Parkinson’s Disease Detection based on Changes of Emotions during Speech

Justyna Skibińska
Brno University of Technology
Brno, Czech Republic
Tampere University
Tampere, Finland
Email: skibinska@feec.vutbr.cz

Radim Burget
Brno University of Technology
Brno, Czech Republic
Email: burgetrm@feec.vutbr.cz

Abstract—Parkinson’s disease (PD) is the neurodegenerative disease which affects 2-3 % of the population beyond 65 years of age in EU. When PD treatment is administered early, it is significantly more effective. Unfortunately, it is quite challenging to detect this disease at its early stage and when the symptoms can be recognized it is usually quite late. For this reason there is big motivation for development more accessible and accurate solutions for the detection of PD. One of the early symptoms is so-called hypomimia. This paper introduces an automatic method, which can objectively detect PD. The method is based on analysis of emotion changes during pronunciation defined speech exercises. We achieved balanced accuracy 69 % using XGBoost algorithm. As the exercise we proposed to use a Czech tongue twister - the difficult to pronounce sentence. The features can be explained and thus it can be used in clinical practice. We identified that the most valuable emotion for PD detection in this case is fear.

I. INTRODUCTION

Parkinson’s disease (PD) is the neurodegenerative disease and affects 2-3 % of the population beyond 65 % [1]. It is supposed that the population aging society will be one of the main problems that not only Europe will face within the next 30 years. At the same time, the number of people who are suffering from the neurodegenerative disease will increase [2]. When Parkinson’s treatment is administered in the early stage, the impairment of health is significantly smaller thanks to the applied treatment [3].

This is the reason why early detection of this disease highly-demanded.

For this purpose, there emerged plenty of new approaches in recent years. Some of them are based on new technologies that bring new opportunities that can offer significantly easier Parkinson disease detection and thus to detect it in its early stage [4].

Unfortunately, it is not an easy task, since for many years the disease has only insignificant visible markers and cognitive skills in a healthy population also vary significantly due to the level of education, age, etc. The most accurate is magnetic resonance imaging (MRI), position emission tomography (PET), or computer tomography (CT), which are unfortunately relatively expensive and because of this, they are rarely used as a preventive screening but rather at an advanced stage of the disease [5], [6]. Because of this, it is desirable to create and use cheaper solutions.

Thanks to the technologies, the detection or measuring the progress of the disease could make visible even some of the markers, which are insignificant at first glance. Thus it can consequently significantly improve the quality of life of Parkinsonians, their family, and also to create communication channels with their physicians [7]. There is also an opportunity, that the numbers of necessary visits of doctors will be reduced with the usage of telemedicine solutions and thus decrease the cost spending for the health system [7], [8].

There are plenty of symptoms of Parkinson’s disease. These symptoms could be divided into two groups: the motor and non-motor symptoms. For the motor symptoms, it could be inter alia included: the freezing of gait, bradykinesia, tremor, dyskinesia, dysphagia. Whereas, to the non-motor, it can be classified for example depression, anxiety, sleep disorders, urinary symptoms, dysarthria, or hypomimia [9], [10], [1].

The aforementioned hypomimia manifests in the reduction and slowness of facial movement (facial bradykinesia). The faces of Parkinsonians remind the so-called 'poker face' [11]. Additionally, the asymmetry of moving facial muscles is observed likewise the stiffness of the muscles is detected. Those symptoms are the reason for an instance of difficulties with expressing emotions [9]. There is also quite an interesting fact, that PD patients have worse ability to recognize the emotions of other people when compared to healthy control (HC) people [12].

The communication allows people to exchange information, ideas, and feelings or emotions [13] as well. In the case of Parkinsonism, this process is disrupted because this disease has also a negative influence on voice tract (dysarthria) [10] likewise additionally cognitive skills [14]. Moreover, the difficulties in communication skills affect also the social well-being of PD [14]. One of the speech exercise regarded as challenging to pronounce is tongue twister because of meeting problems of using correctly the mouth and tongue. It could be assumed that dysarthria would manifest especially during trying to pronounce tongue twister by PD due to the deterioration of articulators [10].

In this paper, we examined the possibility to detect Parkin-
son disease based on changes in emotion during speech exercise. The assumption of this work was dictated by the statement that the expression of emotions by PD is disturbed, they manifest the stiffness of the facial muscles likewise PD has a problem with the correct articulation. The novel approach of the machine learning model succeeded with 0.69 balanced accuracies and gave us answers which features were valuable for the obtained prediction. Moreover, we have done the statistical analysis and found those features, which were different between the group of PD and HC. Additionally, we tested more valuable tongue twister likewise reading aloud long text by participants of the experiment as the chosen speech exercise.

The main contribution of this paper is that we proved that tongue twisters can be used for PD detection and it can be based on facial features and emotion recognition. The found features can be interpreted and explained, so they can be used also in clinical practice.

The rest of the paper is structured as follows. In section, II describes related work. Section III introduces the experiment, i.e. how the experiment was performed (subsection A. Experiment description), describes the data (subsection B. Data description), how the features were extracted (subsection C. Feature extraction) and which metrics were chosen to evaluate created models (subsection D. Metrics). Section IV provides results and discussion (subsections: A. Results, B. Discussion). The last section concludes the paper.

II. RELATED WORKS

One of the most accurate approaches for PD detection is the imaging of the brain using e.g. MRI, CT, or PET [1]. The disadvantage of those methods is they are expensive and for this reason they are not very suitable for preventive screening. For this reason, we focus on cheaper approaches. Some of the symptoms like sleep disorders are challenging to monitor due to carrying out tests with the usage of uncomfortable for patients’ gold standard method - polysomnography [15].

The novel methods that have emerged in recent years include audio analysis [16], video analysis [17], wearable sensor analysis [18]. There also some works that are focused on multimodal analysis [19]. Most of the works that are related to video analysis are focused on emotion analysis (hypomimia) [20], and gait analysis [17].

There are currently only a few works that are focused on the detection of PD using the hypomimia. Due to the common problem of lack of the data in this domain, the approaches usually use just a statistical analysis [21], [22]. Few of the approaches use also automatic emotion detection [23], [24], [20]. These studies focus on the analysis of the differences in expressing the emotion of PD versus HC. The most frequent emotions that are taken into consideration are anger, fear, happiness, sadness, disgust or surprise, and neutral [25], [26].

For a better understanding of hypomimia in PD patients, there are a couple of approaches that have been used so far for facial expression. For example the analysis of electromyography (EMG) records, the affectograms, the Action Units, the Maximally Discriminative Facial Movement Coding System, and methods based on machine learning emotion recognition [25].

For PD analysis, facial features, and action units (AU) were used so far. Such a system was introduced for example in [26]. It was presented by the function of the frequency of AU. The results indicated the statistical significance of PD versus HC. Additionally, the measurements of EMG expressed through the muscle contradiction showed also a statistical difference between studied two groups. Another work focused on AU [24]. The automatic detection of PD was achieved with the usage of a 3D sensor, AU, and linear regression. The quantity of prediction was up to 0.99 Area Under Curve (AUC). Unfortunately, the dataset used is relatively limited.

In 2018, it was introduced a solution that used Mel Frequency Coefficient (MFCC) which were extracted audio and AU from video, which regressed the level of facial expression [20]. The hierarchical Bayesian network was trained on the dataset, which contains records of Parkinsonians. The obtained result for multiclass classification was 0.55 F1-score [20]. Unfortunately, the dataset is not reproducible since the dataset is private.

However, the PD detection approaches based on emotions are subjective. The imitation of emotions at a specific point of time will be at some level biased. Additionally, labeling is also not optimal, mostly done by physicians which are also a subjective process.

A few neural network architectures were proposed so far for the facial emotion recognition (FER) purpose and part of them were trained on the FER2013 dataset. The dataset contains 7 types of emotions: anger, disgust, sadness, surprise, fear, happiness, and neutral. The total number of the data were 35 685 examples [27]. One of the proposed architecture for this purpose was the simple convolutional neural network with the submission of activation function from traditional softmax to linear support vector machine [28]. The achieved accuracy was 71.2 %. Human accuracy on the FER2013 dataset was 65 ± 5 %. Another presented architecture was based on a deep neural network with two convolutional neural layers, max-pooling, and also four Inception layers. The achieved accuracy was 66.4 % [29]. Next, in [30] VGG16 was used together with a soft label constructor to take into consideration possible augmentation of the emotions. They reached 73.73 % accuracy [30].

The conclusion from the recent work is that FER can outperform humans’ abilities in emotional recognition. Thanks to this, the neural network architecture has the potential to improve also PD detection, for example for clinical practice. One of the biggest limitations of the current approaches is still a pretty limited amount of data for the training of machine learning algorithms. For those reasons, analysis of general emotion detection can vary significantly based on a particular activity. In this paper, we propose a technique that is based on pronunciation unified speech exercises and we validate whether they can improve the quality of PD detection.
III. EXPERIMENT

A. Experiment description

The main goal of this experiment has been to evaluate the accuracy of Parkinson’s disease detection based on changes in the expression of emotion during the speech. In particular in face mimic changes during exercises - i.e. during reading selected text by PD and HC.

The scheme of the carried-out experiment is shown in fig. 1. First, the process of feature extraction was done. It is described in more detail in section III-C.

Subsequently, the features were pre-selected using the minimum redundancy maximum relevance (mRMR) algorithm. The 50 most relevant features were chosen. Afterward, the machine learning classifiers were used with the Stratified 10-fold Cross-Validation. The 6 classifiers were trained and the results obtained with the usage of them were compared. The following classifiers were used: Support Vector Machine (SVM), k-nearest neighbors (kNN), Decision Tree (DT), Random Forest (RF), Logistic Regression (LR), and XGBoost.

Results were evaluated using U-Man Whitney tests which validated whether the distributions of the two populations (HC vs. PD) for each feature differ. The false discovery rate (FDR) correction was also calculated to offset the number of type I errors.

B. Data description

We enrolled 45 HC (21 females (mean age 62 ± 9.22), 24 males (mean age 66 ± 9.17)) and 70 PD patients (27 females (mean age 68 ± 8.04), 43 males (mean age 66 ± 7.83)). For female patients mean UPDRS III was 21.6 ± 13.50 and the mean duration of PD in years was 7.2 ± 4.82, whereas for male Parkinsonians mean UPDRS III was 26.3 ± 11.31 and mean years duration of PD in years was 7.9 ± 4.64. The patients and healthy control groups were recorded during pronunciation inter alia: vowels, sentences, and reading long texts. The records of the video were acquired with the sampling frequency of 25 frames per record (FPS).

For this experiment, a Czech sentence (Celý večer se učí sčítat)¹ was selected, and reading them aloud long text was selected. The meaning of the sentence is "He’s been learning to count all night", but rather than its meaning it is important that the sentence is hard to pronounce. The collection of the data was done by neurologists. The experiment was approved by the Ethical Committee of Masaryk University.

This we got in total 70+45 = 115 video records for speech exercise. The total length of the records is various.

C. Feature extraction

The process of obtaining features consists of several steps. General idea was to extract those features that can describe changes of emotion in a certain period, here, dedicated and next this same could be applied to numerically analyze emotions during speech exercise. First, it was used publicly available model of facial emotion recognition (FER) [31] trained on the publicly available dataset: FER 2013 [32]. FER architecture contains two steps of facial emotion recognition to provide the most accurate prediction. It is called Multi-task Cascaded Convolutional Networks (MTCNN) for face detection [33] and architecture for evaluation the intensity of the emotions, they are anger, fear, happiness, sadness, disgust, neutral and surprise [31]. The features were extracted from video records in the form of a seven time-series, which represent dynamics of the facial expressions. The regression values were calculated separately for each frame steamed from the video. An example is shown in fig. 3. Moreover, the results of using FER on a single frame of PD can be seen in fig. 2.

In the next step, changes of emotions during the given time interval are represented numerically and statistical differences between two examined groups, i.e. PD versus HC was calculated. They are minimum, maximum, the amplitude of the course, standard deviation (STD), relevant standard deviation (RSTD), variance, the slope of the time-series, and features computed based on distribution data, i.e.: kurtosis and skewness. Furthermore, also some indicators describing the measure of information disorders were taken into consideration. They were approximate entropy dedicated to medical series and Shannon entropy. This set of features was calculated for all of the types of records of emotions.

D. Metrics

For the evaluation purpose, 4 metrics were calculated. Due to the small imbalance of the dataset, the accuracy balanced and Matthews’s correlation coefficient was computed. Additionally, clinically valuable features of sensitivity and

¹Link for the pronunciation of the sentence: https://bit.ly/2DVPJ5M
specificity were taken also into consideration. The equation of them are presented below:

\[ \text{Sensitivity} = \frac{TP}{TP + FN} \]  
\[ \text{Specificity} = \frac{TN}{TN + FP} \]

\[ \text{Balanced Accuracy} = \frac{\text{Sensitivity} + \text{Specificity}}{2} \]

\[ MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \]

E. XGBoost as classifier

A few classifiers were tested in this work, however, the XGBoost classifier deserves special attention. This algorithm characterizes optimizing specific loss function as function approximation and using regularization techniques. Additionally, XGBoost is computing faster because of applied parallelization and distributed techniques [34]. The equations 4, 5, 6, 7, 8 explain mathematically the used one of the biggest advantages of XGBoost - regularization technique.

This algorithm needs to minimize the function which contains loss function \( l \) and regularization part \( \Omega \) (see Eq. 4).

\[ \mathcal{L} = \sum_{i=1}^{n} [l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i))] + \Omega(f_t) \]  

(4)

However, the obstacle with minimization Eq. 4 is that commonly used optimization techniques need to be performed in the Euclidean space and they are not suitable in this case. For this reason, the second-order Taylor approximation is used for this problem to optimize new obtained function in Euclidean space.

The assumption of this transformation is to use as specific point previous step of iteration \((t-1)\) and loss function \( l \) as the function, which is going to be approximate. After this transformation, the function is presenting in such a shape:

\[ \mathcal{L} = \sum_{i=1}^{n} [g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)] + \Omega(f_t) \]  

(5)

where \( g_i \) and \( h_i \) are the first and second-order gradient, respectively. In the next step, it is removing the constant part (see Eq. 6 and it use at the moment the optimization techniques dedicated to quadratic functions.

\[ \bar{\mathcal{L}}(t) = \sum_{i=1}^{n} [g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)] + \Omega(f_t) \]  

(6)

The main point is to find the weights \( w \) of the nodes and after function’s transformations present as form Eq. 7 and it is equal to Eq. 8.

\[ \mathcal{L}(f_m) \propto \sum_{j=1}^{T_m} [G_{jm} w_{jm} + \frac{1}{2} H_{jm} w_{jm}^2] \]  

(7)

\[ w_{jm} = -\frac{G_{jm}}{H_{jm}}, j = 1, \ldots, T_m. \]  

(8)

Where, \( G_{jm} \) equals to the sum of gradient in the region (leaf) \( j \) and \( H_{jm} \) is to the sum of hessian in region \( j \) for the set of leaves \( m \).

To summarize and highlight the mentioned optimization technique of XGBoost, this classifier is taking into consideration L1 and L2 regularization likewise penalization of the number of leaf nodes [34].
IV. RESULTS AND DISCUSSION

A. Results

The results of the statistical analysis of tongue twister are shown in the tab. I. The 10 most statistically different features between PD and HC were standard deviation (fear_std), variance (fear_variance), maximum (fear_max), range (fear_range) of fear, standard deviation, variance (angry_variance), minimum (angry_min) of anger and approximate entropy of sadness. For all of them, p-value and p-value with FDR correction was below 0.05 value. Additionally, the median and interquartile range were calculated for all of them. According to statistical analysis, the most relevant dependency which allows us to distinguish between Parkinsonian versus healthy person is a manifestation of fear on the face.

Results of prediction by 6 classifiers are shown in the tab. II. The most accurate solution that allows us to detect Parkinson’s disease has been identified as the XGBoost classifier. The accuracy balanced succeed in 0.69, MCC was also the highest for this algorithm and was equal to 0.39. The sensitivity and specificity were 0.71 and 0.67 respectively. The highest sensitivity was achieved for LR (0.73) and specificity 0.73 was achieved for DT.

The interpretability of the model was explained with the usage of SHAP value’s and it is presented in Fig. 4. The most important features according to explainability in descending order are the standard deviation of fear (fear_std), approximate entropy of surprise (surprise_approx_entropy), the variance of fear (fear_variance), maximum of anger (angry_max), range of fear (fear_range), the variance of anger (anger_variance), maximum of fear (fear_maximum), mean of surprise (surprise_mean), mean of fear (fear_mean), approximate entropy of sadness (sad_approx_entropy). The fear_std, fear_variance, angry_max, angry_variance, fear_max, and fear_mean are positively correlated with the prediction of Parkinson’s disease. However, the surprise_approx_entropy and sad_approx_entropy are negatively correlated with Parkinson’s disease.

![SHAP's values obtained from the XGBoost model for the tongue twister.](image_url)

Results of predictions between each particular speech exercise were also compared and are shown in the tab. III. It was evaluated with the same metrics for also reading long texts. It was found that better balanced accuracy was achieved with tongue twister exercise and it was 0.69 versus 0.60 balanced accuracy for reading long text.

B. Discussion

The automatic method of detection of Parkinson’s disease and analyzing changes of emotion for the Czech tongue twister was created. We have created features that occurred to be statistically different for this task. The most valuable feature was identified expression of fear emotion. In this paper, this emotion was expressed mainly due to trying to pronounce difficult sentences. We found that those speech exercises are even better for the purpose of detecting PD than reading long text. The best machine learning algorithm was identified as the XGBoost, this even using an imbalanced dataset. The changes in fear were correlated the best by features like standard deviation, variance, range, or mean. It could mean that this kind of emotion is recognized during this speech exercise and difficulties which meet the PD making them at a certain level anxious in comparison to HC. Moreover, the approximate entropy of sadness and surprise is negatively correlated with Parkinson’s disease, this fact could indicate lower changes in the expression of this emotion for PD. Additionally, the general performance of the XGBoost model was 0.69 balanced accuracy which indicates promising results and shows that changes in the emotion of PD during speech exercise are valuable from the scientific point of view.

V. CONCLUSION

In this study, we presented the novel attitude of detection of Parkinson’s disease based on facial features expression. Firstly, we elaborated on the features, which could differentiate healthy control group versus Parkinsonians. The features describing differences in expressing fear during the time were the most significant from the statistical point of view. The XGBoost classifier outperformed other classifiers and achieved 0.69 balanced accuracy. The chosen by us algorithm allows us to apply SHAP’s value and this same interpret created model. Once again, the features based on fear emotion indicate on positive correlation with Parkinson’s disease. Furthermore, it seems that the selection of speech exercise plays a crucial role in the final performance. The space for future research is to increase the database and likewise test another tongue twister or develop another value for this research speech exercise.

ACKNOWLEDGMENT

The authors gratefully acknowledge funding from European Union’s Horizon 2020 Research and Innovation programme under the Marie Skłodowska Curie grant agreement No. 813278 (A-WEAR: A network for dynamic wearable applications with privacy constraints, http://www.a-wear.eu/). This work does not represent the opinion of the European Union, and the European Union is not responsible for any use that might be made of its content.
### TABLE I  
**Statistical analysis**

| Features          | pVal    | pVal(FDR) | Median (PD) | Median (HC) | IQR (PD)  | IQR (HC)  |
|-------------------|---------|-----------|-------------|-------------|-----------|-----------|
| fear_std          | 0.0034  | 0.0408    | 0.0564      | 0.0350      | 0.0649    | 0.0390    |
| fear_variance     | 0.0034  | 0.0408    | 0.0032      | 0.0012      | 0.0078    | 0.0032    |
| angry_mean        | 0.0038  | 0.0408    | 0.1762      | 0.1276      | 0.1953    | 0.1639    |
| angry_std         | 0.0039  | 0.0408    | 0.0819      | 0.0562      | 0.0605    | 0.0499    |
| angry_variance    | 0.0039  | 0.0408    | 0.0067      | 0.0032      | 0.0097    | 0.0060    |
| fear_max          | 0.0040  | 0.0408    | 0.2950      | 0.2100      | 0.3050    | 0.1800    |
| fear_mean         | 0.0045  | 0.0408    | 0.0988      | 0.0725      | 0.1711    | 0.0719    |
| angry_min         | 0.0050  | 0.0408    | 0.0600      | 0.0400      | 0.0700    | 0.0600    |
| fear_range        | 0.0050  | 0.0408    | 0.2500      | 0.1900      | 0.2700    | 0.1500    |
| sad_approx_entropy| 0.0053  | 0.0408    | 0.3624      | 0.4990      | 0.3005    | 0.2781    |

### TABLE II  
**RESULT OF PREDICTION FOR TONGUE TWISTER THANKS TO THE DIFFERENT ALGORITHMS**

| Classifier       | ACC (balanced) | SEN  | SPE   | MCC  |
|-------------------|----------------|------|-------|------|
| KNN               | 0.63 (0.13)    | 0.69 (0.17) | 0.57 (0.23) | 0.27 (0.27) |
| SVM               | 0.59 (0.13)    | 0.70 (0.16) | 0.49 (0.25) | 0.19 (0.27) |
| Decision Tree     | 0.57 (0.13)    | 0.41 (0.17) | **0.73 (0.24)** | 0.15 (0.27) |
| Random Forest     | 0.62 (0.14)    | 0.68 (0.18) | 0.56 (0.23) | 0.25 (0.29) |
| Logistic Regression| 0.59 (0.13)    | **0.73 (0.16)** | 0.45 (0.24) | 0.18 (0.28) |
| XGBoost           | **0.69 (0.14)** | 0.71 (0.17) | 0.67 (0.22) | **0.39 (0.29)** |

### TABLE III  
**RESULTS OF PREDICTION FOR TONGUE TWISTER AND READING TEXT WITH THE USE OF THE XGBoost**

| Speech exercise | ACC (balanced) | SEN  | SPE   | MCC  |
|-----------------|----------------|------|-------|------|
| Tongue twister  | 0.69 (0.14)    | 0.71 (0.17) | 0.67 (0.22) | 0.39 (0.29) |
| Long text       | 0.60 (0.16)    | 0.66 (0.21) | 0.54 (0.27) | 0.20 (0.34) |

### REFERENCES

[1] W. Poewe, K. Seppi, C. M. Tanner, G. M. Halliday, P. Brandt, J. Volkmann, A.-E. Schrag, and A. E. Lang, “Parkinson disease,” *Nature reviews Disease primers*, vol. 3, no. 1, pp. 1–21, 2017.

[2] *Global Europe 2050*, 2020 (accessed August 12, 2020).

[3] R. Prashanth and S. D. Roy, “Early detection of parkinson’s disease through patient questionnaire and predictive modelling,” *International journal of medical informatics*, vol. 119, pp. 75–87, 2018.

[4] Z. Obermeyer and E. J. Emanuel, “Predicting the future—big data, machine learning, and clinical medicine,” *The New England journal of medicine*, vol. 375, no. 13, p. 1216, 2016.

[5] M. Ghassemi, R. Hilker, J. Rudolf, J. Sobesky, and W.-D. Heiss, “Analysis of facial expressions in parkinson’s disease through video-based automatic methods,” *Journal of neuroscience methods*, vol. 281, pp. 7–20, 2017.

[6] P. G. Vilda, J. Mekyska, A. G. Rodellar, D. P. Alonso, V. R. Biarge, and A. A. Marquina, “Monitoring Parkinson disease from speech articulation kinematics,” *Laues: revista espanola de ciencias del habla*, no. 4, p. 2, 2017.

[7] M. Rajnoha, J. Mekyska, R. Burget, I. Eliasova, M. Kostalova, and I. Rektorova, “Towards identification of hypomimia in Parkinson’s disease based on face recognition methods,” in *2018 10th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, pp. 1–4, IEEE, 2018.

[8] L. Ricciardi, F. Visco-Comandini, R. Erro, F. Morgante, M. Bologna, A. Fassano, D. Ricciardi, M. J. Edwards, and J. Kilner, “Facial emotion recognition and expression in Parkinson’s disease: an emotional mirror mechanism?,” *PLoS One*, vol. 12, no. 1, p. e0169110, 2017.

[9] U. Sioresmi et al., “Tongue twisters in pronunciation class,” in *International Conference on Teacher Training and Education*, Sebela Maret University, 2016.

[10] D. Aarsland, B. Creeve, M. Politis, K. R. Chaudhuri, D. Weintraub, C. Ballard, et al., “Cognitive decline in Parkinson’s disease,” *Nature Reviews Neurology*, vol. 13, no. 4, pp. 217–231, 2017.

[11] Y. Zhang, R. Ren, L. D. Sanford, L. Yang, J. Zhou, L. Tan, T. Li, J. Zhang, Y.-K. Wing, J. Shi, et al., “Sleep in Parkinson’s disease: A systematic review and meta-analysis of polysomnographic findings,” *Sleep Medicine Reviews*, vol. 51, p. 101281, 2020.

[12] A. Gómez-Rodellar, A. Álvarez-Marquina, J. Mekyska, D. Palacios-Alonso, D. Meghraoui, and P. Gómez-Vilda, “Performance of articulation kinetic distributions vs MFCCs in Parkinson’s disease: a deep learning approach,” *IEEE Journal of Health Informatics*, vol. 31, no. 9, pp. 1272–1282, 2016.

[13] E. J. Henderson, E. R. Dorsey, M. S. Okun, N. Okubadejo, P. Chan, J. Andrejack, et al., “Integrated and patient-centred management of Parkinson’s disease: a network model for emerging chronic neurologic care,” *The Lancet Neurology*, 2020.
[21] D. Bowers, K. Miller, W. Bosch, D. Gokcay, O. Pedraza, U. Springer, and M. Okun, “Faces of emotion in parkinsons disease: micro-expressivity and bradykinesia during voluntary facial expressions,” Journal of the International Neuropsychological Society: JINS, vol. 12, no. 6, p. 765, 2006.

[22] L. Ricciardi, A. De Angelis, L. Marsili, I. Faiman, P. Pradhan, E. Pereira, M. Edwards, F. Morgante, and M. Bologna, “Hypomimia in parkinson’s disease: an axial sign responsive to levodopa,” European Journal of Neurology, 2020.

[23] N. Vinokurov, D. Arkadir, E. Linetsky, H. Bergman, and D. Weinshall, “Quantifying hypomimia in parkinson patients using a depth camera,” in International Symposium on Pervasive Computing Paradigms for Mental Health, pp. 63–71, Springer, 2015.

[24] B. Sonawane and P. Sharma, “Review of automated emotion-based quantification of facial expression in parkinson’s patients,” environment, vol. 7, p. 8.

[25] P. Wu, I. Gonzalez, G. Patsis, D. Jiang, H. Sahli, E. Kerckhofs, and M. Vandekerckhove, “Objectifying facial expressivity assessment of parkinson’s patients: preliminary study,” Computational and mathematical methods in medicine, vol. 2014, 2014.

[26] Y. Gan, J. Chen, and L. Xu, “Facial expression recognition boosted by soft label with a diverse ensemble,” Pattern Recognition Letters, vol. 125, pp. 105–112, 2019.

[27] J. Shenk, Facial expression recognition, 2020 (accessed August 11, 2020).

[28] I. J. Goodfellow, D. Erhan, P. L. Carrier, A. Courville, M. Mirza, B. Hamner, W. Cukierski, Y. Tang, D. Thaler, D.-H. Lee, et al., “Challenges in representation learning: A report on three machine learning contests,” Neural Networks, vol. 64, pp. 59–63, 2015.

[29] I. de Paz Centeno, Face detection, 2020 (accessed August 11, 2020).

[30] T. Chen and C. Guestrin, “Xgboost: A scalable tree boosting system,” in Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, pp. 785–794, 2016.