Road Extraction on Remote Sensing Imagery

October 29, 2020

Principal Investigator: Dr. Zhengyuan Zhu

Research assistants: Ricardo Batista, Minsung Jang, and Hao Sun

Student advisor: Xin Zhang

Table of contents

I. Project overview

Goal:

Automated road extraction and change detection

2005 2006

2005 2006 2007 2008 2016 2017

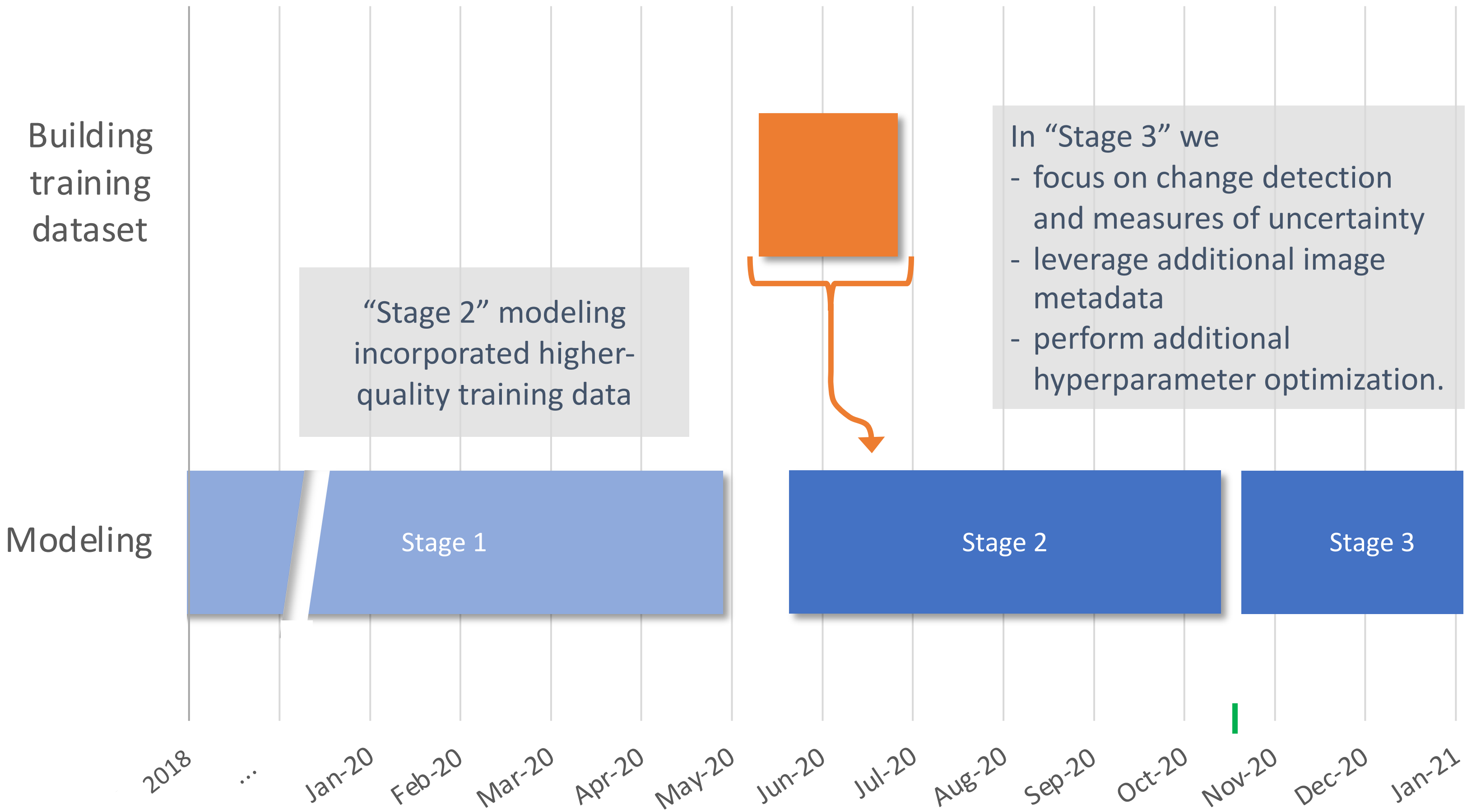
Classification task is non-trivial

Road covered by trees

Not all paved areas are roads

The classifier should take into consideration a sort of context $\Longrightarrow$ *neural nets*

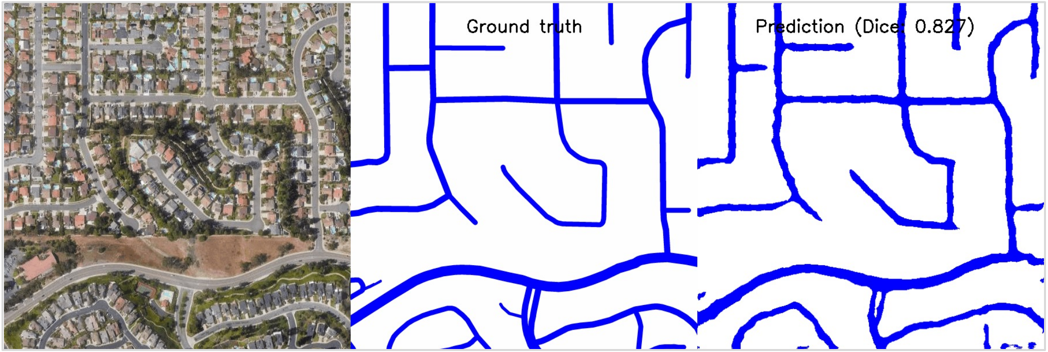
Project timeline



Results summary

We have achieved ≈80% dice score with models trained on 30k images

Model	Dice score
AD-LinkNet	0.78
UNet	0.78
GLNet	0.72



We are now extracting roads on 2005-2017 satellite time series

…

2005

2017

II. Building a training dataset

Why use Google dataset?

The current <u>NRI dataset</u> consisting of 5k remote sensing image and road annotation image pairs …

- is **small** considering task difficulty and modeling limitations,
- contains numerous instances of **poorly annotated** roads from machine learning perspective (see example right), and
- is kept in a specialized server to ensure **confidentiality** but which has limited computing capacity.

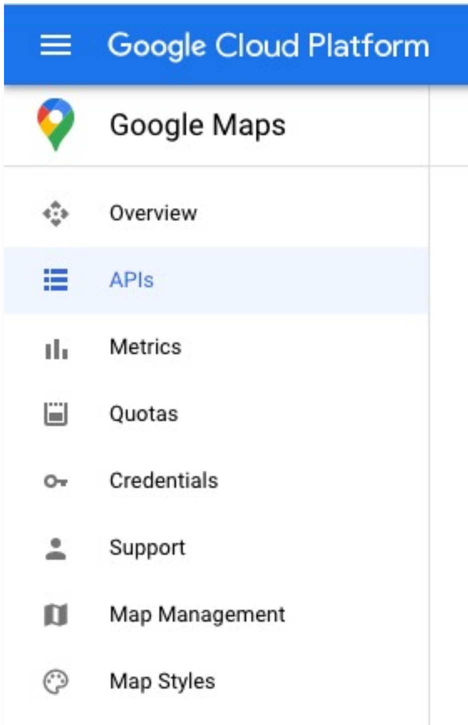
Problematic image type:

coarse annotation

Problematic image type:

omission of major roads (not NRI road?)

Google Maps dataset



- Google provides a comprehensive public **API** for Google Maps
  - The parameters include centroid coordinates, layers (e.g., road) and physical scale (selected as $0.5 \times 0.5\ \mathrm{miles}^2$ to match the NRI dataset)

- Our initial sample consisted of **40k random coordinates** of the contiguous U.S. territory matching the NRI sample

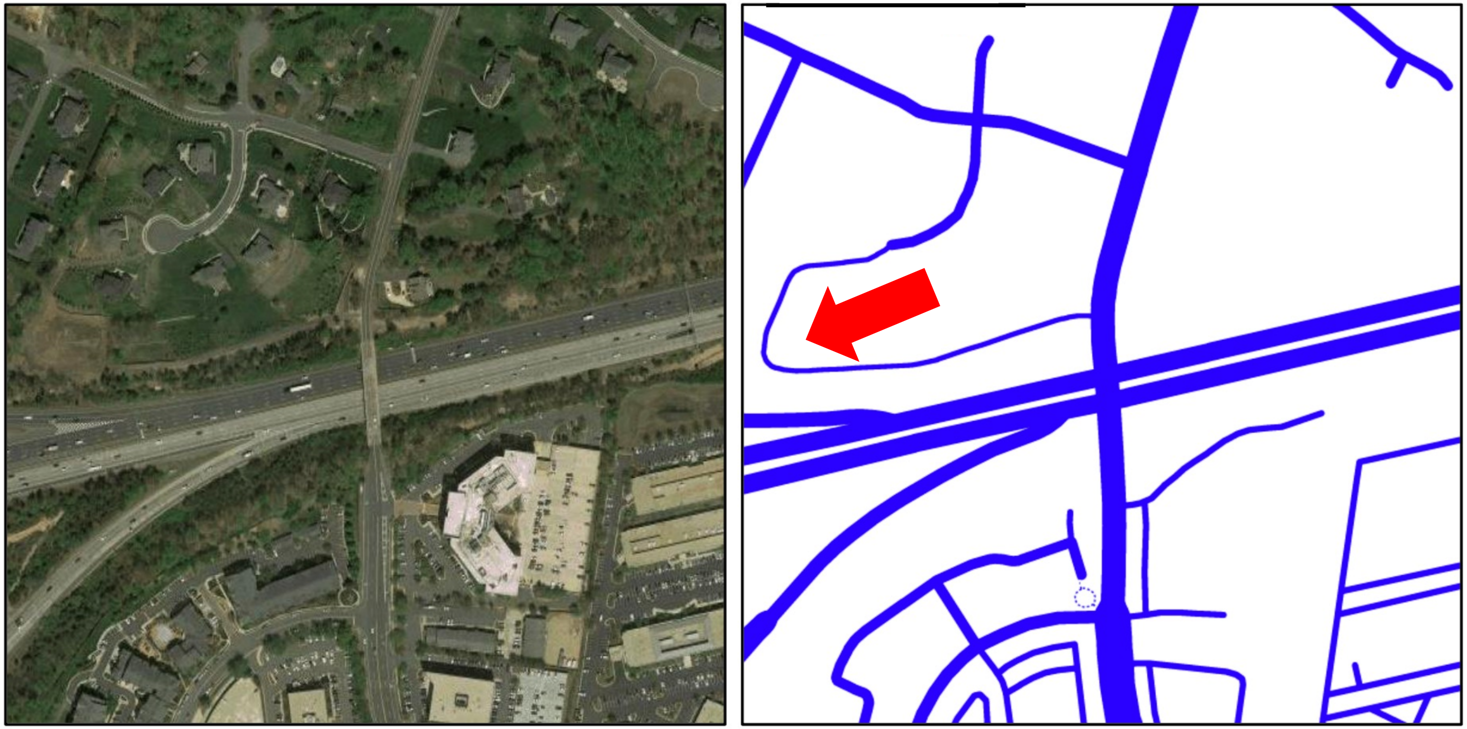
- Though generally accurate, certain image pairs do contain certain **annotation errors**

40k image pairs from Google

Human verification

Training dataset

Example – Image *discarded* since Google created a road that doesn't exist

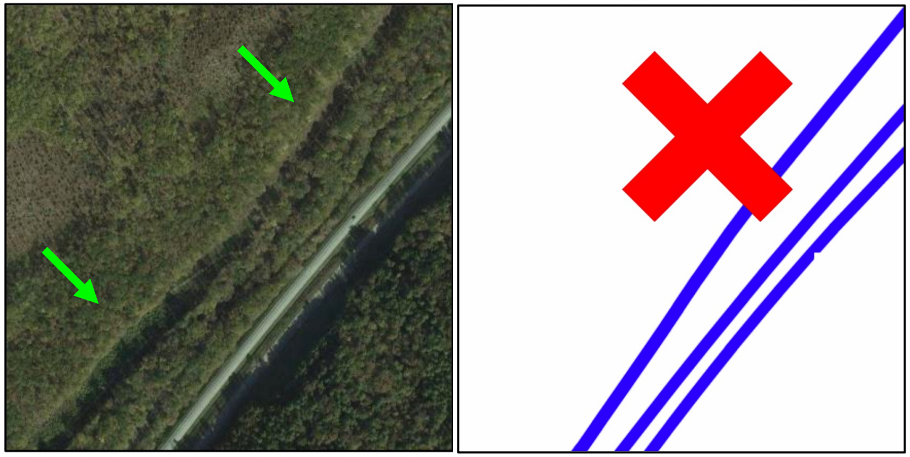


Human verification

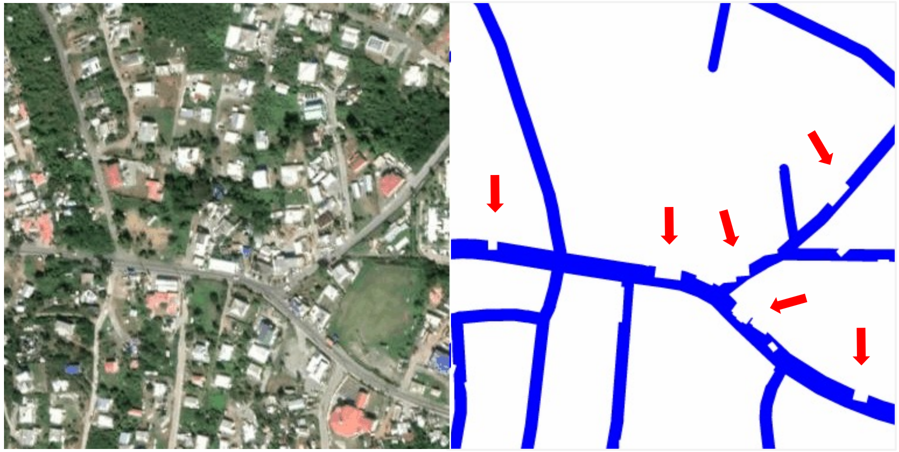
Check 1: Is a main (a.k.a., common use) road missing?

Check 2: Is a road only partially captured?

Check 3: Did Google create a road that DNE?



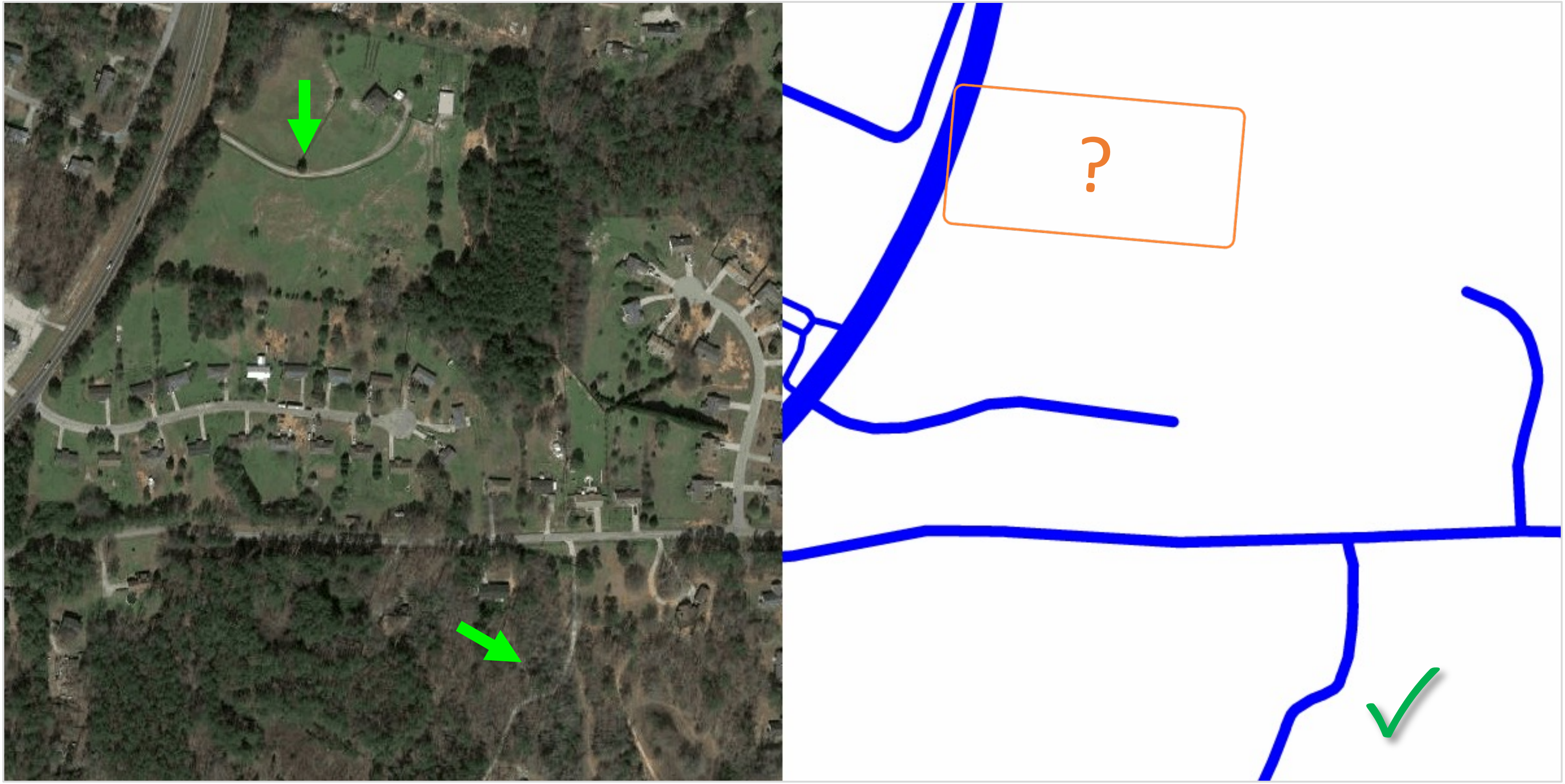
Special cases (i.e., other image errors)



Shortcomings of the training dataset

Treatment of minor roads

Google is inconsistent with its treatment of **minor roads** (e.g., driveways, farm access roads, parking aisles).



Shortcomings of the training dataset

Potential class imbalance

Potential **class imbalance**: there are many satellite images with no roads whatsoever.

III. Training

Modeling overview

Stage 1

Dataset for training

- NRI 5k images (4 states)

Model

- UNet

Stage 2

Dataset for training

- Google 30k images (49 states)

Models

- UNet
- AD-LinkNet
- GLNet

Selected methods

- Model per land use type
- Transfer learning
- Hyperparameter optimization

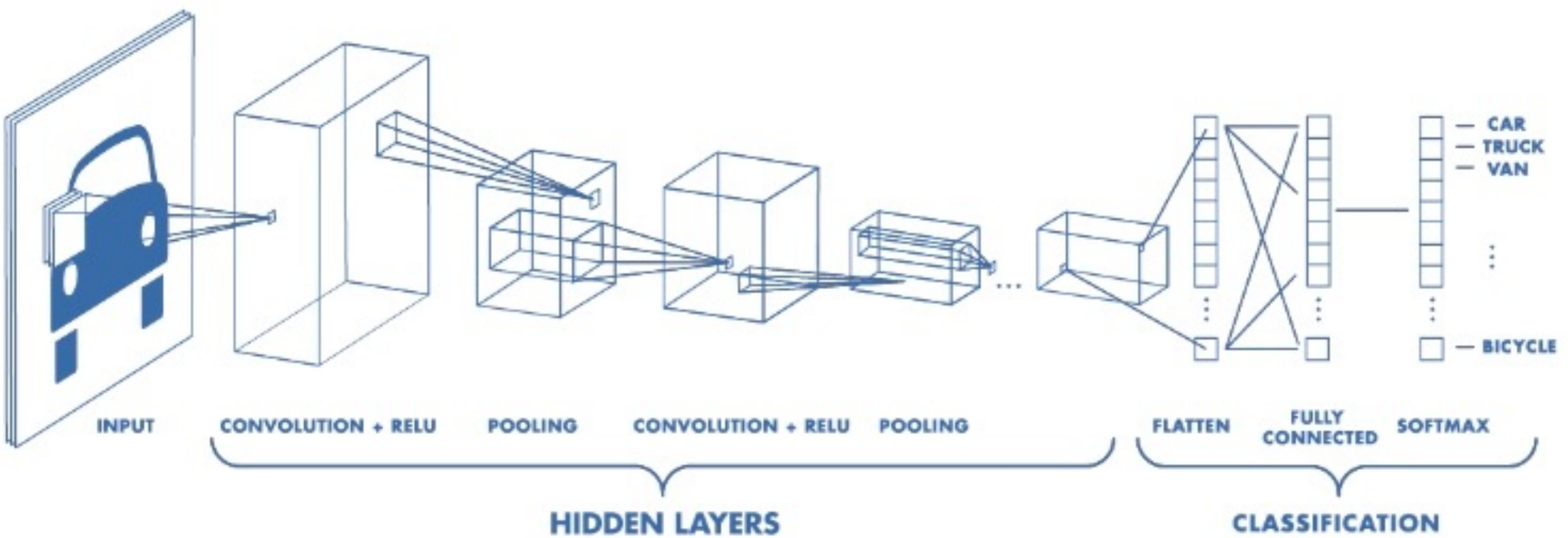
Datasets

- **Original NRI dataset** (4,979 image pairs)
  - IA: 1644, FL: 1165, OR: 672, ND: 1498
- **Google Maps**: (31,981 image pairs)
  - Random coordinates across contiguous U.S. (49 states). As such, includes road-free locations.
  - Filtered from original 38,641 images via visual inspection
- **NRI time series dataset** (size TBD)

2005 2016 2017

Models

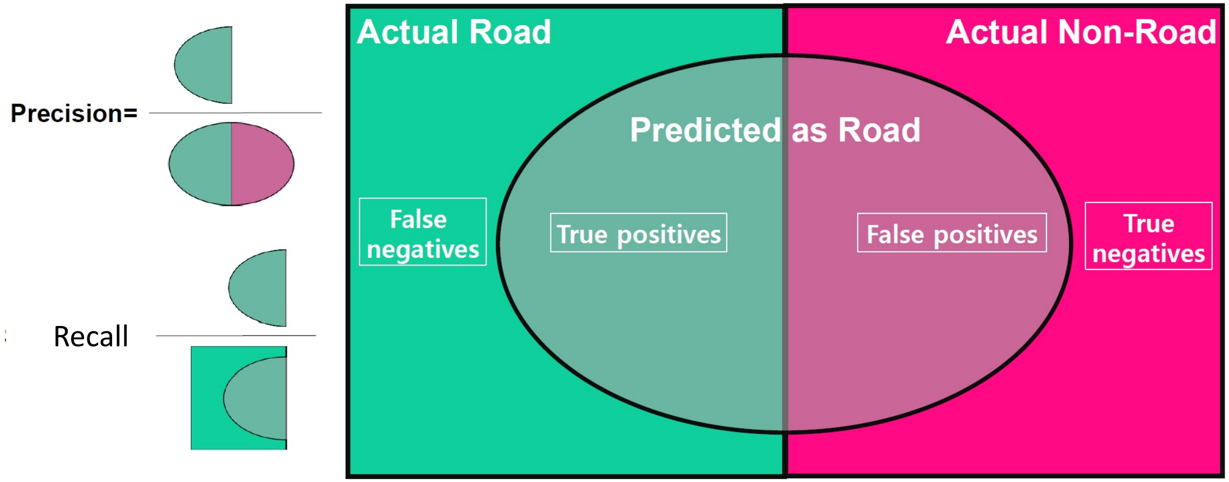
Classical Convolutional Neural Network structure



Source: https://towardsdatascience.com/basics-of-the-classic-cnn-a3dce1225add

Loss function

- We use pixel-level **dice loss** (i.e., 1 − **dice score**) to measure the estimation and prediction performance

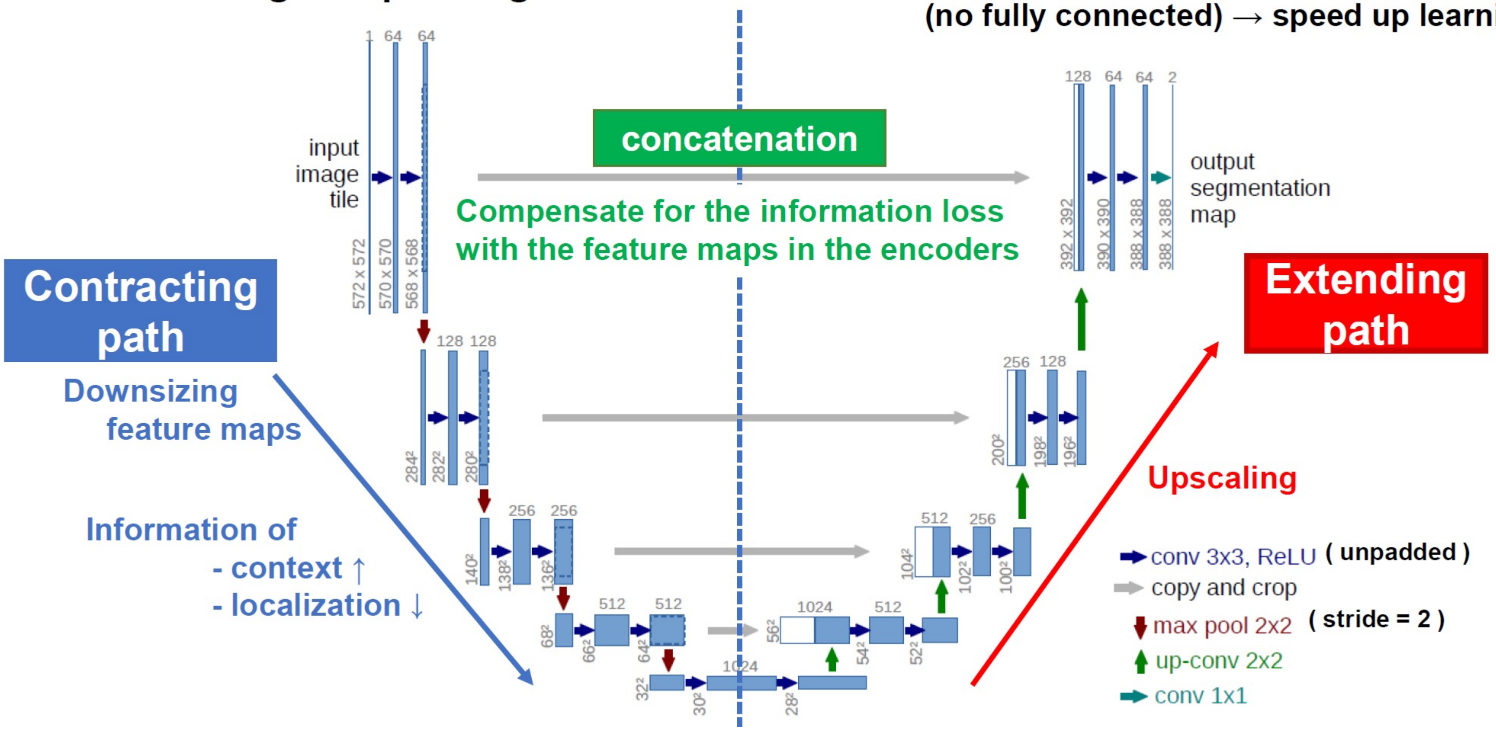


$$Dice\ Score = 2\times\frac{Precision \times Recall}{Precision + Recall}$$ is between 0 and 1

U-Net

- **Contracting – Expanding + Concatenation**

Only convolutional and max pooling layers (no fully connected) → speed up learning



- ➢ Winner of 2015 ISBI challenge for biomedical segmentation
- ➢ The architecture looks like a 'U' shape
- ➢ Left: encoder; Right decoder
- ➢ Pixel-wise Prediction
- ➢ Image in; segmentation out
- ➢ Requires less training images
- ➢ Reduce overfitting by design

Ref. (Ronneberger et al. 2015 https://lmb.informatik.uni-freiburg.de/people/ronneber/u-net/)

AD-LinkNet

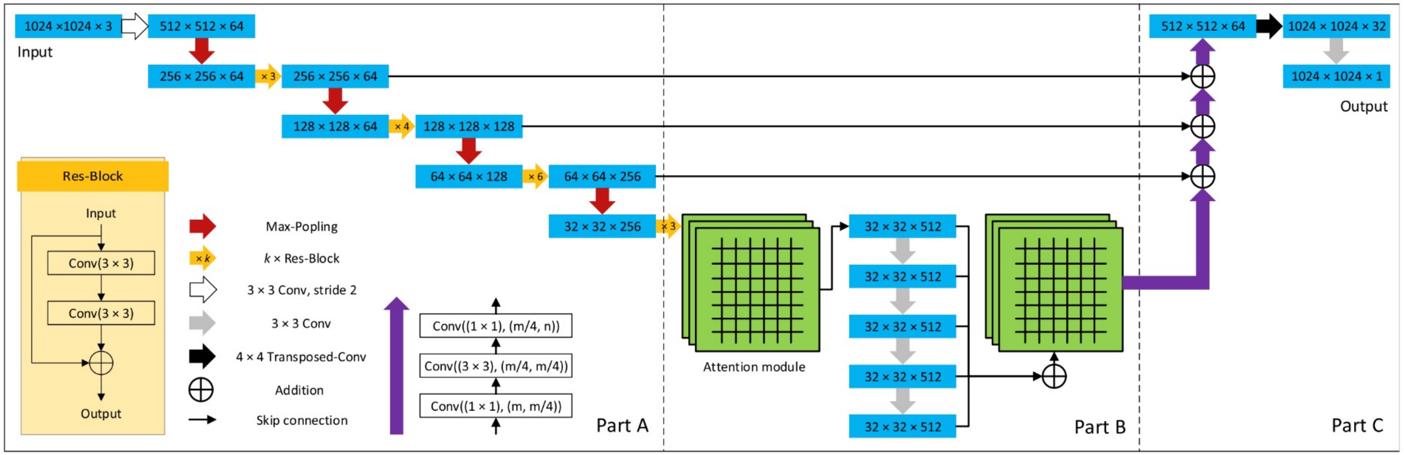


FIGURE 2. The structure diagram of AD-LinkNet.

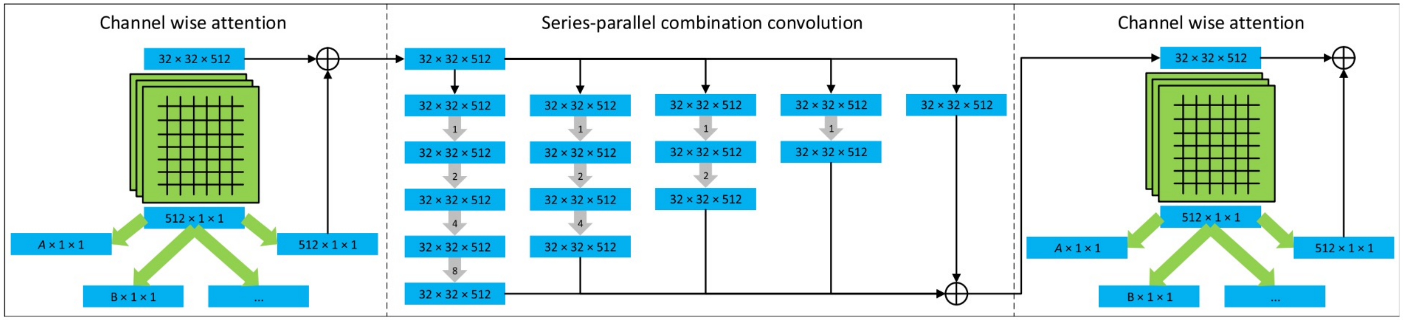
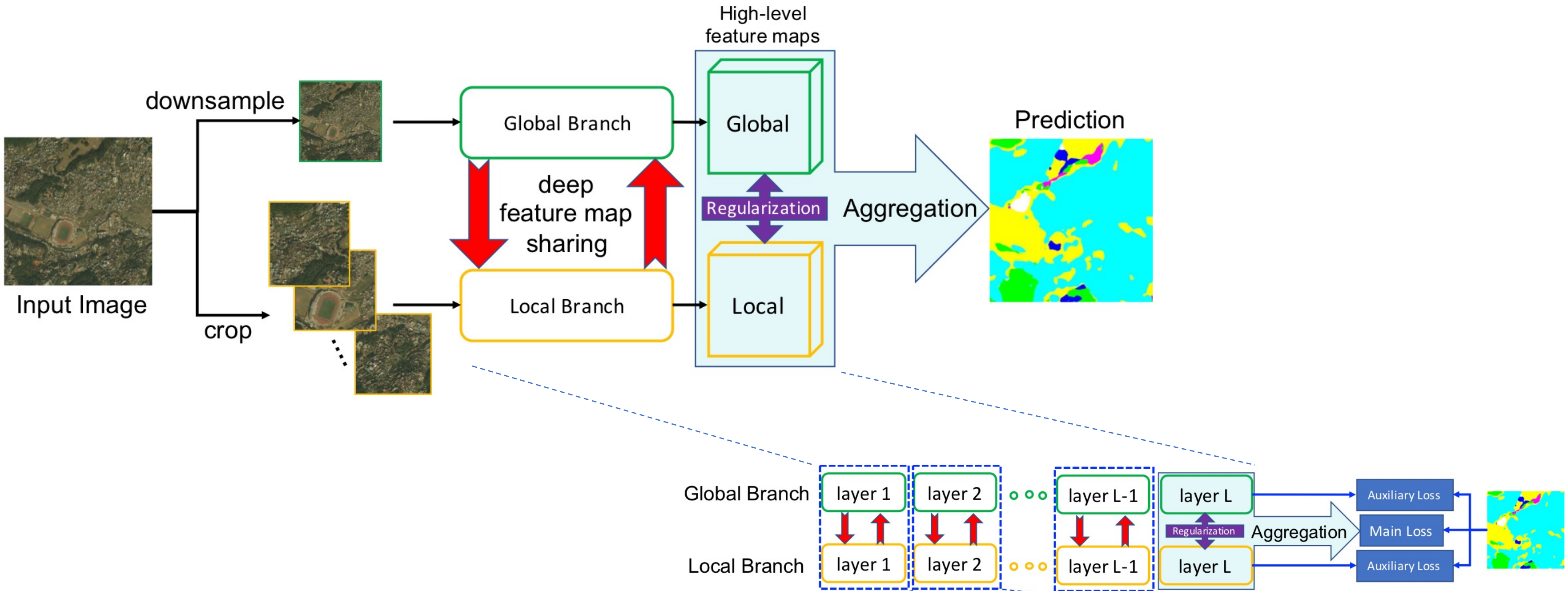


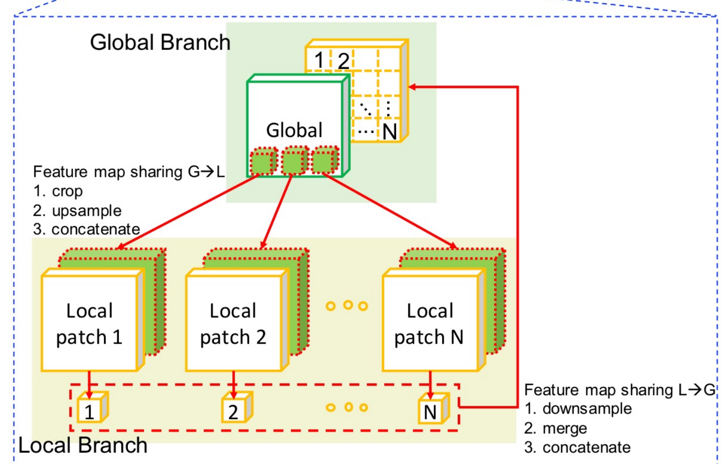
FIGURE 3. Schematic diagram of the AD-LinkNet central part.

- Winner of CVPR's 2018 DeepGlobe road extraction competition
- Short for attention dilation-Linknet
- Serial parallel combination dilated convolution
- Channel-wise attention mechanism

Ref. (Wu, Ming, et al. 2019 IEEE)

GLNet





- Short for Global-Local Network
- Combine global contexts and local high-resolution fine structures
- Memory-efficient for ultra-high resolution images

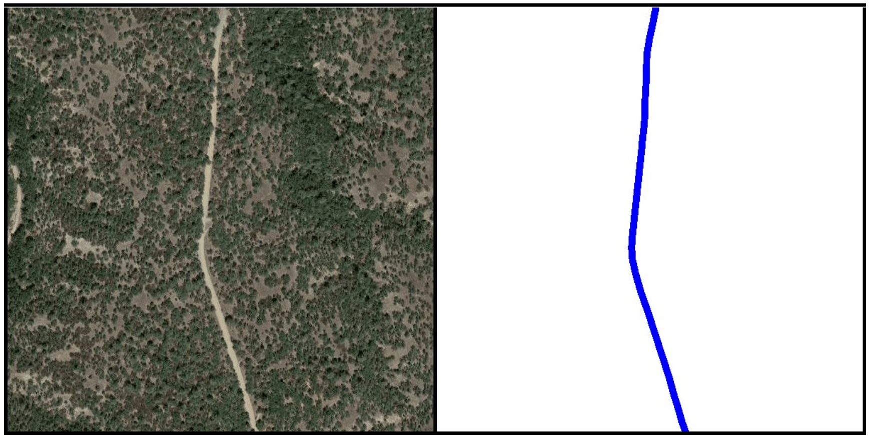
Ref. (Chen, Wuyang, et al, 2020 IEEE)

Methods

Build a model per land use type

- To improve model performance, we classified the images (as either **rural** or **urban**) by land use types and then trained a model on each of the two sets of images.

Rural



Urban

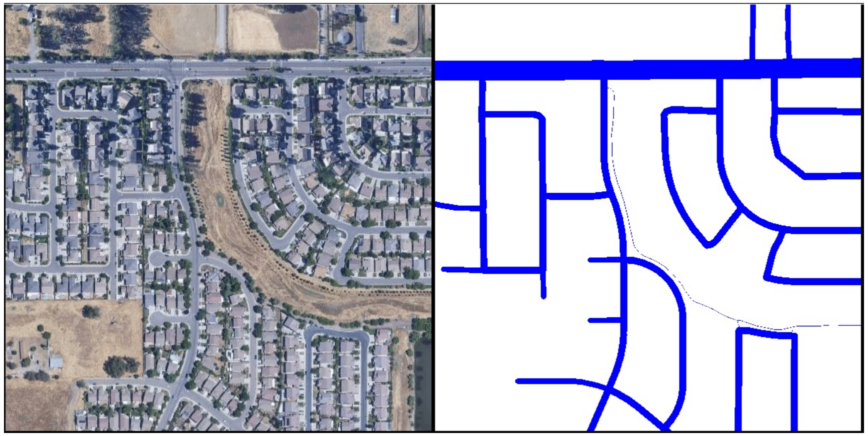
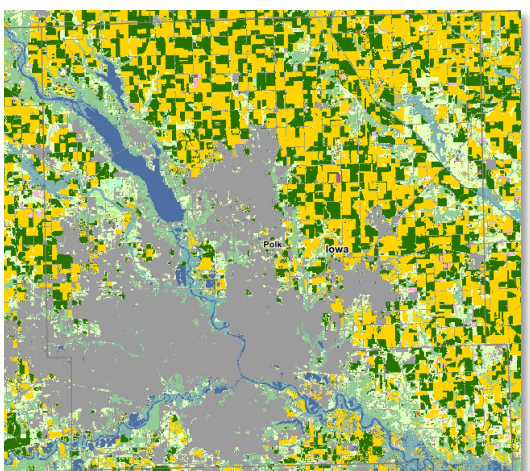


Image classification | Google dataset

- We classified an image as "rural" or "urban" based on the Cropland Data Layer (**CDL**), which assigns land use categories to 30-meter pixels.
- If $\geq 20\%$ of the CDL pixels for the $\approx 0.5 \times 0.5\ \mathrm{mi}^2$ area captured by a Google satellite image is "developed" land, we label it "**urban**." Otherwise, the image is "**rural**."
- Urban ratio (UR) = (number of developed land CDL pixels) / (total CDL pixels)
  - UR > 0.2 → urban type: 7.9% of total dataset (2526 images)
  - UR ≤ 0.2 → rural type : 92.8% of total dataset



UR = 0.0 | UR = 0.2 | UR = 1.0

Google Maps

CDL Image

Image classification | Original NRI dataset

- The metadata for each image includes the **3 most representative land use types**. *(Repeated land use types allowed for a single image.)*
- We classify an image as “**urban**” if either
  1. one of its 3 land use types is “large urban,” or
  2. the list of 3 types consists entirely of “small urban” or “public road.”
- Otherwise, we classify each image as “**rural**”

[Public Road, **Large Urban**, Small Urban]

[Urban]

[**Public Road, Small Urban, Small Urban**]

[Urban]

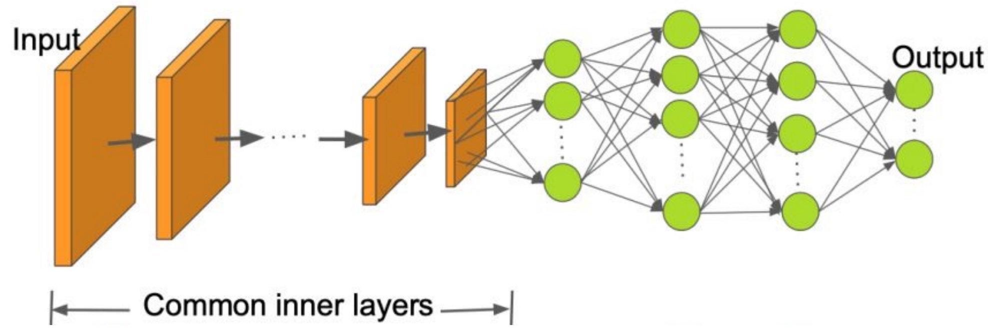
[Cropland, Cropland, Public Road]

[Rural]

Transfer learning

ImageNet
(1M images)

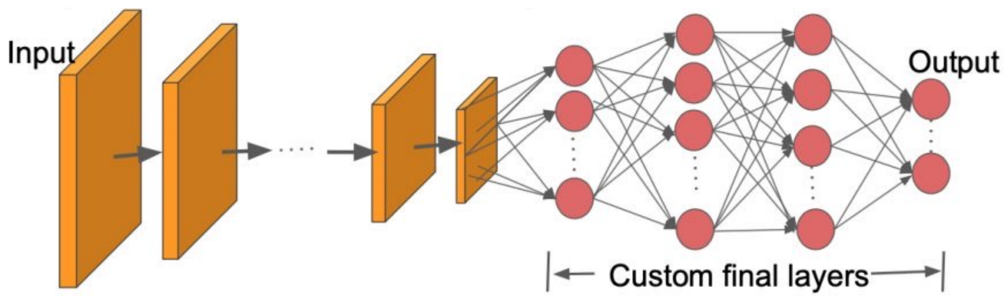
Pretrained
Model



Classification:
Cat, dog, boat,
chair, …

All 30k
Google
Maps images

Custom
Model



Road
annotation

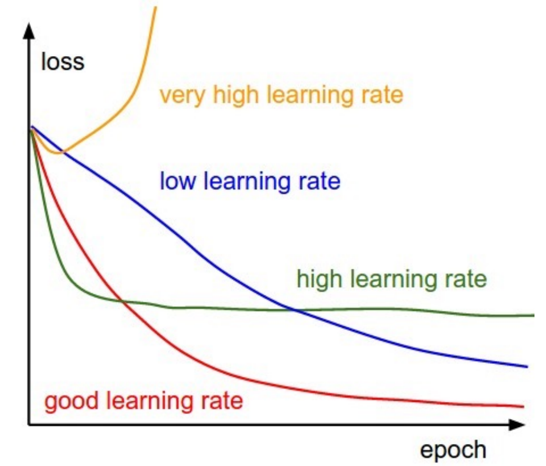
Only the
”urban”
Google images

“Top-level” model: this model has the <u>same architecture</u> as the model immediately above. Additionally, we initialize this model’s weights with the previous model’s <u>fitted weights</u>.

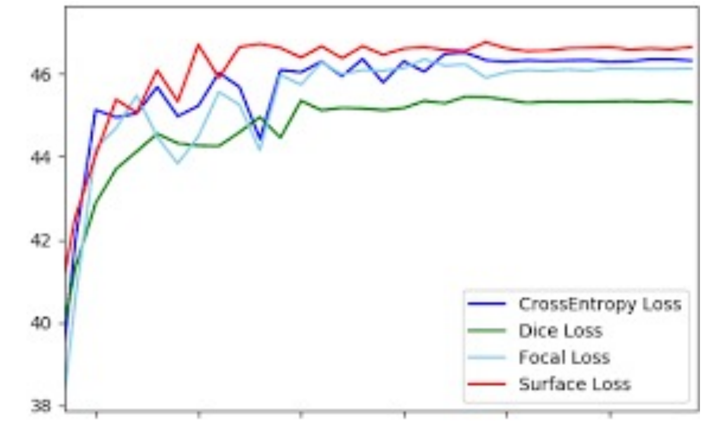
Road
annotation

Hyperparameter optimization

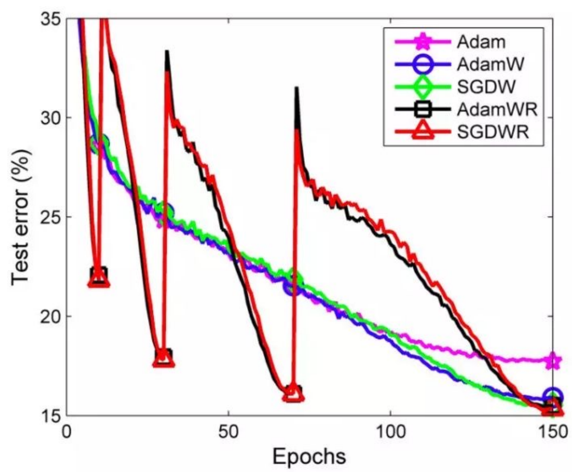
Learning rate



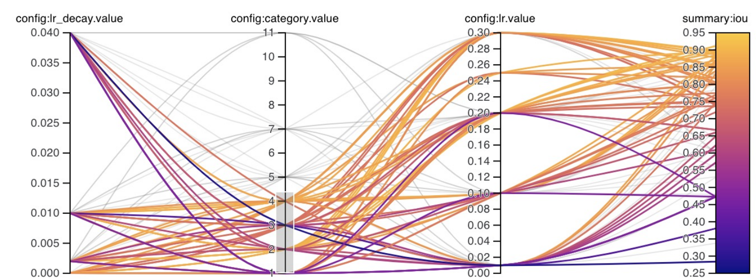
Loss function



Optimizer, weight decay, and restarts



Programmatic hyperparameter optimization

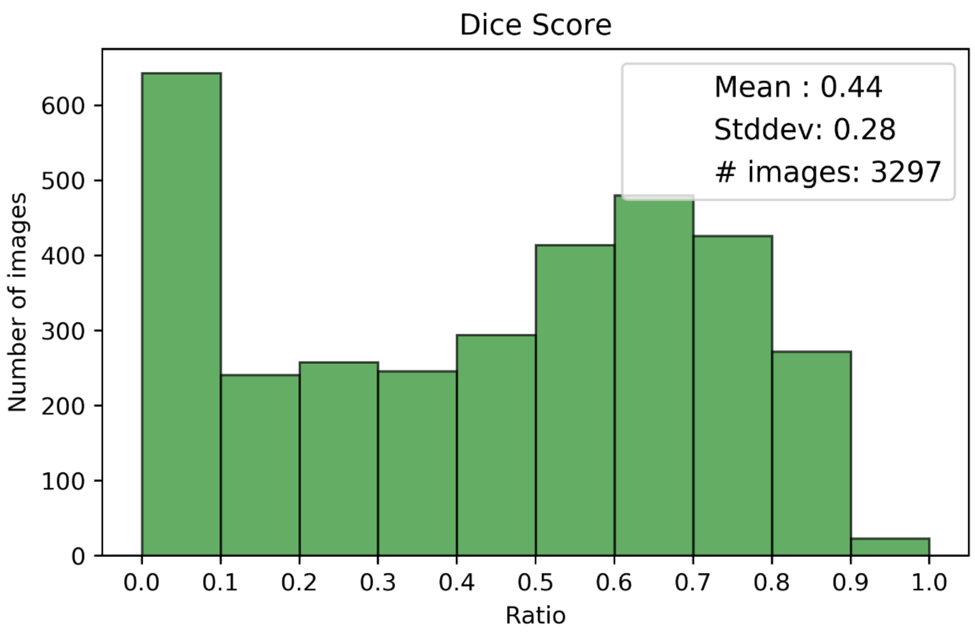


Results

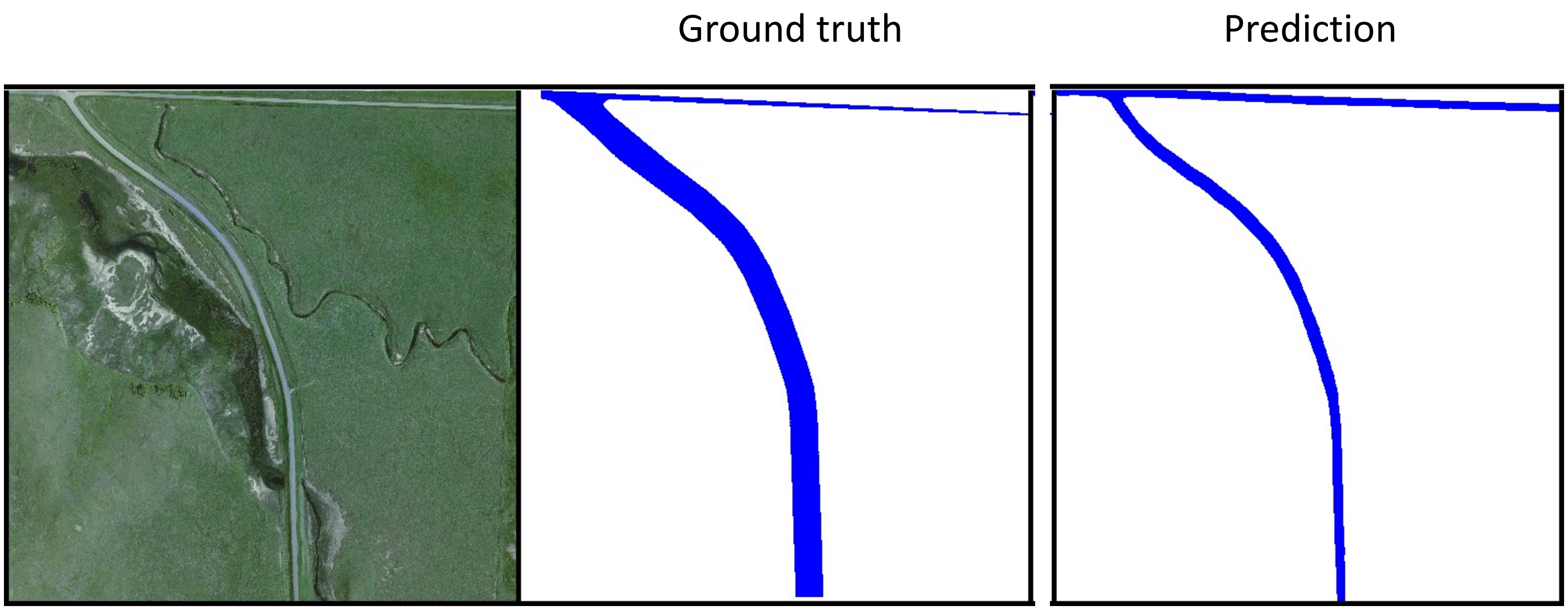
Stage 1

Stage 1 | Original NRI dataset

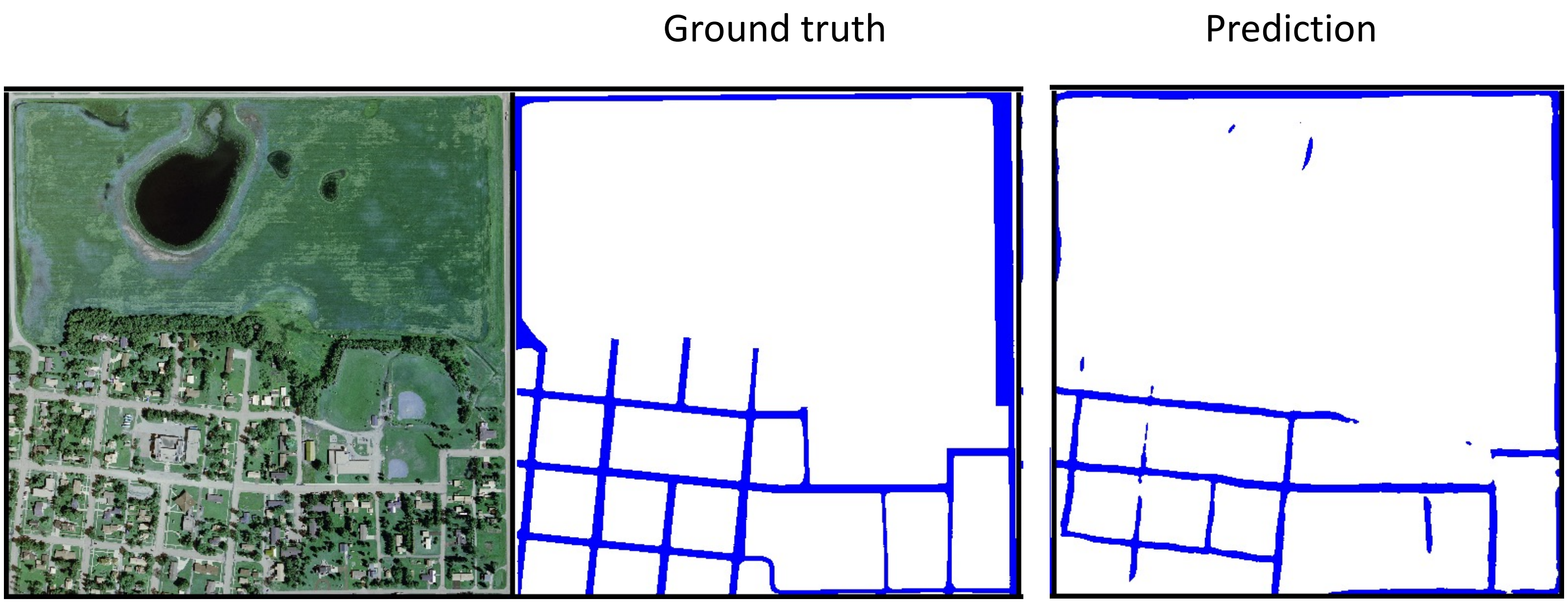
	Rural				Urban				Total			
Model	**Dice**	SE	Prec.	Recall	**Dice**	SE	Prec.	Recall	**Dice**	SE	Prec.	Recall
U-Net	**0.45**	0.28	0.45	0.50	**0.35**	0.21	0.36	0.28	**0.44**	0.28	0.43	0.49



Original NRI dataset | Example 1 (rural)

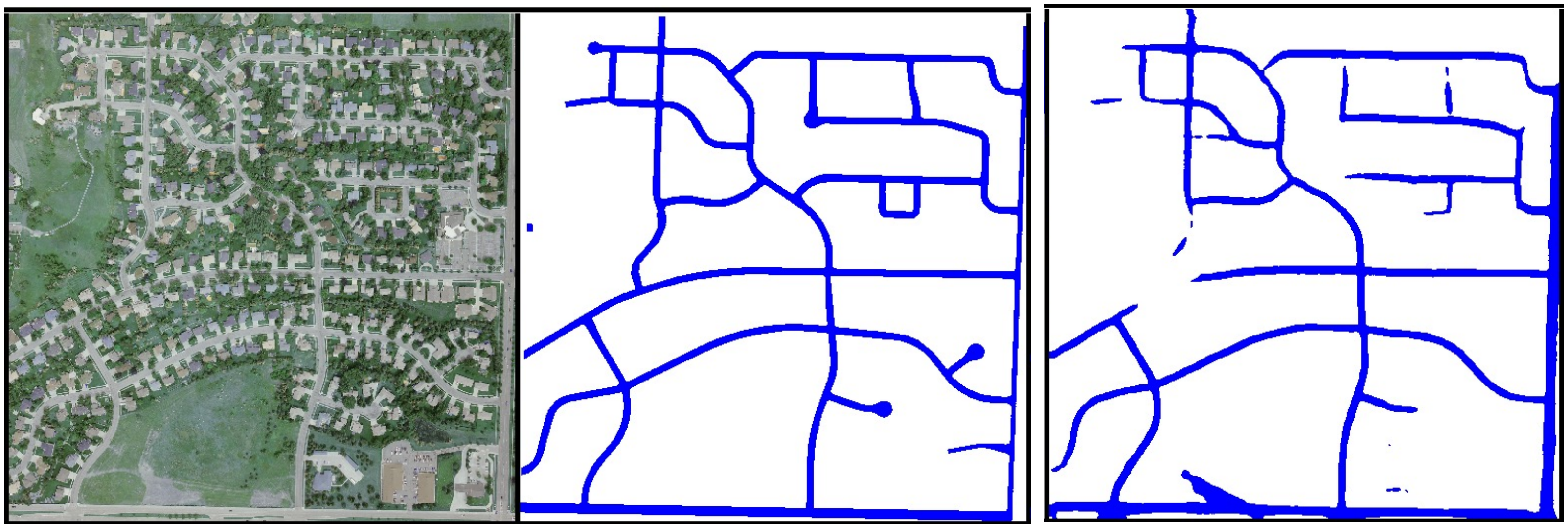


Original NRI dataset | Example 2 (semi-developed)



Original NRI dataset | Example 3 (urban)

Ground truth | Prediction

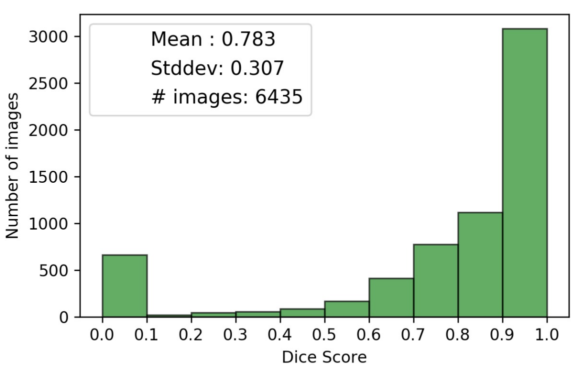


Stage 2

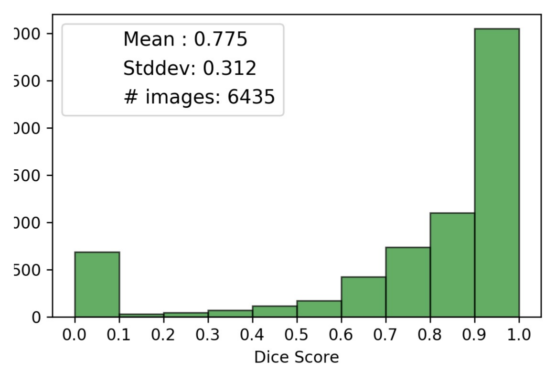
Stage 2 | Google dataset

	Rural				Urban				Total			
Model	**Dice**	SE	Prec.	Recall	**Dice**	SE	Prec.	Recall	**Dice**	SE	Prec.	Recall
AD-LinkNet (TL*)	**0.79**	0.32	0.86	0.83	**0.73**	0.14	0.74	0.75	**0.78**	0.31	0.85	0.82
U-Net (TL)	**0.78**	0.32	0.84	0.84	**0.72**	0.15	0.74	0.74	**0.78**	0.31	0.83	0.83
GLNet (TL)	**0.73**	0.35	0.88	0.75	**0.62**	0.24	0.68	0.69	**0.72**	0.35	0.86	0.75

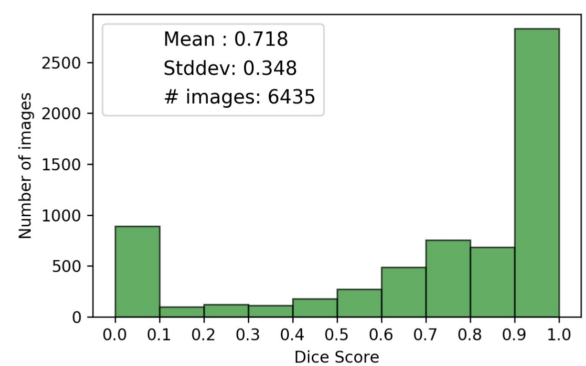
Ad-LinkNet



U-Net



GLNet



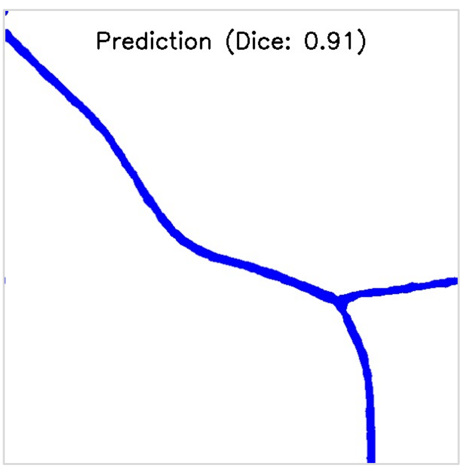
* TL = Transfer learning

Google dataset | Example 1 (rural)

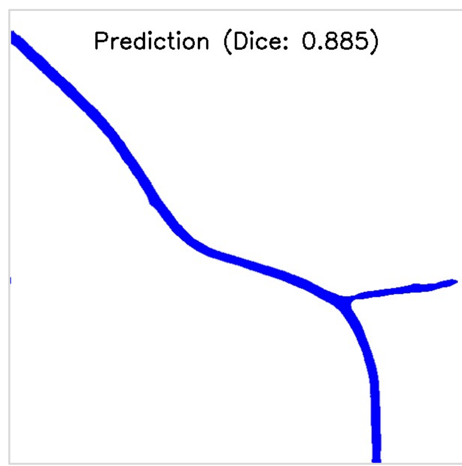
KY_0388

Ad-LinkNet	U-Net	GLNet

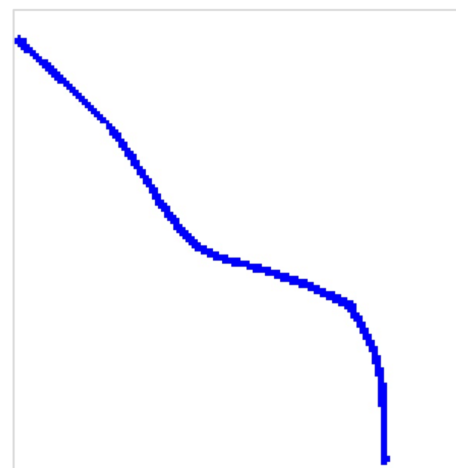
Ad-LinkNet



U-Net

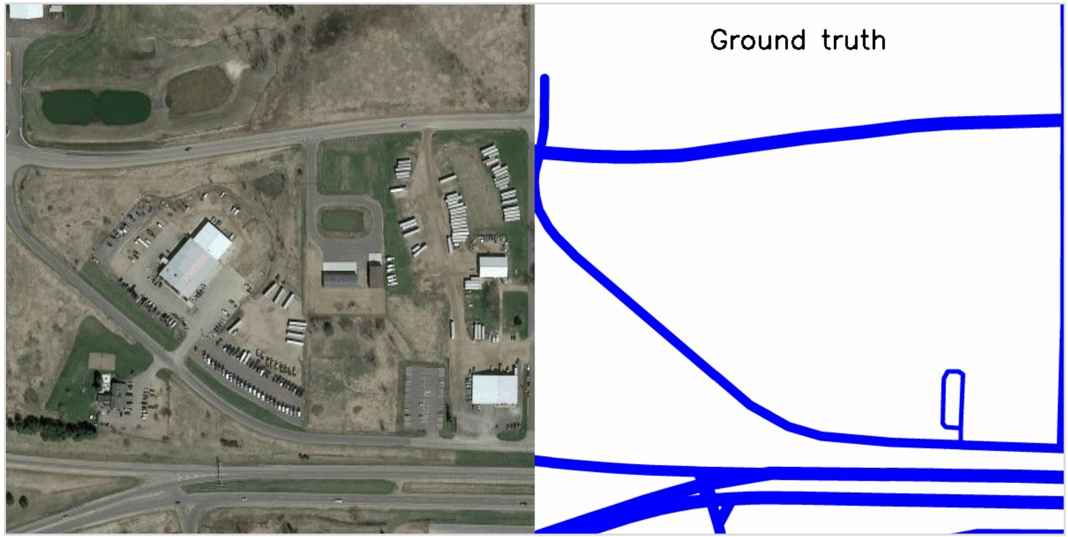


GLNet

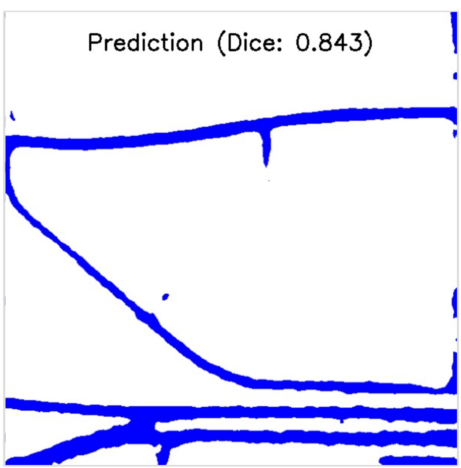


Google dataset | Example 2 (semi-developed)

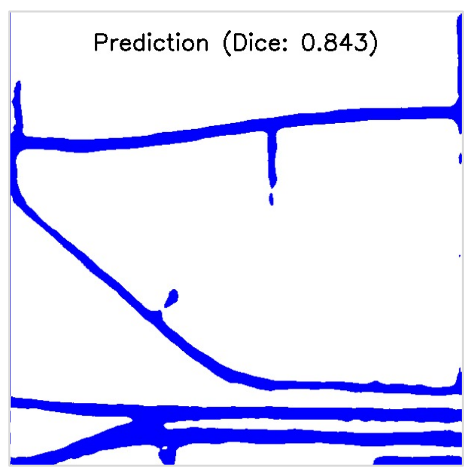
WI_0796



Ad-LinkNet



U-Net

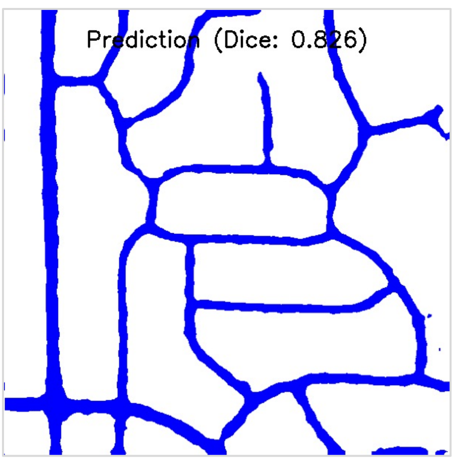


GLNet

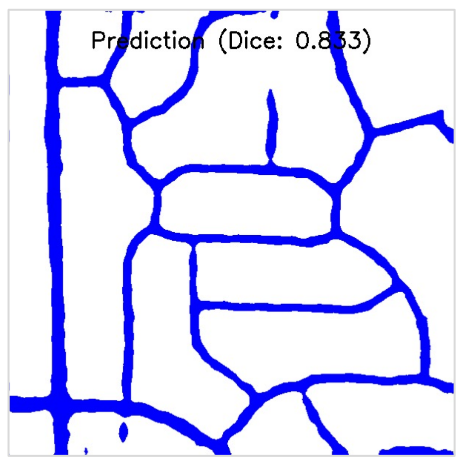
Google dataset | Example 3 (urban)

NJ_0177

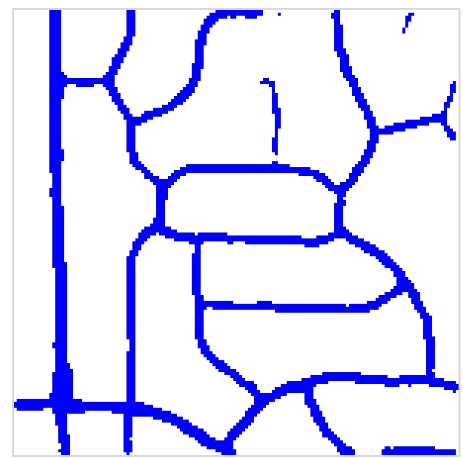
Ad-LinkNet



U-Net



GLNet



Stage 2 | Original NRI dataset

Stage 1: trained on **original NRI dataset**

	Rural				Urban				Total			
Model	**Dice**	SE	Prec.	Recall	**Dice**	SE	Prec.	Recall	**Dice**	SE	Prec.	Recall
U-Net	**0.45**	0.28	0.45	0.50	**0.35**	0.21	0.36	0.28	**0.44**	0.28	0.43	0.49

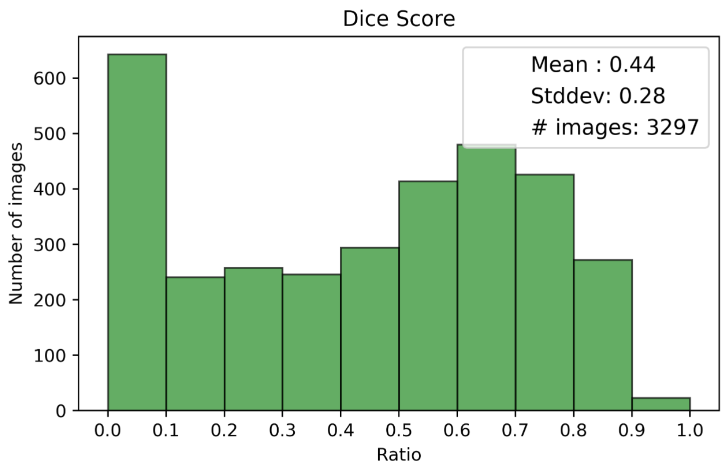
Stage 2: trained on **Google dataset**

	Rural				Urban				Total			
Model	**Dice**	SE	Prec.	Recall	**Dice**	SE	Prec.	Recall	**Dice**	SE	Prec.	Recall
AD-LinkNet (TL*)	**0.56**	0.23	0.63	0.61	**0.51**	0.15	0.49	0.59	**0.56**	0.22	0.62	0.61
U-Net (TL)	**0.54**	0.22	0.57	0.63	**0.48**	0.16	0.47	0.58	**0.54**	0.22	0.56	0.62

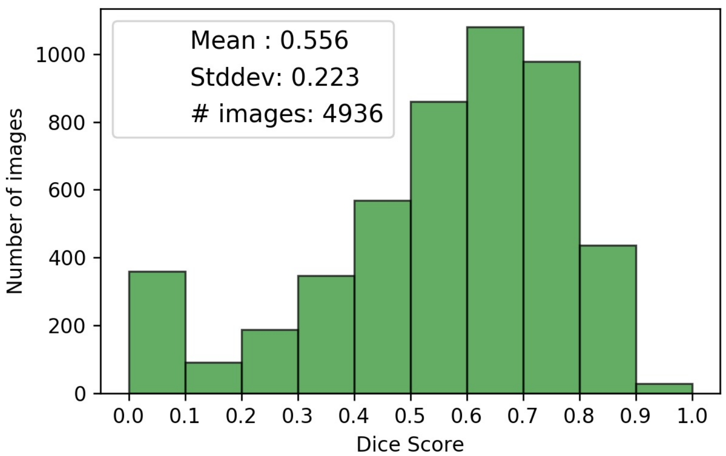
* TL = Transfer learning

Dice score histograms

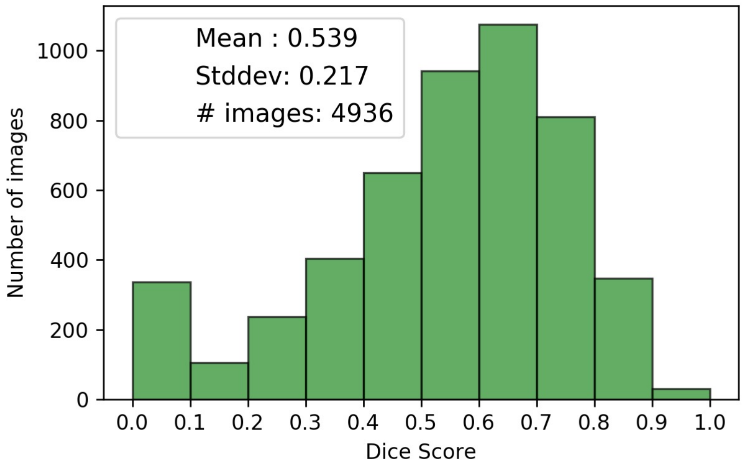
U-Net (Stage 1)



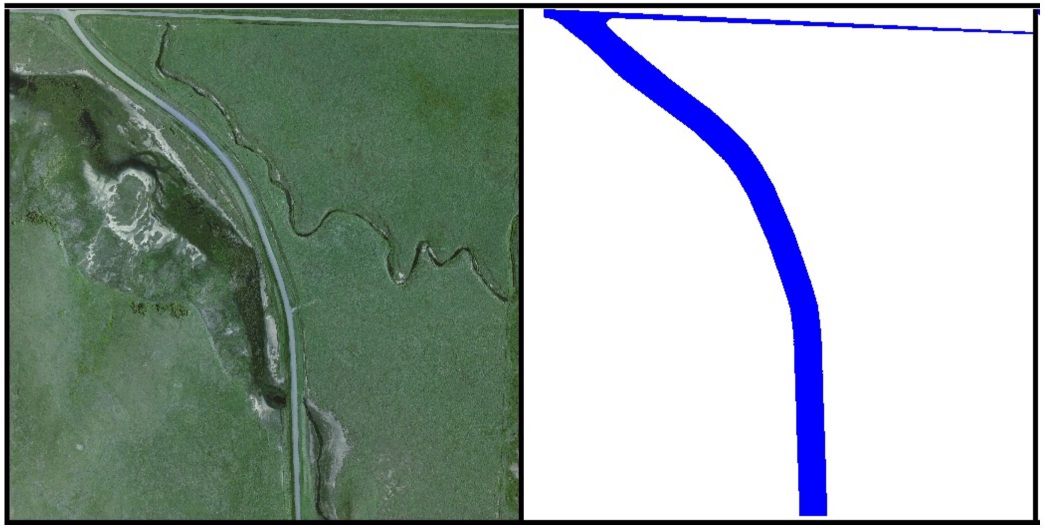
Ad-LinkNet



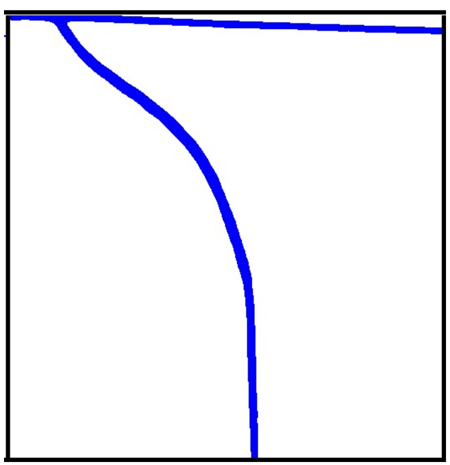
U-Net (Stage 2)



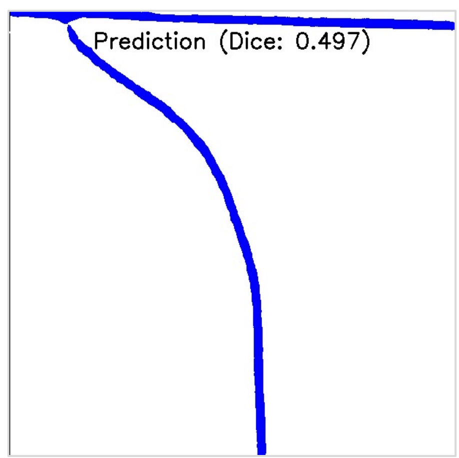
NRI dataset | Example 1 (rural)



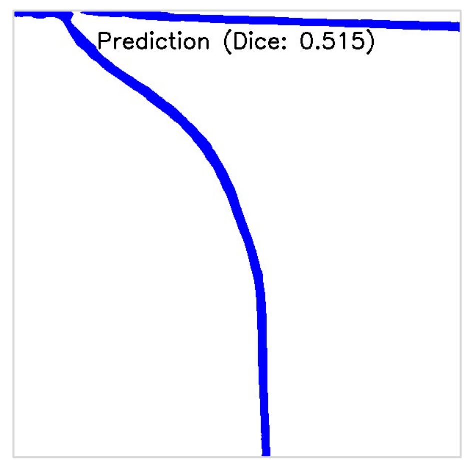
U-Net (Stage 1)



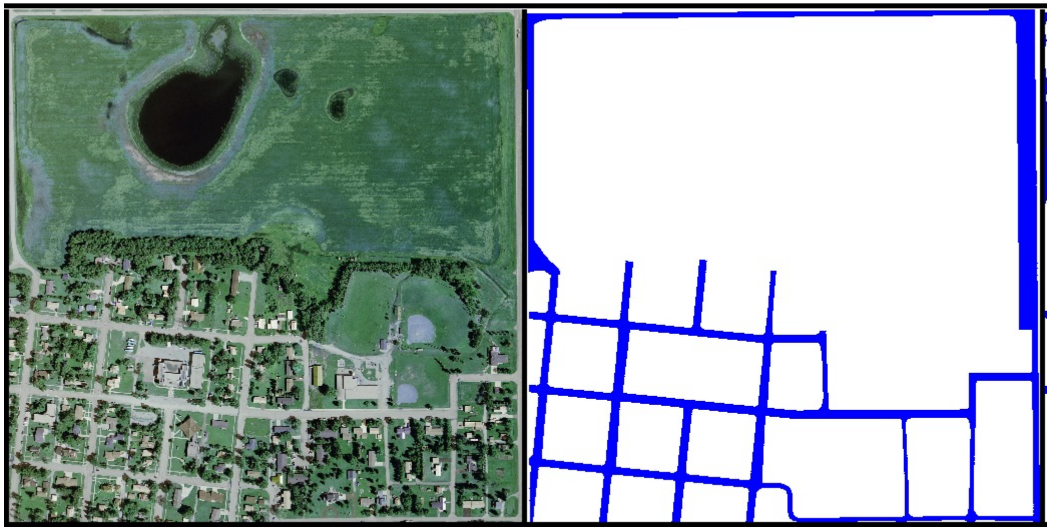
Ad-LinkNet



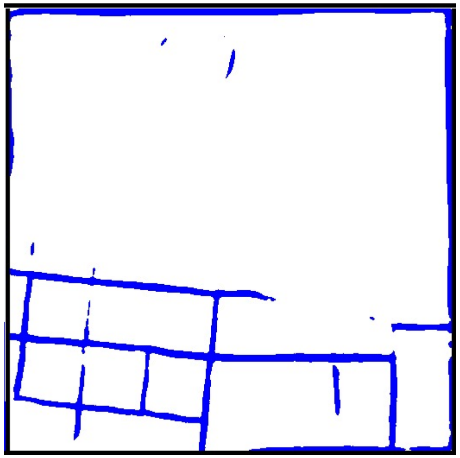
U-Net (Stage 2)



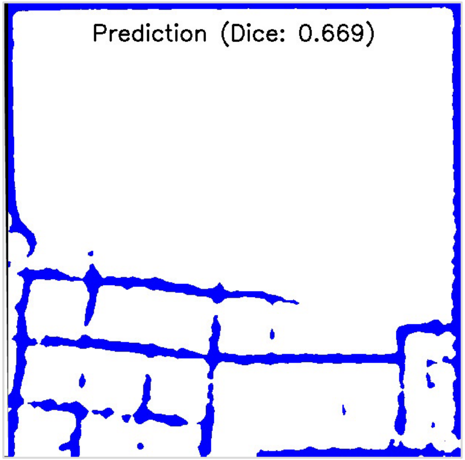
NRI dataset | Example 2 (semi-developed)



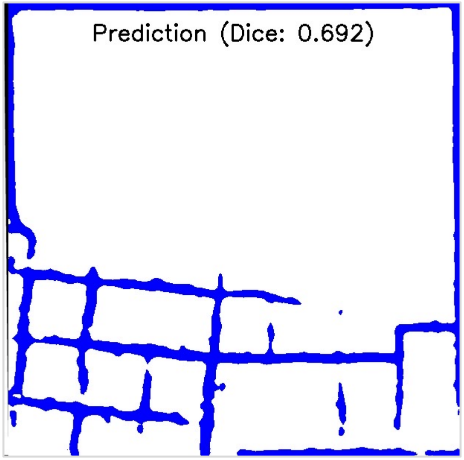
U-Net (Stage 1)



Ad-LinkNet

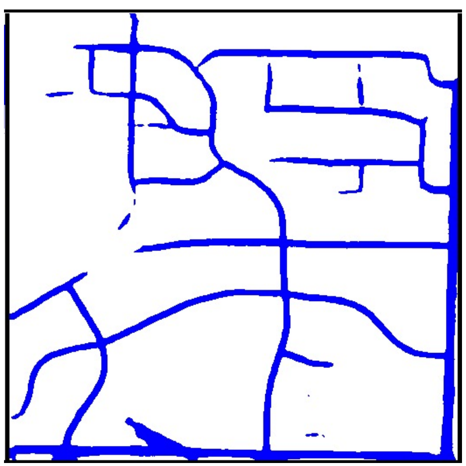


U-Net (Stage 2)

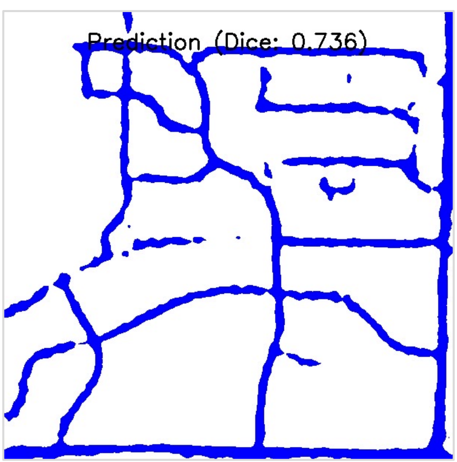


NRI dataset | Example 3 (urban)

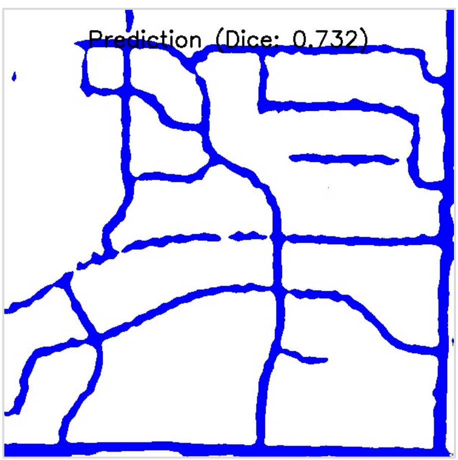
U-Net (Stage 1)



Ad-LinkNet

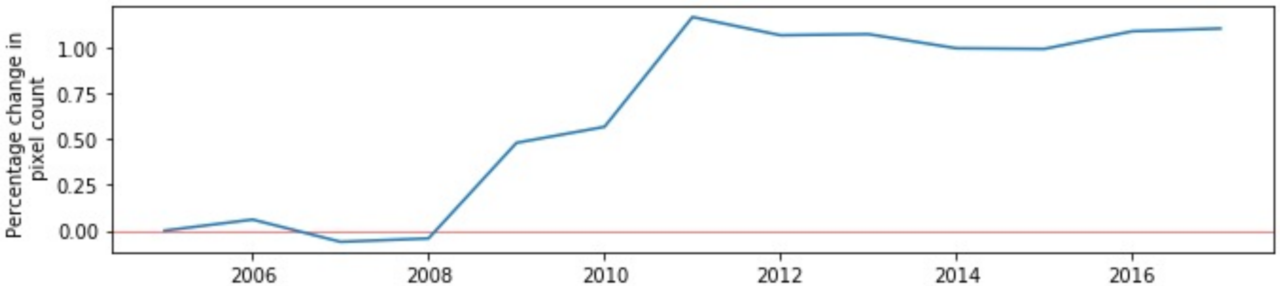


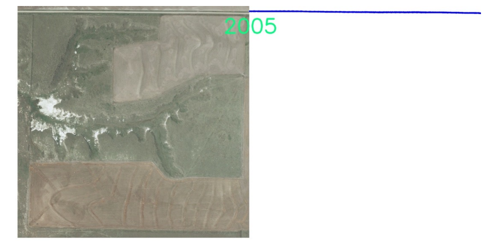
U-Net (Stage 2)

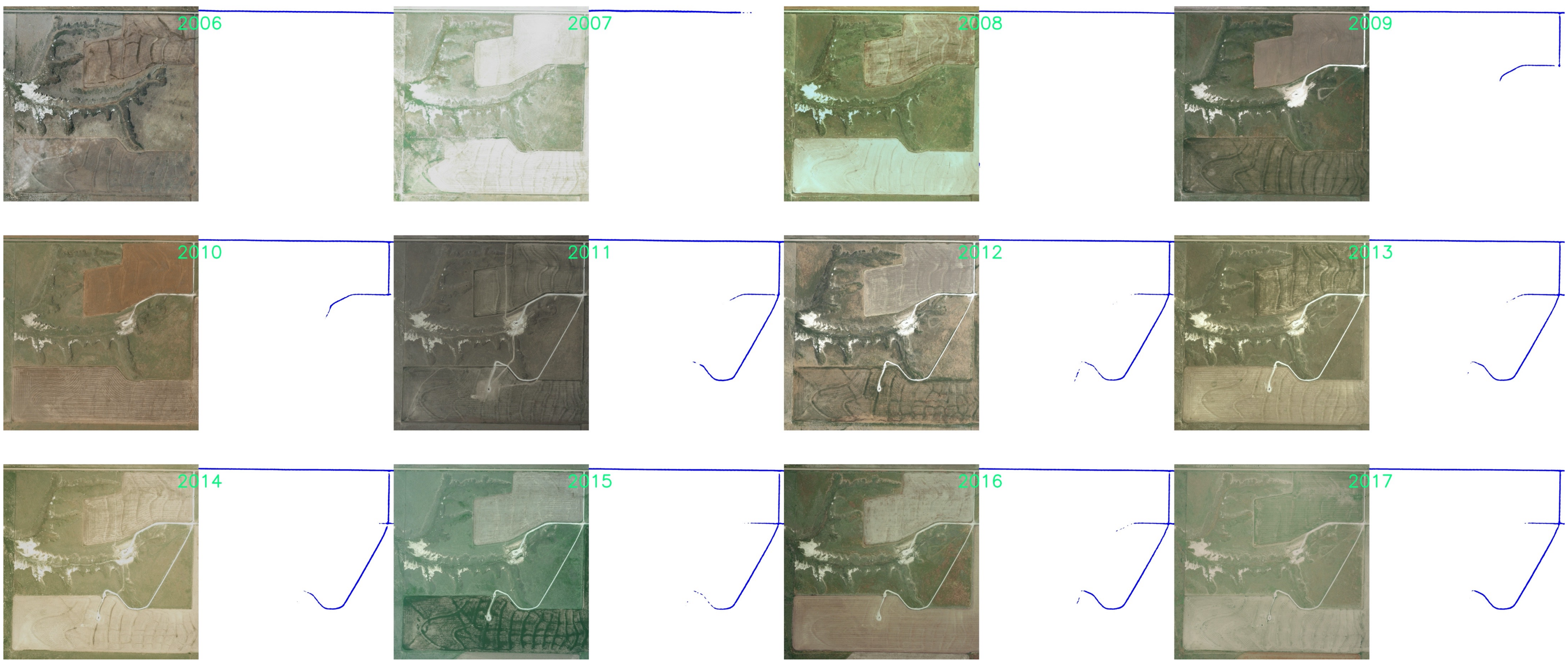


Stage 3

Example 1 Model: AD-LinkNet (transfer learning, rural)

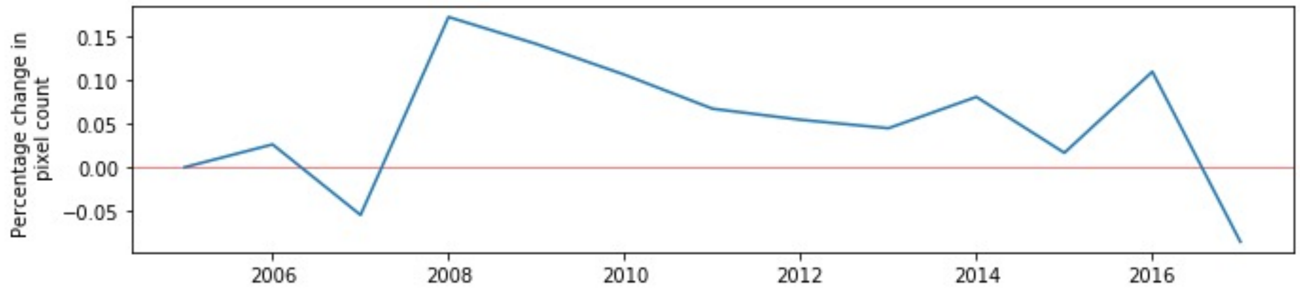


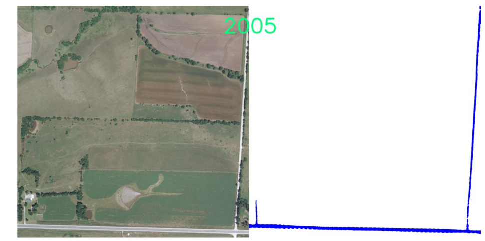


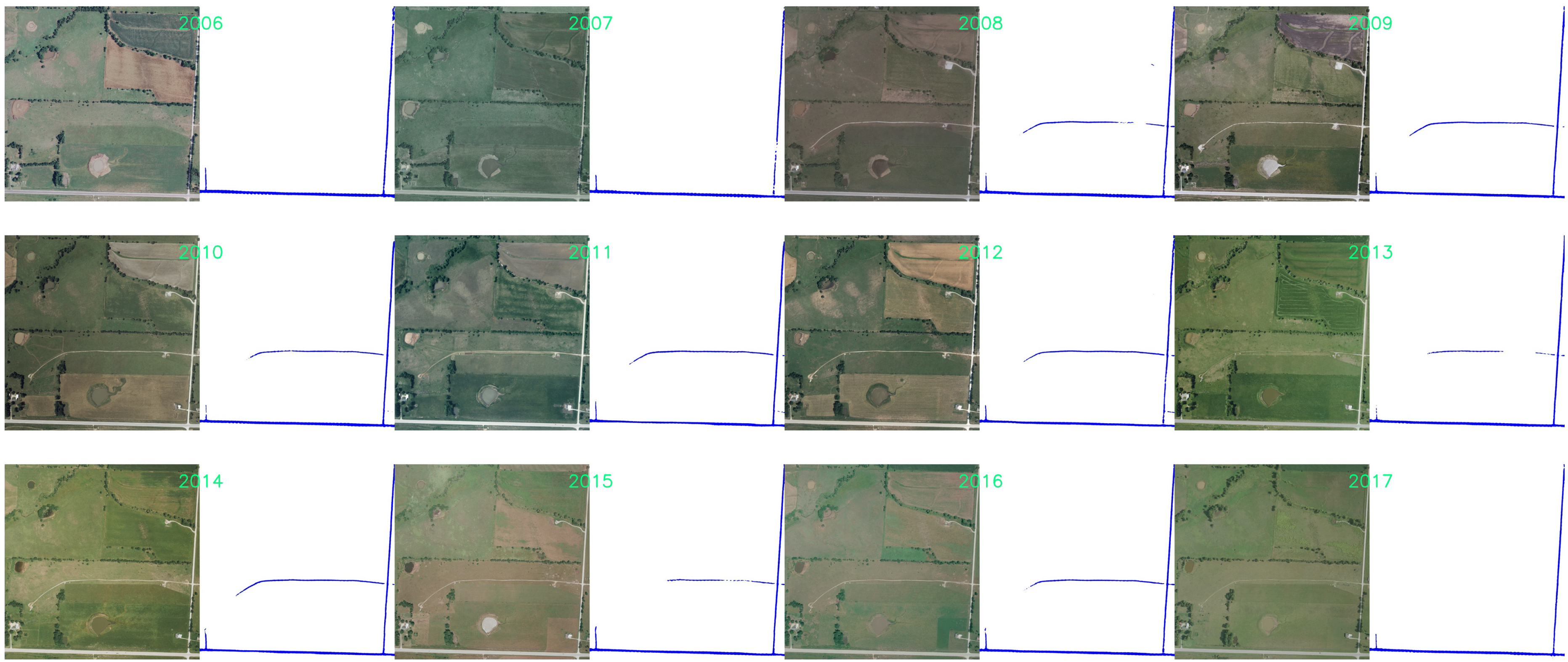


Example 2

Model: AD-LinkNet (transfer learning, rural)

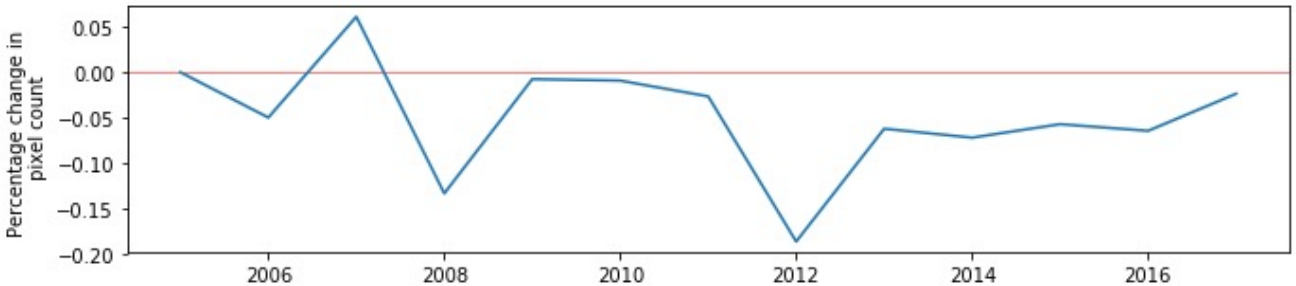


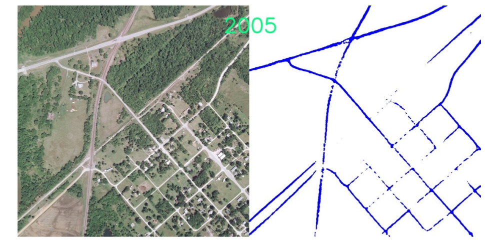


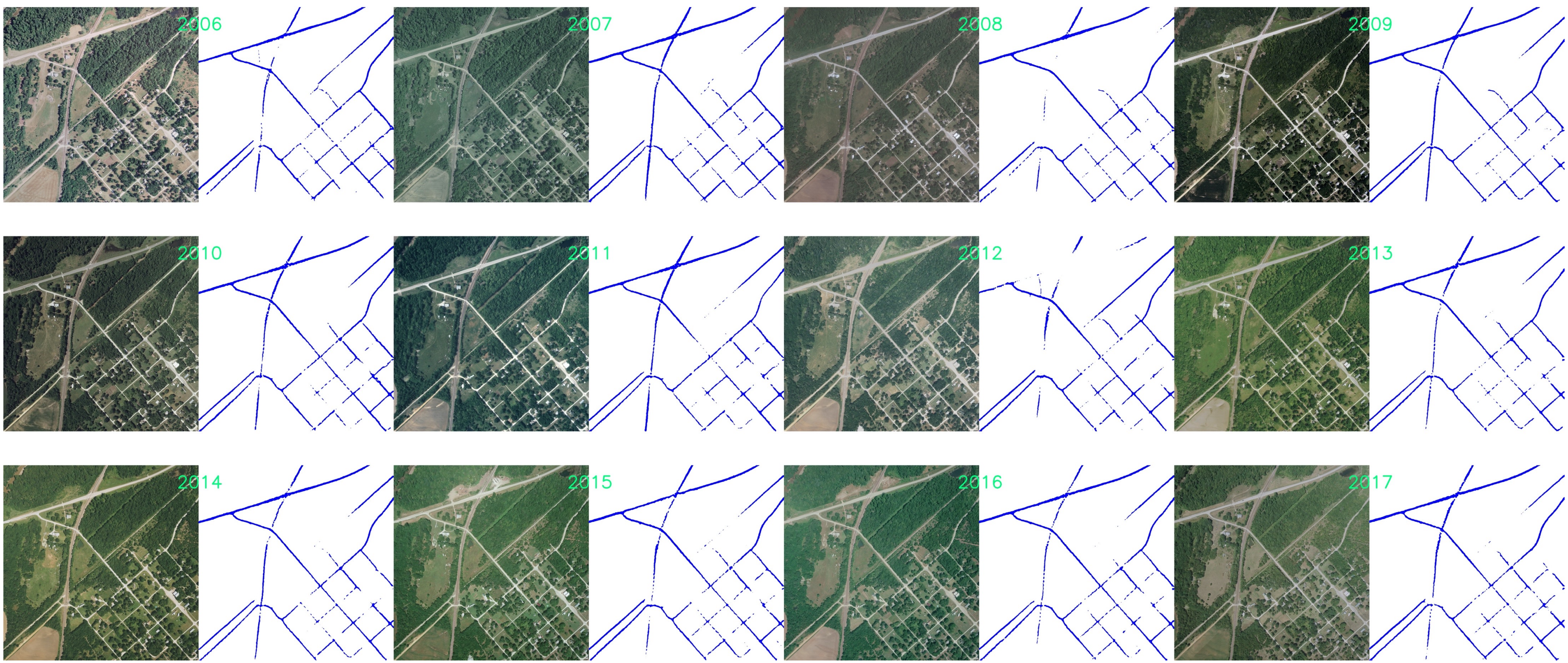


Example 3

Model: AD-LinkNet (transfer learning, rural)







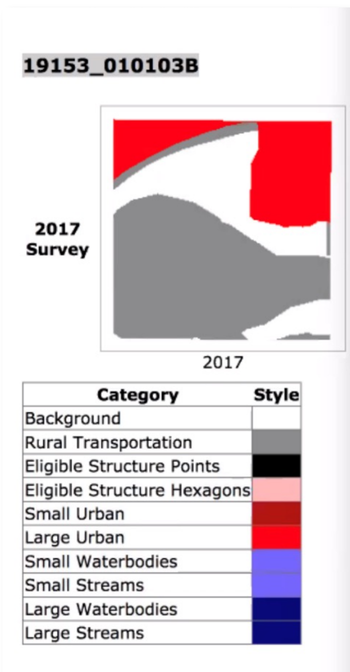
IV. Next steps

Next steps

- Determine change detection thresholds
- **Classification of NRI images by terrain type** (so we can then fit a separate model for each class)
- Conformal prediction
- Hyperparameter optimization
- Consider training on higher resolution images
- Alternate loss functions
- Post-processing
- Run models longer
- Promising tweaks to GLNet
- Deeper version of UNet

Classification of NRI images

- The pre-defined land use types in NRI, similar to CDL for Google help to advance the classification performance
- Select a representative class out of the polygons for certain types



Category	Style
Background	
Rural Transportation	
Eligible Structure Points	
Eligible Structure Hexagons	
Small Urban	
Large Urban	
Small Waterbodies	
Small Streams	
Large Waterbodies	
Large Streams	

Questions?