Signature Fraudulent Detection

Kaarthik Shrinivas V,19BCE1461,Vinod B,19BCE1574,Sai Vignesh,19BCE1389

Vellore Institute of Technology, Chennai.

Abstract:— **Now everything has got digitized. There are many Fraudulent activities happening in several sectors due to fake signatures by criminals for getting unauthorized access to the user whose money/asset they want to access. There is no proper System for Verifying Signatures of the user. So our Project basically proposes a model to solve this problem. In our Project, we get an input signature image from the user/customer and check it with the original signature of the user we need to verify and check whether the signature in the input image is fraudulent or not. We use combination of CNN, SIFT and ORB algorithm along with image processing to solve this problem**

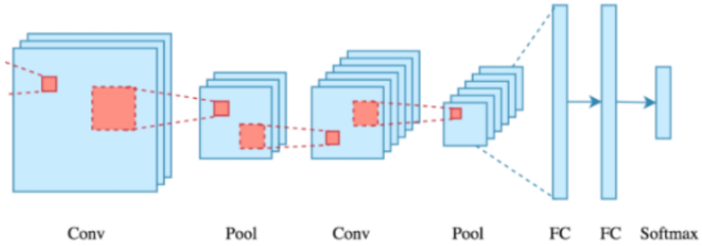
Keywords—**Fraudulent, CNN, SIFT, ORB, Image processing**

1. Introduction

Due to Increased Criminal and Fraudulent Activities that take place in bank, educational and several other sectors due to fraudulent/fake signatures produced by criminals. There is no proper system for verifying these fraudulent signatures. So we thought of proposing a project which could provide a solution for this problem.

In the past few decades, Deep Learning has proved to be a very powerful tool because of its ability to handle large amounts of data. The interest to use hidden layers has surpassed traditional techniques, especially in pattern recognition. One of the most popular deep neural networks is Convolutional Neural Networks.

CNN's were first developed and used around the 1980s. The most that a CNN could do at that time was recognize handwritten digits. It was mostly used in the postal sectors to read zip codes, pin codes, etc. The important thing to remember about any deep learning model is that it requires a large amount of data to train and also requires a lot of computing resources. This was a major drawback for CNNs at that period and hence CNNs were only limited to the postal sectors and it failed to enter the world of machine learning.



In deep learning, a convolutional neural network (CNN/ConvNet) is a class of deep neural networks, most commonly applied to analyze visual imagery. Now when we think of a neural network we think about matrix multiplications but that is not the case with ConvNet. It uses a special technique called Convolution. Now in mathematics convolution is a mathematical operation on two functions that produces a third function that expresses how the shape of one is modified by the other.

Despite the power and resource complexity of CNNs, they provide in-depth results. At the root of it all, it is just recognizing patterns and details that are so minute and inconspicuous that it goes unnoticed to the human eye. But when it comes to understanding the contents of an image it fails. Thus along with CNN we use SIFT and ORB algorithms to process the recognized signature to detect the fraudulent.

The scale-invariant feature transform (SIFT) is a computer vision algorithm to detect, describe, and match local features in images, invented by David Lowe in 1999. Applications include object recognition, robotic mapping and navigation, image stitching, 3D modeling, gesture recognition, video tracking, individual identification of wildlife and match moving.

Major advantages of SIFT are

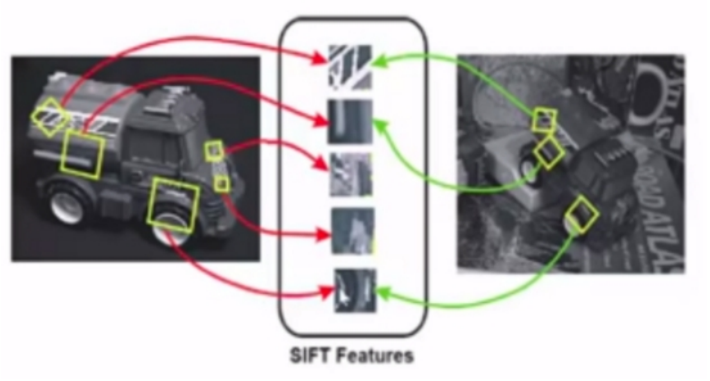
Locality: features are local, so robust to occlusion and clutter (no prior segmentation)

Distinctiveness: individual features can be matched to a large database of objects

Quantity: many features can be generated for even small objects

Efficiency: close to real-time performance

Extensibility: can easily be extended to a wide range of different feature types, with each adding robustness



ORB is a fusion of FAST keypoint detector and BRIEF descriptor with some added features to improve the performance. FAST is Features from Accelerated Segment Test used to detect features from the provided image. It also uses a pyramid to produce multiscale-features. Now it doesn't compute the orientation and descriptors for the features, so this is where BRIEF comes in the role.

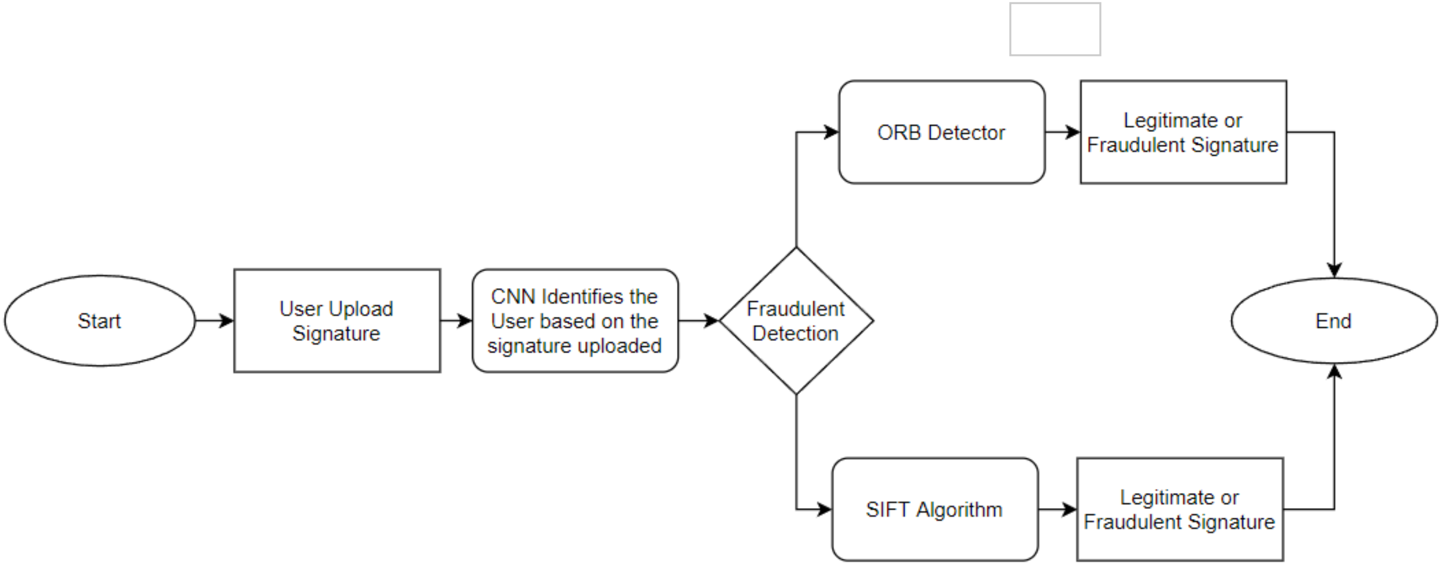
ORB uses BRIEF descriptors but as the BRIEF performs poorly with rotation. So what ORB does is to rotate the BRIEF according to the orientation of keypoints. Using the orientation of the patch, its rotation matrix is found and rotates the BRIEF to get the rotated version. ORB is an efficient alternative to SIFT or SURF algorithms used for feature extraction, in computation cost, matching performance, and mainly the patents. SIFT and SURF are patented and you are supposed to pay them for its use. But ORB is not patented.

2. OBJECTIVES OF THE WORK

Objectives of our project is:

- To maximize the accuracy in detecting a user's signature
- To accurately process the signature for fraudulent detection
- To predict whether the signature is fraudulent or not by comparing with the original signature

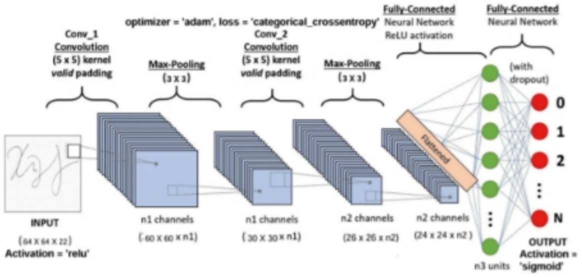
3. PROPOSED WORK


Start
User Upload Signature
CNN Identifies the User based on the signature uploaded
Fraudulent Detection
ORB Detector
Legitimate or Fraudulent Signature
SIFT Algorithm
Legitimate or Fraudulent Signature
End

In our Project we collected the sample images from the dataset which we got from kaggle website.After this we splitted the dataset as training and test part and used the training part to train the CNN model to make the model recognize the user of the input signature.After this we introduced two algorithms SIFT and ORB for checking whether the input signature is fraudulent or not.

Convolution Neural Network(CNN)

Convolutional Neural Networks (CNNs) have tested no-hit in recent years at an outsized variety of image processing-based machine learning tasks. Several different strategies of playacting such tasks as shown in Fig below revolve around a method of feature extraction, during which hand-chosen options are extracted from a picture fed into a classifier to make a classification call. Such processes are solely as sturdy because of the chosen options, which regularly take giant amounts of care and energy to construct. Against this, in CNN, the options fed into the ultimate linear classifier all learned from the dataset. A CNN consists of a variety of layers as shown in Fig. Below, beginning at the raw image pixels, each performs an easy computation and feeds the result to the successive layer, with the ultimate result being fed to a linear classifier. The layers computation area unit supports a variety of parameters that are learned through the method of backpropagation, during which for every parameter, the gradient of the classification loss with relation to that parameter is computed and therefore the parameter is updated to minimize the loss performed. The look of any signature verification system typically needs the answer of 5 sub-issues: data retrieval, pre-processing, feature extraction, identification method, and performance analysis. Off-line signature verification just deals with pictures non-heritable by a scanner or a photographic camera. In an associate degree off-line signature verification system, a signature is non-heritable as a picture. This picture depicts a private sort of humans. The method needs neither be too sensitive nor too rough.



It should have a proper balance between an occasional False Acceptance Rate (FAR) and an occasional False Rejection Rate (FRR).

Signature Forgery Detection :

When the image is finally classified into one among the present classes of the users. We proposed a technique for the detection of the fraudulent signature using two Algorithms which compares input images with the original signatures image.

Algorithms Used :

- SIFT Algorithm
- ORB Algorithm

SIFT Algorithm :

The **scale-invariant feature transform** (**SIFT**) is a computer vision algorithm to detect, describe, and match local *features* in images, invented by David Lowe in 1999. Applications include object recognition, robotic mapping and navigation, image stitching, 3D modeling, gesture recognition, video tracking, individual identification of wildlife and match moving

SIFT keypoints of objects are first extracted from a set of reference images [1] and stored in a database. An object is recognized in a new image by individually comparing each feature from the new image to this database and finding candidate matching features based on Euclidean distance of their feature vectors. From the full set of matches, subsets of keypoints that agree on the object and its location, scale, and orientation in the new image are identified to filter out good matches. The determination of consistent clusters is performed rapidly by using an efficient hash table implementation of the generalised Hough transform. Each cluster of 3 or more features that agree on an object and its pose is then subject to further detailed model verification and subsequently outliers are discarded. Finally the probability that a particular set of features indicates the presence of an object is computed, given the accuracy of fit and number of probable false matches. Object matches that pass all these tests can be identified as correct with high confidence.

(Keypoint Detection using SIFT Algorithm)

ORB Algorithm :

Oriented Fast and Rotated Brief(ORB) Algorithm

ORB is a fusion of FAST keypoint detector and BRIEF descriptor with some added features to improve the performance. FAST is Features from Accelerated Segment Test used to detect features from the provided image. It also uses a pyramid to produce multiscale-features. Now it doesn't compute the orientation and descriptors for the features, so this is where BRIEF comes in the role.

ORB uses BRIEF descriptors but as the BRIEF performs poorly with rotation. So what ORB does is to rotate the BRIEF according to the orientation of keypoints. Using the orientation of the patch, its rotation matrix is found and rotates the BRIEF to get the rotated version. ORB is an efficient alternative to SIFT or SURF algorithms used for feature extraction, in computation cost, matching performance, and mainly the patents. SIFT and SURF are patented and you are supposed to pay them for its use. But ORB is not patented.

The ORB image matching algorithm is generally divided into three steps:

- Feature point extraction
- Generating feature point descriptors
- Feature point matching.

(Keypoint Detection in ORB Algorithm)

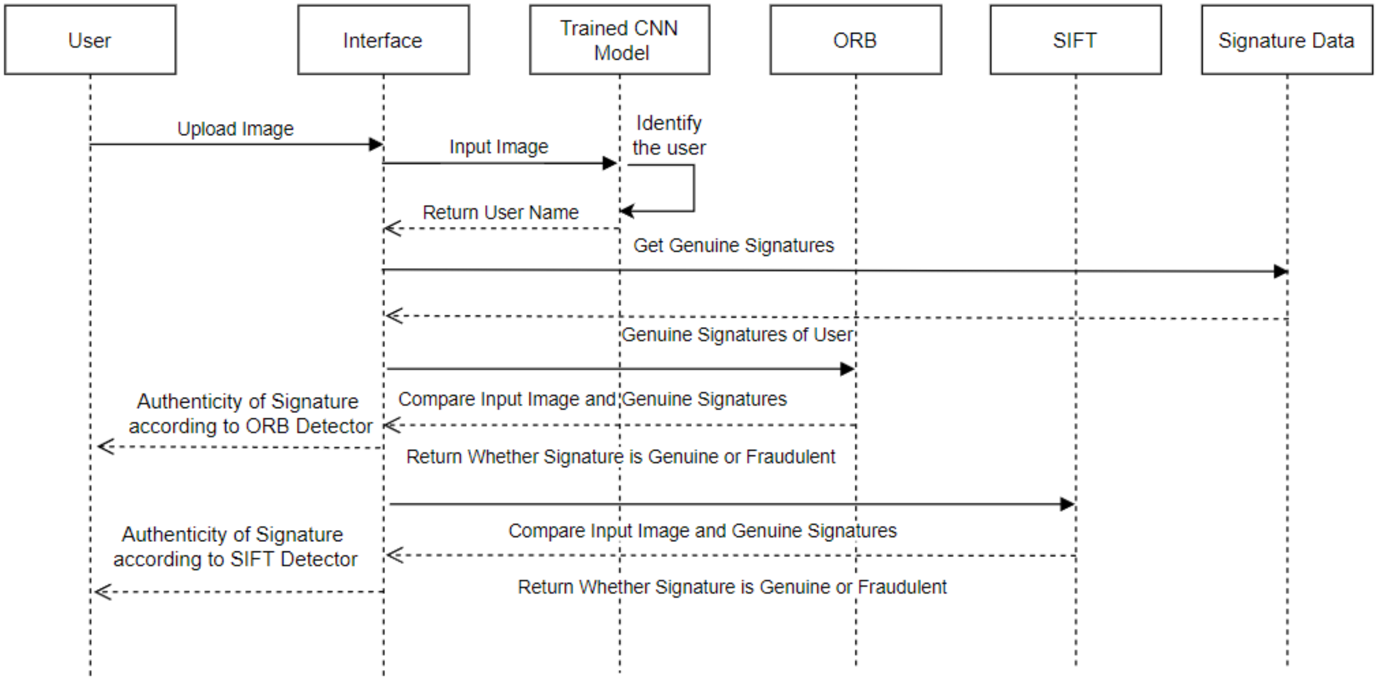

User
Interface
Trained CNN Model
ORB
SIFT
Signature Data
Upload Image
Input Image
Identify the user
Return User Name
Get Genuine Signatures
Genuine Signatures of User
Compare Input Image and Genuine Signatures
Authenticity of Signature according to ORB Detector
Return Whether Signature is Genuine or Fraudulent
Compare Input Image and Genuine Signatures
Authenticity of Signature according to SIFT Detector
Return Whether Signature is Genuine or Fraudulent

Fig. Sequence Diagram of the Proposed system

3.1. ALGORITHM:

STEP 1: User uploads an image.

STEP 2: The CNN Model finds the user for input signature

STEP 3: Later the Image is compared with Genuine Image to determine whether the input image is fraudulent or not.

STEP 4 : This is done using the Algorithms SIFT and ORB

STEP 5: Output is displayed(Genuine - If its not fraudulent,Fraudulent - If the Signature is fraudulent)

4. RESULTS

The Signature Fraudulent detection project is working efficiently as desired. The metrics have been assessed for the all the algorithms and the accuracy for the algorithms have been obtained as follows:
CNN = 95%

The ORB fraudulent detection confusion matrix is as follows:

Actual \ Predicted	True	False
True	53/55	2/55
False	0/5	5/5

The SIFT fraudulent detection confusion matrix is as follows:

Actual \ Predicted	True	False
True	54/55	1/55
False	1/5	4/5

Hence as we know accuracy = (True-positive + True-negative) / Total no.of images.
Accuracy of ORB and SIFT are obtained as follows:
ORB = 96.67 %
SIFT = 96.67 %
As observed SIFT and ORB are similar in accuracy, but ORB detector is best to detect fraudulent signatures as it was able to find the maximum number of fraudulent images compared to SIFT algorithm.

When tested with various sample images the results are as follows:

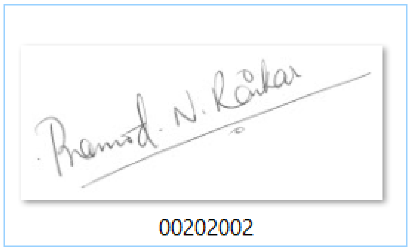
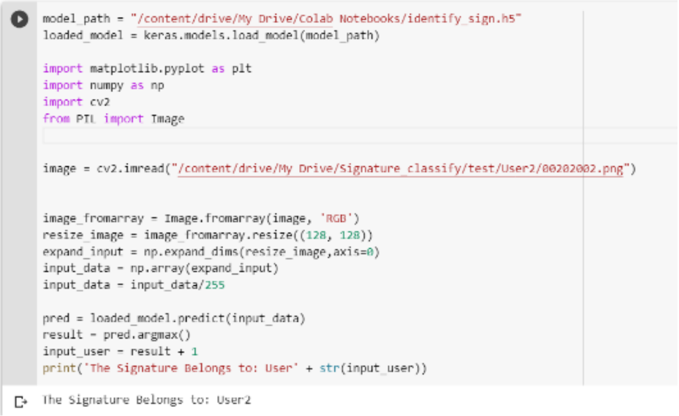


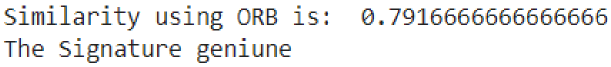
Fig. Query Image - 00X0N00Y
(X - Signed By, N - Sample no., Y - Signature of)



```
model_path = "/content/drive/My Drive/Colab Notebooks/identify_sign.h5"
loaded_model = keras.models.load_model(model_path)

import matplotlib.pyplot as plt
import numpy as np
import cv2
from PIL import Image


image = cv2.imread("/content/drive/My Drive/Signature_classify/test/User2/00202002.png")


image_fromarray = Image.fromarray(image, 'RGB')
resize_image = image_fromarray.resize((128, 128))
expand_input = np.expand_dims(resize_image,axis=0)
input_data = np.array(expand_input)
input_data = input_data/255

pred = loaded_model.predict(input_data)
result = pred.argmax()
input_user = result + 1
print('The Signature Belongs to: User' + str(input_user))
```

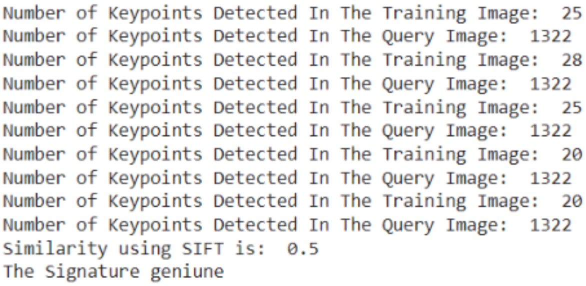
```
The Signature Belongs to: User2
```

Fig. CNN classifies input image as signature of User 2



```
Similarity using ORB is:  0.7916666666666666
The Signature geniune
```

Fig. ORB detector finds signature genuine



```
Number of Keypoints Detected In The Training Image:  25
Number of Keypoints Detected In The Query Image:  1322
Number of Keypoints Detected In The Training Image:  28
Number of Keypoints Detected In The Query Image:  1322
Number of Keypoints Detected In The Training Image:  25
Number of Keypoints Detected In The Query Image:  1322
Number of Keypoints Detected In The Training Image:  20
Number of Keypoints Detected In The Query Image:  1322
Number of Keypoints Detected In The Training Image:  20
Number of Keypoints Detected In The Query Image:  1322
Similarity using SIFT is:  0.5
The Signature geniune
```

Fig. SIFT detector finds signature genuine

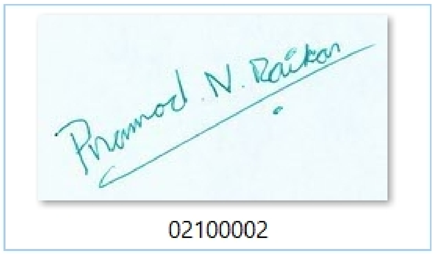
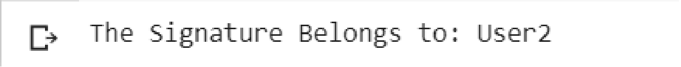
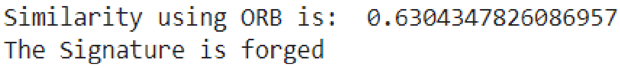


Fig. Forged Query Image



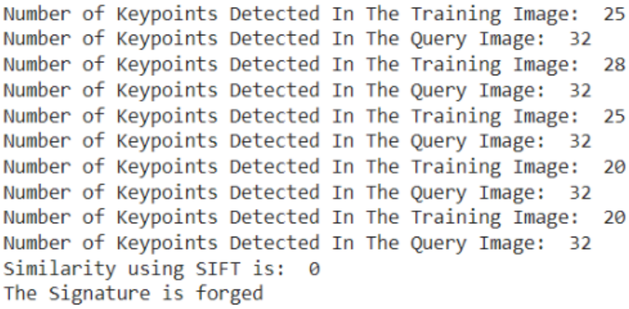
```
The Signature Belongs to: User2
```

Fig. CNN classifies input image as signature of User 2



```
Similarity using ORB is:  0.6304347826086957
The Signature is forged
```

Fig. ORB detector finds signature fraudulent



```
Number of Keypoints Detected In The Training Image:  25
Number of Keypoints Detected In The Query Image:  32
Number of Keypoints Detected In The Training Image:  28
Number of Keypoints Detected In The Query Image:  32
Number of Keypoints Detected In The Training Image:  25
Number of Keypoints Detected In The Query Image:  32
Number of Keypoints Detected In The Training Image:  20
Number of Keypoints Detected In The Query Image:  32
Number of Keypoints Detected In The Training Image:  20
Number of Keypoints Detected In The Query Image:  32
Similarity using SIFT is:  0
The Signature is forged
```

Fig. SIFT detector finds signature fraudulent

5. LIMITATIONS

Limitations of CNN Model :

- CNN does not encode the position and orientation of objects.
- Lack of ability to be spatially invariant to the input data.
- Lots of training data is required.

Limitations of SIFT Algorithm :

- Takes long time (SURF provides similar performance while running faster)
- Generally doesn't work well with lighting changes and blur

6. CONCLUSIONS AND FUTURE WORK

The system successfully recognizes and identifies the signature holder accurately and with the help of signature fraudulent detection algorithms SIFT and ORB it verifies the input image is fraudulent or not.

The planned system is highly economical in recognizing and sleuthing the forgeries at runtime and therefore the responsibility of the system can be magnified by training the extracted features on the Neural Networks by storing the extracted features. Negligible misclassification or error is required in such sensitive applications although it's at the cost of a High Recognition Rate (HRR). Different aim is that the probability of a forgery signature as if it’s a real one is zero. As a future work, we may also aim at increasing the resultant system accuracy by trying new and better parameter coefficients that increases the deviation between real and forged signatures

Accuracy of the model proposed in our project is 95%, 96.67% and 96.67% respectively for CNN, SIFT and ORB. We can further increase the accuracy by introducing more efficient fraudulent detection algorithms.

REFERENCES

[1] Jivesh Poddara, Vinanti Parikha, Santosh Kumar Bhartia,”Offline Signature Recognition and Forgery Detection using Deep Learning”, The 3rd International Conference on Emerging Data and Industry 4.0 (EDI40), April 6 - 9, 2020, Warsaw, Poland

[2] Zagoruyko, S., & Komodakis, N. (2015). “Learning to compare image patches via convolutional neural networks.” In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4353-4361).

[3] Fahmy, M. M. (2010). “Online handwritten signature verification system based on DWT features extraction and neural network classification.” Ain Shams Engineering Journal, 1(1), 59–70.

[4] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). “Imagenet classification with deep convolutional neural networks.” Advances in neural information processing systems (pp. 1097-1105).

[5] Khalajzadeh, H., Mansouri, M., & Teshnehlab, M. (2012). “Persian signature verification using convolutional neural networks.” International Journal of Engineering Research and Technology, 1(2), 7-12.

[6] Batista, L., Granger, E., & Sabourin, R. (2012). “Dynamic selection of generative discriminative ensembles for off-line signature verification.” Pattern Recognition, 45(4), 1326-1340.

[7] Shahane P.R., Choukade A.S., & Diyewar A.N. (2015) “Online biometric authentication mistreatment Matlab.” International Journal Of Innovative analysis in Electrical, Physics, Instrumentation, and management Engineering

[8] Eman Alajrami, Belal A. M. Ashqar, Bassem S. Abu-Nasser, Ahmed J. Khalil, Musleh M. Musleh, Alaa M. Barhoom,Samy S. Abu-Naser, “Handwritten Signature Verification using Deep Learning” International Journal of Academic Multidisciplinary Research (IJAMR), ISSN: 2643-9670, Vol. 3 Issue 12, December – 2019, Pages: 39-44

[9] Jerome Gideon S, Anurag Kandulna, Aron Abhishek Kujur, Diana A, Kumudha Raimond, "Handwritten Signature Forgery Detection using Convolutional Neural Networks", 8th International Conference on advances in Computing and Communication (ICACC-2018)

[10] T. M. Ghanim and A. M. Nabil, "Offline Signature Verification and Forgery Detection Approach," 2018 13th International Conference on Computer Engineering and Systems (ICCES), 2018, pp. 293-298, doi: 10.1109/ICCES.2018.8639420.