Abstract

Title: The application of data mining and machine learning techniques to study eye movement properties in the context of neurodegenerative diseases and changes in the emotional system.

Vision is the dominant sense in humans and other primates. Visual analysis of the environment involves many areas of the brain and is sensitive to changes resulting from neurodegenerative diseases (ND), such as incurable Parkinson's disease (PD). These changes affect the patient's motor skills and are seen in the parameters of eye movements (EM) as well as in various emotions that PD also disrupts. The aim of the presented research is to identify EM parameters, which are innovative PD biomarkers, using data mining (DM) methods and to build classification and prediction models based on machine learning (ML) techniques, allowing for the estimation of PD progress.

The tests were carried out at the Department of Neurology at the Bródnowski Hospital, based on EM attributes calculated from the data obtained. The paper presents EM computational algorithms, data mining methods, and new predictive models of PD progress based on rough set theory (RST), fuzzy rough set theory (FRST), quadratic discriminant analysis (QDA), decision trees (DT), random forest (RF), multilayer perceptron (MLP), and other ML methods. A universal asynchronous and parallel processing pattern for video eye-tracking based on pre-trained convolutional models (CNN) with different architectures and compilation methods was developed. With its use, a video-eye-tracking system was implemented, allowing for online oculometric tests.

DM and ML methods have shown correlations between oculometry and PD neurological symptoms. Algorithms applied in a multidimensional parameter space enabled the finding of EM changes associated with PD symptoms. On this basis, models were built to classify motor symptoms according to the UPDRS scale used in the diagnosis of Parkinson's disease. The built ML models showed very reliable results in EM classification and prediction, even for unequal and small groups of observations.

In the case of emotion classification, the set of ED and ML methods used enabled the prediction of the subject's mood and age based on the analysis of EM parameters. In addition, the parameters related to chaos calculated on the basis of EM showed correlations with emotions. The created model, according to the available literature, for the first time allowed us to determine whether facial expressions express happiness only on the basis of the image of facial activity of the upper part of the face near the eyes.

Also the first time the online system for EM estimation was implemented, enabling remote oculometric tests using consumer computer equipment. Conducted tests proved that even using a webcam with a frequency of 30 frames per second, it is possible to

detect differences in EM latencies that indicate the late age of the examined person or the developing neurodegenerative disease.

This thesis showed that the effect of PD progression can be observed in EM properties using DM and ML methods. New biomarkers have been introduced to assess PD progress and their effectiveness in automatic classification and prediction of PD symptoms has been proven. Analysis of oculometric data using built ML models yielded reliable results and can be used for screening to counteract the effects of PD.