Road Detection from Remote Sensing Imagery

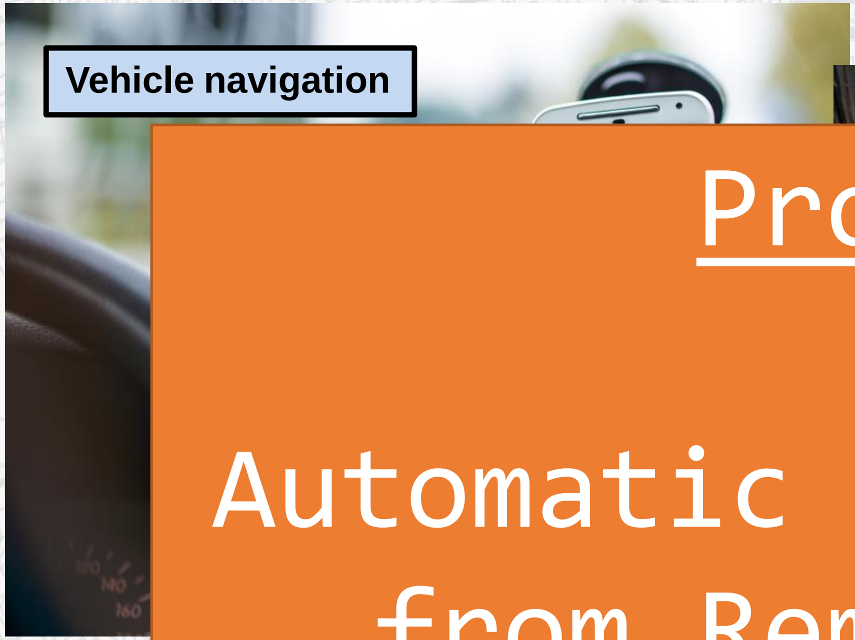
Pantelis Kaniouras

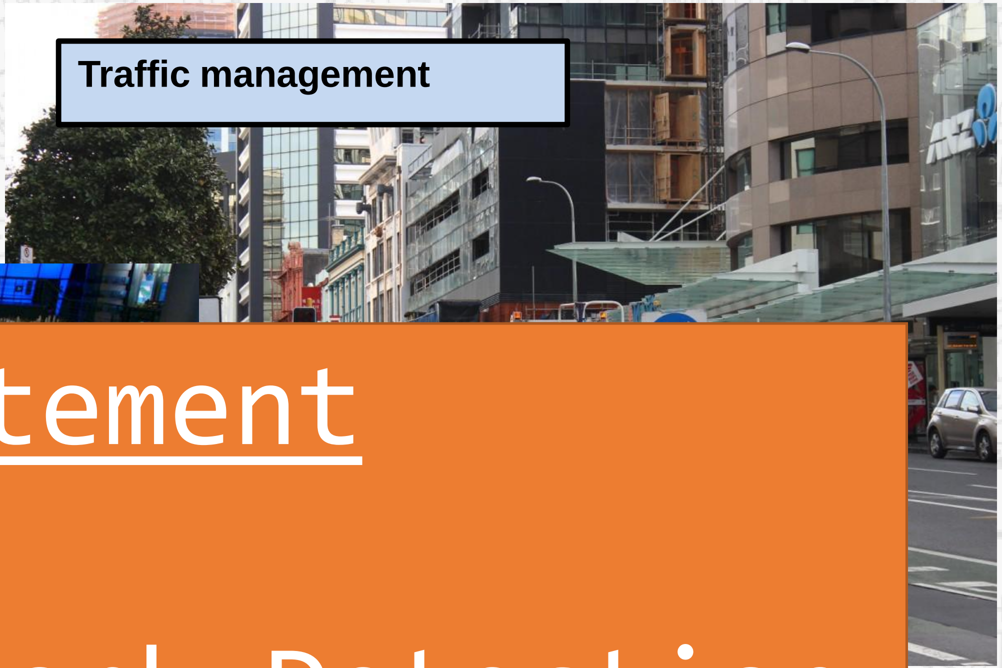
Mentor #1: Liangliang Nan
Mentor #2: Roderik Lindenbergh
Mentor #3: Frido Kuijper (TNO)

Presentation Contents

1. Introduction & Related Work

1.1 Motivation





Problem statement

Automatic Road Network Detection from Remote Sensing imagery

1.2 How is road detection solved?

Conventional methods:

1. l-wise cl fication
2. K methods
3. Mat rphologic s
4. e conto els
5. A combination of the above

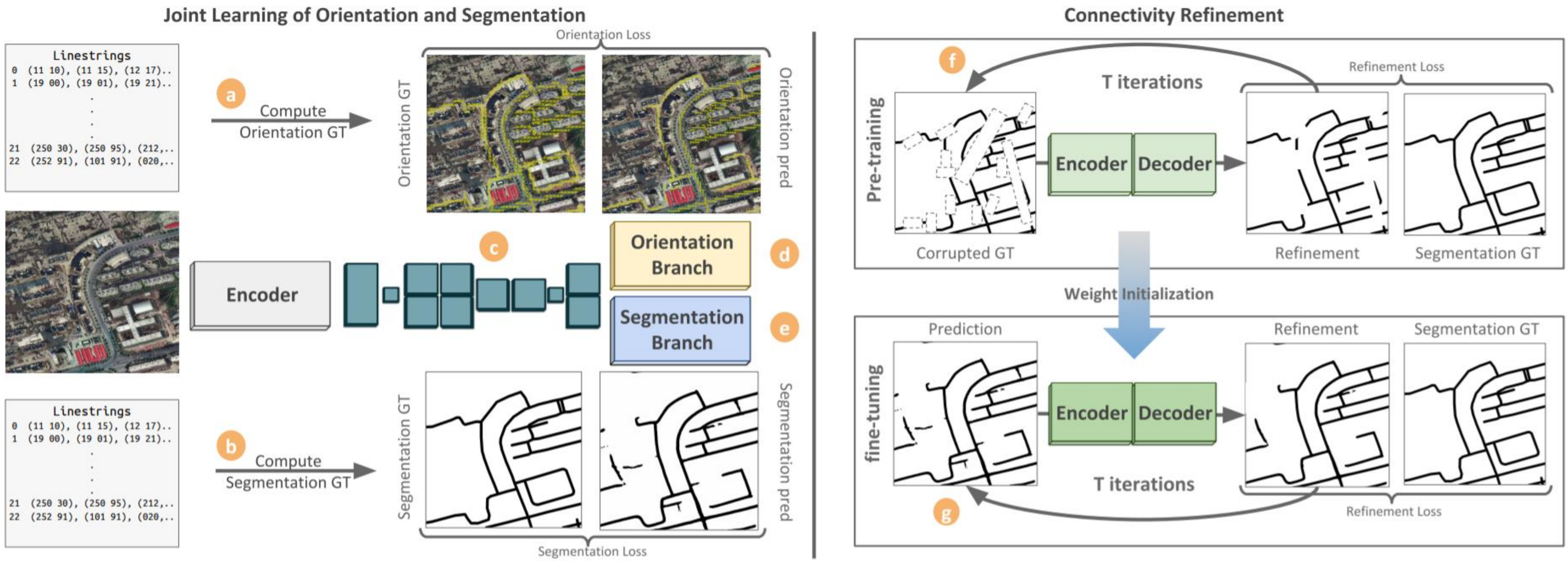
DEEP LEARNING

How to solve road detection using Deep Learning?

Convolutional Neural Networks

1. **Semantic Segmentation**
2. **Vector Graph Extraction**

1.3 Semantic Segmentation: Example work



Retrieved from Improved Road Connectivity by Joint Learning of Orientation and Segmentation

1.4 Vector Graph Extraction: Example work

Incremental Graph Construction:
CNN Training

Retrieved from RoadTracer: Automatic Extraction of Road Networks from Aerial Images

1.5 Errors of Deep Learning Models

Ground truth   Prediction result

1.6 Observations

1. Incorrect ground truth
2. Shadows
3. Similar texture
4. Atmospheric conditions
5. Obstacles blocking sight – occlusion
6. Seasonality
7. Wrong heuristics
8. Prior knowledge not taken into consideration

Prior Knowledge == Road Properties

- ***curvilinear** objects*
- ***continuous** objects*
- ***consistent** (texture)* objects
- ***small ratio** of representation in imagery*

2. Proposed Methodology

2.1 Goal

Find a way to utilize as much prior knowledge as possible and improve the performance of the road detection task

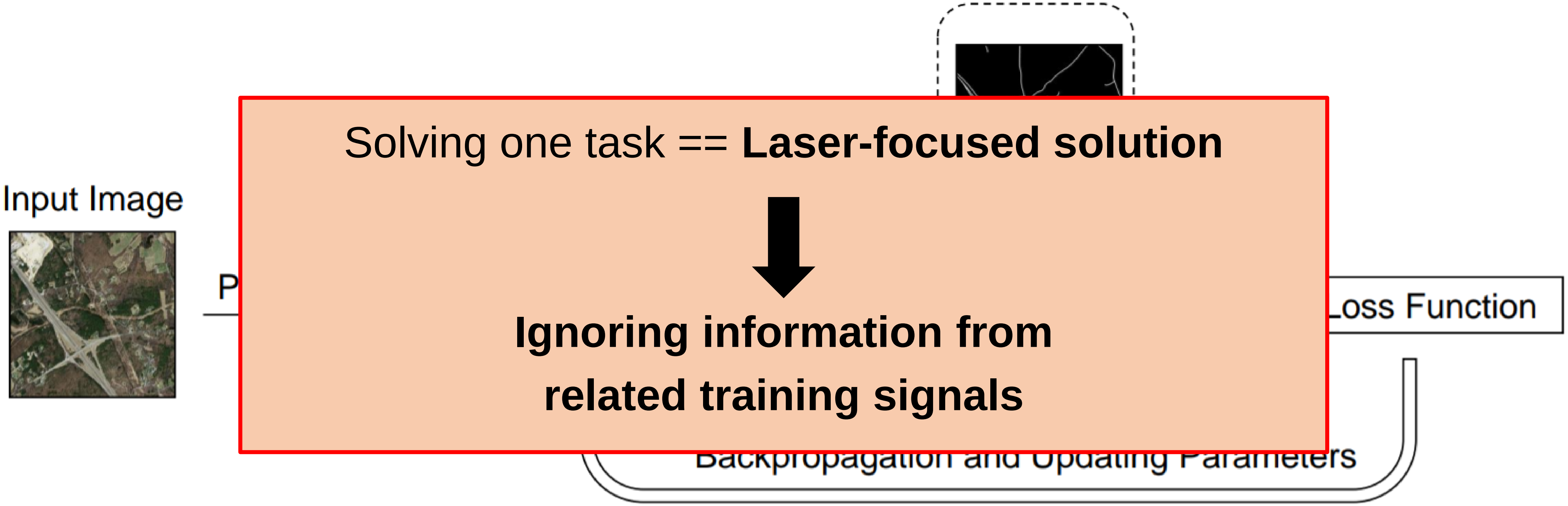
2.2 Research question

Main: ***To which extent is it possible to utilize road properties to improve road detection from remote sensing imagery using deep learning techniques?***

Sub-questions

- *Can prior knowledge be incorporated as a constraint into a deep learning model? If yes, how?*
- *What are the limitations of a model that combines concepts of different models into one, unified model?*
- *Can prior knowledge improve road detection?*

2.3 Conventional Approaches



Solving one task == **Laser-focused solution**

Ignoring information from related training signals

Retrieved from Road Extraction by using atrous spatial pyramid pooling integrated encoder-decoder network and structural similarity loss

2.4 Proposed Method

Shared Knowledge → **Multi-Task Learning (MTL)**

Why does MTL work?

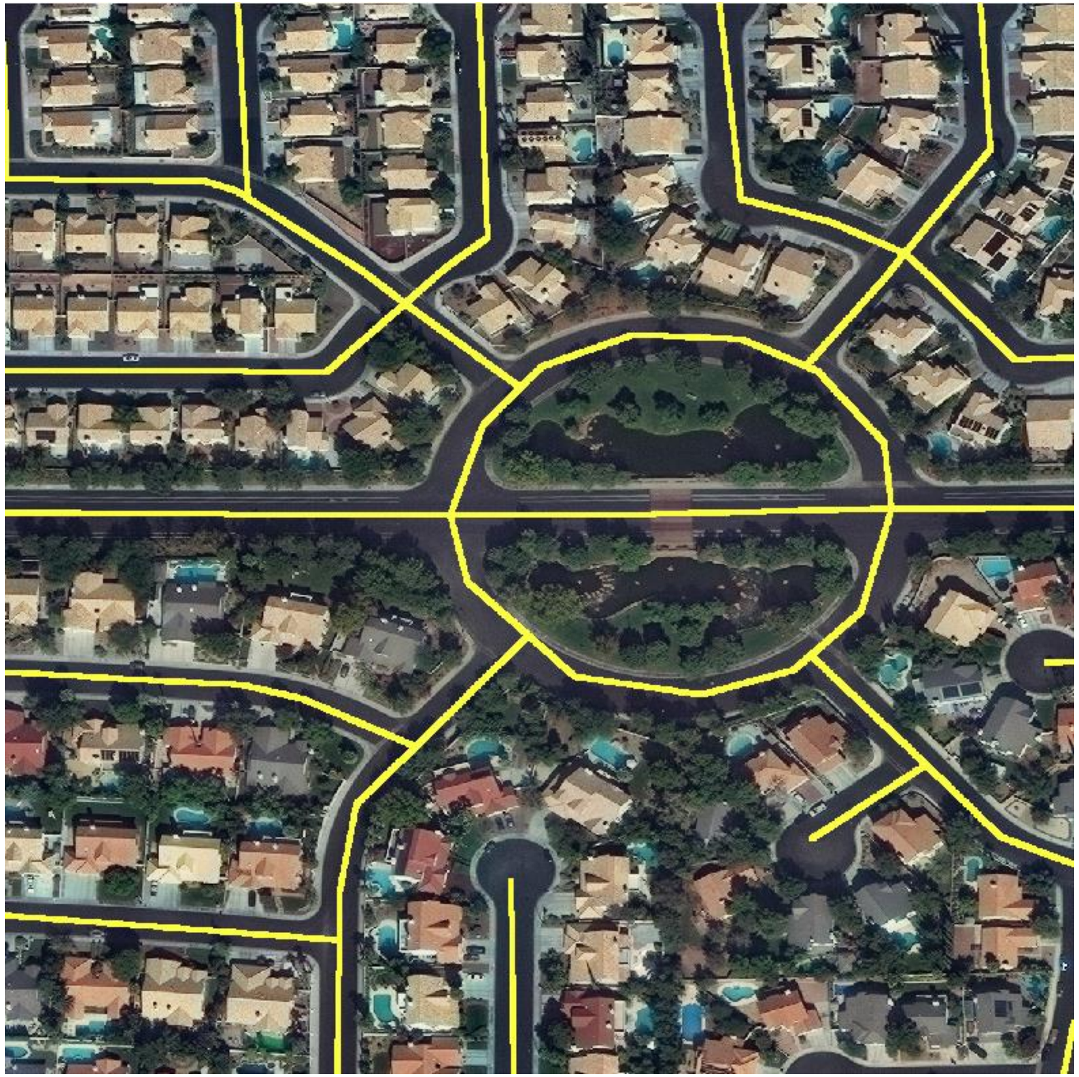
- Implicit data augmentation
- Attention focusing
- Eavesdropping
- Representation bias
- Regularization

2.6 Related tasks

Proposed tasks that might assist the primary task of
Road Detection Learning

2.6.1 Learning tasks

Road Detection Learning Task

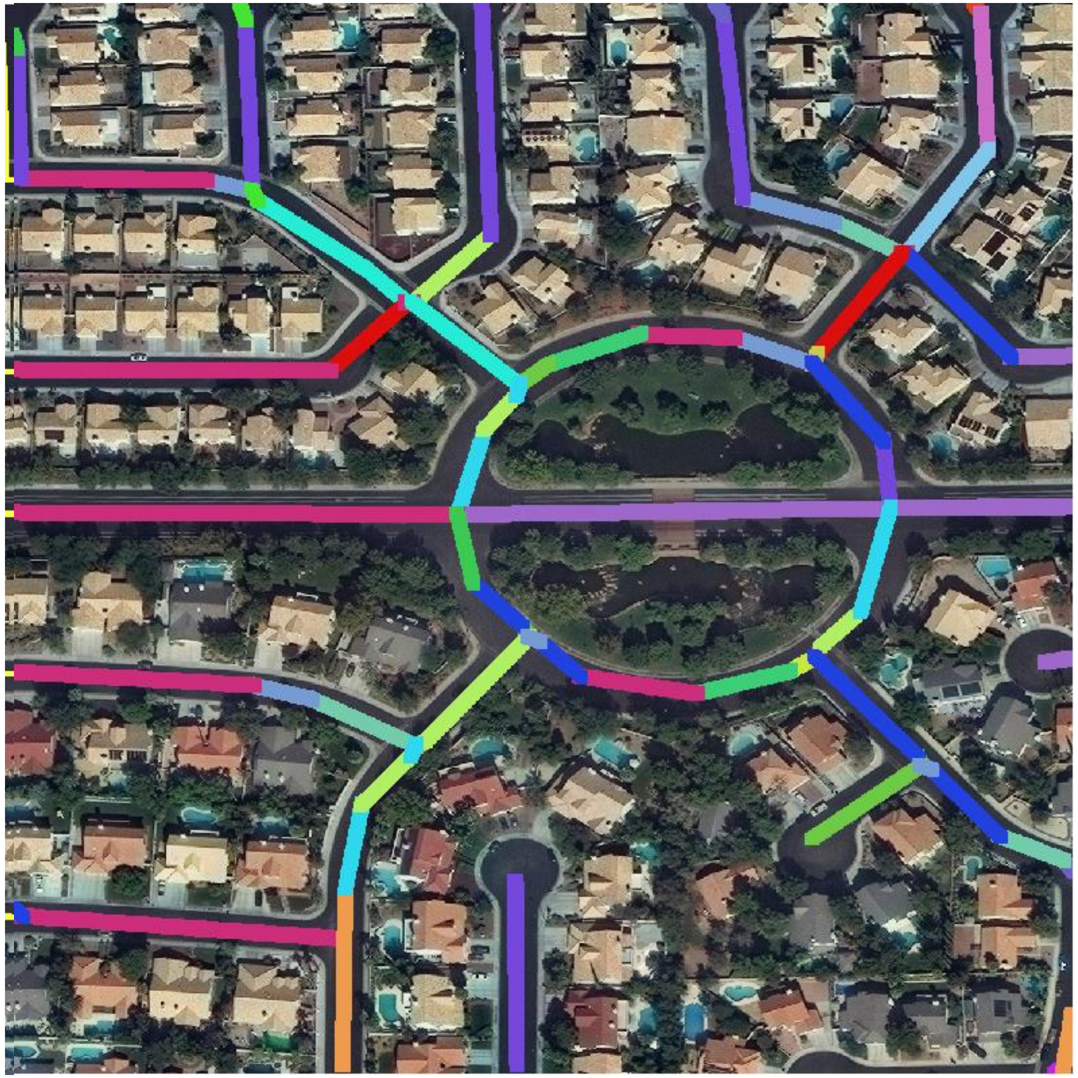


2.6.2 Learning tasks

Road Intersection Learning Task

2.6.3 Learning tasks

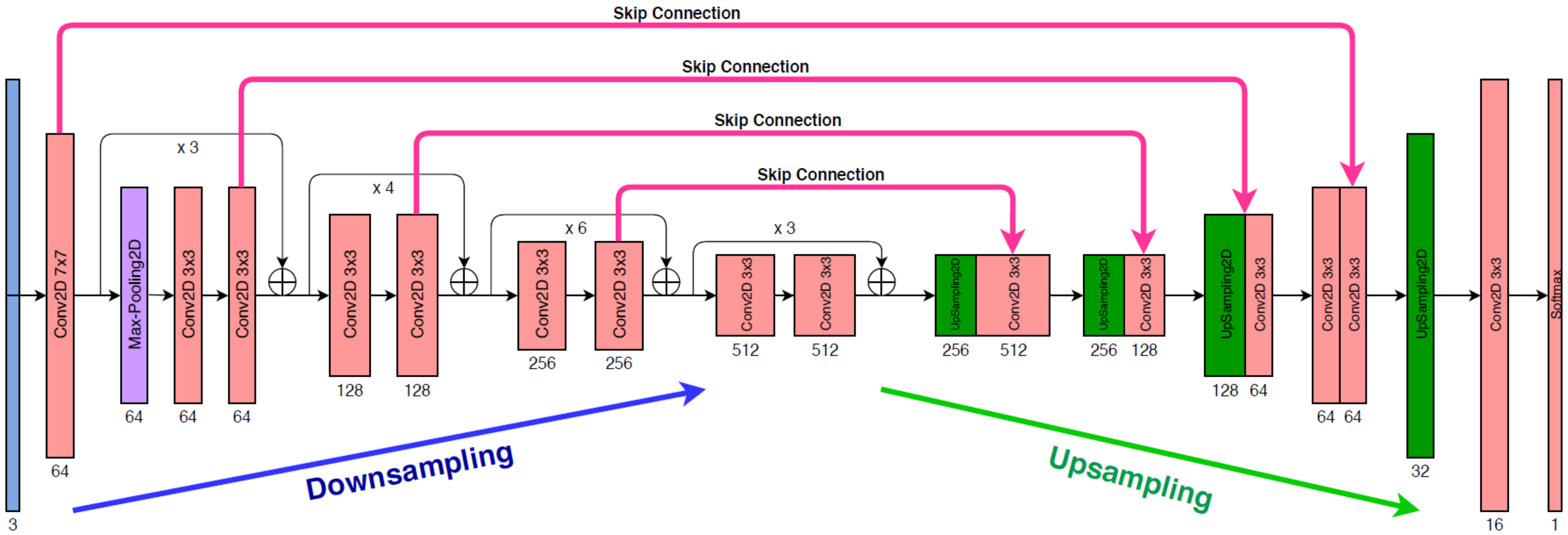
Road Orientation Learning Task



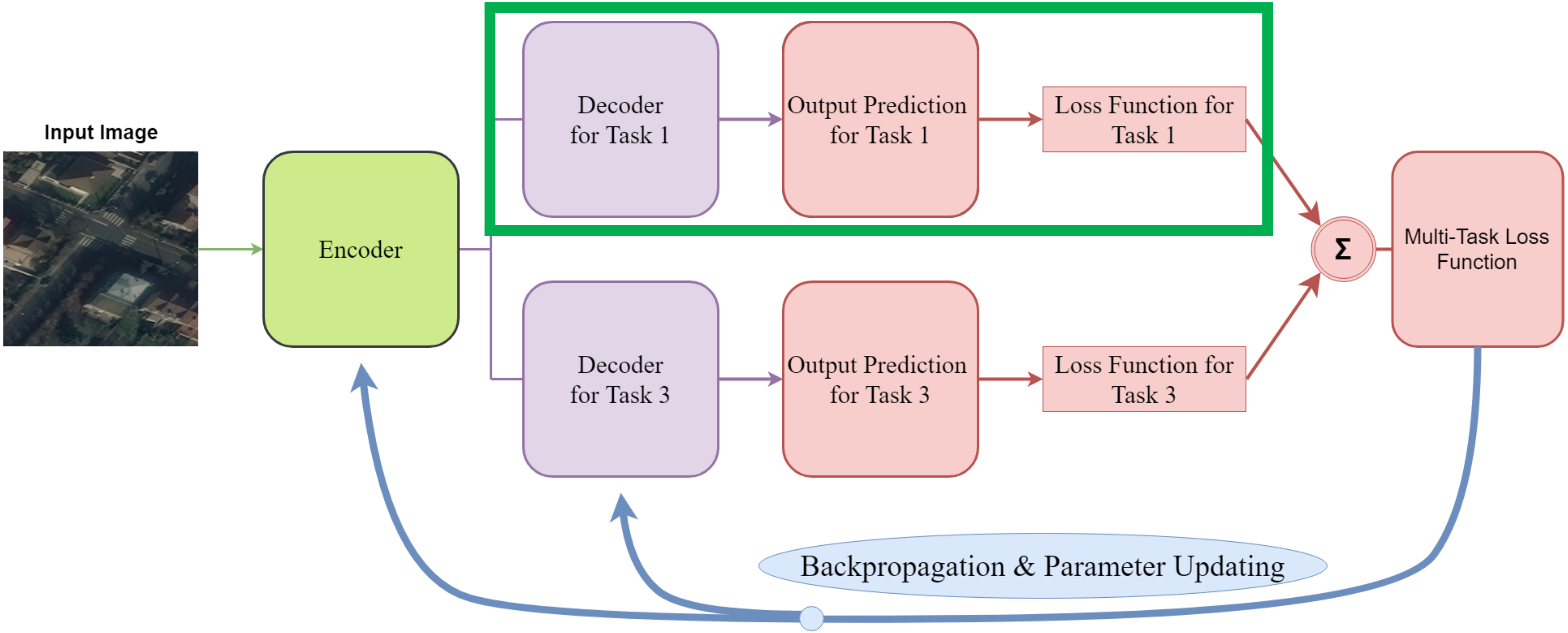
2.6.4 Learning tasks

Gaussian Road Mask Learning Task

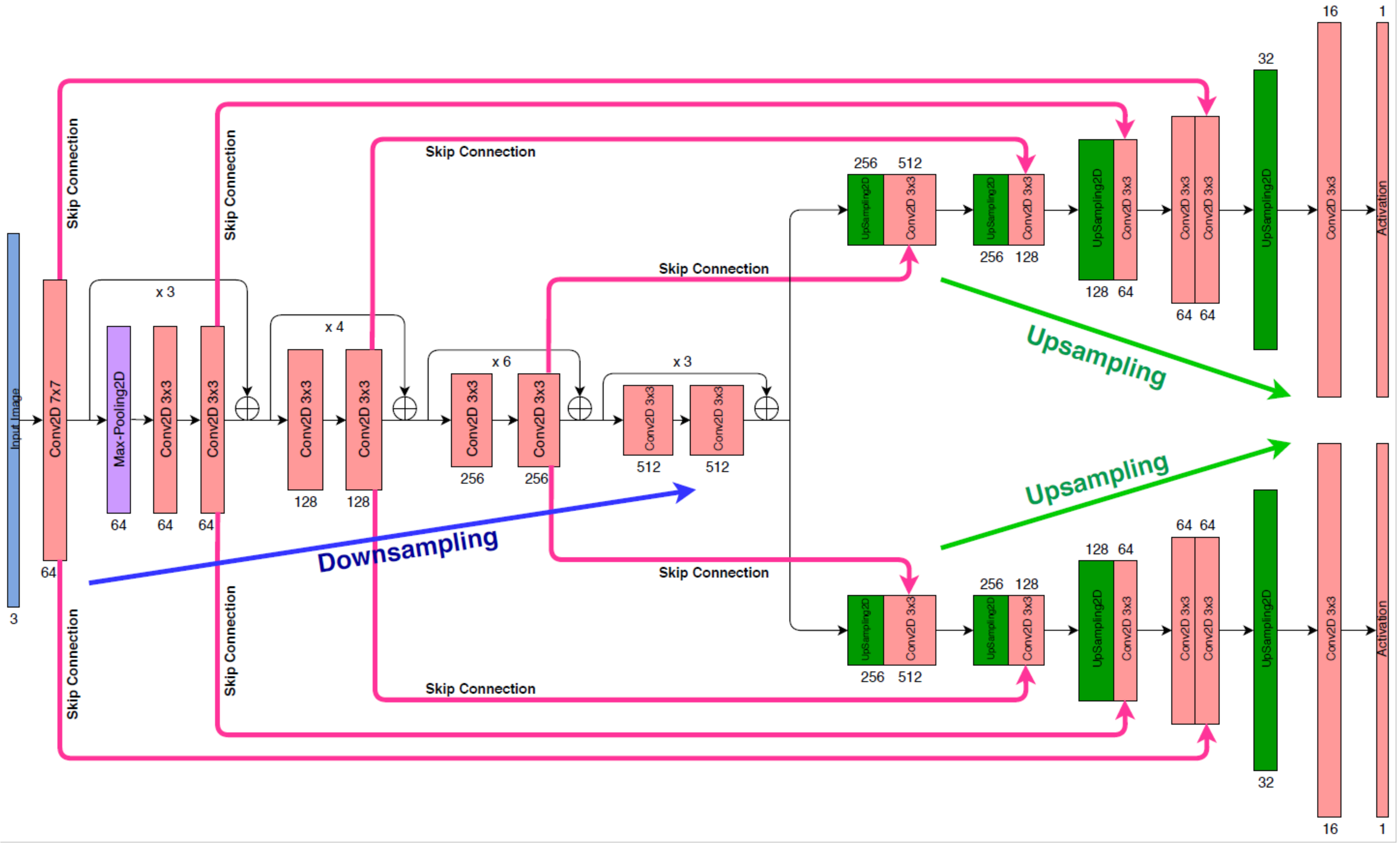
2.7 Network Architecture



2.8 Proposed 3-task-solving model



2.8 Proposed 3-task-solving model



2.9 Loss Function

$$\boxed{Loss = \lambda_1 * CE + \lambda_2 * (1 - IoU)}$$

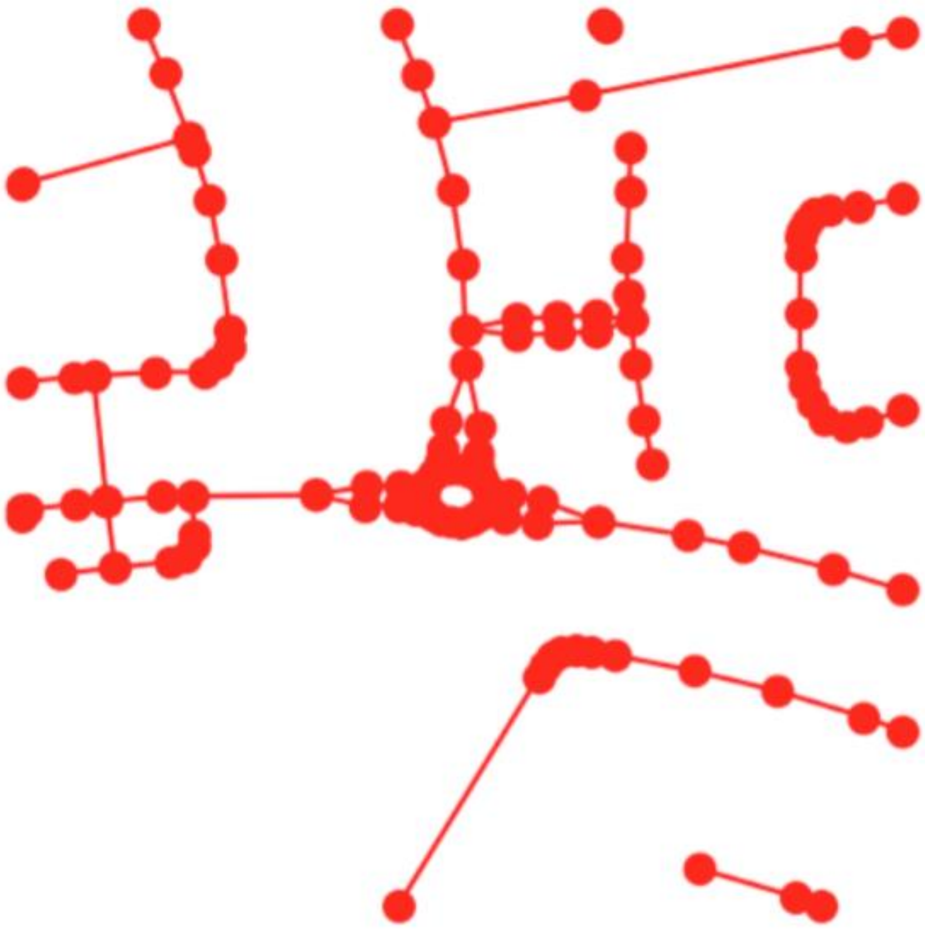
where,

λ_1, λ_2: weight parameters

CE: Cross-Entropy Loss (either Categorical or Binary)

IoU: Intersection over Union

2.10 Dataset: SpaceNet



Retrieved from [SpaceNet: A Remote Sensing Dataset and Challenge Series]

2.11 Implementation details

Hyperparameters

- Learning rate: 0.001
- Reduce on plateau by: 0.1
- Batch size: 15
- Input image size: 256x256
- Pre-trained weights: ImageNet
- Optimizer: Adam

Hardware & Software

- Programming language: Python Programming Language
- GPU: NVIDIA GeForce GTX 1080 Ti (provided by the HPC Cluster)
- Framework: Keras (TesnorFlow as Backend)

2.12 Evaluation Methods

Qualitative evaluation

Visual inspection

Quantitative evaluation

Topology

clDice (centerline-in-mask-dice-coefficient)

Segmentation

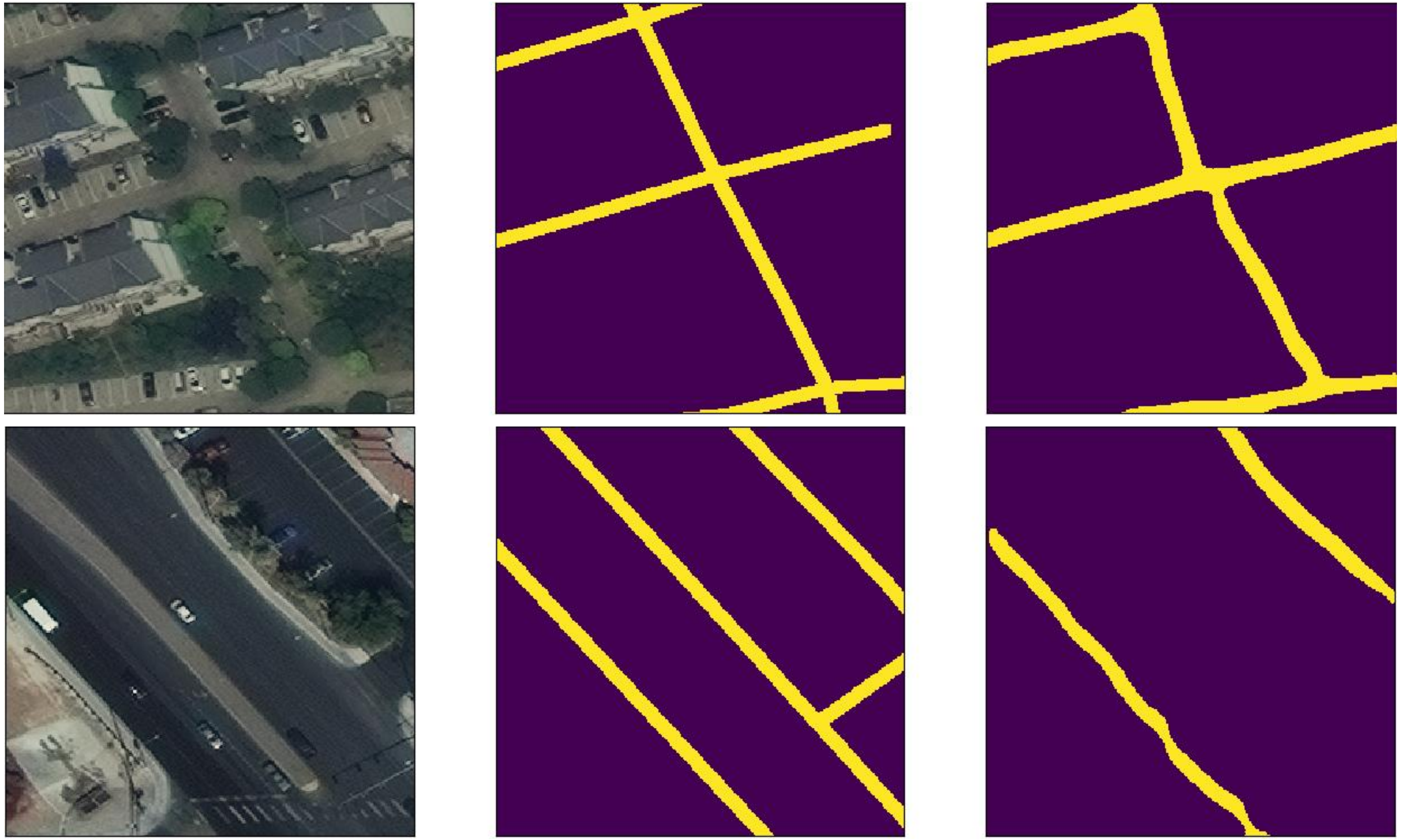
Intersection over Union &
F1-Score

3. Experiments & Analysis

3.4 Results of Single-Task solving models

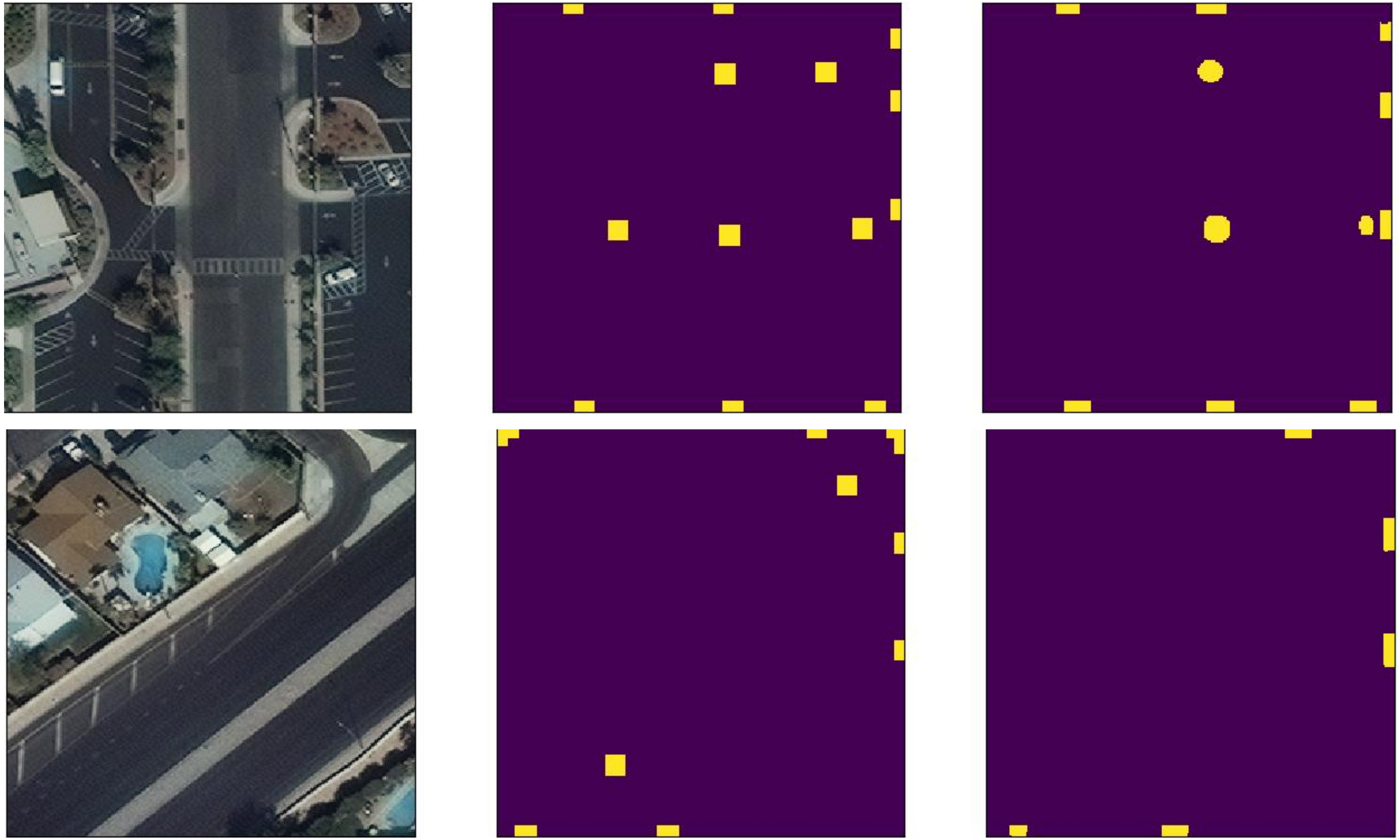
Train proposed tasks on their own and identify errors and limitations

Results: Road Detection Learning task (primary)



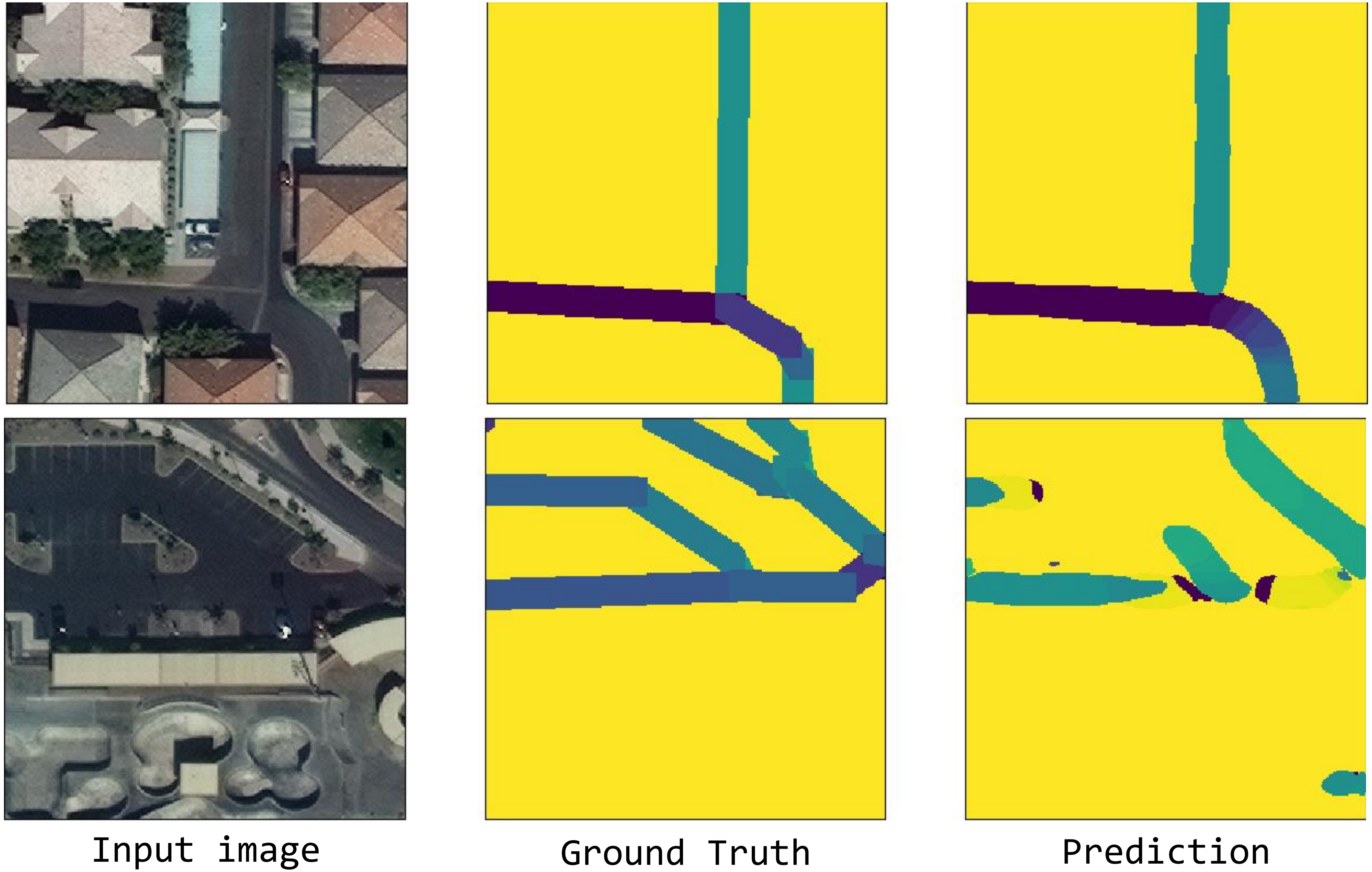
Input image   Ground Truth   Prediction

Results: Road Intersection Learning task

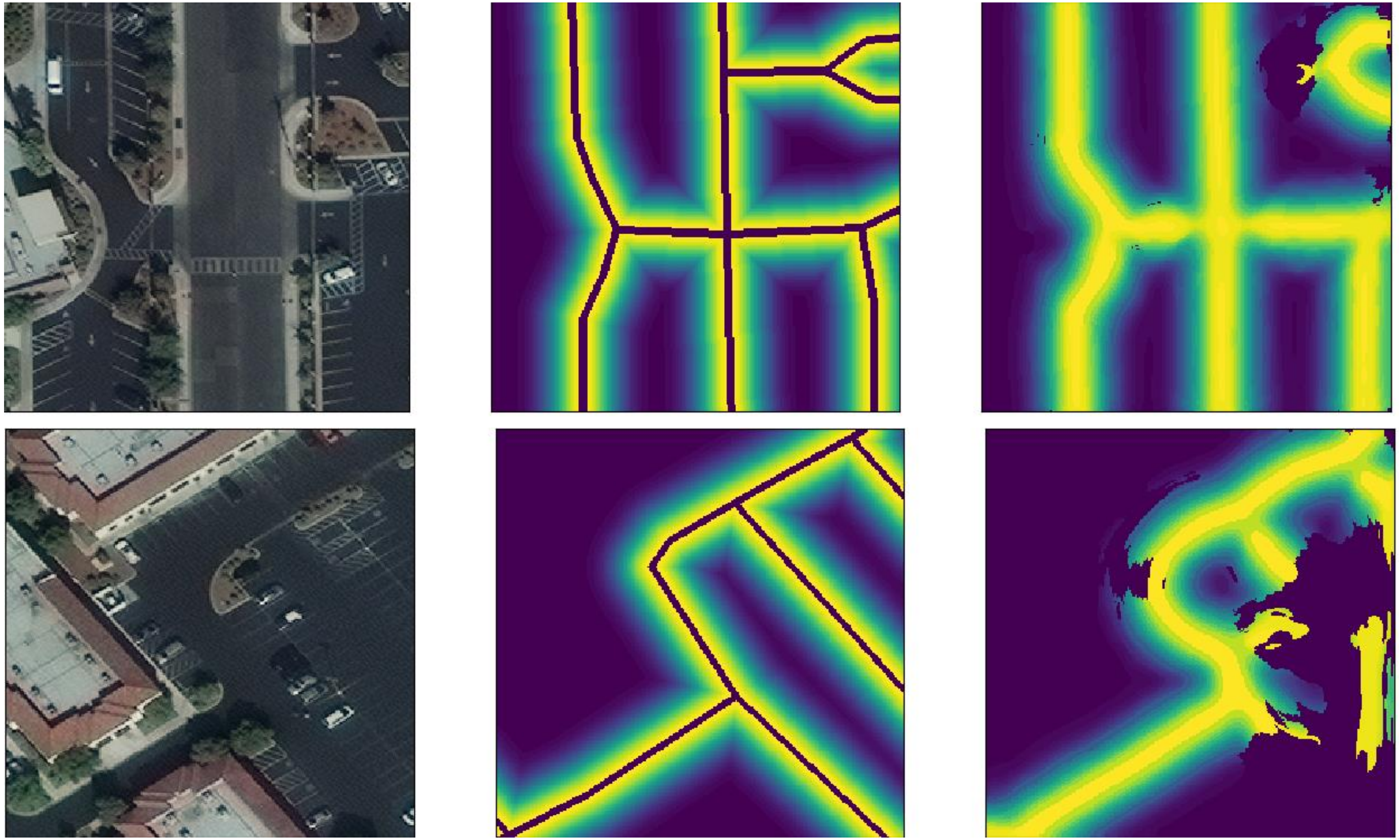


Input image Ground Truth Prediction

Results: Road Orientation Learning task



<u>Results:</u> Gaussian Road Mask Learning task

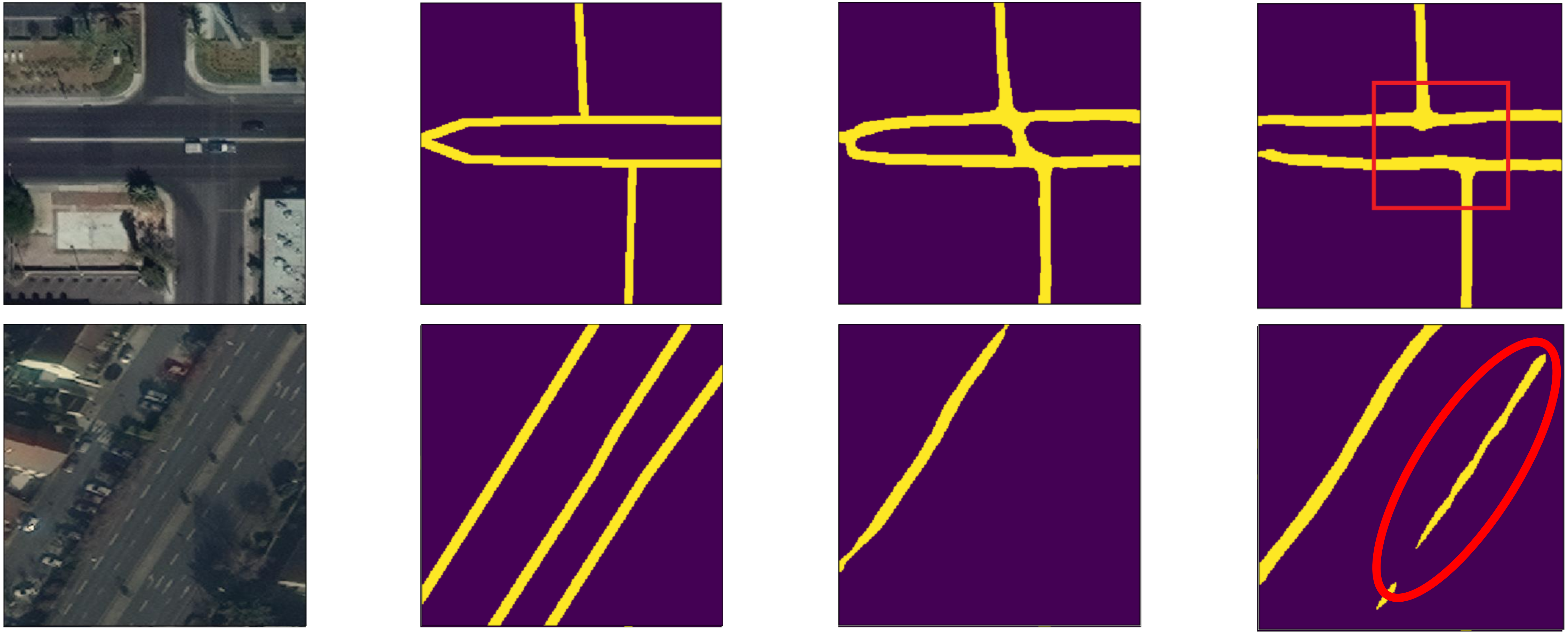


Input image Ground Truth Prediction

3.5 Results of Multi-Task Learning models

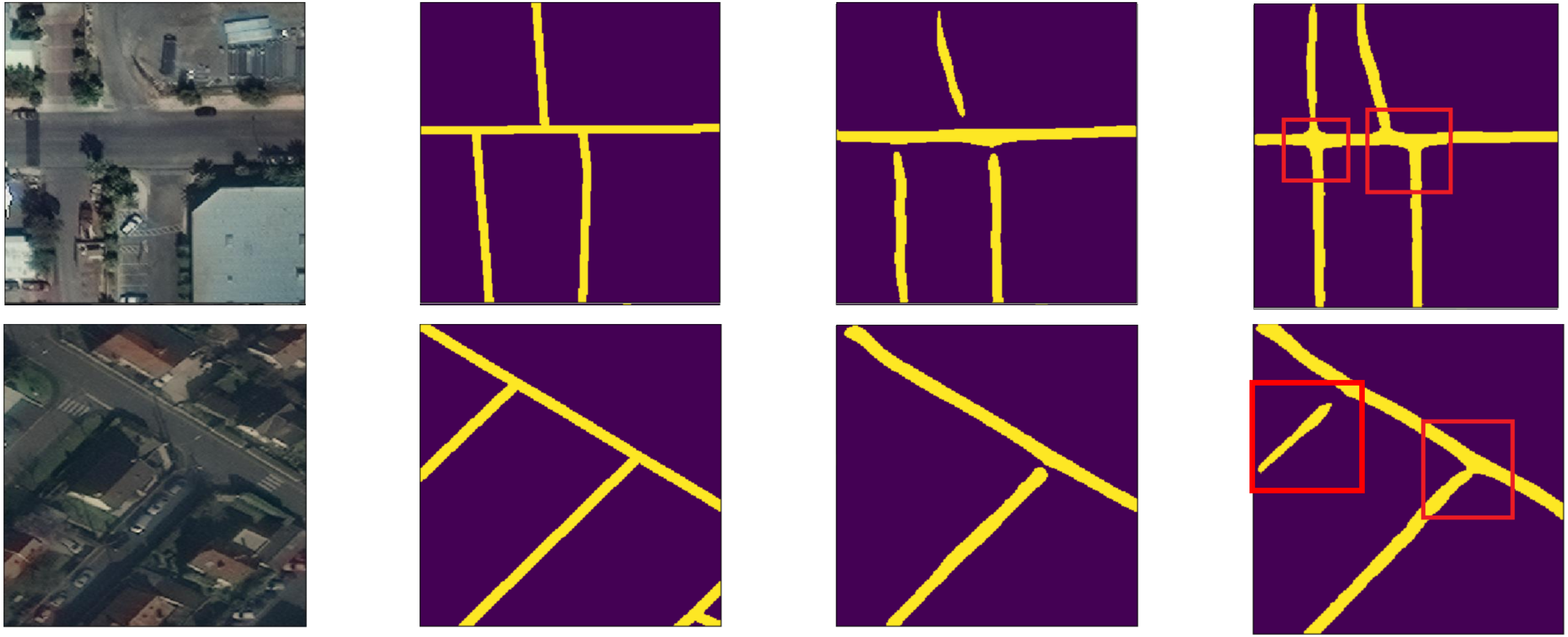
Compare the prediction of the Single-trained Road Detection Learning task
vs
MTL-trained Road Detection Learning task

3.5.1 Results: Visual Inspection



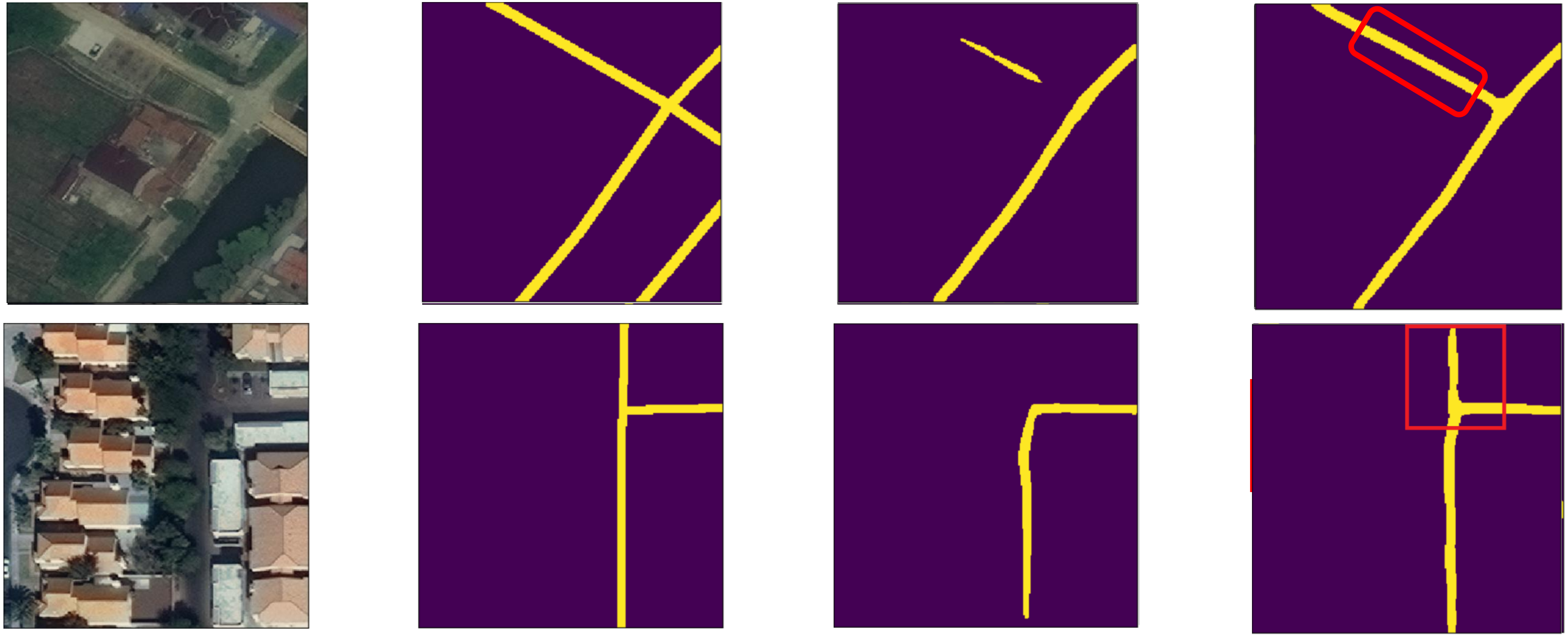
Input image Ground Truth Single-Task Proposed model

3.5.1 Results: Visual Inspection



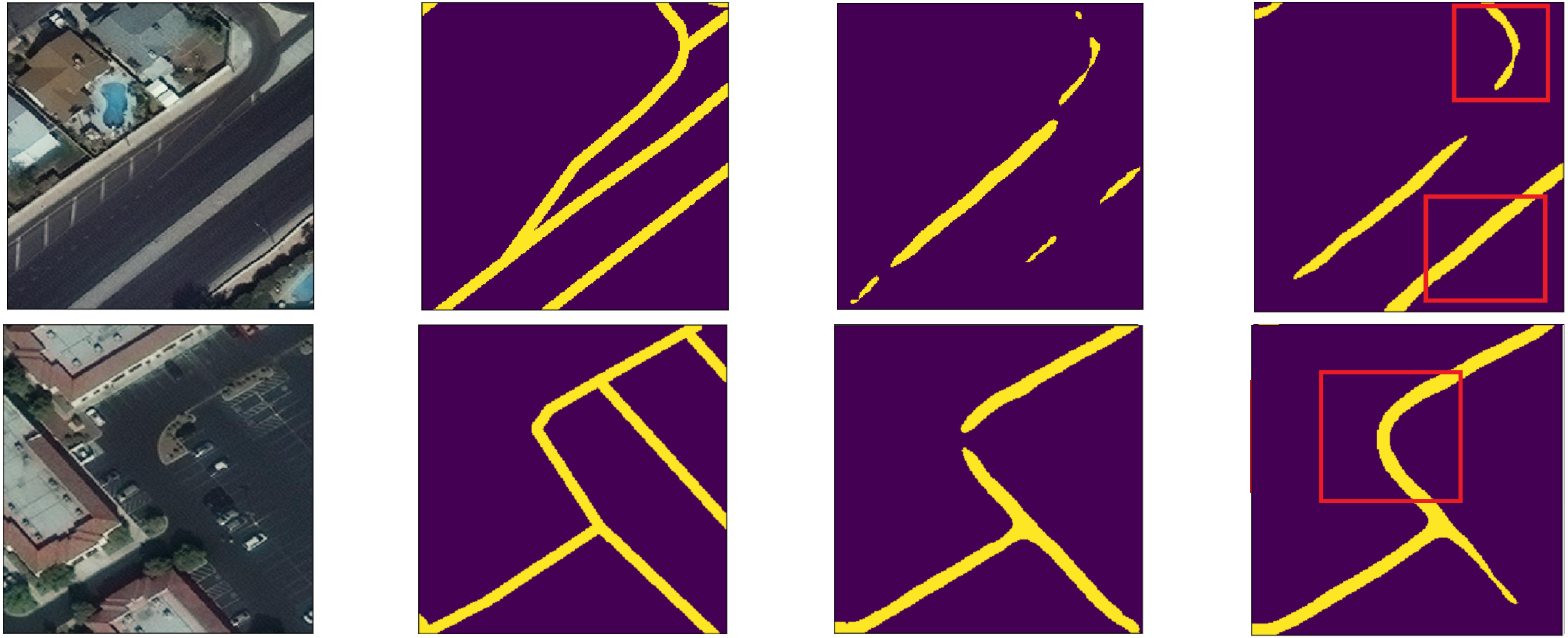
Input image | Ground Truth | Single-Task | Proposed model

3.5.1 Results: Visual Inspection



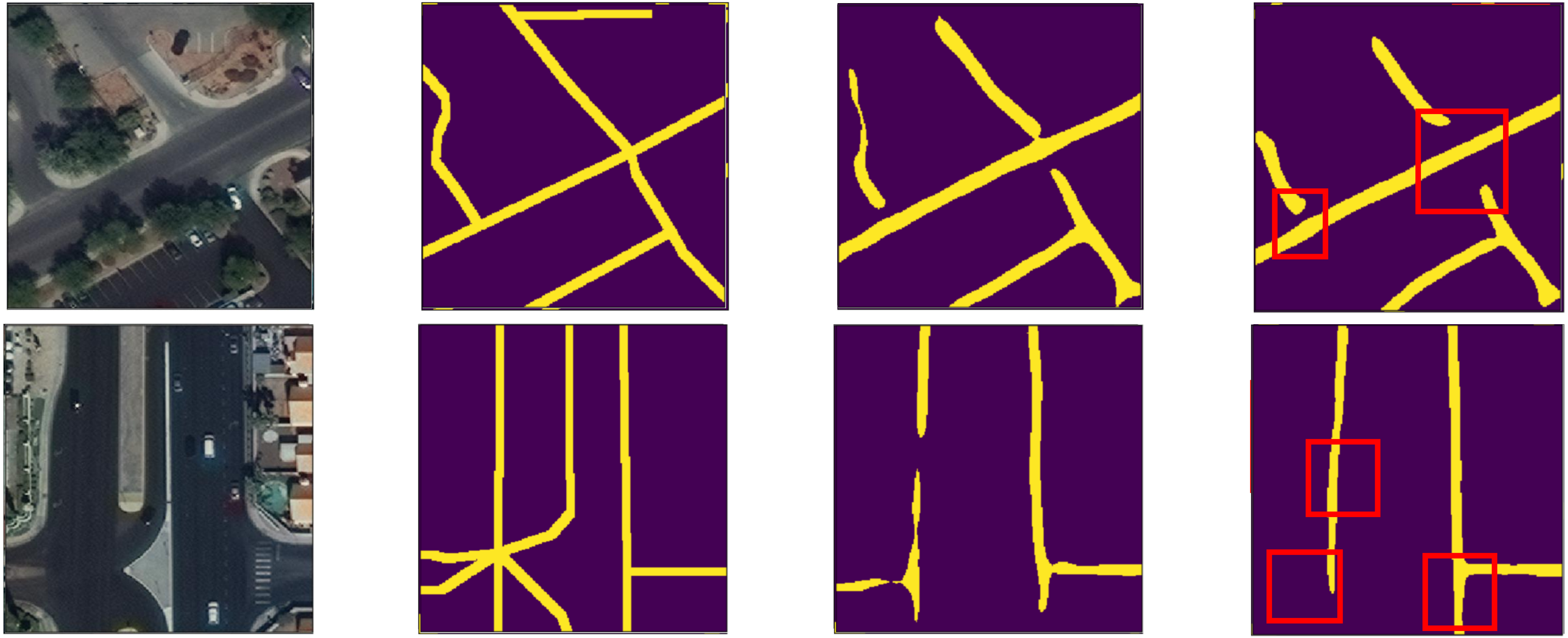
Input image Ground Truth Single-Task Proposed model

3.5.1 Results: Visual Inspection



Input image Ground Truth Single-Task Proposed model

3.5.1 Results: Visual Inspection



Input image | Ground Truth | Single-Task | Proposed model

3.5.2 Results: Quantitative Evaluation

Comparative evaluation of the Road Detection Learning task

TWO	INTERSECTION	ORIENTATION	GAUSSIAN	Mean IoU Score	Mean F1 Score	Mean clDice Score
✓	-	-	-	0.684	0.732	0.671
✓	✓	-	-	0.677	0.725	0.661
✓	-	-	✓	0.679	0.726	0.661
✓	-	✓	-	0.680	0.729	0.676
✓	✓	✓	-	**0.691**	**0.740**	**0.698**
✓	✓	-	✓	0.681	0.727	0.657
✓	-	✓	✓	0.668	0.712	0.621
✓	✓	✓	✓	0.673	0.716	0.621

<u>3.5.2 Results:</u> Quantitative Evaluation

Comparative evaluation of the Road Orientation Learning task,

ORIENTATION	TWO	GAUSSIAN	INTERSECTION	Mean IoU Score	Mean F1 Score
✓	-	-	-	0.891	0.896
✓	✓	-	-	0.680	0.729
✓	✓	-	✓	0.886	0.890
✓	✓	✓	-	0.894	0.898
✓	✓	✓	✓	**0.896**	**0.900**

Gaussian Road Mask Learning task

GAUSSIAN	TWO	ORIENTATION	INTERSECTION	Mean IoU Score	Mean F1 Score
✓	-	-	-	0.199	0.216
✓	✓	-	-	0.211	0.230
✓	✓	✓	-	0.206	0.225
✓	✓	-	✓	**0.213**	**0.232**
✓	✓	✓	✓	0.211	0.229

and

Road Orientation Learning task

INTERSECTION	TWO	ORIENTATION	GAUSSIAN	Mean IoU Score	Mean F1 Score
✓	-	-	-	0.681	0.720
✓	✓	-	-	0.687	0.728
✓	✓	-	✓	0.691	0.729
✓	✓	✓	-	**0.697**	**0.739**
✓	✓	✓	✓	0.601	0.603

4. Conclusions & Discussion

4.1 Conclusions

Pros

- **Better** qualitative & quantitative results
- **Increased** road network connectivity
- Occasionally better segmentation
- **Proof of concept** → Multi-Task Learning can improve accuracy of primary task
- **Additional benefit** → Related tasks can also improve their accuracy when jointly trained

Cons

- MTL takes **longer time** to train
- Needs more powerful **resources**
- More **complex** – difficult to implement, modify or fine-tune
- Still **not the optimal solution** to the problem

4.2 Discussion

Can prior knowledge be incorporated as a constraint into a deep learning model? If yes, how?

Yes

In this project, a model was enforced to solve tasks that preserve desired properties

(e.g. connectivity, geometry, extent)

According to the selected tasks, different aspects of the target object are kept

Discussion

Can prior knowledge improve road detection?

Yes.

Visual and computational evaluation proved that prior knowledge can improve road detection

HOWEVER

it didn’t outperform the single task model by far.

There is still a lot room for improvement

Discussion

What are the limitations of a model that combines concepts of different models into one, unified model?

Increased complexity.

Deep learning models are hard to implement, modify or fine-tune. More complex models will require more time, resources and expertise

Discussion

To which extent is it possible to utilize road properties to improve road detection from remote sensing imagery using deep learning techniques?

Although the proposed methodology managed to increase the performance of the road detection task, problems still occur.

Two reasons:

1. MTL needs to be enhanced
2. Constraints are weaker than loss functions or specialized architectures/modules

Recommendations for future work

- Provide train images with **high road appearance** (e.g. use train images with more road representation than 2%)
- More **data augmentation + test time augmentation** (currently, I am only using image overlap)
- Correct **errors in the dataset** (e.g. manually create better ground truth masks)
- Investigate if **other related tasks** could be helpful (e.g. degree of road’s connectivity)
- Investigate if training over **a specific landscape** is better
- Investigate how **road widths** affect the results
- Investigate how different **input image sizes** affect the model’s performance
- Implement a sophisticated task-weighting scheme (or other MTL enhancing techniques)
- Investigate if other neural network architectures are more suitable
- To achieve optimal network connectivity, transfer everything to the vector domain → implement a procedural model that preserves topology

Thank you for your time