INTERNATIONAL CONFERENCE ON

# APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN TRANSPORTATION

## CO-ORGANIZERS

HO CHI MINH CITY DEPARTMENT OF TRANSPORT

UNIVERSITY OF TRANSPORT & COMMUNICATIONS (UTC)

FENG CHIA UNIVERSITY (FCU)

ASIAN INSTITUTE OF TECHNOLOGY IN VIETNAM

## SPONSORS

HO CHI MINH CITY DEPARTMENT OF SCIENCE & TECHNOLOGY

SUMITOMO CONSTRUCTION MACHINERY Co., Ltd

INNOVATIVE TECHNOLOGY DEVELOPMENT CORPORATION

ISBN: 978-604-76-2102-6

9 786047 621026

International Conference on APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN TRANSPORTATION

HO CHI MINH CITY - VIETNAM

11 DECEMBER 2020

ISBN: 978-604-76-2102-6

CO-ORGANIZERS

SPONSORS

# CONFERENCE PROCEEDINGS

INTERNATIONAL CONFERENCE ON

# APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN TRANSPORTATION

NHÀ XUẤT BẢN GIAO THÔNG VẬN TẢI

# CONFERENCE PROCEEDINGS

## INTERNATIONAL CONFERENCE ON

# APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN TRANSPORTATION

**Ho Chi Minh City, Vietnam**

# COMMITTEE

## Organizing Committee

1. Assoc.Prof. Nguyen Van Hung
2. Dr. Vo Truong Son
3. MSc. Tran Quang Lam
4. Prof. Tien-Yin (Jimmy) Chou
5. Dr. Phung Van Dong
6. Dr. Ngo Chau Phuong
7. Assoc.Prof. Nguyen Duy Tien
8. Dr. Trinh Quang Khai
9. Dr. Tran Xuan Truong
10. Assoc.Prof. Nguyen Canh Minh
11. Dr. Nguyen Quoc Tuan
12. Dr. Nguyen Xuan Tung
13. Assoc.Prof. Le Van Bach
14. Dr. Nguyen Van Binh
15. Dr. Nguyen Thac Quang
16. MSc. Tran Phong Nha

## Program Committee

1. Assoc.Prof. Nguyen Thanh Hai
2. Prof. Le Hung Lan
3. Dr. Bui Ngoc Dung
4. Dr. Nguyen Van Binh
5. Dr. Vu Viet Hung
6. Dr. Dao Thanh Toan
7. Dr. Hoang Van Thong
8. Assoc.Prof. Nguyen Huu Hung
9. Dr. Le Ba Anh
10. Assoc.Prof. Tran Ngoc Hien
11. Dr. Dinh Thi Thanh Binh
12. MSc. Ho Viet Long
13. MSc. Tran Phong Nha
14. MSc. Doan Van Tan
15. Prof. Cheong Siew Ann
16. Prof. Wen-Nung Lie
17. Prof. Tien-Yin (Jimmy) Chou
18. Prof. Dimitar Borisov
19. Prof. Hansung Lee

## Secretariat

1. Dr. Vu Viet Hung
2. MSc. Pham Anh Linh
3. MA. Ngo Thuy Linh
4. MSc. Nguyen Thi Mai Nhung
5. MSc. Doan Minh Huy
6. Assoc.Prof. Hoang Thanh Van

## Table of Contents

| No | Title/ Author | Page |
|---|---|---|
| 1 | **Preface** | 1 |
| 2 | **Development of An Intellectual Monitoring System for Comprehensive Assessment of Urban Road - Transport Environmental Pollutions**<br>*Anna Makarova, Andrey Glushko, Vitaliy Chelnokov, Alexey Matasov, Georgii Priorov* | 3 |
| 3 | **Crack Detection for Asphalt Pavements Using Deep Learning-Based Faster R-CNN**<br>*Van Phuc Tran, Thai Son Tran, Van Phuc Le, Hyun Jong Lee* | 12 |
| 4 | **Applications of Artificial Intelligence in Manufacturing Systems**<br>*Tran Ngoc Hien* | 18 |
| 5 | **Development of A Smart Application for Safe Inland Waterway Transportation Based on Built-In Sensors and GPS**<br>*Duc-Nghia Tran, Ton That Long, Nguyen Canh Minh, Pham Tien Thanh, Duc-Tan Tran* | 26 |
| 6 | **Pose Invariant Face Recognition System**<br>*Nguyen Dong Dong, Bui Ngoc Dung, Hoang Xuan Tung, Tran Nguyen Cac* | 33 |
| 7 | **A Simulation of Intelligent Traffic Control Using Fuzzy Logic Under Mixed Traffic Condition**<br>*Vuong Xuan Can, Nguyen Hoang Van, Vu Trong Thuat, Nguyen Thi An* | 41 |
| 8 | **The Solution Formulation for Monitoring Remote Parameters of Railway Equipment by IoT Model**<br>*Nguyen Van Nghia, Do Viet Dung, An Thi Hoai Thu Anh* | 48 |
| 9 | **Applications of Multi Agent Systems in Urban Congestion Control**<br>*Huan Vu* | 59 |
| 10 | **Optimization for Online People and Parcels Share-A-Ride Taxis with Time Dependent Travel Times**<br>*Nguyen Duong Hung, Nguyen Xuan Hoai, Pham Quang Dung, Nguyen Van Son* | 69 |

# POSE INVARIANT FACE RECOGNITION SYSTEM

**Nguyen Dong Dong[1, 2], Bui Ngoc Dung[1*], Hoang Xuan Tung[3], Tran Nguyen Cac[4]**

[1]*University of Transport and Communications, No 3 Cau Giay Street, Hanoi, Vietnam.*

[2]*Samsung Vietnam Mobile R&D Center, No 1 Pham Van Bach Street, Hanoi, Vietnam.*

[3]*VNU University of Engineering and Technology, 144, Xuan Thuy Street, Hanoi, Vietnam.*

[4]*Research and Development Division, Genexis B.V., Netherlands.*

[*]*Corresponding author: dnbui@utc.edu.vn*

**Abstract:** Face recognition has become popular in security surveillance systems. These systems work well only when the face is in frontal and the system itself has considerable training data. This paper proposes a method of face recognition with various view using Generative Adversarial Networks (GAN). In this method, a pre-trained model of face is used to generate the training data of different pose variations from several frontal faces using GAN. These faces will be used as input of the FaceNet and Support Vector Machine to be trained to perform the task of face recognition. The experiment results demonstrate the accuracy of the proposed method in the case of lacking the training data. This paper focus on the applicability of Generative Adversarial Network in Face Recognition System by applying the GAN-based method in Face Recognition System to increase the number of training samples and detect the faces of different poses that was not possible with the conventional Face Recognition System. The comparison between Face Recognition System using GAN and conventional Face Recognition System was presented.

**Keywords:** face recognition, generative adversarial network, facenet.

## 1. INTRODUCTION

Face recognition is a highlight in image processing field, a computer application which is able to automatically recognize human face through a digital image or a video frame. Face recognition system is used in dealing with security problems, detecting and identifying a person in restricted area. A face recognition system commonly includes three main parts: face detection, facial feature extraction and identification of the face by comparing its facial features with stored information in database. Due to the advances in deep learning, the accuracy of face recognition methods is improved significantly. However, face recognition in practice always comes with challenges like pose, illumination, face expression and maybe their combinations. Besides, it requires a great deal of training data for desired accuracy.

There have been many researches trying to solve the problems of the complex face recognition. In [1], Principal Component Analysis (PCA) was used to extract the features from the face and recognize using neural network. This method overcomes the disadvantage of illumination sensitive of PCA [2]. A multi-view deep network has been used to find the nonlinear discriminant and view invariant representation shared between views [3]. To recognize the pose-invariant face, [4] using GAN to generate the high quality images with 9 different views for training data. Multi-scale local binary pattern features from 27 landmarks were extracted to form the feature vector for pose in variance face recognition [5]. In [6], pose robust feature obtained from the combination of component level and landmark level to overcome the nonlinear intra personal variation [7]. Pose invariance face recognition also be used by Fully-trained Generative Adversarial Networks (FTGAN) for pixel transformation to achieve coarse face transformation, and these faces are refined by key point alignment [7]. Disentangled Representation Learning GAN (DR-GAN) [8] is a method for PIFR (Pose-Invariant Face Recognition) whose purpose is to solve the problem of signficant pose discrepancy between two faces in face recognition, which is not same as the purpose of our proposed system.

In this paper, we propose a new method of face recognition with pose invariance of faces but lack of training images. The system is capable of recognizing faces of different poses and influenced by external agents, while performing facial representations from a fixed posture. Especially, with only a small amount of training data, the system still produces positive recognition results thanks to the ability to re-represent faces from the original data set. To do this, beside the frontal faces, we use GAN to generate fake faces from different poses. This step helps the system recognize the non-frontal faces and increases the training data. These faces were given to the input of FaceNet to extract high quality features, also known as face embedding. Finally, Support Vector Machine (SVM) will be used to train a face identification system. The general framework is shown in figure 1.

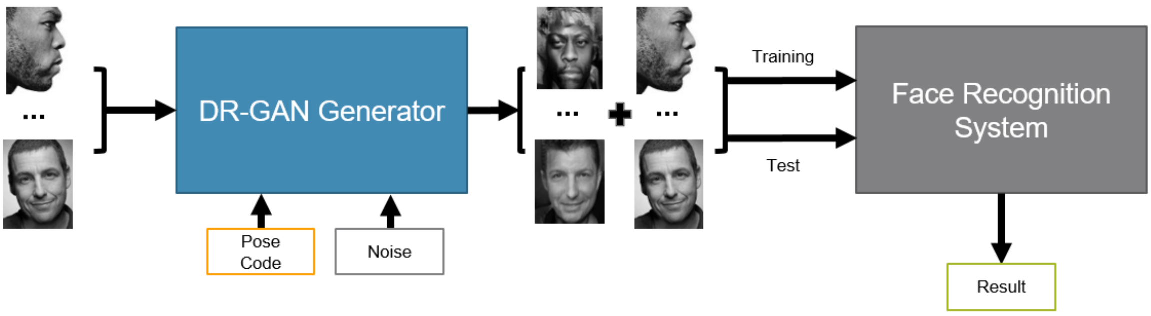


***Figure 1**: DR-GAN Generator works as a face rotator.*

## 2. METHODOLOGY

Generative Adversarial Networks (GAN) [9] consists of two main components. Generator G is a generative model and discriminator D is a distinguishing model. G takes a probability distribution as input and tries to synthesize a realistic image. D takes two input images, one being the real image from training data and another being fake image generated by G. It tries to determine whether the input is a fake or a real one. D and G are trained together, D tries to become more efficient in distinguishing between real and fake images and G tries to generate more realistic images that D would not be able to recognize [9].

The process of face recognition using GAN includes 4 steps:

- Generate frontal face images from original training data.
- Face detection.
- Facial feature extraction.
- Face recognition.

Using the frontal and lateral face images, Disentangled Representation Learning GAN (DR-GAN) [8] generator generates extra frontal face images. Training data includes original images and generated images of which a proportion is used for validation. Training images are processed to detect the faces using Haar-Like feature method with AdaBoost and Cascade Classifier algorithms. These detected faces were given to the input of the FaceNet in order to extract high quality facial features. Finally, SVM will be used to train the face identification system.

### 2.1. Frontal face images generation using DR-GAN.

Input images will be resized to 96x96 and passed into the encoder of generator $G_{enc}$. Output of $G_{enc}$ is a 320-dimension feature vector $f(x)$. $f(x)$ will be combined with pose code *c* and noise vector *z* in order to generate a $(320 + N^p + N^z)$-dimension combined vector which $N^p$ is the total number of pose and $N^z$ is noise vector dimension. A sequence of fractionally-strided convolutions (FConv) in decoder of generator $G_{dec}$ transforms combined vector into a synthetic 96x96-dimension image $\hat{x} = G(x, c, z)$ [8]. The face generation process is shown in figure 2.

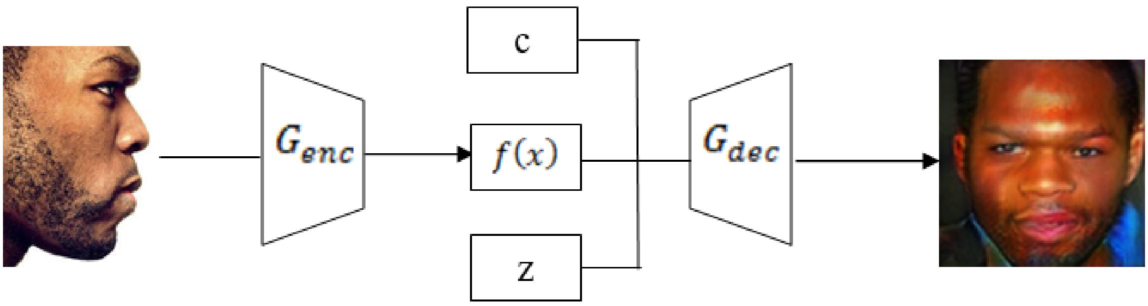


***Figure 2**. Frontal face generation.*

### 2.2. Facial feature extraction.

In FaceNet, input images are transformed into a set of 128-dimension vectors in which Euclidean distance of similar faces is shortest and vice versa. These vectors are called facial feature vector or embedding vector, which are used for training in the face recognition system.

Face recognition model:

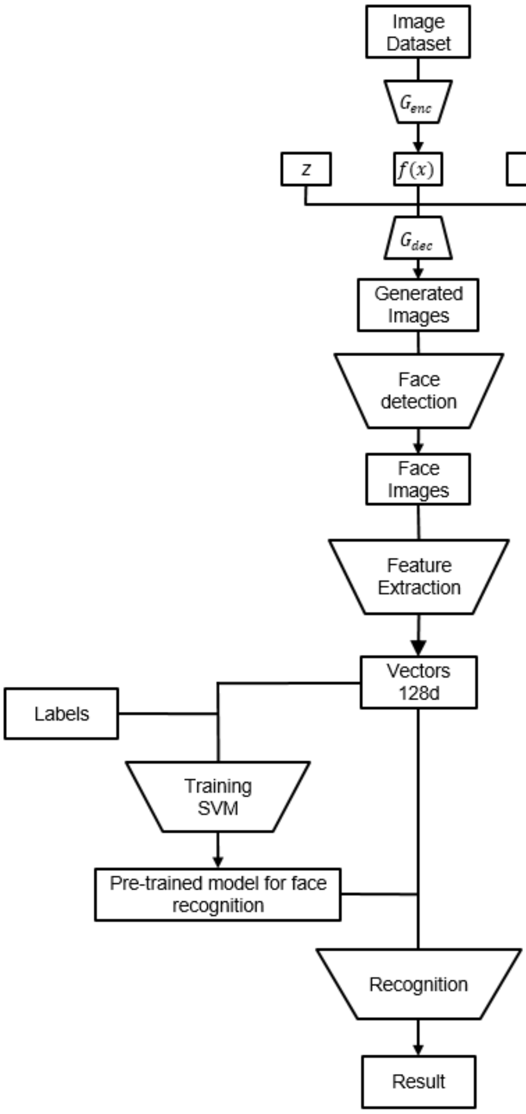


***Figure 3****. Training and recognizing processes*.

Where:

- $Image\ Dataset$ : A small data set includes frontal face images and lateral face images of each individual.
- $G_{enc}$ : encoder of *generator* in DR-GAN.
- $G_{dec}$ : decoder of *generator* in DR-GAN.
- $f(x)$ : 320-dimension feature vector given by encoder.
- $c$ : pose code.
- $z$ : noise vector.
- $Vectors\ 128d$ : 128-dimension facial vector given by FaceNet.

## 3. EXPERIMENTAL AND RESULTS.

### 3.1. Data


***Figure 4**. Some photos in training data set.*

This research study uses a small sample of data in CFP (Celebrities in Frontal-Profile in the Wild) dataset. The sample consists of 10 frontal-face images and 4 lateral-face images of several celebrities, which will be used as input data. The accuracy of system depends largely on the amount of input data. However, through practical training in a small scale, the accuracy of the proposed face recognition system is relatively positive.

### 3.2. Image generation using GAN.

With the sample input, Disentangled Representation Learning GAN (DR-GAN) then generates a new set of images that share several features with the input, and recognizable by the face-detecting methods.

DR-GAN generates fake images look like original images of training data.

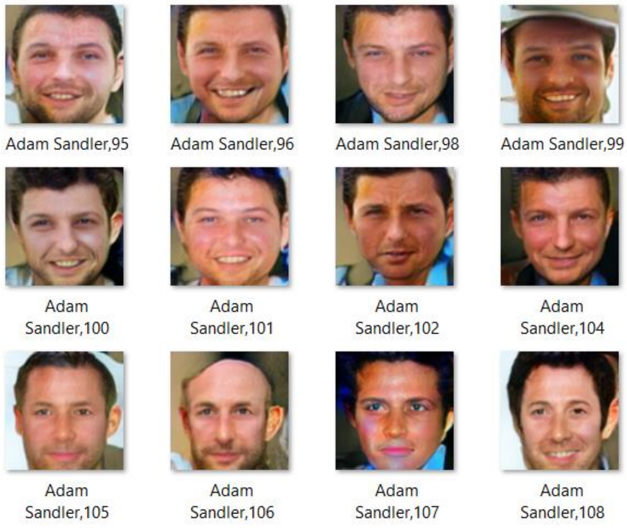


***Figure 5.*** *Generated images.*

In figure 5, the images were generated from Adam Sandler's sample. In general, these generated images have Adam Sandler's facial features. Each image was indexed to name with numerical order for testing purpose.

### 3.3. Face identification.

The output vectors of Facial feature extraction are then identified by the SVM using pre-trained models.

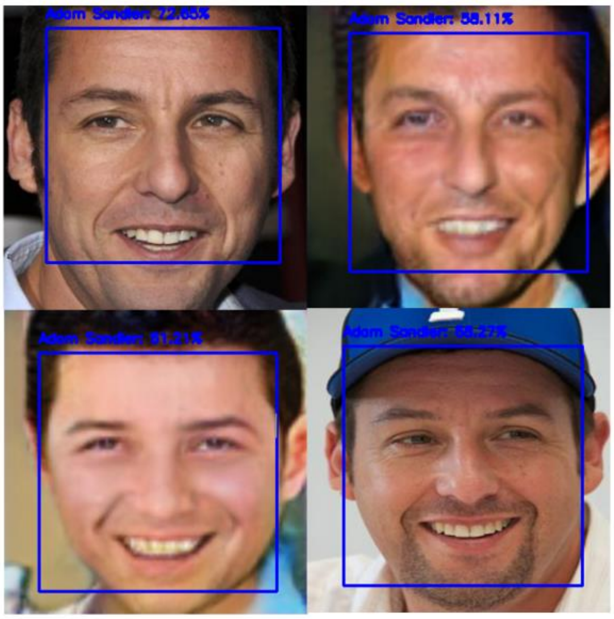

***Figure 6.*** *Results of face identification.*

In figure 5, Adam Sandler's images (both generated and original images) are correctly identified. The bounding box shown face position and the result of identification including name and reliability.

* Result evaluation.

The result of face identification highlights that the proposed system is able to provide a favorable result compared to the conventional system.

With a test sample of 18 images (not included in the training data, including frontal and lateral face images), conventional system is only able to recognize correctly 4 out 18 images (only the frontal ones), with the highest reliability of 90.12%. Whereas, the proposed system using GAN gives better test result with 11 out 18 images (both frontal and lateral) correctly recognized, and the highest reliability is 93.98%.


***Figure 7.*** *The results of the image generation from the training data set and comparison of the identification results between the conventional face recognition system and the proposed one using GAN.*

In figure 7, the training data includes both generated and original images. As the result, conventional face recognition system only is able to identify some individuals by their frontal face. Whereas, the proposed system using GAN can identify someone by both their frontal and lateral faces. Images that are correctly recognized are those with the reliability indicator beneath.

## 4. CONCLUSION

This paper proposed the face recognition framework that can recognize the face at difference poses. Generative Adversarial Network can act as processing tool that generate the synthesis faces in the case lack of training images and poses. This part helps to solve the problem of the pose difference between two faces by fix the pose and learn the representation of features from the face. The results show our proposed method outperformed the conventional method.

## REFERENCES

[1] Mayank Agarwal, Nikunj Jain, Mr. Manish Kumar and Himanshu Agrawal, "Face Recognition Using Eigen Faces and Artificial Neural Network", International Journal of Computer Theory and Engineering, Vol. 2 (4), pp. 1793-8201, August, 2010.

[2] Turk, M., and Pentland, A., "Eigenfaces for recognition", Journal of Cognitive Neuroscience, Vol. 3, pp. 71-86, 1991.

[3] Kan, M., Shan, S., Chen, X.: Multi-view deep network for cross-view classification. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp.4847–4855, 2016.

[4] Avishek Bhattacharjee, Samik Banerjee, and Sukhendu Das, "PosIX-GAN: Generating multiple poses using GAN for Pose-Invariant Face Recognition", ECCV 2018.

[5] Dong, C.; Cao, X.; Fang, W.; Jian, S. Blessing of Dimensionality: High-Dimensional Feature and Its Efficient Compression for Face Verification. In Proceedings of the IEEE Conference on Computer Vision & Pattern Recognition, Portland, OR, USA, 23–28 June 2013, pp. 3025–3032.

[6] Ding, C.; Choi, J.; Tao, D.; Davis, L.S. Multi-Directional Multi-Level Dual-Cross Patterns for Robust Face Recognition. IEEE Trans. Pattern Anal. Mach. Intell. 2016, 38, 518–531.

[7] Weiwei Zhuang, Liang Chen, Chaoqun Hong, Yuxin Liang, and Keshou Wu, "FT-GAN: Face Transformation with Key Points Alignment for Pose-Invariant Face Recognition", electronics, 2019.

[8] Luan Tran, Xi Yin, Xiaoming Liu (2017), Disentangled Representation Learning GAN for Pose-Invariant Face Recognition, Department of Computer Science and Engineering Michigan State University, East Lansing MI 48824.

[9] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In NIPS, 2014.