Human Activity Monitoring for Mental Health Assessment

Design Proposal

Team Members:
Phuoc Nguyen,
Sunny Qin,
Virginia Pohl,
Meskerem Terega,
Xinyi Yan

Advisor: Nikolaos Papanikolopoulos

February 22, 2019

Table of contents

1. Abstract

Experiments done in clinical research shows that mental illness and neuropsychiatric disorders like anxiety, OCD (Obsessive Compulsive Disorder), autism, etc. are critical public health issues, as these mental illnesses are a leading cause of self-suicide, injury, and disease for people around the world. This neuropsychiatric disorder shows behavioral disorder at very early ages for example, according to the National Institute of Mental Health, signs of autism spectrum disorder start in early infancy for some children and some of the symptoms are lack of eye contact, repetitive movements such as hand flapping, rocking, etc. Therefore, early diagnosis and treatment can make a world of difference in the progression of mental illness. The prognosis of such mental disorder requires consistent monitoring for abnormalities in the activity of the person. Considering the availability of such specialized medical doctors to analyze the activity of the at-risk person for an hour compared to the volume of patients, along with the high-cost, there is a strong motivation for investigating technological alternatives to solve the problem. The advancement of devices/software algorithm for aiding such monitoring activities will be an essential step towards effective utilization of the diagnostic resource. In this project, we intend to provide a technique to monitor human activity automatically in real time and assess mental health in an easier way instead of high-cost manual diagnosis. We approach the problem from a computer vision standpoint and we use a camera or smartphone for tracking the activities of a person with a mental disorder in the video frame.

2. Introduction

Mental disorders have a significant impact on individual, families, and society. Mental disorders among children are an important public health issue because of their prevalence, early onset, and impact on the child, family, and community. Based on the Centers For Disease Control and Prevention (CDC) data [1], in the United States about 13 – 20% of children are experiencing mental disorder at an early age. Besides, in 2010, the leading cause of death among children aged from 12 to 17 is suicide which can result from interaction of mental disorder and other factors. Also, Research has shown that mental disorder can interfere with social functioning and negatively impact physical well-being [1], therefore, the early diagnosis will always be better for effective treatments. While approaching towards how to find the type of serious behavioral abnormality at a very young age, there is a strong motivation for investigating a device/ software that monitor the activity of the person to relate to one neuropsychiatric disorders.

According to the National Institute of Mental Health [2], mental disorders are disabilities related mainly with the execution of the neurological system and brain. Some of the examples are anxiety, OCD, hyperactivity disorder, autism and conduct disorder. Earlier diagnosing mental disorders can be made, if there is a system to detect specific symptoms. This project is for early detection of the symptoms. Users can take videos of patients behaviors using smartphones or sensors, and send the videos to the system. The system can detect the specific movements including their angles and frequency, and generate a list contains the mentioned information. Videos will be sorted in an order according to the movement lists and sent to doctors for further diagnosis. In this case, the detection of mental disorders will be brought forward. It can also shorter the time of doctors when diagnosing patients, since their behavior will be provided in prior with brief analysis.

3. Prior Work

Pose detection has been studied for decades. In the past few years, people focused on two main fields, single person pose estimation, and multi-person pose estimation. For single person pose estimation, there are several architectures they used such as tree-structured graphical models [4], non-tree models [5], and convolutional architectures [6]. The basic idea of these is that "combining local observations on body parts and the spatial dependencies between them" can give an inference on the pose estimation [3]. Tree-structured graphical models and non-tree models both work on improving the accuracy of spatial dependencies while using convolutional neural networks(CNNs) can increase the reliability of local observations of body parts. In multi-person pose estimation, most approaches use a top-down strategy [7]. These algorithms, however, are difficult to capture spatial dependencies between different people. Therefore, some proposed bottom-up strategy [8], which associates part detection candidates to an individual. This method can work well on multi-person estimation. However, it may cost much time to run on a single image. There is another method widely used in this field right now, named part affinity fields (PAFs) [9]. This representation is used in OpenPose [10], an open source library which can achieve real-time pose estimation on multiple people. OpenPose can take a variety of inputs, including images and videos, and for every person detected, provide keypoint estimation of their body and/or face. The body keypoints OpenPose estimates are points of major interest along the body, and include the knees, hips, shoulders, elbows and more. OpenPose also draws the vectors between the keypoints, connecting the feet to the knees, and then the knees to the hips and so on. These keypoints can then be further analyzed.

In the last year, there is a group [11] working on a similar topic. Their work includes video processing, OpenPose keypoint extraction, keypoint analysis, and action classification. They use OpenPose to extract key point of a person including 25 body keypoints, 2*21 hand key points, and 70 face key points [10]. For action classification, they apply the convolutional neural network and KTH dataset to classify the movements of the key point. During their work [11], they successfully classify activities related to neuropsychiatric disorders within KTH dataset. Also, their system can work very well at a distance ranging from 1 meter to 20 meters and at any video quality. The limitation of this design is that the dataset they used is not enough for disorder detection. The KTH dataset is limited because it only classifies six human actions and does not include activities related to disorders of interest [12]. Moreover, the design had a three minute time delay between capturing the video and processing it. Ideally, the processing would happen within a few milliseconds, providing virtually real time feedback about the video.

4. Requirement Specifications

4.1 Problem Statement

Clinical visits for diagnosing children with neuropsychiatric disorders can be time consuming and costly and not every parent has access to these. Yet, early detection is becoming even more imperative as the cases of neuropsychiatric disorders rises. The goal of this project is to aid doctors in diagnosing children by enabling parents to take videos of their children and sending the video through a computer vision algorithm that will look for key behaviors. These behaviors are related to neuropsychiatric disorders and the analysis of the video should help with a medical professional making a diagnosis.

4.2 Customer Needs

#	Customer Requirement
1	The system shall include a user interface
2	The system shall be able to detect and track single person
3	The system shall store all data collected (video and analysis)
4	The system shall process the video collected to monitor human behavior
5	The system shall be able to filter between human actions that meet the target behavior and actions that are not the target behavior.
6	The system shall provide real-time feedback about the human behavior
7	The system should be able to detect single person in crowded environment

4.3 Product Design Specification

#	Metric	Units	Ideal Value	Acceptable Range

1	Tracking accuracy	pixels	10	±2
2	Tracking distance	feet	10	9-11
3	Image quality	Megapixels	12	8-12
4	Camera focal length	millimeters	4.3	4-6
5	Response time	milliseconds	4	3-5
6	Movement speed	feet/s	7	±10%
7	Success Rate	%	80	75-100

5. Proposed Design

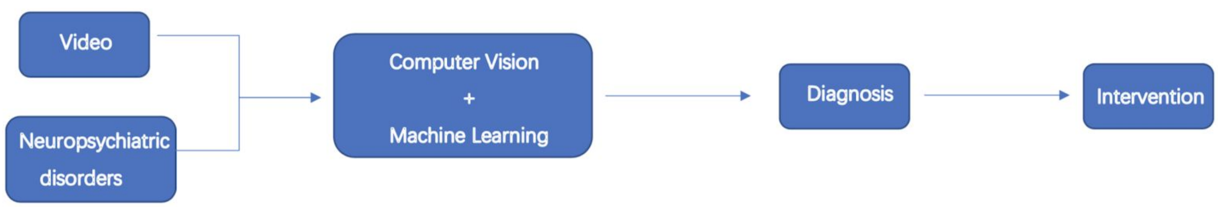

Figure 1. Proposed Design

Our project will base the design on the procedure outlined in Figure 1. Given the real-time video as the input, our design can provide accurate detection of human behaviors. Compared with the typical actions of the neuropsychiatric disorder, our system can help diagnose this disorder based on our detection. After diagnosis, the next step is intervention, which will require medical professionals.

Our system uses computer vision and machine learning algorithms to detect human behaviors. By using these algorithms, our system can achieve feature extraction, behavior classification and evaluation on behaviors. Also, our design should detect real-time movements under several scenarios:

1. A single person in the scene
2. Multi-person in the scene

Compared to prior work, we may provide a new algorithm to improve the accuracy and efficiency on different scenarios' detection. Also, a new dataset may be applied to our system for training to classify more human behaviors.

6. Budget

This project is based in software development and will use video captured from pre-owned phones and cameras. The software platform is available to us at no cost. Therefore, no hardware or software purchases should be necessary. The group does have a $300 budget if this changes.

7. Timeline

Section 1 Familiarize ourselves with image processing	02/04/19	02/14/19
Sub-task 1 - learn basic knowledge by seeing notes	02/04/19	02/07/19
Sub-task 2. use Matlab to do edge detection, edge thinning, edge following and line fitting	02/04/19	02/14/19
Milestone1 - take a picture and finish edge detection by using Matlab	02/14/19	02/14/19
Section 2 Familiarize ourselves with the CAVE and tracking system	02/15/19	03/01/19
Sub-task 1 get connection with medical students	02/15/19	02/18/19
Sub-task 2 go to lab and familiar with CAVE	02/19/19	02/22/19
Sub-task 3 Develop the software that monitors humans and human activities in the CAVE.	02/23/19	03/01/19
Milestone 2- know how to use the equipments in lab	02/22/19	02/22/19
Section 3 Design Work	02/23/19	04/02/19
Sub-task 1 decide one neurodevelopmental disorder	02/23/19	02/25/19
Sub-task 2 look for related human behaviour	02/25/19	03/01/19
Sub-task 3 understand the existing computer vision algorithms for this project	02/25/19	03/01/19
Sub-task 4 Software design		
Sub-task 5 Hardware design	03/14/19	03/14/19
Milestone 3- a report about human behavior of that specific disorder	03/04/19	03/04/19
Section 4 Development Work		
Sub-task 1 develop the hardware in CAVE	03/05/19	03/15/19
Sub-task 2 develop the algorithm in CAVE	03/05/19	03/15/19
Sub-task 3 develop the hardware in public settings	03/16/19	03/29/19
Sub-task 4 develop the algorithm in public settings	03/16/19	03/29/19
Milestone 4- a report about the development of hardware and algorithm	04/02/19	04/02/19

Section 5 Testing with human subjects	04/03/19	04/20/19
Sub-task 1 Hardware: test that hardware is accurately collecting the data	04/03/19	04/08/19
Sub-task 2 Software: test that algorithm can process this data	04/03/19	04/08/19
Milestone 5- a brief report on testing result	04/10/19	04/10/19
Sub-task 3 improve the hardware design	04/11/19	04/18/19
Sub-task 4 improve the software design	04/11/19	04/18/19
Milestone 6- finish all the testing part, write a brief report on how to improve and what has improved	04/20/19	04/20/19
Section 6 Conclude our achievements and prepare to present them	04/21/19	04/30/19
Sub-task 1 poster draft	04/21/19	04/23/19
Milestone 7- poster finished	04/24/19	04/24/19
Sub-task 2 prepare presentation	04/21/19	04/25/19
Milestone 8- presentation	04/25/19	04/25/19
design show	04/30/19	04/30/19

8. Communication

The team is structured with a lead communicator with the advisor and resources. The project group meets on regular class time every Tuesday and Thursday. Also, the group meets every Monday from 2:30 to 4:00 to discuss ideas and to move our project forward and with the advisor Professor Nikolaos Papanikolopoulos on Thursday from 3:30 to 4:00 to be on track and to report where the team is so far on the project. The team also arrange more time if needed to work on things related to the project before the deadline of task the group planned. The group use Google Drive to share documents and files, GitHub to share code, Slack to update the group with a short text, and email as needed.

9. Conclusion

In this paper, we proposed the application of our system for monitoring the activity of a child towards the goal of finding severe behavioral abnormalities. We provide a details of our system design from the video input to intenvation. Our system uses a computer vision and machine learning algorithms to detect human behavior in real-time using the input from the video. In order to detect multiple people in the video, our system uses representation called affinity fields which is used in OpenPose that take a variety of image for every person detected. Our project can have a significant impact on medical diagnosis in neuropsychiatric disorders can make a world of difference in the progression of mental illness.

References

[1] “Mental Health Surveillance Among Children - United States, 2005–2011.” *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 2013.
[2] “Chronic Illness & Mental HealthExternal.” Bethesda, *National Institutes of Health*, National Institute of Mental Health. 2015.
[3] Cao, Zhe, et al. “OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields.” 2018.
[4] P. F. Felzenszwalb and D. P. Huttenlocher, “Pictorial structures for object recognition,” in *IJCV*, 2005.
[5] X. Lan and D. P. Huttenlocher, “Beyond trees: Common-factor models for 2d human pose recovery,” in *ICCV*, 2005.
[6] S.-E. Wei, V. Ramakrishna, T. Kanade, and Y. Sheikh, “Convolu- tional pose machines,” in *CVPR*, 2016.
[7] L. Pishchulin, A. Jain, M. Andriluka, T. Thorma ̈hlen, and B. Schiele, “Articulated people detection and pose estimation: Reshaping the future,” in *CVPR*, 2012.
[8] L. Pishchulin, E. Insafutdinov, S. Tang, B. Andres, M. Andriluka, P. Gehler, and B. Schiele, “Deepcut: Joint subset partition and labeling for multi person pose estimation,” in *CVPR*, 2016.
[9] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, “Realtime multi-person 2d pose estimation using part affinity fields,” in *CVPR*, 2017.
[10] G. Hidalgo, Z. Cao, T. Simon, S.-E. Wei, H. Joo, Y. Sheikh, and Y. Raaj, “Openpose.” *GitHub Repository*, 2018.
[11] E. Low, A. Widjaja, M. Johnson, A.-A. Lawati, C. Patterson, “Senior-Design.” *GitHub Repository*, 2018.
[12] I. Laptev and B. Caputo, “Recognition of human actions.” *KTH*, 2005