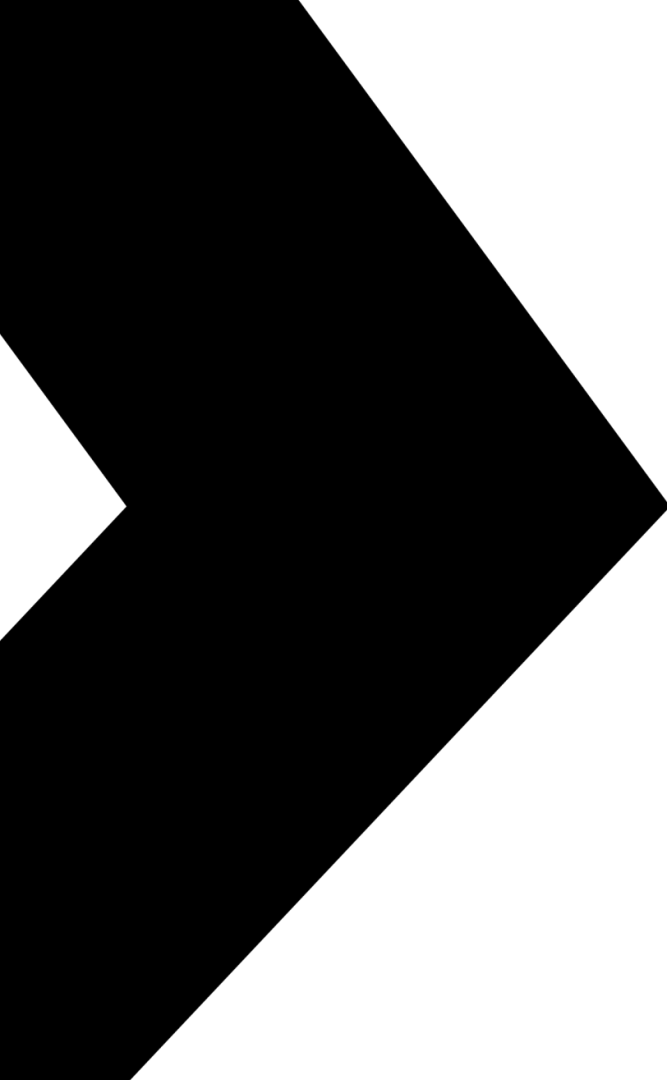
Ground truth



Output

Real-Time Output

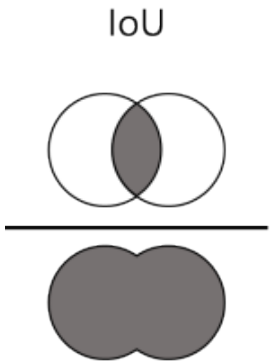
- Video Link

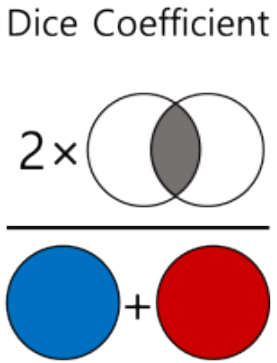


Metrics

Metrics

METRIC	SCORE
Per Pixel Accuracy	75.5%
F1 score	0.639
IoU	0.515





Challenges

- Making model lightweight
- Training a ConvNet on limited GPU memory
- Optimizing hyper-params for performance

Thank You

UNIVERSITY OF MARYLAND FEARLESSLY FORWARD