Capstone Proposal - Machine Learning Engineer Nanodegree

SUBJECT IDENTIFICATION BY EYES MOVEMENT: A MACHINE LEARNING APPROACH

Guilherme Lima
May 10th, 2019

Domain Background

Human identification has always been a very important problem in our world. Not only for knowing if a person is who he says he is (authentication problem), but also for identifying who is that person is (recognition problem). This can be seen in several areas as financial services, security & safety, health care, telecommunication, etc. when referred to giving someone access to resources, personalized services or individual care. [1][2]

Classical approaches as token or knowledge-based identification (e.g. passwords, physical tokens, keys, id card) are not based on any inherent attributes of an individual [2], so they present some significant drawbacks as they can be stolen, forgotten, broken or copied. Once biometrics identification apply technology to accurately identify a person by his intrinsic physical or behavioral characteristics [3], it has been increasingly used for human identification. For example: facial recognition, fingerprints, finger geometry (the size and position of fingers), iris recognition, vein recognition, retina scanning, voice recognition and DNA matching – as physiological identifiers, and also typing patterns, walking gait and other gestures – as behavioral identifiers.

Although eye-based biometrics methods have a long tradition in this field of study (like iris pattern recognition algorithms and retina scanning), an identification method based on eyes movement was first proposed by P. Kasprowski and J. Ober in 2004 [4]. Later, in 2012, Kasprowski and Komogortsev organized the First Eye Movements' Verification and Identification Competition in the IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS 2012).

Problem Statement

The competition proposed in BTAS 2012 was broken into two parts: the main competition (http://www.kasprowski.pl/emvic2012) and an additional competition on Kaggle web service (http://www.kaggle.com/c/emvic) [5]. This project is focused on solving the Kaggle competition: determining how people may be identified based on their eyes movement characteristics using a classification model.

Datasets and Inputs

Dataset provided [6] was collected at Silesian University of Technology, Poland by Dr. Paweł Kasprowski and consists of 978 samples from 37 subjects, all taken with 250Hz frequency using Ober2 movement tracking system. Each sample consists of 2048 measures in a time frame of 8192 ms through the following process [5]:

> *Step stimulus was presented in a form of a jumping dot interpolated on the 3x3 grid. The stimulus consisted of eleven dot position changes giving twelve consecutive dot positions. Subjects were given instructions to follow the dot. First dot appeared in the middle of the screen. After 1600 ms the dot in the middle disappeared and for 20 ms a screen was blank. Subsequently, a dot appeared in the upper right corner. The sequence continued until all locations of the 3x3 grid were visited. Dot movements on the grid were interspersed with dots presented at the central screen location.*

All data required for this project is available at Kaggle's competition site at https://www.kaggle.com/c/2742/download-all, and can also be obtained via Kaggle API (*kaggle competitions download -c emvic*).

Datasets are available in CSV format: 'train.csv' file for training set and 'test.csv' for the test set. The assignment of recordings to each set was made by random stratified sampling in the proportion 65% to 35% (training and test set respectively) [5]. Both files are a text file with one line for every sample, with the following structure:

Table 1 - Datasets structure

Column	Description
class	sample's classification (1 - 37 for training set and '?' in test set)
lx	list of 2048 comma separated values of left eye gaze points on X axis
ly	list of 2048 comma separated values of left eye gaze points on Y axis
rx	list of 2048 comma separated values of right eye gaze points on X axis
ry	list of 2048 comma separated values of right eye gaze points on Y axis

Solution Statement

Due to its nature, the problem will be approached as a classification problem. The aim is to analyze the data provided and build classification models using labeled recordings in the training sets and then try to use those models to classify unlabeled recordings in the testing sets.

Benchmark Model

Kaggle project administrators made available code for benchmarks for the competition (https://github.com/benhamner/emvic): one for random forest implementation, one for support vector machine and one uniform. For each one of these implementations they also provided a file with the results obtained when the benchmark code is applied to the test set.

This project will be compared to random forest and support vector machine benchmarks provided, using the evaluation metrics describe in the next section.

Evaluation Metrics

Predicted probabilities that a subject generated the eye movement sample will be evaluated using two metrics:

- **Accuracy (ACC1)**: computed as the number of records classified correctly to the whole number of records. Correct classification is when correct subject id ('class') is marked by the highest probability. [5]
- **Log loss**: guarantees that solution in which correct subject id are marked by the high probability (not necessary the highest one) get better score. It is defined as:

$$\log loss = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{M} y_{i,j} \log(y_{i,j}\hat{}) \quad (1)$$

  where N is the number of samples, M is the number of subjects, log is the natural logarithm, $y_{i,j}$^ is the posterior probability that the j^{th} subject generated the i^{th} sample, and $y_{i,j}$ is the ground truth ($y_{i,j}$=1 means that the j^{th} subject generated the i^{th} sample, $y_{i,j}$=0 indicates otherwise). [5]

Project Design

In order to find a solution to the problem given, I will carry out the following approach (figure 1):

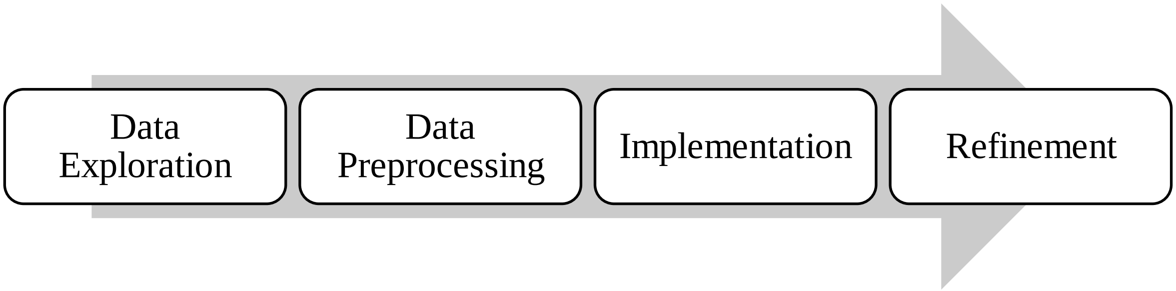

Figure 1 - Theoretical workflow

Data Exploration

The first step after collecting all data provided is to get better understanding of it. Running an exploratory data analysis (EDA) will help to figure out how to better handle features and, specially, to look for abnormalities or characteristics about the data that need to be addressed before model's implementation.

Data Preprocessing

Each sample was the result of recording one's person's eye during about 8 seconds and was taken with a frequency of 250Hz. This means that each sample consists of 2048 single measurements representing where each eye is looking at. Since we have x and y positions for left eye, as also for right eye, in each experiment are collected 2048 x 4 = 8192 integers. This brings two main considerations to keep in mind: feature extraction and dimensionality reduction.

To convert these values into a set of features one should considers techniques which have been successfully used in similar problems [4], so I will consider using methods like Principal Component Analysis (PCA), Fourier transform, cepstrum (inverse Fourier transform of the logarithm of the power spectrum of a signal) [7], or wavelets.

Also, due to the very nature of the problem dealing with eyes movement, I will consider some feature engineering in order to use first and second derivative of position (velocity and acceleration). Along with PCA, summarizing them as an average can help to reduce the number of features.

Implementation

Once I have extracted the most relevant features for identification, I will prepare the classification model. This project will consider classification algorithms between k-Nearest Neighbor, Decision Tree and Support Vector Machines, as suggested by P. Kasprowski and J. Ober [4]. Model implementation will use k-fold cross-validation.

Metrics proposed will be evaluated and results compared to benchmarks provided.

Refinement

As part of an iterative implementation process, I will look for opportunities for hyper-parameters tuning, reevaluate features and implementations in order to refine models and try to improve their performance.

References

[1]Kasprowski, P. (2012), "EMVIC - Eye Movements' Verification and Identification Competition", http://www.kasprowski.pl/emvic2012/biometrics.php. [Accessed: 10-May-2019].

[2]Anil Jain, Lin Hong, and Sharath Pankanti. 2000. Biometric identification. Commun. ACM 43, 2 (February 2000), 90-98. DOI: https://doi.org/10.1145/328236.328110

[3]Rouse, M. (2019). What is biometrics? Available at: https://searchsecurity.techtarget.com/definition/biometrics [Accessed 10 May 2019].

[4]Kasprowski, Pawel & Ober, Józef. (2004). Eye Movements in Biometrics. 248-258. 10.1007/978-3-540-25976-3_23.

[5]Kasprowski, Pawel & Komogortsev, Oleg & Karpov, Alex. (2012). First Eye Movement Verification and Identification Competition at BTAS 2012. 2012 IEEE 5th International Conference on Biometrics: Theory, Applications and Systems, BTAS 2012. 10.1109/BTAS.2012.6374577.

[6]Kasprowski, P., Ober, J. 2004. Eye Movement in Biometrics, In Proceedings of Biometric Authentication Workshop, European Conference on Computer Vision in Prague 2004, LNCS 3087, Springer-Verlag.the IEEE/IARP International Conference on Biometrics (ICB), pp. 1-8.

[7] Rabiner, L. R., Schafer, R. W.: Digital Processing of Speech Signals, Prentice Hall, Englewood Cliffs, NJ (1978)