Available online at www.sciencedirect.com

ScienceDirect

Cognitive Systems Research 52 (2018) 36–48

Cognitive Systems RESEARCH

www.elsevier.com/locate/cogsys

Optimized cuttlefish algorithm for diagnosis of Parkinson's disease

Deepak Gupta [a], Arnav Julka [a], Sanchit Jain [a], Tushar Aggarwal [a], Ashish Khanna [a], N. Arunkumar [b], Victor Hugo C. de Albuquerque [c,*]

[a] *Maharaja Agrasen Institute of Technology, Delhi, India*
[b] *Sastra University, Tanjore, India*
[c] *Graduate Program in Applied Informatics, University of Fortaleza, Fortaleza, CE, Brazil*

Received 14 May 2018; received in revised form 10 June 2018; accepted 12 June 2018
Available online 18 June 2018

Abstract

This paper presents an optimized cuttlefish algorithm for feature selection based on the traditional cuttlefish algorithm, which can be used for diagnosis of Parkinson's disease at its early stage. Parkinson is a central nervous system disorder, caused due to the loss of brain cells. Parkinson's disease is incurable and could eventually lead to death but medications can help to control symptoms and elongate the patient's life to some extent. The proposed model uses the traditional cuttlefish algorithm as a search strategy to ascertain the optimal subset of features. The decision tree and k-nearest neighbor classifier as a judgment on the selected features. The Parkinson speech with multiple types of sound recordings and Parkinson Handwriting sample's datasets are used to evaluate the proposed model. The proposed algorithm can be used in predicting the Parkinson's disease with an accuracy of approximately 94% and help individual to have proper treatment at early stage. The experimental result reveals that the proposed bio-inspired algorithm finds an optimal subset of features, maximizing the accuracy, minimizing number of features selected and is more stable.

Keywords: Parkinson's disease; Optimized cuttlefish algorithm; Feature extraction; Machine learning; Evolutionary algorithms

1. Introduction

Parkinson is a central nervous system disorder caused due to the loss of brain cells. The early symptoms of the disease are the trembling, tremors, stiffness, problem in walking and control of hands. The nerve cells are unable to send signals to each other leading to depression and other nervous system disorders. The detection of Parkinson at early stages is very important as there is no cure present. It leads to change in speech and handwriting patterns which can be useful in identification of the disease at an earlier stage. Recent works has shown that speech patterns of the patients have been proven beneficial to predict the same by finding correlation between min and max frequencies, measures of changing frequency variation etc. Recently a data set was prepared by Erdogdu Sakar et al. in which patients are asked to say only the sustained vowels 'a' and 'o' given in Erdogdu Sakar et al. (2013). Deep learning methods have also being used by many people to determine the Parkinson Disease. Apart from detection by speech pattern a second alternative is to study the handwriting patterns showcased in figures drawn by patients and healthy people. In this the individuals are given a handwriting exam in which they are asked to draw meanders and spirals both in clockwise and anticlockwise direction. Two copies are taken – one in the air and another on paper. A number of factors such as time taken to draw the figure, mean and

* Corresponding author.
E-mail addresses: deepakgupta@mait.ac.in (D. Gupta), ashishkhanna@mait.ac.in (A. Khanna), victor.albuquerque@unifor.br (V.H.C. de Albuquerque).

https://doi.org/10.1016/j.cogsys.2018.06.006
1389-0417/

maximum distance between exam template and handwriting trace, age, right/left handed etc. are then taken into account to determine if the particular individual falls into the category of a patient or a healthy person. Recently a study was conducted by Botucatu Medical School, São Paulo State University – Brazil and data from many individuals was collected and combined to form a data set given in Pereira et al. (2016). In this paper we have used modified Cuttlefish Optimization algorithm to determine whether the person has the disease or not based on the speech data collected in Erdogdu Sakar et al. (2013), Little, McSharry, Hunter, and Ramig (2009) and Pereira et al. (2016).

For the past few years, data is increasing day by day which in turn, is introducing many problems along with it. Data have been increased in terms of number of examples and also in the number of features, due to which it is getting harder to preprocess as well as process that data. Bigger data is more prone to noise which needs to be treated because if not, it could result in decreased performance on the result. Not only this, using that much data can result in higher computational costs and increased complexity. Hence, feature selection plays an important role in model construction in machine learning. In feature selection, aka variable selection, we choose a subset of relevant features from the set of features such that the accuracy of the model (after feature selection) is comparable to the original model (before feature selection). Feature selection reduces the computation cost of the model as well as complexity of the dataset. Along with these, it could be used to avoid the curse of dimensionality and make our model less prone to over fitting. Feature selection can be categorized in filter, wrapper and embedded methods.

Evolutionary algorithms are a part of artificial intelligence which mainly focuses on biological evolution. Biological evolution consists of four main processes namely reproduction, mutation recombination and selection. Unlike traditional optimization techniques, evolutionary algorithms depend on random sampling. These processes are repeatedly applied on the solutions formally stated as population and the fitness function is used to determine the quality of the solution. These solutions change according to the evolutionary process which eventually helps us to find the global solution to the problem. Evolutionary algorithms have been recognized to perform well under different circumstances as it does not assume the underlying fitness landscape. Even simple evolutionary algorithms can solve complex problems easily. Only drawback in evolutionary algorithms is the computation cost factor which can be reduced using fitness function approximation. There are many types of evolutionary algorithms, some of them are genetic algorithms, evolutionary programming, evolutionary strategies (https://en.wikipedia.org/wiki/Evolutionary_algorithm).

The work in the field of evolutionary algorithms and Meta Heuristic algorithms has been increasing in the recent past. These algorithms are mostly genetic based algorithms and nature inspired using the concepts of mutation and survival of the fittest based on the calculation of a fitness function which discards the unimportant solutions and further carries down good solutions down the genetic line optimizing them further. One of the algorithms- Firefly Algorithm proposed by Xin-She Yang uses flashing characteristics of fireflies (Xin-She, 2009). The same author has also proposed another bat algorithm based on bat echoing characteristics (Xin-She, 2010). Many other algorithms include ant colony optimization algorithm (Gonzalez-Pardo, Jung, & Camacho, 2017), gravitational search algorithm based on the relationship between mass and gravity (Pham et al., 2005) have been inspired from laws of nature, bees algorithm by Pham et al. given in Esmat (2009) and CAB algorithm by Erik Cuevas et al. given in Erik, Mauricio, Daniel, Marco, and Guillermo (2012).

Using the above ideas, we have introduced a novel improved and optimized version of traditional cuttlefish algorithm (OCFA) inspired by the color changing behavior of cuttlefish depending on the surroundings and features such as visibility and reflectivity of light, and combined it with feature extraction. OCFA is an improved and optimized version of original CFA (Adel Sabry Eesa, 2013). We have been able to optimize the proposed algorithm using the different fitness functions as compared to the original algorithm. The OCFA has been applied on Parkinson's speech and HandPD datasets for identification of Parkinson's disease at early stage which helps individual to have proper treatment at early stage. The algorithm efficiently extracted all the perpetual attributes and gave an average accuracy of 92%.

Online smart healthcare technologies using IoT has become a big field of research today. Some of the methods have been devised in Elhoseny et al. (2018). One such example is given in Vardhana, Arunkumar, Abdulhay, and Ramirez-Gonzalez (2018) where convolutional neural networks were used for bio-medical image segmentation with hardware acceleration. Use of IoT based technologies can also be used in this smart healthcare system (Rodrigues et al., 2018). OCFA can be implemented online so that it can give a reasonable prediction to the input drawings in real time. Further studies can be conducted on creation of virtual environments using augmented reality and similar concepts to treat the Parkinson's Disease given in Silva et al. (2017) and Oliveira, Fernandes, Pinheiro, Ribeiro, and De Albuquerque (2016), and Gupta, Sundaram, Khanna, Hassanien, and de Albuquerque (2018). When dealing with health care data online, it is vital to secure it. Therefore methods to ensure secure data transmission in IoT based health shall be implemented (Lins, Oliveira, Rodrigues, & De Albuquerque, 2018). New algorithms have been designed and they can be modeled so that they could be applied to the Parkinson's Image dataset. For example, Rebouças Filho, Cortez, Barros, De Albuquerque, and Tavares (2017a) shows an adaptive method applied in the segmentation of lung images and Rebouças Filho, De Albuquerque, and Tavares (2017b) shows a new method to extract features from medical images.