Data Mining

Seminar 03

Face Recognition: A Convolutional Neural-Network Approach

Department of Computer Science
Kristianstad University
Course: DT584C
Master's in computer science

Student: Hazrat Ali <react.dev.se@gmail.com>
Teacher: Dawit Mengistu

Date: 16 January 2020

Table of contents

Introduction

This is a critique of the article “Face Recognition: A Convolutional Neural-Network Approach” authored by Steve Lawrence et al and published in “ IEEE Transactions on neural networks, vol 8 - January 1997”. This article is concerned about face recognition using convolutional network. Convolution neural network is a fully connected neural network. The authors also present comparison of different face recognition approaches.

The authors are IEEE members and they belong to the University of Queensland, St. Lucia, Australia, however C. L. Giles works at Research Institute, Princeton, and the University of Maryland, USA.

This research was supported partly by Australian Research Council (ACT) and the Australian Telecommunications and Electronics Research Board (SL).

The research achieves a considerable research results, however there are several weaknesses in this article which have been mentioned in sections as follows.

Body

Using biometrics devices in personal identification has been increased because it provides very reliable and non intrusive approach. The authors of the article mentions the most famous biometric features include fingerprints, speech, signature dynamics, and face recognition. There are two categories of face recognition systems mentioned in the article.

The first category involved non real time identification of face in a big database. While the second is based on real time identification of a face in a security system, or for access control. The article is about real time face recognition systems.

The dataset used in this research article is taken from ORL database [1] which contains 400 pictures of 40 different people.

Related work

In this section we summarize we summarize the related work mentioned in the article. There have been many methods used for face recognition. We are summarizing them in the sections below.

A. Geometrical features

Many people have explored using geometrical features for face recognition. It uses ratios of distances between a set of geometric features such as nose length and mouth position and chin shape and so on. So systems may extract these features directly and the accuracy of the system is dependant on the precisely measured distances between features.

B. Eigenfaces

An eigenface is a set of eigenvectors which is used in human face recognition. This approach uses a low dimensional representation of face images and can be generated using a statistical process called principal component analysis (PCA).

C. Template matching

It is a technique of for small parts in an image which is then matched to a template image. Feature based template matching means the extraction of image features such as shape, textures, colors, height, length to be matched in the target image [2].

It performs direct correlation of image segments and is effective when the query images have the same scale.

D. Graph matching

Graphs are used to structure information and is used in the field of computer science like pattern recognition and computer vision[3].

E. Neural network approaches

Artificial neural network is a set of algorithms used to recognize patterns. It is modeled after the functioning of the human neural network. It helps to cluster or classify a dataset based on labels. The numerical data are fed in the form of vectors which the machine perceives and cluster it.

The article mentions the experimental results of research[4] in which there 50 principal components are extracted and reduced to five dimensions. A standard multilayer perceptron (MLP) classifier was used on a simple database which contains the images of 20 people. The article does not mention any numeric or concrete results though.

F. Hidden Markov Model (HMM)

“In simpler Markov models, the state is directly visible to the observer, and therefore the state transition probabilities are the only parameters, while in the hidden Markov model, the state is not directly visible, but the output dependent on the state, is visible. Each

state has a probability distribution over the output tokens. [5]" -- (excerpted from wikipedia and therefore in double quotes)

The author mentions the results of a research article which uses ORL database which gives 13% error rate.

System components

A. Overview

In this section the author has described the components of the proposed system. The article uses a local image sampling technique and a technique for partial lightening invariance which is a self organizing map (SOM) and transforms the sample image into a quantized lower dimensional space.

B. Local image sampling

The article makes use of two techniques for local image sampling. The first one creates a vector from the image using intensity values at each point in the window. While the second creates a representation of the local sample and then form a vector by both the intensity of the center pixel and its difference with all other pixels.

C. The self organizing map

The self organizing map (SOM) is a type of artificial neural network that is used for dimension reduction

D. KL Transform

This method is used to remove redundancy in a dataset via principal component analysis (PCA). PCA generates a set of orthogonal axes of the input data in the order of decreasing variance.

E. Convolutional Networks

2-D images can fit into many models during the training but do not generalize well to unseen projects. On the other hand convolutional network achieve some degree of shift using three ideas: local receptive fields, shared weights, spatial subsampling.

A typical convolutional network shown in Figure 1. is composed of several layers each of which contains several planes. The images are firstly normalized and centered and then entered at the input layer. Each unit in a plane receives input from the neighborhood of the previous layer. A feature is considered a plane with a feature detector that is convolved with the local window obtained by scanning a window in the previous layer. These layers are termed a convolutional layers which detects a feature. The network is trained with usual backpropagation gradient descent procedure[13].

System details

The system designed by this research article is shown in the high level diagram (Figure 1).

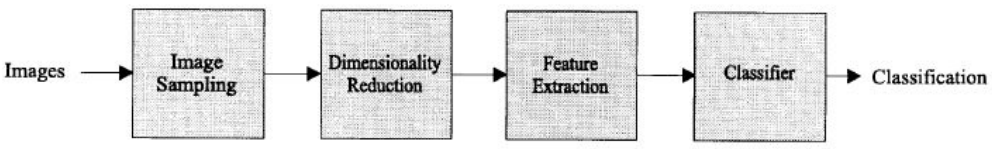

Figure 1 - A high level block diagram of the proposed system

The working of this system is given below.

1. The training set has images extracted from local samples at each step. The window of size 5x5 is moved by four pixels at each step.
2. A self organizing matrix of order 3 dimensions and 5 nodes per dimensions gives 5^3 = 125 nodes, which is trained on the vectors from the previous stage. The three dimensions represent three features.
3. The local images are passed through SOM to create new training and test sets in each step. Each image is represented by three maps which corresponds to a dimension in the SOM. If the image has dimensions 92 x 112 and step size is 4, the maps are 23 x 28.
4. The newly created training set is used for training the convolution neural network.

Experimental Results

The experiments conducted in the article consists of 5 training and 5 test images. The experimental environment consists of 40 classes, 3 dimensions in the SOM, 5 number of nodes per SOM dimension, image extraction with original intensity values. The results of the experiments performed are given below.

1. **Variation in the number of classes in the output:** It was observed in the article that the error rate increases as the number of classes are increased from 10 to 20 to 40. This is depicted in Figure 2 and Table 1.

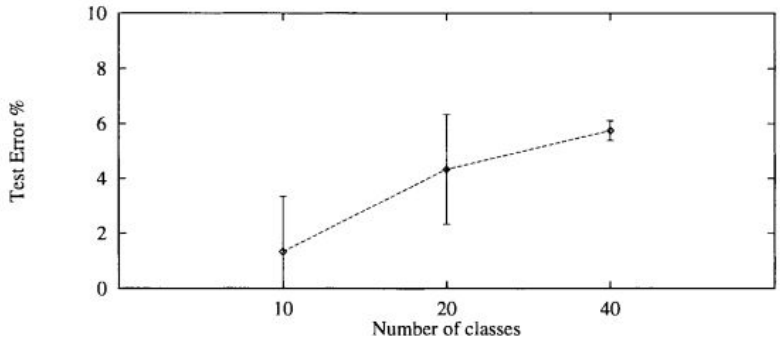

Figure 2 - Error rate is a function of number of classes

Number of classes	10	20	40
Error rate	1.33%	4.33%	5.75%

Table 1 - Error rate is the average of classes in 3 simulations in the face recognition system

2. **Variation in the dimensionality of the SOM :** As the number of dimensions is increased in this experiment the error rate decreases. The system gives best performance at dimension 3 as shown in Figure 3 and Table 2.

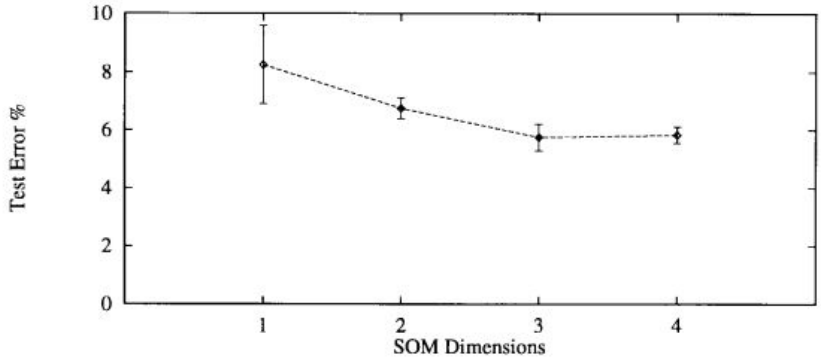

Figure 3 - Error rate is a function of number of dimensions

SOM Dimension	1	2	3	4
Error rate	8.25%	6.75%	5.75%	5.83%

Table 2 - Error rate is the average of dimensions in 3 simulations in the face recognition system

3. **Variation in the quantization level of the SOM :** As the number of nodes per dimension increases the error rate decreases until 8 or 9 nodes which shows the best error rate.

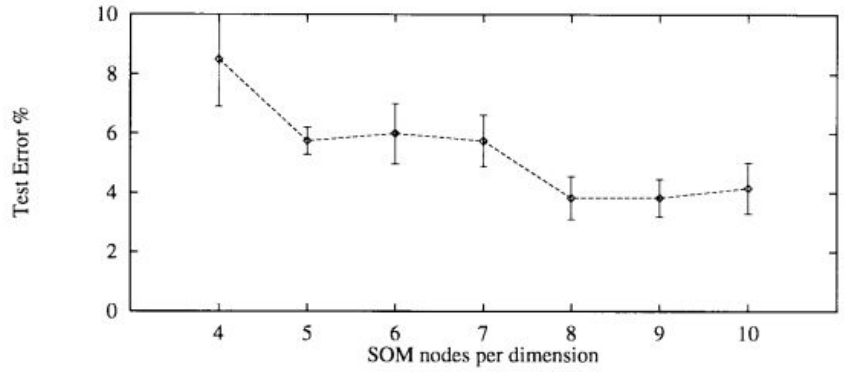

Figure 4 - Error rate is a function of number of nodes per dimension in SOM

SOM Size	4	5	6	7	8	9	10
Error rate	8.5%	5.75%	6.0%	5.75%	3.83%	3.83%	4.16%

Table 3 - Error rate is the average of nodes per dimension in 3 simulations in the face recognition system

4. **Variation in the image sample extraction algorithm:** The article states that using original intensity values give the best performance, Table 4.

Input type	Pixel intensities	Differences w/base intensity
Error rate	5.75%	7.17%

Table 4 - Error rate of the face recognition system with varying image sample representation

5. **Substituting the SOM with the KL transform :** Table 5 below shows the results or replacing the SOM with KL transform. The article observed that only one eigenvector is used instead of 2 or 3, the performance is the best. It states further that quantization of SOM could provide a degree of invariance to minor differences in the image samples, and quantization of PCA projections may improve performance.

Dimensionality reduction	Linear PCA	SOM
Error rate	5.33%	3.83%

Table 5 - Error rate of the face recognition system with Linear PCA vs SOM feature extractions

6. **Replacing the CN with MLP :** In this experiment the author has replaced convolutional network with an MLP and observed performance has affected badly. This is due to the lack of inbuilt invariance in MLP to minor translation and local deformation.

	Linear PCA	SOM
MLP	41.2%	39.6%
CN	5.33%	3.83%

Table 6 - Error rate of the face recognition system with various feature extraction and classification methods. Average of 3 simulations

7. **The tradeoff between rejection threshold and recognition accuracy :** The author gives a histogram when the classifier is correct and when it is wrong for best use case. The figure shows that the classification performance increase if case rejected which are below a certain confidence threshold.

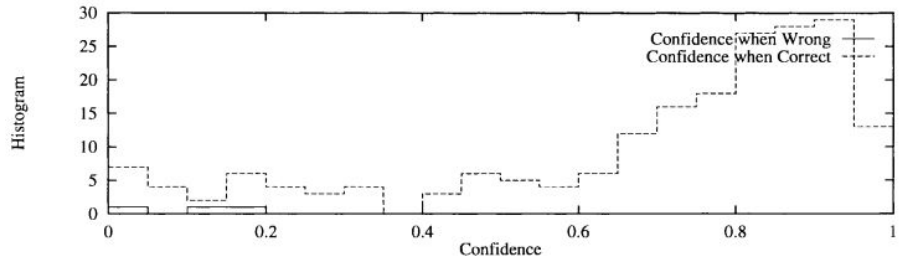

Figure 5 - Histogram of confidence of classifier when it is correct and confidence when it is wrong. Classification is improved when case with low confidence are rejected

8. **Comparison with other known results on the same database :** Here the author gives a summary of the performance of the systems using ORL database. A SOM quantization level of 8 was used a best performance was observed because it performs recognition 500 times faster than others.

System	Error rate	Classification time
Top-down HMM	13%	n/a
Eigenfaces	10.5%	n/a
Pseudo 2D-HMM	5%	240 seconds[1]
SOM+CN	3.8%	< 0.5 seconds[2]

Table 7 - Error rate and performance of various systems

9. **Variation of the number of training images per person :** Here the author changes the number of images per class used in the training set from one to five and for eigenfaces algorithm. It was observed using separate training vectors resulted in better performance. Using 40 to 100 eigenfaces resulted in same performance. It was observed that PCA + CN and SOM+CN are both superior to the eigenfaces technique even if there is only one training image per person. It was conclude finally that SOM+CN method outperforms than PCA+CN method.

Strengths

- An extensive set of experiments and observations have been performed in this article.
- All the results are explained with some explanation and figures.
- A detailed discussion and results analysis section is given.

Weaknesses

- Figures poorly presented and not eligible
- The language used is not simple and easily understandable e.g “There are ten different images of 40 distinct subjects” could be rephrased as “There are ten differenct images of each of the 40 test subjects”. Another “... there are 20 people in the database.”
- Some ambiguous sentences like “...using a query database of 95 images from a total of 685 individuals” could be rephrased as “...95 images were randomly selected from images of 685 individuals”
- Non scientific language has been used in a few places e.g “... They present good results with a database of 87 people …”. The word good means different to different researchers but research require more precise figures for example “..They present results with 87% accuracy”. Another one is “They report good results for discrimination of ten distinctive subjects”
- Some sentences are mentioned without (or before) reference e.g “Samaria also performed extensive tests…”
- Some ambiguities in describing the results e.g “... We note that a system which guesses the correct answer would be right one out of 40 times, giving an error rate of 97.5%”
- No concrete or numerical results are being reported in some places like page 4 - neural network.

Conclusion

In this article we reviewed the article “Face Recognition: A convolutional Neural Network Approach” and presented a summary of the outcomes. We also mentioned some strengths and weaknesses in the article.

We observed that although a very extensive article with many experimentations have been presented with worth mentioning results but there are weaknesses especially in the structure and scientific writing aspect of the article.

References

[1] “The database of faces.” [Online]. Available: http://cam-orl.co.uk/facedatabase.html. [Accessed: 11-Jan-2020].

[2] “Template matching.” [Online]. Available: https://en.wikipedia.org/wiki/Template_matching. [Accessed: 15-Jan-2020].

[3] “Graph matching.” [Online]. Available: https://docs.google.com/document/d/1TguEspzIq-hb4wMYSYDDUOHRAC5V6aog4TcGlqeexP0/edit. [Accessed: 16-Jan-2020].

[4] D. DeMers and G. W. Cottrell, “Non-Linear Dimensionality Reduction,” in *Advances in Neural Information Processing Systems 5*, S. J. Hanson, J. D. Cowan, and C. L. Giles, Eds. Morgan-Kaufmann, 1993, pp. 580–587.

[5] “Hidden Markov Model.” [Online]. Available: https://en.wikipedia.org/wiki/Hidden_Markov_model. [Accessed: 12-Jan-2020].