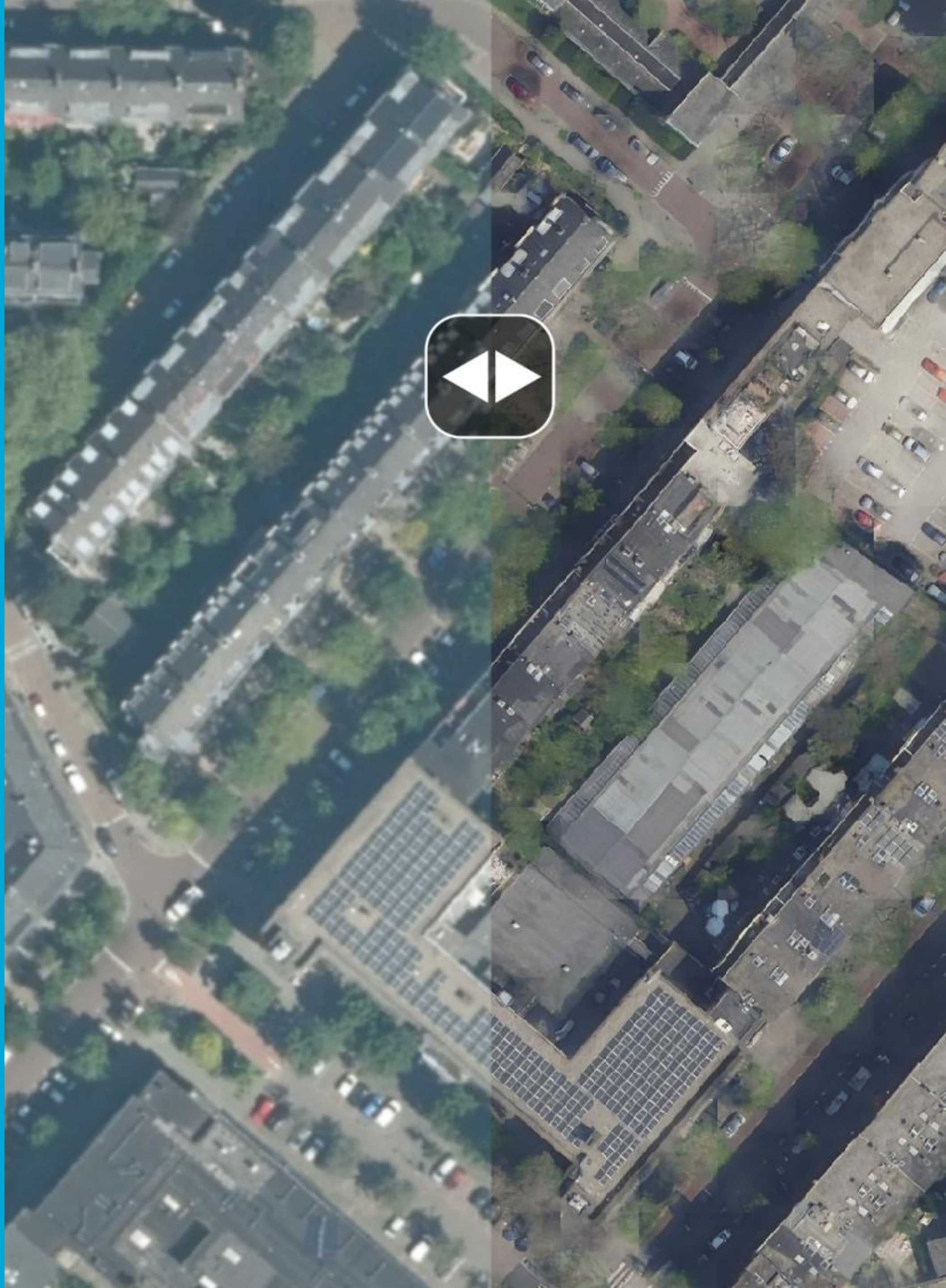
# Super-Resolution for Enhanced Aerial Imagery

P5 Presentation | Michalis Michalas

30-6-2025

Supervisors: Dr.ir. Martijn Meijers
Dr. Azarakhsh Rafiee
Co-reader: Nail Ibrahimli
External Supervisors: Sven Briels
Camilo Caceres



# Contents

01

# Introduction

# Problem Statement

## Applications of High-Resolution Aerial Imagery:

- Urban Planning
- Object Detection
- Environmental Monitoring

## Challenges:

- Acquisition cost
- Sensor noise
- Optical distortions
- Limited availability

***Super-resolution bridges this gap—transforming available LR to HR for actionable insights despite seasonal differences.***

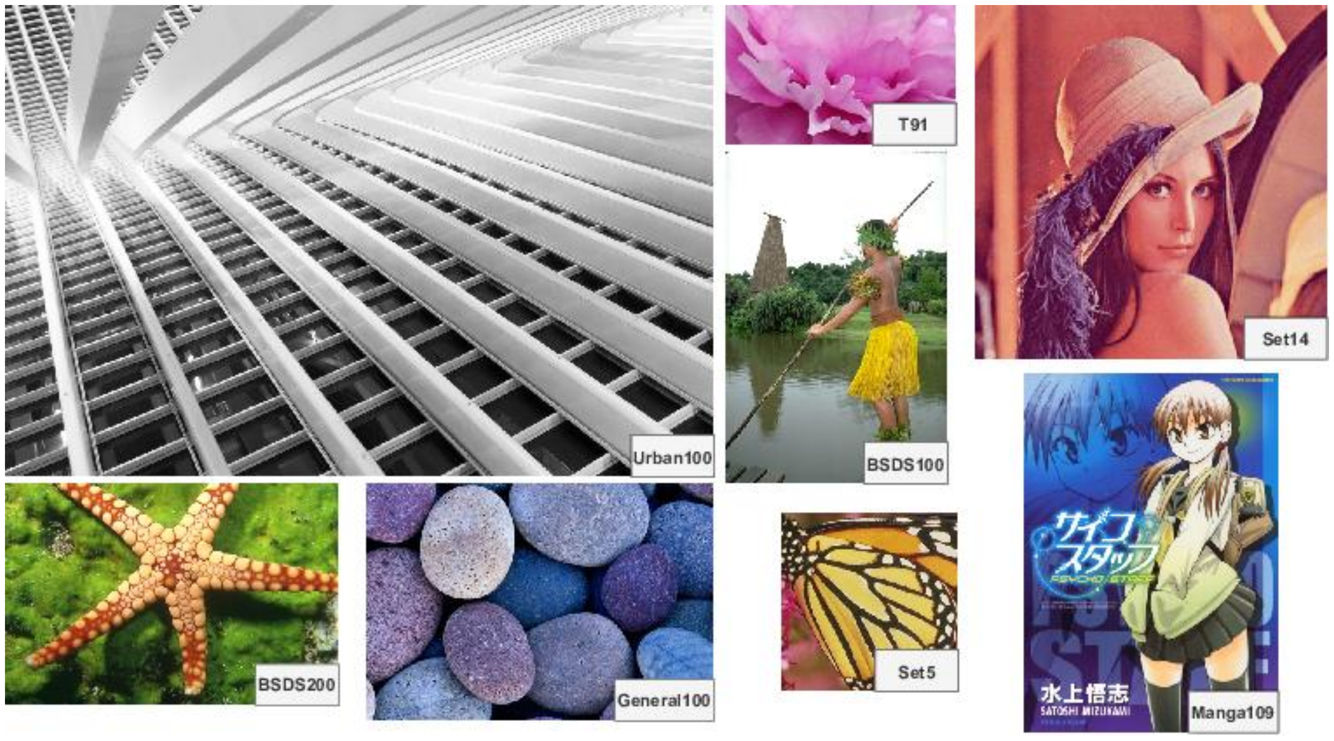
# State of the Art & Research Gap

## State of the Art

- Most SR research focuses on natural or facial imagery
- Few works adapt GANs to aerial imagery or geospatial use cases
- Downstream task performance rarely evaluated

## Research Gap

- Lack of domain-specific architectures for aerial imagery
- Limited integration of SR with geospatial pipelines
- No robust evaluations under real-world, misaligned conditions




Source: Moser et al. (2022) — *Hitchhiker's Guide to Super-Resolution*

*To what extent can GAN-based super-resolution enhance 25 cm aerial imagery to 8 cm, ensuring its applicability for object detection tasks?*

# Research Questions

## Main Question

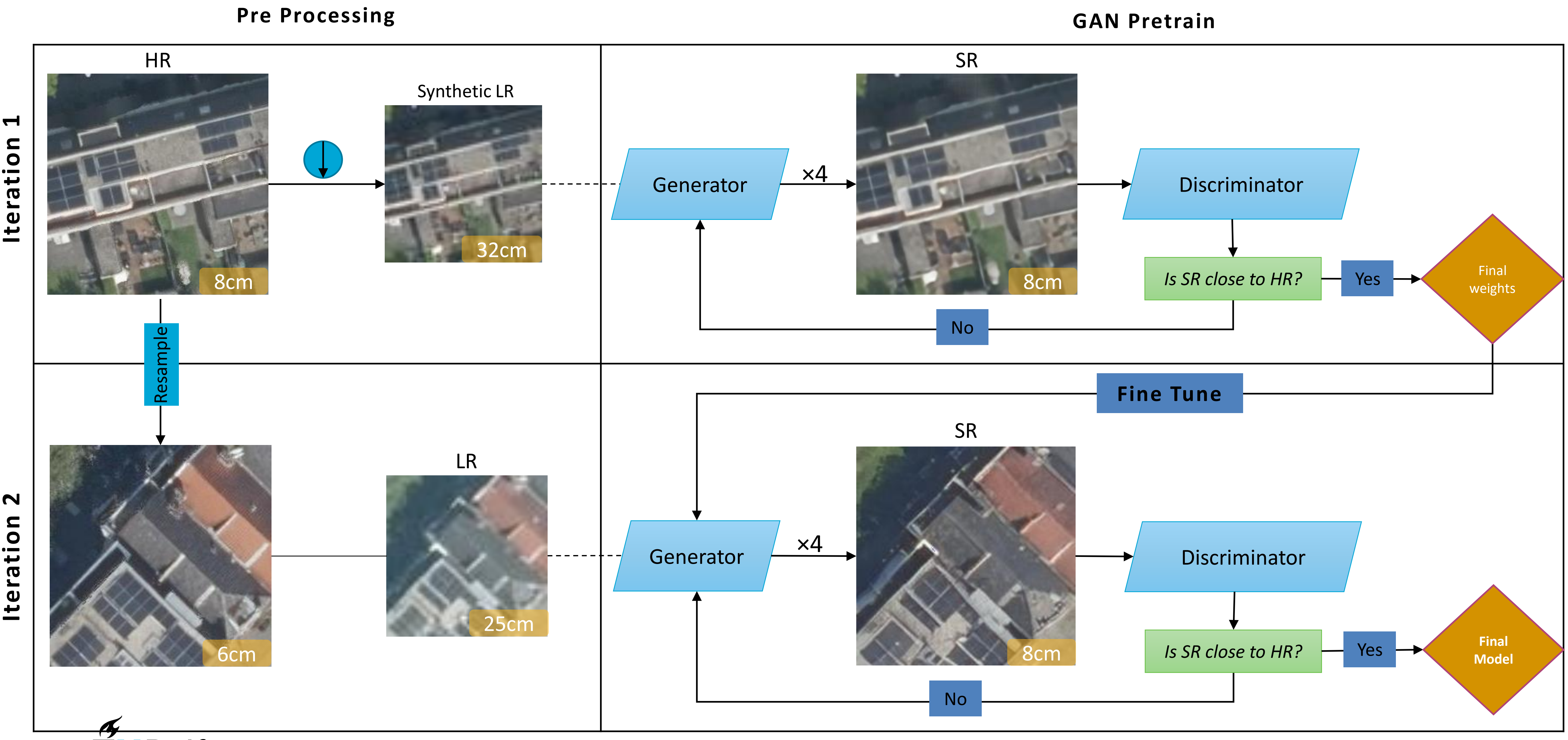
To what extent can GAN-based super-resolution enhance 25 cm aerial imagery to 8 cm, ensuring its applicability for object detection tasks?

## Sub-questions

1. How accurately can a GAN reconstruct 8 cm HR images from 25 cm LR aerial inputs, especially for building edges and solar panels?
2. How do seasonal differences between HR and LR images (e.g., winter HR vs. Summer LR) affect GAN performance, and can domain adaptation via pre-training on synthetic data mitigate these effects?
3. What are the limitations of GANs in preserving geometric fidelity (e.g., artifacts, hallucination) for geospatial use cases?
4. What metrics best assess SR image quality for downstream object detection tasks?

02

# Methodology


Pre Processing
GAN Pretrain
Iteration 1
Iteration 2
HR
8cm
Synthetic LR
32cm
Resample
6cm
LR
25cm
Generator
×4
SR
8cm
Discriminator
Is SR close to HR?
Yes
No
Final weights
Fine Tune
Generator
×4
SR
8cm
Discriminator
Is SR close to HR?
Yes
No
Final Model

03

# Implementation

# Model Architecture

SRGAN adapted by *Ledig et al. (2017)*

Generator enhancements:

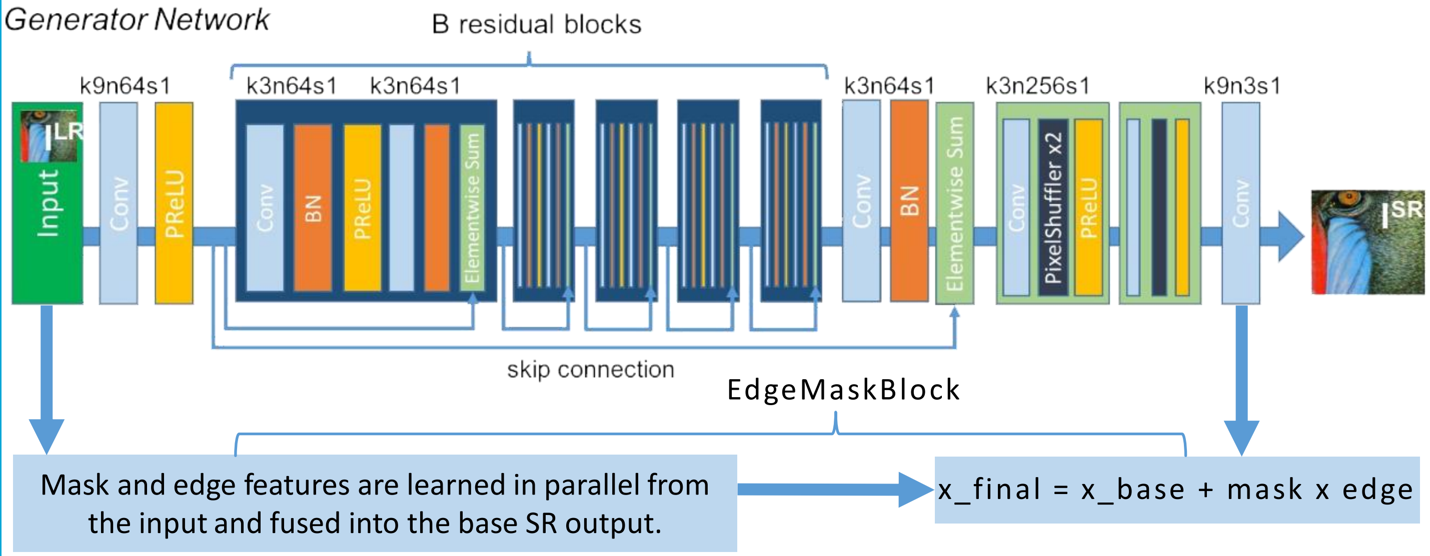
- Edge-aware residual blocks
- Mask-guided refinement module

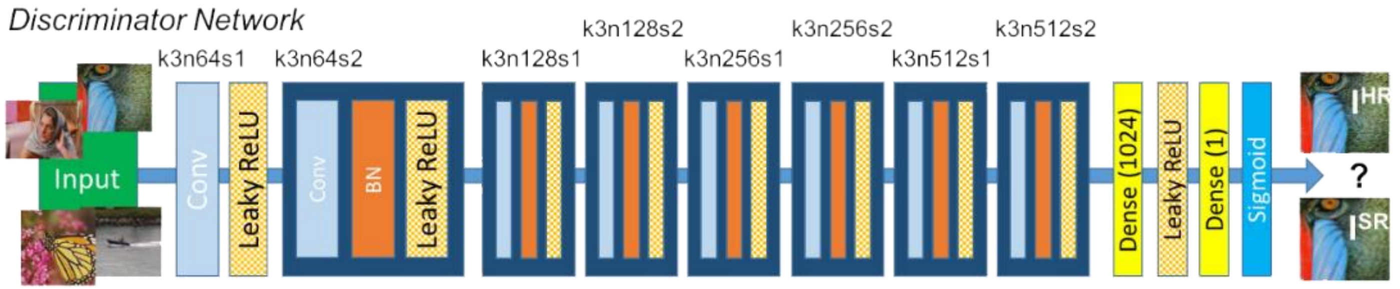
Discriminator:

- Patch based binary classifier

Losses:

- MSE (pretraining)
- Perceptual (VGG features)
- Adversarial (fine tuning)



# Training Pipeline

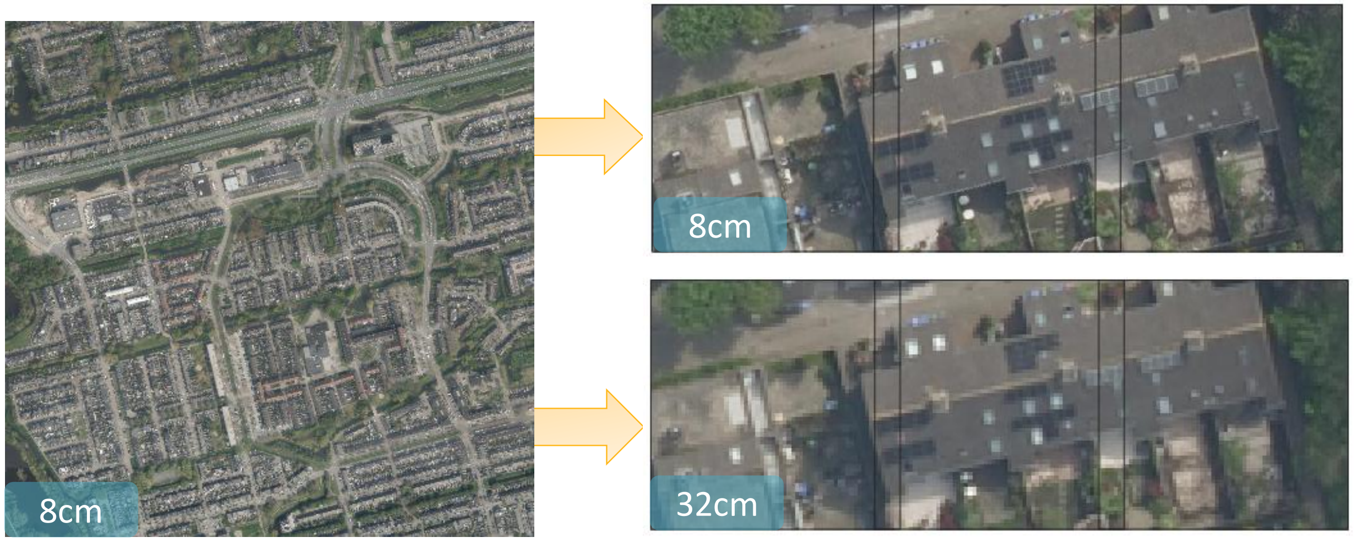
## Tiling Strategy

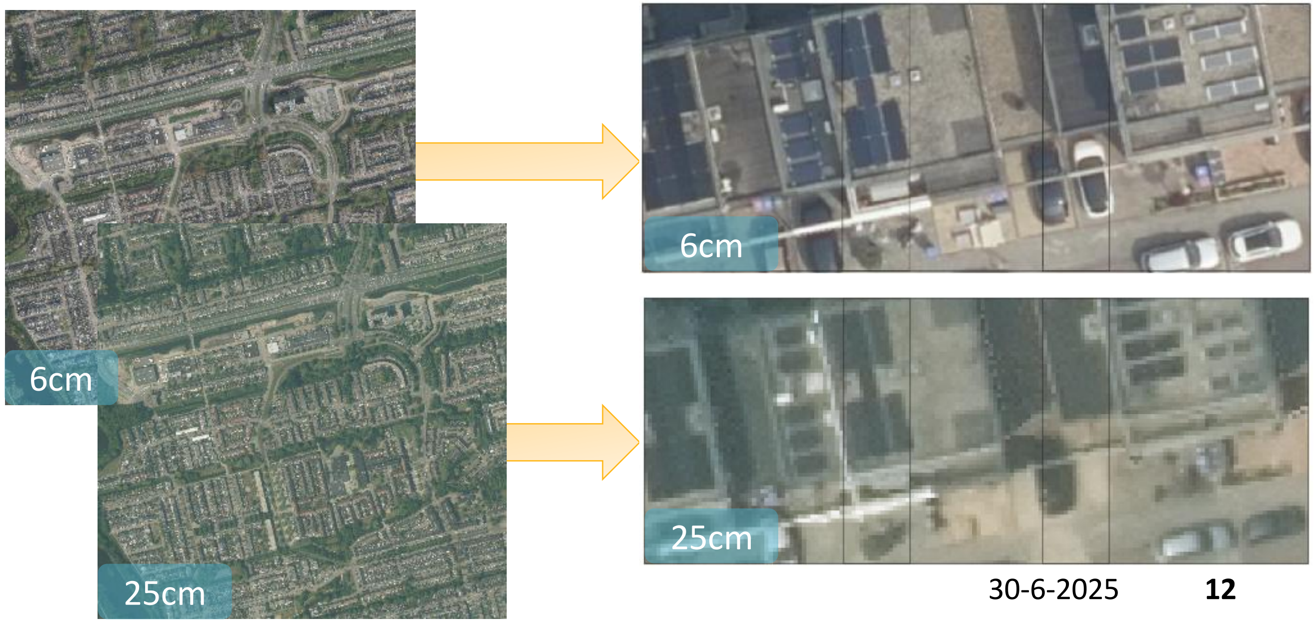
Iteration 1:

- Source: HR raster 8cm
- HR tiles 256 x 256 8cm
- Synthetic LR tiles 64 x 64 32cm (Bicubic)
- Overlap: 10%

Iteration 2:

- Sources:
  - HR raster upsampled to 6cm (from 8cm)
  - LR raster 25cm
- HR tiles 256 x 256 6cm
- LR tiles 64 x 64 25cm
- Overlap: 20%



# Training Pipeline

## Evaluation Strategy

Iteration 1:

- Train/Test split 80/20 with : Delft / Rotterdam

Iteration 2:

- Train test split 80 – 20 with: Delft/ Rotterdam /Utrecht
- Tested generalization to: Den Haag / Zwolle

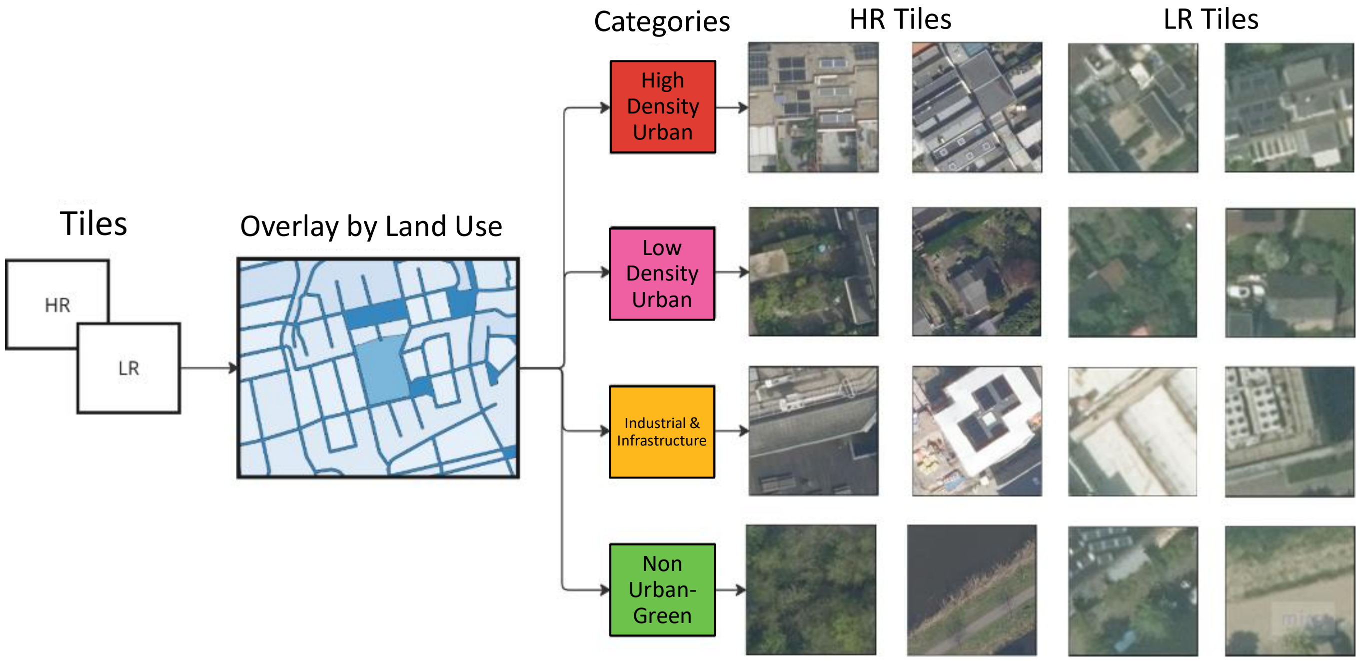
Evaluation Metrics:

- PSNR – Pixel level fidelity
- SSIM – Structural similarity
- LPIPS – Perceptual similarity (deep features)

Downstream Evaluation:

- Segment Anything Model (Meta AI)
- Semantic Segmentation (READAR B.V.)

***Tile Categorization using Urban Atlas:***




| Parameters | Iteration 1 | Iteration 2 |
| --- | --- | --- |
| Samples | 2.300 | 23.800 |
| Pre train epochs | 2000 | 0 |
| Fine train epochs | 4000 | 2000 |

***Comparison Baseline:***

*All evaluations (metrics & downstream tasks) are benchmarked against:*

***Bicubic Upsampling** at matching resolution*

***SRGAN** performance is interpreted relative to this standard baseline.*

04

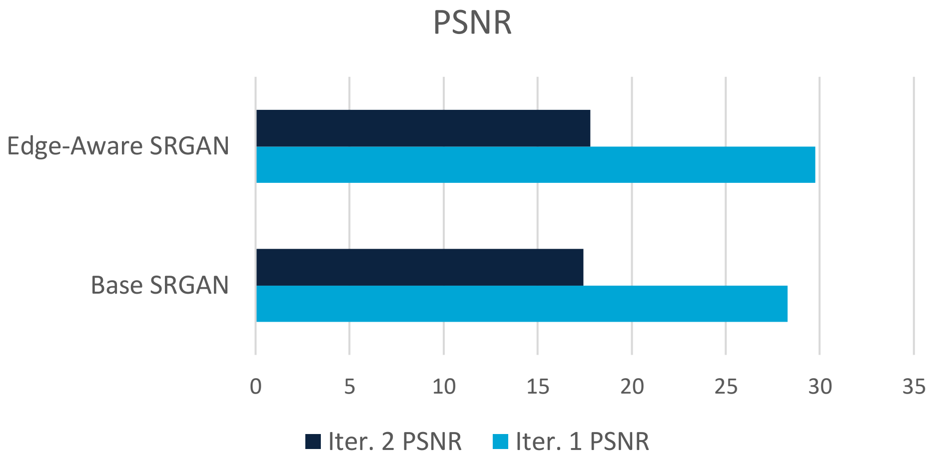
# Results

# Results

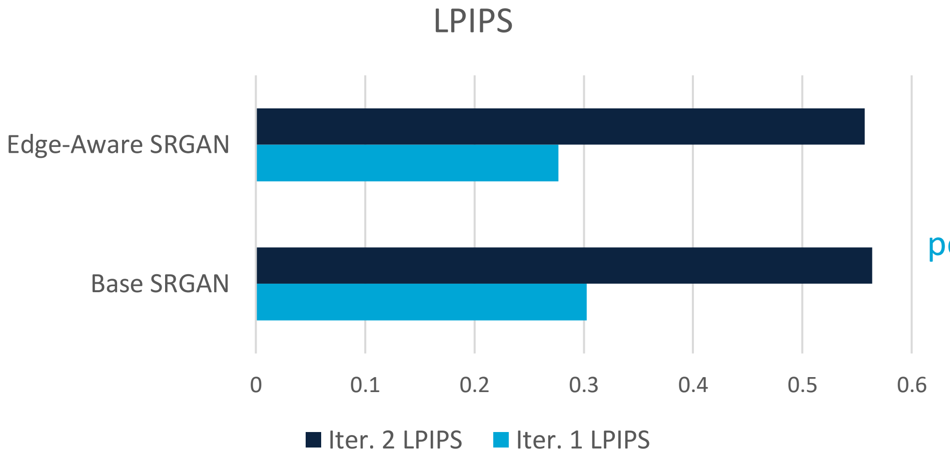
## Base model vs Edge Aware model

### Key Findings

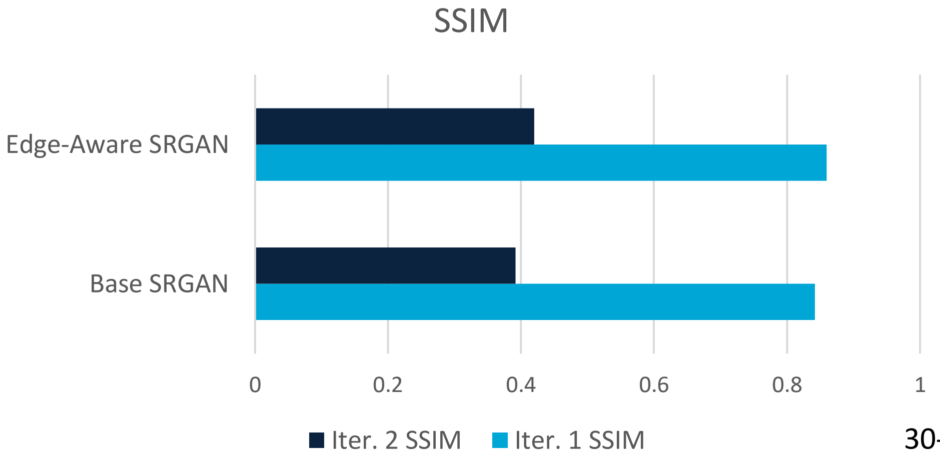
- PSNR improves in both iterations indicating less noise and better pixel-level accuracy
- LPIPS decreases reflecting better perceptual quality
- SSIM increases suggesting stronger structural similarity to GT




↑ PSNR ⟶ Higher pixel-level accuracy




↓ LPIPS ⟶ Lower perceptual difference



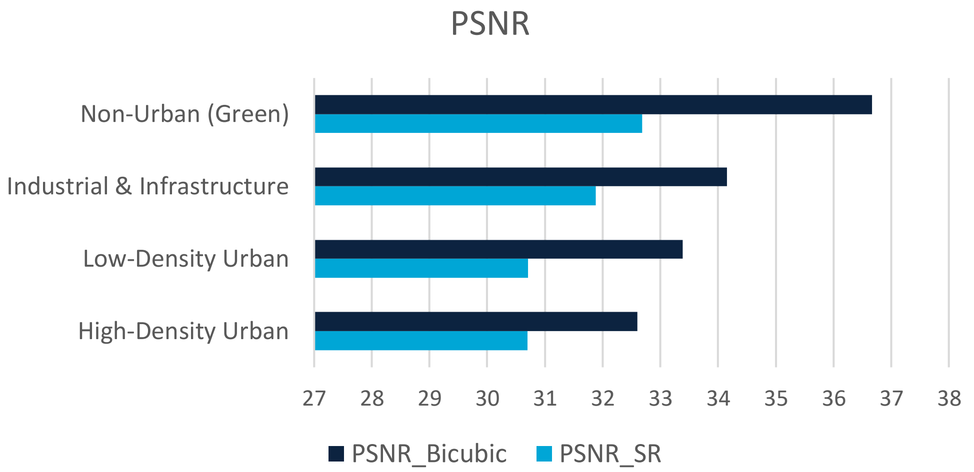

↑ SSIM ⟶ Stronger structural similarity

# Results

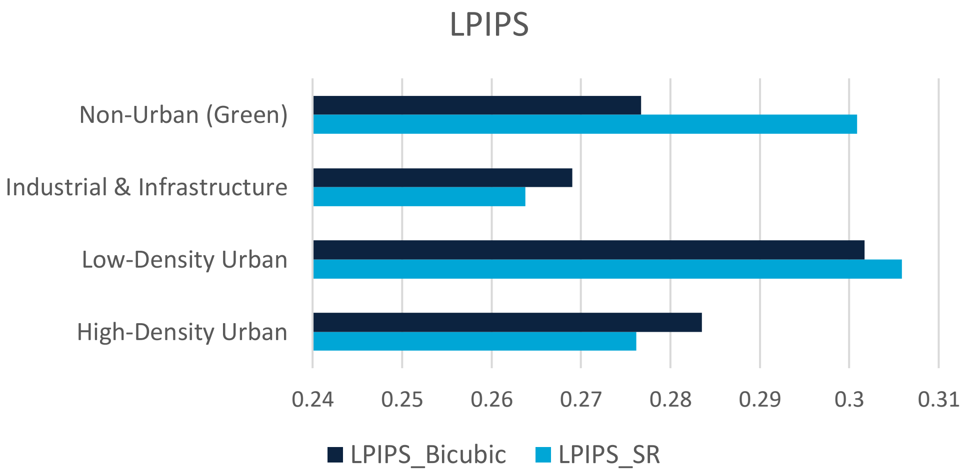
## Iteration 1

### Key Findings

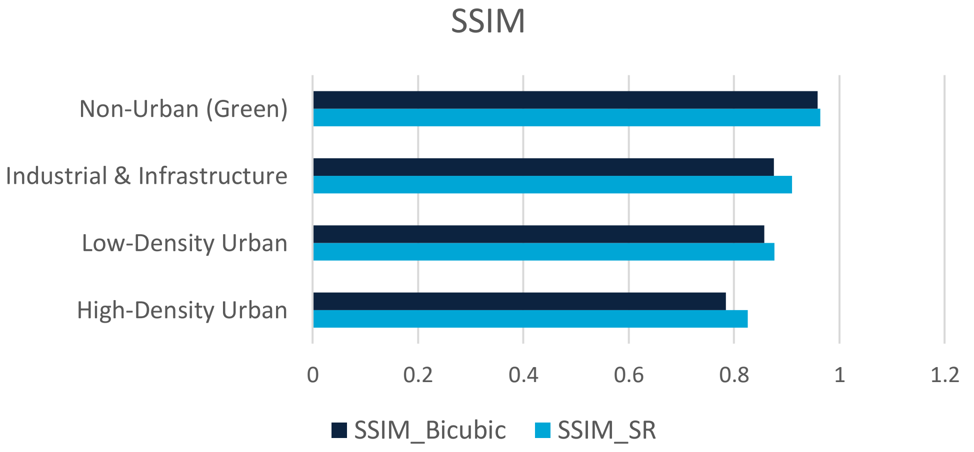
- Bicubic outperforms SRGAN in PSNR due to deterministic nature
- SRGAN achieves better SSIM and LPIPS indicating better structural and perceptual reconstruction of textures and edges.




↑ PSNR ⟶ Higher pixel-level accuracy




↓ LPIPS ⟶ Lower perceptual difference




↑ SSIM ⟶ Stronger structural similarity

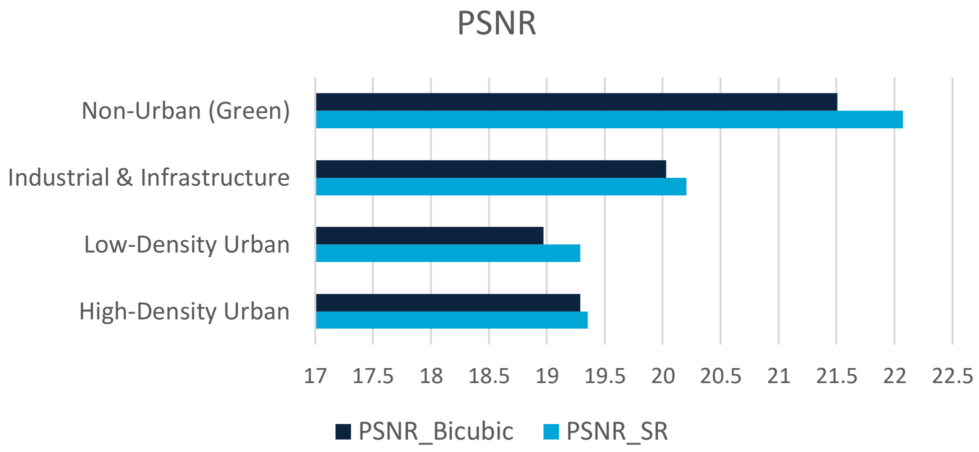
TU Delft

# Results

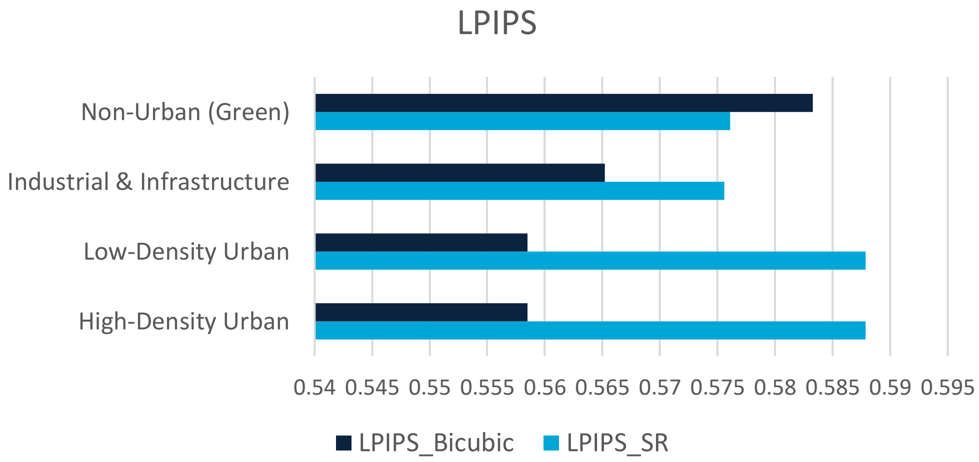
## Iteration 2

### Key Findings

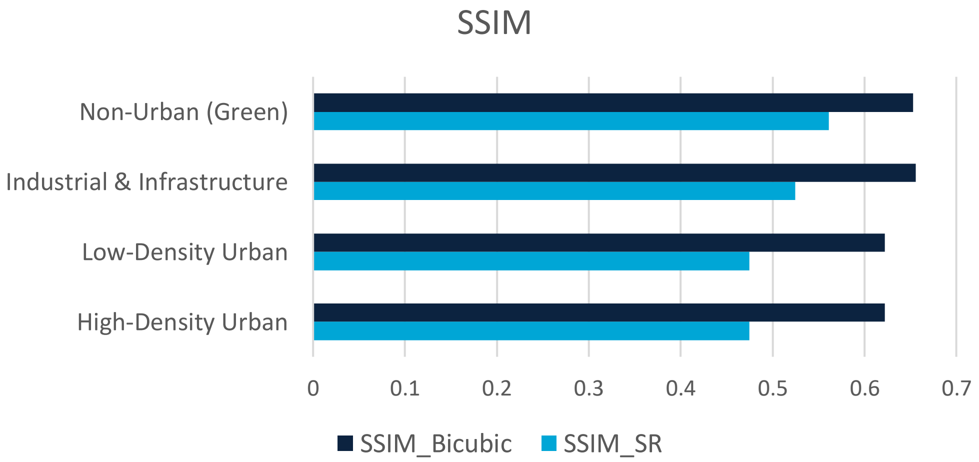
- SRGAN outperforms Bicubic in PSNR showing improved pixel level reconstruction
- Close LPIPS scores
- Noise in SRGAN is penalized in SSIM




↑ PSNR ⟶ Higher pixel-level accuracy




↓ LPIPS ⟶ Lower perceptual difference



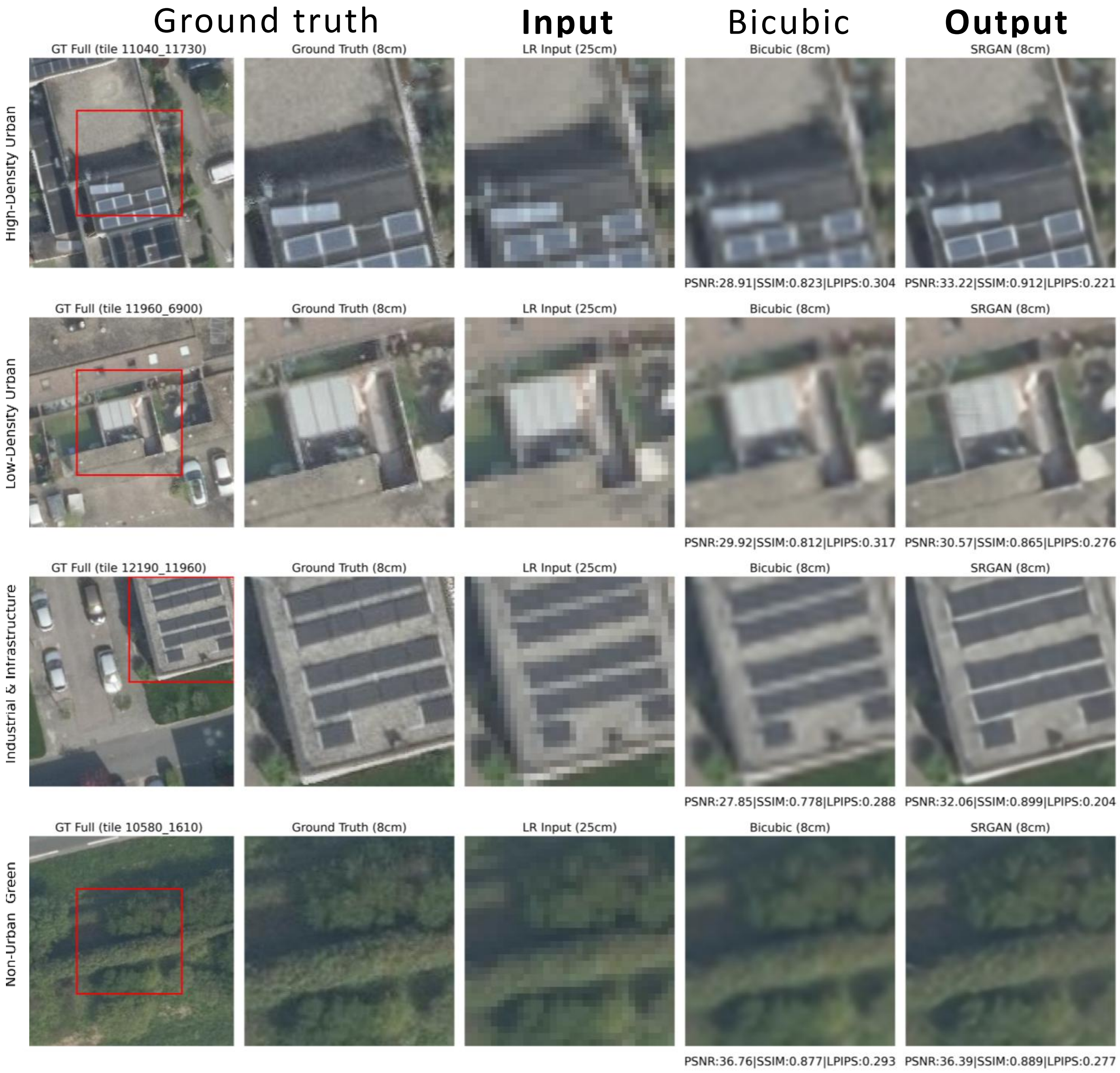

↑ SSIM ⟶ Stronger structural similarity

# Results:

Iteration 1 – Synthetic input

Key Findings

- The model produces sharper and more detailed visual outputs
- Shows improved perceptual fidelity over Bicubic
- Preserves roof textures and natural patterns
- Successfully reconstructs rooftop elements



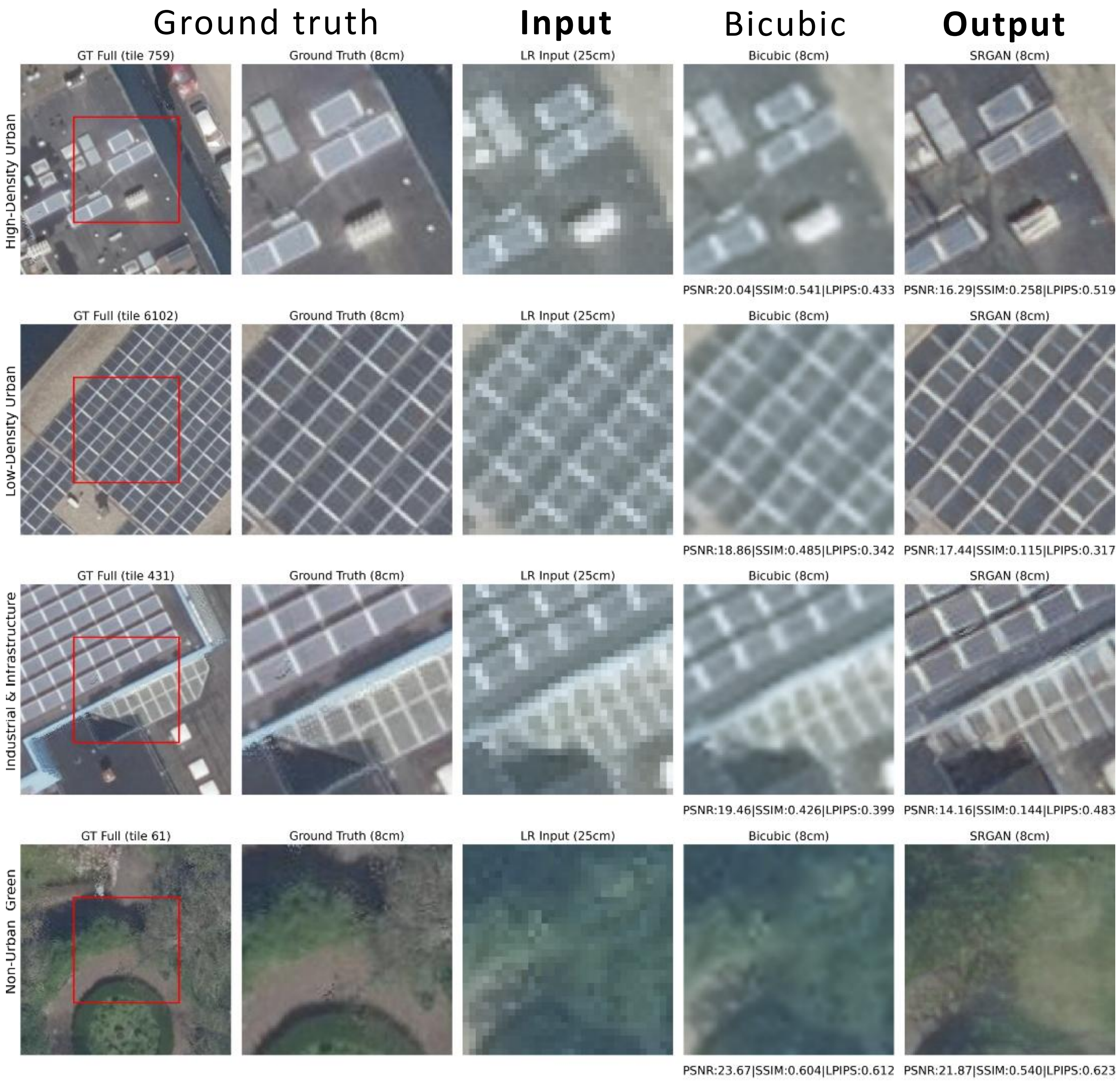

TU Delft

# Results:

Iteration 2 – Real World Input

Key Findings

- Model reconstructs clear building shapes and structural details
- Fine textures like solar panels, roof materials, and shadows are recovered, even when seasons differ
- Model robust to environmental and temporal variation
- Realistic outputs

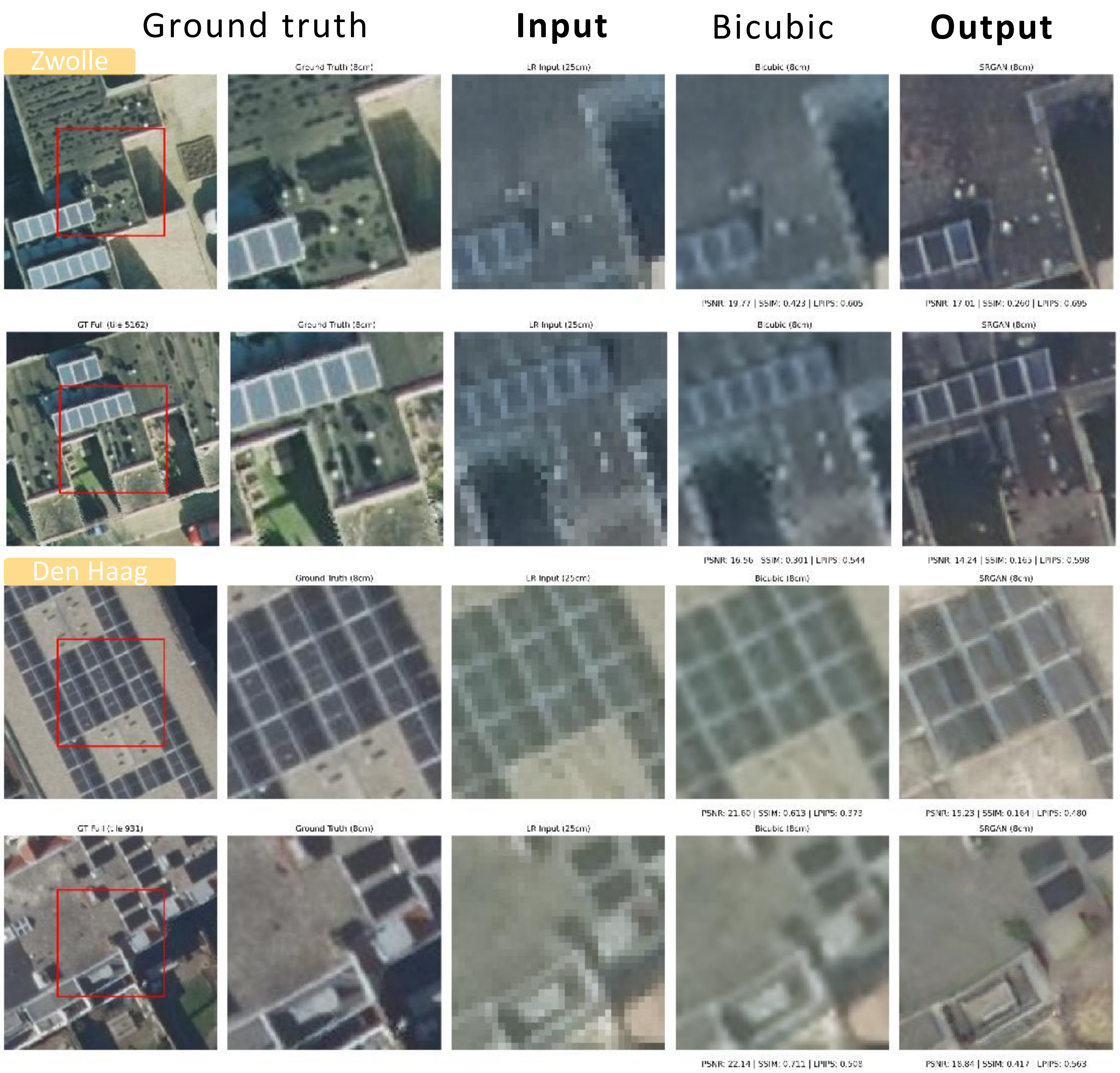


# Results:

Adaptability to new geographical areas

Key Findings

- Good adaptability to unseen cities
- Recovered building outlines, rooftop layers
- Better structure preservation that bicubic
- Strong generalization in urban and industrial areas
- Model struggles with entirely unseen patterns
- Highly dependent on input quality




TU Delft

# Results:

Downstream Tasks

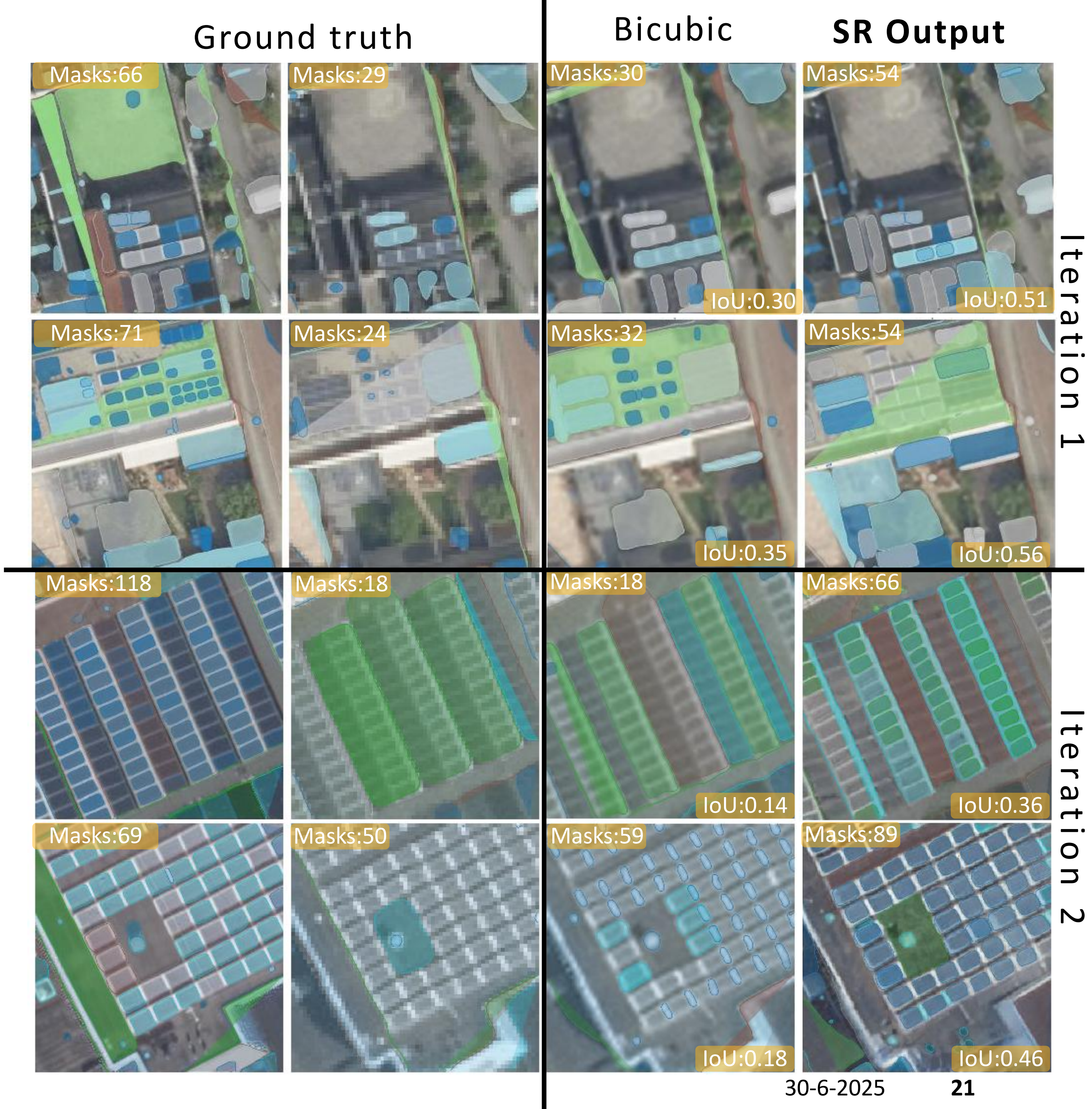
SAM

- Enables more accurate and complete segmentation masks
- Clearly separates fine structures that Bicubic often merges or misses
- Maintains high IoU scores for building footprints across both iterations, even under domain shift (e.g., Iteration 2)

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union}$$

Note: Mask count is not a formal metric—it illustrates how well objects are reconstructed and distinguishable by models like SAM

TUDelft

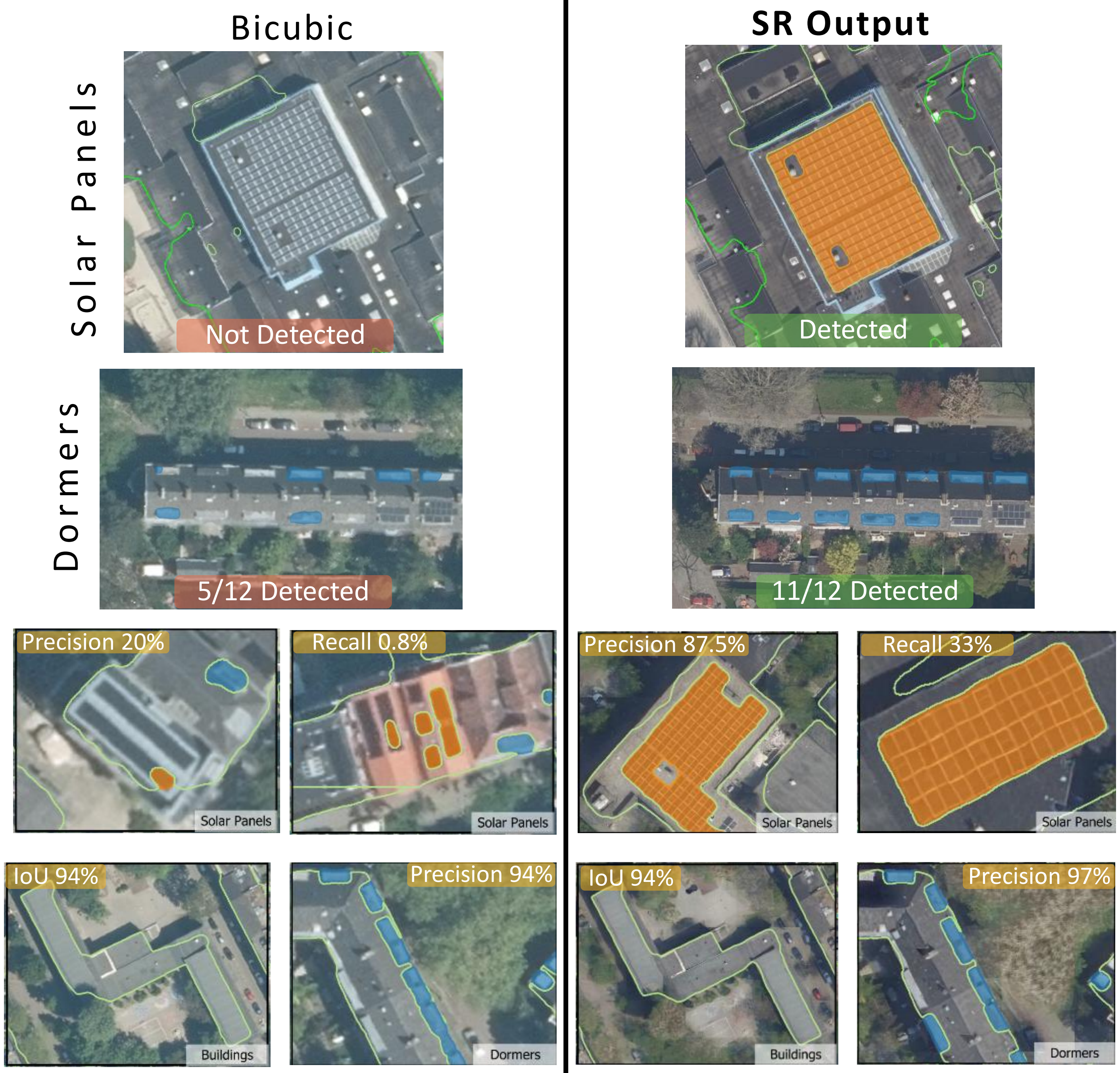


# Results:

## Downstream Tasks

### Semantic Segmentation (Readar B.V.)

- Higher object detection performance, especially for small urban features
- Produces more complete and precise segmentation, whereas Bicubic often fails entirely for key classes
- Detects more PV panels and dormers, with significant improvements in precision and recall
- Manual validation confirms SRGAN avoids false detections and captures more true positives




TU Delft

05

# Conclusions

# Research Questions

## Main Question

To what extend can GAN-based super-resolution enhance 25 cm aerial imagery to 8 cm,

ensuring its applicability for object detection tasks?

- GAN-based SR significantly improves **visual quality, clarity, and structural fidelity**
- Successfully **reconstructs small-scale features** essential for segmentation
- Outperforms Bicubic in **both visual and functional accuracy**

## Sub-questions

1. How accurately can a GAN reconstruct 8 cm HR images from 25 cm LR aerial inputs, especially for building edges and solar panels?
    - The model reconstructs rooftop contours, building edges, and textures with **high fidelity**
    - Outperforms interpolation based methods by **preserving geometry** and **avoiding blurring**

# Research Questions

## Sub-questions

2. How do seasonal differences between HR and LR images (e.g., winter HR vs. Summer LR) affect GAN performance, and can domain adaptation via pre-training on synthetic data mitigate these effects?
    - Seasonal differences introduce **color, shading, and vegetation shifts**
    - **Two phase** training (synthetic + real) mitigates this
    - **Adaptation** is effective but depends on training data diversity
3. What are the limitations of GANs in preserving geometric fidelity (e.g., artifacts, hallucination) for geospatial use cases?
    - **Hallucinations** occur in irregular or underrepresented regions
    - Ghosting appears in less-structured zones due to **limited training data**
4. What metrics best assess SR image quality for downstream object detection tasks?
    - Image quality metrics offer **partial insights**, task aware metrics provide most meaningful evaluation

# Contributions & Limitations

## Scientific Contributions

- Developed a GAN-based SR pipeline for aerial imagery
- Applied a two-phase training strategy for domain adaptation
- Enabled downstream segmentation without retraining

## Limitations

- Limited generalization to unseen cities
- Results likely to improve with more diverse data and full hyperparameter tuning

TU Delft

# Future Work

## Architecture & Training

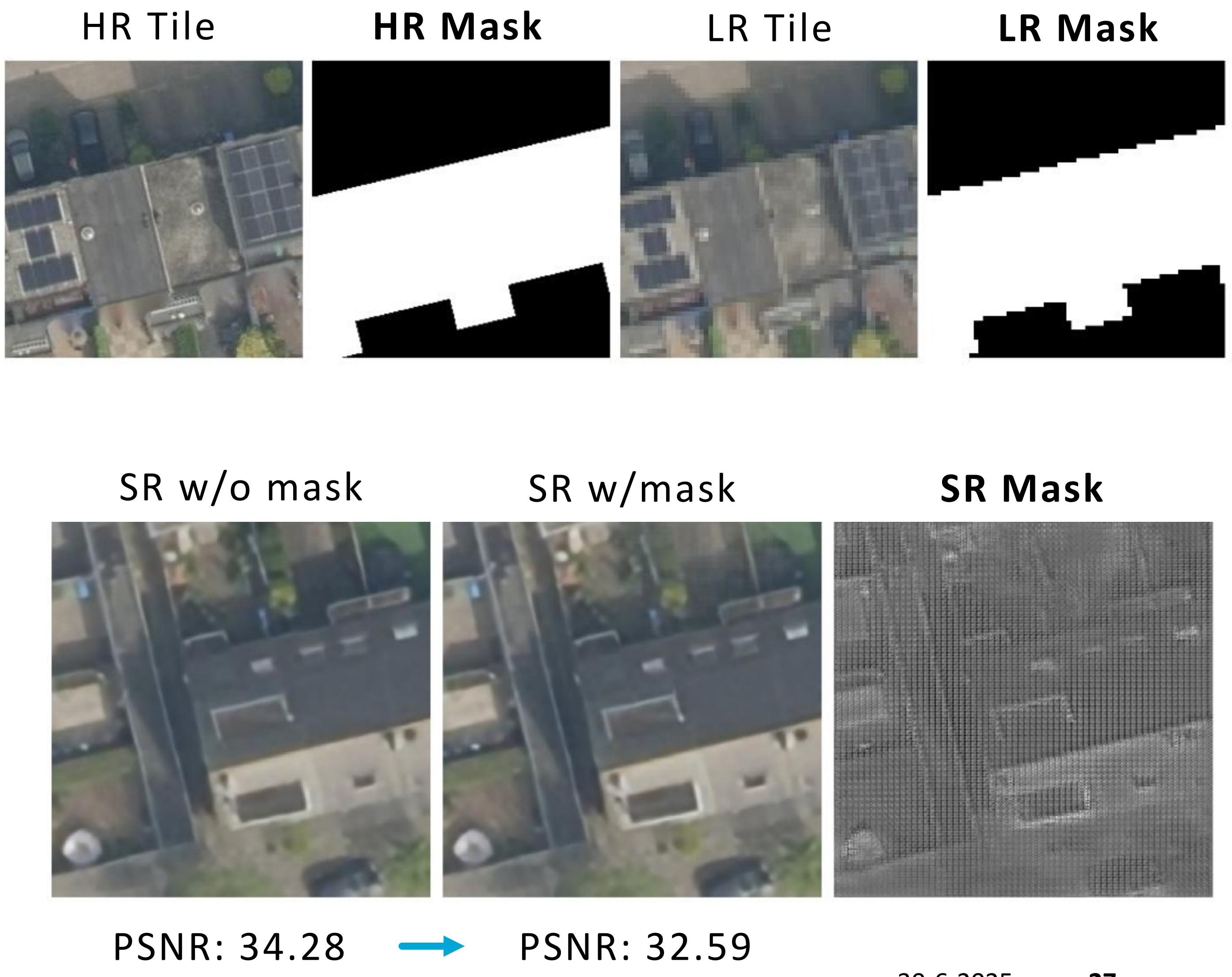
- Introduce edge-aware loss functions
- Use smarter conditioning with semantic masks
- Incorporate data from multiple cities to improve generalization

## Evaluation

- Exclude vegetation areas in metric computation

## Integration

- Develop end-to-end pipelines linking SR to object detection tasks




TU Delft

# Thank You
# For Your Attention!

P5 Presentation | Michalis Michalas

30-6-2025