University of Sindh Journal of Information and Communication Technology
(USJICT)
Volume 7, Issue 1

ISSN-E: 2523-1235, ISSN-P: 2521-5582
Website: http://sujo.usindh.edu.pk/index.php/USJICT/

A Framework for Forgery Detection from the Digital X-Ray Images

1st Samar Abbas Mangi, 2nd Samina Rajper, 3rd Jamil Ahmed Chandio, 4th Noor Ahmed Shaikh

Institute Of Computer Science, Shah Abdul Latif University Khairpur Mirs Sindh
mangisamar@gmail.com, samina.rajper@gmail.com , jameel_lrk@yahoo.com , noor.shaikh@salu.edu.pk

Abstract: Due to high significance of medical images, the medical image analysis has become well recognized research area of computer science and recent advances in computational technology have boost up the process to alter the image, sound and video data collectively referred to as ISV. During the process of ISV capturing to printing, there are fair chances of ISV forgery, which could become leading case of misdiagnosis in medical images such as Digital X-rays (DX), CTs, MRIs , such unnecessary manipulation with medical image may potentially lead to misdiagnosis. Such unauthorized changes to medical images not only effect the accuracy of diagnosis but also effect the cost of treatment. However forgery detection (FD) at early stage would provide additional assistance to doctors to investigate the health condition patients in more effective way. Since the malicious intensions (MI) of human or malfunctioning of DXs would include extra noise into ISVs, which could become of the leading cause of misdiagnosis but the pixel level image analytics could be used to detect such anomalies. This research offers a framework to detect the forgery in DXs and ISVs. This framework includes three key steps: I. Data Preparation, in which the entropy of forgery area be detected by offering significant method. II. Construction of decision Model: that includes in in which support vector machine (SVM) is used to construct the classical model. III. Visualize the Results in which a confusion matrix and other measures are used to detect the forgeries in the DXs. Our proposed approach contributes highest classification accuracy such as 96.90%.

Keywords: Machine Learning; classification, Supervised Learning, Strong Objects, Forgery, X-Rays;

I. Introduction

Recent advancements of digital image processing (DIP) are playing very significant role to interpret the medical images by applying image enhancement techniques (IES), image transformation techniques (IST), image analysis techniques (IAT) and forgery detection in medical image (FDMI) is counted as one of the significant areas of research computer science [1-6]. Since the medical images are frequently prescribed by the doctors to diagnose the various diseases and mostly these noninvasive investigations are very reliable source to diagnose the various diseases but due to the advancement of digital image processing techniques, it is very hard to detect the forgeries from medical images such as sound, videos (ISVs) i.e. Digital X-rays (DX), Digital ECG, Digital EEG, MRIs and many more [7-9]. A little variation or modification in ISV could become the leading cause for misdiagnosis and could become cause for improper treatment which could become leading cause of loss of precious human lives and their unnecessary cost of treatment [10]. An incorrect region of interest (ROI) could be very serious when it effects the diagnosis therefore medical consultants are very careful to perform the interpretation of medical images in shape of report [11-12]. a single X-Ray could interprets various organs and each organ has its own morphologies, appearance and anatomy therefore the use of artificial intelligence techniques could help the doctors during the diagnosis process and forgeries in medical images could also be predicted by using the machine learning techniques [13]. In order to mitigate all above stated problems these papers presents a framework to predict the forgeries from the medical images such as X-Ray chest. Some of the very nice approaches are offering very nice frameworks [14-18] but by considering all above problems. This paper offers a framework to classify the forgery X-Rays by using Machine Learning (ML) techniques. The framework is comprising over three phases where first phase is used to preparer the data and second phase is used to construct the decision model and the third phase is used to interpret the Result of the proposed framework. This paper contributes (a) XRD preparation method (b) highest classification accuracy.

The paper is divided into five sections where section one is reserved for introduction and the section two is placed for literature review. Section three for describing the method and section four for interpreting the results of the conducted experiments. The last section is used for the conclusion and discussion. Type Style and Fonts

Wherever Times is specified, Times Roman or Times New Roman may be used. If neither is available on your word processor, please use the font closest in appearance to Times. Avoid using bit-mapped fonts if possible. True-Type 1 or Open Type fonts are preferred. Please embed symbol fonts, as well, for math, etc.

II. LITERATUE REVIEW

Since the forgery positive patients would face unnecessary treatment therefore the use of machine learning techniques would be useful for doctors to detect the malicious set of objects of ISVs. Although lots of literature available on the forgery detection but yet more accuracies are required to enhance the results. Some of the related literature is presented as given below.

An approach [14] for the detection of the tempering medical images was proposed by using digital image processing techniques where data embedding technique was used with the reversible wavelet. The techniques uses double wavelet transformation methods where first have been used to calculate the pixel quantiles by selecting the regions of the image and second calculates the approximation for possible tempering situation under the medical image, however the accuracy of the approach is not shown by the authors whereas we consider the strong and weak set of objects in ISVs because strong objects would be more useful to predict and to enhance the classification accuracy for ISVs.

An approach [15] proposed for detection of tempering of human faces was proposed to perform the forensics of human face recognition. As per authors it has become problem for humans to detect the forgery in human faces. The computer vision technique were used to detect the forgeries by consider the human faces as an object where collaborative feature learning method has been applied. The deep learning method our.net was used and the classification accuracy was reported as 96.00%, whereas we propose 96.90% accuracy.

A research approach [16] was proposed to protect the digital medical images such as X-Rays where water marking technique have been used to secure the medical images. The authors proposed this method to secure the medical images from the unauthorized access and to secure from the forgeries. Since the forgeries and image editing have become easier because of technological revolution. The authors' concluded that around 90% accuracy have been obtained with water marking system whereas we detect the strong set of objects by using the ROIs.

An approach [17] to perform the medical image analysis was proposed where a pixel level intensities were examined by using the image reconstruction based technique at each spatial location of the image. A photon counter detection is really difficult to estimate because each multi-bin based spatial locations are need to estimate with significant method. The authors used post processing method to detect the abnormalities in the image and convolutional neural network based classification method was used to extract the deepest knowledge of each homogenous pixel patterns under the X-Ray based images and reported classification accuracy was around 0.93% mean approximation for strong correlations whereas we propose 96.90% classification accuracy.

An approach [18] to protect the medical images was proposed by incorporating the JEPG ghost technique. The medical images such as X-Rays, CTs, MRIs and many more are need to be marked with visual cryptography to fix a predicate for originality of medical images. The authors proposed forgery detection algorithm by optimizing the digital image processing such as watershed algorithm, segmentation techniques and image reconstruction technique. The reported accuracy is approximated as 92 to 95 percent was proposed for originality of each medical image whereas we detect the forgery positive and negative ISVs with good accuracy.

An approach [19] was proposed to perform the image analysis to detect the abnormalities from the computed tomography images such as X-Rays, CTs and others. the medical image very significant for every human and need to be secure in many senses such as security , untampered and others whereas recent technological advances have made possible to alter such images , the approach has proposed a deep learning based algorithm to classify the abnormalities under the medical images by using deep neural networks whereas .

III. METHOLOGY

This paper offers a framework to classify the forgery positive and native X-Rays based images by using machine learning methods. The framework is comprising over three phases where first phase is used to preparer the X-Ray data (XRD) and at phase two is used to construct the decision model by using SVM and the result of the framework is shown in section three.

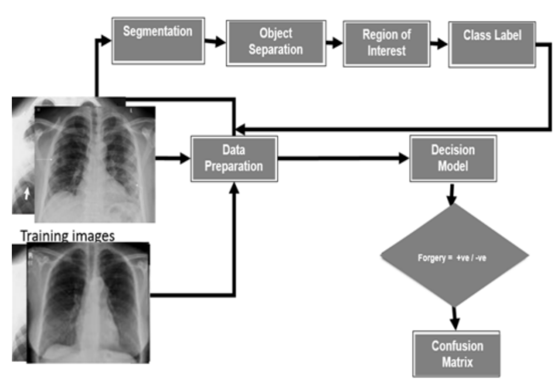

Figure 1: System Architecture for forgery detection for medical images

Dataset: The used dataset was collected from the radiological department of CMCH Larkana. The dataset comprises over 500 XRD images where each image is labelled by the doctors.

A. Phase 1: Data Preparation Layer

In X-Rays data (XRD) the entropy of forgery area be detected by offering significant preprocessing method where object segmentation based data (SBD) , object separated data (OSD) and data acquired through region of interest (ROI) may provide more precise assistance to build the decision models to predict the forgery images. Each XRD is considered as SBD, OSD and ROI by assuming with the notation of $r^1 \leq r^2 \leq x^2 + y^2 \leq (r+\zeta)^2$ and the r term used for radius parameters which lies over the x,y position under the XRD data however they rotated with VGG16 to generate the more models have been used to represent radius [20-22]. There are fair chances to consider special positions so called location x_i, y_i, y i representing the central point of pixels (CPP) as $[x_1, \dots, x_{2r+1}], -r - \zeta \leq x_i \leq r +$ however the size of the each segmented object data (SOD) is required to be collected from the variable r which could be used under the umbrella of the Gaussian function to convolved (GFC) represented by $f_r\ (\vec{x})$ where x ⃗^* could be considered as each point in set of similar pixles to be correlated by the expression such as [23].

$$\left(\Sigma_{(\vec{x})}\left(f_r(\vec{x})\right)^2\right)^{\frac{1}{2}}$$

and each segment of pixels is represented by

$$\bar{I}\vec{x}^* = \left(\Sigma_{x_{\epsilon\Omega}}\left(I_r(\vec{x})\right)^2\right)^{\frac{1}{2}}$$

as likelihood.

$$\mathrm{L}(\vec{x}^*) = \mathrm{I}\,(\vec{x}^*) max_r\{I^\circ f_r(\vec{x}^*)\} = \mathrm{I}\,(\vec{x}^*) max_r\left\{\frac{\Sigma_{\vec{x}} I(\vec{x}) f_r(\vec{x}-\vec{x}^*)}{\bar{I}(\vec{x}^*)\bar{f}_r(\vec{x})}\right\} \quad (1)$$

Since the maximization parameters are being represented in eq. (1) where every corresponded set of pixels is considered for XRDF under the function of $f_r\ (\vec{x})$ and the coordinates are shown as x ⃗^* based on central point estimation. The approximation of weak and strong responses is direly needed to identify the continuity of set of similar pixels where the computational process may consider the shapes and ellipse by keeping the separate back gourd and foreground information under the segmentation techniques and crop the best fit filter operations [24-25].

Our proposed techniques consider the more than two response levels to estimate the intensities of the objects where at first, we record the foreground intensity where the object is detected and visible. at second we record back ground intensities where each object cropping is to considered for automated regions and at the third noisy objects or confusing objects are considered as weak responses as presented in eq.(2) where i ,j and the B are corresponding to second layer scenarios under spatial objects and their row as well as column locations whereas we use pixel level XORAY clusters to consider as weak or strong response of objects in ISV and propose maximum classification accuracy i.e. 96.90%.

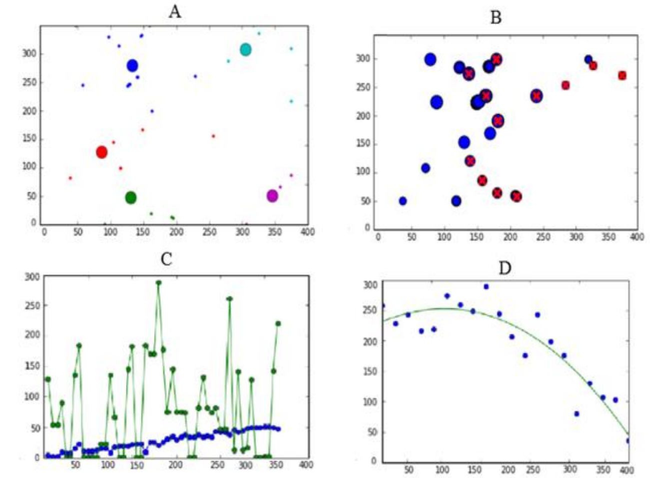

Figure 2: In (A) XORAY clusters are shown (B) weak the weakest candidates are highlighted with red marks (c) A weak as well strong candidates are displayed and (D) Selection of the weak response.

$$A = \sum_{i=1}^{n}\sum_{i=1}^{m} B(i,j) \quad (2)$$

Since the biggest object in X-Ray (BOXRAY) was enquired by using $BOXAY = \sqrt{(x_1 - x_2)^2 (y_1 - y_2)^2}$, where as Shortest object in X-Ray (SOXRAY) was approximated with the use of SOXRAY= $\sqrt{(x_2 - x_1)^2 (y_2 - y_1)^2}$ where the dependent and independent variables were derived through $x_1$and x_2 each re3sidual of decision matrix by considering the acres and the distinguished nodes as passed parameters for every BOXRAY and SOXRAY, where selected regions were $P = evencount + \sqrt{2}\ (oddcount)$ and the areas was could as per eq.(3)

ROI $$\gamma = \frac{A}{p}\frac{4\pi XArea}{p^2} \quad (3)$$

B. Phase 2: Building Of Classification Model

Machine learning based algorithm are very vital tool to construct the decision model. A wide range of supervised and unsupervised methods of artificial intelligence techniques are available. As per nature of problem statement of this research which has to classify the several number of students. A multi- SVM is best suited to recognize the objects. Support vector machine algorithm is intelligent enough to predict various objects by using several number features by considering each feature as vector in the support vector machine network. Hyperplane is the line which is constructed by SVMs to classify the instances. The learning process of SVMs uses class labels to separate the instances by kernel operation such as hyperplane where input data (x) with its cumulative features duly assigned to vectors designated as B0 to Bn. The classification results are presented in next section.

IV. Prepare Your Paper Before Styling

Table 1: Confusion matrix

	Forgery +ve	Forgery -ve
Forgery +ve	612	37
Forgery -ve	34	1611
Measured classification accuracy		96.90%

Table 2: Performance of proposed system

	Forgery +ve	Forgery -ve
Raw images	650	1650
Count of XRD selected using ROIs	649	1645
Observations (Classified)	612	1611
Observations (Non Classified)	37	34
Precession	94.29%	97.93%
Recall	94.73%	97.97%

The confusion matrix [Figure 1] shows that around 612 observation were found for the forgery - Positive X-Rays and approximated 1611 records were classified as forgery-Negative X-Rays. As per detailed results as shown into the [Table2] that around 650 images were belonging to the forgery positive and around 649 images of selected region of interest (ROIs) were chosen for the classification with the class label attribute forgery positive where the 612 images were classified as forgery true positive and approximated 37 images were classified as false negative images. The 94.29% recall measure was estimated for the forgery votive class label attributes whereas around 94.73% estimations were declared as recall estimation for forgery positive class label attribute. The forgery –ve class label consists upon the 1650 X-Rays images where 1645 X-Rays images were selected for classification where the number of 1611 images were estimated as forgery – ve (Forgery free) and misclassification was 34 images, however the 97.93% was estimated as precision measure and the classification 97.75% percent as recall measure and the overall classification accuracy was measured as 96.90% as shown in [Figure 3].

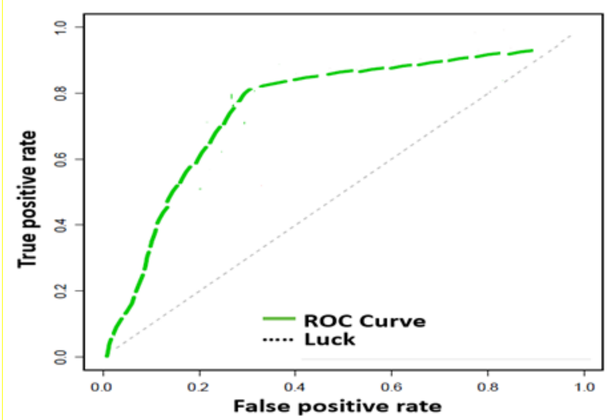

Figure 3: ROC Curve

Conclusion

This proposes a predictive mining techniques to explore the hindsight knowledge of medical datasets such as image, sound and video (ISV). As use case X-Rays images were processed to detect the forgery positive and forgery negative images. Due to recent advances in computational there are faire chances to forgeries such data which could become leading case of misdiagnosis in medical images. Since the malicious intensions (MI) of human or malfunctioning of digital x-rays would include extra noise into the x-ray image which could become of the leading cause of misdiagnosis. However the pixel level image analytics could be used to detect such anomalies. This research offers a framework to detect the forgery in digital x-ray images. The methodology comprises over three stages where the prime objective of data preparation is done at first stage where as the decision model is constructed under the umbrella of build model, meanwhile the third stage is used to visualize the results by using the confusion matrix and other measure`s. The 94.29% recall measure was estimated for the forgery votive class label attributes whereas around 94.73% estimations were declared as recall estimation for forgery positive class label attribute. The forgery –ve class label consists upon the 1650 X-Rays images where 1645 X-Rays images were selected for classification where the number of 1611 images were estimated as forgery – ve (Forgery free) and misclassification was 34 images, however the 97.93% was estimated as precision measure and the classification 97.75% percent as recall measure and contribute highest classification accuracy which is measured as 96.90%.

Acknowledgment

This piece of research work is done under the project, titled "Forgery Detection in Medical Images: A Case of COVID-19 Using Convolutional Neural Network," was conducted at the Institute of Computer Science, Shah Abdul Latif University, Khairpur, Sindh, Pakistan. The project received funding from Sindh HEC.

We would also like to acknowledge the support of the radiological department of CMCH Larkana, Sindh, for providing the dataset for this research. Furthermore, we are indebted to the doctors who assisted in labeling each X-ray image obtained from the department.

REFERENCES

[1] Dixit, A. and Dixit, R., 2022. Forgery detection in medical images with distinguished recognition of original and tampered regions using density-based clustering technique. *Applied Soft Computing*, *130*, p.109652.

[2] Ahmad, M. S., NAWEED, M. S., Waraich, M. M., & Nisa, M. (2017). Texture Analysis Approach to Quantify and Discriminate Normal and Pathological Human Lung CT-Images. Sindh University Research Journal-SURJ (Science Series), 49(2).

[3] Chandio, S., Seman, M. S. A., Koondhar, M. Y., & Shah, A. (2019). Qualitative study on ICT adoption and management practices in public sector universities in Sindh Province, Pakistan. University of Sindh Journal of Information and Communication Technology, 3(3), 149-152.

[4] Ernawan, F., Aminuddin, A., Nincarean, D., Ab Razak, M.F. and Ahmad, F., 2022. Three layer authentications with a spiral block mapping to prove authenticity in medical images. *International Journal of Advanced Computer Science and Applications*, *13*(4).

[5] Chaitra, B. and Reddy, P.V., 2022. Digital image forgery: taxonomy, techniques, and tools–a comprehensive study. *International Journal of System Assurance Engineering and Management*, pp.1-16.

[6] Nikalje, S. and Mane, M.V., 2022, August. Copy-Move and Image Splicing Forgery Detection based on Convolution Neural Network. In *2022 Third International Conference on Intelligent Computing Instrumentation and Control Technologies (ICICICT)* (pp. 391-395). IEEE.

[7] Nataraj, K.R., Harshitha, B.R. and Vidya, S., 2022, November. An Effective Approach for Forgery Detection of Fake COVID-19 Report. In *Proceedings of the International Conference on Cognitive and Intelligent Computing: ICCIC 2021, Volume 1* (pp. 757-773). Singapore: Springer Nature Singapore.

[8] Sadu, C. and Das, P.K., 2022, November. A Detection Method for Copy-Move Forgery Attacks in Digital Images. In *TENCON 2022-2022 IEEE Region 10 Conference (TENCON)* (pp. 1-6). IEEE.

[9] Shashikala, S. and Ravikumar, G.K., 2022. Technology Innovation: Detection of Counterfeit Region in an Image. *ECS Transactions*, *107*(1), p.4517.

[10] Alheeti, K.M.A., Alzahrani, A., Khoshnaw, N. and Al-Dosary, D., 2022, March. Intelligent deep detection method for malicious tampering of cancer imagery. In *2022 7th International Conference on Data Science and Machine Learning Applications (CDMA)* (pp. 25-28). IEEE.

[11] GÜRÜNLÜ, B. and ÖZTÜRK, S., A Novel Method for Forgery Detection on Lung Cancer Images. *International Journal of Information Security Science*, *11*(3), pp.13-20.

[12] Patil, D., Patil, K. and Narawade, V., 2022. A Novel Approach to Image Forgery Detection Techniques in Real World Applications. In *Applications of Artificial Intelligence and Machine Learning: Select Proceedings of ICAAAIML 2021* (pp. 461-473). Singapore: Springer Nature Singapore.

[13] Kaur, S. and Rani, R., 2022. Image Forgery Detection Using Multi-Layer Convolutional Neural Network. In *Advanced Machine Intelligence and Signal Processing* (pp. 855-866). Singapore: Springer Nature Singapore.

[14] Khan, M. J., Yousaf, A., Abbas, A., & Khurshid, K. (2018). Deep learning for automated forgery detection in hyperspectral document images. Journal of Electronic Imaging, 27(5), 053001-053001.

[15] Ghoneim, A., Muhammad, G., Amin, S. U., & Gupta, B. (2018). Medical image forgery detection for smart healthcare. IEEE Communications Magazine, 56(4), 33-37.

[16] Ghai, A., Kumar, P., & Gupta, S. (2024). A deep-learning-based image forgery detection framework for controlling the spread of misinformation. Information Technology & People, 37(2), 966-997.

[17] Agarwal, R., & Verma, O. P. (2020). An efficient copy move forgery detection using deep learning feature extraction and matching algorithm. Multimedia Tools and Applications, 79(11), 7355-7376.

[18] Gill, S. H., Sheikh, N. A., Rajpar, S., Jhanjhi, N. Z., Ahmad, M., Razzaq, M. A., ... & Jaafar, F. (2021). Extended Forgery Detection Framework for COVID-19 Medical Data Using Convolutional Neural Network. Computers, Materials & Continua, 68(3).

[19] Hosny, K. M., Mortda, A. M., Lashin, N. A., & Fouda, M. M. (2023). A new method to detect splicing image forgery using convolutional neural network. Applied Sciences, 13(3), 1272.

[20] Xiao, B., Wei, Y., Bi, X., Li, W., & Ma, J. (2020). Image splicing forgery detection combining coarse to refined convolutional neural network and adaptive clustering. Information Sciences, 511, 172-191.

[21] Muniappan, T., Abd Warif, N. B., Ismail, A., & Abir, N. A. M. (2023). An Evaluation of Convolutional Neural Network (CNN) Model for Copy-Move and Splicing Forgery Detection. International Journal of Intelligent Systems and Applications in Engineering, 11(2), 730-740.

[22] Bayar, B., & Stamm, M. C. (2016, June). A deep learning approach to universal image manipulation detection using a new convolutional layer. In Proceedings of the 4th ACM workshop on information hiding and multimedia security (pp. 5-10).

[23] Rathore, N. K., Jain, N. K., Shukla, P. K., Rawat, U., & Dubey, R. (2021). Image forgery detection using singular value decomposition with some attacks. National Academy Science Letters, 44(4), 331-338.

[24] Heidari, A., Jafari Navimipour, N., Dag, H., & Unal, M. (2024). Deepfake detection using deep learning methods: A systematic and comprehensive review. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 14(2), e1520.

[25] Yang, J., Li, A., Xiao, S., Lu, W., & Gao, X. (2021). MTD-Net: Learning to detect deepfakes images by multi-scale texture difference. IEEE Transactions on Information Forensics and Security, 16, 4234-4245.