DETECTION OF COPY-MOVE FORGERY IN DIGITAL IMAGES USING DIFFERENT COMPUTER VISION APPROACHES

- Name: Younis Abdalla
- Memorial University of Newfoundland and Labrador
- Engineering and Applied Science

- Supervisor Committee
- Dr. M. Tariq Iqbal
- Dr. M. Shehata
- Dr. Minglun Gong

Outline

- Introduction
- Motivations
- Literature review
- Related works
- proposed methods
  - A classic technique using enhanced PatchMatch
  - Deep learning techniques using CNNs model
  - Leveraged deep learning technique using CNNs & GAN
  - Image forgery detection based on deep transfer learning
- Summery and future work
- Contribution

Introduction

- Image forgery detection approaches are many and varied, but they generally all serve the same objectives: detect and localize the forgery.
- Copy-move forgery detection must challenge approach.
- We propose strategies and applications based on the PatchMatch algorithm and deep neural network learning (DNN).
- Test and evaluate our copy-move forgery detector algorithms were presented.

Introduction

Types of digital image forgery:

- Image Retouching
- Image Splicing (Copy-Paste)
- Image Cloning (Copy-Move)
- Image Composition
- Image Enhancement

What is the digital image copy-move forgery ?

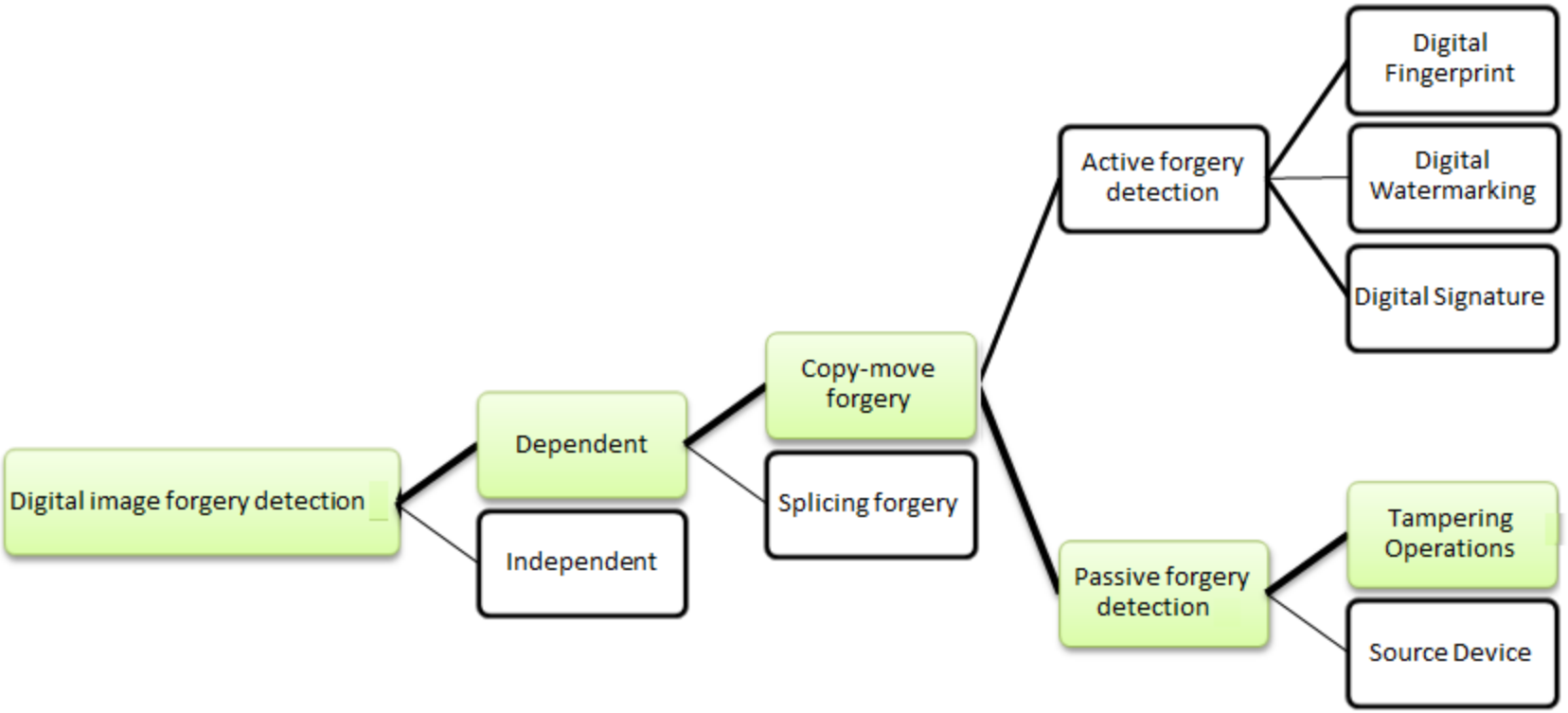
It is a specific type of image manipulation were a part of the image itself is copied and pasted into another part of the same image

Motivations

- There are many problems regarding the authenticity of the digital image in any process.
  - Using images in the courtroom as evidence.
  - In insurance claims for accident verifications.
  - In medically related procedures.
  - Promote fake news ban.
  - Commercial digital images and photography.
- All the above and else, motivated us to dive in this industry to come up with a possible solution

Forgery detection classification Literature review:

- Forgery detection classifications based on the forgery type



Feature extraction Literature review:

- In the classic approach there are three types based features extractions:
  - Fourier-Mellin transform (FMT)
  - Zernike moments (ZM)
  - The polar cosine transforms (PCT)
- In deep learning approaches:
  - CNN develops a feature map from an input image ($f \in R^{h \times w \times d}$)
  - Extracting useful information out of raw pixel values; the information is considered “useful” if it can distinguish among various categories.
  - In the present work, the VGG-16 network was employed

Feature similarity and matching. Literature review:

- The enhanced PatchMatching algorithm runs:
  - Denes-field to find the nearest neighbor field (NN)

$$NN \cong s` = s + \delta(s)$$

  - Post-processing to find patch matching and reduce false alarms.

- In CNN: customized layers to do Self-correlation and Pooling procedure in percentile scheme by preference the vector' scores in percentile ranking.
  - Pearson correlation coefficient $\rho(i, j)$ which, in fact, maintains the feature similarity.
  - The goal here is to match any similar features by correlating all feature together by applying a self-correlation layer:

$$S[i] = [\rho(i, 0), \dots, \rho(i, j), \dots, \rho(i, 255)]$$

(2)(111)

Related work

- Y. Huang et al. (2011): A method called`` Improved DCT-based detection of copy-move forgery in images,`` is suggested in which instead of pixel value comparison “exact match”, quantized DCT coefficients are matched. This method can detect the type of manipulations such as JPEG compression and Gaussian blurring. However, the proposed method fails for any type of geometric transformations of the block such as rotation, scaling, etc.

Related work

- Yaqi Liu et al. (2017): Proposed a copy-move forgery detection based on a convolutional neural kernel network. This is a kind of data-driven local descriptor with a deep convolutional structure. They used only CoMoFoD dataset with JPEG format. However, that was the first direct bridge between the convolutional neural network and copy-move forgery detection.

Related work

- Sri Kaiyan, et al. (2018): Proposed an image forgery detection and localization using GAN and one class-classifier. This algorithm considering one scenario of adding new features patch using GAN to the original image under the assumption of passive forgery for the training task. The experiments show high accuracy of detecting and localization of this type of forgery.

Related work

- D. Sarkar (2018): Presented a comprehensive guide of how to use a transfer learning with real-world applications in Deep Learning. The idea simulates the human`s ability of transfer knowledge across tasks. Therefore, transfer learning is definitely going to be one of the key drivers for machine learning and deep learning success in mainstream adoption in the industry.

Proposed Methods

- We proposed methods that are efficient and fast for detecting Copy-Move Forgery regions even when the copied region was undergone rotation and scaling modification.
  - A classic technique using enhanced PatchMatch
  - Deep learning techniques using CNNs model
  - Leveraged deep learning technique using CNNs & GAN
  - Image forgery detection based on deep transfer learning

First Method

Classic technique using enhanced PatchMatch

Enhanced PatchMatch algorithm. The proposed model:

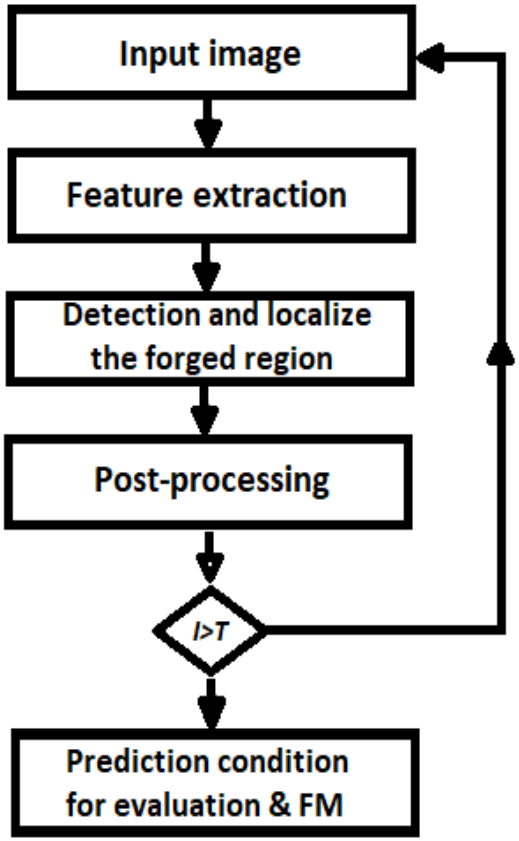
- This algorithm based on an enhanced PatchMatch approach for fast approximation nearest neighbor NN matching $S' = s + \delta(s)$ to find the correspondences between two patches.
- The keys observations were made about this searching for matching patches:
  - First, the offset field initialized using a random guess.
  - Image scan provides offset point propagation for each pixel:

$$\delta(s) = \arg \quad min_{\emptyset \in \Delta^P(s)} D(f(s), f(s + \emptyset))$$

- Random search for new candidate offsets based on the offset points:

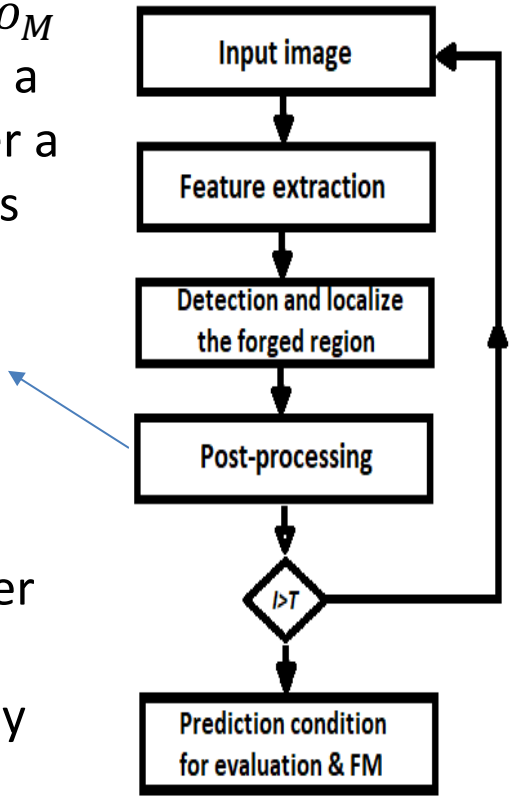
$$\delta_i(s) = \delta(s) + R_i$$

  - Once we find the good guess patch, it is mostly that the all nearby patches will have similar correspondences. Propagation will be done for NN patches.
  - Random guesses search will be done for next NN candidate offset
  - Random guesses will improve gradually.
- The algorithm will propagate good corresponding to neighboring patches and sample nearby image space to find better matching
- Using random initialization makes very few lucky guesses enabling the algorithm to propagate the matching patch's pixel very quickly



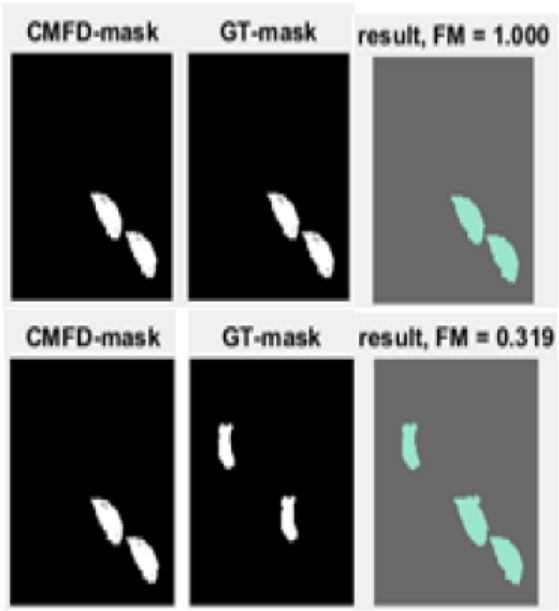
Enhanced PatchMatch algorithm. The proposed model: Post-Processing using Denes Liner Fitting

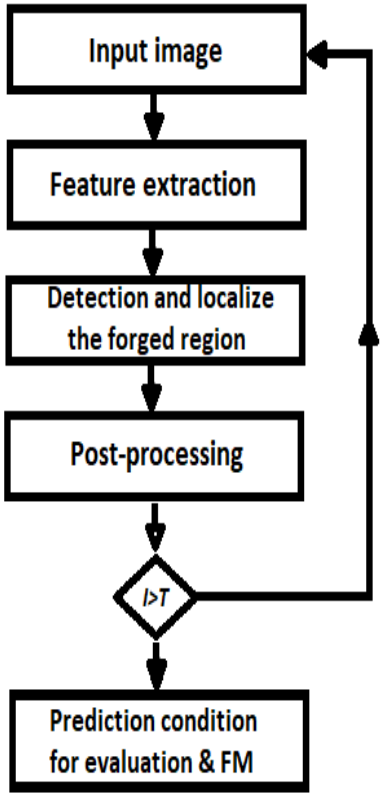
1) Using the median filter with a circular window of a radius of ρ_M
2) Computing fitting errors, using a least-squares linear model over a circular neighborhood of radius ρ_N
3) Bringing $\epsilon^2(S)$ to level T_ϵ^2
4) Deleting regional couples that are actually closer positioned than T_{D2} pixels
5) Deleting of all regions not larger than T_S pixels
6) Mirroring the regions and apply morphological dilation of the elements using a $\rho_D = \rho_M + \rho_N$



2	1	2	3	1	4	1
2	3	4	4	3	1	3
1	11	11	11	11	2	7
3	1	4	4	4	5	1
2	7	1	2	1	2	1

Enhanced PatchMatch algorithm. The proposed model: F-measure Procedure





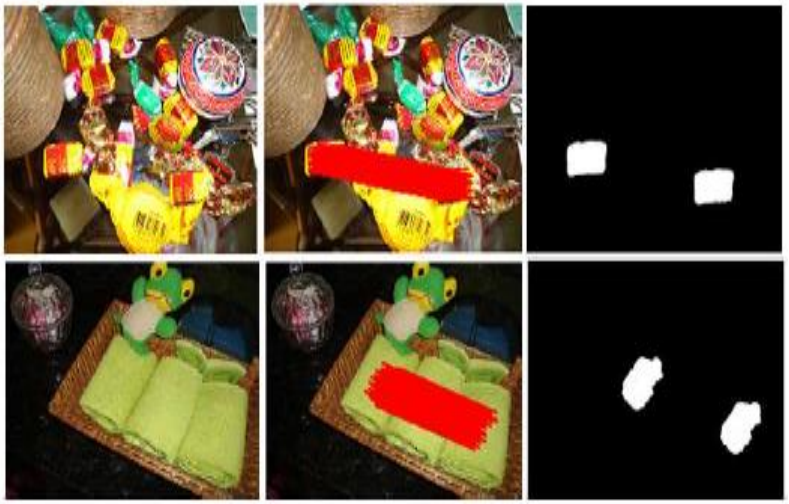
In order to designate F-measure, we first must find which is

1- True positive (TP).
2- False positive (FP).
3- False negative (FN).
4- True negative (TN).

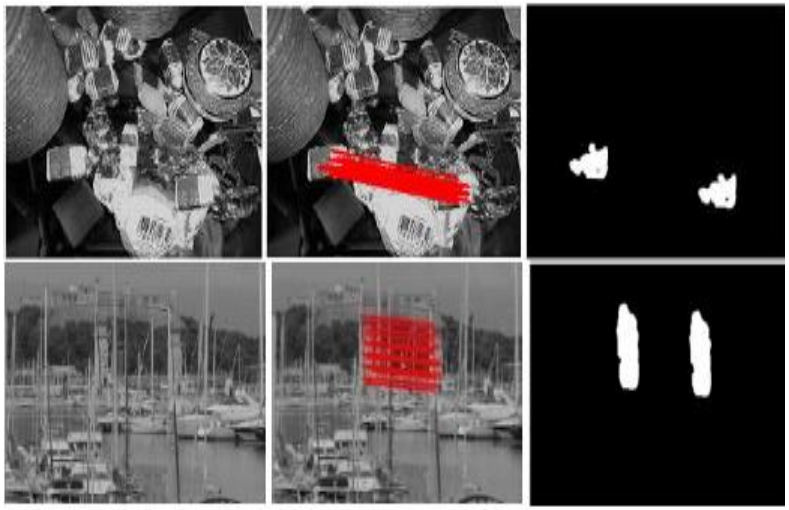
The IEEE F-measure can be written thus:

$$FM = \frac{2|TP|}{2|TP| + |FP| + |FN|}$$

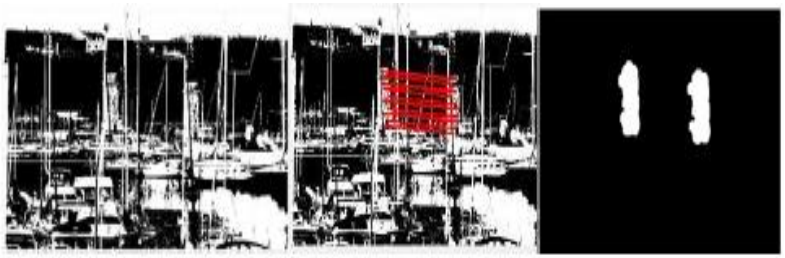
Enhanced PatchMatch algorithm. Experimental results:



Searching for forged images in the GRIP dataset: (a) forged image, (b) offset points,
(c) localization copy-move forgery mask.



Searching for forgery in a gray image: (a) forged image, (b) offset points,
(c) localization copy-move forgery mask.



Searching for forgery in black and white image: (a) forged image, (b) offset points,
(c) localization copy-move forgery mask.

Enhanced PatchMatch algorithm. Experimental results:

Shows the ability of the algorithm to detect the shearing copy-move where the rotation and scaling were applied in same time

Detecting forgery in rotation by 90°

Detecting forgery in rotation by 180°

Detecting forgery in rescaling by 80%

Detecting forgery in rescaling by 120%

Enhanced PatchMatch algorithm.

Model evaluation:

- To objectively evaluate the performance and applicability of the proposed algorithm, we decided to apply the identical function F-measure to at least three algorithms of the-state-of-the-art.
- The other algorithms were outlined by researchers in literature.
- In this examining table, it becomes clear that the conditions are different.
- However, at the end they gave very close FM values.
- The proposed algorithm was examined on many different images.
- The difference of the FM in these experiments is negligible.

	Existing methods			Proposed method	
Par.	SIFT	DyWT+SIFT	PRNU	PatchMatch	E-PatchMatch
TPR	0.953	0.888	0.960	0.539	0.952
TNR	0.791	0.818	0.978	0.993	0.994
FPR	0.046	0.111	0.039	0.460	0.046
FNR	0.029	0.091	0.003	0.105	0.004
Acc	85.5%	85%	97.7%	90.88%	90.5%
FM	0.837	0.842	0.88	0.6885	0.928

FM Parameters and the efficiency of image as given by the proposed algorithm vs. Some Literature

Enhanced PatchMatch algorithm. Conclusion

- The Proposed PatchMatch Algorithm can detect both active and passive forgery
  - Active when we have the ground truth of the forged image
  - Passive when we don't have the ground truth (we use a black sheet to force the algorithm to test whole the image as suspect to be forged and find the copy-move forgery)
- The algorithm can detect copy-move forgery under different geometric operations with high fraction values such as angle larger than 90 degrees, and scale factor over than 80% of the original region.

Second Method

Deep learning techniques using CNNs model

Detect the Copy-move Forgery using CNNs.

Literature review:

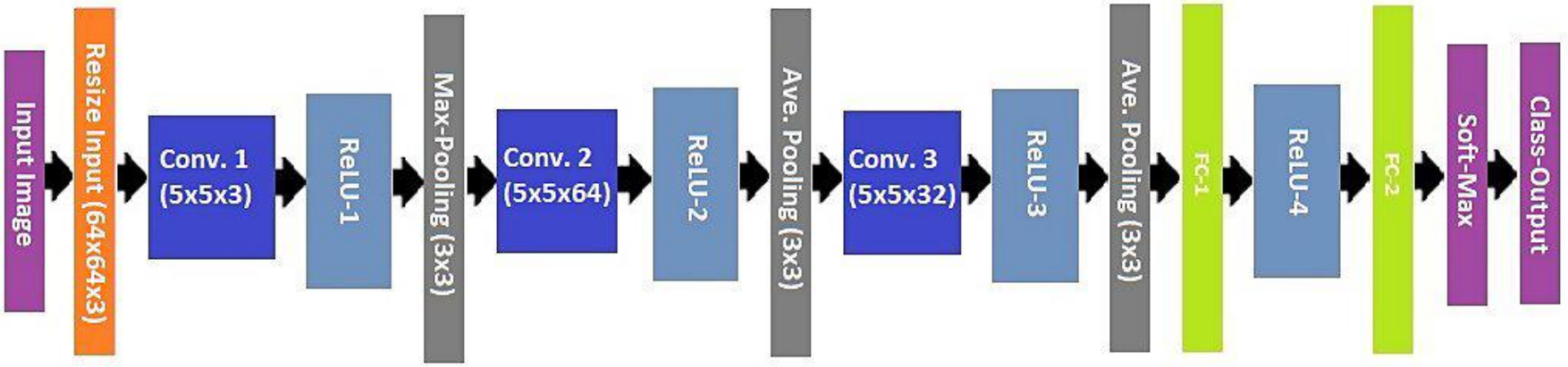
- Convolution operation is actually to sum image pixels in a local region, which would accumulate the duplicate patches of this area to be a large value. This may lead to forged images be more easily detected against pristine images.
- The convolutional in CNN models can exploit the strong spatially local correlation present in input images.
- This local correlation among image pixels is distorted when copy-move is embedded, making it different from the correlation in pristine images. The difference between natural images and distorted images can be effectively captured by CNN models.
- Nonlinear mappings in CNN models make them able to extract rich features for classifying cover images and forged images. These features, which are automatically learned by updating the network, can hardly be designed by hand

Detect the Copy-move Forgery using CNNs.

Related work:

Here summery of some related deep learning works.

CNN	Layer No.	Inventer(s)	Year	Place	Parameters No.	Error rate
LeNet	8	Yann LeCun	1988	First	60 T	N/A
AlexNet	7	Alex K. Hinton, Ilya S.	2012	First	60 M	15.5%
ZFNet	7	Matthew Z, Rob Fergus	2013	First	N/A	14.8%
Google Net	9	Google	2014	First	4 M	6.67%
VGG Net	16	Simonyan, Zisserman	2014	Second	140 M	3.6%
ResNet	152	Kaiming He	2015	First	N/A	3.75%

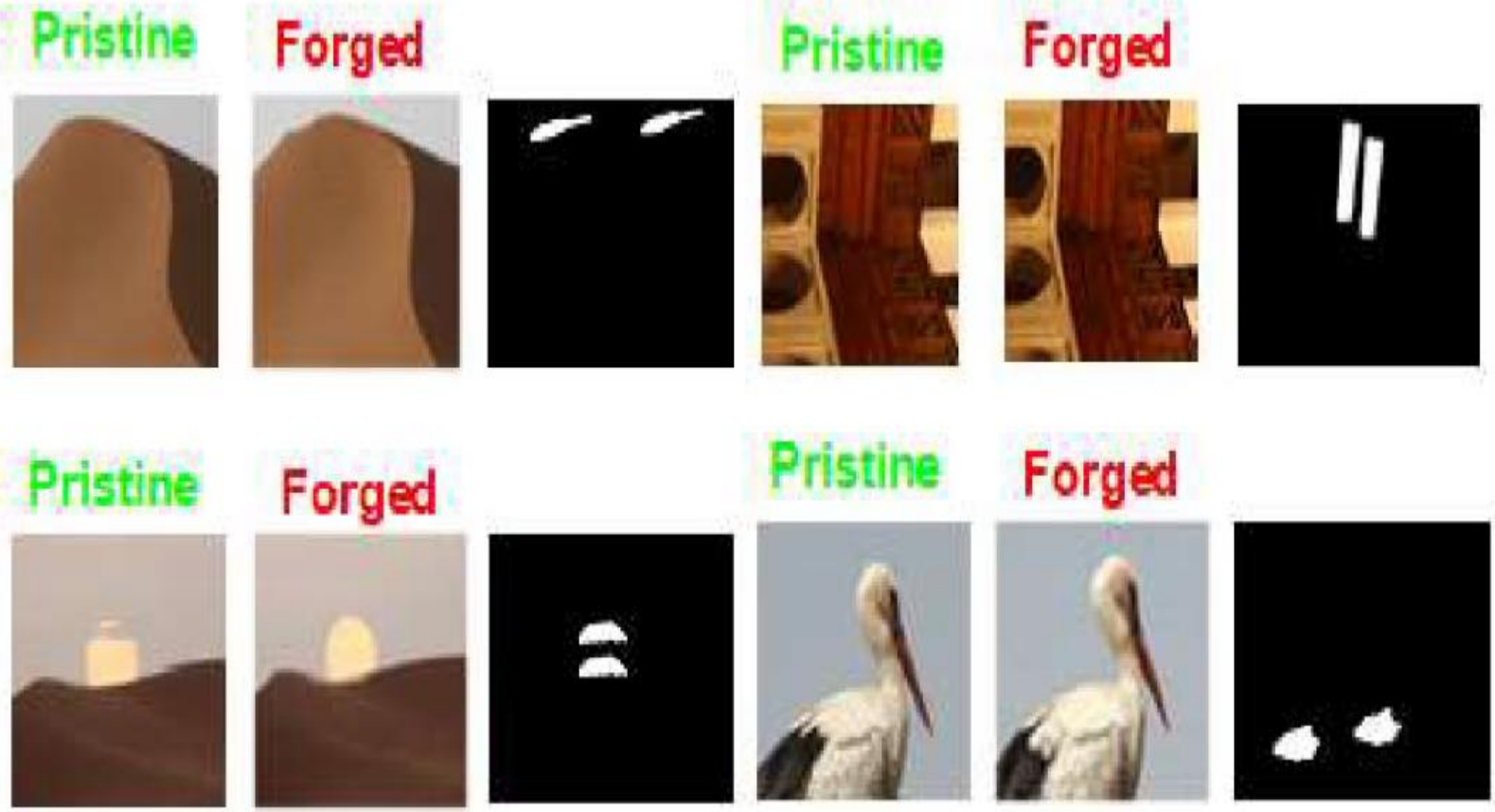


Detect the Copy-move Forgery using CNNs.

The proposed CNN model:

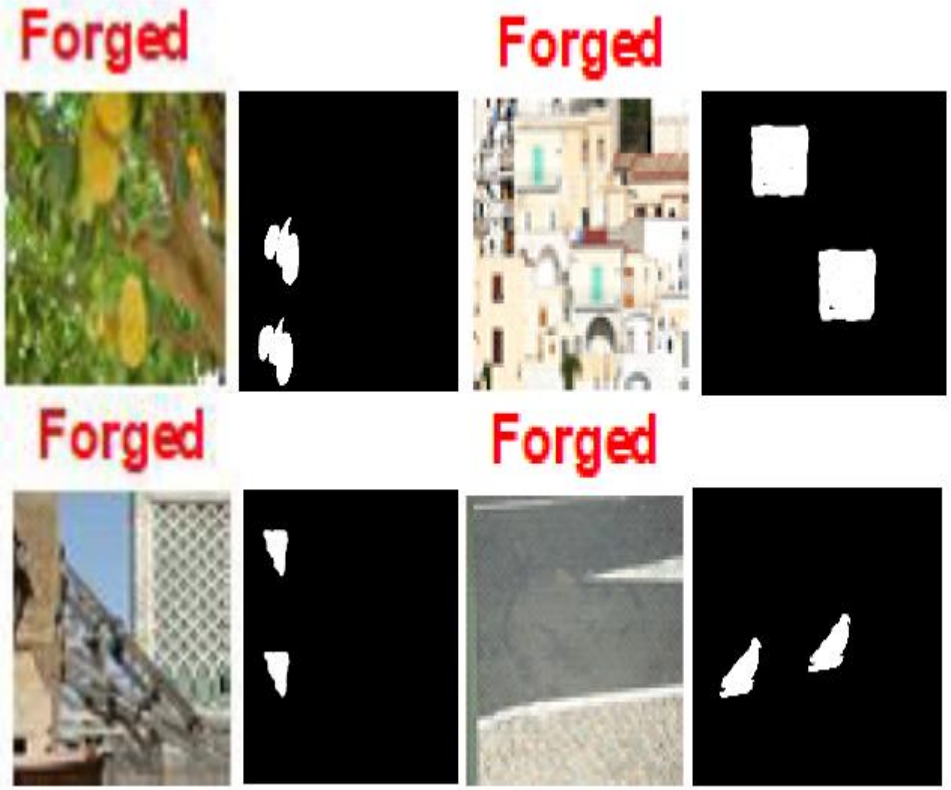
- The proposed CNN will learn how to detect similarities and differences in image features through several of hidden layers.
- Each individual hidden layer will enhance the CNN's learning feature ability in order to increase its detection accuracy.

Detect the Copy-move Forgery using CNNs. Experimental result:



This show some result of using CNN for active CMFD

Detect the Copy-move Forgery using CNNs. Experimental result:



Shows True passive forgery detection i.e. there is no original images (Pristine)

False detection by libeling this forged image as pristine image. The red color on the result refers to false result.

Detect the Copy-move Forgery using CNNs. Experimental result:

Epoch	Iteration	Time Elapsed (sec)	Mini-batch Loss	Validation Loss	Mini-batch Acc.	Validation Acc.	Base Learning Rate
1	1	7.04	0.6870	0.6825	74.00%	87.09%	0.0010
4	20	132.03	0.5228		87.89%		0.0010

TRAINING ON SINGLE CPU.

Epoch	Iteration	Time Elapsed (sec)	Mini-batch Loss	Validation Loss	Mini-batch Acc.	Validation Acc.	Base Learning Rate
1	1	2131.86	0.6895	0.6890	79.00%	84.68%	0.0010
3	30	66732.32	0.5650	0.5602	97.00%	96.10%	0.0010
5	50	112258.92	0.4752		95.00%		0.0010
5	50	123271.12	0.4535		95.00%		0.0010

TRAINING ON SINGLE CPU WITH EXTENDED NETWORK.

Detect the Copy-move Forgery using CNNs.

Conclusion

- On the other hand, a novel neural network-based copy-move forgery detection strategy has been proposed in this work.
- Our CNN learned how to reproduce both forged and pristine outputs in its training phase, enabling copied regions to trigger detection during reconstruction.
- The results for active copy-move detection were good, while the passive detection results were only satisfactory.
- Additionally, overall efficiency was relatively low due to the small size of the experimental dataset utilized in the training phase,
- thus, the proposed CNN algorithm scored very low error rate 0.41 and the accuracy is 96.10%.

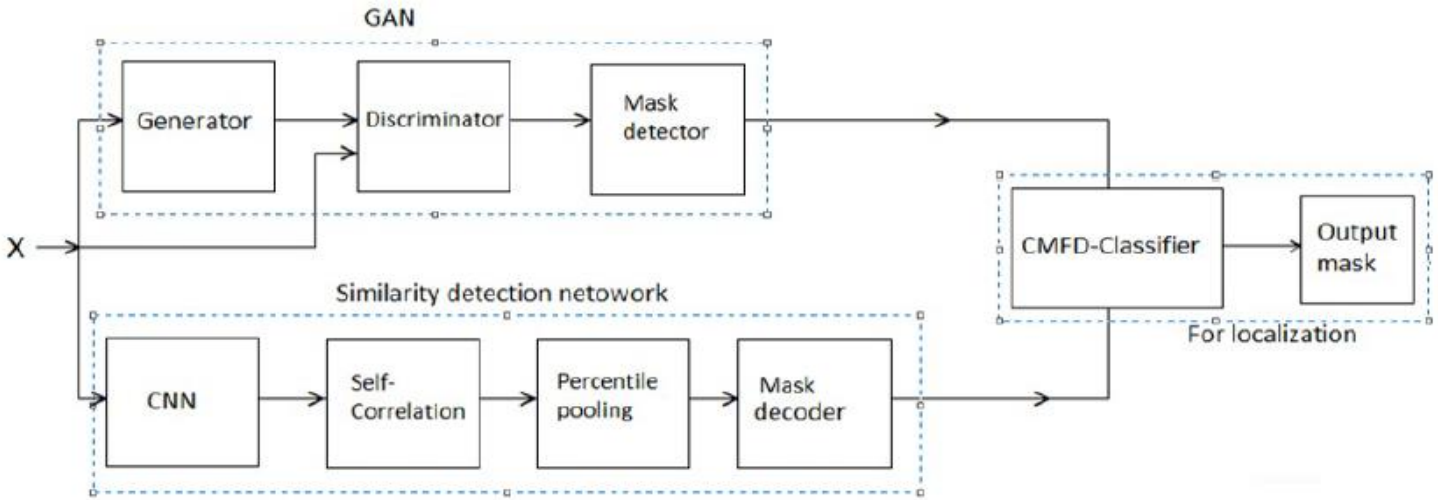
Third Method

Leveraged deep learning technique using CNNs & GAN

Leveraged CNN and GAN Model for copy-move forgery detection.

Literature review:

- The present model investigates copy-move forgery detection using a fusion processing model comprising a deep convolutional model and an adversarial model.
- These have been shown to be highly effective in dealing even with image forgery that derived from generative adversarial networks (GANs).
- The network is developed based on two-branch architecture and a fusion module.
- The two branches are used to localize and identify copy-move forgery regions through CNN and GAN.

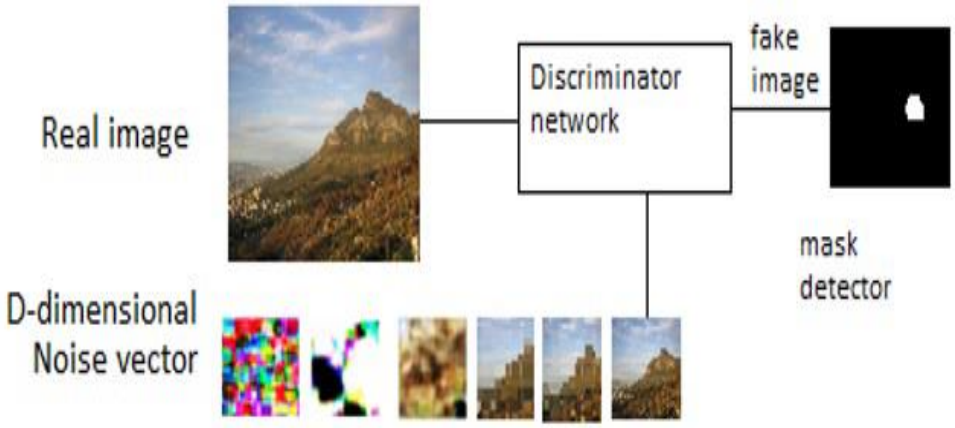


Leveraged CNN and GAN Model for copy-move forgery detection. Literature review:

- CNN have been shown to be highly effective in dealing even with image forgery that derived from generative adversarial networks (GANs).
- In general, GAN-based methods provide the most optimal results in image forgery detection
- The CMFD classifier output is a node which represents the concatenated the sum of the all relevant inputs.

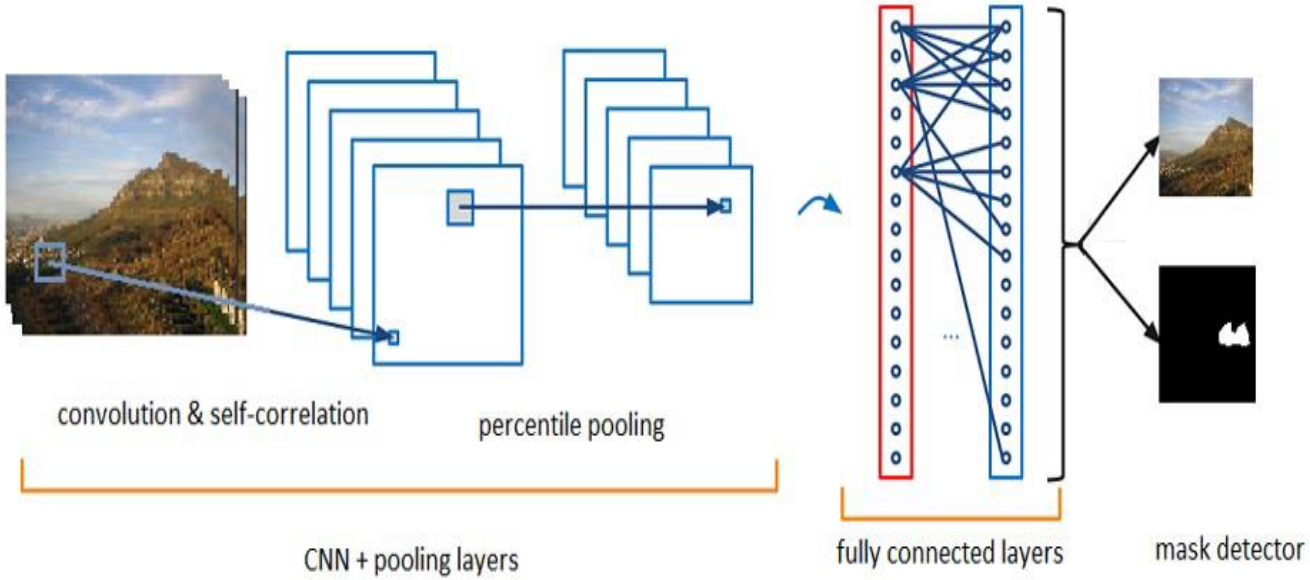
Leveraged CNN and GAN Model for copy-move forgery detection. The proposed model:

- GAN task
  - Building a discriminator network
  - Building a generator network
  - Generating a sample image
- GAN model



Leveraged CNN and GAN Model for copy-move forgery detection. The proposed model:

- CNN tasks
  - Feature extraction
  - Similarity detection
- CNN model

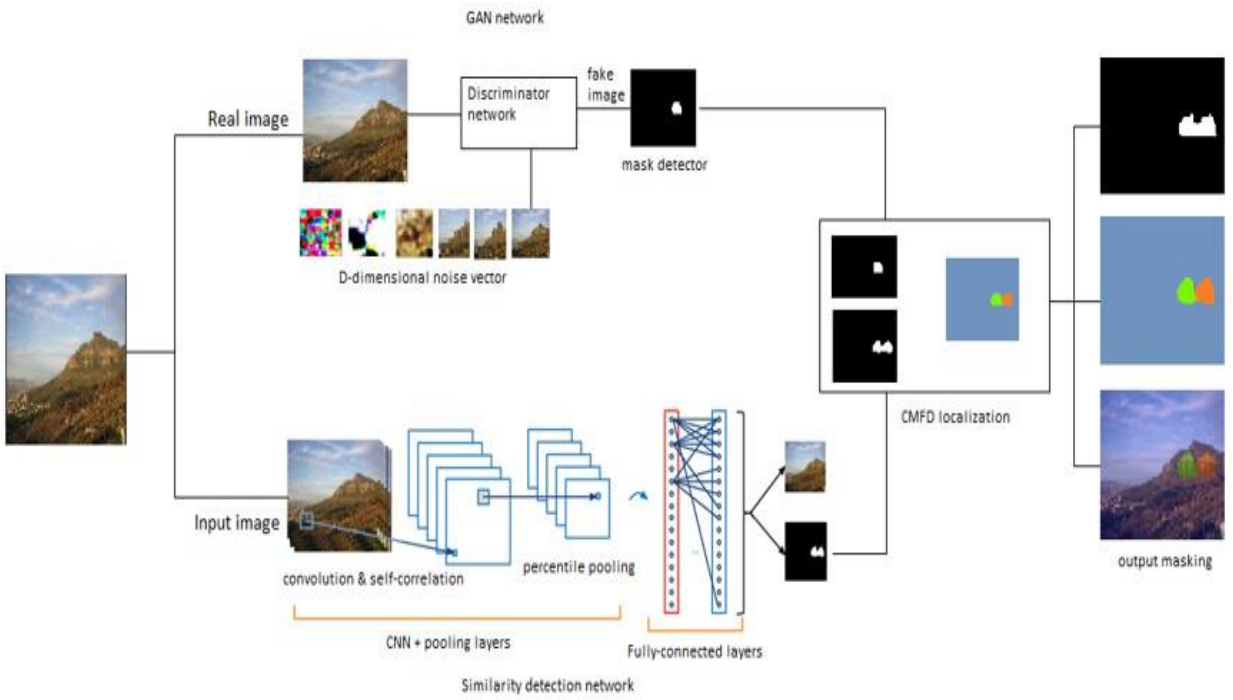


Leveraged CNN and GAN Model for copy-move forgery detection. The proposed model:

- Merging network tasks
  - render a final decision on the copy-move
  - make the copy-move transaction localized
  - make a distinction between the original source and the targeted regions of a potentially forged image

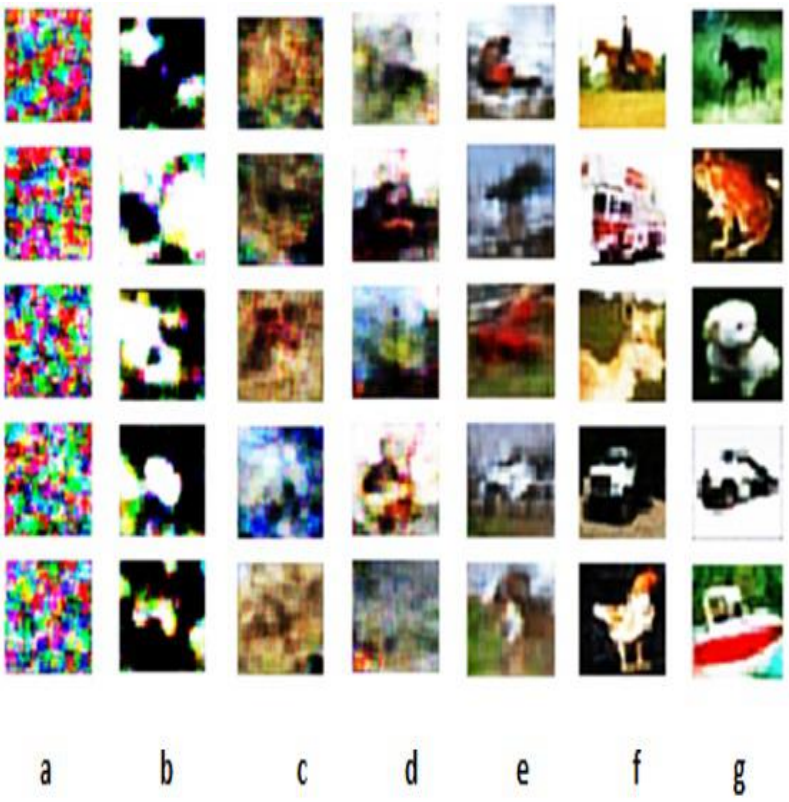
Leveraged CNN and GAN Model for copy-move forgery detection. The proposed model:

- Fully model architecture



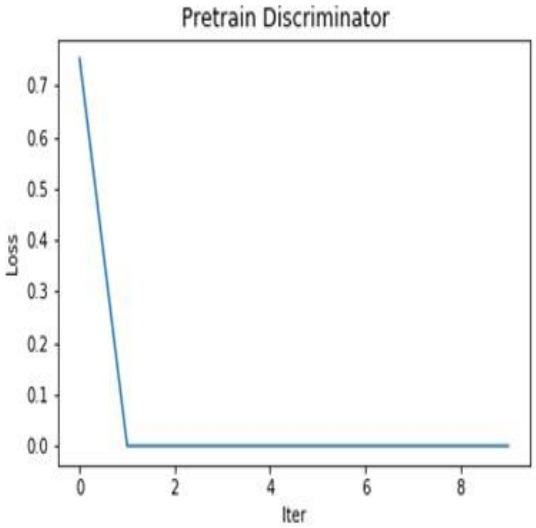
Leveraged CNN and GAN Model for copy-move forgery detection. Experimental results:

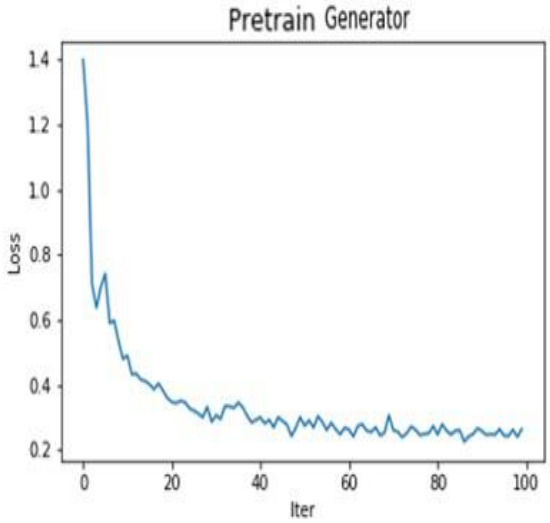
- GAN model training in forgery detection

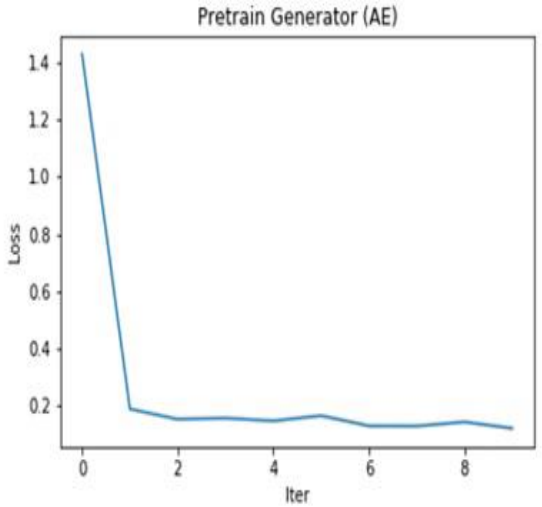


Leveraged CNN and GAN Model for copy-move forgery detection. Experimental results:

- Loss function

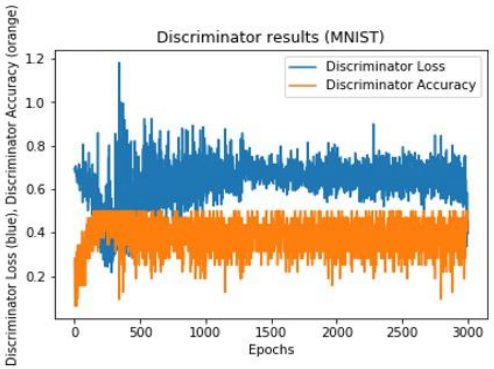


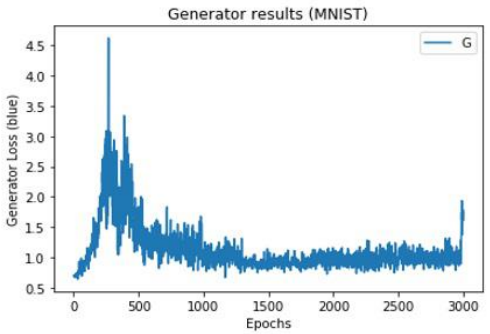




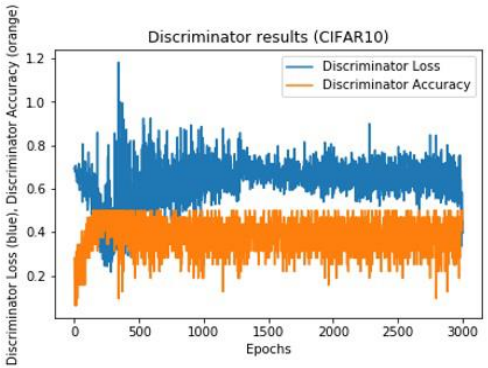
Leveraged CNN and GAN Model for copy-move forgery detection. Experimental results:

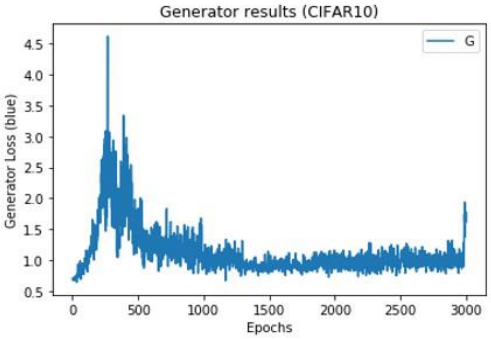
- MNIST dataset





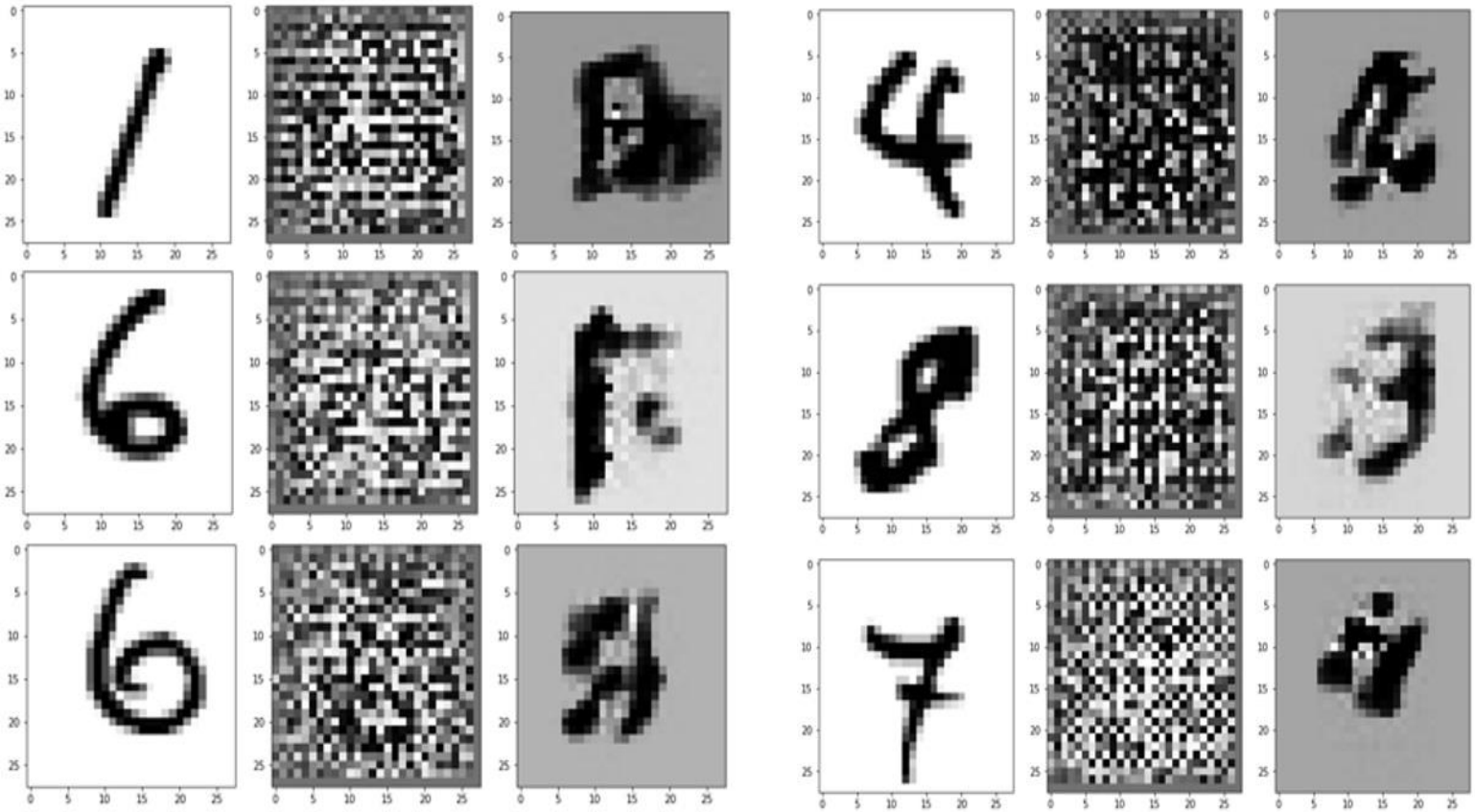
- CIFAR10 dataset





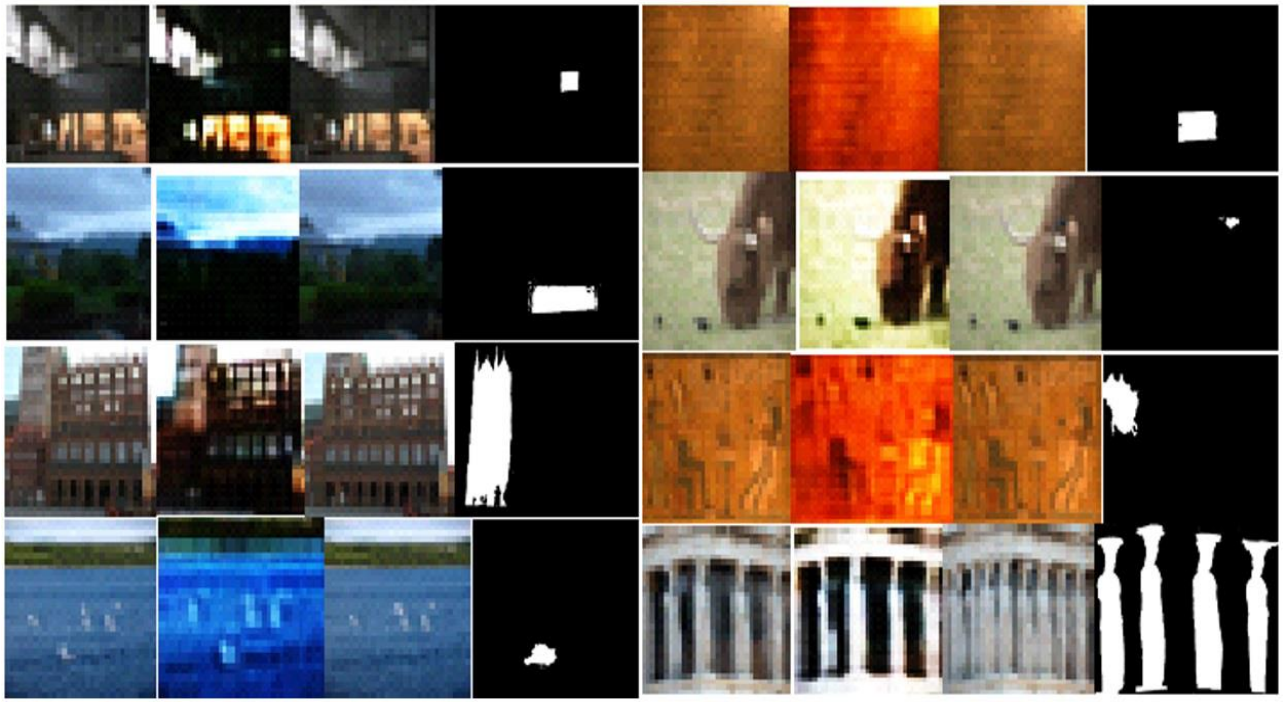
Leveraged CNN and GAN Model for copy-move forgery detection. Experimental results:

- GAN output training using MNIST dataset



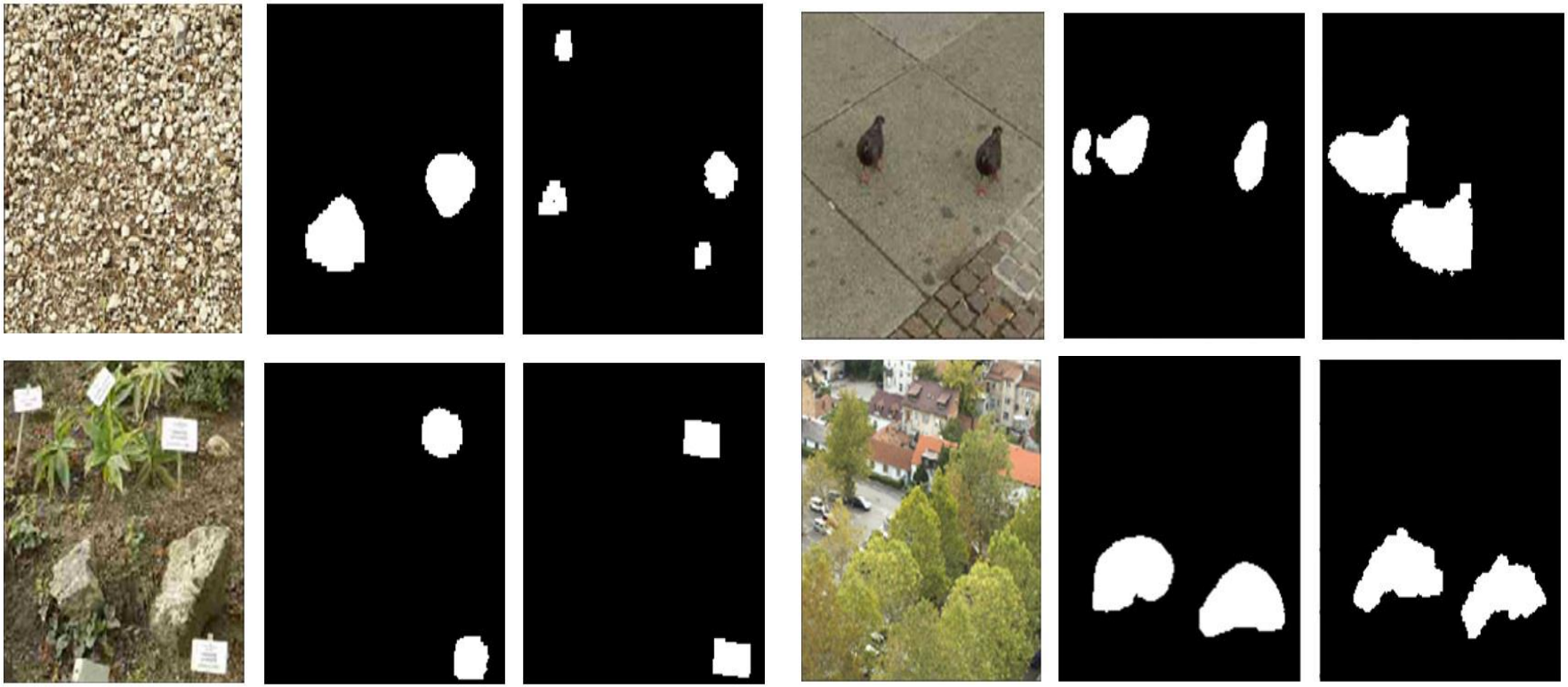
Leveraged CNN and GAN Model for copy-move forgery detection. Experimental results:

- Detection of forgery using GAN model:



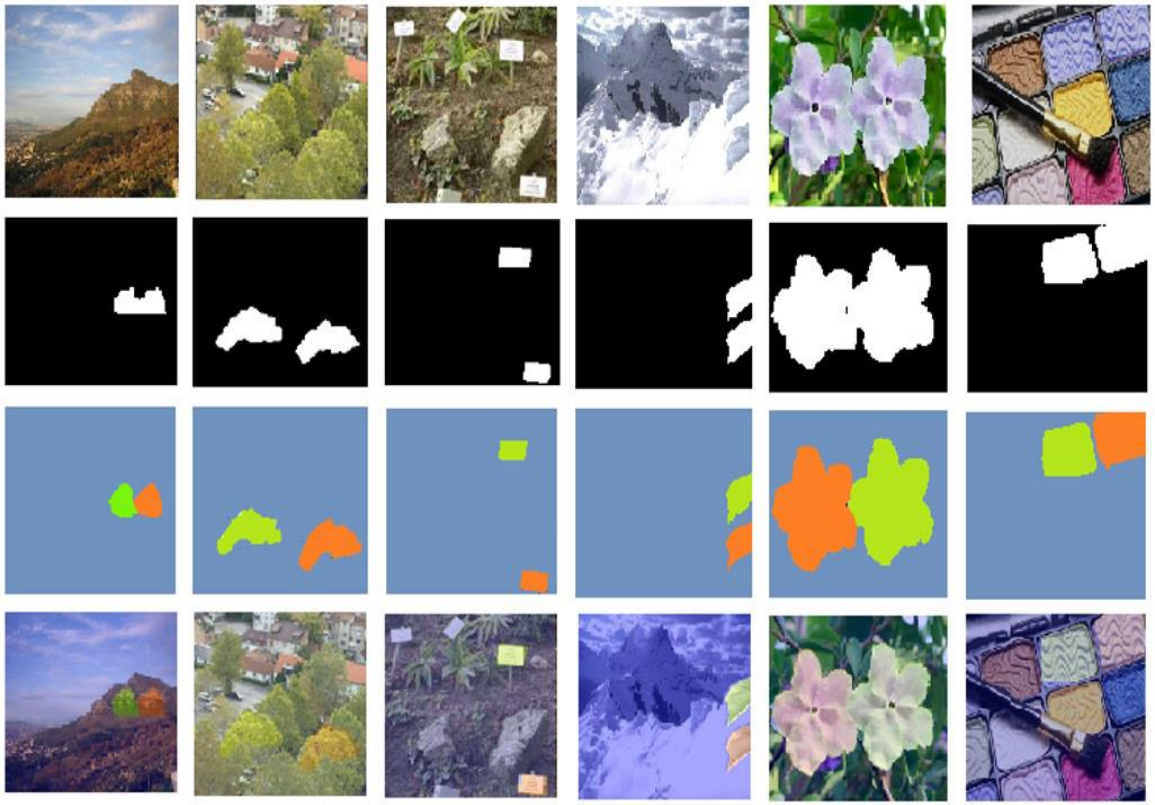
Leveraged CNN and GAN Model for copy-move forgery detection. Experimental results:

- CNN training model results:



Leveraged CNN and GAN Model for copy-move forgery detection. Experimental results:

- Fully model results:

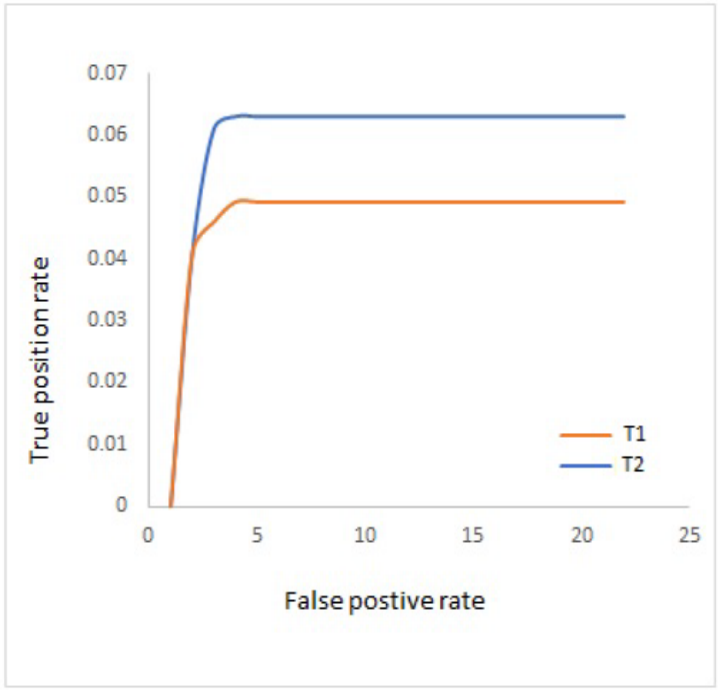


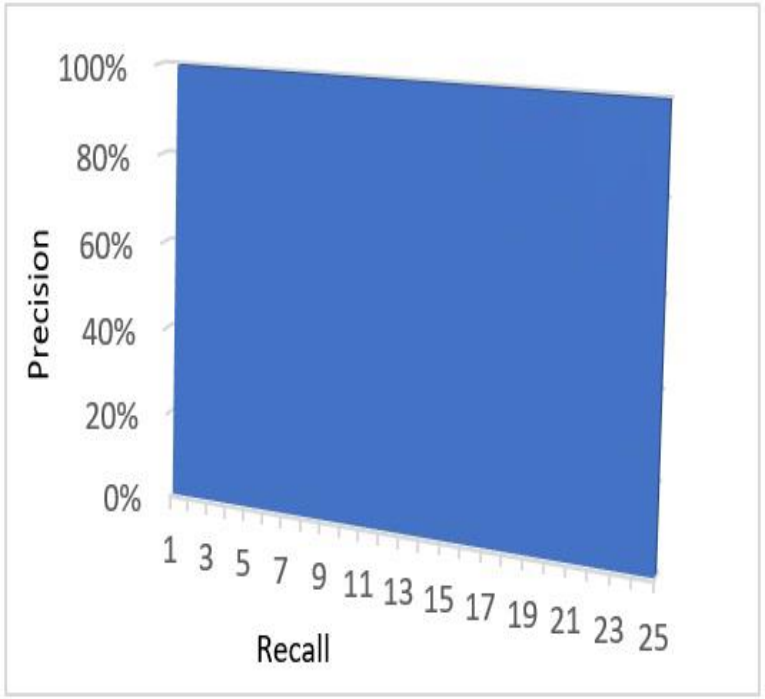
Leveraged CNN and GAN Model for copy-move forgery detection. Experimental results:

- Model evaluation:

Algorithm	[37]	[94]	[23]	[106]	[64]	Proposed
F1	0.4926	0.5439	0.5943	0.6055	0.6318	0.8835
Precision	0.5734	0.5390	0.5440	0.5662	0.5927	0.6963
Recall	0.4939	0.8327	0.8020	0.8040	0.8220	0.8042

Leveraged CNN and GAN Model for copy-move forgery detection. Experimental results:





Leveraged CNN and GAN Model for copy-move forgery detection. Conclusion

- The test results clearly indicate the significantly high accuracy of the proposed method in applying the dataset in the discernment of localization and tampering detection.
- The rationale for the proposed CMF solution in the present work is testing the potential for training auto-encoders to acquire a representation of image patches originating in pristine images.
- The test results indicate a high level of accuracy for localization as well as detections, though the system has no prior knowledge of forgeries.
- The model succeeded in its assigned CDFD task, giving performance results in the 93% to 97% range.

Forth Method

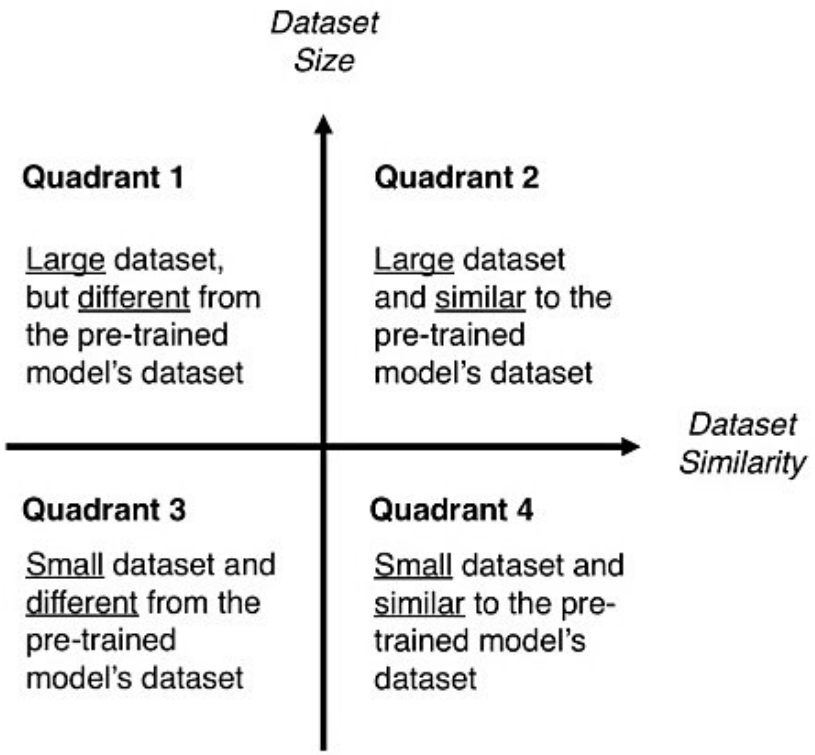
Image forgery detection based on deep transfer learning

Image Forgery Detection Based on Deep Transfer Learning. Literature review:

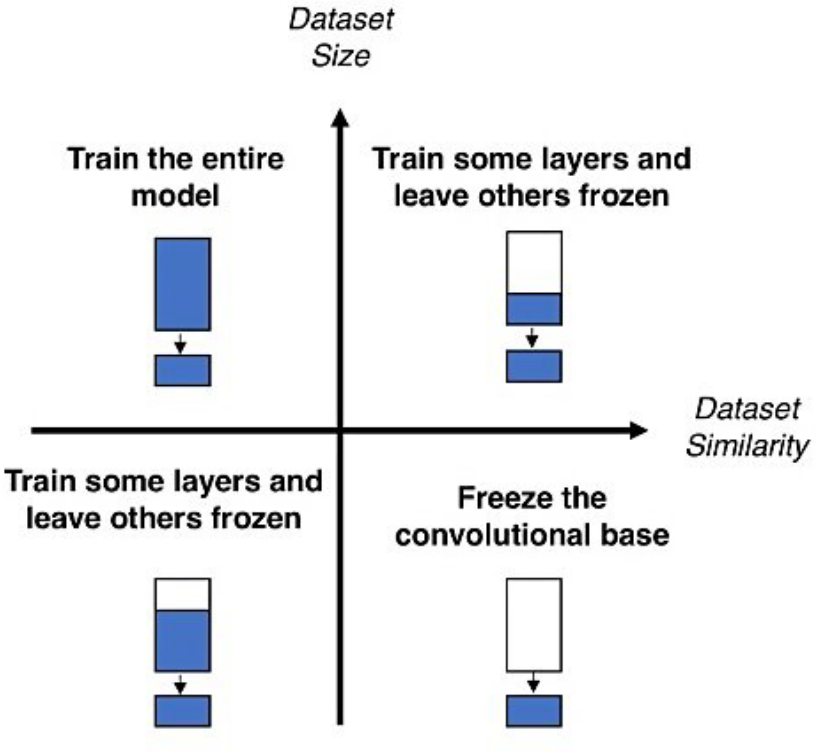
- This method applies prior knowledge that has been transferred to the new model from previous steganalysis models.
- Additionally, because CNN models generally perform badly when transferred to other databases, transfer learning accomplished through
- knowledge transfer allows the model to be easily trained for other databases.
- The various models are then evaluated using image forgery techniques such as shearing, rotating, and scaling images.

Image Forgery Detection Based on Deep Transfer Learning. Literature review:

Size similarity matrix for pretrained model's datasets



Decision map for fine tuning pretrained model



Im, D.J., Kim, C.D., Jiang, H., Memisevic, R.: Generative adversarial metric,

Image Forgery Detection Based on Deep Transfer Learning. The proposed approach:

- Transfer learning approaches:

We demonstrated in this work in the aforementioned strategies, as follows:

  - Instance transfer
  - Relational-knowledge transfer
  - Parameter transfer
  - Feature-representation transfer

Image Forgery Detection Based on Deep Transfer Learning.

The proposed model:

- The baseline training model and the transfer learning model are layered in the same architectures.

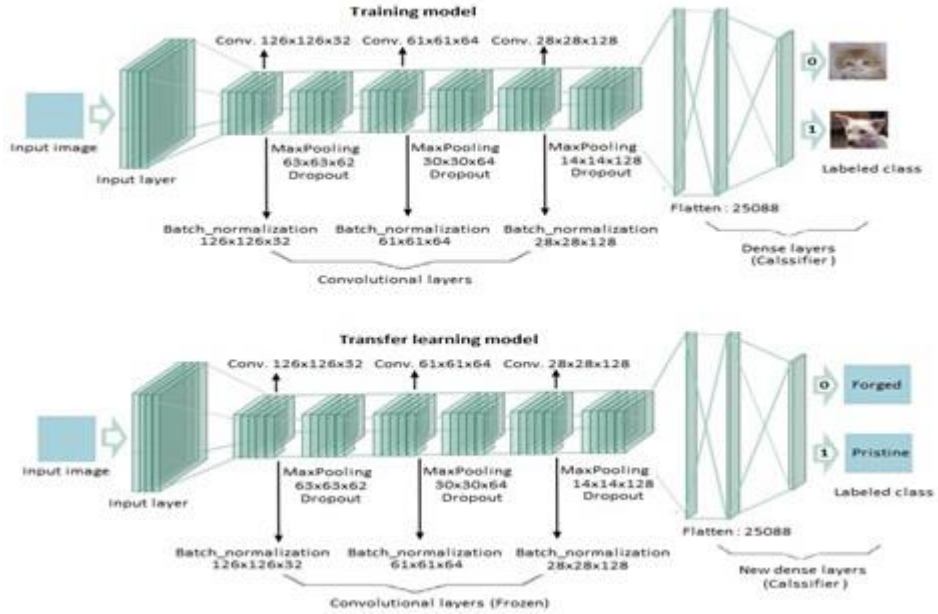


Image Forgery Detection Based on Deep Transfer Learning.
Result and discussion:

- The results of this new approach provided:

1. Training data generator with old dataset

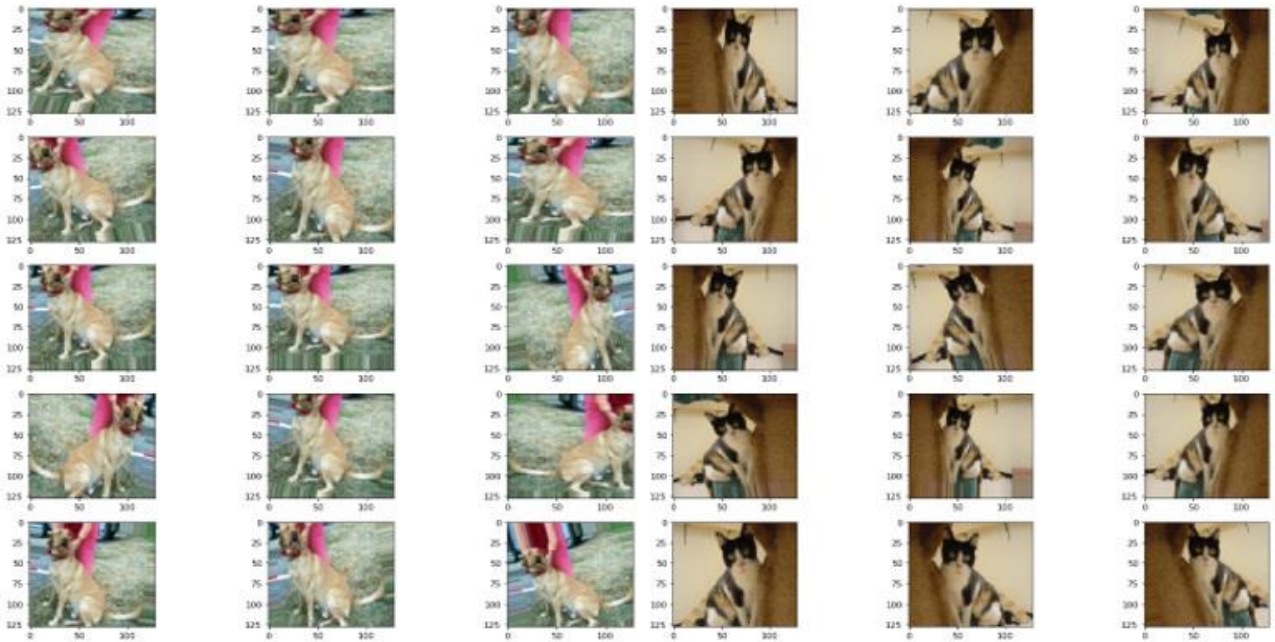
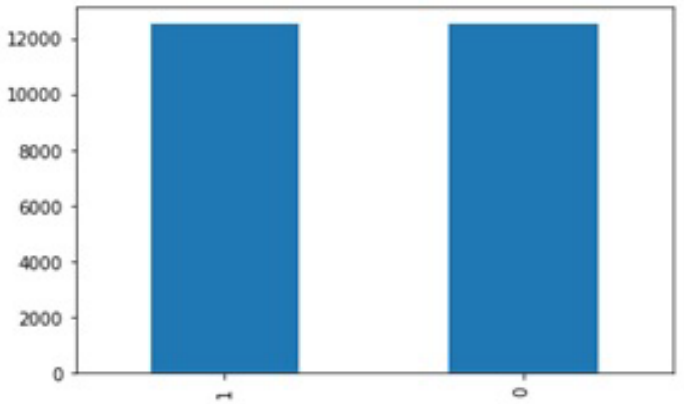


Image Forgery Detection Based on Deep Transfer Learning. Result and discussion:

2. The two categories that used in the training task presents the training result map to a dog is 1 and cat is 0



3. The validation result map to model using the original dataset: dog is 1 and cat is 0

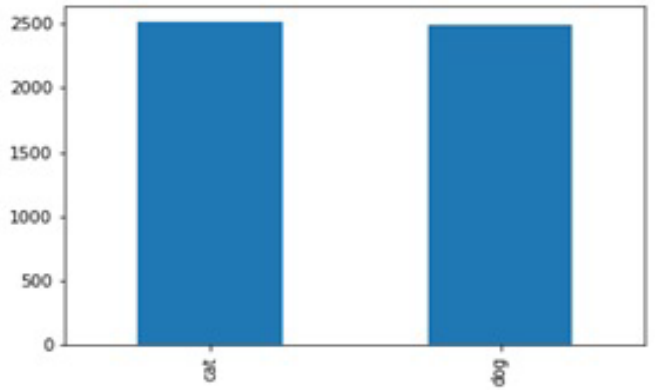


Image Forgery Detection Based on Deep Transfer Learning.

Result and discussion:

4. Create Testing Generator based on the original dataset

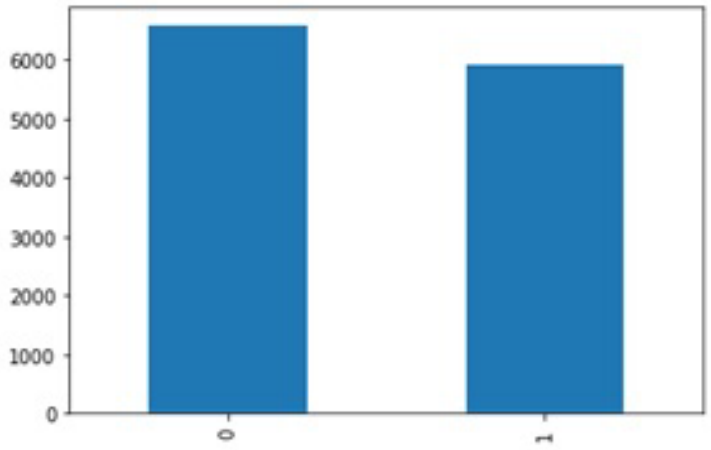
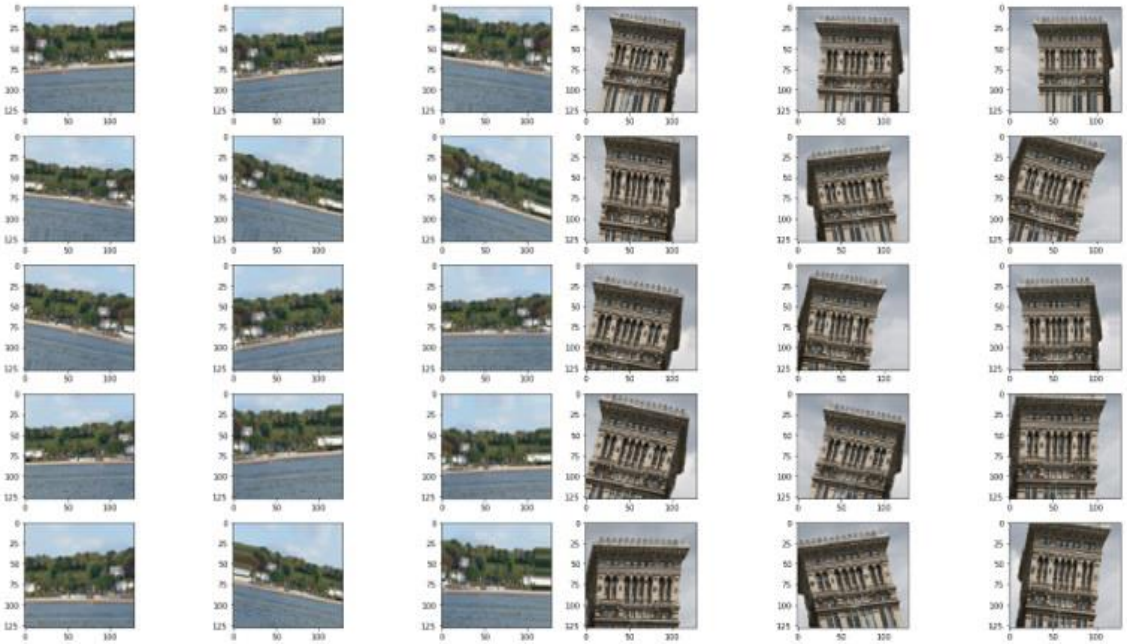


Image Forgery Detection Based on Deep Transfer Learning.

Result and discussion:

5. This figure shows examples of the data generator work using the new dataset



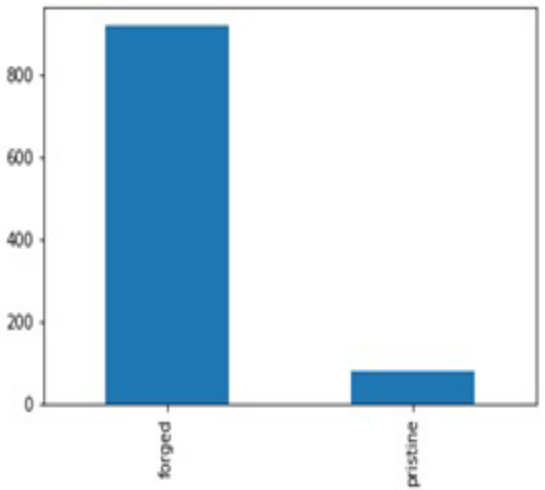
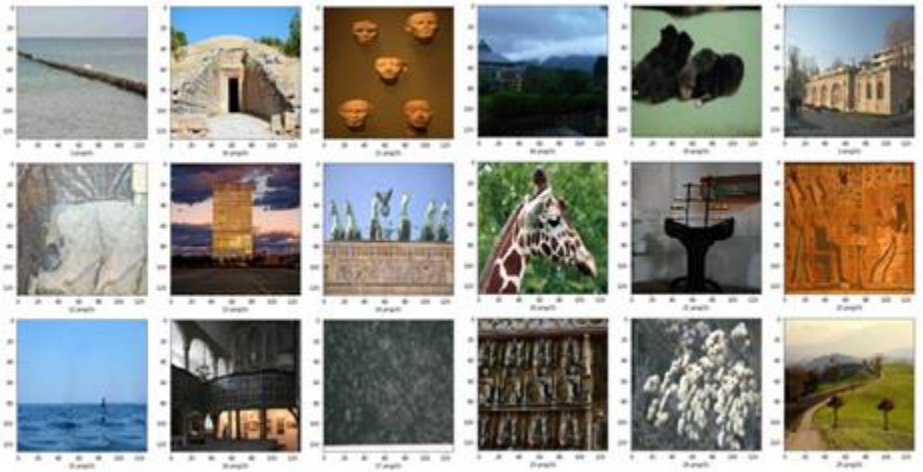


Image Forgery Detection Based on Deep Transfer Learning.

Result and discussion:

6. The predicted result with images using the new dataset



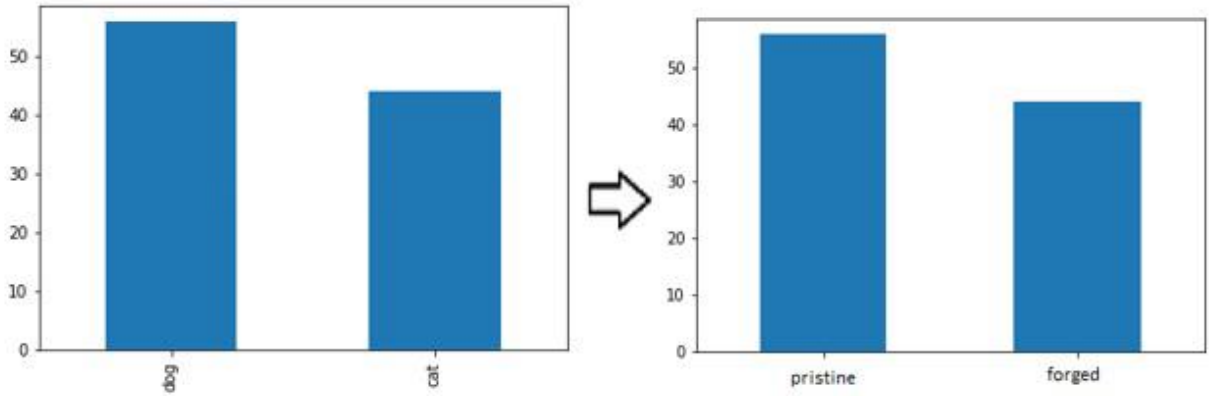


Image Forgery Detection Based on Deep Transfer Learning. Result and discussion:

7. Model evaluation

The loss and the validation accuracy of the both models

The model	F1-Score	Loss	Acc	Val-loss	Val-Acc.
Trained model	0.89	0.2683	88.93%	0.2122	91.29%
Transfer model	0.93	0.2583	92.94%	0.2347	94.89%

The evaluation based on the validation accuracy between two closer

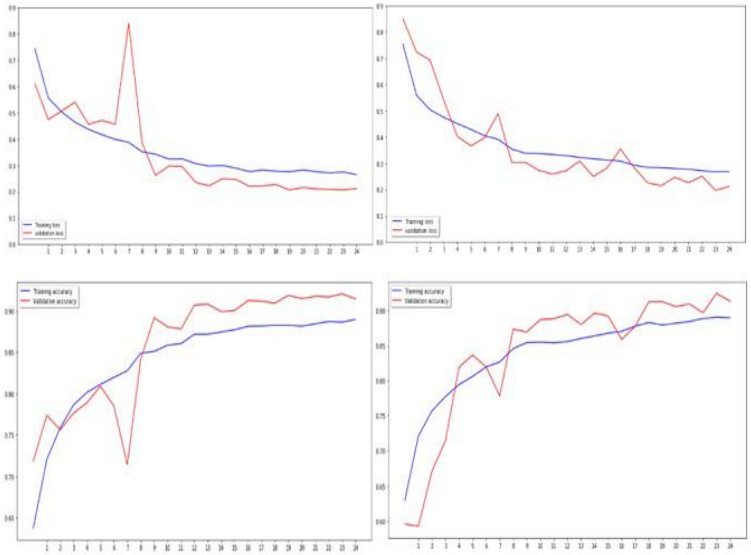
The Model	Training data	Validation accuracy
[2]	MSCOCO	77%
	ImageNet subset	93%
Presented model	Dogs -and- Cats	91.29%
	Pristine -and- Forged	94.89%

[2] R. V. G. B. Soares, " Inductive Transfer," In: Sammut C., Webb G.I.(eds) Encyclopedia

Image Forgery Detection Based on Deep Transfer Learning. Result and discussion:

Training trial

- Training loss vs. validation loss in different training trial
- Training accuracy vs. validation accuracy



New dataset

- Training loss vs. validation loss in different training trial
- Training accuracy vs. validation accuracy

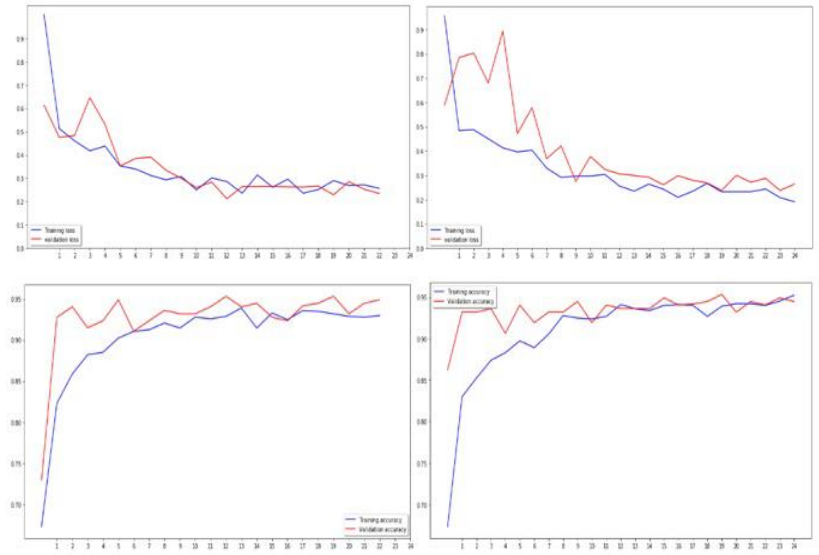


Image Forgery Detection Based on Deep Transfer Learning. Conclusion

The experimental results, which show an image manipulation detection has validation accuracy of over 94.89%, indicate that the proposed transfer learning approach successfully accelerates CNN model convergence but does not improve image quality.

Summery and future work

- PatchMatch algorithm exhibits optimal efficiency in the enhanced PatchMatch approach, and F-score generally showing higher FM versus the state-of-the-art.
- Using the convolutional neural network (CNN) architecture approach is promising and shows well enhancement of a copy-move forgery detection efficiency.
- Deep convolution learning algorithm compounds GAN with CNN was an optimal solution to distinguish between source and target patch in CMF images.
- In addressing CMF problem, the transfer learning model proposed a novel approach to training image forgery detection models.

Summery and future work

For future work:

- To introduce new trends in the research field of image forgery detection, we will look deeply at deep learning and neural networks to develop a new method that can improve the accuracy of copy-move forgery detection of digital images.
- We also need to look at improving multi-copy-move forgery detection (MCMFD) algorithms to reduce the lack of detection using single copy-move forgery detection algorithms.

Contribution

The main contribution of this work is to enhance the capability of detecting and localizing copy-move forgery, by:

- We proposed new algorithms to detect the copy-move forger based on classic technique (Enhanced PatchMach)
- Deep learning-based methods were proposed using CNN VGG16 to enhance the classic technique above. This model can solve the most existing copy-move forgery image databases.

Contribution

- The CNN & GAN model was an extensive version model of CNN and GAN models spritely for forgery detection which was leveraged to diagnosing the type of copy-move forgery as one unit to achieve more robustness.
- Finally, we used deep transfer learning for digital image forgery detection by using fewer resources to achieve higher efficiency forgery detection.
- Experimental results showed that the proposed methods have an impressive performance in the copy-move forgery detection.

Publications

- Conference papers:
  - Fall Detection – Vision-Based Indoor Environment, “IEEE. NECEC conference” (2014)
  - Using MSER Algorithm to Characterize an Active Camera Movement, “IEEE. NECEC conference” (2015)
  - Hand Gesture Detection and Pose Estimation Using Image Processing A Survey , “IEEE. NECEC conference ” (2016)
  - Black Ice detection using Kinect, “IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE)” (2017)
  - Copy-Move Forgery Detection Based on Enhanced Patch-Match, “IEEE. NECEC conference” (2017)
- Journal papers:
  - Copy-Move Forgery Detection Based on Enhanced Patch-Match, “International Journal of Computer Science Issues” (2017)
  - Fusion approach system of copy-move forgery detection, “American Journal of Computer Science and Engineering Survey” (2018)
  - Convolutional Neural Network for Copy-move Forgery Detection , Under revision, Submitted (2018)

Thank You

Q & A