Diagnosis and Screening of Velocardiofacial Syndrome by Evaluating Facial Photographs Using a Deep Learning-Based Algorithm

Rong-Min Baek
Seoul National University Bundang Hospital
Anna Cho
Seoul National University Bundang Hospital
Yoon Gi Chung
Seoul National University Bundang Hospital
Yonghoon Jeon
ezCaretech Co. Ltd
Hunmin Kim
Seoul National University Bundang Hospital
Hee Hwang
Kakao Healthcare, Pangyo, Gyeonggi-do, South Korea
Yujin Myung (surgene@snu.ac.kr)
Seoul National University Bundang Hospital

Article

Keywords:

Posted Date: March 6th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2601333/v1

Additional Declarations: No competing interests reported.

Abstract

Purpose: The early detection and diagnosis of rare genetic diseases are crucial for maintaining patient health and well-being. However, the diagnosis of said diseases can be challenging owing to their rarity, limited clinical experience of physicians, and specialized facilities required for diagnosis. In recent years, deep learning algorithms have been investigated as a potential strategy for the efficient and accurate diagnosis of these diseases. Herein, we used the deep learning algorithm of multitask cascaded convolutional neural networks (MTCNN) to develop a face recognition model for the diagnosis of velocardiofacial syndrome (VCFS).

Methods: We trained the model on a publicly available labeled face dataset and evaluated its performance. Subsequently, we analyzed the binary classification performance of diagnosing VCFS using the most efficient face recognition model. The facial images of 98 patients with VCFS and 91 non-VCFS controls who visited Seoul National University were used to create training and test sets. Moreover, we analyzed whether the classification results matched the known facial phenotype of patients with VCFS.

Results: The facial recognition model showed high accuracy, ranging from 94.03% to 99.78%, depending on the training dataset. In contrast, the accuracy of the binary classification model for the diagnosis of VCFS varied from 80.82% to 88.02% when evaluating with photographs taken at various angles of the patient depending on the structure. When only evaluating frontal photographs, the accuracy was 95.00%. Moreover, the importance level analyzed through the gradient-weighted class activation mapping heatmap showed the characteristic parts of perinasal and periorbital area to be consistent with the conventional facial phenotypes of VCFS.

Conclusion: We attempted to diagnose the patients' genetic syndrome using MTCNN-based deep learning algorithms only with the photos of the faces of patients with VCFS. We obtained high accuracy, and deep learning-based diagnosis has been conformed to be greatly helpful for medical staff in the early detection and diagnosis of children with rare genetic diseases, enabling them to provide treatment in a timely manner.

Introduction

Velocardiofacial Syndrome (VCFS) is a genetic disorder characterized by a complex phenotype, which includes cleft palate or velopharyngeal insufficiency, cardiac anomalies, characteristic facial features, and intellectual disabilities [1, 2]. It is caused by a microdeletion of chromosome region 22q11.2, occurring mostly as a de novo event [3]. VCFS is the most frequent chromosomal microdeletion syndrome, with an estimated incidence of 1 in 3,000–6,000 live births [4].

Current research has established that the size of the deletion in VCFS can result in various clinical presentations, including cardiac anomalies, palatal abnormalities, immunological abnormalities, endocrine problems, genitourinary and gastrointestinal complications, and multiple neurodevelopmental disorders [3]. Owing to its phenotypic variability, VCFS has various names, including DiGeorge Sequence,

CATCH 22 Syndrome, Conotruncal Anomaly Face Syndrome, and 22q11.2 Microdeletion Syndrome, all of which refer to the same genetic condition.

Owing to the wide range of clinical phenotypes associated with VCFS, it is crucial to diagnose and detect patients early on to minimize long-term health problems. Methods for the genomic testing of VCFS include chromosomal microarray (CMA) and targeted deletion analysis techniques, including fluorescence in situ hybridization (FISH), quantitative polymerase chain reaction, and multiplex ligation-dependent probe amplification [5, 6]. While all of these tests have high sensitivity, it is important for clinicians to suspect VCFS and perform the necessary corresponding tests. In particular, when severe symptoms, such as complex cardiac anomalies or convulsions due to hypocalcemia, do not present in the neonatal period, diagnosis may be delayed until childhood or adolescence, despite the presence of various clinical phenotypes.

Currently, only few studies have utilized computerized diagnostic support for the screening and diagnosis of rare genetic syndromes. One notable example is the Face2Gene study [7], which used a cascaded deep convolutional neural network (CNN)-based method to detect Noonan syndrome. Furthermore, one multi-national study that developed a point-of-care screening tool achieved a diagnostic accuracy of 88% [8]. As technology continues to rapidly improve, the development of computer-based screening tools to aid in the diagnosis of rare genetic diseases is expected to show significant progress. In this study, the authors developed a facial detection model using only a frontal photograph of the patient through multitask cascaded convolutional neural networks (MTCNN) and applied it to the early diagnosis and detection of VCFS.

Methods

Dataset

This study was approved by the Institutional Review Board of the Seoul National University Bundang Hospital (approval number: B-2102-666-106) and adhered to all ethical guidelines. A total of 1626 face images were collected from 98 patients with VCFS and 91 non-VCFS controls who visited the Seoul National University Bundang Hospital between May 2003 and September 2020. Heterozygous deletion at chromosome 22q11.2 was confirmed in all patients with VCFS by CMA or targeted deletion analysis with FISH. Non-VCFS facial photos were obtained from cases without facial phenotypic abnormalities. All cases were photographed in six different poses, including the front, front with head raised, left and right profiles, and left and right 45º views. Written consent was obtained from all patients in both groups for the use of their standardized medical photographs for research purposes.

Image preprocessing

All face images in this study were transformed into colored square images at dimensions of 112x112x3 pixels. We used MTCNN to extract the face regions from the images. MTCNN was used to perform consecutive face detection and alignment through the use of three types of CNN modules [9]. In the face

detection step, the MTCNN algorithm placed bounding boxes on the images to extract the face regions. In the face alignment step, the algorithm used the five facial landmarks within the bounding boxes to rotate the face regions as much as possible to horizontally align them, thus producing pre-normalized face images. After the MTCNN process, the pre-normalized face images were fully normalized to have a mean and standard deviation of 0.5 for each color channel. To increase the diversity of data, random horizontal flip was applied to the images with a ratio of 0.5 for image augmentation.

Face recognition

We constructed two face recognition models using publicly available large-scale datasets: the CASIA-WebFace dataset, with 453,580 images for 10,575 subjects [10] and the MS-Celeb-1M dataset, with 3,923,399 images for 86,876 subjects [11]. These datasets were used to learn general facial representation. The CNN architecture used was residual neural networks with 100 layers (ResNet-100), which has been widely used in many deep learning-based face recognition studies [12] because of its ability to effectively solve the degradation problem in deep neural networks [13].

We trained and validated the two CNN-based face recognition models, naming them the CASIA and MS-Celeb-1M face recognition models. We performed face verification as an external validation to identify if two images belonged to the same group. For external validation, we used three publicly available datasets that were not used in the training period: the labeled faces in the wild (LFW) dataset, with 13,233 images for 5,749 subjects [14], in-the-wild age database (AgeDB), with 16,488 images for 568 subjects [15], and celebrities in frontal-profile in the wild (CFP-FP) dataset, with 7,000 images for 500 subjects [16].

The additive angular margin loss function was used to maximize the face class separability of two images belonging to different groups [17]. The feature dimension used was 512, the feature scale was 32, and the angular margin was 0.5. Performance evaluation was based on the accuracy of face verification (Figure 1).

Binary syndrome classification

We constructed two binary syndrome classification models using the CASIA and MS-Celeb-1M face recognition models. We modified the output layers of the ResNet-100 architecture to comprise two successive fully-connected layers, batch normalization [18], and the rectified linear unit (ReLU) and dropout [19] components for binary classification. Subsequently, we fine-tuned the modified output layers using VCFS and non-VCFS images.

In the image preprocessing step, 185 VCFS and 58 non-VCFS images were excluded as MTCNN failed to extract their face regions. Accordingly, 1,105 images from 98 patients with VCFS and 521 images from 91 non-VCFS patients were included in the VCFS and non-VCFS groups, respectively. In the VCFS group, 868 images from 88 patients and 52 images from 10 patients were assigned to the training and test datasets, respectively, and in the non-VCFS group, 442 images from 81 patients and 21 images from 10 patients

were assigned to the training and test datasets, respectively. To handle the data imbalance between VCFS and non-VCFS images, we used random over-sampling method.

We separately trained two binary classification models—the CASIA and MS-Celeb-1M classifiers—using a five-fold cross-validation method. The models were evaluated based on the angles of face images: multiple angles, including frontal angle, 45º angle, and lateral hemiface; and frontal angle only. In the final layer of our modified output layers, we used a softmax activation function to output the probability of a given image being classified as either indicating VCFS or non-VCFS (Figure 2).

The performance of the binary classification models was evaluated using various metrics, including the F1-score, accuracy, and area under the receiver operating characteristic curve (AUROC) for the multiple angles using 52 VCFS and 21 non-VCFS images and accuracy and AUROC for the frontal angle only using 10 VCFS and 10 non-VCFS images.

Moreover, we generated two-dimensional colored heatmaps to visualize which face regions contributed to the classification decisions in our classifiers using gradient-weighted class activation mapping (Grad CAM) [20], which calculates the gradients of the last convolutional layers using the partial derivatives of the predicted class scores, providing the importance levels of image pixels in terms of deciding whether the corresponding images indicate VCFS or non-VCFS.

Implementation details

We used Python version 3.7 with PyTorch version 1.4.0 library [21] and four 12-GB NVIDIA TITAN V graphics processing units with the CUDA version 10.1 programming interface. In the face recognition phase, we used the stochastic gradient descent algorithm as an optimization process, with an initial learning rate of 0.05, Nesterov momentum of 0.9, and weight decay of 0.0005. We used the cross-entropy loss function as the cost function, as well as a mini-batch size of 256.

In the fine-tuning phase, we used a He Uniform Initializer [22] for the ReLU activation function and the Adam optimizer [23] with cosine annealing and warm restarts [24] for the optimization process, with an initial learning rate of 0.0003. We used cross-entropy loss function as the cost function, a mini-batch size of 128, and early stopping, with a patience of 50, to avoid overfitting.

Results

Performance of Facial Recognition Models

The performance of our CASIA and MS-Celeb-1M face recognition models were evaluated through face verification using the LFW, AgeDB, and CFP-FP datasets. Using the CASIA face recognition model, we obtained an accuracy of 99.52%, 94.57%, and 95.89% for the LFW dataset, AgeDB, and CFP-FP dataset,

respectively. Moreover, we obtained an accuracy of 99.78%, 97.93%, and 94.03% for the LFW dataset, AgeDB, and CFP-FP dataset, respectively, using the MS-Celeb-1M face recognition model.

Binary Classification of VCFS and non-VCFS images

In this study, the performance of the CASIA and MS-Celeb-1M binary classification models were evaluated using 10 VCFS and non-VCFS cases. The CASIA classifier achieved an F1-score and accuracy of 88.02% and 87.67%, respectively, when using images captured from multiple angles, whereas the MS-Celeb-1M classifier yielded an F1-score and accuracy of 81.43% and 80.82%, respectively. Additionally, it has been observed that the accuracy of face recognition was higher using images captured from the frontal angle, with an accuracy of 95.00% for both the CASIA and MS-Celeb-1M classifiers, compared to that using multiple angles. Furthermore, the performance of face verification and classification is detailed in Table 1, and the Receiver Operating Characteristic (ROC) curves for the classifiers are presented in Figure 1. The area under the ROC curve (AUROC) was found to be 0.93 and 0.91 for the CASIA and MS-Celeb-1M classifiers, respectively, when using images captured from multiple angles, and 1.00 and 0.98 for the CASIA and MS-Celeb-1M classifiers, respectively, when using images captured from the frontal angle (Figure 3).

Table 1. Performance of the face recognition models and VCFS classifiers.

Model number			1	2	3	4
Structure			ResNet100	ResNet100	ResNet100	ResNet101
Train data			CASIA	MS-Celeb-1M	K-Face	Asian-Celeb
						CASIA(asian)
						K-Face
Face verification accuracy (%)	Test data	LFW	99.52	99.78	70.78	98.23
		AgeDB-30	94.57	97.93	51.1	88.22
		CFP-FP	95.89	94.03	64.2	89.13
Classification accuracy (%)	Multiple angle*	F1-Score	88.02	81.43	67.5	90.34
		Accuracy	87.67	80.82	71.23	90.41
	Frontal image	Accuracy	95	95	65	85

*Multiple angle : Facial photographs that were taken from front, lateral 45 degrees and lateral hemiface were all included in the classification.

We demonstrated the use of Grad CAM heatmaps in 10 patients with VCFS patients to visualize the regions of the face that contributed most to the classification decisions made by the CASIA and MS-

Celeb-1M classifiers. The importance levels of these regions are represented by color, with low contributions represented by blue and high contributions represented by red. The stronger the correlation between the face regions and VCFS, the higher the importance levels. Our results revealed that the face regions identified by Grad CAM as having high importance levels showed good agreement with the facial phenotypes characteristic of patients identified to have VCFS, which were carefully examined by physicians (Figure 4).

Discussion

VCFS is a chromosomal microdeletion disorder that is recognized as the most prevalent among this class of syndromes. It is often characterized by a specific facial phenotype; however, clinicians may find it challenging to perform a diagnosis based on facial features alone, as there is a significant degree of symptomatic variability, and the facial features of VCFS are less pronounced compared to those of other chromosomal syndromes, including Down or Williams syndrome. The diagnosis of VCFS is established in a proband through the identification of a heterozygous deletion at chromosome 22q11.2. Most patients with VCFS (85%) have a 2.54-Mb deletion encompassing approximately 40 genes [25]. Recently, most individuals with a 22q11.2 recurrent deletion are identified through CMA performed in the context of evaluation for developmental delay, intellectual disability, or autism spectrum disorder. CMA using oligonucleotide or SNP arrays can detect recurrent deletion with 100% sensitivity, and the ability to size the deletion depends on the type of microarray used and density of probes in the 22q11.2 region [5].

There are over 180 clinical features associated with VCFS. The most common among these is congenital heart disease, which is found in approximately 70% of children with VCFS according to a study by Young et al. [26]. Palatal anomalies and accompanying velopharyngeal insufficiency are relatively common in patients with VCFS, and early surgical intervention and speech therapy for speech disability are particularly important for future speech development and social behavior, as noted in the studies of Lachman et al. [27]. VCFS is also associated with various anomalies, including vascular, ocular, cranial, and limb abnormalities, as well as possible psychiatric disorders [28]. Given the wide range of clinical features associated with VCFS, it is important to detect the patients' genetic abnormalities early on to prevent the development of additional abnormalities.

The use of machine learning systems to aid in the diagnosis of rare genetic diseases with typical facial phenotypes has been explored in previous research. For instance, in a study conducted by Gurovich et al. [7], the researchers were able to achieve a high level of accuracy in the diagnosis of patients with Cornelia de Lange Syndrome using facial phenotypes with the aid of the deep CNN-based platform DeepGestalt. They also demonstrated that the genotype of Noonan Syndrome could be distinguished solely based on facial features using deep learning techniques.

The significance of this study lies in its demonstration that the facial phenotype of VCFS can be accurately classified with a CNN-based face recognition model, achieving an accuracy rate of 95%. Given that VCFS is relatively common among multiple anomaly syndromes, utilizing two-dimensional medical

photography in conjunction with this model may aid in the early diagnosis and prevention of various abnormalities in affected individuals. To the best of our knowledge, this study is the first to be conducted in an Asian population among a genetically anomalous cohort, making it an important contribution to the field.

This study also provides qualitative confirmation that the CNN-based face recognition model used in this study is able to effectively identify the same facial features that clinicians typically evaluate during the diagnostic process for VCFS. Specifically, the heatmap generated by the model substantially overlapped with the facial areas that clinicians traditionally assess, which serves as a “retrospective validation” of previously established reference standards for the diagnosis of VCFS using machine learning systems and suggests that the facial recognition model developed in this study is well-aligned with the clinically recognized facial phenotypes of patients with VCFS.

The results of this study demonstrate the potential utility of using machine learning technology in the diagnosis of rare genetic diseases such as VCFS, particularly in primary medical institutions that may not have extensive experience with these diseases. The use of the CNN-based face recognition model in this study has the potential to enable more efficient and rapid diagnosis of VCFS, especially for primary physicians who may have limited experience with rare diseases. Furthermore, this technology can serve as a powerful tool for screening and diagnosing patients in areas where genetic diagnosis is not feasible owing to practical limitations, including economically underdeveloped countries or small population areas with limited access to large medical institutions.

Nevertheless, this study has some limitations. As with other machine learning studies, the accuracy of the model is limited by the quality and diversity of the input data. In this study, the face photos used for training were all taken from the front and had a homogeneity of backgrounds and facial expressions, which may not generalize well to photographs taken under different conditions. Additionally, the study was conducted using a sample of Asian subjects, specifically Koreans, and additional research using different racial and ethnic populations is necessary to confirm the generalizability of these results.

Furthermore, it is crucial to consider the ethical implications of this technology. As this model is based on the photographs of patients, particularly children, it is essential to implement robust security measures to protect the personal information of patients. One potential solution to this issue could be the use of blockchain-based technology to ensure the secure handling of personal medical information.

Conclusion

In this study, we evaluated the accuracy of a face recognition technology using various types of MTCNN for the binary classification of VCFS. We obtained good results and confirmed that the heatmap of the region used for diagnosis matched the clinically recognized facial phenotypes of VCFS. Our findings suggest that the use of the deep learning-based diagnosis of VCFS may have the potential to aid medical staff in regions where rare diseases are not common. This diagnostic approach could also help detect and respond to rare diagnoses early on in situations where genetic diagnosis is not available.

Nevertheless, further studies to confirm the generalizability of these results and consider the ethical and practical implications of this technology are warranted.

Declarations

Author contributions

R.M.B., A.C., Y.G.C. performed data interpretation, statistical analysis, and wrote the manuscript. Y.J., H.K., H.H. performed data interpretation. Y.J.M. performed data interpretation and statistical analysis and wrote the manuscript. All authors revised, reviewed, and approved the final manuscript.

Acknowledgement

The authors have obtained informed consent for publication of the images from the patients that facial images were included in the figure. This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HR22C1363).

Competing interests

All authors declare no financial or non-financial competing interests.

Data availability

The datasets generated and/or analysed during the current study are not publicly available due to hospital and instutional review board policies, but are available from the corresponding author on reasonable request.

References

1. Stevens, C. A., Carey, J. C. & Shigeoka, A.O. Di George anomaly and velocardiofacial syndrome. *Pediatrics.* **85**, 526-530 (1990).
2. Rozas, M. F. et al. Association between phenotype and deletion size in 22q11.2 microdeletion syndrome: systematic review and meta-analysis. *Orphanet. J. Rare. Dis*. **14,** 195 (2019).
3. McDonald-McGinn, D. M., et al. 22q11.2 deletion syndrome. *Nat. Rev. Dis. Primers.* **1**, 15071 (2015).
4. Botto, L. D., et al. A population-based study of the 22q11.2 deletion: phenotype, incidence, and contribution to major birth defects in the population. *Pediatrics* **112**, 101-107 (2003).
5. Mantripragada, K. K., et al. DNA copy-number analysis of the 22q11 deletion-syndrome region using array-CGH with genomic and PCR-based targets. *Int J Mol Med.* **13**, 273-279 (2004).
6. Antshel, K. M., et al. Behavior and corpus callosum morphology relationships in velocardiofacial syndrome (22q11. 2 deletion syndrome). *Psychiatry Res.* **138**: 235-245 (2005).

7. Gurovich, Y., et al, Identifying facial phenotypes of genetic disorders using deep learning. *Nat Med.* **25**, 60-64 (2019).
8. Porras, A. R., et al. Development and evaluation of a machine learning-based point-of-care screening tool for genetic syndromes in children: a multinational retrospective study. *Lancet. Digit. Health.* **3**, e635-e643 (2021).
9. Zhang, K., et al. Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters.* **23**, 1499-1503 (2016).
10. Yi, D., et al. Learning face representation from scratch. arXiv preprint arXiv:1411.7923, 2014.
11. Guo, Y. et al. Ms-celeb-1m: A dataset and benchmark for large-scale face recognition. in European conference on computer vision. (Springer, 2016).
12. Masi, I. et al. Deep face recognition: A survey. in 2018 31st SIBGRAPI conference on graphics, patterns and images (SIBGRAPI). *IEEE.* 471-478 (2018).
13. He, K., et al. Deep residual learning for image recognition. in *Proceedings of the IEEE conference on computer vision and pattern recognition.* 770-778 (2016).
14. Huang, G. B. et al. Labeled faces in the wild: A database forstudying face recognition in unconstrained environments. in Workshop on faces in 'Real-Life'Images: detection, alignment, and recognition. (2008).
15. Moschoglou, S. et al. Agedb: the first manually collected, in-the-wild age database. in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 51-59 (2017).
16. Sengupta, S. et al. Frontal to profile face verification in the wild. in 2016 IEEE Winter Conference on Applications of Computer Vision (WACV). *IEEE.* (2016).
17. Deng, J. et al. Arcface: Additive angular margin loss for deep face recognition. in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2019).
18. Ioffe, S. & Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. in International conference on machine learning. PMLR. (2015).
19. Srivastava, N. et al. Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res*. **15**: 1929-1958 (2014).
20. Selvaraju, R. R. et al. Grad-cam: Visual explanations from deep networks via gradient-based localization. *Proceedings of the IEEE international conference on computer vision.* (2017).
21. Paszke, A. et al. Pytorch: An imperative style, high-performance deep learning library. *Adv. Neural. Inf. Process. Syst.* 32 (2019).
22. He, K., et al. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. *Proceedings of the IEEE international conference on computer vision.* (2015).
23. Kingma, D. P. & Adam J. L. Ba. A method for stochastic gradient descent. *ICLR: International Conference on Learning Representations.* (2015).
24. Loshchilov, I. & Hutter F. SGDR: Stochastic Gradient Descent with Warm Restarts. *arXiv: Learning.* (2017).

25. McDonald-McGinn, D. M. & Zackai, E. H. Genetic counseling for the 22q11.2 deletion. *Dev. Disabil. Res. Rev*. **14**, 69-74 (2008).
26. Young, D., Shprintzen, R. J. & Goldberg, R. B. Cardiac malformations in the velocardiofacial syndrome. *Am. J. Cardiol.* **46**, 643-648 (1980).
27. Myung, Y., Ahn, T., Kim, B. K., Jeong, J. H., Baek, R. M. Clinical significance of the levator veli palatini muscle in velocardiofacial syndrome patients: implications in velopharyngeal incompetence and pharyngeal flap surgery. *Cleft Palate Craniofac J.* **55**, 521-527 (2018).
28. Lachman, H. M., et al., Association of codon 108/158 catechol-O-methyltransferase gene polymorphism with the psychiatric manifestations of velo-cardio-facial syndrome. *Am. J. Med. Genet*. **67**, 468-472 (1996)

Figures

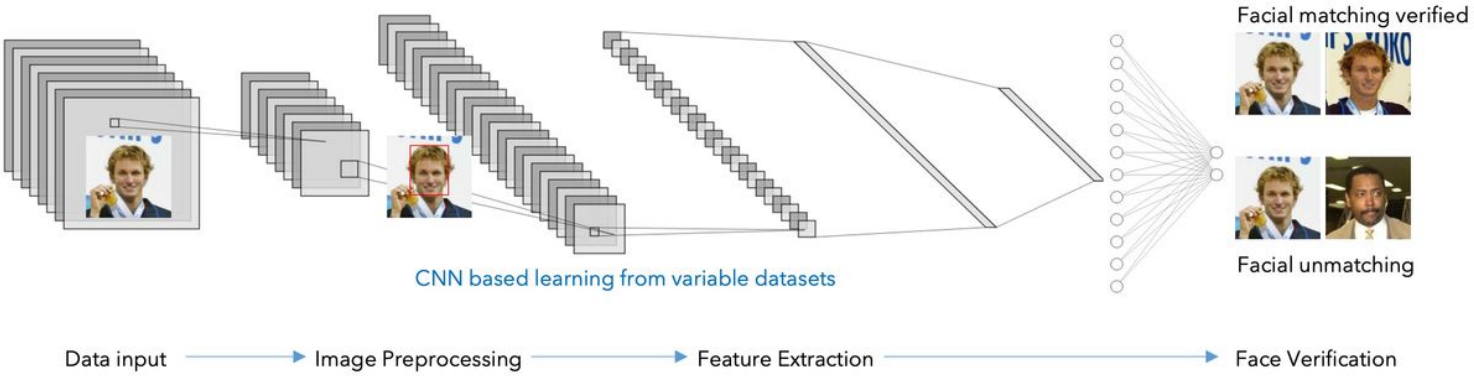


Figure 1

Diagram of the face recognition model construction using large-scale datasets for human facial images. Resnet-100 was used for the architecture of cascaded convolutional networks (CNN). Training and validation was performed with the CASIA-WebFace and MS-Celeb-1M datasets.

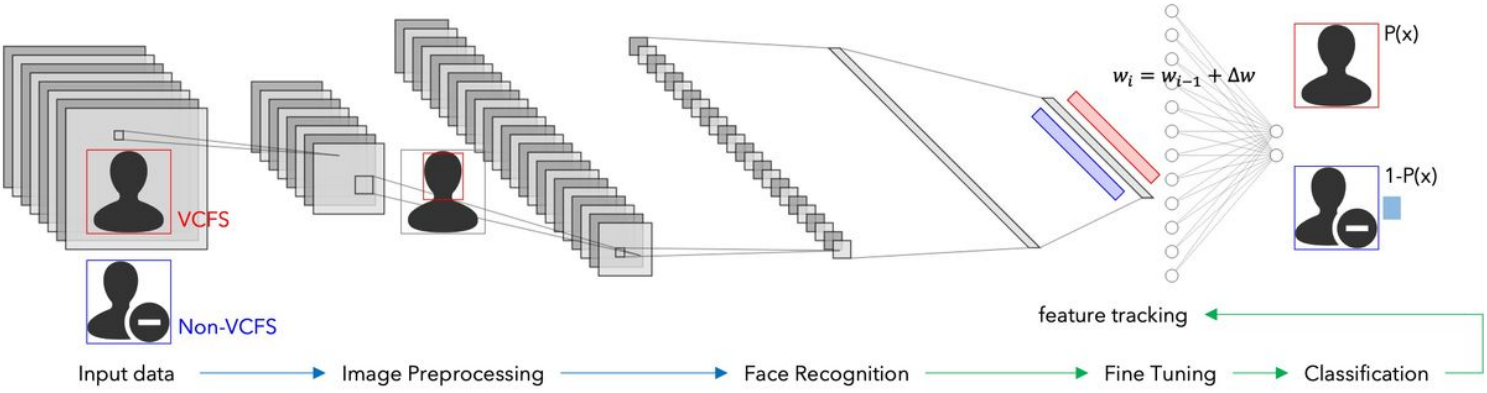


Figure 2

Structure of binary classification model construction. MTCNN-based algorithm was applied for facial recognition and image sorting. The fully connected layer was transferred to binary classification type, with training and parameter optimization using VCFS and normal face images.

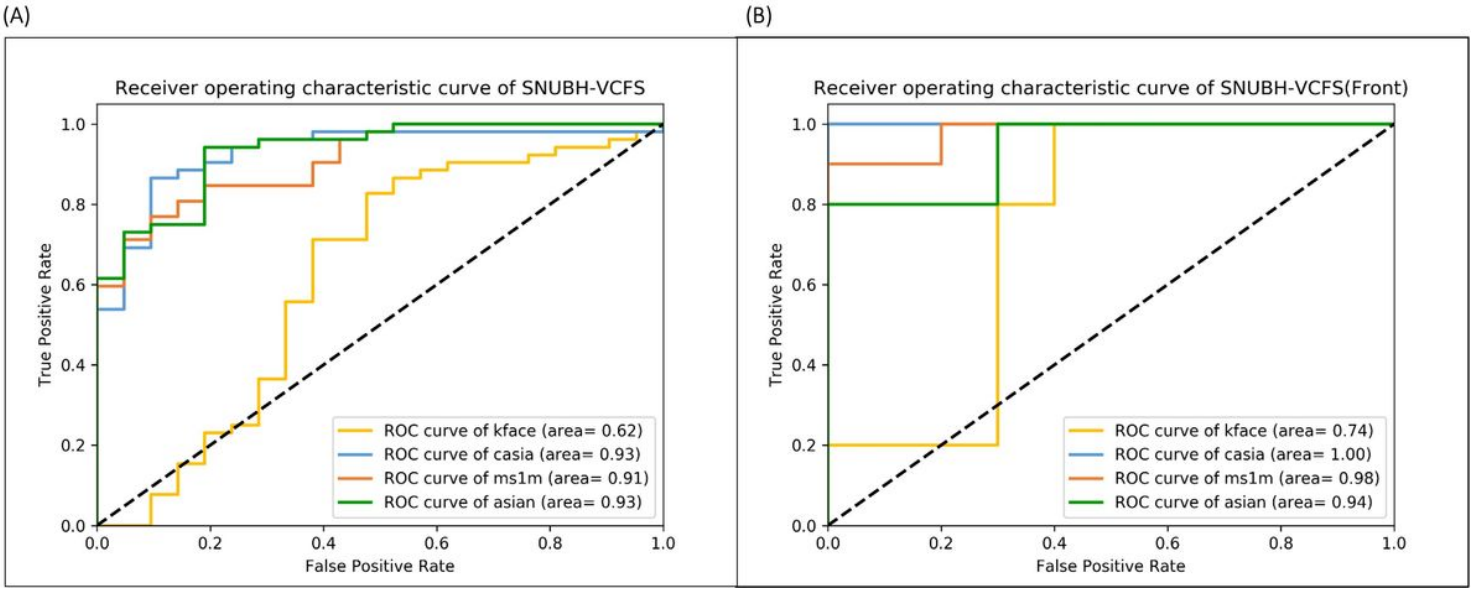


Figure 3

Receiver operating characteristic (ROC) curve of the facial classification models. (A) Evaluation of multiple angle facial images and (B) frontal angle facial images. Accuracy was higher in evaluating frontal images only, and different results are shown between the two training datasets.

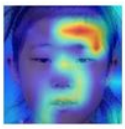
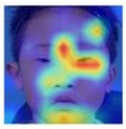
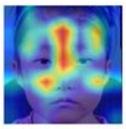
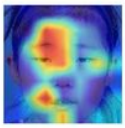
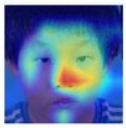
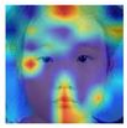
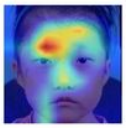
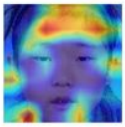
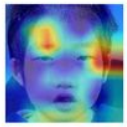
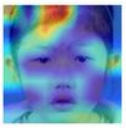
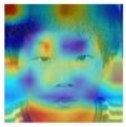
Model 1										
VCFS (%)	96.47	99.98	99.95	95.14	99.37	99.63	99.74	99.80	99.39	99.96
Model 2										
VCFS (%)	97.10	99.37	99.94	97.23	99.89	69.35	99.64	99.95	6.39	100.00
Model 3										
VCFS (%)	99.72	99.99	87.44	99.97	99.84	96.17	99.87	94.42	99.18	89.43

Figure 4

Two-dimensional colored heatmaps. Facial regions that contributed to the decision of binary classification between VCFS and non-VCFS was mapped by gradient-weighted class activation. Most of the regions of interest include perinasal and periorbital areas, which are consistent with the known facial phenotypes of VCFS.