



Article Towards Automatic Detection of Social Anxiety Disorder via Gaze Interaction

Sara Shafique ¹, Iftikhar Ahmed Khan ^{1,}*, Sajid Shah ², Waqas Jadoon ¹, Rab Nawaz Jadoon ¹ and Mohammed ElAffendi ²

- ¹ Department of Computer Sciences, COMSATS University Islamabad, Abbottabad Campus, TOBE Camp, Abbottabad 22060, Pakistan
- ² EIAS Data Science Lab, College of Computer and Information Sciences, Prince Sultan University, Riyadh 11586, Saudi Arabia
- * Correspondence: iftikharahmed@cuiatd.edu.pk; Tel.: +92-331-5720248

Featured Application: This work presented a method to detect Social Anxiety Disorder (SAD) objectively Haar Cascade Classifiers. This method, therefore, could be used by clinicians and psychologists as added aiding method to detect SAD in patients.

Abstract: Social anxiety disorder (SAD) is an extreme fear of underperformance in various social situations. It is necessary to detect people with or without SAD for counseling and treatment. A few manual techniques in the existing literature show the possibility of SAD detection from gaze interaction. However, an automated prediction of SAD is scarce. In this research, an automatic technique to predict SAD using gaze interaction/avoidance is proposed, where a custom application was developed that used the Haar Cascade classifier to predict gaze interaction/avoidance. The experiments were conducted on 50 participants in a live environment using the developed application. SAD classes were predicted by using decision tree classifiers from the created gaze dataset. The results proved that SAD could be predicated with an overall accuracy of 80%. Furthermore, four classes of SAD (Mark, Moderate, Severe, Very Severe along with 'No SAD') could be predicted with an accuracy of 80%, 70%, 90%, 80%, and 80%, respectively. The research proved the possibility to predict SAD using computer-based methods without human intervention. Furthermore, it created the possibility of aiding a subjective Liebowitz Social Anxiety Scale (LSAS) with an objective technique described in this research.

Keywords: social anxiety disorder (SAD); LSAS scale; automatic prediction; haar cascade classifier; decision tree classifier; SAD classes

1. Introduction

Humans show various social behaviors. Gaze interaction is one example of nonverbal social communication behaviors. Gaze interaction helps people judge someone to be attentive, attractive, credible, socially competent, or socially dominant, and hence is an essential factor in many social situations [1]. The opposite of gaze interaction is gaze avoidance, which is associated with several psychiatric conditions of social situations like Social Anxiety Disorder (SAD). SAD is a fear of social situations or underperformance and is the fourth most common psychological disorder [1]. Furthermore, about 13% of people suffer from SAD during their lifetimes [1].

In healthy social communication, gaze interaction is necessary. Some social situations prevalently depend on gaze interaction, like appearing in an interview for a job or presenting to an audience [2]. According to McCroskey [3], social anxiety is "An individual's level of fear or anxiety connected with either anticipated or real communication with other persons". A person having SAD may not be able to live a quality life. Here, it is essential to distinguish between shyness and SAD, where SAD could be a cause of disturbance in



Citation: Shafique, S.; Khan, I.A.; Shah, S.; Jadoon, W.; Jadoon, R.N.; ElAffendi, M. Towards Automatic Detection of Social Anxiety Disorder via Gaze Interaction. *Appl. Sci.* **2022**, *12*, 12298. https://doi.org/10.3390/ app122312298

Academic Editor: Feng Liu

Received: 25 October 2022 Accepted: 28 November 2022 Published: 1 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). normal routine. About 80% of people with SAD also experience other psychiatric conditions like specific phobia, major depressive disorder (MDD), and dependency on alcohol. The problems associated with the disorder demand proper treatment.

If left untreated, SAD could cause various social negative behaviors like isolation, cyberbullying, harmful, social academic, and psychological loneliness, problems in making friends, depression, and school refusal [4]. If left untreated, it has an increased risk of persistence into adulthood after 11 years of age. There are various significant methods to treat SAD, i.e., cognitive behavioral therapy (CBT), exercising, yoga, etc. Computers and the latest machine-learning algorithms can also aid in all the above-mentioned methods either directly or indirectly.

In recent years, computers have been effectively used to measure social and mental disorders. The reliance on information and communication technologies (ICT) is increasing because of the following two main reasons:

- 1. To avoid the patients' fear: Despite SAD being the cause of various sociological damages, only half of the adults seek treatment [5], because they fear being negatively evaluated by a healthcare professional which, by itself, is a symptom of SAD.
- 2. Financial issues: SAD imposes extensive economic costs on individuals. The treatment costs are significantly higher for people that have SAD as compared to other chronic physical diseases. The studies by Sugano et al. [4] and Pampouchidou et al. [5] analyzed the data from the Netherlands' national mental health survey and estimated an annual mean cost of €11,952 per person. In addition, literature such as [6,7] also suggests that there is a rising cost of the diagnosis of psychological disorders.

The solution to mentioned problems could be solved by developing ICT-oriented applications that could automatically predict patients' SAD. There exist some examples of the use of ICTs to predict psychological disorders. For example, machine-learning algorithms are used to predict depression among senior citizens [8], an automated video-based recognition of depression and anxiety symptom severity [5], or to improve and predict the state of healthcare [9]. Machine learning is used for many other applications within social contexts, e.g., social media presence and analysis on Twitter or YouTube [10–12], or general applications of image or language processing [13–17]. However, research in using machine-learning techniques to predict SAD is scarce.

The existing studies on the prediction of SAD mostly used gaze interaction [8,18]. The proposed methods used live remote communication or manual coding to measure gaze interaction. There are multiple challenges present in the existing research related to SAD prediction. Firstly, techniques to detect gaze interaction use live remote communication. A remote video conversation via software like Skype is entirely different from live communication e.g., Face to Face (F2F), presentations, or an interview [18]. Secondly, the existing research that used live F2F communication heavily relied on humans to watch out for gaze interaction/avoidance and then manually record it [8,19]. Thirdly, no public dataset exists related to the prediction of SAD. A dataset is important and could prove to be a primary intellectual output of research [1]. Therefore, the aim of this research is also to prepare a gaze interaction and avoidance dataset to predict SAD. In addition, SAD is further categorized into four major classes as per the LSAS scale [9]. The proposed work is the pioneer study conducted to find a relationship between gaze interaction and SAD classes.

The rest of the paper is organized as follows: Section 2 describes related work conducted in ICT based context of SAD, while Section 3 describes methods and materials col-lected or used in this research. Section 4 explains the experimental setup and describes the analysis and results. Section 5 concludes the paper and explains the implications and po-tential future directions of the research.

2. Related Work

SAD is a disorder with a marked and persistent fear of social situations, which often is the cause of embarrassment or humiliation [20]. Individuals can experience social stress in a range of contexts, such as giving a public speech, speaking to a superior in a structural environment, or when engaged in unexpected interpersonal communication settings. Mostly, those who are socially anxious account for themselves as being shy when involved in an interpersonal setting or feeling some apprehension to communicate with others [21].

The research [20] explains that people with SAD have common symptoms such as:

- 1. Self-conscious in front of other people
- 2. Worrying for days or weeks before an event involving other people as well
- 3. Stay away from places where people involved
- 4. Tough time making friends and keeping friends
- 5. Some physical symptoms when they are with other people, like blushing, nausea, trembling, heavy sweating, and having a tough time talking.

If untreated, SAD can construct various kinds of psychopathology spanning internalizing and externalizing conditions [22]. The externalizing conditions incorporate a variety of disinhibited or externally focused behavioral symptoms, including aggression, conduct problems, delinquent behavior, oppositionality, hyperactivity, and attention problems. The internalizing conditions include a variety of over-inhibited or internally focused symptoms, including anxiety, fear, sadness/depression, social withdrawal, and somatic complaints [22]. Thus, SAD is an example of an internalizing condition. Another epidemiological study by Blalock [2] estimates that about 80% of people with SAD also experience a comorbid psychiatric condition such as other anxiety disorders and major depressive disorder (MDD).

Psychiatrists use the LSAS to predict someone's SAD condition. The LSAS is a short questionnaire developed in 1987 by psychiatrist Michael Liebowitz. However, this subjective scale depends on the correct answers provided by the subjects. Jahedi et al. [23] described some disadvantages of subjective scales such as:

- Subjective scale filling is a time-consuming task because it includes a proper understanding of the questionnaire and an exact answer for each question; otherwise, it will predict incorrect results.
- 2. It is an arduous task for children, semi-literate, and non-literate people to fill the subjective scale, as they cannot read and understand it properly.
- 3. A subjective scale needs aid in the filling and is thus prone to other people's bias.
- The subjective scale may be prone to produce incorrect results due to manual evaluation.

A subjective measurement can be improved by replacing it with some objective measures. With the advancement in machine-learning techniques, automatic gaze estimation has improved. Recent studies have proposed the automatic detection of gaze interaction with eye trackers or wearable cameras [4]. However, there are many central issues in detecting gaze interaction with such wearable devices. Firstly, it is hard for participants to wear eye trackers or cameras because they feel more self-conscious if they have psychological disorders [24]. Secondly, gaze interaction largely depends on the conditions of interaction [25]. For example, talking with close family and friends is a different kind of interaction than talking with strangers.

To overcome the shortcomings, the researchers have designed interfaces that help in the video annotation of gaze interaction manually. Manually here means a human annotates a video by tracking the gaze of the participants and inputting gaze avoidance or interaction as a coded value [8]. Recent studies also contributed to new instruments that efficiently browse various videos like sports [26], education [27], first-person videos [28], and surveillance [29]. These works used techniques based on computer vision and formulated modern GUI support for video browsing like-colored video timelines [30], content-aware fast-forwarding [31], and direct manipulation [32]. However, work related to real-time automatic gaze analysis to detect gaze for predicting SAD is scarce.

Researchers like Langer et al. [8] and Howell et al. [18] suggested that SAD is associated with gaze avoidance. Farabee et al. [19] estimated gaze interaction in a live situation where women saw the videos of the confederate. The videos were pre-recorded, but the women participants had no idea about it. For them, it was live remote communication. However,

one of the drawbacks of the study was the use of human coders to code gaze avoidance or gaze interaction. Such coding could be biased because of the human coder's involvement. Assuming a coder is not biased, there is still a possibility of human error due to loss of concentration, over-coding, or under-coding of the values. Such errors might result in a prediction error.

Weeks et al. [33] measured the gaze interaction in a live communication where all participants took part in a role-play task, which was a conversation with an unknown female (whom they were told was also a study participant). Two video cameras were fixed on the wall of the test room. A stopwatch was used to figure out the total number of seconds in which the participants maintained gaze interaction with their confederates during each minute of the communication. The drawback of this study was again the involvement of a human who used a stopwatch to measure the time of gaze interaction or gaze avoidance. Hence, the same flaws, as explained in an earlier paragraph, might have crept into the study as well.

Howell et al. [18] measured gaze avoidance by setting up a live remote environment. The participants interacted with the interviewer through a webcam. However, Langer et al. [8] argued that such a live remote environment setup might not be able to produce real anxiety situations in social environments as in live F2F communication and might not have produced the correct results. There are some other shortcomings suggested by Hebert et al. [34]. For example, the authors measured the gaze behavior in live communication, where participants talked with their real friends. A video from live communication was recorded. However, the problem is that talking to friends is what most of us do in normal daily life and thus might not be a trigger for SAD. Secondly, the coders used manual detection of gaze interaction while watching the video. As said earlier, the use of a manual technique to measure gaze interaction is not an efficient approach as human coders could miss an essential gaze interaction or avoidance moment. Furthermore, such coding might be prone to the coder's bias, who ultimately knew the purpose of the experiment. Therefore, there is a need for a technique to automate the measurement of gaze interaction in live communication to detect SAD.

Table 1 summarizes the discussion provided in the related work section. The first column shows the reference from where the parameters are extracted. The second column shows the use of a live environment or not, while the third column explains whether manual or automated coding is used in the research. The fourth column explains whether the research compares the sub-classes of SAD with gaze interaction or not. The fifth column supplies information about the production of any public dataset. The last column states whether any relationship was proved between the SAD and gaze interaction in the research or not. The table highlight research gaps such as no usage of automatic prediction of gaze in a live environment. Furthermore, no study proved a relationship between gaze interaction/avoidance with SAD classes. In addition, the lack of a dataset in SAD prediction is also a gap that this study aims to fill in.

Table 1.	Gaze	interaction	detection	method	to	measure	the	SAD	in	а	live	communic	ation
studies c	omparis	on.											

Ref	Live Environment	Automatic Coding	Relationship with SAD Subclasses	Dataset of SAD	SAD—Gaze Avoidance Relationship Proved
[33]	1	×	×	×	×
[8]	1	×	×	×	1
[18]	×	1	×	×	1
[19]	\checkmark	×	×	×	×

3. Materials and Methods

Accurate measurement and estimation are important in learning social behaviors such as the prediction of someone's personality [4]. However, this is a more serious job

when someone's health is involved. At the same time, the cost of such a prediction should also be affordable for everyone. To achieve "affordability" in this research, we decided to capture live F2F communication with the participants and for this purpose, an interview was set up with the participants. A built-in P.C. camera was used to capture the live F2F communication of the participants.

To automate the prediction process, the Haar Cascade classifier was identified as a method to detect human faces and their gaze interaction/avoidance. The method was proposed by Viola and Jones in 2001 [35] and is being used in the latest face detection, recognition, and gaze interaction prediction research. Examples include [36–38], etc. Besides, the algorithm is open-source and is easily available as linked libraries for languages like Python.

A detailed understanding of SAD and finding its relationship with gaze interaction could help health planners to design and implement the right healthcare delivery system. However, studies like [8,18,33] conducted experiments to predict only the presence or absence of SAD. As per the LSAS scale [9], SAD is further divided into four classes according to severity level. Treatment of SAD depends on the severity of the SAD as everyone can have a different severity level. Therefore, by extending the prediction of SAD into classes, psychiatrists can drill down the SAD cases and diagnose and prescribe the treatments accordingly. Therefore, to fulfill our third goal, we extended the procedure to include the prediction of SAD in four classes. A class showing the absence of SAD was also included. The fourth goal is to create a dataset with data collected from the selected participants. The first step toward data collection was to measure the SAD of the participants.

The participants were evaluated for SAD with an LSAS SAD evaluation scale before data collection. The choice of a suitable psychological evaluation scale for this purpose is discussed in the next subsection. The participants filled out the scale as per their subjective assessment of themselves. The information about the selection of the participants that engaged in the subjective data collection is provided in the second sub-section. A summary of the filtered participants for the second phase of automated data collection is provided in the "Selected participants" sub-section. The design and development of an application for automated data collection are discussed in the last sub-section.

3.1. Psychological Scale Selection

There are various scales available to measure SAD, but the most used is LSAS (Liebowitz Social Anxiety Scale) [39]. The scale is evaluated for its psychometric properties as per standards set in the Structured Clinical Interview for DSM Disorders (SCID) [40]. It is a semi-structured interview guide that explains the process of the assessment of current and lifetime Axis disorders. Some examples of these disorders are anxiety, mood disorders, psychosis, and substance use disorders. The latest version of SCID is called SCID-V. It is considered a standard instrument for the valuation of Axis psychopathology, as explained by the Diagnostic and Statistical Manual-V (DSM-V) [41]. SCID-V also has the LSAS [9], which is a standard clinician-administered interview to assess social anxiety symptoms.

The LSAS has shown exceptional divergent and convergent validity [18]. Scores on the LSAS can distinguish between patients with SAD and those without SAD [33]. It can also classify the participants into four more classes as follows. A score of less than 55 on the LSAS scale means that the participant belongs to the "No SAD" group. A participant with a score exceeding 55 on the LSAS scale means that he/she belongs to a SAD group. It further has criteria or thresholds for the symptoms of SAD. It includes the severity and length which are needed for a symptom to qualify for the diagnosis. The total score shows the severity of the individual's SAD in terms of 1) "No SAD" (\leq 55), 2) Marked (\geq 55 and \leq 65), 3) Moderate (\geq 65 and \leq 80), 4) Severe (\geq 80 and \leq 95), and 5) Very Severe (\geq 95). The comparison of SAD subclasses can be useful for three reasons:

- To better understand the relationship between SAD and gaze interaction by comparing all classes as each class might have a different kind of gaze interaction pattern,
- 2. For correct prediction of SAD using gaze interaction for each class separately, and

3. To supply effective treatment of SAD for each class as each SAD person in a different class might not need the same kind of treatment.

3.2. Criteria for the Selection of Participants

As explained in the related work section, subjective assessment scales have various disadvantages. However, to predict SAD automatically, there is a requirement to have subjective assessment data of the participants to overcome some of the disadvantages like difficulty in filling out scale by children and non-literate. The following four conditions were applied to make the results correct:

- (1) Participants above the age of 18 were considered.
- (2) The educational qualification of all participants was taken above 10th grade.
- (3) For each class of SAD, the number of male and female participants was taken equally.
- (4) The number of participants selected was greater than the existing literature review on SAD prediction.

The experimenter informed participants about the experiment verbally and in the form of written information on the provided consent form. It was necessary to fill out the form to show the participants' volunteer participation agreement for the experiment. One of the participants supplied informed consent to publish the information/image(s) in an online open-access publication. All the methods were conducted following World Health Organization (WHO) relevant guidelines and regulations. The process of data collection, taking consent, and the ethical procedure followed was approved by the COMSATS University Islamabad, Research, and Evaluation Committee (CUI-REC) co-located in COMSATS University Islamabad, Abbottabad Campus, Pakistan. The committee also looks after the ethical issues in research.

After the consent, the participants were provided with the printed LSAS scale. The experimenter also supplied instructions on filling it out accurately. After filling out the scale, the participants returned it to the experimenter. The filled scales by 180 participants were manually analyzed one by one, and the final score of each participant was computed. Each of the participants was grouped based on the final score. The grouping was based on the LSAS threshold values to declare someone with "No SAD" or to any of the four classes: Marked, Moderate, Severe, and Very Severe. Most of the participants i.e., 66 had a "No SAD" group, 43 had Moderate SAD, 39 were with Marked SAD, 10 were identified as having Severe SAD, and 22 had "Very Severe" SAD as per an analysis of LSAS scores. In the phase, we investigated 180 participants, among which 77.2% were male and 22.8% were female. As there were only 10 participants in the severe class, therefore, it was decided to collect data from only 50 participants (10 from each class) in the second phase.

The percentage distribution of the data collection at various locations is illustrated in Table 2 below.

S. No	Location	Percentage of Participants	Description				
1.	Home	10%	The data was collected at the locations where it was most convenient				
2.	University	12.8%	for the participants. This step was especially important to collect data from women as due to social and cultural barriers most women could				
3.	College	77.2%	not travel across to various locations for the experiment				

Table 2. Data collected at various locations per participants' convenience.

3.3. Final Selection of the Participants

As per our planned experimental setup section (c.f. "Experimental setup"), the procedure of collecting the data from a single participant could take about an hour, including consent agreement and delivering setting instructions. Furthermore, the non-availability of some participants for the experiment resulted in the wastage of the time slot and reallocation to another time slot. Besides, the experiment was conducted at three separate locations (c.f. Table 2) and was a time-consuming process. Moreover, we found only 10 participants in the severe SAD class after analyzing 180 participants in Phase 1. Therefore, it was decided to collect data from only 50 participants (10 from each class) in the second phase. This process helped us to save experimentation time as well as to experiment with enough participants. Various existing research conducted similar experiments with a lesser or almost equal number of participants [34].

All participants with an LSAS score range of "Severe SAD" were therefore selected for the final experiment. This group had 5 male and 5 female participants. The participants from the rest of the classes were selected randomly. All the remaining groups were selected with an equal number of male and female participants as well. This process was done to avoid any influence of gender as a covariate.

3.4. Development of the Application

OpenCV 3.4.3 and Python version 2.7.14 were used to develop the application. We applied Haar cascade Classifier to the live videos of the participants. The developed application works in a sequence of the following steps (c.f. Figure 1).



Figure 1. Process of gaze detection with the developed application.

Step 1: Each participant's interaction with the interviewer was captured using an RGB camera in the form of a video.

Step 2: Each frame from the video was input into a Haar cascade classifier to detect a face in the frame.

Step 3: By using the feature extraction, gaze was detected.

Step 4: The gaze pupil was marked as a bounding box, as shown in Figure 2.

Step 5: A Calibration file was created in Python to set the reference values for each participant's gaze interaction. The calibration feature was used as described [42]. For example, corresponding eX and eY positions were stored as the horizontal and the vertical coordinates, and eW and eH were captured as the height and width of the bounding box. Reference values were unique for each participant and were figured out according to the participant's gaze bounding box coordinates.

Step 6: Selected reference values were compared with the participant's gaze values in real time.

- 1. If the real-time video values of a participant were within the limits of selected reference values of gaze interaction, then it was considered gaze interaction.
- 2. If the real-time video values of a participant were out of the limits of the selected reference value of gaze interaction, then gaze avoidance was considered.



Figure 2. Gaze interaction and bounding box formation.

Step 7: The gaze avoidance was further divided into four subclasses (gazing above, gazing below, gazing right, gazing left) by detecting coordinate values of the bounding box during gaze avoidance (c.f. Figure 3). The decision to collect these data was taken to enrich the dataset, as well as to test any possible links between these gaze behaviors with SAD.



Figure 3. Gaze avoidance. (a) Looking left, (b) looking above, (c) looking right, (d) looking below.

3.5. Experimental Setup

The experiment was conducted in the Fall 2021 semester from 6 September 2021 to 31st December 2021. As explained earlier (c.f. Table 2), three separate locations were used to collect data. The experiment was conducted in a separate room for each location. In college and university, a room for the execution of the experiment was made available. Participants who opted to take the experiment in their homes or near their homes were invited for a conversation session in a nearby primary school. It was ensured that no external noise could disturb the participants. The room had a table and two chairs. A laptop facing the interviewer was placed on the table with an external camera mounted for automatic detection of gaze via the proposed application. The examiner sat on one chair, and participants were called to sit on the other chair for the interview. Figure 4 shows the experimental setup of the proposed work.



Figure 4. Experimental setup.

3.5.1. Conversation Used for Experimental Study

After the scale-filling phase, the participants were invited for a conversation session as per the setup explained in the prior section. It took an average of 5 days to analyze the collected scale data and to inform the finally selected participants for the experiment. The conversation session was planned to be 5 min without any break. The session was video recorded for later use and to create a dataset of SAD. The conversation used for this study was the same as used by [8] and is described in Table 3 below. The observer was asked to act like an interviewer who wanted to know about the personality of the participants.

Conversation Domain	Task Explanation	Time Frame
Introduction Conversation	The observer and participant were introduced to each other.	1 min
Social Support Conversation	The participants were asked to choose something that they want to change about themselves.	2 min
Conflict Conversation	The participants were asked to tell something about their best friends and what they want to change in them.	Minutes

 Table 3. The conversation contents and procedure.

3.5.2. Data Preparation

As the first step of the analysis, data were pre-processed to prepare a dataset. In the recorded real-time conversation, gaze interaction was measured by using the developed application (c.f. development of the application). The data labeling process as explained in the latter part of this section was used to predict whether to consider the extracted data values of a participant's gaze as gaze interaction or avoidance. We collected 5 min of video from 50 participants. We took 6 frames per second (fps) to collect data from each frame. This makes 1800 gaze values for each participant in a 5-min conversation. The collected data were used to create a balanced dataset of gaze interaction/avoidance to detect SAD in real-time environments. Our dataset holds 90,000 values of gaze interaction/avoidance values against interactions of 50 Participants. This dataset could prove to be useful in the future to find the relationship between SAD with other combinations of SAD symptoms like trembling voice to predict SAD. The dataset has three parts: "Annotation", "Calibration", and "Data".

The annotations include:

- 1. Detected eye landmark's location on the whole image coordinate in pixels.
- 2. The target position of gaze in the screen coordinate.
- 3. 3D gaze target position compared to the camera.
- 4. The assessed 3D left gaze center in the camera coordinate system.
- 5. The assessed 3D right gaze center in the camera coordinate system.
- 6. Labeling in terms of gaze interaction and gaze avoidance.
- 7. Gaze avoidance labeling is further extended in terms of looking above, looking below, and looking toward left and right.

Besides, there is also a "Calibration" file for each participant, which has:

- 1. The intrinsic parameter of the laptop camera.
- 2. The projection matrix of the camera.
- 3. The position of the image plane in the camera coordinate.
- 4. The laptop screen size.
- 5. The screen height and width in pixels.
- 6. The screen height and width are in millimeters.

The "Data" file had:

- 1. Gaze prediction graph for an individual participant.
- 2. Score file that has the scores of subjective scale and automatic gaze prediction.
- 3. The recorded file of real-time gaze prediction.

The data file can be found in the Supplementary Materials of this manuscript.

4. Analysis, Results, and Discussions

This section includes a comparison of SAD classes in terms of gaze interaction/avoidance. The interaction and avoidance data for each class were computed separately. As can be seen from Figure 5, participants in the "No SAD" class had the highest gaze interaction, with an average percentage corresponding to 82.9%. This shows that people who do not have SAD are more likely to keep gaze interaction. The gaze interaction average percentage tends to decrease as the severity of the SAD tends to increase. As Figure 5 shows, the marked,

moderate, severe, and very severe average gaze interaction percentages decreased to 64.7%, 46.6%, 31%, and 27.4%, respectively. The largest gaze interaction percentage of a participant in the "No SAD" group was 88%, in the Marked group it was 70.50%, in the Moderate group it was 53.70%, in the Severe group it was 37.80%, and in the 'Very Severe' group it was 34.60%. An overall analysis revealed a linear relationship between gaze interaction and increasing severity of SAD.





The next analyses were conducted to analyze the relationship between subgroups such as females vs. males. In addition, data analysis in the context of parameters like accuracy, precision, and recall was also conducted.

4.1. Relationship between Gender-Based Gaze Interaction/Avoidance and SAD Classes

The data were divided based on gender groups and processed for each SAD class separately. Figure 6 describes the relationship between SAD and the gaze interaction of female participants. The graph in Figure 7 describes the relationship between SAD and the gaze interaction of male participants. From the results, the gaze interaction trend becomes negative as SAD class severity increases in both male and female cases. Thus, it can be said that the SAD prediction method, as described in this research, applies to all genders.









4.2. Is SAD Predictable via Gaze Avoidance sub-Behaviors?

The gaze avoidance data were further categorized into looking below, looking above, looking right, and looking left behaviors (c.f. Section 3.4.). It was hypothesized that participants with SAD would show dominant looking-down behavior. This finding could give us a possibility of detecting SAD, even from a sub-behavior of gaze avoidance. Figure 8 shows female gaze avoidance behavior further sampled into the four sub-behaviors. In the gaze avoidance category, 35%, and 34.47% of the sub behavior in 'severe' and 'very severe' participants were "looking below" behavior. Similarly, Figure 9 shows male gaze avoidance behavior, which shows that looking below behavior intensity is lesser in males, i.e., 20% and 20.4% in case of severe and 'very severe' classes. However, this sub-behavior still forms major gaze avoidance behavior in males as well. A paired sample t-test conducted on the joint gender data showed a significant difference between looking below behavior and all other avoidance behaviors, as shown in Table 4. Furthermore, all the participants showed this behavior with medium or large effect size as Cohen's D column in Table 4 shows. Figures 8 and 9 also show the trend that as SAD class severity increases, looking below behavior also increases, which is not the case for all other gaze avoidance sub-behaviors. This is a sign that it might be possible to predict SAD by only looking below sub behavior, which in turn supports the hypothesis that looking below behavior is associated with SAD, as well as the severity of SAD.

Table 4. Pair sample independent t-test to show the difference between looking below sub-behavior and all other sub-behaviors.

Test Pair	t	Df	Sig. (2-tailed)	Cohen's D
Looking Below—Gaze Interaction	-7.427	49	0.000	-1.05
Looking Below—Looking Above	4.054	49	0.000	0.57
Looking Below—Looking Right	2.630	49	0.011	0.37
Looking Below—Looking Left	4.096	49	0.000	0.58

4.3. SAD Prediction Using Gaze Interaction Dataset

A decision tree classifier was used to predict SAD. The decision tree does not require computationally intensive data processing as well as it does not require assumptions of the distribution of data. Similarly, this algorithm is useful for finding hidden patterns in a dataset. It can exclude unimportant features and its accuracy is comparable to other classification techniques [43]. The data collection is simple with a non-normal distribution of data. Therefore, a decision tree classifier was used to predict SAD. The data were fed



to the decision tree, which classified data into five classes (one No SAD class and four SAD classes).

Figure 8. Comparison between gaze avoidance behavior of females.



Figure 9. Comparison between gaze avoidance behavior of males.

The classifier was used to predict an overall accuracy as well as to predict the accuracy of the sub-classes as well. The earlier literature only differentiated between the SAD group (A group of participants that have a social anxiety disorder) and the 'No SAD' group (A group of participants that have no social anxiety disorder). In this way, the sub-class classification was a novel process executed in this research. The results show that the overall accuracy of this method was 80%. The confusion matrix of the dataset prediction of SAD is illustrated in Figure 10.

From the confusion matrix, the precision and recall scores were calculated as illustrated in Table 5. As can be seen from the data in Table 5, the recall percentage of 'No SAD', Marked, Moderate, Severe, and very severe groups were 80%, 70%, 90%, 80%, and 80%, respectively. The third column refers to recall, i.e., the percentage of total relevant results of each class that is correctly classified by the algorithm. The precision scores of 'No SAD', Marked, Moderate, Severe, and very severe groups were 73%, 88%, 75%, 80%, and 89%, respectively.

		Predicted Classes						
	-		0	1	2	3	4	
	0]]	8	1	0	1	0]	
	1	[2	7	1	0	0]	
Actual Classes	2	[1	0	9	0	0]	
	3	[0	0	1	8	1]	
	4	[0	0	1	1	8]]	

Figure 10. Confusion matrix of dataset. Class 0, 1, 2, 3, and 4 stand for 'No SAD', Marked, Moderate, Severe, and Very Severe groups, respectively.

Table 5. Recall, precision score of SAD classes.

Class	Recall	Precision
0—No SAD	0.80	0.73
1—Marked	0.70	0.88
2—Moderate	0.90	0.75
3—Severe	0.80	0.80
4—Very Severe	0.80	0.89

These results show that there is a possibility of replacing a subjective LSAS scale of SAD prediction with an objective method defined in this research with an accuracy of 80%. As shown earlier (c.f. Section 2—Related Work), the subjective scale has many disadvantages, including a biased response from the participants that could affect the accuracy of the SAD prediction with the LSAS scale. There is a possibility of replacing that scale with the process described in this research. Otherwise, the process could be used along with the LSAS scale for a more correct assessment and prediction of SAD.

5. Conclusions, Implications, and Future Directions

This study investigated the possibility of automatic social anxiety disorder (SAD) prediction. The experiments were conducted in a live environment with human participants by using a custom-developed application. The application used the Haar Cascade classifier. SAD classes were predicted by using decision tree classifiers from the created gaze dataset. The results proved that SAD could be predicated with an overall accuracy of 80%. Furthermore, each class of SAD could be predicted with accuracies of 80%, 70%, 90%, 80%, and 80%, respectively. The possibility of replacing a subjective scale (LSAS) with an objective technique used in this research is discussed.

The results suggest that individuals with SAD make lower gaze interactions during a social situation conversation. The results also suggest that as SAD class severity increases, the gaze interaction decreases. This study contributed in the following ways: (1) detected gaze interaction automatically in a real-time conversation using a simple RGB camera, (2) annotation (labeling gaze in real-time) is done in terms of gaze interaction and gaze avoidance to find the relationship between SAD and gaze interaction, (3) pioneered in finding a relationship between gaze interaction and SAD classes, (4) we created first gaze dataset that could be used in future analyses of SAD individually or along with a combination of some other SAD symptoms.

Findings from the current study provide further support that individuals with elevated SAD are likely to engage in lower gaze interaction, which may have negative impacts on social functioning and contribute to the maintenance of SAD symptoms. Importantly, this support was derived using several novel methodological approaches that addressed weaknesses from earlier studies extending the external and ecological validity of this body of research. Our results confirmed that the larger trait of social anxiety severity was

inversely associated with the frequency and duration of gaze interaction throughout the totality of a conversation task. Finally, our results supply implications for psychotherapy. Given that gaze interaction is an essential part of meaningful human interaction, actively targeting improvement in gaze interaction as a part of SAD treatment will likely increase the quality of life for SAD patients. It is also essential to consider the nature of the social stimulus (e.g., positive; ambiguous), duration, and CBT skill development stage when preparing in vivo exposures to boost positive corrective experiences and/or successful self-regulation.

The results of the current study should be viewed considering its limitations. The results of the study revealed that the female participants showed "looking below" behavior more than their male counterparts in each class, i.e., 5% as compared to 4.6% in the "No SAD" class, 11.2% vs. 7.4% in the "Marked" class, 16.4% vs. 14% in the "Moderate" class, 35% vs. 20% in the "Severe" class, and 35.5% vs. 20.4% in the "Very Severe" class. Therefore, there is a possibility of future research that decides whether this difference is a natural difference across the world, or if it is more associated with Asian culture generally and specifically in the Muslim world. Furthermore, there is a need to separate SAD from traditional shyness that is part of the eastern and Muslim cultures.

A relationship between gaze interaction with SAD classes was found with 50 participants (10 in each class). In the future, a larger sample size can be used to predict and deeply analyze the relationship between gaze and SAD within its classes. Arguably, bigger samples are essential for detecting existing relationships between self-reported indices. Overall, further expansion of this type of research to larger and more diverse samples will help in advancing the objective assessment of psychological disorders. Only a decision tree classifier was used for SAD prediction. Tables 6 and 7 contain multi-class confusion matrix data for our SAD classes. The results of the tables could be used to filter out the 20% of the cases that might not be predictable from the data. However, to do that, there are only 5 features available in the present dataset to predict SAD. In the future work, kernel tricks method can be used to engineer more features to increase classification accuracy. In addition, more complex decision classifiers such as Random Forest can also be used to increase classification accuracy. The prediction can be evaluated using many other machine-learning algorithms (e.g., ANN, SVM, Bayes, etc.) for classification tasks to find the best one for the task.

Class	ТР	FN	FP	TN
0—No SAD	8	2	3	37
1—Marked	7	3	1	39
2—Moderate	9	1	3	37
3—Severe	8	2	2	38
4—Very Severe	8	2	1	39

Table 6. Multi-class confusion matrix output for TP, FP, FN, and TN.

 Table 7. Multi-class confusion matrix output.

Class	Specificity	Positive Likelihood Ratio	Negative Likelihood Ratio	Positive Predictive Value	Negative Predictive Value	Accuracy
0—No SAD	92.50%	10.67	0.22	72.73%	94.87%	90.00%
1—Marked	97.50%	28.00	0.31	87.50%	92.86%	92.00%
2—Moderate	92.50%	12.00	0.11	75.00%	97.37%	92.00%
3—Severe	95.00%	16.00	0.21	80.00%	95.00%	92.00%
4—Very Severe	97.50%	32.00	0.21	88.89%	95.12%	94.00%

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/app122312298/s1.

Author Contributions: Conceptualization, I.A.K. and S.S. (Sajid Shah); Data curation, S.S. (Sara Shafique) and I.A.K.; Funding acquisition, M.E.; Investigation, S.S. (Sara Shafique), S.S. (Sajid Shah), and W.J.; Methodology, I.A.K., R.N.J. and M.E.; Resources, R.N.J.; Software, S.S. (Sara Shafique), S.S. (Sajid Shah) and W.J.; Validation, W.J.; Writing—original draft, S.S. (Sara Shafique) and I.A.K.; Writing—review & editing, R.N.J. and M.E. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no funding.

Institutional Review Board Statement: All the methods were conducted following World Health Organization (WHO) relevant guidelines and regulations. The process of data collection, taking consent and, the ethical procedure followed was approved by the COMSATS University Islamabad, Research and Evaluation Committee (CUI-REC) co-located in COMSATS University Islamabad, Abbottabad Campus, Pakistan. The committee also looks after the ethical issues in research.

Informed Consent Statement: The experimenter informed participants about the experiment verbally and in the form of written information on the provided consent form. It was necessary to fill out the informed consent form to show the participants' volunteer participation agreement for the experiment. One of the participants supplied informed consent to publish the information/image(s) in an online open-access publication (c.f. Figure 3).

Data Availability Statement: The data can be found attached to the manuscript Supplementary Materials.

Acknowledgments: This work has been supported by EIAS Data Science Lab, Prince Sultan University, KSA. The authors would like to thank EIAS Data Science Lab and Prince Sultan University for their encouragement, support, and the facilitation of resources needed to complete the project.

Conflicts of Interest: The authors declare no conflict of interest. Furthermore, the funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

- Dekker, R. The importance of having data sets. In Proceedings of the IATUL Conferences, Porto, Portugal, 23 May 2006; Available online: https://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1760&context=iatul (accessed on 1 November 2022).
- Blalock, D.V.; Kashdan, T.B.; Farmer, A.S. Trait and Daily Emotion Regulation in Social Anxiety Disorder. Cogn. Ther. Res. 2016, 40, 416–425. [CrossRef]
- 3. McCroskey, J.C. The Communication Apprehension Perspective. In *Avoiding Communication: Shyness, Reticence, and Communication Apprehension;* Daly, J.A., McCroskey, J.A., Eds.; SAGE: Beverly Hills, CA, USA, 1984; pp. 13–38.
- Sugano, Y.; Fritz, M.; Bulling, X.A. It's Written All over Your Face: Full-Face Appearance-Based Gaze Estimation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 21–26 July 2017; pp. 51–60.
- Pampouchidou, A.; Pediaditis, M.; Kazantzaki, E.; Sfakianakis, S.; Apostolaki, I.A.; Argyraki, K.; Manousos, D.; Meriaudeau, F.; Marias, K.; Yang, F.; et al. Automated Facial Video-Based Recognition of Depression and Anxiety Symptom Severity: Cross-Corpus Validation. *Mach. Vis. Appl.* 2020, 31, 30. [CrossRef]
- Nina, E.; Leibing, E.; Leichsenring, F.; Manfred, B.; Herpertz, S.; Hoyer, J.; Konnopka, A.; Salzer, S.; Strauss, B.; Wiltink, J.; et al. The costs of social anxiety disorder: The role of symptom severity and comorbidities. J. Affect. Disord. 2014, 165, 87–94. [CrossRef]
- Hassan, F. Psycho-Social Impacts of Mental Disorders in India and Emergent Needs of Evidence Based Policies for Public Health. J. Adv. Res. Sci. Soc. Sci. 2018, 1, 83–90.
- Langer, J.K.; Lim, M.H.; Fernandez, K.C.; Rodebaugh, T.L. Social Anxiety Disorder Is Associated with Reduced Eye Contact during Conversation Primed for Conflict. *Cogn. Ther. Res.* 2017, 41, 220–229. [CrossRef]
- Heimberg, R.G.; Horner, K.J.; Juster, H.R.; Safren, S.A.; Brown, E.J.; Schneier, F.R.; Liebowitz, M.R. Psychometric Properties of the Liebowitz Social Anxiety Scale. *Psychol. Med.* 1999, 29, 199–212. [CrossRef]
- 10. Khan, M.U.; Bukhari, S.M.; Maqsood, T.; Fayyaz, M.A.; Dancey, D.; Nawaz, R. SCNN-Attack: A Side-Channel Attack to Identify YouTube Videos in a VPN and Non-VPN Network Traffic. *Electronics* **2022**, *11*, 350. [CrossRef]
- Sarwar, R.; Zia, A.; Nawaz, R.; Fayoumi, A.; Aljohani, N.R.; Hassan, S.U. Webometrics: Evolution of social media presence of universities. *Scientometrics* 2021, 126, 951–967. [CrossRef]
- 12. Hassan, S.U.; Aljohani, N.R.; Idrees, N.; Sarwar, R.; Nawaz, R.; Martínez-Cámara, E.; Ventura, S.; Herrera, F. Predicting literature's early impact with sentiment analysis in Twitter. *Knowl. -Based Syst.* **2020**, *192*, 105383. [CrossRef]
- 13. Hassan, H.; Bashir, A.K.; Ahmad, M.; Menon, V.G.; Afridi, I.U.; Nawaz, R.; Luo, B. Real-time image dehazing by superpixels segmentation and guidance filter. *J. Real-Time Image Process.* **2021**, *18*, 1555–1575. [CrossRef]

- 14. Riaz, F.; Rehman, S.; Ajmal, M.; Hafiz, R.; Hassan, A.; Aljohani, N.R.; Nawaz, R.; Young, R.; Coimbra, M. Gaussian mixture model based probabilistic modeling of images for medical image segmentation. *IEEE Access* **2020**, *8*, 16846–16856. [CrossRef]
- Iqbal, S.; Hassan, S.U.; Aljohani, N.R.; Alelyani, S.; Nawaz, R.; Bornmann, L. A decade of in-text citation analysis based on natural language processing and machine learning techniques: An overview of empirical studies. *Scientometrics* 2021, 126, 6551–6599. [CrossRef]
- Zaman, F.; Shardlow, M.; Hassan, S.U.; Aljohani, N.R.; Nawaz, R. HTSS: A novel hybrid text summarization and simplification architecture. *Inf. Processing Manag.* 2020, 57, 102351. [CrossRef]
- Khan, U.A.; Khan, I.A.; Din, A.; Jadoon, W.; Jadoon, R.N.; Khan, M.A.; Khan, F.G.; Khan, A.N. Towards a Complete Set of Gym Exercises Detection Using Smartphone Sensors. *Sci. Program.* 2020, 2020, 6471438. [CrossRef]
- Howell, A.N.; Zibulsky, D.A.; Srivastav, A.; Justin, W. Weeks. Relations among Social Anxiety, Eye Contact Avoidance, State Anxiety, and Perception of Interaction Performance during a Live Conversation. *Cogn. Behav. Ther.* 2016, 45, 111–122. [CrossRef]
- Farabee, D.J.; Ramsey, S.L.; Cole, S.G. Social Anxiety and Speaker Gaze in a Persuasive Atmosphere. J. Res. Personal. 1993, 27, 365–376. [CrossRef]
- Langer, K.J.; Tonge, N.A.; Piccirillo, M.; Rodebaugh, T.L.; Thompson, R.J.; Gotlib, I.H. Symptoms of Social Anxiety Disorder and Major Depressive Disorder: A Network Perspective. J. Affect. Disord. 2019, 243, 531–538. [CrossRef]
- Maryam, M.; Mashhadi, A.; Shamloo, Z.S.; Leahy, R.L. The Effectiveness of Group Emotional Schema Therapy on Emotional Regulation and Social Anxiety Symptoms. *Int. J. Cogn. Ther.* 2019, 12, 16–24.
- 22. Willner, J.C.; Gatzke-Kopp, L.M.; Bray, B.C. The Dynamics of Internalizing and Externalizing Comorbidity across the Early School Years. *Dev. Psychopathol.* 2016, 28 Pt 1, 1033. [CrossRef]
- Jahedi, S.; Mendez, F. On the Advantages and Disadvantages of Subjective Measures. J. Econ. Behav. Organ. 2014, 98, 97–114. [CrossRef]
- Montgomery, K.; White, C.; Thompson, A. A Mixed Methods Survey of Social Anxiety, Anxiety, Depression and Wig Use in Alopecia. *BMJ Open* 2017, 7, e015468. [CrossRef] [PubMed]
- Hamilton, A.F.C. Gazing at Me: The Importance of Social Meaning in Understanding Direct-Gaze Cues. *Philos. Trans. R. Soc. B Biol. Sci.* 2016, 371, 20150080. [CrossRef] [PubMed]
- 26. Tank, D. A Survey on Sport Video Summarization. Int. J. Sci. Adv. Res. Technol. 2016, 2, 189924626.
- Kitamura, N.; Shishido, H.; Enomoto, T.; Kameda, Y.; Yamamoto, J.-I.; Kitahara, I. Development of Multi-View Video Browsing Interface Specialized for Developmental Child Training. In Proceedings of the 12th Asia Pacific Workshop on Mixed and Augmented Reality (APMAR), IEEE, Ikoma, Japan, 28–29 March 2019; pp. 1–8.
- Nakazawa, A.; Mitsuzumi, Y.; Watanabe, Y.; Kurazume, R.; Yoshikawa, S.; Honda, M. First-Person Video Analysis for Evaluating Skill Level in the Humanitude Tender-Care Technique. J. Intell. Robot. Syst. 2019, 98, 1–16. [CrossRef]
- Ghatak, S.; Rup, S.; Majhi, B.; Swamy, M.N.S. An Improved Surveillance Video Synopsis Framework: A HSATLBO Optimization Approach. *Multimed. Tools Appl.* 2019, 79, 1–33. [CrossRef]
- Olsen, W.A.; Seymour, D.; Findlay, M.A.; Watson, H.S.; Mozell, W.M. Video Service with Automated Video Timeline Curation. U.S. Patent 10,068,614, 4 September 2018.
- Abibouraguimane, I.; Hagihara, K.; Higuchi, K.; Itoh, Y.; Sato, Y.; Hayashida, T.; Sugimoto, M. CoSummary: Adaptive Fast-Forwarding for Surgical Videos by Detecting Collaborative Scenes Using Hand Regions and Gaze Positions. In Proceedings of the 24th International Conference on Intelligent User Interfaces, Marina del Ray, CA, USA, 17–20 March 2019; pp. 580–590.
- Shugrina, M.; Zhang, W.; Chevalier, F.; Fidler, S.; Singh, K. Color Builder: A Direct Manipulation Interface for Versatile Color Theme Authoring. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, ACM, Glasgow, UK, 4–9 May 2019; p. 456.
- 33. Weeks, J.W.; Howell, A.N.; Goldin, P.R. Gaze Avoidance in Social Anxiety Disorder. Depress. Anxiety 2013, 30, 749–756. [CrossRef]
- 34. Hebert, E.A.; Dugas, M.J. Behavioral Experiments for Intolerance of Uncertainty: Challenging the Unknown in the Treatment of Generalized Anxiety Disorder. *Cogn. Behav. Pract.* **2019**, *26*, 421–436. [CrossRef]
- 35. Viola, P.; Jones, M.J. Robust real-time face detection. Int. J. Comput. Vis. 2004, 57, 137–154. [CrossRef]
- Windy, D.; Hilmawan, S.V. Implementasi Face Recognition secara Real-time dengan Metode Haar Cascade Classifier menggunakan OpenCV-Python. UG J. 2022, 16, 51–59.
- Shamrat, F.J.M.; Majumder, A.; Antu, P.R.; Kumar, S.; Barmon, I.N.; Ranjan, R. Human Face Recognition Applying Haar Cascade Classifier. In *Pervasive Computing and Social Networking*; Springer: Singapore, 2022; pp. 143–157.
- Premaladha, J.; Reddy, M.S.; Reddy, T.H.K.; Charan, Y.S.S.; Nirmala, V. Recognition of Facial Expression Using Haar Cascade Classifier and Deep Learning. In *Inventive Communication and Computational Technologies*; Springer: Singapore, 2022; p. 335.
- 39. Liebowitz, M.R. Social phobia. *Mod. Probl. Pharmacopsychiatr.* **1987**, 22, 141–173.
- 40. LaBruzza, A.L. Using DSM-IV: A Clinician's Guide to Psychiatric Diagnosis; Jason Aronson: Lanham, MD, USA, 1994.
- 41. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5*®); American Psychiatric Pub.: Washington, DC, USA, 2013.
- 42. Katrychuk, D.; Griffith, H.; Komogortsev, O. A Calibration Framework for Photosensor-Based Eye-Tracking System. In ACM Symposium on Eye Tracking Research and Applications; Association for Computing Machinery: New York, NY, USA, 2020; pp. 1–5.
- Chakure, A. Decision Tree Classification. *Medium* 2019, 1, 1–10. Available online: https://builtin.com/data-science/classification-tree (accessed on 1 November 2022).