Computational study of psychosis symptoms and facial expressions

Supriya Vijay
Carnegie Mellon University
supriyav@cs.cmu.edu

Tadas Baltrušaitis
Carnegie Mellon University
tbaltrus@cs.cmu.edu

Luciana Pennant
Harvard Medical School
lpennant@mclean.harvard.edu

Dost Öngür
Harvard Medical School
dongur@partners.org

Justin T. Baker
Harvard Medical School
jtbaker@partners.org

Louis-Philippe Morency
Carnegie Mellon University
morency@cs.cmu.edu

ABSTRACT

Nonverbal behavior is an essential part of face-to-face clinical interactions, especially in mental health settings. Facial expressions can reveal information about a person's emotions, mental state and social intentions, and is routinely used by both patients and their doctors in a variety of ways that can impact healthcare outcomes. In our work we demonstrate how automated tools of facial expression analysis can help in assessing a number of clinical scales that measure severity of psychosis symptoms. We analyze nonverbal behaviors during semi-structured clinical interviews and find interesting results at the whole session level and when looking at individual responses to clinicians questions. We demonstrate the importance of such analysis and highlight a number of behavioral indicators related to various psychosis symptoms.

Author Keywords

Facial expressions, Schizophrenia, Nonverbal behavior, Clinical scales

ACM Classification Keywords

J.3. Life and medical sciences: Health informatics

INTRODUCTION

Nonverbal behavior plays an important role in human communication [3] and is an essential part of face-to-face clinical interactions, especially in mental health settings. Facial expressions, gaze, body gestures and vocal prosody can reveal information about a person's emotions, mental state and social intentions, and is routinely used by both patients and their doctors in a variety of ways that ultimately can impact healthcare outcomes [11].

While the medical community is increasingly embracing the important role nonverbal communication plays in clinical settings [11], a major challenge for reliable and effective behavioral and mental health care is the lack of objective markers of illness. Although astute clinicians can capture key nuances in patient behavior and thought process, a mental healthcare system which relies solely on expert human judgment to pick up on key diagnostic and prognostic features of illness is (a) costly, (b) extremely difficult to monitor for efficacy, and (c) leads to idiosyncratic clinician behavior that leads to highly variable clinical outcomes.

The focus of our work is on creating novel methods for examining clinical behavior by identifying behavioral indicators relevant to various symptoms. Application of such novel approaches to the study of psychiatric populations could revolutionize the mental healthcare system, providing a needed method to collect objective behavioral data to aid clinical decision making and prediction of clinical outcomes. Importantly, the goal of this work is not to replace clinicians, but rather to augment their abilities, making practioners more effective, and to begin to provide much needed evidence to scaffold and help shape clinical practices.

In our work we are especially interested in identifying behavior indicators relevant to certain symptoms. We do this by analyzing the behavior of a patient during a semi-structured clinical interview. This analysis is done both at the *whole session* level and by breaking down the session into responses to particular questions asked by the clinician. We refer to the latter method as *context* based as it looks at behaviors at different contexts/stages of an interview. Our work demonstrates the importance of looking both at behavior overall and during specific contexts when assessing psychosis symptoms. Furthermore, we identify the relative importance of parts of structured interview at assessing symptoms and make a number of observations of interesting behaviors in our dataset.

RELATED WORK

Bedi et al. [6] looked at early psychosis onset detection in high-risk youths. They analysed transcripts of clinical interviews and were successful at predicting later psychosis development based on language coherence from a sample of 34 youths. Our work, on the other hand, looks at a sample already suffering from psychosis and instead looks at nonverbal behaviors as predictors of psychosis symptoms.

Stratou et al.[15] looked at the effects of gender on nonverbal visual behaviors of sufferers of PTSD and depression. However, they only looked at differences between healthy individuals and PTSD sufferers and did not look at the severity of the illness.

Figure 1: Example of the recording setup of the clinician an the patient during the semi-structured interview in our dataset.

A large portion of work looking at predicting mental health scales and symptoms has concentrated on depression analysis. Cohn et al. [7] looked at using facial expressions and acoustic features to detect depression. Alghowinem et al. [2, 1] investigated use of eye and head pose tracking technologies for depression detection and the ability of approaches to generalise across cultures. Finally, work by Girard et al. [10] compared automatic and manual techniques for depression analysis and demonstrated suitability of automatic techniques for various behavioral hypothesis testing.

DATASET

We collected a dataset of 18 adult patients (4 females) currently hospitalized on the inpatient psychiatric units at a major psychiatric hospital. Each patient underwent a semi-structured naturalistic clinical interview, similar to the daily clinical encounter with their treatment team (e.g., clinical rounds). Each interview lasted approximately 10-15 minutes and used scripts we have modeled on routine clinical interactions to extract an array of psychiatric attributes needed to establish mental status. Before the interview started each patient was left on their own for about 2 minutes with the recording equipment running. Typical structure of the interview can be seen in Figure 2.

After the interview each patient underwent a secondary interview where they were administered three clinical scales: the Positive And Negative Syndrome Scale of Schizophrenia (PANSS) [12] , Brief Psychiatric Rating Scale (BPRS)[14], and The Montgomery-Åsberg Depression Rating Scale (MADRS)[13]. PANSS assesses the following three groups of symptoms: *positive symptoms* - an excess and distortion of normal functions (e.g. hallucination and delusions); *negative symptoms* - a diminution or loss of normal functions (e.g. blunted affect and emotional withdrawal); *general symptoms* - such as depression and anxiety. BPRS is a clinical scale used to measure psychiatric symptoms such as depression, anxiety, hallucinations and unusual behavior. MADRS is a ten item diagnostic questionnaire which psychiatrists use to measure the severity of depressive episodes in patients with mood disorders.

Interviews occurred in a specialized interview room that was set up with high quality webcams recording 1280×960 pixels at 30 frames per second and head-mounted microphones to record the behavior both of the patient and the clinician. Example images captured by the recording setup can be seen in Figure 1. Each of the interviews was manually transcribed at utterance level, including both the patient's and the clinician's speech. Transcripts also included the timing of each utterance (beginning and end) and annotations of the exact question asked out of possible 13 during the semi-structured interview (for a full question list see Figure 2). Note that not all of the patients were asked the exact same questions and that the question ordering was not always identical.

METHODOLOGY AND EXPERIMENTS

In our work we are interested in studying facial behavior indicators and their predictive power for various clinical scales. We do this by building predictive models using various different indicators and analyzing their predictive power. We are especially interested in *context* - looking at the behaviors of a patient when they respond to specific questions asked by a clinician or when they are left in the room on their own. This is inspired by similar models that look at behaviors at different times in a structured interview to lead to better predictions [8]. In our work we concentrated on facial expressions such as brow raises, frowns, smiles etc. In the following section we will describe our method for extracting context specific behavior and using them to predict clinical scales.

Question segmentation Using the transcriptions of the interviews, we segmented the frames of each session based on the start and end times of the patient's utterances of answers to the specific questions that were asked. We also segmented the time the patient spent on their own.

Features To analyze the facial expressions of the patient in the clinical interviews we used the OpenFace toolbox [5]. It contains an implementation of a state-of-the-art facial Action Unit (AU) recognition system [4]. Facial Action Units are based on the Facial Action Coding System (FACS), which identifies visually discernible facial muscle actions and is a common way to describe facial expressions [9]. We used a subset of 20 most reliable AUs recognised by OpenFace. We used OpenFace to extract the AUs from the whole recording session of a patient (the time spent alone in the interview room and during the interview with the clinician). This provided us with AU presence and intensity scores at each video frame of the recording.

AU features were computed independently for each answer segment by computing the mean and standard deviations of intensities of each Action Unit (this lead to 40 a dimensional input feature vector per question). When a certain question was not asked of a participant we used the data from the whole session to fill in the missing values.

Approach For predicting clinical scales we trained a linear kernel Support Vector Regressor (SVR) for each question/context and the whole session. This led to 15 predictors for each clinical scale (alone time, 13 questions, and the whole session).

For determining the C parameter for the SVR models we used leave-one-out validation on the training data (maximizing the Pearson Correlation Coefficient). For testing we used leave-one-subject-out testing, i.e. leaving 17 participants for train-

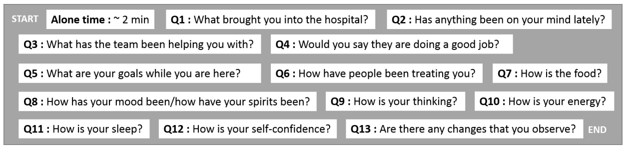


Figure 2: Typical flow of the semi-structured clinical interview. Note that not all questions were asked of all patients.

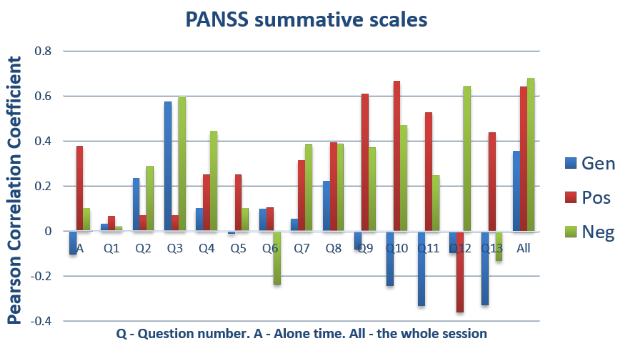


Figure 3: Discriminative power of behaviors at different contexts for the PANSS summative scales. Observe how different questions have different predictive powers for different scales and how some questions contribute very little (such as Q1, Q5 and Q6).

ing/validation in each fold. All of our experiments were done in a person independent manner.

RESULTS AND DISCUSSION

We compared the predictive power of different types of behavior indicators: whole session, question level (context), and alone time. The results are shown in Figures 3 and 4. The results clearly indicate the different predictive powers of different contexts within the semi-structured interview when considering the cumulative clinical scales.

It is interesting to note that different questions are the most predictive of each of the scales. For example the negative symptoms as identified by PANSS scale (blunted affect, social withdrawal etc.) are best predicted by the question about self-confidence, while the positive symptoms (grandiosity, hyperactivity, delusions etc.) by the question about the patient's energy.

We would also like to point out some observations of individual AU contributions to the prediction of the clinical scales. We visualized a subset of AUs and their correlations with the clinical scales in Figure 5 and Figure 6

Our findings of differences in AU5 (eye widening, see Figure 6b) for patients with more severe depression symptoms is in line with work from Alghowinem et al. [2], who found that the average distance between the eyelids was significantly smaller in depressed subjects. Another of our findings that is in line with previous research can be seen in Figure 6a. Stratou et al. [15] also found correspondences between symptoms of depression and the activation of AU4 (brow lowerer).

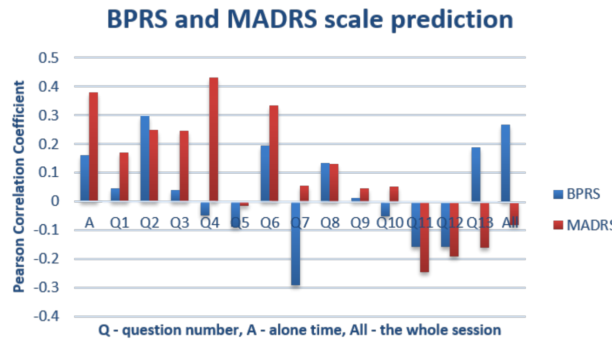


Figure 4: Discriminative power of behaviors at different contexts for the BPRS and MADRS summative scales. Observe how time on their own and response to Q4 (question about the team) are indicative of MADRS depression scale.

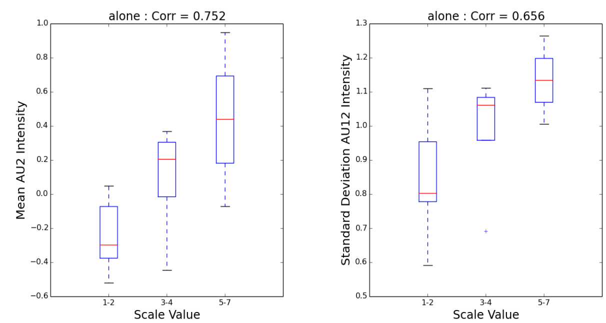
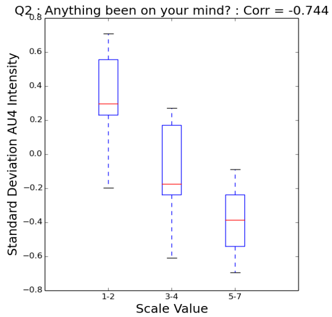


Figure 5: Behavior of subjects with unusual thought content, defined as unusual, odd, strange, or bizarre thought content in BPRS. Note that subjects with more severe symptoms raise their brows more (AU2) and have a greater variability in smiling behavior (AU12) when they are alone in the interview room.

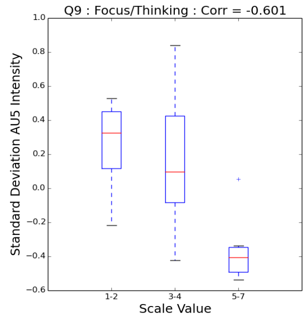
An interesting observation is that AU12 is negatively correlated with the PANSS Negative summative scale (Figure 6d). AU12 is the pulling of lip corners which is most associated with smiling behavior. The negative symptoms describe a lack or normal function, such as blunted affect and social withdrawal. Also note that Q12 is the most predictive for negative symptom scales.

Figure 6c shows that the mean of the intensities of AU2 (outer brow raiser) it positively correlated with delusions. This could be because subjects with PANSS positive symptoms (delusions, hyperactivity, hallucinations, etc.) are more expressive.

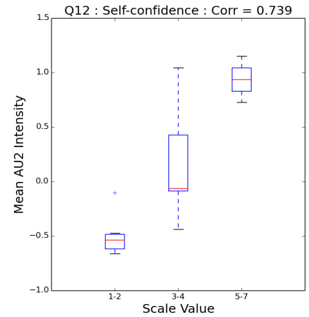
Finally, we note that patients with unusual thought content (a scale in BPRS defined as unusual, odd, strange, or bizarre thought content) are expressive even when not interacting with a clinician and show more brow flashes (AU2) and greater variability in smiles (see Figure 5). Interestingly,



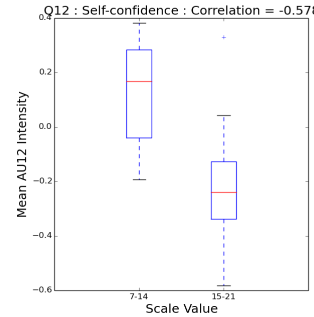
(a) BPRS Depression



(b) PANSS G6 Depression



(c) PANSS P1 Delusions



(d) PANSS Negative Cumulative Scale

Figure 6: Behavior indicators of a subset of analyzed clinical scales

the same nonverbal behaviors were not indicative of unusual thought content in other parts of the interview. These findings indicate the importance of looking at the context of behavior.

CONCLUSIONS

Our experiments show that automatically detected facial Action Units can be used to assess a number of psychosis symptoms from three clinical scales - PANSS, BPRS and MADRS. We demonstrate the importance of analysing the behaviors at question level (with context) when identifying nonverbal behavior indicators.

In the future we will explore fusion mechanisms for integrating reactions to particular questions into one predictive model. We also plan on incorporating language and acoustic features as part of analysis. In addition, we plan to use the video recordings of the clinician to gain a better insight into the nature of the dyadic interaction.

REFERENCES

1. Alghowinem, S., Göcke, R., Cohn, J. F., Wagner, M., Parker, G., and Breakspear, M. Cross-Cultural Detection of Depression from Nonverbal Behaviour. *FG* (2015).
2. Alghowinem, S., Göcke, R., Wagner, M., Parker, G., and Breakspear, M. Eye movement analysis for depression detection. *ICIP)* (2013), 4220–4224.
3. Ambady, N., and Rosenthal, R. Thin Slices of Expressive behavior as Predictors of Interpersonal Consequences : a Meta-Analysis. *Psychological Bulletin 111*, 2 (1992), 256–274.
4. Baltrušaitis, T., Mahmoud, M., and Robinson, P. Cross-dataset learning and person-specic normalisation for automatic Action Unit detection. In *FERA Challenge, in conjunction with FG* (2015).
5. Baltrušaitis, T., Robinson, P., and Morency, L.-P. OpenFace: an open source facial behavior analysis toolkit. In *IEEE WACV* (2016).
6. Bedi, G., Carrillo, F., Cecchi, G. A., Slezak, D. F., Sigman, M., Mota, N. B., Ribeiro, S., Javitt, D. C., Copelli, M., and Corcoran, C. M. Automated analysis of free speech predicts psychosis onset in high-risk youths. *NPJ Schizophrenia 1*, May (2015).
7. Cohn, J. F., Kruez, T. S., Matthews, I., Yang, Y., Nguyen, M. H., Padilla, M. T., Zhou, F., and De la Torre, F. Detecting depression from facial actions and vocal prosody. In *ACII*, IEEE (sep 2009), 1–7.
8. Devault, D., Artstein, R., Benn, G., Dey, T., Fast, E., Gainer, A., Georgila, K., Gratch, J., Hartholt, A., Lhommet, M., Lucas, G., Marsella, S., Morbini, F., Nazarian, A., Scherer, S., Stratou, G., Suri, A., Traum, D., Wood, R., Xu, Y., Rizzo, A., and Morency, L.-p. SimSensei Kiosk : A Virtual Human Interviewer for Healthcare Decision Support. *2014 International Conference on Autonomous Agents and Multi-Agent Systems. International Foundation for Autonomous Agents and Multiagent Systems*, 1 (2014), 1061–1068.
9. Ekman, P., and Friesen, W. V. *Manual for the Facial Action Coding System*. Palo Alto: Consulting Psychologists Press, 1977.
10. Girard, J. M., Cohn, J. F., Mahoor, M. H., Mavadati, S. M., Hammal, Z., and Rosenwald, D. P. Nonverbal Social Withdrawal in Depression : Evidence from manual and automatic analyses. *Image and Vision Computing 32*, 10 (2013), 641–647.
11. Hall, J. A., Harrigan, J. A., and Rosenthal, R. Nonverbal behavior in clinicianpatient interaction. *Applied and Preventive Psychology 4*, 1 (1995), 21–37.
12. Kay, S. R., Fiszbein, A., and Opler, L. A. The positive and negative syndrome scale (panss) for schizophrenia. *Schizophrenia bulletin 13*, 2 (1987), 261.
13. Montgomery, S. A., and Asberg, M. A new depression scale designed to be sensitive to change. *The British journal of psychiatry 134*, 4 (1979), 382–389.
14. Overall, J. E., and Gorham, D. R. The brief psychiatric rating scale. *Psychological reports 10*, 3 (1962), 799–812.
15. Stratou, G., Scherer, S., Gratch, J., and Morency, L.-P. Automatic nonverbal behavior indicators of depression and PTSD: the effect of gender. *Journal on Multimodal User Interfaces 9* (2015), 17–29.