Anonymizing facial images to improve patient privacy

To minimize the risks of inappropriately disclosing facial images of patients, we developed the digital mask to erase identifiable features while retaining disease-relevant features needed for diagnosis. The digital mask has shown the ability to evade recognition by human researchers and existing facial-recognition algorithms, and improves patients' willingness to share medical information.

This is a summary of:
Yang, Y. et al. A digital mask to safeguard patient privacy. *Nat. Med.* https://doi.org/10.1038/s41591-022-01966-1 (2022).

Published online:
15 September 2022

The problem

Privacy protection of facial images has attained prominence over the past decade owing to the digitalization of medical records and concerns about data breaches. Facial images are useful for identifying signs of disease; however, these inevitably record patient biometric identifiers. Thus, the first challenge is to separate biometric identity from medical information that can be derived from facial images.

Additionally, concerns about image breaches have hampered the development of digital health technology (for example, deep learning assistance) that depends on broad participation in medical data collection. The second challenge is to improve the willingness of healthcare users to share their facial images and to reduce concerns about the misuse of facial-recognition technology.

The solution

We set out to develop an alternative procedure for sharing and recording facial images, and to provide an additional data format for privacy protection.

As periocular biometrics are one of the most distinctive subsets of individual biometric information[1], we focused on four pathological ocular manifestations that involve more than ten abnormal phenotypes. We developed the digital mask (DM), which inputs an original video of a patient's face and outputs a de-identified three-dimensional (3D) reconstructed video based on the complementary use of deep learning and 3D reconstruction. Deep learning achieves feature extraction from various facial parts, while 3D reconstruction automatically digitizes the shapes and motions of 3D faces, eyelids and eyeballs based on the extracted facial features[2–4] (Fig. 1). Converting DM-reconstructed videos back to the original videos is extremely difficult because most of the necessary information is no longer retained in the set of digital representations that constitute this mask.

Experiments were then conducted to validate the efficiency of the DM. First, to assess reconstruction performance of the DM, we quantitatively evaluated the error between original videos and DM videos at the pixel level. Error was extremely low, ranging from 0.81% to 1.61%. Second, we evaluated the DM in clinical practice. The independent diagnoses from the original videos and the diagnosis from the DM-reconstructed videos were highly consistent ($k > 0.8$). We also compared the DM with the traditional de-identification method of cropping and found that the risk of being identified was decreased in the masked patients. Third, we evaluated the willingness of patients to share videos processed anonymously by the DM. Over 80% of patients believed that the DM can alleviate privacy concerns and expressed an increased willingness to share their personal information. Finally, we confirmed that the DM can also evade artificial intelligence–powered facial-recognition algorithms.

The implications

'Protecting privacy' does not equate to 'absolute removal of identity characteristics'[5]. One of the most important principles of privacy protection is balancing disclosure risk against data utility. Therefore, the purpose of the DM is to provide an approach to health-information disclosure that de-identifies protected health information as much as possible, without compromising the ability of clinicians to reach a diagnosis.

In addition to its potential utilization in research and routine clinical practice, the DM could be applied to telemedicine, including online automatic diagnosis and patient triage for more-efficient healthcare delivery. Furthermore, the DM can obtain quantitative parameters (such as the degree of eyeball rotation, eyelid shape parameters, and rotation frequency), which might help diagnosis in the future. Additionally, many other non-ocular disorders involve facial manifestations, and we propose that with further development, the DM has the potential to be applied in, for example, otorhinolaryngology, neurology, and oral and maxillofacial surgery.

One limitation of our study is that the reconstruction of conjunctival hyperemia, eyelid edema and abnormal tissue growth, such as ocular tumors, remains challenging owing to insufficient model capacity. We intend to improve the DM by including a sufficiently large sample of abnormal cases for detailed analysis, or by constructing an extra sub-model on top of the existing model. In addition, the risk that the DM might be compromised still remains, and we hope to formulate relevant rules of technology security in the future.

Ruixin Wang and Haotian Lin, State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangzhou, Guangdong, China.

EXPERT OPINION

"This manuscript, which proposes an alternative approach to imaging anonymization, holds a great deal of value for the medical community because facial imaging is nearly impossible to anonymize. For this reason, I think this is a positive contribution to the literature." **Charlotte Tschider. Loyola University Chicago, Chicago, IL, USA.**

FIGURE

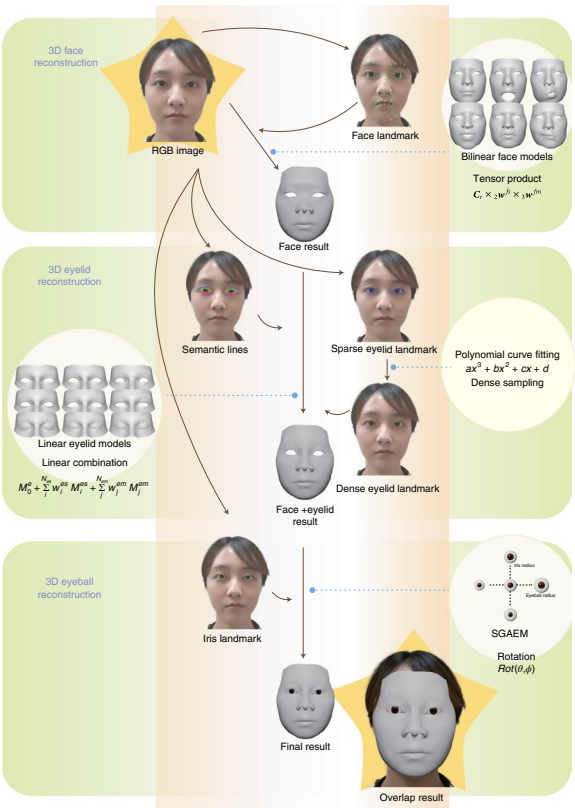

Fig. 1 | Development of the DM. The DM uses RGB (red-green-blue) images (individual frames from a video) as input and outputs 3D reconstructed meshes. For a particular frame, the algorithm first extracts 2D facial landmarks from the RGB image and fits a set of face-model weights for 3D face reconstruction. The algorithm then extracts 2D eyelid landmarks and 2D semantic lines, and fits eyelid model weights for 3D eyelid reconstruction. Finally, the algorithm extracts 2D iris landmarks and solves eyeball rotation for 3D eyeball reconstruction. SGAEM, simplified geometry and appearance eyeball model.

BEHIND THE PAPER

Our work was born in the context of the COVID-19 pandemic. To meet the explosion in demand for tele-medicine, our Zhongshan Ophthalmic Centre built an internet hospital. As remote healthcare for ophthalmic diseases requires a large amount of digital facial-related medical information, we realized that healthcare organizations must take more responsibility for protecting privacy and enable patients to uphold the right to decide how their data is used. The team of Qionghai Dai and Feng Xu is a leader in the field of face reconstruction. We met at an academic conference many years ago and became good friends and collaborators. After an in-depth discussion in a regular exchange meeting, I clearly remember that we were all very excited because we had found a promising solution to the problem, and the technology matched an important application scenario. This project has been a pleasant and successful medicine–engineering collaboration. **H.L.**

REFERENCES

1. Mason, J. et al. An investigation of biometric authentication in the healthcare environment. *Array* **8**, 100042 (2020).
This paper reports that periocular biometrics can be used to assist in building robust identity-verification systems.

2. Cao, C., Weng, Y., Zhou, S., Tong, Y. & Zhou, K. Facewarehouse: A 3d facial expression database for visual computing. *IEEE Trans. Vis. Comput. Graph.* **20**, 413–425 (2013).
This paper reports a representative bilinear model of face reconstruction.

3. Wen, Q., Xu, F., Lu, M. & Yong, J.-H. Real-time 3D eyelids tracking from semantic edges. *ACM Trans. Graph.* **36**, 193 (2017).
This original article proposes a technique for 3D eyelid reconstruction.

4. Wen, Q., Xu, F. & Yong, J.-H. Real-time 3D eye performance reconstruction for RGBD cameras. *IEEE Trans. Vis. Comput. Graph.* **23**, 2586–2598 (2016).
This research article proposes a method for reconstruction of 3D eye performance.

5. Health Insurance Portability and Accountability Act of 1996. Pub. L. No. 104-191, 110 Stat (1996).
The US federal statute that defines US privacy rule standards.

FROM THE EDITOR

"The diagnostic use of facial images by digital medicine poses risks for patient privacy. This study develops a solution for overcoming these risks by algorithmically removing patient-identifying information from facial images while retaining sufficient information for accurate diagnosis. This type of privacy-preserving methodology can facilitate public acceptance of facial imaging for use in digital medicine." **Editorial Team**, ***Nature Medicine***