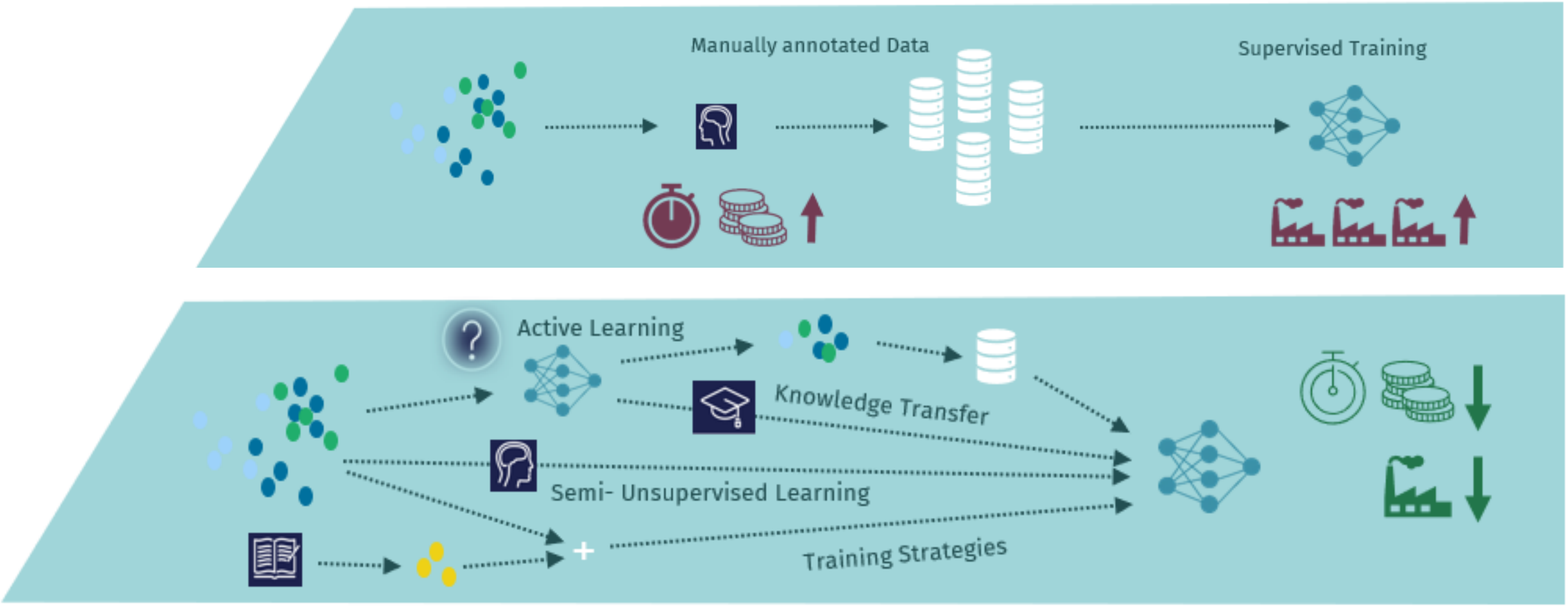
KIDELTA LEARNING
Scalable AI for Automated Driving

Final Event | March 09, 2023

Didactics

Marius Bachhofer

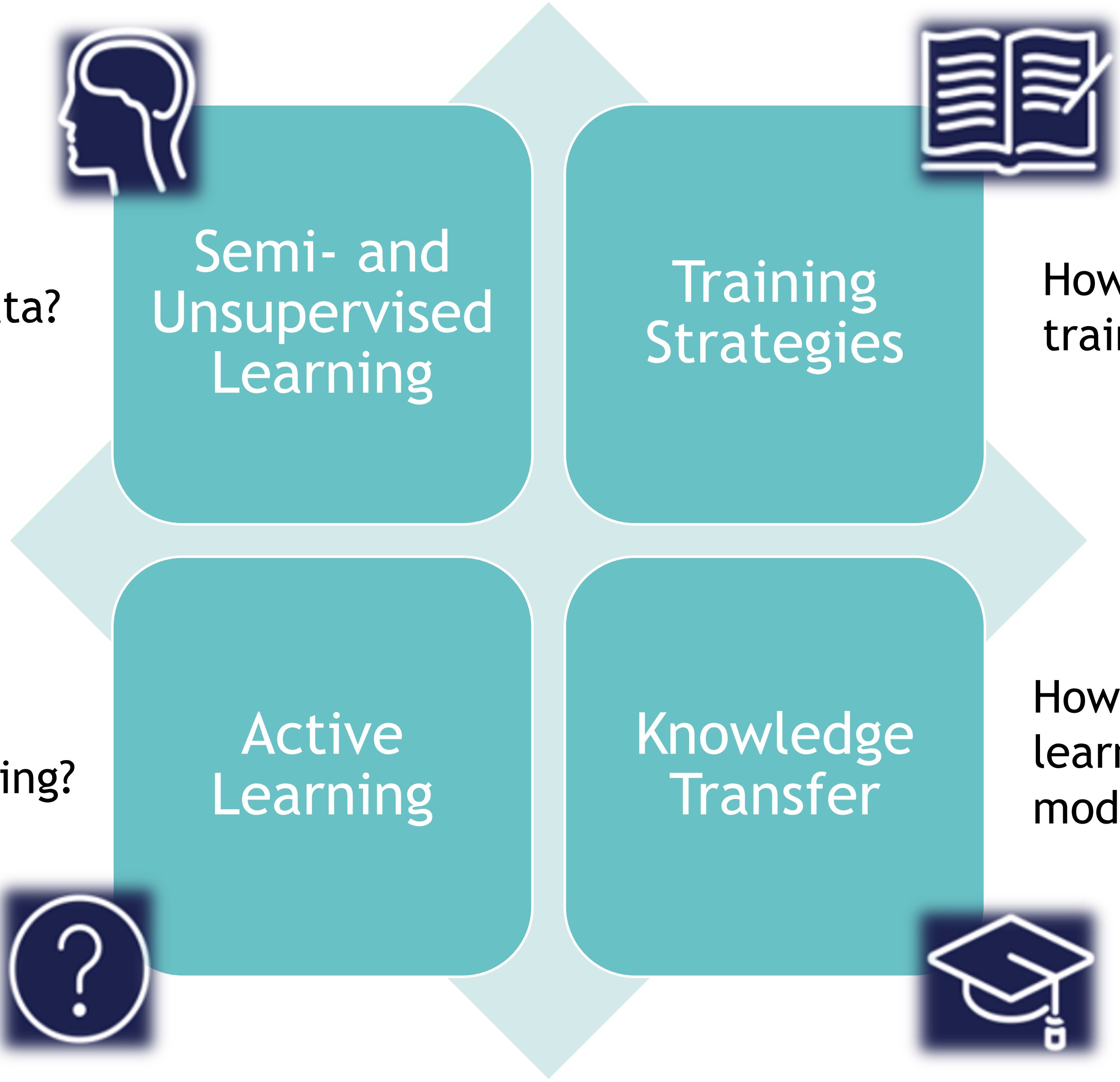
Didactics
In a Nutshell



Didactics

How to use unlabeled Data?

How to improve the existing training methods?



What data to use for training?

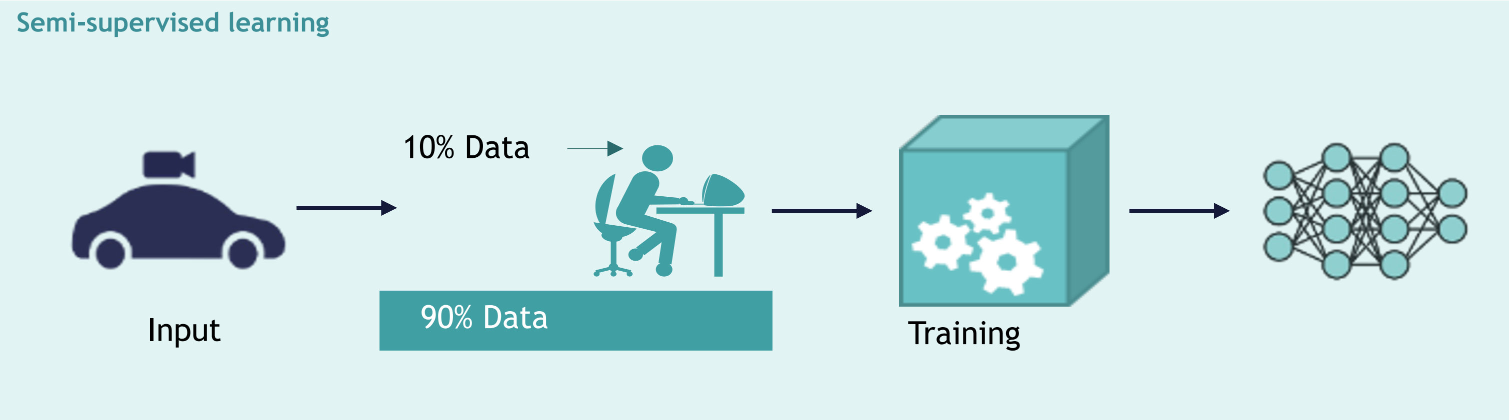
How to transfer already learned knowledge to new models?

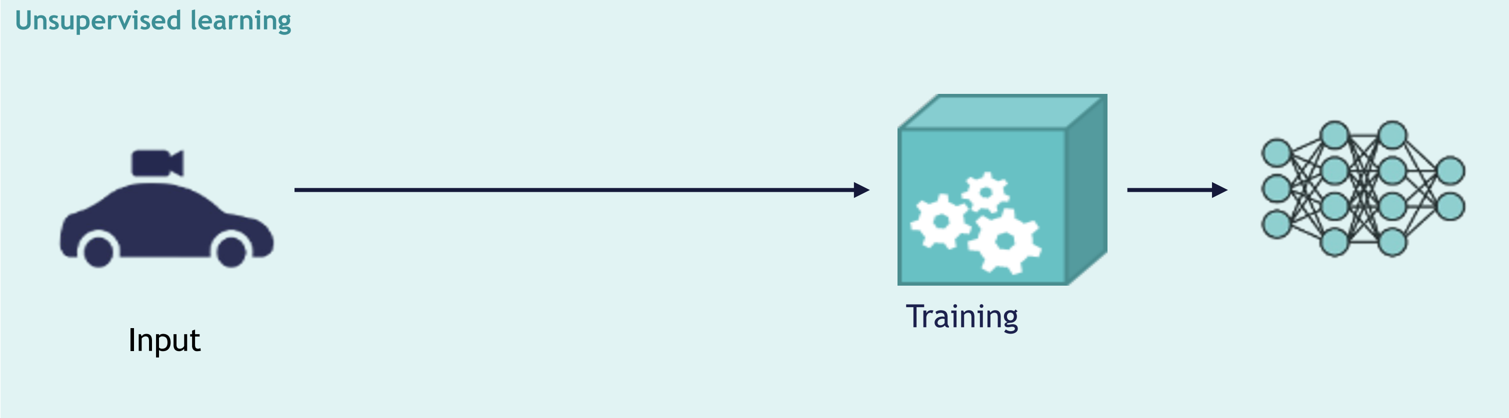
How to use unlabeled Data?

Semi- and Unsupervised Learning

Motivation:

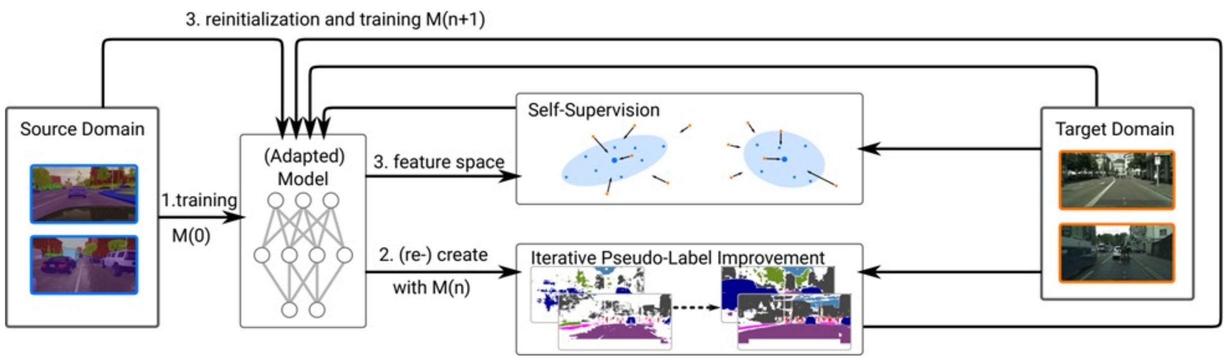
- Labeled data is expensive!
- Unlabeled data is cheap!



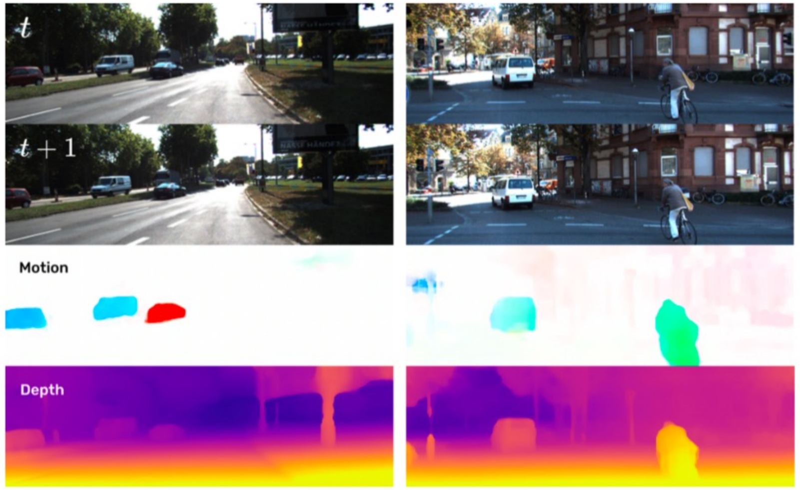


How to use unlabeled Data?

Semi- and Unsupervised Learning



An Iterative Model for Domain Adaptation



M3: Monocular Self-Supervised Depth, Pose and Motion

Other Example Contributions:

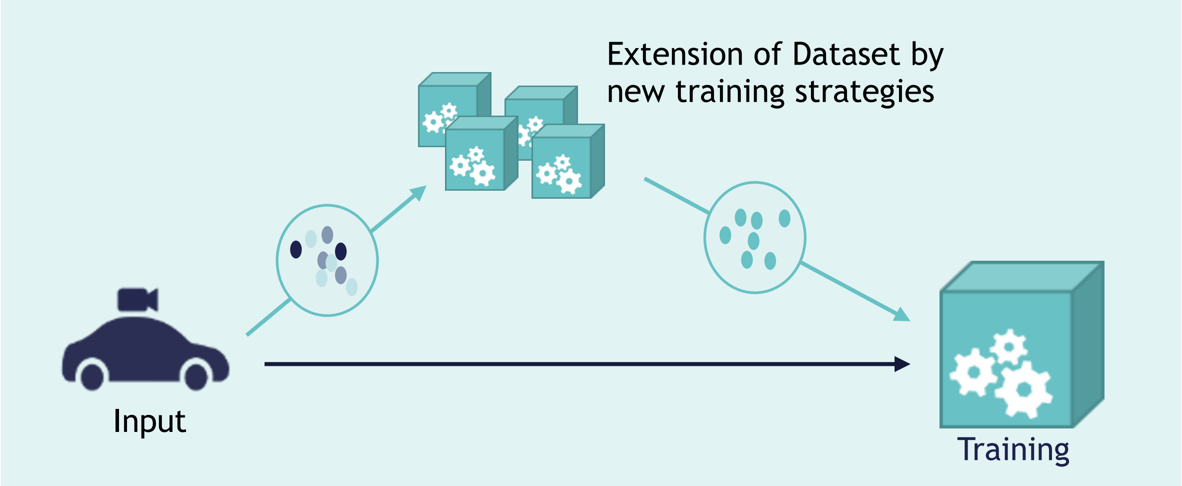
- Self-supervised Learning of 3D Human Body Pose
- Towards Unsupervised Class-Incremental Learning in Semantic Segmentation
- And many more …

How to improve existing methods?

Training Strategies

Motivation:

- effective usage of resources!
- Achieve tolerance of data variation, e.g. compression, data augmentation and anomalies
- Reusable and transferable models in different settings

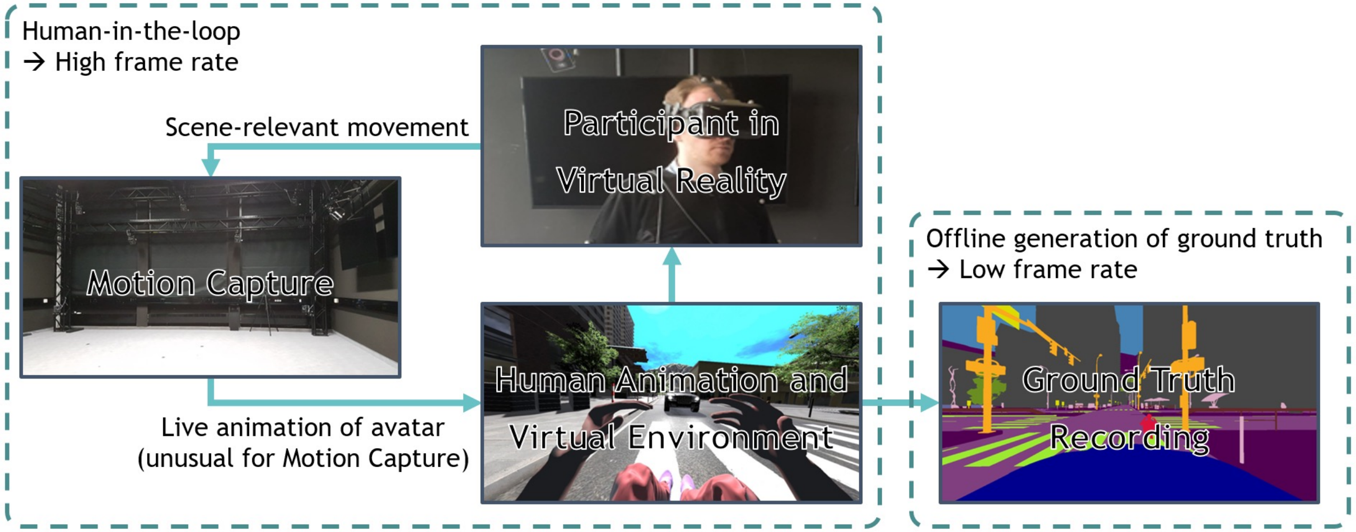


Example Contributions:

- Evaluate the influence of data augmentation methods
- Investigate hierarchical data

How to improve existing methods?

Training Strategies



Co-Simulation and (near) live Corner Case detection

- Recording of relevant human interaction data with Motion Capture lab
- Dangerous situations without potential for physical harm
- Live detection of rare poses during recording

Data augmentation with anisotropic diffusion

- Avoiding overfitting
- Cheap extension of the data → Significant increase of mAP

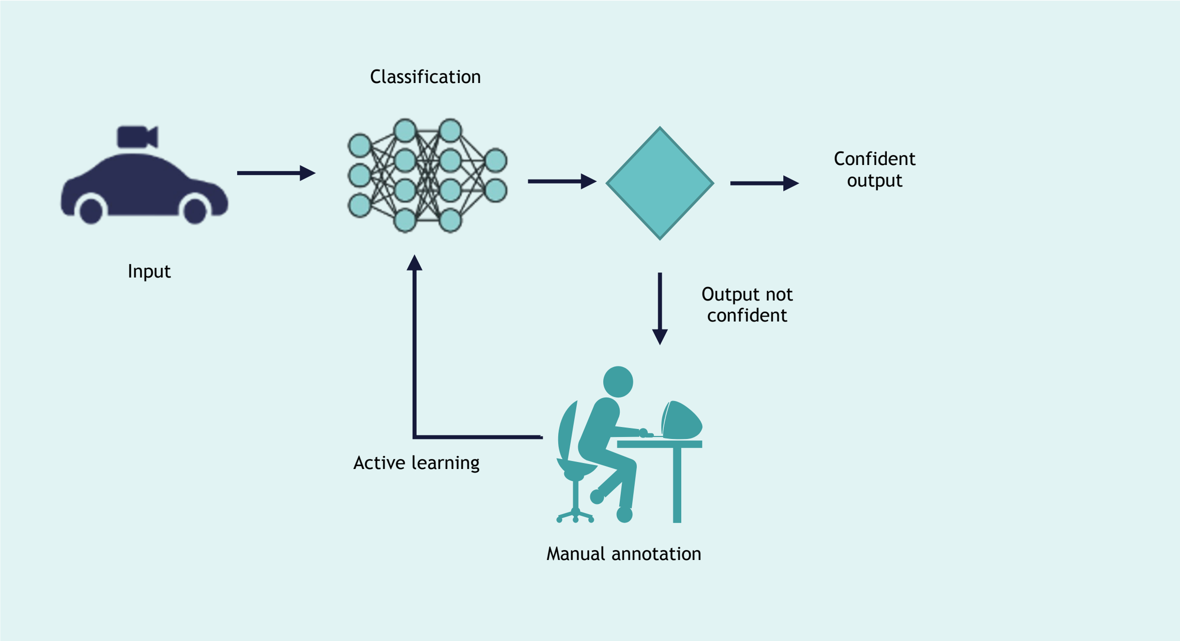
And many more …

What data to use for training?

Active Learning

Motivation:

- Reduce labeling cost by choosing the right subset for labeling!
- Involve the training process for better dataset acquisition



Example Contributions:

- Active Learning on semantic segmentation
- Active Learning on Point Clouds
- A synthetic Oracle for Active Learning
- And many more …

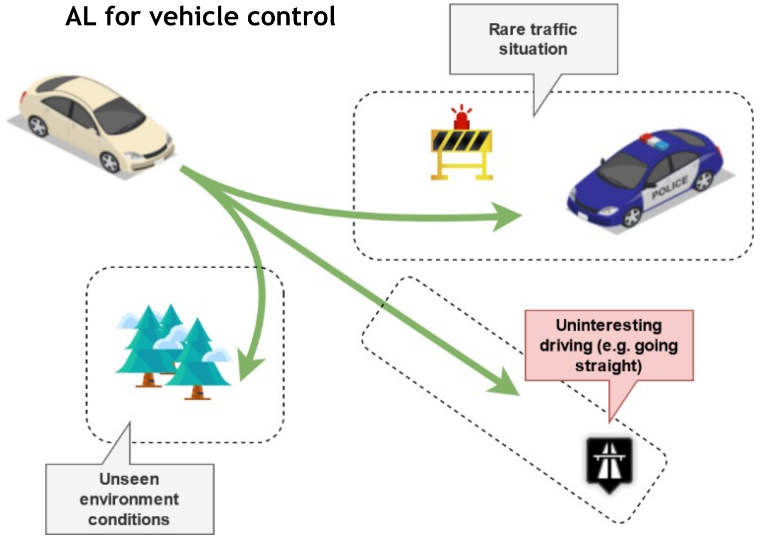
What data to use for training?

Active Learning ?

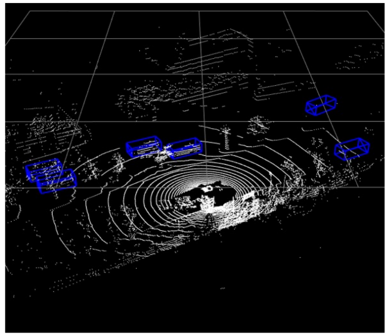
Findings:

- High demands to resources!
- Sophisticated Acquisition Strategies necessary
  - Complex retraining

AL for vehicle control



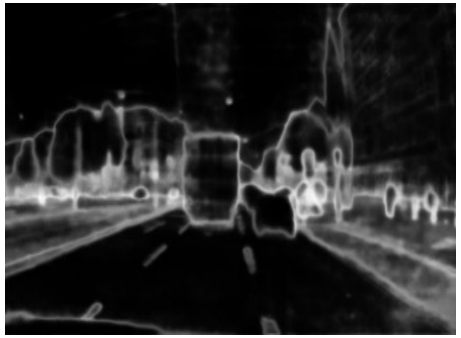
AL for PointCloud Data Selection



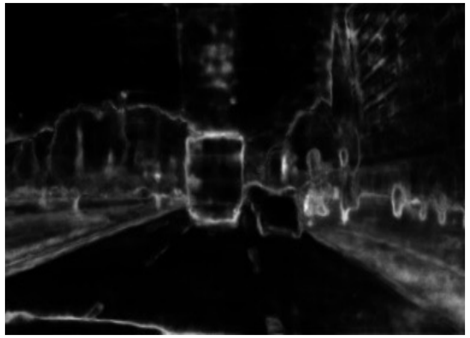
AL for Kamera Data Selection

AL for Data Generation

Entropy



Mutual Information

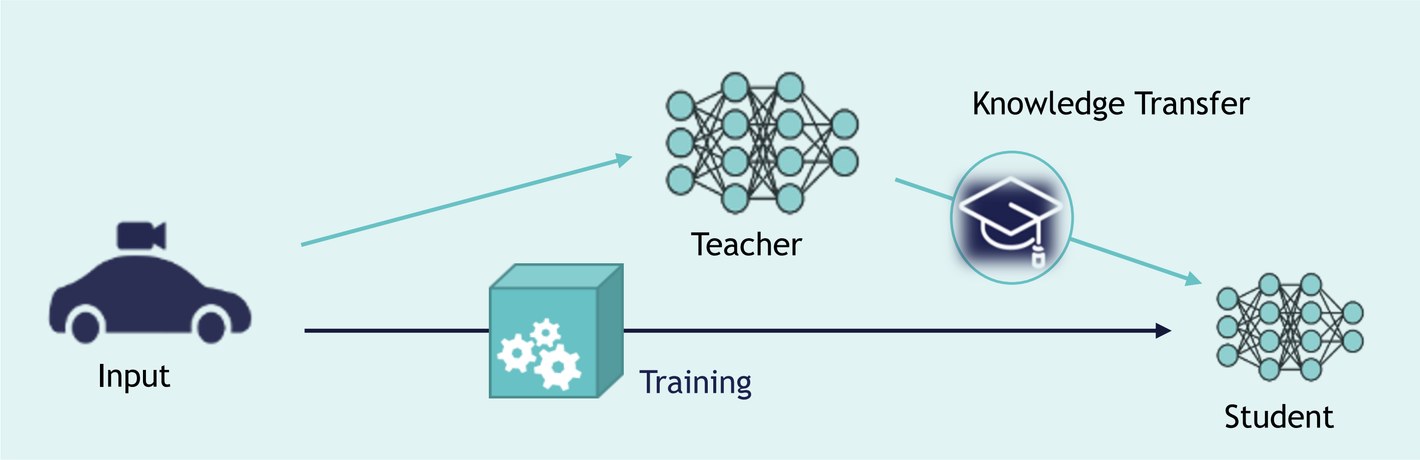


Why to transfer knowledge to new models?

Knowledge Transfer

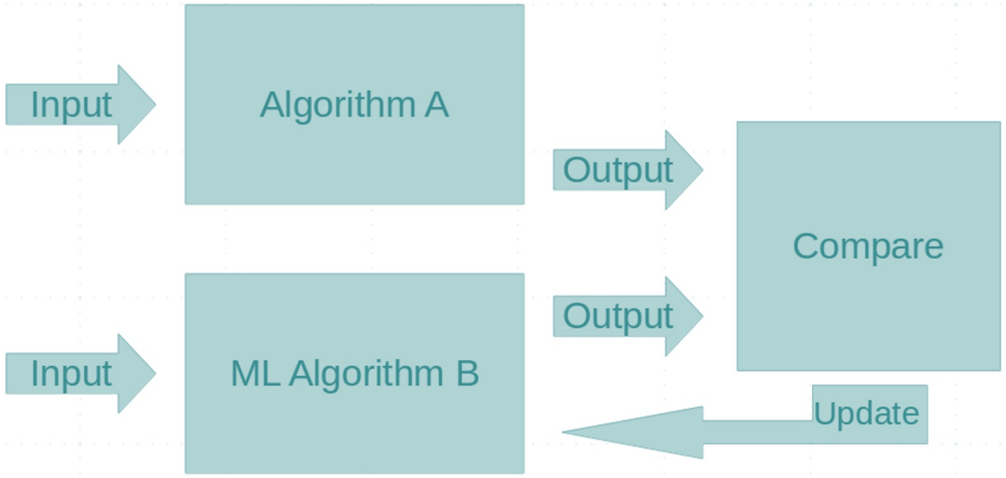
Motivation:

- Automatic labeling
- Model compression
- Surrogate networks for classical approaches
- Cross-modal knowledge distillation



How to transfer learned knowledge?

Knowledge Transfer



Example Contributions:

- Multi-Task Teacher-Student
- Teacher Student Networks with Multi-Scale training
- Knowledge Transfer between Neural Sparse Voxel Fields
- Knowledge Transfer from Classical Algorithms

Learnable Dublicate Removal using Graph Neural Network for Single Stage Object Detector

- Higher Average Precision and significant faster inference
- Ent-to-end training possible
- Last hand engineered component removed

SSD raw After pre-filtering Final detections

Diverse Topics in Didactics:

TP3 Topics	
Sensors	Camera
	Lidar
	Radar
	Stereo Cam
	Metadata
Task	Obj.-Detection
	Classification
	Segmentation
	Depth estimation
	Human Pose Estimation
	Trajectory estimation
	Trajectory Planning
	Anomaly Det.
	Instance Segmentation
	Sensorimotor Control
Datasets	nuScenes
	Cityscapes
	BDD100k
	Toyota DDAD
	KITTI
	SemanticKITTI
	Synthetic
	CARLA
	Multi dataset combination
	A2D2
	GTSDB
	Synthia-Seq
Domain Change	Domain Adaptation (generic)
	Task (e.g. new classes, different class specification)
	Image Level (e.g. neural style transfer)
	Representation Level (e.g. aligning distributions in deep layer)
	Catastrophic forgetting
	Synthetic
	Sensor
	Highres → Lowres
	Lowres → Highres
	Location and Time
	Environment
	New Output Domain
Supervision	Supervised
	Unsupervised
	Semi-supervised
	weakly-superv.
	Federated Learning
	Active Learning
	Self-Supervised
	Reinforcement Learning

Base Idea	GAN-based approach
	CNN
	3D CNN (e.g. VoxelNet)
	Graph Neural Networks
	One/ Few Shot
	(Variational) Autoencoder
	Bayesian approach for uncertainty estimation
	Teacher Student
	Online Learning
	Adversarial attacks
	Meta Learning
	Generation of Training Data
	Pre-Training Proxy/Pseudo Task
	Data Augmentation / Transformation
	Invariant / Shared Features
	Corner Cases / Anomaly Detection
	Hierarchical Data Compression
	Smarter ways and usage of data augmentation
	Information Theory (Shannon, KL, Wasserstein, ...)
	Single-Task Single-Teacher-Single-Student
	Multi-Task Teacher-Student
	Surrogate Networks
Target Hardware	NVIDIA Turing T4
	NVIDIA Jetson nano
	ZF ProAI
	Intel Movidius Myriad X
Quantization	8bit
	4bit
	binary
	other
Pruning	post processing
	simultaneous training and pruning
	manual pruning
	pruning automation
	NAS automation
	other
Tensor compression	heuristic rank selection
	algorithmic rank selection
	AI based rank selection
	other
Teacher→ Student	Large Network → Small Network
	Location Specific → Location Generalization
	Classical Algorithm → ML-Algorithm
	Ensemble or Multiscale → Single Network

KI DELTA LEARNING

Scalable AI for Automated Driving

Marius Bachhofer | ZF Friedrichshafen AG | marius.bachhofer@zf.com

KI Delta Learning is a project of the KI Familie. It was initiated and developed by the VDA Leitinitiative autonomous and connected driving and is funded by the Federal Ministry for Economic Affairs and Climate Action.

KI FAMILIE

Supported by:

Federal Ministry for Economic Affairs and Climate Action

on the basis of a decision by the German Bundestag

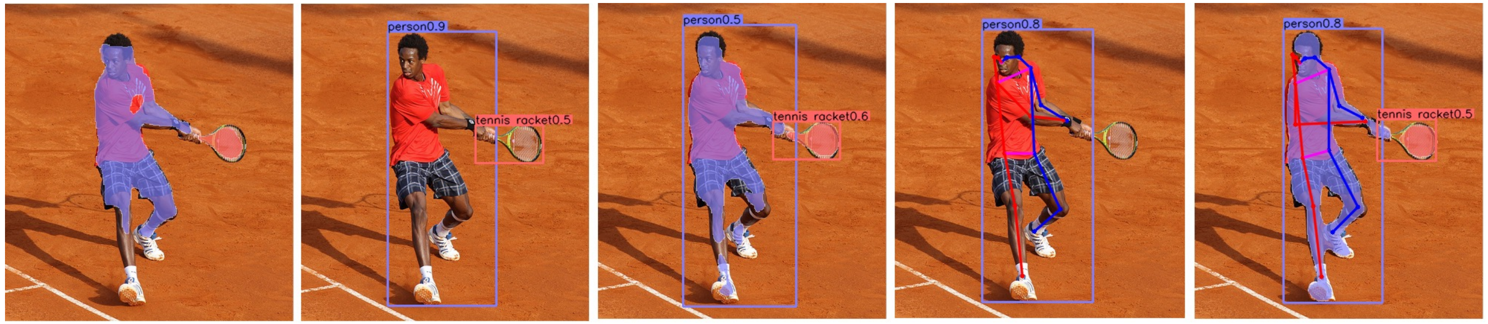
www.ki-deltalearning.de @KI_Familie KI Familie

KIDELTA LEARNING

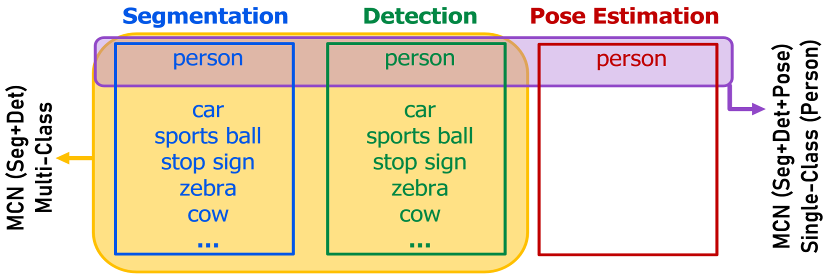
Scalable AI for Automated Driving

Pitch

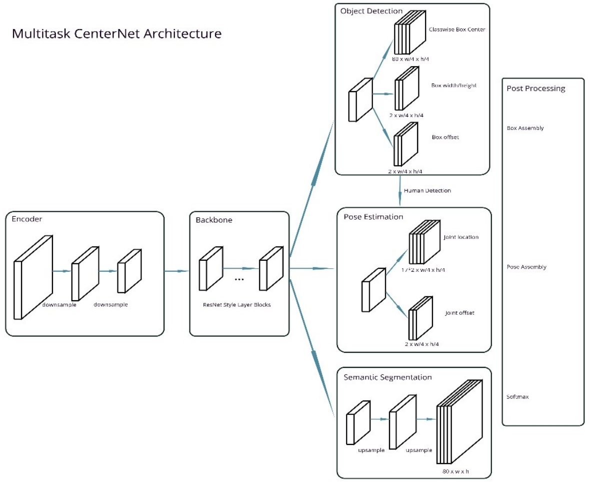
Shared Backbones with MultiTask CenterNet



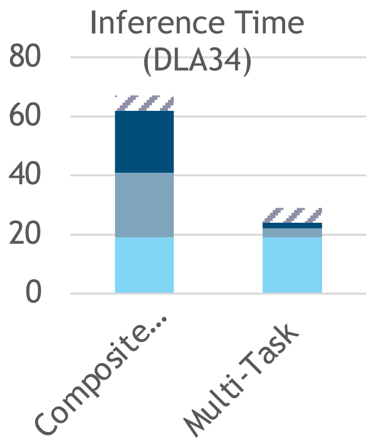
Network setups trained with various amounts of heads for the tasks segmentation, detection and human pose estimation.



Tasks and classes used in training setups.



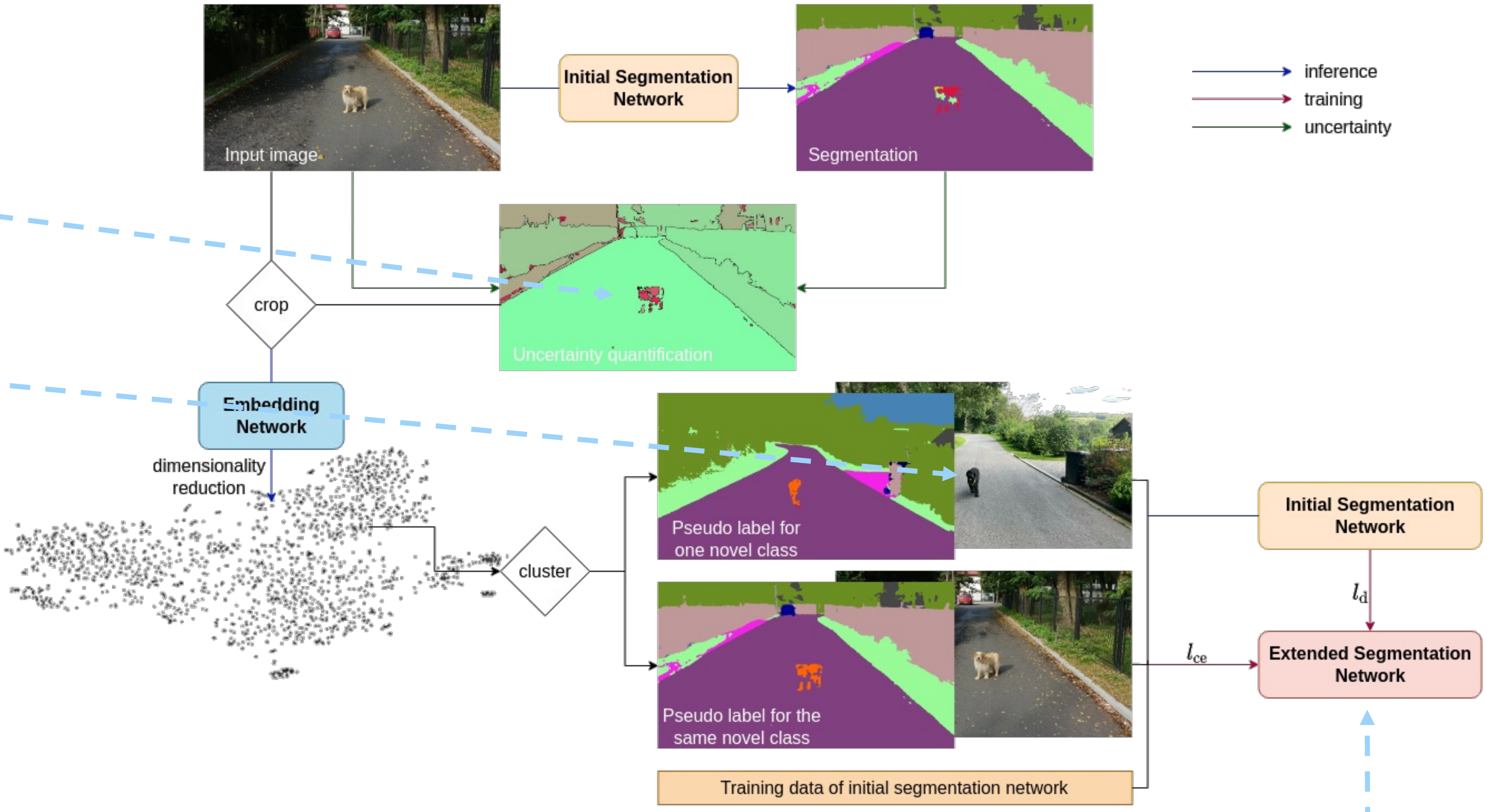
- Solve **several vision tasks** at once
- **Latency** is vastly reduced while performance stays the same or even gets exceeded
- **Robustness** of automotive perception systems is increased with more tasks



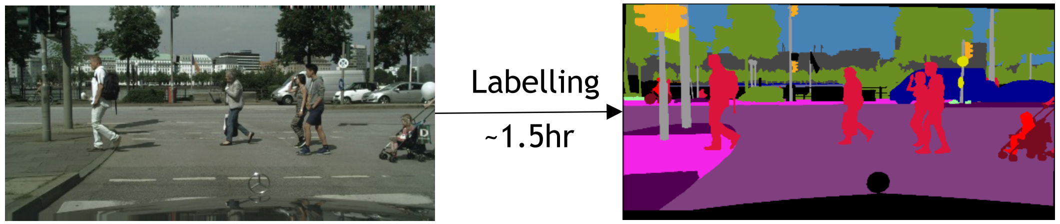
Latency comparison

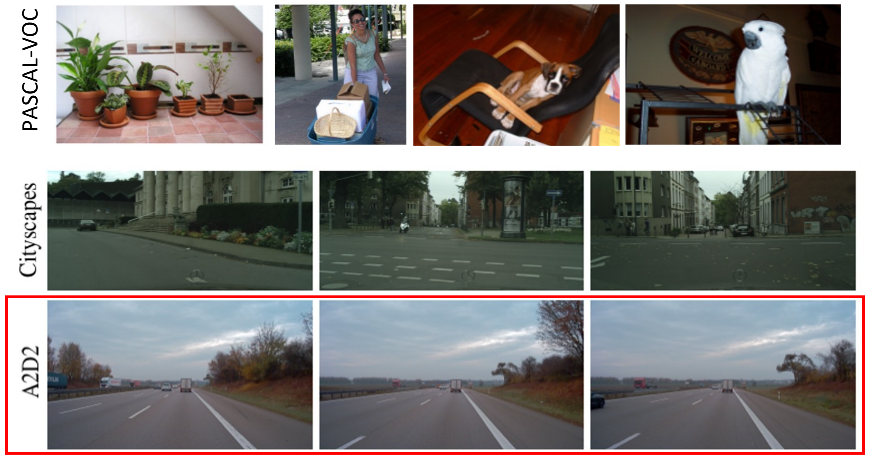
How Can Neural Networks Discover the World?

- What is unknown?
- Have I seen something similar before?
- Give it a name and recognize it from now on.



Problem: Which is the best active learning method for a dataset?

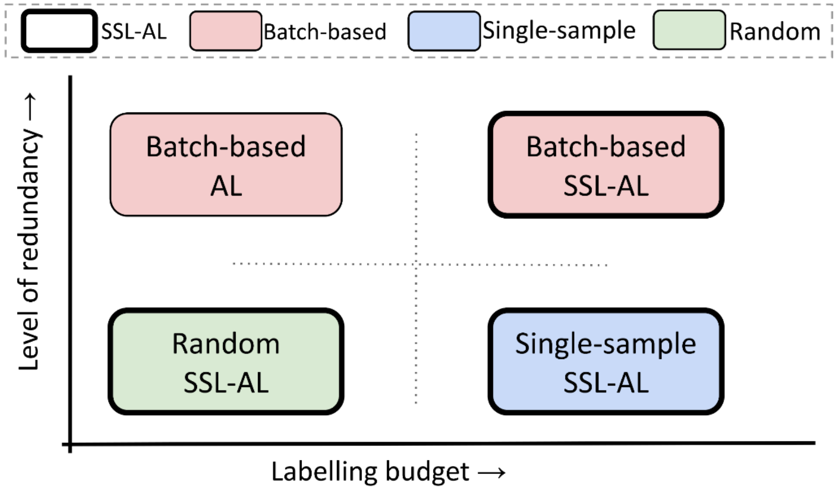




- Current benchmarks focus on diverse data
- Real data is very redundant

Key findings

- *Diverse datasets* → Single-sample AL method
- *Redundant datasets* → Batch-based AL method
- SSL integrates well with batch-based AL method



➔ Our findings apply to realistic datasets

MGiaD: Multigrid in all Dimensions

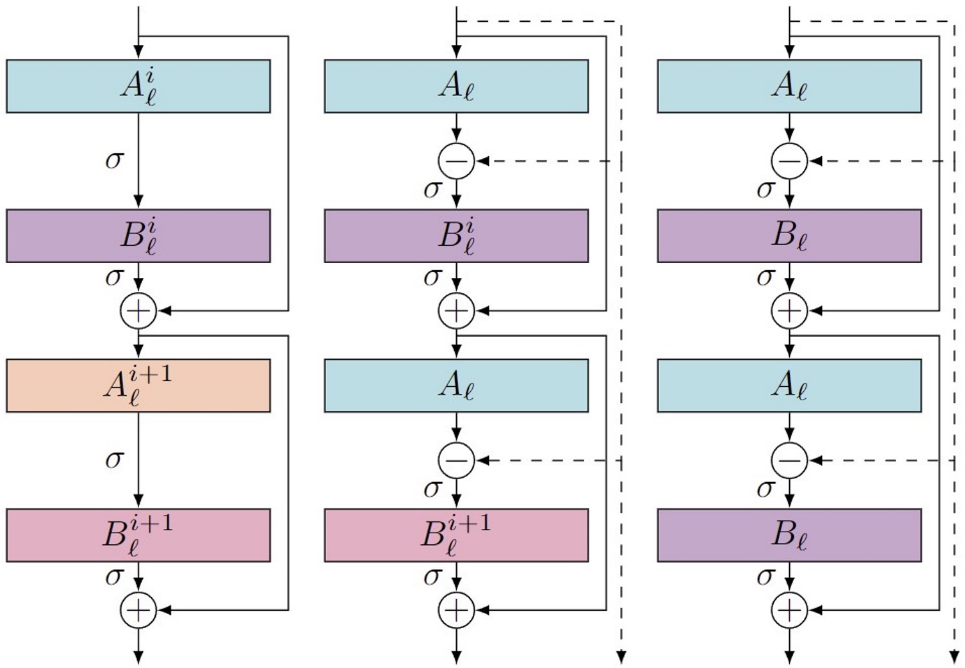
Problem: Overparameterization (inefficient use of weights), e.g. hard hardware limitations

Goal: Efficient and robust convolutional neural networks (CNNs) architectures:
less weights and maintain accuracy

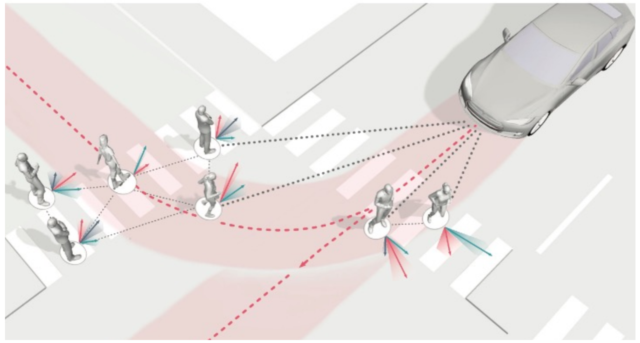
Approach:

- Reusing weight tensors
- Hierarchy of grouped convolutions
- Mathematically motivated by iterative methods

➡ Improved parameter-accuracy trade-off

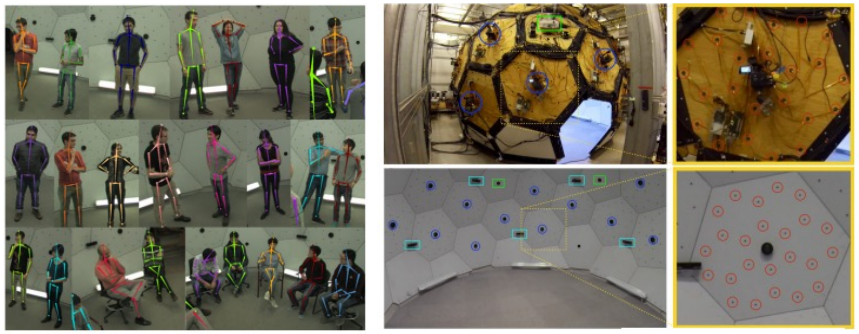


5 Unsupervised Learning of 3D Human Body Pose



3D Trajectory Prediction
[Ivanovic et al., ICRA2020]

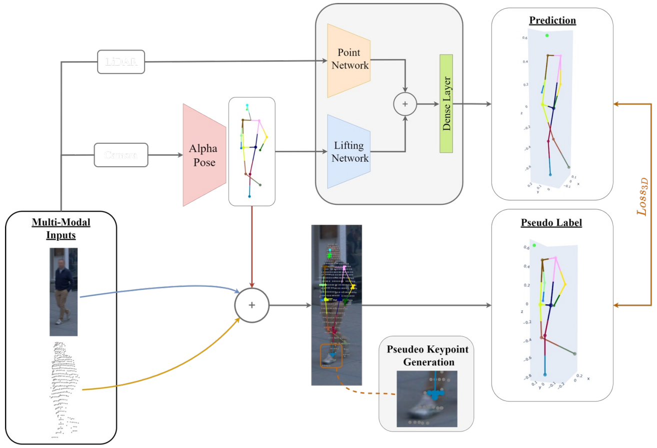
Action & Gesture Anticipation
[Wiederer et al., IROS2020]



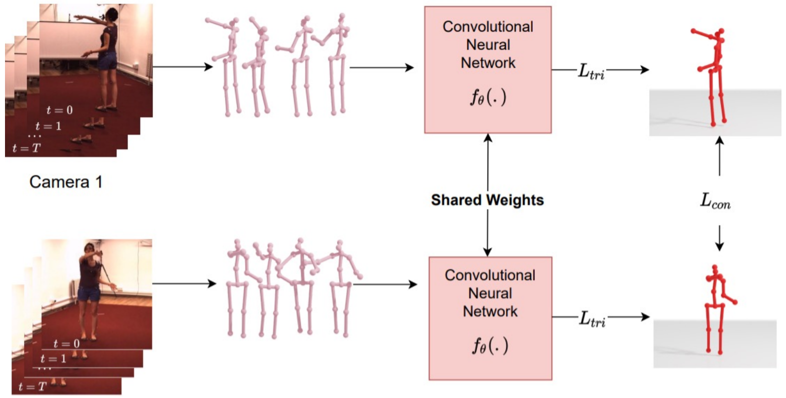
Motion Capture Systems
[Joo et al., TPAMI2016]

2D GT, 3D GT, Indoor

Unsupervised, real-world



LIDAR Supervision



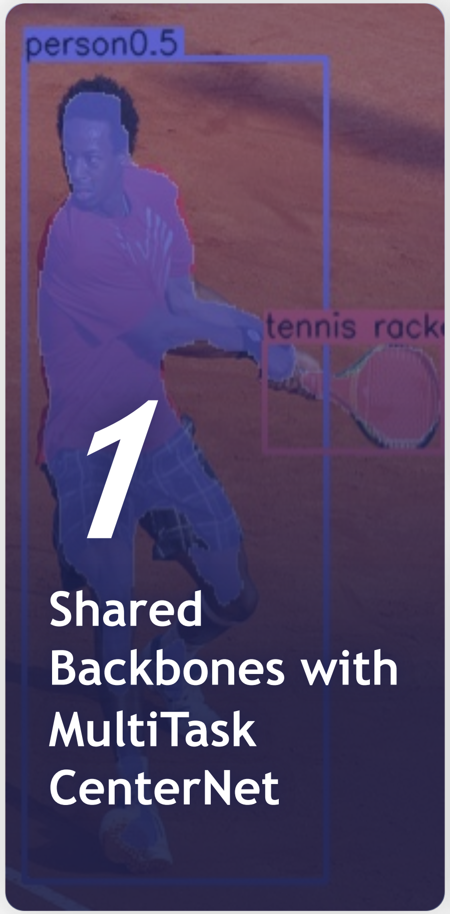
Multi-view Supervision

KI DELTA LEARNING

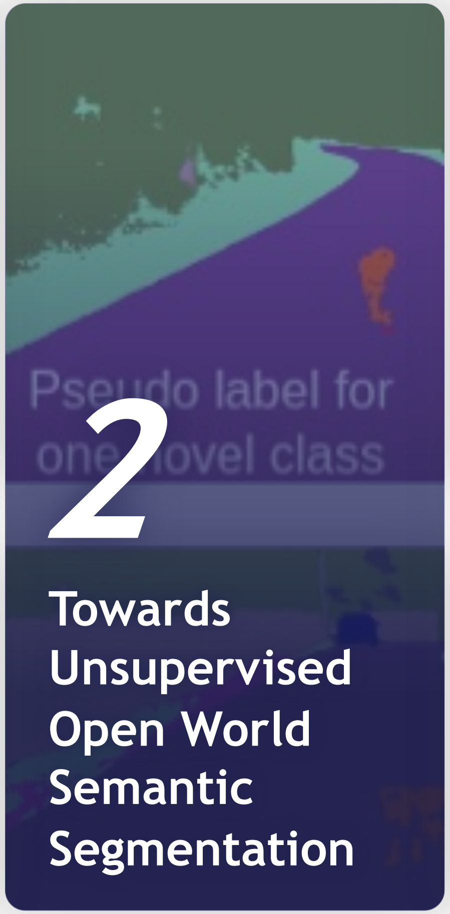
Scalable AI for Automated Driving

Vote

Voting

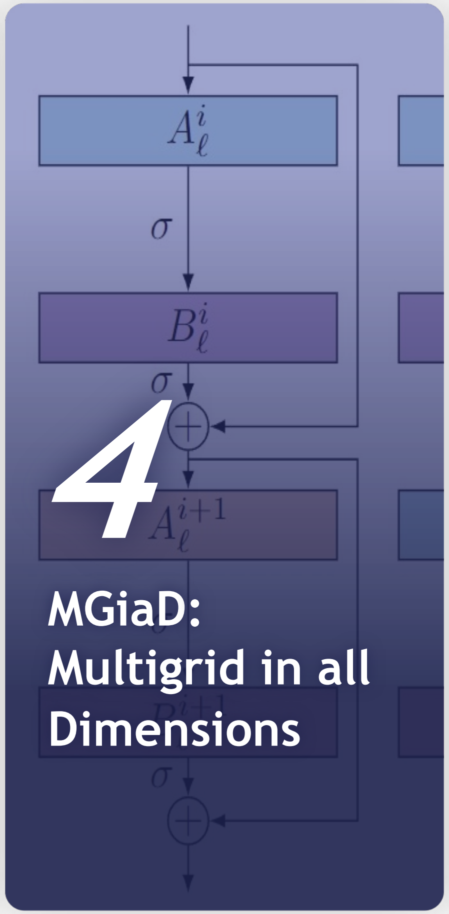


1 Shared Backbones with MultiTask CenterNet



2 Towards Unsupervised Open World Semantic Segmentation

3 Active Learning for Semantic Segmentation in Realistic Driving Scenarios



4 MGiaD: Multigrid in all Dimensions

5 Unsupervised Learning of 3D Human Body Pose