UTILIZING IMAGE SEGMENTATION FOR THE EXTRACTION OF HEART RATE VARIABILITY IN THE PURSUIT OF UPDATING MENTAL HEALTH THERAPEUTIC VALIDITY PROCESSES

by

Lauren E. Johnson

A thesis submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Master of Science in
Electrical Engineering

Charlotte

2019

Approved by:

_____

Dr. James M. Conrad

_____

Dr. Chen Chen

_____

Dr. Douglas Markant

ABSTRACT

LAUREN E. JOHNSON. Utilizing Image Segmentation for the Extraction of Heart Rate Variability in the Pursuit of Updating Mental Health Therapeutic Validity Processes. (Under the direction of DR. JAMES M. CONRAD)

Image and signal processing modalities play critical roles in the diagnosis and treatment efficacy of medical complications in people today. However, these tools have not yet been utilized to aid in the clinical diagnosis or treatment of neurocognitive disorders, such as Attention Deficit Hyperactivity Disorder (ADHD). The subjective nature of diagnosing neurodevelopmental disorders like ADHD has many merits. Yet, there is inadequate training for many doctors that can diagnosis ADHD without being a specialist in psychological disorder or disabilities, which leads to a concern for such subjective processes arising from those struggling to find a diagnosis and therapy to aid them in daily life. Therefore, a need for objective measurements to aid in the treatment and diagnoses processes exists. The objective of this work is to create an inexpensive method to be utilized in doctors' offices and research to aid in the treatment efficacy and eventual diagnosis processes of neurodevelopmental disorders, beginning with ADHD. This thesis work analyzes the potential for various image and signal processing modalities, and highlights the viability of a computer vision modality. Part of the goal and inspiration for this work is to eventually utilize various measurements from a single camera, beginning with the physiological measurement of heart rate variability (HRV). HRV provides great insight into the autonomic nervous system, which is affected in persons with neurodevelopmental disorders. This work therefore focuses on improving upon previous works that came close to a distanced distinction of HRV detection for a person via a face camera, and compares the results of these algorithms to other HRV capturing methods.

DEDICATION

The author dedicates this work to the late Dr. Bharatkumar S. Joshi, who strived for using engineering principles to better the medical field and world around them. Dr. Joshi's legacy for excellence in engineering and joy for learning and sharing knowledge is recognized by the author as a constant source of inspiration to this day.

Additionally, this work is dedicated to family and friends of the author who strive for the inclusivity of people with all abilities, and for the physical/mental health of all individuals in all shapes and forms.

ACKNOWLEDGEMENTS

The author would like to acknowledge the support and assistance of their advisor, Dr. James M. Conrad, and the other members of the advisory committee, Dr. Chen Chen and Dr. Douglas Markant. Additional acknowledgement goes to other supporters of the project. Dr. Andrew Willis taught many of the courses that led to the author's pursuits in signal processing and medical technology.

Volunteer and personal experiences with early childhood education for individuals with various mental and physical capabilities also led to a distinct interest in aiding others. From these experiences and educational backgrounds in engineering and cognitive science, the author pursued the following topic to aid in the process of finding the right diagnosis and treatment plan for all persons in a faster and objective manner.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Psychology and Medical Related Terms

AAP	The American Academy for Pediatrics
ACE	Adverse Childhood Experience
ADHD	Attention Deficit Hyperactivity Disorder
ANS	Autonomic nervous system
ASD	Autism Spectrum Disorder
CEM	Cognitive-energetic model
CHADD	The society Children and Adults with ADHD
$\mathrm{Da}_{Ri/Ti}$	Dopamine receptor/transmitter (respectively), i = number
DAv	Delay aversion
DDM	Drift-diffusion model of decision-making
DMN	Default mode network
DSM-5	The Diagnostic and Statistical Manual of Mental Disorders, 5th edition
ECG	Electrocardiogram
HFA	High-functioning ASD
HRV	Heart Rate Variability
MAHNOB-HCI	Dataset created for emotion tracking and determination
NSCH	National Survey of Children's Health
PNS	Parasympathetic nervous system
PRV	Pulse Rate Variability
SR	State regulation
TPD	Temporal processing deficits

Method and Signal Related Terms

AC	Alternating current; a signal's non-zero (positive) frequency components

DC	Direct current/component; a signal's time average (mean) or dependent value
EVM	Eulerian video magnification method
FFT	Fast Fourier transform
HF	High frequency range
IBI	Interbeat interval
iPPG	Imaging or video-imaging PPG signal
KLT	Kanade-Lucas tracking algorithm
LF	Low frequency range
OUGP	Ornstein-Uhlenbeck third-order Gaussian process filtering method
PBV	Blood-volum pulse method
Pk-Pk	Peak-to-peak interval
PP	Pulse-to-pulse intervals
PPG	Photoplethysmography
RGB	Red, green, and blue channels of visible spectrum cameras
ROI	Region of Interest
rPPG	Remote PPG methodology
RR	Interval between consecutive heart beats
ULF$*$	Ultra-low frequency range
VLF	Very-low frequency range

Analytics and Units

α	Skin-tone value or metric of varying skin types or pigmentation; used in the EVM method
λ_c	Cut-off frequency
Hz	Hertz, unit of measure for frequency
r	Pearson's correlation coefficient
MAE	Mean absolute error

MPE	Mean percentage error
RMSE	Root mean squared error

CHAPTER 1: INTRODUCTION

Since the days of the ancient Romans, physicians began recognizing the relationship of cranial injuries with disruption to one's mental capabilities, whom the Roman physician Galen would term as the lesion method [1]. Each branch or practice within the medical field now is able to work with psychologists to find links between brain regions and mental functions, leading to the development of such fields as seen in cognitive science. Cognitive scientists utilize concepts from the disciplines of linguistics, artificial intelligence, philosophy, psychology, neuroscience and artificial intelligence [2].

With cognitive science findings, the scientific and medical communities are recognizing the associative nature of the "mind-body" connection. A new example of this recognition is seen with the findings that gingivitis may be the cause of plaque developing on the brain in Alzheimer's patients, meaning that preventing gingivitis with good oral health care could help prevent the development of Alzheimer's [3]. Yet just as good "physical" health affects the mind - such as lack of food or sleep can cause emotional outbursts and improper concentration - so can the mind help affect the body. These connections between the physical and mental states of a person aid in many diagnoses to this day, particularly for neurological disorders or disabilities.

Some research has gone into using physical signals for objectifying mental health related diagnoses, such as using speech analysis to detect Parkinson's disease [4]. However, there are limited numbers of this type of research, and thus far has mostly focused on neurodegenerative disorders. In an effort to continue searching for objective measurements to add to the research community's discoveries of determining and adapting one's mental health, the author began looking into various neurocognitive disorders. The results of this investigation are covered in the following paragraphs as well as in the author's paper on the topic [5].

1.1 Purpose

Individuals with Attention Deficit Hyperactivity Disorder (ADHD) - a neurodevelopmental disorder - are affected by symptomology like inattention [6, 7], hyperactivity [8], and impulsivity [9], which affect everyday performances such as learning and academic/work performance [6, 8, 10]. While seemingly high functioning, these persons struggle compared with their non-disordered peers; they are more likely to have lower grade point averages, higher dropout and unemployment rates, and more difficulties with their concentration, memory, and time management skills [11, 12, 13].

The Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5), describes the medical classification of ADHD as a persistent pattern of inattention and/or hyperactivity-impulsivity that interferes with functioning or development [14]. Two subtypes of ADHD exist: inattentive and hyperactive/impulsive. The inattentive diagnosis primarily focuses on lack of attention or sustained attention, while the hyperactive diagnosis involves excessive activity through talking, fidgeting, and inability to remain still during leisurely activities.

The DSM-5 gives descriptors of symptomology and requirements to be met for ADHD to be the proper diagnosis. Yet, it does not give any indication on correct methods or tasks to use during the diagnosing process, beyond incorporating educational and parental/guardian information and observations. The American Academy for Pediatrics (AAP) also follows suit of the DSM by providing guidelines which lack specificity for the actual process for diagnosing and treating children with ADHD [15]. Typically a psychologist would utilize a general neuropsychological examination which assesses multiple domains, such as intelligence, visuo-spatial, and attention [16]. Sources speak about the tasks utilized during such examinations, but there are multiple options for each cognitive domain in a neuropsychological evaluation, adding to the subjectivity and potential biasing of the evaluator. However, such biasing is limited for experts (psychologists and psychiatrists) compared to general-practitioners

or pediatricians who primarily follow the guidelines provided by the AAP and DSM. According to the society Children and Adults with ADHD (CHADD), utilizing census data collected from the 2016 National Survey of Children's Health (NSCH), only 14% of children in the U.S. were diagnosed by an expert [18].

Since, these criterion are met through subjective measures, misdiagnoses become a concern for many patients, doctors and researchers. Studies have validated this trend for misdiagnosing populations such as children with differing subtypes of ADHD [8]. Other studies indicate the trend for misdiagnosing a patient with ADHD who in fact has another condition, such as:

- autism spectrum disorder (ASD) [19]
- scotopic sensitivity syndrome (SSS; also known as Meares-Irlen syndrome), whose characteristics include distortions of print when reading, such as the text appearing to move or vibrate, along with reduced word recognition and decreased ability to maintain reading over longer periods of time [20, 21]
- environmental factors like overcrowded classrooms or relative maturity generalizations [22]
- at least one of the adverse childhood experience (ACE) scores [23], as well as many others beyond the scope of this thesis [24]

It is noteworthy to emphasize that while these factors could be separate from ADHD, the authors do not claim they could not be separate from ADHD, merely that other causes or diagnoses provide more encompassing explanations for many individuals. This also includes a comorbid diagnosis of ADHD with one or more learning disabilities or other cognitive disorders.

A recent study [25] utilizing census data collected from the 2016 NSCH found that of an estimated 6.1 million U.S. children 2-17 years of age, 9.4% had received an ADHD diagnosis at one point; only 5.4 million, or 8.4% of all U.S. children were indicated as

having ADHD at the time of the survey. While this does not mean the 11.6% difference of children were neurotypical and received an ADHD diagnosis, it does provide insight into the rate of misdiagnois of ADHD when a different diagnosis was appropriate. These rates also do not account for those who received diagnoses in adulthood, or diagnoses that excluded ADHD when a comorbid diagnosis with other cognitive or neurological disorders should have been given. Studies also indicate concern with whether treatments prescribed to patients with ADHD (ie, stimulant medications or non-pharmacological therapy) effect said patients' cognitive performance, along with their biological structures like the cardiovascular system [26].

1.2 Objective of this Work

The purpose of this work is to create an inexpensive method that could be utilized in doctors' offices and research to aid in creating more objective measurements for evaluation in the treatment efficacy of neurodevelopmental disorders, beginning with ADHD. The work uses image processing over other methods for the purpose that such a method will be helpful for assessing patients who do not like other people or items touching them (e.g., children with ASD). This also aids with creating a methodology that could be quickly and inexpensively implemented within even a typical general practitioner's office. This work focuses on updating previous work on pulse rate variability (PRV) [27, 28] and remote heart rate monitoring for physical activity assessment [29]. Additionally, this work is utilizing emotion tracking databases for full range of physiological data recorded from multiple modes of medical and non-medical devices [30].

1.3 Contribution

This thesis works toward creating more objective, inexpensive models for physiological measurements to be utilized in research and clinical treatment efficacy programs for neurodevelopmental disorders by implementing a heart rate variability (HRV) im-

age processing model that utilizes data from a typical webcam. By beginning with HRV, a physiological signal that provides insight for cognitive and physical function, the method can be implemented for multitudinous diagnosis or therapeutic/treatment processes in the future. In using a typical webcam that does not need to touch the individual and records facial data of a person, this increases the modularity of the method for future work such that other measurements recorded through the images would be available.

1.4 Organization

This thesis is assimilated into five chapters. Chapter 1 illustrates the motivation for the presented work along with a brief introduction to its relevant topics. Chapter 2 reviews the background information of signal processing for treatment efficacy of ADHD, cognitive and medical standings on HRV, and previous works on HRV models of various measurement modalities. Much of the content in Chapter 2 was also mentioned in the author's paper [5]. Chapter 3 covers the algorithms and methodologies utilized to implement this work. Chapter 4 analyzes the results of this work. Chapter 5 discusses this work's conclusions drawn from the full breadth of this thesis, along with future work to follow.

CHAPTER 2: BACKGROUND

The purpose of this chapter is to review literature, also discussed in [5], utilizing technology for treatment efficacy and diagnosis studies in order to provide an informed system design which will reduce misdiagnosis rates and improve the general diagnosis and treatment processes. The reasons for neurocognitive/neurodevelopmental misdiagnoses often originates from a lack or misguided understanding for the underlying cognitive mechanisms resulting in these various conditions; these are compounded upon each other along with a misunderstanding of each condition's diagnosis and treatment efficacy resources. Knowing these complications, caused by focusing on symptomology, leads to the following discussion for a more detailed road of understanding the mechanisms leading to said symptomology.

2.1 Objective Data for Therapy Validity for ADHD patients

Much of the information in this section comes from the author's publication [5], with adjustments for this thesis.

2.1.1 Cognitive Background

As ADHD relates to neurodevelopment, the nature of the brain must be discussed and regions which could impact this disorder.

Studies have indicated a link between dopamine receptors and executive functioning (EF), which there are fewer of (in number or functionality) in persons with ADHD [31]. Cognitive models explaining ADHD also help support this link with the dopaminergic system. One model explains how participants with ADHD prefer to choose options which hasten time to reward, known as delay aversion (DAv), which predicted many hyperactive subtypes explaining their impulsive or reward sensitive behaviors [32]. Another model explains inhibitory control deficits, or executive dysfuntion and impulsivity [33] which was the more historical view of ADHD. As the overall dopaminergic system influences multiple brain regions (e.g. amygdala, portions of the basal gan-

glia, prefrontal cortex, etc.) effecting cognitive abilities like working memory, reward, attention and motor control, the connection to ADHD seems essential [1].

A sole dopamine related explanation excludes parts of the ADHD spectrum. In fact, the noradrenergic (norepinephrine) system also affects arousal, attention and working memory in certain brain regions (e.g. hypothalamus, thalamus, etc.), which helps explain more ADHD symptoms, i.e. emotional regulation. A cognitive model that supports the inclusion of the noradrenergic system is the cognitive-energetic model (CEM). The CEM involves arousal mechanism disruptions, which effects sustained attention, in people with ADHD [34].

There is also the multi-pathway model, which shows that inhibitory deficits, DAv, and temporal processing deficits (TPD) are all dissociable yet influential toward describing ADHD subtypes [35, 36]. TPD also incorporates the noradrenergic system as it also transmits to the temporal lobe, along with the dopaminergic system, thus adding to the validity of the heterogeneous multi-pathway model. A new additional model is the drift-diffusion model of decision-making (DDM), where slower event rates or reaction times was indicative of higher inhibitions and slower information processing [37, 38]. The DDM shows consistency with the multi-pathway model supporting multiple facets for ADHD brains.

Many of the aforementioned studies utilized reasoning based on medical imaging modalities to back up their claims for the cognitive models of ADHD. Said research aided in the discovery of an understanding of cognition for persons with ADHD. Therefore, numerous studies are looking at different imaging and signal processing modalities to find a more objective way of not only researching, but understanding treatment efficacy with stimulant medications. Genetic advancements and testing have also been employed for determining if a person has or is predisposed to developing ADHD. However, this paper does not cover the topic as the authors define genetic testing as not being apart of image or signal processing modalities (for some reference

to genetic or neuroscience studies please see [31]). The following sections discuss current research focused on developing treatment and diagnostic aids for ADHD, as well as any concerns with each modality.

2.1.2 Medical Imaging and Signal Processing

The meta-analysis by Sridhar et al.[39] covers most of the medical modalities utilized at discovering differences between neurotypical persons and those with ADHD. Primary modalities include magnetic resonance imaging (MRI), functional magnetic resonance imaging (fMRI), electroencephalography (EEG), single photon emission computed tomography (SPECT), positron emission tomography (PET), and discussion of the genetic markers underlying ADHD. The studies discussed by Sridhar et al. help provide a greater understanding of the differences of ADHD and neurotypical brains, particularly the effects of iron levels affecting dopaminergic regions of the brain found in PET/SPECT studies the authors reviewed. The EEG studies Sridhar et al. covered generally mentioned high accuracy rates for measurements from the frontal lobe regions and improved predictability of treatment efficacy when also utilizing machine learning techniques.

Resting fMRI studies showing the blood-oxygen levels and low-frequency default mode network (DMN; default brain activities used for comparison in fMRIs) were also indicative of a vast effect on ADHD brains, showing many brain regions that differed from the default neurotypical brain [39]. Other multi-modal studies, combining fMRI and typical structural MRI studies also support the vast areas impacted by ADHD. Such indications were discovered by utilizing machine learning techniques to create better classifiers which found that ADHD participants have reduced gray matter volume primarily in the frontal lobe, and differing visio-spatial and DMN connectivity versus their neurotypical counterparts [40, 41, 42]. Volume size in areas like the amygdala and hippocampus also contribute to some ADHD symptoms related to those regions [43]. A primary stimulant, methylphendiate, used in treating ADHD

was used in fMRI studies and shown to aid significantly in mediating attentional control complication for inhibitory tasks [44].

These modalities helped show the complex and variable regions that ADHD affects the brain. As such varying modalities (from fMRI to EEGs) show this, it goes to prove even further that multi-deficit models are more indicative of ADHD than any single deficit model. Many of these modalities are utilized in research to support cognitive models of ADHD and cognition in general, but other techniques should be considered if they can validly provide diagnosis or treatment efficacy of ADHD.

2.1.3 Other Image and Signal Modalities

The following section reviews imaging and signal modalities that do not look at the brain, but rather at other physiological attributes indicative of brain behaviors.

2.1.3.1 Imaging

A link between microsaccades, small fixational saccades, and prestimulus anticipation has been established [45, 46]. The ability to inhibit microsaccades when expecting a stimuli, something persons with ADHD struggle with, is correlated with improved task performance [45]. Gottlieb et al's paper [47] backs up these claims by showing that information-seeking cognitive models use gaze patterns, as they differ when acquiring a task and when learning. Their paper also mentions that gaze is utilized in monitoring uncertainty or predictability of the surrounding environment.

Two varying techniques for measuring microsaccades during computerized tasks of attention were accomplished by Fried et. al. [48], and Danker et. al. [49]. Each utilized a different form of continuous performance test (CPT) which involved looking for targets and/or non-targets in a region of interest (ROI) as a test of attention for a participant. Each of these have some merits, yet both have some pitfalls in their methodology as well.

One inconsistency with [49] findings was that nearly a third of the Danker et al.

ADHD participants showed better sustained attention performance than the median of all their participants. This variability for the ADHD population on saccade effects could be due to either medication taken by participants (as the authors did not screen for this conflict), or their ability to hyper-focus on some task. However, another reason is that those participants did not suffer from poor sustained attention as much as other symptoms of ADHD.

Danker et al. also discussed that their findings supported ADHD participants having fewer microsaccades and more saccades than neurotypical participants. This contradicts the findings of Fried et al. who used a different rendition of an attentional testing task. A possible explanation for these conflicting discoveries could be that Danker et al. presented their targets and non-targets centrally (the shapes and colors took up the center of the screen), while Fried et al. presented their targets and non-targets peripherally (9°, vertically away from center). Studies confirm that ADHD participants struggle with fixation tasks and show excessive saccadic intrusions when instructed versus not instructed to fixate [50, 51].

2.1.3.2 Physiological

Another cognitive model that requires mentioning, is the state regulation deficit (SRD) model. The SRD resembles similarities to the other multiple deficit models, as it discusses the regulation ability of the energetic states of arousal, activation and effort. According to a study by van der Meere [52], this model helps explain studies showing slow response times in persons with ADHD as the SRD model indicates the deficiencies of effortful maintenance for optimal activation levels required for maintaining a consistent state of motor preparation. Additionally, much of state regulation pertains to physiological data, such as heart-rate, breath-rate, eye movements, etcetera - each of which may require various measuring techniques to record.

An interesting investigation by van der Meere involved looking at the heart rate variability (HRV) of persons with ADHD, captured with an electrocardiogram (ECG)

modality. The author discussed that midfrequency band of 0.03 to 0.15 Hz involves modulation by the sympathetic and parasympathetic (vagal) influences, and lower amplitudes of these frequencies indicates a more effective compensatory SR. The author found that participants with ADHD and HFA exhibited larger amplitudes than their control peers for the midfrequency band particularly during slow stimulus presentations over fast presentations. Thus, ADHD participants allocated less effort in the slow condition and had difficulty remaining motivated.

Recent studies have also investigated the relation between ADHD and the cardiovascular system. One study showed the use of a remote HRV monitor (ie, ECG) such that a longer evaluation which followed the patient through routines could be evaluated for treatment efficacy [53]. Another use for such remote diagnostics for treatment efficacy comes from studies indicating a longer adverse effect on cardiovascular systems for persons taking ADHD stimulant medications [26].

A benefit of a camera modality for eye-tracking is that other physiological factors can also be determined with a camera. Studies have proven that heart-rate can be measured by visible-spectrum cameras. Effectively, this means that measuring HRV would also be plausible with the camera (discussed further in Section 2.2, thus adding to the utility of the visible-spectrum imaging modality [54].

2.1.4 Decision Matrix

Based on the discussed technological modalities, the authors have summarized their findings in Table 2.1 with some adjustments from the original found in [5]. The table implements the multiattribute utility theory (MAUT), also known as a decision matrix [55]. The categories for ranking include:

- the general accessibility of cost and safety (modalities which could be harmful to subsets of persons, will be penalized),
- time (length of process; mostly for procedure and analysis),

Table 2.1: Results of Review of Technology Trends for ADHD Diagnosis and Treatment Efficacy

Modality	**Information**		**Safety**	**Price**	**Time**	**Total**
	Spectral	***Temporal*2***				
PET & SPECT	4	2	1	2	3	12
EEG	3	10	4.5	3	3.5	24
MRI	5	2	3	2	3	15
fMRI	3.5	9	3	2	3	20.5
Physiological	1.5	8	5	4.5	4.5	23.5
Computer Vision[a]	1.5	9	5	5	4.5	**25**

[a] Computer Vision refers to the eye-movement and HRV factors being measured with a visible-spectrum camera.

- how much information into the ADHD process a modality can provide in terms of spatial and temporal resolution

Each category will be given a rating of 1 (poor) to 5 (high). However, the temporal information column is given a weight of two as ADHD cognitive functionality can only be represented well via temporal data during cognitive tasks.

The results of Table 2.1 show that modalities which provide brain analysis are given higher scores for spatial and temporal information. Modalities which can provide more details into specific brain functionality during tasks give higher temporal resolution and understanding. The visible-spectrum camera modality had the best total score, based on its accessibility making it more cost-effective and reducing the time spent for data collection and analysis, along with adding the potential for other physiological measurements than typical physiological capturing methods. While other modalities provide better insight into the brain effects of ADHD on a spatial level, measurements of eye movements and HRV captured with computer vision techniques, still provide valuable indicators for treatment efficacy, and thus, diagnosis of ADHD as well. Therefore, this modality will be utilized for the primary modality of this work.

2.2 History of Heart Rate Variability

There are studies about the physical impact of working the body (i.e., exercising) on HRV and autonomic nervous system (ANS) [29, 56], but what about the gymnastics of the mind? The remaining portions of this chapter will review the deeper connections of HRV with ANS and cognitive capabilities. From there previous work on creating remote HRV measuring modalities are covered, along with any datasets that accompany these works as well.

2.2.1 Physiological Signals and Their Implications

To understand the connection between ADHD and HRV, the ANS and its subsystems must first be reviewed. The ANS involves maintaining the body's homeostasis (especially internal), meaning it pertains of all systems related to automatic or involuntary responses of the body. The ANS breaks down into the two subsystems known as sympathetic and parasympathetic as depicted in Figure 2.1. The sympathetic system is responsible for the secretion of adrenaline and noradrenaline (norepinepherine), which aids in accelerating HR while dilating (ie, relaxing) the bronchi of the lungs. On the other hand, the parasympathetic nervous system (PNS) conserves and restores energy by slowing the HR and constricting the bronchi primarily through the vagus nerve; vagal tone, therefore, is directly associated with proper heart and lung strengthening, as in, aerobic exercise [57].

Each of these systems continues to compliment each other not just at the abstract level of expending or conserving energy, but also down at the neurotransmitter level. Serotonin, dopamine, norepinephrine, and epinephrine receptors are all of the general monoamine, where the latter three are catecholamines as they all are one step from each other in structure. These receptor systems are each linked with some form of arousal and attention (whether excitatory or inhibitory) for cognitive and physical systems, and therefore fall under the sympathetic system. Acetylcholine systems

are inversely related to the monoamine systems for the physiological structures, and the primary receptors of the PNS are the muscarinic acetylcholine (ie, cholinergic) subtype.

As the ANS subsystems suggest, the physiological relation to neurotransmitter systems that closely effect cognitive abilities often hindered by ADHD, such connections can be utilized to aid in understanding ADHD's effect on more than cognitive performance. Though many physiological measures, such as breath rate, show a connection with the ANS, only heart related measures are connected with the dopaminergic system. This further supports the conclusions made in [5] which found HRV to be an informative metric for the treatment efficacy of persons with ADHD, yet now we must define exactly what that entails.

One final correlation to the ANS, is its role in the fear-neurocircuitry system. This system includes regions such as the hippocampus and ANS - regions previously discussed as being negatively influenced in persons with ADHD. Due to the fear-neurocircuitry's succinct relation to emotional responses, emotion elicitation trials would serve as a good performance metric for establishing the differences in different emotion tasks which could infer the difference experienced by patients with ADHD during varying attentional/cognitive tasks.

The 1996 Taskforce [58] expressly defines HRV as the oscillation within the time interval between consecutive heartbeats (RR). Essentially, HRV does not simply look at the beats per minute as HR does, but in fact looks primarily at the interbeat intervals (IBI) and its variability. The HRV signal is further broken down into time and frequency domain components for full analysis (full details are covered in Table 2.2). While all of these measurements aid in the understanding of HRV, the frequency domain has particular relevance to physiological health. Firstly we define several frequency ranges as the following: ultra-low (ULF), very-low (VLF), low-frequency (LF) and high-frequency (HF), and additionally the range ultra-low-star (ULF$*$). The

Table 2.2: Standard HRV measurements as determined by the 1996 Task Force on HRV. Adapted from [58, 60]

Variable	Units	Description
Time Domain, Statistical Measures		
SDNN[a]	ms	Standard deviation of all NN intervals
SDANN[a]	ms	Standard deviation of the averages of NN intervals in all short-term (typically, 5 min) segments of the entire recording
RMSSD[a]	ms	Root mean squared of the sum of the squares of differences between adjacent NN intervals
SDNN index	ms	Mean of the standard deviations of all NN intervals for all short-term (typically, 5 min) segments of the entire recording
SDSD	ms	Standard deviation of differences between adjacent NN intervals
NN50 count		Number of pairs of adjacent NN intervals differing by more than 50 ms in the entire recording
pNN50	%	NN50 count divided by the total number of all NN intervals
Time Domain, Geometric Measures		
HRV triangular index[a]		Total number of all NN intervals divided by the height of the histogram of all NN intervals measured on a discrete scale with bins of 1/128 s
TINN	ms	Baseline width of the minimum square difference triangular interpolation of the highest peak of the histogram of all NN intervals
Differential index	ms	Difference between the widths of the histogram of differences between adjacent NN intervals measured at selected heights
Logarithmic index	ms^{-1}	Coefficient ϕ of the exponential curve $ke^{\phi t}$, which is the best approximation of the histogram of absolute differences between adjacent NN intervals
Frequency Domain, Short-term Recordings (2-5 min)		
Total power	ms^2	Variance of all NN intervals ($\approx\leq 0.4Hz$)
ULF*[b]	ms^2	Power in the ULF range ($0.0005 \leq f \leq 0.003Hz$)
VLF	ms^2	Power in VLF range ($f \leq 0.04Hz$)
LF	ms^2	Power in LF range ($0.04 \leq f \leq 0.15Hz$)
LF norm	n.u.	LF power in normalized units: $100 * LF/(Totalpower - VLF)$
HF	ms^2	Power in HF range ($0.15 \leq f \leq 0.4Hz$)
HF norm	n.u.	HF power in normalized units: $100 * HF/(Totalpower - VLF)$
LF/HF		Ratio LF/HF
Frequency Domain, Long-term Recordings (24 h)		
Total power	ms^2	Variance of all NN intervals ($\approx\leq 0.4Hz$)
ULF	ms^2	Power in the ULF range ($f \leq 0.003Hz$)
VLF	ms^2	Power in the VLF range ($0.003 \leq f \leq 0.04Hz$)
LF	ms^2	Power in the LF range ($0.04 \leq f \leq 0.15Hz$)
HF	ms^2	Power in the HF range ($0.15 \leq f \leq 0.4Hz$)
α		Slope of the linear interpolation of the spectrum in a log-log scale ($f \leq 0.01Hz$)

[a] These measurements are the most indicative values of HRV for statistical calculations (in the time-domain) according to [58].

[b] This value was added for reference to work done by [59], but was not defined by the Task Force of 1996.

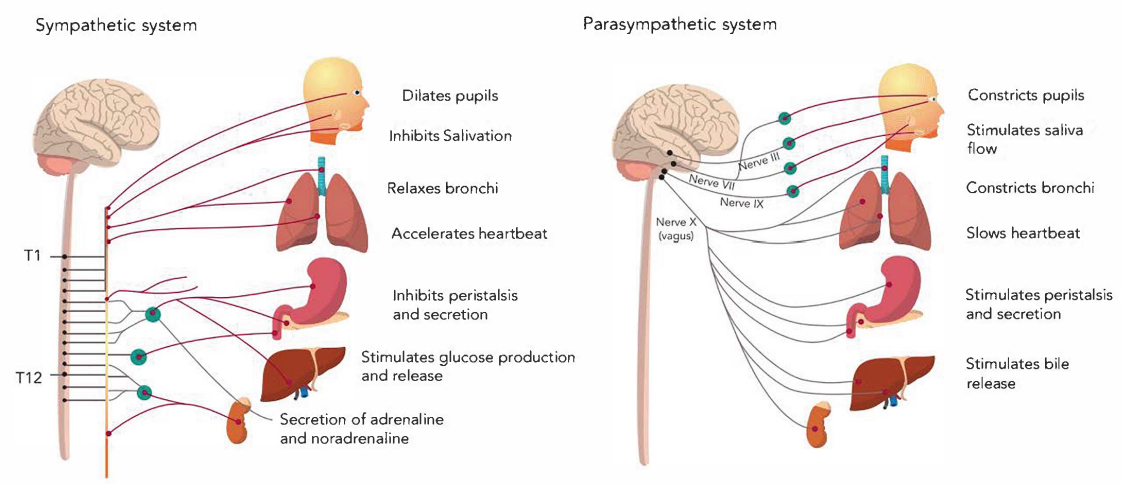

Figure 2.1: Sympathetic (left) and Parasympathetic (right) Anatomy and Effects on Organ Systems. Adapted from [57].

1996 Task Force on HRV [58] mentions that ULF and VLF ranges are typically only looked at for 24 hour recordings, while the LF/HF ratio is only reviewed for short-term (2-5 minutes) measurements. LF/HF ratio is not modulated by physiological mechanisms that are considered stationary, and therefore during longer recordings this measurement can be determined by breaking up the recording into shorter segments (eg, 2-5 minutes) then taking the average over all segments.

In a 2015 article [60], the work from the 1996 Task Force [58] was reviewed in an attempt to update some of the HRV analysis techniques. The frequency bands for the frequency domain HRV measurements, were certainly unchanged from the two articles which span nearly 20 years. Yet, this contradicts the ranges utilized by recent work [57] that cited the 2015 [60] article. However, most researchers maintain the original ranges from the 1996 Task Force for their analyses, whether working with designated HRV measuring devices or working toward developing new methods to utilize other modalities such as optical sensors.

Another crucial measurement to understand is pulse rate variability (PRV). The PRV signal is obtained by measuring the pulse-to-pulse (PP) intervals from the peaks of RRs; this refers to utilizing a peak-to-peak (Pk-Pk) analysis. While this PRV

measurement seemingly appears to be a surrogate for HRV, as some studies have indicated [28], this serves as more of an interpolation of the HRV signal than being a true representation or replacement.

2.2.2 Modalities and Methods

One of the primary techniques many articles reference (such as [57]) pertains to photoplethysmography (PPG) which utilizes optical signal processing to detect the blood volume changes in the microvascular tissue bed just under the skin's surface. The final waveform includes two components; the first being the pulsatile physiological signals that are correlated with the blood volume changes with each heart beat, and the second component is the baseline with various LF portions related to respiration, sympathetic activity and thermoregulation. As [52, 53] and their references reported lower parasympathetic and higher sympathetic responses in ADHD patients, the expectation would be to see such responses in all forms of HRV measurements and modalities.

Some PPG methods include using phone cameras that analyze heart-rate and HRV by having the user place a finger on the camera and remaining still [61]. Researchers also utilize the accelerometer in phones or wearable HR measuring devices (eg, Fitbits, Apple Watches) to remove motion artifacts from the signals [57, 62]. The next step in this evolution was toward an optical method of HRV measurement that did not involve contact to a user's body.

Remote photoplethysmography (rPPG) enables contactless heart-rate monitoring using optical sensors, such as a regular video camera [29]. In a recent study utilizing the rPPG method [63], a general overview of previous work on utilizing rPPG for HR measurements was given. One confusing aspect about this method however is the existence of a similarly named method - imaging/video-imaging photoplethysmography (iPPG). The only distinction from each method is that iPPG is a subclass of rPPG that specifically works with RGB spectrum cameras, while rPPG could use

methods from the near-infrared spectrum [64], or five color channeled cameras [28]. However, these terms seem to be interchangeable [28]. Beyond the terminology used to reference working with light detecting sensors, the next step in using this method for heart related measurements goes to the type of method for extrapolating such data from imaging modalities.

2.2.2.1 Methods Designed for HR and Variability Estimation

Primarily, the previous work will begin with the Voila-Jones [65] method to detect the face in video frames, followed by some form of region-of-interest (ROI) selection for the specific areas of the skin to include for the signal. Some work has been done to aid in this ROI selection process, including: using discriminative response map fitting [66, **?**], and adaptive patch selection [67]. Other works employ the Kanade-Lucas Tracking (KLT) algorithm as well to improve the processing speed of their algorithms [68]. After this first phase, comes the second phase of chrominance estimation and rPPG signal extraction, which could involve various algorithms.

There are many known algorithms for rPPG signal extraction, which could involve various additional methods used alongside them. Simplistic algorithms for rPPG signal extraction include: only looking at the green channel (G), the green subtracted from the red channel (G-R), independent component analysis (ICA), principal component analysis (PCA). Most work to this point employs the pulsalite blood volume (ie, blood volume pulse; PBV) method [69]. PBV uses the dependency of the output response of a camera on the skin pigmentation of a person, the light source's emission spectrum and internal RGB filters of the camera in order to calculate the RGB components of the blood volume vector. Not all implementations of "PBV" utilize each of the three dependencies, yet the concept for removing the background/external dependencies of the rPPG signal generally refers to the PBV method. However, other methods add additional metrics to use in calculations for enhanced signal retention.

The plane-orthogonal-skin method (POS) [70] builds on the PBV method. POS

projects the temporally normalized signal that is orthogonal to the normalized RGB space and skin-tone, making it skin-tone independent. A method known as sub-band rPPG (SB), improves upon POS by creating frequency bins, or sub-bands, using fast Fourier transform (FFT) to be independently analyzed across each color channel which is critical for HR estimation during physical activity [29]. This work was expanded into SB with continuous wavelet transform (SB-CWT) instead of using FFT for the signal decomposition [63]. Eulerian video magnification (EVM) produces motion or color magnifications (using spatial or temporal filtering, respectively) with first-order Taylor series common in optical flow or tracking analyses, like that of KLT [54]. Other methods include: Self-Adaptive Matrix Complex (SAMC) [71], full video pulse (FVP) extraction [64], etcetera.

After interpolating the skin-tonal variations into an impulse for each video frame, methods continue to diverge for exactly how to filter the signals to their final forms. Most variations pertain to different types of bandpass filters, how many filters, and data interpolation or noise/motion artifact removal. The work referenced in this section typically looked at determining rPPG methods with fitness videos, though some looked at HR estimation using conditions where a participant is sitting while performing cognitive tasks [**?**]. With these methods toward HR rPPG estimation, the direction now turns toward efforts for using rPPG methods for HRV signal extraction.

The biggest advancement toward HRV estimation via optical sensors has been the pulse rate variability (PRV) method [27, 28]. In a recent study [28], a sumarization of additional PRV/rPPG methods was provided with the PRV signal determined from peak-to-peak intervals of a post-process signal (ie, filtering applied after the initial data recording). The HRV metrics determined by the 1996 Task Force always are utilized in these studies, though most only use some of the metrics. For example, the only study referenced that calculated for key HRV frequency and time domain metrics (refer to Table 2.2) and tested participants in stress conditions utilized a five

color channeled camera, increasing the monetary cost of the method. While HRV measurement through visible-spectrum cameras currently has limited research, the work that has been achieved thus far provides a great step to HRV-rPPG.

2.2.3 Datasets

Through each of the aforementioned studies, the research had to include data for comparison. Some work created their own datasets that were not made publicly available, yet many worked with data that has been made publicly available (either to all or solely to academic researchers). In a recent study [63], the authors mention multiple datasets suitable for rPPG studies. One such dataset being the MMSE/MMSE-HR dataset created by [71], with the HR subset being especially made for the use in rPPG research advancements.

Another dataset heavily utilized being the well referenced MAHNOB-HCI Tagging database, which is a multimodal database recorded in response to affective stimuli with the goal of emotion recognition and implicit tagging research [30]. This database includes face videos, audio signals, eye gaze data, EEG signals and peripheral nervous system physiological recordings, such as ECG recordings. The MAHNOB-HCI dataset has been used for emotion research, but also in rPPG work. In [68] the author uses this dataset for HR estimation using the color camera recordings and ECG data for ground truth. This study even mentioned using the dataset for any potential future work on HRV implementation for cameras.

A fully public dataset comes from Physionet, and was expressly implemented in the recent PRV studies [27, 28]. Though there are even more potential datasets that could be utilized for imaging HRV algorithm testing, those explicitly mentioned here have been worked with for such work making them great testbench datasets.

CHAPTER 3: METHOD

This chapter reviews the methods and algorithms utilized in this work. First, a more detailed review of previous methods relative to this work will be discussed for better understanding toward the choices of the final method implemented. Such review also bears weight on any oversights of past works, whether they be explicit mistakes or implicit for potential user misinterpretation. These reviews then lead into the final method for this work, followed by the metrics used for the final analysis of the method's outcomes.

3.1 Previous Methods

All methods tend to utilize the Voila-Jones algorithm for detecting the face in videos/images [65]. Further ROI editing and selection occurs from there before moving into additional steps for selection of the iPPG signal. Yet all rPPG methods always need to have the unadulterated RGB signal, which is calculated as:

$$C_{face_i} = \frac{1}{mn} \sum_{j=1}^{m} \sum_{k=1}^{n} I_{i_{jk}} \tag{3.1}$$

in which C_{face_i} denotes the i-th spatial RGB color channel signal (ie, raw RGB signal); $I_{i_{jk}}$ represents the pixel intensity value I at each j-th row and k-th column location for every i-th color channel, where i increments by one for each possible color in ascending order (ie, R = 1, G = 2, and G = 3). These rPPG methods begin to vary from here, whether by amplifying skin-tone color changes, removing illumination variations of the environment for noise reduction, or some other slight alteration to filtering the iPPG signal.

EVM for heart-rate extraction, for example, utilizes temporal filtering to amplify the skin-tone fluctuations [54]. One drawback though for their open-source method comes from needing to determine the correct alpha (α) coefficient for each skin-type and individual, to pass into their algorithm. The authors indicate that α may be

forced to zero in order to amplify lower frequencies for wavelengths below a spatial cutoff frequency (λ_c); this could mostly be useful for motion amplification, such as determining breath rate. For pulse amplification their α values tend to be between 100-120 for adults going from skin-type I-III (ie, fair to dark) respectively. Yet for infants the α variable was higher (150) for lighter skin-toned children; this is more than likely due to the faster heart-rate for infants and young children. Inversely realted to the effects of skin-type and age on α is their effects on λ_c, where the adult with lighter skin-tone had the highest λ_c (1000) and the infant had the lowest (600). These amplifications with the EVM method make HR easy to distinguish, however, there are also skin-tone independent methods that significantly reduce the concern of debugging individual trials proper α values.

Other research for HR/PRV extraction with rPPG methods look to eliminate any dependency on factors such as environment lighting. In one such study the raw RGB signal is first corrected for the background luminescence before temporal filtering and normalization.For illumination rectification of a raw RGB signal $C_{face_i}(j)$, the equation is defined as:

$$C_{IR_i} = C_{face_i} - hC_{bg_i} \tag{3.2}$$

such that $C_{bg_i}(j)$ is the background luminescence at each j-th video frame or time increment, and h being a weighting filter. The Normalized Least Mean Squared (NLMS) filter was employed to first find the optimal h to eliminate motion artifacts of the PPG signal, which is defined by:

$$h_i(j+1) = h_i(j) + \frac{\mu C_{IR_i}(j) C_{bg_i}(j)}{C^H_{bg_i}(j) C_{bg_i}(j)} \tag{3.3}$$

with the stepsize coefficient of μ (found to be 0.0003 in [68]), $C^H_{bg_i}(j)$ as the Hermitian transpose of $C_{bg_i}(j)$, and the normalizing factor of $C^H_{bg_i}(j)C_{bg_i}(j)$ being the input energy. From equation 3.3, the optimal h can now be applied back into equation 3.2,

followed by temporal filtering and normalization of the signal.

Many other studies work with a variation of rPPG algorithms, all of which begin with normalization to illiminate the dependency on α, unlike EVM [72, 63, 73]. This equation makes the iPPG signal become skin-tone independent by detrending the DC component. Temporally normalizing the raw RGB-signal C (ie, AC/DC-1) is defined as:

$$\overline{C_i} = \frac{C_i}{\mu(C_i)} - 1 \tag{3.4}$$

where $\overline{C_i}$ denotes the zero-mean color variation of the i-th channel, and $\mu(\cdot)$ represents a typical moving average operator.

3.1.1 PRV Algorithms and Limitations

When comparing rPPG methods alongside an rPPG PRV implementation [28], the authors decided to implement ICA methods over those like EVM based on findings from [74] that ICA provided better results with the HRV metrics. However, the filter parameters from [74] are not fully explained, so the findings could be due to their filtering methods used in conjunction with EVM and ICA rather than the fault of those rPPG methods themselves.

An additional important distinction refers back to the discussion of what HRV and PRV signals actually define and represent, back from section 2.2. Though the authors of this PRV-iPPG method [27, 28] run their analyses using the HRV time and frequency domain metrics, does not mean that the PRV method usurps the HRV method. This point is crucial to understand going further, as the Pk-Pk interval variations do not equate to methods which look into additional interval changes of a raw signal. Therefore, after an initial band-pass filter, instead of up-sampling the video's PPG signal from the sampling frequency of the camera to the sampling frequency of the ECG signal (ie, the cubic interpolation step of this PRV method resampling to 1 kHZ from 15 Hz), the reference signal should be down-sampled to

the rate of the video signal rate for better analytical comparison.

One such method which could aid in the full HRV analysis is the Ornstein-Uhlenbeck third-order Gaussian process (OUGP) filter [59]. This filter serves as an improvement of the authors' original detrending filter algorithm known as the smoothing by Gaussian process priors (SGP) [75]. In more recent HRV studies [68], detrending filters are a part of the temporal processing steps of the overall rPPG design. However, these PRV implementations analyze the iPPG signal in an opposite way by drawing conclusions from information which was highly interpolated signals.

The primary reason for the PRV method to up-sample their video signal comes from the recommendation to record HR information at a sampling rate of 250 to 500 Hz or higher [58]. Such sampling rates are due to a typical HR max of 240 bpm, so a minimum of 250 Hz (ie, 1/250 cycles per second) would ensure capture of such information. Yet, Nyquist's theorem states that in order to capture a signal well, the sampling frequency should be twice that of the frequency desired for capture and analysis, which explains the 500 Hz max. Since most ECG type of devices record at 1 kHz for optimal signal understanding, the choice to up-sample the 15 Hz video rate (ie, 15 fps) to 1 kHz is understood. However, by interpolating/resampling the video signal by such a large factor (ie, a factor of 66.667) the information inferred for the HRV analyses could be incorrect based on this large sampling rate gap interpolating information that may not exist.

Other concerns regarding this PRV work relates to the utilization of the PhysioNet dataset [27, 76]. This PhysioNet dataset only provides preprocessed video signals, which is bad for testing other ROI algorithms, including the calculation of background illumination for luminescence correction. The work also never explicitly covers the frequency ranges they utilized for their HRV analyses. Though the authors cite the 1996 Taskforce, not redefining the work calls into question if the correct ranges were used.

The suggested sampling rate ranges for HRV analyses are also backed up by [61, 77], though their source for an accurate sampling rate mentioned the minimum frequency as 200 Hz instead of 250 Hz. This discrepancy could be due to quoting a source that came before the 1996 Taskforce publication, or a difference in the methodology and technology of that era. Altini et al vary from Rodriguez by interpolating the video signal only the 180Hz from 30 Hz; this means the interpolation increases by a factor of 6. The authors choose this frequency as their design uses typical PPG methods with an iPhone camera against a user's finger, and the iPhone has the capacity to compute at 180Hz easily.

The iPhone PRV method also only does the interpolation step for the frequency domain HRV analyses and not for the time domain [61]. Upon further investigation of the rPPG-PRV work [27, 28], this interpolation step was never mentioned as being used for only the frequency domain analyses. The rPPG-PRV publications do state that the Kubios HRV analysis software is utilized for their frequency domain analyses while the time domain metrics are calculated in MATLAB. The authors of the rPPG-PRV papers could have realized that the interpolated signals are aligned in time; yet, this does not however clear up the new concerns for the results from these PRV methods since the time domain analyses are done differently. However, the interpolation step may not be a problem if the reference signal is down-sampled instead of up-sampling the video signal.

3.2 Method and Algorithms

Firstly, this work utilizes the MAHNOB-HCI dataset [30] for testing the HRV method. As this dataset is well cited and publicly accessible (to academic researchers), this already creates an ideal scenario for continued research. MAHNOB-HCI also utilizes a typical RGB-spectrum camera facing the participant and provides physiological measurements from ECG to EEG signals and eye-gaze tracking. The trials also relate to emotion based cognitive tasks, making an all around ideal dataset for this work

and its usage for future work in expanding to other physiological signal extraction with computer vision techniques, as well as other modalities for testing (see Figure 5.1).

Before continuing into the final algorithm, this dataset's full utility must be further broken down. The MAHNOB-HCI dataset culminates thirty participants over forty emotion trials each. Each odd trial (ie, every other trial) is a neutral emotional trial that lasts approximately ten to twenty seconds each and is meant to bring the subject back to baseline in between emotionally charged trials of approximately two minutes each; the charged trials are elicited from video clips that would bring a person joy or some other emotion. There are a total of three fear trials from different climatic moments of classic horror films (ie, *The Shining* and *Silent Hill*) which are tested against each subject and is important to note for the work of this thesis. Only the subjects who gave permission for publication and had zero errors during trials were utilized during testing of this thesis. The final flow diagram for the algorithm can be found in Figure 3.1, though the following subsections will provide greater detail about each step of this process.

3.2.1 Algorithm of rPPG Methods

Using Viola-Jones method to detect the face, which leads into using the KLT algorithm for improved processing speed of the Viola-Jones algorithm. From there, the ROI of the face is further simplified by detecting the eyes region, such that the final ROI of just the lower portion of the face where eye blinking contributing to noise in the iPPG signal is eliminated. Figure 3.2 demonstrates this ROI selection process, for a subject whose forehead region would be less informative due to the obstruction of the skin by the participant's hair. This is done in order to improve the quality of the raw skin-average signal as some subjects have hair on their forehead or beards/mustaches, and thus keeps the algorithm modular for the various potential interference of hair impeding the ROI. A video sequence is then saved with the edited

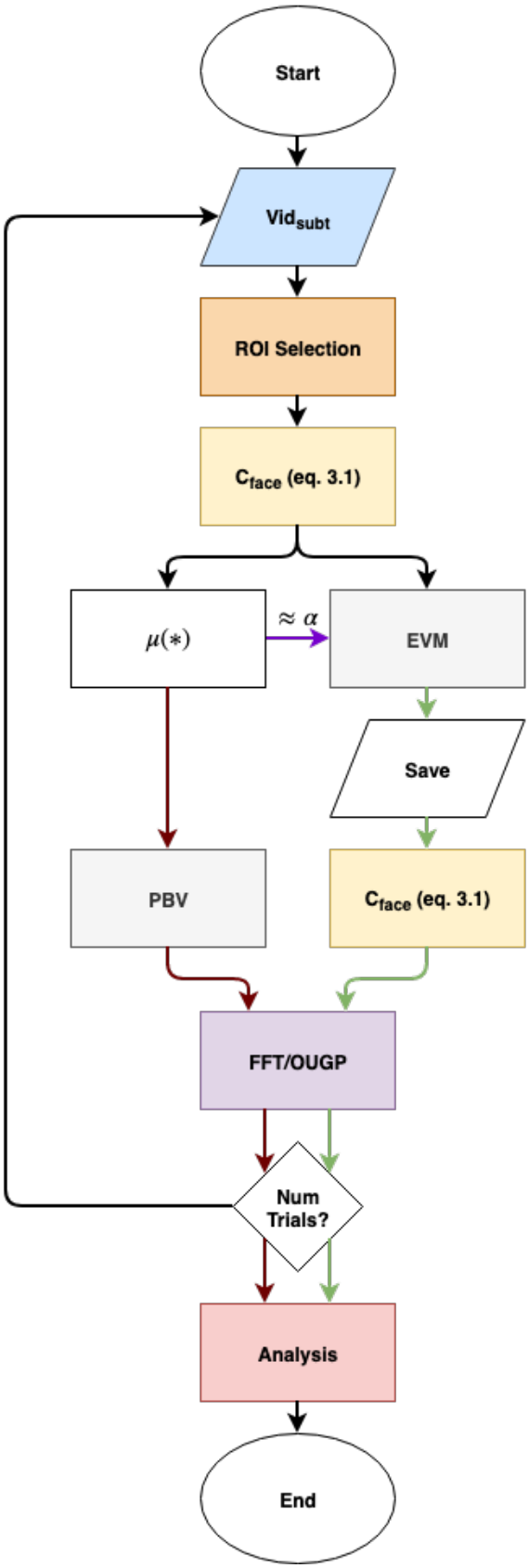

Figure 3.1: The iPPG-HRV algorithm flow diagram of this thesis.

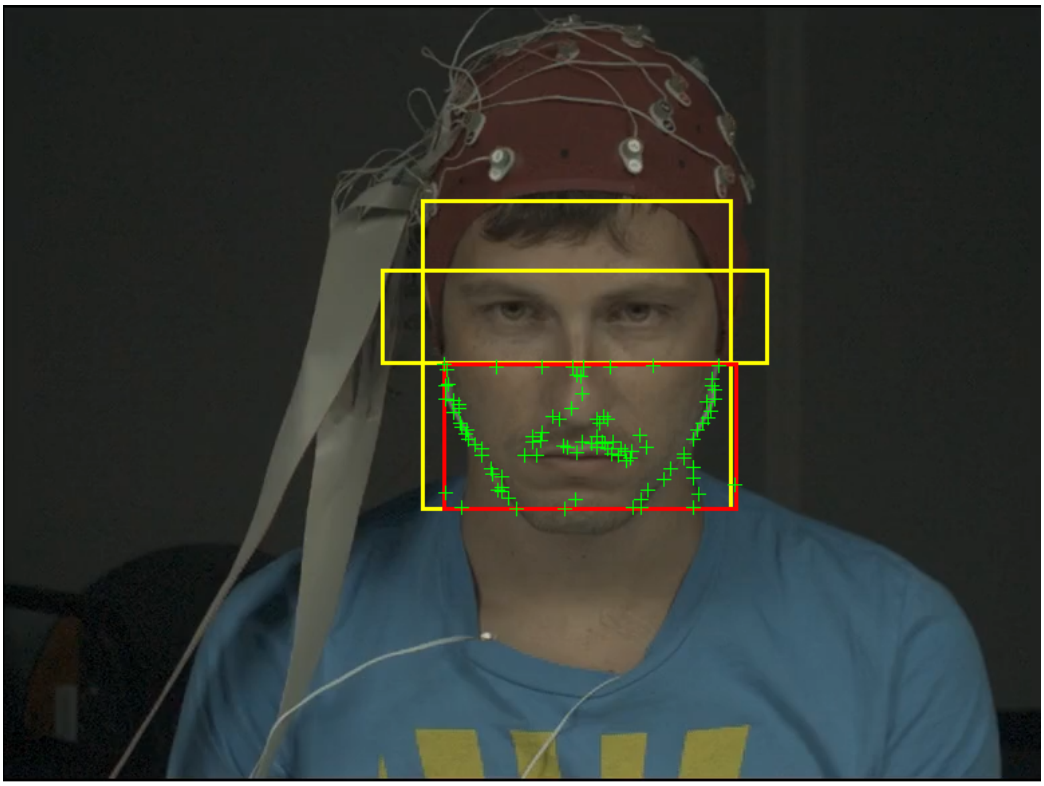

Figure 3.2: ROI selection process, where the yellow regions indicate the ROIs for the face and eyes detected with the Voila-Jones algorithm. Green markers indicate the points of interest found with the Voila-Jones to KLT methodology after minor editing of the selected ROI (red box).

ROI information. The background signal of the video is also calculated during this step for each trail's video, which is used for later calculations as this signal shows the illumination variations within the video.

The next step will load the saved ROI edited videos and the following processes are applied to each video sequence individually. From these videos, all pixels of non-zero values are then averaged for each frame of the video to obtain the raw iPPG signal resulting in $C_i(j)$. Following this, two paths emerge for the remaining steps: EVM and non-EVM methods. First, each trial's background illumination signal is now subtracted from the raw iPPG signal to reduce noise from the illumination changes in the environment just as from equations 3.2 and 3.3 giving the variable $C_I R_i$.

Second, using the moving average (denominator of equation 3.4: $\mu(C_i)$) which can

then be utilized to calculate the α or skin-tone factor for the EVM method. This signal is henceforth referred to as C_{EVM_i} to represent the usage of the EVM method on the video.

The algorithm next synchronizes the iPPG and ECG signals with each other before any other processing.

3.2.2 Synchronization and Filtering

Initial confusion with the variable provided in the metadata for each of the MAHNOB-HCI trials created complications for the ground truth ECG signal's comparison to the iPPG signal. The values of such variables were also not matching the true length of the signal, such as a 10 second video clip with metadata variables $vidEndSmp - vidBeginSmp$ giving 21 seconds for the video clip length. Yet, detailed investigation resulted in discovering a trigger signal before the physiological recordings began was time-stamp for when the recording started between each signal; this serves as indicator for the beginning of each ECG recording with the beginning of the video recording. Upon removing this section of the trigger signal, then the ECG signal is trimmed to match the same length as the iPPG signal length. To achieve this, a time vector was calculated for each signal as being the number of samples divided by the sampling rate of each respective signal. Following this the ECG data was linearly interpolated as it was sampled at the higher rate; the interpolated signal then would have values which matched those of the iPPG signal's time vector.

After determining the synchronized positions for the ECG and iPPG signals, the iPPG signal was filtered in various ways before the final HRV analysis could be preformed. As previously mentioned, butterworth filters are useful for real-time applications; yet, for simplicity since this application is a post-processing performance of this rPPG implementation, an ideal band-pass filter was applied for the typical heart-rate frequencies of [0.7 4] Hz (ie, 42 to 240 bpm).

Most studies now either interpolate the iPPG signal or use a moving average filter,

followed by an HRV metric analysis. The OUGP filter [59] is used for the smoothing of the iPPG signal, along with filtering for the various frequency domain metrics by their respective frequency limits.

3.3 HRV Metrics and Statistical Analyses

For full comprehension in the performance of the method outlined in Figure 3.1, certain metrics must be calculated and reviewed. The HRV metrics mentioned in Table 2.2 are calculated according to their respective definitions. All measures are calculated in MATLAB, where only the four key time domain metrics and short-term designations of the frequency domain are calculated.

Additionally, several statistical measures are analyzed for better comparison to previous works' results. These methods are as follows:

- *Pearson's correlation coefficient (r)* is a statistical parameter used to measure the linear association between two continuous variables C_{rPPG_j} and C_{ref_j}, which is defined by:

$$r = \frac{\sum_{j=1}^{n}(C_{rPPG_j} - \overline{C_{rPPG_j}})(C_{ref_j} - \overline{C_{ref_j}})}{\sqrt{\sum_{j=1}^{n}(C_{rPPG_j} - \overline{C_{rPPG_j}})^2 \sum_{j=1}^{n}(C_{ref_j} - \overline{C_{ref_j}})^2}} \tag{3.5}$$

  where r is an interval that ranges from -1 to $+1$; a value of $+1$ indicates a perfect positive association, 0 means no association and -1 indicates a perfect negative association.

- *Root Mean Squared Error (RMSE)* is a common measure of the differences between two continuous variables, but in comparison with MAE, RMSE it punishes large errors, defined by:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(C_{rPPG_j} - C_{ref_j})^2} \tag{3.6}$$

where C_{rPPG_j} is the predicted value and C_{ref_j} is the observed value.

- *Mean absolute error (MAE)* calculates the absolute average errors' magnitude via the formula of:

$$MAE = \frac{1}{n}\sum_{j=1}^{n}|C_{rPPG_j} - C_{ref_j}| \tag{3.7}$$

where C_{rPPG_j} denotes the j-th average value, C_{ref_j} for the reference average and n represents the number of facial recordings.

- *Mean percentage error (MPE)* measures the average percentage error where the estimated values (C_{rPPG_j}) differ from the reference values (C_{ref_j}):

$$MPE = \frac{100\%}{n}\sum_{j=1}^{n}\frac{C_{rPPG_j} - C_{ref_j}}{C_{ref_j}} \tag{3.8}$$

Each of these methods will also be calculated for the individual rPPG method (ie, EVM, etcetera), along with each i-th color channel and the sum of the color channels.

CHAPTER 4: RESULTS

Application of the aforementioned knowledge and rPPG-HRV method lead to the following results presented in this chapter. Low to abstract explanations are also given throughout for enhanced understanding of the implications for the current rendition of this method.

4.1 Overview

Initial simulations of this implementation began with running a handful of trials for the primary MAHNOB-HCI subject, who typically represents the general method of interest across most sources that utilize this dataset. In honor of this tradition, most of the steps for the method were tested on this subject, such as the ROI selection process of Figure 3.2. From the ROI selection and iPPG extraction stages, indicated in the flow diagram of Figure 3.1, to the analysis processes followed.

To indicate the different effects the various HRV frequency domain metrics have on the time domain signals, Figure 4.1 depicts each of the frequency ranges had on the overall signal. Each subplot in the depiction was calculated on neutral trials of the ECG signal as these HRV methods are primarily applied to such medical signals in current practice. From these graphs time domain representations of each range's effect on the overall HR signal. This figure shows how the ULF$*$ method really lacks informative representation of RR intervals and in general explains why this metric is rarely calculated for short-term recordings such as these from the MAHNOB-HCI dataset. However, additional metrics should be analyzed before drawing any conclusions. The figures and tables that follow will aid in showing other HRV metrics of interest.

Measurements for the frequency domain PSD responses of each frequency range, LF/HF ratio plots, and time domain plots are provided for the comparison between ground truth ECG recordings and the EVM and PBV methods. These can all be

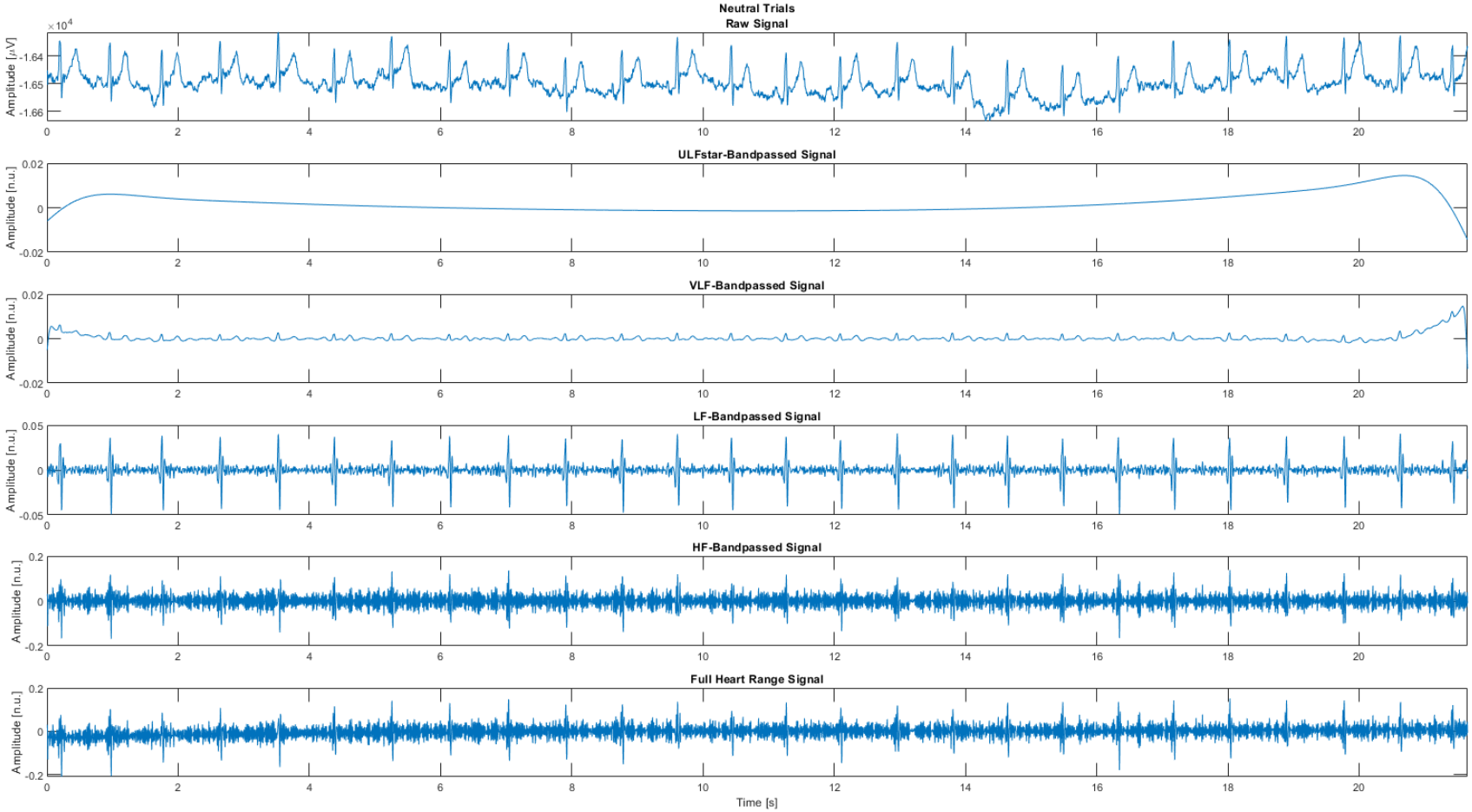

Figure 4.1: These plots each depict the ECG signal for a neutral trial. The top signal shows the raw signal in its original form in μV of a subject, where each subsequent graph depicts the normalized HR signal for all frequency domain filters in order from least to greatest (ULF$*$, VLF, LF, HF, Total band of all possible HR frequencies).

found in Figures 4.1 to 4.4. Further statistical measurements of the neutral (ie, trials in between emotion elicitation trials with non-applicable emotion elicitation) and fear elicitation trials for the primary subject are given in Table 4.2. Yet before discussing these figures, some general data needs to be reviewed.

A synopsis of the HR statistical calculations across the two types of trials of interest and two rPPG methods can be found in Table 4.1. First looking at the neutral trials, the PBV method outperforms the EVM implementation in error calculations for all channels. In the fear trials, the EVM method did show one decent calculation, though its blue channel continued to show its poor performance for each trial type. On the other hand, the PBV method did fairly well again for the neutral trials. The best performing color channels for the EVM and PBV methods across each trail type was the total rgb and red (R), respectively; therefore, these color channels were analyzed in the graphical methods for further investigation into their differences. Additionally, these results indicate that looking only at the green channel of an image, as many methods have done before, may not have been as informative as previously thought.

Continuing to the values depicted for the time domain analyses in Table 4.2, the general understanding of these values can be confusing though generally, the better SDNN and RMSSD values are smaller. This therefore further vindicates the better performance of the rgb channel for the EVM method and red channel of the PBV method. The average rr values are more confusing from just the table; however, in Figure 4.2 the first ninty beats of Neutral and Fear trials gives a better indication for the greater swing, or higher difference detection between consecutive RR intervals. Referring back to Table 4.1, the PBV method has already been shown to indicate less caluclation errors for the HR measurements. Combining this knowledge with the plots of Figure 4.2 and rr values of Table 4.2 the PBV method continues to outperform EVM.

From the tachogram plots of Figure 4.2, a better sense of the two rPPG meth-

Table 4.1: HR statistical analyses for the neutral and fear conditions, where the ground truth, mean HR for the neutral and fear trials was determined as 110.03 and 119.99 BPM, respectively.

Emotion	rPPG	channel	Mean (BPM)	MAE (BPM)	MPE (%)	RMSE (BPM)
Neutral, ecg=110.03 (BPM)	**EVM**	R	85.13	25.60	22.97	25.60
		G	88.01	23.09	20.75	23.09
		B	114.51	36.87	-34.57	36.87
		rgb	**85.93**	**24.74**	**22.21**	**24.74**
	PBV	**R**	**117.02**	**9.47**	**-7.51**	**9.47**
		G	119.42	11.44	-9.76	11.44
		B	117.90	19.42	-16.47	19.42
		rgb	119.32	17.94	-16.88	17.94
Fear, ecg=119.99 (BPM)	**EVM**	R	111.77	8.21	6.84	8.21
		G	113.62	6.37	5.31	6.37
		B	215.35	95.36	-79.48	95.36
		rgb	**117.24**	**2.75**	**2.29**	**2.75**
	PBV	**R**	**119.27**	**0.72**	**0.60**	**0.72**
		G	117.90	2.09	1.74	2.09
		B	116.70	3.28	2.74	3.28
		rgb	134.69	14.70	-12.25	14.70

Table 4.2: Time domain analyses for the neutral and fear conditions.

Emotion	rPPG	channel	rr (avg)	SDNN	RMSSD
Neutral	**EVM**	R	0.0085	0.0072	0.0060
		G	0.0092	0.0079	0.0057
		B	0.0699	0.0592	0.0362
		rgb	**0.0090**	**0.0077**	**0.0059**
	PBV	**R**	**0.0920**	**0.0807**	**0.0488**
		G	0.1438	0.1214	0.0731
		B	0.1155	0.1003	0.0599
		rgb	0.0964	0.0825	0.0503
Fear	**EVM**	R	0.0111	0.0056	0.0065
		G	0.0178	0.0089	0.0079
		B	0.0634	0.0317	0.0199
		rgb	**0.0139**	**0.0069**	**0.0072**
	PBV	**R**	**0.1983**	**0.0992**	**0.0597**
		G	0.2193	0.1096	0.0675
		B	0.1771	0.0886	0.0544
		rgb	0.0183	0.0092	0.0054

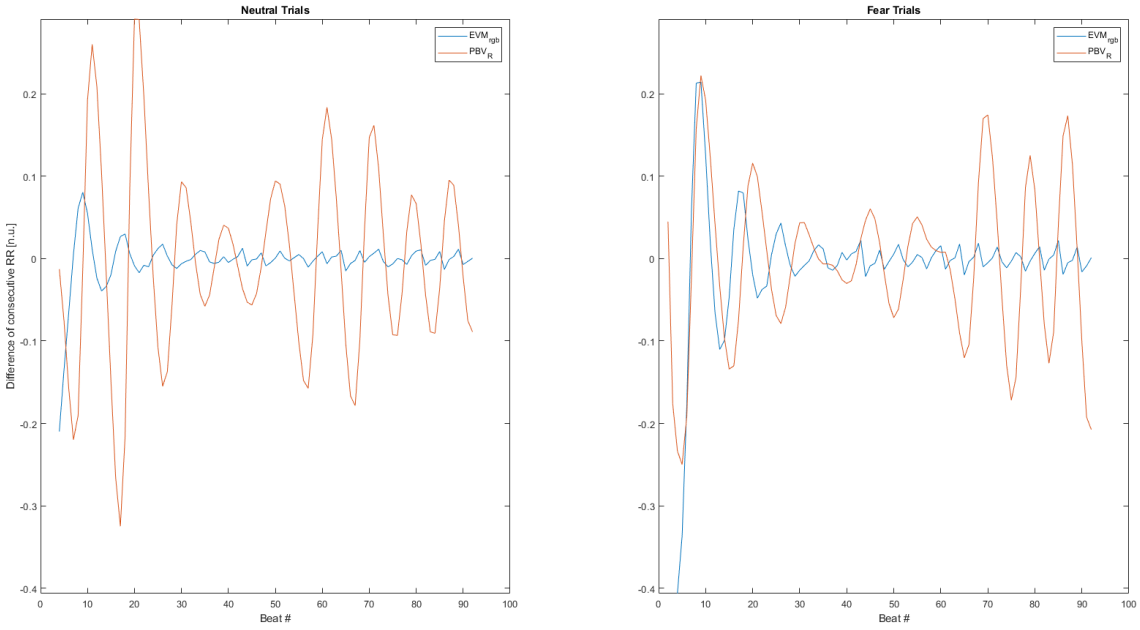

Figure 4.2: Time domain analysis signal plots for the rPPG methods in both emotion trial types. The left graph is for neutral trials for both EVM and PBV methods, while the right depicts the same RR intervals for the first 90 beats of the fear trials.

ods in both analyzed emotional elicitation trials began forming. Moving toward the frequency domain metrics of Table 4.3, the values are uninformative from just the summary this table depicts. Looking at the power spectral density (PSD) plots of Figure 4.3 provides a clearer picture of the table's values. For each of the emotion trials the PBV method provides higher values than the ground truth method.

A better understanding of these values comes from the [58] paper, in which approximate correspondence between the time and frequency domain metrics. The SDNN and RMSSD of the time domain values are correlated with the total power (TP) and HF power values, respectively. Therefore, in looking at the TP value and the HF value and response, we can continue to understand that the PBV method is performing better than the EVM in the emotion trials.

An additional metric to consider heavily is the PSD plot of the LF/HF responses, which can be found in Figure 4.4. The graphs were created in the FFT spectrum and better illustrate the difference between the neutral and fear trails. Just as with the

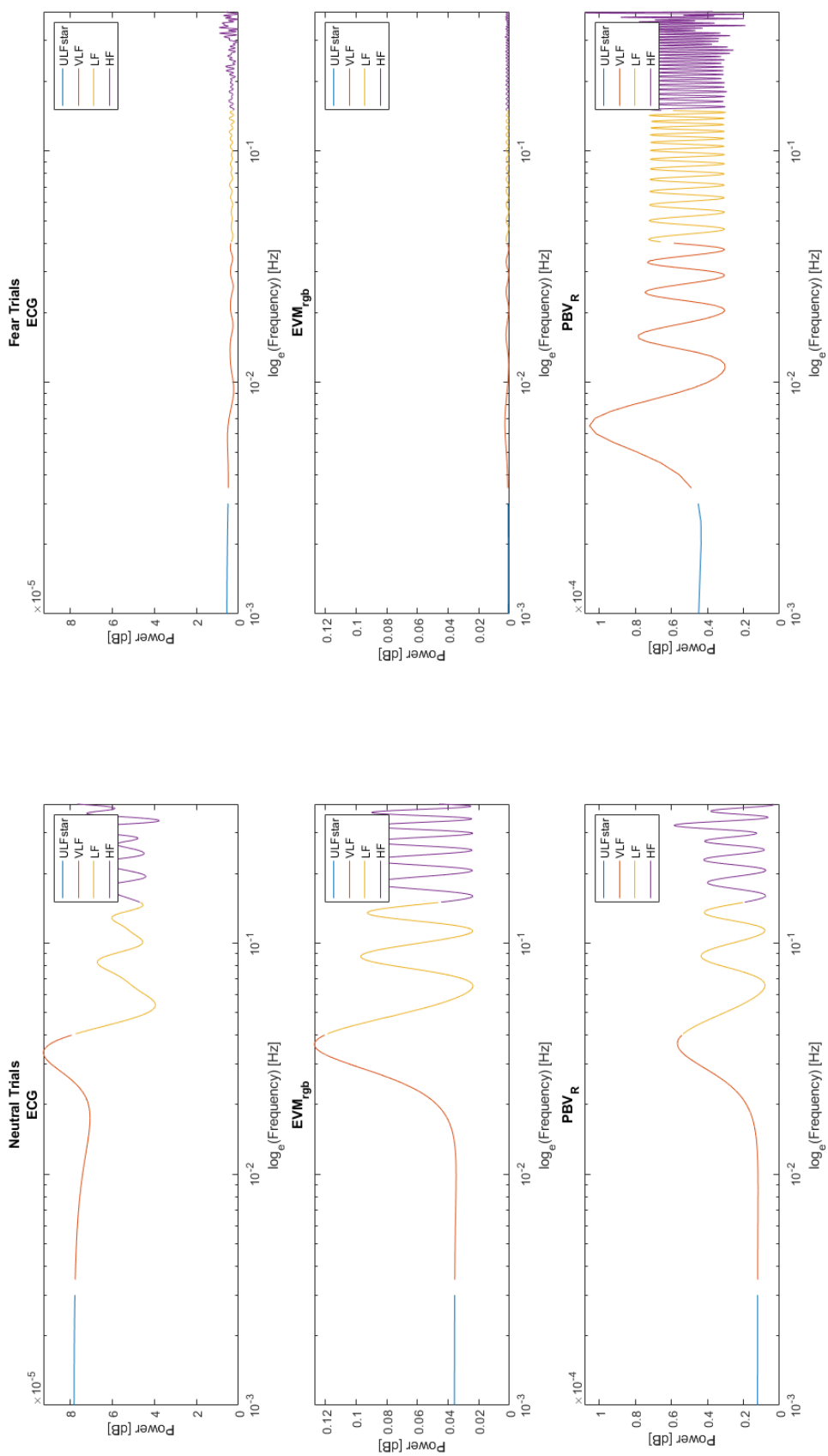

Figure 4.3: Frequency domain analysis signal plots for all methods across the neutral and fear trials.

Table 4.3: HR statistical analyses for the neutral and fear conditions, where the ground truth, mean HR for the neutral and fear trials was determined as 110.03 and 119.99 BPM, respectively.

Emotion	rPPG	channel	VLF	LF	HF	LF/HF	TP
Neutral, ecg=110.03 (BPM)	**EVM**	R	0.002643	-376.15	-812.42	0.463	0.00064
		G	0.002203	-492.22	-1070.57	0.460	0.00080
		B	0.000005	-241.14	-683.70	0.353	0.00000
		rgb	**0.002501**	**-393.34**	**-853.35**	**0.461**	**0.00066**
	PBV	**R**	**0.000017**	**-166.40**	**-366.02**	**0.455**	**0.00000**
		G	0.000001	-136.48	-317.87	0.429	0.00000
		B	0.000005	-136.36	-308.39	0.442	0.00000
		rgb	0.000003	-337.71	-727.30	0.464	0.00000
	ecgN		**0.000001**	**-190.07**	**-470.28**	**0.4157**	**0.0000**
Fear, ecg=119.99 (BPM)	**EVM**	R	0.000110	-280.54	-673.40	0.417	0.00000
		G	0.000060	-281.01	-685.13	0.410	0.00000
		B	0.000032	-280.29	-696.66	0.402	0.00000
		rgb	**0.000053**	**-279.56**	**-669.18**	**0.418**	**0.00000**
	PBV	**R**	**0.000002**	**-276.28**	**-656.16**	**0.421**	**0.00000**
		G	0.000003	-279.81	-654.96	0.427	0.00000
		B	0.000000	-279.33	-681.30	0.410	0.00000
		rgb	0.000000	-290.89	-939.79	0.310	0.00000
	ecgF		**0.000000**	**-279.78**	**-716.28**	**0.391**	**0.00000**

graphs in Figure 4.3, the ECG and EVM trials highlight the difference between the two emotional trials where the fear signals have a much lower amplitude. Though the PBV graphs may seem to be varying from the range of the true values, they still are closer to the values of the ECG data, whose LF and HF values are much closer to those of the PBV compared to the EVM. The current data though does not seem as varying from each other for the frequency domain metrics as would be desired, which leads to a discussion of these findings and the viability of these rPPG methods for determining differing cognitive states.

4.2 Discussion

Though the above results provide insight into the general outcomes of the proposed rPPG-HRV methods, they do not give a broader understanding as to why the results are as they are. These remaining sections for this chapter will provide substantial insight into the implications of this method and any of the trade-offs for the current

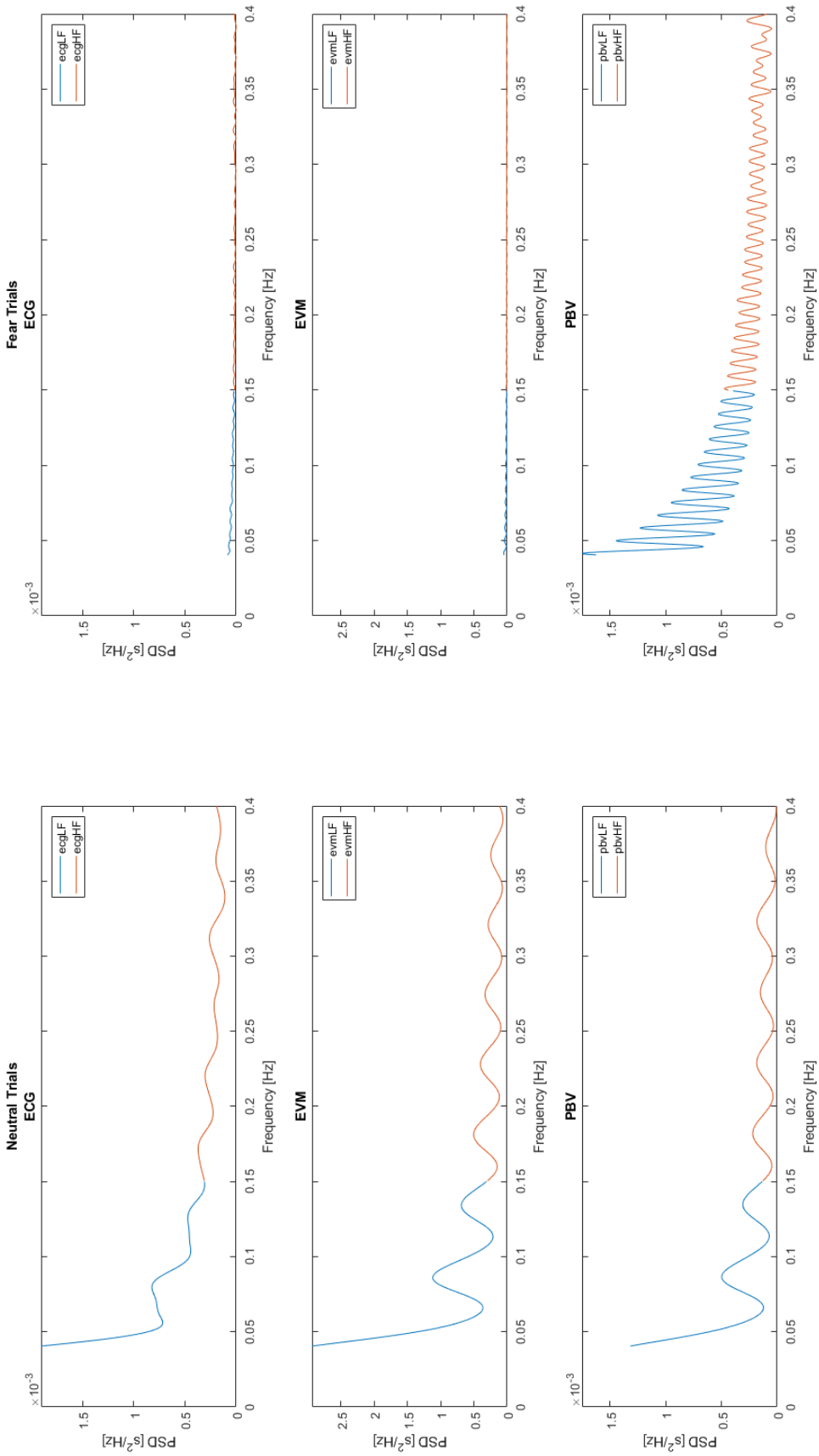

Figure 4.4: LF/HF ratio comparison plots in the FFT spectrum for all methods across the neutral and fear trials.

implementation.

4.2.1 Complications and Drawbacks of Current Setup

The Voila-Jones Algorithm struggles with data that has a subject's bare chest or neck region intersecting with the edges of the video frame. During such a scenario, the algorithm seems to crash and consider the ROI of the full face to go past the actual size of the video frame. This scenario immediately shows how quick the entire method could crumble and give terrible data just based on this odd limitation of the Voila-Jones face detection algorithm. Not only does the ROI selection in the current method have this bounding box exceeding the true frame size potential for complication, it also is simplistic at the moment in detecting the difference between skin and non-skin. While the post-processing filters help remove this issue by removing luminescence and other DC component variances from the raw iPPG signal, a more robust alternative for skin-patch selection and tracking over time would be more beneficial to eliminating errors caused at the beginning of a method's rPPG extraction.

While a general viability has been established for the PBV implementation of the tested methods, the lack of difference between the two emotional trials is worth discussing. One potential issue would be the current method analyzes emotion trials that are much longer than neutral trials. Additionally, the dataset currently implemented did not include individuals with ADHD, nor did it test said individuals pre- and post-treatment processes. This means that the variation difference between trials showing greater distinctions should not be inferred as quickly. This further indicates that the anticipated larger distinctions of the LF/HF ratio do not necessarily correlate to the higher ratio distinctions of those with ADHD pre- and post-treatment processes, which leads to other implications of this thesis.

4.2.2 Implications of Emotion Elicitation Tasks and rPPG-HRV

Multiple types of emotionally charged trials were available in the MAHNOB-HCI dataset. However, as this work aims to understand how rPPG-HRV methods might perform for treatment efficacy of ADHD patients, the neutral and fear trials were of greater value for such an investigation. Neutral trials were considered as a form of baseline signal, while the fear trials are heavily linked to the fear-neurocircuitry system of which the ANS is tightly entangled.

Currently, a key value for distinguishing a difference between the two emotion subsets, namely the LF/HF ratio, seems minuscule. A possible explanation may be from the fewer fear based trials compared to the neutral ones, which come from two classic horror film moments which may have been known by the subjects whether they had seen the film or merely heard of them. Additionally, as the subjects have more neutral trials than fear based ones and the length of each also varies extremely, results may be skewed and potentially less indicative of true understanding. Yet, the PBV method performed extremely well across these trials, especially compared to the EVM and a difference between the LF and HF values separately is very distinct. These distinctions between higher or lower LF and HF magnitudes between the two emotional trial types further proved the relevance of the PBV method whose results matched the ground truth far closer than the EVM. Therefore, based upon the current results, this indicates a general validity for using the PBV method in continued research.

CHAPTER 5: CONCLUSIONS

Multiple contributions were achieved in this thesis. Firstly, the state of current research toward developing objective measurements for the understanding and treatment efficacy process of neurotypical individuals and those with ADHD. In reviewing this background, visible spectrum image and signal processing methods were determined as the most viable options for the goals of this thesis based on their safety, modularity, and in-expense.

Additionally, the correlation between neurocognitive disorders like ADHD, physiological measurements and ANS - a part of the fear-neurocircuitry system - was surveyed. Subsequently, a review of the current efforts in rPPG methodologies led to the combination of various rPPG algorithms, noise removal and filtering, to enhance the iPPG signal extraction for improved HRV detection and analysis These survey culminations grant substantial information which not only aided in the choices of method design, from a RGB camera to the filters utilized, but also provided the basis of knowledge for the design and analysis of this rPPG-HRV method.

Further contributions include the analysis and comparison of different video based trials for different types of emotional states, which are highly correlated to affecting the cognitive state of the mind. The results of this thesis also showed that channels other than solely the green channel of a camera provide better results in some of the rPPG methods and should therefore not be overlooked so easily. With these additions to the current rPPG repertoire, a new system for collecting and understanding physiological attributes is attainable. Not only will these aid in the ability to analyse performance of individuals in the physical realm as past rPPG methods have focused on, but also mental understanding; this was achieved by investigating an emotion elicitation dataset for comparison of neutral and emotional trials. Thus, future improvements to the current method are now discussed for furthering the potential gains and achieving the long-term aims of this thesis.

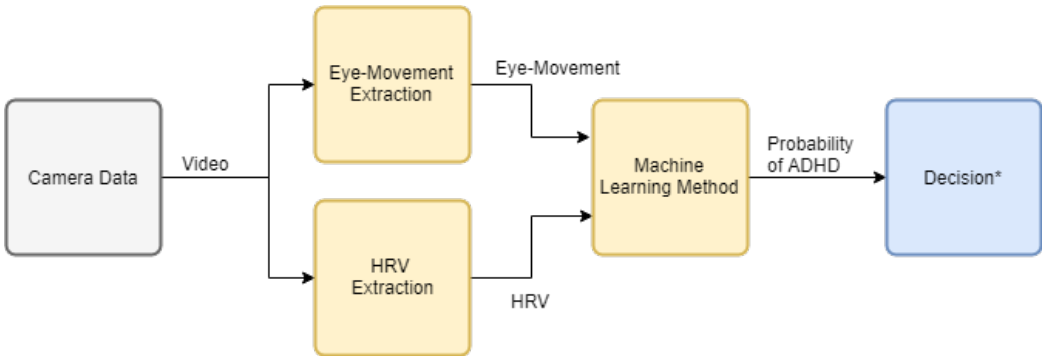

Figure 5.1: Block diagram of system design for objective measurements to be utilized in diagnosis and treatment efficacy of ADHD [5]

.

5.1 Future Work

In regards to the HRV method, future work could test various combinations of ROI selection, rPPG methods, illumination and motion artifact correction, as well as filtering techniques. One recent study working with the MAHNOB-HCI dataset utilized adaptive patch selection for improved ROI detection and estimation [67]. Since different patch regions (eg, left and right cheek, chin/mouth, forehead) are tracked throughout this method, further investigation for the usefullness of varying regions for individuals could be analysed across multiple trials to see how even each patch is affected during different cognitive tasks and under the influence of pharmaceutical treatments.

Additional investments toward adding multiple parameters for informative decisions would include eye-gaze tracking, as depicted in Figure 5.1. The MAHNOB-HCI dataset already includes ground-truth information on gaze data, yet testing on other datasets, particularly ones which indicate any neurocognitive or neurodevelopmental disorder, would enhance the overall method's robustness. These are but a few steps on the path of utilizing image segmentation for the extraction of HRV for updating mental health, therapeutic validity processes. For example, other neurocognitive/developmental disorders affect the ANS, like ASD and PTSD. Even though this

thesis focused primarily on ADHD for inspiration, this is only one segment in the signal or season in life, which can continue improving; whether by success or failure, the segment goes to the next time interval and the season will change.

REFERENCES

[1] M. T. Banich and R. J. Compton, *Cognitive Neuroscience*. Cambridge: Cambridge University Press, 4 ed., 2018.

[2] D. Gentner, "Psychology in cognitive science: 1978-2038," *Topics in Cognitive Science*, 2010.

[3] S. S. Dominy, C. Lynch, F. Ermini, M. Benedyk, A. Marczyk, A. Konradi, M. Nguyen, U. Haditsch, D. Raha, C. Griffin, L. J. Holsinger, S. Arastu-Kapur, S. Kaba, A. Lee, M. I. Ryder, B. Potempa, P. Mydel, A. Hellvard, K. Adamowicz, H. Hasturk, G. D. Walker, E. C. Reynolds, R. L. M. Faull, M. A. Curtis, M. Dragunow, and J. Potempa, "Porphyromonas gingivalis in Alzheimer's disease brains: Evidence for disease causation and treatment with small-molecule inhibitors," *Science Advances*, vol. 5, no. 1, 2019.

[4] S. B. Appakaya and R. Sankar, "Effectiveness of Speech Analysis in Classification of Neurodegenerative Diseases: A study on Parkinson's Disease," in *SoutheastCon 2018*, (St. Petersburg), IEEE, 2018.

[5] L. E. Johnson and J. M. Conrad, "A Survey of Technologies Utilized in the Treatment and Diagnosis of Attention Deficit Hyperactivity Disorder," in *IEEE UEMCON 2018*, 2018.

[6] G. C. Burgess, B. E. Depue, L. Ruzic, E. G. Willcutt, Y. P. Du, and M. T. Banich, "Attentional Control Activation Relates to Working Memory in Attention-Deficit/Hyperactivity Disorder," *Biological Psychiatry*, vol. 67, pp. 632–640, 2010.

[7] M. J. Kofler, M. D. Rapport, J. Bolden, D. E. Sarver, and J. S. Raiker, "ADHD and Working Memory: The Impact of Central Executive Deficits and Exceeding Storage/Rehearsal Capacity on Observed Inattentive Behavior," *Journal of Abnormal Child Psychology*, vol. 38, pp. 149–161, 2010.

[8] M. D. Rapport, J. Bolden, M. J. Kofler, D. E. Sarver, J. S. Raiker, and R. M. Alderson, "Hyperactivity in Boys with Attention-Deficit/Hyperactivity Disorder (ADHD): A Ubiquitous Core Symptom or Manifestation of Working Memory Deficits?," *Journal of Abnormal Child Psychology*, vol. 37, no. 4, pp. 521–34, 2009.

[9] J. S. Raiker, M. D. Rapport, M. J. Kofler, and D. E. Sarver, "Objectively-Measured Impulsivity and Attention-Deficit/Hyperactivity Disorder (ADHD): Testing Competing Predictions from the Working Memory and Behavioral Inhibition Models of ADHD," *Journal of Abnormal Child Psychology*, vol. 40, pp. 699–713, 2012.

[10] M. D. Rapport, R. M. Alderson, M. J. Kofler, D. E. Sarver, J. Bolden, and V. Sims, "Working Memory Deficits in Boys with Attention-deficit/Hyperactivity Disorder (ADHD): The Contribution of Central Executive and Subsystem Processes," *Journal of Abnormal Child Psychology*, vol. 36, no. 6, pp. 825–37, 2008.

[11] G. J. Dupaul, L. L. Weyandt, S. M. O 'dell, and M. Varejao, "College Students With ADHD Current Status and Future Directions," *Journal of Attention Disorders*, vol. 13, no. 3, pp. 234–250, 2009.

[12] M. R. Dvorsky and J. M. Langberg, "Predicting Impairment in College Students With ADHD: The Role of Executive Functions," *Journal of Attention Disorders*, vol. 1, 2014.

[13] S. Woltering, Z. Liu, A. Rokeach, and R. Tannock, "Neurophysiological differences in inhibitory control between adults with ADHD and their peers," *Neuropsychologia*, vol. 51, pp. 1888–1895, 2013.

[14] American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition.* American Psychiatric Association, 5 2013.

[15] American Academy of Pediatrics, "ADHD: Clinical Practice Guideline for the Diagnosis, Evaluation, and Treatment of Attention-Deficit/ Hyperactivity Disorder in Children and Adolescents," *Pediatrics*, vol. 128, no. 5, 2011.

[16] Neuro Assessment and Development Center, "Neuropsychological Assessment."

[17] Children and Adults with Attention-Deficit Hyperactivity Disorder, "Understanding ADHD - Data and Statistics," 2018.

[18] S. N. Visser, J. R. Holbrook, M. L. Danielson, and R. H. Bitsko, "Diagnostic experiences of children with attention-deficit/hyperactivity disorder," *National Health Statistics Reports Number*, vol. 81, 2015.

[19] A. M. Smith, *Determining the Efficacy of the BASC-2 PRS in the Differential Diagnosis of Autism and Attention-Deficit/Hyperactivity Disorder.* PhD thesis, Texas Woman's University, 2015.

[20] S. J. Loew, K. Watson, and Watson Kenneth, "The Prevalence of Symptoms of Scotopic Sensitivity/Meares-Irlen Syndrome in Subjects Diagnosed with ADHD: Does Misdiagnosis Play a Significant Role?," *Hrvatska Revija za Rehabilitacijska Istrazivanja*, vol. 49, pp. 64–72, 2013.

[21] G. L. Robinson, "Coloured lenses and reading: A review of research into reading achievement reading strategies and causal mechanisms," *Australian Journal of Special Education*, vol. 18, pp. 3–14, 1994.

[22] P. C. Ford-Jones, "Misdiagnosis of attention deficit hyperactivity disorder: 'Normal behaviour' and relative maturity," *Paediatrics & Child Health*, vol. 20, no. 4, pp. 200–202, 2015.

[23] N. B. Harris, "How childhood trauma affects health across a lifetime," 2014.

[24] E. J. Sonuga-Barke, M. Kennedy, R. Kumsta, N. Knights, D. Golm, M. Rutter, B. Maughan, W. Schlotz, and J. Kreppner, "Child-to-adult neurodevelopmental and mental health trajectories after early life deprivation: the young adult follow-up of the longitudinal English and Romanian Adoptees study," *The Lancet*, vol. 389, no. 10078, pp. 1539–1548, 2017.

[25] M. L. Danielson, R. H. Bitsko, R. M. Ghandour, J. R. Holbrook, M. D. Kogan, and S. J. Blumberg, "Prevalence of Parent-Reported ADHD Diagnosis and Associated Treatment Among U.S. Children and Adolescents, 2016," *Journal of Clinical Child & Adolescent Psychology*, vol. 47, no. 2, pp. 199–212, 2018.

[26] J. Martinez-Raga, A. Ferreros, C. Knecht, R. de Alvaro, and E. Carabal, "Attention-deficit hyperactivity disorder medication use: factors involved in prescribing, safety aspects and outcomes.," *Therapeutic advances in drug safety*, vol. 8, pp. 87–99, 3 2017.

[27] A. Melchor Rodriguez and J. Ramos Castro, "Pulse rate variability analysis by video using face detection and tracking algorithms," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 5696–5699, IEEE, 8 2015.

[28] A. M. Rodríguez, J. Ramos-Castro, M. Rodríguez, R.-C. Biomed, and E. Online, "Video pulse rate variability analysis in stationary and motion conditions," *BioMedical Engineering Online*, 2018.

[29] W. Wang, A. C. den Brinker, S. Stuijk, and G. de Haan, "Robust heart rate from fitness videos," *Physiological Measurement*, vol. 38, pp. 1023–1044, 2017.

[30] M. Soleymani, J. Lichtenauer, T. Pun, and M. Pantic, "A Multimodal Database for Affect Recognition and Implicit Tagging," *IEEE Transactions on Affective Computing*, vol. 3, no. 1, 2012.

[31] S. DiMaio, N. Grizenko, and R. Joober, "Dopamine genes and attention-deficit hyperactivity disorder: a review.," *J Psychiatry Neurosci*, vol. 28, no. 1, pp. 27–38, 2003.

[32] E. J. S. Sonuga-Barke, E. Taylor, S. Sembi, and J. Smith, "Hyperactivity and Delay Aversion-I. The Effect of Delay on Choice," *Journal of Child Psychology and Psychiatry*, vol. 33, pp. 387–398, 2 1992.

[33] L. F. Koziol, D. E. Budding, and D. Chidekel, *ADHD as a Model of Brain-Behavior Relationships*. SpringerBriefs in Neuroscience, New York, NY: Springer, 2013.

[34] J. A. Sergeant, "Modeling Attention-Deficit/Hyperactivity Disorder: A Critical Appraisal of the Cognitive-Energetic Model," *Biological Psychiatry*, vol. 57, pp. 1248–1255, 6 2005.

[35] E. Sonuga-Barke, P. Bitsakou, and M. Thompson, "Beyond the Dual Pathway Model: Evidence for the Dissociation of Timing, Inhibitory, and Delay-Related Impairments in Attention-Deficit/Hyperactivity Disorder," *Journal of the American Academy of Child and Adolescent Psychiatry*, vol. 49, no. 4, pp. 345–355, 2010.

[36] P. de Zeeuw, J. Weusten, S. van Dijk, J. van Belle, and S. Durston, "Deficits in Cognitive Control, Timing and Reward Sensitivity Appear to be Dissociable in ADHD," *PLoS ONE*, vol. 7, no. 12, 2012.

[37] C. Huang-Pollock, R. Ratcliff, G. Mckoon, Z. Shapiro, A. Weigard, and H. Galloway-Long, "Using the Diffusion Model to Explain Cognitive Deficits in Attention Deficit Hyperactivity Disorder," *Journal of Abnormal Child Psychology*, vol. 45, no. 1, pp. 57–68, 2017.

[38] S. Ziegler, M. L. Pedersen, A. M. Mowinckel, and G. Biele, "Modelling ADHD: A review of ADHD theories through their predictions for computational models of decision-making and reinforcement learning," *Neuroscience & Biobehavioral Reviews*, vol. 71, pp. 633–656, 12 2016.

[39] C. Sridhar, S. Bhat, U. Rajendra Acharya, H. Adeli, and G. M. Bairy, "Diagnosis of attention deficit hyperactivity disorder using imaging and signal processing techniques," *Computers in Biology and Medicine*, vol. 88, pp. 93–99, 2017.

[40] S. Ghiassian, R. Greiner, P. Jin, and M. R. G. Brown, "Using Functional or Structural Magnetic Resonance Images and Personal Characteristic Data to Identify ADHD and Autism formerly the Alberta Innovates Centre for Machine Learning (AICML), Project Title: Patient Specific fMRI-based Psychiatric Diagnosis and," *PLoS ONE*, vol. 11, no. 12, 2016.

[41] L. Zou, J. Zheng, C. Miao, M. J. Mckeown, and Z. J. Wang, "3D CNN Based Automatic Diagnosis of Attention Deficit Hyperactivity Disorder Using Functional and Structural MRI," *IEEE Access*, vol. 5, pp. 23626–23636, 2017.

[42] B. Sen, N. C. Borle, R. Greiner, and M. R. G. Brown, "A general prediction model for the detection of ADHD and Autism using structural and functional MRI," *PLoS ONE*, vol. 13, no. 4, 2018.

[43] M. Hoogman, J. Bralten, D. P. Hibar, M. Mennes, M. P. Zwiers, and et al, "Subcortical brain volume differences in participants with attention deficit hyperactivity disorder in children and adults: a cross-sectional mega-analysis," *The Lancet Psychiatry*, vol. 4, no. 4, pp. 310–319, 2017.

[44] M. D. Rosenberg, S. Zhang, W.-T. Hsu, D. Scheinost, E. S. Finn, X. Shen, R. Todd Constable, X. Chiang, S. R. Li, X. Marvin, and M. Chun, "Methylphenidate Modulates Functional Network Connectivity to Enhance Attention," *Journal of Neuroscience*, vol. 63, no. 37, pp. 9547–9557, 2016.

[45] E. Betta and M. Turatto, "Are you ready? I can tell by looking at your microsaccades," *NeuroReport*, vol. 17, pp. 1001–1004, 7 2006.

[46] Z. M. Hafed, L. P. Lovejoy, and R. J. Krauzlis, "Modulation of microsaccades in monkey during a covert visual attention task.," *The Journal of Neuroscience*, vol. 31, pp. 15219–30, 10 2011.

[47] J. Gottlieb, P. Y. Oudeyer, M. Lopes, and A. Baranes, "Information-seeking, curiosity, and attention: Computational and neural mechanisms," *Trends in Cognitive Sciences*, vol. 17, no. 11, pp. 585–593, 2013.

[48] M. Fried, E. Tsitsiashvili, Y. S. Bonneh, A. Sterkin, T. Wygnanski-Jaffe, T. Epstein, and U. Polat, "ADHD subjects fail to suppress eye blinks and microsaccades while anticipating visual stimuli but recover with medication," *Vision Research*, vol. 101, pp. 62–72, 2014.

[49] Y. Danker, L. Shalev, M. Carrasco, and S. Yuval-Greenberg, "Prestimulus Inhibition of Saccades in Adults With and Without Attention-Deficit/Hyperactivity Disorder as an Index of Temporal Expectations," *Psychological Science*, vol. 28, no. 7, pp. 835–850, 2017.

[50] T. D. Gould, T. M. Bastain, M. E. Israel, D. W. Hommer, and F. X. Castellanos, "Altered performance on an ocular fixation task in attention-deficit/hyperactivity disorder," *Biological Psychiatry*, vol. 50, 2001.

[51] D. P. Munoz, I. T. Armstrong, K. A. Hampton, and K. D. Moore, "Altered control of visual fixation and saccadic eye movements in attention-deficit hyperactivity disorder," *Journal of Neurophysiology*, vol. 90, pp. 503–514, 2003.

[52] J. van der Meere, "State regulation and attention deficit hyperactivity disorder," *Attention Deficit Hyperactivity Disorder*, pp. 413–433, 2005.

[53] T. Yüksel and Ã. Özcan, "Heart rate variability as an indicator of autonomous nervous system activity in children with attention deficit hyperactivity disorder," *Anadolu Psikiyatri Dergisi*, vol. 19, no. 5, pp. 493–500, 2018.

[54] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. Freeman, "Eulerian Video Magnification for Revealing Subtle Changes in the World," *ACM Transactions on Graphics (Proceedings SIGGRAPH 2012)*, vol. 31, no. 4, 2012.

[55] P. Goolkasian, T. Carmichael, M. Croy, B. Davis, M. Faust, M. Hadzikadic, H. Lipford, A. Raja, and L. Van Wallendael, *Cognitive Science: An Interactive Approach*. National Social Science Press, 2011.

[56] J. P. Sturmberg, J. M. Bennett, M. Picard, and A. J. E. Seely, "The trajectory of life. Decreasing physiological network complexity through changing fractal patterns," *Front. Physiol*, vol. 6, p. 169, 2015.

[57] N. Singh, K. J. Moneghetti, J. W. Christle, D. Hadley, D. Plews, and V. Froelicher, "Heart Rate Variability: An Old Metric with New Meaning in the Era of using mHealth Technologies for Health and Exercise Training Guidance. Part One: Physiology and Methods," *Arrhythmia & Electrophysiology Review*, vol. 7, no. 3, p. 193, 2018.

[58] Task Force of The European Society of Cardiology and The North American Society of Pacing and Electrophysiology, "Heart rate variability: Standards of measurement, physiological interpretation, and clinical use," *European Heart Journal*, vol. 17, pp. 354–381, 1996.

[59] A. C. Fisher, â. A. Eleuteri, â. D. Groves, and â. C. J. Dewhurst, "The Ornstein-Uhlenbeck third-order Gaussian process (OUGP) applied directly to the unresampled heart rate variability (HRV) tachogram for detrending and low-pass filtering,"

[60] R. Sassi, S. Cerutti, F. Lombardi, M. Malik, H. V. Huikuri, C.-K. Peng, G. Schmidt, and Y. Yamamoto, "Advances in heart rate variability signal analysis: joint position statement by the e-Cardiology ESC Working Group and the European Heart Rhythm Association co-endorsed by the Asia Pacific Heart Rhythm Society," *Europace*, vol. 17, pp. 1341–1353, 2015.

[61] M. Altini, "Heart Rate Variability using the phone's camera," 2014.

[62] N. Singh, K. J. Moneghetti, J. W. Christle, D. Hadley, V. Froelicher, and D. Plews, "Heart Rate Variability: An Old Metric with New Meaning in the Era of Using mHealth technologies for Health and Exercise Training Guidance. Part Two: Prognosis and Training," *Arrhythmia & Electrophysiology Review*, vol. 7, no. 4, p. 1, 2018.

[63] M. Finžgar and P. Podržaj, "A wavelet-based decomposition method for a robust extraction of pulse rate from video recordings," *PeerJ*, vol. 6, no. e5859, 2018.

[64] W. Wang, A. C. Den Brinker, G. D. Haan, D. J. McDuff, J. R. Estepp, A. M. Piasecki, E. B. Blackford, S. Tulyakov, X. Alameda-Pineda, E. Ricci, L. Yin, J. F. Cohn, and N. Sebe, "Advancements in noncontact, multiparameter physiological measurements using a webcam," *Biomedical Optics Express*, vol. 63, no. 3, pp. 702–715, 2016.

[65] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," pp. I–511–I–518, Institute of Electrical and Electronics Engineers (IEEE), 8 2005.

[66] A. Asthana, S. Zafeiriou, S. Cheng, and M. Pantic, "Robust discriminative response map fitting with constrained local models," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2013.

[67] Z. Wang, X. Yang, and K.-T. Cheng, "Accurate face alignment and adaptive patch selection for heart rate estimation from videos under realistic scenarios," 2018.

[68] X. Li, J. Chen, G. Zhao, and M. Pietikäinen, "Remote heart rate measurement from face videos under realistic situations," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4264–4271, 2014.

[69] G. de Haan and A. van Leest, "Improved motion robustness of remote-PPG by using the blood volume pulse signature," *Physiological Measurement*, vol. 35, pp. 1913–1926, 2014.

[70] W. Wang, A. C. den Brinker, S. Stuijk, and G. de Haan, "Algorithmic Principles of Remote PPG," *IEEE transactions on bio-medical engineering*, vol. 64, pp. 1479–1491, 7 2017.

[71] S. Tulyakov, X. Alameda-Pineda, E. Ricci, L. Yin, J. F. Cohn, and N. Sebe, "Self-Adaptive Matrix Completion for Heart Rate Estimation from Face Videos under Realistic Conditions," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[72] W. Wang, A. C. Den Brinker, S. Stuijk, G. De Haan, . M. Z. Poh, D. J. Mcduff, and R. W. Picard, "Amplitude-selective filtering for remote-PPG," *Biomedical Optics Express*, vol. 8, no. 3, pp. 405–410, 2017.

[73] I. R. Tayibnapis, Y. M. Yang, and K. M. Lim, "Blood volume pulse extraction for non-contact heart rate measurement by digital camera using singular value decomposition and Burg algorithm," *Energies*, 2018.

[74] K. Alghoul, S. Alharthi, H. Al Osman, and A. El Saddik, "Heart Rate Variability Extraction from Videos Signals: ICA vs. EVM Comparison," *IEEE Access*, vol. 5, pp. 4711–4719, 2017.

[75] A. Eleuteri, A. C. Fisher, D. Groves, and C. J. Dewhurst, "An Efficient Time-Varying Filter for Detrending and Bandwidth Limiting the Heart Rate Variability Tachogram without Resampling: MATLAB Open-Source Code and Internet Web-Based Implementation," *Computational and Mathematical Methods in Medicine*, vol. 2012, 2012.

[76] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet," *Circulation*, vol. 101, 6 2000.

[77] D. J. Plews, B. Scott, M. Altini, M. Wood, A. E. Kilding, and P. B. Laursen, "Comparison of heart-rate-variability recording with smartphone photoplethysmography, polar H7 chest strap, and electrocardiography," *International Journal of Sports Physiology and Performance*, 2017.

APPENDIX: ACKNOWLEDGEMENT OF MAHNOB-HCI DATASET

Portions of the research in this paper uses the MAHNOB Database collected by Professor Pantic and the iBUG group at Imperial College London, and in part collected in collaboration with Prof. Pun and his team of University of Geneva, in the scope of MAHNOB project financially s upported b y t he E uropean R esearch Council under the European Communityâs 7th Framework Programme (FP7/20072013) / ERC Starting Grant agreement No. 203143.