DETECTING UNAUTHORIZED PERSON USING CNN AND OPENCV TECHNIQUES

J.Hemavathi, Dr. N. Syed Siraj Ahmed

1PG Scalar, 2Assistant Professor.
1Master of Computer Applications,

1Madanapalle Institute of Technology and Science, Madanapalle, Andhra Pradesh, India, 517352.

ABSTRACT

Unauthorized person, which means when someone enters into your device, system or your house without your permission is said to be unauthorized access. CNN methods are being used to handle these unauthorized access problems with the help of open cv (open source computer vision library). It is a open source computer vision and machine learning software library. Apart from that we also use 3D modelling, face detection, visual authentication techniques for capturing images of unauthorized person we also installed camera with memory facility. This project demonstrates to increase the capability of the devices we have constructed. Our proposed method integrates a better approach, intended to advance the cooperativeness of the explore operation. In this work, we develop the application with a device to eradicate the unauthorized access of unknown persons into our premises. Our application can be able to alert the persons whenever any unknown person is trying to enter into our premises.

KEYWORDS: Access, Unauthorized, Application, Deep Learning, Face Detection.

1.INTRODUCTION

Intelligent systems appear more and more in people's lives, and often need to be identified when using intelligent systems. Traditional methods of identification mainly identify individuals with some personal characteristics, such as identity documents, such as documents and keys, which have obvious shortcomings. They are easily forgotten, lost or faked. If you use some of the personal characteristics to identify the effect will be quite good, such as: face recognition, fingerprinting and so on. In terms of algorithms, there are sharing parameters between the convolution layer and the convolution layer of CNN. The advantage of this is that the memory requirements are reduced, and the number of parameters to be trained is correspondingly reduced. The performance of the algorithm is therefore improved. At the same time, in other machine learning algorithms, the pictures need us to perform preprocessing or feature extraction. However, we rarely need to do these operations when using CNN for image processing. This is something other machine learning algorithms cannot do. There are also some shortcomings in depth learning. One of them is that it requires a lot of samples to construct a depth model, which limits the application of this algorithm. Today, very good results have been achieved in the field of face recognition and license plate character recognition, so this topic will do some simple research on CNN-based face recognition technology.
2. Convolution neural network

2.1 Convolutional neural network introduction: With the development of convolutional neural networks, the achievements made in various competitions are getting better and better, making it the focus of research. In order to improve the training performance of the forward BP algorithm, an effective method is to reduce the number of learning parameters. This can be done by convolution of the spatial relationship of the neural network. Convolutional neural network, the network structure is proposed, it minimizes the input data pretreatment. In the structure of convolution neural network, the input data is input from the initial input layer, through each layer processing, and then into the other hierarchy, each layer has a convolution kernel to obtain the most significant data characteristics. The previously mentioned obvious features such as translation, rotation and the like can be obtained by this method.

2.2 Convolution neural network basic structure Neural network can be divided into two kinds, biological neural network is one of them, and artificial neural network is another kind. Here mainly introduces artificial neural network. An artificial neural network is a data model that processes information and is similar in structure to the synaptic connections in the brain. Neural network is composed of many neurons; the output of the previous neuron can be used as the input of the latter neuron. The corresponding formula is as follows:

\[ h_{W,b}(x) = f(W^T x) = f\left(\sum_{i=1}^{3} W_i x_i + b\right) \]

This unit is also called Logistic regression model. When many neurons are linked together, and when they were layered, the structure can now be called a neural network model. Figure 1 shows a neural network with hidden layers.

![Figure 1. Neural Networks.](image)

In this neural network, X1, X2, X3 are the input of the neural network. +1 is the offset node, also known as the intercept term. The leftmost column of this neural network model is the input layer of the neural network, the rightmost column of which is the output layer of the neural network. The middle layer of the network model is a hidden layer, which is fully connected between the input layer and the output layer. The values of all the nodes in the network model cannot be seen in the training sample set. By observing this neural network model, we can see that the model contains a total of 3 input units, 3 hidden units and 1 output unit. Now, use nl to represent the number of layers in the neural network, and the number of layers in this neural network is 3. Now mark each layer, the first layer can be
expressed by $L_1$, then the output layer of the neural network $L_1$, its output layer is $L_{nl}$, in this neural network, the following parameters exist:

$$(W, b) = (W^1, b^1, W^2, b^2)$$

is the connection parameter between the $j$th cell of layer 1 and the $i$th cell of layer $l+1$, and $b_i$ is the offset of the $i$th cell of layer $l+1$. In this neural network model, set $a_i(l)$ to represent the output value of the first few cells in this layer. Let $l$ denote this layer and $i$ the first few cells in this layer. Given that the set of parameters $W$ and $b$ have been given, we can use the formula $h_{w,b}(x)$ to calculate the output of this neural network. The calculation of forward propagation is as shown in equation (3). Neural network training methods and Logistic regression model is similar, but due to the multi-layered neural network, but also the need for gradient descent + chain derivation rule.

### 3. CNN Model

**Construction and Training**

3.1. **CNN model**

At present, the typical architecture of neural network is divided into the following categories: LeNet5, AlexNet, ZF Net, GooLeNet, and VGGNet, the following will LeNet5 architecture for a detailed analysis. LeNet5 is a CNN classic structure that existed long ago, and it is mainly used in the recognition of handwritten fonts. It contains a total of seven layers of structure, except for the input layer, each of the other has training parameters, and each layer contains a plurality of Feature Maps, we can extract the input features through a convolution kernel.

### 2. AN OVERVIEW

**Step1: convolutional operation**

The first building block in our plan of attack is convolution operation. In this step, we will touch on feature detectors, which basically serve as the neural network's filters. We will also discuss feature maps, learning the parameters of such maps, how patterns are detected, the layers of detection, and how the findings are mapped out.

**The Convolution Operation**

![Input Image and Feature Detector Diagram](image-url)
Step (1b): Relu Layer

The second part of this step will involve the Rectified Linear Unit or Relook. We will cover Relook layers and explore how linearity functions in the context of Convolutional Neural Networks.

Not necessary for understanding CNN's, but there's no harm in a quick lesson to improve your skills.

Step 2: Pooling Layer

In this part, we'll cover pooling and will get to understand exactly how it generally works. Our nexus here, however, will be a specific type of pooling; max pooling. We'll cover various approaches, though, including mean (or sum) pooling. This part will end with a demonstration made using a visual interactive tool that will definitely sort the whole concept out for you.

Step 3: Flattening

This will be a brief breakdown of the flattening process and how we move from pooled to flattened layers when working with Convolutional Neural Networks.
Step 4: Full Connection

In this part, everything that we covered throughout the section will be merged together. By learning this, you'll get to envision a fuller picture of how Convolutional Neural Networks operate and how the "neurons" that are finally produced learn the classification of images.

Summary

In the end, we'll wrap everything up and give a quick recap of the concept covered in the section. If you feel like it will do you any benefit (and it probably will), you should check out the extra tutorial in which Soft ax and Cross-Entropy are covered. It's not mandatory for the course, but you will likely come across these concepts when working with Convolutional Neural Networks and it will do you a lot of good to be familiar with them.

3. LITERATURE REVIEW

Literature survey is the most important step in software development process. Before developing the tool it is necessary to determine the time factor, economy and company strength. Once these things are satisfied, ten next steps are to determine which operating system and language can be used for developing the tool. Once the programmers start building the tool the programmers need lot of external support. This support can be obtained from senior programmers, from book or from websites. Before building the system the above consideration are taken into account for developing the proposed system.


With the transition of facial expression recognition (FER) from laboratory-controlled to challenging in-the-wild conditions and the recent success of deep learning techniques in various fields, deep neural networks have increasingly been leveraged to learn discriminative representations for automatic FER. Recent deep FER systems generally focus on two important issues: overfitting caused by a lack of sufficient training data and expression-unrelated variations, such as illumination, head pose and identity bias. In this paper, we provide a comprehensive survey on deep FER, including datasets and algorithms that provide insights into these intrinsic problems. First, we introduce the available datasets that are widely used in the literature and provide accepted data selection and evaluation principles for these datasets. We then describe the standard pipeline of a deep FER system with the related background knowledge and suggestions of applicable implementations for each stage. For the state of the art in deep FER, we review existing novel deep neural networks and related training strategies that are designed for FER based on both static images and dynamic image sequences, and discuss their advantages and limitations. Competitive performances on widely used benchmarks are also summarized in this section. We then extend our survey to additional related issues and application scenarios. Finally, we review the remaining challenges and corresponding opportunities in this field as well as future directions for the design of robust deep FER systems.

Facial Expression Recognition (FER) can be widely applied to various research areas, such as mental diseases diagnosis and human social/physiological interaction detection. With the emerging advanced technologies in hardware and sensors, FER systems have been developed to support real-world application scenes, instead of laboratory environments. Although the laboratory-controlled FER systems achieve very high accuracy, around 97%, the technical transferring from the laboratory to real-world applications faces a great barrier of very low accuracy, approximately 50%. In this survey, we comprehensively discuss three significant challenges in the unconstrained real-world environments, such as illumination variation, head pose, and subject-dependence, which may not be resolved by only analysing images/videos in the FER system. We focus on those sensors that may provide extra information and help the FER systems to detect emotion in both static images and video sequences. We introduce three categories of sensors that may help improve the accuracy and reliability of an expression recognition system by tackling the challenges mentioned above in pure image/video processing. The first group is detailed-face sensors, which detect a small dynamic change of a face component, such as eye-trackers, which may help differentiate the background noise and the feature of faces. The second is non-visual sensors, such as audio, depth, and EEG sensors, which provide extra information in addition to visual dimension and improve the recognition reliability for example in illumination variation and position shift situation. The last is target-focused sensors, such as infrared thermal sensors, which can facilitate the FER systems to filter useless visual contents and may help resist illumination variation. Also, we discuss the methods of fusing different inputs obtained from multimodal sensors in an emotion system. We comparatively review the most prominent multimodal emotional expression recognition approaches and point out their advantages and limitations. We briefly introduce the benchmark data sets related to FER systems for each category of sensors and extend our survey to the open challenges and issues. Meanwhile, we design a framework of an expression recognition system, which uses multimodal sensor data (provided by the three categories of sensors) to provide complete information about emotions to assist the pure face image/video analysis. We theoretically analyse the feasibility and achievability of our new expression recognition system, especially for the use in the wild environment, and point out the future directions to design an efficient, emotional expression recognition system.


Facial expression recognition is one of the very important research topics in computer vision. Studies on nonverbal communication have shown that 55% of intentional information is conveyed through facial expressions. Expression recognition has recently found a lot many applications in medical and advertising industries. In this paper we have proposed a parallel Convolutional Neural Network (CNN) structure for detection of expression from frontal faces. The CNNs are trained on two most important subfacial patches. The overall feature vector will be the features concatenated from the parallel models. We have experimentally found applying such a strategy provides better results than the models which take the entire facial image. We have also compared our performance with other benchmark CNN structures like AlexNet and VGG16.


Facial expression recognition has been an active research area over the past few decades, and it is still challenging due to the high intra-class variation. Traditional approaches for this problem rely on hand-crafted features such as SIFT, HOG and LBP, followed by a classifier trained on a database of images or videos. Most of these works perform reasonably well on datasets of images captured in a controlled condition, but fail to perform as good on more
challenging datasets with more image variation and partial faces. In recent years, several works proposed an end-to-end framework for facial expression recognition, using deep learning models. Despite the better performance of these works, there still seems to be a great room for improvement. In this work, we propose a deep learning approach based on attentional convolutional network, which is able to focus on important parts of the face, and achieves significant improvement over previous models on multiple datasets, including FER-2013, CK+, FERG, and JAFFE. We also use a visualization technique which is able to find important face regions for detecting different emotions, based on the classifier’s output. Through experimental results, we show that different emotions seems to be sensitive to different parts of the face.


In recent years, with the development of artificial intelligence and human–computer interaction, more attention has been paid to the recognition and analysis of facial expressions. Despite much great success, there are a lot of unsatisfying problems, because facial expressions are subtle and complex. Hence, facial expression recognition is still a challenging problem. In most papers, the entire face image is often chosen as the input information. In our daily life, people can perceive other’s current emotions only by several facial components (such as eye, mouth and nose), and other areas of the face (such as hair, skin tone, ears, etc.) play a smaller role in determining one’s emotion. If the entire face image is used as the only input information, the system will produce some unnecessary information and miss some important information in the process of feature extraction. To solve the above problem, this paper proposes a method that combines multiple sub-regions and the entire face image by weighting, which can capture more important feature information that is conducive to improving the recognition accuracy. Our proposed method was evaluated based on four well-known publicly available facial expression databases: JAFFE, CK+, FER2013 and SFEW. The new method showed better performance than most state-of-the-art methods.

4. ARCHITECTURE
SYSTEM TESTING

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, sub-assemblies, assemblies and/or a finished product. It is the process of exercising software with the intent of ensuring that the software system meets its requirements and user expectations and does not fail in an unacceptable manner. There are various types of test. Each test type addresses a specific testing requirement.

TYPES OF TESTS

Unit testing

Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. It is the testing of individual software units of the application. It is done after the completion of an individual unit before integration. This is a structural testing, that relies on knowledge of its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly defined inputs and expected results.

Integration testing

Integration tests are designed to test integrated software components to determine if they actually run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields. Integration tests demonstrate that although the components were individually satisfaction, as shown by successfully unit testing, the combination of components is correct and consistent. Integration testing is specifically aimed at exposing the problems that arise from the combination of components.

Functional test

Functional tests provide systematic demonstrations that functions tested are available as specified by the business and technical requirements, system documentation, and user manuals.

Functional testing is centered on the following items:

Valid Input : identified classes of valid input must be accepted.
Invalid Input : identified classes of invalid input must be rejected.
Functions : identified functions must be exercised.
Output : identified classes of application outputs must be exercised.
Systems/Procedures: interfacing systems or procedures must be invoked.

Organization and preparation of functional tests is focused on requirements, key functions, or special test cases. In addition, systematic coverage pertaining to identify Business process flows; data fields, predefined processes, and successive processes must be considered for testing. Before functional testing is complete, additional tests are identified and the effective value of current tests is determined.

**SYSTEM TEST**

System testing ensures that the entire integrated software system meets requirements. It tests a configuration to ensure known and predictable results. An example of system testing is the configuration-oriented system integration test. System testing is based on process descriptions and flows, emphasizing pre-driven process links and integration points.

**White Box Testing**

White Box Testing is a testing in which the software tester has knowledge of the inner workings, structure and language of the software, or at least its purpose. It is used to test areas that cannot be reached from a black box level.

**Black Box Testing**

Black Box Testing is testing the software without any knowledge of the inner workings, structure or language of the module being tested. Black box tests, as most other kinds of tests, must be written from a definitive source document, such as specification or requirements document. It is a testing in which the software under test is treated, as a black box. You cannot “see” into it. The test provides inputs and responds to outputs without considering how the software works.

6.1 Unit Testing:

Unit testing is usually conducted as part of a combined code and unit test phase of the software lifecycle, although it is not uncommon for coding and unit testing to be conducted as two distinct phases.

**Test strategy and approach**

Field testing will be performed manually and functional tests will be written in detail.

**Test objectives**

- All field entries must work properly.

- Pages must be activated from the identified link.

- The entry screen, messages and responses must not be delayed.
Features to be tested

- Verify that the entries are of the correct format
- No duplicate entries should be allowed
- All links should take the user to the correct page.

6.2 INTEGRATION TESTING

Software integration testing is the incremental integration testing of two or more integrated software components on a single platform to produce failures caused by interface defects.

The task of the integration test is to check that components or software applications, e.g. components in a software system or - one step up - software applications at the company level - interact without error.

Test Results: All the test cases mentioned above passed successfully. No defects encountered.

6.3 Acceptance Testing

User Acceptance Testing is a critical phase of any project and requires significant participation by the end user. It also ensures that the system meets the functional requirements.

Test Results: All the test cases mentioned above passed successfully. No defects encountered.

5. MODULES

System

User

- 1.1 Capture Images:
  Capture the images using open CV platform

- 1.2 Train Images:
  We will using Convolutional neural network algorithm to train the model.

- 1.3 Notification:
  When the training is completed, a notification will be displayed.

- 1.4 Person Identification:
  Whenever our model detects a known person, it recognizes him/her.

2.1 Enter Your Name:

The user needs to enter his/her first name.
• 2.3 Capture Images:

The user requests the application to capture images of his/her face without disguise.

• 2.4 View Results:

• If known person is detected system will show you the name of that person or else it will displayed as Unknown.

PROPOSED SYSTEM:

We propose a method in which a device is built by simply providing facial recognition of the person and the device that can be considered a useful system since it helps to reduce the limitations obtained from existing system. By providing support to these types of devices, we will prevent ourself from the unknown persons.

FLOW OF THE PROJECT:

ADVANTAGES:

• High efficiency.
• Time Saving.
• Inexpensive.
• Low complexities.
6. CONCLUSION

In this project we have proposed a UGI based application which is used to recognize the persons by using DL techniques. We used an algorithm named Convolutional neural network which termed to be easily identify the response to capture, Images we have used a webcam by using a computer vision technique. In this project we have successfully developed an application, that which can detect and recognize the faces.

7. FUTURE SCOPE

Features are typically distinct points in an image and the descriptor gives a signature, so it describes the key point that is considered. It extracts the local neighborhood around that point so a local image patch is created and a signature from this local patch is computed.

8. REFERENCES


