An Object Detection and Classification Model for Crime Evidence Analysis Using YOLO

Conference Paper · May 2020

3 authors:

Francisca Onaolapo Oladipo
Thomas Adewumi University

Emeka Ogbuju
Federal University Lokoja

Femi Alayesanmi
Federal University Lokoja

Proceedings on Big Data Analytics
& Innovation (Peer-Reviewed),
Volume 3, 2018, pp.18-26

An Object Detection and Classification Model for Crime Evidence Analysis Using YOLO

ALAYESANMI FEMI SAMSON, FRANCISCA OLADIPO &
EMEKA OGBUJU
Federal University Lokoja, Nigeria

ABSTRACT Object detection is a key aspect of digital forensics of visual-based evidence from video surveillance systems and forensic photographs. The process of digital forensics can be very difficult and may require highly technical analysis of voluminous contents during a forensic investigation as every image and video collected for evidence from a particular crime scene provides a concrete visual documentation of the crime scene. In this study, an object detection model based on *You Only Look Once* (YOLO) Convolutional Neural Network (CNN) architecture was developed to detect objects at a crime scene without human involvement or any external control. The aim is to detect objects at a crime scene without human involvement or any external control thereby introducing an optimized approach for crime evidence analysis. Using cross-industry process for data mining, the CNN was trained on a dataset of 5 classes of objects with 1, 173 custom images common to indoor crime scenes. The result of the system gave an average accuracy of 67.84% at 0.013 confidence thresholds after training on an Intel(R) Core (TM) i72720QM CPU with a processor speed of 2.20GHz. The model was deployed on an Android-based forensic case documentation mobile application to review its effectiveness in the problem domain in real time. The study concludes that a crime scene evidence analysis is possible through an optimized approach as it provides a second-eye reviewer to forensic personnel.

Keywords: YOLO, Object detection, Neural network architecture, Image processing, Classification, Crime evidence

1. Introduction

Object detection is recognizing an instance of object classes over a wide range of image data using computational techniques or the nude eye. Object detection and image processing have been a frequent research over the years due to their numerous practical applications. Joseph *et al.* (2017) developed an object detection model with a novel convolutional neural network called YOLO (You Only Look Once) which is a state-of-the-art approach for real-time object detection in digital images. The architecture of the model involves a single convolutional neural network to classify boundary boxes of objects in a resized input image of 448 x 448. Likewise, Rayson *et al.* (2017) developed a real-time end to end automatic license plate recognition system in support of the state-of-the-art YOLO convolutional neural network algorithm for object detection. Although Rayson *et al.* (2017) did not create the model from scratch for object detection, they retrain CR-NET, YoloV2 and Fast Yolo which are pre-trained state-of-the-art models for object detection (that is transfer learning), and they obtained a good level of accuracy. Another novel approach for object detection and localization in digital images for crime scene investigation was developed by Surajit *et al.* (2017). The approach presents a Faster R-CNN (Region-based Convolutional Neural Network) to optimize existing deep learning algorithms for object detection. Faster R-CNN is a pre-trained convolutional neural network which builds on the architecture of Region Proposal Network for object detection in digital images. Hong-Wei (2015) also developed a model during the 2015 Emotion Recognition in the Wild contest based on prediction facial expression in static digital images. This model has an application on crowd analysis over a video surveillance camera for digital forensics. The research utilizes a transfer learning approach on a pre-trained deep convolutional neural network using a small dataset of static image from movies. Another application of convolutional neural network for art painting identification to detect copyrighted images used without content provider's permission for commercial purposes was developed by Yiyu *et al.* (2017). A comprehensive description of the development of other novel method adopted for object detection in images using Deep Neural Networks was provided by Christian *et al.* (2011).

The aim of this study is to optimize crime scene evidence analysis during a criminal investigation and provide a second-eye reviewer to forensic personnel. Certain evidence patterns usually found at most crime scenes are very useful for the reconstruction process and other forensic investigations, some of these patterns are bloodstain pattern, glass fracture pattern, fire burn patterns, dead victim faces and position, furniture position pattern, injury wound pattern, and so on (Davide *et al.,* 2011; Bhatt, 2017). This study utilizes a transfer learning approach on the state -of-the-art YOLOV2 object detection model based on the Convolutional Neural Network (CNN) algorithm to detect relevant objects at indoor crime scenes for evidence analysis. The study progressed a bit further to use this model in a mobile application for crime scene documentation. The application has two features which include case detail management (evidence documentation) and crime scene object detection for evidence analysis in real time.

2. Methodology

The selected methodology for this project is the Cross-Industry Standard Process for Data Mining (CRISP-DM) due to its sequential and iterative approach to problem-solving by applying data science and machine learning algorithms (Antonio & Jose, 2016). CRISP-DM is known as the standardized approach for data science projects. It follows a systematic development and delivery process from business understanding to model deployment. The ten (10) steps of the CRISP-DM (Figure 1) as it applied to the study are presented in sections 2.1 detailing the datasets use in the model creation and section 2.2 detailing the object-detection approach employed.

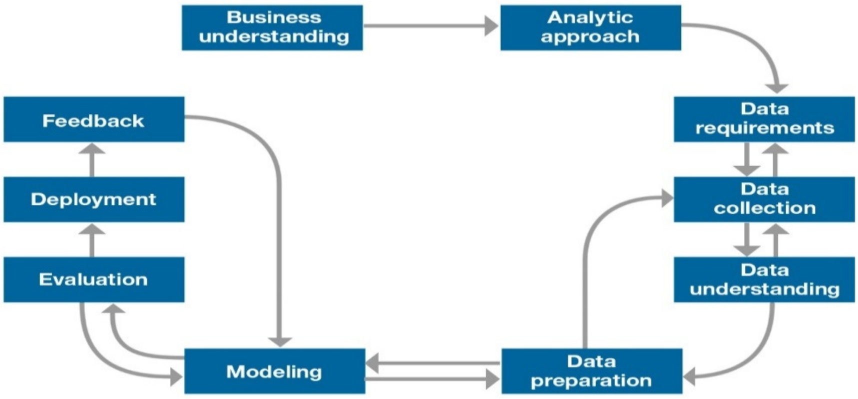


Figure 1: Ten Steps of the Cross Industry Process for Data Mining (source: www-01.ibm.com/common/ssi/cgi-bin)

2.1 Data and Model Creation

As an image classification problem with the aim of object detection for crime evidence analysis, the study was centred on developing and training convolutional neural network for object detection from digital forensic images using the YOLO algorithm. The dataset were obtained using the Google Advanced Search Tool based on their relevance to an indoor environment especially a bedroom. It contains 1,173 images including 5 classes of object common to indoor crime scene. Figure 1 shows a sample of the images collected as follows: bed (382), laptop (163), pc monitor (137), person (292) and sofa (200).

Figure 2: Sample Images of the Crime Evidence Analysis Dataset

The images acquired are not necessarily of the same quality, although they have the same resolution similar to the approach used by Rayson *et al.* (2017). Objects in these images are also of different positions, color, shapes, and sizes to expand the scope of the model while training. Every image in the dataset has the following annotation available in a text file (.xml): the class of object, the height of the image, the width of image, identification, and position of the objects in the image. Figure 1 also shows the boundary box of different types of object. The dataset were pre-processed to a .png file extension. The images were prepared by creating annotations for every image collected in the dataset. Annotations are corresponding ".xml" files which gives a detailed description of corresponding images in the dataset; such details like the position of the desired object in the image using boundary boxes. The dataset was split to 80% for training and 20% for testing. The split is done such that each set has the same number of the individual object class. The model was created on the training set and evaluated using the testing set. The deployment phase implements the trained model on an Android-based mobile application for crime scene evidence analysis and documentation.

2.2 Object Detection Approach

Specifically the Tiny-YOLO was used and the parameters were tuned to enhance its performance for the required task. The Tiny-YOLO architecture is shown in Table 1. The CNN was trained to detect five classes of object in much less time by changing the number of filters in the final convolutional layer to match the number of classes (C) in the dataset. The YOLO model use anchor boxes (A) predict boundary boxes as output with four coordinates (x, y, w, h), confidence and C class probability, and therefore the equation for the number of filters is given as:

$$\textit{Number of filters} = (C + 5) \times A \qquad (1)$$

where

C = 5 (5 class of objects)

A = 5 anchor boxes

Hence, the number of filters in the object detection task must be 50 as depicted in Equation 1.

Table I: Tiny-YOLO Architecture

	Layers	**Filters**	**Size**	**Input**	**Output**
0	conv	16	3 x 3/1	416 x 416 x 3	416 x 416 x 16
1	max		2 x 2/2	416 x 416 x 16	208 x 208 x 16
2	conv	32	3 x 3/1	208 x 208 x 16	208 x 208 x 32
3	max		2 x 2/2	208 x 208 x 32	104 x 104 x 32
4	conv	64	3 x 3/1	104 x 104 x 32	104 x 104 x 64
5	max		2 x 2/2	104 x 104 x 64	52 x 52 x 64
6	conv	128	3 x 3/1	52 x 52 x 64	52 x 52 x 128
7	max		2 x 2/2	52 x 52 x 128	26 x 26 x 128
8	conv	256	3 x 3/1	26 x 26 x 128	26 x 26 x 256
9	max		2 x 2/2	26 x 26 x 256	13 x 13 x 256
10	conv	512	3 x 3/1	13 x 13 x 256	13 x 13 x 512
11	max		2 x 2/1	13 x 13 x 512	13 x 13 x 512
12	conv	1024	3 x 3/1	13 x 13 x 512	13 x 13 x 1024
13	conv	1024	3 x 3/1	13 x 13 x 1024	13 x 13 x 1024
14	conv	50	1 x 1/1	13 x 13 x 1024	13 x 13 x 50
15	detection				

Source: https://pjreddie.com/darknet/yolo/.

3. Results

From the forgoing as shown in the Introduction section, CNN have played a major role in real-time object detection for crime evidence analysis but this study proposed a more optimized approach using the YOLOv2 object detection on a mobile app with real-time object detection and case documentation feature for crime evidence analysis. This study had successfully implemented an object detection system. The result achieved an average accuracy of 67.84% after training with 97 epochs (7000 iterative steps) on an Intel(R) Core (TM) i72720QM CPU with a processor speed of 2.20GHz. The trained and tested object detection CNN has 24 convolutional layers with 3 fully connected layers; Figure 2 depicts a Graph visualization of the network

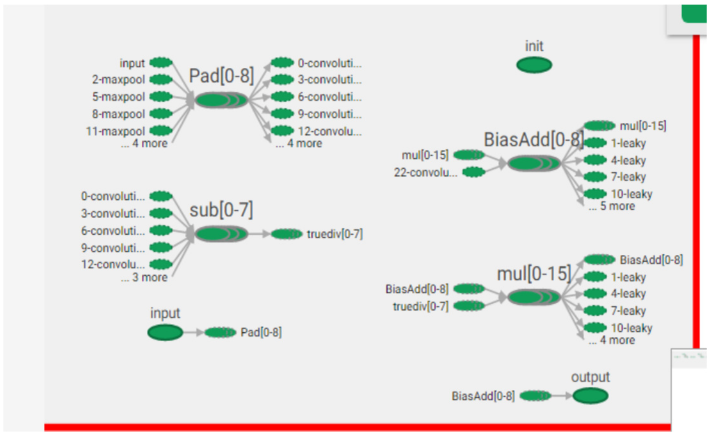
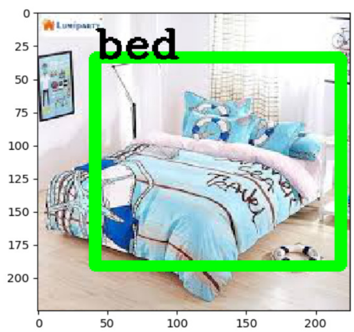


Figure 3: A 76 Node Graph Visualization of the Network on TensorBoard

The test set of the dataset was used to evaluate the object detection of the system. The output of different confidence thresholds on the dataset was initially evaluated. At 1.0, no object was detected hence the confidence threshold was reduced by 2 until all five objects could be detected at a confidence threshold of 0.013. With this threshold, 90% of correct detection on the test set was achieved with an average time of 5.89 seconds.



4. Discussion

The results obtained in this study had shown a promising future for crime evidence analysis. Crime scene reconstruction starts with the identification of shreds of evidence at a crime scene. The aim is to find out the people or things involved within the environment of the crime which may be a pointer to apprehending culprits. The implication of this study is to help the jury system in investigation of crimes so that there won't be instances of implicating innocent people. The system will greatly assist in decision making and focus on prosecuting the actual criminals involved in the crime. Having obtained a promising result, the study went further to evaluate the accuracy of the model on a video file and an image not in the Crime Evidence Analysis dataset used in the training. This is to validate its accuracy in a real-life crime scene. As depicted in Figure 3, the model detected the image clearly.

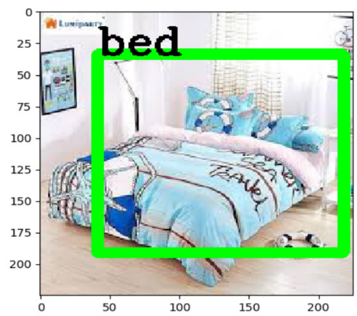


Figure 4: Object Detection Output on a Sample Image

The level of the intelligence exhibited by the system is actually needed whenever a crime scene analysis is required.

5. Conclusion

This study had been able to meet its aim by successfully developing a machine learning forensics through a transfer learning on YOLO using a custom dataset collected and prepared in the study. The model was trained on an Intel(R) Core (TM) i72720QM CPU with a processor speed of 2.20GHz achieving an accuracy of 67.84%. This accuracy would have been much better if the model was trained on a GPU; the training on Intel(R) Core (TM) i72720QM CPU was a major limitation. The study also went further to evaluate the accuracy of the model on the test set having an equal number of images on every class of object to validate the achieved accuracy. The model was deployed on a mobile app which allows users to create an account, document key details of physical criminal cases, view previously created cases, analyse these cases and take pictures at crime scenes (object detection module). However, the study had a limitation in the integration of the model into the mobile app for effective end-user operations. Notwithstanding, tThe tool app can be used to keep records of physical criminal cases digitally and the model for detection of object found at an indoor crime scene for evidence analysis. Though this system was not designed to prevent crime, the knowledge of its existence will add to the reduction of crime in the nation.

Correspondence
Emeka Ogbuju
Department of Computer Science
Federal University Lokoja
Kogi State, Nigeria
Email: emeka.ogbuju@fulokoja.edu.ng

References

Antonio, J. T. B., & José, C. R. (2016). *Data Mining Methods Applied to a Digital Forensics Task for Supervised Machine Learning.* Seville: Department of Languages and Computer Systems, University of Seville.

Bhatt, P. (2017). Machine learning forensics: A new branch of digital forensics. *International Journal of Advanced Research in Computer Science,* 217-222.

Christian, S., Alexander, T., & Dumitru, E. (n.d.). *Deep Neural Networks for Object Detection.* Google, Inc.

Davide, A., Giorgio, G., & Fabio, R. (2011). *Machine Learning in Computer Forensics (and the Lessons Learned from Machine Learning in Computer Security).* Cagliari, Italy: University of Cagliari.

Hong-Wei, N., Viet, D. N., Vassilios, V., & Stefan, W. (2015). *Deep Learning for Emotion Recognition on Small Datasets Using Transfer Learning.* Singapore: University of Illinois at Urbana-Champaign.

Rayson, L., Evair, S., Luiz, A., Luiz, S., Gabriel, R. G., William, R. S., & David, M. (2017). *A Robust Real-Time Automatic License Plate Recognition Based on the YOLO detector.* Belo Horizonte, MG, Brazil: Department of Computer Science, Federal University of Minas Gerais.

Surajit, S., E., F., Enrique, A., & Laura, F.R. (2017). Object Detection for Crime Scene Evidence Analysis Using Deep Learning. ICIAP 2017, Part II, LNCS 10485, 14-24.

Yiyu, H., & Jongweon, K. (2017). Art painting identification using a convolutional neural network. *International Journal of Applied Engineering Research* 12(4), 532 -539.