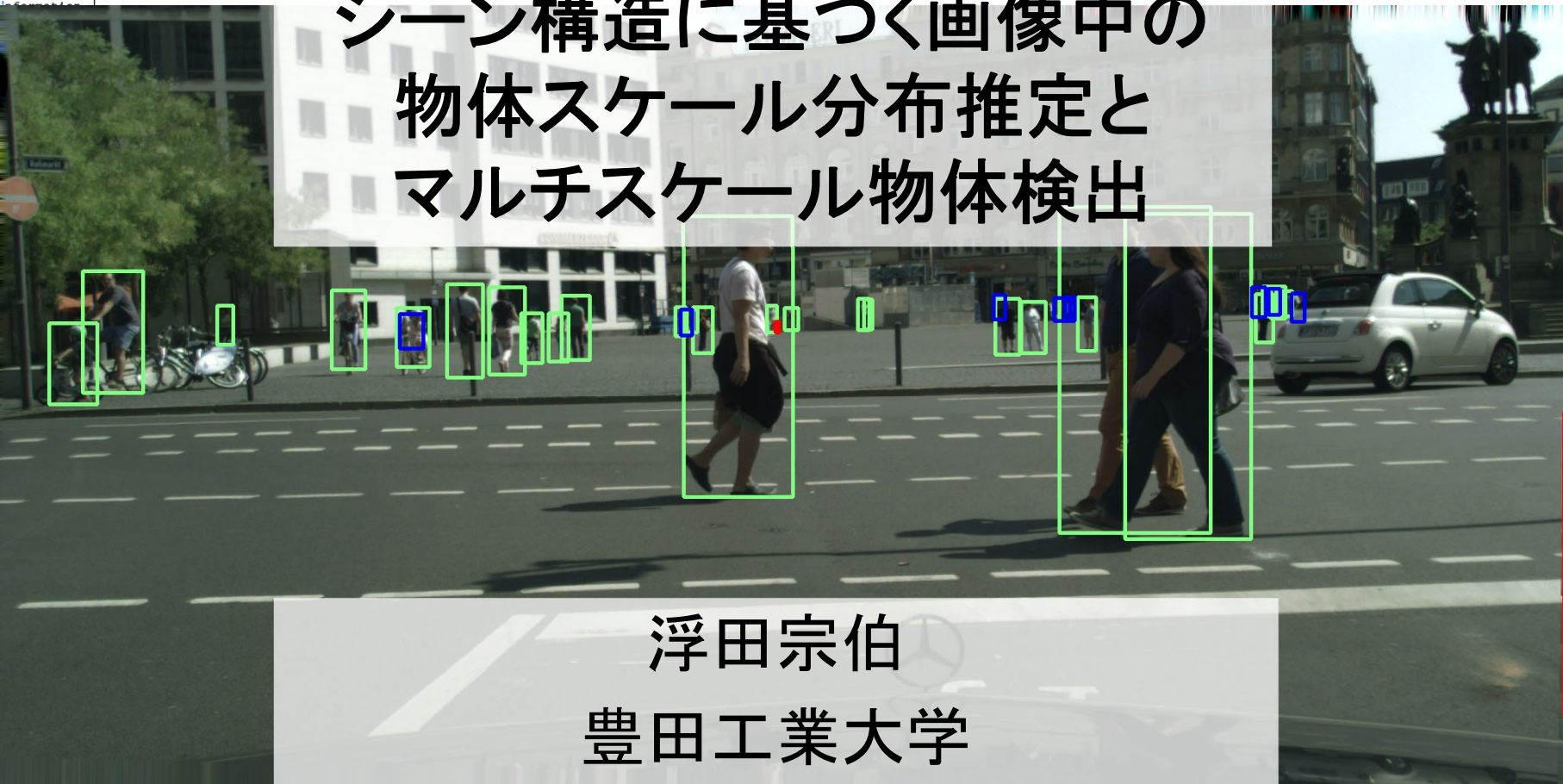


# シーン構造に基づく画像中の 物体スケール分布推定と マルチスケール物体検出



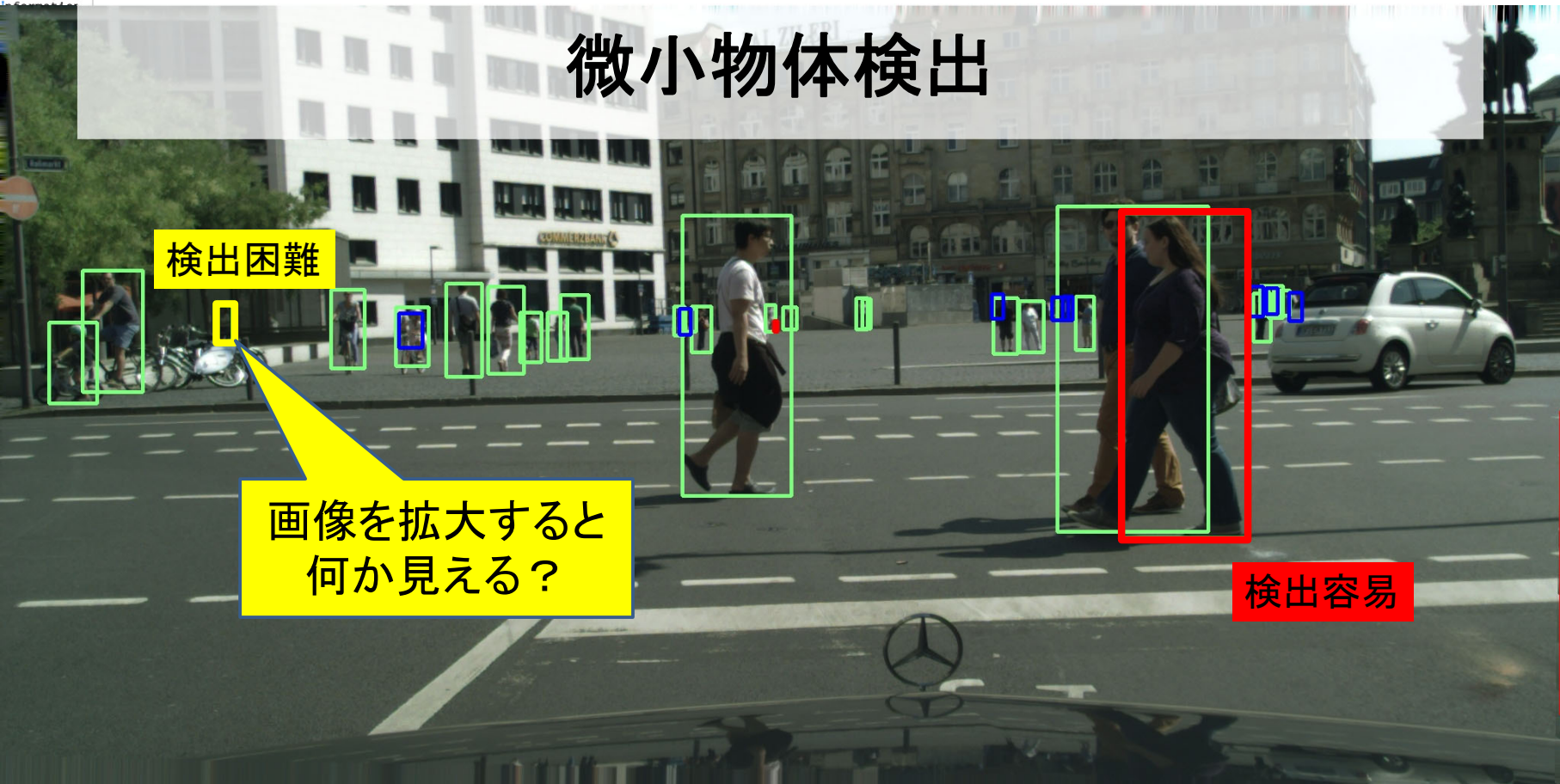
浮田宗伯  
豊田工業大学

# 微小物体検出

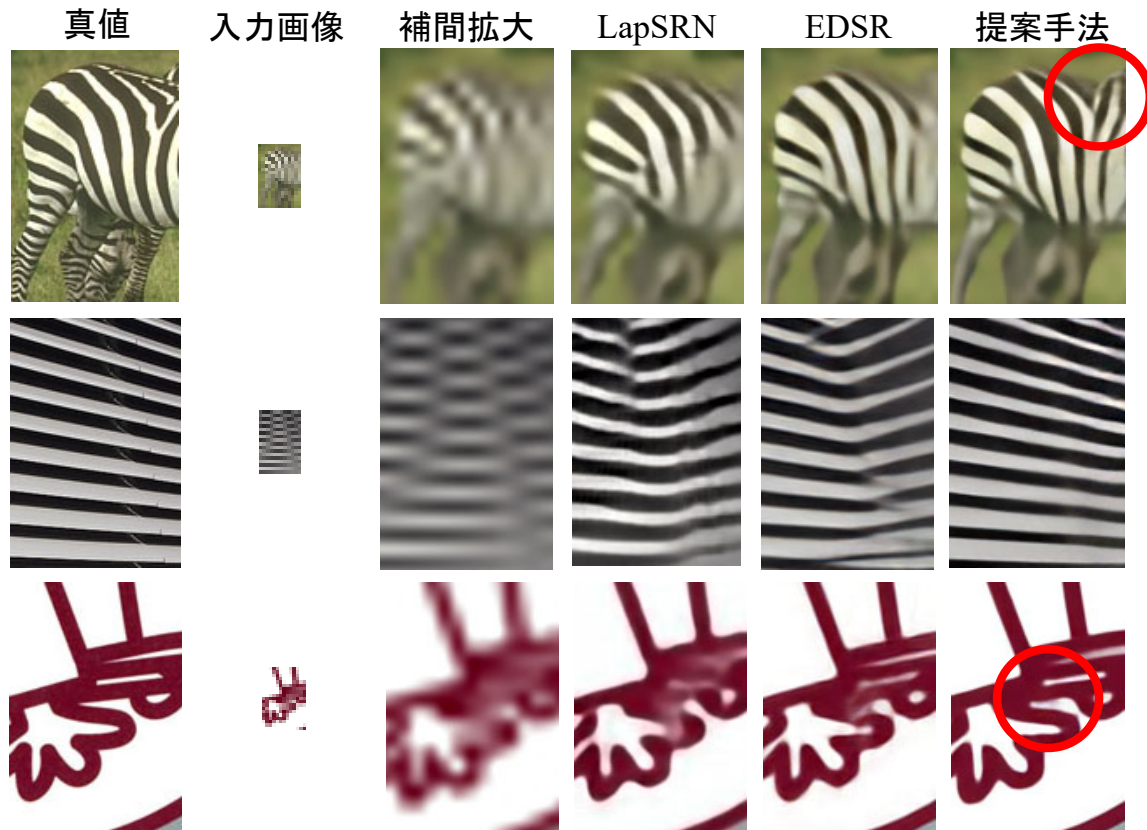
検出困難

画像を拡大すると  
何か見える？

検出容易



# 画像超解像



# 動画の超解像

## ～ 単一画像超解像と複数画像超解像の統合～



補間拡大

TOFlow

VSR-DUF

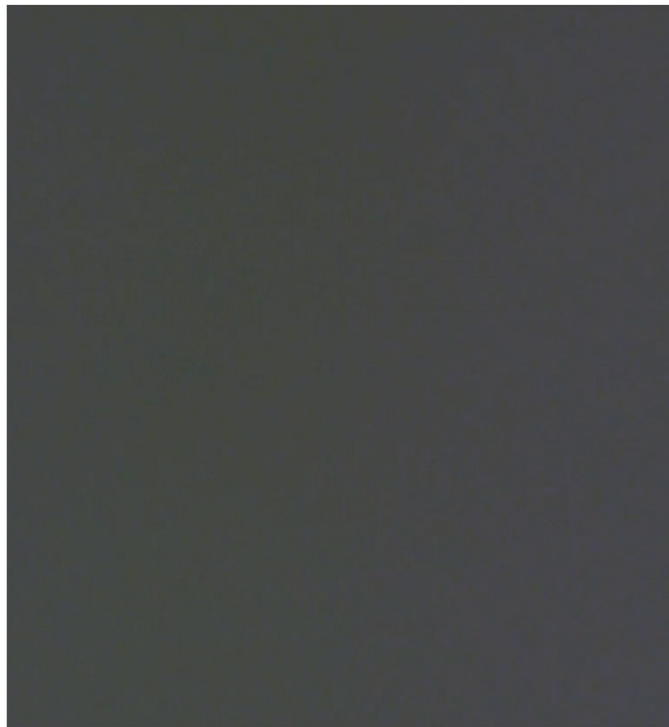
提案手法  
(3近傍フレーム)

提案手法  
(6近傍フレーム)

真値

# 時空間の超解像

## ～空間超解像と時間超解像の相互補助～



補間計算結果



提案手法結果

# 超解像と物体検出の統合学習



高解像度画像  
(元画像)



低解像度画像



補間拡大画像



超解像と物体検出  
の独立学習



提案手法

手法	縦横4倍	縦横8倍
高解像画像(元画像)	75.8	75.8
低解像画像	41.7	16.6
補間拡大画像	41.3	11.2
独立学習	41.9	10.6
統合学習(提案手法)	62.2	37.5

物体検出成功率

20画素

ドローンの安全飛行のための遠方鳥検出



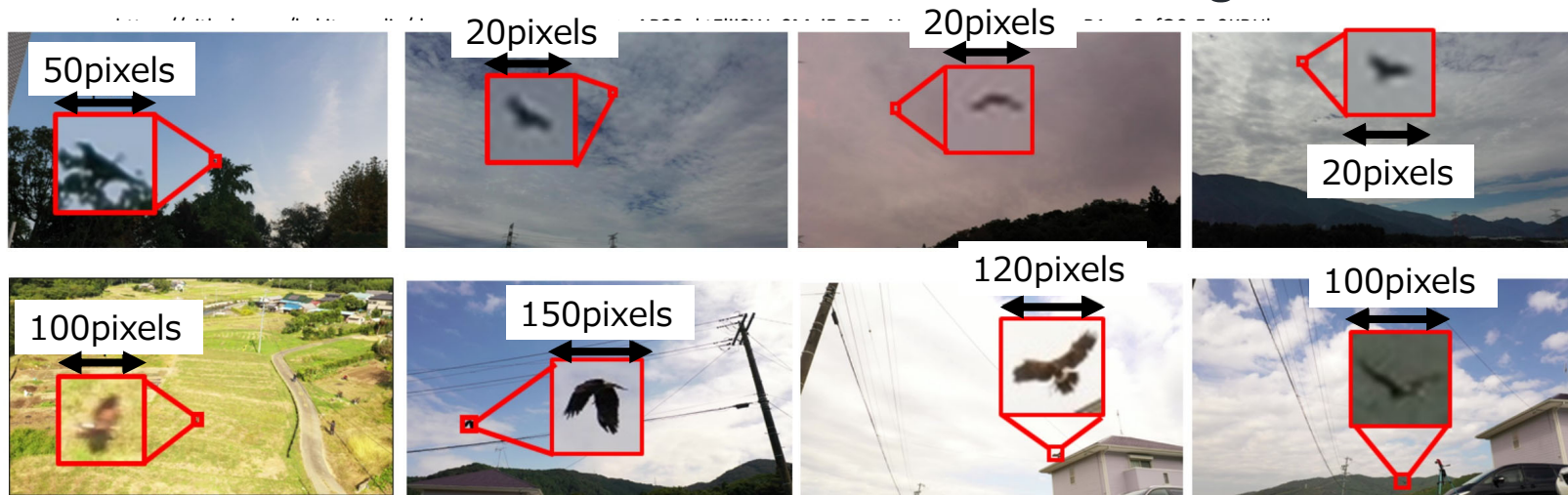
浮田宗伯

豊田工業大学

知能情報メディア研究室

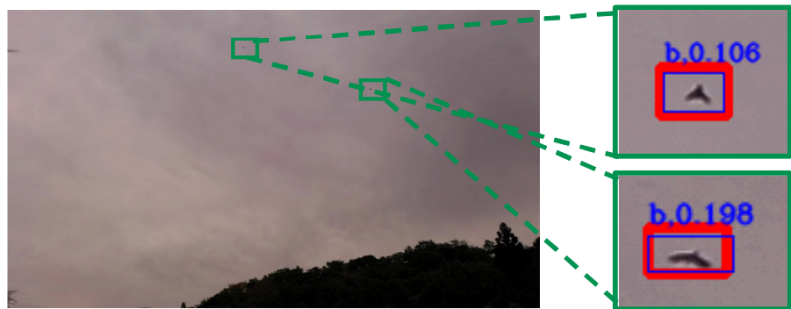
# 提案データセット

- 種々の鳥を撮影
  - 48,395 images with 62,106 instances
  - 10~150pix size of birds in 4K images
- Distant Bird Detection Dataset for Safe Drone Flight

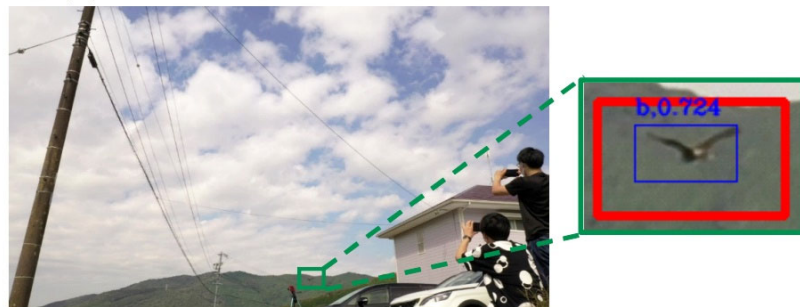




# 実験結果：定性評価



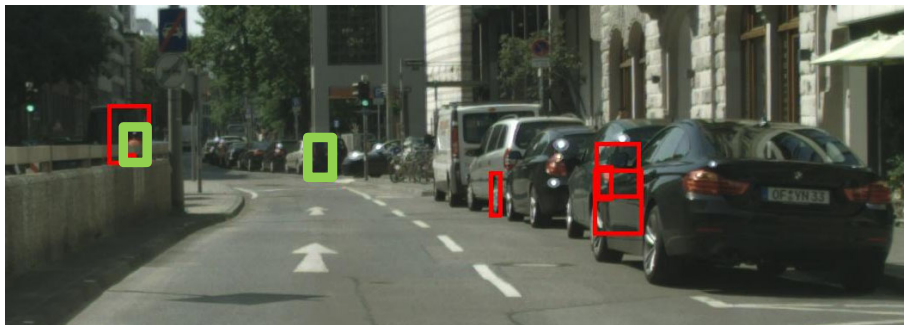
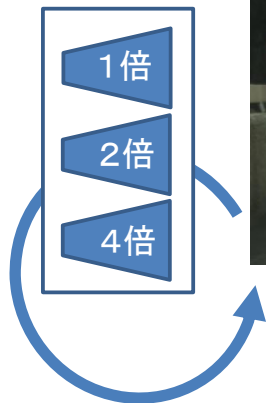
(1) True-positives in the sky background



(2) True-positive in the forest background

# シーン構造に基づく画像中の 物体スケール分布推定と マルチスケール物体検出

# 複数スケールの物体検出(微小物体検出含む)を そのまま一般シーンに適用すると...



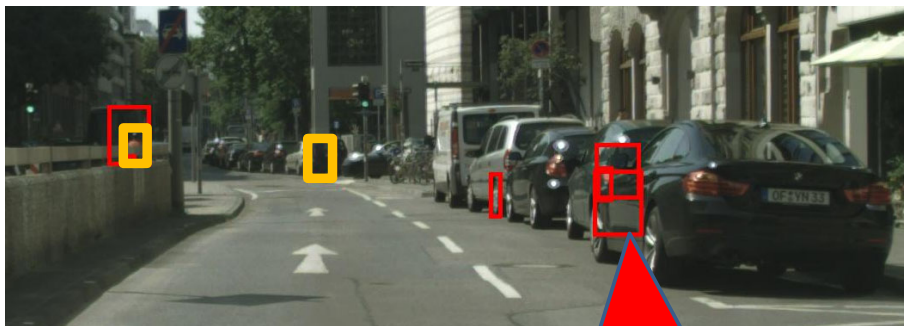
人検出の結果:

 正検出

 誤検出



# 複数スケールの物体検出(微小物体検出含む)を そのまま一般シーンに適用すると...



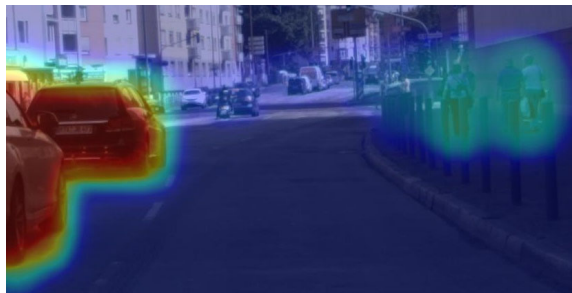
-  正検出
-  誤検出

「車のスケール」と  
「検出された人のスケール」が  
不整合

# 研究目標： シーン構造から予測される 各スケールにおける**物体存在確率マップ**



## Region-Dependent Scale Proposals (RDSP)



超解像1倍スケール



超解像2倍スケール

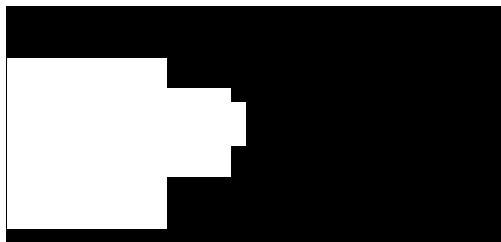


超解像4倍スケール

# RDSP用の学習データセットの生成法

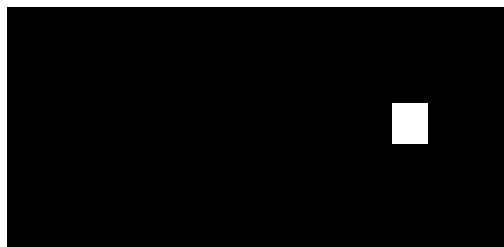


(a) Image



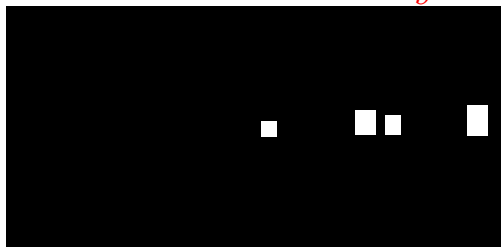
(b) Factor 1

$$64 \leq H_b$$



(c) Factor 2

$$32 \leq H_b < 64$$



(d) Factor 4

$$H_b < 32$$

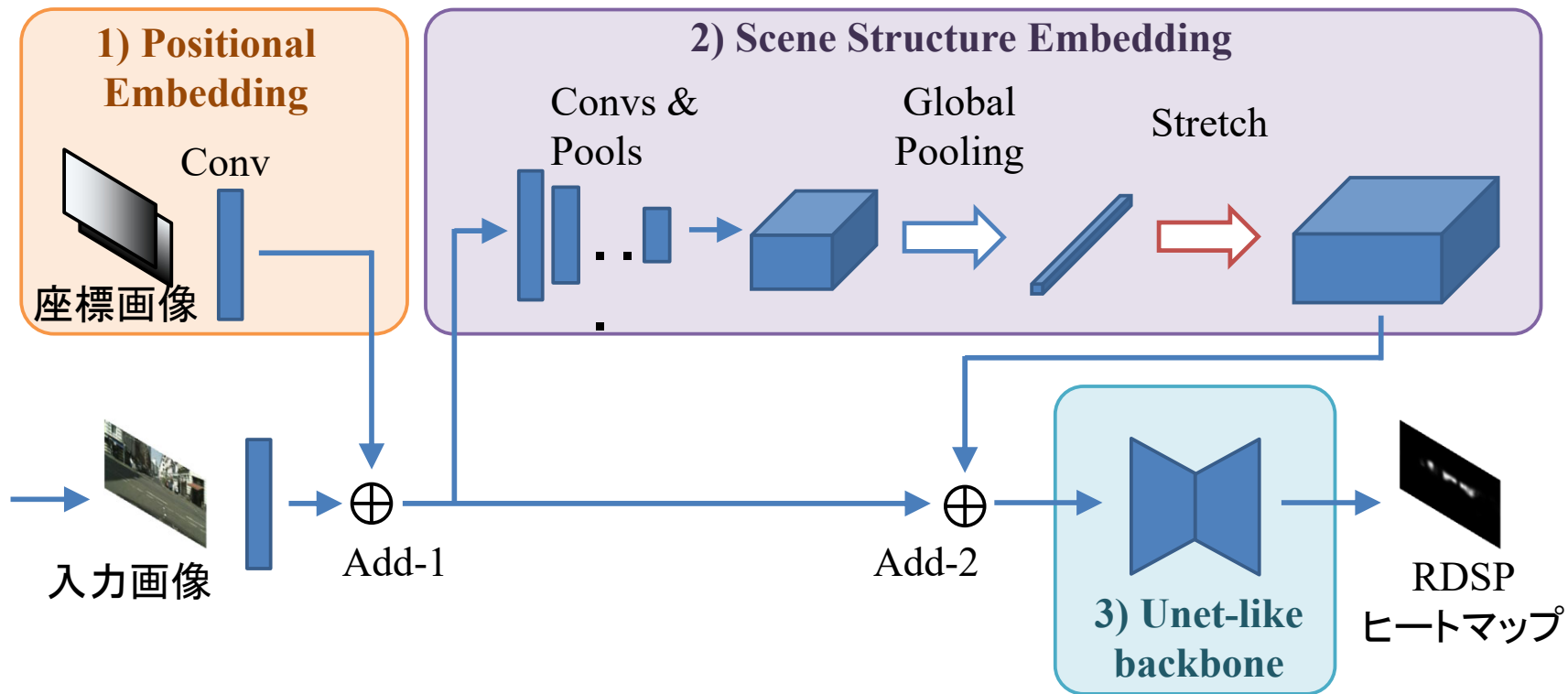
アノテーションされている物体中、  
下記物体のバウンディングボックス  
をRDSP学習に利用:

Person, Rider, Car Truck, Bus,  
On rails, Motorcycle, Bicycle,  
Caravan, Trailer

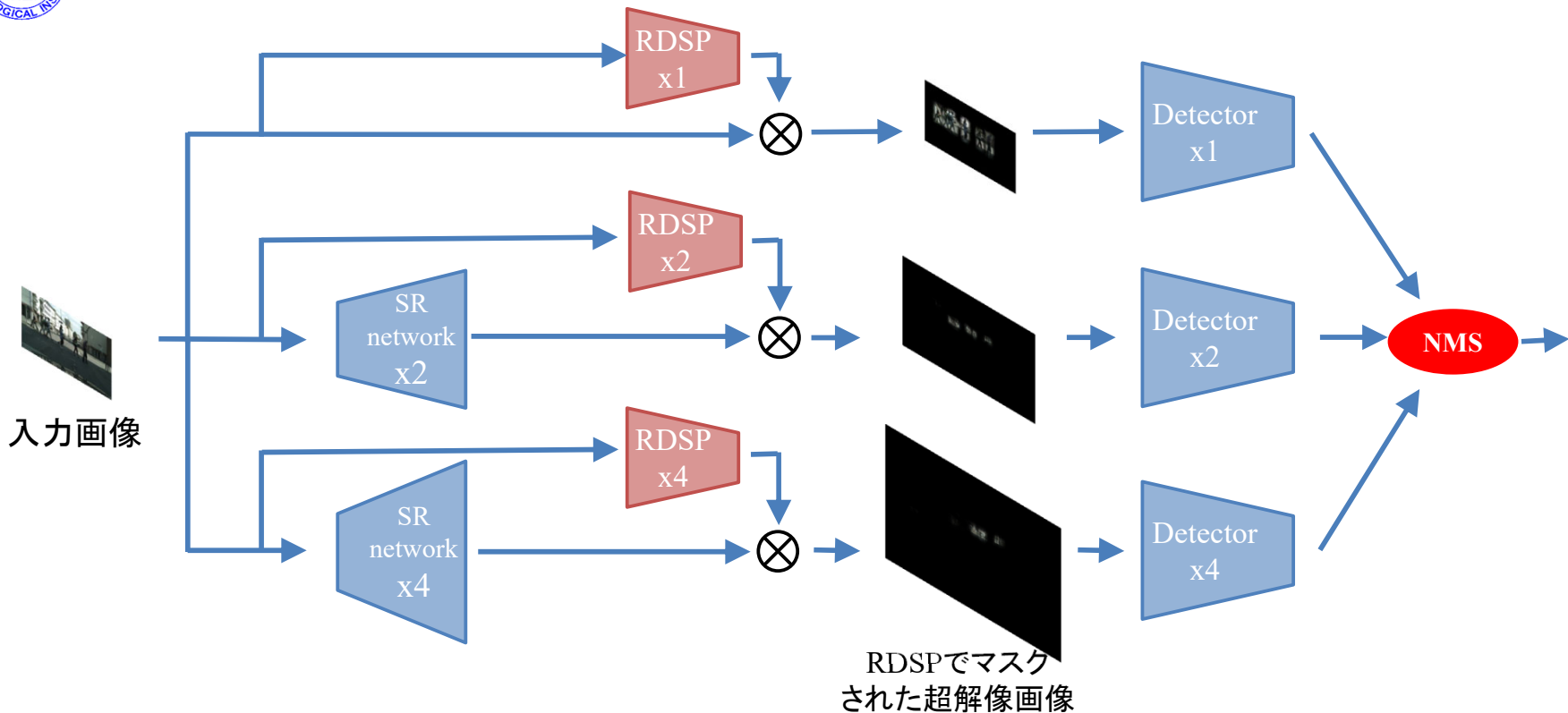
Cityscapesデータセットの例 (2048 x 1024 画素の画像に対するスケール設定):

$H_b$  = 物体のバウンディングボックスの縦サイズ

# RDSPネットワークの詳細



# RDSPを含んだ提案物体検出器

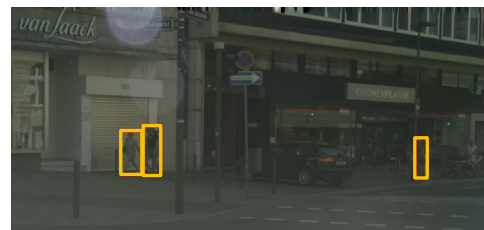
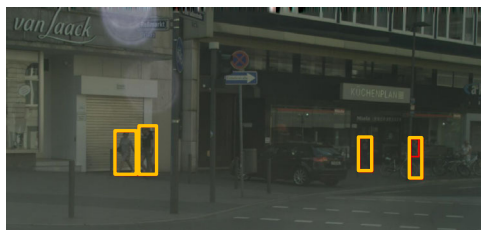
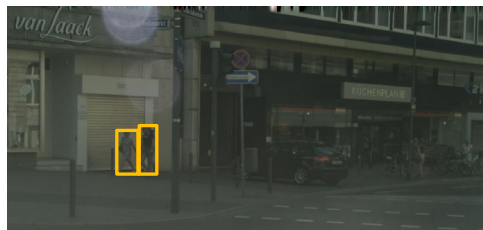
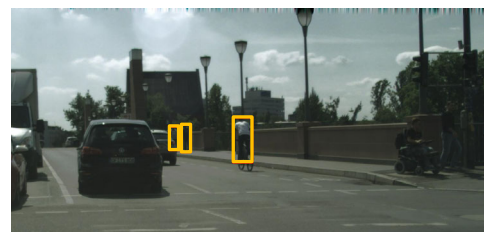
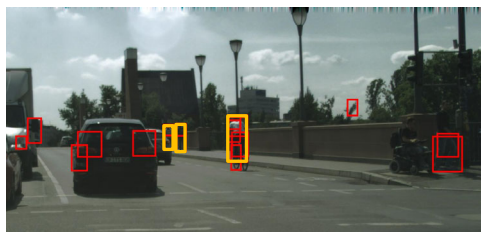
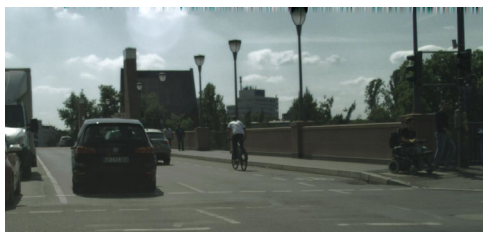
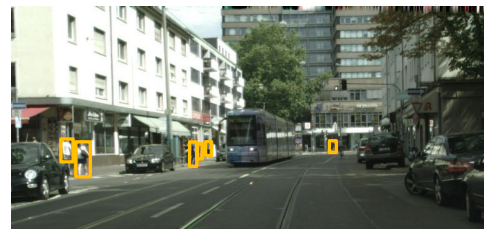




# 実験結果

□ 正検出 □ 誤検出

物体検出器に  
SSDを利用



(a) 従来手法  
(超解像無し)

(b) 従来手法  
(超解像あり, RDSP無し)

(c) 提案手法  
(超解像あり, RDSPあり)

# 定量評価

## ～多様な物体検出器に対する汎用性～

Model		AP	AP <sub>50</sub>	AP <sub>75</sub>	AP <sub>S</sub>	AP <sub>M</sub>	AP <sub>L</sub>
One-stage detector	SSD [2]	35.9	63.9	35.9	15.9	44.5	<b>58.3</b>
	SSD + Ours	<b>37.8</b>	<b>66.2</b>	<b>38.1</b>	<b>17.6</b>	<b>47.1</b>	57.8
	RetinaNet [3]	37.6	66.6	38.1	20.0	45.3	<b>58.3</b>
	RetinaNet + Ours	<b>39.1</b>	<b>68.5</b>	<b>39.8</b>	<b>21.7</b>	<b>46.6</b>	57.5
Two-stage detector	Faster R-CNN [4]	36.0	61.1	37.5	17.8	43.6	<b>58.2</b>
	Faster R-CNN + Ours	<b>38.2</b>	<b>65.8</b>	<b>40.2</b>	<b>21.0</b>	<b>45.6</b>	57.3
Anchor-free detector	FCOS [51]	36.4	65.5	36.5	18.8	44.5	<b>54.1</b>
	FCOS + Ours	<b>37.3</b>	<b>67.0</b>	<b>37.4</b>	<b>21.1</b>	<b>44.8</b>	53.3
	CenterNet [49]	38.8	66.1	39.2	18.4	47.2	<b>62.6</b>
	CenterNet + Ours	<b>40.0</b>	<b>67.6</b>	<b>41.1</b>	<b>20.2</b>	<b>48.2</b>	62.2

# まとめ

- 概要

- 画像超解像を利用した微小物体検出
- スケールに応じた物体存在領域を推定し、スケールの異なる物体検出の精度向上

- 今後の研究

- 基礎：超解像を利用した物体検出精度の向上
- RDSP：撮影画像をそのまま利用するのではなく、セマンティックセグメンテーション、モノデプス推定などの結果も利用