How to Extract Roads from Satellite Images ?

Why extract roads from satellite images

- Road Extraction and land cover mapping, in general, from satellite imagery, is an important tool for monitoring and efficient map generation for an intelligent transportation system, automobile navigation and emergency support in times of natural disasters.

- Automating road extraction plays an important role in dynamic spatial development and plays an important role in large scale mapping and urban planning.

How can we do this ?

Image segmentation methods for automated road extraction and can be broadly classified into -

- **classical computer vision algorithms**: Traditional image processing techniques for image segmentation such as thresholding and edge detection have been successful in mapping outlines of the roads but require extensive manual intervention
- **machine learning algorithms:** we have seen a wide adoption of deep learning technologies owing to acceleration in development of computing technologies and promising scientific research in machine learning in the last decade.

How are we going to do it ?

We use deep learning methods trained with labelled satellite images for automated labelling of road pixels on aerial images with semantic labels.

Dataset

We use Massachusetts Roads Dataset consists of 1171 aerial images of the state of Massushussets:

- Size: 1500 x 1500 pixels
- Number of Images: 804 labelled images in the training set which is split into a set of 643 images for training and 161 for validation.

Methodology

Method

We use deep learning based semantic segmentation network architectures to train satellite images labelled with masks at a pixel level.

A) Data Preparation
B) Data Augmentation
C) Loss Functions
D) Schedulers and Optimizers
E) Network Architectures.

Data Preparation

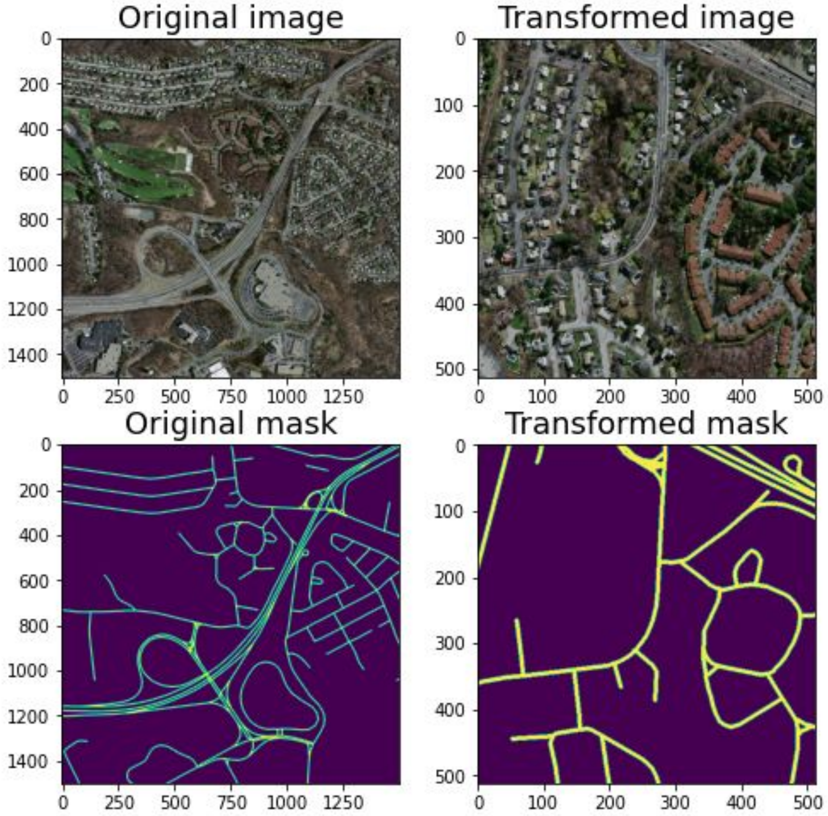
Normalisation : The pixel intensity values in the images, present in 8-bit format, were normalised to bring the pixel values in the range 0-1 by dividing by 255.

Centering : Further, the images processed by subtracting per-channel mean from pixel values calculated on the training set

Standardisation : Finally, the values obtained by subtracting per-channel mean were divided by per-channel standard deviation calculate on the training set

Data Augmentation

- Random Crop - Images are randomly cropped at 512 x 512 pixel size
- Random Horizontal Flip (with probability 0.5)
- Random Vertical Flip (with probability 0.5)
- Random Rotate (with probability 0.5)
- Transpose (with probability
- Random Shift-Scale-Rotate
- Random Brightness and Contrast added/removed
- Random Gamma transformations
- Random Blur (with probability of 1%)

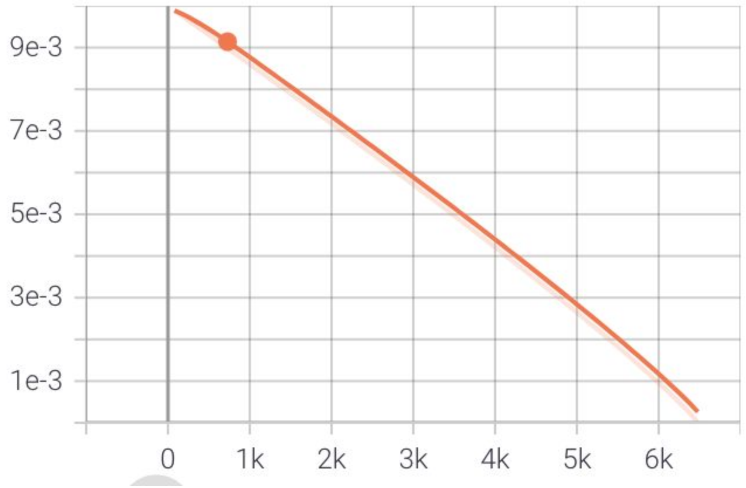


Loss Functions

- Categorical Cross Entropy Loss
- Weighed Categorical Cross Entropy Loss
- Focal Loss
- Dice Loss

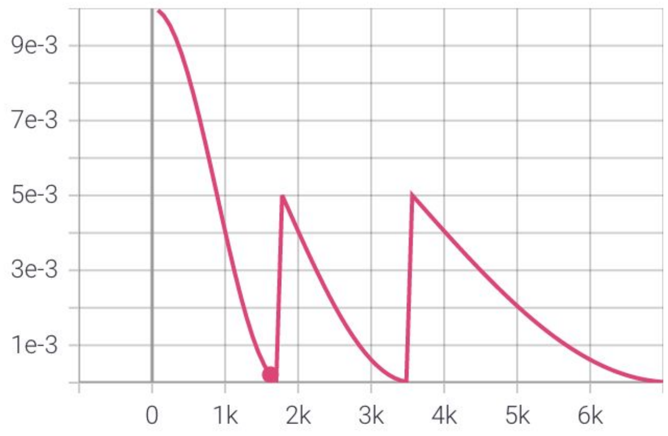
Scheduler and Optimizers

1) Stochastic Gradient Descent (Optimiser) with Polynomial LR (Scheduler)



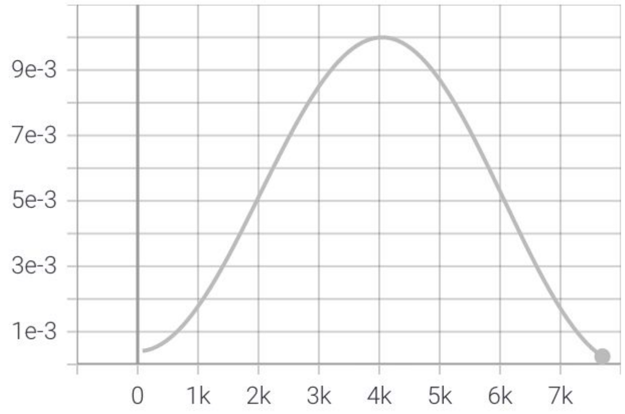
Scheduler and Optimizers

2) Stochastic Gradient Descent (Optimiser) with Warm Restarts (Scheduler)



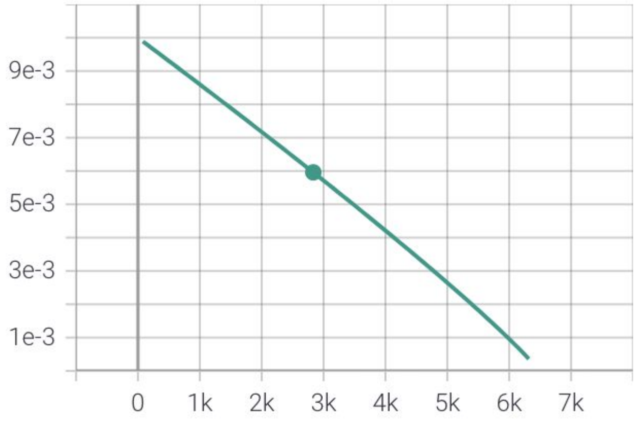
Scheduler and Optimizers

3) Stochastic Gradient Descent (Optimiser) with One Cycle LR (Scheduler)



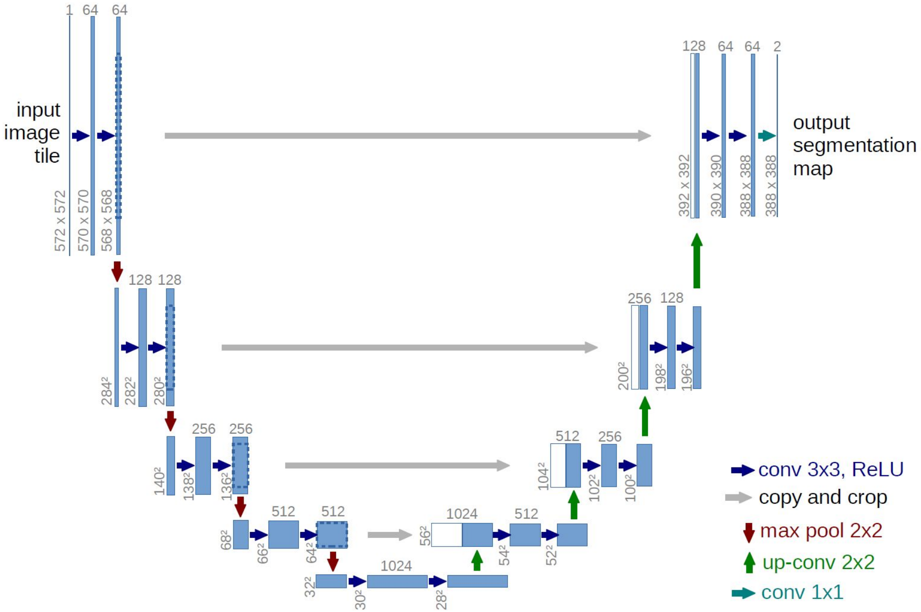
Scheduler and Optimizers

4) Adam Optimiser (max lr decaying by polynomial factor)



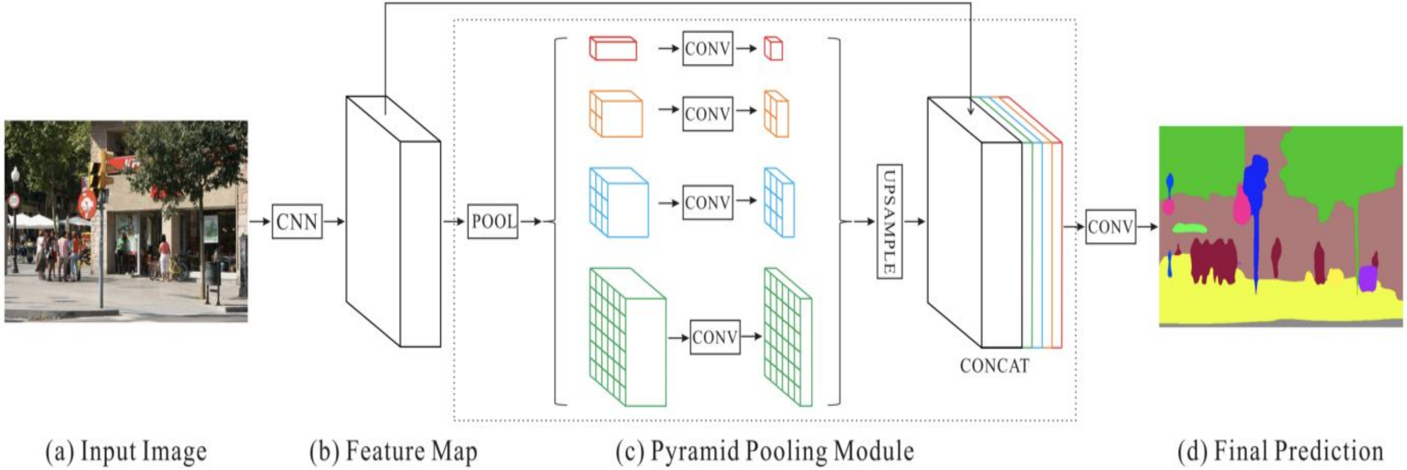
Network Architectures

- U-Net



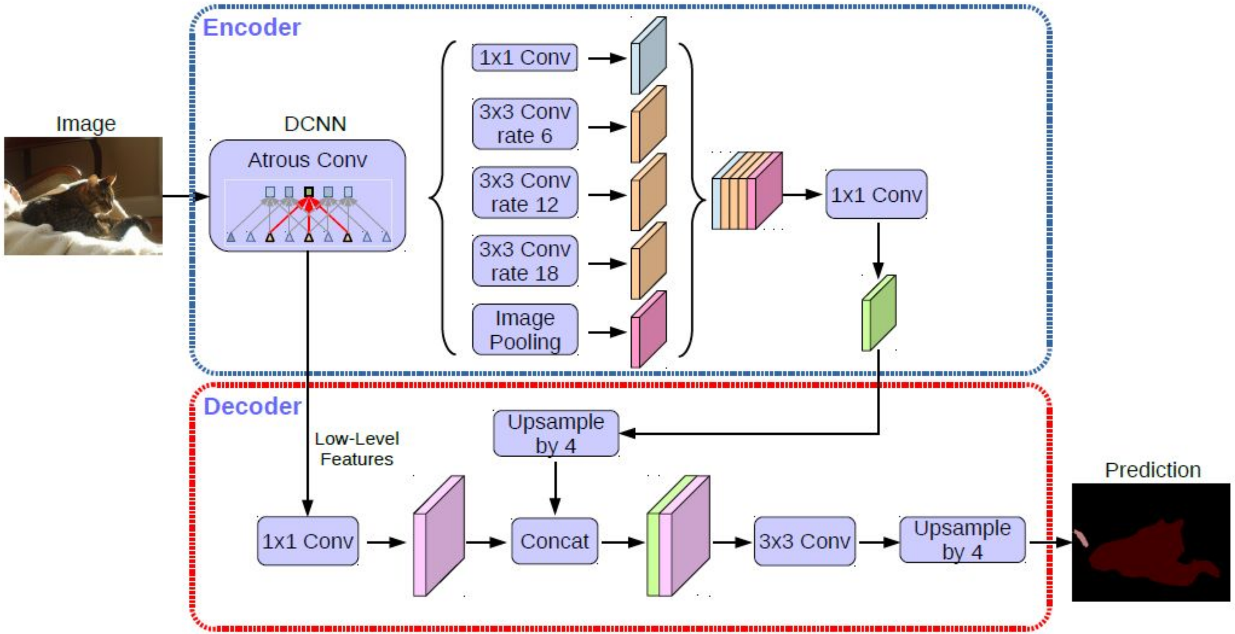
Network Architectures

- PSPNet (Resnet-50 backbone)



Network Architectures

- Deeplabv3+



Metrics

- Intersection over Union (IOU)
- Pixel Accuracy
- Dice Score (or F1 score)
- Precision-Recall

Experiments and Results

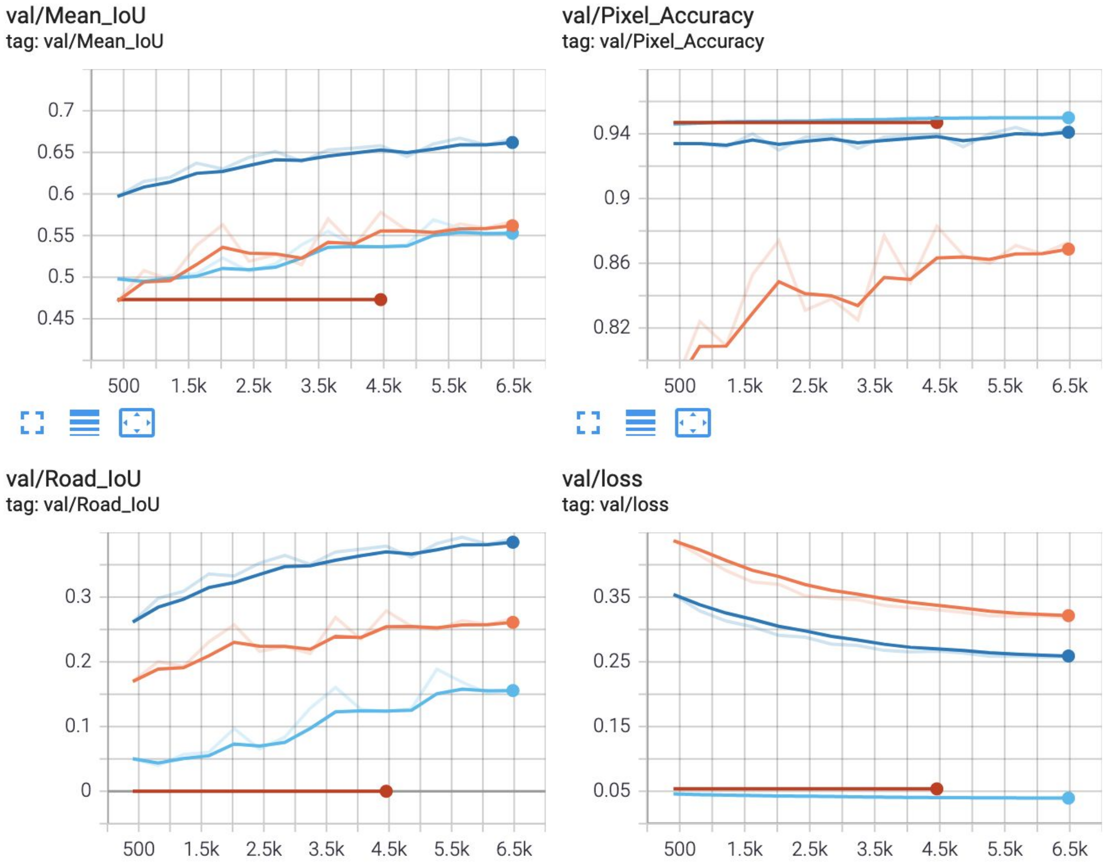
Choice of Loss Function

Experiments

Over Experiment 1, Experiment 2 (weighted cross entropy loss with ratios 1:20 and $1/\sqrt{20}$), Experiment 3 (dice loss) and Experiment 4 (focal loss), the loss functions were varied with a Unet with Resnet-50 backbone.

Results

- During an initial analysis, it was observed that the ratio of pixels belonging to the road and background is 1:20 (class imbalance). The weighing of the cross entropy loss was experimented with and it was observed that using inverse square root of pixel distributions exceeded results on the Class IOU (Road IOU) being monitored
- It was found to be performing much better than Dice Loss and Focal Loss. It can be argued that weighed cross entropy loss performs better than dice loss and focal loss because the former optimised for the overall segmentation IOU
- focal loss has been reported to perform better when the class imbalance is of the order of 1:1000



Exp. 1 (Weighted CE - 1:20) - Orange, Exp. 2 (Weighed CE - 1:5) - Blue,
Exp. 3 Dice Loss, Exp. 4 - Focal Loss

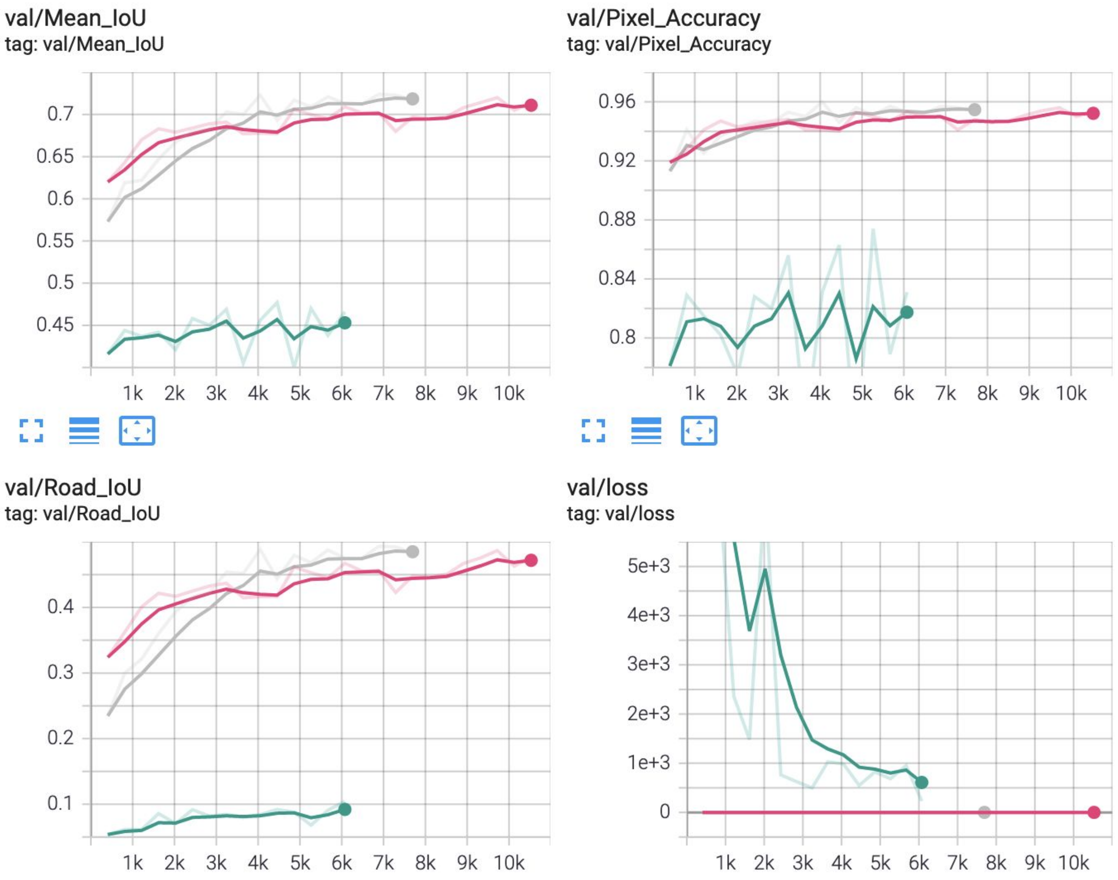
Choice of Optimizer and Scheduler

Experiments

Over Experiment 5, Experiment 6 and Experiment 7, the performance of Stochastic Gradient Descent with Cosine Annealing Warm Restarts (cyclic LR) , SGD with One CycleLR (superconvergence) and Adam were compared using Resnet-50 backbone of Unet and cross entropy loss weighted in 1:5 proportions

Results

- Warm Restarts Scheduler (Exp 5; Pink) converges much faster than One Cycle LR policy.
- With careful selection of hyperparameters, both the schedulers seem to converge to, approximately, similar values.
- The performance of Adam pales in comparison to the above experiment (probably due to high max LR).



Exp 5 : Pink - SGD with Warm Restarts, Exp 6: Grey - SGD with One Cycle LR and Exp 7 : Grey : Adam optimizer - Validation Plots

Choice of Network Architecture

Experiments

Over Experiment 5, Experiment 8 and Experiment 9, using SGD with Warm Restarts Scheduler and Weighed CE Loss (1:5), performance of UNet, PSPNet and Deeplabv3+ (all Resnet-50 backbones) were compared in Experiments 5, 8 and 9 respectively

Results

- We observe that performance of Deeplabv3+ far exceeds UNet and PSPNet on class-wise IOU calculated, pixel accuracy, dice score and recall

Overall Summary

Table 1: Experiment Summary - Table 1

Experiment	Architecture	Loss	Optimizer Scheduler
Exp 1	Unet-Resnet50	Weighted CE (1:20)	SGD with Poly LR scheduler
Exp 2	Unet-Resnet50	Weighted CE (1:5)	SGD with Poly LR scheduler
Exp 3	Unet-Resnet50	Dice Loss	SGD with Poly LR scheduler
Exp 4	Unet-Resnet50	Focal Loss	SGD with Poly LR scheduler
Exp 5	Unet-Resnet50	Weighted CE (1:5)	SGD with Cosine Annealing Warm Restarts
Exp 6	Unet-Resnet50	Weighted CE (1:5)	SGD with OneCycle LR
Exp 7	Unet-Resnet50	Weighted CE (1:5)	Adam
Exp 8	PSPNet	Weighted CE (1:5)	SGD with Cosine Annealing Warm Restarts
Exp 9	Deeplabv3+	Weighted CE (1:5)	SGD with Cosine Annealing Warm Restarts

Table 2: Experiment Results - Table 2

Experiment	Road IOU	Pixel Accuracy	Dice Score	Precision	Recall
Exp 1	0.283	0.847	0.440	0.296	0.859
Exp 2	0.396	0.927	0.565	0.488	0.678
Exp 3	8.4e-13	0.929	8.4e-13	1.00	8.45
Exp 4	0.2098	0.934	0.336	0.600	0.248
Exp 5	0.502	0.945	0.667	0.583	0.783
Exp 6	0.497	0.932	0.657	0.579	0.781
Exp 7	0.104	0.839	0.187	0.144	0.2761
Exp 8	0.4528	0.9380	0.6222	0.5423	0.737
Exp 9	0.5101	0.9450	0.6742	0.574	0.8201

Visual Analysis

Visualization of Predictions and Comparison.
Image (left), Ground Truth (center) and Prediction (right)

Post-processing (Discussion only)

We can observe in the visualisation of predictions, there are certain “open holes”and breaks in a continuous stretch of road as documented in the above predictions. To obtain a road network, it could be an important exercise to fill in the blind spots in the mask extracted and thus obtain the masks.

As noted in the paper, *City-Scale Road Extraction from Satellite Imagery v2*, the authors use morphological transformations (erosion and dilation) in image processing for post-processing. The process of erosion involves removal of foreground object and process of dilation involves adding foreground object around the boundary.While, erosion is used for diminish the features, dilation is used for accentuate features.

a) Morphological Opening - The opening operation erodes an image and then dilates the eroded image. This could be useful for noise removal

b) Morphological Closing - The closing operation dilates an image and then erodes the dilates image.. It could be useful for closing small holes in the pixels of the road.

Conclusion

- Satellite imagery analysis is crucial for intelligent remote monitoring and building system to generate efficient routes. The use cases for the technology spans across humanitarian to military.methods to road segmentation maps from satellite images has been discussed.
- It was observed that using carefully tuned weights in cross entropy loss, hyperparameters in the optimiser and scheduler (SGD with warm restarts) with Deeplabv3+ yields the best performance.
-