AUTHOR

Kevin Hoxha
K.Hoxha@student.tudelft.nl

SUPERVISORS

***Supervisor**: Girish Vaidya*
***Responsible Professor**: Marco Zuñiga Zamalloa*

Exploring the Spatial Characteristics of MARS

Assessing the Impact of Neural Net Depth Increase and PointNet Architecture Integration on MARS Performance

AFFILIATIONS

Delft University of Technology

1. Introduction

- In modern workplaces, screen exposure to colleagues poses significant privacy risks.
- VideowindoW's transparent screens, coupled with advanced pose reconstruction using mmWave radars [1], offer promise in addressing these challenges.

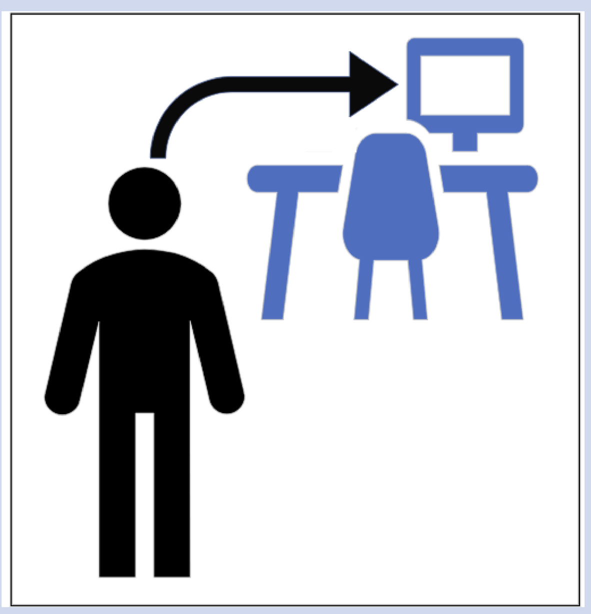

Figure 1: Illustration of screen exposure in the workplace

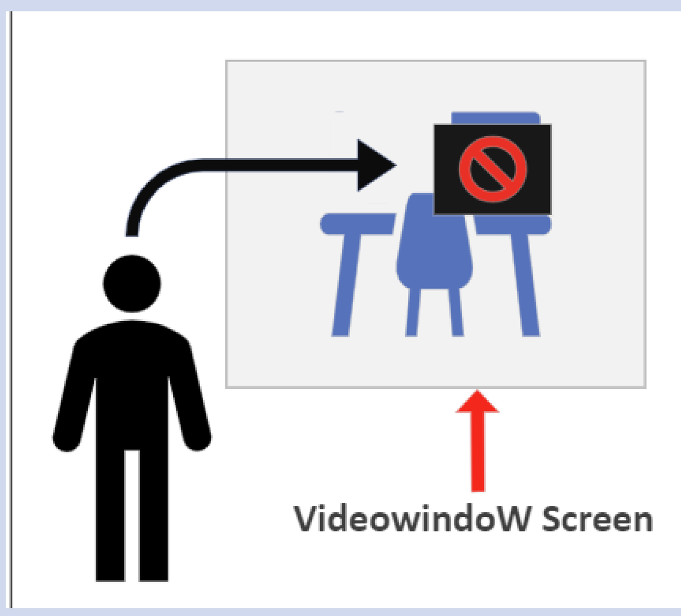

Figure 2: Illustration of VideowindoW's partial screen blocking

- Two state-of-the-art solutions for mmWave-based pose reconstruction:

1. **MARS** [2] is a low-cost, accurate 3D pose reconstruction system that estimates the location of 19 key joints of the human body with a unique movement dataset, as shown in **Figure 3**.
2. **PointNet** [3] is a neural network architecture tailored for processing raw point clouds to classify and segment objects.

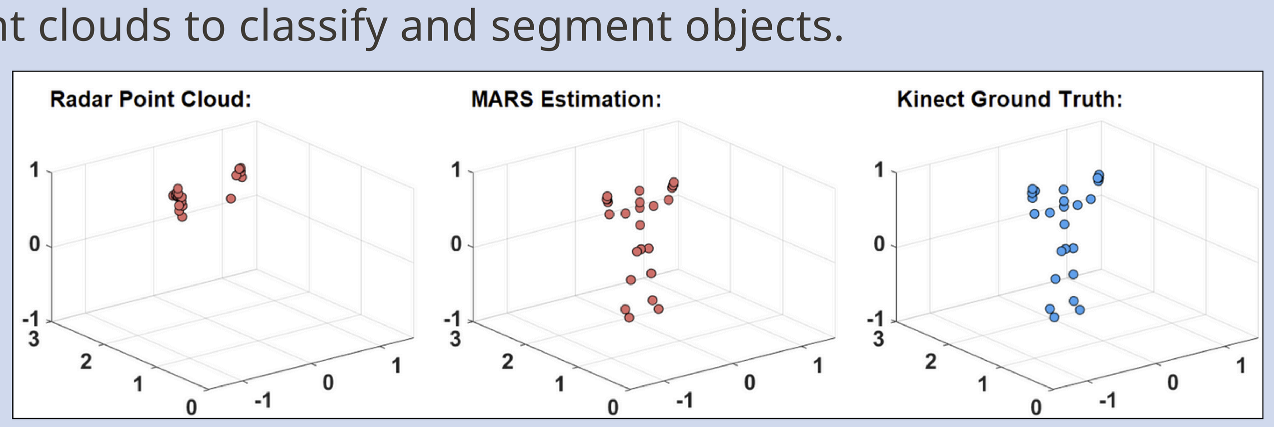

Figure 3: MARS 3D human pose estimation through using mmWave radar point cloud

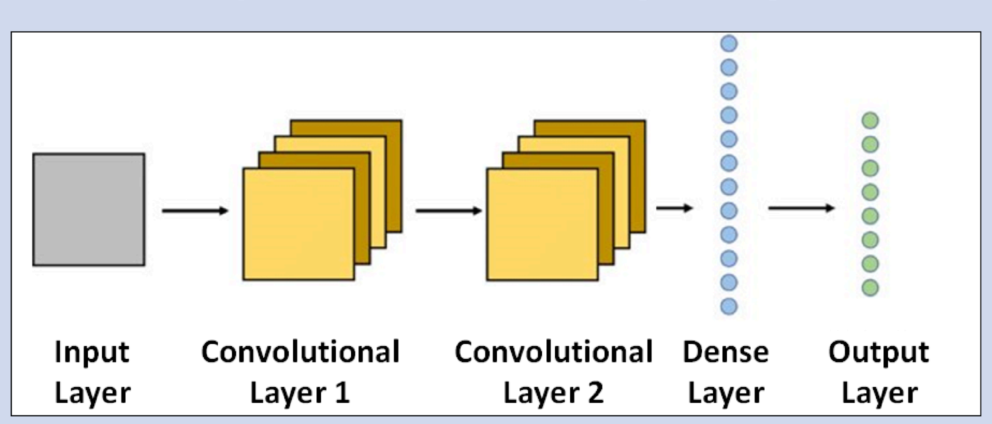

Figure 4: The convolutional neural network in the MARS architecture

2. Objective

- **Subquestion 1:** What is the performance impact of increasing the depth of the CNN in the MARS architecture?
- **Subquestion 2:** What is the performance impact of integrating PointNet instead of using the current MARS architecture?

3. Methodology

Subquestion 1

- Models with 1, 2, 3, and 4 convolutional layers are trained 10 times each with 150 epochs.
- All model parameters and batch size are identical to the MARS paper [2].
- The performance metrics are averaged and plotted.
- The same procedure is conducted to determine the optimal number of dense layers.

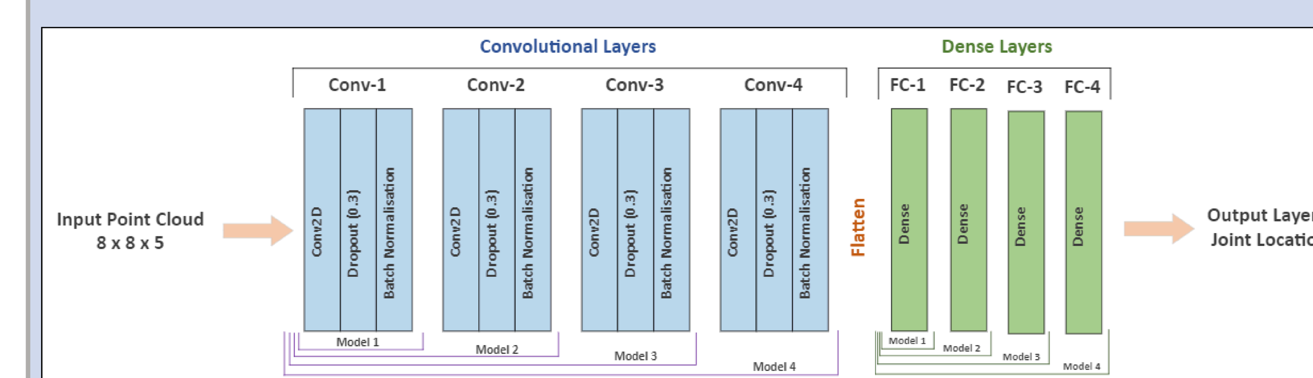

Figure 5: Architecture of the MARS models with increased number of layers

Subquestion 2

- PointNet is adapted for regression.
- The MARS dataset point clouds are preprocessed to fit PointNet.
- The CNN in MARS is replaced with PointNet.
- The PointNet-integrated model is trained, evaluated, and compared with the results from the MARS paper [2].

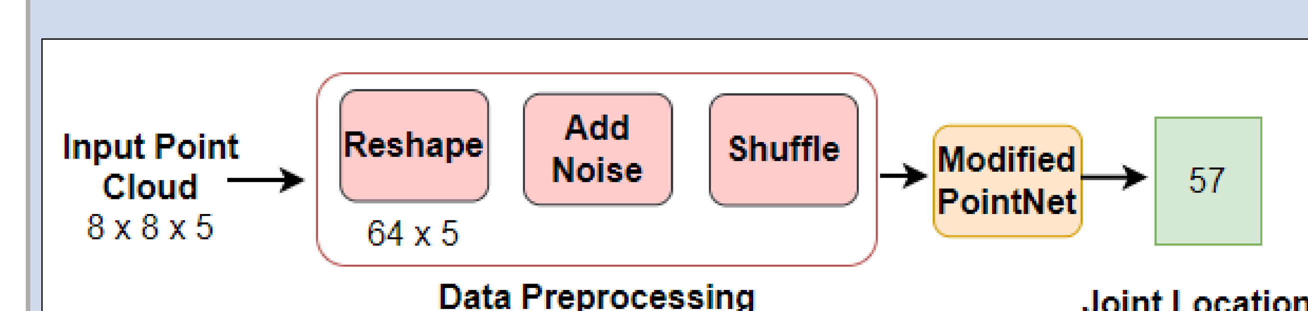

Figure 6: MARS architecture after PointNet integration

4. Experiment Results

Subquestion 1:

- The optimal number of convolutional layers is **2**, since it produces the lowest MAE and RMSE, as can be seen in Figure 7. Models with more layers overfit the training data, while models with fewer layers underfit the data.
- The optimal number of dense layers is **2**, since it produces the lowest MAE and RMSE, as seen in Figure 8. Models with more layers overfit the training data and introduce unnecessary complexity in the model and in the training time.
- This setup modestly improves the **MAE** by 3.2 % and the **RMSE** by 0.09 % compared to the default MARS configuration.

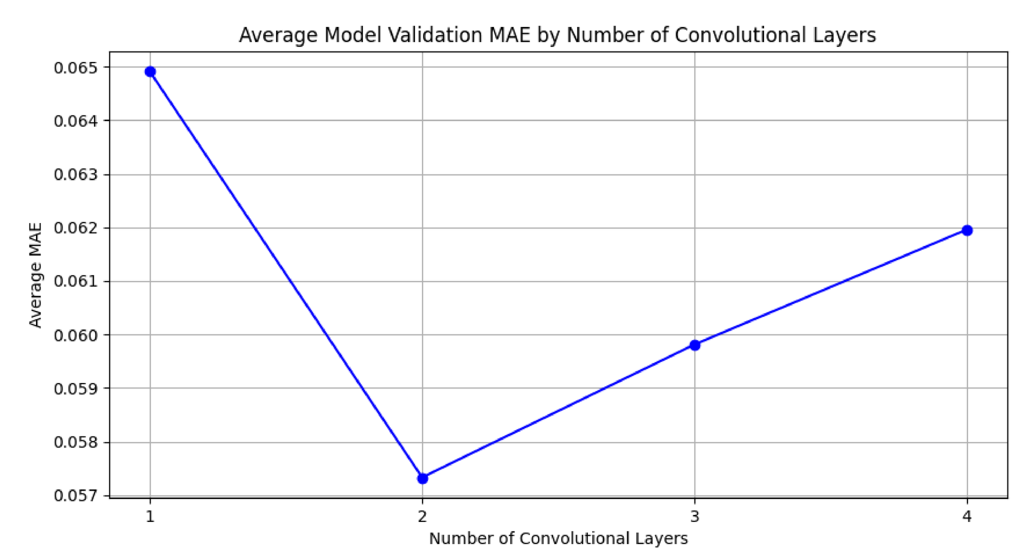

Figure 7: Validation MAE across convolutional layers

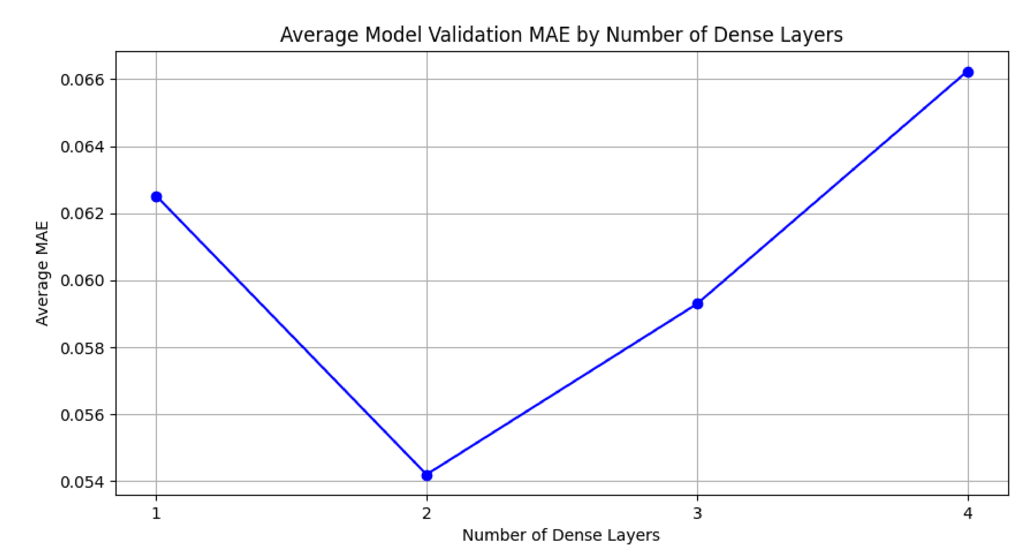

Figure 8: Validation MAE across dense layers

Subquestion 2:

- **Effectiveness of PointNet**: PointNet performs 32.4% worse on the **MAE** and 27.9% worse on the **RMSE.**
- **Limitation of PointNet**: PointNet for regression tasks does not capture local structural details.
- **Comparison with CNNs**: Research also indicates that traditional CNN architectures, like those in MARS, generally perform better than PointNet for human pose estimation tasks [4].

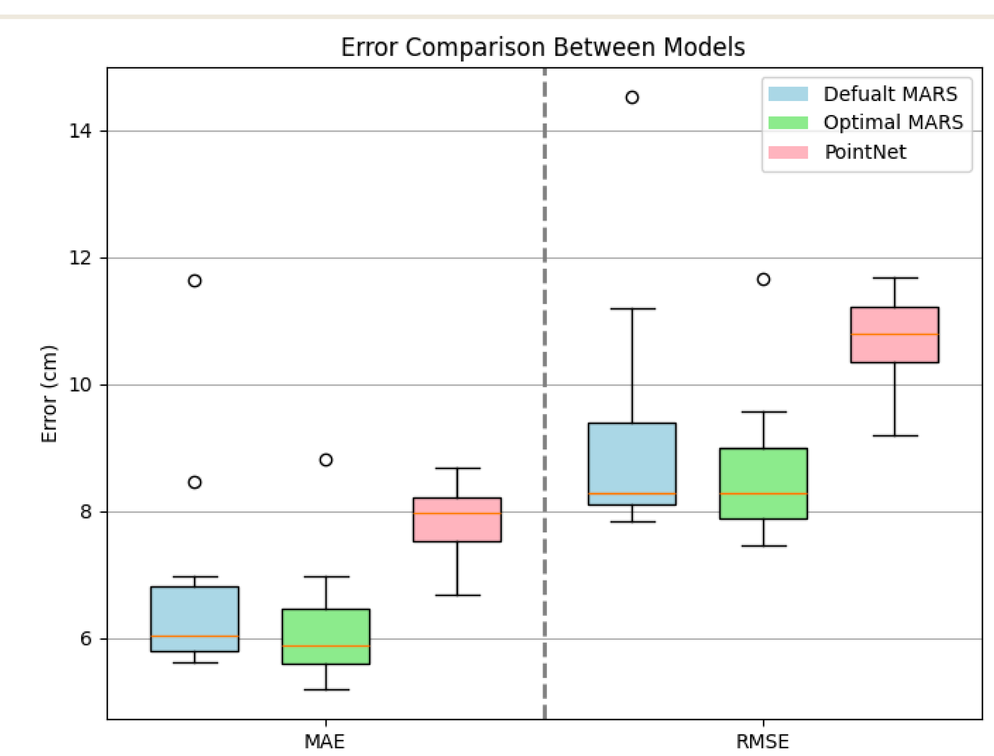

Figure 9: Validation MAE and RMSE across models

5. Limitations & Conclusions

- **Limitations:**
  - The MARS dataset does not have complex movements.
  - All combinations of convolutional and dense layers could not be tested due to the large runtimes.
- **Optimal MARS Configuration**: The best configuration for MARS consists of two convolutional layers, two dense layers, and an output layer.
- **Performance Improvement**: This setup slightly enhanced the accuracy of MARS and produced more consistent results with fewer outliers.
- **PointNet Integration**: Incorporating PointNet into MARS did not yield performance gains due to PointNet's inability to preserve local structure information.

6. References

[1] C. Iovescu, 'The fundamentals of millimeter wave sensors', 2017. [Online]. Available: https://api.semanticscholar.org/CorpusID:30625501.
[2] S. An and U. Y. Ogras, "MARS: mmWave-based Assistive Rehabilitation System for Smart Healthcare", ACM Transactions on Embedded Computing Systems, vol. 20, no. 5s, pp. 1–22, Sep. 2021, doi: 10.1145/3477003
[3] C. R. Qi, H. Su, M, Kaichun and L. J. Leonidas. (2016). PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation.
[4] Y. Zhou, H. Dong and A. El Saddik, 'Learning to Estimate 3D Human Pose From Point Cloud', IEEE Sensors Journal, vol. PP, pp. 1–1, Jun. 2020. doi: 10.1109/JSEN.2020.2999849.