Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 186 (2021) 546–554

Procedia Computer Science

www.elsevier.com/locate/procedia

14th International Symposium "Intelligent Systems"

Parkinson's Disease Classification and Medication Adherence Monitoring Using Smartphone-based Gait Assessment and Deep Reinforcement Learning Algorithm

Armando de Jesús Plasencia Salgueiro[a,*], Yulia Shichkina[b], Arlety García García[c], Lynnette González Rodríguez[a]

[a]Institute of Cybernetics, Mathematics and Physics, Calle 15 No. 551, Vedado, Havana 10400, Havana, Cuba
[b]Saint Petersburg Electrotechnical University "LETI", ul. Professora Popova 5, 197376, St. Petersburg, Russian Federation
[c]Universidad de la Isla de la Juventud "Jesús Montané Oropesa", Nueva Gerona, Cuba

Abstract

For the diagnosis and classification of Parkinson's patients, the unified scale of Parkinson's patients (Unified Parkinson's Disease Rating Scale – UPDRS) is used, which requires the patient to perform a series of tests among which the biomarkers of speech, the facial expression, the hand movement and walking analysis are considered, after which the doctor diagnoses the patient whether or not he has Parkinson's disease according to the score obtained. The work proposes a system for monitoring patients with the use of cell phones and their automatic classification according to the data collected by them. The system starts from the budget that Parkinson's patients have different abnormalities when walking if they do not follow the required medication. The cell phone collects the data passively while the patient has his cell phone in his pocket. After that, the data preprocessor helps to extract the walking cycles that this Parkinson-related biomarker contains. The algorithm proposed for classification and Medication Adherence Monitoring is the Deep Reinforcement Learning. With this work we demonstrate the feasibility of using cell phones to monitor the biomarker walking in Parkinson's patients and the possibility of Passive Medication Adherence Monitoring and Dynamic Treatment Regimes.

Keywords: Parkinson's disease; biomarkers monitoring; medication adherence; deep reinforcement learning; smatphone

* Corresponding author. Tel.: +53 53851773.
E-mail address: armando@icimaf.cu

1877-0509 © 2021 The Authors. Published by Elsevier B.V.

10.1016/j.procs.2021.04.175

1. Introduction

The Parkinson's disease is a degenerative disorder of slow progress of the nervous system caused by the lack of the levels of dopamine which can provoke uncontrolled involuntary movements of the body and psychological affections.

The doctors usually diagnose the disease clinical way according with their abilities and experience. However, due to the subjectivity of the diagnosis in occasions diagnoses and erroneous treatments can take place. For the development of a practical, low cost and general diagnosis system of the symptoms of the patients of Parkinson it is necessary to define the methods firstly to carry out, the sources of data information and the algorithms of machine learning that allow carrying out the analysis of the symptoms of the illness should come [1]. The use of the artificial intelligence techniques in the field of the medicine and in particular in the precocious diagnosis of Parkinson's disease it has demonstrated to be very effective and efficient [2]

A dedicated medication adherence for the Parkinson's disease (PD) patients allows the physicians to accurately adjust the treatment based on the patient's clinical response. In contrast, non-adherence due to missed, mistimed or extra doses leads to increase in parkinsonism, e.g., motor fluctuations.

In real practice, achieving a high medication adherence among PD patients is a difficult task. Firstly, disease progressing is one relevant factor. In the early stage, patients are diligent in taking prescribed drugs 3-4 times per day. However, as the treatment progresses, they are required to take drugs 6-10 times per day. Greater regimen complexity is observed to profoundly impair medication adherence, which drops sharply with each incremental dose in daily life. Secondly, as a non-motor PD symptom, depression is considered as a significant risk factor for medication adherence. The non-adherence aggravates depression while depression fuels non-adherence. Lastly, cognitive impairment, e.g., dementia, affects at least 40% of PD patients. Dementia impairs the patient's working memory and executive function, resulting in missed dose or overdosage in daily life. To summarize, the non-adherence caused by PD symptoms presents a knowledge gap between patients and physicians, i.e., patients cannot recall or validate the medicine intake events due to which physicians cannot proactively suggest improvements in the treatment process [3].

A practical approach to facilitate medicine intake detection would serve as the first milestone for improving medication adherence. Once the physicians have assigned a drug schedule to the patients, how can they verify the day-to-day occurrences of medicine intake events?

(Zhang H. et al., 2019) [3] Proposed an AI-care strategy which can assist in the PD treatment by taking advantage of mobile technologies, it leverages a commercial of-the-shelf smartphone to monitor medication adherence without the user's awareness. The periods and frequency of intake activities are compared against the drug schedule assigned by the physicians to determine the mistimed, missed and extra doses. Afterward, the results from medicine intake detection, along with the medication effectiveness measurement, are transferred to the clinics automatically.

Human Activity Recognition - HAR - has emerged as a key research area in the last years and is gaining increasing attention by the pervasive computing research community (see picture below, that illustrates the increasing number of publications in HAR with wearable accelerometers), especially for the development of context-aware systems. There are many potential applications for HAR, like: elderly monitoring, life log systems for monitoring energy expenditure and for supporting weight-loss programs, and digital assistants for weight lifting exercises.

The article is dividing in three parts, 1. Introduction, 2. Methods in order to provide the theoretical foundations and the requisites of the main platforms, and 3. Results, where we expose the selected platforms, their characteristics and how to implement them.

2. Methods

2.1. Modeling the Parkinson's tremor

Control theory is a mathematically oriented discipline within engineering that concerns the design and analysis of systems for the regulation. Control theory has a broad scope, encompassing problems as diverse as system identification, state and parameter estimation, analysis of nonlinear feedback control systems, and optimal control.

The analogy between mechanical control and regulation in biological systems was articulated in Norbert Wiener's 1948 book *Cybernetics: or Control and Communication in the Animal and the Machine.* This analogy has proven fruitful in many domains of physiology (including computational neuroscience) [4].

Parkinson's disease (PD) is a central nervous system disorder with vast symptoms. The disease is caused by degeneration or malfunctioning of basal ganglia (BG). This portion is composed of different parts and its main function is movement control. PD is caused by decreased dopamine secretion from ''substantia nigra pars compacta''. The common symptom is tremor, which can be subdivided into physiological and pathological tremors. Physiological tremor is a low-amplitude oscillation that, with suitable recording techniques, can be demonstrated in almost all normal subjects [5].

The internal relationships among the BG components as well as the relationships with external blocks are shown in Fig. 1.

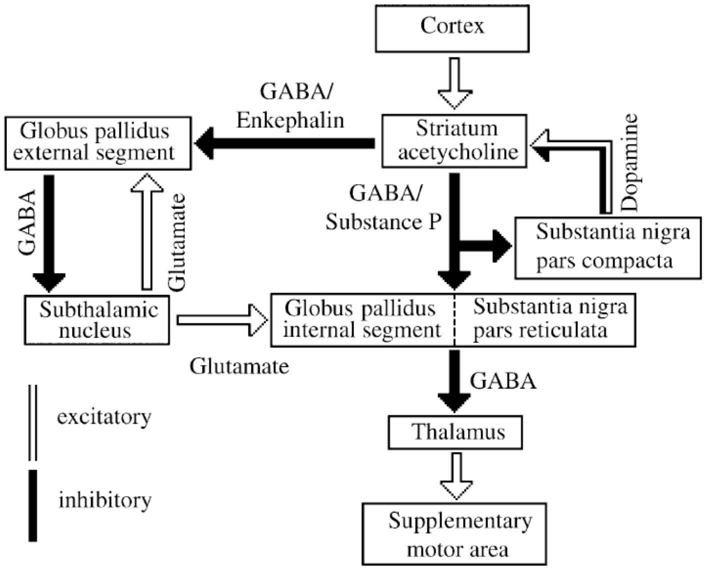

Fig. 1. Conceptual model of BG [5]

One routes in PD treatment is the medical treatment. So, we should restore the failure internal feedback loop between "substantia nigra pars compacta" and "Striatum acetycholine" with and external feedback loop that can feed with exact doses of dopamine to the organism in require time that can simulate the function of internal loop (Dynamic Treatment Regimes – DTR).

PD is a Mental Disease. Mental diseases are characterized by a long-term period of clinical treatments that usually require adaptation in the duration, dose, or type of treatment over time.

One goal of healthcare decision-making is to develop effective treatment regimens that can dynamically adapt to the varying clinical states and improve the long-term benefits of patients. *Dynamic treatment regimens (*DTRs), provide a new paradigm to automate the process of developing new effective treatment regimens for individual patients with long-term care. A DTR is composed of a sequence of decision rules to determine the course of actions (e.g., treatment type, drug dosage, or reexamination timing) at a time point according to the current health status and prior treatment history of an individual patient. Unlike traditional randomized controlled trials that are mainly used as an evaluative tool for confirming the efficacy of a newly developed treatment, DTRs are tailored for generating new scientific hypotheses and developing optimal treatments across or within groups of patients. Utilizing valid data generated, an optimal DTR is capable of optimizing the final clinical outcome of particular interest can be derived [6].

2.2. Reinforcement learning (RL)

To build this external loop we propose the use of reinforcement learning (RL), a subfield in machine learning. In RL an agent chooses an action at each time step based on its current state, and receives an evaluative feedback and the new state from the environment. The goal of the agent is to learn an optimal policy (i.e., a mapping from the states to the actions) that maximizes the accumulated reward it receives over time. Therefore, agents in RL do not receive direct instructions regarding which action they should take, instead they must learn which actions are the best through trial-and-error interactions with the environment. This adaptive closed-loop feature renders RL distinct from traditional supervised learning methods for regression or classification, in which a list of correct labels must be provided, or from unsupervised learning approaches to dimensionality reduction or density estimation, which aim at finding hidden structures in a collection of example data. Moreover, in comparison with other traditional control-based methods, RL does not require a well-represented mathematical model of the environment, but develops a control policy directly from experience to predict states and rewards during a learning procedure. Since the design of RL is letting an agent controller interact with the system, unknown and time-varying dynamics as well as changing performance requirements can be naturally accounted for by the controller. Lastly, RL is uniquely suited to systems with inherent time delays, in which decisions are performed without immediate knowledge of effectiveness, but evaluated by a long-term future reward [6].

The design of DTRs can be viewed as a sequential decision-making problem that fits into the RL framework well. The series of decision rules in DTRs are equivalent to the policies in RL, while the treatment outcomes are expressed by the reward functions. The inputs in DTRs are a set of clinical observations and assessments of patients, and the outputs are the treatments options at each stage, equivalent to the states and actions in RL, respectively. Apparently, applying RL methods to solve DTR problems demonstrates several benefits. RL is capable of achieving time-dependent decisions on the best treatment for each patient at each decision time, thus accounting for heterogeneity across patients. This precise treatment can be achieved even without relying on the identification of any accurate mathematical models or explicit relationship between treatments and outcomes. Furthermore, RL driven solutions enable to improve long-term outcomes by considering delayed effect of treatments, which is the major characteristic of medical treatment. Finally, by careful engineering the reward function using expert or domain knowledge, RL provides an elegant way to multi-objective optimization of treatment between efficacy and the raised side effect [6].

In the RL set-up, an autonomous agent, controlled by a machine learning algorithm, observes a state s_t from its environment at timestep t. The agent interacts with the environment by taking an action at in state s_t. When the agent takes an action, the environment and the agent transition to a new state s_{t+1} based on the current state and the chosen action. The state is a sufficient statistic of the environment and thereby comprises all the necessary information for the agent to take the best action, which can include parts of the agent, such as the position of its actuators and sensors. In the optimal control literature, states and actions are often denoted by x_t and u_t, respectively.

The best sequence of actions is determined by the rewards provided by the environment. Every time the environment transitions to a new state, it also provides a scalar reward r_{t+1} to the agent as feedback. The goal of the agent is to learn a *policy* (control strategy) π that maximizes the expected *return* (cumulative, discounted reward). Given a state, a policy returns an action to perform; an *optimal policy* is any policy that maximizes the expected return in the environment. In this respect, RL aims to solve the same problem as optimal control. However, the challenge in RL is that the agent needs to learn about the consequences of actions in the environment by trial and error, as, unlike in optimal control, a model of the state transition dynamics is not available to the agent. Every interaction with the environment yields information, which the agent uses to update its knowledge. This *perception-action-learning* loop is illustrated in Fig. 2 [7].

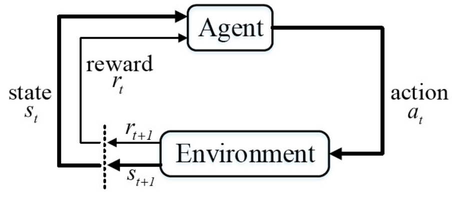

Fig. 2. The agent actuates with the mean. In anyone of the states of the mean, the agent takes an action that changes the state and returns a recompense [7]

Formally, RL can be described as a Markov decision process (MDP), which consists of:

_ A set of states ***S***, plus a distribution of starting states $p(s_0)$.
_ A set of actions ***A***.
_ Transition dynamics T (s_{t+1} /s_t; a_t) that map a state-action pair at time t onto a distribution of states at time t + 1.
_ An immediate/instantaneous reward function ***R***(s_t, a_t, s_{t+1}).
_ A discount factor γ €[0; 1], where lower values place more emphasis on immediate rewards.

A key concept underlying RL is the Markov property is that only the current state affects the next state, or in other words, the future is conditionally independent of the past given the present state. This means that any decisions made at s_t can be based solely on s_{t-1}, rather than {s_0; s_1,…, s_{t-1}}.

Although this assumption is held by the majority of RL algorithms, it is somewhat unrealistic, as it requires the states to be *fully observable*.

Challenges in RL

It is instructive to emphasize some challenges faced in RL:

_ The optimal policy must be inferred by trial-and-error interaction with the environment. The only learning signal the agent receives is the reward.
_ The observations of the agent depend on its actions and can contain strong temporal correlations.
_ Agents must deal with long-range time dependencies:

Often the consequences of an action only materialize after many transitions of the environment. This is known as the (temporal) credit assignment problem.

2.3. Deep Reinforcement Learning

Integration of deep neural networks into RL is a key factor in the success of Deep RL (DRL). It can automatically abstract and extract high-level features and semantic interpretation directly from the input data, avoiding complex feature engineering or delicate feature hand-crafting and selection for an individual task [6]

Many of the successes in DRL have been based on scaling up prior work in RL to high-dimensional problems. This is due to the learning of low-dimensional feature representations and the powerful function approximation properties of neural networks. By means of representation learning, DRL can deal efficiently with the curse of dimensionality, unlike tabular and traditional non-parametric methods.

For instance, convolutional neural networks (CNNs) can be used as components of RL agents, allowing them to learn directly from raw, high dimensional visual inputs. In general, DRL is based on training deep neural networks to approximate the optimal policy π*, and/or the optimal value functions *V*, Q** and *A**.

Although there have been DRL successes with gradient free methods, the vast majority of current works rely on gradients and hence the backpropagation algorithm. The primary motivation is that when available, gradients provide a strong learning signal. In reality, these gradients are estimated based on approximations, through sampling or otherwise, and as such we have to create algorithms with useful inductive biases in order for them to be tractable.

The other benefit of back propagation is to view the optimization of the expected return as the optimizations of a stochastic function. This function can comprise of several parts—models, policies and value functions—which can be combined in various ways. The individual parts, such as value functions, may not directly optimize the expected return, but can instead embody useful information about the RL domain. For example, using a differentiable model and policy, it is possible to forward propagate and backpropagate through entire rollouts; on the other hand, inaccuracies can accumulate over long time steps, and it may be pertinent to instead use a value function to summarize the statistics of the rollouts [7].

3. Results

3.1. Exist Application

The System of [3] builds on the fact that Parkinsonism gait abnormality responds to the medication. It works by putting a smartphone in the pocket with no special requirements, a universal behavior existing in our daily routine. A smartphone can continuously sense the gait information, and the active participation of the users is not required.

In this way, the system enables passive sensing. The system consists of a smartphone end and a cloud-server end, where the smartphone collects and transmits the raw gait data to the cloud server, and the cloud-server is then responsible for analyzing the data and feeding back the results to the smartphone. Based on a drug schedule assigned by the healthcare provider, the smartphone reminds the user to take medicine or informs the next medication time. Through this method, the system helps the patients to avoid the missed, mistimed, or extra doses.

3.2. Components of the Medication Adherence Monitoring System

Data Collector: Data collection are conducted in a nonclinical daily-life environment (e.g., at home or office). system utilizes the built-in inertial sensors (i.e., accelerometer and gyroscope) of a smartphone to collect gait information which responses to medicine medication for PD patients.

3.2.1. Inertial Sensors in a Smartphone

Accelerometer and gyroscope are two types of built-in inertial sensors measuring the inertial dynamics in three directions, namely the X, Y and Z axis.

Accelerometer: The three-axis accelerometer is built on the basic principle of acceleration, and is used to measure the orientation of a smartphone's acceleration (including the gravity) related to the surface of the Earth.

The accelerometer can gauge the orientation of a stationary item with respect to Earth's surface. In our study, the three-axis accelerometer measures the change of smartphone's linear velocity, and thereby reflects the movement of the PD patients.

Gyroscope: Although the accelerometer gauges the acceleration along with a particular direction, it provides little lateral orientation information with only the reference of gravity direction. Instead, the built-in three-axis gyroscope senses the angular velocity alone with one direction in the three-dimensional space. In our study, a PD patient generates both accelerations and rotations in different directions while walking. Correspondingly, gyroscope combined with accelerometer together provides us a powerful array of gait information.

Medicine Intake Detector: The system in [3] implements a medicine intake detector to achieve the monitoring of medication adherence. To begin with, a concatenation layer is responsible for data fusion from different sensors. Afterward, a residual network consisting of convolutional layers and fully-connected layers achieves the medicine intake detection. The architecture of intake detector is a multi-view convolutional neural network (MVCNN) containing the modules of multi-view data fusion, feature extractor, and medicine intake predictor (Fig. 3).

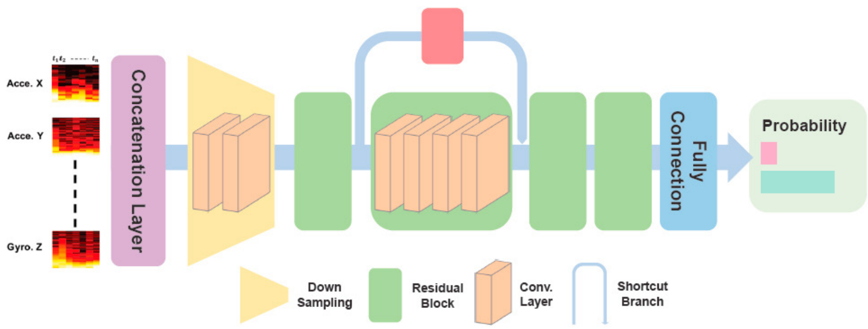

Fig. 3. The medicine intake detection in [3] is achieved by a multi-view CNN architecture. A concatenation layer carries out the data fusion from different sensors. A residual network consisting of four residual blocks achieves the feature extraction. A fully-connection layer and a SoftMax activation function perform the prediction.

This system marks a closer step towards passive medication adherence monitoring of PD in daily life [3]. However, it exhibits some limitations. First, current data collection is active and each participant is instructed by an APP to walk twenty steps. Second, it builds on the phenomenon that impaired gait responds to dopaminergic therapy. This performance may degenerate on subjects whose gait impairment are inconspicuous. Is necessary to develop a person-center protocol.

3.2.2. Deep reinforcement learning algorithm

The proposed Parkinson's Disease Classification and Medication Adherence Monitoring Using Smartphone-based Gait Assessment is implemented, for the intake detection, a deep reinforcement learning scheme to deal with complex situations dynamically, turning the classification task into a game. Also, was introduced a selective attention mechanism into the reinforcement learning scheme to focus on the crucial dimensions of the data [8] by the conception of Multi-Modality Sensor Data Classification with Selective Attention. This mechanism helps to capture extra information from the signal and thus it is able to significantly improve the discriminative power of the classifier. Considering that the signals in different categories may have different inter-dimension dependency is used a LSTM (Long Short-Term Memory) to exploit the latent correlation between signal dimensions. By exploring the dependency in sensor data is implemented a weighted average spatial LSTM (WAS-LSTM) classifier.

The advantages of use these methods for the intake detection are:

- The use of deep reinforcement learning to automatically select the most distinguishable features, called focal zone, for multimodal sensor data of different sensor types and combinations. The design of a novel objective function as the award in reinforcement learning task to optimize the focal zone. According to the authors [8], the new reward model saves more than 98% training time of the deep reinforcement learning.
- The use of Weighted Average Spatial LSTM classifier to capture the cross-dimensional dependency in multimodal sensor data.
- The creation of a selective attention mechanism for sensor data classification using the spatial information only. The proposed method is insensitive to sensor types since it is capable of handling multimodal sensor data.
- The proposed to use algorithm is shown in Fig. 4. The main focus of the algorithm is to exploit the latent dependency between different signal dimensions. To this end, the proposed approach contains several components: 1) the replicate and shuffle processing; 2) the selective attention learning; 3) the sequential LSTM-based classification.

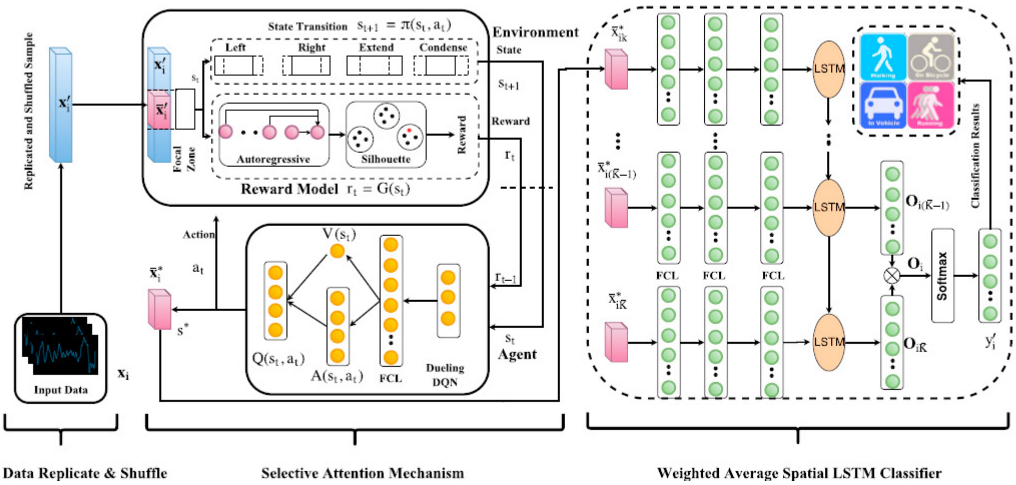

Fig. 4. Flowchart of the proposed approach. The focal zone $\bar{x}_i$ is a selected fragment from x_i' to feed in the state transition and the reward model. In each step *t*, one action is selected by the state transition to update s_t based on the agent's feedback. The reward model evaluates the quality of the focal zone to the reward r_t. The dueling DQN (Deep Q Networks DQN) is employed to find the optimal focal zone $\bar{x}_i^*$ which will be feed into the LSTM based classifier to explore the inter-dimension dependency and predict the sample's label y_i'. FCL denotes Fully Connected Layer. The State Transition contains four actions: left shifting, right shifting, extend, and condense [8].

The main focus of the proposed approach is exploiting the latent relationship between sensor signal dimensions. The signals belonging to different categories are supposed to have different inter-dimension dependent relationships which contain rich and discriminative information. This information is critical to improve the distinctive signal pattern discovery.

In practice, the sensor signal is often arranged as 1-D vector, which is less informative for the limited and fixed element arrangement. The elements order and the number of elements in each signal vector can affect the element dependency. In many real-world scenarios, the multimodal sensor data are associated with the practical placement where the optimal dimension sequence varies with the sensor types and combinations. To amend these drawbacks was proposed three techniques. First, replicate and shuffle the input sensor signal vector on dimension-wise in order to provide as much latent dependency as possible among feature dimensions. Second, introduce a focal zone as a selective attention mechanism, where the optimal inter-dimension dependency for each sample only depends on a small subset of features. The focal zone is optimized by deep reinforcement learning which has been proved to be stable and well performed in policy learning. Third, propose the WAS-LSTM classifier by extracting the distinctive inter-dimension dependency.

3.3. Data sets for experiments

- Daphnet Freezing of Gait Data Set (https://archive.ics.uci.edu/ml/machine-learning-databases/00245/). This dataset contains the annotated readings of 3 acceleration sensors at the hip and leg of Parkinson's disease patients that experience freezing of gait (FoG) during walking tasks.
- Wearable Computing: Accelerometers' Data Classification of Body Postures and Movements, (http://groupware.les.inf.puc-rio.br/har#dataset). Containing 165,633, one of which is invalid. There are five target activities, Sitting, Sitting Down, Standing, Standing Up, Walking.

3.4. Medication Management System

Adherence to medications is an important indicator of the quality of medication management and impacts on health outcomes and cost-effectiveness of healthcare delivery. We say above that the results from medicine intake detection, along with the medication effectiveness measurement, are transferred to the clinics automatically.

First, we need an information system for monitoring the status of patients with Parkinson’s disease. A possible architecture of that system is showed in Fig. 5.

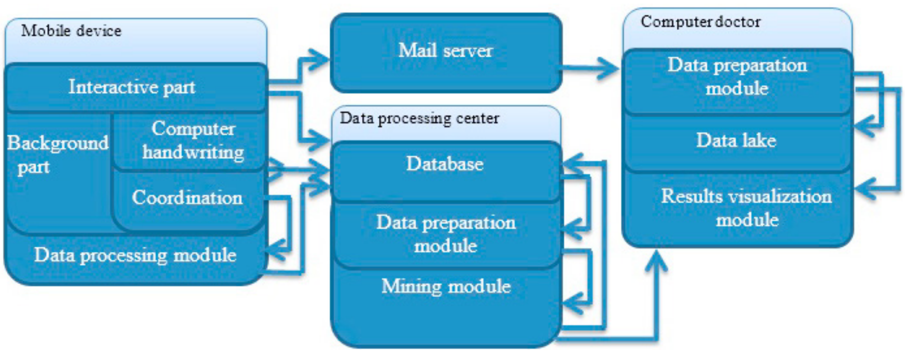

Fig. 5. Architecture of the information system for monitoring the status of patients with PD [9]

More specifically, there are two open source Medication Management System that can be used to receive, process and display the information obtained to the output of Deep Reinforcement Learning algorithms for the decision making. These systems are AdhereR [10] and MediPi [11].

- AdhereR, is a package for the widely used open source statistical environment R, designed to support researchers in computing Electronic healthcare data (EHD)-based adherence estimates and in visualizing individual medication histories and adherence patterns. AdhereR implements a set of functions that are consistent with current adherence guidelines, definitions and operationalizations. The package is freely available for use and its implementation facilitates the integration of medication history visualizations in open-source clinical decision support systems CDSS platforms.

- MediPi is an HSCIC-built telehealth system demonstrator, comprising hardware and software components enabling a secure, patient friendly and low-cost solution for monitoring of chronic medical conditions in the home. MediPi hardware's is a Raspberry Pi2 Linux based platform within a 7" touchscreen enclosure. The MediPi software is written in Java for maximum flexibility and can be deployed on any device – Android, PC, iOS, etc. Built with open-source software – making a flexible, configurable and extensible system allowing integration of any suitable medical device, either directly via USB or wirelessly.

4. Conclusions

The Parkinson's disease is a degenerative disorder of slow progress of the nervous system caused by the lack of the levels of dopamine which can provoke uncontrolled involuntary movements of the body and psychological affections.

In the article were proposed a system that make possible realize a *Dynamic treatment regimens (*DTRs) for PD patients, provide a new paradigm to automate the process of developing new effective treatment regimens for individual patients with long-term care. The system consists in a platform AdhereR or MediPi type using and smartphone for patient's gate data collection and a Deep Reinforcement Learning algorithm for classification according the unified scale of Parkinson's patients (Unified Parkinson's Disease Rating Scale - UPDRS) and monitoring medication adherence without the user's awareness. This work demonstrates the feasibility of using cell phones to monitor the biomarker walking in Parkinson's patients and this will be allowing the physicians to accurately adjust the treatment based on the patient's clinical response.

Acknowledgements

The Project has the subvention of the National Program "Automation, Robotics and Artificial Intelligence" of the Ministry of Science and Technology of Cuba. The reported study was funded by RFBR and CITMA according to the research project №18-57-34001.

References

[1] Plasencia A. et al. (2019). Methods, information sources and algorithms for the analysis of symptoms and support to the Parkinson disease patients. In: *IX Taller Internacional de Cibernética Aplicada*.

[2] Muthumanickam S. et al. (2018). Parkinson's Disease Detection And Classification Using Machine Learning And Deep Learning Algorithms – A Survey. *International Journal of Engineering Science Invention (IJESI)*, 7(5) ver. 1, pp.56–63.

[3] Zhang, H., Xu, C., Li, H., Rathore, Aditya Singh, Song, C., Yan, Z., Li, D., Lin, F., Wang, K. and Xu, W. (2019). PDMove: Towards passive medication adherence monitoring of parkinson's disease using smartphone-based gait assessment. In: *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 3, p23. doi: 10.1145/3351281.

[4] Thomas, P.J., Olufsen, M., Sepulchre, R. et al. (2019). Control theory in biology and medicine. *Biological Cybernetics*, 113, pp.1–6. doi: 10.1007/s00422-018-00791-5.

[5] Haeri, M., Yashar Sarbaz and Shahriar Gharibzadeh (2005). Modeling the Parkinson's tremor and its treatments. *Journal of Theoretical Biology*, 236, pp.311–322. doi:10.1016/j.jtbi.2005.03.014.

[6] Yu, C., Liu, J. and Nemati, S. (2020). Reinforcement learning in healthcare: A survey. [online] arXiv:1908.08796v4 [cs.LG]. Available at: https://arxiv.org/abs/1908.08796v4 [Accessed 24 Apr. 2020].

[7] Plasencia, A., Shichkina, Y., Suárez, I. and Ruiz, Z. (2019). Open source robotic simulators platforms for teaching deep reinforcement learning algorithms. *Procedia Computer Science*, 150, pp.162–170. doi: 10.1016/j.procs.2019.02.031.

[8] Xiang Zhang, Yao, L., Chaoran Huang, Sen Wang, Mingkui Tan, Long, G. and Can Wang (2018). Multi-modality sensor data classification with selective attention. In: *International Joint Conferences on Artificial Intelligence Organization*, pp.3111–3117.

[9] Shichkina Y., Kataeva G., Irishina Y. and Stanevich E. (2020). The architecture of the system for monitoring the status in patients with Parkinson's disease using mobile technologies. In: *Intelligent Distributed Computing XIII*, pp.531–540. doi: 10.1007/978-3-030-32258-8_62.

[10] Dima A. and Dediu D. (2017). Computation of adherence to medication and visualization of medication histories in R with AdhereR: Towards transparent and reproducible use of electronic healthcare data. *PLOS ONE*, 12(4), e0174426. doi: 10.1371/journal.pone.0174426.

[11] Robinson R. (2016). *MediPi Telehealth Demonstrator*.