visuAAL

Privacy-Aware and Acceptable Video-Based Technologies and Services for Active and Assisted Living

ESR12. Measuring dementia behaviours through depth sensors

Joint visuAAL-GoodBrother conference

Alicante, Spain

18.06.2024

Irene Ballester Campos

Computer Vision Lab, TU Wien

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 861091".

Universitat d'Alacant
Universidad de Alicante
Project Coordinator

RWTH AACHEN UNIVERSITY

Stockholm University

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

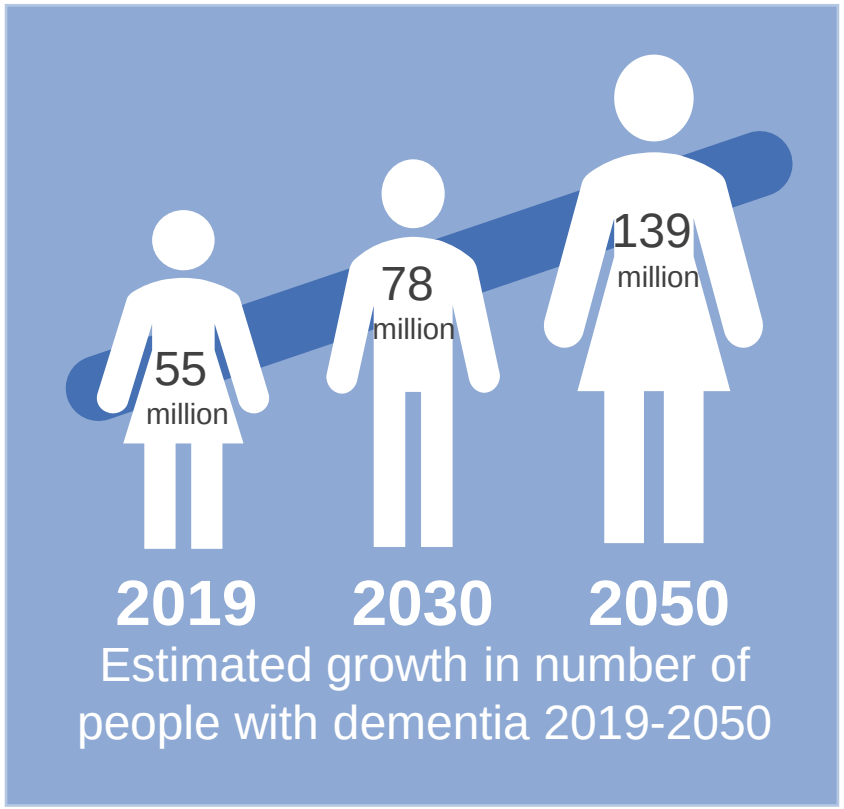
TU WIEN

What's dementia?

- Syndrome in which there is a **deterioration in cognitive functioning beyond what might be expected from normal ageing** [1]
- One of the **major causes of dependency** among older people [2]

Why dementia?

- Behavioural changes **strongly correlated** with the degree of functional and cognitive impairment [2].
- **Behavioral and Psychological Symptoms of Dementia (BPSD):** agitation, aberrant motor behaviour, anxiety, irritability, depression, apathy, delusions, changes in sleep or appetite [3].



Data source: WHO [1]

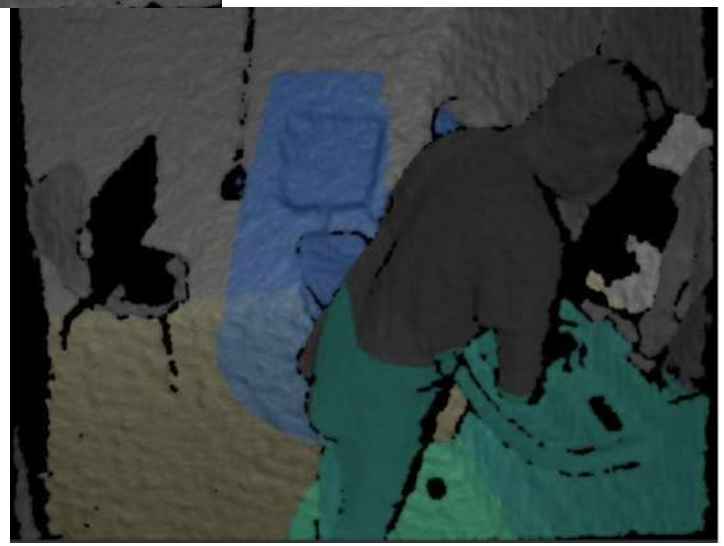
[1] World Health Organization https://www.who.int/news-room/fact-sheets/detail/dementia (accessed April 28, 2024)
[2] **Global status report on the public health response to dementia**. World Health Organization (2021)
[3] Joaquim Cerejeira, Luısa Lagarto, and Elizabeta Blagoja Mukaetova-Ladinska. "Behavioral and psychological symptoms of dementia". In: *Frontiers in neurology 3 (2012*), p. 73.

Data modality: depth

One of the main concerns: PRIVACY

RGB

Depth



AI for behaviour analysis from unobtrusive sensor data

Goal: Development of AI methods for **measuring** the **behaviours** of care home residents with **dementia** using **unobtrusive sensors (depth maps)**

In order to:

1. **Unobtrusive Remote Patient Monitoring**
2. **Provide assistance with ADLs**

AI for behaviour analysis from unobtrusive sensor data

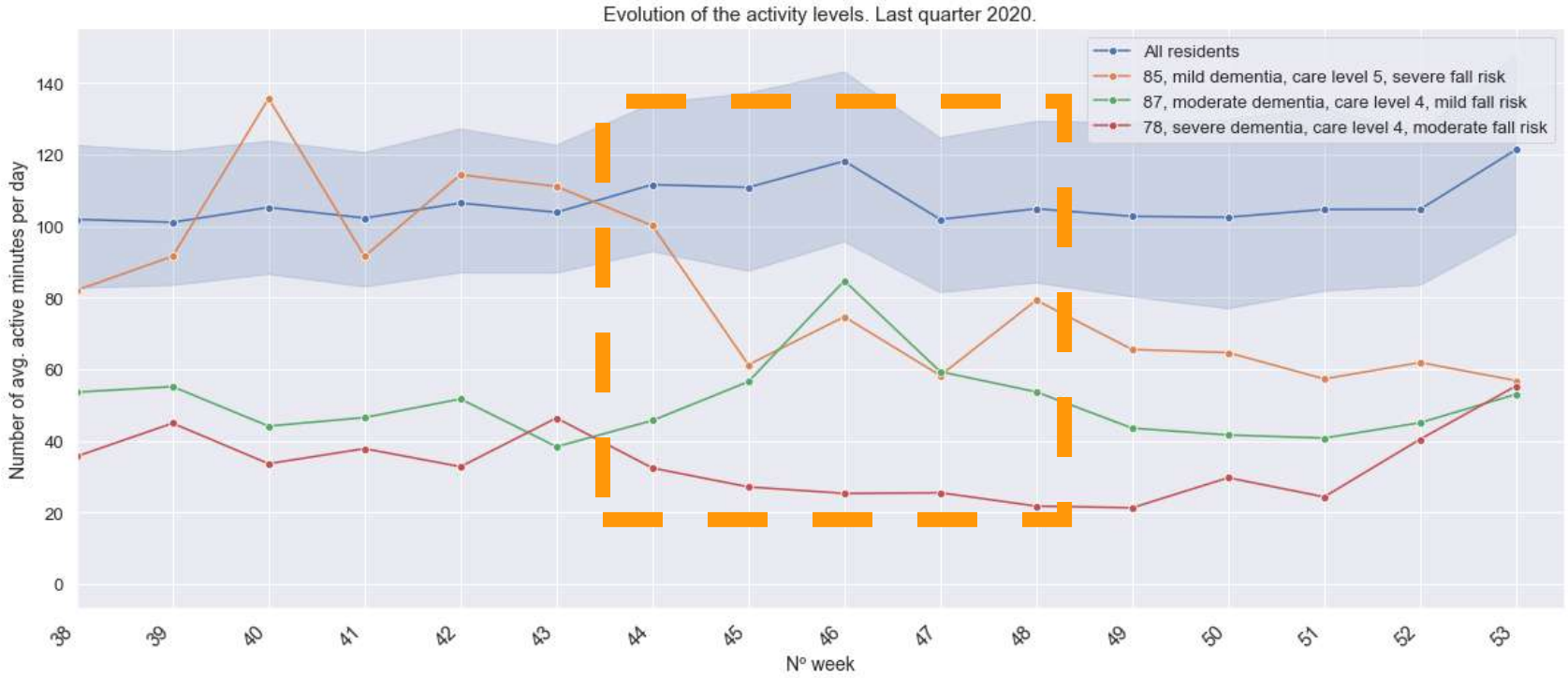
Goal: Development of AI methods for **measuring** the **behaviours** of care home residents with **dementia** using **unobtrusive sensors (depth maps)**

In order to:

1. **Unobtrusive Remote Patient Monitoring**
2. **Provide assistance with ADLs**

Ultimate goal

Detect and measure functional and behavioural changes indicative of dementia



Research Questions

RQ1. Different inputs for behaviour measurement

RQ2. Robust performance for real-world HAR

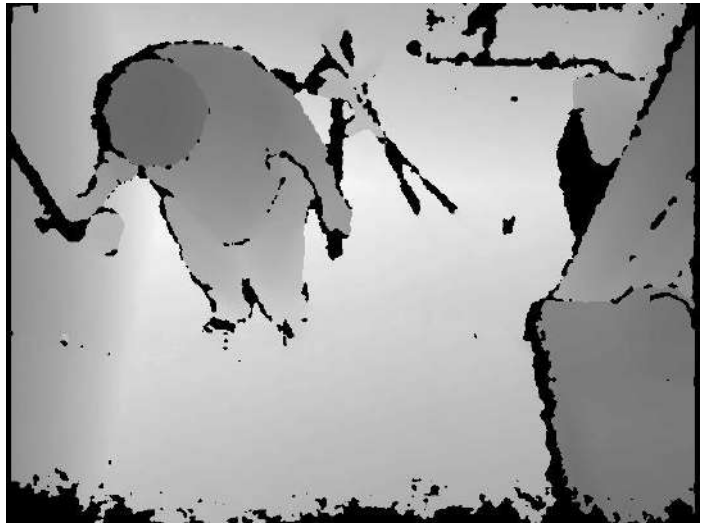
Research Questions

RQ1. Different inputs for behaviour measurement

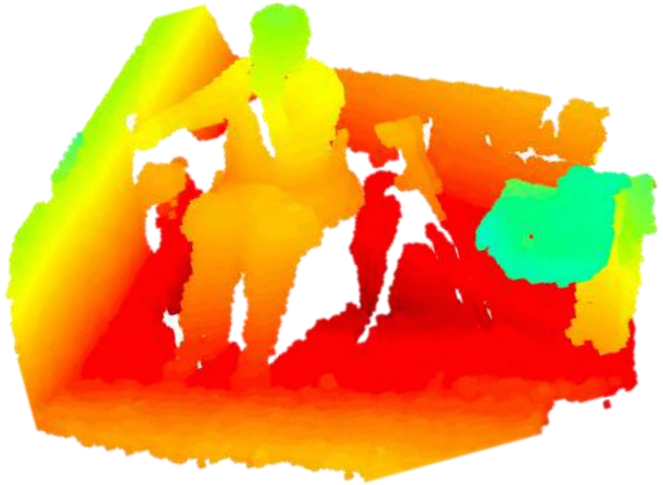
RQ2. Robust performance for real-world HAR

Different inputs

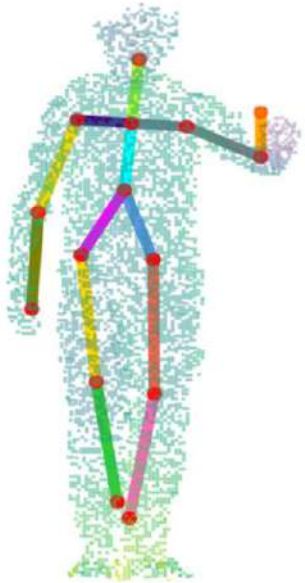
(Raw) depth images



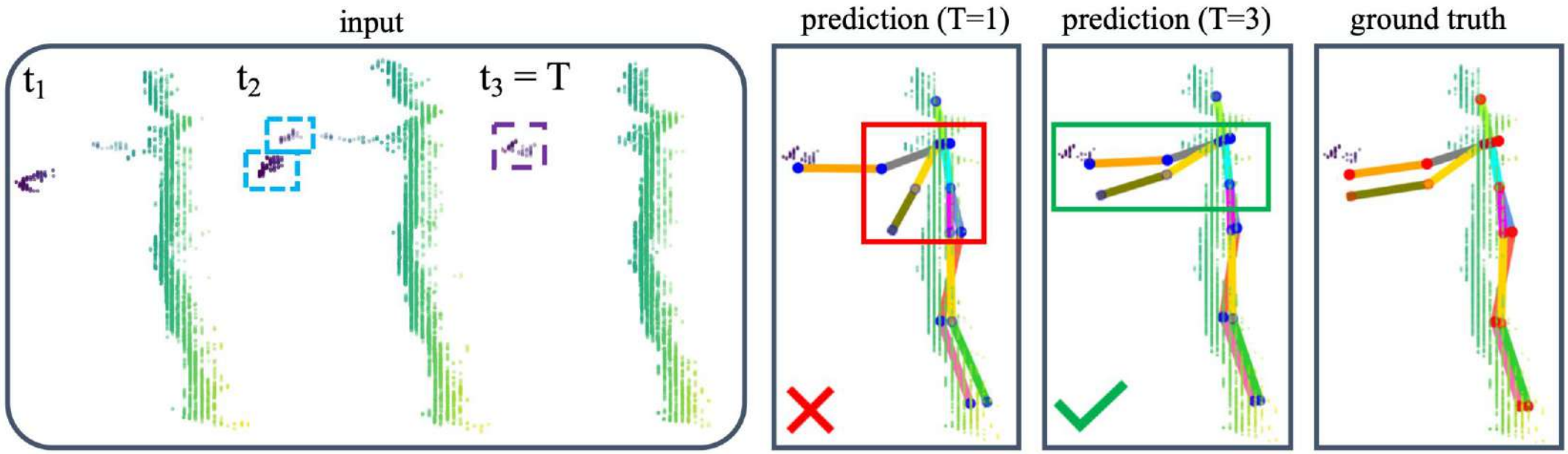
Point clouds from depth



Skeletons from depth/point clouds

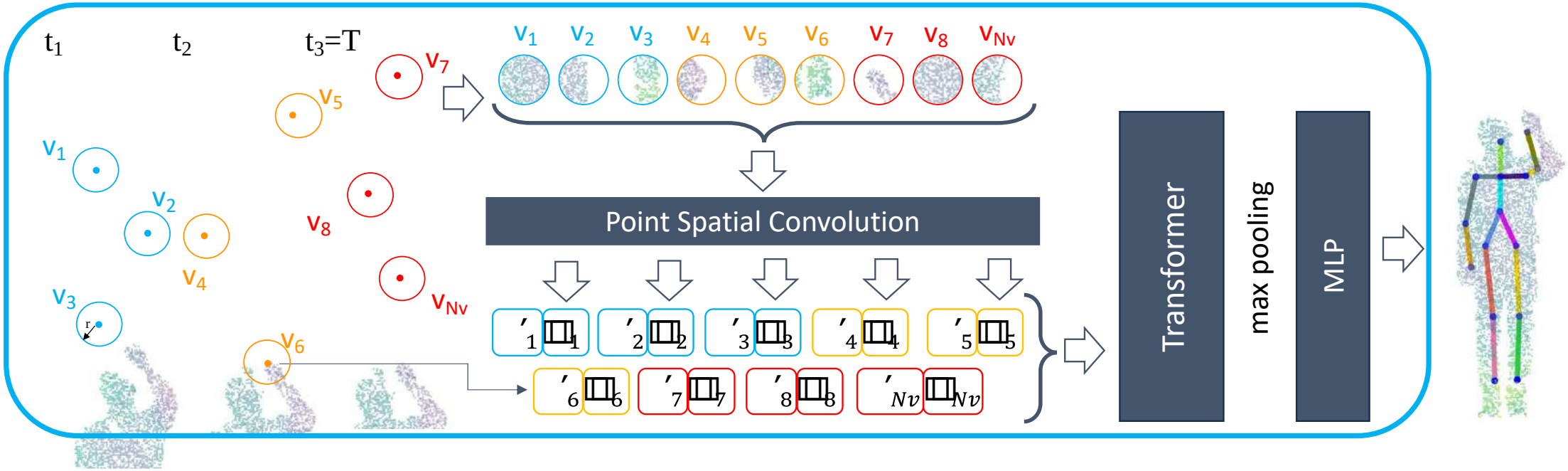


Motivation: Sequence information helps with occlusions and noise



SPiKE: Method

- How to process sequences of point clouds?
  - Spatial point convolution in local areas
  - Transformer + MLP for prediction of 3D joints

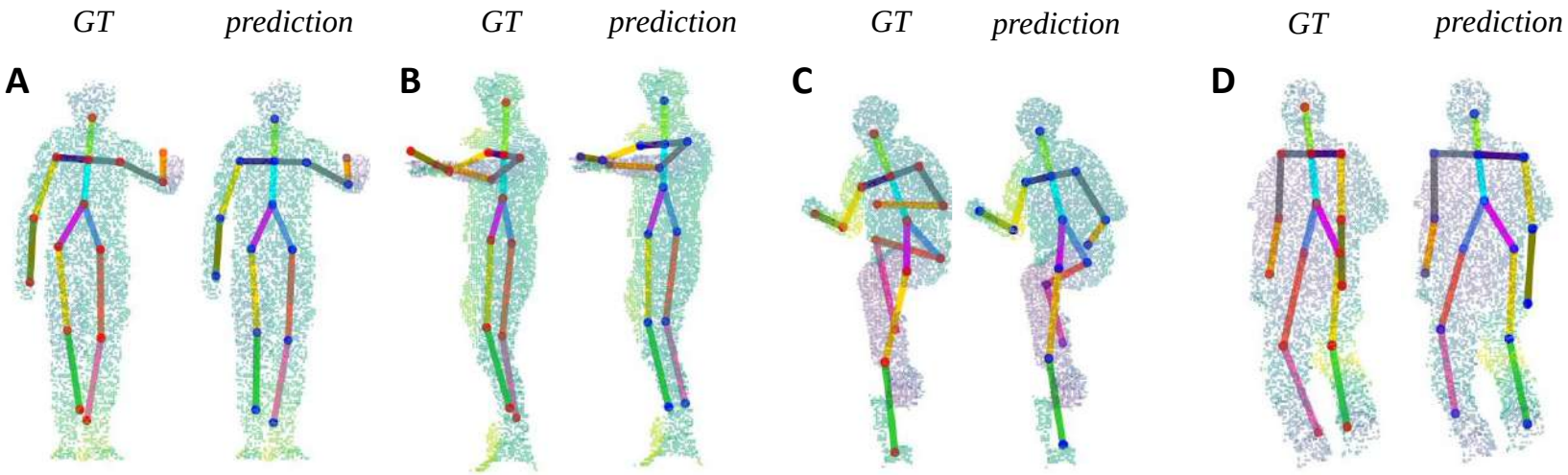


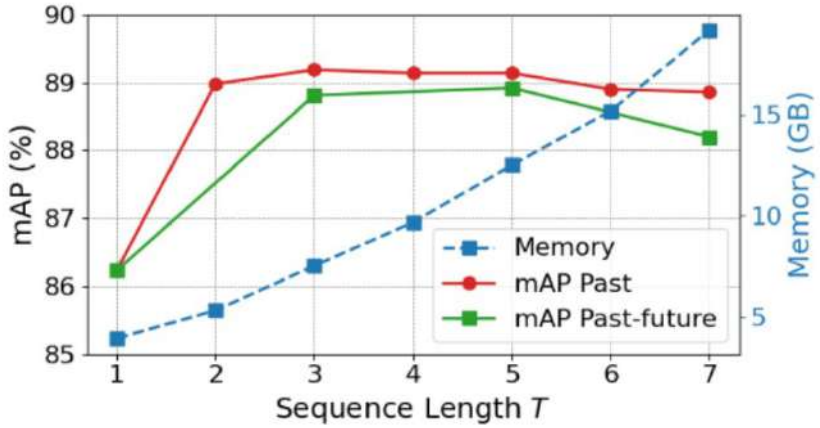
SPiKE: Comparison with SOTA

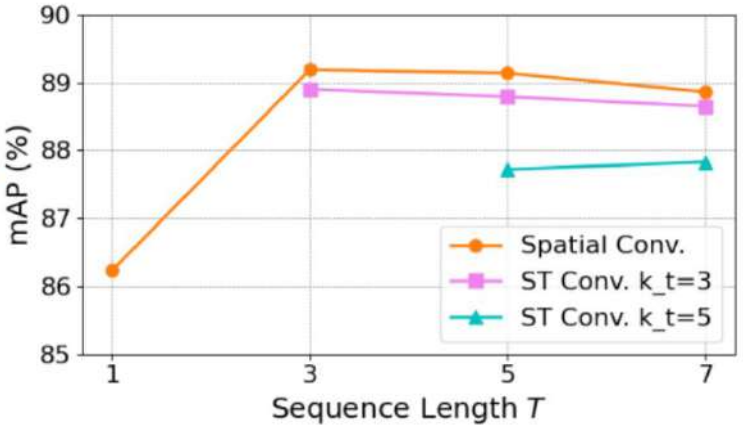
Table 1. Comparison with the state-of-the-art for ITOP front-view (0.1m mAP)

Method	Single Modality Methods						Multimodal Methods		
	V2V	A2J	WSM*	Zhou et al.	DECA	SPiKE (Ours)	WSM	AdaPose	HRNet+ RefiNet
	2018	2019	2020	2020	2021	-	2020	2021	2023
Modality	voxels	depth	points	points	depth	points	depth+points		
Head	98.29	98.54	-	96.73	93.87	98.42	98.15	98.42	-
Neck	99.07	99.20	-	98.05	97.90	99.47	99.47	98.67	-
Shoulders	97.18	96.23	-	94.38	95.22	97.48	94.69	95.39	-
Elbows	80.42	78.92	-	73.67	84.53	81.64	82.80	90.74	-
Hands	67.26	68.35	-	54.95	56.49	71.71	69.10	82.15	-
Torso	98.73	98.52	-	98.35	99.04	99.24	99.67	99.71	-
Hips	93.23	90.85	-	91.77	97.42	93.68	95.71	96.43	-
Knee	91.80	90.75	-	90.74	94.56	91.56	91.00	94.41	-
Feet	87.60	86.91	-	86.30	92.04	84.30	89.96	92.84	-
Upper B.	-	-	-	80.10	83.03	88.75	-	-	80.8
Lower B.	-	-	-	89.60	95.30	89.85	-	-	88.1
Mean	88.74	88.00	75.64*	85.11	88.75	**89.19**	89.59	**93.38**	84.2

SPiKE: Qualitative results and ablations



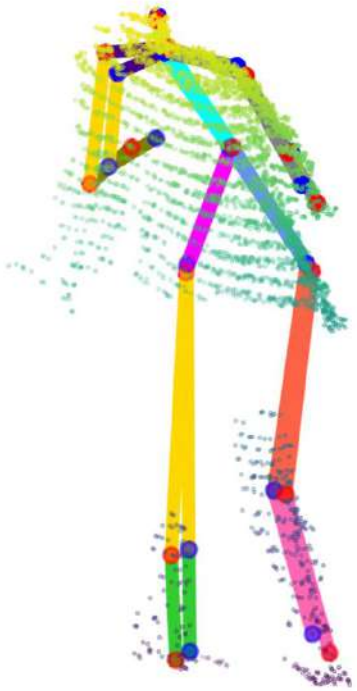




SPiKE: Conclusions and future steps

1. **Sequence information is useful**, but up to a certain sequence length
2. Spatial Convolutions help to preserve spatial structure (useful for HPE)
3. **SOTA in ITOP**

- Future work:
  How does SPiKE perform with **real-world data**?



Research Questions

RQ1. Different inputs for behaviour measurement

RQ2. Robust performance for real-world HAR

Domain gaps

vs

Benchmarks

Real-world data

2.5D
(self)-occlusions
complex, natural activities
unbalanced classes
frontal vs. tilted angles
sparsity and noise

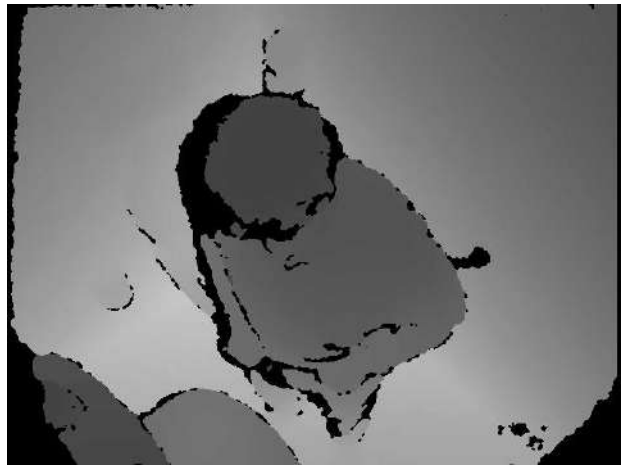
Bathroom Activities Dataset (BAD)

Description

- 50 full sequences
- **19 subjects with dementia** using the toilet
- 8 classes: walking around, undressing, sitting down, sitting on the toilet, standing up, dressing, washing hands
- 2 different locations:
  - **BAD1**: 3 subjects (36k frames)
  - **BAD2:** 16 subjects (21k frames)
- **Unbalanced** dataset:
  E.g.: in BAD2: sitting: 8k frames vs. sitting down: 679 frames

BAD1 – 3 subjects

BAD2 – 16 subjects



visuAAL

Domain Gaps

Performance gap between benchmarks and real-world datasets

Performance gap between real-world scenes

Human Activity Recognition				
P4T [2]	Tested on (acc.)			
Trained on	*MSRAction3D*	*NTU 60 (cs)*	*BAD1 (cs)*	*BAD2 (cs)*
MSRAction3D	90.94	-	-	-
NTU 60 (cs)	-	90.2	-	-
BAD1 (cs)	-	-	53.15	7.23
BAD2 (cs)	-	-	7.90	63.63

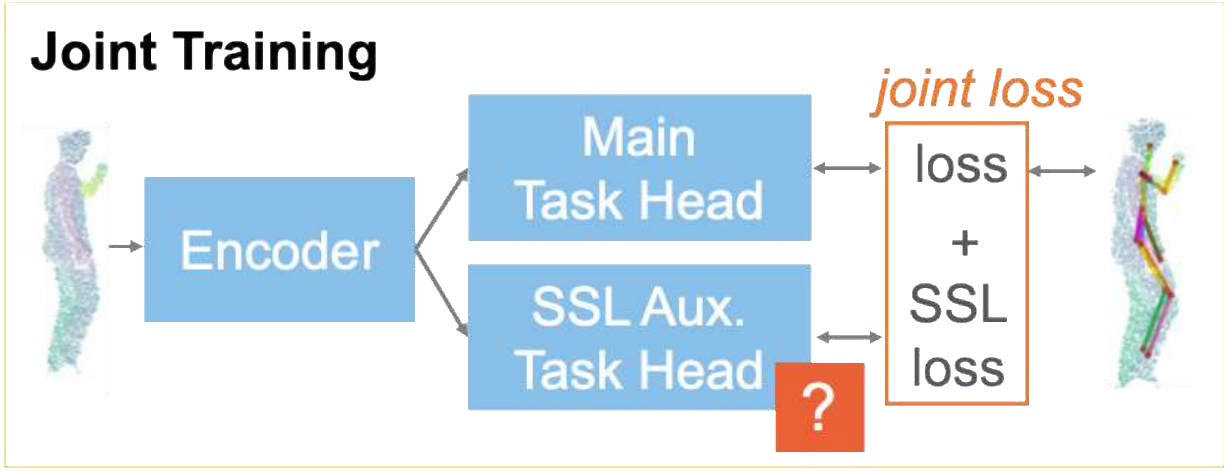
Human Pose Estimation		
SPiKE	Tested on (mAP@0.1m)	
Trained on	*ITOP-SIDE*	*ITOP-TOP*
ITOP-SIDE	89.19	24.75*
ITOP-TOP	36.81*	81.58*

Performance gap between views

**Work-in-progress*

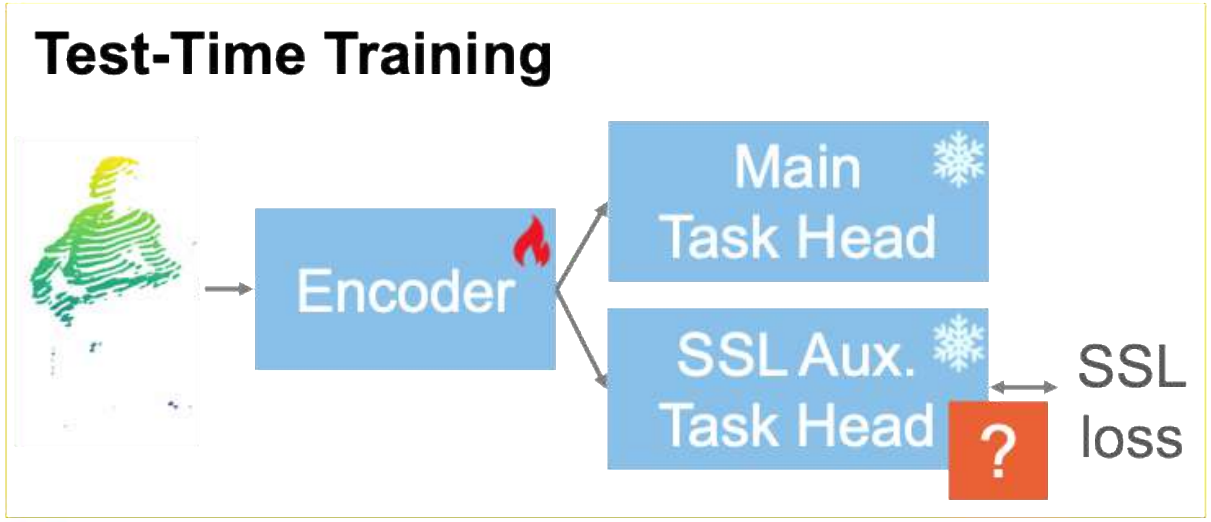
[2] Fan, H., Yang, Y., & Kankanhalli, M. (2021). Point 4d transformer networks for spatio-temporal modeling in point cloud videos. CVPR 2021

Test-Time Training for Domain Adaptation in Point Cloud Sequences

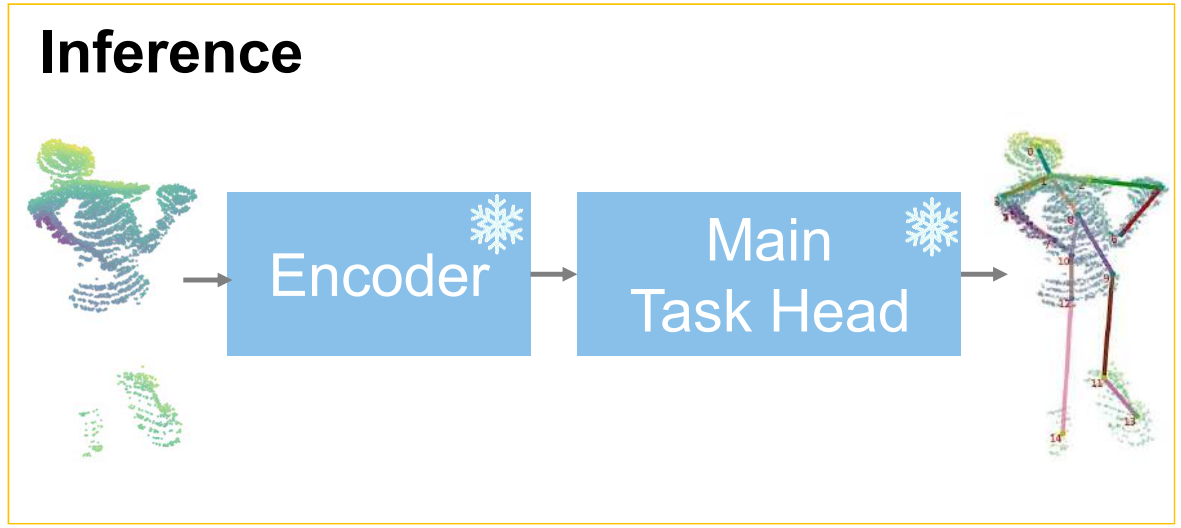


Source domain

Target domain



Target domain



Conclusions

- **Take away message**
  - Very promising achievements in benchmarks, how do they behave with real-world data?
  - **Need for evaluation in real-world scenarios and (potentially) domain adaptation for AAL applications**

- Future work:
  - Pre-training strategies
  - Self-supervised learning
  - Continual learning

Thank you!

Irene Ballester Campos

TU Wien

irene.ballester@tuwien.ac.at

Measuring dementia behaviours through depth sensors

Irene Ballester Campos
irene.ballester@tuwien.ac.at

visuAAL

Innovative Training Network on Privacy-Aware and Acceptable Video-Based Technologies and Services for Active and Assisted Living

@visuAAL_ITN

https://www.visuaal-itn.eu/

This project has received funding from the European Union's Horizon 2020 research and Innovation programme under the Marie Skłodowska-Curie grant agreement No 861091 and from the WWTF under the project number ICT20-055.