

Multimodal Behavioral Markers Exploring Suicidal Intent in Social Media Videos

Ankit Shah*

aps1@andrew.cmu.edu
Carnegie Mellon University

Mahmoud Al Ismail*

mahmoudi@andrew.cmu.edu
Carnegie Mellon University

Vaibhav*

vvaibhav@andrew.cmu.edu
Carnegie Mellon University

Jeffrey Girard

jgirard2@andrew.cmu.edu
Carnegie Mellon University

Vasu Sharma*

vasus@andrew.cmu.edu
Carnegie Mellon University

Louis Philippe Morency

morency@cs.cmu.edu
Carnegie Mellon University

ABSTRACT

Suicide is one of the leading causes of death in the modern world. In this digital age, individuals are increasingly using social media to express themselves and often use these platforms to express suicidal intent. Various studies have inspected suicidal intent behavioral markers in controlled environments but it is still unexplored if such markers will generalize to suicidal intent expressed on social media. In this work, we set out to study multimodal behavioral markers related to suicidal intent when expressed on social media videos. We explore verbal, acoustic and visual behavioral markers in the context of identifying individuals at higher risk of suicidal attempt. Our analysis reveals that frequent silences, slouched shoulders, rapid hand movements and profanity are predominant multimodal behavioral markers indicative of suicidal intent¹.

KEYWORDS

Multimodal behavior markers, suicidal ideation, social media, healthcare analysis, multimodal predictive modelling

ACM Reference Format:

Ankit Shah, Vaibhav, Vasu Sharma, Mahmoud Al Ismail, Jeffrey Girard, and Louis Philippe Morency. 2019. Multimodal Behavioral Markers Exploring Suicidal Intent in Social Media Videos. In *2019 International Conference on Multimodal Interaction (ICMI '19), October 14–18, 2019, Suzhou, China*. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3340555.3353718>

*First four authors contributed equally to this research.

¹Code: https://github.com/ankitshah009/icmi_19_suicidal_intent_detection

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ICMI '19, October 14–18, 2019, Suzhou, China

© 2019 Association for Computing Machinery.

ACM ISBN 978-1-4503-6860-5/19/10...\$15.00

<https://doi.org/10.1145/3340555.3353718>

1 INTRODUCTION

Suicide is ranked as the tenth leading cause of death across the world [30]. Prompt treatment and attention towards individuals expressing suicidal intent can mitigate the danger posed to their lives. Nowadays, individuals with suicidal intent often receive the needed support from their social circles, psychiatric help or through social media platforms [16]. Social alienation has been noted in the literature to be a trait frequently observed in individuals expressing suicidal intent and such individuals often use social media platforms to stream their stories and express their suicidal intent [24].

Previous research on identifying behavioral markers of suicidal intent [11, 18, 22] typically focuses on controlled environments, using, for example, dyadic interviews between a health care provider and the participant. This motivates us to study suicide relevant behavioral markers expressed in the videos on social media platforms. We explore the use case of screening videos with a higher risk of suicidal ideation using multimodal behavioral markers.

In this paper, we present the first study of behavioral markers of expressed suicidal intent on social media videos. In order to explore this task, we create a new annotated dataset utilizing social media videos. We perform statistical hypothesis testing to understand the relative importance of behavioral markers and confirm our research hypothesis as stated in the literature review. Finally, we perform multimodal predictive modeling as a way to study the joint impact of behavioral markers.

2 RELATED WORK

Klonsky et al. [17] proposed a three-stage theory beginning from the development of suicide ideation to the subsequent progression of suicide attempts. Often individuals in any of these stages show suicidal behaviors and carry the potential risk of death by suicide. Nock et al. [20] identified several risk factors for suicidal behaviors, namely, demographic, psychiatric, psychological and biological. Additional risk factors identified by Joint et al. [6] include mental or emotional disorders, found in up to 90% of suicide victims at the time

of death. The study conducted by Hendin et al. [14], shows pre-suicidal individuals exhibit one or more of the following nine affective states: desperation, hopelessness, rage, anxiety, feelings of abandonment, loneliness, guilt, humiliation, and self-hatred.

Previous research to identify suicide relevant behaviors [1, 3, 15, 26, 28, 31] focus on uni-modal or bi-modal analysis. Scherer et al. [25, 26] used speech features like reduced vowel space and normalized amplitude quotient (NAQ) to differentiate between individuals with distress and suicidal tendency. Some studies [11, 18] used visual features such as gaze aversion and eyebrow-raising to identify differences between the suicidal and nonsuicidal groups. Pestian et al. [22] combined linguistic and acoustic features. The nature of the videos used in these studies consisted of dyadic interviews between the patients and a clinician. This setup differs from the kinds of video present on social media as addressed in our paper. To the best of our knowledge, this is the first study to present tri-modal behavioral markers related to suicidal intent with a focus on social media videos.

The widespread usage of social media has resulted in research works [10, 16, 19] to understand the relationship between suicidal ideation and social media. Dunlop et al. [10] show that online discussion forums were associated with increases in suicidal ideation while Luxton et al. [19] show that social media can pose a threat to some vulnerable groups of society. Some uni-modal studies [21, 27] have examined behavioral markers of depression in social media by leveraging the posts on Reddit and Twitter.

3 DATASET

One of the major challenges associated with social media videos is the *unconstrained environment*. Since the video discussion are not probed using specific semi-structured interviews, we find a high degree of variability in content as well as the quality of recording (e.g. different camera angles or microphone quality). Thus, we collect and annotate videos from social media and create a new dataset.

We focus our study on social media website videos from widely popular sites such as Youtube. We explore the use case of screening online videos for higher risk of suicidal intent. First, we collect a set of high-risk videos using queries with phrases associated with suicidal intent and high distress. Second, we annotate each video to be in one of the two primary categories: (a) *suicide*: individuals with suicidal intent and (b) *distress*: depressed individuals without any suicidal or self-harm tendencies. Each video is assigned to one of these categories by 2 different annotators. We observe a high inter-annotator agreement ($\alpha = 0.84, \kappa = 0.841$) for these two labels - suicide or distress. Our exploratory dataset consisted of 90 annotated where each video contains manual transcription. These videos were selected to contain only one

person talking in front of the camera. Videos were clipped to a maximum of 5-minute duration.

4 MULTIMODAL BEHAVIORAL MARKERS

We present the multimodal behavioral markers in our study and how these behavioral markers are operationalized in our experiments. Based on prior research and inspection of the dataset, we select behavioral markers frequently observed in the expression of suicidal intent on social media. We present such markers for each of the modalities and explain their representation in our analysis.

Verbal: For the verbal modality, we hypothesize that the expression of suicidal intent typically involves the use of ‘stronger’ language with frequent mentions of self-harm and strong negative emotion. To study such verbal markers, we extracted using LIWC [29] software performed on the video transcripts. This gives us 16 features belonging to categories like linguistic, social, affective and psychological processes. Each of these is a continuous feature and represents the percentage usage for a particular category based on proprietary dictionaries. Refer [29] for details on the various features. For the purpose of predictive modelling, we also compute the mean pooled BERT [9] features ² for the all the sentences as additional 768-dimension feature for predictive modelling.

Acoustic: Impulsivity and emotional instability have often been cited as key behavioral markers observed in individuals with suicidal intent [23]. These behavior markers often manifest themselves as high pitch variations and longer and more frequent pauses during speech. Hence we hypothesize that such acoustic features are important for studying the expression of suicidal intent. To study these acoustic features, we utilize OpenSmile [13] to extract the speech features. We extract F0, F1-F3, F1-F3 bandwidths, loudness, spectral flux and average duration of voiced and unvoiced segments. Each feature has been averaged over the length of the video resulting in scalar features. Variance in pitch has been captured as the standard deviation in the F0 values in voiced segments.

Visual: Impulsivity and emotional instability in individuals also cause visual manifestations in their behaviors [23]. For example, rapid hand movements and frequent pose changes are representative of an agitated mental state [8]. The frequent shift in eye gaze and eye gaze aversion [12] are also important behavioral markers since they represent social alienation and avoidance which have been shown to be psychological features often observed in individuals with suicidal intent [24]. Elongated periods of negative sentiment based facial expressions and frequent changes in facial expression are other visual behavior markers which could be important indicators of suicidal intent. We use OpenPose [4] to extract the pose features. OpenPose extracts full body 25

²Sentence vector corresponding to [CLS] token

keypoints per frame as shown in figure 3. The frames are extracted at 5fps and each of the pose key-point coordinates are normalized. We then compute the average displacement of each keypoint over the video and use it to represent that keypoint. The OpenFace toolkit [2] is used to extract Facial action units and eye gaze information. Previous research showed that Action Units 1, 4, 10, 15 and 17 as important markers for negative sentiment expressions like sadness, anger and disgust [32]. We use mean and standard deviation of the frame level probabilities for each action unit computed over the entire video as features. Additionally, we extract gaze features as pupil location and gaze angle. The average and standard deviation of these gaze features are computed over the video which help us study gaze aversion and frequently shifting gaze.

5 STATISTICAL HYPOTHESIS TESTING

In this section, we present our statistical hypothesis testing method designed to study the effect of behavior markers in isolation. We use a standardized mean difference test to understand how the hypothesized behavioral markers are used during the expression of suicidal intent. The standardized mean difference is computed as the difference in the values of different features for the primary target classes, i.e. suicidal intent expression and psychological distress, without expression of suicidal intent. We use non-parametric boot-strap with 2000 iterations and a 95% confidence interval for this analysis. We present the results of standardized mean difference analysis for all the three modalities: verbal, acoustic and visual.

Verbal: We hypothesized that verbal content of the language used by individuals with suicidal intent could be used to identify the individuals with a higher risk of suicidal ideation. To validate this hypothesis, we study the verbal content in speech used by these individuals. In particular, we utilize pronoun, verb, adj, positive emotion (posemo), negative emotion (negemo), anger, anxiety, sad, social, informal, swear and death related LIWC word clusters for this analysis. Based on the standardized mean difference test with the target variable distress/suicide, we found that only anger, negemo and death related word clusters were statistically different among these groups, which are all either affective or psychological features. We hypothesized swear to be statistically significant as well but that wasn't observed in Figure 1.

Acoustic: Previous work found features like formant frequencies and formant bandwidths to be significant for suicidal speech [7]. Our analysis of social media videos does not show statistically significant features when considered in isolation (see Figure 2).

Visual: Figure 3 shows the standardized mean difference test for 6 different body parts. These results show that motion information for keypoints 1, 2, 4, 5 and 7, corresponding

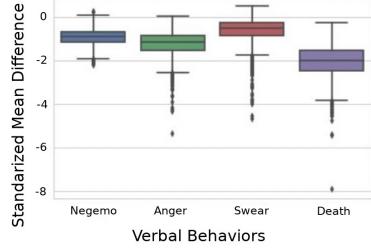


Figure 1: Results of the standard mean difference test for 'distress vs suicide' as the target variable for various linguistic features. The intervals must be entirely on one side of 0 for the difference to be statistically significant between the two groups

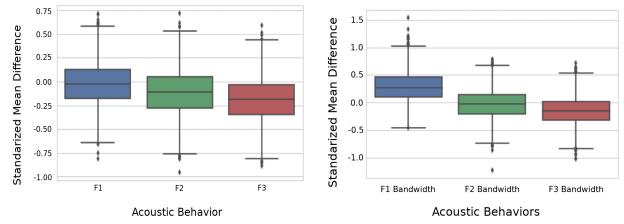


Figure 2: Box plots for the standardized mean difference test for 'distress vs. suicide' using the chosen speech features.

to shoulders, hands and head as seen from Figure 3 comes out as statistically significant. This implies that the average movement of shoulders, head and hands seems to be different between suicidal and distress videos. We believe motion in these keypoints can be indicative of slouched shoulders and agitated hand movements which could serve as important features in detecting suicidal individuals. In the second part of our analysis where we studied facial behavioral markers, we found the standardized mean difference between the target classes and gaze aversion and action units to be statistically insignificant.

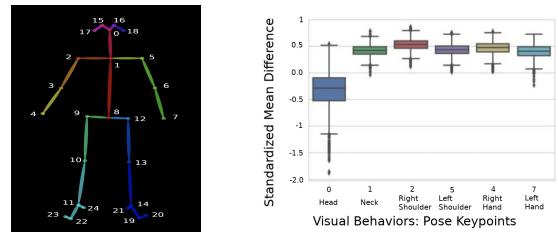


Figure 3: Left: Locations of 25 body keypoints extracted from OpenPose. Right: Box plots for the standardized mean difference test for 'distress vs. suicide' using the extracted pose features.

6 MULTIMODAL PREDICTIVE MODELLING

In this section, we present our multimodal predictive modelling method which is designed to study behavioral markers in social media videos. We perform experiments using uni-modal, bi-modal and tri-modal combinations of modalities. We present in our paper experiments with the XGBoost [5] classification model. XGBoost was used for our analysis since it is a powerful ensemble based model and performs well on classification tasks while retaining the interpretability of the model. Interpretability of XGBoost comes from the fact that it's actually an ensemble of tree based classifiers which construct an ensemble of decision trees based on information gain criterion. They have the inherent advantage of being able to compute the relative importance of the features which are used to construct the trees. Importance is calculated for a single decision tree by the amount that each attribute split point improves the performance measure, weighted by the number of observations the node is responsible for. The feature importances are then averaged across all of the decision trees within the model to generate the overall feature importances. We use these feature importances to identify the relative importance of our behavioral markers and study their prevalence in the expression of suicidal intent. We perform 5 fold cross validation using stratified splits and report area under the ROC curve (AUC).

Prediction results: Figure 4 shows the performance for different modalities. We observe that the verbal modality is most important for this classification task followed by acoustic and visual. Fusion of all modalities results in better performance. Overall, fusion of acoustic and language features provide better AUC across different classifiers which point to the complementary information contained in these modalities.

Behavior Analysis: Table 1 shows the feature importance as obtained after training the XGBoost classifier. This gives us insight into the explainability of the learned models. For the acoustic features, we observe the length of unvoiced segments to be the most important feature which signifies that **silences** and longer pauses in speech are useful features. **Standard deviation of F0** is the next most highly ranked acoustic feature; which represents larger pitch variations in individual speech. For the visual features, average displacement of right hand, neck, right shoulder, left hand and left shoulder are found to be the most important (in that order). This is consistent with our findings in the hypothesis testing where standardized mean difference shows a similar trend. Thus, **slouched shoulders** and **rapid hand movements** are significant in detecting suicidal behavior. For the language features, the LIWC features derived from the dictionaries corresponding to **death, profanity and swear**

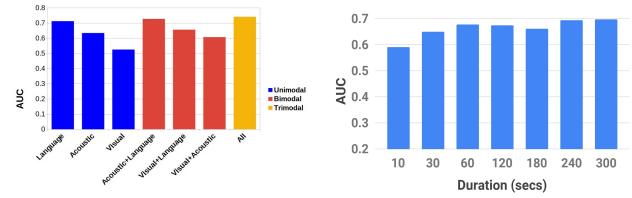


Figure 4: Left: Results for XGBoost on automatic features. Right: Variation of performance of Random Forest classifier using all automatic features with duration of video context used (in seconds)

words and **anger** related words are found to be most important. Thus, strong speech related to swearing or frequent mentions of self-harm are important in identifying individuals with suicidal intent.

Features (Importance Score)		
Language	Acoustic	Visual
death (0.189)	unvoiced segments (0.175)	left hand (0.058)
swear (0.134)	std F0 (0.097)	torso (0.050)
anger (0.107)	max F0 (0.082)	left shoulder (0.046)
fuckpast (0.093)	F2 (0.079)	right shoulder (0.044)

Table 1: Feature importance scores calculated using our predictive model to identify the most important features from each modality.

A dimension which frequently remains unexplored in previous research is the length of video needed to perform the prediction. For timely intervention, it is critical that the model uses the least amount of video length for prediction. Figure 4 shows the variation of performance of a Random Forest classifier using features from all modalities with duration of video context used. Note that the Random Forest classifier performs equally well on this task. We see the performance saturates after about 60 seconds of video context which could imply that 60 seconds is enough for prediction.

7 CONCLUSION

In this work, we study multimodal behavioral markers of suicidal intent in social media videos. We first study these markers using a statistical hypothesis testing then perform multimodal predictive modelling. We observed key features from each modality which are commonly seen in individuals expressing suicidal intent on social media videos. We found the use of death related words and profanity to be a common verbal behavioral marker. Also, we observe larger pitch variations and longer silences to be important behavioral markers in speech while frequent hand movements and slouched shoulders are important visual markers. We hope these insights could guide future work.

8 ACKNOWLEDGEMENT

We thank Keith Maki and Alexandria Vail for their inputs and suggestions for the paper.

REFERENCES

[1] Amayas Abboute, Yasser Boudjeriou, Gilles Entringer, Jérôme Azé, Sandra Bringay, and Pascal Poncelet. 2014. Mining twitter for suicide prevention. In *International Conference on Applications of Natural Language to Data Bases/Information Systems*. Springer, 250–253.

[2] Tadas Baltrušaitis, Amir Zadeh, Yao Chong Lim, and Louis-Philippe Morency. 2018. Openface 2.0: Facial behavior analysis toolkit. In *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*. IEEE, 59–66.

[3] Wassim Bouachir, Rafik Gouiaa, Bo Li, and Rita Noumeir. 2018. Intelligent video surveillance for real-time detection of suicide attempts. *Pattern Recognition Letters* 110 (2018), 1–7.

[4] Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, and Yaser Sheikh. 2018. OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields. In *arXiv preprint arXiv:1812.08008*.

[5] Tianqi Chen and Carlos Guestrin. 2016. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16)*. ACM, New York, NY, USA, 785–794. <https://doi.org/10.1145/2939672.2939785>

[6] Joint Commission et al. 2016. Detecting and treating suicide ideation in all settings. *Sentinel Event Alert* 56 (2016), 1–7.

[7] et al. Cummins, Nicholas. 2015. A review of depression and suicide risk assessment using speech analysis. *Speech Communication* 71 (2015): 10–49. (2015).

[8] N. Dael, M. Mortillaro, and K.R. Scherer. 2012. Emotion expression in body action and posture, Vol. 12(5). 1085–1101.

[9] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805* (2018).

[10] Sally M Dunlop, Eian More, and Daniel Romer. 2011. Where do youth learn about suicides on the Internet, and what influence does this have on suicidal ideation? *Journal of child psychology and psychiatry* 52, 10 (2011), 1073–1080.

[11] Naomi Eigbe, Tadas Baltrušaitis, Louis-Philippe Morency, and John Pestian. 2018. Toward Visual Behavior Markers of Suicidal Ideation. In *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*. IEEE, 530–534.

[12] Naomi Eigbe, Tadas Baltrušaitis, Louis-Philippe Morency, and John Pestian. 2018. Toward Visual Behavior Markers of Suicidal Ideation. *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)* (2018), 530–534.

[13] Florian Eyben, Martin Wöllmer, and Björn Schuller. 2010. Opensmile: the munich versatile and fast open-source audio feature extractor. In *Proceedings of the 18th ACM international conference on Multimedia*. ACM, 1459–1462.

[14] John T. Maltsberger Hendin, Herbert and Katalin Szanto. 2007. The role of intense affective states in signaling a suicide crisis. *The Journal of Nervous and Mental Disease* 195.5 (2007), 363–368.

[15] Shaoxiong Ji, Celina Ping Yu, Sai-fu Fung, Shirui Pan, and Guodong Long. 2018. Supervised Learning for Suicidal Ideation Detection in Online User Content. *Complexity* 2018 (2018).

[16] Carolina A Klein. 2012. Live deaths online: Internet suicide and lethality. *Journal of the American Academy of Psychiatry and the Law Online* 40, 4 (2012), 530–536.

[17] E David Klonsky and Alexis M May. 2015. The three-step theory (3ST): A new theory of suicide rooted in the “ideation-to-action” framework. *International Journal of Cognitive Therapy* 8, 2 (2015), 114–129.

[18] Eugene Laksana, Tadas Baltrušaitis, Louis-Philippe Morency, and John P Pestian. 2017. Investigating facial behavior indicators of suicidal ideation. In *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*. IEEE, 770–777.

[19] David D Luxton, Jennifer D June, and Jonathan M Fairall. 2012. Social media and suicide: a public health perspective. *American journal of public health* 102, S2 (2012), S195–S200.

[20] Matthew K Nock, Guilherme Borges, Evelyn J Bromet, Christine B Cha, Ronald C Kessler, and Sing Lee. 2008. Suicide and suicidal behavior. *Epidemiologic reviews* 30, 1 (2008), 133–154.

[21] Ted Pedersen. 2015. Screening Twitter users for depression and PTSD with lexical decision lists. In *Proceedings of the 2nd workshop on computational linguistics and clinical psychology: from linguistic signal to clinical reality*. 46–53.

[22] John P Pestian, Michael Sorter, Brian Connolly, Kevin Bretonnel Cohen, Cheryl McCullumsmith, Jeffry T Gee, Louis-Philippe Morency, Stefan Scherer, Lesley Rohlf, and STM Research Group. 2017. A machine learning approach to identifying the thought markers of suicidal subjects: a prospective multicenter trial. *Suicide and Life-Threatening Behavior* 47, 1 (2017), 112–121.

[23] Evyn M Peters, Lloyd D Balbuena, Steven Marwaha, Marilyn Baetz, and Rudy Cecil Bowen. 2016. Mood instability and impulsivity as trait predictors of suicidal thoughts. *Psychology and psychotherapy* 89 4 (2016), 435–444.

[24] Jessica D Ribeiro and Thomas E Joiner. 2009. The interpersonal-psychological theory of suicidal behavior: Current status and future directions. *Journal of clinical psychology* 65, 12 (2009), 1291–1299.

[25] Stefan Scherer, Louis-Philippe Morency, Jonathan Gratch, and John Pestian. 2015. Reduced vowel space is a robust indicator of psychological distress: A cross-corpus analysis. In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 4789–4793.

[26] Stefan Scherer, John Pestian, and Louis-Philippe Morency. 2013. Investigating the speech characteristics of suicidal adolescents. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 709–713.

[27] Judy Hanwen Shen and Frank Rudzicz. 2017. Detecting Anxiety through Reddit. In *Proceedings of the Fourth Workshop on Computational Linguistics and Clinical Psychology – From Linguistic Signal to Clinical Reality*. Association for Computational Linguistics, Vancouver, BC, 58–65. <https://doi.org/10.18653/v1/W17-3107>

[28] Han-Chin Shing, Suraj Nair, Ayah Zirikly, Meir Friedenberg, Hal Daumé III, and Philip Resnik. 2018. Expert, Crowd sourced, and Machine Assessment of Suicide Risk via Online Postings. In *Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic*. 25–36.

[29] Yla R Tausczik and James W Pennebaker. 2010. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of language and social psychology* 29, 1 (2010), 24–54.

[30] Peeter Värnik. 2012. Suicide in the world. *International journal of environmental research and public health* 9, 3 (2012), 760–71.

[31] Andrew Yates, Arman Cohan, and Nazli Goharian. 2017. Depression and self-harm risk assessment in online forums. *arXiv preprint arXiv:1709.01848* (2017).

[32] Yongdong Zhang, Weiming Dong, Bao-Gang Hu, and Qiang Ji. 2018. Classifier Learning with Prior Probabilities for Facial Action Unit Recognition. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2018), 5108–5116.