Centre for Sensors, Instruments and Systems Development
UNIVERSITAT POLITECNICA DE CATALUNYA
Shaping light to your needs

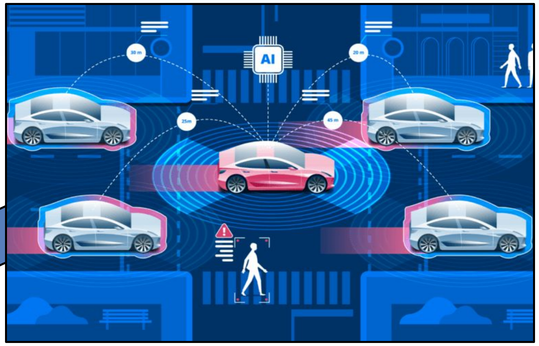
Pedestrian Detection in 3D Point Clouds using Deep Neural Networks

Òscar Lorente Corominas

Advisors: Josep R. Casas, Santiago Royo
CD6, UPC - ETSETB: Escola Tècnica Superior d'Enginyeria de Telecomunicació de Barcelona
Degree Thesis · 2020

1. Introduction (I)

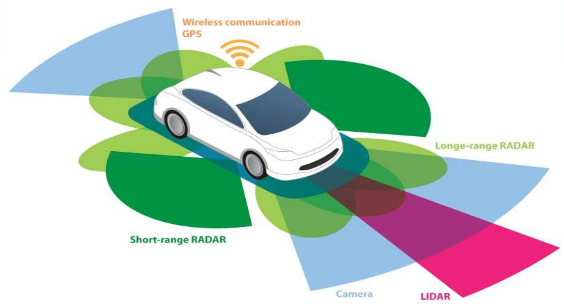
1. Introduction (I)



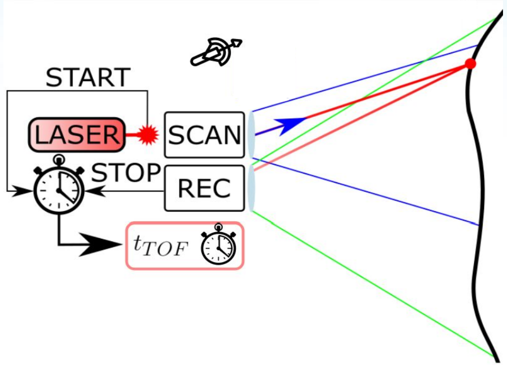
WE NEED MORE INFORMATION

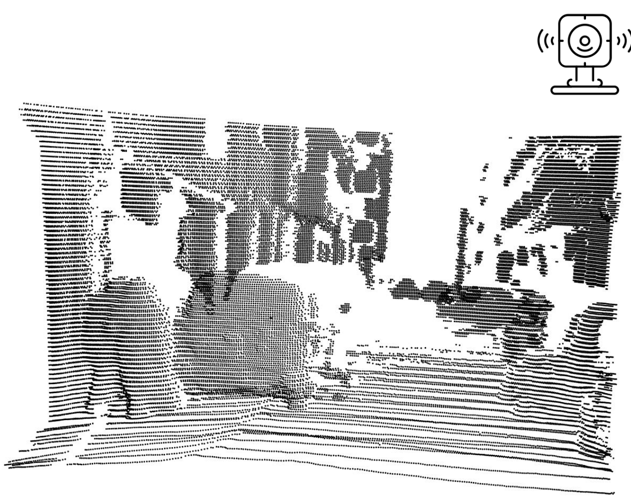
1. Introduction (II)

Combine sensors



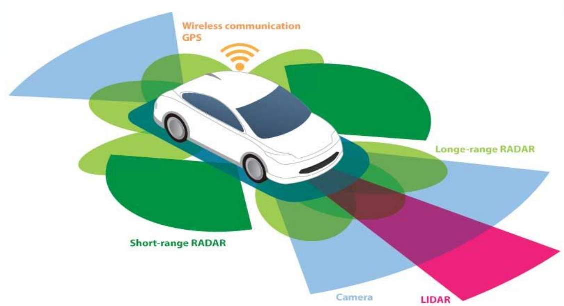
Time Of Flight: **LIDAR**



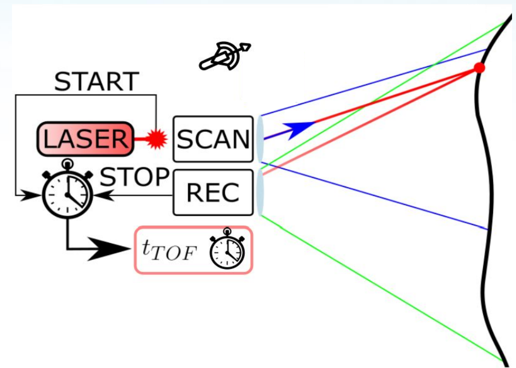


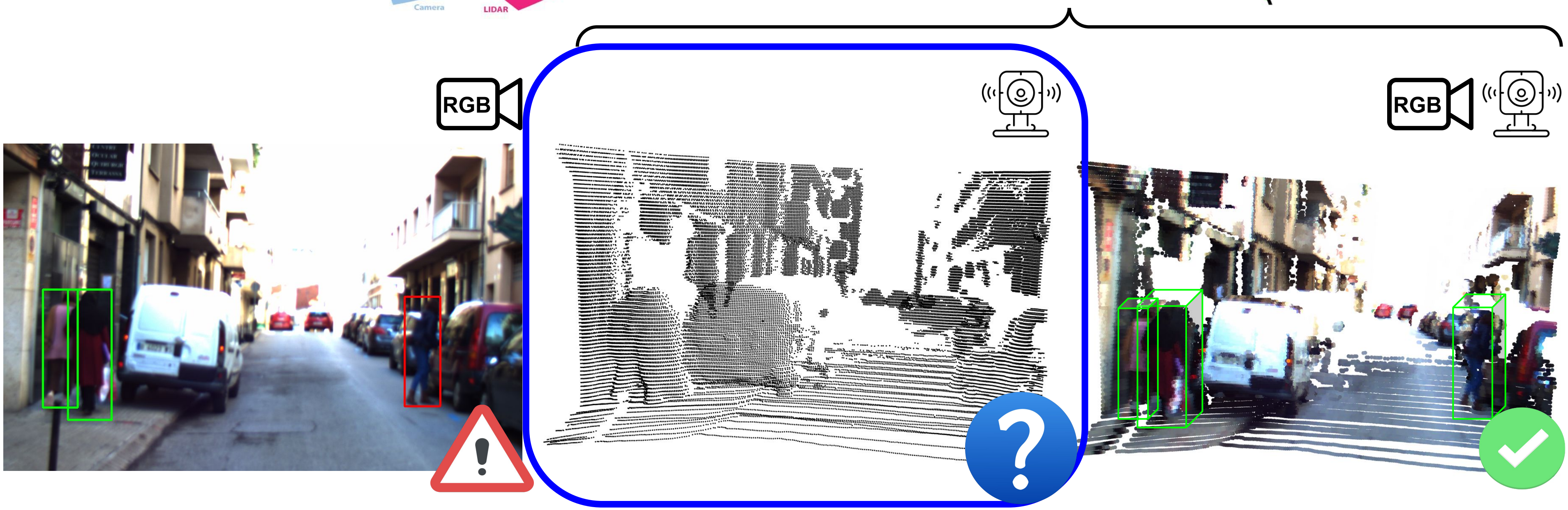
1. Introduction (II)

Combine sensors



Time Of Flight: **LIDAR**

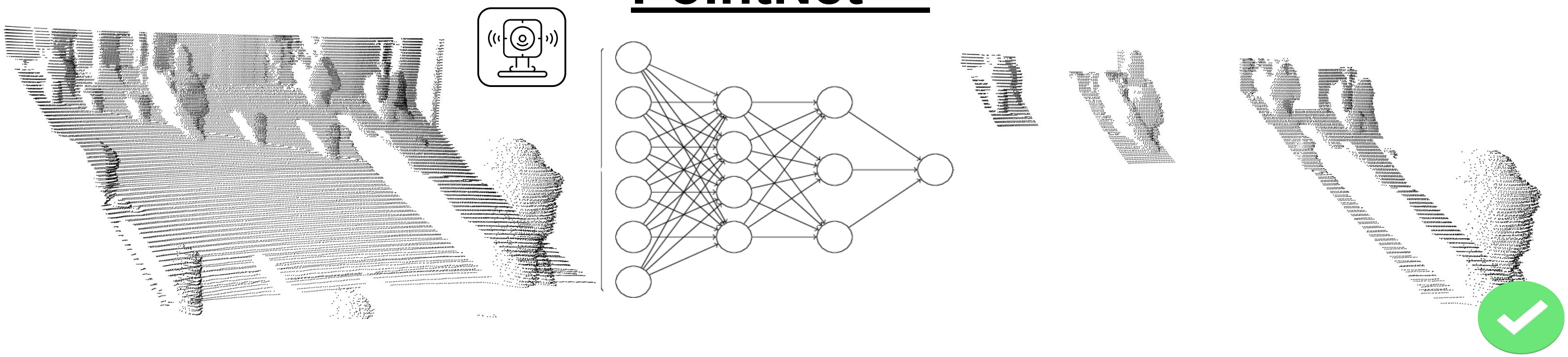




2. Objectives (I)

- Pedestrian detection system in point clouds using Deep Neural Networks

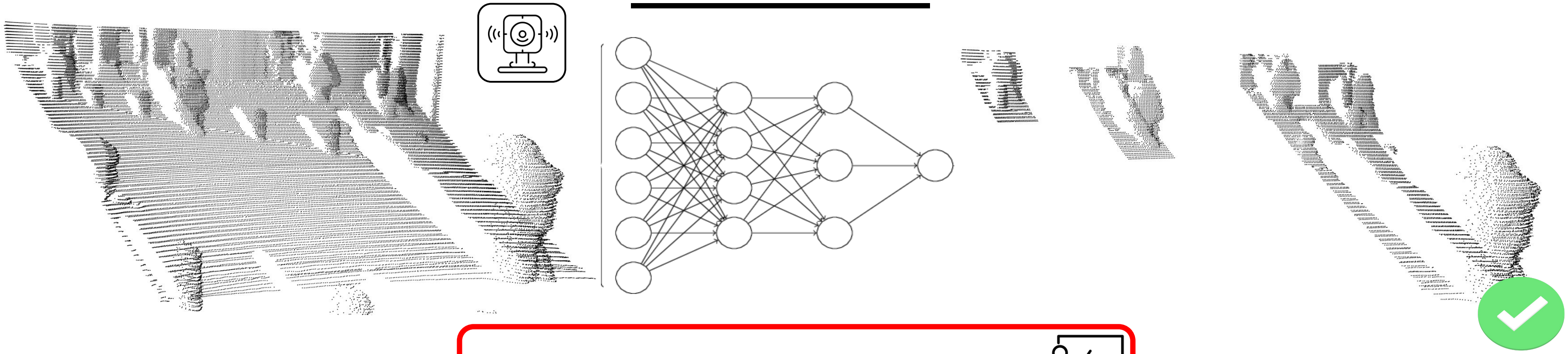
PointNet++



2. Objectives (I)

Pedestrian detection system in point clouds using Deep Neural Networks

PointNet++



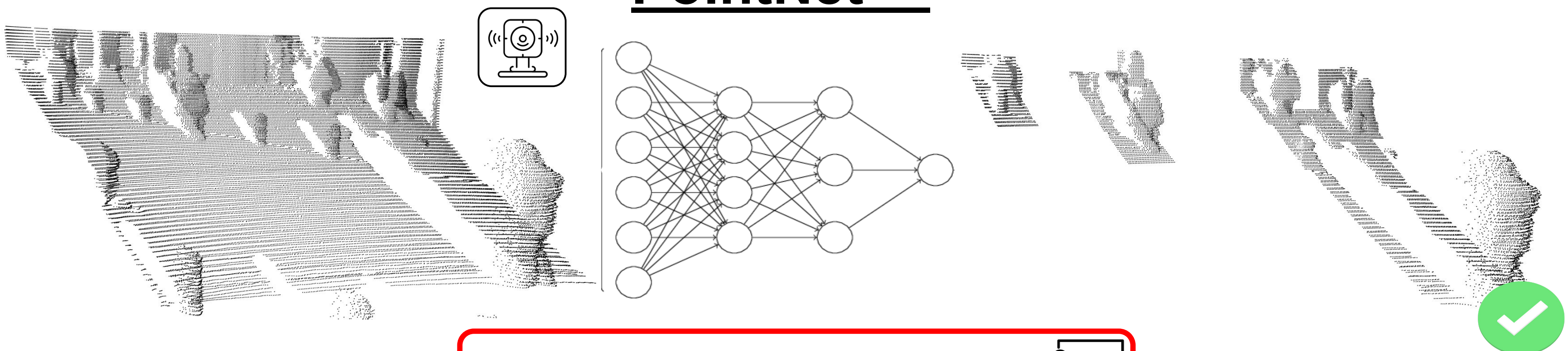
How do we train the network?

> **Labeled pedestrian and non-pedestrian 3D point clouds** ← **Manual labeling**

2. Objectives (I)

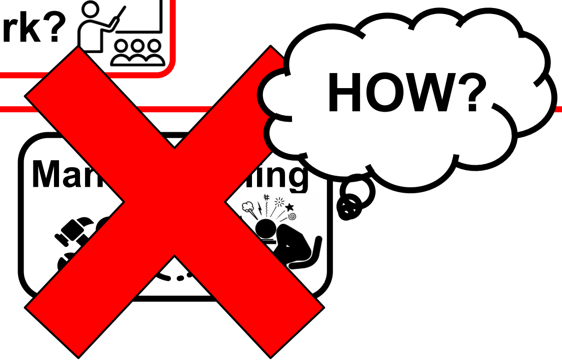
Pedestrian detection system in point clouds using Deep Neural Networks

PointNet++



How do we train the network?

Labeled pedestrian and non-pedestrian 3D point clouds



2. Objectives (II)

Pedestrian detection system in point clouds using PointNet++

System to generate a dataset with ground truth in point clouds

Labeled pedestrian and non-pedestrian 3D point clouds

2. Objectives (II)

Pedestrian detection system in point clouds using PointNet++

System to generate a dataset with ground truth in point clouds

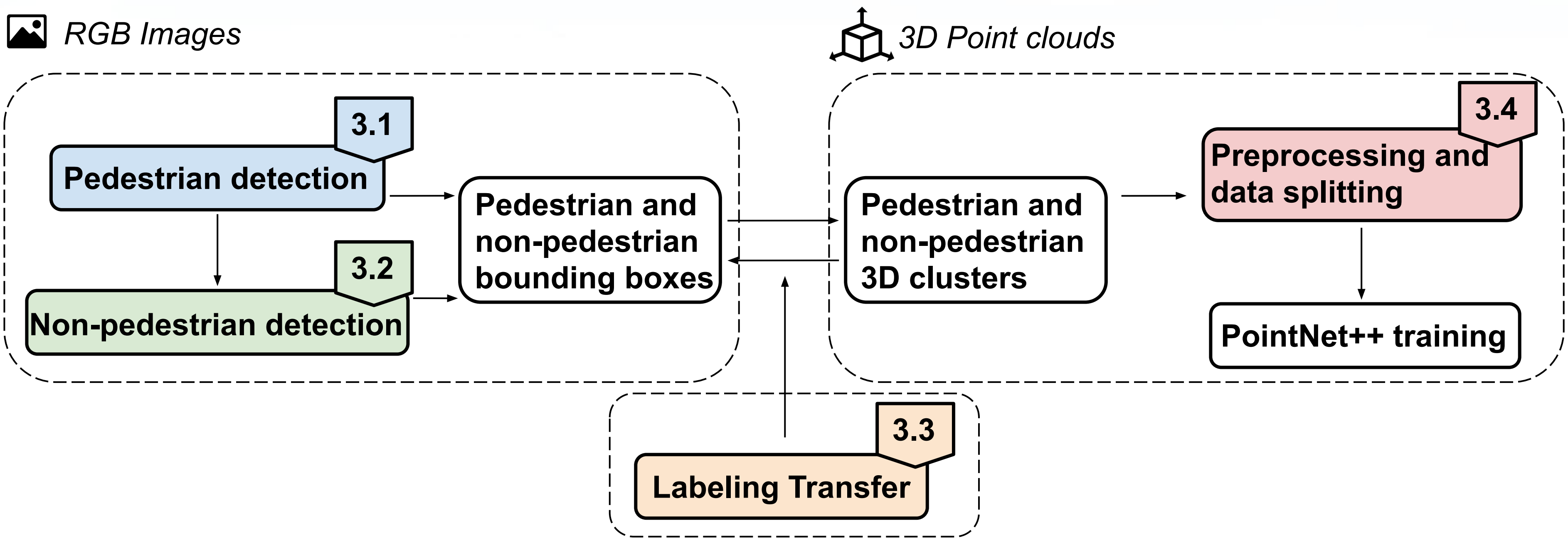
Labeled pedestrian and non-pedestrian 3D point clouds

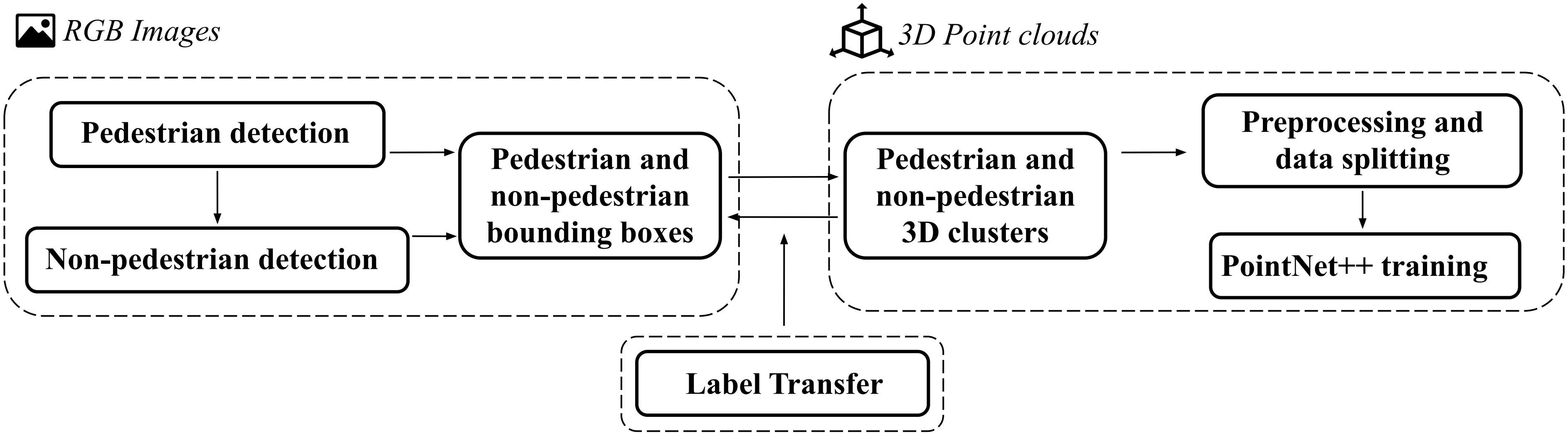
Detect pedestrians and non-pedestrians in RGB images

Transfer labels from RGB images onto 3D point clouds

Contents

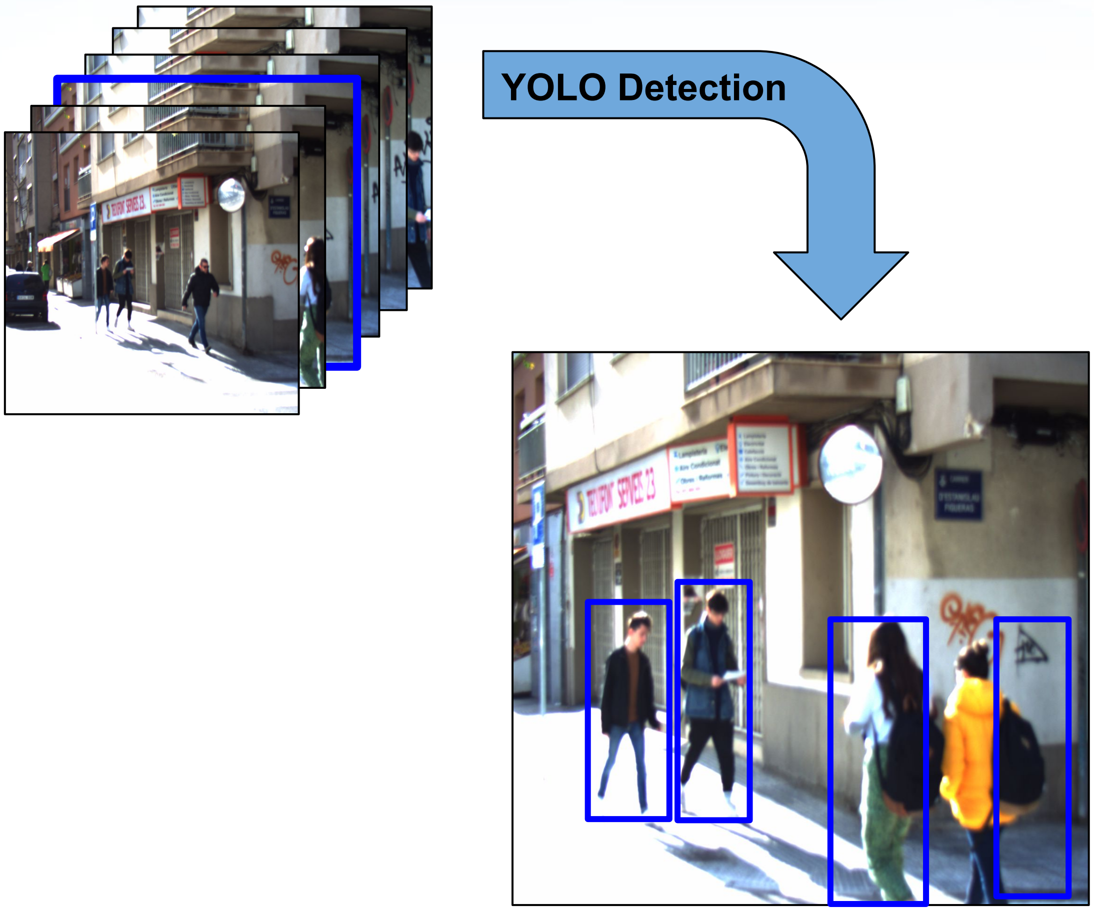
3. Methods


RGB Images
3D Point clouds
3.1
Pedestrian detection
3.2
Non-pedestrian detection
Pedestrian and non-pedestrian bounding boxes
Pedestrian and non-pedestrian 3D clusters
3.4
Preprocessing and data splitting
PointNet++ training
3.3
Labeling Transfer


RGB Images
Pedestrian detection
Non-pedestrian detection
Pedestrian and non-pedestrian bounding boxes
3D Point clouds
Pedestrian and non-pedestrian 3D clusters
Preprocessing and data splitting
PointNet++ training
Label Transfer

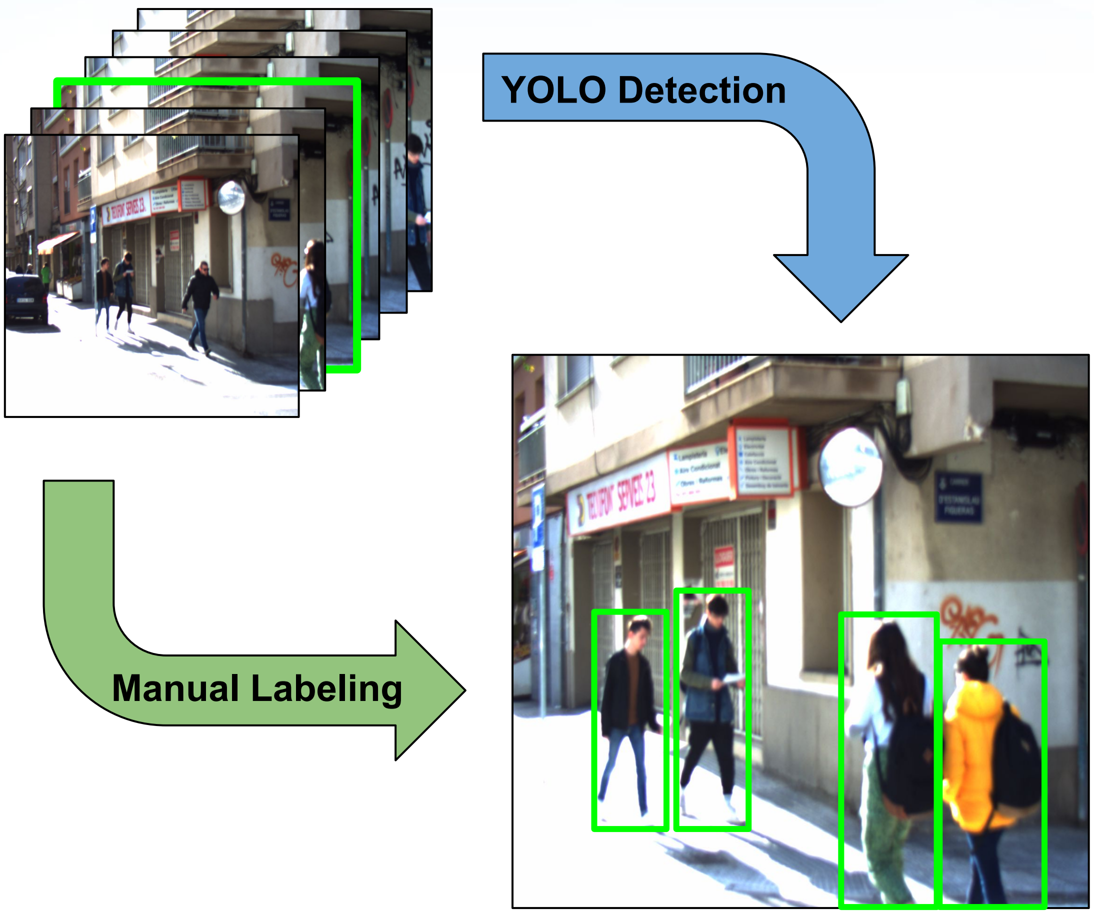
3.1. Pedestrian Detection in RGB Images (I)

YOLO: You Only Look Once



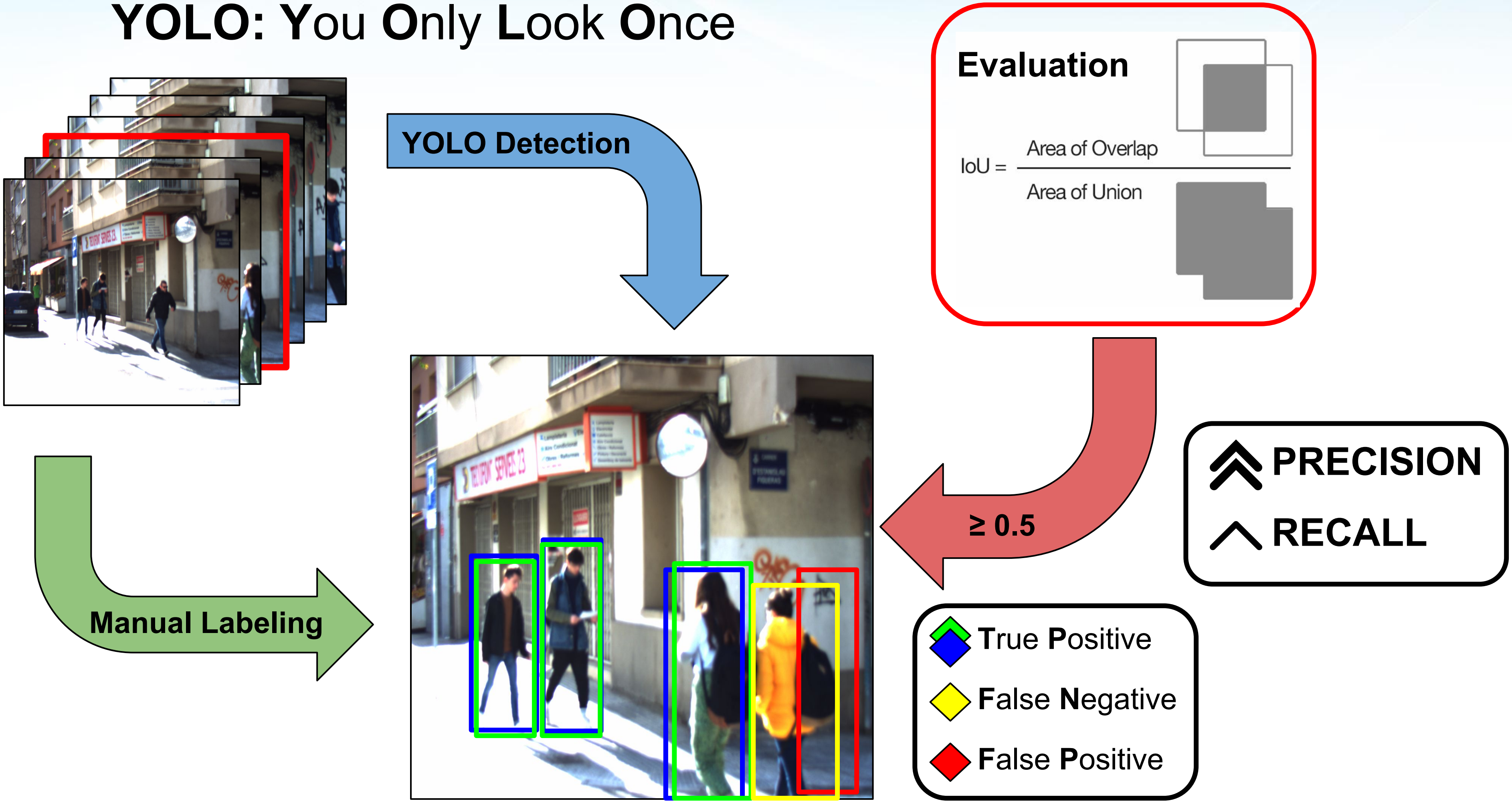
3.1. Pedestrian Detection in RGB Images (II)

YOLO: You Only Look Once



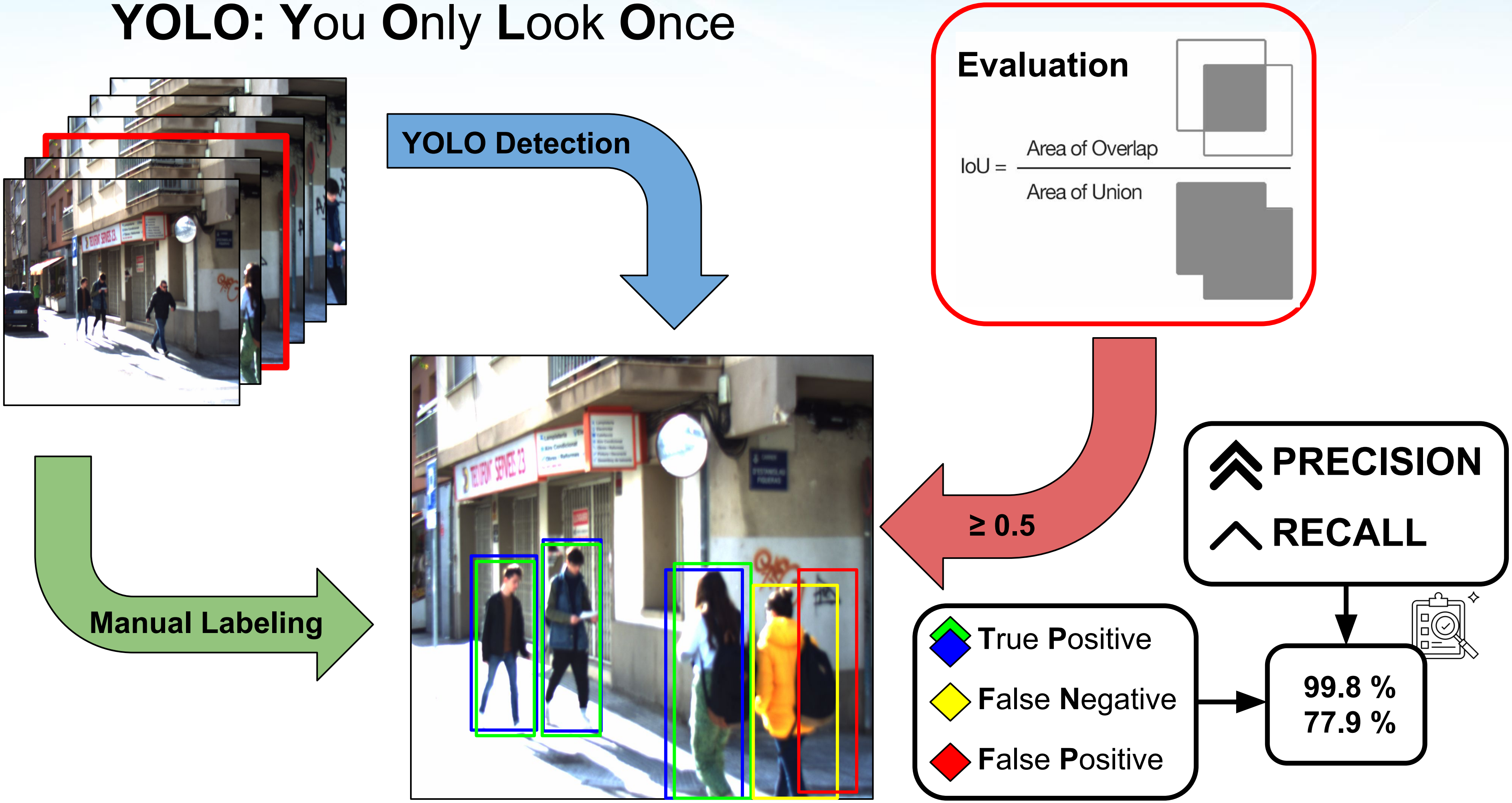
3.1. Pedestrian Detection in RGB Images (III)

YOLO: You Only Look Once



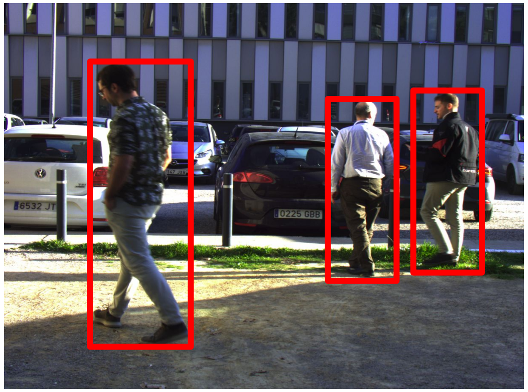
3.1. Pedestrian Detection in RGB Images (III)

YOLO: You Only Look Once



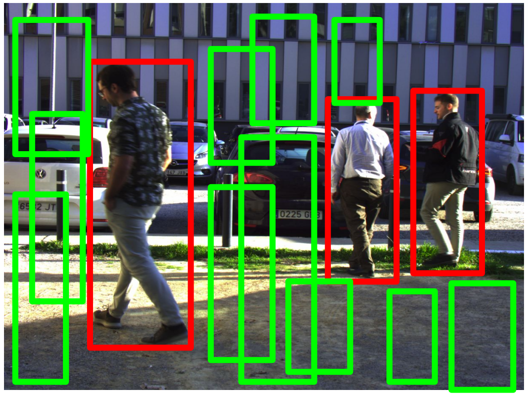
3.2. Non-Pedestrian Detection in RGB Images (I)

Pedestrian bounding boxes



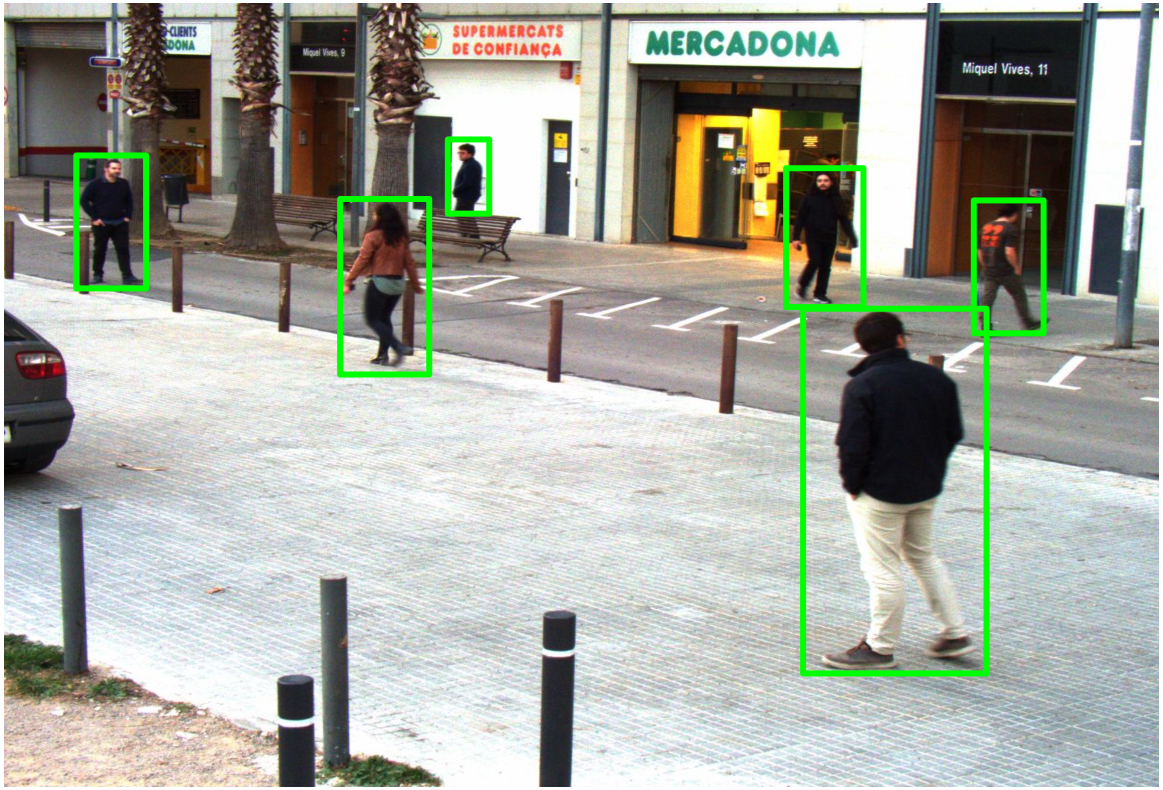
Label non-pedestrians

Non-pedestrian bounding boxes



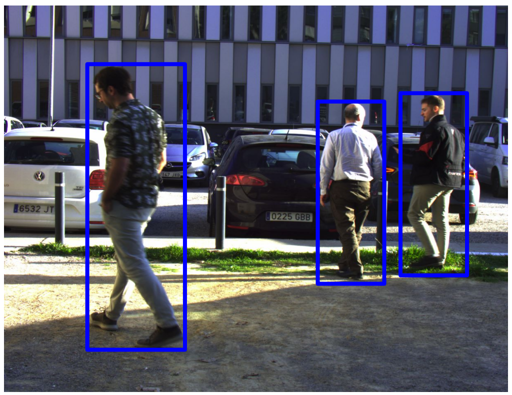
11.5 % Pedestrians
88.5 % Non-pedestrians

Similar size and shape statistics


SUPERMERCATS
DE CONFIANÇA
MERCADONA
Miquel Vives, 11
Miquel Vives, 9

3.2. Non-Pedestrian Detection in RGB Images (I)

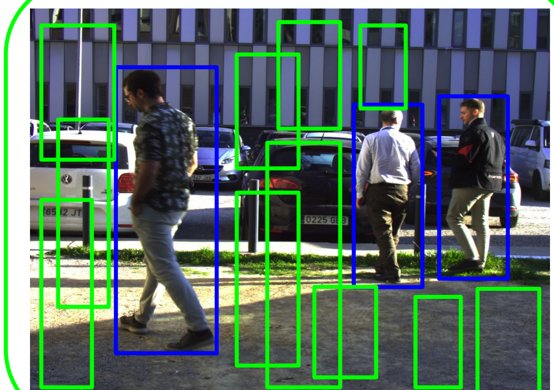
Pedestrian bounding boxes



13,232

Label non-pedestrians

Non-pedestrian bounding boxes



102,130

11.5 % Pedestrians
88.5 % Non-pedestrians

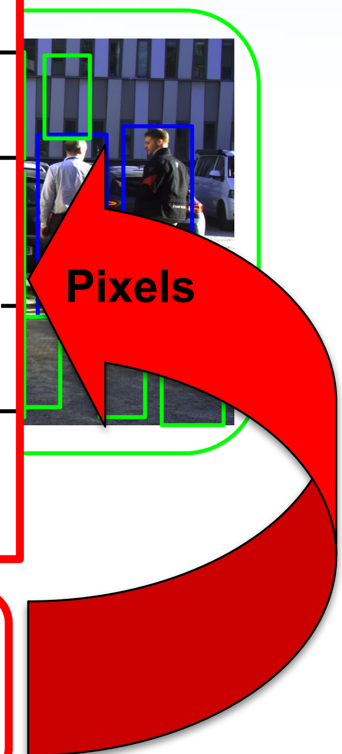
Similar size and shape statistics

3.2. Non-Pedestrian Detection in RGB Images (II)

Pedestri

Bounding box class	Statistics	Width	Height	Ratio
Pedestrian	Mean	95.3	266.3	2.9
	Std. deviation	56.9	145.6	1.0
Non-pedestrian	Mean	111.1	302.5	2.8
	Std. deviation	37.9	102.3	0.9

unding boxes

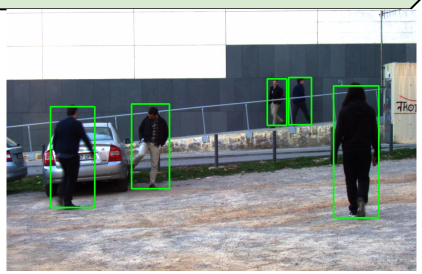


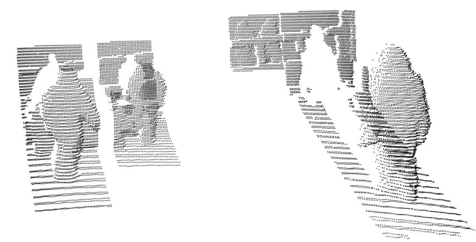
11.5 % Pedestrians
88.5 % Non-pedestrians

Similar size and shape statistics

3.3. Labeling Transfer (I)

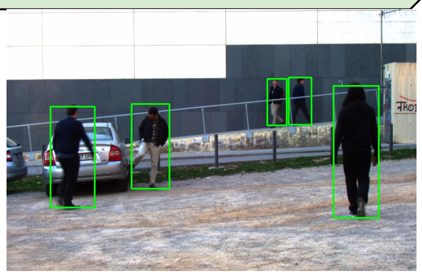
Detect pedestrians and non-pedestrians in RGB images → **Project 3D point clouds onto the RGB images** → **Transfer pedestrian and non-pedestrian labels**





3.3. Labeling Transfer (I)

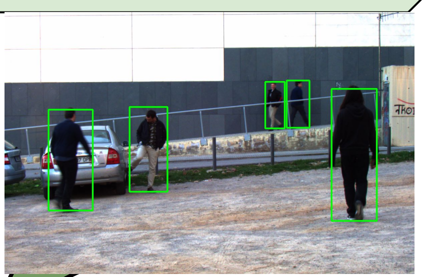
Detect pedestrians and non-pedestrians in RGB images → **Project 3D point clouds onto the RGB images** → **Transfer pedestrian and non-pedestrian labels**



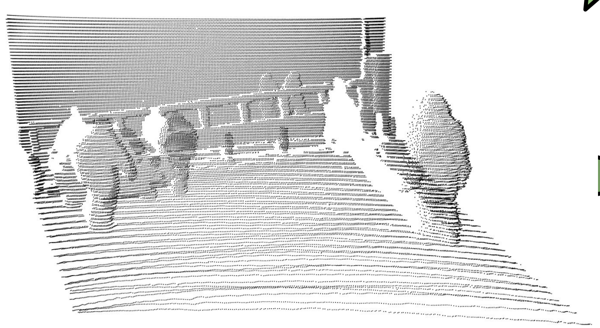
We need registered and synchronized RGB images and 3D point clouds

3.3. Labeling Transfer (I)

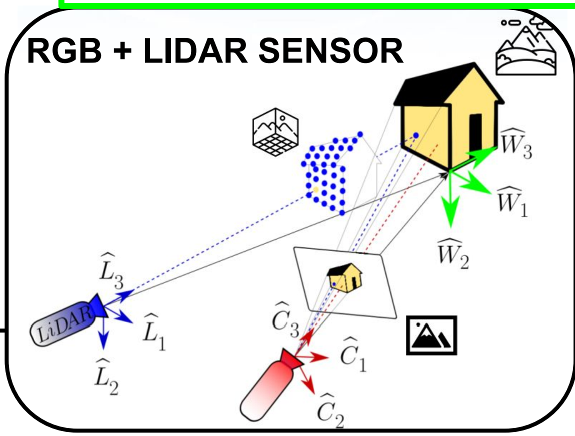
Detect pedestrians and non-pedestrians in RGB images → **Project 3D point clouds onto the RGB images** → **Transfer pedestrian and non-pedestrian labels**



We need registered and synchronized RGB images and 3D point clouds



Projection Matrices



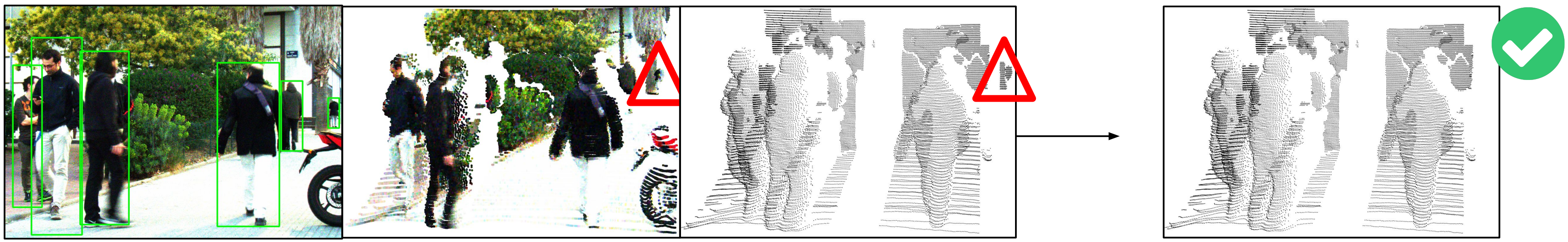
3.3. Labeling Transfer (II)

Problems

LIDAR cannot capture some elements properly

Constraints

Minimum # points: 1024



3.3. Labeling Transfer (III)

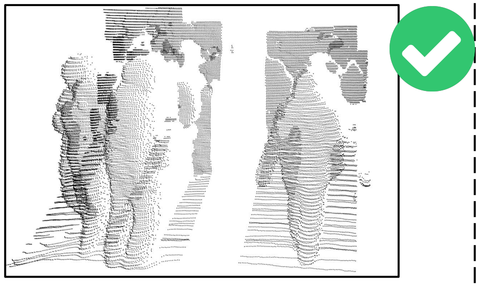
Problems

LIDAR cannot capture some elements properly

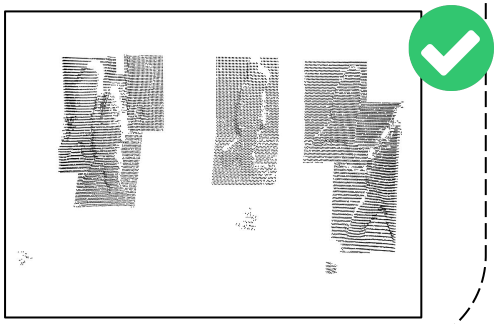
LIDAR Field Of View (FOV) < Camera resolution

Constraints

Minimum # points: 1024



Minimum area: 70%



3.4. Pedestrian Detection in 3D Point Clouds (I)

Labeled pedestrian and non-pedestrian point clouds

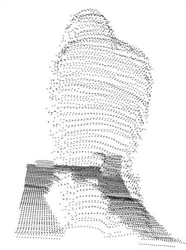
3.4. Pedestrian Detection in 3D Point Clouds (I)

Labeled pedestrian and non-pedestrian point clouds

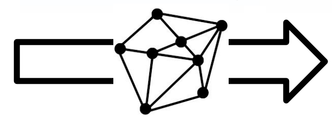
Preprocessing

- Downsampling
- Normalization

Original: 7063 points



Farthest Point Sampling



Downsampled: 1024 points

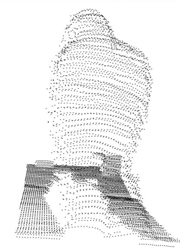
3.4. Pedestrian Detection in 3D Point Clouds (II)

Labeled pedestrian and non-pedestrian point clouds

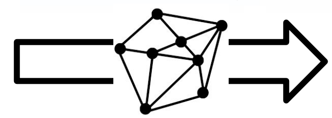
Preprocessing

- Downsampling
- Normalization

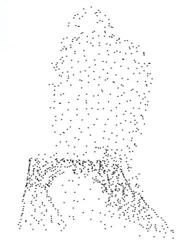
Original:
7063 points



Farthest Point Sampling



Downsampled:
1024 points



Dataset splitting

10 Scenes

DATASET

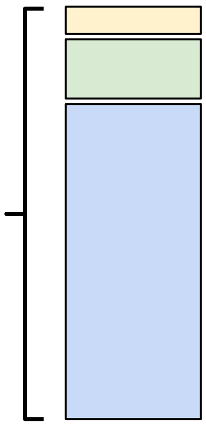
Scene #10

TEST

Scenes #1-9

20% VALIDATION

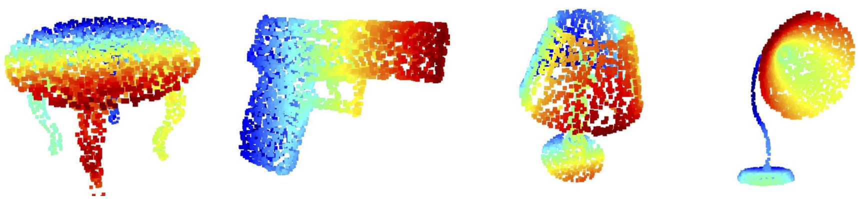
80% TRAINING



4. Previous Experiments

PointNet++ datasets: ModelNet40

Charles R. Qi et al., **<< Pointnet++: Deep hierarchical feature learning on point sets in a metric space >>**, Standford University, 2017

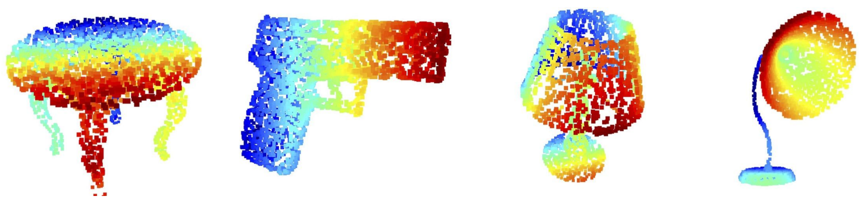


	ModelNet40	Our datasets
Source	CAD models	LIDAR sensor
Points density	Uniform	Not uniform
# classes	40	2
Balanced dataset	Yes	No

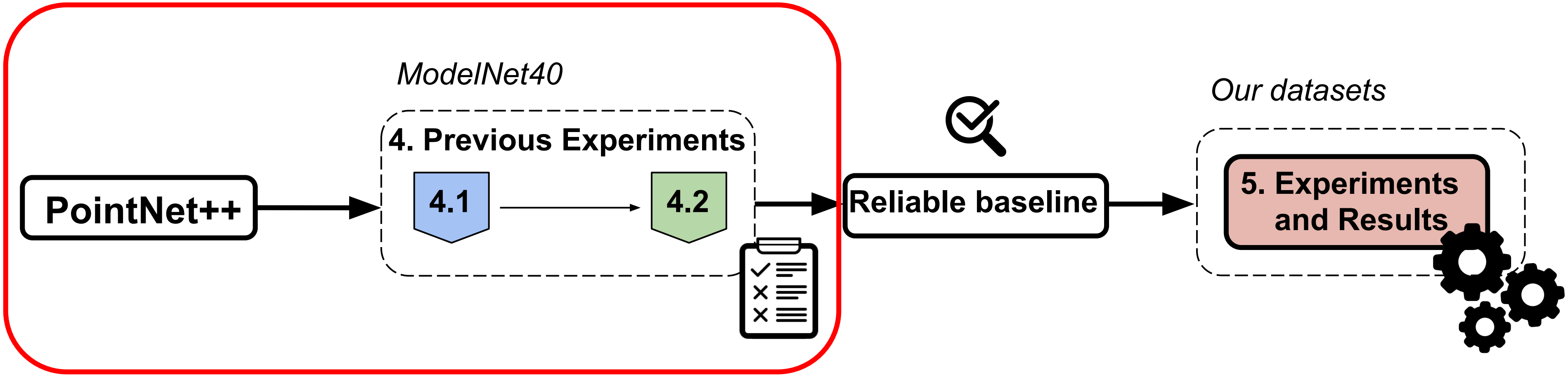
4. Previous Experiments

PointNet++ datasets: ModelNet40

Charles R. Qi et al., **<< Pointnet++: Deep hierarchical feature learning on point sets in a metric space >>**, Standford University, 2017

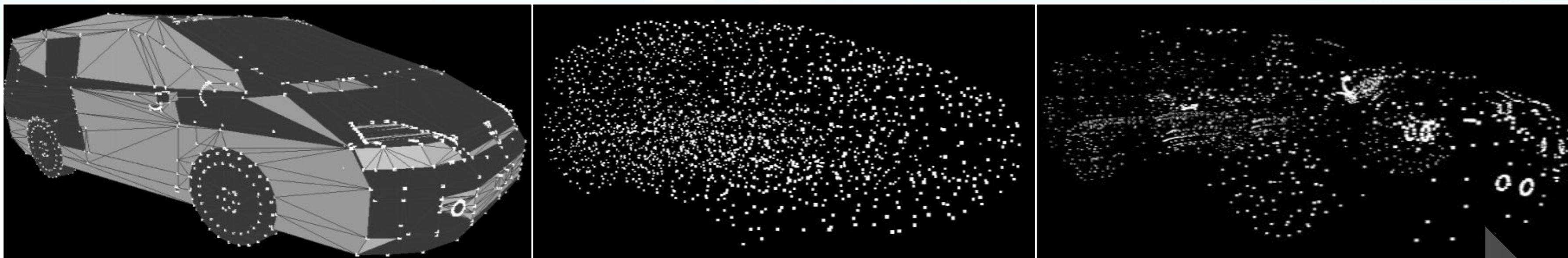


		ModelNet40	Our datasets
4.1	**Source**	CAD models	LIDAR sensor
	Points density	Uniform	Not uniform
4.2	**# classes**	40	2
	Balanced dataset	Yes	No

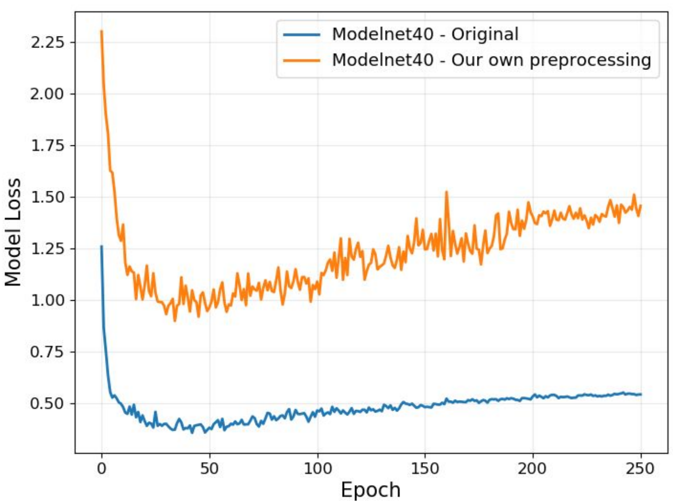


4.1. ModelNet40: Different Preprocessing

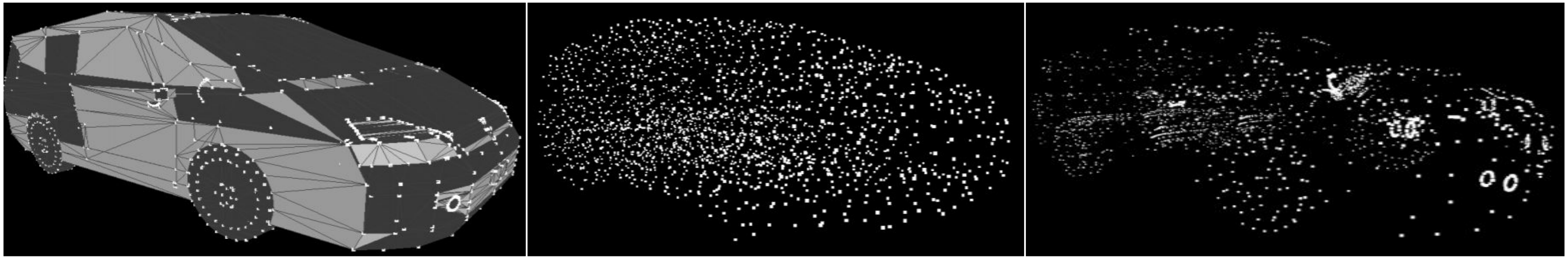
Original *PointNet authors preprocessing* *Our own preprocessing*



Less interpretability



Experiment	Accuracy	Average class accuracy
Original	88.9	86.6
Our own preprocessing	**74.1**	**62.6**



4.2. ModelNet40: Binary Classification

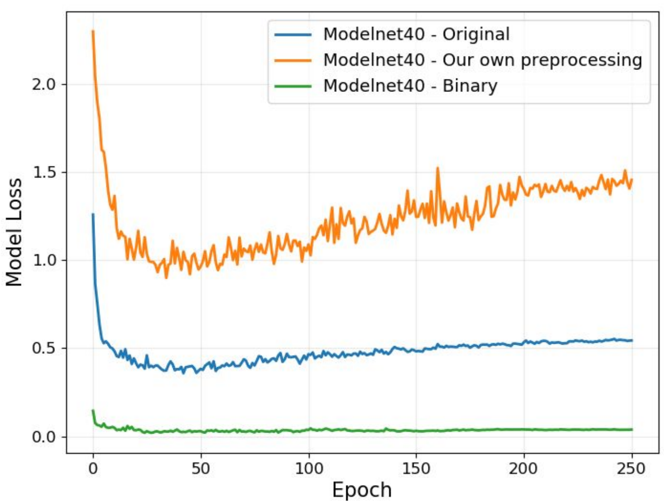
Original

40 classes
Balanced dataset

Binary - Imbalanced

2 classes
Imbalanced dataset

Precision	Recall
98.0	96.4

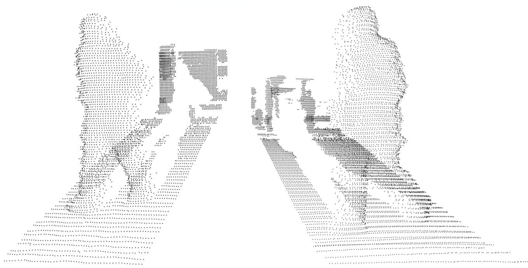


Experiment	Accuracy	Average class accuracy
Original	88.9	86.6
Our own preprocessing	74.1	62.6
Binary	**99.4**	**98.1**

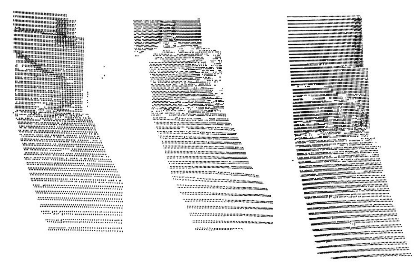
5. Experiments and Results

Our datasets

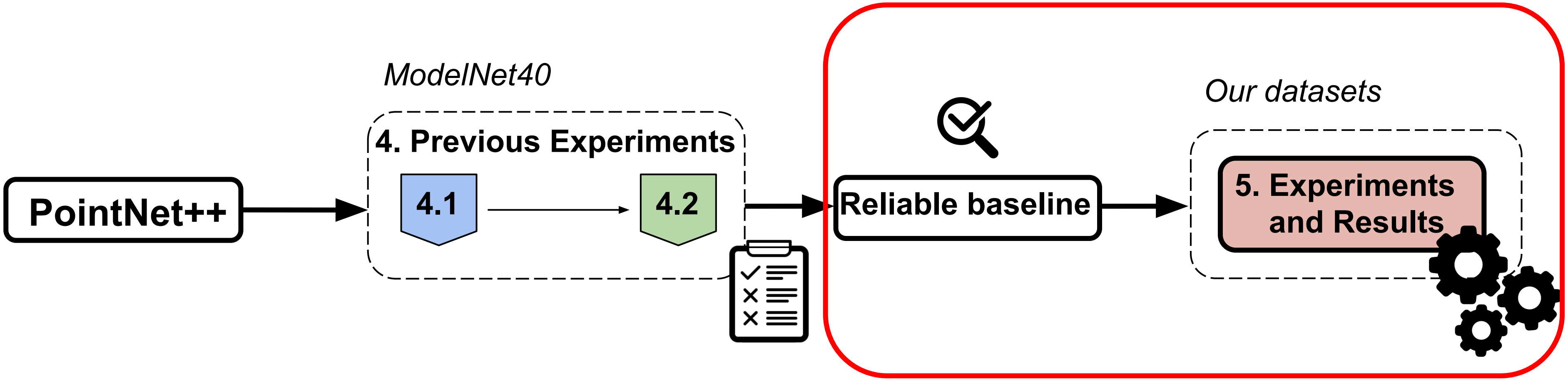
Pedestrian clusters



Non-pedestrian clusters



	Our datasets
Source	LIDAR sensor
Points density	Not uniform
# classes	2
Balanced dataset	No



5.1. Batch size vs. number of point clouds

Our datasets

	ModelNet40	Our datasets
Number of point clouds	12,308	87,536

Experiment	Precision	Recall
Batch size: 32	96.6	32.5
Batch size: 64	95.4	29.9
Batch size: 128	78.0	31.9
Less training clusters	93.2	31.9

5.1. Batch size vs. number of point clouds

Our datasets

	ModelNet40	Our datasets
Number of point clouds	12,308	87,536

Baseline

Experiment	Precision	Recall
Batch size: 32	**96.6**	**32.5**
Batch size: 64	95.4	29.9
Batch size: 128	78.0	31.9
Less training clusters	93.2	31.9

5.1. Batch size vs. number of point clouds

Our datasets

	ModelNet40	Our datasets
Number of point clouds	12,308	87,536

Baseline

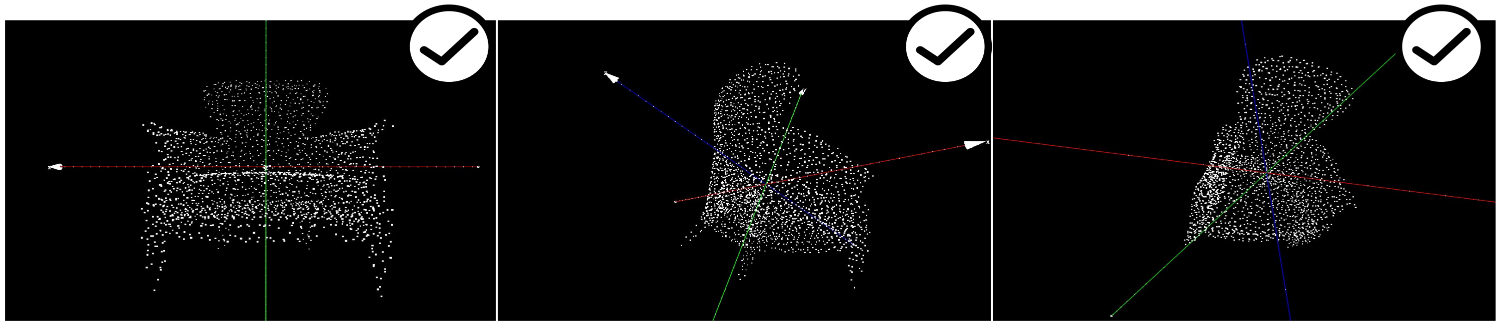
Experiment	Precision	Recall
Batch size: 32	**96.6**	**32.5**
Batch size: 64	95.4	29.9
Batch size: 128	78.0	31.9
Less training clusters	93.2	31.9

YOLO	**99.8**	**77.9**

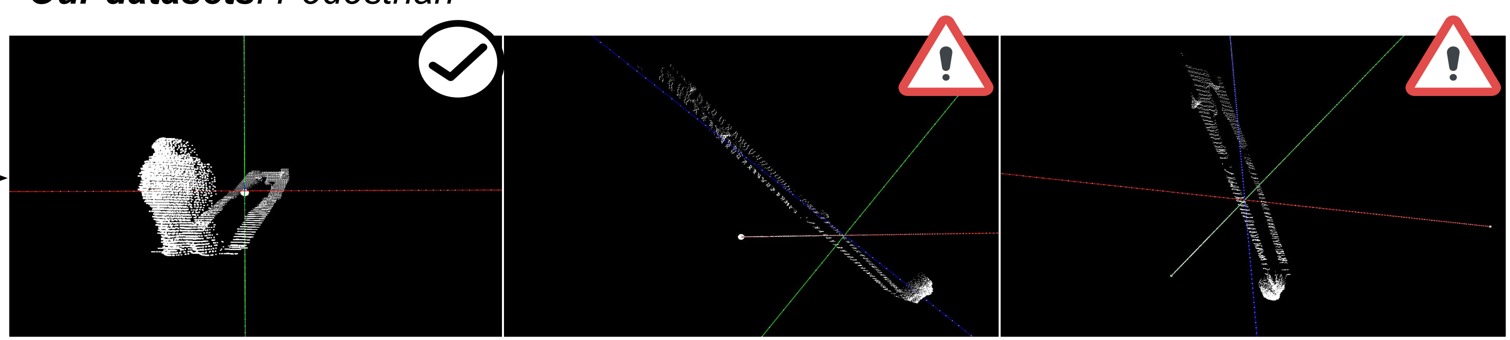
5.2. Without data augmentation (I)

<u>Data augmentation</u> → **ROTATION**

Modelnet40: *Chair*



Our datasets: *Pedestrian*



5.2. Without data augmentation (II)

<u>Our datasets</u>

	Experiment	Precision	Recall
	Batch size: 32	**96.6**	**32.5**
	Batch size: 64	95.4	29.9
	Batch size: 128	78.0	31.9
	Less training clusters	93.2	31.9
*	**Without data augmentation**	**99.1** ↑	**98.6** ↑↑

5.2. Without data augmentation (II)

Our datasets

Baseline

Experiment	Precision	Recall
Batch size: 32	96.6	32.5
Batch size: 64	95.4	29.9
Batch size: 128	78.0	31.9
Less training clusters	93.2	31.9
Without data augmentation	**99.1** ↑	**98.6** ↑↑

YOLO	**99.8**	**77.9**

5.3. Non-pedestrians with less overlap

Our datasets

	↓↑	*
Non-pedestrian overlap	~70%	~20%

	Experiment	Precision	Recall
	Batch size: 32	96.6	32.5
	Batch size: 64	95.4	29.9
	Batch size: 128	78.0	31.9
	Less training clusters	93.2	31.9
↓↑	**Without data augmentation**	**99.1**	**98.6**
*	**Non-pedestrians with less overlap**	**97.1** ↓	**97.7** ↓

5.4. Multi Scale Grouping (MSG) Model

Our datasets

		*
PointNet++ architecture	Single Scale Grouping	Multi Scale Grouping

	Experiment	Precision	Recall
	Batch size: 32	96.6	32.5
	Batch size: 64	95.4	29.9
	Batch size: 128	78.0	31.9
	Less training clusters	93.2	31.9
	Without data augmentation	**99.1**	**98.6**
	Non-pedestrians with less overlap	97.1	97.7
*	**MSG model**	**99.4** ↑	**92.5** ↓

5.5. Batch Size: 32 + Without data augmentation

Our datasets

Experiment	Precision	Recall
Batch size: 32	96.6	32.5
Batch size: 64	95.4	29.9
Batch size: 128	78.0	31.9
Less training clusters	93.2	31.9
Without data augmentation	**99.1**	**98.6**
Non-pedestrians with less overlap	97.1	97.7
MSG Model	99.4	92.5

6. Conclusions

Pedestrian detection system in point clouds using Deep Neural Networks

6. Conclusions

Pedestrian detection system in point clouds using Deep Neural Networks

PointNet++ can help YOLO

	Precision	Recall
YOLO	99.8	77.9
PointNet++	**99.1**	**98.6**

6. Conclusions

Pedestrian detection system in point clouds using Deep Neural Networks

PointNet++ can help YOLO

	Precision	Recall
YOLO	99.8	77.9
PointNet++	**99.1**	**98.6**

System to generate a dataset with ground truth in point clouds

6. Conclusions

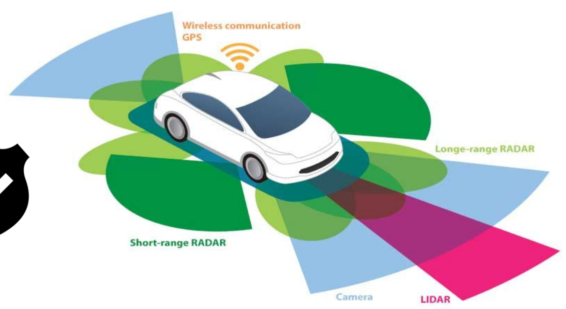
Pedestrian detection system in point clouds using Deep Neural Networks

PointNet++ can help YOLO

	Precision	Recall
YOLO	99.8	77.9
PointNet++	**99.1**	**98.6**

System to generate a dataset with ground truth in point clouds

LIDAR sensors → safety, reliability



7. Contributions

- Pedestrian detection in RGB images with YOLO
- YOLO evaluation
- Non-pedestrian detection in RGB images
- Labeling transfer onto 3D point clouds
- Preprocessing and data splitting in 3D point clouds
- Pedestrian detection in 3D point clouds with PointNet++

8. Future Work

- ➔ Scanning strategy to detect pedestrians in point clouds
- ➔ Real-time implementation
- ➔ PointNet++ parameters optimization
- ➔ Strategy to combine point clouds with RGB images

CD6 UPC

Centre for Sensors, Instruments and Systems Development
UNIVERSITAT POLITECNICA DE CATALUNYA

Shaping light to your needs

Thank you for your attention!

EPIC European Photonics Industry Consortium

SECPhO Southern European Cluster in Photonics and Optics

CIT UPC CENTRE D'INNOVACIÓ I TECNOLOGIA
UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

tecnio catalonia ACCIÓ Generalitat de Catalunya Fins a 31/12/2019

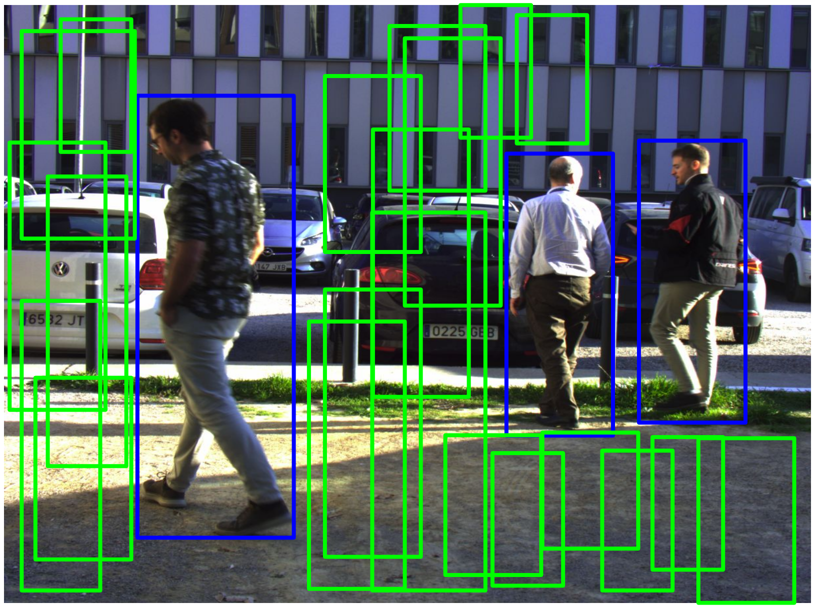
AENOR Empresa Registrada UNE-EN ISO 9001 ER-0016/2012

CERTIFIED IQNet MANAGEMENT SYSTEM

YOLO Evaluation

Scene #	GT boxes	YOLO detections	TP	FP	FN	Precision	Recall
1	104	96	96	0	8	100	92.3
2	117	95	95	0	22	100	81.2
3	127	101	101	0	26	100	79.5
4	168	140	140	0	28	100	83.3
5	122	96	96	0	26	100	78.7
6	118	109	108	1	10	99.1	91.5
7	190	145	145	0	45	100	76.3
8	213	160	160	0	53	100	75.1
9	184	146	145	1	39	99.3	78.8
10	131	63	63	0	68	100	48.1
Total	**1474**	**1151**	**1149**	**2**	**325**	**99.8**	**77.9**

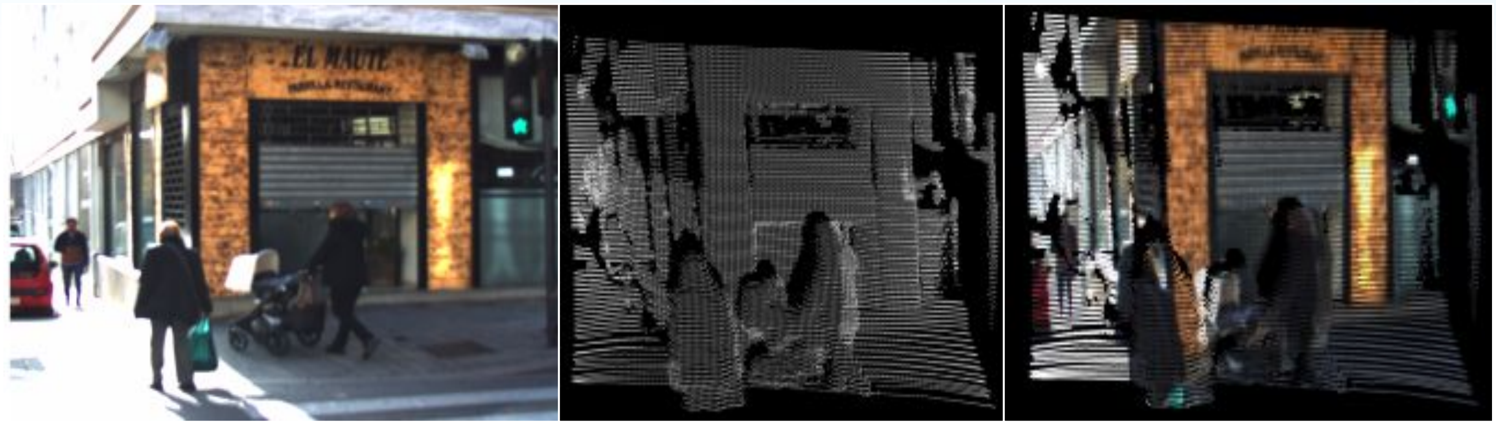
Non-pedestrian bounding boxes with more overlap



Non-pedestrian bounding boxes with less overlap

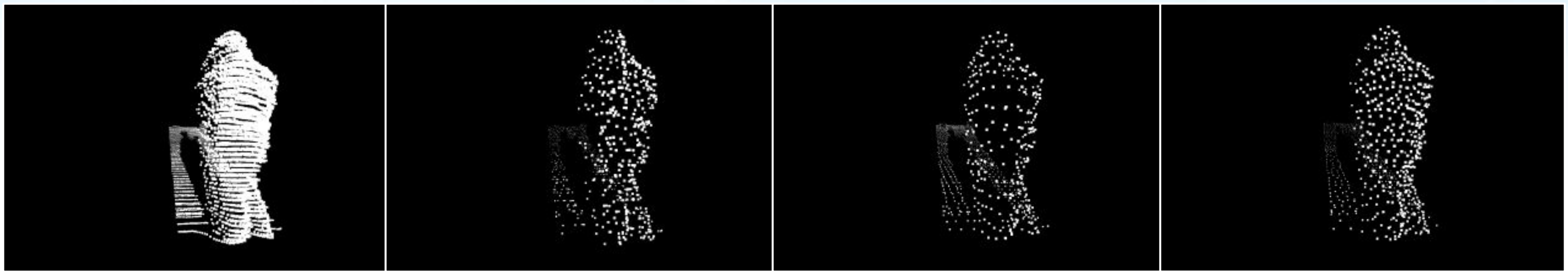
Labeling transfer

Highway: Ego-motion effect

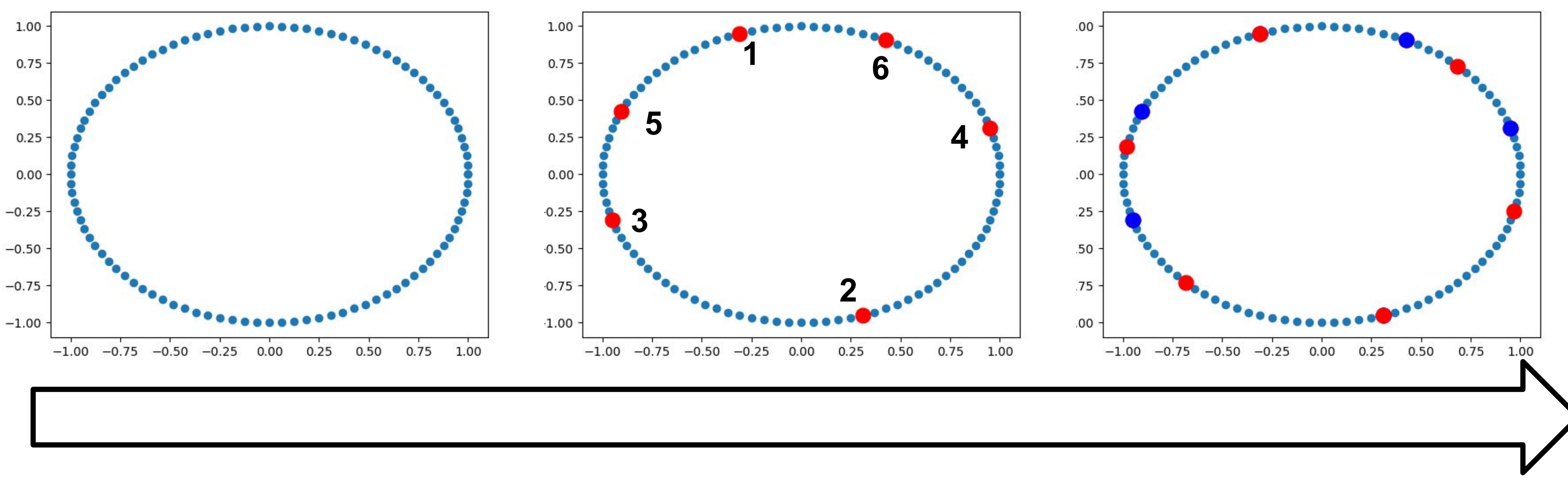


Indoor: Calibration issues

Original	Random	Voxel grid	Farthest Point Sampling



Farthest Point Sampling

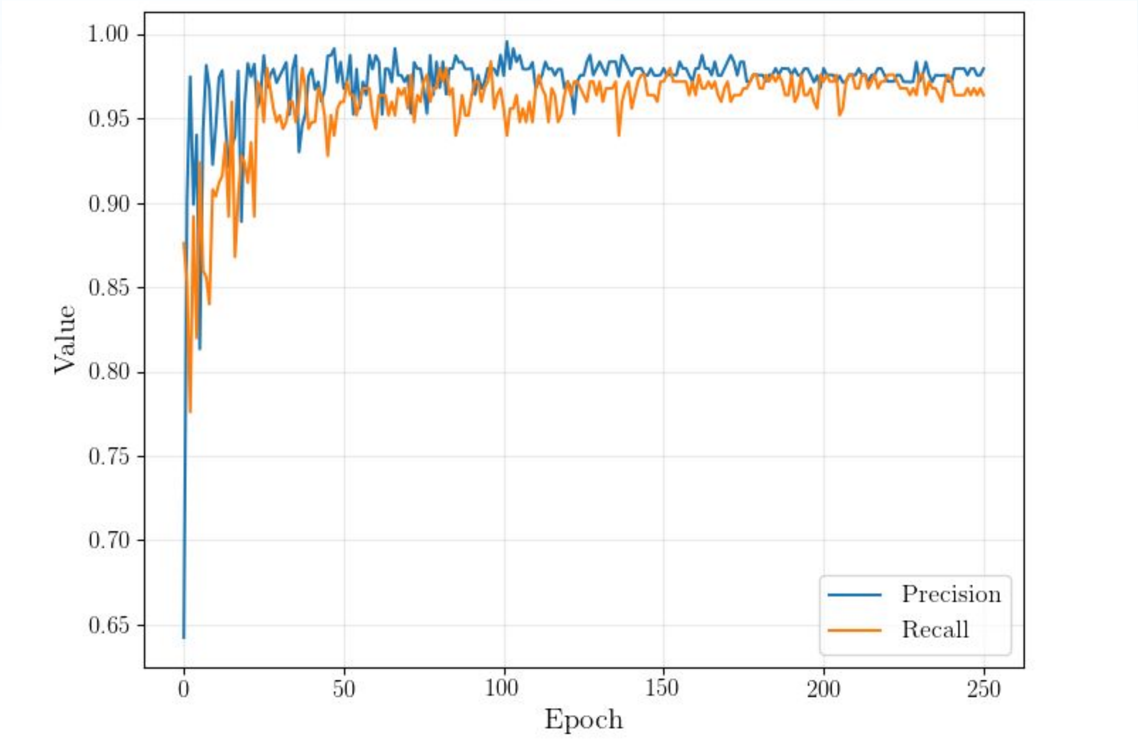


1 2 3 4 5

6 7 8 9 10

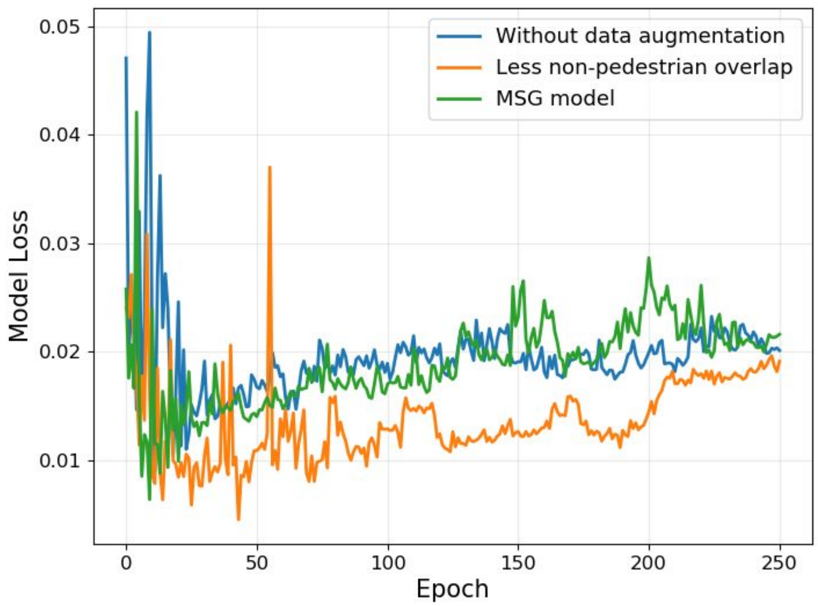
Dataset splitting	Clusters		
Dataset	**Pedestrians**	**Non-Pedestrians**	**Total**
Training	6,932	60,388	67,320
Validation	1,733	15,098	16,831
Test	345	3,040	3,385
Total	**9,010**	**78,526**	**87,536**

Modelnet40 - Binary classification

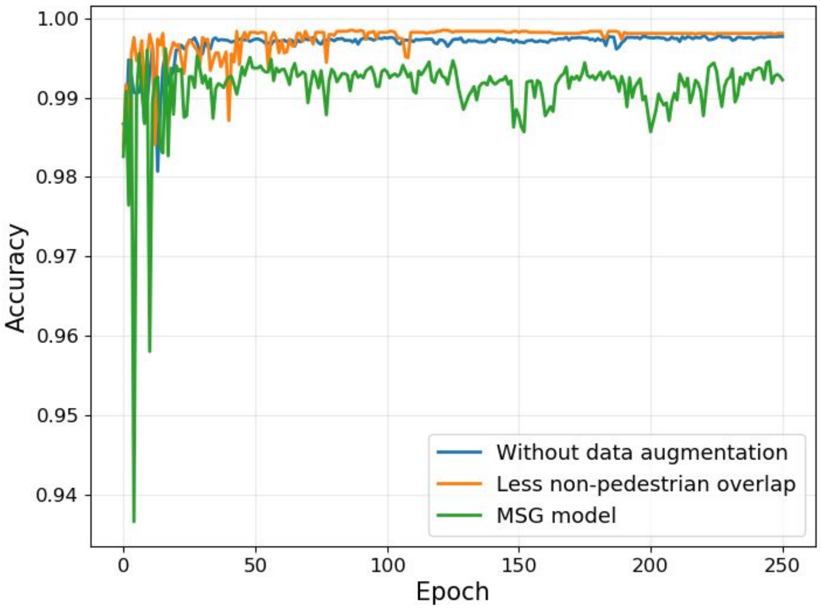


Our datasets

Model Loss

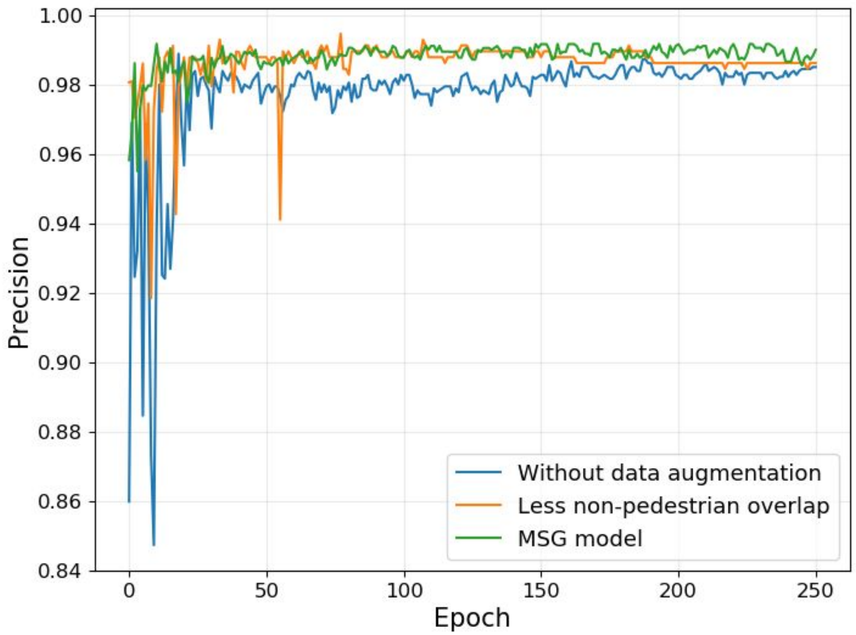


Accuracy



Our datasets

Precision



Recall