

BBM 415 - Fundamentals of Image Processing

Laboratory Assignment 4

Murat Çelik, 21827263
Department of Computer Engineering
Hacettepe University
Ankara, Turkey
`b21827263@cs.hacettepe.edu.tr`

January 7, 2023

1 Introduction

Before the existence of Deep Learning, different techniques were used for image classification. One of them is to search a part of a picture in the other picture with the sliding windows algorithm with different techniques. The general name of this problem is called template matching.

1.1 Experiment Details

One of these techniques is the normalized cross correlation method. Today I will solve the classification problem with NCC using American Sign Language data. Below I have put into the NCC algorithm for all images with each template image. This algorithm finds and scores the best match between two images. After finding the best window, I assigned a score for each template. Then I chose the best score to determine which template was closest for that image. Thus, the best result of our picture was from the class of that template.

The experiment was carried out with templates in 26 different classes on 780 pictures. Below we can see the code of these stages, the overall score and results.

1.2 Experiment Results

The success rate is seen as 10%. It can be said that the results are quite behind compared to today's algorithms. It can be said that this algorithm gives better results than expected for the old periods.

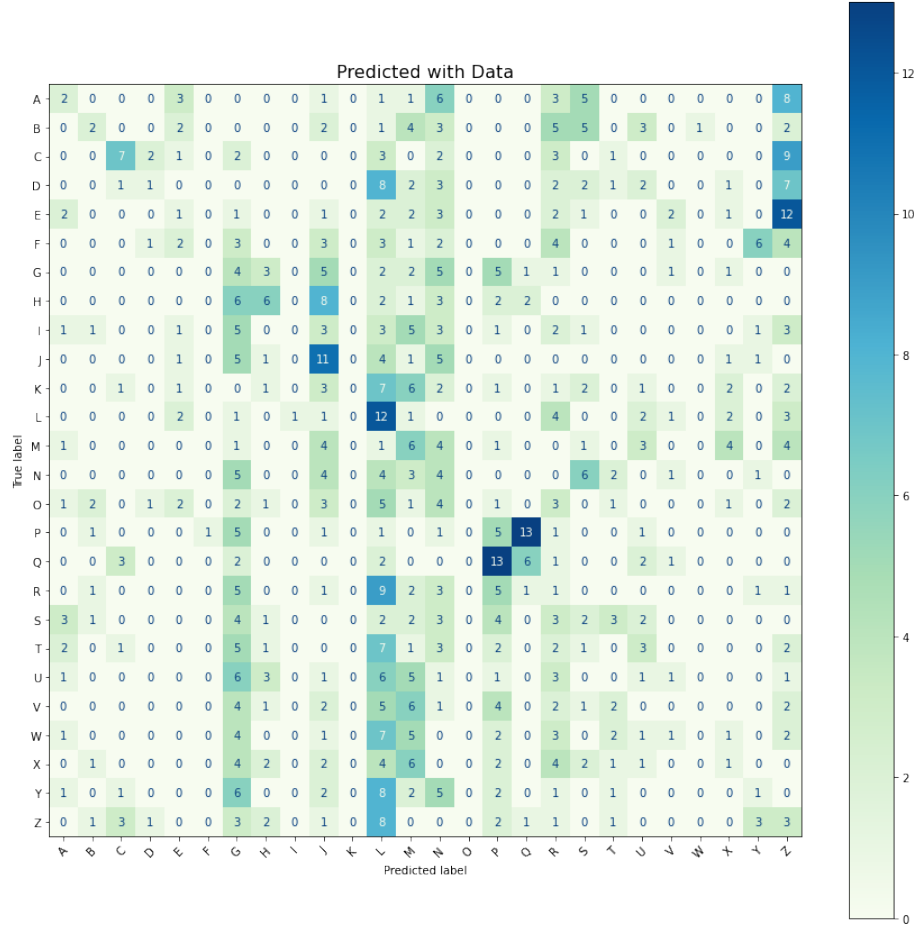


Figure 1: Confusion matrix with 781 results

For Figure 1, it is observed that the classes G, L, P, Q, Z are the most mentioned classes in estimation. The main reason for this is likely to be due to the proximity of the letters to each other.

It is observed that there are 13 matches each due to the similarities between the P and Q data. But when you look at it, the prediction of Q as Q appears to be 6 predicts. This can be explained by the fact that the signs are similar to each other.

Again, it seems that the letter E is less predicts. The main reason for this is that the fingers are not used in the letter E representation, which shows that it is difficult to distinguish.

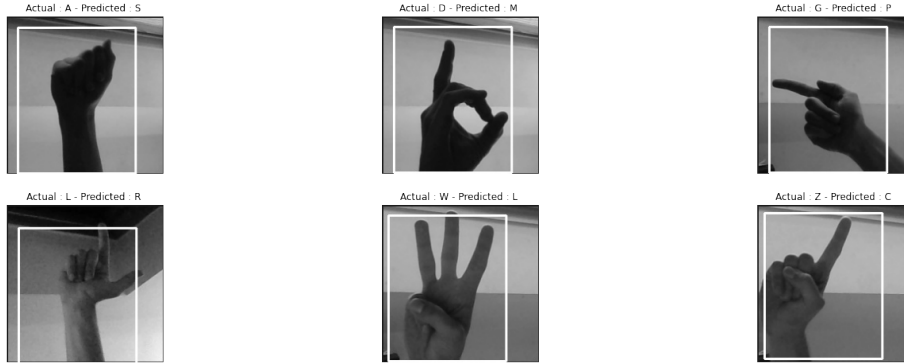


Figure 2: Wrongly predicted images

We see the wrong results above Figure 2. There can be many reasons for getting an error. One of them may be the contrast difference between the main image and the template. Different options may have been estimated due to the shape of the sign. There may have been distortions in the estimation due to the lack of sharp distinctions between the background and the hand. It is normal to draw wrong conclusions for such reasons. It is a difficult problem for the computer to find the difference between the pictures, which are made up of 0s and 1s. Being able to solve this problem with a certain success with image matching is also fruitful.

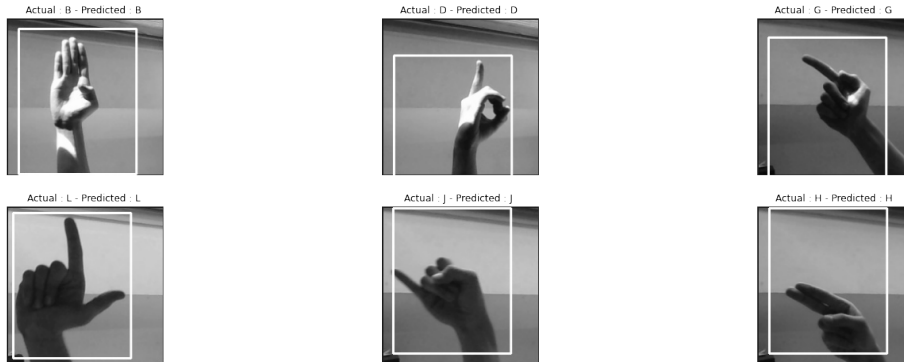


Figure 3: Correctly predicted images

Above we see the correct results on Figure 3. There are certain reasons why the results are correct. When we first examine the pictures above, it can be said that good results are obtained with the contrast difference between the hand and the background. Another reason is that letters have a unique structure. For example, the letters H and L have a different structure from other letters. Results within 10% success can be observed in this way.

2 Histogram of Oriented Gradients (HOG)

The histogram of oriented gradients (HOG) is a feature descriptor used in computer vision and image processing for the purpose of object detection. The technique counts occurrences of gradient orientation in localized portions of an image.



Figure 4: Binary-Original Images and HOG Computation

HOG calculation is performed on all images. Another classification is made with the results. Above is a test to see if the original image should be converted to binary image before HOG calculation. It was decided to make it with the original image.

2.1 HOG Results

We can see that the success has decreased from 10% to 7%. We can say that the processing speed has decreased, but at this point, success is more important.

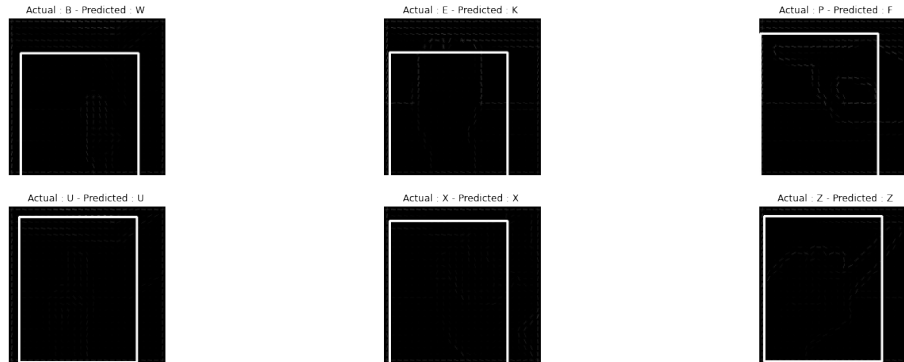


Figure 5: Result of HOG Computation

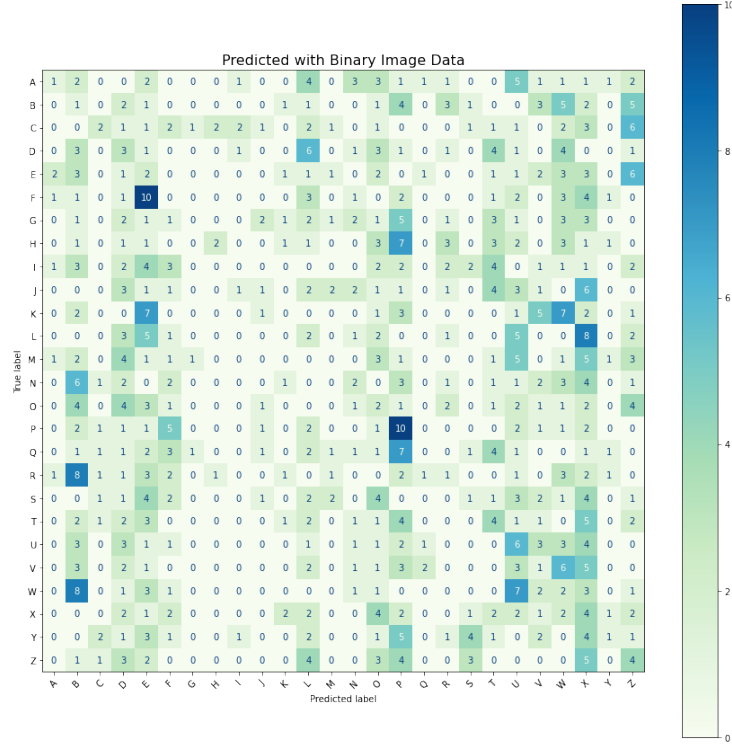


Figure 6: Confusion matrix with 781 results on HOG Computation

For Figure 6, it seems that the estimates are directed towards 5 letters. "D,E,P,W,X". The reason for the general success here is that the contrast difference in the picture is proportional to the effect on the gradients.

3 Conclusion

We have come to a position where we can easily solve many different problems with artificial intelligence technology. By spending less time than the problem we have tested above, we could go to the solution with algorithms with near-perfect results. The value of old work comes from the fact that it is the basis of new work. Thus, with a solid basic knowledge, we have the chance to think more broadly for the solution of these technologies. Thus, we experienced both the importance of image processing and its use.

4 references

Template Matching
Histogram of oriented gradients