GaitID-2-SquatID: Deep transfer learning for human kinematics

Vinay Uday Prabhu [*1] **Stephanie Tietz** [*2]

Abstract

In the quest for more secure and varied methods of human identification, researchers have already found noteworthy results using bipedal gait analysis through accelerometric data. Using a gait-trained model, we show that transfer learning can successfully transverse different kinematic modalities and sensor modalities at the same time. Not only did the model transfer successfully from the domain of bipedal gait to the domain of body-weight squats, but it also crossed sensor domains. The gait model was originally trained on data from smartphone sensors while the squat data was collected from from inertial measurement unit (IMU) sensors. With 92% top-1 accuracy, the model classifies humans remarkably well based on their squat motions using IMU sensors and requiring only a small amount of data.

1. Introduction

Performing longitudinal studies of a cohort of athletes or patients becomes challenging, especially when they share the same medical equipment or sensors. Currently, it is being done manually via careful registration and tracking. In this paper, we address this challenge by performing machine-learning-aided automated identification of the cohort members by harnessing deep transfer learning. Researchers have enjoyed a lot of success recently in identifying humans at a large scale by performing deep learning on accelerometric gait data such as the data found in `GaitNet` (See (Prabhu & Whaley, 2009; Vinay Uday Prabhu & Whaley, 2018; Sprager & Juric, 2015)). Using these gait-trained model(s), we sought to discover if we can use transfer learning to bring human classification into the *squats domain*.

Studies that analyze gait and other human motion generally rely on video capture or inertial measurement unit (IMU) sensors (Prabhu & Whaley, 2009; Ahmadi et al., 2014). IMU sensors, like the one used in this study, have accelerometers and gyroscopes as well as additional sensors such as magnetometers, depending on the model. Smartphones have also been used as IMU sensors due to the fact that they have accelerometers and usually gyroscopes as well. In fact, the accelerometric gait data used to capture the data found in `GaitNet` was from smartphones (Prabhu & Whaley, 2009). IMU sensors have also been used previously with squat data in particular.

Physiotherapists and fitness coaches both work on squat mechanics with their athletes and patients, agreeing for the most part on what a ”correct” squat should look like. They can also recognize common incorrect squats, or deviations, in their clients. Both professions often deal with groups, which makes real-time feedback more difficult, and athletes/clients also perform these exercises when alone. To this end, researchers used IMU sensors to help classify these movements (O'Reilly et al., 2015). After standardizing the correct squat form and some deviations, the researchers had different people do the same mechanics for each variation and achieved a multiclass model accuracy of 56.55% - finding the *squat variation* amongst the data. Prior to our work, to the best of our knowledge, no one has seen if a model could instead find the *person* amongst the data. We wanted to see if the model could locate common artifacts across all of a user's squat variations - artifacts that were unique to just one person. By achieving this goal, the model could assist physiotherapists and coaches by automatically identifying the user from a group of clients or athletes.

Instead of trying to train a brand new model, our investigation turned to transfer learning. Recognizing that human classification through motion exists and was already established in the `GaitNet` studies, we decided to use a model from that study - `DeepGaitID` - as the pretrained model for our exploration. Much to our surprise, we found that the gait model does transfer well to a different modality of human kinematics. In addition to the high accuracy from the model and the low amount of training data required, the one facet of our results that was especially promising to note was the multidimensional transmodal versatility. While the gait models were trained on real-world tri-axial accelerometric gait data emanating from the IMU sensors in commercial smartphones, the squat data was collected using

[*]Equal contribution [1]UnifyID AI Labs, Redwood City, CA [2]Stanford University, Stanford, CA, USA. Correspondence to: Stephanie Tietz <stietz@stanford.edu>.

Proceedings of the Time Series Workshop, ICML 2019, 36^{th} International Conference on Machine Learning, Long Beach, CA, 2019.

an off-the-shelf IMU sensor device `metamotionr`[1]. Thus, the transfer learning is happening across different kinematic and sensor modalities at once.

The main contributions of the paper are as follows:

1. Introduce a new tri-axial accelerometric dataset (`Squat-20`) for the time-series community

2. Disseminate the first ever research effort to identify humans on the basis of their squatting motion

3. Successfully showcase deep transfer learning across two different modalities of human motion - bipedal gait and bodyweight squats

4. Open source the data and code associated with the experiments

In Section 2, we present the dataset collection procedure as well as detailed explanations into the type of data collected. Then in Section 3, we present the transfer learning procedure and results. Finally in Section 4, we conclude the paper and cover some of the directions in which we are currently extending this work.

2. `Squat-20`: Dataset description

The data collected from participants was in the form of bodyweight squats, which are squats that are performed without any equipment (no extra weight). We asked the volunteers to do eight different variations of a squat: ”correct” squats - those done with proper mechanics - as well as seven popular deviations (which we hope to use as a further classification environment in the future). The variation of squat movements performed can be seen in Table 1 along with the corresponding squat form ID used in the data.

Because there was not a constraint on squat experience, we noticed that some participants' data was more ”noisy” than others. Visually, this noise was discernible to the trainer on-site as the more amateur participants, though accurately performing the squat with the instructed mechanics, often hesitated in their movements. These hesitation marks presented a challenge to classifying squat form using traditional batch classification algorithms as the unique squat form characteristics appeared to be harder to discern. We open up this dataset for outside exploration, and we will also continue our own exploration into joint userID and squat form identification. This noise, however, did not hinder participant-based classification.

[1] https://mbientlab.com/metamotionr/

Squat Form Variations

SquatFormID	Description
0	Correct
1	Knees pass over toes
2	Knees move towards each other
3	Knees move away from each other
4	Heels up during movement
5	Hips shift to the left
6	Hips shift to the right
7	Improper hip flexion

Table 1. Every variation of the squat movement along with the corresponding form ID# used in the data

2.1. Squat Form and Deviations

A correct squat is one that maximizes power efficiency through a person's global and local centers of gravity. As one deviates away from this ideal model, the movement becomes more inefficient and can often cause injury. Participants were instructed to follow the guidelines established by the National Strength and Conditioning Association (NSCA) (Coburn & Malek, 2012). With feet shoulder-width apart, the participant was told to sit back and let their knees slowly bend while keeping a flat back and their chest up. Their heels remained on the floor with the knees in line with their toes. When beginning their ascent, the participants made sure to extend their hips and knees at the same rate and to keep their chest up, again with their heels down and knees moving in the same line as their toes.

Some incorrect variations occur due to deviations from the body's global center of gravity. For example, a user may shift their hips to the left or right while descending into their squat (SquatFormID 5,6). Another deviation in this category occurs when a user shifts their mass too far forward, often seen through the user's heels raising off the ground (SquatFormID 4) or their knees passing far over their toes (SquatFormID 1). Improper hip flexion occurs when the user drops their chest during descent and/or leads the ascent by straightening their knees, inefficiently shifting their center of gravity during the movement (SquatFormID 7).

Other deviations displace local centers of gravity in the joints, specifically the knees. If the knees move towards each other or away from each other during the squat (SquatFormID 2,3), the user is away from their maximum potential which is when the knees remain tracked in line with the toes.

The set of deviations used for data collection is not comprehensive, but does reflect seven of the most popular deviations seen as verified by a physiotherapist and a personal trainer.

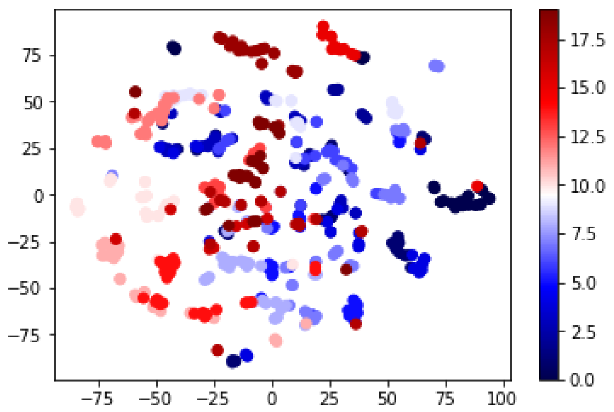

Figure 1. tSNE visualization of the 80-dimensional features extracted from the pre-softmax layers of the `DeepGaitID` classifier

2.2. Data Collection

Twenty healthy participants agreed to contribute data for this study. Volunteers varied widely in age, from 22 to 60 years old with a median age of 27 and a standard deviation 9.35. They also varied in their history of squat performance from those with little to no experience up through advanced athletes.

After the participant signed a consent form, an Mbient Lab MetaMotionR IMU sensor was placed on their back at the L4 vertebra. Each participant was instructed on the expected form for the correct squat and each deviation. They performed ten correct squats and three to five of each deviation under the supervision of a certified personal trainer and fitness coach.

Accelerometric and gyroscopic data were collected at a frequency of 100Hz during each of the squats. Accelerometric magnitudes for each participant over their squat variations can be seen in Figure 3. Data was then labeled according to squat form ID (from 0 to 7) as well as participant ID (from 0 to 19).

3. Transfer learning procedure and results

The `GaitNet` dataset is the largest accelerometric human gait dataset ever compiled (Prabhu & Whaley, 2009; Vinay Uday Prabhu & Whaley, 2018). The entire dataset is a $1.2e6 \times 4 \times 100$ tensor and contains tri-axial accelerometric data collected from 1000 volunteers in 150+ countries. Each gait cycle matrix A is of size 4×100. The 4 axes are x, y, z plus the *magintude* axis. The dimension of the temporal axis is the end-result of resampling all gait cycles to size 100. That is, $A = [a_x(t), a_y(t), a_z(t), a_{mag}(t) = \sqrt{a_x^2(t) + a_y^2(t) + a_z^2(t)}]; t = 1, .., 100$. One of the models used to analyze this dataset was a deep CNN architecture named `DeepGaitID` that achieved 63% top-1 accuracy on the 1000-class GaitNet dataset. We chose this model to serve as our pretrained model for transfer learning.

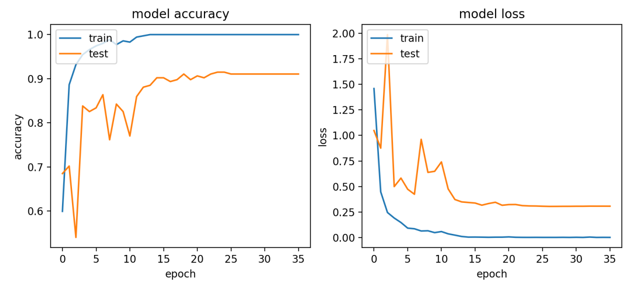

Figure 2. The epoch-wise plot of the train/test accuracy/loss

Transfer learning works on a spectrum. Generally in classification problems, the last layer of the original model is removed and replaced with a layer specific to the new data. Other layers can be frozen or left unfrozen. If all pretrained layers are frozen, the weights from those layers are used to extract features from the new data as that data is passed through to the added layer(s). Strong results come from this scenario if there is a high correlation between the original and new data. On the other end of the spectrum, all of the layers can remain unfrozen. In this case, the pretrained model acts as a weights-initializer for the new data (Pan & Yang, 2010). For this study, we chose to retrain all layers of the original model. Some of the benefits of transfer learning, which we saw in this study, are that less data is required to achieve significant results and that the model learns faster than when it trained the original data (Pan & Yang, 2010).

In order to investigate the potential of transfer learning using this model, we removed the 1000-node softmax layer from the `DeepGaitID` model(, froze every layer,) and passed the `Squat-20` tensors to obtain 80-dimensional feature vectors. Figure 1 shows the t-SNE visualization of the features obtained colored according to the user-id. Clusters emerged, dividing users reasonably well, which sets the stage for transfer learning.

We constructed a `DeepSquatID` CNN by inheriting all the pre-softmax layers of the `DeepGaitID` model and introducing a new softmax layer with 20 nodes. We then retrained this model on the `Squat-20` dataset using categorical cross-entropy loss and the `rmsprop` optimizer. To provide regularization, we used early stopping and reduced learning rate on plateau (factor=0.1, $\varepsilon = 1e-4$) strategies. We also performed label smoothing as an additional regularization pre-processing step with $\epsilon = 0.1$.

SquatID Classification Report

ParticipantID	precision	recall	f1-score	support
0	1.00	1.00	1.00	14
1	0.81	0.81	0.81	16
2	1.00	1.00	1.00	4
3	0.94	0.94	0.94	16
4	0.92	1.00	0.96	11
5	0.94	0.94	0.94	18
6	1.00	0.90	0.95	10
7	0.81	0.94	0.87	18
8	0.89	0.94	0.92	18
9	0.83	0.91	0.87	11
10	1.00	1.00	1.00	12
11	0.88	0.93	0.90	15
12	1.00	0.75	0.86	12
13	0.67	1.00	0.80	8
14	0.90	0.75	0.82	12
15	1.00	1.00	1.00	6
16	0.67	0.67	0.67	3
17	0.67	0.50	0.57	8
18	1.00	1.00	1.00	11
19	1.00	0.75	0.86	12
macro avg	0.90	0.89	0.89	235
weighted avg	0.90	0.90	0.90	235

Table 2. The classification report of the Squat-ID classification problem

In order to ensure fast enrollment of the athletes/patients, we performed a rather frugal train-test split of $0.6 : 0.4$ which left us with 352 tensors in our training set and 235 tensors in our testing set. Even with this small amount of data, our model achieved a 92% top-1 accuracy and a weighted F1 score of 0.90 as seen the classification report in Table 2. In Figure 4, the class-wise confusion matrix reveals a low number of misclassifications and high accuracy across the classes. For further exploration into this study, a *colab* notebook[2] showcasing the obtained results has been duly open-sourced.

4. Conclusion and future work

In this paper, we were able to showcase a successful transfer learning experiment that entailed using a deep CNN model pretrained on the state-of-the-art `GaitNet`. Though the original dataset contained accelerometric gait data collected from commercial phones, we were able to transfer human classification into the domain of squat exercise signatures emanating from a commercial off-the-shelf IMU sensor kit. This is currently a work in progress and we are extending this work in the following two directions:

1. Trying to replicate the results with tri-axial gyroscopic data.
2. Performing participant classification in conjunction with squat-type classification. A naive attempt at using the `DeepGaitID` CNN and trying to predict the squat-type rather than user-ID yielded an accuracy of $\sim 45\%$.

[2] https://github.com/vinayprabhu/GaitID-2-SquatID

References

Ahmadi, A., Mitchell, E., Destelle, F., Gowing, M., OConnor, N., Richter, C., and Moran, K. Automatic activity classification and movement assessment during a sports training session using wearable inertial sensors. *2014 11th International Conference on Wearable and Implantable Body Sensor Networks*, 2014.

Coburn, J. and Malek, M. (eds.). *NSCA's Essentials of Personal Training*. National Strength and Conditioning Association, 2012.

O'Reilly, M., Whelan, D., Chanialidis, C., Friel, N., Delahunt, E., Ward, T., and Caulfield, B. Evaluating squat performance with a single inertial measurement unit. *2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, 2015.

Pan, S. and Yang, Q. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22 (10):1345–1359, 2010.

Prabhu, V. U. and Whaley, J. Gaitnet 1.0 : State-of-the-art accelerometric gait based identification of humans. In *Big Data in Precision Health Conference*. Stanford, 2009.

Sprager, S. and Juric, M. Inertial sensor-based gait recognition: A review. *Sensors*, 15(9):22089–22127, 2015.

Vinay Uday Prabhu, D. W. and Whaley, J. Simuni: Sampling impostors from misfit universal background models in accelerometric gait biometric verification. In *BayLearn: Bay Area Machine Learning Symposium*, 2018.

Figure 3. Visualization of the variations in a_{mag} across different users performing squats [User-(userID) (`squatformID`)]

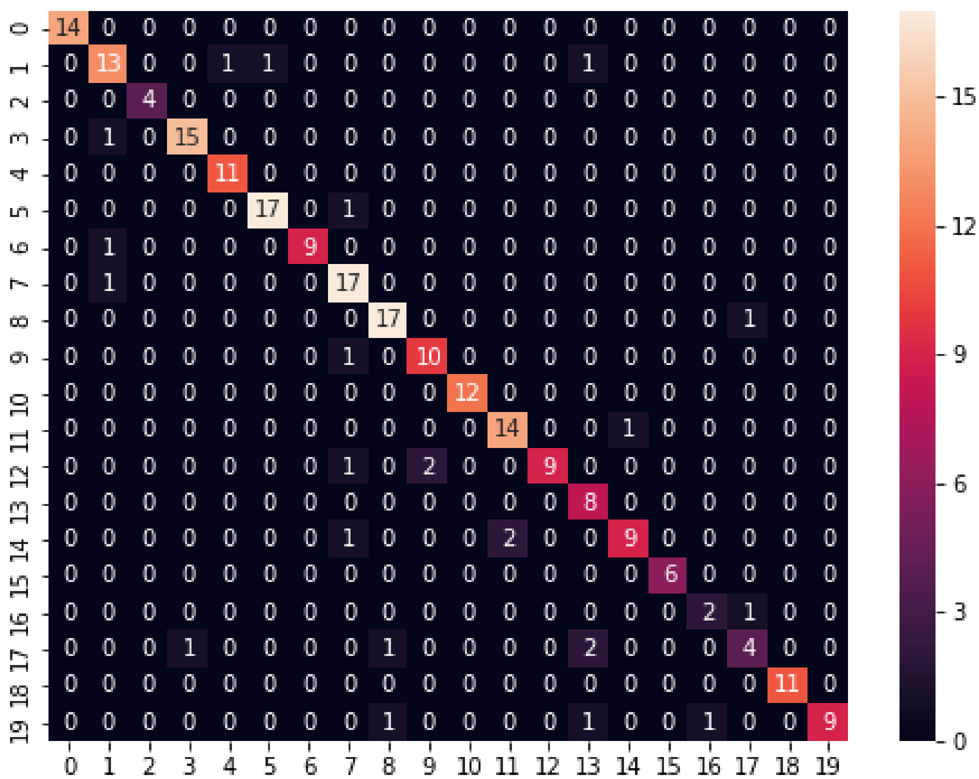

Figure 4. Confusion matrix for the `Squat-20` dataset