Technische Universität München
Lehrstuhl für Datenverarbeitung
Prof. Dr.-Ing. Klaus Diepold

Monty Matlab: Group 5

Authors: *Tobias Hofmann*, *Tobias Krug*, *Rayene Messaoud*, *Alexander Schölles*

***Abstract*—Machine Learning (ML) is widely used in many human behavior classification applications. This work deals with binary classification of acceleration data, recorded with an IMU-sensor of any smartphone. We show how to exploit characteristics of normal and unnormal walking. Different approaches, feature selections and hyperparameters are discussed. In addition to the classification algorithm, we identify diverse data, preprocessing and a straightforward Graphical User Interface (GUI) to guide the development as crucial elements to yield a high accuracy. Additionally, different ML techniques are evaluated and discussed. To determine the most accurate approach, a MATLAB implementation for training, testing and validation is carried out to determine the best performing model. A stacked model of a Support Vector Machine (SVM), a Random Forest (RF) and three Long Short-Term Memory (LSTM) networks yields the highest balanced accuracy with 99.5196 % on the test data. The SVM and RF use an amount of 102 features per acceleration dimension to learn the walking patterns. The three LSTM networks operate with 36, 213 and 378 hidden units to track different dynamics of human gait.**

***Keywords*—Classification, Gait Recognition, MATLAB, Neural Networks, Artificial Intelligence, Data Engineering**

I. Introduction

Analysing the walking behavior of a person is considered easy for the human eye. In contrast several challenges arise if a computer is tasked with this decision. Firstly, a feasible sensor, e.g. an IMU sensor in a smartphone, needs to be monitored. Secondly, preprocessing is needed before feeding any algorithm. At last, the computer requires the right algorithms and parameters to yield high accuracy - for training, testing and validation accuracy.

We present a ML approach and a GUI that allows preprocessing, training/testing and inspection of recorded gait data. The resulting model is able to distinguish between normal and unnormal gait. IMU-based classification of human movement data is well discussed in literature. Most of the works use a ML approach, like SVM, RF or Neural Networks to recognize diseases, e.g. Parkinson. In general very high accuracies can be reached [1, 2, 3].

This work aims to develop a combined ML model to reach high accuracies for binary classification of acceleration time series data from recorded gaits. This paper focuses on the algorithm itself and not on the preprocessing or the GUI. In contrast to other works building on feature dependent algorithms alone, we employ a stack of SVM, RF and three LSTM networks. To reach the goal of high testing and validation accuracy, optimal parameters for data split and hyperparameters of applied ML algorithms need to be determined iteratively. The impact of said features, hyperparameters and their combination are evaluated and discussed. The paper presents the ML approaches and the selected features. Further, the best combination of individual classifiers and their hyperparameters are outlined and results are discussed.

II. Method and Discussion

The pipeline is depicted in Figure 1. It consists of a first stage, wherein the raw data is trimmed to remove irrelevant data at the start and end. The same stage automatically splits the data according to the selected training ratio of 70:30. From then on, one part of the data is used as training data to train the actual classifier ensemble depicted in Figure 2 and the other is used as test data to rate the trained models[1].

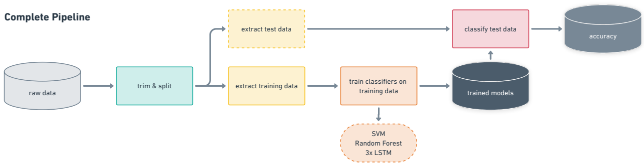

Fig. 1. Flowchart of the complete data processing pipeline

The model itself consists of five classifiers, each with specific settings regarding whether to sort the axes by their mean to compensate for different orientations during data acquisition, which features to use, how many iterations to train and how many hidden units to configure in the LSTM layer of the Neural Networks (NNs). Figure 2 shows how the SVM and RF are feed with features, whereas the LSTMs are directly working on the time-series data. Depending on the settings, our proposed ensemble operates either in 1-stage or 2-stage stacking scenario. Our test yielded, that the 1-stage stack provides superior testing accuracy, while still working reasonably well on validation data.

The search for the optimal model was done in an iterative manner. The optimization target was chosen as the mean

[1] Validation data and training/testing data were acquired by different people to support the calculation of reliable validation/generalization figures.

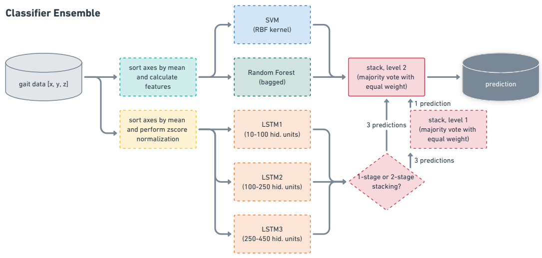

Fig. 2. Flowchart of classifier ensemble architecture

balanced accuracy of the model when applied on training, testing and validation data. In total eleven cycles with an amount of 234 models were trained and evaluated. The results of the 10th cycle are shown in table I. The goal was to determine the impact of nondeterminism introduced by MATLAB's LSTM training procedure by training 20 different models with the final settings. Table I summarizes the results of this cycle and shows a remarkable mean balanced accuracy on test data of almost 99 % for the final settings. The entry point for this cycle was determined by choosing the most promising model from the previous 9 optimization cycles. There we applied a parameter sweep across all settings. The final model has a unique combination of outstanding balanced accuracy on test data, good validation performance, fast classification (as SVM and RF operate with the same feature settings which allows a combined feature extraction), and short training time.

Tab. I. STATISTICS ON THE BALANCED ACCURACY OF 20 TRAINED MODELS WITH FINAL SETTINGS

Metric	Max [%]	Min [%]	Mean [%]	Std [%]
mean of mean (train, test, val.)	96.73	95.23	96.02	0.235
training	100	99.92	99.96	0.030
testing	99.52	97.85	98.82	0.408
validation	90.71	86.72	89.29	0.927

Subsequently, the model with the highest testing accuracy was chosen as it promises to perform well on data acquired by other people with similar movements while still supporting generalization as is indicated by its validation accuracy. The model reached a balanced accuracy on training data of 99.9693 %, on test data of 99.5196 % and on validation data of 90.7076 %.

The feature-based models SVM and RF are fed with feature vectors instead of the preprocessed data. Before assessing the features, the axes of acceleration data are sorted by mean value to tackle the problem of varying phone orientation during data recording. Besides basic statistical features like mean and variance, advanced features like the Shannon Entropy or the Wavelet Leader Estimation are evaluated and combined in a feature vector.

While the SVM model trains with a maximum number of 200 iterations using randomsearch, the RF classifier uses a bagging ensemble technique with 100 iterations.

Further the optimization process yielded the improvement by introducing multiple LSTM networks in parallel - each adapting to different time horizons of human gait dynamics. Therefore three networks are placed in parallel with 36, 213 and 378 hidden units respectively. The training was done with a minibatch size of 20 and an epoch limit of 30. Similarly to the feature extraction, an axes-sorting algorithm was implemented and contributed to a higher accuracy.

Figure 3 depicts the confusion matrix of the final model. The classification procedure runs almost instantaneously while maintaining the mentioned accuracy figures. This supports potential applications in real-time systems like gait classification in phones, smart-watches or similar.

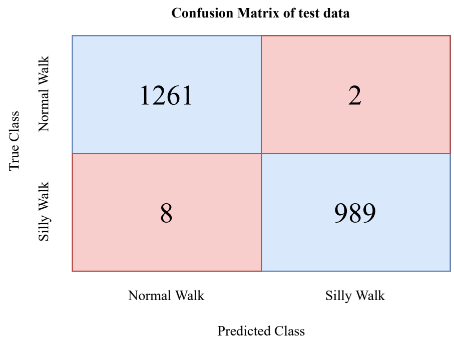

Fig. 3. Confusion Matrix of the final model on the test data

III. Conclusion

To conclude, we show a data-driven approach that was derived during an extensive optimization procedure. We yield a powerful model to classify acceleration data from human gait recorded with smartphones. The combination of LSTM models, which work with time series data, and feature-based models, namely SVM and RF, achieves an accuracy of more than 99 %. Empirical research by the authors showed that 1-stage stacking is the optimal way to implement this combination, while 2-stage stacking pretends to perform better on generalization - indicated by the higher validation accuracy.

We estimate further potential by enhancing the axes-sorting algorithm. This could be done by coordinate transformation in case 6- or 9-axes IMU-sensor-data is available. This technique may improve performance and reduce the false positive rate. This would allow new applications e.g. to recognize Parkinson or other gait-affecting conditions.

References

[1] Caramia, C.; Torricelli, D.; Schmid, M.; Munoz-Gonzalez, A.; Gonzalez-Vargas, J.; Grandas, F.; Pons, J. L.: *IMU-Based Classification of Parkinson's Disease From Gait: A Sensitivity Analysis on Sensor Location and Feature Selection*, In: *IEEE Journal of Biomedical and Health Informatics 22 (6)*, S. 1765–1774, 2018. http://dx.doi.org/10.1109/JBHI.2018.2865218. – DOI 10.1109/JBHI.2018.2865218

[2] Ajani, O. S.; El-Hussieny, H.: *An ANFIS-based Human Activity Recognition using IMU sensor Fusion.* In: *2019 Novel Intelligent and Leading Emerging Sciences Conference (NILES)* Bd. 1, 2019, S. 34–37

[3] Cuzzolin, F.; Sapienza, M.; Esser, P.; Saha, S.; Franssen, M. M.; Collett, J.; Dawes, H.: *Metric learning for Parkinsonian identification from IMU gait measurements*, In: *Gait & Posture 54*, 127-132, 2017. http://dx.doi.org/https://doi.org/10.1016/j.gaitpost.2017.02.012. – DOI https://doi.org/10.1016/j.gaitpost.2017.02.012. – ISSN 0966–6362