Predicting Depression Tendency based on Image, Text and Behavior Data from Instagram

Yu Ching Huang[1], Chieh-Feng Chiang[2] and Arbee L. P. Chen[3]
[1]*Department of Computer Science, National Tsing Hua University, Hsinchu, Taiwan*
[2]*Center of Technology in Education, China Medical University, Taichung, Taiwan*
[3]*Department of Computer Science and Information Engineering, Asia University, Taichung, Taiwan*

Keywords: Depression Detection, Social Media, Deep Learning.

Abstract: Depression is common but serious mental disorder. It is classified as a mood disorder, which means that it is characterized by negative thoughts and emotions. With the development of Internet technology, more and more people post their life story and express their emotion on social media. Social media can provide a way to characterize and predict depression. It has been widely utilized by researchers to study mental health issues. However, most of the existing studies focus on textual data from social media. Few studies consider both text and image data. In this study, we aim to predict one's depression tendency by analyzing image, text and behavior of his/her postings on Instagram. An effective mechanism is first employed to collect depressive and non-depressive user accounts. Next, three sets of features are extracted from image, text and behavior data to build the predictive deep learning model. We examine the potential for leveraging social media postings in understanding depression. Our experiment results demonstrate that the proposed model recognizes users who have depression tendency with an F-1 score of 82.3%. We are currently developing a tool based on this study for screening and detecting depression in an early stage.

1 INTRODUCTION

Depression is an important mental health care issue and is leading causes of disability. The World Health Organization (WHO) estimated that nearly 300 million people worldwide suffer from depression (Organization et al., 2017). By the year 2020, depression will be the second leading cause of death after heart disease. However, without observable diagnostic criteria, the sign of depression often goes unnoticed. Early detection of depression is essential to provide appropriate interventions for preventing fatal situations. With the rapid development of communication and network technologies, people are increasingly using social media platforms, such as Facebook and Instagram, to share thoughts and emotions with friends. As a result, social media contains a great amount of valuable information. The behavior, language used and photo posted on social media may indicate feelings of worthlessness, helplessness and self-hatred that characterize depression. Social media resources have been widely utilized to study mental health issue, where current studies relied more on analyzing text data (De Choudhury et al., 2013a; De Choudhury et al., 2014; De Choudhury et al., 2013b) rather than on analyzing visual ones, such as images of Instagram and Snapchat. Since the launch of Instagram in 2010, the photo-and-video based social media has rapidly increased its number of users to over 600 million. Moreover, psychologists have noted that imagery can be an effective medium for expressing negative emotions (Andalibi et al., 2017; De Choudhury and De, 2014; Manikonda and De Choudhury, 2017). Instagram is chosen as the primary platform in this study because of its extremely large image dataset.

In this paper, we aim to predict one's depression tendency by analyzing both text and image on Instagram. Additionally, we propose a more robust prediction model using deep learning techniques.

In short, this work employs an effective mechanism to collect depressive and non-depressive user accounts. Three sets of features are extracted including text-related features, image-related features and other behavior features. For inference, a deep learning classifier is trained to predict whether a user has a depression tendency. Note that our goal is not to offer a diagnosis but rather to make a prediction on which users are likely suffering from depression.

Huang, Y., Chiang, C. and Chen, A.
Predicting Depression Tendency based on Image, Text and Behavior Data from Instagram.
DOI: 10.5220/0007833600320040
In *Proceedings of the 8th International Conference on Data Science, Technology and Applications (DATA 2019)*, pages 32-40
ISBN: 978-989-758-377-3

Our main contributions in this paper are as follows:

- We employ an efficient data collection mechanism not requiring users to do any screening test. It is more efficient than previous research of using the CES-D (Center for Epidemiological Studies Depression) questionnaire, which should take around 10 minutes to complete and 5 minutes to score.
- We introduce a deep learning model combining text, image and behavior as features to predict a user's depression tendency. This can be the first approach on conducting a research using both text and image Instagram posting data for predicting depression.

The remainder of this paper is organized as follows: In Section 2, we provide a brief overview of the background of our approach. Specifically, we first briefly review recent mental health research using social media data; subsequently, we review CNN (Convolutional Neural Networks) *transfer learning* (Pan and Yang, 2010) for feature extraction. In Section 3, we describe the goal of our predicting depression tendency of users. In Section 4, our approach is described, including how data is collected and preprocessed, and how features are extracted. In Section 5, we show the performance of our model. And finally, we summarize this paper and propose future work in Section 6.

2 RELATED WORK

Depression is the most common mental illness in the world, which has been found to have the positive correlation with the risk of early death. Nowadays, data on social media has been widely utilized for studies on mental health issues. An early work (Park et al., 2012) showed that people post about their depression and even their treatment on Twitter. The advent of social media presents a new opportunity for early detection and intervention in mental disorder. We propose a predictive model utilizing a popular image-based social media, Instagram, as a source for depression detection. Instagram members currently contribute almost 100 million new postings per day (Instagram, 2016), which makes Instagram one of the most popular social networks worldwide.

In the following, we first review recent research on mental health using social media data that are relevant to this paper; subsequently, we review feature extraction techniques through CNN.

Mental health issue has been widely studied, including major depressive disorder (Chen et al., 2018a), post-traumatic stress disorder (De Choudhury et al., 2014), and bipolar disorder. Identifying depression from social media, the majority of previous studies are mostly based on the analysis of textual contents from publicly available data. M. De Choudhury et al. (De Choudhury et al., 2013b) examined Twitter postings of individuals suffering from major depressive disorder to build statistical models that predict the risk for depression. As a similar study on early detection, Chen et al. (Chen et al., 2018b) identify users who suffer from depression by measuring eight basic emotions as features from Twitter postings.

Besides, emoticons and images have also been utilized for detecting positive and negative sentiments. Kang et al. (Kang et al., 2016) introduce a multimodel method for analyzing tweets on Twitter, in order to detect users with depressive moods. Extracting sentiments from images and texts allows a continuous monitoring on user's moods. Andalibi et al. (Andalibi et al., 2015) fetched users' photos on Instagram and found that patients with depression posted more photos than others, and those photos tended to be darker, grayer with low saturation. Andalibi et al., Kang et al., and Reece et al. (Andalibi et al., 2015; Kang et al., 2016; Reece and Danforth, 2017) analyzed images from social media using handcrafted features. In our work, features are extracted automatically using deep learning techniques instead of being identified manually by domain experts. Moreover, we utilize not only images from Instagram but also text data to build our model, which is different from the Andalibi et al., Kang et al., and Reece et al. (Andalibi et al., 2015; Kang et al., 2016; Reece and Danforth, 2017) focusing on analyzing images only.

For extracting features from images, recently, CNN has become a very popular tool that automatically extracts and classifies features. Training a new network requires a sufficiently large number of labeled images. However, the data for recognizing users of depression on Instagram is not big enough. That leads to another approach: we applied transfer learning (Pan and Yang, 2010) on CNN. Transfer learning which aims to learn from related tasks has attracted more and more attention recently. Research of this field begins with Thrun et al. (Thrun, 1996) that discussed about the role of previously learned knowledge for generalization, particularly when the training data is scarce. Yosinski et al. (Yosinski et al., 2014) studied extensively the transferability of features pre-learned from *ImageNet* dataset by employing different fine-tuning strategies on other datasets, and demonstrated that the transferability of features decreases when the distance between the base task and target task increases. Oquab et al. (Oquab et al.,

2014) showed how image representations learned with CNN on large-scale annotated datasets can be efficiently transferred to other visual recognition tasks, and how a method is designed to reuse layers trained on the ImageNet dataset to compute the mid-level image representation for images. Despite differences in image statistics and tasks in the two datasets, the transferred representation leads to significantly improved results for classification. Our work is primarily motivated by Yosinski et al. (Yosinski et al., 2014) and Oquab et al. (Oquab et al., 2014), which comprehensively explored feature transferability of deep convolutional neural networks and learnt visual features automatically.

Therefore, with the present work we aim to (1) collect user-level data from Instagram by applying an effective data collection approach which is more efficient and realistic than traditional questionnaire survey; (2) build a multi-model for depression tendency prediction by extracting potential signals from textual and visual contents, and social behaviors through Instagram postings as features; (3) examine and demonstrate the effectiveness of these features for identifying users suffering from depression by deep learning techniques; and (4) apply a CNN-based model to automatically learn features from images.

3 TASK DESCRIPTION

The main approach of predicting depression tendency is to build a classifier recognizing people who may have depression tendency. Specifically, the task is formulated as follows. Given a set of p users $U = \{u_1, u_2 \ldots, u_p\}$, where each user is associated with n postings. There are sets of text and image features in each posting, $TEXT^{u_i} = \{text_1^{u_i}, \ldots, text_n^{u_i}\}$ and $IMAGE^{u_i} = \{image_1^{u_i}, \ldots, image_n^{u_i}\}$. Moreover, there is a set of behavior features $BEHAVIOR = \{b_1^{u_i}, \ldots, b_n^{u_i}\}$. The goal is to learn a model for predicting the depression tendency of a user, where the set of depression tendency is labeled as $D = \{D_1, D_2, \ldots, D_p\}$. In this work, D is set to $\{Yes, No\}$ as a binary label implying that each user is classified as with depression tendency or without.

4 METHOD

4.1 Data Collection

To train our model, we require information from two different types of users: users with depression tendency and users without. To achieve this purpose, we employed a data collection mechanism to efficiently collect these users. In order to identify users with depression from Instagram, we first crawled Instagram postings with hashtags related to depression (e.g., #憂鬱症(depression), #自殺(suicide), #憂鬱(depressed)). This approach is validated in the previous literature: a recent work (De Choudhury et al., 2016) suggests that the use of the word *depression* increases one's chance of using the word *suicide* in the future, which is one of the symptoms of depression. Then we collect the users whose postings have depression related hashtags to our depression-user pool. From this pool, we select the self-reported users who explicitly state, in their Instagram bio, that they suffer from a mental illness as our depressive users. In other words, we select a depressive user by checking if his/her bio contains any keywords related to depression. For non-depressive users, we crawled postings with hashtags related to the word *happy*. Users with these postings are collected in our non-depressive-user pool. From this pool, users with any word related to depression in their bio or in their postings are filtered out. According to the research (Chen et al., 2018a), we double check their postings in order to make sure there is no overlap between the two groups, namely the depressive users and non-depressive users. This selection method may be biased but we have to make sure there is no noise that may weaken the classification model. After the users have been identified, we manually double check and label them into depressive users and non-depressive users. Moreover, we do the same filtering process on the followings/followers of the users in the above two categories to widen the dataset. Each user account in this research is public account. Each posting in this dataset is public and contains images, captions, likes, comments, and hashtags, if tagged. Moreover, this dataset is only for academic use. In the future if we want to work with medical institutions, we will have to pay more attention to user's privacy.

4.2 Data Pre-processing

For data cleaning and pre-processing, users having more than 50% postings containing hyperlinks are removed. We also exclude inactive users, private accounts, business accounts and official accounts.

4.3 Feature Extraction

In this work, we focused on three main types of features, i.e. text, image and behavior to be detailed in the following.

4.3.1 Text Features

We generated posting representation using *Word2vec* (Mikolov et al., 2013) as our text features. This approach includes two major steps, i.e. post segmentation step and word representation step. The post segmentation step segments sentences in each posting into words. The Word Representation step collects words in all postings and converts them into vectors. Suppose a user u_i has the a set of n postings, where each posting contains containing $TEXT^{u_i} = \{text_1^{u_i}, \ldots, text_n^{u_i}\}$ and images $IMAGE^{u_i} = \{image_1^{u_i}, \ldots, image_n^{u_i}\}$. Let $w_j^{i,m}$ be the j-th word in the m-th posting of user u_i, and $l_{i,m}$ be the length of the m-th posting of user u_i. We represent the m-th posting using $l_{i,m}$ words, i.e. $text_m^{u_i} =< w_1^{i,m}, w_2^{i,m}, \ldots, w_{l_{i,m}}^{i,m} >$. Each word $w_j^{i,m}$ is projected to a vector $v_j^{i,m} \in \mathbb{R}^{d * l_{i,m}}$,where d is the dimension of the word vectors. The value of d dimension is chosen based on a parameter tuning experiment (see Section 5). Finally, we concatenate all word vectors in all the postings of each user u_i into a vector set with exactly L vectors, where $L = \sum_m l_{i,m}$. Each user has different posting number, if the length of the vector set is fewer than L, we fill it with zero. This vector set is our posting representation which is used as text feature for our prediction model.

In the post segmentations step, we removed hyperlinks and stop words, and extracted hashtags from the text in a posting. Since emoticons have become a true digital "language," we converted every emoticon to text. Next, we adopted the segmentation tool Jeiba (Sun, 2012), one of the most popular Chinese word segmentation tools, to segment sentences into words. However, there are words which Jeiba cannot segment correctly such as 安柏寧(Alprazolam), a kind of medicine used to treat anxiety disorders. In order to solve this problem, a customized dictionary is used along with Jeiba. We crawled the medicine list from Tcdruginfo website which provides drug information for the Chinese general public around the world.

In the word representation step, we used *Word2vec* (Mikolov et al., 2013) to learn the vector representation of words. *Word2vec* is able to find the associations between words from a great amount of text and project words with similar meanings to have similar vector representations.

We used the text in the postings not taken in the training or testing data to train our *word2vec* model, totaling 407,116 Instagram postings. Then, according to this well pre-trained model, we projected each word in the training and testing postings into a word vector representation. Finally, we aggregated the word vector representations of each posting from a user to generate a vector set with L vectors to be used as the text feature for a user.

4.3.2 Image Features

According to (Andalibi et al., 2015; Reece and Danforth, 2017), the photos posted on Instagram can show signs of depression and can be used for early screening and detection. We generate image representations using deep convolutional neural networks. In the field of image feature extraction, typical approaches extract a set of "hand-crafted" features. However, handcrafted features are not flexible enough, and it is hard to design an efficient method to generate handcrafted features. We use CNN, a successful machine learning technique for image classification, to automatically learn image representations. A disadvantage of training a new CNN is that the training stage requires a large number of labeled images. Fortunately, transfer learning can solve this problem, and we employ transfer learning to a pre-trained CNN model to extract features, as illustrated in Fig 1. The key idea of this approach is that the internal layers of CNN act as an extractor of mid-level image representations, which can be pre-trained on a large dataset (i.e. the ImageNet), and then be re-used on depressive and non-depressive image classification. First, we pre-train *VGG16* network (Simonyan and Zisserman, 2014) on the ImageNet. The ImageNet dataset (Deng et al., 2009) is a large-scale ontology of images organized according to the hierarchy of WordNet (Miller, 1995). It contains more than 14 million images with 1000 categories. Second, we transfer the pre-trained weights parameters to the target task, remove the output layer of the pre-trained model and add two fully connected layers for classification that output the probabilities of an image. Then the depressive degree is generated. We use our dataset to fine-tune the classifier. After the classifier is fine-tuned with our dataset, we then use the first fully connected layer's output in this classifier as the feature vector of an image (Oquab et al., 2014). This is an approach using CNN to learn image embedding. Suppose a user u_i has the a set of n posts of images $IMAGE^{u_i}$=$\{image_1^{u_i}, \ldots, image_n^{u_i}\}$, each 224*224 image is extracted into a 64 dimensional feature vector. We concatenate all feature vectors for each user into a vector set with L vectors. If the length of the vector set is less than L, we fill it with zero vectors. This vector set is the image features for our prediction model.

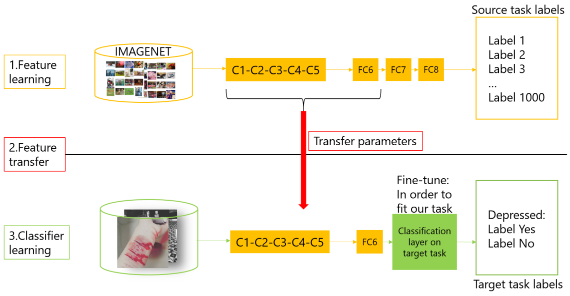

Figure 1: Transfer learning of a CNN.

4.3.3 Behavior Features

How users behave on social media can be demonstrated by features of social behaviors and writing behaviors. The social behaviors include the time the users post, the posting frequency, the number of likes their post gets, etc. The writing behaviors come from how the users write in their posts. According to the research (Ramirez-Esparza et al., 2008), depressive users tend to focus on themselves and detach from others. First, we extract nine features as social behaviors: *post time, number of posts, post on which day of a week, post on workday or weekend, posting frequency, following, followers, likes, and comments*. The post time is extracted because (Nutt et al., 2008) shows that depressive users often have sleeping disorder and seek for help online. Second, we obtain six features to identify a user's writing behaviors in each posting: *total word counts, number of first-person pronouns* (i.e. the number of the term "I" used per posting), *counts of hashtag related to depression (from the depression hashtags list), hashtag counts per posting, emoji counts, and absolutist word counts*. According to the research (Al-Mosaiwi and Johnstone, 2018), people who are depressive tend to use an absolutist word like "always," "complete," and "definitely." The ranges of the values for these features are different which may cause difficulty in building a robust and effective classifier (Grus, 2015). We apply the technique of min-max normalization to normalize the feature values into [0, 1] before the training and testing processes. With these features, we are able to distinguish between depressive users and non-depressive users. Finally, we leverage these three sets of extracted features to build the classifier for predicting depression tendency.

4.4 Prediction Model

After feature extraction, three sets of features are obtained. As shown in Figure 2, we aggregate them to represent a user. Next, we employ a CNN to construct a classifier model trained to predict depression tendency in two classes. Our CNN network is composed of five layers, three convolution layers each with 5 x 5 max-pooling and ReLU activation function, followed by two fully connected layers. The last fully connected layer uses Softmax function. Softmax function ensures that the activation of each output unit sums to 1, and therefore we can take the output as a set of conditional probabilities. These probabilities are the basis for the predicted labels. When the probability is larger than 0.5, the label is set to 1.

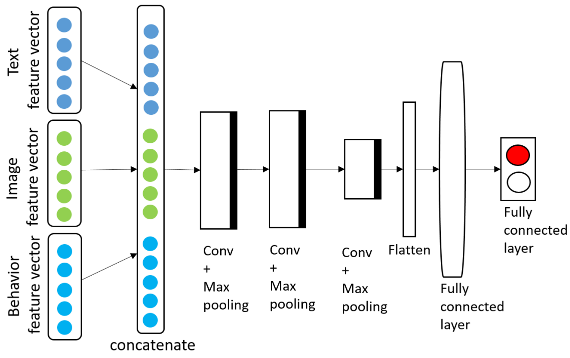

Figure 2: Prediction Model's Architecture.

5 EXPERIMENTS

In this section, we introduce the dataset collected from Instagram, the feature normalization process, and the details of parameter tuning. Then we show the experimental results of the proposed transfer learning method along with the automatic feature extraction process on the images. Finally, the performance of our model is presented.

5.1 Dataset

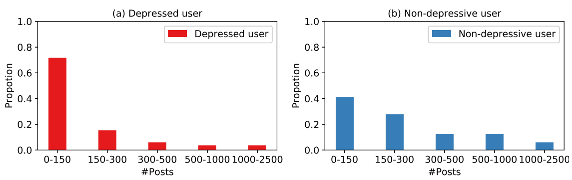

Figure 3: Distribution of the numbers of postings for the two user groups.

We collected 512 users from Instagram through our efficient data collection process, which include 117 users who have depression tendency and 395 non-depressive users. We collected all of their postings on Instagram, where each posting contains information such as images, captions, likes, comments, and hashtags, if tagged. The distribution of the numbers of postings for the two user groups is shown in Fig 3 where X-axis shows the number of postings and Y-axis shows the corresponding propotion of the users. We observe that most users in both groups have the number of postings between 0-150 and fewer users have over 1000 posts. Most depressed users have between 0-150 postings whereas most non-depressive users have between 0-300 postings. Note that users having less than 50 postings are considered inactive users and excluded in our dataset.

5.2 Normalization Process

The range difference of the behavior features may cause difficulty in building a robust and effective classifier (Grus, 2015). Therefore, we apply min-max normalization to normalize all feature values into [0, 1] before the training and testing processes. The min-max normalization is executed as follows:

$$MinMaxNormaliztion(x^b) = \frac{x_i^b - min(x^b)}{max(x^b) - min(x^b)}$$

where x_i^b is the i-th data item in the b-th feature,e.g.,the i-th data item in the behavior feature "comments."

5.3 Parameter Tuning

When building a machine learning model, we define our model architecture by exploring a range of possibilities. The tuning of the hyper-parameters of the deep neural network model depends on the dataset used. Our model is trained with backpropagation using Adam optimizer. We adopt cross-entropy as our loss function. In our experiments, the word embedding is trained on 407,116 Instagram postings. According to our experiments, we choose the word vector dimension as 400 and the length of the vector set L as 1000.

5.4 Evaluation of Image Feature Extraction Model

In this subsection, we evaluate the image feature extraction model. In order to extract features that are useful for identifying depressed and non-depressive people, we pre-trained a model on ImageNet then fine-tuned this model by our dataset. The dataset used here contains 1155 depressive images and 1156 non-depressive images, including images crawled from Instagram with hashtags related to depression and images from a well-known dataset Emotion6 (Panda et al., 2018).

Our model classified images properly. As shown in Fig 4, this image has a hashtag #憂鬱症(depression), and our model predicts this image as depressive with probability 0.903. Our model also predicts the image shown in Fig 5 as depressive with probability 0.01.

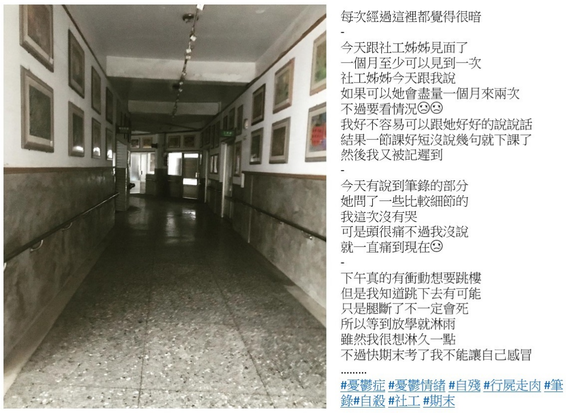

Figure 4: Depressive photo.

Figure 5: Non depressive photo.

5.5 Model and Feature Analysis

In this subsection we discuss the performance of the *Word2vec* model and the three sets of features.

Word2vec Model Evaluation. We evaluate our word embedding matrix built from a pre-trained *Word2vec* model which affects the overall performance of our model. As shown in Table 1(a) and Table 1(b), we respectively calculate the similarity of two words and find the top closest words for a given word to see whether the *Word2vec* model has been properly trained. We can see the words "depressed" and "happy" are not similar, whereas the words "sorrow" and "sad" are similar. This result indicates that our model is well built.

Feature Analysis. We built a model using the combination of the three sets of features to figure out the importance of various features, and the results are shown in Table 2. The precision and recall tend to be steady after running the model ten times. The pre-

Table 1: Word2Vec Model Evaluation.

(a) Similarity.

Word A, Word B	Similarity
憂鬱(depressed), 開心(happy)	0.17
悲傷(sorrow), 難過(sad)	0.71

(b) Top Five Similar Words.

Word	Top five similar words
憂鬱(depressed)	煩悶(unhappy) 焦慮(anxious) 鬱悶(depressive) 煩躁(irritable) 沈重(down)
開心(happy)	高興(glad) 盡興(enjoy) 興奮(exited) 歡喜(joy) 逗(funny)

Table 2: Feature analysis.

Model	Text	Behavior	Image	Text+ Behavior	Text+ Image	Image+ Behavior	Text+Image+ Behavior
Precision	0.770	0.710	0.811	0.793	0.880	0.823	0.888
Recall	0.595	0.528	0.741	0.604	0.748	0.744	0.767
F1-score	0.671	0.605	0.774	0.685	0.808	0.781	0.823

cisions using the text features only, the behavior features only, and the image features only are 77%, 71% and 81.1%, respectively. From here we can see that adopting transfer learning with VGG16 as an image feature extractor, the extracted features can effectively identify depressive photos. Moreover, using both image and text features makes the model more robust than using the pair of image and behavior features or the pair of behavior and text features. We found that the user's behavior features on Instagram contribute less significantly on distinguishing between a depressive and non-depressive user, but still improve the performance of the whole model.

5.6 Performance of our Model

In this subsection, we present the performance of our model where each user is classified into two labels,i.e. having depression tendency or not. The dataset contains 117 users with depression tendency and 395 users without depression tendency, resulting in the total number of 512 users. We use 80% of the dataset as training data, and the remaining 20% as validation data. We use precision, recall and F1-score to evaluate the model. The precision of our model is 88.8% and the recall is 73.5%, implying that our model can identify both depression and non-depression classes well. We train our model with batch size 32 through 20 epochs, and it is evident that the classifier has good performance according to the receiver-operator characteristic (ROC) curve, where the area under the ROC curve (AUC) is larger than 0.5. This means our model can distinguish depressive users and non-depressive users well. We compare our method with other studies on depression detection in Fig 6. We evaluate each method with precision and recall. Compared with Mitchell et al. (Mitchell et al., 2009), our model performs better although we have less depressive users. We also have better precision and recall than Reece et al. (Reece and Danforth, 2017), where handcrafted features were used from Instagram photos. This indicates that the approach of automatically learning features from images and building a deep learning prediction model not only outperforms the method with handcrafted features but also reduces the manpower cost. Compared with the work of, Wu et al. (Wu et al., 2018) using posting, behavior, and living environment data from Facebook for depression detection, this work takes image data as the extra information. The slight improvement on performance may be because of the smaller number of users than that of the work by Wu et al. (Wu et al., 2018). The positive experiment result of our model demonstrates that text, image and behavior on Instagram provide useful signals that can be utilized to classify and predict whether a user has depression tendency. Among these factors, image is the most important factor for predicting depression tendency. This indicates that the markers of depression are observable in Instagram user's images. Images provide an effective way of express-

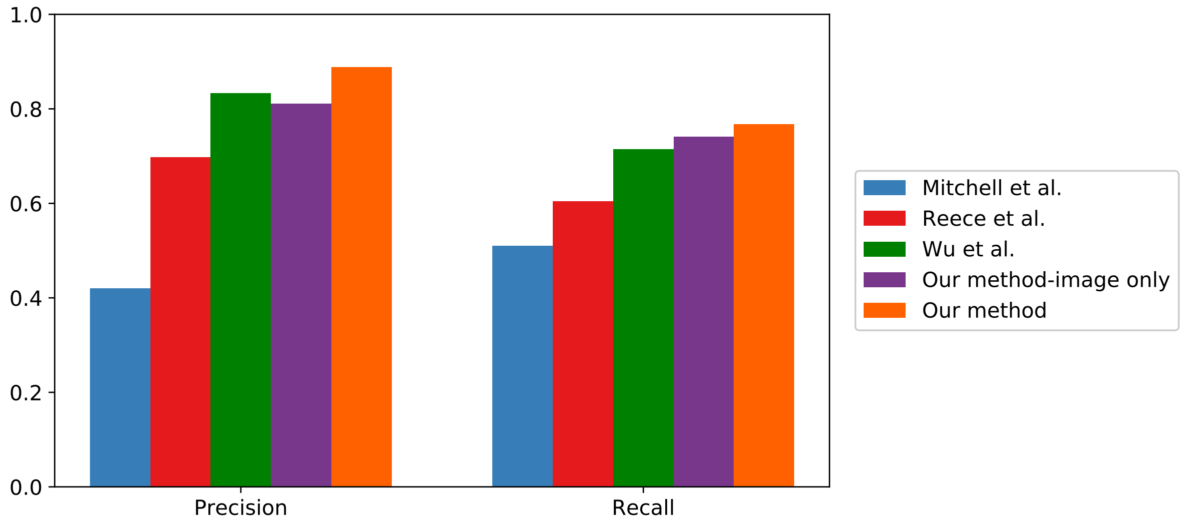

Figure 6: Comparison with other studies.

ing negative feelings. The experiment results show that modeling textual and visual features together performs much better than individual features. Although the user's behavior features on Instagram are less significant, they still help improve the performance of prediction.

6 CONCLUSION

Mining social media activities in order to understand mental health has been gaining considerable attention recently. In this paper, we demonstrate that it is possible to use online social media data for depression detection. We propose a machine learning approach for predicting depression tendency utilizing text, image and social behavior data. We use transfer learning to pre-train a CNN model for automatically extracting features from images, further combined with text and behavior features to build a deep learning classifier. The F-1 score of the classifier is 82.3% and the area under ROC curve (AUC) is larger than 0.5. Combining image and text makes the model more robust than using text only. In the future, we can take this research result to further work with medical institutions. Under the premise of privacy protection, by integrating the Instagram data of the clinic visiting patients, it is highly expected that the predictive precision can be enhanced and greatly contribute to the realization of tools for early screening and detection of depression, which further expands the potential value of this research.

REFERENCES

Al-Mosaiwi, M. and Johnstone, T. (2018). In an absolute state: Elevated use of absolutist words is a marker specific to anxiety, depression, and suicidal ideation. *Clinical Psychological Science*, page 2167702617747074.

Andalibi, N., Ozturk, P., and Forte, A. (2015). Depression-related imagery on instagram. In *Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative work & social computing*, pages 231–234. ACM.

Andalibi, N., Öztürk, P., and Forte, A. (2017). Sensitive self-disclosures, responses, and social support on instagram: The case of# depression. In *CSCW*, pages 1485–1500.

Chen, X., Sykora, M., Jackson, T., Elayan, S., and Munir, F. (2018a). Tweeting your mental health: an exploration of different classifiers and features with emotional signals in identifying mental health conditions.

Chen, X., Sykora, M. D., Jackson, T. W., and Elayan, S. (2018b). What about mood swings: Identifying depression on twitter with temporal measures of emotions. In *Companion of the The Web Conference 2018 on The Web Conference 2018*, pages 1653–1660. International World Wide Web Conferences Steering Committee.

De Choudhury, M., Counts, S., and Horvitz, E. (2013a). Predicting postpartum changes in emotion and behavior via social media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3267–3276. ACM.

De Choudhury, M., Counts, S., Horvitz, E. J., and Hoff, A. (2014). Characterizing and predicting postpartum depression from shared facebook data. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, pages 626–638. ACM.

De Choudhury, M. and De, S. (2014). Mental health discourse on reddit: Self-disclosure, social support, and anonymity. In *ICWSM*.

De Choudhury, M., Gamon, M., Counts, S., and Horvitz, E. (2013b). Predicting depression via social media. *ICWSM*, 13:1–10.

De Choudhury, M., Kiciman, E., Dredze, M., Coppersmith, G., and Kumar, M. (2016). Discovering shifts to suicidal ideation from mental health content in social media. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, pages 2098–2110. ACM.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. (2009). 34imagenet: A large-scale hierarchical image database. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 248–255. Ieee.

Grus, J. (2015). *Data science from scratch: first principles with python*. " O'Reilly Media, Inc.".

Instagram (Accessed July 26, 2016). *Instagram (2016) Instagram press release*. Instagram.

Kang, K., Yoon, C., and Kim, E. Y. (2016). Identifying depressive users in twitter using multimodal analysis. In *Big Data and Smart Computing (BigComp), 2016 International Conference on*, pages 231–238. IEEE.

Manikonda, L. and De Choudhury, M. (2017). Modeling and understanding visual attributes of mental health disclosures in social media. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 170–181. ACM.

Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Miller, G. A. (1995). Wordnet: a lexical database for english. *Communications of the ACM*, 38(11):39–41.

Mitchell, A. J., Vaze, A., and Rao, S. (2009). Clinical diagnosis of depression in primary care: a meta-analysis. *The Lancet*, 374(9690):609–619.

Nutt, D., Wilson, S., and Paterson, L. (2008). Sleep disorders as core symptoms of depression. *Dialogues in clinical neuroscience*, 10(3):329.

Oquab, M., Bottou, L., Laptev, I., and Sivic, J. (2014). Learning and transferring mid-level image representations using convolutional neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1717–1724.

Organization, W. H. et al. (2017). Depression and other common mental disorders: global health estimates. 2017. *Links*.

Pan, S. J. and Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10):1345–1359.

Panda, R., Zhang, J., Li, H., Lee, J.-Y., Lu, X., and Roy-Chowdhury, A. K. (2018). Contemplating visual emotions: Understanding and overcoming dataset bias. In *European Conference on Computer Vision*.

Park, M., Cha, C., and Cha, M. (2012). Depressive moods of users portrayed in twitter. In *Proceedings of the ACM SIGKDD Workshop on healthcare informatics (HI-KDD)*, volume 2012, pages 1–8. ACM New York, NY.

Ramirez-Esparza, N., Chung, C. K., Kacewicz, E., and Pennebaker, J. W. (2008). The psychology of word use in depression forums in english and in spanish: Texting two text analytic approaches. In *ICWSM*.

Reece, A. G. and Danforth, C. M. (2017). Instagram photos reveal predictive markers of depression. *EPJ Data Science*, 6(1):15.

Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Sun, J. (2012). 'jieba'chinese word segmentation tool.

Thrun, S. (1996). Is learning the n-th thing any easier than learning the first? In *Advances in neural information processing systems*, pages 640–646.

Wu, M. Y., Shen, C.-Y., Wang, E. T., and Chen, A. L. (2018). A deep architecture for depression detection using posting, behavior, and living environment data. *Journal of Intelligent Information Systems*, pages 1–20.

Yosinski, J., Clune, J., Bengio, Y., and Lipson, H. (2014). How transferable are features in deep neural networks? In *Advances in neural information processing systems*, pages 3320–3328.