



OPEN ACCESS

EDITED BY
Matteo Ferrara,
University of Bologna, Italy

REVIEWED BY
Christian Rathgeb,
Darmstadt University of Applied
Sciences, Germany
Robert Goldberg,
UCLA Health System, United States
Ahmad Saaed Mohammad,
Al-Mustansiriya University, Iraq

*CORRESPONDENCE
Luuk Spreeuwers
l.j.spreeuwers@utwente.nl

SPECIALTY SECTION
This article was submitted to
Computer Vision,
a section of the journal
Frontiers in Computer Science

RECEIVED 31 May 2022
ACCEPTED 05 July 2022
PUBLISHED 03 August 2022

CITATION
Lestriandoko NH, Veldhuis R and
Spreeuwers L (2022) The contribution
of different face parts to deep face
recognition.
Front. Comput. Sci. 4:958629.
doi: 10.3389/fcomp.2022.958629

COPYRIGHT
© 2022 Lestriandoko, Veldhuis and
Spreeuwers. This is an open-access
article distributed under the terms of
the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution
or reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

The contribution of different face parts to deep face recognition

Nova Hadi Lestriandoko^{1,2}, Raymond Veldhuis¹ and Luuk Spreeuwers^{1*}

¹Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, Enschede, Netherlands, ²National Research and Innovation Agency, Jakarta, Indonesia

The development of face recognition improvements still lacks knowledge on what parts of the face are important. In this article, the authors present face parts analysis to obtain important recognition information in a certain area of the face, more than just the eye or eyebrow, from the black box perspective. In addition, the authors propose a more advanced way to select parts without introducing artifacts using the average face and morphing. Furthermore, multiple face recognition systems are used to analyze the face component contribution. Finally, the results show that the four deep face recognition systems produce a different behavior for each experiment. However, the eyebrows are still the most important part of deep face recognition systems. In addition, the face texture played an important role deeper than the face shape.

KEYWORDS

face exploitation, face component contribution, deep face recognition, face morphing, average face, face texture, face geometry

Introduction

The face is one part of the human body that is commonly used in biometrics and forensic identification. It contains discriminative information, allowing people to be recognized. Automatic face recognition systems exploit this information to compare face images to other face images or face models. The main steps in facial recognition are face detection, face alignment, feature extraction, and classification.

Some challenges affect Face Recognition (FR) performance, for instance, partial occlusion. Recently, the COVID-19 pandemic affected the world and created a new normal that covering the face with a mask is an obligation. Extreme occlusion, particularly covering a large area of the face (for example, by a balaclava or hazmat suit), also is a major challenge in face recognition. It leaves only a small area of the face visible and thus significantly decreases recognition performance.

In recent years, the study of the face part's contribution to recognition has been studied extensively as part of solving FR challenges. The masking of a certain face area and partial occlusion were commonly used to demonstrate a decrease in performance. Most of the works investigated the impact of masking face regions on classical face recognition methods. Only a few studies were published on the impact of deep

learning-based face recognition. Radji et al. (2015) analyzed the impact of eyebrows and eyes removal on traditional FR by blurring those areas. Hofer et al. (2021) blacked out certain face areas (eyes, nose, and mouth) and analyzed how three state-of-the-art face detection methods (MTCNN, Retinaface, and DLIB) were affected. They also concluded that the region around the nose plays an important role in face detection. Lestriandoko et al. (2019) analyzed the different frequency band contributions to recognition using DWT and LBP. They also presented a block-based area contribution. Furthermore, there were some works on partial occlusion to improve face recognition of partly occluded faces. Zhao et al. (2021) improved FaceNet based on the attention mechanism to address the partial occlusion challenge. Scarf occlusion, glasses, and 10–40% partial face occlusion were used to show how the proposed methods performed. Damer et al. (2020) compared the effect of wearing masks on three deep face recognition methods: ArcFace, SphereFace, and a commercial system COTS (MegaMatcher 11.2 SDK from the vendor Neurotechnology). The experimental results showed that wearing a mask decreased the performance of open face recognition systems (ArcFace and SphereFace). However, the COTS still worked well without any significant performance drop. Malakar et al. (2021) reconstructed the occluded area using PCA, then classified the features using FaceNet. They improved performance by up to 15%. Furthermore, Mohammad et al. (2019) solved one of the challenges, that is, face with eyeglasses, using an eyeglasses frame removal method to improve face recognition on datasets containing eyeglasses. An overview of the current face recognition system response to occluded areas and how they solved the occlusion challenges was published by Zeng et al. (2021). The researcher classified the techniques of face recognition under occlusion into three categories: occlusion robust feature extraction (ORFE), occlusion aware face recognition (OAFR), and occlusion recovery-based face recognition (ORecFR). Another study (Zhang et al., 2020) reviewed the methods of occlusion in 2D face detection, 2D face recognition, and 3D face recognition.

Although some methods successfully localized certain parts of the face, and even retrained the network using this area, the understanding of the contribution of face components for deep face recognition remains limited. This article attempts to answer the research question: In terms of location, what is the important information in the face for recognition? The goal is to obtain knowledge of the face component's contribution to deep face recognition by removing certain information in the image. We used the average face morphing to remove certain information in the face. The face components' contribution is obtained by observing the response of four deep face recognition systems: DLIB (King, 2009), FaceNet (Schroff et al., 2015), ArcFace (Deng et al., 2019), and a commercial FR system FaceVacs (FaceVACS Engine 9.4.0., 2020).

The main contributions of this article are as follows.

1. An advanced way to select parts without introducing artifacts using the average face and morphing.
2. A comprehensive analysis of the face component contribution.
3. A comparative evaluation of the response of four deep face recognition systems.

The structure of the article is organized as follows. Section Related works presents related work. Then, the proposed methods are described in Section Proposed methods. Next, Section Experimental results reports the experimental results. Finally, the conclusion is presented in Section Discussion and conclusions.

Related works

Block-based analysis

Understanding the impact of facial characteristics is relevant for the further development of face recognition. Some research in this area has highlighted the role of certain face features for recognition, such as the eyes, eyebrows, ears, mouth, and nose. As in previous research, the face parts analysis for the classic method was performed by applying block-based analysis. Tome et al. (2013) separated the face into 15 facial regions and used PCA-SVM for feature extraction and classification. Their results showed that the nose area has significant discriminative power that gives near to full-face performance. However, the separation of face parts was not optimal because the nose region still contained a small part of the eyes and eyebrows. A similar method was presented by Juefei-Xu and Savvides (2011), which used eyebrows as a stand-alone biometric for recognition. Researchers compared the importance of eyebrows, eyes, and full face. Their results showed that eyebrows are a promising biometric for identifying people. Furthermore, Ahonen et al. (2004), Caifeng et al. (2005), Nikisins and Greitans (2012), and Loderer et al. (2015) analyzed the importance of the area (blocked-based area) of the face for recognition or detection using LBP. In detail, Lestriandoko et al. (2019) also compared the discriminative area of LBP between the clear dataset and the dataset containing face expression. They found that the dataset with face expression variation produced a smaller area of the discriminative component, especially on the face component shape. In contrast, almost the whole face component was meaningful for LBP on the clear dataset. Still, in the context of face components, face geometry, and texture also play a role in face recognition and detection.

Geometry analysis

Some previous works also analyzed the role of face geometry in recognition. Most of them used face geometry for expression

or emotion recognition. Lin et al. (2021) proposed an emotion recognition system based on the face component geometry (face landmarks) mask. A deep learning Convolutional Neural Network (CNN) is used for classification. Oztel et al. (2018) proposed facial expression recognition using eyes and eyebrows geometry only. They claimed their method was robust to lower face occlusions and motion. Furthermore, García-Ramírez et al. (2018) proposed mouth and eyebrow segmentation and interpolated polynomial feature extraction. In that study, the mouth and eyebrows geometry played an important role in emotion recognition.

Masking and removing face components

The other ways to analyze the contribution of face parts are face part removal and masking (face occlusion). Spreeuwers et al. (2014) defined 30 mask regions by covering and excluding certain kinds of variation, for example, excluding eyes and mouth. For each region, a separate classifier is designed. By the use of Fixed FAR Vote Fusion, they successfully improved the recognition performance. In the same area, Trigueros et al. (2018) trained a CNN architecture using various occlusions with various intensities, noise types, and noise levels to find out which part of the face is more important for facial recognition. The resulting knowledge was then used to improve the CNN robustness against face occlusion. Furthermore, Sadr et al. (2003) described a novel approach to determining the importance of facial parts. The eyes and eyebrows were removed by the Photoshop clone function, and they observed how the recognition performance decreased. The results demonstrated that the eyebrows influence the performance as significantly as the eyes. Radji et al. (2014) described a similar way of determining the importance of eyes and eyebrows on FR using various traditional methods, that is, PCA, SVD, DWT, DCT, and DWT-SVD. They showed that removal of the eyebrows decreased recognition performance. However, there was a limitation to their experiments in that the used datasets were too small so any conclusions would lack validity.

Face manipulation using morphing

Face manipulation using morphing was commonly used to attack the human identity on the recognition system. The morphing produced smooth changes of face shape and texture, seamless, and low distortion. Qin et al. (2021) proposed a method of morphing attacks on partial face components. They compared the vulnerability of the face recognition system to morphing attacks on nine partial face components and showed that the nose and eyes produced the biggest impact. Furthermore, Midtun et al. (2017) presented the use of average faces to anonymize faces in journalism. The whole face morphing successfully removed the main identity. However, the

non-facial attributes, for example, hair style and clothes, still gave enough information for the reader to recognize the person.

Proposed methods

We proposed six advanced ways to analyze the face component contribution as follows, related to the shape and geometry of the face parts.

- a) **NDT**: Non Discriminative Texture
- b) **iNDT**: Inverse Non-Discriminative Texture
- c) **NDTS**: Non Discriminative Texture and Shape
- d) **iNDTS**: Inverse Non-Discriminative Texture and Shape
- e) **NDS**: Non Discriminative Shape
- f) **iNDS**: Inverse Non-Discriminative Texture and Shape

For each method, except for NDS and iNDS, the morphing was applied to single and multi-components of the face below.

1. Eyebrows (numbers 18–27 of DLIB points)
2. Eyes (numbers 37–48 of DLIB points)
3. Nose (numbers 28–36 of DLIB points)
4. Mouth (numbers 49–60 of DLIB points)
5. Eyebrows-Eyes
6. Eyebrows-Nose
7. Eyebrows-Mouth
8. Eyes-Nose
9. Eyes-Mouth
10. Nose-Mouth
11. Eyebrows-Eyes-Nose
12. Eyebrows-Eyes-Mouth
13. Eyebrows-Nose-Mouth
14. Eyes-Nose-Mouth
15. Eyebrows-Eyes-Nose-Mouth

Initially, the mean of all registered datasets was calculated as the average face that will be used in all experiments to remove the discrimination information on the face. Next, we applied the morphing using six advanced ways (NDT, iNDT, NDTS, iNDTS, NDS, and iNDS) to a particular area of the face, for example, the eyebrows-nose. Then, the verification was established on four deep face recognition systems. For similarity matrix calculation, both of gallery and probe were the modified face. Thus, we did not take into account the contribution of specific face parts for the similarity score. Finally, False Non-Match Rate (FNMR), False Match Rate (FMR), and Detection Error Tradeoff (DET) were used to analyze the response of deep face recognition systems to face components removal.

Average face calculation

The average face was generated from the FRGCv2 dataset containing 2,000 controlled images from 100 individuals, facial expression, slightly pose, male and female gender, various races,

skin tones, and ages. Figure 1 shows the method of average face calculation. First, all images of the FRGCv2 dataset were aligned based on five landmark points, that is, four points eyes corner, and a nose peak point, using DLIB alignment. Then, we extracted the 68 face landmarks for the whole dataset. The mean points were obtained by averaging the 2,000 face landmarks of the dataset. That mean point was used to re-align the dataset to the same 68 landmarks using the morphing. This step affected the shape of the face and produced some weird faces. However, the process goal was to generate a perfect and clear average face. Thus, the method required that all images were in the precise landmarks.

Face morphing

Face morphing is an image manipulation method to combine two separate faces image into a new image. Normally, the face morph uses weighted landmark points to control the desired face. However, we used a modified face morphing to change the discrimination information in a specific face area in the experiments.

All methods used similar steps, that is, morphing points target calculation, creating the Delaunay triangles, and image transformation. The seamless method, introduced by Pérez et al. (2003), was used at the end of the morphing process to make the cropping border smoother. It adapted the texture and

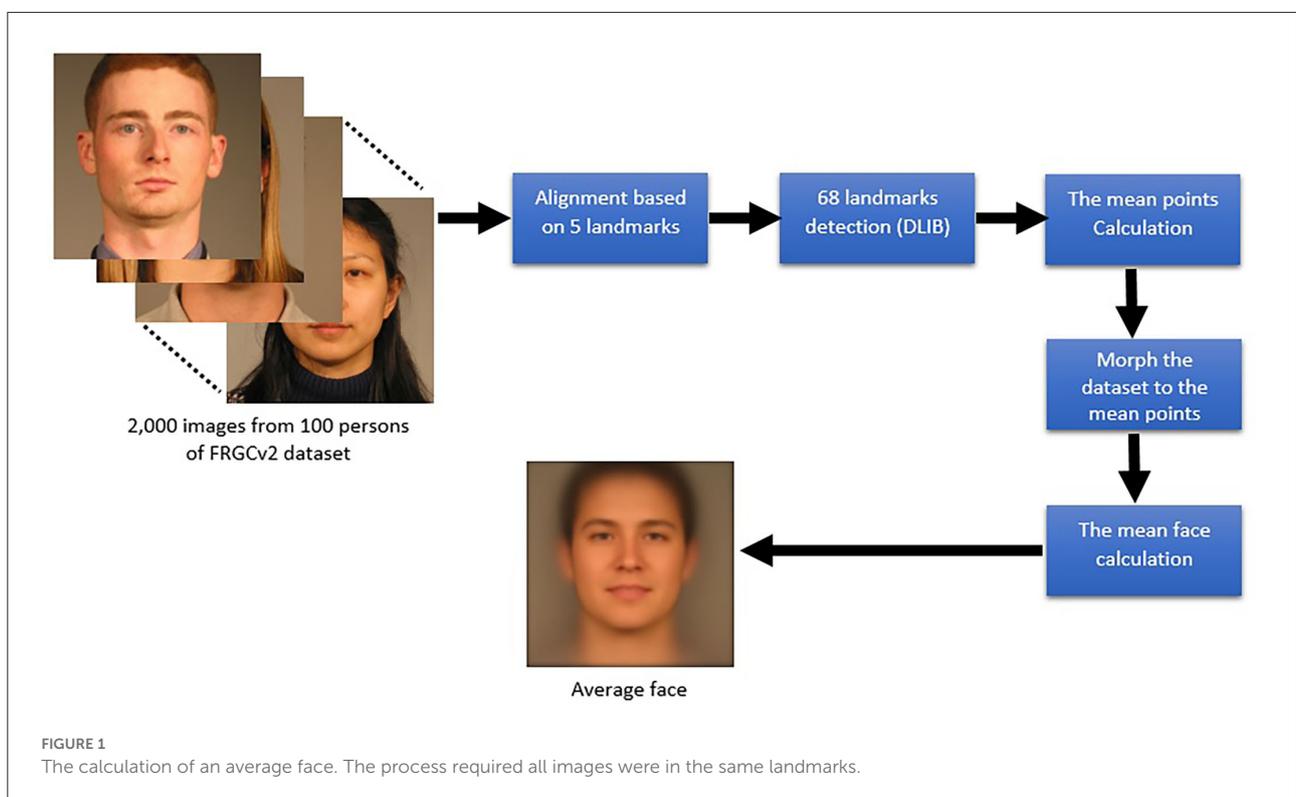
illumination of the object into the color of the background. Moreover, the differences between the six experiments, as shown in Table 1, were in the use of input and target image, the use of facial landmarks, and the morphing target.

Non-discriminative texture and inverse non-discriminative texture

The NDT method replaced the face component using the morphed average face. The goal is to remove the inner face component without changing the shape/geometry. The steps of NDT, as shown in Figure 2, are described as follows.

1. Morph partial part of the average face geometry to the original face landmarks.
2. Replace the original face part with the average face component.
3. Smooth the contrast between cropped part and background using a seamless method.

iNDT has similar steps to NDT. However, iNDT inversed the use of average face and original face. The aim is to exclude only the discrimination information of specific face components. In other words, the iNDT removes all discrimination information in the face, except the detail or texture of particular face components. This method replaced the average face component



using the morphed original face parts without changing the shape/geometry. The method is arranged as follows.

1. Morph partial part of the original face geometry to the average face landmarks.
2. Replace the average face part with the original face component.
3. Smooth the contrast between cropped part and background using a seamless method.

Non-discriminative texture and shape and inverse non-discriminative texture and shape

This method removed the original face components' geometry and texture using average face parts replacement on the original face. The aim is to exclude the discrimination information of the face components completely, that is, remove both the face part shape and pattern. The NDTS steps, as shown in Figure 3, are determined as follows.

1. Specify the larger area of a specific face component that covers the face component geometry.
2. Morph the original face component geometry only (not the whole geometry) to the average face.
3. Replace the original face part with the average face component.
4. Smooth the contrast between cropped part and background using a seamless method.

For the iNDTS, the use of geometry morphing was inversed. The aim is to remove the average face shape and texture using the original face parts replacement on the average face. This method removed all discrimination information in the face except the shape and details of specific face components. The steps are as follows.

1. Specify the larger area of a specific face component that covers the face component geometry.
2. Morph the average face component geometry to the original face.
3. Replace the average face part with the original face component.
4. Smooth the contrast between cropped part and background using a seamless method.

Non-discriminative shape and inverse non-discriminative shape

The NDS experiment removed only the original face shape using geometry face morphing to the average face, without changing the texture. The method is similar to the alignment

TABLE 1 The differences in face morphing methods.

Methods	Input	Target	Morph area	Landmark
NDT	Original image	Average image	Inner texture	Face parts
iNDT	Average image	Original image	Inner texture	Face parts
NDTS	Original image	Average image	Shape and texture	Face parts
iNDTS	Average image	Original image	Shape and texture	Face parts
NDS	Original image	Average image	Shape	Whole face
iNDS	Average image	Original image	Shape	Whole face

process for face registration. Thus, the original face landmarks were changed into the average face landmarks. The steps, as shown in Figure 4, are defined as follows.

1. Extract the face landmarks.
2. Morph the original face geometry to the average face geometry.

On the other hand, iNDS morphed the average face geometry into the original face geometry. Therefore, the shape of the average face mimicked the shape of the original face. The steps are described as follows.

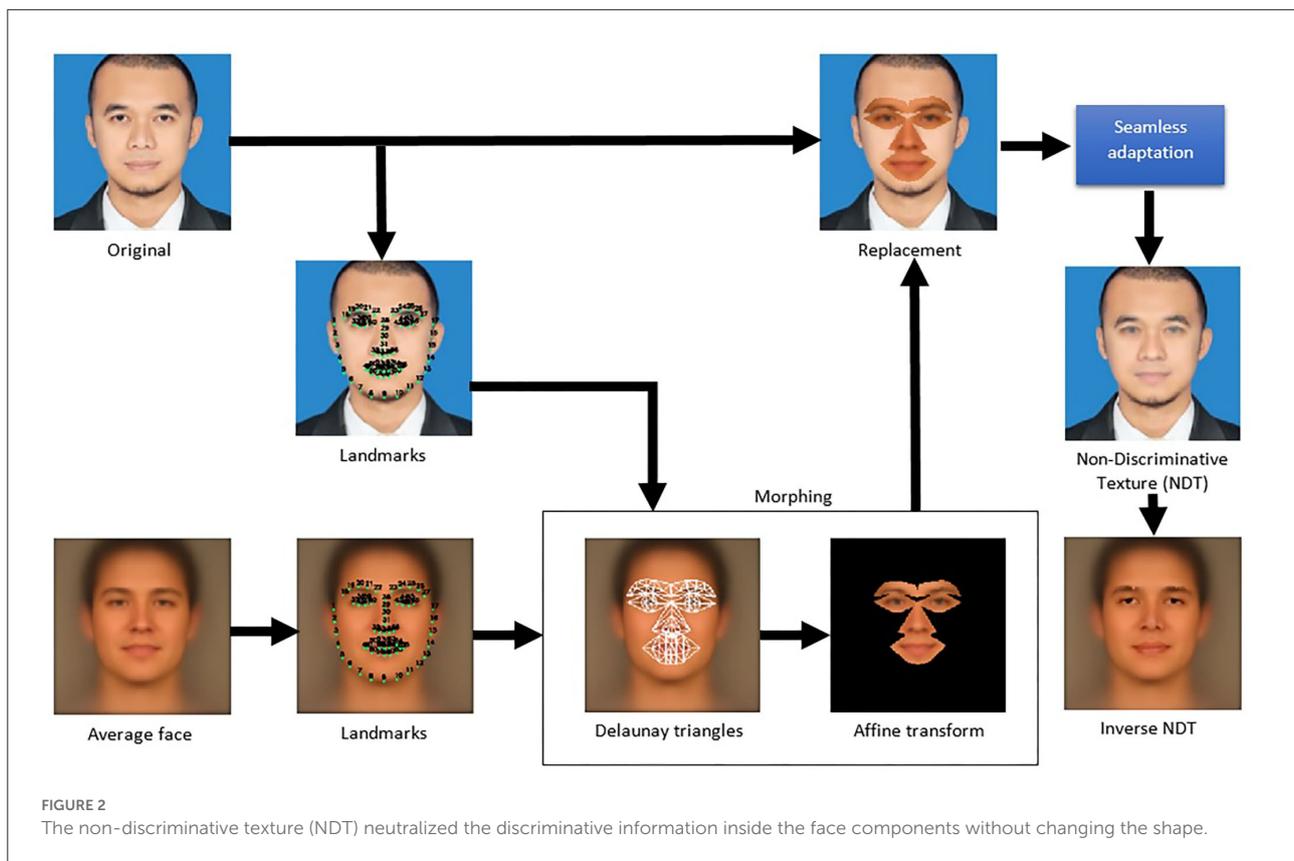
1. Extract the face landmarks.
2. Morph the average face geometry to the original face geometry.

Experimental results

Dataset and preprocessing

We used 2,000 samples from the FRGCv2 dataset (Phillips et al., 2005) containing controlled images, 20 images per person, 100 individuals, and registered to a 60% scale which included the amount of background. The cropped face images were resized to 150×150 pixels. Moreover, the dataset also contained facial expressions.

We also ensured that all images are aligned by DLIB alignment before applying average face calculation and morphing experiments. The alignment was based on five face landmarks that are a part of 68 DLIB face landmarks. These five points are the four points of the left and right eye corners and the nose peak. Before feature extraction, the face recognition systems detect and crop the face automatically. DLIB provides two face detection methods: HOG+Linear SVM (Dalal and Triggs, 2005) and MMOD CNN (King, 2015). For the experiments, we used the HOG+Linear SVM method to detect faces and the DLIB shape predictor to extract the 68 face landmarks. On the other hand, FaceNet used MTCNN (Zhang et al., 2016) and ArcFace used SSD-face (Liu et al., 2016) to detect the face. Both of them are based on a Convolutional Neural Network (CNN) to detect multiple faces in the image.



Non-discriminative texture

Based on the single-part analysis, NDT only changed the small area of the face component, especially, the inner detail of the face component. Figure 5 showed that the dataset with texture neutralization on a single face component produced a slight difference in each one's performance that it was difficult to see which one was the most important or less important. But, the graph still indicated that the texture of eyebrows and nose was the most important component for DLIB and FaceNet. They produced the highest increase of error in performance compared to the whole face. The eyebrows texture also had an important role in ArcFace and Facevacs. On the other hand, the mouth and eyes texture was a noise on FaceNet. They produced lower errors than the original face. It indicated that FaceNet was more sensitive to facial expression, for example, eyes blink, smile, or laugh.

The multi-part analysis supported the single-part analysis and showed the order of face part contribution. The order of the highest contribution on DLIB was eyebrows-eyes-nose-mouth, eyebrows-nose-mouth, and eyebrows-eyes-nose, respectively. However, the order was difficult to be defined after the eyebrows-eyes-nose line. They were overlapping and crossing each other. Overall, the combination of eyebrows and nose produced the highest increase of error on DLIB.

Based on the response of FaceNet, the multi-part analysis produced different behavior compared to DLIB, ArcFace, and Facevacs. The highest contribution was eyebrows-nose, eyebrows-eyes-nose-mouth, eyebrows-eyes-nose, and then other combinations. Although there was a crossing-overlapping line in FaceNet results, we still can see that the eyebrows-nose was the most important part for FaceNet. On the other hand, the eyes and mouth still contained discriminative information, but they contributed to the FaceNet recognition system as a noise.

For the ArcFace multi-part observation, the eyebrows-eyes-nose-mouth and eyebrows-eyes-nose produced the highest contribution. The other face parts combination produced a similar error and overlapping-crossing each other. Overall, the eyebrows and nose were a meaningful part of ArcFace.

Based on the Facevacs results, the lines were less overlapping and crossing each other. The order of multi-part contribution was eyebrows-eyes-nose-mouth, eyebrows-eyes-nose, eyebrows-nose-mouth, eyebrows-eyes-mouth, eyes-nose-mouth, followed by eyebrows-eyes, eyebrows-nose, and eyebrows-mouth. The order indicated that eyebrows contained the most discriminative information for Facevacs. Moreover, the combination of eyes-nose, eyes-mouth, and nose-mouth contained less discrimination information than eyebrows only.

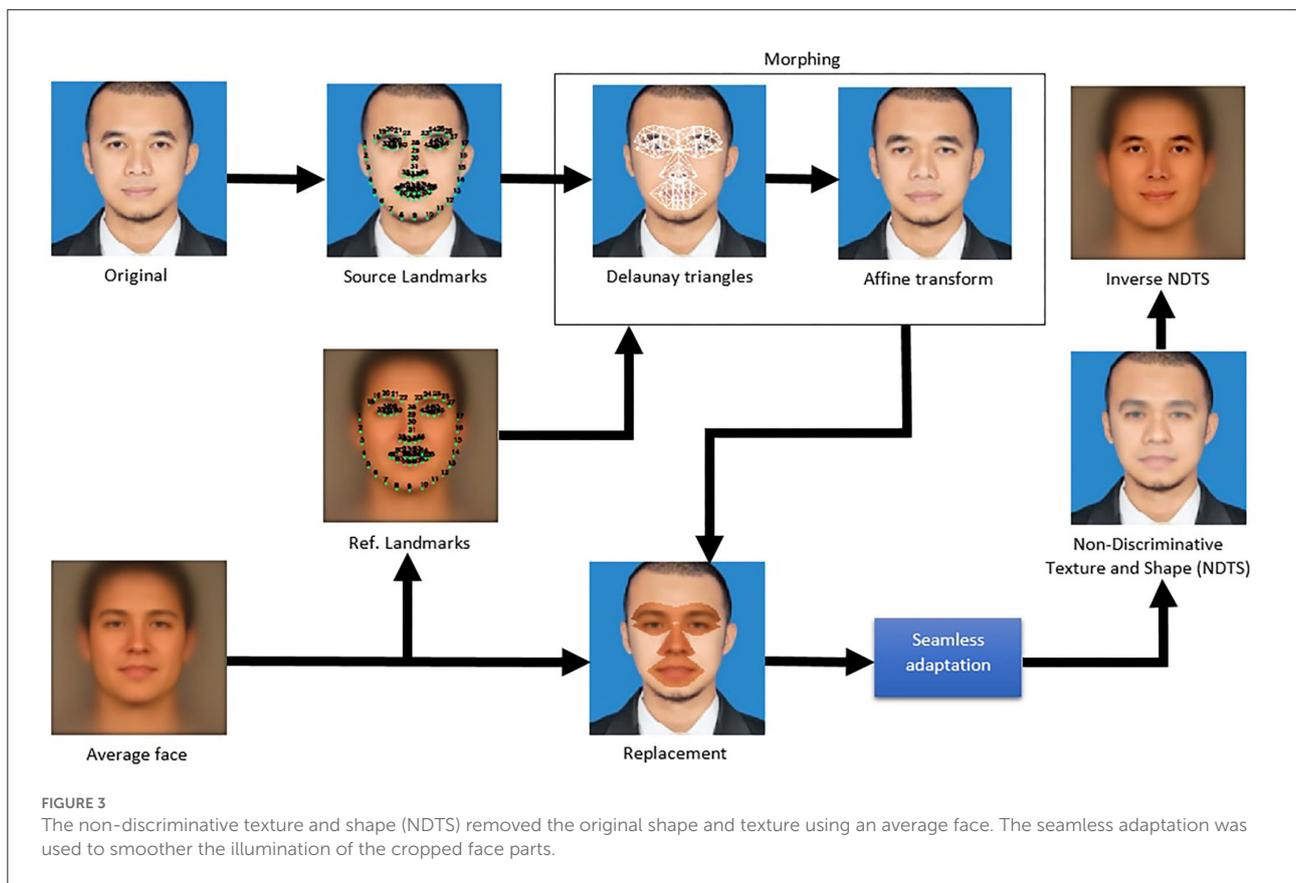


Figure 6 showed the behavior comparison in FNMR@0.01%FMR between four deep FRs. Each FR system produced different behavior. ArcFace and Facevacs were less sensitive to texture neutralization or texture removal. On the contrary, DLIB and FaceNet responded to the NDT well. The face parts played a similar role in all FR, except the mouth. The reason was the dataset contained facial expressions. Overall, the eyebrows' texture is an important component of all FR. The NDT experiment also indicated that the remaining face component still has a big contribution to recognition, for example, face geometry, face periphery, jaw, chin, forehead, hair, ear, and background.

Inverse non-discriminative texture

Vice versa with the NDT experiment, the iNDT observed the discrimination information in the face by neutralizing the whole face except the certain area in the face parts, namely, eyebrows, eyes, nose, and mouth. In other words, the iNDT experiment remained the original face component on the average face. We observed the performance of FR systems using the DET curve. Figure 7 shows the behavior of all FR systems. In contrast to the

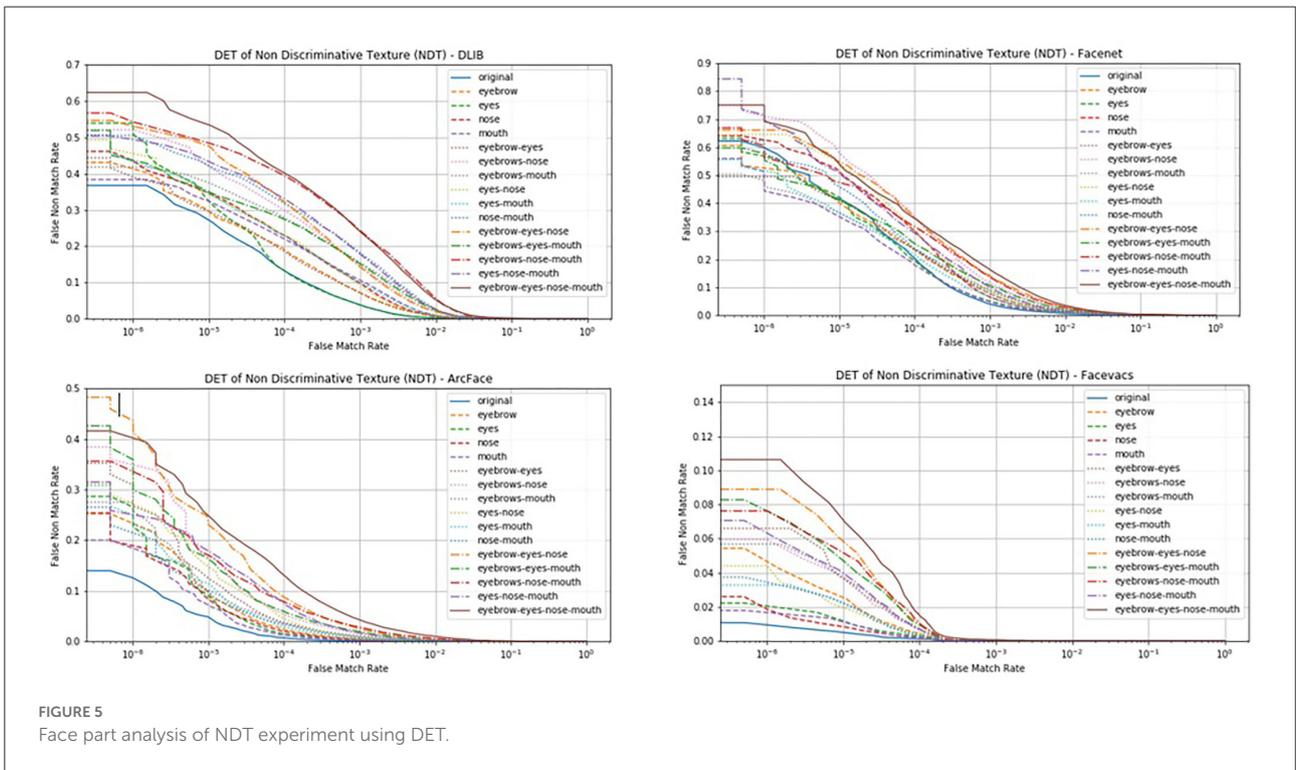
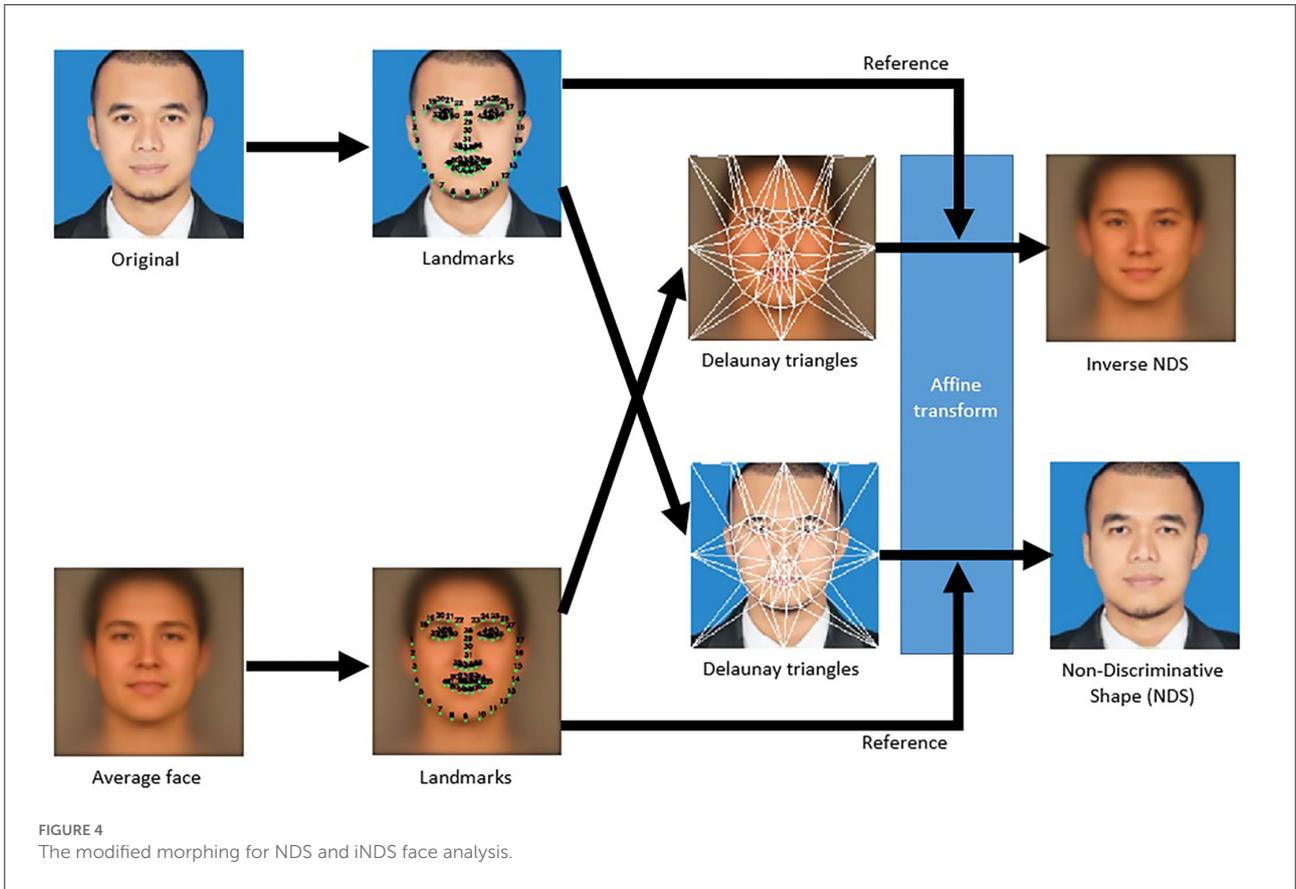
NDT experiment, the lower the error the higher the contribution of the face part.

Based on the DLIB response to the iNDT, eyebrows and nose texture played an important role. Both the individual part and multi-part analysis showed a lower error on the combination containing eyebrows and nose. The eyebrows-nose also played a more important role for DLIB than eyebrows-eyes-mouth and eyes-nose-mouth.

The different behavior on iNDT was shown by FaceNet. The eyebrows-eyes-nose-mouth performance was similar to eyebrows-eyes-nose. It indicated that the mouth was less important than others. The face parts containing eyebrows also produced a lower error than the face parts without eyebrows. Thus, the eyebrows played an important role in the FaceNet recognition system.

Based on ArcFace response to the iNDT, the results were similar to DLIB. The eyebrows produced a lower error than others. The combination containing eyebrows also produced lower error than the face part combination without eyebrows. Thus, the eyebrows texture contained important discriminative information for ArcFace, followed by nose, eyes, and mouth, respectively.

The commercial FR system Facevacs also had a similar response to ArcFace and DLIB. However, the eyebrows-nose



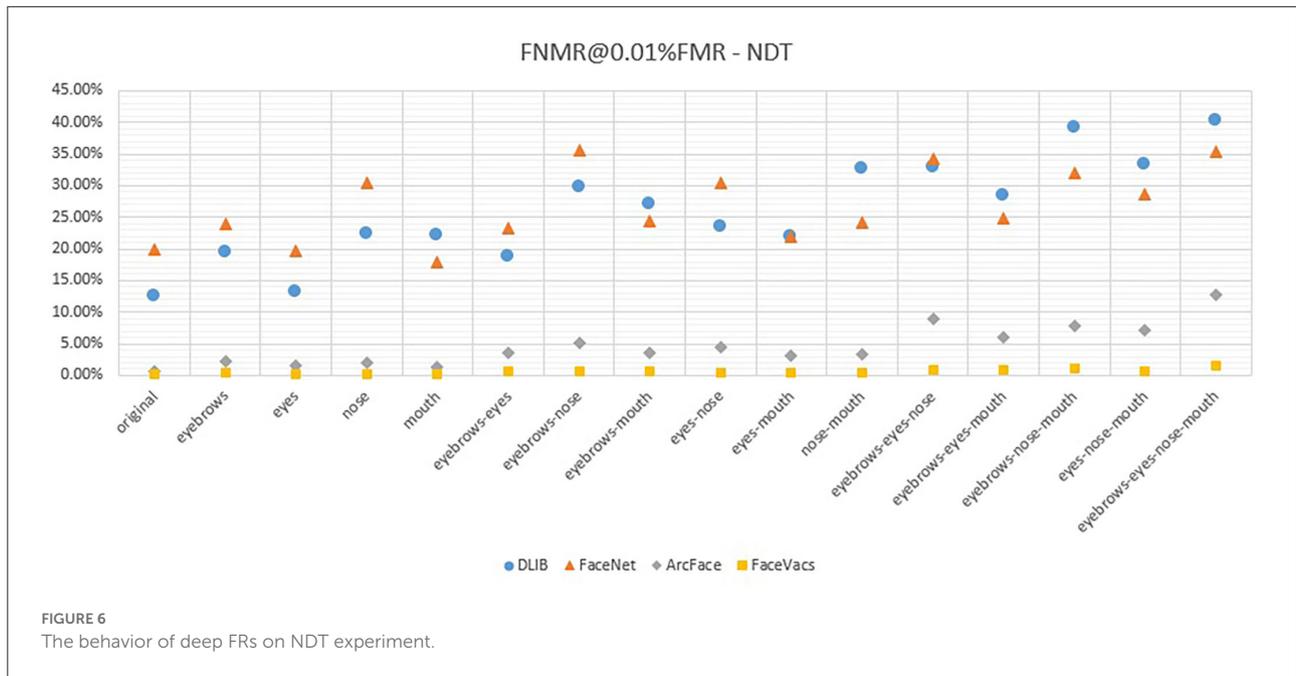


FIGURE 6 The behavior of deep FRs on NDT experiment.

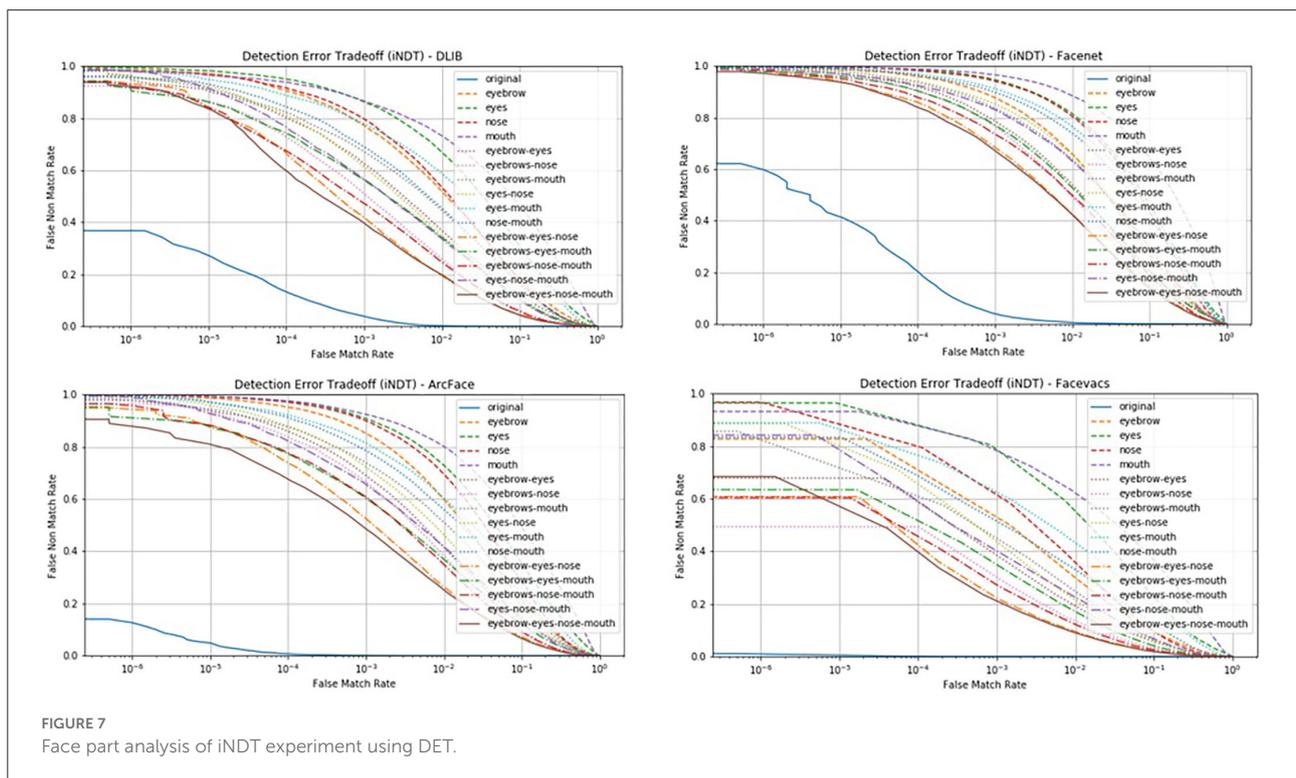
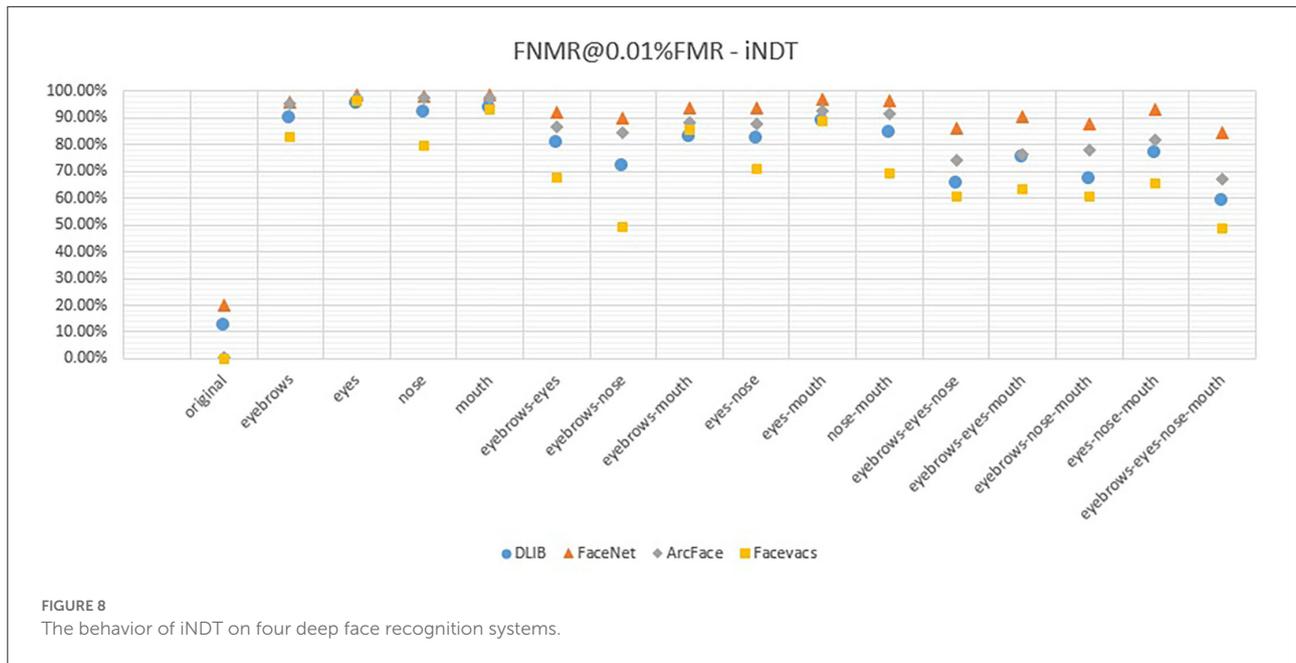


FIGURE 7 Face part analysis of iNDT experiment using DET.

produced a better performance than eyebrows-eyes-mouth and eyes-nose-mouth. It indicated that eyebrows and nose texture contain strong discriminative information. The individual results showed similar behavior, the eyebrows, followed by the nose, played an important role in Facevacs.

The comparison of the iNDT response of four deep FRs is shown in Figure 8. We presented the behavior at FNMR@0.01%FMR. Based on the graph, the joining of eyebrows and nose texture produced the best performance on all FRs. However, the Facevacs produced a good recognition on



eyebrows-nose only, followed by DLIB and ArcFace. In contrast, FaceNet produced the highest FNMR and did not work well on individual face parts or multi-part texture recognition.

Non-discriminative texture and shape

The NDTs experiment removed the discrimination information in the face part, especially the face part geometry/shape and texture. The shape and pattern removal completely excluded a certain face part using the average. We observed how the deep face recognition systems responded to the face without one or multi-parts. Figure 9 presented the DET of face part removal analysis using the NDTs method. The graph pointed out that the higher increase of error, the higher the face part contribution to recognition.

Based on the DLIB graph, the individual face part analysis showed that the nose shape and texture neutralization produced the highest error, followed by mouth, eyebrows, and eyes, respectively. Furthermore, the multi-parts analysis showed the order of contribution, namely, eyebrows-eyes-nose-mouth, eyebrows-nose-mouth, eyes-nose-mouth, eyebrows-eyes-nose, eyebrows-eyes-mouth, and so forth. For the two-part analysis, nose-mouth produced the highest error, but the eyebrows-nose overcame at $FMR = 10^{-4}$. It indicated that the nose shape and texture contained the most important discrimination information for DLIB.

Next, the FaceNet graph showed different behavior compared to others. The mouth and eyes-mouth produced the error less than the original image. It indicated that the mouth was still a noise for FaceNet. Moreover, eyebrows and

nose played an important role for FaceNet. The face part combination containing eyebrows and nose produced a higher error than others.

Based on the ArcFace graph, the most important individual face part was the eyebrows, followed by the nose, eyes, and mouth, respectively. In addition, the multi-parts graph showed the order as follows: eyebrows-eyes-nose-mouth, eyebrows-eyes-nose, eyebrows-nose-mouth, eyebrows-nose, eyebrows-eyes-mouth, and so forth. That order indicated that eyebrows and nose played an important role in the ArcFace recognition system, then followed by other face parts. Moreover, the eyebrows-nose produced a higher error than eyebrows-eyes-mouth and eyes-nose-mouth, which means eyebrows-nose contained higher discriminative information than eyebrows-eyes-mouth and eyes-nose-mouth.

Based on the Facevac results, the eyebrows were still the most important part, followed by the nose, eyes, and mouth, respectively. The behavior was similar to ArcFace. However, the multi-parts combination produced a small difference of error, for example, eyebrows-eyes-nose, eyebrows-eyes-mouth, and eyebrows-nose-mouth that overlap and cross each other. But they still had a higher contribution than eyes-nose-mouth. The eyebrows-eyes, eyebrows-nose, and eyebrows-mouth also produced a higher contribution than eyes-nose-mouth. So, the shape and texture of eyebrows was the most valuable part for Facevac.

Overall, the NDTs experiment pointed out that the shape and texture of the nose were the most important part of DLIB. On the other hand, the shape and texture of eyebrows were the most important part of FaceNet, ArcFace, and Facevac.

Figure 10 showed the behavior comparison of NDTs in FNMR@0.01%FMR between four deep FRs. Each FR system produced different behavior. The Facevacs was less sensitive to shape and texture removal. Although the eyebrows, eyes, nose, and mouth were neutralized, Facevacs still worked well

to recognize the dataset. On the contrary, ArcFace, DLIB, and FaceNet were more sensitive to the NDTs experiment. ArcFace and FaceNet had similar behavior. On the other hand, DLIB produced different behavior. Mouth shape and texture became more important for DLIB. Similar to the NDT experiment,

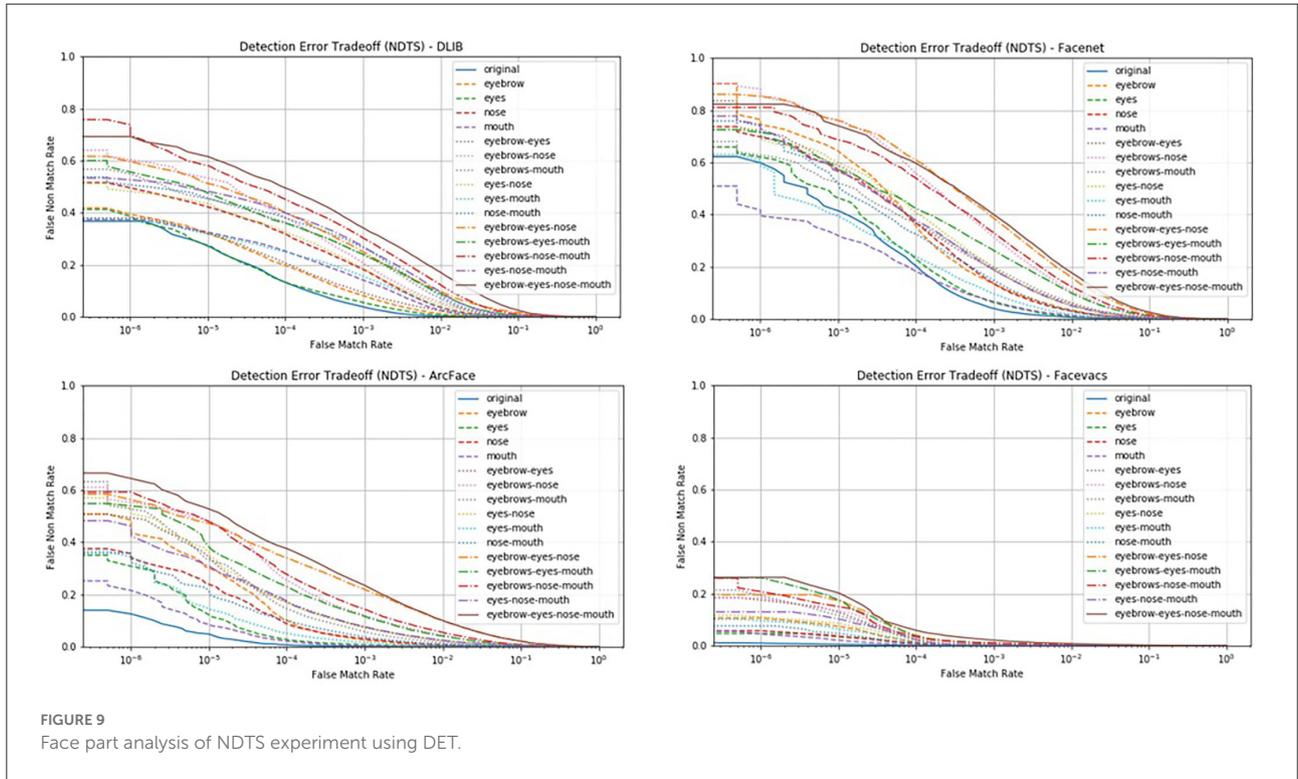


FIGURE 9 Face part analysis of NDTs experiment using DET.

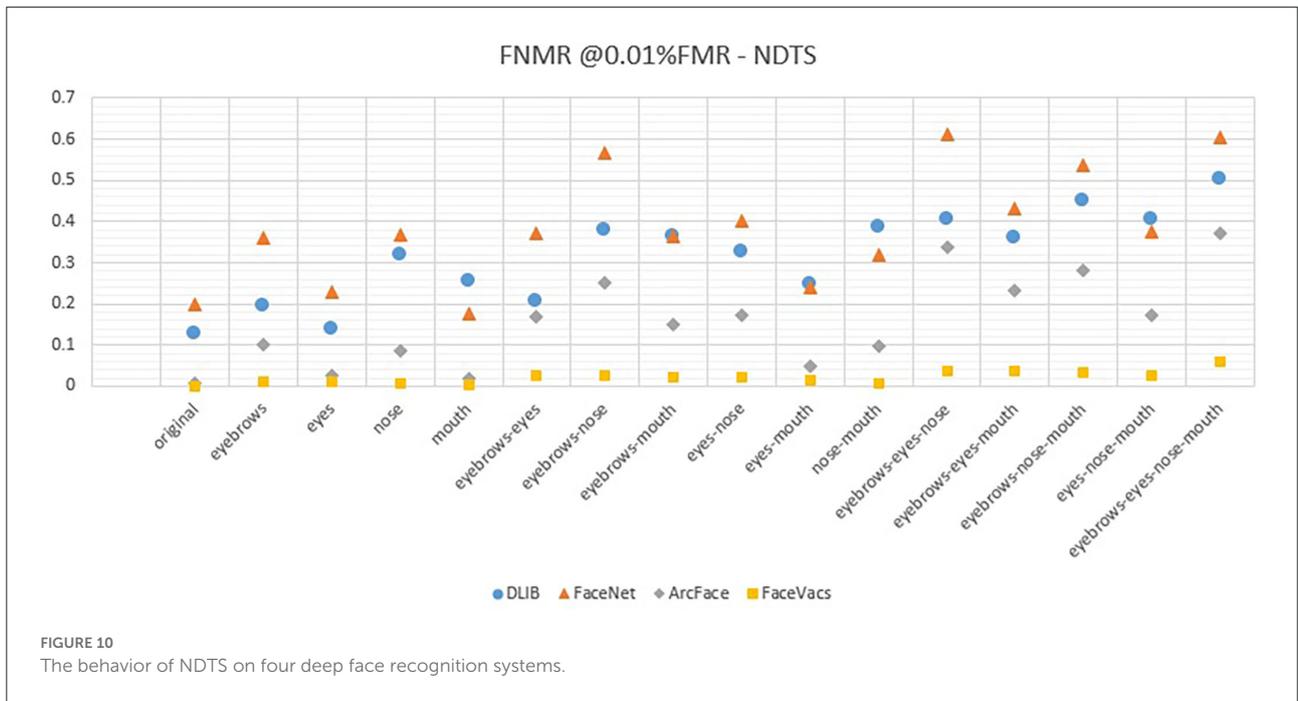


FIGURE 10 The behavior of NDTs on four deep face recognition systems.

the NDTs experiment also indicated that the remaining face component still has a big contribution to recognition.

Inverse non-discriminative texture and shape

Then iNDTS remained the original face component on the average face. The difference between iNDT and iNDTS was the iNDT remained only the texture of the face part, and iNDTS remained both the shape and texture of the face part. The response of deep FRs to the iNDTS is shown in Figure 11. For the inverse experiments, the lower the error, the higher the contribution of face parts.

The DLIB produced the lowest error on eyebrows-nose and eyebrows-eyes-nose. The next lowest error were eyebrows-eyes-nose-mouth, eyebrows-nose-mouth, eyebrows, eyebrows-eyes, respectively. That order means that the most important discrimination information was on the eyebrows and nose. In addition, the eyebrows-nose had a better performance than eyebrows-eyes-nose and eyebrows-eyes-nose-mouth. The indication was the sensitivity of DLIB to facial expression in the mouth and eyes because the iNDTS experiment removes the

whole discrimination information in the face except the certain face part.

FaceNet produced a similar behavior to DLIB. The eyebrows were the most important face part. Furthermore, the eyebrows-eyes-nose produced a better performance than eyebrows-eyes-nose-mouth. The behavior indicated that the mouth affected the FaceNet as a noise. It means that FaceNet was sensitive to facial expressions.

Based on the ArcFace results, the eyebrows played the most important role, followed by the eyes, nose, and mouth, respectively. Moreover, the order of multi-parts performance indicated that ArcFace was more robust to facial expression. The shape and texture of the mouth still played a role in recognition.

Next, based on the Facevacs results, the eyebrows were the most important part for recognition. Even, the standalone eyebrows produced a better performance than eyebrows-eyes-mouth, eyebrows-mouth, eyes-nose, eyes-mouth, and nose-mouth. The eyebrows-eyes-nose also produced a better performance than eyebrows-eyes-nose-mouth. The behavior indicated that Facevacs was also sensitive to facial expression.

Figure 12 shows the comparison of deep FRs response to iNDTS experiment. Based on the graph, all FRs produced similar behavior at FNMR@0.01%FMR. The eyebrows played the most important role in recognition. On the contrary, the mouth increased the FNMR on both individual parts and multi-parts.

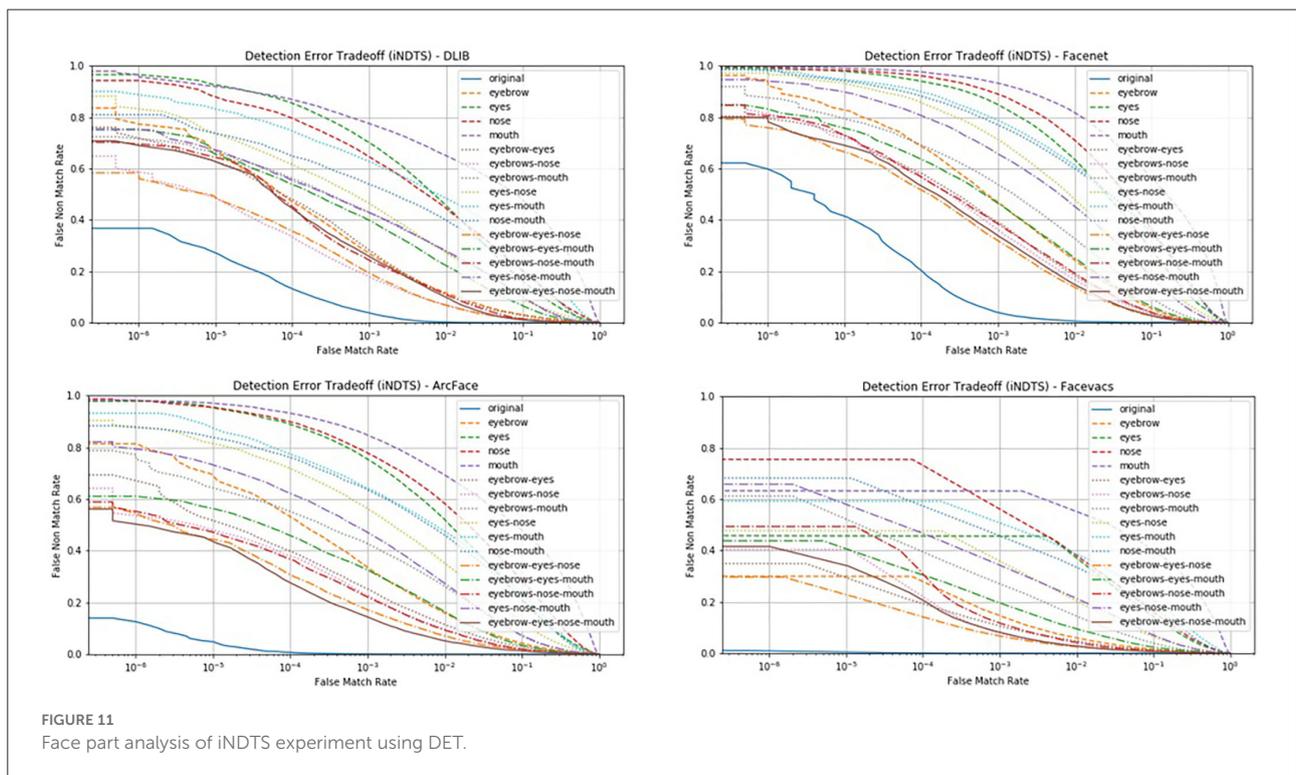


FIGURE 11 Face part analysis of iNDTS experiment using DET.

Non-discriminative shape and inverse non-discriminative shape

The NDS experiment morphed the face geometry/shape to an average face shape. On the other hand, the iNDS experiment morphed the average face geometry into the original face shape. The results of the NDS and iNDS experiments are shown in

Figure 13. The graph showed that the texture removal produced a higher FNMR than shape removal on four deep FRs. Thus, the texture played an important role in recognition. Furthermore, Figure 14 showed the impact of the shape and texture of the face part on recognition. The shape of eyebrows affected the recognition performance higher than others. The shape of eyebrows decreased the FNMR of iNDS significantly on both

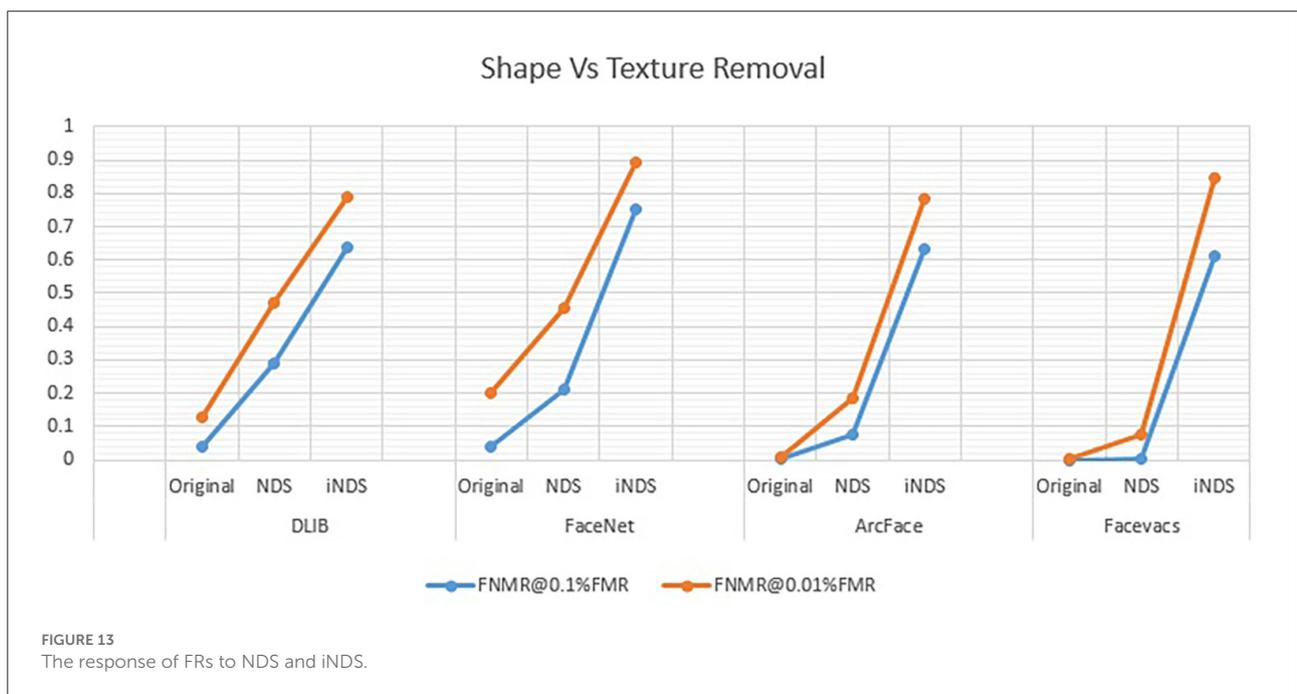
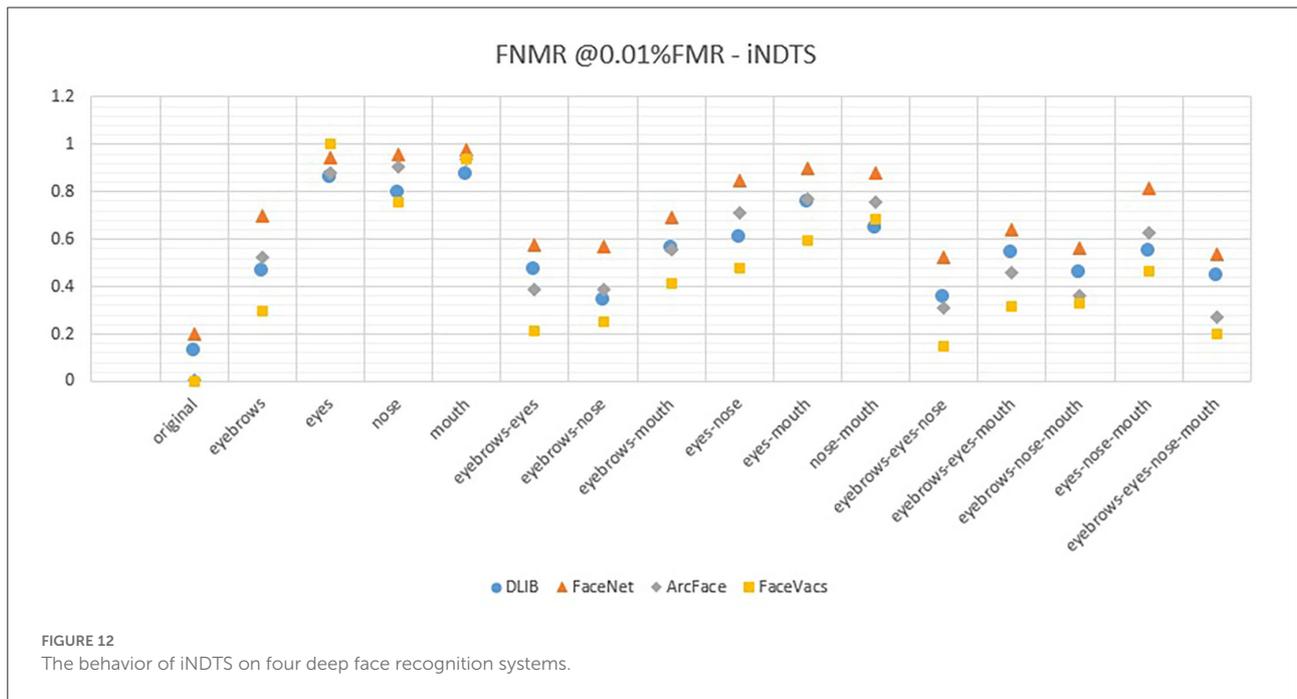




FIGURE 14 The impact of texture and shape of face parts.

individual parts and multi-parts. On the other hand, the shape of the eyes, nose, and mouth was less important to recognition.

Discussion and conclusions

This work presented an advanced way of face analysis using the morphing and average face to address the lack of knowledge on what parts of the face are important, especially on deep face recognition systems. The lacks were the primitive ways to mask the face parts produce artifacts, most of the attention is on the eyes and eyebrows, and no extensive analysis in Deep Learning Face Recognition systems. We introduced six advanced ways to neutralize the discrimination information in the specific area of the face, namely, NDT, iNDT, NDTs, iNDTs, NDS, and iNDS. The face parts selection was done carefully by morphing without introducing artifacts.

Based on the experimental results, the deep FRs produced different behaviors in each experiment. The NDT and NDTs removed the discrimination information in certain face parts. The NDT pointed out that the eyebrows texture contained the most discriminative information for recognition. On the other hand, the facial expression in the dataset affected the FaceNet that threatened the mouth texture as a noise. The NDT

experiment also indicated that the remaining face component still has a big contribution to recognition, for example, face geometry, face periphery, jaw, chin, forehead, hair, ear, and background. Furthermore, the NDTs showed that the shape and texture of the nose were important to DLIB. For FaceNet, ArcFace, and FaceVacs, the eyebrows' shape and texture still played the most important part in recognition.

The inverse of NDT and NDTs, that is, iNDT and iNDTs, observed the role of certain face parts only without the remaining face components. They neutralized the whole face using the average face and replace a certain face part using the original face parts. The iNDT showed that the texture of eyebrows played an important role in recognition. Moreover, the iNDTs pointed out that the texture and shape of eyebrows contained the most discriminative information for recognition. The iNDTs also showed that the ArcFace was more robust to facial expression.

Furthermore, the contribution of texture and shape to recognition was analyzed by NDS and iNDS experiments. The iNDT and iNDTs also strengthened the role of texture and shape in recognition. Based on the experimental results, the texture was more important than the shape. However, the shape of the eyebrows still contained significant discriminative information for recognition.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

Ethics statement

Ethical review and approval was not required for the current study in accordance with the local legislation and institutional requirements. Written informed consent was obtained from the individual(s) for the publication of any identifiable images or data included in this article.

Author contributions

NL contributed to drafting the work or revising it critically for important intellectual content, the acquisition, analysis, or interpretation of data for the work. LS and RV contributed substantial contributions to the conception or design of the work and provide approval for publication of the content. All authors contributed to the article and approved the submitted version.

References

- Ahonen, T., Hadid, A., and Pietikainen, M. (2004). "Face recognition with local binary patterns," in *Computer Vision - ECCV 2004. Lecture Notes in Computer Science, Vol. 3021*, eds T. Pajdla and J. Matas (Berlin; Heidelberg: Springer), 469–481. doi: 10.1007/978-3-540-24670-1_36
- Caifeng, S., Shaogang, G., and McOwan, W. P. (2005). Robust facial expression recognition using local binary patterns. *IEEE Int. Conf. Image Process.* 2, 370–373. doi: 10.1109/ICIP.2005.1530069
- Dalal, N., and Triggs, B. (2005). "Histograms of oriented gradients for human detection," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (San Diego, CA: IEEE), 886–893. doi: 10.1109/CVPR.2005.177
- Damer, N., Grebe, J. H., Chen, C., Boutros, F., Kirchbuchner, F., and Kuijper, A. (2020). "The effect of wearing a mask on face recognition performance: an exploratory study," in *BIOSIG 2020 - Proceedings of the 19th International Conference of the Biometrics Special Interest Group*, eds A. Brömme, C. Busch, A. Dantcheva, K. Raja, C. Rathgeb, and A. Uhl (Bonn: Gesellschaft für Informatik e.V.), S. 1–10.
- Deng, J., Guo, J., Xue, N., and Zafeiriou, S. (2019). "ArcFace: additive angular margin loss for deep face recognition," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (Long Beach, CA: IEEE), 4685–4694. doi: 10.1109/CVPR.2019.00482
- FaceVACS Engine 9.4.0. (2020). *Commercial-Off-The-Shelf (COTS) FRS (Cognitec FaceVACS-SDK Version 9.4.0)*. Available online at: <http://www.cognitec.com/technology.html> (accessed October 1, 2020).
- García-Ramírez, J., Olvera-López, J., Pineda, I., and Martín-Ortiz, M. (2018). Mouth and eyebrow segmentation for emotion recognition using interpolated polynomials. *J. Intell. Fuzzy Syst.* 34, 1–13. doi: 10.3233/JIFS-169496
- Hofer, P., Roland, M., Schwarz, P., Schwaighofer, M., and Mayrhofer, R. (2021). "Importance of different facial parts for face detection networks," in *IEEE International Workshop on Biometrics and Forensics* (Rome: IEEE), 1–6. doi: 10.1109/IWBF50991.2021.9465087
- Juefei-Xu, F., and Savvides, M. (2011). "Can your eyebrows tell me who you are?," in *5th International Conference on Signal Processing and Communication Systems* (Honolulu, HI), 1–8. doi: 10.1109/ICSPCS.2011.6140879
- King, D. E. (2009). Dlib-ml: a machine learning toolkit. *J. Mach. Learn. Res.* 10, 1755–1758.
- King, D. E. (2015). *Max-Margin Object Detection*. Available online at: <http://arxiv.org/abs/1502.00046>
- Lestriandoko, N. H., Spreewers, L. J., and Veldhuis, R. N. J. (2019). "Multi-resolution face recognition: the behaviors of local binary pattern at different frequency bands," in *Proceedings of Symposium on Information Theory and Signal Processing in the Benelux*, (Leuven), 63–70.
- Lin, S., Yue, Y., and Zhu, X. (2021). "Emotion recognition using representative geometric feature mask based on CNN," in *IEEE 4th International Conference on Information Systems and Computer Aided Education* (Dalian: IEEE), 257–261. doi: 10.1109/ICISCAE52414.2021.9590797
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., et al. (2016). "SSD: single shot multibox detector," in *Computer Vision - ECCV 2016. ECCV 2016. Lecture Notes in Computer Science, Vol 9905*, eds B. Leibe, J. Matas, N. Sebe, and M. Welling (Cham: Springer).
- Loderer, M., Pavlovičová, J., Oravec, M., and Mazanec, J. (2015). "Face parts importance in face and expression recognition," in *International Conference on Systems, Signals and Image Processing* (London), 188–191. doi: 10.1109/IWSSIP.2015.7314208
- Malakar, S., Chiracharit, W., Chamnongthai, K., and Charoenpong, T. (2021). "Masked face recognition using principal component analysis and deep learning," in *18th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology* (Chiang Mai: IEEE), 785–788. doi: 10.1109/ECTI-CON51831.2021.9454857
- Midtun, J., Tessem, B., Karlsen, S., and Nyre, L. (2017). "Realistic face manipulation by morphing with average faces," in *Norsk Informatikkonferanse* (Oslo).

Funding

The research described in this article was supported by the Research and Innovation in Science and Technology Project (RISET-Pro) of the Ministry of Research, Technology, and Higher Education of the Republic of Indonesia (World Bank Loan No.8245-ID).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor MF declared a past co-authorship with the authors RV and LS.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Mohammad, A.S., Rattani, A., and Derakhshani, R. (2019). Eyebrows and eyeglasses as soft biometrics using deep learning. *IET Biometr.* 8, 378–390. doi: 10.1049/iet-bmt.2018.5230
- Nikisins, O., and Greitans, M. (2012). “A mini-batch discriminative feature weighting algorithm for LBP - Based face recognition,” in *International Conference on Imaging Systems and Techniques* (Manchester), 170–175. doi: 10.1109/IST.2012.6295521
- Ozcel, I., Yolcu, G., Öz, C., Kazan, S., and Bunyak, F. (2018). iFER: facial expression recognition using automatically selected geometric eye and eyebrow features. *J. Electron. Imag.* 27, 023003. doi: 10.1117/1.JEL.27.2.023003
- Pérez, P., Gangnet, M., and Blake, A. (2003). Poisson image editing. *ACM Trans. Graph.* 22, 313–318. doi: 10.1145/882262.882269
- Phillips, P. J., Flynn, P. J., Scruggs, T., Bowyer, K. W., Chang, J., Hoffman, K., et al. (2005). “Overview of the face recognition grand challenge,” in *IEEE Conf. Comp Vision and Pattern Recognition* (San Diego, CA: IEEE), 947–954. doi: 10.1109/CVPR.2005.268
- Qin, L., Peng, F., Venkatesh, S., Ramachandra, R., Long, M., and Busch, C. (2021). Low visual distortion and robust morphing attacks based on partial face image manipulation. *IEEE Transact. Biometr. Behavi. Identity Sci.*, 72–88. doi: 10.1109/TBIOM.2020.3022007
- Radji, N., Cherif, D., and Azrar, A. (2014). “Effect of eyes and eyebrows on face recognition system performance,” in *International Image Processing, Applications and Systems Conference (Sfax)*, 1–7. doi: 10.1109/IPAS.2014.7043297
- Radji, N., Cherif, D., and Azrar, A. (2015). “Importance of eyes and eyebrows for face recognition system,” in *3rd International Conference on Control, Engineering and Information Technology (Tlemcen)*, 1–6. doi: 10.1109/CEIT.2015.7233088
- Sadr, J., Jarudi, I., and Sinha, P. (2003). The role of eyebrows in face recognition. *Perception* 32, 285–293. doi: 10.1068/p5027
- Schroff, F., Kalenichenko, D., and Philbin, J. (2015). “FaceNet: A unified embedding for face recognition and clustering,” in *IEEE Conference on Computer Vision and Pattern Recognition* (Boston, MA: IEEE), 815–823. doi: 10.1109/CVPR.2015.7298682
- Spreeuwens, L. J., Veldhuis, R. N. J., Sultanali, S., and Diephuis, J. (2014). “Fixed FAR vote fusion of regional facial classifiers,” in *International Conference of the Biometrics Special Interest Group* (Darmstadt), 1–4.
- Tome, P., Blázquez, L., Vera-Rodríguez, R., Fierrez, J., Ortega-García, J., Expósito, N., et al. (2013). “Understanding the discrimination power of facial regions in forensic casework,” in *International Workshop on Biometrics and Forensics* (Lisbon), 1–4. doi: 10.1109/IWBF.2013.6547306
- Trigueros, D. S., Meng, L., and Hartnett, M. (2018). Enhancing convolutional neural networks for face recognition with occlusion maps and batch triplet loss. *Image Vis. Comput.* 79, 99–108. doi: 10.1016/j.imavis.2018.09.011
- Zeng, D., Veldhuis, R., and Spreeuwens, L. (2021). A survey of face recognition techniques under occlusion. *IET Biom.* 10, 581–606. doi: 10.1049/bme2.12029
- Zhang, K., Zhang, Z., Li, Z., and Qiao, Y. (2016). Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Process. Lett.* 23, 1499–1503. doi: 10.1109/LSP.2016.2603342
- Zhang, Z., Ji, X., Cui, X., and Ma, J. (2020). “A survey on occluded face recognition,” in *The 9th International Conference on Networks, Communication and Computing (ICNCC 2020), December 18–20, 2020* (Tokyo; New York, NY: ACM), 10.
- Zhao, Y., Wang, L., Tan, M., Yan, X., Zhang, X., and Feng, H. (2021). “Face recognition with partial occlusion based on attention mechanism,” in *International Conference on Electronic Information Engineering and Computer Science* (Changchun), 562–566. doi: 10.1109/EIECSS3707.2021.9587947