ORIGINAL RESEARCH

Use of Facial Recognition Software to Identify Disaster Victims With Facial Injuries

John Broach, MD, MPH, MBA, FACEP; Rothsovann Yong, MD, MA; Mary-Elise Manuell, MD, MA, FACEP; Constance Nichols, MD, MS, FACEP

ABSTRACT

Objective: After large-scale disasters, victim identification frequently presents a challenge and a priority for responders attempting to reunite families and ensure proper identification of deceased persons. The purpose of this investigation was to determine whether currently commercially available facial recognition software can successfully identify disaster victims with facial injuries.

Methods: Photos of 106 people were taken before and after application of moulage designed to simulate traumatic facial injuries. These photos as well as photos from volunteers' personal photo collections were analyzed by using facial recognition software to determine whether this technology could accurately identify a person with facial injuries.

Results: The study results suggest that a responder could expect to get a correct match between submitted photos and photos of injured patients between 39% and 45% of the time and a much higher percentage of correct returns if submitted photos were of optimal quality with percentages correct exceeding 90% in most situations.

Conclusions: The present results suggest that the use of this software would provide significant benefit to responders. Although a correct result was returned only 40% of the time, this would still likely represent a benefit for a responder trying to identify hundreds or thousands of victims. (*Disaster Med Public Health Preparedness.* 2017;11:568-572)

Key Words: disaster medicine, disaster victim identification, facial recognition, mass casualty incidents

After large-scale disasters, victim identification frequently presents a challenge and a priority for responders attempting to reunite families and ensure proper identification of deceased persons.[1] After events that affect a large geographic area as well as those that generate a large number of casualties, such as the Indian Ocean tsunami in 2004, the challenges of victim identification are more difficult as a result of difficult travel conditions and a very large number of victims.[2,3] Family reunification is especially important for parents seeking their children, because children are particularly vulnerable after a disaster and may suffer significant ill effects of prolonged separation from their families.[4]

The challenges posed by large numbers of casualties is rightly the subject of an entire field within disaster response known as disaster victim identification or DVI.[5] This science involves a variety of techniques including analysis of personal effects, forensic odontology, DNA analysis, and more.[6,7] Although positive identification by family members alone is not considered conclusive victim identification for deceased victims, photographs of victims are frequently used to help family members of victims locate and identify their loved ones.[8] Indeed, advances in digital photography and easy transmission of data over the Internet and via text, SMS, and e-mail have made rapid transmission of digital images ubiquitous, even in disaster zones.[9] This has aided in the process of victim identification in disasters such as the 2010 Haiti earthquake when the Israeli emergency medical response team used a wireless network and laptops to allow people to view images of victims in order to aid in locating loved ones, in particular missing children.[10]

Although digital images can be captured and transmitted quickly and with very high resolution and fidelity, an unaided responder or the family member of a victim may still have to manually view many photographs to find a picture of a loved one. This can be very upsetting if family members need to view many images of injured people and it is certainly time consuming, especially in large-scale events where there may be thousands of victims and concerned family members.[2]

Facial recognition software has become sufficiently advanced and inexpensive that it is used in various

commercial applications and for uses as routine as "tagging" friends in group photos on social media platforms.[11] This availability of high-performing facial recognition software suggests that commercially available products may be of use to disaster responders in reducing the number of photographs that would need to be manually viewed in order to successfully reunite victims with families. This may save time and resources and spare victims and responders the trauma of viewing large numbers of victim photos in the search for a "match." In order to be used in this way, however, commercially available software would need to demonstrate the ability to recognize victims even if they had injuries to the face, which might be expected to hinder the performance of the software.

The purpose of this investigation was to determine whether currently commercially available facial recognition software can successfully match victim photographs with photos of the same people in an uninjured state. This would be the situation that could exist after a disaster in which family members submitted photographs of loved ones to a central database and software was used to narrow the number of potential matches for further review and confirmation by human responders.

METHODS

Study subjects were selected from volunteers among faculty and staff members in the Department of Emergency Medicine at the University of Massachusetts Memorial Medical Center and from a group of high school students participating in a series of health service career demonstrations through the Massachusetts Area Health Education Centers (MassAHEC) Health Occupations of America at The University of Massachusetts Medical School (UMass). All study photos were taken between January and March 2012.

All adult subjects provided informed consent to participate. Many of the participants from MassAHEC were under 18 years of age and thus parental consent was obtained for their participation. Prior to the day of the MassAHEC event, details of the study were provided in writing to the parents of potential participants and parents were given an opportunity to call the principal investigator with any questions regarding their child's potential participation.

The consent process and research protocol were reviewed and approved by the Institutional Review Board of UMass.

Prior to the study, the investigators determined injury patterns that would be considered "mild," "moderate," and "severe" based on extent, number, and combination of injuries. Twenty-five victim moulage template sets, each containing a unique combination of mild, moderate, and severe injury patterns were created so that on the day of the photos being taken moulage could be quickly applied.

TABLE 1

Facial Injury Severity Criteria	
Severity	Criteria
Mild	1 injury $\leq 3\,cm \times 2\,cm$
Moderate	1 or more injuries: $>3\,cm \times 2\,cm \leq 9\,cm \times 2\,cm$
Severe	1 or more injuries: $>9\,cm \times 2\,cm$ or area $>25\%$ of face

The victim moulage sets were created by dividing the face into 4 major zones. Zone I included the forehead, zone II included the eyes, zone III included the nose, and zone IV included the chin. The zones were further subdivided into right and left side. Seven different types of injuries ranging from a 3-cm linear laceration to an avulsion wound of the nose were assigned to the different zones and subdivisions to create unique injury patterns of varying severity. These victim moulage template sets were numbered and randomly assigned to each participant. The criteria for the various levels of injuries as well as an example of an assigned injury pattern are shown in Table 1 and Figure 1, respectively.

All study photos were taken with an iPhone 4s (Apple, Cupertino, CA), iPad 2 (Apple), and camera (Nikon D3100 with Nikon DX SWM VR aspherical 0.28 m/0.92 ft lens; Nikon Inc, Melville, NY) to simulate various devices that might be used to photograph disaster victims in an actual event. All photos were included in the comparison library.

Each participant then had a "no-moulage" photo taken with each of 3 devices. No-moulage photos were those taken on the day of the study before any facial injury moulage was applied. They were taken against a white background and were intended to serve as a control, representing an ideal prestudy photo (ie, taken in a well-lit environment, head on, without background interference, and in good focus).

After the no-moulage photo, mild, moderate, and severe injury photos were taken with each of the devices. Following this, all moulage was removed. This meant that each participant submitted one prestudy photo, had 3 no-moulage photos taken, and a total of 9 moulage photos taken.

For the purposes of this study, the prestudy photo is considered to be that which was either submitted by the participant prior to the study or a scanned image of a photo ID. This means that all prestudy photos were images taken before the subject's participation in the study. Participants were only instructed to submit a photo in which their face was included in a "facing forward" orientation. Otherwise, any type of photo could be submitted, including one taken either recently or in the more distant past and including other people or not. Some subjects elected to use images scanned from their government or work identification cards. This was

FIGURE 1

Injury Patterns for Victim I Moulage Template Set.
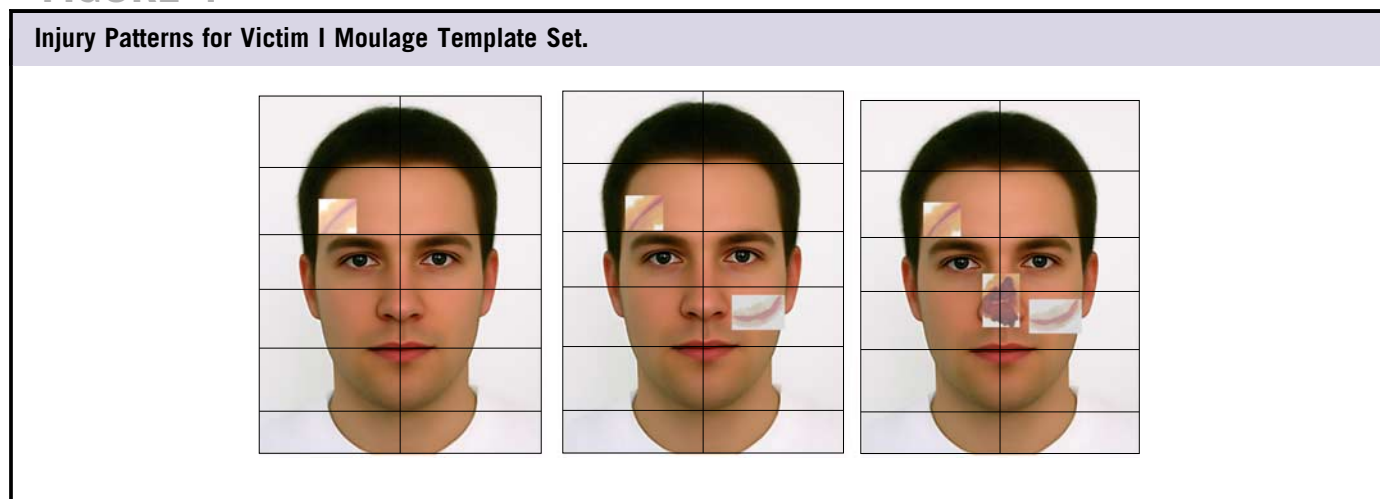
also an acceptable way to submit a prestudy photo. Each prestudy photo was entered into iPhoto (iPhoto 11 version 9.4.2; using Mac Desktop OS X version 10.8.2; Apple) and was matched by using this program against all study photos.

The analysis of these data included an assessment of the percentage of times that both the prestudy and "no-moulage" photos matched the moulaged photos over all degrees of injuries. For each prestudy photo and each no-moulage photo, the photo was compared against the library of study moulage photos and it was recorded whether or not a correct match was made and how many incorrect matches were returned. This was meant to simulate the process that a responder would go through if he or she was attempting to match a submitted photo to a library of disaster victims. Statistical analysis for this investigation was completed by using Statistics Calculator, StatPac Version 4.0 (StatPac, Inc, Northfield, MN).

RESULTS

A total of 136 subjects were initially identified to be a part of this study. Thirty subjects were excluded from the study; 7 had corrupt data when analysis was attempted, 7 were minors who initially expressed desire to participate but did not have consent forms signed by their guardians, 1 was not able to be scheduled for study photographs to be taken, and 15 did not submit a prestudy photo for comparison. A total of 106 participants had all needed data for analysis and were included.

Table 2 presents the demographic breakdown of the subjects in this study. Gender, race, age, presence or absence of facial hair, presence or absence of eyeglasses in the prestudy photo, and type of image used as a prestudy photo were recorded. The study included an equal number of male and female participants and participants of ages that spanned a range of between 15 to older than 65 years. We included a variety of races although participants were primarily white.

TABLE 2

Demographic Characteristics of the Study Subjects[a]

Characteristic	No. (%)
Gender	
Male	53 (50)
Female	53 (50)
Race	
Asian or Pacific Islander	14 (13.2)
Black	6 (5.7)
Hispanic	3 (2.8)
White	81 (76.4)
Multiracial	2 (1.9)
Age, years	
15-25	28 (26.4)
26-35	21 (19.8)
36-45	29 (27.4)
46-65	27 (25.5)
$\geq$66	1 (0.94)
Facial hair	
Yes	17 (16.0)
No	89 (84.0)
Glasses	
Yes	15 (14.2)
No	91 (85.8)
Type of image used for analysis	
Photograph	84 (79.2)
Government-issued ID	3 (2.8)
Other ID	19 (17.9)

[a]n = 106.

Both prestudy and no-moulage photos were considered "well photos" and moulaged photos were considered "injured photos." Each well photo was input into the facial recognition software and the correct matches and incorrect matches were recorded as well as whether each photo analyzed matched the minimal, moderate, or severe injury patterns.

Table 3 presents findings related to percentage correct matches and average numbers of returned incorrect photos for

TABLE 3

Correct Matches and Average Number of Returned Incorrect Photos

Camera	Match Mild (%)	Match Moderate (%)	Match Severe (%)	Mean Percentage of Correct Returned, %	Average Number of Incorrect
Prestudy photo	48.11	48.11	36.79	44.34	5.91
No-moulage photo	94.39	92.52	87.85	91.59	16.47
IPad					
Prestudy photo	46.22	42.45	31.13	39.93	6.15
No moulage photo	96.26	95.33	91.59	94.39	20.72
iPhone					
Prestudy photo	44.34	42.45	48.11	44.97	6.09
No-moulage photo	94.39	96.26	95.33	95.33	21.01

TABLE 4

Comparisons of Male and Female Subjects

	Male Percentage Correct, %	Female Percentage Correct, %	P Value[a]	Male Number of Incorrect	Female Number of Incorrect	P Value[b]
Camera	58.49	30.19	0.0041	8.75	3.06	0.0031
iPhone	58.49	31.44	0.0061	8.42	3.77	0.0110
iPad	59.12	30.19	0.0034	9.08	3.23	0.0014

[a]Two-sample t-test between percentages.
[b]Independent group t-test between means.

both types of well photos with respect to various levels of moulage in the injured photos. For all types of devices, prestudy photos matched mild and moderate moulage 42% to 49% of the time and severe moulage 31% to 49% of the time. The only statistically significant pair-wise comparison among correct match results was between the mobile phone photos and digital tablet photos for severe moulage with a P value of 0.048. It is unclear why the mobile phone obtained significantly better results on this measure. The rest of the results represent a non-statistically significant trend toward severe moulage being more difficult to match with prestudy photos and with no-moulage photos but no statistically significant differences were found between results for the 3 different devices studied or on the basis of severity of injury.

An average of the percentage correct matches of mild, moderate, and severe injury photos is represented as the mean of the percentages returned for each device and labeled as "mean percentage correct returned." Functionally, this was a measure of the expected results if this technology were deployed in real-world situations. The results of this study suggest that a responder could expect to get a correct match between submitted photos and photos of injured patients between 39% and 45% of the time depending on which device was used to take the photos of injured patients and a much higher percentage of correct returns if submitted photos were of optimal quality with percentages correct exceeding

90% in most situations. In addition, a relatively low number of incorrect results was returned in each situation, between 5 and 7 for prestudy photos and between 16 and 22 for no-moulage photos.

Analysis of other measured variables did not reveal statistically significant differences when subjects were compared by race, age, presence or absence of facial hair, presence or absence of eyeglasses, or type of prestudy photo used. However, both the likelihood of a correct return and the average number of incorrect returns were affected by the gender of the subject. Males were significantly more likely to match prestudy photos to injured photos than women but also had a larger number of incorrect returned images. Functionally, this meant that men were more likely to be correctly identified by the software but had more incorrect images returned with each correct result. Table 4 summarizes the results of comparisons of male and female subjects.

DISCUSSION
While the overall utility of this technology was made clear in this investigation, we obtained several unexpected results. It is unclear why more incorrect results were returned when using idealized "no-moulage" photos taken on the day of the study when compared to the submitted prestudy photo. It is unclear whether this was related to the similar background used in all study photos or to some other factor.

It was also expected that more severe facial injury patterns would limit the utility of the facial recognition software. However, we found no statistically significant difference among the levels of injury. It is unclear whether this was due to a negligible effect of injury on the performance of the software or whether different types of injury might have created more significant differences. No well established criteria for severity of facial injury in disasters exist and thus injury patterns were designed on the basis of the clinical experience in emergency medicine of the study authors.

Limitations

This study included several limitations which may limit its generalizability to real-world conditions. First, all photos were taken under ideal lighting conditions, which may have artificially enhanced the performance of the facial recognition program. In addition, injuries were simulated but other conditions such as cleanliness of the face, presence of perspiration, etc, were not assessed.

Another significant limitation of this study was that no nonsubjects were entered in the program. In a real-world situation, it is likely that loved ones looking for a family member who was not actually a victim would also submit photos and this was not accounted for in our investigation. Future investigations could attempt to determine how many "false-positive" identification results would be obtained from a library of injured photographs.

Finally, and perhaps most significantly, this study used a small number of participants and is therefore limited in terms of statistical power and generalizability. Future investigations using larger numbers of subjects would be helpful in determining the ultimate utility of this technology if it were deployed in a real-world disaster situation.

CONCLUSIONS

This study demonstrated that commercially available facial identification software may be of benefit to disaster responders trying to identify a large number of victims, even if these victims have significant facial injuries. The results of this study suggest that the use of this software would provide significant benefit to responders. Although a correct result was returned only 40% of the time, this would still likely represent a benefit for a responder trying to identify hundreds or thousands of victims. Given that several incorrect results were returned with each correct result, a manual verification would still be needed

to ensure proper identification of victims. A correct result was returned on a much more consistent basis (roughly 90% of the time) using idealized "well photos," which suggests that improving the quality of photos submitted would also likely improve the utility of this technology.

About the Authors

University of Massachusetts Medical School/University of Massachusetts Memorial Medical Center, Worcester, Massachusetts (Drs Broach and Nichols); Department of Emergency Medicine, Lowell General Hospital, Lowell, Massachusetts (Dr Yong); and Urgent Care, Harrington HealthCare System, Southbridge, Massachusetts (Dr Manuell).

Correspondence and reprint requests to John Broach, MD, MPH, MBA, FACEP, UMass Medical School/UMass Memorial Medical Center, 55 Lake Avenue North, Worcester, MA 01655 (e-mail: john.broach@umassmemorial.org).

REFERENCES

1. Federal Emergency Management Agency (FEMA). Post-Disaster Reunification of Children: A Nationwide Approach. https://www.fema.gov/media-library/assets/documents/85559. Published November 2013. Accessed March 22, 2017.
2. Morgan OW, Sribanditmongkol P, Perera C, et al. Mass fatality management following the South Asian tsunami disaster: case studies in Thailand, Indonesia, and Sri Lanka. *PLoS Med.* 2006;3(6):e195. https://doi.org/10.1371/journal.pmed.0030195.
3. Tsokos M, Lessig R, Grundmann C, et al. Experiences in tsunami victim identification. Experiences in tsunami victim identification. *Int J Leg Med.* 2006;120(3):185-187. doi: 10.1007/s00414-005-0031-4
4. Nager A. Family reunification: concepts and challenges. *Clin Pediatr Emerg Med.* 2009;10(3):195-207.
5. Byard R, Winskog C. Potential problems arising during international disaster victim identification (DVI) exercises. *Forensic Sci Med Pathol.* 2010;6(1):1-2. https://doi.org/10.1007/s12024-009-9141-5.
6. InterPol. Disaster Victim Identification (DVI). http://www.interpol.int/INTERPOL-expertise/Forensics/DVI. Accessed March 22, 2017.
7. Wright K, Mundorff A, Chaseling J, et al. A new disaster victim identification management strategy targeting "near identification-threshold" cases: experiences from the Boxing Day tsunami. *Forensic Sci Int.* 2015;250 (May):91-97. https://doi.org/10.1016/j.forsciint.2015.03.007.
8. Broughton D, Allen E, Hannemann R, et al. Reuniting fractured families after a disaster: the role of the National Center for Missing & Exploited Children. *Pediatrics.* 2006;117(5). https://doi.org/10.1542/peds.2006-0099S.
9. Bikker J. Disaster victim identification in the information age: the use of personal data, post-mortem privacy and the rights of the victim's relatives. *Scripted.* 2013;10:1. https://script-ed.org/article/asas/.
10. Levy G, Blumberg N, Kreiss Y, et al. Application of information technology within a field hospital deployment following the January 2010 Haiti earthquake disaster. *J Am Med Inform Assoc.* 2010;17:626e630.
11. McHugh M. Facebook can recognize you don't show your face. *Wired Magazine.* http://www.wired.com/2015/06/facebook-can-recognize-even-dont-show-face/. Published June 24, 2015. Accessed March 22, 2017.