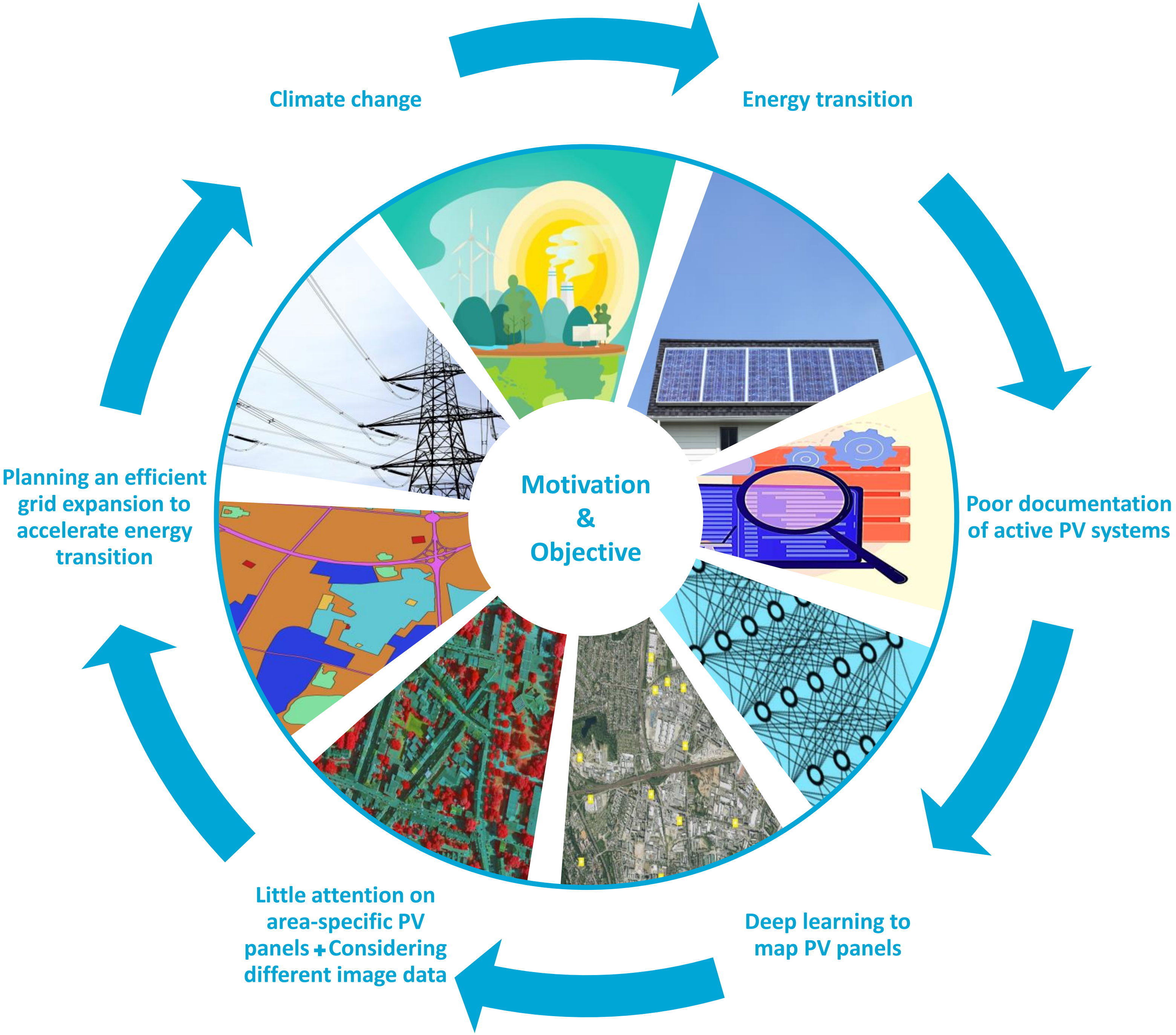
MSc Geomatics for the Built Environment | Thesis presentation

Automated rooftop solar panel detection through Convolutional Neural Networks

Simon Pena Pereira

1st supervisor: Dr. Azarakhsh Rafiee
2nd supervisor: Dr. Stef Lhermitte
Co-reader: Dr. Roderik Lindenbergh

TU Delft


Climate change
Energy transition
Poor documentation of active PV systems
Deep learning to map PV panels
Little attention on area-specific PV panels + Considering different image data
Planning an efficient grid expansion to accelerate energy transition
Motivation & Objective

TU Delft

Introduction - Research Questions

Main question:

To what extent is a CNN with U-Net architecture suitable for detecting PV panels on rooftops in aerial images?

Sub-questions:

- What is the impact of different land use types on the detection of PV panels?
- How is the correlation between roof and panel color affecting the detection of PV panels?
- What is the effect of adding near-infrared data to aerial images on the detection of PV panels?
- How sensitive is the model towards lower resolutions with regard to the panel size?

Theoretical background

Convolutional Neural Network

→ Extracts high-level semantic information from images

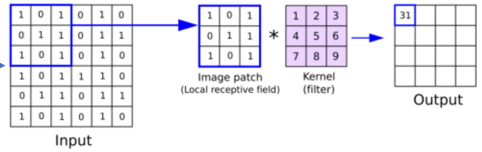
Semantic Segmentation

→ Classified image in which each pixel is associated with a class

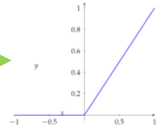
U-Net architecture

- contracting path (left)
- expansive path (right)

Convolutional layer 3 x 3

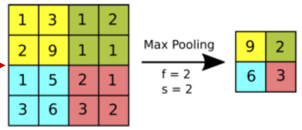


ReLU



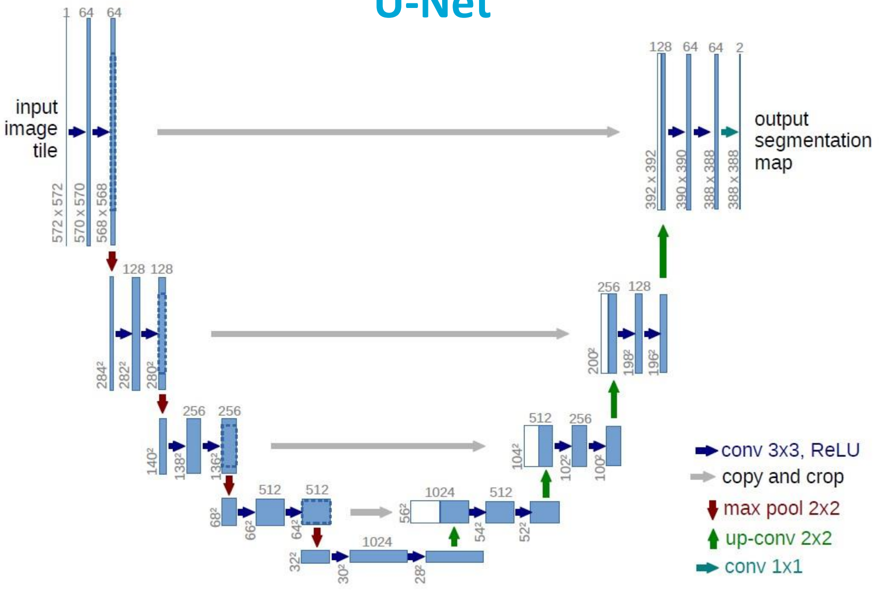
ReLU(x) = max(0, x)

Max pooling 2 x 2



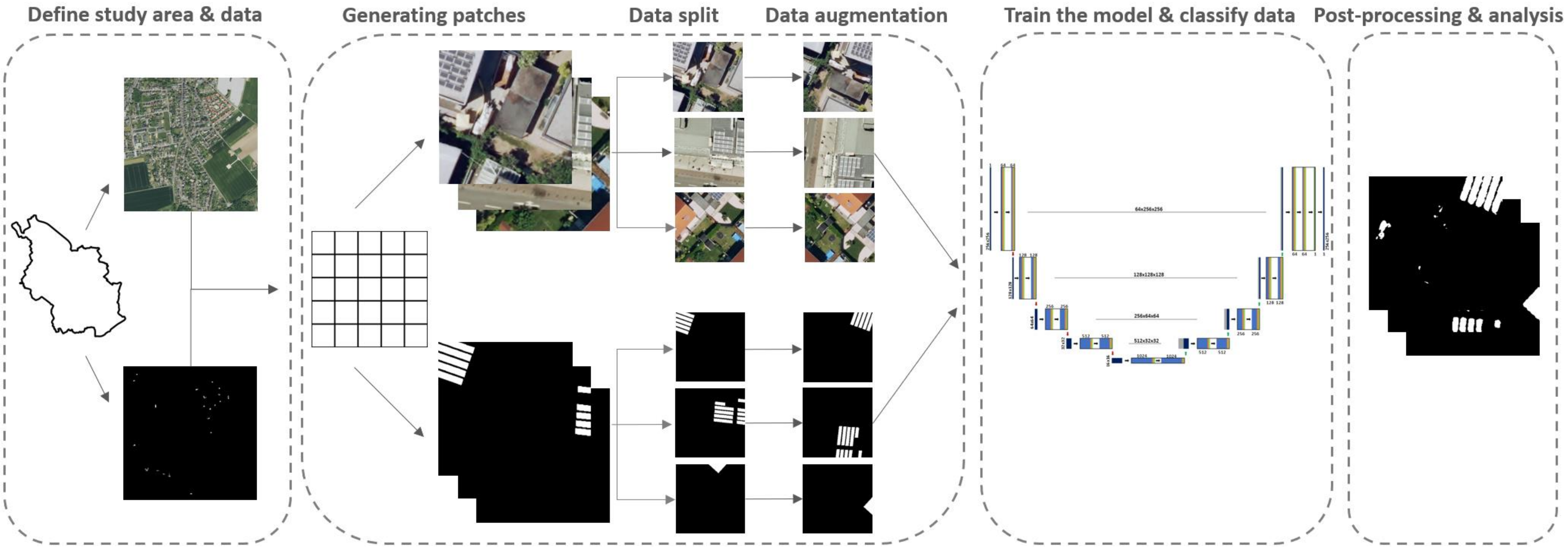
(Reynolds, 2019)

U-Net

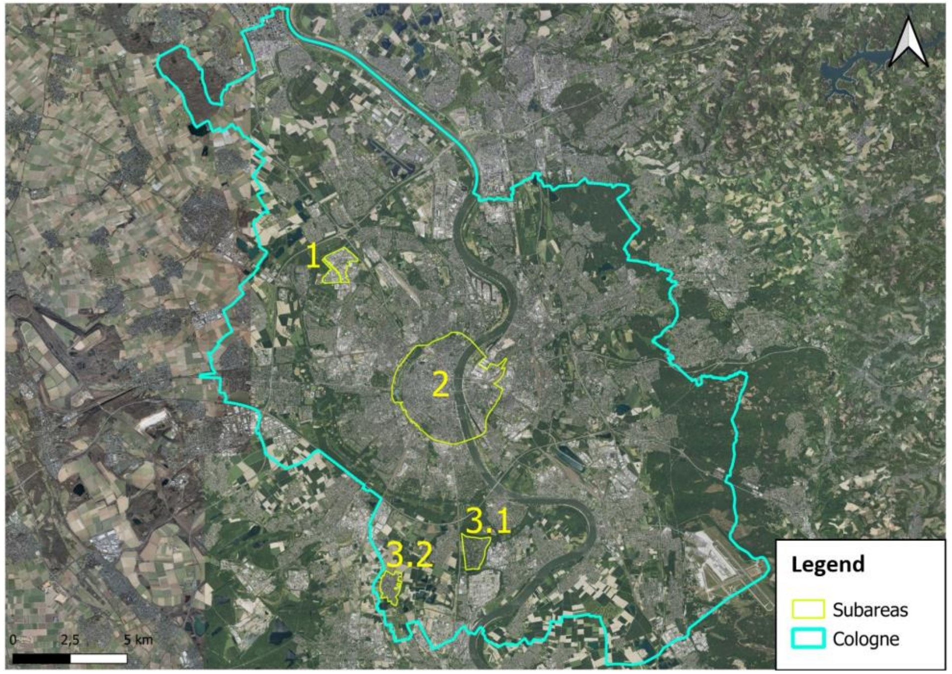


(Ronneberger et al., 2015)

Methodology - Overview



Technical Implementation – Define Study Area



	commercial	city center	suburbs	total
Area (km^2)	1,393	7,694	2,227	**11,314**
Buildings	638	20,998	5,055	**26,691**
Buildings/km^2	456	2,729	2,270	**2,359**

1. Commercial area:
- 295 PV panels
- White-greyish roof

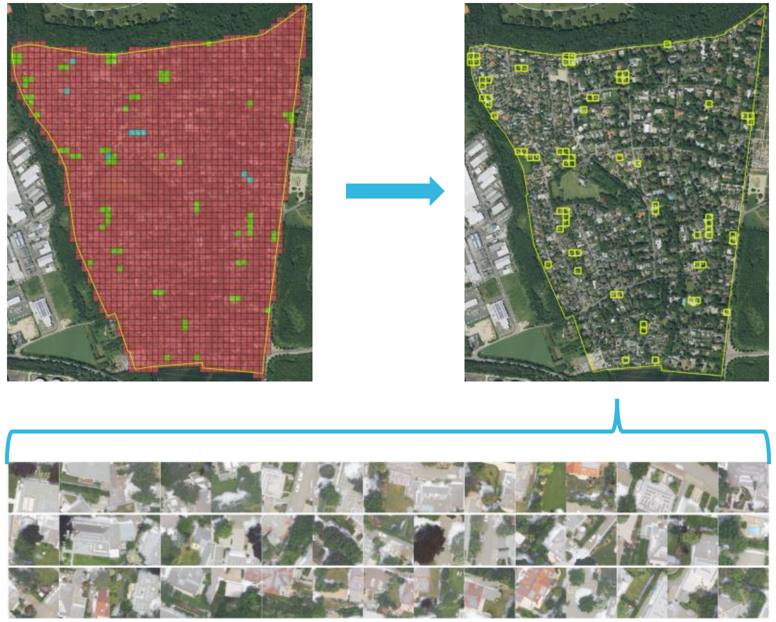
2. City center:
- 40 PV panels
- Greyish roof

3. Suburbs:
- 28 PV panels
- Red roof

	commercial	city center	suburbs	total
Buildings with PV panels	31	62	78	**171**
PV panels	7,994	2,431	2,083	**12,508**
Mean PV panels/building	258	39	26	**73**
Buildings (with PV panel) mean size (m^2)	1,364	418	140	**410**

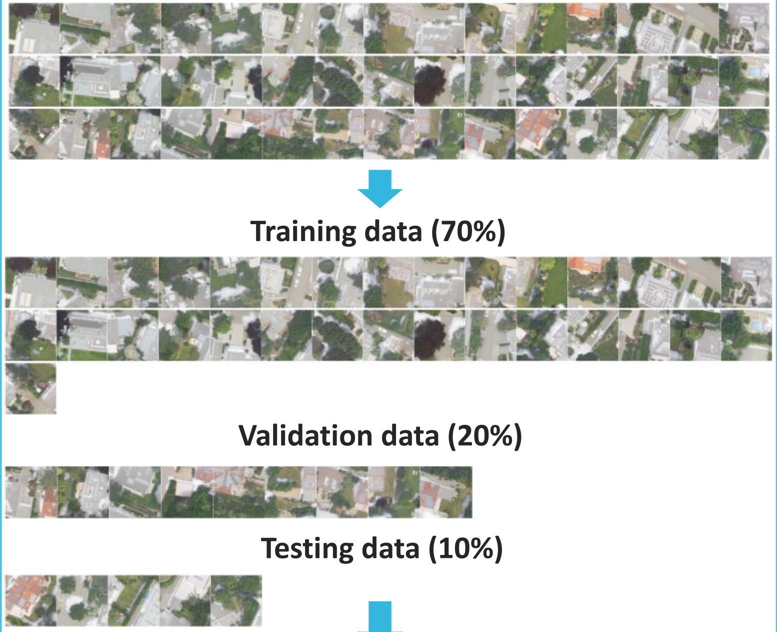
Technical Implementation – Pre-processing steps

Generating patches of 256 x 256 pixels



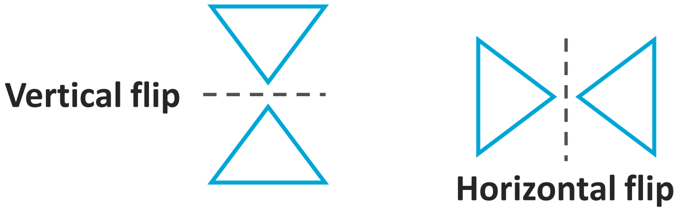
Resolution	commercial	city center	suburbs	total
Number of patches (10 cm)	100	100	100	300
Number of patches (20 cm)	50	77	73	200

Data split



Custom data set for TensorFlow

Data augmentation



Flipping training data randomly (70%)

Flipping validation data randomly (20%)

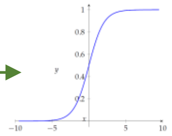
Technical Implementation – U-Net configuration

Weights → **He uniform**

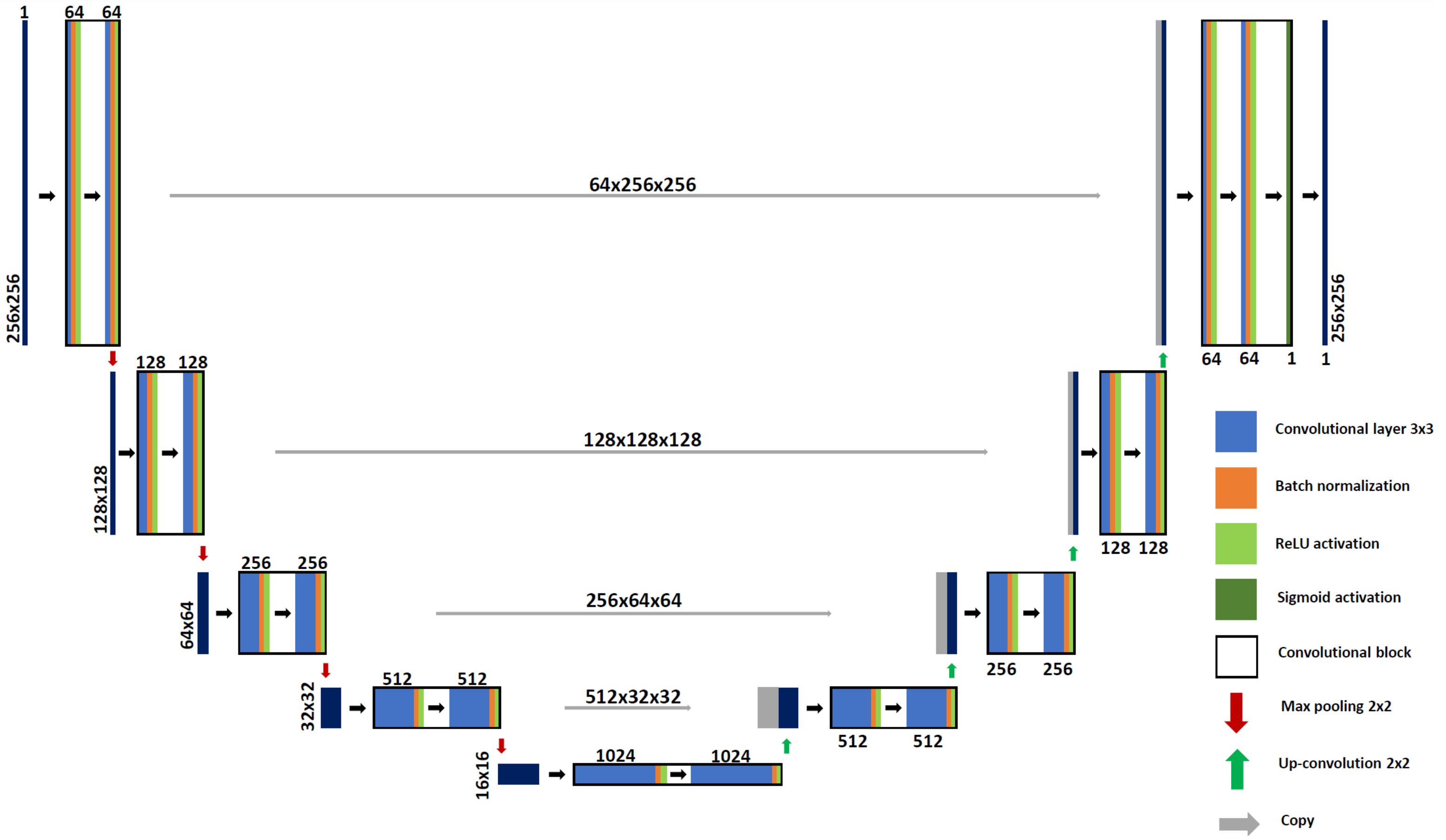
Zero-padding →

$$\begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix} * \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 21 & 17 & 0 \\ 0 & 25 & 14 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Sigmoid →

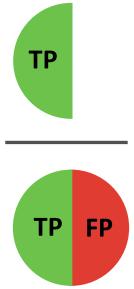


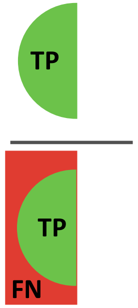
→ **Probabilities [0,1]**



Technical Implementation – Model Evaluation

Quantitative analysis

Precision = 

Recall = 

Harmonic average → F1-Score = $\frac{2 * TP}{2 * TP + FP + FN}$

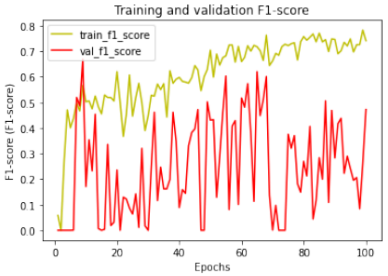
Intersection over Union = $\frac{\textit{Overlap}}{\textit{Union}}$ =  = $\frac{TP}{TP + FP + FN}$

Qualitative analysis *(Visual assessment)*

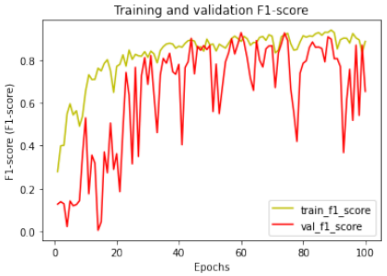
- Comparing images, labels, predicted probabilities, and prediction masks per sample and between areas
- Analysis of mean reflectance per rooftop
  - Labels
  - True Positives
  - True Negatives
  - False Positives
  - False Negatives

Technical Implementation – Loss function & learning rate

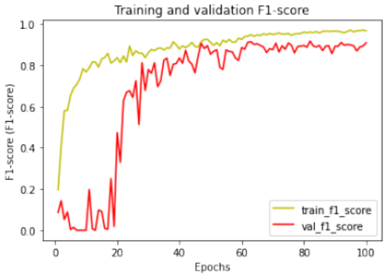
Comparison of binary cross-entropy (BCE) and focal loss (FL)



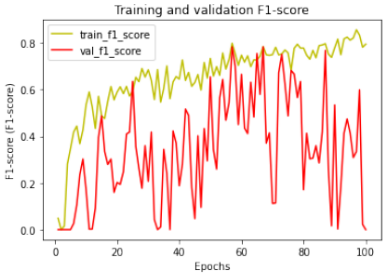
(a) Learning rate = 0.01



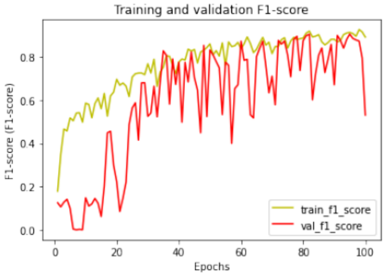
(b) Learning rate = 0.001



(c) Learning rate = 0.0001



(d) Learning rate = 0.01



(e) Learning rate = 0.001

(f) Learning rate = 0.0001

Area	loss function	accuracy (%)	precision (%)	recall (%)	F1-score (%)	IoU (%)
all areas	BCE	98.87	**95.32**	87.29	91.13	91.25
all areas	FL	**99.21**	94.36	**93.76**	94.06	93.97

City center (100 images):

loss	LR	accu. (%)	prec. (%)	recall (%)	F1-score (%)	IoU (%)
BCE	1e-2	96.91	84.74	22.63	35.72	59.31
BCE	1e-3	98.11	90.49	55.93	69.13	75.45
BCE	1e-4	**99.23**	93.97	85.1	89.31	89.95
FL	1e-2	96.21	0	0	0	48.1
FL	1e-3	98.04	**98.39**	49.17	65.57	73.39
FL	1e-4	98.90	84.57	**86.76**	85.65	86.88

Taken parameters:

- Binary cross-entropy (BCE)
- Learning rate: 0.0001
- Optimizer: adaptive moment estimation (Adam)

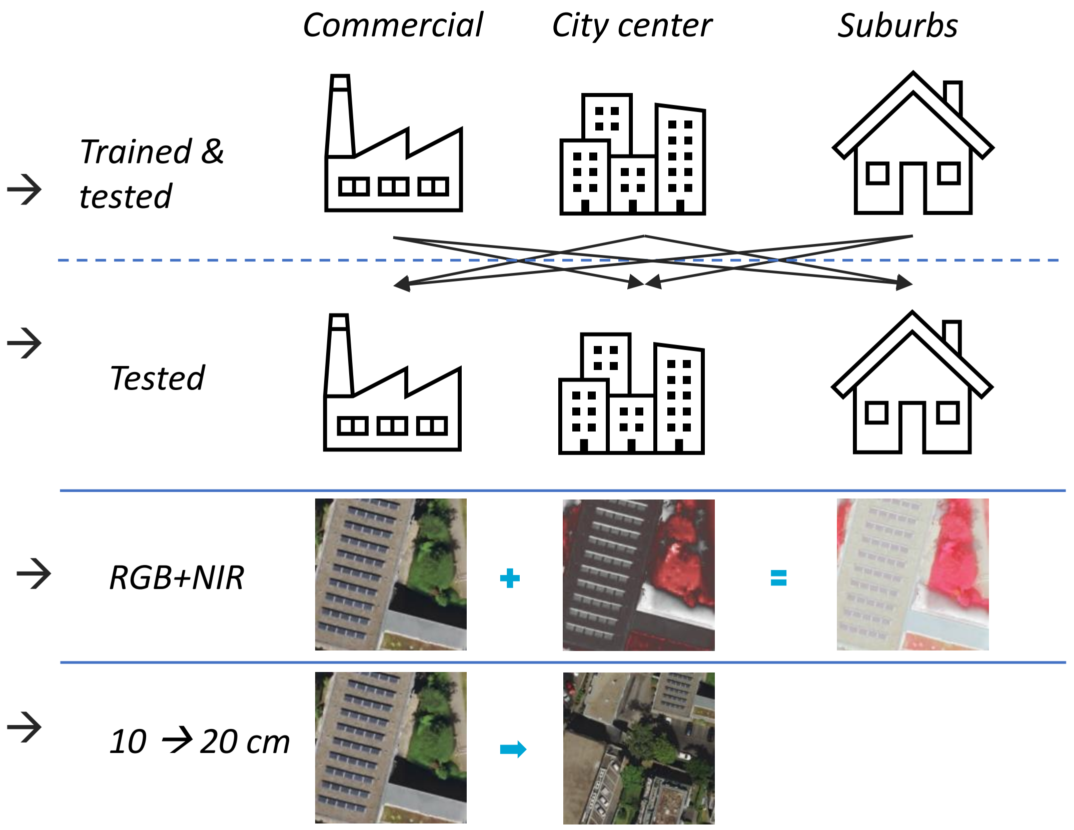
Training and testing experiments

1. Training and evaluating a U-Net within the same area based on TrueDOPs at a resolution of 10 cm with RGB channels

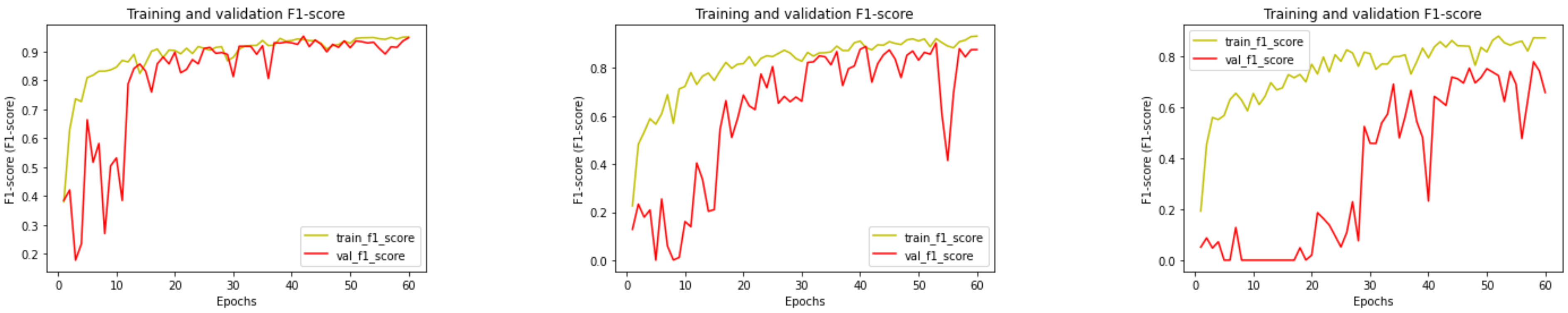
2. Evaluating the U-Net's performance based on **cross-validation**

3. Evaluating the U-Net's performance by training and evaluating with **Near-infrared** (NIR) data

4. Training and assessing the U-Net on **lower-resolution** TrueDOPs



Results – General classifications of all areas



(a) commercial area (b) city center (c) suburbs

Area	precision (%)	recall (%)	F1-score (%)	IoU (%)
commercial	89.40	**91.5**	**90.44**	88.96
city center	89.1	85.59	87.31	88.25
suburbs	**97.86**	60.66	74.89	78.96
all areas	91.64	88.74	90.16	**90.36**

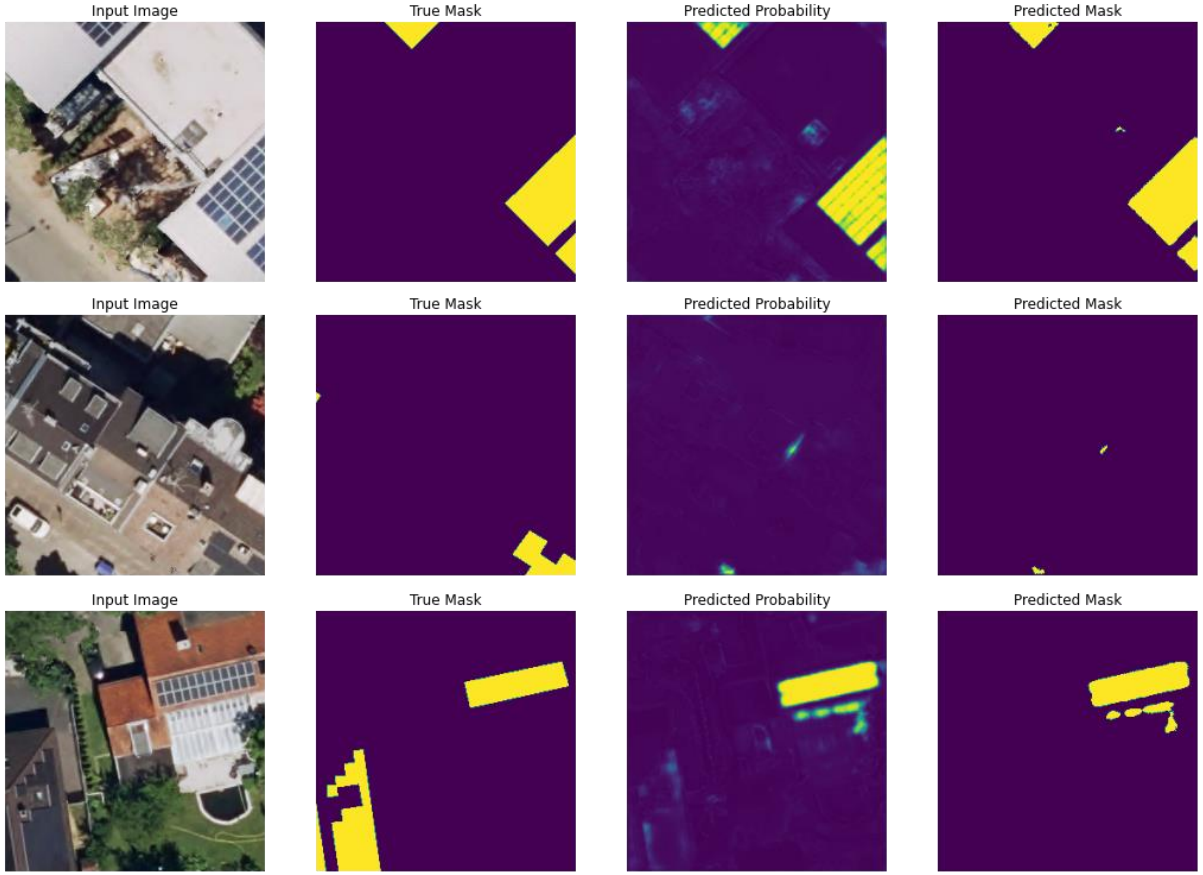
- **Best overall results:** Commercial- and all areas
- **Poorest result:** Suburbs

Results – General classifications of all areas

Commercial area

City center

Suburbs



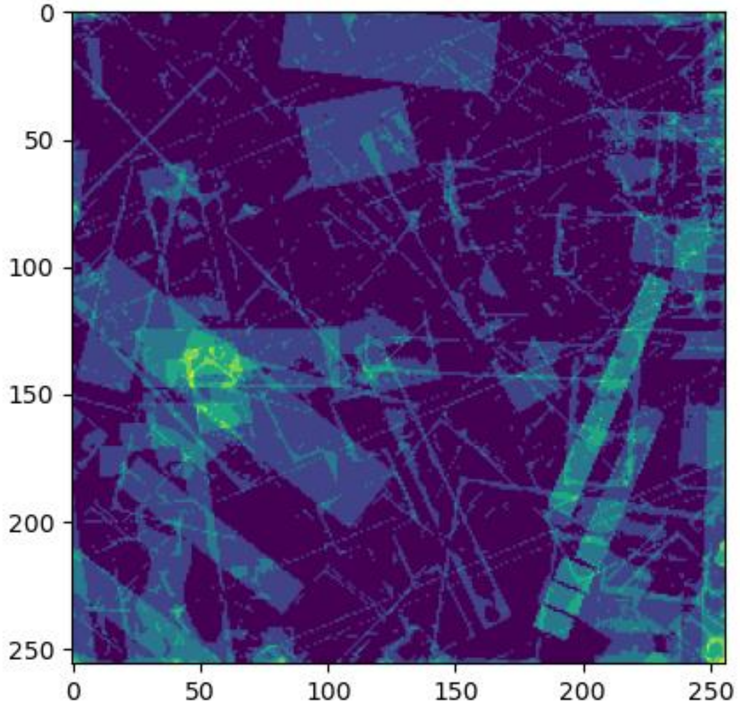
TU Delft

Results – Analysis of potential artifacts

Does the U-Net produce artifacts at the patches' edges?

→ Heat map of all False Negative Classifications
(False Negatives → Not detected PV panels)

→ No systematic error can be found



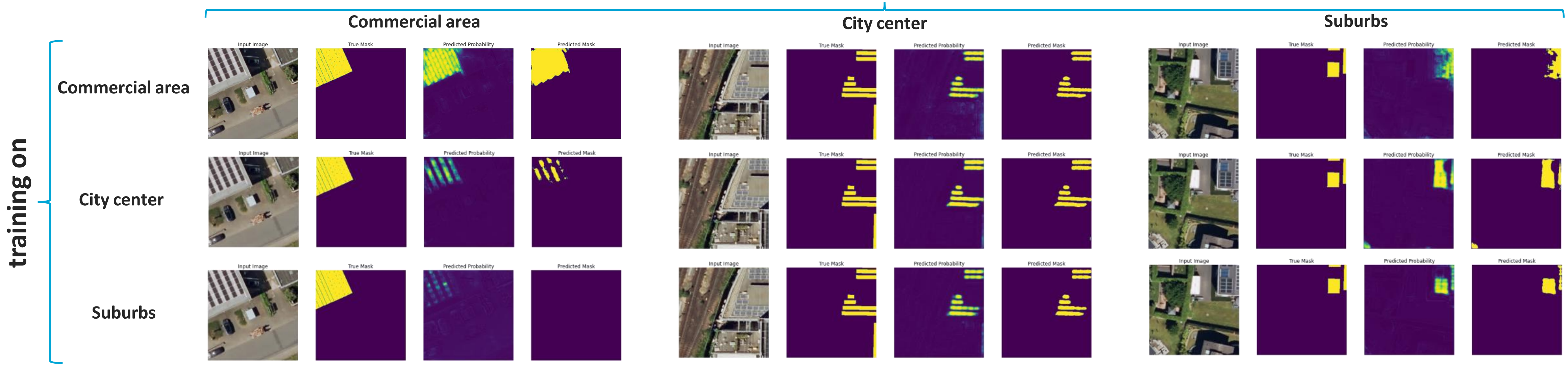
Results – Cross-validation: commercial area, city center, and suburbs

F1-score:

trained/predicted	commercial	city center	suburbs
commercial	90.44%	**72.89%**	61.85%
city center	**59.82%**	87.31%	**77.73%**
suburbs	48.52%	63.49%	74.89%

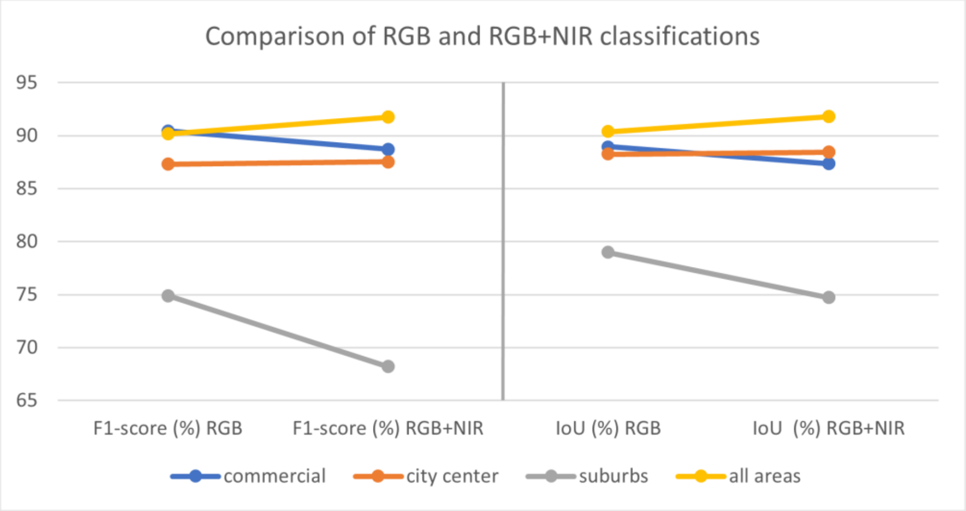
- **Best results:** City center
- **Poorest result:** Suburbs

prediction for



Results – Classification based on TrueDOPs including NIR data

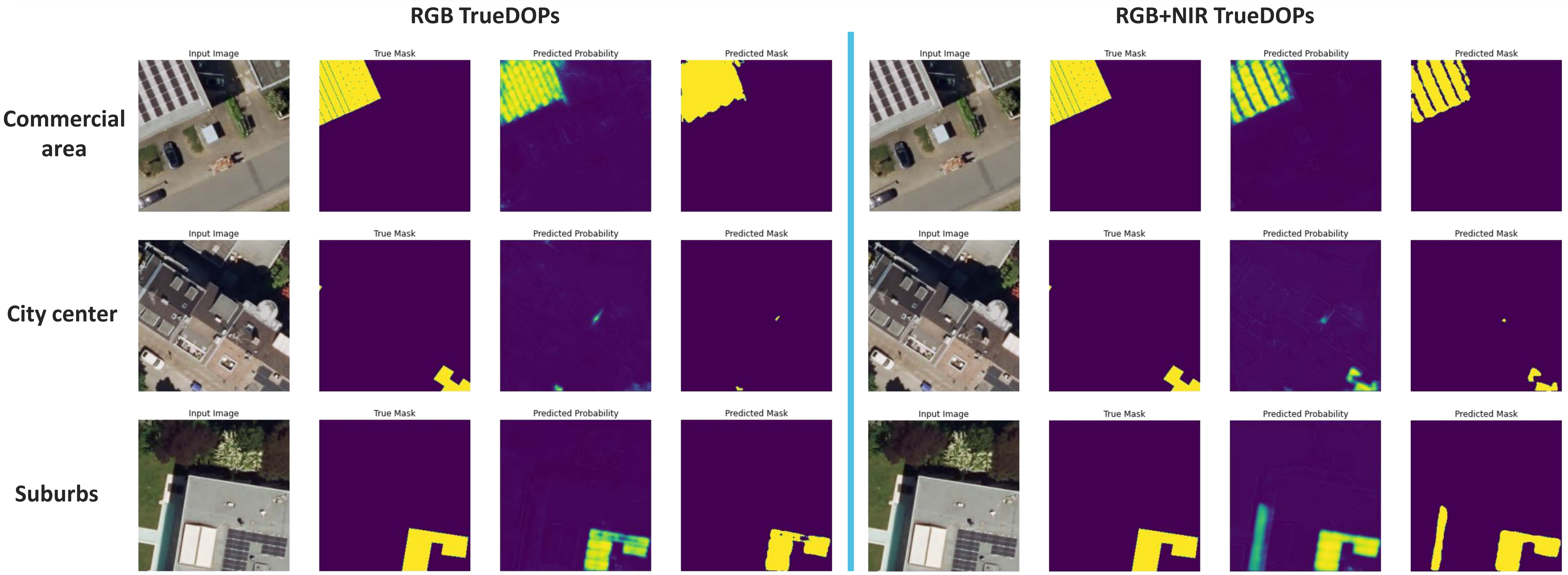
Area	precision (%)	recall (%)	F1-score (%)	IoU (%)
commercial	93.91	84.07	88.72	87.35
city center	92.07	83.41	87.53	88.45
suburbs	**96.81**	52.65	68.21	74.71
all areas	94.06	**89.55**	**91.75**	**91.81**



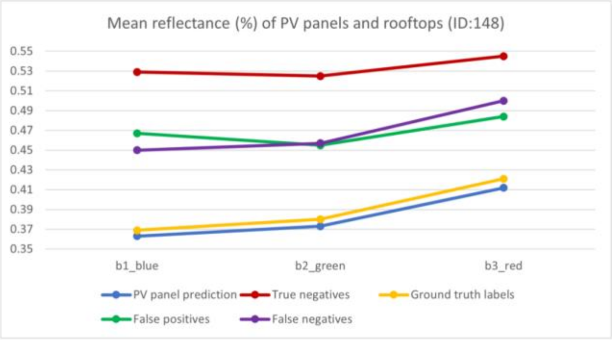
→ Negative impact on the classification of suburb images

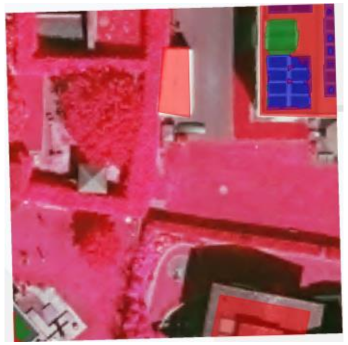
TUDelft

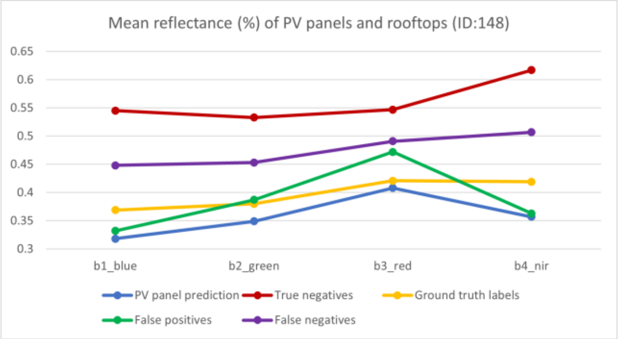
Results – Classification based on TrueDOPs including NIR data

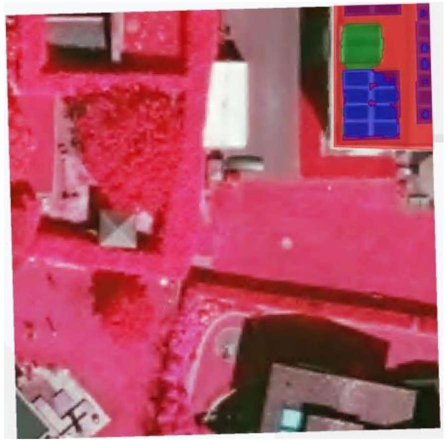


Results – Classification based on TrueDOPs including NIR data

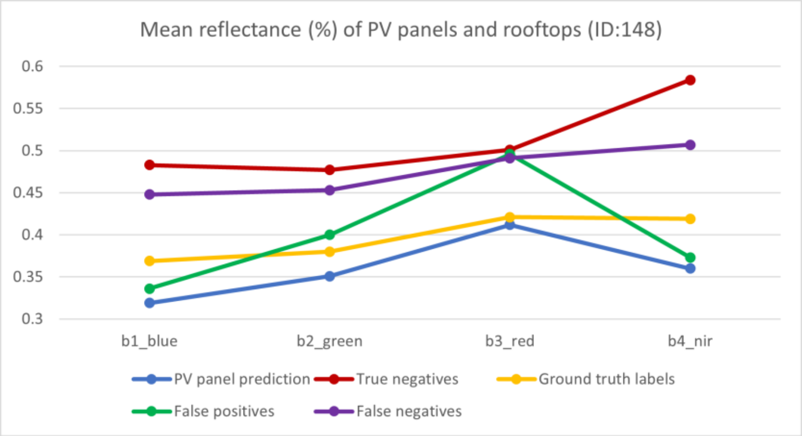








False color (NIR, R, G)

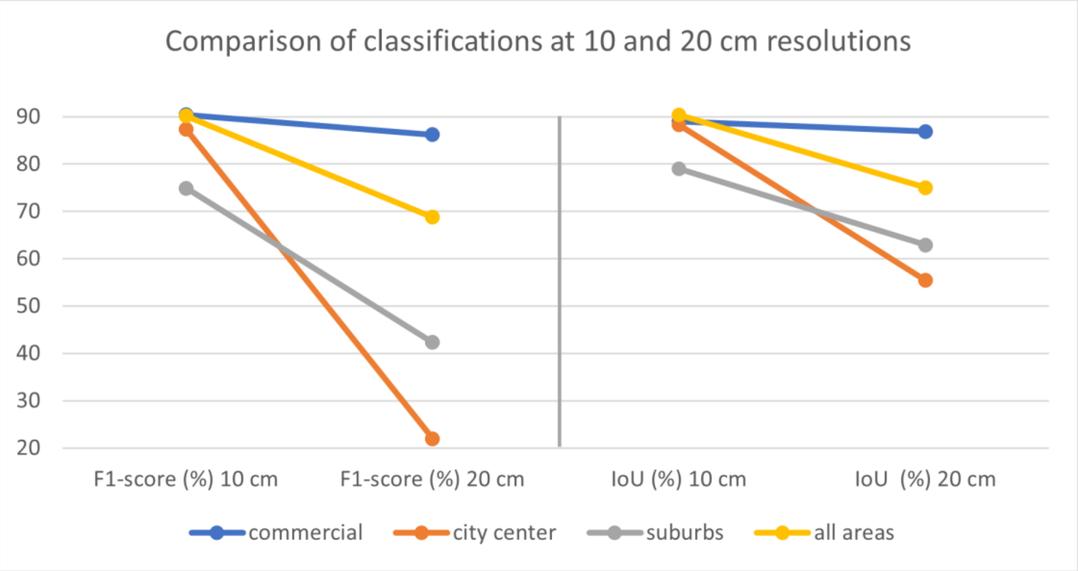


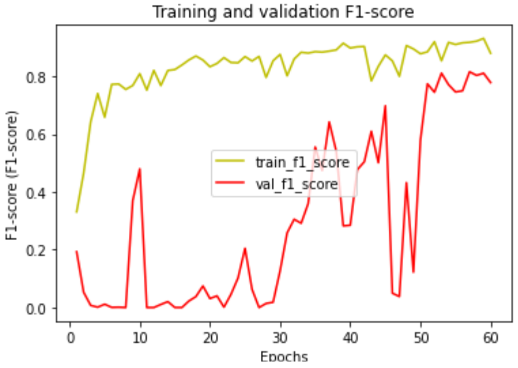
→ Mean reflectance indicates similarities between PV panel and ground truth data

TU Delft

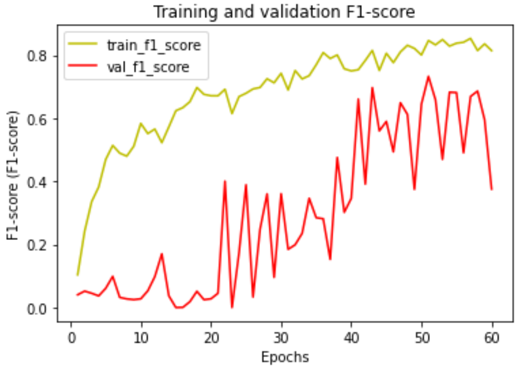
Results – Classification of lower-resolution TrueDOPs

Area	precision (%)	recall (%)	F1-score (%)	IoU (%)
commercial	87.29	**85.17**	**86.22**	**86.89**
city center	**93.12**	12.47	22	55.4
suburbs	85.46	28.12	42.32	62.89
all areas	77.09	62.09	68.78	75.04





(a) commercial area



(b) suburbs

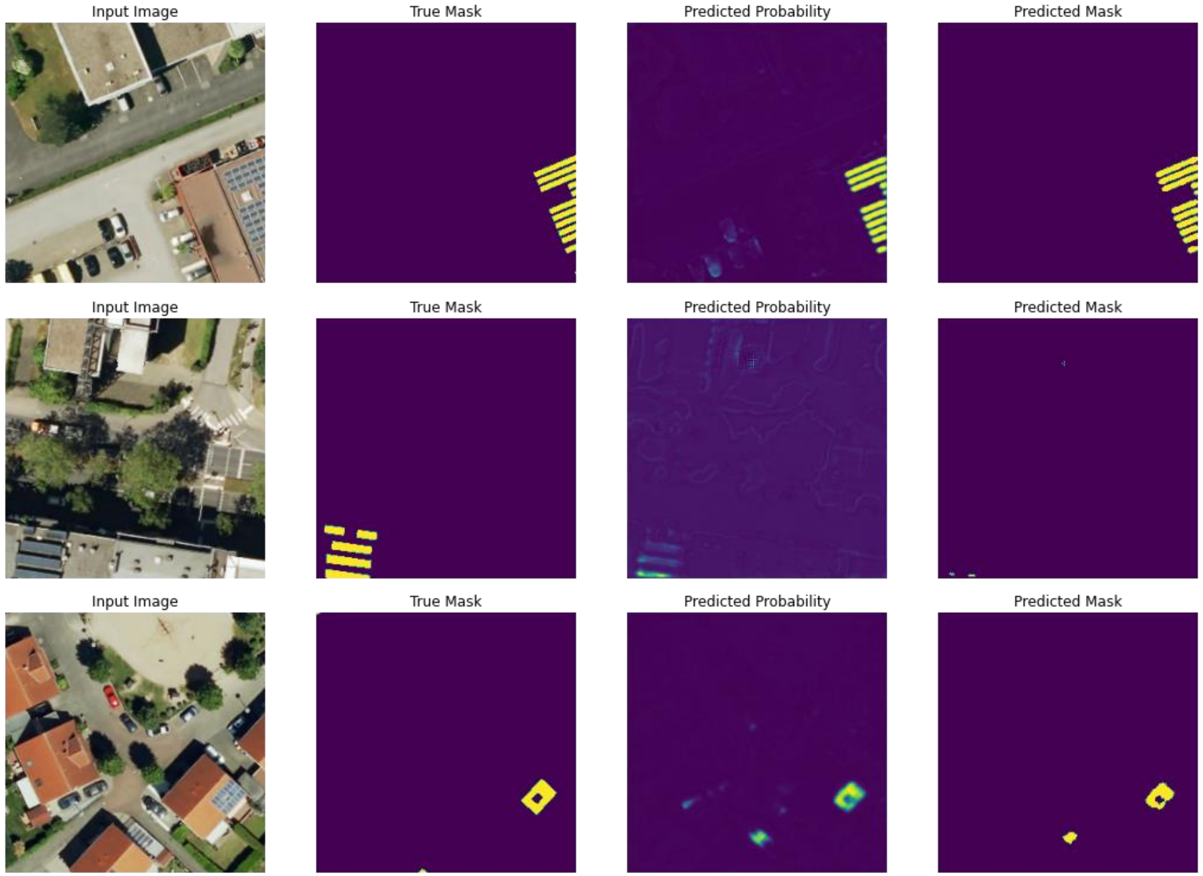
- Barely any impact on training in the commercial area
- Performance drop for all areas, city center, and suburbs
- Notably: Low recall score for city center/suburbs

Results – Classification of lower-resolution TrueDOPs

Commercial area

City center

Suburbs



Discussion - Hyperparameters

Weight initialization	Epochs	Loss function	Learning rate
• Transfer-learning could prevent from fluctuations in the training • Pre-trained weights based on RGB channels	• No early-stopping • Strongly depends on the model's performance and the number of input images • Preventing different regions from over- or underfitting	• Binary cross-entropy outperformed Focal loss • Weighted loss functions to address class imbalance • Class imbalance is not present in all areas	• Depends on the number of input images and on the performance
→ randomly by He uniform to allow comparison between RGB and NIR	→ 60 epochs	→ Binary cross-entropy	→ 0.0001 due to few training images

TU Delft

Discussion – Quantitative & Qualitative Results

RGB classifications	Cross-validations	Near-infrared	Lower-resolutions
• Higher precision than recall score except for the commercial area • Heterogeneous rooftops cause more false negatives	• Performance drops when validating the model in a different region than where it was trained *(Jong et al. 2020)* • Similar effect on a local level, especially between commercial areas and suburbs	• Rarely examined in research • Mixed results	• Lower precision due misclassification of small objects • Significant drop in recall scores for heterogeneous areas with class imbalances

TU Delft

Limitations

- **Collecting ground truth data:** Only annotations of high confidence
- **Amount of input data:** Little training data
- **Data augmentation:** No changes of brightness, contrast, saturation, or hue
- **Output format:** PNG instead of TIFF

Conclusions – Research Questions

Sub-questions:

- **What is the impact of different land use types on the detection of PV panels?**

Answer: ***+ Commercial area:*** *Homogeneity of commercial areas + large PV systems → facilitate detection*

- Suburbs: *Greater variation of rooftops + smaller PV systems → poor classification results*

- **How is the correlation between roof color and panel color affecting the detection of PV panels?**

Answer: ***+ Commercial area:*** *High contrast → facilitates detection*

- Suburbs: *Low contrast between black roofs and black PV panels → impairs detection rate*

Conclusions – Research Questions

Sub-questions:

- **What is the effect of adding near-infrared data to aerial images on the detection of PV panels?**

Answer: ***+ All areas:*** *Slight improvement; might be caused by inconsistency of training*

- Suburbs: *Causing more false negatives*

- **How sensitive is the model towards lower resolutions with regard to the panel size?**

Answer: **+** *Barely any effect when detecting large PV systems*

- *Sensitive towards lower-resolution images with small PV systems*

Conclusions – Research Questions

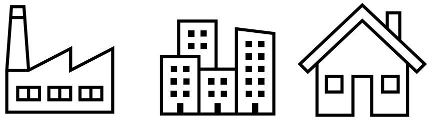
Main question:

To what extent is a CNN with U-Net architecture suitable for detecting PV panels on rooftops in aerial images?

Answer:

- A U-Net is suitable for classifying PV panels on RGB TrueDOPs at 10 cm spatial resolution in patches of 256 x 256 pixels
- It works better for homogeneous surroundings with white or greyish rooftops and large PV systems

Contribution

Land use types	Rooftop colors	NIR	Change of resolution
			
→ Emphasizes the impact of differences in land use types and their characteristics on the detection rate	→ A better understanding of: • Importance of contrast • Visual pattern of PV frames • precision and recall	→ No significant improvement in the model's performance	→ Importance of proportion between image dimensions, spatial resolution, and the PV system

TUDelft

Future Work

→ Adapt the composition of training data and the hyperparameter to the urban and architectural properties of the area of interest as well as to the PV system sizes

- **Additional data:** Height data or building footprint; If available, include thermal infrared imagery
- **Classes:** PV panels and Solar Thermal Collectors; Black and Blue PV panels
- **Amount of training data:** Data augmentation & Synthetic training data
- **Weights:** Transfer learning should be considered for RGB images
- **Regularization:** Appropriate number of epochs should be chosen manually; Batch normalization; Dropout

Thank You!