UNIVERSITY OF MALAYA

Candidature Defence

Forged text detection method in Video, Natural Scene and Document Images

Lokesh Kishor Nandanwar

WMA190005 (17207738/1)

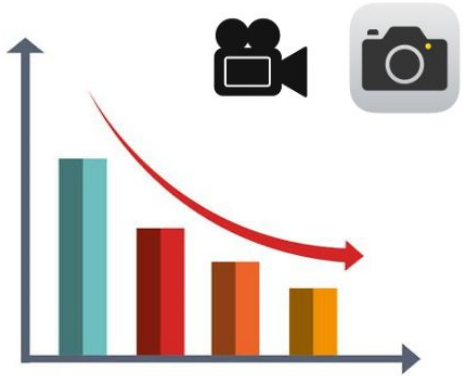
UNIVERSITY OF MALAYA

Content Overview

1. Introduction
2. Literature Review
3. Problem statement
4. Research Questions
5. Research Objectives
6. Scope of Study
7. Works Completed
8. Work ongoing
9. Significance of the Work
10. Summary of Study
11. Publications
12. Conclusion

UNIVERSITY OF MALAYA

Introduction



Rate of visual media **consumption** has increased the rate of **crimes and frauds**.

Low cost digital imaging devices available with advanced features along with **free* and cheap** digital storage services.

Easy to manipulate the visual media using latest and advanced low-cost **editing softwares**.

UNIVERSITY OF MALAYA

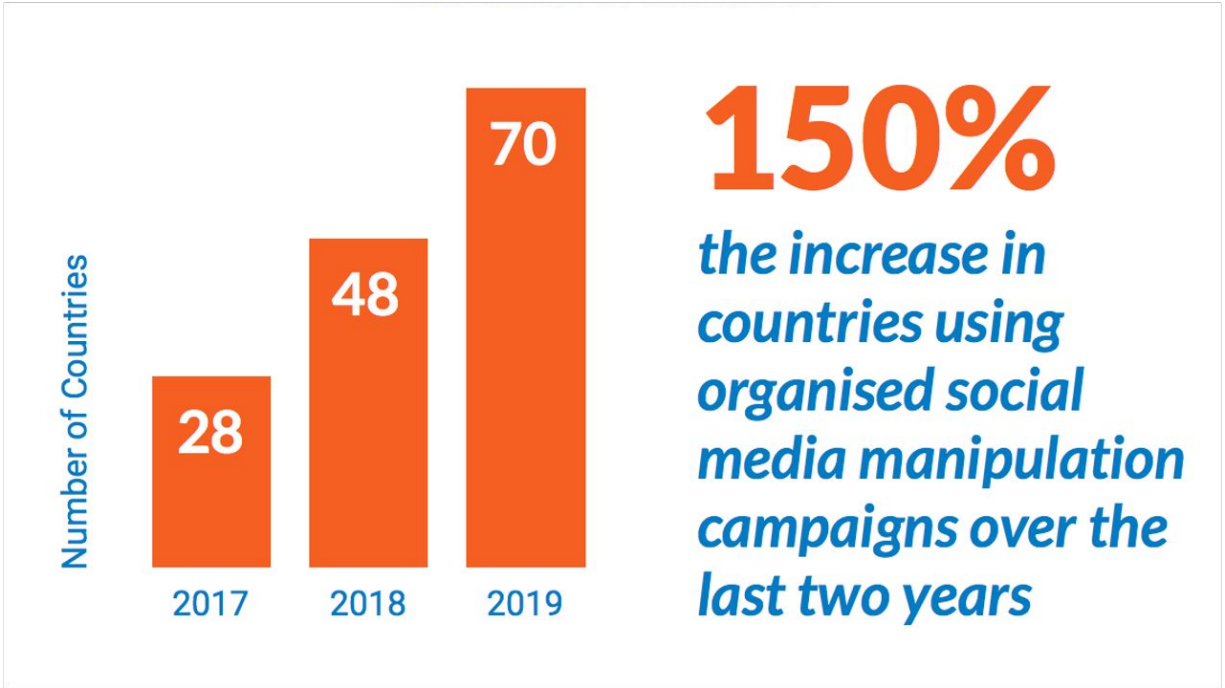
Introduction

- ➔ The credibility of digital multimedia content is no longer be taken for granted.
- ➔ Altering, tampering and forging content is a serious threat for forensic applications such as:
  1. Forging property, insurance, certificates, banking documents.
  2. Creating fake suicide notes and fake answer scripts.
  3. Image manipulation in medicine, justice, news reporting and accounting professions, etc

UNIVERSITY OF MALAYA

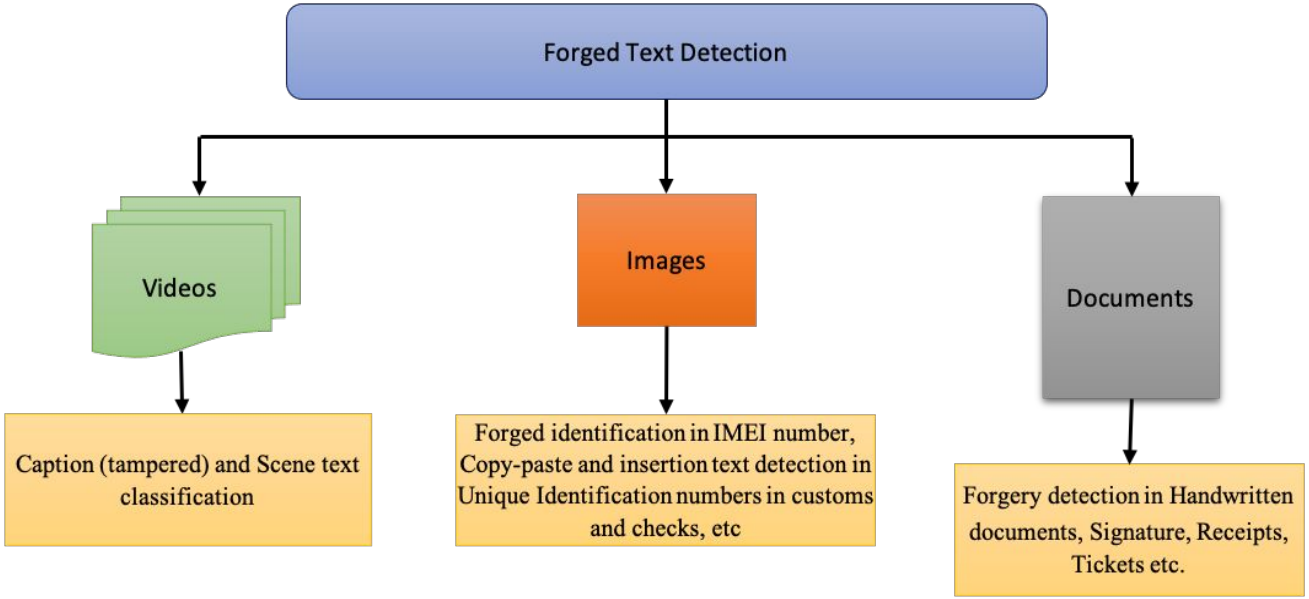
Introduction

Hers's how social media misinformation/crimes/frauds increased in the world



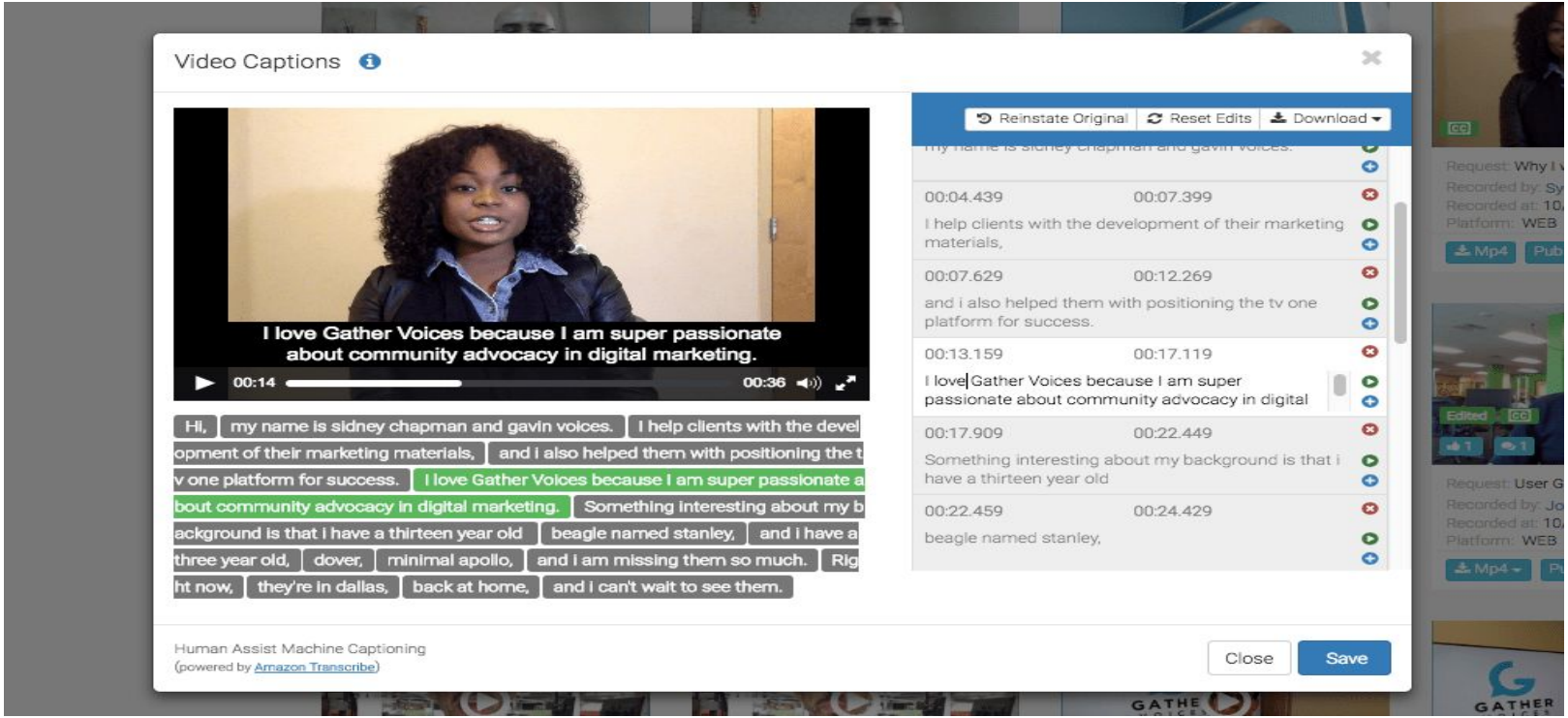
UNIVERSITY OF MALAYA

Our Focus : Forged Text Detection



UNIVERSITY OF MALAYA

Applications: Video annotation or video understanding at semantic level



UNIVERSITY OF MALAYA

Forged text detection in Videos

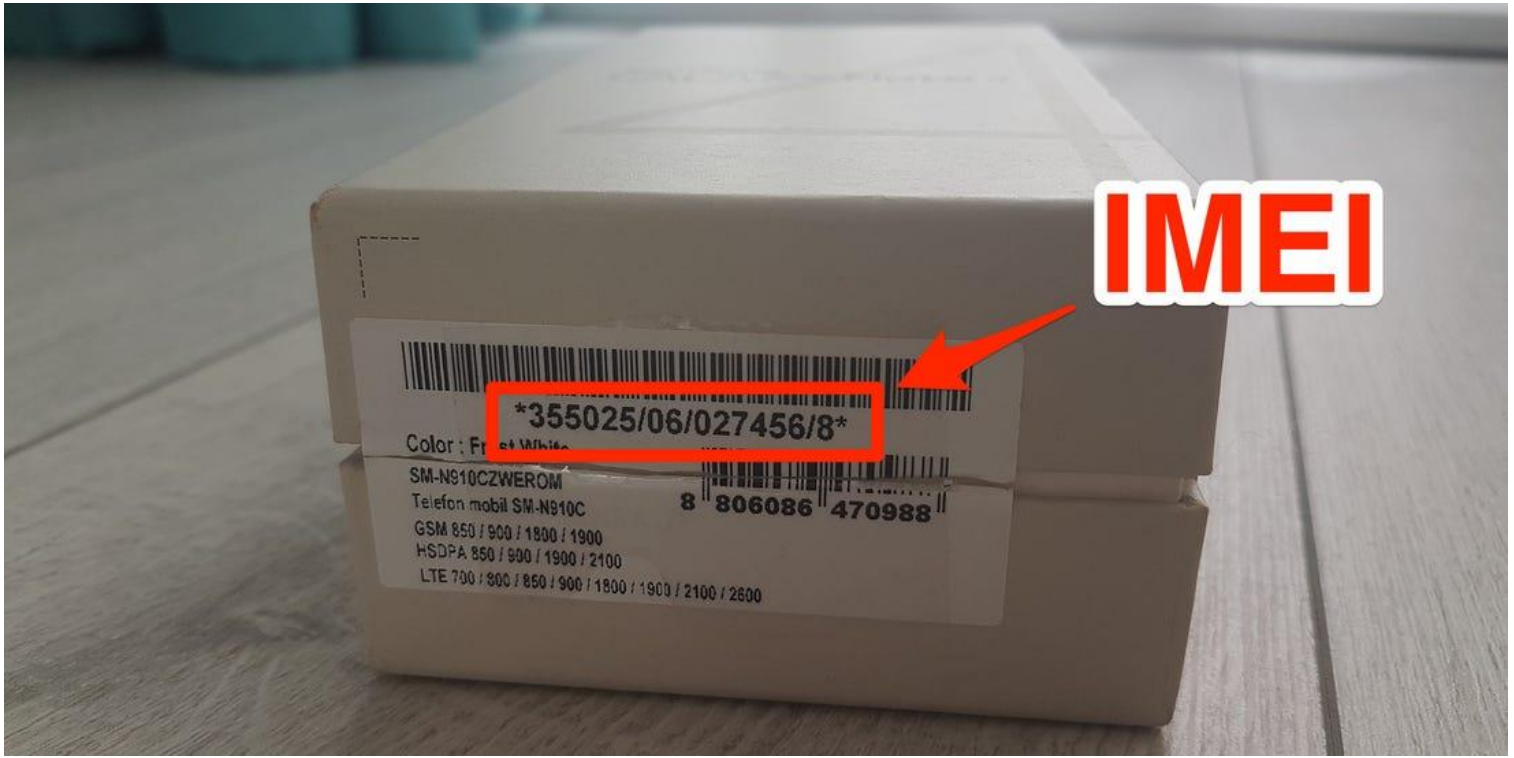
(a) Scene text image image

(b) Caption (tampered) text

- ➔ Caption text is used to **annotate** at semantic level.
- ➔ Scene text is used to **understand the content**.
- ➔ Caption text is **inserted text (Tampered)**.
- ➔ Applications in News reporting, teleshopping, Cooking shows, Defence discoveries, Social Media etc.

UNIVERSITY OF MALAYA

Applications: IMEI number forgery detection to Avoid smuggling and illegal selling, second handle mobile selling



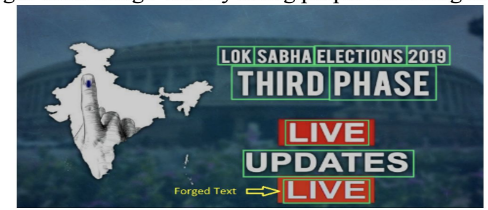
UNIVERSITY OF MALAYA

Forged text detection in still images

(a) Original and forged IMEI numbers created using copy-paste operations are marked by a green and red color, respectively.

Source image Target image

(b) The word "MirrorNow" in the source image is copied and pasted on Target image. Text detection method segment the words in source and target video image well by fixing proper bounding boxes.



Target image

(c) The word "LIVE" is inserted using Photoshop software for creating forged text in target image. Text detection method segment original and forged text in the natural scene image by fixing proper bounding boxes.

- ➔ **Forensics** and Forgery identification.
- ➔ Copy-paste and insertion operation are used.
- ➔ Hard to notice the difference between the original images and the forged ones.
- ➔ Detect smart phones for **stealing and smuggling** them illegally.
- ➔ Second handle mobile selling

UNIVERSITY OF MALAYA

Applications: Avert breaching of airport security (Air ticket forgery).

Fake e-ticket cases at airports in 2018 highest in four years; agencies mull alternatives

PTI | Dec 30, 2018, 05.19 PM IST

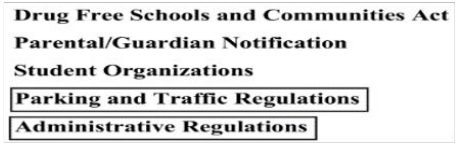
NEW DELHI: Incidents of fake e-tickets usage to gain illegal entry into Indian airports were highest in 2018 in past four years, prompting security agencies to moot biometric or barcode-based access system for passengers.

While security officials ruled out any terror-like or extreme sabotage threat in these recorded incidents, they expressed concern over the potential of this menace being misused in future to breach the airport security.

As per a CISF data accessed by PTI, a total of 140 incidents (about 26 per cent more) of illegal entries of passengers using fake or cancelled e-tickets were intercepted till early December as compared to 111 such incidents registered last year. The comparative figures for 2016 were 74 while for 2015 it was 43.

UNIVERSITY OF MALAYA

Forged text detection in documents



Copy-paste operation

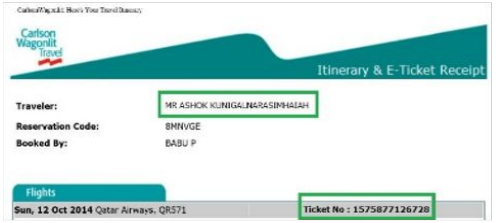
Insertion (imitation) operation

(a) Illustration of sample forged PDF document images by copy-paste and insertion operations. Note: altered texts are enclosed by bounding boxes, which appear to be genuine text in terms of font, color and size.

Original handwritten word

Forged handwritten word

(b) Illustration of sample forged Handwritten document images by insertion of characters. These can also be seen also be evident in case of forged signatures in documents.



(c) Original and forged Air-ticket traveler name created using copy-paste operation are marked by green and red color respectively.

- **Avert breaching** of airport security (Air ticket forgery).
- Reduce crimes in **Forged property documents** for ill intentions.
- **Fake suicide note** detection in crimes.
- Detect fake certificates.

UNIVERSITY OF MALAYA

Literature Review:

The methods are classified in three broad categories:

1. Forged text detection in Videos
2. Forged text detection in still images
3. Altered text detection in documents

UNIVERSITY OF MALAYA

1. Forged text detection in Videos

Method	Objective	Concept	Drawbacks	Multimedia formats supported
Shivakumara et al. 2014	Separation of graphics and scene text in video	• Works based on the fact that caption text has high contrast and clarity, while scene text does not.	• Not robust to features as it is based on contrast and clarity. • Not adequate for text level forgery. • Not effective in case of documents	• Videos • Images
Xu et al 2014	Graphics and scene text classification in video.	• based on contrast and clarity • Extracting distinct features through distribution of Eigen values.	• Not robust to features as it is based on contrast and clarity. • Not adequate for text level forgery. • Not suitable for documents	• Videos • Images
Roy et al. 2016	Tampered features for scene and caption text classification in video frames	• DCT coefficients to differentiate caption text from scene text.	• Not effective for Documents images • Poor performance for complex images	• Videos • Images
Bhardwaj and Pankjakshan 2016	Image overlay text detection based on JEPG truncation error analysis.	• Extracts tampered features through truncation errors given by a color filter array for detecting caption text in video	• Not adequate for forged text caused by copy-paste and insertion operations • Poor Performance for Documents images.	• Videos • Images
Chen et al. 2016	Automatic detection of object-based forgery	• Frame Manipulation Detector and Forgery Identification	• Focus on visual content and not text in videos frames • Not robust for forged text caused by copy-paste and insertion operations	• Video
Feng et al 2017	Digital video forensic	• motion adaptive frame deletion detection		• Video
Amiano et al. 2018	Video copy-move detection and localization	• patch match based dense field algorithm		• Video • Images
Fadi et al 2019	Inter-frame forgery detection	• Use of spatio-temporal information		• Video
Ghosh et. al 2019	Presence of graphical text in scene images	• Based on CNN • Edited text and text in natural scene images as a graphical text for classification	• Method does not consider caption and scene text in video images • Not robust as Graphical text can also be present as caption or scene text.	• Video • Images

UNIVERSITY OF MALAYA

Implementation of Existing Methods

Caption text (Forged/Edited) | Scene text(Original)

(a) Roy et. al [18] (2016) classifies Caption text as scene text and vice-versa due of weak feature extraction as distortion is not noticeable

Caption text detected as Scene text | Scene text detected as Caption text

(b) Ghosh et. al [20] (2019) based on CNN failed to detect the forgery classifies caption text as scene text and scene text as caption text due to blurriness and consistency in shape respectively.

Caption text detected as Scene text | Scene text detected as Caption text

(c) Fadi et. al [15] (2019) based on spatio-temporal information failed to detect the original and tampered text, classifies caption text as scene text due to shadow in text and scene text as caption text due to no character shapes and less distortions

UNIVERSITY OF MALAYA

2. Forged text detection in Still Images

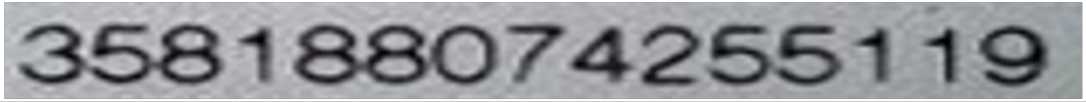
Method	Objective	Concept	Drawbacks	Multimedia formats supported
Pun et al. 2015	Image forgery detection based on matchings	• adaptive over-segmentation and features point matching.	• Based on visual features and not the text information • Not robust at pixel level forgery.	• Images
Yang et al. 2017	Copy-move forgery detection	• Based on hybrid features. • An improved matching algorithm		• Images
Shivakumara et al. 2018	Detecting forged IMEI numbers based on color space and a fusion approach	• The variance of each color space (RGB) is used to obtain a fused image for each input image. • Features based on connected components are extracted from Canny and Sobel edge images of input and fused images for forged IMEI number detection	• If a forged image does not contain sufficient distortion, to be observed in Canny or sobel, the method doesn't work well. • Sensitive to complex background. • Template based classification	• Images • Documents
Kundu et al. 2019	Fourier spectrum for classifying forged handwriting text from original, blurred and noised handwriting text images.	• Extract feature from the Fourier spectrum and the features fed to neural network classifier for classification.	• Performance degrades for character level forgery • Not suitable for small forged operations.	• Images • Documents

UNIVERSITY OF MALAYA

Implementation of Existing Methods

Copy-paste forgery at character level, "3" at 6th position from left is copy-pasted

(a) Yang et al. [22], (2017) Hybrid method failed to detect the forgery by copy-paste operation due to minute distortions at pixel level in IMEI images



Original Image

(b) Shivakumara et al. [25], (2018) failed to detect the original image because of noise in the image and classified it as forged in IMEI images.

Original forged

(c) Shivakumara et al. [25], (2018) detects blurred original text as forged and classifies forged image as original due to very less deformation in text image.

Original forged

(c) Kundu et al. [26], (2019) detects both as original due to consistency in shape and uniform background even though it has unnoticeable distortions due to insertion operations.

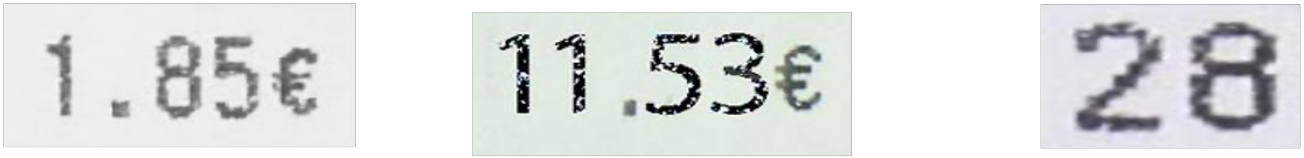
UNIVERSITY OF MALAYA

3. Forged text detection in documents

Method	Objective	Concept	Drawbacks	Multimedia formats supported
Halder and Garain 2010	Color features based approach for determining ink age in printed documents.	• Color features for printed text images. • Use of ink quality based features	• Not suitable to image forgery detection, only for age estimation • Not robust for documents with noise/blur or degradations.	• Handwritten Documents
Barboza et al 2013	Color based model to determine document ages for forensic purposes	• Uses ink quality of handwritten document images captured at different intervals of time. • Identifies a given image as old or new	• Poor performance for printed documents text. • Not robust to pixel level forgery detection.	• Handwritten Documents
Khan et al. 2015	Automatic ink mismatch detection	• Analyses the ink of different pens to find fraudulent documents • Effective for handwritten documents	• Ink features not robust for printed documents.	• Handwritten Documents
Luo et al. 2015	Localized forgery detection in hyperspectral document images	• Explores ink quality in the hyperspectral domain for fraud document identification.	• Not effective on printed texts since when digitized, the quality of handwritten document ink changes are very low.	• Handwritten Documents
Raghunandan et al. 2016	Fourier coefficients for Identifying fraud handwriting documents	• Fourier coefficients for studying the quality of handwriting documents. • Quality-based features • If a document suffers from poor quality, it is considered as an original one else a fraud one.	• Quality-based features not robust for documents affected by adverse factors, such as distortions, noises, blur, and forgery operations. • The method does not work at the text line or word levels and requires the full document	• Handwritten Documents
Wang et al. 2017	Fourier-residual for printer identification from document images.	• Extracts features from residuals given by the Fourier transform for printer identification.	• The primary goal of this method is to identify printers rather than forged/tampered document images.	• Printed Documents
Khan et al. 2018	Automated forgery detection in multispectral document images	• Method explores ink matching based on fuzzy k-means clustering • Partition the spectral responses of ink pixels in handwritten notes into different clusters	• Not suitable for printed text where we can see low changes in ink when digitized.	• Handwritten Documents
Berenguel et al. 2019	Detecting counterfeit documents	• Based on a deep learning model • Expect some abrupt changes in the background texture of the document	• Not suitable for documents with plain background	• Printed Documents

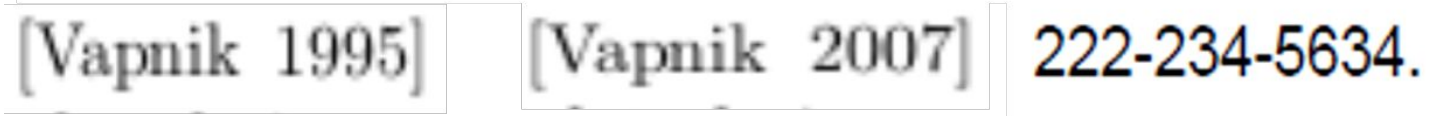
UNIVERSITY OF MALAYA

Implementation of Existing Methods



Original | Insertion forgery | Copy-paste forgery

(a) Wang et al. [35], (2017) fails to detect the forgery of characters in words and classifies both Price Receipts images as original.



Original | Copy-paste forgery | Insertion forgery

(b) Berenguel et al. [28], (2019), based on deep learning not able to detect the forgery in document images at word level when distortion is not noticeable.

Original | blurred | forged

(c) Kundu et al. [26], (2019) fails to detect the character level forgery classifies forge image as original and detects blurred original text as forged in air ticket images.

UNIVERSITY OF MALAYA

Summary of Review:

1. Methods are **good when there are clear differences** between forged and genuine text.
2. **Fails at a minute difference** at the pixel level or character levels.
3. Most use images/documents that **do not suffer from degradations, noises, blur, poor quality, and ageing** for forgery detection.
4. Scope limited to **single multimedia** type only.

List of Challenges in Forged Text detection

Videos	Works well when the images are forged at word level but not at character level.
	Not robust to clutter background and degraded text
	Fails when the color and texture of the images varies arbitrarily
Still Images	The methods are not robust to the images affected by different resolution, contrast and blur.
	When the text contains irregular shape/sized characters, the performance of the method degrades.
	Sensitive to causes affected by perspective distortion
Documents	The success of the method depends on text or foreground information
	May not work well for the images of clutter background
	Sensitive to degradations and ageing

UNIVERSITY OF MALAYA

Problem Statement

Development of a **unified robust system** that can detect forged text in **video, still images and document images** in images affected by **distortion, noise, blur** environment.

UNIVERSITY OF MALAYA

Proposed Datasets

1. **Caption-Scene text classification Dataset (Video):** We detect caption and scene text images from the Concert, Recipe, Craft, Teleshopping and Yoga action class images, which consists of **2814 text images** including caption and scene texts.
2. **Forged Video text dataset (Video):** Our dataset includes 171 forged images by copy-paste operation, 215 images by insertion operation and 386 images are original images, which totally gives **772 images** for experimentation.
3. **IMEI forgery dataset (Still images):** Contains of 500 forged and 500 original text from IMEI images. Total of **1000 images**.
4. **Air tickets dataset (Documents):** Contains of 500 forged and 500 original text images with total **1000 images**.
5. **Forged Documents dataset (Documents):** Contains 110 altered text line images and the same number of original text line images, which gives a total of **220 text line images** for experimentation.

UNIVERSITY OF MALAYA

Standard Datasets

1. **Roy et. al dataset (Video):** This dataset consists of **900 caption and 650 scene texts** from video images, which gives a total of 1550 text images for experimentation.
2. **Bharadwaj et. al Dataset (Video):** Consists of two sets of images of different resolutions, namely, set-1 that contains 1280 × 720 pixels and set-2 that contains 1920 × 1080 pixels images. Set-1 provide 2233 and Set-2 provide 2415, which gives total 4648 images for experimentation.
3. **ICPR FDC 2018 Dataset (Still images):** Dataset from **ICPR 2018 Fraud Detection Contest (FDC)**. Provides 301 altered samples extracted from the ground truth given in the dataset and we select 527 original (unaltered) samples, which gives a total of 828 images for experimentation.
4. **Altered Handwriting dataset (Documents):** It includes four classes of **Original, Altered, Blurred and Noisy text** with tampering done at the word level. Each class contains 200 images, it gives a total of 800 images for experimentation.

Evaluation Metrics

1. **Confusion Matrix:** Also known as an **error matrix**, is a specific table layout that allows **visualization of the performance** of an algorithm in Statistical Classification.
2. **Average Classification Rate:** Classification rate is defined as the number of images classified correctly by the proposed method divided by the actual number of images. The average classification rate is defined as the **mean of diagonal elements** of the confusion matrix.
3. **Recall, Precision, F-measure:** **Precision** quantifies the number of positive class predictions that actually belong to the positive class. **Recall** quantifies the number of positive class predictions made out of all positive examples in the dataset. F-Measure provides a single **score** that balances both the concerns of **precision** and **recall** in one number.

UNIVERSITY OF MALAYA

Research Objectives

1. To develop a new method for forged text detection in video images through classification of tampered text and natural scene text.
2. To propose a new method for forged text detection in natural scene images by exploring Fourier spectrum analysis.
3. To explore a new method for detecting altered text in the document images based on fusion and reconstruction of the images.
4. To design and develop a unified method for detecting forgery in video, natural scene and document images.

UNIVERSITY OF MALAYA

Research Questions

1. What is the way to investigate a method for forged detection through classification of tampered text and scene text in videos?
2. How to explore and employ Fourier spectrum analysis to detect the forgery in natural images?
3. By what means one can approach fusion and reconstruction methods to spot the changes caused by forgery in document images?
4. How to develop a unified method for detecting forgery which can adapt to multiple multimedia formats such as video, natural scene and document images?

UNIVERSITY OF MALAYA

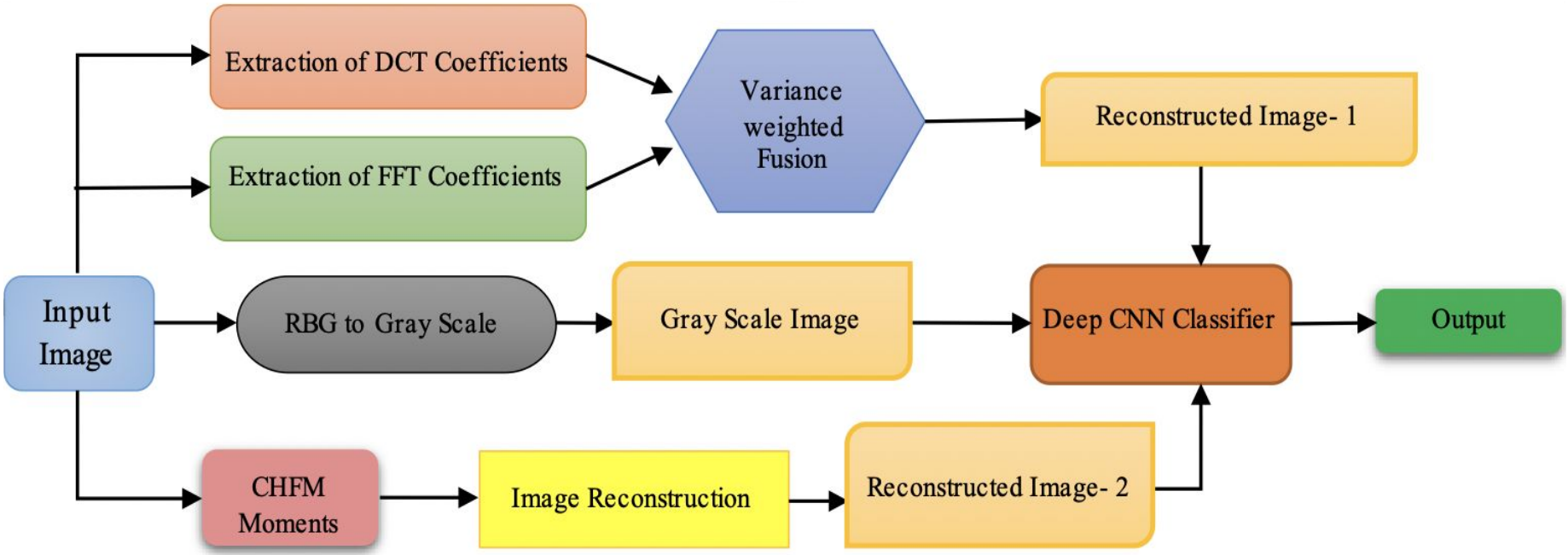
Research Objective-1

RO1: To develop a new method for forged text detection in video images through classification of tampered text and natural scene text.

RQ1: What is the way to investigate a method for forged detection through classification of tampered text and scene text in videos?

UNIVERSITY OF MALAYA

Methodology: Caption and Scene text classification in Video



UNIVERSITY OF MALAYA

(a) DCT and FFT Coefficients for Reconstructed Image-1

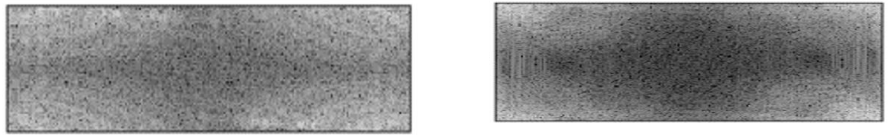
1. $$Var(x,y) = \frac{1}{M \times M} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [abs(p(i,j) - \mu)]^2$$

2. $$\omega_{I_{FFT}} = \frac{Var_{I_{FFT}}}{Var_{I_{FFT}} + Var_{I_{DCT}}}$$

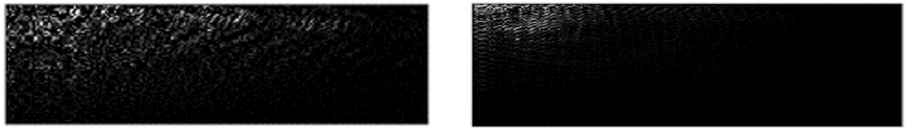
3. $$\omega_{I_{DCT}} = \frac{Var_{I_{DCT}}}{Var_{I_{FFT}} + Var_{I_{DCT}}}$$

4. $$F_{coeff} = I_{FFT}\ o\ \omega_{I_{FFT}} + I_{DCT}\ o\ \omega_{I_{DCT}}$$

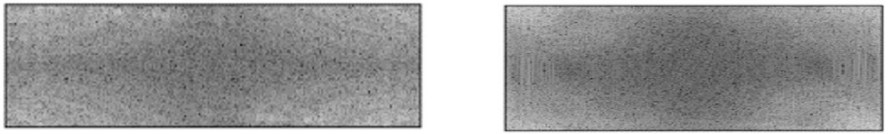
(a) Input: caption and scene text images



(b) Fourier coefficient of caption and scene text



(c) DCT coefficient images of caption and scene



(d) Fusion of DCT and Fourier coefficient

UNIVERSITY OF MALAYA

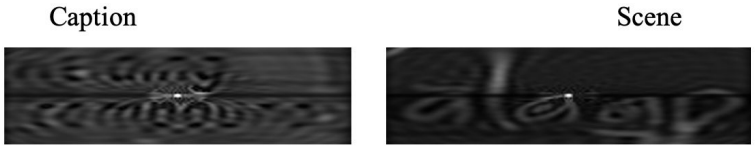
(b) Chebyshev-Harmonic-Fourier-Moments (CHFM) for Image Reconstruction-2

1. CHFM Basis Function :

$$\hat{f}(x_i, y_k) = \sum_{n=0}^{n_{max}} \sum_{m=-m_{max}}^{m_{max}} M_{nm} K(r, \theta)$$

2. Image Reconstruction from CHFM :

$$M_{nm} = \frac{1}{2\pi} \int_0^{2\pi} \int_0^1 f(r, \theta) K_{nm}^*(r, \theta) r dr d\theta$$



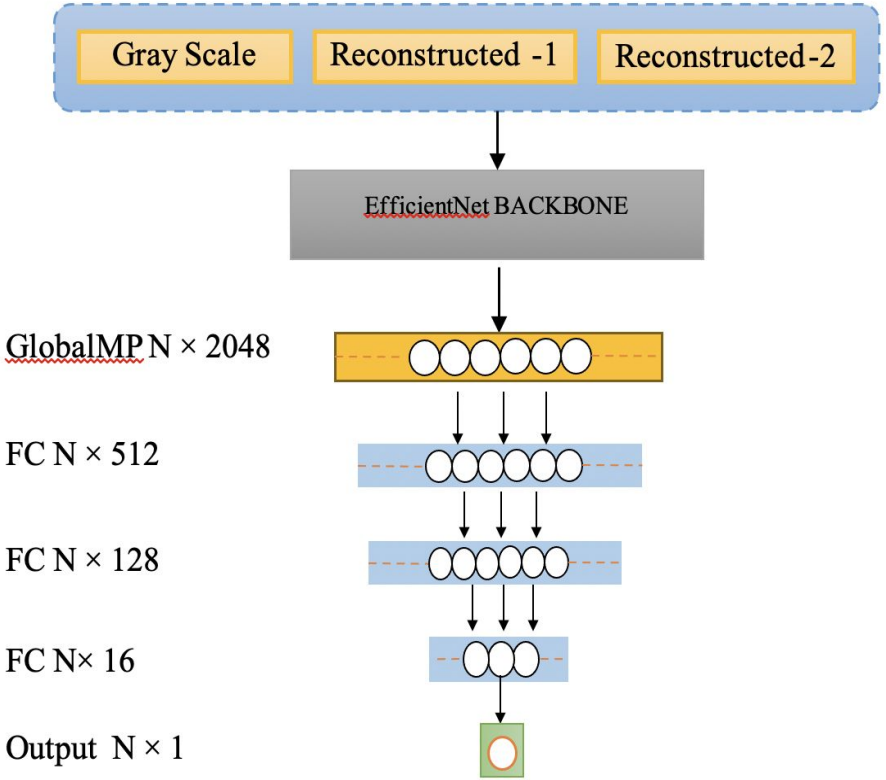
(a) Reconstructed Images 2 (CHFM Moments Image) for Caption and Scene Text images.

(b) Canny edge images of Reconstructed Image-1 by DCT-FFT fusion for caption and scene text

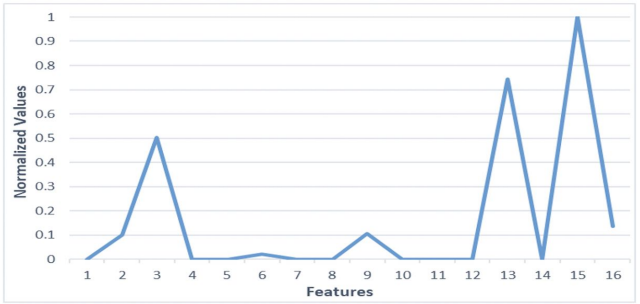
(c) Canny edge images of Reconstructed Images- 2 by CHFM for caption and scene text.

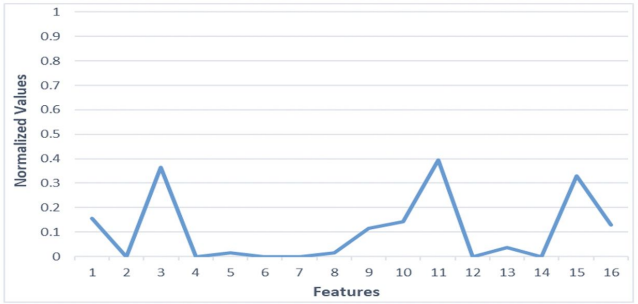
UNIVERSITY OF MALAYA

(c) Deep CNN for Classification of Caption and Scene Text



Final layer vector for Caption and Scene Images





UNIVERSITY OF MALAYA

Experimental Results

Dataset Name	Types	Input Image		Input + Reconstructed image-1		Input+ Reconstructed image-2		Proposed	
		Scene	Caption	Scene	Caption	Scene	Caption	Scene	Caption
Our Dataset	Scene	76.6	23.4	94.5	5.5	89.1	10.9	92.15	7.85
	Caption	20.6	79.4	14.4	85.6	9.6	90.4	5.5	94.5
	ACR	78.0		90.05		89.75		**93.32**	
Roy et. al Dataset	Scene	82.5	17.5	93.3	6.7	91.3	8.7	94.7	5.3
	Caption	11.9	88.1	8.9	91.1	5.2	94.8	3.8	96.2
	ACR	85.3		92.2		93.5		**95.45**	

Ablation Study of the Proposed method

Methods	Actions	Recipe		Concert		Crafts		Teleshopping		Yoga	
	Classes	Scene	Caption	Scene	Caption	Scene	Caption	Scene	Caption	Scene	Caption
Proposed	Scene	94.8	5.2	90.6	9.4	99	1	92.2	7.8	-	-
	Caption	6.4	93.6	6.1	93.9	0	100	1.5	98.5	1.0	99.0
	ACR	**94.2**		**92.25**		**99.5**		**95.35**		**99.0**	
Ghosh et. al (CNN)	Scene	60.15	39.85	61.61	38.39	64.01	35.99	61.03	38.97	-	-
	Caption	18.7	81.3	8.1	91.9	7.3	92.7	13.4	86.6	6.98	93.02
	ACR	70.7		76.7		78.3		73.8		93.0	
Roy et al.	Scene	64.2	35.7	62.4	37.5	66.6	33.3	63.8	36.1	-	-
	Caption	32.5	67.4	35.2	64.7	33.5	66.4	34.5	65.4	33.2	66.7
	ACR	65.8		63.55		66.5		64.6		66.7	

Comparison with latest existing methods on our dataset

UNIVERSITY OF MALAYA

Related Publication

ICPRAI 2020

Springer Link

International Conference on Pattern Recognition and Artificial Intelligence

ICPRAI 2020: Pattern Recognition and Artificial Intelligence pp 80-92 | Cite as

A New DCT-FFT Fusion Based Method for Caption and Scene Text Classification in Action Video Images

Authors | Authors and affiliations

Lokesh Nandanwar, Palaiahnakote Shivakumara, Suvojit Manna, Umapada Pal, Tong Lu, Michael Blumenstein

Conference paper

First Online: 09 October 2020

106 Downloads

Part of the Lecture Notes in Computer Science book series (LNCS, volume 12068)

Abstract

Achieving better recognition rate for text in video action images is challenging due to multi-type texts with unpredictable backgrounds. We propose a new method for the classification of captions (which is edited text) and scene texts (which is part of an image in video images of Yoga, Concert, Teleshopping, Craft, and Recipe classes). The proposed method introduces a new fusion criterion-based on DCT and Fourier coefficients to extract features that represent good clarity and visibility of captions to separate them from scene texts. The variances for coefficients of corresponding pixels of DCT and Fourier images are computed to derive the respective weights. The weights and coefficients are further used to generate a fused image. Furthermore, the proposed method estimates sparsity in Canny edge image of each fused image to derive rules for classifying caption and scene texts. Lastly, the proposed method is evaluated on images of five above-mentioned action image classes to validate the derived rules. Comparative studies with the state-of-the-art methods on the standard databases show that the proposed method outperforms the existing methods in terms of classification. The recognition experiments before and after classification show that the recognition performance rate improves significantly after classification.

Keywords

Caption text | Scene text | Fusion | Caption and scene text classification | Action image recognition

10th DAS 2020

Springer Link

International Workshop on Document Analysis Systems

DAS 2020: Document Analysis Systems pp 512-528 | Cite as

A New Common Points Detection Method for Classification of 2D and 3D Texts in Video/Scene Images

Authors | Authors and affiliations

Lokesh Nandanwar, Palaiahnakote Shivakumara, Ahlad Kumar, Tong Lu, Umapada Pal, Daniel Lopresti

Conference paper

First Online: 14 August 2020

109 Downloads

Part of the Lecture Notes in Computer Science book series (LNCS, volume 12116)

Abstract

Achieving high quality recognition result for video and natural scene images that contain both standard 2D text as well as decorative 3D text is challenging. Methods developed for 2D text may fail for 3D text due to the presence of pixels representing shadow and depth in the 3D text. This work aims at classification of 2D and 3D texts in video or scene images such that one can choose an appropriate method in the classified text for achieving better results. The proposed method explores Generalized Gradient Vector Flow (GGVF) for finding dominant points for input 2D and 3D text line images based on opposite direction symmetry. For each dominant point, our approach finds distance between neighbor points and plots a histogram to choose points which contribute to the highest peak as candidates. Distance symmetry between a candidate point and its neighbor points is checked and if a candidate point is visited twice, a common point is created. Statistical features such as the mean and standard deviation of the common points and candidate points are extracted to feed to Neural Network (NN) for classification. Experimental results on dataset of 2D-3D text line images and the dataset collected from standard natural scene images show that the proposed method outperforms exiting methods. Furthermore, recognition experiments before and after classification show recognition performance improves significantly as a result of applying our method.

Keywords

Gradient Vector Flow | Edge points | Candidate points | 2D text | 3D text | Text recognition | Video/scene images

25th ICPR 2020

Gmail — Lokesh Nandanwar <lokeshnandanwar150@gmail.com>

Decision on ICPR 2020 submission 475

The PaperCept Conference Management System <iapr@papercept.net> Sun, Oct 11, 2020 at 9:28 PM
Reply-To: Nicu Sebe <sebe@disi.unitn.it>
To: Shivakumara Palaiahnakote <shiva@comp.nus.edu.sg>, Shivakumara Palaiahnakote <hudempsk@yahoo.com>
Cc: lokeshnandanwar150@gmail.com, raghavendra.ramachandra@ntnu.no, lutong@nju.edu.cn, umapada@isical.ac.in, lopresti@cse.lehigh.edu, badrul@um.edu.my

Message from The PaperCept Conference Management System

Message originated by Nicu Sebe

Dear Dr. Shivakumara Palaiahnakote,

After careful consideration of reviewers scores, we are pleased to inform you that the paper

Submission number: 475
Title: Local Gradient Difference Based Mass Features for Classification of 2D-3D Natural Scene Text Images
Authors: Lokesh Nandanwar, Shivakumara Palaiahnakote*, Raghavendra Ramachandra, Tong Lu, Umapada Pal, Daniel Lopresti, Nor Badrul Anuar

you have submitted to ICPR 2020 has been ACCEPTED for presentation and inclusion in the proceedings. The type of presentation (Oral/Poster) will be decided soon (around the end of October).

Attached to this e-mail message you will find the reviews that were received for your manuscript (these may not be present in case your manuscript was only evaluated by the corresponding AC and the Track chairs) together with a report from the Area chair in charge of your submission.

We need to receive the camera ready version for inclusion in the Conference proceedings by October 19, 2020 at the latest.

To upload the camera ready version of the manuscript, the corresponding author must log in to the submission site, enter the workspace as author and follow the action link. Only the corresponding author is able to do this. Only minor changes will be allowed to the version approved and these addressing the reviewers' suggestions for improvement.

Please note that at least one of the authors must register for the conference paying a regular registration fee (not student fee) before submitting the camera-ready copy of his/her manuscript(s). Papers submitted without registration will not be included in the conference program or in the proceedings. Instructions for registration will be given on the conference webpage in the following days.

Sincerely,

Kim Boyer
Brian Lovell
Marcello Pelillo
Nicu Sebe
Rene Vidal
Jingyi Yu

Program Chairs, ICPR2020

Decision: Accepted as Revised paper.
Decision on Supplementary material: Accepted

Final submission deadline October 15, 2020.

UNIVERSITY OF MALAYA

Related Publication

International Journal of Pattern Recognition and Artificial Intelligence (IJPRAI 2020)

International Journal of Pattern Recognition
and Artificial Intelligence
Vol. XX, No. X (2006) 1–19

World Scientific
www.worldscientific.com

A NEW HYBRID METHOD FOR CAPTION AND SCENE TEXT CLASSIFICATION IN ACTION VIDEO IMAGES

[1]Lokesh Nandanwar, [1]Palaiahnakote Shivakumara, [2]Umapada Pal, [3]Tong Lu and [4]Michael Blumenstein

[1]*Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur, Malaysia. Email: lokeshnandanwar150@gmail.com, shiva@um.edu.my*

[2]*Computer Vision and Pattern Recognition Unit, Indian Statistical Institute, Kolkata, India. Email: umapada@isical.ac.in*

[3]*National Key Lab for Novel Software Technology, Nanjing University, Nanjing, China. Email: lutong@nju.edu.cn*

[4]*University of Technology Sydney, Australia. Email: michael.blumenstein@uts.edu*

Abstract. Achieving a better recognition rate for text in the action video image is challenging due to multi-type text with unpredictable actions in the background. We propose a new method for the classification of caption (which is edited text) and scene text (which is part of a video image) in video images. This work considers five action classes, namely, Yoga, Concert, Teleshopping, Craft, and Recipe, where it is expected both the text plays a vital role in understanding the video content. The proposed method introduces a new fusion criterion-based on DCT and Fourier coefficients to obtain the reconstructed images for caption and scene text. The fusion criterion involves computing the variances for coefficients of corresponding pixels of DCT and Fourier images, and the same variances are considered as the respective weights. This step results in Reconstructed image-1. Inspired by the special property of Chebyshev-Harmonic-Fourier-Moments (CHFM) that has the ability to reconstruct a redundancy-free image, we explore CHFM for obtaining the Reconstructed image-2. The reconstructed images along with the input images are passed to Deep Convolutional Neural Network (DCNN) for classification of caption and scene text. Experimental results on five action classes and comparative study with the existing methods demonstrate the proposed method is effective. In addition, the experiments of different recognition methods on the output of classification show that the proposed classification improves recognition performance significantly after classification.

Keywords: Caption text, Scene text, Fusion, DCT coefficients, Chebyshev-Harmonic-Fourier-Moments, Caption and Scene text classification, Action image recognition.

Dear Authors of Selected Paper in ICPRAI 2020:

Paper # 137: Lokesh Nandanwar, Shivakumara Palaiahnakote, Umapada Pal, Tong Lu,
Daniel Lopresti, Bhagesh Seraogi and Bidyut. B. Chaudhuri
A New Method For Detecting Altered Text in Document Images

As indicated on the conference website

- http://www.icprai2020.com or
- https://users.encs.concordia.ca/~icprai20/),

outstanding papers are being selected for inclusion in a Special Issue in IJPRAI (Int. J. of Pattern Recognition and Artificial Intelligence):

https://www.worldscientific.com/page/ijprai/submission-guidelines

to be Guest Edited by Profs. Yue Lu, Nicole Vincent, and Ching Suen:

Please let us know by June 15th whether you and your co-authors are interested in preparing a revised and improved version of your paper by Aug. 1st, to be included as a paper in the Special Issue. Please note that your original conference submission should be extended by 30% in length in the following format for IJPRAI:

https://www.overleaf.com/latex/templates/international-journal-of-pattern-recognition

Once we hear a positive reply from you, we will invite you to submit your paper to our special Easychair website for handling your paper and for review purposes. For your information, the tentative publication schedule is as follows:

June 8 - Notice of tentative acceptance sent to Authors
June 15 - Deadline for Authors to send confirmation of participation
Aug. 1 - Submission of papers via Easychair (please specify it is for the

UNIVERSITY OF MALAYA

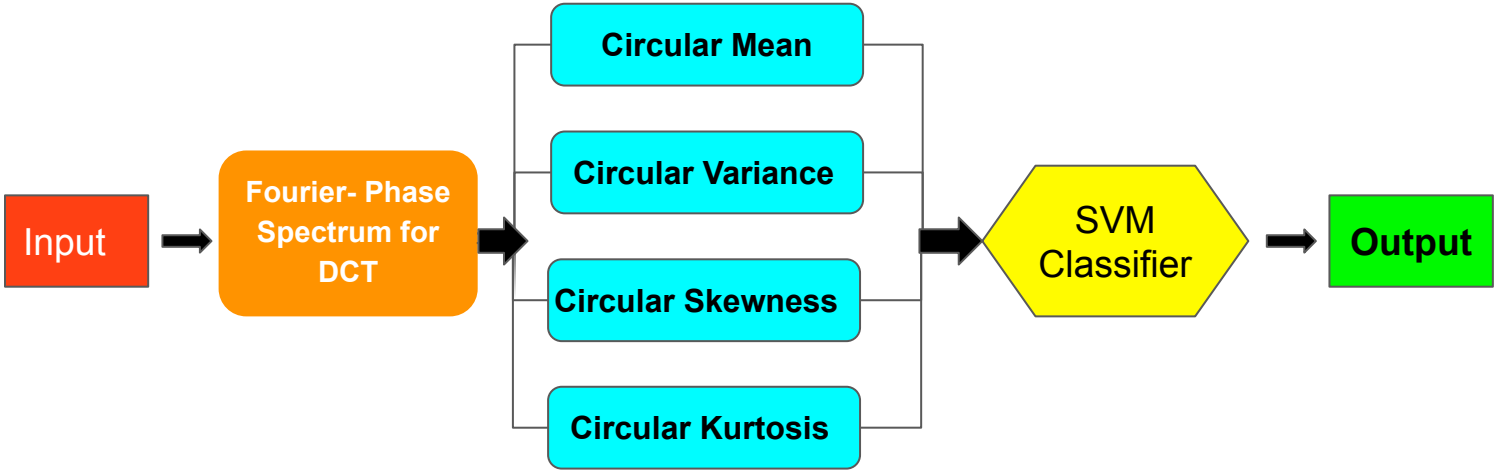
Research Objective-2

RO2: To propose a new method for forged text detection in natural scene images by exploring Fourier spectrum analysis.

RQ2: How to explore and employ Fourier spectrum analysis to detect the forgery in natural images?

UNIVERSITY OF MALAYA

Methodology: Forged IMEI and Air ticket number detection in Images



UNIVERSITY OF MALAYA

(a) Phase Spectrum for DCT

1. Calculating Real DCT Transforms

$$F_{sc}(u,v) = C_1C_2 \sum_{m=1}^{p}\sum_{n=1}^{q} f(m,n) sin\left[\frac{\pi u}{p}(m+0.5)\right] cos\left[\frac{\pi v}{q}(n+0.5)\right]$$

$$F_{ss}(u,v) = C_1C_2 \sum_{m=1}^{p}\sum_{n=1}^{q} f(m,n) sin\left[\frac{\pi u}{p}(m+0.5)\right] sin\left[\frac{\pi v}{q}(n+0.5)\right]$$

$$F_{cc}(u,v) = C_1C_2 \sum_{m=1}^{p}\sum_{n=1}^{q} f(m,n) cos\left[\frac{\pi u}{p}(m+0.5)\right] cos\left[\frac{\pi v}{q}(n+0.5)\right]$$

$$F_{cs}(u,v) = C_1C_2 \sum_{m=1}^{p}\sum_{n=1}^{q} f(m,n) cos\left[\frac{\pi u}{p}(m+0.5)\right] sin\left[\frac{\pi v}{q}(n+0.5)\right]$$

$$C_i = \begin{cases} \sqrt{\frac{2}{p}} & if\ 2 \le u \le p \\ \sqrt{\frac{1}{p}} & if\ u = 1 \end{cases} \qquad i = 1, 2.$$

2. Combining the above four real transforms

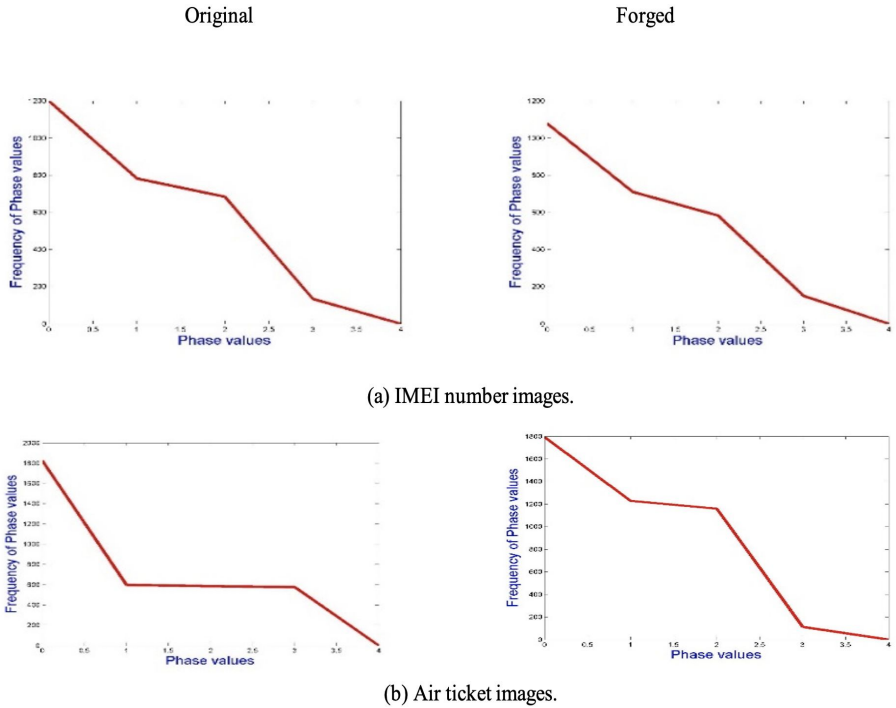
$$\{F_{cc}(u,v) - F_{ss}(u,v)\} - j\{F_{cs}(u,v) - F_{sc}(u,v)\}$$

3. Getting complex fourier transform for DCT

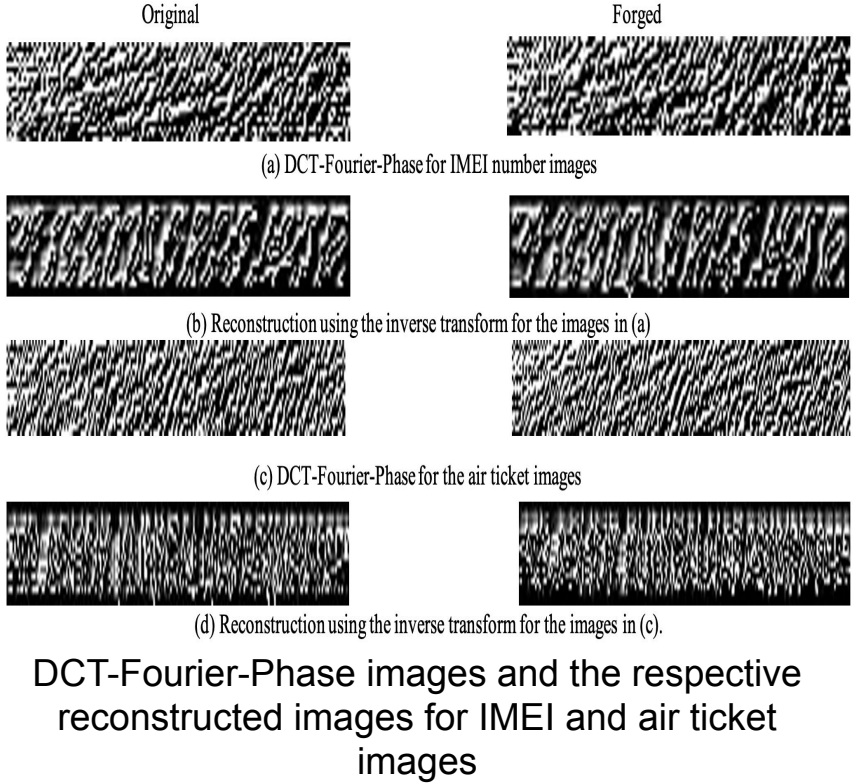
$$F(u,v) = C_1C_2 \sum_{m=1}^{p}\sum_{n=1}^{q} f(m,n) e^{-j\frac{\pi u}{p}(m+0.5)} e^{-j\frac{\pi v}{q}(n+0.5)}$$

UNIVERSITY OF MALAYA

(a) Phase Spectrum for DCT



Fourier Phase for IMEI number and air ticket images



DCT-Fourier-Phase images and the respective reconstructed images for IMEI and air ticket images

UNIVERSITY OF MALAYA

(b) Phase-Statistics for Forged Detection

1. **Circular Mean (CM):** Computes the circular mean in a rectangular based coordinate system
2. **Circular Variance (CV):** This feature is established to describe the distribution of data around the mean
3. **Circular Skewness (CS):** Circular skewness provides information about the shape and symmetry of phase data,
4. **Circular Kurtosis (CK):** Indicates how tall and sharp the peak is in the phase data distribution.

Finally Pass all four features in SVM for final output prediction

UNIVERSITY OF MALAYA

Experimental Results

Methods	IMEI Number Dataset						Air Ticket Dataset					
	Original			Forged			Original			Forged		
	R	P	F	R	P	F	R	P	F	R	P	F
Proposed	**0.83**	0.78	**0.80**	0.76	**0.82**	0.79	**0.77**	**0.81**	**0.79**	**0.82**	0.78	**0.80**
Roy, Shivakumara et al. (2016)	0.63	0.61	0.62	0.60	0.61	0.60	0.67	0.65	0.66	0.65	0.66	0.65
Bhardwaj and Pankajakshan (2016)	0.06	0.52	0.11	**0.94**	0.50	0.65	0.08	0.33	0.13	0.82	0.47	0.60
(Wang, Shivakumara et al. 2017)	0.44	0.39	0.41	0.33	0.38	0.35	0.65	0.82	0.73	0.86	0.71	0.78
(Elkasrawi and Shafait 2014)	0.32	0.34	0.33	0.40	0.37	0.39	0.74	0.66	0.70	0.63	0.71	0.67
Shivakumara et al. (2018)	0.73	**0.82**	0.77	0.85	0.75	**0.80**	0.69	0.46	0.55	0.63	**0.80**	0.70

UNIVERSITY OF MALAYA

Experimental Results

Methods	Roy et al dataset						Bhardwaj et al. dataset					
	Original (scene)			Forged (caption)			Set-1(Low resolution)			Set-2(High resolution)		
	R	P	F	R	P	F	R	P	F	R	P	F
Proposed	**0.88**	0.73	**0.80**	0.68	**0.85**	0.75	**0.82**	**0.87**	**0.84**	**0.92**	**0.88**	**0.90**
Roy, Shivakumara et al. (2016)	0.75	0.71	0.73	0.63	0.68	0.65	0.62	0.58	0.60	0.66	0.63	0.64
Bhardwaj and Pankajakshan (2016)	0.56	0.46	0.51	0.40	0.51	0.45	0.64	0.62	0.63	0.68	0.65	0.67
Wang, Shivakumara et al. (2017)	0.75	0.77	0.76	0.78	0.76	0.77	0.56	0.69	0.62	0.62	0.48	0.54
Elkasrawi and Shafait (2014)	0.83	0.68	0.74	0.61	0.78	0.69	0.47	0.54	0.51	0.40	0.34	0.37
Shivakumara et al. (2018)	0.54	**0.96**	0.69	**0.98**	0.68	**0.80**	0.73	0.70	0.72	0.80	0.74	0.76

Results on Roy et. al and Bharadwaj et. al datasets

Methods	Original			Forged		
	R	P	F	R	P	F
Proposed	0.78	**0.92**	0.84	**0.90**	0.75	0.81
Roy, Shivakumara et al. (2016)	0.88	0.39	0.54	0.60	**0.94**	0.73
Bhardwaj and Pankajakshan (2016)	0.81	0.27	0.41	0.56	0.93	0.70
(Wang, Shivakumara et al. 2017)	**0.87**	0.84	**0.86**	0.85	0.88	**0.86**
(Elkasrawi and Shafait 2014)	0.65	0.62	0.63	0.64	0.67	0.65
Shivakumara et al. (2018)	0.65	0.92	0.76	0.86	0.50	0.63

Results on ICPR FDC dataset

UNIVERSITY OF MALAYA

Related Publication

Expert Systems with Applications: Journal Published

ELSEVIER

Expert Systems with Applications

Volume 164, February 2021, 114014

DCT-phase statistics for forged IMEI numbers and air ticket detection

Lokesh Nandanwar [a], Palaiahnakote Shivakumara [a], Swati Kanchan [b], V. Basavaraja [c], D.S. Guru [c], Umapada Pal [b], Tong Lu [d], Michael Blumenstein [e]

Show more

https://doi.org/10.1016/j.eswa.2020.114014

Get rights and content

Highlights

- A new phase technique for detecting forgeries in IMEI and air tickets.
- It derives phase spectrum using DCT to find suspicious regions.
- We compute phase statistics to study the effect introduced by forgery.

Abstract

New tools have been developing with the intention of having more flexibility and greater user-friendliness for editing the images and documents in digital technologies, but, unfortunately, they are also being used for manipulating and tampering information. Examples of such crimes include creating forged International Mobile Equipment Identity (IMEI) numbers which are embedded on mobile packages and inside smart mobile cases for illicit activities. Another example of such crimes is altering the name or date on air tickets for breaching security at the airport. This paper presents a new expert system for detecting forged IMEI numbers as well as altered air ticket images. The proposed method derives the phase spectrum using the Discrete Cosine Transform (DCT) to highlight the suspicious regions; it is unlike the phase spectrum from a Fourier transform, which is ineffective due to power spectrum noise. From the phase spectrum, our method extracts phase statistics to study the effect of distortions introduced by forgery operations. This results in feature vectors, which are fed to a Support Vector Machine (SVM) classifier for detection of forged IMEI numbers and air ticket images. Experimental results on our dataset of forged IMEI numbers (which is created by us for this work), on altered air tickets, on benchmark datasets of video caption text (which is tampered text), and on altered receipts of the ICPR 2018 FDC dataset, show that the proposed method is robust across different datasets. Furthermore, comparative studies of the proposed method with the existing methods on the same datasets show that the proposed method outperforms the existing methods. The dataset created will be available freely on request to the authors.

UNIVERSITY OF MALAYA

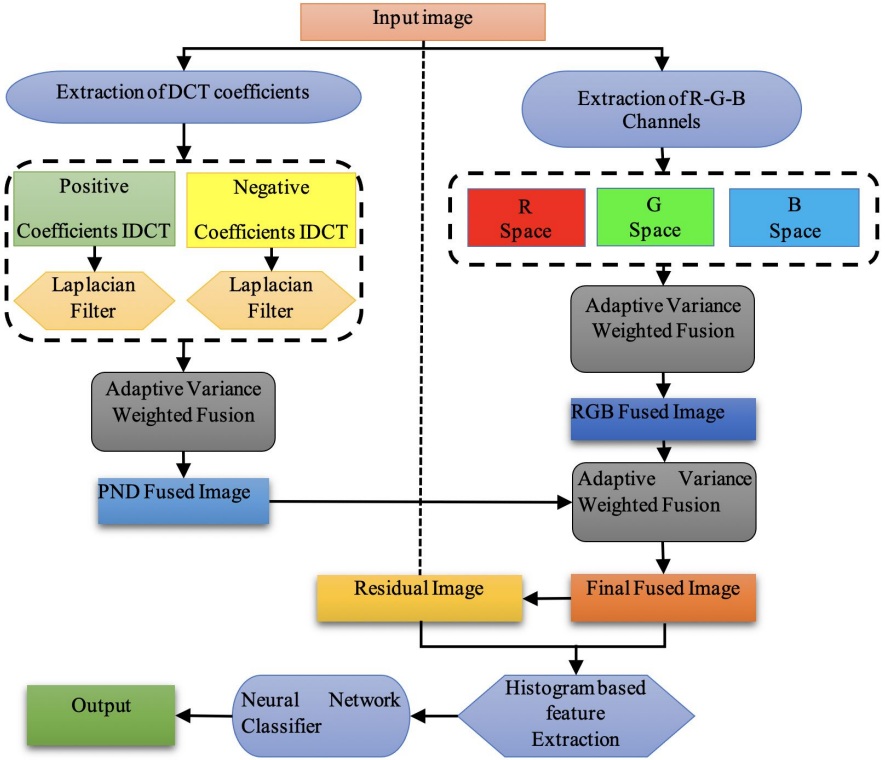
Research Objective-3

RO3: To explore a new method for detecting altered text in the document images based on fusion and reconstruction of the images.

RQ3: By what means one can approach fusion and reconstruction methods to spot the changes caused by forgery in document images?

UNIVERSITY OF MALAYA

Altered text and handwriting detection in Documents



Adaptive Variance Weighted Fusion

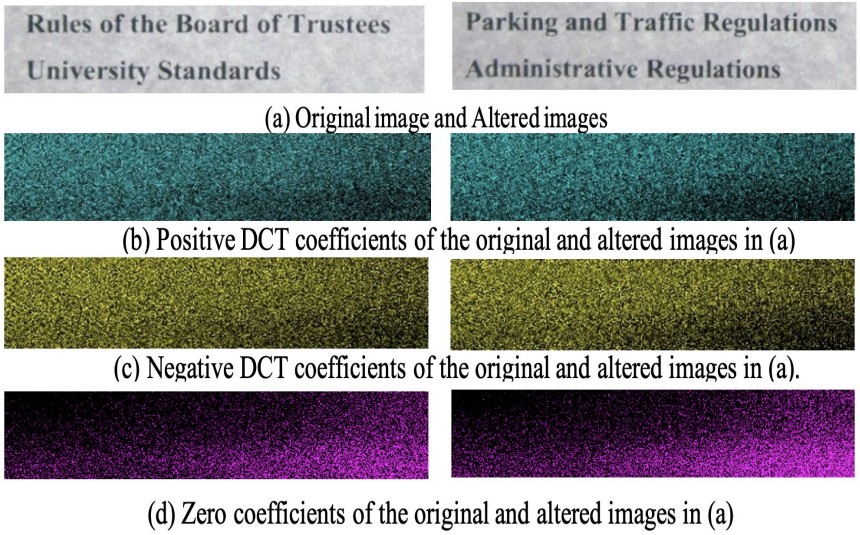
$$Var_i(x,y) = \frac{1}{M \times N} \sum_{a=1}^{M} \sum_{b=1}^{N} [abs(I_i(x+a, y+b) - \mu]^2$$

$$w_i = \frac{Var_i(x,y)}{\sum_{a=1}^{n} Var_a(x,y)}$$

$$\mathcal{F} = \sum_{i=1}^{n} I_i \times w_i$$

UNIVERSITY OF MALAYA

(a) Positive and Negative DCT Coefficient Fusion (PNDF)



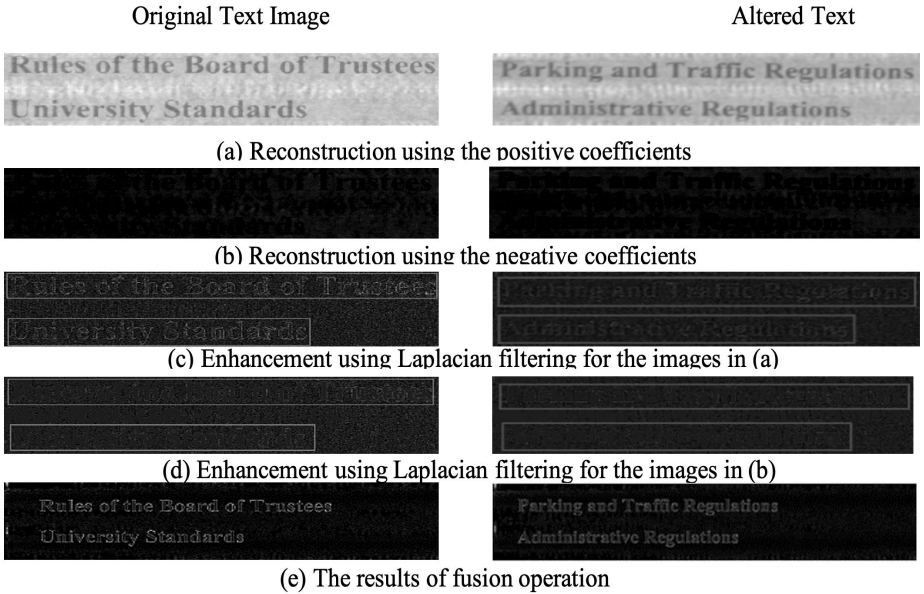
(a) Original image and Altered images

(b) Positive DCT coefficients of the original and altered images in (a)

(c) Negative DCT coefficients of the original and altered images in (a).

(d) Zero coefficients of the original and altered images in (a)

Positive, negative and zero coefficient distributions of the original and altered images.



(a) Reconstruction using the positive coefficients

(b) Reconstruction using the negative coefficients

(c) Enhancement using Laplacian filtering for the images in (a)

(d) Enhancement using Laplacian filtering for the images in (b)

(e) The results of fusion operation

The steps for fusing the positive and negative DCT

UNIVERSITY OF MALAYA

(b) R G B Color Channels Fusion (RGBF)

Original

(a) R channel

(b) G channel

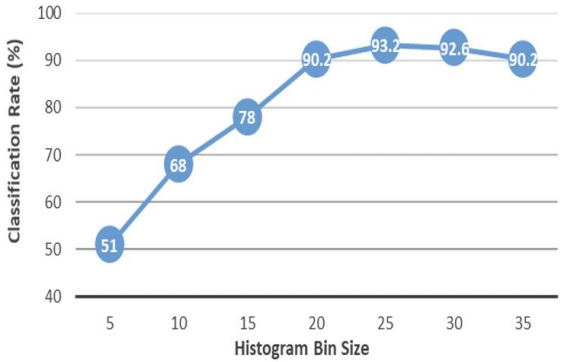
(c) B channel

(d) RGB fusion

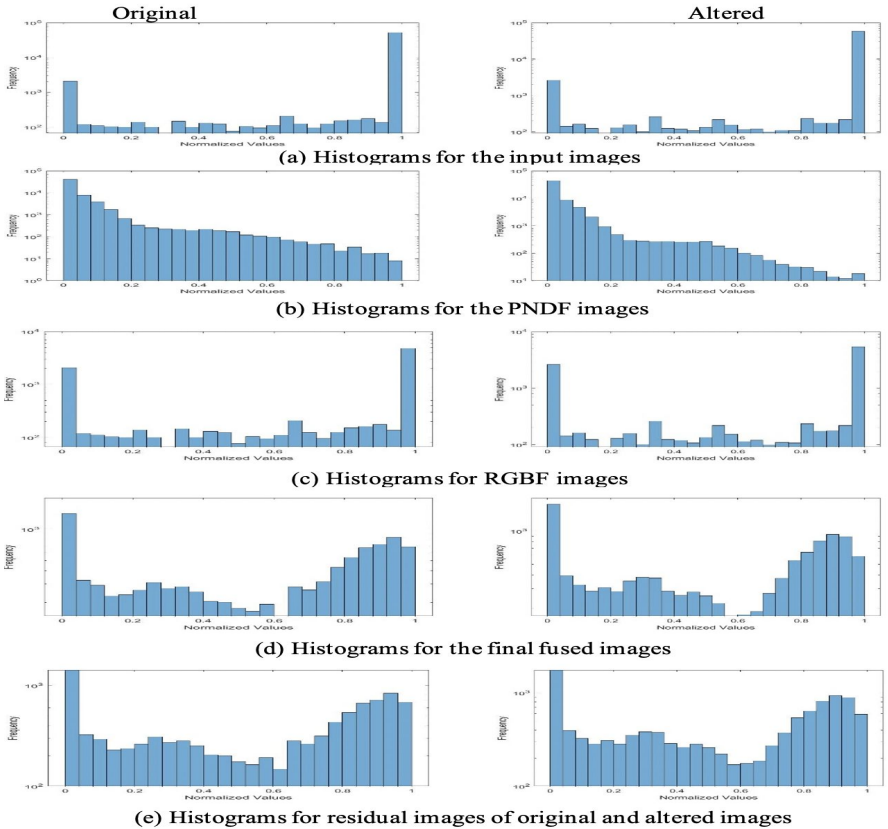
(e) Proposed the final fusion (PNDF + RGBF)

UNIVERSITY OF MALAYA

(c) Features Extraction

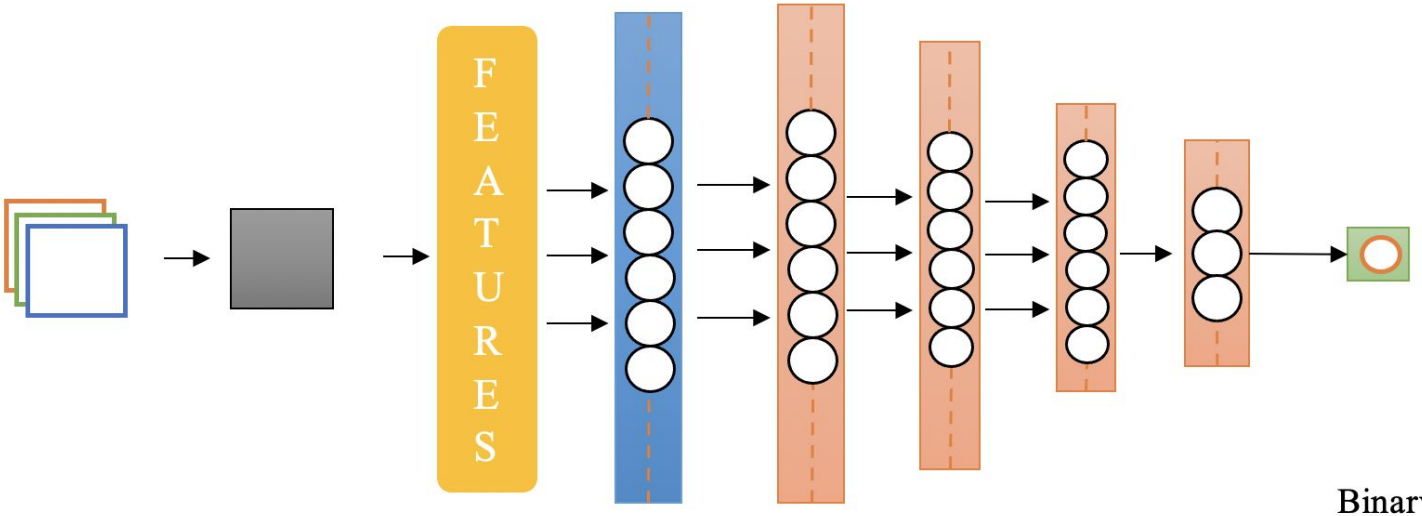


Study to Determining the feasible number of bins for histogram to extract features.



UNIVERSITY OF MALAYA

(c) DNN for Altered Text Detection



UNIVERSITY OF MALAYA

Experimental Results

Results on Altered documents dataset, ICPR FDC and IMEI number dataset

Methods	Dataset	Own Dataset		ICPR 2018		IMEI dataset	
		Original	Altered	Original	Altered	Original	Altered
Kundu et. al	Original	60.0	50.0	90	10	57.8	42.2
	Altered	50.0	50.0	72.5	27.5	41.8	58.2
	ACR	55.0		78.3		58.0	
Wang et al.	Original	80	20	84.66	8.67	83.2	16.8
	Altered	13.4	86.6	7.95	89.33	25.6	74.4
	ACR	83.3		86.99		78.8	
Shivakumara et al.	Original	60	40	92	8	82.2	17.8
	Altered	35	65	49.44	50.66	18	82
	ACR	62.5		71.33		82.1	
Proposed	Original	95.4	4.5	96.7	3.3	89	12.6
	Altered	9.1	90.9	8.0	92.0	10.2	84
	ACR	**93.2**		**94.35**		**86.5**	

Confusion matrix and average classification rate of the proposed and existing methods on Altered handwriting dataset (in %). O, A, B, N denote class names as Original, Altered, Blurred and Noise, respectively and ACR is average classification rate of the classes.

C	Kundu et. al.				Wang et al.				Shivakumara et. al				Proposed method			
	O	A	B	N	O	A	B	N	O	A	B	N	O	A	B	N
O	85.7	4.8	10.0	0.0	57.8	25.0	7.8	9.4	77.0	23.0	0.0	0.0	85	15	0	0
A	20.0	70.0	9.5	0.0	25.0	71.4	1.8	1.8	40.0	50.0	0.0	10.0	25	70	5	0
B	15.8	15.8	63.2	5.2	9.0	1.8	78.2	11.0	22.0	0.0	78.0	0.0	0	0	100	0
N	0.0	0.0	10.0	90.0	14.3	8.0	1.5	76.2	0.0	0.0	10.0	90.0	2.5	0	0	97.5
ACR	77.5				70.1				73.75				**88.12**			

UNIVERSITY OF MALAYA

Related Publication

Transactions on Information Forensics and Security (IFS)

Gmail — Lokesh Nandanwar <lokeshnandanwar150@gmail.com>

Decision: MAJOR REVISION (RQ) - T-IFS-11738-2020

Transactions on Information Forensics & Security <onbehalfof@manuscriptcentral.com> Fri, Oct 23, 2020 at 3:39 PM
Reply-To: emanuele.maiorana@uniroma3.it
To: lutong@nju.edu.cn
Cc: emanuele.maiorana@uniroma3.it, lokeshnandanwar150@gmail.com, shiva@um.edu.my, hamidjalab@um.edu.my, rabhaibrahim@tdtu.edu.vn, raghavendra.ramachandra@ntnu.no, umapada@isical.ac.in, lutong@nju.edu.cn, lopresti@cse.lehigh.edu, patrizio.campisi@uniroma3.it

23-Oct-2020

Prof. Tong Lu
National Key Lab for Novel Software Technology, Nanjing University, China
163 Xianlin Avenue, Qixia District
Nanjing 210023
Nanjing
Jiangsu
China
210023

Paper:T-IFS-11738-2020, "Deep-Conformable Moments for Altered Handwriting Detection"

Dear Prof. Tong Lu,

I am writing to you concerning the above referenced manuscript, which you submitted to the IEEE Transactions on Information Forensics & Security.

Based on the enclosed set of reviews, your manuscript requires a MAJOR REVISION (RQ).

A major issue of the paper is given by its questionable novelty. Reviewers have pointed out that the proposed method is basically a combination of already proposed approaches without a proper justification for such ensemble. More importantly, it is not clear which novelty is proposed with respect to other works already proposed by the authors.
A proper, detailed, and convincing motivation and description of the novel aspects of the proposed approach has to be therefore necessarily provided.

The description of the proposed approach has also to be significantly rewritten, being it too hard to be read and understood as it is.

There are also severe concerns regarding the considered cases of handwriting samples to be discriminated. A proper justification of the taken choices has to be provided in this regard, considering that the author themselves write that there hard hard scenarios which could be dealt with, without then facing them in the paper.

Clarifications regarding the outcomes reported in the paper have to be also provided.

Failing in providing the required modifications and justification would imply the rejection of the paper.

Your revised manuscript must be submitted back to ScholarOne Manuscripts https://mc.manuscriptcentral.com/tifs-ieee no later than 6 weeks from the date of this letter together with a required point-by-point reply that explains how you addressed the reviewers' comments. If we do not receive your revised manuscript within 6 weeks from the date of this letter, your manuscript will be considered withdrawn.

After you finish revising your manuscript, please log into your Author Center at https://mc.manuscriptcentral.com/tifs-ieee to upload the new file(s) to your submission. You will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Revision."

UNIVERSITY OF MALAYA

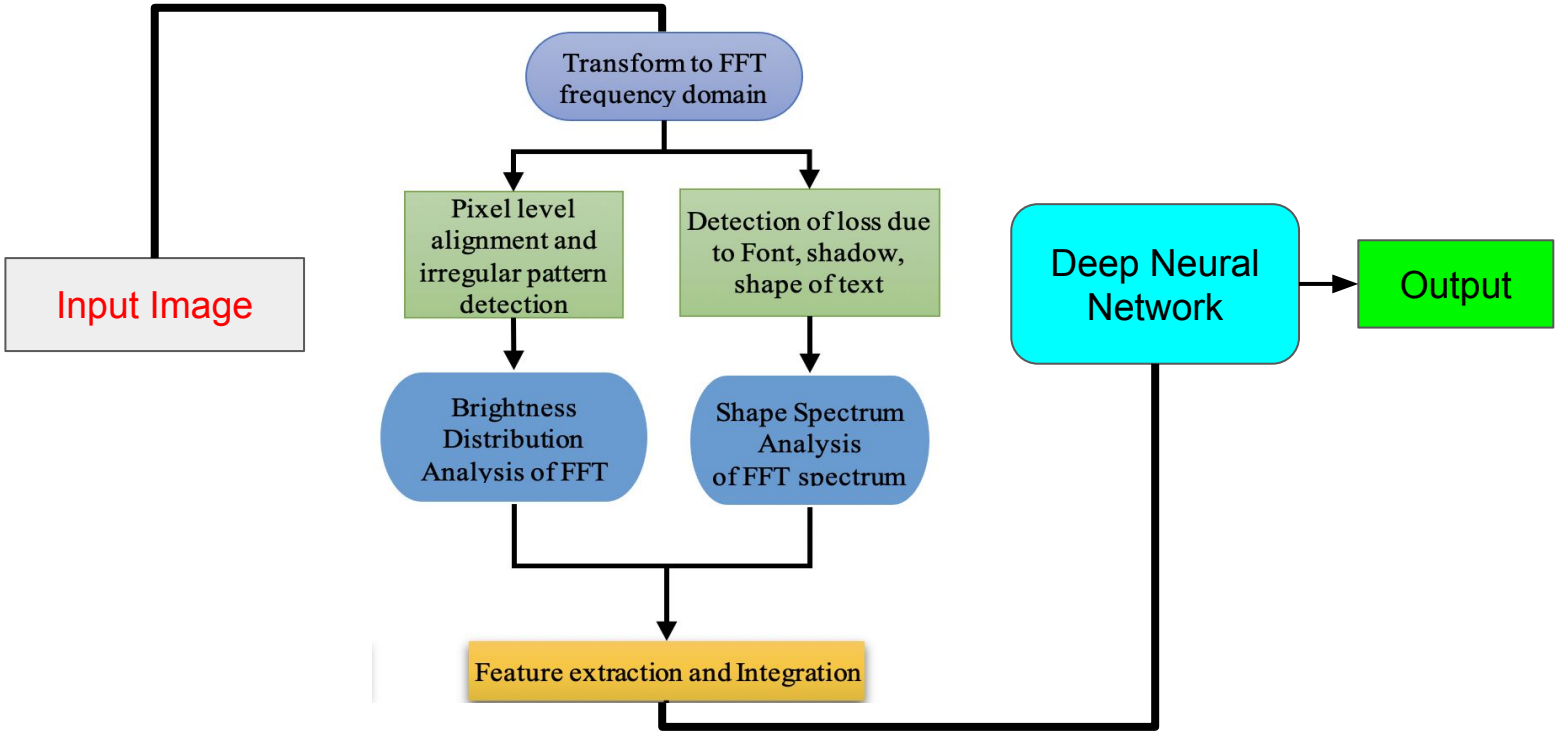
Research Objective-4

RO4: To design and develop a unified method for detecting forgery in video, natural scene and document images.

RQ4: How to develop a unified method for detecting forgery which can adapt to multiple multimedia formats such as video, natural scene and document images?

UNIVERSITY OF MALAYA

Work Ongoing: Unified Method for forgery detection in Video, Scene and Document images



UNIVERSITY OF MALAYA

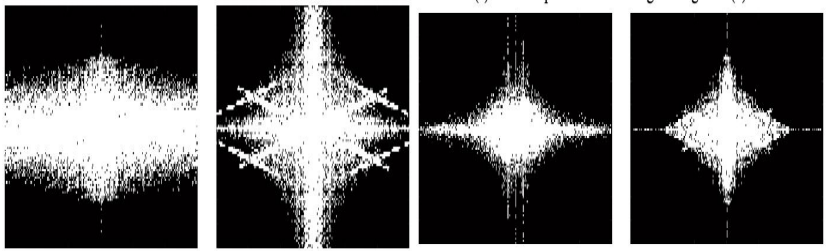
(a) Brightness Distribution Analysis

(a) Original images for copy-paste and insertion operations.

(d) Forged text corresponding to the above images in (a).

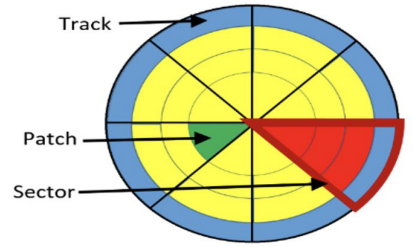
(b) Fourier spectrum of the images in (a).

(e) Fourier spectrum of the forged images in (d).

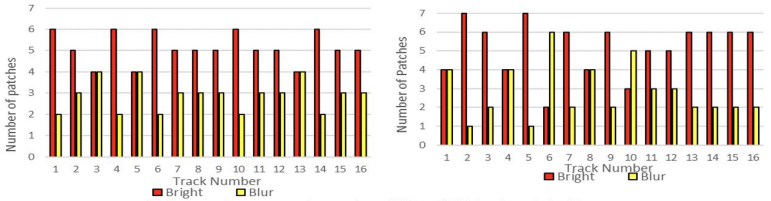


(c) Binary form for Fourier spectrum in (b)

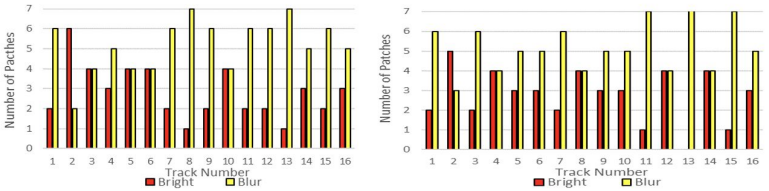
(f) Binary Fourier spectrum of forged images shown in (e)



(a) Constructing patches (green color) for the sectors (red color) and tracks (blue color) of the spectrum



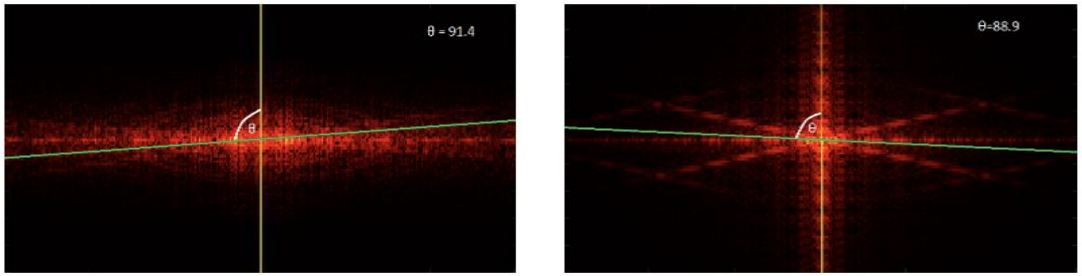
(b) Bar graphs to count the number of SP and BP for the original images



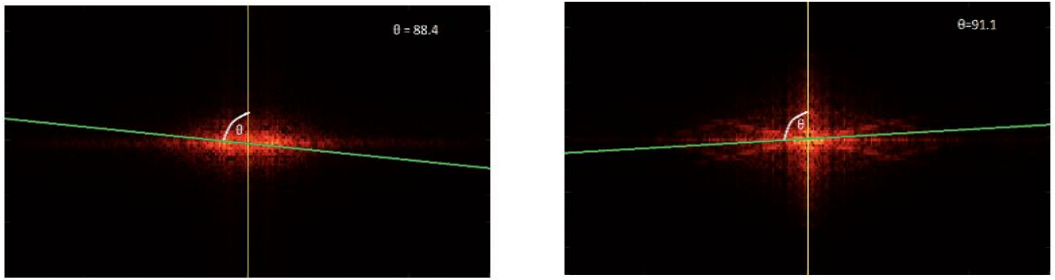
(c) Bar graphs to count the number of SP and BP for the forged images

UNIVERSITY OF MALAYA

(b) Spectrum Shape Analysis



(a) Principal axis (green color line) for the original images of copy-paste and insertion operations.



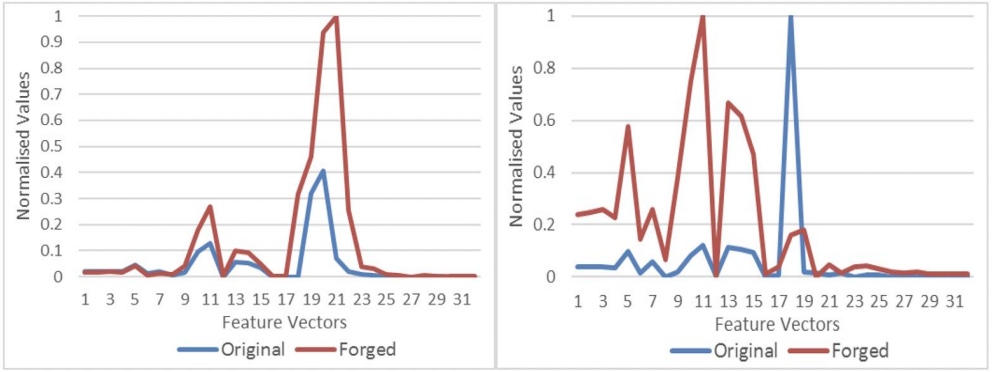
(b) Principal axis (green color line) for the forged text attacked by copy-paste and insertion operations.

Principal axes (green color line) for the original text and the forged text attacked by copy-paste and insertion operations

Feature Extraction: PCA angle for each track in the spectrum

UNIVERSITY OF MALAYA

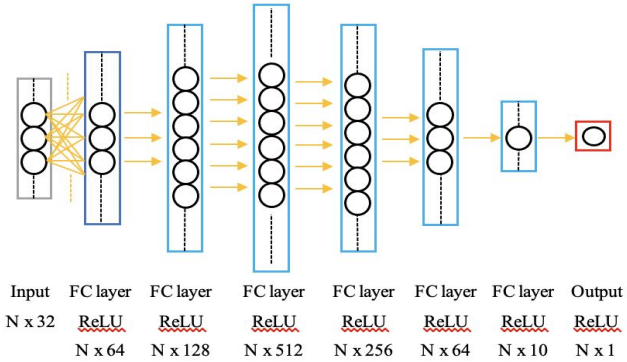
(c) Features Aggregation and Deep NN classifier



(a) Feature vectors of original and forged text image for copy-paste operation.

(b) Feature vectors of original and forged text image for insertion operation.

Normalized feature vector Input



Deep Neural Network Classifier

UNIVERSITY OF MALAYA

Work Remaining

1. Experiments and Evaluation on our and standard datasets on Unified method.

2. Experiments to check the robustness of the method by effects of noise and blur classes.

3. Adding more classes to the existing Four-class Altered handwritten documents dataset by introducing noise and blurriness to forged images.

UNIVERSITY OF MALAYA

Related Publication

IET Image Processing Journal

Gmail — Lokesh Nandanwar <lokeshnandanwar150@gmail.com>

Decision on your Paper - IET Image Processing

RVT -Review Management System <editorialoffice_1@iet-review.rivervalleytechnologies.com> Fri, Oct 23, 2020 at 9:06 PM
Reply-To: iet_ipr@theiet.org
To: lokeshnandanwar150@gmail.com

Dear Mr Lokesh Nandanwar,

IPR-2020-0590

Forged Text Detection in Video, Scene and Document Images

Thank you for submitting your paper to IET Image Processing. The peer review process is now complete and the conclusion is that, while the paper is relevant and potentially worthy of publication, significant revisions are needed. My decision is therefore to provisionally accept your paper subject to major revisions.

The reviewer comments are given below. Please make sure that you address all the comments when revising the manuscript. Please be aware that your paper may be declined if the changes cannot be easily identified. If the referees are not convinced that their concerns have been addressed this will, at best, result in a request for further modifications which will delay publication, at worst it could result in rejection.

For this reason, when resubmitting your revised manuscript, please deal fully with the reviewers' comments and also detail how you have dealt with them in a covering letter. This covering letter should be included in both the 'response to decision letter' dialogue box, as well as uploaded as an additional file. If you do not agree with a reviewer's comment you must state this and explain carefully why you do not agree.

When you have completed your revisions please:

- Upload your revised paper in both PDF and source file format. Note that figures will be used as supplied and that these should therefore be of high quality.

- Please make sure that references are formatted within the IET's house style. Sample references are available from the online Author Guide at IET Digital Library Author Guide

To upload your revised paper please go to https://www.iet-review.rivervalleytechnologies.com/ and the paper should be listed in your Tasks and say "Ready for Major Revision". Then select View Article and you will find the Start Major Revision button here. You will then have the option to Edit the article details and files from the previous version of the paper.

Please submit source files (.doc, docx, or .tex. files for text and .eps, .tif or jpeg files for figures). If your paper has been prepared using LaTeX, please also upload a single .pdf file of the paper together with the LaTeX source file and the figures. Ensure that you delete any files that do not form part of the revision before completing the submission.

Please note that, in order to publish your paper as quickly as possible, we ask that you submit the revision of your paper within 28 days of receiving this email. If you need more time, it is important that you inform the IET office before the deadline so that they may be able to grant an extension. If you are having trouble resubmitting your paper, please contact the IET office directly by replying to this email.

I look forward to receiving your revised version.

Yours sincerely

Professor Farzin Deravi
IET Image Processing

IEEE Transactions on Multimedia (TMM)

From: **IEEE Transactions on Multimedia** <onbehalfof@manuscriptcentral.com>
Date: Wed, Sep 30, 2020 at 10:45 AM
Subject: Decision: MAJOR REVISION (RQ) - MM-011454, 3DTDS: 3D Video Text Detection System
To: <shiva@um.edu.my>
Cc: <jiebo.luo@gmail.com>, <jluo@cs.rochester.edu>, <caoxiaochun@iie.ac.cn>

29-Sep-2020

Dr Palaiahnakote Shivakumara
University of Malaya
B-2-18, Block B,
Annex Building,FSKTM, UM,
Kuala Lumpur-50603, Malaysia
Kuala Lumpur
Malaysia
50603

Paper:MM-011454 3DTDS: 3D Video Text Detection System

Dear Dr Shivakumara,

I am writing to you concerning the above referenced manuscript, which you submitted to the IEEE Transactions on Multimedia. (**See note below about attachments).

Based on the enclosed set of reviews, it was recommended that the manuscript be REVISED AND RESUBMITTED (RQ). The reviews are quite constructive and helpful. Please address all the raised issues in the revision. Note that there is typically only ONE round of major revision.

We hope you will be able to implement the comments of the reviewers. Your revised manuscript must be submitted back to Manuscript Central https://mc.manuscriptcentral.com/tmm-ieee no later than 6 weeks from the date of this letter to be further considered for publication in the IEEE Transactions on Multimedia. If we do not receive your revised manuscript within this specified time, your manuscript will be considered withdrawn.
* PLEASE NOTE: deadline extension requests that exceed 1 week must be submitted to the Editor-in-Chief for approval. The AE and admin cannot approve extension requests greater than 1 week.

Please be sure to upload the revised manuscript through the account of the author who submitted the original version. This can be done by clicking the "Create a Revision" link next to the manuscript's entry.

Please note that it is the T-MM policy that resubmitted papers can only be accepted or rejected either an AQ, A or R decision may be granted). If the reviewers and associate editor are not satisfied that all of their concerns have been addressed, the only option for the paper is rejection. As such, please ensure that all comments are addressed, and include a detailed description of your responses and changes to the manuscript based on the reviewer comments.

Please note that resubmitting your manuscript does not guarantee eventual acceptance, and that your resubmission will be subject to re-review by the reviewers before a decision is rendered.

Please remember that the Associate Editor should only decide RQ (major revision) once during the peer review process of any paper. Subsequent decisions after one RQ should be AQ, A, or R only.

* If you have any questions regarding the reviews, please contact the Associate Editor who managed your paper. All other inquiries should be directed to the Admin.

Best regards,

Scope and Limitations of the Study

1. Classification of tampered text (Caption) and natural text in Video images
2. Limited to detect tampering at text level and not in the scope to detect the forgery at visual context level or general image forgery.
3. Method can be used to detect text forgery in still images such as IMEI number images, and in other images with copy-paste and insertion operations.
4. Not in scope of the method to detect the forgery through splicing which is a result of two source images for forgery
5. Can be helpful in detection of forgery in plain PDF, Handwritten texts and printed text documents.
6. Not in scope to detect forgery in decorated documents such as banking insurances, where there are pictures, logo embedded in text.

UNIVERSITY OF MALAYA

Conclusion of the Study

1. Hybrid Method for Caption and Scene Text Classification in Action Video Images using new fusion concept to integrate the advantages of Fourier and DCT coefficients to obtain the reconstructed image.
2. Forged IMEI Numbers and Air Ticket Detection using the magnitude of DCT and the phase of Fourier, which are combined in a new way for generating a phase-spectrum.
3. A method for Detecting Altered Text in Document and handwriting Images using DCT coefficients fused image and RGB fused image which exploits frequency domain and spatial domain simultaneously for feature extraction.
4. Unified method of Forged Text Detection in Video, Scene and Document Images using FFT coefficients and Fourier spectrum shape analysis.

UNIVERSITY OF MALAYA

All Publications

	Proceedings	Title	Status
1	Expert Systems with Applications Journal (**ESWA**)	DCT-Phase Statistics for Forged IMEI Numbers and Air Ticket Detection	Published
2	10th IAPR workshop on Document Analysis Systems (**DAS 2020**)	A New Common Points Detection based Method for Classification of 2D and 3D Text in Video/Scene Images	Published
3	25th International Conference on Pattern Recognition (**ICPR 2020**)	Chebyshev-Harmonic-Fourier-Moments and Deep CNNs for Detecting Forged Handwriting	Accepted
4	25th International Conference on Pattern Recognition (**ICPR 2020**)	Local Gradient Difference Based Mass Features for Classification of 2D-3D Natural Scene Text Images	Accepted
5	International Conference on Pattern Recognition and Artificial Intelligence (**ICPRAI 2020**)	A New Method For Detecting Altered Text in Document Images	Published
6	International Conference on Pattern Recognition and Artificial Intelligence (**ICPRAI 2020**)	A New Method for Caption and Scene Text Classification in Action Video Images	Published
7	IET Image Processing Journal	Forged Text Detection in Video, Scene and Document Images	Under Revision
8	IEEE Transactions on Information Forensics and Security (**IFS**)	Deep Conformable Moments based Technique for Altered Handwriting Detection	Under Revision
9	IEEE Transactions on Multimedia (**TMM**)	3DTDS: 3D Video Text Detection System	Under Revision
10	International Journal of Pattern Recognition and Artificial Intelligence (**IJPRAI 2020**)	A New Method For Detecting Altered Text in Document Images	Submitted
11	International Journal of Pattern Recognition and Artificial Intelligence (**IJPRAI 2020**)	A New Hybrid Method for Caption and Scene Text Classification in Action Video Images	Submitted

UNIVERSITY OF MALAYA

Conclusion

1. We proposed novel methods along with the unified method for forgery detection.
2. The proposed methods explores characteristics of brightness distribution, shapes of spectrum, Phase statistics for extracting features in DCT, FFT domains.
3. Proposed methods integrates the features and take advantages of deep neural networks.
4. To validate the proposed method, we conduct experiments on our proposed four datasets along with four Standard datasets
5. We Show that the proposed method is effective and generic in each type of multimedia format.
6. Multimodal and robust to different situations, it can be implemented in real time environments in future.

UNIVERSITY OF MALAYA

Thanks!

Questions?