The Blind Date: Improving the Accessibility of Mobile Dating Applications for Individuals with Visual Impairments

Meredith Moore, Corey Heath, Troy McDaniel, Sethuraman Panchanathan
Center for Cognitive Ubiquitous Computing
Arizona State University
Tempe, AZ, USA
{mkmoore7, corey.heath, troy.mcdaniel, panch}@asu.edu

***Abstract*—Online dating platforms such as Tinder, Bumble, and Coffee Meets Bagel have become not only culturally accepted, but incredibly common methods of meeting new people. However, many of these online dating platforms user profiles rely heavily on visual data. This reliance creates a barrier to access for individuals with visual impairments. Using voice-over systems with any of these applications leads to a very confusing interaction that withholds a lot of the available data from users with visual impairments. We propose a method of providing users with visual impairments access to the data encoded in the images of other users profiles. To learn what information people are most interested in knowing about the person in the dating profile, we surveyed both individuals with and without visual impairments to learn what visual features are important in an individual's dating profile. We then built a series of classifiers for the most popular features that were identified by the survey. Because the APIs for the dating applications mentioned above are not public, we created an application that works via screenshots. When the user would like to know more about a given profile, they just need to activate the native screenshot functionality on their phone, and our application will find the image in the screenshot folder on the phone, send it to our web service where the image is securely analyzed and output text is formed. This text is sent back to the phone and read out, providing the user with a coarse idea of what the image contained and helping them to make a more informed decision as to whether they are interested in the individual in the profile or not.**

***Index Terms*—accessibility, computer vision, blind, date**

I. Introduction

Online dating applications have become a very common way for people to meet each other, however, in most online dating platforms, the majority of the data is visual. This creates a barrier to access for individuals with visual impairments and prevents them from being able to fully utilize online dating platforms for the social tools that they are. Recently there have been a few significant advancements in assistive technologies for individuals with visual impairments. Tools like Microsoft's Seeing AI [1] have provided individuals with visual impairments some of the functions that generally require sight. While Seeing AI can help individuals with visual impairments with tasks such as reading/object detection, Seeing AI is mainly camera-based and doesn't work within other applications.

[1]https://www.microsoft.com/en-us/ai/seeing-ai

II. Related Work

This project aligns with research in automatic image description (AID). The aim of automatic image description is to output descriptive text provided the input of an image. In an ideal system, the goal would be a full understanding of an image [1]. Most AID systems use a combination of computer vision techniques and natural language processing to put together a coherent description of the image [2], [3]. While these systems have come a long way and are able to describe an image with relative accuracy, the kind of information that an individual is interested in knowing in a dating profile is different from the kind of information most automatic image description models learn. While image captioning and automatic image description based systems are related, the scope of these systems is a little bit too large for our specific application. Though it is important to know what is going on overall in an image, the features that an individual looks at in a dating profile have more to do with the physical characteristics of the individual portrayed in the user profile and are generally finer grain details than image description [4].

Previous computer vision work that focuses on describing physical characteristics of people are generally found in the facial recognition field [5]–[8]. Facial recognition is the process of recognizing a given face as a particular person. Facial recognition systems can be used for security as a validation method. Some of the more common and naive features that many of these facial recognition algorithms are looking at our age, ethnicity and gender [6], [9], [10]. There are several different algorithms for attaining decent features for each of these attributes of a human face, however, there has been little research in defining other human facial characteristics without getting into the field of facial expression recognition or emotion recognition. We focused on classifying hair length, face shape, smiling or not, gender, and the environment where the photo was taken (indoor/outdoor). Currently, to the best of the authors' knowledge, no accessibility applications have been created to function with common dating apps. However, researchers are going beyond screen readers and examining ways to make mobile usage that relies on visual information

978-1-7281-2723-1/19/$31.00 ©2019 IEEE

more accessible, particularly with photography features [11]–[13].

III. Survey

In order to obtain a better understanding of how both blind and sighted individuals experience online dating, we conducted two surveys. This need-finding survey is important for the development of person-centered accessible technologies. The first survey was administered to sighted individuals, and the second survey was administered to blind individuals. In both surveys, we asked participants to choose from a list which features were the most important to their decision to show interest in a given profile. In the survey that was given to individuals with visual impairments, we asked a few more questions specific to the online dating world for individuals with visual impairments.

A. Methodology

We used Google Forms to administer two different surveys. We chose to use two different surveys because we wanted to ask more in-depth questions about the dating app experience for individuals with visual impairments. Both surveys asked the participants to choose from a list which features were the most important in deciding whether or not to show interest in a given profile.

B. Participants

In total, we surveyed 82 participants on their experience using online dating apps. Of those 80 participants, 50 were sighted, and 32 were blind/low vision. All participants were at least 18 years old. The blind population averaged 37 years old, while the sighted population averaged 21 years old.

C. Results

When asked if they used online dating applications, 94% of the sighted participants indicated that they had used online dating applications before, while 76.7% of the blind group indicated that they used online dating applications.

1) Online Dating with a Visual Impairment: There are some simple accessibility issues that need to be fixed on most mobile dating applications. Things like labeling buttons, better HTML tags, and making it easy to navigate the application via screen reader could vastly improve the general accessibility of these applications. One of the respondents explained this lack of accessibility as:

> I find most online dating apps have accessibility issues or are completely inaccessible with screen readers... I'm also discouraged from online dating by having to enlist sighted assistance.

To show the reader how an individual with a visual impairment would experience a dating profile, watch the video provided below [2].

Beyond some relatively simple engineering fixes, When asked what kind of visual information blind individuals would

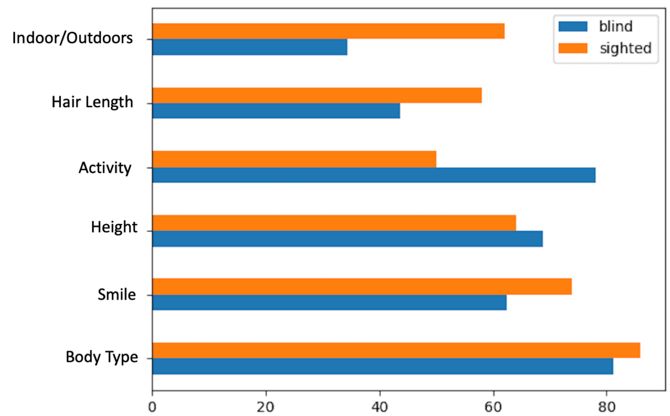

Fig. 1. Total (blind and sighted) response rate to the question 'What physical features are important in a profile?'

like to have access to, one of the most common answers was 'everything that's available to sighted people.' While this is a lofty goal, is a great one to keep in mind and aspire to. One blind participant noted that it would be very useful to know what kinds of activities/objects are in the photos that could help start a conversation (i.e. a book or football that could help initiate a conversation). There were a couple of individuals who would be interested in being able to hear the voice of a potential date.

Individuals with visual impairments generally look for primarily non-visual features, mainly in the potential date's bio. Blind users are more likely to use dating applications that have some text-based features embedded. When asked which dating platform is currently the most accessible 7 of the 14 respondents indicated that OKCupid was the most accessible. Tinder was mentioned by 4 respondents and seems to be more accessible than Bumble. It seems like the basic accessibility features are present in OKCupid and Tinder (labeled buttons, clear navigation, screen-reader compatibility).

When we asked individuals with visual impairments to rate how important physical appearance of a potential date is to them, 15 of the 30 respondents indicated that physical appearance was either very important or important.

2) Feature Preferences: Overall, the features that were identified as important the most frequently (by both sighted and blind individuals) are shown in Figure 1. These features are body shape (84.15%), smile (69.51%), height (65.85%), activity (60.98%), hair length (52.44%), and whether the picture was indoors or outdoors (51.22%). Figure 1 also shows the difference between the percent of blind and sighted individuals who chose each feature. Blind individuals were more likely to choose all of the physical features, while sighted individuals chose fewer physical features. Based on this analysis, these are some of the features that we chose to include in the system.

IV. Approach

Our goal was to create an application that could retrieve information from photos on a potential date's Tinder profile.

[2]https://youtu.be/FIqLmBEgsDs

The front-end interface of the application was developed for an Android mobile phone. This mobile app manages the network permissions needed to transfer data and launches a background thread that monitors the user's photo directory. When a new screenshot is added to the photo directory, the background thread sends the image to a cloud-hosted server that performs the image description tasks.

To use the application, the user navigates their mobile device using a screen reader to launch the application. Once the application is running, the user can run Tinder, and browse potential matches with the screen reader relating the text in the dating profile. To perform the image description, the user takes a screenshot using the native interface for their phone. The application will then give vocal updates about the process. Once the web call is complete, the image description is delivered to the user through the native Android text-to-speech.

The back-end processing is undertaken on a cloud-hosted server. Once the server receives the image, it performs the recognition tasks and builds a natural language text result. Based on the results of the service, the image description tasks that were selected are indoor/outdoor detection, gender recognition, hair length, face shape, and smile detection.

A. Indoor/Outdoor Detection

Detecting whether the photograph was taken indoors or outside was implemented using K-nearest neighbors. This was undertaken using the red-green-blue (RGB) pixel values. Only sub-samples of the image were utilized for classification. These samples consisted of 100x100 pixel squares at the corners of the image. This was done to avoid variation due to photos subjects, which can reasonably be assumed to be in the center of the frame. The sub-samples were compared to the labeled dataset and the class prediction was based on a K value of 10. Frobenius norm was used to compare the image sections. This method achieved an accuracy of 83% based on leave-one-out validation on the training dataset.

B. Gender Detection

While Tinder profiles contain information regarding age and gender, it only provides the info for the person whose profile is being checked. One could infer more knowledge about a person by knowing whom he/she hangs out with. This information could be extracted from group photos on an individual's profile. To do this it was necessary to detect each person in the frame and extract the characteristic features for them all. Gender classification was one important characteristic and hence each detected face was passed through the gender classifier which outputs whether the face is of a male or a female. For detecting the gender from a face we need to learn the distinctive features from both male-female classes. This was undertaken using the Fisherfaces method [14], which uses a linear discriminant analysis approach A Leave One Out Cross Validation on Yale-Face-database gives 98.2% recognition rate. The OpenCV [15] implementation was used for this project.

Fig. 2. A demonstration of how the feature extraction process for the classifier that determined hair length

C. Hair Length

We have trained a hair length classifier to classify a person into one of two classes Long Hair or Short Hair. The task involves two steps: preprocessing of data to collect the features of vectors and training a classifier based on this data. For the preprocessing step, GrabCut [16] along with information extracted from facial recognition algorithms was used to identify the shape of the individual's hair in training images. This is demonstrated in figure 2. Using the extracted information, a support vector machine (SVM) classifier was trained and achieved 89% accuracy on a validation set. The main limitation of this approach is that the model is sensitive to the background color near an individual's head. If the background was similar to a person's hair color, the model may incorrectly classify him/her as having long hair.

D. Face Shape

From the survey, we found that body shape is one of the key features that both sighted and blind individuals consider when judging a potential date's profile. It is very difficult to identify body shape from profile pictures as the majority of dating profile pictures focus on the face of the individuals. Because of this, we chose to assume that face shape is an indication of the body type and based on this assumption, we can classify the face as either long or not.

To classify, the face was first cropped from the profile picture using OpenCV. After cropping out the facial area, the eyes were identified using OpenCV and the distance between the two extremes of the eyes were calculated. An SVM model was trained using the height of the image and the width of the identified facial area as input features. An accuracy of 65% was achieved using this approach on a subset of the training data.

E. Smile Detection

After using OpenCV to detect faces, we used Haar-Cascades [17] to locate the left and right corners of the mouth. We evaluated the ratio of the distance between the left and right corner of the mouth and the width of the face. We then trained an SVM on this feature to classify whether or not a given face was smiling. To compare our results, we also used OpenCV's native function using Haar-Cascades to detect smiles.

V. Discussion

We identified a lack of accessibility for individuals with visual impairments when using online dating systems as most of the content is visually encoded. We surveyed both individuals with and without visual impairments on what physical characteristics were important in determining whether or not they were interested in a potential date. While there were differences between the features that blind and sighted groups used, the two groups were largely in agreement about the top 6 features of importance–body type, smile, height, activity, hair length, and the location of the photo (indoors v outdoors).

From this observation, we assessed which of these features were the most reliable and built a system that will for example output:

> "The photo was taken inside, the picture is of a woman who is smiling with long hair and has a long face."

To view a demonstration of this system, see the video below [3]. While we didn't quite achieve the task set out for us by blind users–to provide all of the information that is available to sighted individuals, this description provides significantly more information about the individual than was accessible before.

A. Shortcomings

This project presents a prototype application and is subject to several limitations. Most noticeably from a user perspective is the time taken to retrieve information regarding an image. Currently, the application takes 2-3 seconds on average to detect an image after the user has taken the screenshot. This interface method was chosen to circumnavigate security issues of having applications capture information from a user's screen; however, it requires using an agent to monitor a file location, which is not efficient. After receiving a screenshot, the application needs to send the image data through a web call to run the algorithms. Each algorithm is run independently, sequentially. This takes an additional 1-2 seconds. This gives a total user wait time of 3-5 seconds until the information is returned. This could be improved upon by continuously monitoring the screen to detect when a compatible image is present. A visual wake words [19] approach could be utilized to detect the image, then the system could notify the user that information is available.

The application would benefit from a different approach to capturing images. Ideally, the approach would detect a profile image and begin the classification process without the need for user initiation. This would allow the application to process and cache results, then return them when the user issues a command.

Improvements are also needed in order to make classification more efficient. Utilizing a cloud-based classification process provides the opportunity to employ distributed computing approaches. Tasks such as gender classification and indoor/outdoor detection could be implemented in parallel to reduce processing time. Several tasks, such as hair length and smile classification, used similar feature vectors based on initial face detection and would need to be run in parallel after feature extraction.

Additional classification tasks would need to be added to in order to provide a better description of what the photo depicts. Generalized image description [18] techniques could be used to provide an overview of the depiction, before providing more specific information on the individuals. Identifying specific information on individuals, such as visible tattoos and piercings, could be relevant for potential matches. Additionally, object detection, particularly of equipment used for recreation, could provide a means of inferring an individual's personality from the images.

VI. Conclusion

Mobile dating applications rely heavily on visual data and as such, are not very accessible to individuals with visual impairments. We developed a system that provides a textual description of images from mobile dating platforms using computer vision techniques. The solution relies on the user to screenshot an image after which the image is sent to a web service where the image is processed for a predetermined set of features. These features are then output as a text description of the image. Our system represents a feasible prototype for a solution to accessible mobile dating interfaces however it does have several limitations. Relying on the native phone screenshot capabilities allowed us to bypass the android application security however the method for taking screenshots may not be the most effective interface for our user base. With limited customization options for triggering screenshots, this could limit the use. This also caused a reliance on a file monitor which caused a noticeable delay in the system. The algorithms that were used to classify images relied on binary classes. This was done to keep the project achievable within a semester time limit. Expanding to multi-class classifiers and deep learning solutions is necessary to provide a greater degree of information to the user. Developing these algorithms within the domain of dating profile pictures allows for the use of assumptions on photo contains that can be useful for classification and verification. Similarly, specific information related to the looks and personality of the individual who included it in their profile is valued over a general description of the photo. These two aspects of the problem set it apart from the current research into descriptive text for images.

[3] https://youtu.be/Ivk0$_9GOgfA$

REFERENCES

[1] R. Bernardi et al., Automatic Description Generation from Images: A Survey of Models, Datasets, and Evaluation Measures, *ArXiv160103896 Cs* 2016.

[2] J. Johnson, A. Karpathy, and L. Fei-Fei, DenseCap: Fully Convolutional Localization Networks for Dense Captioning, In *Proc. of IEEE CVPR* pp. 45654574 2016.

[3] J. Devlin et al., Language Models for Image Captioning: The Quirks and What Works, *ArXiv150501809 Cs*, 2015.

[4] P. Sinha, B. Balas, Y. Ostrovsky, and R. Russell, Face Recognition by Humans: Nineteen Results All Computer Vision Researchers Should Know About, In *Proc. of IEEE*, 94(11), pp. 19481962, 2006.

[5] N. Bellustin et al., Instant Human Face Attributes Recognition System, *Int. J. Adv. Computer Science Applications* 3, 2011.

[6] S. Baluja and H. A. Rowley, Boosting Sex Identification Performance, *Int. J. Computer Vision* 71(1), pp. 111119, 2007.

[7] N. Pinto, J. J. DiCarlo, and D. D. Cox, How far can you get with a modern face recognition test set using only simple features?, In *Proc. of IEEE CVPR* pp. 25912598, 2009.

[8] E. Hjelms and B. K. Low, Face Detection: A Survey, *Computer Vision and Image Understanding* 83(3), pp. 236274, 2001.

[9] N. Kumar, A. C. Berg, P. N. Belhumeur, and S. K. Nayar, Attribute and simile classifiers for face verification, In *Proc. of IEEE International Conference on Computer Vision*, pp. 365372, 2009.

[10] P. Viola and M. J. Jones, Robust Real-Time Face Detection,*Int. J. Comput. Vis.* 57(2) pp. 137154, 2004.

[11] D. Adams, et al. "Blind photographers and VizSnap: A long-term study." In *Proc.of ASSETS*, 2016.

[12] J.P. Bigham, et al. Vizwiz: Nearly real-time answers to visual questions. In *Proc. of UIST 10*, pp. 333-342, 2010.

[13] C., Jayant,et al. Supporting blind photography. In *Proc.of ASSETS* pp. 203-210 2011.

[14] P.N. Belhumeur, et al. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis Machine Intelligence* 7, pp. 711-720, 1997.

[15] G. Bradski, The OpenCV Library, *Dr. Dobb's Journal of Software Tools*, 2000.

[16] C. Rother, V. Kolmogorov, and A. Blake. Grabcut: Interactive foreground extraction using iterated graph cuts. *ACM transactions on graphics (TOG)* 23(3), pp. 309-314, 2004.

[17] R. Padilla, et al. Evaluation of haar cascade classifiers designed for face detection, *World Academy of Science, Engineering and Technology* 64, pp. 362-365 2012.

[18] L. Zhou, et al. Grounded video description. In *Proc. of IEEE CVPR*, pp.6578-6587 2019.

[19] A. Chowdhery, et al. Visual Wake Words Dataset. *arXiv:1906.05721*, 2019.