



<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper presented at *2022 International Symposium on iNnovative Informatics of Biskra, ISNIB 2022, Biskra, 7 December 2022 through 8 December 2022*.

Citation for the original published paper:

Maamouli, K., Benhamza, H., Djeflal, A., Cheddad, A. (2022)

A CNN based architecture for forgery detection in administrative documents

In: *2022 International Symposium on iNnovative Informatics of Biskra, ISNIB 2022*

Institute of Electrical and Electronics Engineers (IEEE)

<https://doi.org/10.1109/ISNIB57382.2022.10076089>

N.B. When citing this work, cite the original published paper.

©2022 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:bth-24483>

A CNN based architecture for forgery detection in administrative documents

Maamouli Khadidja ¹
khadidja.maamouli@univ-biskra.dz

Benhamza Hiba ²
hiba_benhamza@hotmail.fr

Djeffal Abdelhamid ²
abdelhamid.djeffal@univ-biskra.dz

Abbas Cheddad ³
abbas.cheddad@bth.se

¹ *Computer science department
Biskra University*

² *Computer science department
LESIA Laboratory
Biskra, Algeria*

³ *Blekinge Institute of Technology
Karlskrona, Sweden*

Abstract—The use of digital documents is knowing a widespread in different daily administrative and economic transactions. Simultaneously, the forgery of many documents becomes a crime that costs billions to states and companies. Several researchers tried to develop techniques that automatically detect forged documents using machine learning and image processing. With the immense success of deep learning applications, we employ, in this work, a convolutional neural network architecture that uses a gathered dataset of forged and authentic administrative documents. The results obtained on our dataset of 493 documents reached 73.95 % accuracy and 97.3

Key words- Forgery detection, Image processing, Deep learning

I. INTRODUCTION

A scanned administrative document may now be readily changed and amended thanks to advances in digital technology, image processing and editing software. It is challenging for humans to determine whether or not a document has been falsified visually. The prevalence of digitally modified counterfeits in popular media and over the Internet is rapidly increasing. It implies significant flaws and undermines the legitimacy of digital administrative documents. We now live in a time when digital data security, such as pictures and videos, is more vital than ever. The art of tampering with visual content is no longer limited to specialists, allowing an inexperienced person to effortlessly alter image content and its meaning without leaving any trace.

Forgery is defined as the use of a fraudulent document, signature, or other imitation of a valuable entity to deceive another person. Those who perpetrate forgeries are frequently charged with fraud. Contracts, identity cards, and legal certificates are all examples of documents that can be forged. Forged papers are created in order to acquire illicit short-term or long-term benefits. This poses a significant danger to the nations many legal, economic, and security sectors. Due to the lack of a publicly accessible dataset for experimentation purposes, particularly for administrative documents, many research efforts have been made to solve the problem of document forgery by developing and proposing models, approaches, and techniques for detecting forgery from documents. This paper provides a brief overview of digital documents and their methods, the principles of deep learning

method, specifically Convolutional Neural Network (CNN), the main existing works related to image and forgery detection, a detailed description of the proposed method, and finally, the results.

II. ADMINISTRATIVE DIGITAL DOCUMENTS AND THEIR FORGERY DETECTION METHODS

The administrative document refers to a document and the information contained in a document created, received, or maintained by a court for recording administrative, financial, administrative, or managerial functions, policies, decisions, procedures, or the organization of operations or other court activities, subject to exceptions, and used to identify a person or his right or permission [4].

Digital image forgery has drawn increased interest from the scientific community in recent years. Because of this, scientific research on the subject has significantly increased, and various methods and tactics, notably passive ones, have been proposed to detect photo fraud. Verifying a digital images legitimacy is the aim of picture forgery detection. The two picture authentication methods are active and blind or passive [4].

A. Passive methods

1) Without any signature or watermark of the original picture of the sender, passive or blind forgery detection approaches employ the received image to judge its validity or integrity. It is founded on the notion that digital forgeries may not leave any visible signs of tampering but do leave behind statistical aberrations. This method is popular since it does not require prior knowledge of the image. It is based on the notion that any alteration to the picture would affect its consistency or statistics [7]. There are two categories of passive methods:

1) *Forgery independent techniques*: This includes methods such as image retouching and lighting inconsistencies.

2) *Forgery dependent techniques*: This includes methods such as image splicing, CMFD, and JPEG compression properties [9].

1) Copy-move forgery detection (CMFD): Most research on the subject of picture fraud detection is focused on copy-move forgery. One or more areas are copied and

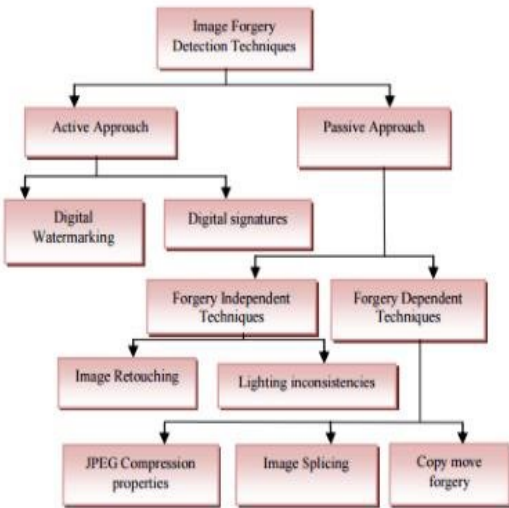


Fig. 1. Categorization of image forgery detection.

pasted within the same picture in copy-move forgery. It is easy to merge these sections into the backdrop because their texture and colour are similar. CMFD is commonly used in document forgery to hide or add information. In most cases, the CMFD method begins with a pre-processing phase to improve picture attributes. The photos may then be separated into blocks and transformed into grayscale hues. After pre-processing, a feature extraction phase is employed to capture information regarding the properties of the image regions of interest. A matching step follows feature extraction, which looks for similarities between two or more features in the picture. Finally, the results of the CMFD method may be displayed to show and locate any tampered locations in the fabricated image [1].

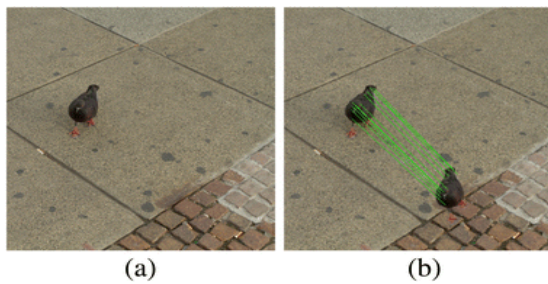


Fig. 2. Original picture vs fake picture: CMFD.

2) Splicing forgery detection: Splicing forgery is the process of integrating parts from many photographs into a single image in order to make a fake image. Even if no post-processing is done, the tampering traces are invisible and difficult to follow. It can be found by looking for the splicing border, the effect of splicing on picture statistics, or the direction of light incident on the image surfaces. Splicing detection is a challenging

problem involving examining composite regions using several techniques. Rapid transitions between the several sections that are merged and their backgrounds offer helpful traces for splicing detection in the image under consideration [14].

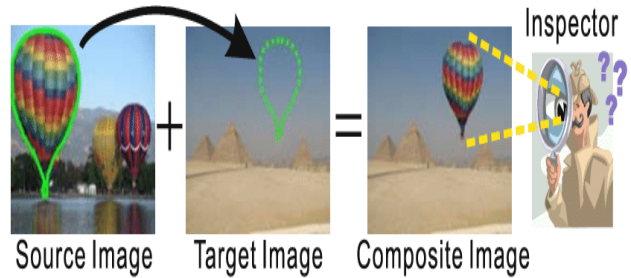


Fig. 3. Original picture vs fake picture: splicing.

- 3) Imitation forgery detection: This type of forgery is generally used for document forgery where there is text to add or modify by trying to find the exact font properties of a document. By looking for the identical font attributes of a document, the fraudster may be able to insert or substitute information. Consequently, the final forged digital document comprises words of various types, font sizes, and other characteristics [6].
- 4) Image retouch detection: The photographs are altered less in digital image editing. Some aspects of the image are improved as a result of this. It is used to enhance or improve specific aspects of the picture. Before merging two images, retouching may necessitate rotating, resizing, or stretching one of them. This is a prevalent and well-publicised sort of picture alteration. In image editing, cloning a portion of a picture is also quite popular. Because there is no drastic shift in the different areas of the picture, detection is extremely tough.

B. Active methods

Active forgery detection approaches, such as digital watermarking or digital signatures, insert a known authentication code into the picture content before it is transferred, allowing the recipient to verify its validity. These approaches can only examine pre-processed photos, but they provide a significantly greater level of assurance [8]

- 1) Steganography: Steganography is an active approach that has been used for picture authentication and integrity verification in several studies. Steganography is a communication system that embeds a secret message into a digital image known as the cover image. A stego picture is sent from a transmitter to a recipient. The secret message recovered from the stego picture can then be viewed by the receiver [4].
- 2) Cellular automata: Cellular automata form a method for extracting statistical information from a digital image using the lower, upper and singular value decomposition. One may employ singular value decomposition, lower and upper decomposition, and one-dimensional cellular

automata to produce a cypher key. This key includes picture features and is tied to digital images; hence any tiny change in the digital image content will change the key value without exception [16].

- 3) A digital watermark: is a marking concealed inside a noise-tolerant signal like audio, video, or picture data. It is usually used to determine who owns the copyright to a signal. The act of hiding digital information in a carrier signal is known as "watermarking". The concealed information should but is not required to have a relationship to the carrier signal. Digital watermarks can be used to validate the carrier signal's validity or integrity and identify the signal's owner. Its commonly used to track down copyright infringements and authenticate banknotes [5].

III. DEEP LEARNING

Image interpretation is essential for detecting administrative fraud. Although the interpretation of images by traditional machine learning algorithms relies mainly on features developed by experts, computer vision is the best application of machine learning. Today, deep learning has taken an enormous step forward in all fields.

A. Definition

Deep Learning (DL) is a class of machine learning techniques for training neural networks with several internal layers and a large number of parameters. In machine learning, DL is the most extensively utilized approach. It is also the most powerful classification approach, which justifies its use in practically all machine learning applications and domains [15].

B. Type of neural networks

- Perceptron: is a binary classifier that is a supervised learning system that divides data into two groups.
- Feed Forward Neural Network: is an artificial neural network wherein connections between the nodes do not form a cycle.
- Recurrent Neural Network (RNN): is designed to save the output of a layer, RNN is fed back to the input to help in predicting the outcome of the layer.
- Convolutional neural networks: Deep networks, called convolutional neural networks (CNN), are incredibly well suited for signal and image processing applications. They exhibit every trait of neural networks. In order to create them, processing layers are stacked down to the levels that do regression or classification. Sharing and connecting parameters, as well as using fewer convolution layers, facilitate learning and network functioning. This method, CNN, is what we investigate in this research [10].

IV. RELATED WORK: DL AND FORGERY DETECTION

Most previous research that dealt with forgery took advantage of natural images' contents in their research. This distinguishes our study as it takes administrative papers for forgery detection. However, as an initial step, we use methods

that reveal forgery in natural images as a starting point for our study.

- Copy Move and Splicing Image Forgery Detection using CNN: Devjani Mallick et al. used deep learning using CNN to detect forgery in images with copy-move and splicing image forgery detection. They used two famous CNN architectures, VGG 16 and VGG19, with a dataset of 1000 pictures. They reported an accuracy of 71.6%, using VGG16, but when using VGG19, they gained 72.9% accuracy.
- An efficient method for image forgery detection based on trigonometric transforms and deep learning: Faten Maher AlAzrak et al. used 220 images for their study 110 tampered with 110 original; they used different methods to detect a forgery in their pictures which are listed below [2]:
 - DDCT & DWT with accuracy 90%
 - DDCT with accuracy 80%.
 - DDCT & DWT with accuracy 100%.
 - DST with accuracy 60%.
 - DST & DWT with accuracy 94.4%.
 - DDFT with accuracy 60%.
 - DDFT & DWT with accuracy 95%.
 - DDFT with accuracy 60%.
 - DDFT & DWT with accuracy 100%.
 - No transform 80%.
- Image Forgery Detection Using Deep Learning by Re-compressing Images :By S. S. Ali et al. proposed a robust deep learning-based technique for identifying picture forgeries in the setting of double image compression. The difference between a picture's original and recompressed versions is utilized for training the model. The difference between the fabricated and original images is found by recompressing the forged picture. Due to the difference in the source of the forged section and the original part of the picture, the forged part is now emphasized. The suggested method is simple, and its results show that it outperforms existing methods. The experiment's findings are encouraging, with an overall validation accuracy of 92.23% [3].
- Automated image splicing detection using deep CNN learned features and ANN-based classifier: In the paper of S.Nath et al., a blind picture splicing detection method is proposed, which employs a backbone of deep convolutional residual networks, followed by a fully connected classifier network that distinguishes between authentic and changed images [13].
- A robust copy-move forgery classification using end-to-end convolution neural network: S. Kumar et al. describe a robust copy-move forgery classification using an end-to-end convolution neural network. They introduce a deep neural network-based approach for classifying photographs depending on whether they include any copy-move forgeries with satisfactory outputs. The following project aims to categorize all photographs with copy-

move forgeries that have been resized, rotated, or compressed at various levels. They suggested a new CNN model with a 93-95% accuracy with similar or hybrid datasets [11].

V. DESIGN OF A DEEP LEARNING ARCHITECTURE FOR FORGERY DETECTION

We will present our convolutional neural network design for classifying administrative documents. We explain the general architecture of our CNN-based classification model. The detailed operations of our model will then be presented in the following parts.

A. General architecture

Various images are used as inputs in our categorization model. First, we create the dataset collection and preprocessing. Following the datasets separation into three parts, we take the first two partitions and validation and perform training. In order to create a model that we can utilize, our system first collects characteristics from the provided images and learns their representations. The final dataset component is then used to wrap up the testing procedure, after which the validation metrics are computed. Two scenarios are possible based on the results: If the findings are confirmed, we will have a model that can be used to spot administrative document fabrication. If the model is unsatisfactory, we analyze the parameters, make changes to the training process, and continue through all the initial phases until we have a satisfactory model.

B. Detailed architecture

Our system goes through four steps in order to develop a deep learning model for detecting and classifying documents whether they are forged or authentic:

- 1) Data collection: Benhamza et al. [4] produced the dataset that we used in our study. It includes files in pdf format. This dataset comprises of, 232 authentic images and 373 forged ones. We divided it into 3 subset (train, validation, test)
- 2) Training phase: During training, we execute a forward pass batch and transmit it across the network. Once we obtain the output, we compare the expected output to the actual labels. Once we know how closely the expected values match the actual values, we change the weights in the network to more closely match the anticipated values to the actual values (labels). We repeat this procedure for each batch until we have covered all of the samples in our training group. This is for one batch. During the training phase, we repeat as many times as necessary to achieve the appropriate degree of accuracy.
- 3) Test phase: The test phase follows the training phase, and it is at this point that a CNN model is ready to categorize images.
- 4) Use phase: We will have a model that we may apply in other domains to identify the forgery of administrative documents if the model has successfully passed the test phase.

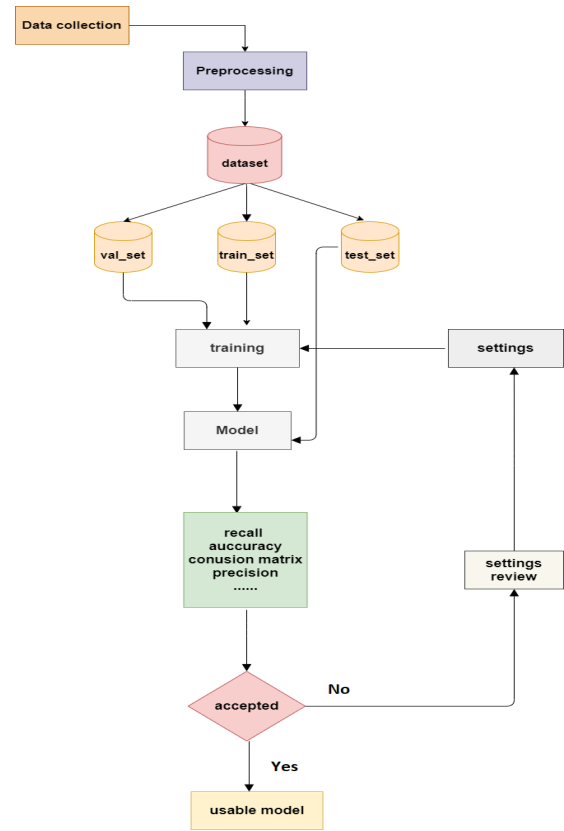


Fig. 4. General architecture.

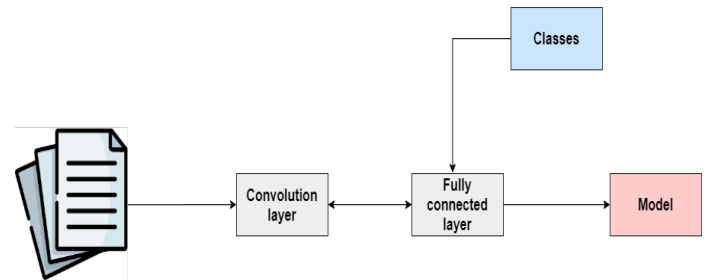


Fig. 5. Training phase

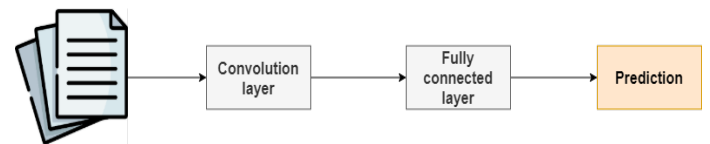


Fig. 6. Test phase

VI. RESULTS AND DISCUSSION

After implement our deep learning CNN model with 9 layers, we train it with our dataset which splitted into 3 subsets: training, validation and testing subset. The following table show us the number of images in each folder:

To visualize the performance of our deep learning CNN over time during training, we created:

	Training	Validation	Testing	Total
Forged	260	75	38	373
Authentic	163	44	24	231
Total	423	119	62	604

TABLE I
PROPOSED STRUCTURE

- An "accuracy" plot on the train "acc" dataset over the training epochs (the orange curve is the validation accuracy and the blue curve represents the validation loss).

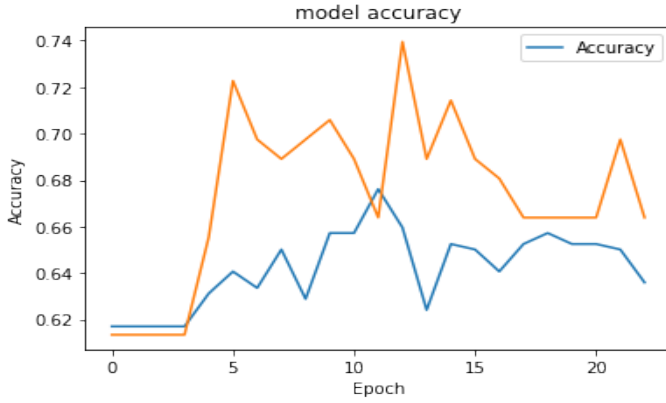


Fig. 7. Model accuracy

- A "loss" graph on the train "loss" dataset over the training epochs (the orange curve is the validation loss).

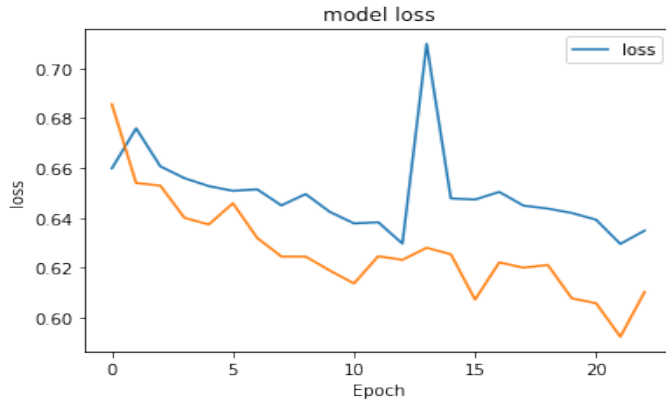


Fig. 8. Model loss

- A "val-accuracy" graph on the validation dataset Val acc over the training epochs.
- A "val-loss" plot on the "val-loss" validation dataset over the training epochs.

In deep learning, the control point is the stored model weights when there is an improvement in classification accuracy on the validation dataset. These weights can be used in validation, which can be used to make predictions as is or as a basis for further training. In our case, the best validation value accuracy (val-accuracy) is reached in epoch 23.

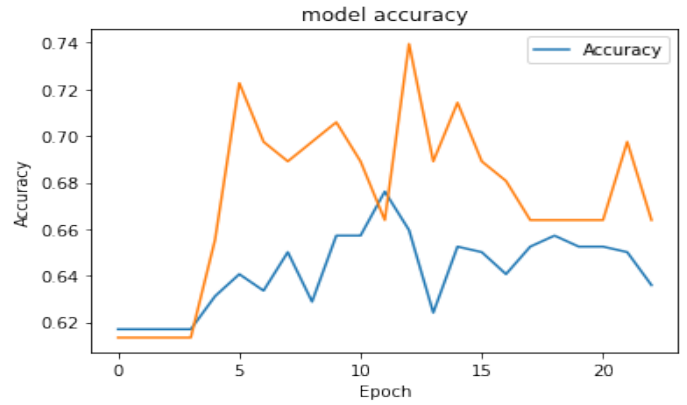


Fig. 9. Model val accuracy

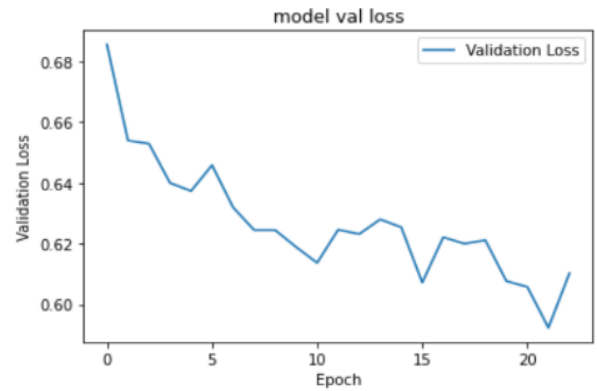


Fig. 10. Model val loss

- Training accuracy accuracy 0.6359
- Training loss loss is 0.6349.
- Training precision is 0.6359
- Training AUC 0.6829
- Validation accuracy Val accuracy 0.7395
- Validation loss Val loss 0.6102
- Validation precision is 0.6639
- Validation AUC 0.7409

	Acc	Val_acc	Recl	Val_recl	Precs	Val_precs
CNN	63.59%	73.95%	64%	66.4%	63.59%	66.39%

TABLE II
TABLE OF RESULTS OF THE PROPOSED ARCHITECTURE.

In the previous table, we note that all values recorded in training have increased in validation. In the test phase, we will use the third subset to test how good our CNN model is doing on new images from the dataset. After testing our model on the third subset, we will obtain the following result:

Confusion matrix: The confusion matrix is a table that is often used to describe the performance of a classification model.

	Accuracy	Precision	Recall
Metrics	67.7%	66%	97.3%

TABLE III
MODEL TESTING ARCHITECTURE.

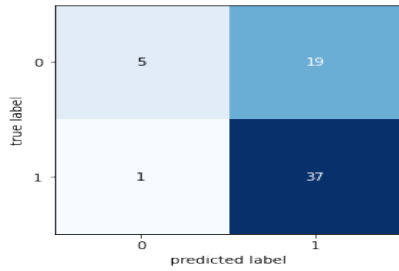


Fig. 11. Confusion matrix

- The model could correctly predict 37 images out of 38 from the forged test folder.
- The model could correctly predict 5 images out of 24 from the authentic test folder

A. Comparison of our work with previous works

Methods	Accuracy	Recall
Our proposed architecture	73.95%	97.3%
VGG19	72.9%	-
VGG16	71.6%	-
Machine learning (SVM)	88%	-

TABLE IV
TABLE OF COMPARISON OF OUR WORK WITH PREVIOUS WORKS.

Although all previous works on image forgery detection were characterized by the same image content and large datasets, our system obtained close and comparable results. The proposed system can attract an administrator's attention to a set of presented documents for more verification to detect possible forgery.

VII. CONCLUSION

The forgery of digital papers has emerged as one of the most well-known crimes due to the growing use of digital documents in administration. We have researched a number of relevant works as well as their approaches and strategies for identifying forged digital documents to combat this crime. We have presented a system based on automatic learning via the convolutional neural network approach by proposing an architecture that can detect forged documents.

We have used a training database of 423 samples representing authentic and forged administrative documents to validate our proposed method. We trained our CNN architecture on those images, and we obtained a precision of 73.95 % and a recall 97.3%.

For future work, we suggest some ideas that can improve our system, such as:

- Using other datasets for training, validation, and testing.
- Increase the number of samples of authentic documents.

- Build custom CNN models for each document type and after classifying them from the first model which is the general model, we step the image for the custom model to confirm our result and minimize the defect.

REFERENCES

- [1] Nor Bakiah Abd Warif, Ainuddin Wahid Abdul Wahab, Mohd Yamani Idna Idris, Roziana Ramli, Rosli Salleh, Shahabuddin Shamshirband, and Kim-Kwang Raymond Choo. Copy-move forgery detection: survey, challenges and future directions. *Journal of Network and Computer Applications*, 75:259–278, 2016.
- [2] Faten Maher Al_Azrak, Ahmed Sedik, Moawad I Dessowky, Ghada M El Banby, Ashraf AM Khalaf, Ahmed S Elkorany, et al. An efficient method for image forgery detection based on trigonometric transforms and deep learning. *Multimedia Tools and Applications*, 79(25):18221–18243, 2020.
- [3] Syed Sadaf Ali, Iyyakutti Iyappan Ganapathi, Ngoc-Son Vu, Syed Danish Ali, Neetesh Saxena, and Naoufel Werghi. Image forgery detection using deep learning by recompressing images. *Electronics*, 11(3):403, 2022.
- [4] Hiba Benhamza, Abdelhamid Djeflal, and Abbas Cheddad. Image forgery detection review. In *2021 International Conference on Information Systems and Advanced Technologies (ICISAT)*, pages 1–7. IEEE, 2021.
- [5] Oussama Benrhouma, Houcemeddine Hermassi, Abd El-Latif, A Ahmed, and Safya Belghith. Chaotic watermark for blind forgery detection in images. *Multimedia Tools and Applications*, 75(14):8695–8718, 2016.
- [6] Romain Bertrand, Petra Gomez-Krämer, Oriol Ramos Terrades, Patrick Franco, and Jean-Marc Ogier. A system based on intrinsic features for fraudulent document detection. In *2013 12th International conference on document analysis and recognition*, pages 106–110. IEEE, 2013.
- [7] Gajanan K Birajdar and Vijay H Mankar. Digital image forgery detection using passive techniques: A survey. *Digital investigation*, 10(3):226–245, 2013.
- [8] RA Dobre, RO Preda, and AE Marcu. Improved active method for image forgery detection and localization on mobile devices. In *2018 IEEE 24th International Symposium for Design and Technology in Electronic Packaging(SIITME)*, pages 255–260. IEEE, 2018.
- [9] Navpreet Kaur Gill, Ruhi Garg, and Er Amit Doegar. A review paper on digital image forgery detection techniques. In *2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, pages 1–7. IEEE, 2017.
- [10] Jean-Claude Heudin. *Comprendre le deep learning: une introduction aux réseaux de neurones*. Science-eBook., 2016.
- [11] Sanjeev Kumar and Suneet K Gupta. A robust copy move forgery classification using end to end convolution neural network. In *2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO)*, pages 253–258. IEEE, 2020.
- [12] Devjani Mallick, Mantasha Shaikh, Anuja Gulhane, and Tabassum Maktum. Copy move and splicing image forgery detection using cnn. In *ITM Web of Conferences*, volume 44, page 03052. EDP Sciences, 2022.
- [13] Souradip Nath and Ruchira Naskar. Automated image splicing detection using deep cnn-learned features and ann-based classifier. *Signal, Image and Video Processing*, 15(7):1601–1608, 2021.
- [14] VV Nath, GKS Gaharwar, and RD Gaharwar. Comprehensive study of different types image forgeries. 2015.
- [15] SEBTI MOHAMED RIAD. Diabetic retinopathy detection based on retinography images.
- [16] A Pahlavan Tafti, MV Malakooti, M Ashourian, and S Janosepah. Digital image forgery detection through data embedding in spatial domain and cellular automata. In *The 7th International Conference on Digital Content, Multimedia Technology and its Applications*, pages 11–15. IEEE, 2011.