International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-10 Issue-10, August 2021

Students Roll Call at Vietnam Universities with Convolutional Neural Networks (Cnn)

Phan Thi Ha, Trinh Thi Van Anh

Abstract: ***In this article, we introduce the construction and training of a CNN model for student face recognition for automatic roll call. Collection of training and testing data, data preprocessing, model training and result evaluation are demonstrated and explained in detail. The model has been tested on the faces of students in a university in Vietnam and yields positive result, which can be applied in practice for automatic roll call. At the same time, the authors have built a learning management system for classes at a university in Vietnam, including automatic roll call.***

Keyword: ***Deep Learning, Convolutional Neural Networks (CNN), ResNet, k-NN, TMCNN.***

I. INTRODUCTION

Currently, classroom management at universities in Vietnam has not been unified and takes a lot of time, such as managing grades, entering grades, calculating attendance grades, sending course materials to students. Particularly, taking student attendance is still still manual and affecting students' reception of knowledge and the quality of students at school. In this context, the authors have built and developed an automatic roll call and class management system, which is an application of the face recognition in the field of image processing.

A face recognition system usually consists of three basic steps:

- Preprocessing: the system receives a still image and then processes the image for better quality, such as re-brightness, noise reduction, ... This helps to simplify the process of extracting facial features. Preprocessing is usually quite simple and fast, so there is no need to use complicated and time-consuming algorithms.
- Feature extraction: Feature extraction is a technique that uses algorithms to extract information that has unique characteristics of a person..
- Face recognition: after the features have been extracted, it will be transferred into the recognition block to classify objects.

Manuscript received on August 17, 2021.
Revised Manuscript received on August 23, 2021.
Manuscript published on August 30, 2021.
* Correspondence Author
Dr. Phan Thi Ha*, Lecturer, Faculty of Information Technology, Posts and Telecommunications Institute of Technology (PTIT), HaNoi, Vietnam, and Computing Fundamental Department, FPT University, Hanoi, VietNam. Email: hapt@ptit.edu.vn
ThS. Trinh Thi Van Anh, Lecturer, Faculty of Information Technology, Posts and Telecommunications Institute of Technology (PTIT), HaNoi, Vietnam, and Computing Fundamental Department, FPT University, Hanoi, VietNam. Email: vanh22@yahoo.com

To serve face recognition, the authors used Deep Learning [1,4] with a convolutional neural network (CNN) model [2,4,3,5,6,10], specifically ResNet101 to classify and get feature vectors, k-NN algorithm [7, 8,9] for identification.

II. MODEL OVERVIEW OF FACE RECOGNITION BASED ON CNN FOR ROLL CALL PROBLEM

A. Overview model of face recognition for roll call problem

Figure [1] represents of the overview model of face recognition based on CNN that the author's team built.

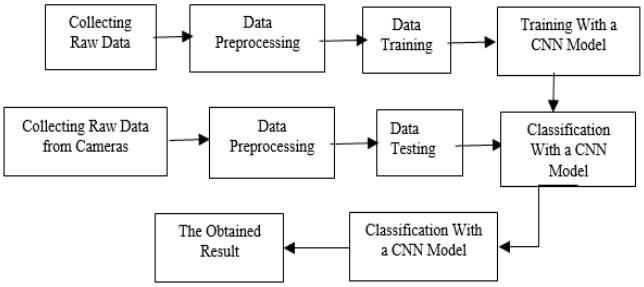

Figure [1]. Overview model of face recognition

Build Training Data

Collecting raw data directly from taking pictures from devices (e.g. phones, cameras) or collecting them online, these images will typically be colored images. This work is done by the author's team through 3 steps:

First step:

Source 1: Collecting photos of some famous Vietnamese singers, actors, models, players, etc. via Google Image (estimated about 5000 photos)

Source 2: Taking pictures from the phone of some students at school (estimated about 500 photos).

Second step:

Use Window's Paint 3D application to crop the face with a 1:1 ratio

Third step:

Use Windows' Paint 3D application to resize the image to 128x128.

The raw data set used as training data includes more than 5500 images of human faces. Each folder stores photos of a person and each person will have many photos at different shooting angles.

Retrieval Number: 100.1/ijitee.J944408101021
DOI: 10.35940/ijitee.J9444.08101021
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication

Data Preprocessing

With the input of colored images (RGB color system), the authors perform the data preprocessing steps: Resize the image to the 112x112, face detection. These preprocessing steps are performed in Python language. Here are the steps in the face recognition problem:

First step: Use Opencv [12] to resize the image to 112x112 using the resize function:

img = cv2.resize(cv2.imread(im),(112,112))

Second step: Face detection by using the MTCNN model [11], we can easily detect the face. The output of MTCNN is the position of the face and points on the face such as eyes, nose, mouth.

Build Test Data

Data collection: The data obtained to construct the test data is taken directly from the camera lite (real time). With OpenCV [9], we can easily turn on the camera to get image data through function VideoCapture.

Data preprocessing: The data to make the test dataset only needs to go through the face detection step without having to be resized to 112x112. With MTCNN, we can also easily detect faces.

B. Model training

From the data obtained after the process of MTCNN, faces can be identified and assigned names to each image set. Call this dataset X={X_j } where X_j is the data of the j^{th} person.

Set the ResNet101 architectural model as R (Wi, L, O), where Wi is the weight at the ith layer, L is the arcface loss function, and O is the Adam optimal function.

We have

$$L = -\frac{1}{N}\sum_{i=1}^{N}\log\left(\frac{e^{\cos\theta_{yi}}}{\sum_{k}^{n} e^{\cos\theta_{yk}}}\right)$$

$$W_t = W_{t-1} - \mu\frac{\hat{m}}{\sqrt{\hat{v}}+\epsilon}$$

In which N(batch size) = 20, n (class number) = 152, $\theta_{yk} = \frac{W_{yk}X_k}{\|W_{yk}\|\|X_k\|}$, W_t is the weight in the t^{th} update, $\hat{m}_t = \frac{m_t}{1-\beta_1^t}; \hat{v}_t = \frac{v_t}{1-\beta_2^t}$, μ(learning rate)=0.001, Iterations= 10000, epoch=10000.

Device: CPU 16 Core hyper-threading, Ram 60 GB, GPU Nvidia Tesla v100

Training rate: 400 sample / 1 second

Training time: 10000*5500/400 = 137 500 seconds, approximately 39 hour.

III. BUILDING A CLASS MANAGEMENT AND ROLL CALL SYSTEM

A. Agents (Participants) and system function

The system has 3 main agents involved: administrators, lecturers and students, all of which share some functions called user function. In addition, each agent has additional specific functions. Specifically:

- **User**

- Login: the user logs in when there is an account to access the system. The system processes and redirects to the corresponding user screen
- Logout: user selects the logout button to log out of the system
- Change password: users change password when needed
- Update information: users update personal information when personal information changes

- **Administrator**

- Has the full function of the user
- Course management: the administrator chooses subject management to add, edit and delete courses
- Lecturer management
- Student management
- Class management

- **Lecturer**

- Has the full function of the user
- View timetable: lecturer selects the timetable to view the teaching schedule
- Manage teaching materials: Lecturer chooses teaching materials management such as adding, editing, deleting
- Grade management and automatic roll call: lecturer chooses grade management to perform component grading for students and take attendance.

- **Student**

- View timetable: student selects a timetable to view their individual class schedule.
- View component grades: student selects component grades to view component grades and view attendance.
- View documents: student selects documents to view course materials.

B. Overview system architecture and entity class diagram

Figure [2] depicts the overview architecture of the system; the system manages 3 main work blocks:
Admin Sever: Functions that serve administrators;
Lecturer Service: Functions that serve lecturers;
Student Service: Functions that serve students;
All 3 of these work blocks use the same data source stored at Database. Here, we use the database management system MySQL for centralized data management.

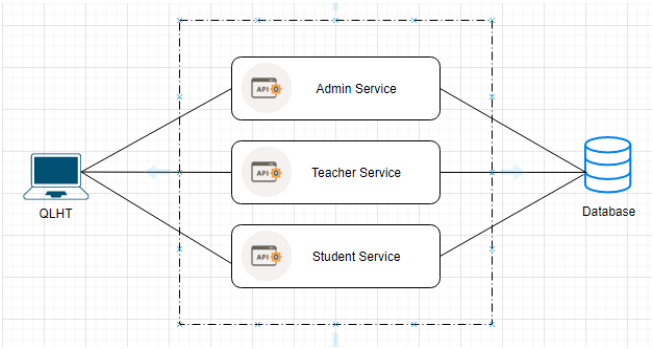

Figure [2]. Overview architecture of the system

Figure [3], is an entity class diagram built to serve the system's program writing process

Retrieval Number: 100.1/ijitee.J944408101021
DOI: 10.35940/ijitee.J9444.08101021
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication

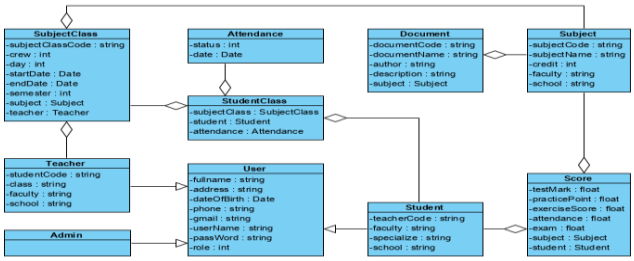

Figure [3]. Entity Class Diagram for software system development

IV. EVALUATION AND RESULT

In this part, we would show the evaluating accuracy of student roll call system base on CNN model with training text and test built in part II. Besides, we also present result of student roll call system.

A. Evaluating the model with the students' dataset

To evaluate the accuracy of the roll call system, we used independent test and training data (presented in part II). The results obtained are shown in Table [4]. Accuracy is calculated using the accuracy_score function [12]

Table [4]. Reliability of identification

Test kit size	Training data	Reliability
275	5225	98,55%
550	4950	98,55%
1100	4400	98,00%

In reality, it is relatively difficult to collect human face images as datasets. The average person only has about 40 photos of faces. This number is quite limited. The more images a person has to make datasets, the higher the probability of correct identification. However, with this test system, facial recognition also gives quite a positive result..

B. Result of system

In this section, we mainly demonstrate the system's automatic roll call function. The camera will automatically recognize the student's face and save the information to the log.txt file with 3 attributes: student code, date, time, Figure [4]. A student can appear multiple times in the log.txt file with different real-times, Figure [5].

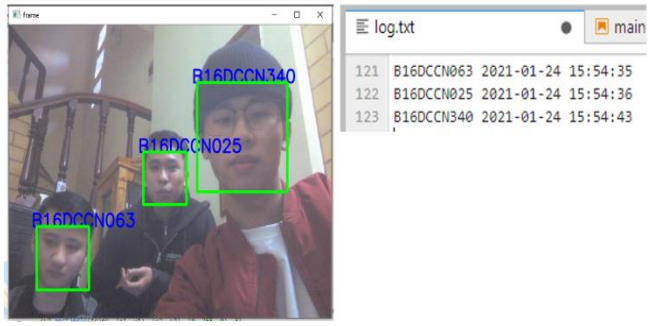

Figure [4]. Camera recoginizing students' faces

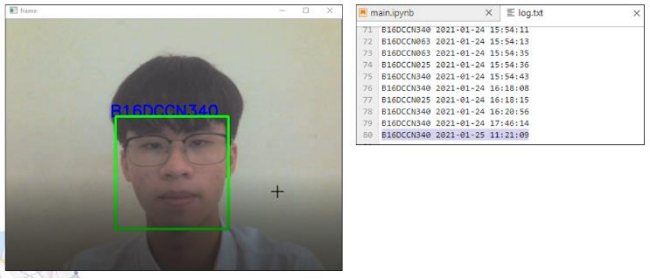

Figure [5]. Camera recoginizing student's face

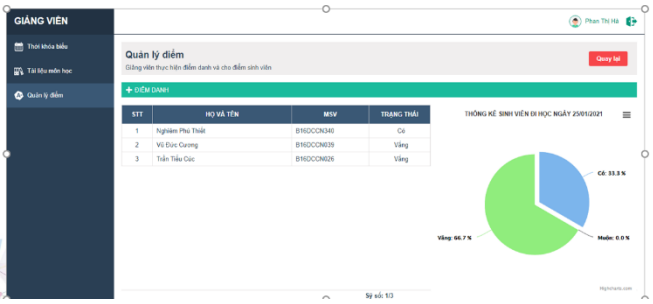

Figure [6]. Student roll call list

Then, the system will match the students in the log.txt file accordingly and on the roll call screen will appear result "Yes" if that student is in the class list. If not in the list, that student will be marked as "Absent" in the log file, Figure [6].

In this section, we mainly describe the automatic roll call function of the system. The camera will automatically recognize the student's face and save the information to the log.txt file with 3 attributes: student code, date, time, Figure [5]. A student can appear many times in the log.txt file with different real times, and many faces can be recognized at one time, Figure [5].

V. CONCLUSION

The article presents the construction of a CNN model for the face recognition problem, building a training and testing data set for the model including more than 5500 raw images from Vietnamese students. The model test shows that the recognition results are quite positive, which can be used for student roll call in university classes in Vietnam. At the same time, the article also introduces the class management system and automatic roll call at universities, which is a very useful and necessary for administrators, lecturers and students at universities in Vietnam.

REFERENCES

1. A. L. Beam, Deep Learning 101, 2017.
2. C. H. R. K. &. R. Samer, Image Recognition Using Convolutional Neural Networks., Cadence Whitepaper, 2015.
3. Xiaoyue Jiang , Abdenour Hadid , Yanwei Pang , Eric Granger, Xiaoyi Feng , Deep Learning in Object Detection and Recognition, Springer,2019
4. Lan Goodfellow, Yoshua Bengio, Aaron Courville, Deep Learning, The MIT Press, 2016
5. Le Lu, Xiaosong Wang, Gustavo Carneiro, Lin Yang, Deep Learning and Convolutional Neural Networks for Medical Imaging and Clinical Informatics, Springer, 2019
6. H. E. Geoffrey, I. Sutskever and K. Alex, ImageNet Classification with Deep Convolutional, 2012.

7. Jo, Taeho, Machine Learning Foundations, Springer, 2021.
8. T. M.Mitchell, Machine Learning, McGraw-Hill, 1997.
9. T. H. Vũ, Machine Learning cơ bản, 2017.
10. Thi Ha Phan, Duc Chung Tran, Mohd Fadzil Hassan, Vietnamese character recognition based on CNN model with reduced character classes, Bulletin of Electrical Engineering and Informatics Vol. 10, No. 2, April 2021, pp. 962~969
11. Kaipeng Zhang, Zhanpeng Zhang, Zhifeng Li, Yu Qiao, Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks, IEEE, page 1499 1503, 2016
12. https://github.com/skvark/opencv-python.

AUTHORS PROFILE

Dr. Phan Thi Ha, is currently a lecturer at the Faculty of Information Technology at Posts and Telecommunications Institute of Technology (PTIT) in Vietnam, and Computing Fundamental Department, FPT University, Hanoi 10000, Vietnam as well. She received a B.Sc.in Math & Informatics, a M.Sc. in Mathematic Guarantee for Computer Systems and a PhD. in Information Systems in 1994, 2000 and 2013, respectively. Her research interests include machine learning, natural language processing and mathematics applications.Email: hapt@ptit.edu.vn, hapt27@fe.edu.vn and hathiphan@yahhoo.com

ThS. Trinh Thi Van Anh, is currently a lecturer and a researcher at the Faculty of Information Technology at Posts and Telecommunications Institute of Technology (PTIT) in Vietnam and Computing Fundamental Department, FPT University, Hanoi 10000, Vietnam as well. She received a B.Sc. in Electronics-Telecommunications in 1993 (HUST), M.Sc. in Computer Science in 1998 (HUST). Her research interests include machine learning, NLP and opinion mining. Email: vanh22@yahoo.com, anhttv@ptit.edu.vn and anhttv20@fe.edu.vn