ENGINEERING IN ADVANCED RESEARCHSCIENCE AND TECHNOLOGY

ISSN 2352-8648
Vol.03, Issue.01
May-2022
Pages: -266-272

A SURVEY ON IMAGE FORGERY DETECTION USING CONVOLUTION NEURAL NETWORK

[1]POLNATI KALYANI, [2]S N V S S S T MURTHY
[1] M. Tech., Dept. of CSE, Prasiddha College of Engineering Technology, AMALAPURAM,. AP, India.
[2]Associate Professor, Dept. of CSE, Prasiddha College of Engineering Technology, AP, India

Abstract:- The means of socialism and connectedness, nowadays, is achieved by uploading their images on social media. The availability of personal pictures on social media has also made readily available for other people, and they can manipulate those private pictures. Recent advances in media generation techniques have made it easier for attackers to create forged images. So, the concept image forensic come into the picture to check the authenticity of these images. Image forgery detection using traditional algorithms takes much time to find forgeries. The new emerging methods for detection of image forgery use a deep neural network algorithm. A hybrid deep learning and machine learning based approach is used in this study for passive image forgery detection. This proposed model classifies the images in to forged and not forged categories using convolution neural network algorithm.

Keywords- Image Tampering, Block-based approaches, Key points based approaches, Deep Learning

INTRODUCTION:-

Nowadays, images play a vital role in several fields like medical, education, digital forensics, sports, scientific research, news media, etc. and they are used as one of the main sources of information. Due to software like Photoshop,GIMP, Coral Draw and android applications like photo hacker, it is very easy to create a forged image. The genuineness of image becomes very crucial in the cases where the image is used as a proof in court of law. Image manipulation is any type of operation that is performed on digital images by using any software; it is also referred as image editing. Image forgery is the technique to modify the content of an image which contradicts with some

fact happened in past. Image tampering is a type of image forgery which replaces some content of an image with new content. If the new content is copied from the same image itself then it is called copy-move tampering and if the new content is copied from different image then it is called image Splicing. The image manipulation detection approaches can be classified into two: (i) Active and (ii) Passive. In active approach, additional information (such as digital watermark) is embedded in the image during the image acquisition stage or at some later stage by authorized person. The active approach uses this embedded information for manipulation detection. The passive approaches do not depend on the additional information for forgery detection. These approaches are also referred as “blind approaches” as the approaches do not use additional information for forgery detection. The passive approaches extract the features from the image and use these features for forgery detection. The passive approaches can be classified he forgery type independent approaches aim to detect other forgeries such as compression and re sampling. Passive forgery type dependent approaches for forgery detection can be categorized into: (i) copy-move and (ii) splicing. The images which are manipulated by these two Approaches are very hard to identify by human. Therefore, it becomes very essential to detect these two kinds of forgeries and also it will be useful for digital image forensics. Unlike active image forgery approaches, e.g., watermarking, passive techniques are more useful, but they are more challenging. Wang et al. [1] used the tampering clues for classification of passive approaches for tampering detection. Generally digital forgery doesn't leave any visual clues of what is tampered with but it may change some statistics of an image and based on this belief these techniques work on an image. Some copymove detection approaches are evaluated by Christlein et al.[2]and some cut-paste detection techniques are evaluated by Zambpglou et al. [3]. In context of this fact, different image tampering detection approaches have been suggested by researchers.

Related Work:-

In deep learning, convolutional neural networks (CNN) are a type of deep neural networks applied to classify the images, image & video analysis which consists of a Input layer, hidden layers and output layer. Layers present in between the input and output layers are called hidden because the activation function and final convolution hide their inputs .CNN pre-trained model called VGG-16 (Visual Geometry Graph) is a 16 layers deep with 13 convolutional layers, 2-fully connected and 1 softmax used for classification and detection.VGG architecture is given in Fig.1.The VGG-16 architecture consists of blocks, each block consists of 2D convolution layer and Max pooling layers.VGG a ConvNet accepts only 224*224 RGB image. Preprocessing layer takes the RGB image with pixel values in the range of 0-225 and subtracts the mean image which is calculated over the entire Image Net training set.VGG16 has a total of 138 million parameters. It's worth noting that all of the conv kernels are 3x3, while the maxpool kernels are 2x2 with a stride of two. Pupillometry dataset consists of four classes. The class labels are no disease, mild, moderate, severe, extremely severe. Each class consists of up to hundreds of images. Existing system is made up of two separate Support Vector Machines each represents the left eye and right eye of infant's images. The outputs of both the classifiers are combined by an OR logical operator to predict the final classification result. The limitations are choosing the right kernel is necessary for svm, more computational power is required to detect the disease images. This model is not supported for checking more and more images and faces more ambiguity. In our system, we used neural networks consists of 16 layers where the input layer accepts the image only in RGB format with pixel values in the range of 0-255, data pre-processing was done to resize the image and rescaling to transform the every pixel value. Images from the training dataset are fed to the CNN model i.e.VGG-16.In output layer activation function used is softmax function which predicts the final classification result from the both eye models. This process was done two times separately for the both left and right eye models. As we used neural networks, they are more powerful than machine learning algorithms have less computational power, and no ambiguity in detecting the retinal disease and achieved correct results when more and more images added and tested other than the images from pupillometry dataset achieved an accuracy of 90.07% .

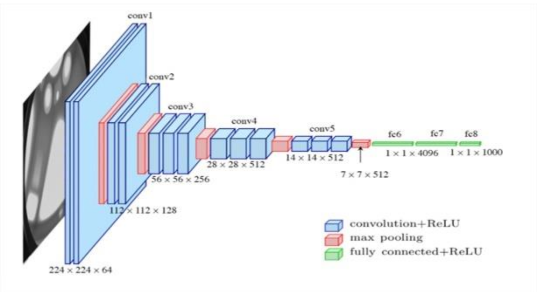


Fig.1.VGG-16 Architecture

Literature Survey:-

Object detection with discriminatively trained part-based models

Authors:- P. F. Felzenszwalb, R. B. Girshick, D. Mcallester, and D. Ramanan

We describe an object detection system based on mixtures of multi scale deformable part models. Our system is able to represent highly variable object classes and achieves state-of-the-art results in the PASCAL object detection challenges. While deformable part models have become quite popular, their value had not been demonstrated on difficult benchmarks such as the PASCAL data sets. Our system relies on new methods for discriminative training with partially labeled data. We combine a margin-sensitive approach for data-mining hard negative examples with a formalism we call latent SVM. A latent SVM is a reformulation of MI--SVM in terms of latent variables. A latent SVM is semi convex, and the training problem becomes convex once latent information is specified for the positive examples. This leads to an iterative training algorithm that alternates between fixing latent values for positive examples and optimizing the latent SVM objective function.

Example-based learning for view-based human face detection,

Authors:- K. K. Sung and T. Poggio

We present an example-based learning approach for locating vertical frontal views of human faces in complex scenes. The technique models the distribution of human face patterns by means of a few view-based "face" and "no face" model clusters. At each image location, a difference feature vector is computed between the local image pattern and the distribution-based model. A trained classifier determines, based on the difference feature vector measurements, whether or not a human face exists at the current image location. We show empirically that the distance metric we adopt for computing difference feature vectors, and the "no face" clusters we include in our distribution-based model, are both critical for the success of our system.

Pedestrian detection: An evaluation of the state of the art

Authors:- C. Wojek, P. Dollar, B. Schiele, and P. Perona

Pedestrian detection is a key problem in computer vision, with several applications that have the potential to positively impact quality of life. In recent years, the number of approaches to detecting pedestrians in monocular images has grown steadily. However, multiple data sets and widely varying evaluation protocols are used, making direct comparisons difficult. To address these shortcomings, we perform an extensive evaluation of the state of the art in a unified framework. We make three primary contributions: 1) We put together a large, well-annotated, and realistic monocular pedestrian detection data set and study the statistics of the size, position, and occlusion patterns of pedestrians in urban scenes, 2) we propose a refined per-frame evaluation methodology that allows us to carry out probing and informative comparisons, including measuring performance in relation to scale and occlusion, and 3) we evaluate the performance of sixteen pretrained state-of-the-art detectors across six data sets. Our study allows us to assess the state of the art and provides a framework for gauging future efforts. Our experiments show that despite significant progress, performance still has much room for improvement. In particular, detection is disappointing at low resolutions and for partially occluded pedestrians.

Detection of spicules on mammogram based on skeleton analysis

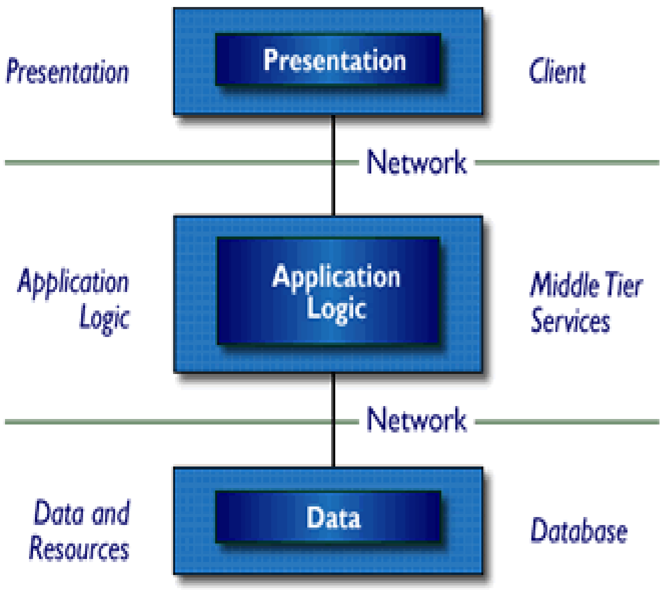
Authors:- H. Kobatake and Y. Yoshinaga

Existence of spicules is one of important clues of malignant tumors. This paper presents a new image processing method for the detection of spicules on mammogram. Spicules can be recognized as line patterns radiating from the center of tumor. To detect such characteristic patterns, line skeletons and a modified Hough transform are proposed. Line skeleton processing is effective in enhancing spinal axes of spicules and in reducing the other skeletons. The modified Hough transform is applied to line skeletons and radiating line structures are obtained. Experiments were made to test the performance of the proposed method. The system was designed using 19 training images, for which one normal case was recognized to be star-shaped. The other case were recognized correctly. Other experiments using 34 test images were also performed. The correct classification rate was 74%. These results shows the effectiveness of the proposed method.

SYSTEM ARCHITECTURE

The current application is being developed by taking the 3-tier architecture as a prototype. The 3-tier architecture is the most common approach used for web applications today. In the typical example of this model, the web browser acts as the client, IIS handles the business logic, and a separate tier MS-SQL Server handles database functions.

Although the 3-tier approach increases scalability and introduces a separation of business logic from the display and database layers, it does not truly separate the application into specialized, functional layers. For prototype or simple web applications, the 3-tier architecture may be sufficient. However, with complex demands placed on web applications, a 3-tiered approach falls short in several key areas, including flexibility and scalability. These shortcomings occur mainly because the business logic tier is still too broad- it has too many functions grouped into one tier that could be separated out into a finer grained model.



Three-Tier Architecture

Tier 1: the client contains the presentation logic, including simple control and user input validation. This application is also known as a thin client. The client interface is developed using ASP.Net Server Controls and HTML controls in some occasions

Tier 2: the middle tier is also known as the application server, which provides the business processes logic and the data access. The business logic/ business rules can be written either with C#.Net or VB.Net languages. These business runes will be deployed as DLL's in IIS web server.

Tier 3: the data server provides the business data. MS-SQL server acts as Tier-3, which is the database layer.

These are some of the advantages of three-tier architecture:

It is easier to modify or replace any tier without affecting the other tiers.

Separating the application and database functionality means better load balancing.

Adequate security policies can be enforced within the server tiers without hindering the clients.

The proposed system can be designed perfectly with the three tier model, as all layers are perfectly getting set as part of the project. In the future, while expanding the system, in order to implement integration touch points and to provide enhanced user interfaces, the n-tier architecture can be used.

Conclusion:-

This paper proposed a image forgery detection approach using CNN based pre-trained Alex Net model to extract deep features, without investing much time in training. The proposed approach also exploits the SVM as a classifier. Compared to the previous work on MICC-F220 dataset, the best accuracy of image forgery detection achieved is 93.94%. In this paper, MICC-F220 dataset en comprising of 220 images of forged and non-forged images are classified using SVM Classifier. Performance of the deep features extracted from a pre-trained Alex Net based model is quite satisfactory, even in the presence of rotational and geometrical transformation and also compared the results of the given approach with the existing state-of-the-art approaches. In the future, we plan to work on various benchmark image forgery datasets and to compare the performance with the existing approaches.

REFERENCES

[1] P. F. Felzenszwalb, R. B. Girshick, D. Mcallester, and D. Ramanan, "Object detection with discriminatively trained part-based models," IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 9, p. 1627, 2010.

[2] K. K. Sung and T. Poggio, "Example-based learning for view-based human face detection," IEEE Trans. Pattern Anal. Mach. Intell., vol. 20, no. 1, pp. 39–51, 2002.

[3] C. Wojek, P. Dollar, B. Schiele, and P. Perona, "Pedestrian detection: An evaluation of the state of the art," IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, no. 4, p. 743, 2012.

[4] H. Kobatake and Y. Yoshinaga, "Detection of spicules on mammogram based on skeleton analysis." IEEE Trans. Med. Imag., vol. 15, no. 3, pp. 235–245, 1996.

[5] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," in ACM MM, 2014.

[6] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in NIPS, 2012.

[7] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime multi-person 2d pose estimation using part affinity fields," in CVPR, 2017.

[8] Z. Yang and R. Nevatia, "A multi-scale cascade fully convolutional network face detector," in ICPR, 2016.

[9] C. Chen, A. Seff, A. L. Kornhauser, and J. Xiao, "Deepdriving: Learning affordance for direct perception in autonomous driving," in ICCV, 2015.

[10] X. Chen, H. Ma, J. Wan, B. Li, and T. Xia, "Multi-view 3d object detection network for autonomous driving," in CVPR, 2017.

[11] A. Dundar, J. Jin, B. Martini, and E. Culurciello, "Embedded streaming deep neural networks accelerator with applications," IEEE Trans. Neural Netw. & Learning Syst., vol. 28, no. 7, pp. 1572–1583, 2017.

[12] R. J. Cintra, S. Duffner, C. Garcia, and A. Leite, "Low-complexity approximate convolutional neural networks," IEEE Trans. Neural Netw. & Learning Syst., vol. PP, no. 99, pp. 1–12, 2018.

[13] S. H. Khan, M. Hayat, M. Bennamoun, F. A. Sohel, and R. Togneri, "Cost-sensitive learning of deep feature representations from imbalanced data." IEEE Trans. Neural Netw. & Learning Syst., vol. PP, no. 99, pp. 1–15, 2017.

[14] A. Stuhlsatz, J. Lippel, and T. Zielke, "Feature extraction with deep neural networks by a generalized discriminant analysis." IEEE Trans. Neural Netw. & Learning Syst., vol. 23, no. 4, pp. 596–608, 2012.

[15] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in CVPR, 2014.

[16] R. Girshick, "Fast r-cnn," in ICCV, 2015.

[17] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in CVPR, 2016.

[18] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards realtime object detection with region proposal networks," in NIPS, 2015, pp. 91–99.

[19] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," Int. J. of Comput. Vision, vol. 60, no. 2, pp. 91–110, 2004.

[20] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in CVPR, 2005.

[21] R. Lienhart and J. Maydt, "An extended set of haar-like features for rapid object detection," in ICIP, 2002.

[22] C. Cortes and V. Vapnik, "Support vector machine," Machine Learning, vol. 20, no. 3, pp. 273–297, 1995.

[23] Y. Freund and R. E. Schapire, "A desicion-theoretic generalization of on-line learning and an application to boosting," J. of Comput. & Sys. Sci., vol. 13, no. 5, pp. 663–671, 1997.

[24] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, “Object detection with discriminatively trained part-based models,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, pp. 1627–1645, 2010.

[25] M. Everingham, L. Van Gool, C. K. Williams, J. Winn, and A. Zisserman, “The pascal visual object classes challenge 2007 (voc 2007) results (2007),” 2008.

[26] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[27] N. Liu, J. Han, D. Zhang, S. Wen, and T. Liu, “Predicting eye fixations using convolutional neural networks,” in CVPR, 2015.

[28] E. Vig, M. Dorr, and D. Cox, “Large-scale optimization of hierarchical features for saliency prediction in natural images,” in CVPR, 2014.

[29] H. Jiang and E. Learned-Miller, “Face detection with the faster r-cnn,” in FG, 2017. [30] D. Chen, S. Ren, Y. Wei, X. Cao, and J. Sun, “Joint cascade face detection and alignment,” in ECCV, 2014..