# Enhancing Smartwatch Personal Identification


Authors
Ege Yümlü - eeyumlu@gmail.com

Supervisors
David Tax, Arman Naseri Jahfari, Ramin Ghorbani

Affiliations
Delft University of Technology, Faculty of Electrical Engineering, Mathematics, and Computer Science


A Comparative Study of Loss Functions in Smartwatch Data-based Personal Identification Using Auto-encoders

## Objective

The objective of the paper is to compare loss functions for outlier detection in personal identification using auto-encoders on smartwatch data. The aim is to find the most accurate loss function and understand the strengths and weaknesses of different approaches. This research contributes to the development of better methods for outlier detection in personal identification using smartwatch data.

## 1. Introduction

Smartwatches are equipped with sensors for continuous monitoring, making them sources of data for analysis tools. However, accurately identifying individuals based on smartwatch data can be challenging due to the presence of outliers [1].

Autoencoder is a neural network that can do unsupervised learning. They compress data to a lower dimensional representation, to later decode it. The difference resulting from this process is the error identified which is used to identify outliers with a threshold

The role of auto-encoders in anomaly detection by using the identified error includes Covid-19 detection using heart rate [2], identifying surface defects in the manufacturing process [3], and anomaly detection for tools under noises [4]. The choice of loss function used varies across papers. As a result, the main research question is :

**“What is the most accurate loss function for outlier detection in personal identification using auto-encoders for smartwatch data?”**

## 2. Methodology

To perform the analysis of different loss functions, the data will be preprocessed first. The heart rate is the feature that will be extracted. These will then be fed to the built autoencoder with the determined structure will be built.

After the autoencoder implementation, the different loss functions will be used , parameters optimized respectively. Then, the evaluation on given split of datasets will be analyzed for accuracy, F1 score, and other visualizations.

Root mean square will be used as a baseline, the other loss functions explore different ideas such as regularization, cosine distance, and contrastive scaling. This will give insight to which loss function is better suited for the given task.

## 3. Results and Findings

The accuracy per individual for all loss functions has been graphed for easier comparison. This can be found below:

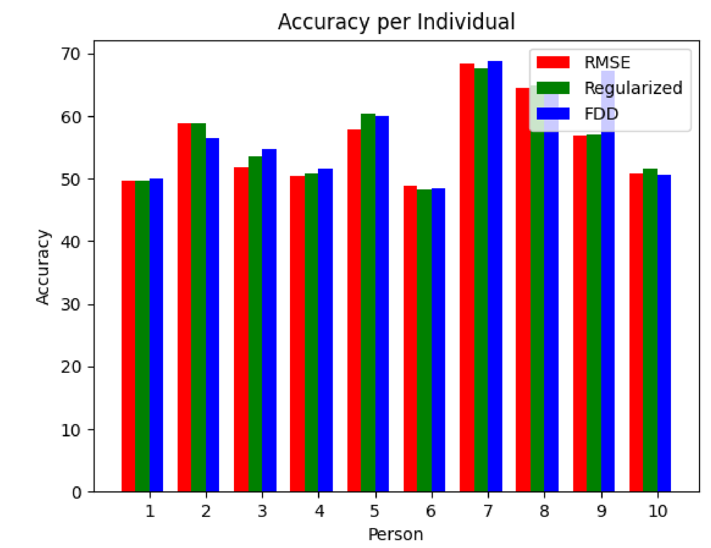


Fig 1: Accuracy Table for All Loss Functions over Individuals

The accuracy does not differ by a substantial amount. To investigate, a step filter was applied to the dataset. The accuracy dropped by 1-2 percent. Nevertheless, showed some individuals' accuracy had a substantial change. To explore, frequency graph of reconstruction error per label were plotted to see the ability of the autoencoder to separate the classes.

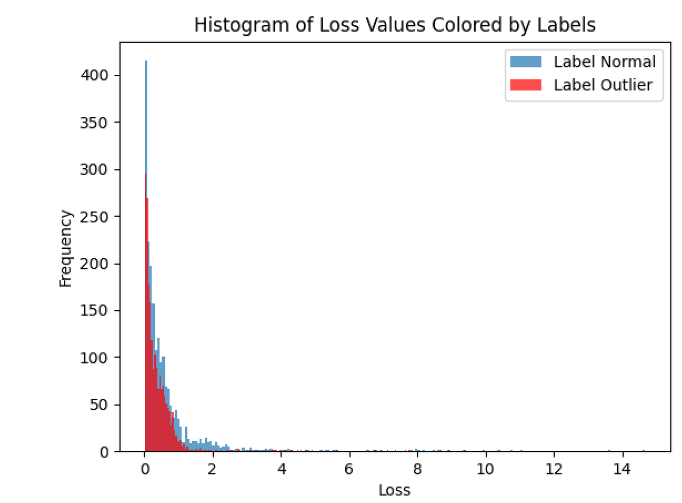


Fig 2: Evaluation Set Frequency Graph with No Training and for Individual 7

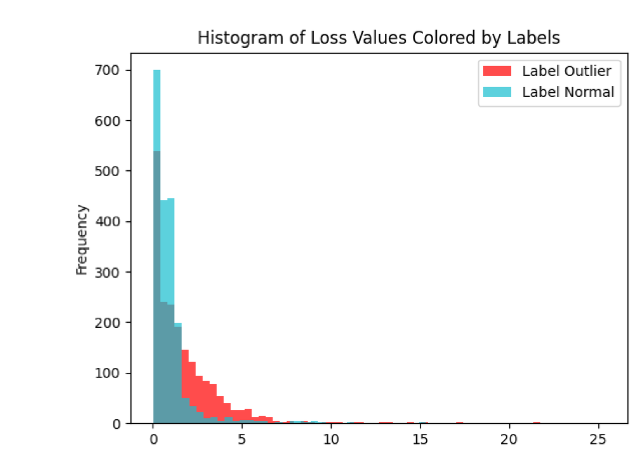


Fig 3: Evaluation Set Frequency Graph with No Training and Step Filter for Individual 7

After further exploration, as seen by the graph which compares the non-trained data for the individual 7 with and without heart rate, it becomes evident that the accuracy of the identification is affected more by the characteristics of the data rather than the underlying model itself. Non-trained versions show a similar structure to their respective accuracy.

## 4. Conclusion

The auto-encoder struggled to capture individual characteristics, resulting in non-separability of classes in the evaluation set. Performance was slightly better than random guessing, but it did not yield significant accuracy in personal identification. The variation was primarily influenced by data characteristics. In conclusion, the auto-encoder algorithm in its current form lacks efficiency and effectiveness in outlier detection in terms for personal identification.

## 5. Limitations

Simple architecture and limited data sources are notable limitations. Future research should explore complex models, expand datasets for improved generalizability, and refine data preprocessing techniques. Additionally, determining the optimal loss function using labeled data is crucial for independent evaluation of outlier detection performance.

## References


[1] CHEN Xiao-Yong, YANG Bo-Xiong, ZHAO Shuai, DING Jie, SUN Peng, and GAN Lin. Intelligent health management based on analysis of big data collected by wearable smart watch. Cognitive Robotics, 3:1–7, 2023

[2] S. Liu, J. Han, E. L. Puyal, S. Kontaxis, S. Sun, P. Locatelli, J. Dineley, F. B. Pokorny, G. D. Costa, L. Leocani, A. I. Guerrero, C. Nos, A. Zabalza, P. S. Sørensen,M. Buron, M. Magyari, Y. Ranjan, Z. Rashid, P. Conde, C. Stewart, A. A. Folarin, R. J. Dobson, R. Bailon, S. Vairavan, N. Cummins, V. A. Narayan, M. Hotopf, G. Comi, B. Schuller, and R. A. D. A. R. . N. S. Consortium. Fitbeat: Covid-19 estimation based on wristband heart rate using a contrastive convolutional auto-encoder. Pattern Recognition, 123, 2022. Cited By :10.

[3] D. . Tsai and P. . Jen. Autoencoder-based anomaly detec- tion for surface defect inspection. Advanced Engineering Informatics, 48, 2021. Cited By :35.

[4] S. Yan, H. Shao, Y. Xiao, B. Liu, and J. Wan. Hybrid robust convolutional autoencoder for unsupervised anomaly detection of machine tools under noises. Robotics and Computer-Integrated Manufacturing, 79, 2022. Cited By :34.