International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Facial Recognition in the Opening of a Door using Deep Learning and a Cloud Service

Gautier Abou Loume [1], Alphonse Binele Abana*[1], Emmanuel Tonye[1], Yvan Kabiena[2]

Submitted: 22/**07**/**2022** **Accepted:** 25/**09**/**2022**

Abstract: We propose an intelligent access control device to overcome certain limits of reliability of traditional systems such as systems based on knowledge of the user (password, PIN code, etc.) or hardware (badge, magnetic card, etc.). Indeed, traditional systems have, among other limitations, such as forgetfulness, theft, loss and falsification. Our system consists of a biometric access control application based on facial recognition and embedded in a Raspberry Pi nano-computer controlling the automatic opening of a door. This device performs the following actions when a person approaches the door: motion detection, real-time shooting of the scene thanks to a connected camera, face detection on the captured image followed by 'facial identification, opening of the door in the event of recognition and notification by email and SMS using SaaS type Cloud services to the owner of the device. It thus makes it possible to make a door automatic and intelligent, to improve the reliability of physical access control systems and consequently to improve the safety of people, goods and services.

Keywords: *A*ccess control, deep learning, facial recognition, IoT.

1. Introduction

The IoT (Internet of Things) is considered as a real revolution in the digital world; it has many applications, in particular: the intelligent transport network, the intelligent electrical network, intelligent agriculture, telemedicine, the smart city and in this case the smart home. In this sense, it is therefore necessary to develop physical access control systems to a building, a house, a room, a place, a machine or specific equipment (safe, vehicle, etc.) compatible with the Internet. Objects. These systems must not only be more secure but also automatic and intelligent. To this end, there are three automatic physical access control techniques: those based on the knowledge of the user (PIN code, password, etc.), those based on hardware (badge, key, card, etc.) and those that rely on biometrics.

Traditional knowledge-based and hardware-based access control systems do not meet these security requirements. On the one hand, the security code (password, PIN code) can be forgotten or guessed and on the other hand, the material (badge, card, etc.) is subject to theft, loss or falsification. This is why biometric systems appear as an alternative to these traditional systems because of their much higher reliability, their universality and also their uniqueness per individual. In this study, we are particularly interested in the facial recognition system which is one of the most used biometric techniques and the most natural way to identify an individual for humans. However, far from being a panacea for physical access control, biometric systems in general face many difficulties. In the particular case of facial recognition, the following difficulties can be listed: the influence of variations in pose, change of lighting, concealment or facial expressions, for example [1]. Automatic physical access control systems, in this case facial rebirth, are generally implemented at a door level. We will then speak of a connected or intelligent door in the jargon of the Internet of Things (IoT).

In this work, we realize an application of facial recognition causing the opening of a door, and exploiting a Raspberry Pi nano-computer as well as Cloud Computing services. The Raspberry Pi is at the heart of this one-door restrictive access control process. Our concern is to improve the reliability of physical access control systems. Our choice of the facial biometric system was guided by its many advantages: high acceptance, non-intrusiveness, reduced size, ergonomics, easy use, contactless use, low implementation cost and also by the fact that it is the method of natural identification for humans [2]. The facial recognition access control system also responds to concerns about social distancing, and limitations on contact with animals and objects resulting from the advent of the Covid-19 pandemic crisis, because its use is contactless. The solution we propose falls within the framework of biometric and automatic physical access control, which can be used

[1] University of Yaounde I, National Advanced School of Enginering, Department of Electrical and Telecommunications Engineering, Cameroon.
*ORCID ID : **0000-0003-2891-347X***
[2] University of Douala, National Advanced School of Engineering of Douala, Department of Telecommunication and Information and Communication Technologies Engineering, JAPAN
*ORCID ID : **0000-0003-2891-347X***
** Corresponding Author Email: binelabana@gmail.com*

for IoT applications.

2. Facial recognition techniques

Facial recognition, also called face recognition, involves identifying one or more people automatically in photos or videos by analyzing and comparing shapes.[4] In the literature, the classification of facial recognition techniques is practically that of image recognition. Thus, a distinction is made between global or holistic methods, local methods and hybrid methods. The figure below illustrates different facial recognition approaches with examples of the algorithms that implement them.

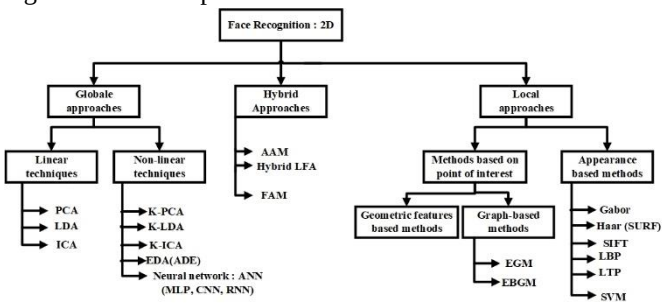

Fig. 1. Facial recognition approaches [5]

Table 1. Meaning of acronyms for different facial recognition approaches

Approach	Advantages	Disadvantages
Global	- Simplicity and speed of implementation - Applicable to low resolution images.	- Sensitivity to variations in illumination, pose, and facial expression. - Large storage capacity, need for sufficient data representative of faces
Hybrids	- Benefits of Global and Local Approaches	- Disadvantages of global and local approaches
Local	- Robustness to variations in illumination, pose and facial expression - Additional information based on local parts - Choice of the most suitable classifier for each type of local feature	- Difficulties in extracting points of interest - Need for good quality images

The meanings of the acronyms used in Fig.1 are listed in Table 2.

Table 2. Meanings of acronyms for different facial recognition approaches

Letter	Acronym	Meaning	Acronym	Meaning
A	AAM	Active Appearance Model	ANN	Artificial Neural Network
C	CNN	Convolutional Neural Networks		
E	EBGM	Elastic Buch Graph Matching	EDA	Exponential Discriminant Analysis
	EGM	Elastic Graph Matching		
F	FAM	Flexible Appearance Model		
I	ICA	Independant Component Analysis		
K	K-ICA	Kernel Independant Component Analysis	K-LDA	Kernel Linear Discriminent Analysis
	K-PCA	Kernel Principle Component Analysis		
L	LBP	Local Binary Pattern	LDA	Linear Discriminent Analysis
	LFA	Local Feature Analysis	LTP	Local Ternary Pattern
M	MLP	Multilayer Perceptron		
P	PCA	Principle Component Analysis		
R	RNN	Recurrent Neural Network		
S	SIFT	Scale-Invariant Feature Transform	SURF	Speeded Up Robust Features
	SVM	Support-Vector Machine		

Global methods and local methods are practically complementary. Hybrid methods, which are either a combination of global and local techniques or new techniques based on statistical models, improve the performance of facial recognition [1].

3. Scientific and technological approaches used

3.1. Deep learning

Deep Learning is a branch of Machine Learning that is based on the principle of artificial neural networks (ANN), however used on a much larger scale [7].

An artificial neural network is a collection of formal neurons associated in layers and operating in parallel. [8].

The formal neuron also called artificial neuron is the elementary unit of the ANN which represents and simulates the functioning of the biological neuron. It performs the

weighted sum of its inputs before passing it to an activation function to produce an output result [8].

In this project we use a particular model of neural networks called convolutional neural network (CNN) for the facial recognition algorithm because they are suitable for image processing. [7][9].

3.2. Transfer learning

It is a learning technique, mainly used in Deep Learning, to train a neural network to a certain task from a model already trained on a similar task. This form of learning has the advantage of requiring less data (thousands instead of millions) and allows much faster learning, going from hundreds or thousands of hours of calculations to a few hours or even a few minutes. [7] In this work, we use the pre-trained CNN VGG-16 network for face recognition.

3.3. Deep learning model used

We use VGG-16 (Visual Geometry Group 16 layers), a neural network of 16 deep layers (13 layers of convolutions and 3 fully connected) [10] to build our model using the Transfer Learning technique. The model takes as input a multichannel image (color image: three channels, namely Red, Green and Blue) of size 224×224 pixels [11].

For this, we will first retrieve all the layers and weights of the pre-trained model VGG16. Only the last dense layer, the classification layer will be replaced by a classifier adapted to our problem. In our case, it will consist of 7 classes (initially 1000) which represent the total number of individuals (people) in our dataset. It is this layer that will make the recognition thanks to a match score. This is Transfer Learning by feature extraction strategy [7].

In addition, we use for our model:

The softmax activation function [7] for the last fully connected layer:

$$f_j(z) = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}}, \quad \forall\, j \in [1;K]\ avec \sum_{j=1}^{K} f_{j(z)} = 1 \qquad (\boldsymbol{E.1})$$

The cross-entropy loss function [7] adapted to our multiple classification problem (multi class).

$$H(p,q) = -\sum_{i=1}^{n} (\, p(i) \log_2(q(i)) \qquad (\boldsymbol{E.2})$$

The Adam optimization function (abbreviation of Adaptive moment estimation), which is a stochastic gradient descent method [12].

And the metric accuracy which merges with the recognition rate in the case of facial identification.

$$acc = RR = \frac{Number\ of\ correctly\ identified\ images}{Total\ number\ of\ images} \qquad (\boldsymbol{E.3})$$

The three-dimensional (3D) representation of the architecture of VGG-16 used in this work is given by fig. 2. It presents the feature extraction from the face image captured (in color: three channels) of size 224×224 pixels.

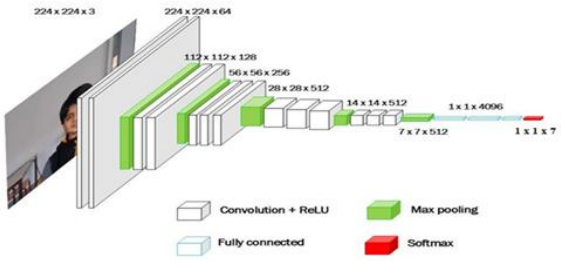

Fig. 2. 3D architecture of the neural network used [13] (modified by the author)

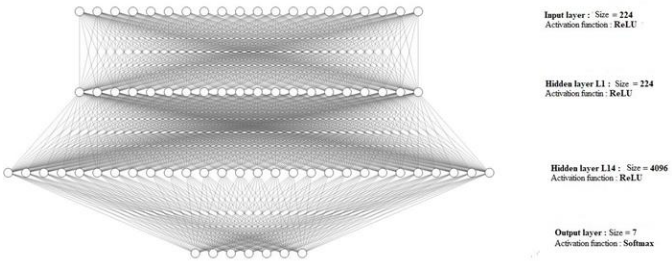

Fig 3. Overview of the 2D architecture of the neural network

Figure 3 presents an overview of the two-dimensional (2D) architecture of the model used.

The size of the data taken as input to our neural network is 224. And the output is a vector of size 7 which represents a probability distribution for the 7 classes of our dataset. We have $= [f_1\ f_2\ f_3\ f_4\ f_5\ f_6\ f_7\ ]$, with f_j , $1 \leq j \leq 7$: the probability that the predicted class is individual j (f being the Softmax function) of our dataset.

For example, in our digital experimentation, we have : $Z = [0.00\ 0.01\ 0.00\ 0.04\ 0.01\ 0.92\ 0.02]$. Thus, we can predict with a probability of 0.92 that the captured face corresponds to person 6 in our dataset.

The hidden layer L1 has a size of 224 and an activation function, the ReLu function [7]. This function is the same for all the other layers of convolutions (12 others). It also applies to other fully connected layers except the last one.

$\boldsymbol{ReLu}: f(x) = \max(0,x)$ (E4)

The activation function used at the output of the neural network is the Softmax function.

3.4. Model training

The dataset used for training the model includes seven (07) different individuals each having twenty (20) images. Which makes a total of 140 images. These images are distributed as follows:

- For the training data set: we use three quarters (3/4) of images per individual, i.e. fifteen (15). Which makes a total of 105 images.

- For the trial or test dataset, we use a quarter of the images per individual, i.e. 5. A total of 35 images are intended to evaluate our learning model.

In addition, we use the Data Augmentation technique on all the face images to artificially increase our training data [7][14]. The other essential hyper parameters for training our model are:

- The batch size equal to 5. It allows us to use less memory and also accelerates learning. This choice of a small number of batch is justified by the technical constraints of the Raspberry Pi embedded system which is somewhat limited in memory (RAM) and processing capacity (CPU).
- The epoch number is 10. Thus, our complete dataset will be presented ten (10) times to our model.

The size of the input images to train our model conforms to that of VGG-16, a three-channel image of size 224 x 224 pixels. It has a total of 14,890,311 parameters of which only 175,623 will be trained.

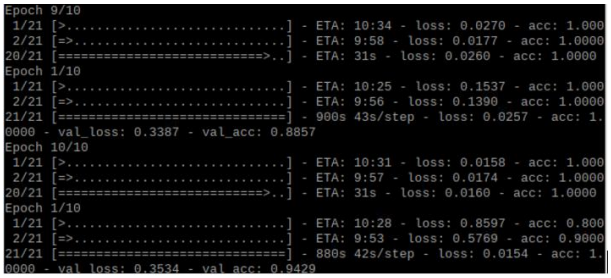

Fig. 4. An overview of model training results

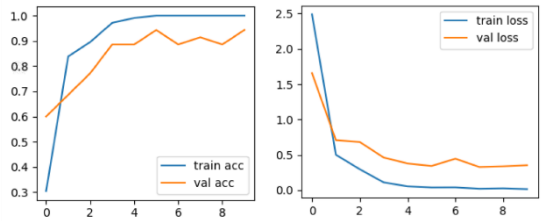

Fig.5 *Losses and accuracies for the ten "epochs" of learning and validation*

Figure 5 illustrates the loss (loss) and accuracies (acc) curves for the training (train) and validation (val) of our model. At the end of the learning for the ten epochs, we obtain with this model an optimal accuracy of 100% for the training and 94.29% for the validation. This test accuracy is practically greather than that of VGG-16 on ImageNet which is 92.7%. We also record a minimal loss of 0.0154 (0.0154 for epoch 10) for training and 0.3291 (0.3534 for epoch 10) for testing. This allowed us to validate this learning model.

4. Device modeling

4.1. Architecture of ΠCRF

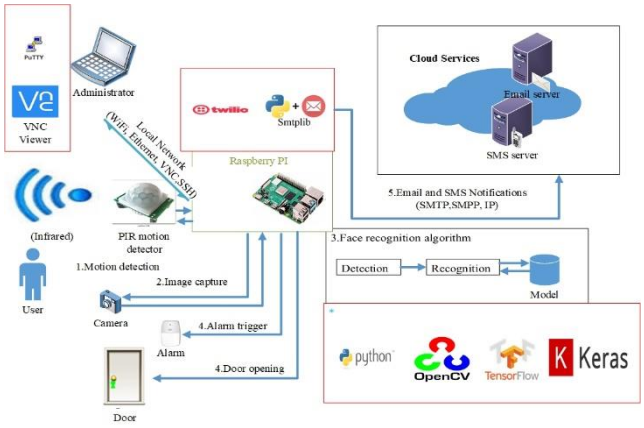

Fig. 6. System architecture

The Fig. 6 shows the architecture of the ΠCRF device. When a person arrives, he approaches the door, there is:

(1) Motion detection of the person using the PIR module;

(2) Shooting of the face of the person who wishes the door to be opened using a connected camera. This image is transmitted to the Raspberry Pi Nano-computer;

(3) Facial identification based on a Transfer Learning model of a VGG-16 convolutional neural network (CNN). This algorithm is embedded and executed by the Raspberry Pi;

(4) Automatic door opening upon face recognition;

(5) Notification by SMS and Email through servers hosted in the Cloud.

A computer connected to the same local area network (LAN) as the Raspberry Pi is used to manage, monitor, control or maintain the tool remotely.

4.2. Device use case diagram

Fig. 7 highlights the interactions between the different internal actors (average user, administrator), external actors (SMSC and mail servers), and the ΠCRF system.

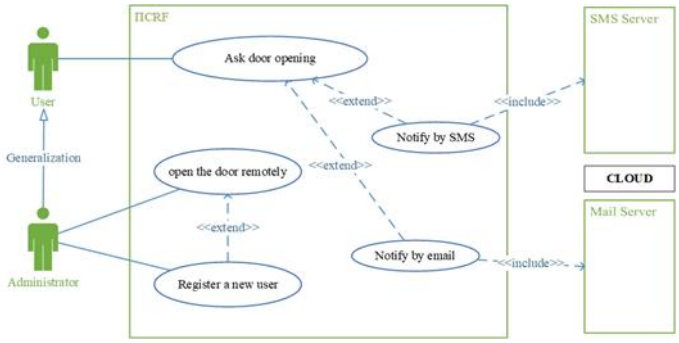

Fig. 7. Diagram of the use cases of our system

4.3. User sequence diagram

Fig. 8 highlights the sequence of interaction between a user and the various subsystems of the device (PIR motion detector, camera, door simulated by the servomotor, SMSC

server and mail server) on the one hand and the administrator on the other hand.

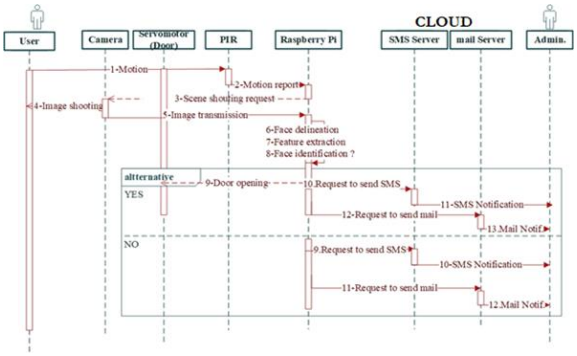

Fig 8. *User real-time facial recognition sequence diagram*

4.4. Facial recognition subsystem

A facial recognition system generally includes the following steps:

- Face detection;
- Feature extraction;
- Classification and Decision.

Feature extraction and classification are performed by a special convolutional neural network (CNN) VGG-16:

4.4.1. Delineation of the facial area

When there is motion detection (by the PIR sensor), our facial recognition system takes as input an image captured by a camera connected to the Raspberry PI nano-computer and detects the face. We use for this purpose cascaded Haar classifiers [1].

4.4.2. Feature extraction

Our machine learning algorithm calculates a digital representation of the detected face as a feature vector. This feature vector is built automatically in the context of Deep Learning [15] and therefore of the CNNs used in this work.

4.4.3. Classification

The face descriptors from the previous step will be used to train the CNN-based classifier. The models or classes characterizing the faces of the individuals of interest resulting from the learning are then calculated from a database of face images of the people to be recognized. It is therefore the model that will be used in the process of recognizing a new face captured by our system on the basis of comparison of similarity between the extracted parameters and the parameters previously recorded in the system.

4.4.4. Decision

The softmax function of the last layer of the CNN network allows us to predict the face class (identity of the person) with a given probability called match score. For better inference we use a recognition threshold value (value between 0 and 1). There is no standard value for this threshold. It depends on the purpose of the facial recognition system.

It is the decision of the identification or not of the face which is used by the Raspberry Pi nano-computer to control the selective opening of the door. Thus, the door will open when the face is recognized and it will remain closed otherwise.

4.5. Cloud services

The cloud services implemented in this work fall into the category of SaaS (Software as Service).

4.5.1. SMS Notification

Fig. 9. SMS notification subsystem architecture.

ΠCRF uses the SMPP (Short Message Peer-to-Peer) communication protocol to transmit SMS to the administrator. The tool connects to the SMS gateway hosted in the cloud.

Fig. 10 represents the architecture of our SMS notification system.

4.5.2. Email notification

For e-mail notifications, the device uses a mail server based on the SMTP protocol (Simple Mail Transfer Protocol). The consultation of the mails can be carried out thanks to the protocol POP (Post Office Protocol) or to the protocol IMAP (Interactive Mail Access Protocol).

4.6. Device production flowchart

The scientific approach adopted for the realization of our solution is as follows.

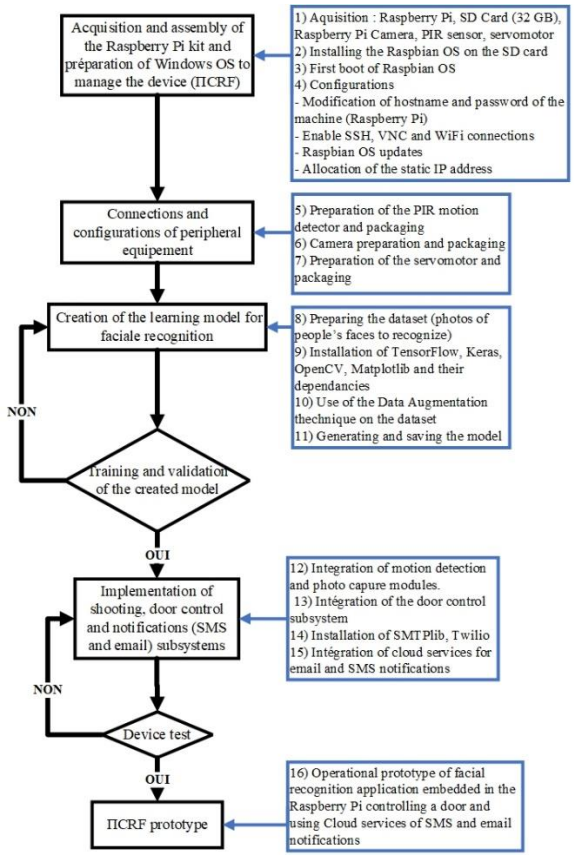

Fig 10. Flowchart of the steps of our methodology

5. Conclusion

The IICRF device is a biometric access control solution. It uses a facial recognition algorithm based on Deep Learning and Transfer Learning, and embedded in a Raspberry Pi nano-computer to control the opening of a door. It also uses Cloud services of SMS and email notifications. It thus makes it possible to make a door automatic and intelligent, to improve the reliability of physical access control systems and consequently to improve the safety of people, goods and services. It is therefore compatible with the Internet of Things (IoT).

References

[1] Souhila GUERFI ABABSA, Authentification d'individus par reconnaissance de caractéristiques biométriques liées aux visages 2D/3D, PhD thesis in signal and image processing, Evry Val d'Essonne University, 2008.

[2] Pawan Kumar Tiwari, Mukesh Kumar Yadav, R. K. G. A. . (2022). Design Simulation and Review of Solar PV Power Forecasting Using Computing Techniques. International Journal on Recent Technologies in Mechanical and Electrical Engineering, 9(5), 18–27. https://doi.org/10.17762/ijrmee.v9i5.370

[3] Nyemo KODJOVI KOUMADI, Authentification automatique du propriétaire d'un téléphone mobile, Master's thesis in Computer Science, University of Quebec at Chicoutimi (UQAC)", 2018.

[4] Sudhakar, C. V., & Reddy, G. U. . (2022). Land use Land cover change Assessment at Cement Industrial area using Landsat data-hybrid classification in part of YSR Kadapa District, Andhra Pradesh, India. International Journal of Intelligent Systems and Applications in Engineering, 10(1), 75–86. https://doi.org/10.18201/ijisae.2022.270

[5] Manitra Tsilavina RAZAFIMANDIMBY, Détection et reconnaissance de visages, Master II dissertation in Applied Mathematics, Computer Science and Statistics, University of ANTANANARIVO, 2016.

[6] Agarwal, D. A. . (2022). Advancing Privacy and Security of Internet of Things to Find Integrated Solutions. International Journal on Future Revolution in Computer Science &Amp; Communication Engineering, 8(2), 05–08. https://doi.org/10.17762/ijfrcsce.v8i2.2067

[7] Electronic Identification. Reconnaissance Faciale : son fonctionnement et sa sécurité. https://www.electronicid.eu/fr/blog/post/reconnaissance-faciale/fr (Accessed April 14, 2022)

[8] Abdelmalik OUAMANE, Reconnaissance Biométrique par Fusion Multimodale du Visage 2D et 3D, PhD Thesis in Electronics, University Mohamed Khider – Biskra, 2015.

[9] Patil, P. ., D. D. . Waghole, D. V. . Deshpande, and D. M. . Karykarte. "Sectoring Method for Improving Various QoS Parameters of Wireless Sensor Networks to Improve Lifespan of the Network". International Journal on Recent and Innovation Trends in Computing and Communication, vol. 10, no. 6, June 2022, pp. 37-43, doi:10.17762/ijritcc.v10i6.5622.

[10] Bouzit Dhikra, Reconnaissance de visage basée sur une approche triangulaire, Master's thesis in Computer Science, University of May 8, 1945 – Guelma, 2019.

[11] Pierre BONAZZA, Système de sécurité biométrique multimodal par imagerie, dédié au contrôle d'accès, Thesis in Computer Science and Image Instrumentation, University of Burgundy Franche-Comté (UBFC), 2019.

[12] Abdelkrim AAZZAB, Mimoun BENZAOUAGH, et Ahmed ABRIANE , "Application des réseaux de neurones artificiels pour la classification : cas des défaillances d'entreprises", Moroccan Journal of Business Studies(MJBS), vol. 1, no. 2, pp. 1-7, 2018.

[13] Bolivar CHACUA, Iván GARCÍA-SANTILLÁN, Paul ROSERO, et al.," People Identification through Facial Recognition using Deep Learning", ResearchGate, 2019.

[14] Hamid OUANAN, Ahmed GAGA, Omar DIOURI, Mohammed OUANAN, et B. AKSASSE, "Development of Deep Learning-Based Facial Recognition System", ResearchGate, 2020.

[15] Alaria, S. K., A. . Raj, V. Sharma, and V. Kumar. "Simulation and Analysis of Hand Gesture Recognition for Indian Sign Language Using CNN". International Journal on Recent and Innovation Trends in Computing and Communication, vol. 10, no. 4, Apr. 2022, pp. 10-14, doi:10.17762/ijritcc.v10i4.5556.

[16] Pascal MONASSE and Kimia NADJAHI, "Classify and Segment visual data, Practical work: Implement your first neural network with Keras". In: OpenClassrooms. . Updated October 21, 2021. Available on: https://openclassrooms.com/fr/courses/4470531-classez-et-segmentez-des-donnees-visuelles/5097666- tp-implement-yourfirst-neural-network-with-keras (Accessed March 13, 2022)

[17] Yann LECUN, et al. "Efficient backprop." Neural networks: Tricks of the trade.Springer Berlin Heidelberg, pp.9-48, 2012.

[18] Rizwan AHMED KHAN, Crenn Arthur, Alexandre MEYER, Saida BOUAKAZ.A novel database of children's spontaneous facial expressions (LIRIS-CSE). Barrett Hodgson University - Université Claude Bernard Lyon1, Pakistan - France, 2019.

[19] Lars ANKILE, & Morgan HEGGLAND, et Kjartan KRANGE, "Application of Facial Recognition using Convolutional Neural Networks for Entry Access Control.", ResearchGate, 2020.

[20] Francis CHARETTE MIGNEAULT, Conception de système de reconnaissance de visages spatio-temporelles sur vidéos à partir d'une seule image de référence, Master's thesis in Automated Production Engineering, University of Quebec, 2017.