International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

Watermark Detection Using Deep Learning

[1]S Sushma, [2]Ch Sai Sindhu, [3]K Leela Sowmya, [4]N Nikith Raj Kumar, [5]T Surya

[1]Assistant Professor, Department of Information Technology,Aditya Engineering College (A), Surampalem, Andhra Pradesh, India.
Sushma.sunkara@aec.edu.in
[2, 3, 4, 5]Student, Department of Information Technology, Aditya Engineering College (A), Surampalem, Andhra Pradesh, India.
18A91A1252@aec.edu.in, 18A91A1230@aec.edu.in, 18A91A1240@aec.edu.in, 18A91A1255@aec.edu.in

Abstract
The main aim of our project "Watermark Detection using Deep learning" is to detect the watermark in an image and display a bounding box around the watermark as a result. Watermarks are used to claim the ownership. Watermarks can be of any type, maybe text, symbols, etc. In our approach, we used CNN algorithm for image classification and YOLO algorithm for watermark detection.

Keywords: Convolutional neural networks, You Only Look Once.

Introduction

This project is "Watermark detection using Deep learning". Watermarks are used to cover content and to claim power of an asset. Without watermarks, precious digital means can be susceptible to content theft or unauthorized use. Watermarks can be a type of text, patterns, and any symbols. Watermarks has a quality of robustness. They can repel with any kind of attacks like cropping, rotation, etc. For the purpose of removal of watermarks, we need to detect them.

K. Zebicche [1] et al., presented an approach that is efficient for watermarking that resists against attacks like cropping segmentation attacks inherently robust, and yields imperceptible watermarks. This technique has been applied on discrete wavelet transform (DWT) and discrete cosine transform (DCT). Based on the ML scheme and the concerted of the two estates, the optimum decoder has been modelled by the generalized Gaussian distribution. Laith Alzubaidi

[2] et al., proposed a title for his paper as reviews of deep learning concepts like CNN architectures, its challenges, its application, and its future directions. In the field of machine learning, deep learning became the most widely used approach. It attained an outstanding result on various tasks like cognitive work. Further, he presented convolutional neural networks (CNN) and described its architecture and their main features. Shrey Srivastava [3] et al., made a comparison on 3 major algorithms in image processing. They are faster region based convolutional neural networks (Faster R-CNN), single shot detection (SSD), You Only Look Once (YOLO). His aim is to get the most efficient and most efficient of those three algorithms. The evaluation of performance for these three algorithms is made by their strengths based on parameters such as

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

accuracy, F1 score, and precision limitations are analysed. Finally, the result is YOLO-v3 performs best out of three algorithms taken. Akram Zeki [4] et al., uses an ISB bit watermarking approach. He considered a point to find the two parameters threshold units in getting the amount of strengths and weaknesses for the watermarking techniques. The issue is related to the normalized cross correlation (NCC) and peak signal to noise ratio (PSNR). These were functions of watermarking approach. He utilized their new watermarking approaches for adding 4 watermarks in intermediate significant bits for 6 image files one on one by changing the present image pixelswith the new pixels and, on the similar hand, new pixels were kept closer to old pixels. Here the updated approach is robust against image processing operations such as compression, blurring, filtering, and noise, which alter the intensity of the pixels instead of their locations.

Existing model:
In the existing model, Multilayer perceptron (MLP) is used. It's a feedforward neural network. A multilayer perceptron contains of three layers of nodes at least. Those are input layer, hidden layer and output layer. Grounded of the error calculations for different confirmation set, the inheritable algorithm selects the stylish watermarking intensity threshold.

Disadvantages of existing model
1. The perpetration to reach the real- time response is delicate.
2. MLP includes too numerous parameters, because it's completely connected, performing in redundancy and inefficiency.

Proposed model:
In utmost of the image processing operations, CNN outperforms state of art results. CNN is specialized in image processing. It contains a grid-suchlike configuration. A digital image can be visualized as two dimensional data.
YOLO algorithm is used to detect objects (in real-time). This algorithm recognizes colorful objects in a picture. This algorithm propagates forward for single time through neural networks for detecting objects.

Advantages of proposed model:
1. Without any human supervision, CNN can automatically detect the important features.
2. CNN is also computationally efficient.
3. CNN model can be able to run on any kind of device.
4. This makes them widely used, since pooling operations and special convolution are used and parameter sharing is performed.
5. YOLO is best known for its accuracy, speed, and learning capabilities.

Literature study
Frederic Cerou [5] et a., introduced this paper. According to this paper they presented the algorithm for estimating the probabilities for rare events. The algorithm which is used is much faster and also more accurate than Monte Carlo estimator. And the algorithm is suitable for watermarking problems because errors are related to different kinds of rare events. Hence, this algorithm is much more useful for estimating the probability of false or error in Zero-bit watermarking.
Pedro Comesana [6] et al., introduced this paper. By this paper we gained the difference between advanced editing tools and Digital Watermarking. Advanced editing tools are used in olden day and also used by non-trained users whereas digital watermarking is used by trained people and it is very secure. In this paper they used some attacks for breaking the bows watermarking system are sensitivity attacks and key guessing attacks. By comparing the attacks, the outputs show the good efficiency for black box detector.

Ning Chen [7] introduced this paper. He overcome the problem by strong water marking algorithm for audio is called as zero watermarking algorithm. In zero-marking algorithm is efficient for audio

signal. In zero-watermarking is efficient for the audio signal on protection come the same host of audio into the multiple embedding of watermarking and key are not itself for host. The result in three ways. There are original host, original watermark, Extracted watermarking without the attacked.

Renjie Zhu [8] et at., described about the safe neural networks watermarking scheme which opposes the attacks like forgery. The property by neural network with possible of trigger set in that we need check out the samples of forged trigger and forged labels are the attacker to allowed into the inner mechanism. And the proposed the trigger is watermarking in opposition to forging attack and we cannot replace the owner's watermarking.

Christian Rey [9] et al., researched on the analysis for the authentication of image of the algorithms of watermarking. It's a multimedia document traffic and authentication is content of the opposed to the duplication and specifically images, and also affected. The algorithm is flexibility in the process of watermarking. And using this model dependent to allows the content on the image. They examine is authentication are used they error-correcting codes.

Mauro Barni [10] et al., proposed by robust watermarking of ccartographic images. In this using technique is TBGN-based (Text-Based Geometric Normalization) and it's performing the cropping attack of image. The algorithm didn't need to snap the text containing part just robustness of text to be presented in digital map. The context to be used the cartographic image watermarking, and it's exiting in different part image.

Srdjan Stankovic [11] et al., proposed an analysis of time frequency and their applications on digital watermarking techniques. This algorithm used for digital image, digital audio and video is presented. The lead of the time-frequency is mainly DCT and Fourier and signal domain. In this proposition of multidimensional formation and apply on videos of digital one, audio of digital model, and images of digital ones in watermark. The theory of exiting watermarking procedure in neither Fourier nor DCT domain.

Athanasios Nikolaidis [12] et al., discussed about the provincial exaggeration contrary watermarking schemes of images that relies on feature extraction. They are opposed for being typical local distortion attacks and with the ahead in position of various approaches. The usage on multibit watermarking technique would be formation of DCT domain, SIFT-based version of watermarking. The signal processing attacks to the against in this algorithm.

Stefano Giovanni Rizzo [13] et al., described a watermarking scheme of fine-grain for intellectual property security. In this watermarking generated used by the hash function, original text and watermark to the secret key. The capacity of embedding in watermarking is 0.632 bits and characters. We can replacement with the symbols of neither unicode symbols very similar nor identical to common of watermark. And we shouldn't allow to the copy and paste using this algorithm with the techniques of fine-grain protection of text watermark.

Chun-Hsien Chou [14] et al., described a scheme of watermarking on color image, which satisfies those essentials of robustness, oblivious detection, and transparency. Through the alterations in quantization indices and perceptually lossless color quantization, watermark transparency is attained. The proposed scheme of watermarking describes its simplicity in the implementation and in its computation.

Tiziano Bianchi [15] et al., came with a solution for fingerprint collusions. In this the server randomly collect the fingerprints of the clients and encodes them using a different projection matrix. Another solution is tardos code has been assigned to each client in safe manner. By

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

these two solutions comparing and averaging they conform that which solution is working very accurate in accessing the fingerprints without fingerprint collusions.

Methods Used

Convolutional Neural Networks(CNN)

Convolutional Neural Network (CNN) is a neural network type that specialises in processing of data which is like a grid-type architecture, comparable to that of an image. A digital image is made up of a grid-like arrangement of pixels with pixel values indicating how bright and what hue each pixel should be.

Three layers are common in convolutional neural networks:

a completely connected subcaste, a pooling subcaste, and a convolutional subcaste.

Working of CNN

They give volume-based advice and employ multi-channeled visuals. CNNs cannot celebrate flat images that just contain range and height, which humans can see. CNNs blend the three colours to form the colour diapason humans experience since digital colour images have red-blue-green (RGB) garbling.

The convolutional layer, which is the fundamental structure block and conducts the most of the computational hard lifting, is the first subcaste in a CNN network. Pollutants or the kernels are being utilized for convolving images or data. Pollutants were minor units which are used to apply the sliding window for the data. The depth of image would be similar to input. For an instance, if the value of RGB's depth is 4, the sludge of depth 4 has to be applied for the image. This method brings the taking of the element on element product for the contaminants in an image and casting these precise units for each of the sliding movement. The two dimensional matrix could be an affair for the difficulty with the three dimensional sludge with colour.

The activation layer, which uses the ReLu (Remedied Linear Unit), is an alternative. In this stage, we use the treatment function to increase CNN non-linearity. Images are made up of disparate objects that are not directly related to one another.

Third, is the pooling Subcaste, which involves down slice of features. It's applied through every subcaste in the 3d volume. Generally, there are hyperparameters within this subcaste.

a. The dimension of spatial extent, which is the value of n, from which we can create a single value by combining N cross and point representations and charts.

b. Stride refers to how many features the sliding window skips throughout the range, while height refers to how bright and colourful each pixel should be. A typical pooling Subcaste employs a non-overlapping 2 cross 2 maximum sludge with a stride of 2. A maximum sludge would yield the highest value in the region's features.

The Completely linked Subcaste, which involves Flattening, is in its early stages. The complete pooled point chart matrix is transubstantiated into a single column, which is then supplied to the neural network for processing. We put these attributes together to make a model using the entirely connected layers. To classify the situation, we'll use an activation function similar to softmax or sigmoid.

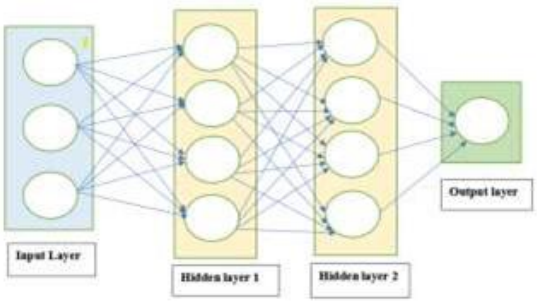


Fig 3.1.1.1: CNN layers

CNN Architecture in Watermark Detection

The proposed architecture has 16 layers including input layer, 5 convolution layers, each convolution layer has Exponential Linear Unit(elu) as activation layer, and connected to MaxPolling Layer. After Flattening layer used 2 drop out

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

layer of value 0.3 and 0.5 respectively. Fig 3.1.2 gives the summary of each layer of model.

Data:
1. No of Training data set: 6121 images
2. No of Validation data set: 681 images
Cost Function: binary_crossentropy
Optimizer: Adam with default learning rate of 0.001 Epocs: 20

Callbacks:
1. Model Check point: Save the model file if the val_loss parameter improved.
2. Early Stopping: Stop the training if there is no improvement in val_loss by delta of 0.001 for a patience of 4 epocs.

Achieved Accuracy at Epoc #14:

loss	acc	val_loss	val_acc
0.177200	0.932048	0.214623	0.914706

You Only Look Once (YOLO)
YOLO is one of the object detection algorithms. It is a feedforwarding neural network algorithm. Hence, its name is abbreviated as You Only Look Once. That is, this algorithm follows only a single forward propagation. In a single run, the prediction in the entire image will be done.

YOLO in Watermark Detection
• Used pre-trained model – yolov5s
• Batch size – 32
• No. of Epochs – 100
• No. of classes – 1

Label annotations:

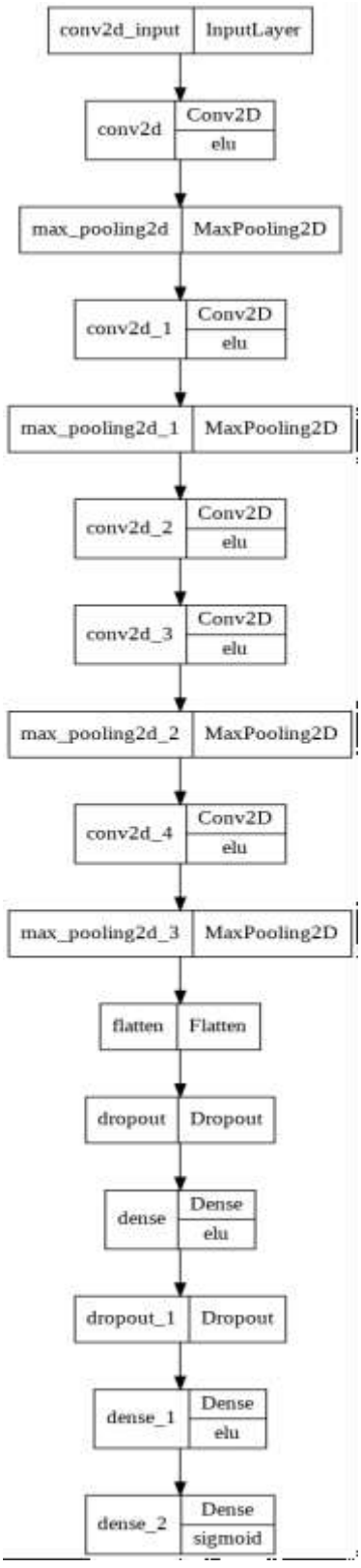


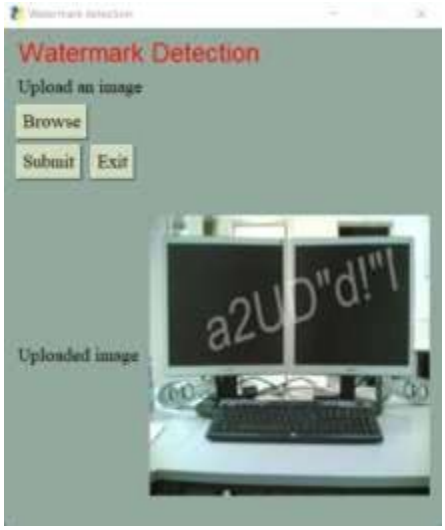
Fig 3.1.2 Model Architecture plot

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

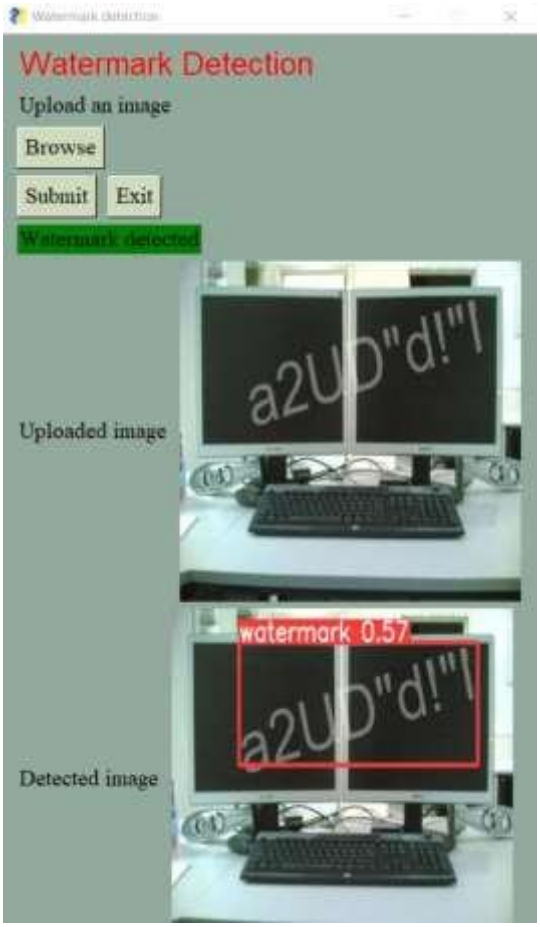
Fig 3.2.1.1: annotations

Screenshots

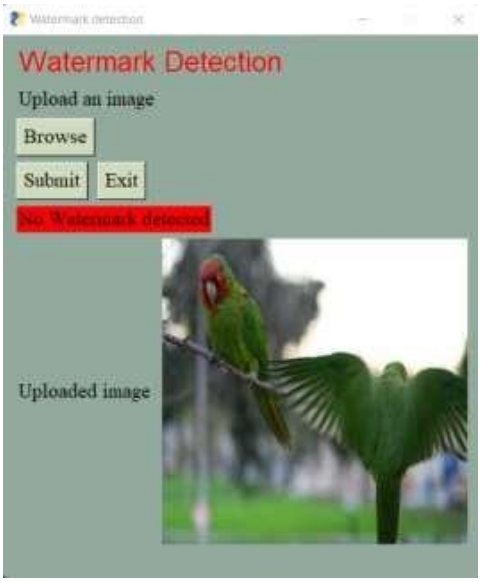
Screen 4.1: Screenshot of application



Screen 4.2 Screenshot after uploading an image



Screen 4.3 Screenshot of detected image with bounding box across watermark.



Screen 4.4 Screenshot of non-watermark image.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

Conclusion

We present a method for detecting watermarks in images using a CNN classifier and YOLO detection in this project. In the previous paradigm, achieving a real-time response is delicate, and it also leads in redundancy and inefficiency. CNN outperforms advanced outcomes for nearly on all image processing procedures. YOLO algorithm is an algorithm which recognizes and detects colored items in a photograph. It's noted for its delicacy, quickness, and learning ability. Our project has been successfully completed with 91.4% accuracy. In conclusion, we believe that the method proposed in this project has a potential to detect the watermark in the chosen image efficiently.

Future Scope

The issue of eliminating watermarks begins with detecting them, hence this could be the future scope. Any watermark removal algorithm would most likely confuse text in photos with watermark if there was no solid medium to evaluate whether or not an image had a watermark. As a result, properly and effectively separating photos with a watermark is critical. This is accomplished via identifying the image's watermark.

References

1. Zebbiche, K., Khelifi, F., & Bouridane, A. (2008). An efficient watermarking technique for the protection of fingerprint images. EURASIP journal on information security, 2008, 1-20.
2. Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., ... & Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. Journal of big Data, 8(1), 1-74.
3. Shrey, S., Divekar, A. V., Chandu, A., Ishika, N., Ved, K., & Pattabiraman, V. (2021). Comparative analysis of deep learning image detection algorithms. Journal of Big Data, 8(1).
4. Zeki, A., Abubakar, A., & Chiroma, H. (2016). An intermediate significant bit (ISB) watermarking technique using neural networks. SpringerPlus, 5(1), 1-25.C.
5. Cérou, F., Furon, T., & Guyader, A. (2008). Experimental assessment of the reliability for watermarking and fingerprinting schemes. EURASIP Journal on Information Security, 2008(1), 414962.B. Boashash, Time-Frequency Analysis and Processing, Elsevier,Amsterdam, The Netherlands, 2003.
6. Comesana, P., & Pérez-González, F. (2007). Breaking the BOWS watermarking system: key guessing and sensitivity attacks. EURASIP Journal on Information Security, 2007, 1-8.
7. Chen, N., & Zhu, J. (2007). A robust zero-watermarking algorithm for audio. EURASIP Journal on Advances in Signal Processing, 2008, 1-7.
8. Zhu, R., Zhang, X., Shi, M., & Tang, Z. (2020). Secure neural network watermarking protocol against forging attack. EURASIP Journal on Image and Video Processing, 2020(1), 1-12.
9. Rey, C., & Dugelay, J. L. (2002). A survey of watermarking algorithms for image authentication. EURASIP Journal on Advances in Signal Processing, 2002(6), 1-9.
10. Barni, M., Bartolini, F., Piva, A., & Salucco, F. (2002). Robust watermarking of cartographic images. EURASIP Journal on Advances in Signal Processing, 2002(2), 1-12.