EARLY DIAGNOSING PARKINSON'S DISEASE VIA A DEEP LEARNING MODEL BASED ON AUGMENTED FACIAL EXPRESSION DATA

Yintao Zhou[1], Meng Pang[1*], Wei Huang[1*], Binghui Wang[2]
[1]Nanchang University, China; [2]Illinois Institute of Technology, USA

ICASSP 2024 KOREA

南昌大学 NANCHANG UNIVERSITY

Background

- The considerable expense associated with in-vivo PD diagnosis.
- PD patients exhibit evident expression disorder, and their ability to express emotions appears less natural compared to normal individuals.
- The suboptimal performance of current facial expressions-based PD diagnosis methods due to two main factors:
  1) the limited size of the training data on PD patients' facial expressions;
  2) the prediction models using conventional machine learning techniques are incapable of extracting high semantic features for classification.

Fig. 1. A comparison between the PD patients' masked faces and the faces of the normal individuals depicting six basic emotions.

Method

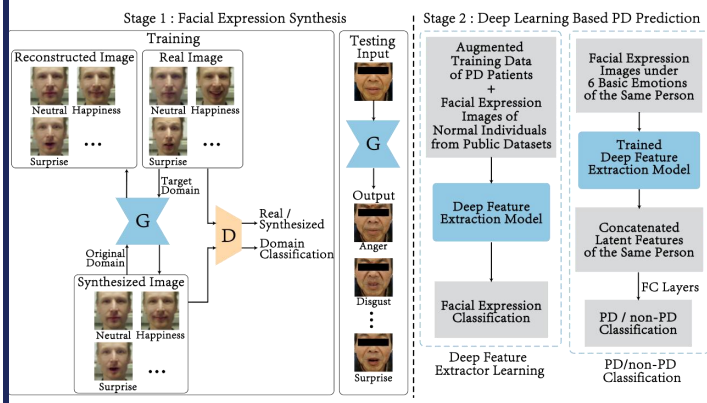

Fig. 2. The flow-chart of the proposed PD diagnosis method.

Stage 1: Facial Expression Synthesis

We design an effective MDFES-GAN to generate multi-domain identity-preserved facial images of PD patients depicting six basic emotions. Four meticulously crafted losses are integrated to optimize the MDFES-GAN model:

Adversarial Loss

To generate realistic-looking facial expression images in multiple domains, the adversarial loss in StarGAN is introduced into our MDFES-GAN model:

$$L_{adv} = E_x[D_{src}(x)] - E_{x,c}[D_{src}(G(x,c))] - \lambda_{gp} E_{\hat{x}}[(\|\nabla_{\hat{x}} D_{src}(\hat{x})\|_2 - 1)^2]$$

Domain Classification Loss

Apart from discerning the authenticity of an image, the discriminator is designed to predict the domain label of the given image as well:

$$L_{cls} = E_{x^*,c^*}[-\log D_{cls}(c^*|x^*)]$$

Reconstruction Loss

To ensure that the generator transfer only the domain-specific part of the input images, a cycle consistency loss is adopted for image reconstruction:

$$L_{rec} = E_{x,c,c'}[\|x - G(G(x,c),c')\|_1]$$

Identity Preserving Loss

In order to preserve crucial identity information while generating facial expression images, we introduce an identity preserving loss:

$$L_{ip} = E_{x,c}[1 - \frac{F(x) \cdot F(G(x,c))}{\|F(x)\|_2 \|F(G(x,c))\|_2}]$$

Stage 2: Deep Learning Based PD Prediction

Deep Feature Extractor Learning

We first learn a deep feature extractor accompanied with an expression classifier using a combination of the augmented training data of PD patients and the normal individuals' facial expression images from public datasets.

PD/non-PD Classification

We train a PD/non-PD classifier which leverages the deep feature extractor as the backbone, with the original d × 6 final FC layer replaced by two new FC layers for fine-tuning.

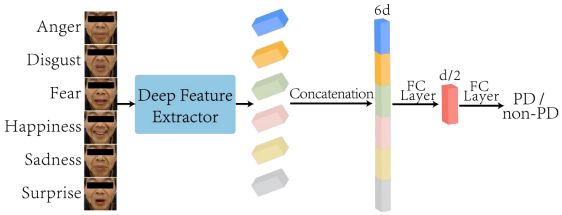

Fig. 3. The pipeline for classifying PD/non-PD.

Experimental Results

Facial Expression Synthesis Evaluation

- MDFES-GAN successfully generates realistic-looking facial expression images of PD patients while preserving their personal identity.
- The preference results show that more synthesized images by MDFES-GAN are perceived to be of high quality than those by StarGAN across six different emotions.

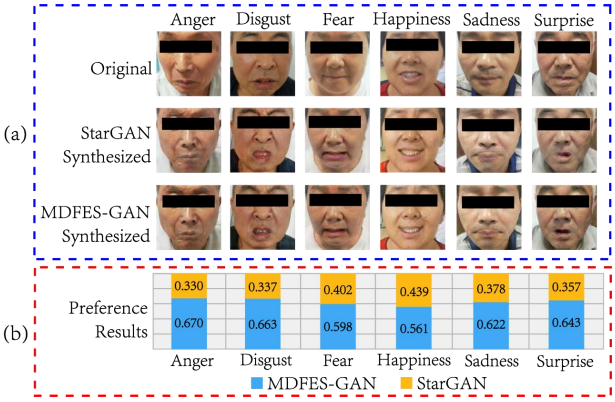

Fig. 4. (a) Comparison between the synthesized facial expression images by StarGAN and our MDFES-GAN. (b) The preference results that answer the question "which image is of higher quality?".

PD Diagnosis Performance

- Our method obtains a high PD diagnosis accuracy of 97.55%, surpassing the performance of the five deep learning methods and significantly outperforming the conventional machine learning-based GLCM+SVM
- Our method delivers an improvement of 6.78% over "Ours w/o synthesized facial expression data", which suggests that introducing the synthesized PD patients' premorbid facial expression images via MDFES-GAN can further facilitate the training of the PD prediction model.

Table 1. Comparison of PD diagnosis accuracy (%) with other methods for PD diagnosis.

Methods	Accuracy (%)
GLCM+SVM [27]	67.71
SwinTransformer-Small [28]	81.90
ConvNeXt-Tiny [29]	80.94
CoAtNet-0 [30]	90.16
EfficientNet-B0 [31]	82.81
EfficientNetV2-Small [32]	89.87
Ours w/o synthesized facial expression data	90.77
Ours	**97.55**

Conclusion

We proposes a novel in-vitro PD diagnosis method that leverages facial expression data augmentation and deep neural network prediction. Empirical studies have demonstrated the efficacy of MDFES-GAN in multi-domain facial expression synthesis, along with the promising performance of the proposed approach in PD diagnosis.

Acknowledgement: This work is jointly supported by NSFC No.62271239 and Jiangxi Double Thousand Plan No.JXSQ2023201022, by Jiangxi Provincial Natural Science Foundation (20232BAB212025) and Jiangxi Provincial High-level and Urgently Needed Overseas Talent Programs (20232BCJ25024).