UTILIZATION AND OPTIMIZATION OF DEEP LEARNING ALGORITHMS IN CLASSIFYING LITTERING IN A COMPLEX ENVIRONMENT

A Thesis Presented to the Faculty, College of Computer Studies
Laguna State Polytechnic University
Los Baños, Campus
Los Baños, Laguna

In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Computer Science
Specialized in Intelligent Systems

PERILLA, IAN CARLO R.
MANA-AY, JERIKA G.
RECTO, ERICSON Z.

JONARDO R. ASOR

JULY 2024

APPROVAL SHEET

This thesis entitled "**UTILIZATION AND OPTIMIZATION OF DEEP LEARNING ALGORITHMS IN CLASSIFYING LITTERING IN A COMPLEX ENVIRONMENT**", prepared and submitted by **Ian Carlo R. Perilla, Jerika G. Mana-ay,** and **Ericson Z. Recto**, in partial fulfillment of the requirements for the degree, **Bachelor of Science in Computer Science specialized in Intelligent System**, is hereby recommended for acceptance and approval.

JONARDO R. ASOR
Adviser

Approval and Acceptance by the Committee on Oral Examination with a grade of ___ on ___.

GENE MARCK B. CATEDRILLA
Member

CAROLINA R. JOVAL
Member

JOCELYN O. PADALLAN
Member

JONARDO R. ASOR
Research Coordinator

Accepted as partial fulfillment of the requirements for the degree of **BACHELOR OF SCIENCE IN COMPUTER SCIENCE**.

JEFFERSON L. LERIOS, DIT
Dean, College of Computer Studies

OFELIA B. MANINGAS, Ed.D.
Chairperson, Research and Development Office

____
Date Signed

RESEARCH CONTRIBUTION NO.	

ACKNOWLEDGEMENT

The researchers would like to express their deepest gratitude to the Almighty God for providing strength, wisdom, and guidance throughout this research journey.

Special thanks to their dedicated adviser, Mr. Jonardo R. Asor, whose valuable insights and unwavering support played a crucial role in the success of this thesis.

The researchers thank the esteemed faculty members, Mr. Gene Marck B. Catedrilla, Ms. Carolina R. Joval, and Ms. Jocelyn O. Padallan, for their continuous encouragement and scholarly contributions that enriched their academic experience.

Thanks to their families for their understanding, patience, and constant encouragement. Their support has been a pillar of strength, fueling their perseverance.

Lastly, the researchers acknowledge themselves for the determination, hard work, and collaborative spirit invested in this endeavor. Together, they have navigated challenges and celebrated achievements, making this research a collective triumph.

The Researchers

DEDICATION

The researchers dedicate this study to their parents, whose love, inspiration, and words of support have empowered them to develop resilience, embrace opportunities, and lead joyful lives. The guidance from their parents imparts perseverance and helps them navigate challenging moments. They extend this dedication to their affectionate and supportive parents,

Mr. and Mrs. Perilla

Mr. and Mrs. Mana-ay

Mr. and Mrs. Recto

Foremost, the researchers dedicate this to the Almighty God, who provides them with the guidance, faith, and strength essential for the success of this study.

Completing the thesis work was only possible with the inspiration, support, and encouragement they received.

I.C.R.P

J.G.M

E.Z.R

ABSTRACT

The study explores the use of deep learning algorithms and the challenges related to technology adoption in littering detection. By effectively employing a camera as a surveillance tool to monitor specific littering activities from a distance, the collected diverse datasets encompassing plastic bags, trash bags, plastic bottles, and more underwent preprocessing using Roboflow to ensure the quality of the images. The dataset comprises 3,000 images of Humans and 6,000 images of Trash. Divide it into three sets: allocate 70% of the data to the training set, assign 20% to the test set, and reserve the remaining 10% to validate the model's predictions. A comprehensive analysis comparing three leading deep learning algorithms—YOLOv8, ResNet-50, and SSD-MobileNetv2—covers two critical dimensions. The first dimension involves assessing their performance in accurately classifying littering incidents, while the second focuses on evaluating their computational efficiency for real-time applications. The findings indicate YOLO as the optimal choice for real-time littering detection, boasting an 82% accuracy in classifying Trash and Humans, with an impressive 90ms inference speed during real-time applications. On the other hand, SSD MobileNetv2 is the second-best model to perform real-time litter detection with an accuracy of 60% and 200ms inference time.

Keywords: Deep Learning, Real-time Littering Detection, ResNet-50, SSD-MobileNetv2, YOLOv8

TABLE OF CONTENTS

Preliminaries **Page**

LIST OF TABLES

LIST OF FIGURES

DEFINITION OF TERMS

Anaconda	A free, open-source tool called Anaconda Python enables users to build and run Python programs. The developer is Continuum.io, a Python programming business.
Bounding boxes	Are rectangular or square-shaped markers drawn around an object in an image to define its boundaries.
Cartesian Coordinates	It is a mathematical system used to define the position of a point in a two-dimensional or three-dimensional space. It consists of two or three numerical values, denoted as (x, y) or (x, y, z), representing the distances along orthogonal axes. The x-axis, y-axis, and, optionally, the z-axis intersects at the origin, providing a systematic way to specify the location of a point to these axes.
Computer Vision	It is a branch of computer science that works on giving machines the ability to recognize and comprehend objects and people in pictures and videos.
Conceptual Framework	A theoretical structure of model provides a systematic and organized approach to understanding a particular phenomenon or solving a problem.
Evaluation Matrix	It is a systematic tool to assess and compare options or alternatives based on predefined criteria. This matrix typically consists of rows representing the alternatives and columns representing the criteria
Hyperparameters	Are the configuration settings that guide a machine learning model’s learning process, influencing its behavior and performance.
Imutils	A Python library provides a collection of convenience functions to simplify everyday tasks such as resizing, rotating, displaying images, and working with OpenCV in a more user-friendly manner.
Matplotlib	A charting library included NumPy, an extensive data numerical processing resource for Python. Matplotlib embeds charts in Python programs via an object-oriented API.

NumPy	It is a crucial Python package for scientific computing. This Python library provides various functions for fast manipulation of arrays, including mathematical, logical, shape, and manipulation functions. These include sorting, selection, I/O, discrete Fourier transforms, introductory linear algebra, basic statistical operations, random simulation, and more. It also contains multidimensional array objects and various derived objects such as masked arrays and matrices.
Open CV	Stands for Computer Vision Open-Source Library. It is the most widely used, well documented, and widely shared computer vision library. A collection of open-source computer vision algorithms is called OpenCV.
Pixels	It carries specific color and intensity information, collectively forming the visual representation of an image on a screen or in a digital file.
PyCharm	It provides a comprehensive set of tools and features to assist Python developers in writing, debugging, and managing their code.
Roboflow	It is an open-source platform that provides various datasets of annotated images, including those related to trash and humans. The study utilizes Roboflow to enhance the efficiency of training the models for detection.
Sanitary Landfills	Are engineered waste disposal sites designed to minimize environmental impacts while safely managing and containing solid waste.
Scikit-learn	It is an open-source machine learning library for the Python programming language. It provides many tools and algorithms for classification, regression, clustering, dimensionality reduction and more.
TensorFlow	It is an open-source machine learning framework developed by the Google Brain team. It provides a comprehensive platform for building and deploying machine learning models, especially deep learning models.
Video Surveillance System	It is a network made up of recorders, monitors, and displays. Analog or digital cameras are available with a range of options to experiment with, including resolution, frame rate, color type, and more.

Visual Data	It refers to information presented in a visual format, typically in the form of images, videos, or graphics. It encompasses the visual representation of data, often captured through cameras, sensors, or other imaging devices.
Xyxy dist. function	It refers to a bounding box representation using the coordinates of two diagonally opposites corners: (x_1, y_1) and (x_2, y_2). This notation succinctly denotes the top-left and bottom-right corners of the bounding box, simplifying the expression of object locations.

CHAPTER I

INTRODUCTION

Background of the Study

Waste pollution has been a growing problem worldwide, and it has become a significant concern in many countries due to its detrimental effects on the environment and public health. The Philippines is no exception, and it has been grappling with waste pollution for decades. The problem has become more severe in recent years due to the rapid growth of urban areas and industrialization, increasing the amount of waste generated. (Goswami, 2021), Waste pollution has been evident in the Philippines since the industrial revolution. However, the lack of effective policies and laws has resulted in mass pollution, and the country is now facing a situation where there is much more garbage than dumpsites. As reported by (ABS-CBN news, 2021), the Department of Environment and Natural Resources (DENR) recorded 362,000 metric tons of waste generated in the country in 2021 alone. According to a report by DENR 2018, which further compounds the waste management issue. Moreover, only 22 percent of municipalities have landfills, which leads to the problem of illegal dumping, also known as littering.

Making an area unsightly with garbage or inappropriately disposing of waste in known as littering. Littering generates pollution, a considerable environmental danger, and has become a growing source of worry in many nations. A study (Pamintuan, 2019) stated an increase in waste after typhoon Ondoy from previous years. A study by (Grobler et al., 2022) suggests that people litter due to their carelessness. In addition, they also blamed the need for infrastructure, such as waste bins, or the lack of discipline of others around them. In their findings, lack of education is eminent in the study. It is essential to be cautious about spreading false or misleading information, as it can harm the environment and lead to mistakes that make cause harm. Lack of consciousness or education is one of the causes of littering (Zhang, et al., 2022). Another study by (Maharoof et al., 2022) discusses the factors that affect youth's littering pattern. The visibility of litter and the group where they belong can induce littering.

Littering not only affects the appearance of cities but also poses severe threats to public health and the environment. The DENR reported in 2019 that they collected 50 metric tons of trash from Luneta Park, a popular tourist spot in Manila, on December 25 and 26. Littering not only affects national parks but can also damage bodies of water in the area. (Meijer et al., 2021), he found that the Pasig River contributed to 6.43% of the pollution worldwide because of its waste plastics. The Pasig River is considered the second-longest river in the Philippines. Seven billion plastic debris end up in bodies of water such as the ocean, which comes from rivers like Pasig. In addition, according to (Environment.cenn.org, 2023), 100,000 mammals are killed by these plastics every year.

Interventions and campaigns are needed to address littering problems in our society (Carvalho & Mazzon, 2019). One such intervention is the implementation of Republic Act No. 9003, also known as the Ecological Solid Waste Management Act. This provision requires the bureau to assists the National Solid Waste Management Commission in implementing solid waste management plans and creating policies to achieve the National Ecology Center's (NEC's) objectives. The NEC is responsible for information dissemination, consultation, education, and training of various local government units on ecological waste management (DENR, 2019). The National Solid Waste Management Commission has established the NEC to educate and replicate environmental advocacy program. This department comprises inventors, academicians, youth, and businesses who can assist NEC in its functions.

Former president (Marcos SR., 1975) at this moment ordered and decreed: “Any person who shall litter or throw garbage, filth, or other waste matters in public places, such as roads, canals, esteros or parks, shall suffer an imprisonment of not less than five days nor more than one year of a fine of not less than One Hundred Pesos (PHP 100.00) nor more than Two thousand Pesos (PHP 2,000.00), or both such fine and imprisonment at the discretion of the court or tribunal, without prejudice to the imposition of a higher penalty under any other law or decree”. MMDA Regulation Number 99-006, or the Anti-Littering Law, prohibits littering and throwing garbage or waste in public places. Under the said regulation, violators shall be penalized by administrative fine of Five Hundred Pesos (PHP 500.00) or community service for one day. Failure to pay administrative fines, the violator shall be penalized after conviction by a fine of One Thousand Pesos (PHP 1,000.00) or imprisonment of three to seven days of arrest, menor, or both at the court’s discretion.

Despite efforts to address littering, it remains a prevalent environmental problem. (Franceschini, 2019) conducted a study and found that the oceans contain a maximum of 12.7 million tons of litter. Similarly, (Kalnasa M. et al., 2019) estimated that the beaches of Macajalar Bay have a total of 4.82 million litter items. Furthermore, according to a report conducted by (GAIA, 2019) and quoted by (Cruz, 2021), the Filipinos consume around 160 million packets and 90 million bags made of plastic per day, totaling 90 billion worth of plastic per year. Thus, making recycling impossible (Renaud et al., 2018) and increasing the amount of litter in the country. Solving illegal littering is crucial to maintaining a healthy and sustainable environment. Proper waste disposal practices, public education campaigns, and enforcement of existing laws can help reduce littering. Addressing the issue of littering can have a positive impact on public health, the environment, and the economy.

This study employed a different algorithm to detect litters, which is achievable through computer vision algorithms. One such algorithm is the Residual Neural Network-50 (ResNet-50), widely recognized and considered one of the best in the world for computer vision (Zhang, 2021). This algorithm has also demonstrated high accuracy in a littering detection system employed by (Husni et al., 2019).

One of the study's objectives is to compare the ResNet-50 algorithm with the You Only Look Oncev8 (YOLOv8) algorithm, renowned for its speed as one-stage detection algorithm (Korchani & Sethom, 2021). YOLO has proven effective in detecting microplastics in bodies of water (Jia et al., 2023). Another algorithm is Single Shot Detection, which combines with Mobile Netv2 (Balmik et al., 2023).

The three algorithms underwent training to differentiate between trash and other objects and accurately categorize different types of trash. Subsequently, the algorithms underwent testing to evaluate their effectiveness in monitoring and detecting litterers. Following testing, they implemented a camera to oversee real-time videos and provide accurate outputs and footage, assisting authorities in identifying suspects in littering crimes. To determine which algorithm has the highest accuracy and efficiency, the algorithms would be compared with one another.

Objectives of the Study

This study aims to provide a robust model for detecting instances of littering within intricate and challenging environments. Thus, this study specifically aims to:

1. To integrate visual data from cameras and apply the Law of Cosines to calculate the distance between trash and humans to create a comprehensive surveillance model capable robust detection of littering in complex environments.
2. To develop a model that integrates time monitoring and recording for securing instances of littering for documentation and analysis.
3. To evaluate the YOLO model in terms of Accuracy and Speed in detecting littering instances in real-time.

Scope and Limitations of the Study

This study focuses on three specific algorithms: YOLOv8, ResNet-50, and SSD-MobileNetv2. The emphasis would be on utilizing these algorithms' most stable, accurate, and up-to-date versions. The training process for these three algorithms involved using images from the collected dataset. A camera would record real-time videos. The accuracy of the model detecting acts of littering would be recorded and documented to evaluate its effectiveness. Furthermore, the implementation would take place within the campus of Laguna State Polytechnic University-Los Baños campus, with a particular focus on densely populated areas.

The study's limitations include the number of rubbish types constrains the model's accuracy. It can distinguish new forms of trash. The devices hardware capacity and algorithm size may compromise the model's effectiveness. As littering happens in split seconds, the study prioritized speed over accuracy. The model only records a few seconds of the act to prevent memory shortage, which can diminish its effectiveness. Computational capacity may also be reduced due to bias inflicted by datasets, as the models rely on pictures and not videos or simulations. The models cannot use videos due to limitations in size and hardware capacity.

Significance of the Study

Department of Environment and Natural Resources (DENR). The department can benefit from the study by implementing the algorithm in its monitoring and environmental protection plan. As the spearhead of the resource protection program, it can utilize the system to reduce the damage caused by solid waste that affects the environment. Additionally, the department can assist the LGU by providing data on potential littering hotspots.

National Solid Waste Management Commission (NSWMC). This commission can incorporate the algorithm into its solid waste management plan. It can also help create a system that identifies different types of litter using the algorithm. The commission would able to utilize this for surveillance of illegal dumping sites and maintain a record of events related to littering.

Local Government Unit (LGU). This study can assist by creating solid evidence for people caught in littering. It is a support system capable of detecting and developing new routines to identify littering perpetrators. Additionally, it aids in implementing laws regarding proper waste disposal. A system to monitor the area eases their work and helps create a new ordinance that can leverage the algorithm's potential.

Waste Management Officer. The system can help keep their hallways clean and reduce problems related to littering. It provides them with ease in their work and establishes a new pattern of waste disposal areas that can minimize garbage in specific locations. They can also identify new waste disposal sites to reduce littering further.

Future Researchers. They can use the study as a reference to create similar or related research to aid in waste disposal. Additionally, the system can assist them in differentiating between humans and garbage. Future researchers can develop new types of datasets that can detect the act of littering and determine the areas where littering is prominent.

CHAPTER II

THEORETICAL FRAMEWORK

Related Literature

This chapter discusses the study's associated literature and previously completed investigations whose key findings support and validate the provided reasons in for building the model. It goes over several systems and their functions. The study emphasized the importance of numerous literature and studies in creating the "Utilization and Optimization of Deep Learning Algorithm in Classifying Littering in a Complex Environment."

Littering

Littering, as defined by (Chaudhary, 2021), is a purposeful action described as 'keeping litter in hand while occupying an area and placing it in the area when leaving'. In addition, (Ojedokun, 2022) stated that littering, the improper disposal of waste, has become a pervasive and costly problem worldwide. As a result, such behaviors may endanger the environment. (Franceschini, 2019) cited a United Nations Environment Program (UNEP) research that found an annual global influx of between 4.8 and 12.7 million tons of trash into the waters.

Assessing the environmental impact of various carrier bag options, (Cinavancik-Uslu et al., 2019) used littering as an indicator and provided a comprehensive evaluation of the environmental impacts of different carrier bag options. This information can assist policymakers and consumers in making more informed decisions. (Amadei et al., 2022) used littering to estimate the plastic footprint of the European Union.

People often discard their trash inappropriate places due to different reasons. Some reasons for their behavior include peer influence and social groups. Littering impacts people's desire to engage in anti-littering conduct through factors such as mindset, personal standard, perception of behavioral control, and moral duty, according to research by (Singh et al., 2021). (Buitrago, 2020) believes litter, mainly marine littering, to be an issue beyond borders, with enormous and permanent environmental costs that indirectly reduce the property's value. Improper waste disposal can also result in health problems, as noted by (Somaroo, 2015) and cited by (Sasman, 2021), especially in low-income countries facing difficulties in illegal and improper dumping practices that negatively impact health. However, quantifying is challenging (Sasman, 2021, cited Zariba, 2016).

Trash

Domestic and industrial discharges contribute to air pollution and eutrophication, resulting in air quality, land degradation, and adverse effects on flora and fauna. Throughout history, municipal bodies have maintained clean roads, managed city garbage, and ensured proper disposal (Priyadarshi et al., 2020).

In developing countries, solid waste management presents a significant challenge for authorities in small and large cities due to the increasing generation of solid waste and the strain it places on municipal budgets. Furthermore, the need to understand the various factors impacting the waste management system compounds the issue (Abdel-Shafy et al., 2018). People create billions of tons of garbage yearly, which is likely to rise significantly, exacerbating waste management challenges. By 2050, the World Bank predicts 70% rise in worldwide urban garbage from 2016, reaching 3.4 billion tons of solid waste (Mariyam, 2022).

Image Detection

Image detection, as defined, refers to the technique of finding and locating objects and patterns inside a picture or image (Córdova et al., 2022). People utilize it in various applications, such as identifying faces, recognizing objects, and performing medical imaging. To effectively identify and categorize objects in new images, one must evaluate the image using machine learning techniques such as deep learning neural networks. These networks are trained on annotated photo datasets to classify the objects in the image. Similarly, (Korchani et al., 2021) stated that image recognition uses computer vision and deep learning approaches to identify objects and events in real time video feeds. Moreover, image detection identifies objects or other elements within digital images. It focuses on a specific application that uses high-resolution aerial images and deep learning methods like YOLO for image identification to find litter and solid waste on the streets, as mentioned by (Ulloa-Torrealba, 2023) in his article.

PlastOPol is a recently discovered dataset used to assess how well various object recognition networks perform when combined with computer vision techniques to identify litter. According to the study (Córdova et al., 2022), YOLOv8 is the quickest network for completing this assignment. The study emphasizes the crucial role of image detection in creating effective litter detection system using deep learning and computer vision techniques. Their (Korchani et al., 201) highlights the importance of image detection in various areas, including face identification, trash detection, and object detection.

(Husni et al., 2021) discussed training machine learning model to distinguish between various features or patterns in an image and categorize or label it accurately. The importance of image detection in several applications, including face identification and litter detection, is covered in the study of (Korchani et al., 2021). (Ulloa-Torrealba et al., 2023) presented an article demonstrating how they used image detection software to identify and categorize litter in high-resolution street photos. The authors created a map of the area's litter using deep learning based on the YOLOv8 algorithm. This map can help identify places where trash is more prevalent, resulting in more effective clean-up procedures and waste management regulations.

Litter Detection

Litter detection, as described, refers to identifying and locating litter or waste in an environment through various technologies such as computer vision, machine learning, and sensors (Córdova et al., 2022). According to a recent study by (Korchani et al., 2021), video surveillance cameras and computer vision algorithms can identify improper waste disposal in public spaces. This technology can help track individuals who litter and ensure proper waste management. The study by (Ulloa-Torrealba et al., 2023) defines *litter* detection as identifying and detecting litter objects in an environment using techniques such as image analysis, machine learning, and remote sensing.

Research by (Jia et al., 2023) also highlights that litter detection enables efficient and accurate monitoring of plastic pollution in water bodies, which is essential for developing effective strategies to manage plastic waste and reduce its environmental impact. In their article, (Husni et al., 2021) aim to develop a real-time monitoring system based on image classification methods to detect littering activity. The project aims to help researchers develop low-technology. These simple solutions could reduce some of the tasks that volunteers, and citizen scientists perform in collecting and monitoring environmental debris by using litter detection. The goal of this project is to increase scientists' access to data, which ultimately impact specific management and policy decisions.

Researchers have evaluated deep learning systems to assist in identifying litter and trash based on images and have assessed their performance in identifying litter in real-world scenarios and cluttered picture backgrounds (Córdova et al., 2022). The litter detection approach in the research conducted by (Korchani & Sethom, 2021) involves monitoring the individual with an ambiguous form and checking the existence of a waste container.

Real-time Detection

Real-time detection is the ability of an object detection algorithm to perform detection and classification tasks at a speed that is acceptable for practical applications, such as video surveillance systems, as defined by (Wilson et al., 2022). In the study by (Valarmathi et al., 2023), real-time detection refers to the ability to detect objects, specifically humans, in a video stream or sequence without significant delay. According to (Kundu, 2023), real-time detection refers to an object detection algorithm's ability to perform object stream or sequence without significant delay. According to (Kundu, 2023), real-time detection refers to an object detection algorithm's ability to perform object detection and recognition at video frame rates or faster.

According to (IRJET, 2022), YOLO models are employed in the study to locate objects in footage and real-time video feed. After being trained and evaluated on a dataset of various object images, the models exhibited good accuracy and quick inference times. Therefore, YOLO models are ideal for real-time detection applications. The article (IJSRCSEIT, 2020) explains the process of tracking objects in a video stream using real-time object detection. To monitor items in real-time, the study combines the YOLO (You Only Look Once) object identification technique with the SORT (Simple Online and Real-time Tracking) tracking algorithm. (Kundu, 2023) states that the YOLO technique uses real-time detection for object recognition. The developers of YOLO intended it to detect objects in real-time by analyzing images and videos within the same amount of time. Based on the study by (Wilson et al., 2022) examines the precision and speed of the YOLO approach for object detection using real-time detection.

The article in (IRJET, 2022) emphasizes the advantages of utilizing YOLO models for accurate and quick detection. Similarly, (IJSRCSEIT, 2020) discussed the significance of real-time object tracking in live video streams. According to (Valarmathi et al., 2023), real-time detection is helpful in emergencies because it can increase the efficiency and precision of human action recognition. Real-time detection can assist rescue teams in promptly locating and rescuing catastrophe victims, minimizing damage and casualties.

Deep Learning

Deep models can be referred to as ‘neural networks with deep structures’, as stated by (Zhong-Qui Zhao et al., 2019). According to (Junghwan Lee et al., 2022), deep learning has advanced many scientific fields such as computer vision, speech recognition, and natural language processing.

Two types of deep learning object detection exist two-stage and one-stage. SSD (Liu et al., 2016; Lu et al., 2021), YOLO (Redmon & Farhadi, 2018), and ResNet models (Lin et al., 2018) provide instantaneous detection using regression. These models quickly regress bounding boxes and class probabilities, making detecting more efficient.

Properly utilizing of deep learning algorithms, such as CNN, YOLO, and R-CNN, bridged the gap in litter detection (Cordova, 2022). The models were processed and trained using algorithms to recognize litter. A study (Politikos, 2021) used deep learning to predict different kinds of trash on ocean floors, achieving a mean average precision of 62%. Another research conducted by (Garin et al., 2021) suggested that deep learning and technological tools can aid in monitoring macro litters in the ocean. Additionally, one of the datasets used the deep learning algorithm R-CNN for litter determination (Proenca, 2020).

Machine Learning

Machine Learning (ML) has become a prevalent tool researchers and data analyst utilize in significant corporations worldwide. Researchers have recognized the ML algorithms, a subset of Artificial Intelligence (AI), in handling the challenges posed by the massive influx of big data (Suhaimi et al., 2020). Specifically, supervised machine learning, a type of ML, has gained considerable attention for its ability to establish relationships between input and output variables. In this approach, input variables, also known as features or ‘X variables’, are associated with an output variable referred to as the target or ‘y variable (Ali, 2022). By employing human supervision, algorithms are trained and accurately labeled to perform tasks effectively.

Numerous studies have utilized machine learning for detecting litter. (Korchani, 2021) utilized a machine learning algorithm to develop a program that can detect litterers. They employed the Convolutional Neural Network algorithm for computer vision (Zhang, 2021) in detecting garbage.

Convolutional Neural Network

Convolutional neural networks (CNNs), as defined by (Zhang, 2021), are the most widely used algorithm and have been applied extensively in various fields, including computer vision, speech processing, and face recognition. The complex sequence of cells forming the visual cortex in a cat's brain-inspired the structure of CNNs. Experiments using CNN for image features obtained good values for precision and recall, as reported in the study (Alfarrarjeh, 2018), cited by (Ping, 2020). Good feature extraction is necessary for achieving highly accurate results in studies.

(Ping, 2020) conducted a study where they trained one of the most extensive convolutional neural networks using subsets of ImageNet. The study yielded the most outstanding results reported on these datasets so far. In the study by (Husni et al., 2021), a real-time litter monitoring system was created, utilizing CNN as an image classification algorithm to predict littering in a mini garden. (Sharma, 2021) used CNN to detect human movements with high accuracy by training it on Channel State Information (CSI).

Using CNN as an algorithm in the study could be beneficial. As shown by (Husni et al., 2021), CNN has a higher accuracy rate than other algorithms in litter predictions. The study (Korchani, 2021) used CNN to suggest the optimal ways to detect and recognize garbage regarding recognition performance. To detect people who litter in Bogor Botanic Gardens (Rizki

2018) used CNN. A comparative study conducted by (Ali, 2023) using four iterations of CNN showed a positive result in trash recognition. CNN is, therefore, suitable for image-based classification for littering and trash detection, and its high accuracy yield gives it a good reputation for object classification.

You Only Look Once

You Only Look Once (YOLO) is a quick object identification neural network that can achieve real-time performance by partitioning an input image into grid cells. YOLO predicts each cell's bounding box and object categorization (Córdova et al., 2022). Because of its quick training stage and higher performance compared to prior versions, YOLO is the most promising technique for object recognition. In order to achieve this, we use genetic programming to modify the set of previous boxes. YOLO operates as one-stage detector, as highlighted by (Korchani & Sethom, 2021) simultaneously training bounding box coordinates and corresponding class label probabilities. This design choice makes YOLO faster than two-stages detectors like R-CNN, necessitating an external region proposal algorithm. Yolov8, the latest version, utilizes a feature extraction network based on Darknet-53 and predicts bounding boxes at three scales to enhance detection accuracy. YOLO finds wide applications in various fields, including autonomous driving, surveillance, and robotics.

YOLO is a real-time object identification system based on deep neural networks commonly used for identifying macroplastics in water bodies, according to (Jia et al., 2023). Furthermore, (Ulloa-Torrealba et al., 2023) present YOLO as an object detection technique that conducts object recognition and classification concurrently utilizing a CNN architecture. YOLO splits a picture into grid cells and guesses bounding boxes and class probabilities in each one as a single-stage detector, confirming its reputation for remarkable speed and precision in object detection tasks.

Residual Neural Network

This deep learning architecture addresses the vanishing gradient problem by introducing shortcut connections between layers, enabling the training of intense neural networks (Hussain et al., 2020). To enhance ResNet's performance with limited data, (Wang et al., 2022) modified the architecture, addressing a common challenge in the practical application of deep learning.

ResNet-50 is one of the backbone models in the proposed technique for litter identification from digital images using machine learning in the paper by (Liu et al., 2023). The study conducted by (Wang et al., 2022) presented several modifications to the original ResNet architecture, including a new bottleneck structure and a dense connection, aimed at reducing the number of parameters and improving the model's efficiency. They evaluated the performance of their modified ResNet models on several benchmark datasets for image classification tasks and compared them with other state-of the-art models.

To move information from one layer to the next without changing it, ResNet created skip connections, sometimes referred to as shortcut connections. Deep ResNets are built by putting residual blocks, allowing for the formation of networks with up to a hundred layers. This method learns parameters from initial network activations (Sharma, 2021).

Single Shot Detector

(Yin et al., 2021) emphasized the Single Shot MultiBox Detector (SSD) as a 19-object detection algorithm that is an end-to-end, single-stage approach. It incorporates the regression concepts from YOLO and integrates the anchor box mechanism found in Faster RCNN. (Huang, 2021) SSD is a technique for identifying targets of different sizes using feature maps of multiple scales through multi-scale target detection. The process entails selecting candidate frames of various sizes for discrete feature maps in the following pyramid, with VGG165 as the backbone network.

According to a study (Zha, 2020), the SSD algorithm performed well in detection accuracy and speed. As stated, (Bai, 2021), SSD is a new approach for detecting end-to-end, just like traditional target detection algorithms. Preserving high accuracy and real-time performance while displaying great detection results for small targets and multiple objectives improves the speed significantly.

According to a study (Li, 2019), SSD is a cutting-edge object identification system that can identify things in photos using just one deep neural network. SSD does away with feature resampling and bounding box suggestions, in contrast to conventional object identification techniques. Instead, it employs distinct tiny convolutional filters on many feature maps, resulting in a notable enhancement in speed and precision. According to the analysis (Ding, 2021), SSD matches targets of different shapes with anchors of varying sizes and aspect ratios to identify targets of varied forms. However, we cannot change the dimensions and form of the tiny infrared target as they are primarily predetermined. Anchor waste occurs when an anchor's size is too big or tiny to fit the objective.

MobileNet

In the study by (Howard et al., 2018), MobileNet is introduced as a convolutional neural network architecture that utilizes depth-wise separable convolution, except for its initial layer. Unlike traditional convolutional networks, depth-wise separable convolution reduces the computational requirements for image convolution. The authors created MobileNetv2, which enables users to perform object detection and classification tasks.

The architecture of MobileNet comprises two components: depth-wise convolutional and 1x1 pointwise convolutional (Suharto et al., 2020). Inception-ResNet-v2, another significant design, combines the benefits of residual and inception networks, showcasing superior feature extraction capabilities for BCH images using convolutional neural networks (CNNs) (Ogundokun et al., 2022).

To further highlight the benefits of depth-wise separable convolutions in MobileNet, (Wang et al., 2022) emphasize the reduction in parameters and computational complexity without a significant loss in classification accuracy. Based on MobileNet, the authors propose three improved models, known as Dilated MobileNet or Dilated Convolutional MobileNet models, which incorporate receptive local field expansion in shallow layers.

Euclidean Distance

As defined by (Hoffman et al., 2019), Euclidean distance calculates the square root of the sum of squared coordinate differences to calculate the distance between two locations in a two-dimensional space. This measure, widely used in domains such as image processing and machine learning, has applications in spatial modelling. (Cardarili et al., 2019) Moreover, (Behrens et al.,2018) define *Euclidean distance* as the shortest straight-line distance in Euclidean space, extensively used in spatial modeling and machine learning.

In the study by (Jia et al., 2023), Euclidean distance is employed to calculate the distance between the center points of the detected objects and the corresponding ground truth labels in the training data. This distance serves as a measure of the accuracy of the detection algorithm. The lower the Euclidean distance between the detected object and the ground truth label, the more accurate the detection algorithms are. In their study, (Behrens et al., 2018) used Euclidean distance fields to generate training data for machine learning models. Specifically, they used Euclidean distance fields to model the distance of each point in a 3D space to the nearest surface.

The algorithm uses the Euclidean distance to calculate the distance between the predicted and actual bounding boxes. The prediction is regarded as true positive if the Euclidean distance between the two bounding boxes is smaller than a given threshold. Therefore, Euclidean distance is a crucial metric for evaluating the accuracy of the detection algorithm, and it is employed to evaluate the suggested model's performance (Jia et al., 2023). (Komarasamy, 2021) conducted a study using the Euclidean algorithm to determine the distance between people during the COVID-19 pandemic in the context of littering. In the present study, we can implement the algorithm to measure the distance between humans and trash, which can serve as variable to trigger alerts when someone is littering.

Polyfit

Employing tracker such as the Global Positioning System (GPS) and Received Signal Strength Indicator can address the significant challenge in distance estimation for littering detection and surveillance. In a study by (Malic, 2020), an Octave function was utilized, callable in Python through the Matlab function.

This function fits a given dataset to a polynomial equation, creating a curve that accurately passes through the points and enabling real-time tracking of objects in the study, like humans and trash. Similarly, road lane detection has benefited from this function, as demonstrated by (Huu, 2022), who used it to detect white lines on the road, predicting the distance between the lines to assist the model in staying on track for self-driving applications.

In the study, Polyfit allows the calculation of polynomial fitting equations, which is essential for curve fitting and distance estimation models. The study can use it to compare the accuracy of fitted models and evaluate the results.

Law of Cosines

In the study, distance estimation is essential, and to achieve a model capable of it, the study utilized the Law of Cosines as the foundation for distance estimation. This geometric equation, commonly used in satellite location systems for determining distances between ground stations and satellites (Frieman, 2020), would be applied similarly in the model, where the camera's position is analogous to satellite positioning.

The Law of Cosines aids real-time detection by calculating the distance between trash and humans. The Joint Doppler and Ranging Law of Cosines

(JDR-LOC) is an industry standard for satellite movement detection and positioning (Jun, 2019), further affirming the study's reliance on the Law of Cosines for distances estimation.

Algorithm Analysis

This analysis exhibits the algorithms used to determine distances between humans and trash in detecting littering. This part of the study aims to demonstrate how YOLOv8 integrate various algorithms to enable the model to perceive spatial data captured by the camera. This part explains the theoretical side of the study and how it can be applied to the model to improve its capabilities. YOLOv8 is a state-of-the-art object detection model but cannot do spatial calculations, which the algorithms below provide.

Euclidean Distance Algorithm

To calculate the distance between two points, one needs to square the x and y coordinates of both points, then add the results together and find the square root of the sum. This formula is also like the Euclidean distance algorithm. The formula is also like the Euclidean distance algorithm. The algorithm finds the shortest path between the two points in a cartesian coordinate plane. The formula can be written as:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

It is usually used in spatial analysis problems like Geographic Information System (GIS) by finding the hypothenuse of a box by employing the distance formula to the maximum and minimum coordinates.

Furthermore, the algorithm is also used in different machine learning models, such as K-Means, to determine the shortest distance between the centroid. It is an algorithm that finds the path between two points in a straight line (Suwanda, 2020). The method utilizes the Pythagorean theorem, which is a way of solving the hypothenuse of a triangle given the length of two legs I pursuit of finding the value of the most extended leg. The formula used is:

$$c^2 = a^2 + b^2$$

The sum of a and b is the given shorter leg being squared equal to the c, which is the hypotenuse. The algorithm applies the theorem with a box's maximum and minimum xy coordinates.

In addition, a study about spatial modeling found a more practical use of the algorithm as it is independent and can be easily computed by (GIS) (Behrens, 2018). It also indicated in the study that it has a 95% confidence interval, implying it has excellent accuracy. In distance estimation, the algorithm also garnered 95% accuracy in breaching human's social distancing during the pandemic (Ahmed, 2021).

Furthermore, the algorithm was also used in different spatial distance studies, such as micro aerial vehicles (MAV) n motion planning (Han, 2019). This algorithm outperforms other methods in terms of calculating distance. It is also helpful in the health sector, as K-Means uses Euclidean distance to determine the distance between centroids (Arminarahmah, 2021).

Table 1. Application of Euclidean Distance Algorithm

Application	**Algorithm**
GIS	Euclidean Distance
Spatial Mapping	Euclidean Distance
Distance Estimation	Euclidean Distance

Table 1 shows the summary of the algorithm application. The implementation of the algorithm solves broad types of problems, from distance estimations of coordinates inside a model to distance estimations in real-time. The algorithm can help in mapping unexplored areas such as oceans. It can improve healthcare by determining the growth rate of tumors and helping diagnose patients.

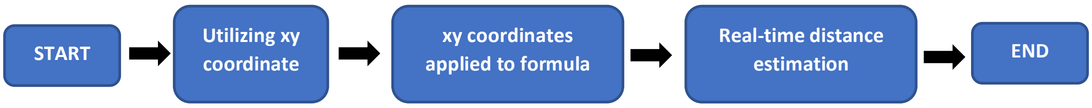


Figure 1. Architecture of Euclidean Distance

The flexibility of the distance estimator technique makes it suitable for resolving various spatial analysis problems. Its primary use is to identify limited regions inside a plane. However, it can also help to extract insights from spatial datasets for data analysis jobs. Futhermore, the algorithm's real-time monitoring capabilities allow applications like tracking changing item locations, which is essential in domains like robotics and surveillance. Because of its flexibility to adjust to shifting spatial configuration, it is helpful for dynamic spatial analysis, which has applications in fields like crowd surveillance and traffic management. The approach is also helpful in geographic applications, where it helps with route planning and disaster response. In addition, accurately measuring distances in various situations is crucial for creating models and conducting simulations.

Advantages

It always finds the shortest path between two objects. Distance estimation during experiments is not a problem because the focus been solely on finding a single solution. It is also faster since it only uses two variables. It has a low time complexity of $o\ (\log b)$.

Disadvantages

The ability to compute the distance depends on th date given by the model. As the algorithm depends on the dataset provided, it can calculate distances only within the set limits of the algorithm. Thus, abnormal sizes of objects can cause problems in terms of distance estimation as they may appear nearer to the camera instead. It may also be a problem in estimating distance if a third party or another object appears or disappers in the camera. The system may miscalculate the distance between objects.

Law of Cosines

The LOC scheme is beneficial when creating an adequate navigation infrastructure since it requires a minimum number of navigation nodes—possibly just one. Users using current proximity link radios on planetary worlds such as the Moon and Mars serve as an exemplary case (Lee, 2019). Please remember that this technique can be applied to any surface or object, not limited to a single planet. This technology can be used in cameras to monitor human behavior. This algorithm in distance measurements can be used on a non-moving high altitude to monitor the distance of two objects simultaneously.

In order to improve the algorithm further, we can use the Euclidean distances as a reference point for calculating the distance of the objects in sight of the camera. The algorithm can be modified to determine the distances of objects and be able to like them accordingly. In addition, The Global Positioning System currently utilizes it to track locations. References satellites cover huge distances with low constraints. In addition, it uses Doppler measurements, which lessen inaccuracy between the reference point and the user point (Lee, 2020). This further implicates the remarkable ability of the algorithm in distance estimation. The algorithm is also based on the Pythagorean theorem as it is also a modified version of the equation, which utilizes instead of the angle of the triangles to find the length of the opposite side of the angle of the reference point.

While the Law of Cosines is good at determining distance, it is still has problems in terms of velocity error. The reference point and user should first known to create calculations. With the angle and initial distance, the model would not work and would have substantial error margins (Lee, 2019).

Table 2. Application of Euclidean Distance Algorithm

Application	Algorithm
GPS	LOC
Doppler Radar	LOC
Joint Doppler Radar Law of Cosines (JDR-LOC)	LOC

These are applications of LOC in different studies that involves distance estimation. Most of them are used mainly for an extended range of distances. Thus, it can be reliablein terms of spatial analysis. Satelite systems use most of them, implying their suitability for communication between users, as seen on smartphones.

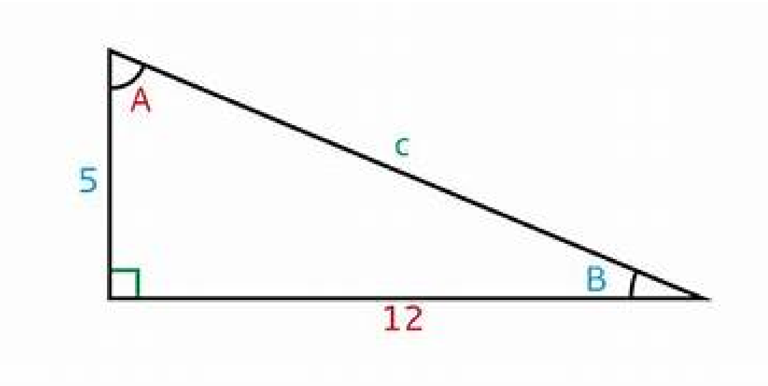


Figure 2. Law of Cosines

The diagram in Figure 2 illustrates the use of the Law of Cosines (LOC) to determine the length of the third side ‘c’ or distance in a triangle. The algorithms use all the information, employing the sides ‘a’ and ‘b’ and the angle ‘c’. The algorithm calculates side ‘c’ length by applying the LOC. The approach also uses the Side-Angle-Side (SAS) theorem to calculate the length of the remaining side of the triangle using the unknown values of side ‘b’, angle ‘c’, and side ‘a’. With this all-encompassing method, the algorithm can effectively extract the geometric relationships inside the triangle, making it easier to determine distance accurately by combining LOC and SAS concepts.

Advantages

Over the years, this technology has gained a solid reputation, especially in satellite application and communications, due to its outstanding precision. Its performance history of consistency serve as more evidence of its dependability. Notably, it sets itself apart from other guidance systems with a low mistake rate. This feature increases its attractiveness and usefulness in a variety of technical applications. It iis essential in situations where accuracy and dependability are critical, including satellite operations and communication networks.

Disadvantages

Using the Law of Cosines (LOC) presents several drawbacks. Its complex mathematical calculations can increase computational demands, impacting processing time. Accuracy issues may ariise when trash and humans are nearby, introducing errors due to minor measurement inaccuracies. The LOC assumes a fixed camera orientation, potentially causing problems in scenarios with variable angles or movements. Sensitivity to noise, such as lighting variations, may affect the reliability of distance estimates. Additionally, the limited adaptability of the LOC to diverse environmental conditions and configurations can hinder its robustness in different littering scenarios.

Synthesis and Relevance of the Study

Littering remains evident in the Philippines despite government efforts, nad trash remains a prevalent problem. This issue negatively impact us and our environment, leading to problems such as floods and plastic pollution. With technological advancements, machine learning algorithms are gaining recognition in our country. This study aims to assist the community in detecting litterers and mitigating the ecological impact of trash.

Researchers have conducted similar studies to examine how plastics affect oceans. Using cameras to recognize and capture evidence of littering saves time and aids waste management officers in monitoring specific areas. Furthermore, it enhances the area’s aesthetic value, as monitoring helps catch perpetrators and identify areas with high litter volume. Additionally, it has an indirect economic impact by providing solid evidence to penalize those who litter. Other researchers can use this study to develop a more robust detection system incorporating facial recognition to strengthen the evidence.

Therefore, this study can benefit various aspects of littering management. These technologies can be helpful in this field and serve as a reference for bomb detection units, potentially saving lives. Moreover, it can help locate lost items in a specific are and return valuable possessions to their rightful owners.

Table 3. Matrix of Reviewed Related Literature

AUTHOR & YEAR	TITLE	METHODS/ ALGORITHM USED	FINDINGS
Rangel-Buitrago, N., Williams, A., Costa, M., & de Jonge, V. (2020)	Curbing the inexorable rising in marine litter: An Overview	LMP	This study is a comprehensive overview of the current situation and trends of marine litter, including its source, types, impacts, and current efforts to address the issue. It highlights the urgent need to address the growing problem of marine litter in order to protect the health of ocean ecosystems and the well-being of human communities that depend on them.

Sasman, M., Dolan, C., Villegas, D., Eyob, E., & Barrett, C. (2021)	The Influence of Marginalization on Cultural Attitudes and Trash Disposal Practices in Esfuerzo de Paraíso of the Dominican Republic: A Qualitative Interview Study	Interviews	According to the study, the residents of Esfuerzo de Paraso in the Dominican Republic have a strong sense of belonging to a community and cultural identity, but their waste disposal habits are impacted by a variety of variables, including poverty, a lack of services and infrastructure, and marginalization. The study underlines the need of addressing littering and waste management challenges in marginalized populations in a culturally appropriate manner.
Husni, N. et al. (2021)	Real-Time Littering Activity Monitoring Based on Image Classification Method	CNN, CNN-LSTM	Real-time monitoring of littering activity can be achieved using a combination of image processing techniques. The study developed an image classification model that can accurately identify littering activities in real-time, with an overall accuracy of 92.5%. The model can be integrated into a surveillance system to help enforce anti-littering laws and encourage responsible waste management behavior.

Córdova, M. et al. (2022)	Litter Detection with Deep Learning: A Comparative Study	CNN, Faster RCNN, Mask-R CNN, EfficientDet, RetinaNet, and YOLO-v5	The use of deep learning algorithms can achieve high accuracy in litter detection from images. Among the deep learning models tested, YOLOv3 and RetinaNet achieved the highest performance in litter detection. The developed models can potentially be integrated into litter monitoring systems and help improve waste management practices.
Korchani, B., & Sethom, K. (2021)	Real-Time Littering Detection for Smart City using Deep Learning Algorithm	CNN, SSD, YOLO	The paper suggests utilizing deep learning techniques to create a real-time litter monitoring system for smart cities. The suggested method detects items that are polluting photos and categorizes them into specified classifications using a convolutional neural network (CNN).

Jia T. et al., (2023)	Deep learning for detecting microplastic litter in water bodies: A review	YOLO, SSD, CNN, Inception ResNet, Shuffle- Xception, Improved Mask R-CNN, VGG16	The study provides a comprehensive review of deep learning methods used to detect microplastic litter in water bodies. The authors examined a range of deep learning methodologies, including long short-term memory (LSTM) networks, convolutional neural networks (CNN), and deep belief networks (DBN). It demonstrates how deep learning techniques can quickly and accurately detect macroplastic litter in water bodies, making them a valuable tool for environmental management and monitoring.
Cardarilli, G. et al. (2020)	N-Dimensional Approximation of Euclidean Distance	Manhattan distance	The study introduces a technique to estimate the Euclidean distance between two points in an n- dimensional space using an approximation algorithm. The precision and effectiveness of their method highlights its versatility in solving various problems in data analysis, machine learning, and computer vision.

You L. et al., (2022)	GPU-accelerated Faster Mean Shift with Euclidean distance metrics	Mean-shift	The study proposed a GPU-accelerated method for performing mean shift segmentation using Euclidean distance metrics. The proposed approach might be used in real-time image and video processing jobs.
IRJET (2021)	Object Detection Using YOLO Models	YOLO	The study investigates and evaluates object detection using the YOLO (You Only Look Once) paradigm. The results show that the YOLO model outperforms other cutting-edge object identification methods in terms of accuracy and speed.
International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT) (2020)	Object Tracking by Detection using YOLO and SORT	YOLO and SORT	The study presents a new approach for tracking objects by merging two deep learning algorithms, namely You Only Look Once (YOLO) and Simple Online Real-time Tracking (SORT). The YOLO algorithm is employed to identify objects, whereas SORT is employed to track them based on the detected objects.

Ping, P. et al. (2020)	Smart Street Litter Detection and Classification Based on faster R-CNN and Edge Computing	R-CNN and EDGE Computing	The study uses an object detection method called faster R-CNN o detect and classify litter on the street. The potential application of this system is in managing litter and ensuring clean streets in smart city environments.
Naf'an E., Sulaiman R., Mohamad Ali N. (2023)	Optimization of Trash Identification on the house Compound Using a Convolutional Neural Network (CNN) and Sensor System	CNN	The study proposed a novel system that integrates a low-cost wireless sensor network and a CNN to identify and classify waste materials into recyclable and non-recyclable categories. It can contribute to an effective waste management strategy that can reduce the environmental impact of waste, promote recycling, and encourage sustainable practices.
Yin Q., Yang W., Ran M., Wang S. (2021)	FD-SSD: An improved SSD object detection algorithm based on features fusion and dilated convolution	SSD	The algorithm improves the feature extraction process of SSD. It demonstrates the effectiveness of the proposed algorithm in real- world application.

Li Y. et al., (2020)	Multi-block SSD based on small object detection for UAV railway scene surveillance	SSD	The method improves the detection accuracy of small objects in complex scenes by combining multi-block feature maps with the SSD algorithm. It demonstrates the effectiveness and superiority of the algorithm for small detection in UAV railway scene surveillance.
Politikos et al., (2021)	Automatic detection of seafloor marine litter using towed camera images and deep learning	CNN, Deep learning	The author explores the application of deep learning techniques to detect marine litter on the seafloor. Towed camera images are utilized for data collection, and deep learning algorithms are employed for automatic detection.
Proença et al., (2020)	TACO: Trash Annotations in Context for Litter Detection	R-CNN Deep Learning	A dataset containing images of annotated trash that is used to train R-CNN algorithm.

Komarasamy (2021)	Minimising The Spread of Covid-19 Using YOLO V3 Algorithm	YOLO	This work proposes a system that utilizes surveillance cameras to monitor the movement of people and calculate social distancing using Euclidean distance. It focuses on maintaining social distancing specifically with COVID-positive individuals, improving accuracy by 98% in predicting social distancing with infected individuals and identifying protocol violators.

In Figure 3, the matrix of reviewed literature focusing on littering detection, several key themes emerge related to popular deep learning models such as YOLO, ResNet, and SSD. The literature highlights the effectiveness of YOLO in object detection tasks, emphasizing its ability to process images in real-time with high accuracy. The inclusion of ResNet in the reviewed literature points to its significance in addressing the challenges of training deep neural networks. ResNet's residual learning framework facilitates the training of deeper networks, contributing to improved litter detection performance by mitigating the vanishing gradient problem. Furthermore, the literature delves into the applications of SSD, emphasizing its advantages in terms of speed and accuracy. SSD's multi-box detection system allows for efficient localization and classification of litter objects, making it a promising choice for real-time littering detection systems.

Overall, the matrix of reviewed literature provides insights into strengths and nuances of these deep learning models, shedding light on their contributions to advancing litter detection technology.

Conceptual Framework

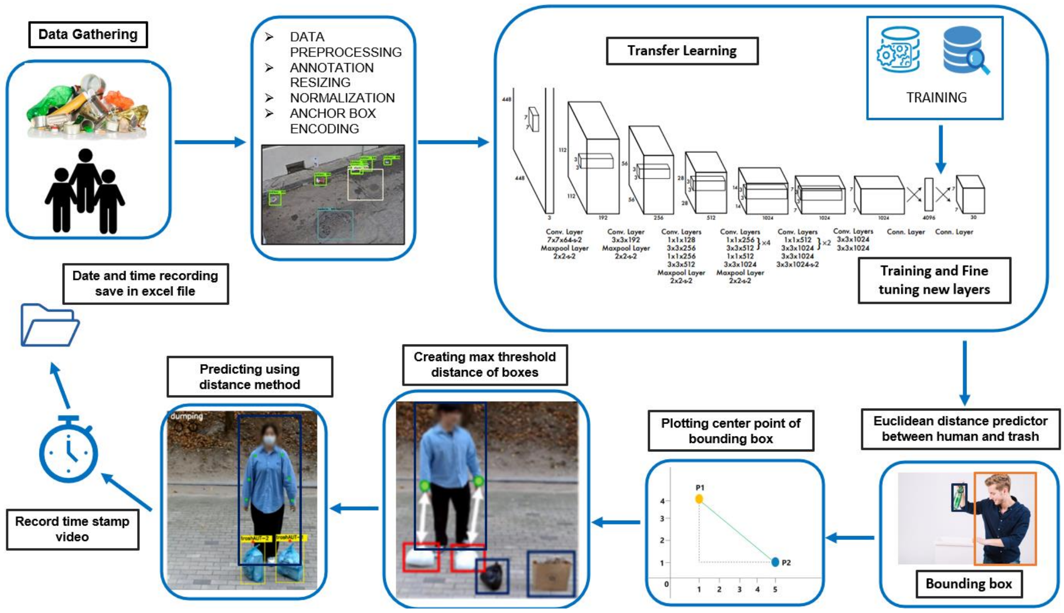


Figure 3. Conceptual Framework

The model used two kinds of data, primary and secondary data, which would be processed. The images annotated and given bounding boxes. The data would be set to numerical values through annotation by determining the point of bounding boxes. First, the study used the images to customize pre-trained models by adding extra layers to the frozen model weights. After detecting the boxes, the model used Euclidean distance to determine the distance between the human and the trash. The box's center would be acquired and converted to numerical values. Then, the system used the numerical values' distance to create a certain threshold, determining whether someone is littering by detecting if humans and trash detected reached the maximum distance. The time module record the time of littering occurrence and save it in an xlsx file, which anyone can use to review whether someone is littering.

CHAPTER III

METHODOLOGY

Materials

The following are the materials needed to complete the study. These materials gathered and utilized as aids for the research process.

Programming Language

Python

The study used Python as a programming language to develop a system that can process images captured by the camera in real-time. Python's imaging library, OpenCV, employed to preprocess the images and extract features that aid in identifying litter. The study used machine learning libraries such as Scikit-Learn and TensorFlow to train a model to detect litter in the camera images. Finally, integrated libraries like NumPy and Matplotlib analyzed the collected data and visualize the outcomes.

Software

This part of the study reveals all the software used in the study. Furthermore, it also explains and demonstrates the functionality of each software.

Google Colab

The study utilized Google Colab to store the code, data, and documentation in a single location. This approach facilitates the replication of the research by others and simplifies the process of reproducing the analysis in the future. Google Colab is a free, cloud-based platform that allows users to write and execute Python code collaboratively, making it versatile for various needs.

Anaconda

The study used Anaconda to create interactive data visualization, write and run code snippets, and document all the progress and findings of the study. Creating a consistent environment for the system ensures that the code can be easily transferred and replicated across various computers and platforms. It also includes other libraries that can aid in data processing.

PyCharm

It offers a wide range of features that greatly enhance the development process. With advanced code completion capabilities, developers can write code faster and with fewer errors. The built-in debugger allows for efficient troubleshooting and fixing of issues in Python scripts, ensuring smoother development. Furthermore, PyCharm's seamless integration with version control systems simplifies collaboration and enables effective code management throughout the study.

Frameworks

These are collections of pre-written code that serve as the foundation for software applications. The upcoming sections would list the utilized frameworks and explain their individual uses in the study.

TensorFlow

It provides powerful tools for building and evaluating the model that detects and classifies litter. The study used TensorFlow to analyze large amounts of image and video data, extract meaningful features, and identify patterns and trends that can help to understand littering behavior better.

Matplotlib

It is a necessary library for the study's visualization and plots. Researchers can efficiently exhibit the findings of their investigation in the form of graphs, charts, and figures by utilizing Matplotlib. With its simple syntax and abundant documentation, Matplotlib enables researchers to present their findings visually appealing and instructively

Numerical Python

A core Python module that provides efficient and robust tools for working with arrays and performing numerical computations. Researchers use NumPy in the study to pre-process and modify picture data due to its efficient array manipulation features. These routines enable picture array operations such as cropping, resizing, filtering, and other changes.

Open-Source Computer Vision

The library can process images and videos powerfully and perform tasks such as object detection, feature extraction, and user interface creation. It has a broad set of functions and algorithms, making it suitable for various computer vision tasks. Researchers can use OpenCV for image modification, object identification, feature extraction, and designing user-friendly application interfaces.

Imutils

A Python toolkit, Imutils, used in the study to simplify image and video processing tasks. It has features for resizing, rotating, cropping photos, and processing video feeds. The study may use Imutils to focus on detecting litter in camera footage while optimizing the underlying image and video processing tasks.

Scikit-Learn

A powerful Python package provides a wide range of machine learning tools and methods for picture categorization. It improves model efficiency and streamlines workflow by offering feature extraction, data preparation, and model assessment functionalities. The study use Scikit-Learn to apply classification methods effectively, enhancing the model's picture classification skills.

Hardware

Table 4. Hardware Specification

Hardware	Specification
Processor	Intel ® Core ™ i3-1115G4 @ 3.00GHz
RAM	8.00 Gigabytes
Solid State Drive	300 Gigabytes
System Type	64-bit Operating System X 64-based processor

Laptop Specification

This study utilized laptops as the primary hardware. The hardware requirements mentioned in Table 4 served as prerequisites for the equipment devices used. The equipment usually had specific operating system compatibility or general working framework requirements. These considerations were taken into account to ensure the model’s smooth operation and compatibility during the implementation phase of the study.

Data

Table 5. Dataset

Data	Number of Images	Description
Humans	3,000	The human images in .jpg format features a diverse collection of high-resolution photographs depicting individuals in various poses and environments. The dataset is designed for object detection, and human activity analysis, comprehensively representing human diversity.
Trash	6,000	The trash images in .jpg format comprises a comprehensive collection of high-quality photographs capturing various types of litter in diverse environmental settings. This dataset is tailored for tasks related to litter detection, environmental monitoring, and waste management, offering researchers and developers a valuable resource for training and evaluating machine learning models in the context of trash recognition.

The study utilized the dataset sourced from the open-source website Roboflow. Roboflow provides diverse datasets of annotated trash, convertible to various formats. Additionally, the platform enables the uploading and annotation of datasets. The dataset includes images from Roboflow and images captured in different locations, featuring trash and humans—the two classes under detection in the study.

Research Design

This study employed an experimental research design, manipulating one or more variables to observe their effect on another variable. This design establishes cause-and-effect relationships between the independent and dependent variables (Harris D., 2019). The researcher manipulated the input and measured the model's output accuracy and speed to determine the effectiveness of the three algorithms utilized: YOLOv8, ResNet-50, and SSD-MobileNetv2. The study utilized various data collection methods such as observation, computer simulation, and experimentation to achieve the research objectives.

Data Collection Instrument

Actively searching and filtering various datasets available on the internet. Specifically, they focus on identifying datasets that contain humans and trash images in .jpg format. This search led them to discover Roboflow, an open-source platform providing diverse datasets of annotated trash. Additionally, the study utilizes a collection of human images that encompassing variations in gender, color, size, and other attributes.

The following steps outline the process undertaken to gather and construct the datasets required for the study.

1. The data is examined and filtered to determine its relevance to the study. Carefully scanning the dataset to identify images and annotations of humans and trash objects in various environments.

2. During the investigation, it was explored whether the dataset had been utilized in previous studies to assess the credibility of the authors and the sample. After verifying the authenticity of the data, the analysts identified the data repositories from which it could be downloaded and cloned.

3. The study used the cloned data to train and test the model's accuracy in distinguishing between trash and humans.

Model Development

Model development is a crucial aspect when creating an algorithm for the study. Various algorithms suitable for image recognition of trash were filtered and narrowed down to ensure the selection of appropriate models.

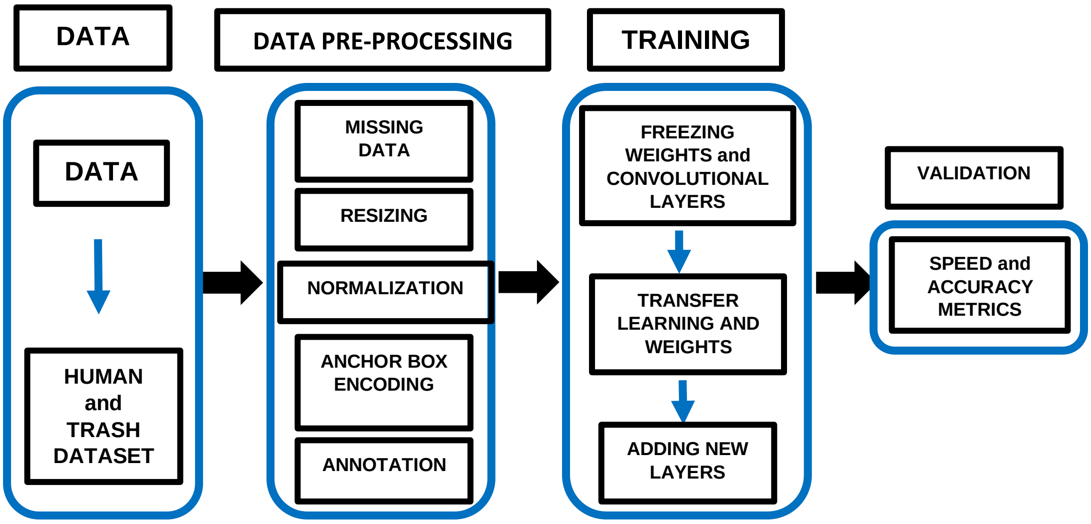


Figure 4. Model Development

Figure 4 illustrates the baseline for the algorithms that used in this investigation. Various strategies employed to preprocess the dataset using different designs and pipelines to fine-tune the dataset incorporated within the model.

Data Pre-processing

The collected data underwent processing to ensure the quality of the images. Resizing the photos would contribute to create a homogeneous dataset. Removing missing data and applying augmentation techniques would also enhance real-time detection.

Data Split

The dataset would be divided into train, test, and validation segments. The training set comprise 70% of the data, the test set comprise 20% of the dataset, and the remaining 10% validate the model's predictions.

Solution

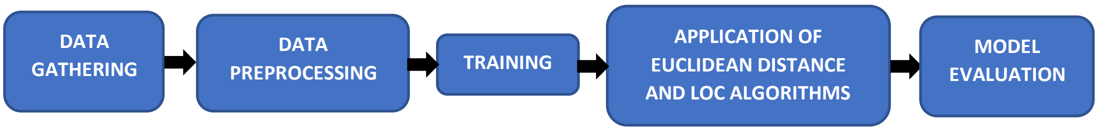


Figure 5. Solution for Littering Detection

Following the training phase on the dataset containing images of humans and trash, plans are in place to enhance the model by incorporating advanced algorithms for spatial awareness. This integration would enable the model to identify individual instances of litter and understand the spatial relationships between humans and trash in a given environment. By leveraging this spatial awareness, the model aims to provide a more nuanced and context-aware litter detection system capable of discerning intentional littering behaviors and improving overall accuracy in varied scenarios.

Experiment

The experimentation would primarily focus on trials to identify the optimal values for achieving high accuracy in recognizing illegal garbage disposal. Hyperparameters would be adjusted to obtain the desired results. When considering hyperparameters, one should consider the learning rate, batch size, number of epochs, confidence threshold, weight decay, optimizer, and Intersection over Union (IOU).

Learning rate. This technique facilitates faster learning for the algorithm by updating its parameters during training, which helps the model to adapt quickly.

Batch size. Generate a training set of a desirable size to achieve accuracy in detecting objects in images, mainly trash, and humans.

Number of Epochs. To train the dataset, the trainer needs to process it a certain number of times, which determines the required number of iterations.

Confidence Threshold. These parameters filter the precision and recall of the detected trash. They can be adjusted to increase the number of detected littering instances while potentially reducing accuracy or to enhance accuracy by reducing the number of detectable littering actions.

Weight Decay. It limits the size of the weights by penalizing the models, reducing the possibility of overfitting the data, and making them more straightforward and robust.

Optimizer. It updates the model's parameters repeatedly depending on the gradients of the loss function concerning those parameters. The gradient shows the sharpest rise or descending direction, and the optimizer utilizes this information to alter the settings to decrease the loss.

IOU. It calculates the overlap between the expected and actual bounding boxes of rubbish and people.

Table 6. List of Hyperparameter Controlled Variables

HYPERPARAMETER	YOLOv8	SSD-MobileNetv2	RESNET-50
Learning rate	0.001	0.08	0.04
Batch size	16	8	8
Number of epochs	100	100	100
Confidence Threshold	0.25	0.25	0.25
Weight decay	0.0005	0.00004	0.00004
Optimizer	SGD	SGD	SGD
IOU	0.7	0.6	0.6

Model Evaluation

The accuracy matrix uses two tests, Intersection over Union (IOU) and Mean Average Precision (mAP), to determine the model's accuracy by obtaining the actual positive and negative values for litter detection. The mAP utilizes the data from IOU to compute the average precision of models. It uses inference time and frames per second (fps) for speed. Inference time refers to the time the model needs to predict, while fps represents the number of frames the model can process.

Accuracy Matrix

Intersection over Union

The threshold is a value that indicates the proximity of bounding boxes to each other, ranging from zero to one. Typically, two bounding boxes are involved: one representing the ground truth or annotated box and the other predicted by the model. In this case, the model detects trash and humans, represented by the annotated boxes. If the value equals or exceeds the threshold, it is considered a positive. If a result does not meet the minimum threshold requirement, it is considered a false positive, per (Tsoi, 2019).

$$IOU = \frac{Area\ of\ Intersection}{(Ground\ Truth\ Area + Predicted\ Box\ Area) - Area\ of\ Intersection}$$

The formula for IOU involves dividing the area of the intersection by the sum of the ground truth are and the predicted box area, subtracted by the area of the intersection. IOU would help the study by assessing whether the model obtained a true positive or negative value.

Mean Average Precision

An evaluation metric commonly used for object detection. It is suitable for this study as trash considers the trade-off between precision and recall, considering both false positives (FP) and false negatives (FN). This characteristic makes mAP a suitable metric for most detection applications (Shah, 2022).

To calculate the mAP, the models must generate prediction scores and use them to determine precision and recall. Below is the formula for precision and recall.

$$Precision = \frac{TP}{TP + FP}$$

To calculate the precision, we divide the number of correct predictions by the sum of incorrect and correct forecast. In the study context, precision represents the model's accuracy in detecting objects, specifically littering and other objects. A higher precision rate indicates a higher level of accuracy for the model and vice versa.

$$Recall\ (TPR) = \frac{TP}{TP + FN}$$

Recall depends on the ratio of true optimistic predictions to the sum of accurate positive and false pessimistic predictions. Recall measures the model's ability to detect all instances of genuine litter correctly. The study also derive mAP as the average accuracy estimated for each class, including categories such as persons and rubbish.

$$Average\ Precision\ (AP) \int_{r=0}^{1} p(r)dr$$

The obtain precision values from the precision-recall curve at different confidence thresholds to calculate the average class's average precisions. Then, sum up the precision values corresponding to recall values ranging from zero to one with a step size of 0.1. The resulting value represents the average precision.

$$mAP = \frac{1}{k}\sum_{i}^{k} AP_i$$

The mean average precision (mAP) is the mean of all precision values across all classes. API stands for Average Precision, which is determined separately for every class, where k is the total number of classes.

F1-score

$$f1 = \frac{2 * (Precision * Recall)}{Precision + Recall}$$

It evaluates the performance of two classifiers using precision and recall, with the F1 score reaching a maximum of 1 and a minimum of 0. Precision and recall contribute equally to the F1 score as a percentage, calculated using the specified procedure.

Accuracy

$$Accuracy = \frac{(TP + TN)}{TP + FP + TN + FN}$$

The study by (Kumar V. et al., 2020) defines *accuracy* as the measure of how well three different machine learning techniques can predict clinical outcomes after shoulder arthroplasty. It refers to the ability of the machine learning models to correctly predict the outcome of the surgery based on the input data.

Confusion Matrix

This matrix compiles the model's performance based on the test data. The study used a multiclass confusion matrix, which helps classify trash that is evident in the study. According to (Markoulidakis, 2021) research, the confusion matrix is a matrix that provides a mix of expected vs. actual class occurrences, and it plays an integral part in evaluating algorithm performance.

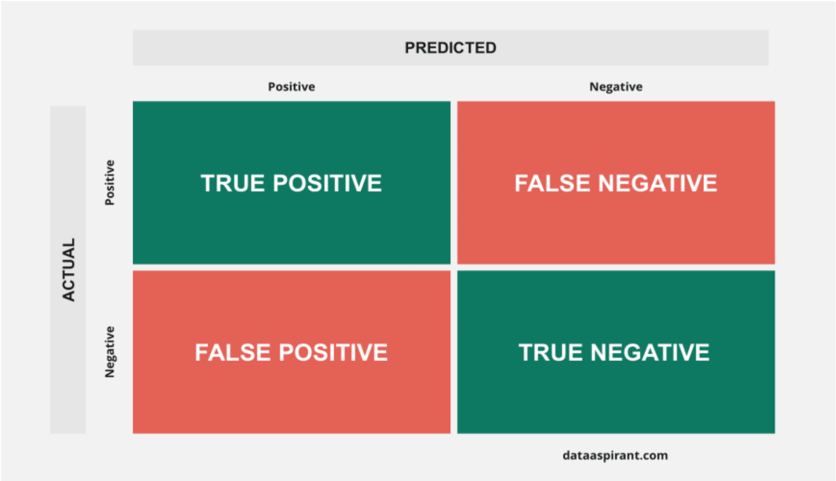


Figure 6. Confusion Matrix

Speed Matrix

Inference time. It refers to the time needed for the model to predict or infer the occurrence of littering. The study used the model's results, which YOLOv8 (Canu, 2023) demonstrated by analyzing the inference time through processing video frames and predicting the objects within each frame.

Frames per second (fps). It is the number of frames a model can process in real-time or within a video. The study utilized this metric to evaluate the model's ability to handle a large volume of images in seconds efficiently.

CHAPTER IV

RESULTS AND DISCUSSION

The investigation conducted in this study centered on achieving three primary objectives within the domain of littering detection utilizing visual data from cameras. The study presents the results obtained through meticulous exploration and experimentation linked to each objective. The following discussion is an in-depth examination and interpretation of these results, aiming to extract insights, emphasize significance, and explore potential future advancements in littering detection and real-time video analysis.

Objective 1. Integrating camera visual data with the Law of Cosines to compute distances between trash and humans results in a viable surveillance model for litter detection in complex environment.

Integration of Visual Data from Cameras

Figure 7. Visual Data from Camera

Using OpenCV, one can capture from the camera and create a visual representation of the surroundings. The system converts the raw data into frames, each comprising pixels. In this case, the video pixel dimensions are 1152 x 640 pixels, as shown in Figure 7.

Figure 8. Cartesian Coordinates of Pixels

Figure 8 shows that the model used the pixels to convert coordinates represented as numerical data similar to Cartesian coordinates by obtaining the width and height of the pixels. In the study (Gollapudi S., 2019), pixels from OpenCV can be used to convert Cartesian coordinates with a 1:1 ratio.

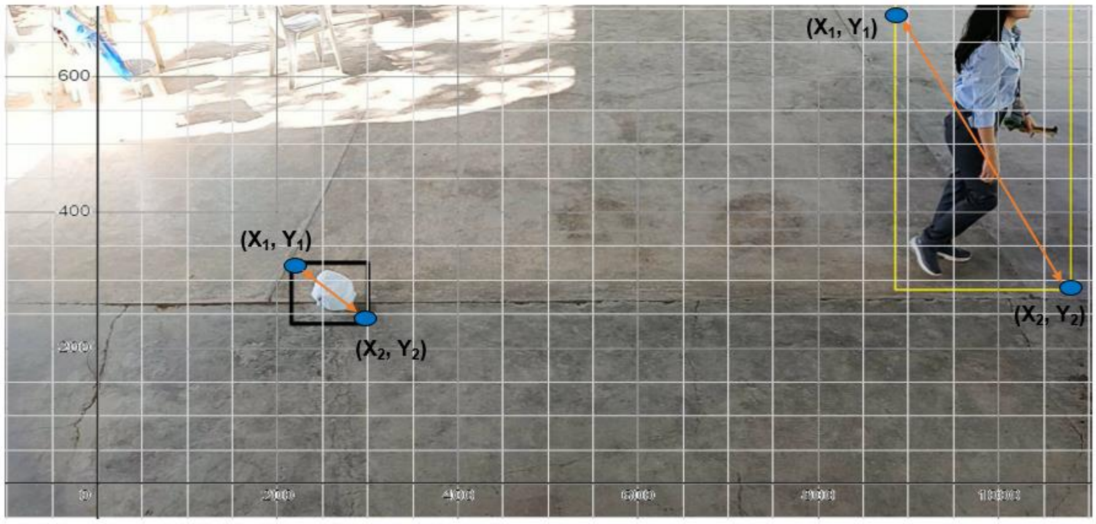


Figure 9. Computed Euclidean Distance using dist. function in Python

As shown in Figure 9, the dist. function provides the coordinates of the two points of the bounding boxes. The bounding box coordinates are in the format '(x_1, y_1, x_2, y_2) where (x_1, y_1) is the top-left corner, and (x_2, y_2) is the bottom-right corner. $d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$ is the formula used by dist.

function to calculate the pixel distance between objects. The distance obtained would be poly-fitted to convert the pixel size to the approximate distance between human and trash. The model utilized these coordinates to determine instance of littering.

Application of Law of Cosines to Calculate the Distance Between Trash and Humans

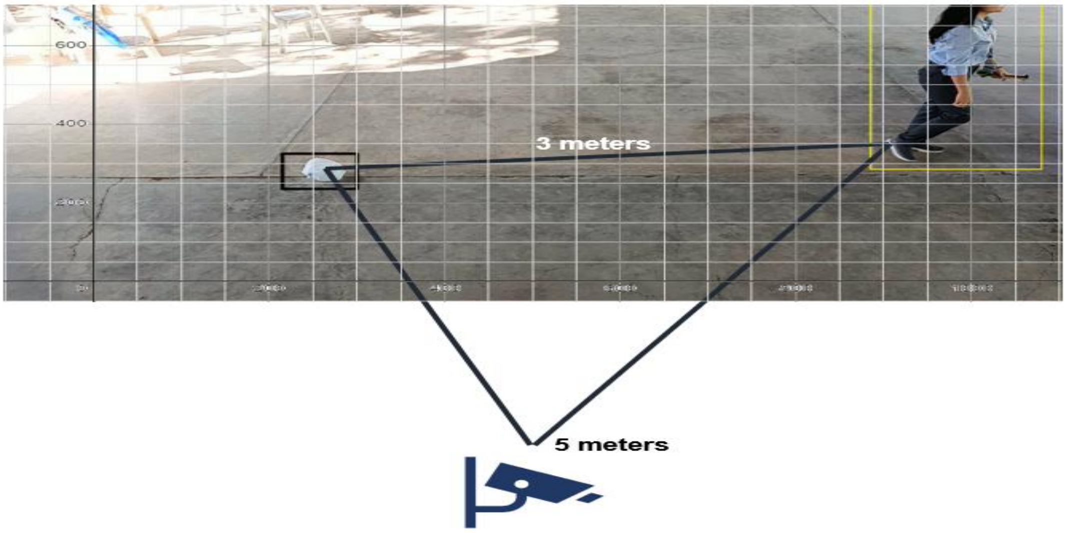


Figure 10. Law of Cosines Application

Figure 10 illustrates the application of the LOC to calculate the distance between Trash and Humans in the study. The model utilized the distance between the camera and the two detected objects and inputted these distances into $c = \sqrt{a^2 + b^2 + 2\, a\, b\, cos\gamma}$ the LOC formula. Wherein 'c' is the opposite side, which in the study is the distance between the two classes is equal to the square root of the sum of squared distance between camera and trash (a), camera and human (b) subtracted by the twice of ab multiplied to cosine γ. The model compares which length is more significant than between a and b, which switches the values to prevent miscalculations.

Objective 2. The developed model integrates time monitoring and recording to secure instances of littering for documentation and analysis.

Time Monitoring and Recording

	A	B	C	D
1	2023-10-21 13:17:35			
2	2023-10-21 13:17:56			
3	2023-10-25 11:27:55			
4	2023-10-25 11:27:55			
5	2023-10-25 11:27:55			
6	2023-10-25 11:27:55			
7	2023-10-25 11:27:55			
8	2023-10-25 11:27:55			
9	2023-10-25 11:27:55			
10	2023-10-25 11:27:55			
11	2023-10-25 11:27:55			
12				
13				

Figure 11. Date and Time Recording saves in xlsx file

The model utilized the 'open workbook' function to open an Excel file and save the time and date of the littering instances, as shown in Figure 11. Furthermore, OpenCV integrates timestamping into the output video, enabling one to rewind the video to the exact date and time of the littering occurrence.

Image and Video Recording

Figure 12. Snapshot of Littering being recorded by the Model

The model would identify the two classes, and if a person departs or throws garbage in the frame, the color of the trash bounding box would change to red. This process would generate photos that can help identify littering. The mentioned timestamp can also serve as reference for video replay to identify the person responsible for the trash in the areas, as illustrated in Figure 11.

Objective 3. To evaluate the YOLO model in terms of Accuracy and Speed in detecting littering instances in real-time.

Confusion Matrix Score of YOLO model

Table 7. Confusion Matrix of YOLO model

YOLO model	
257	88
56	455

Table 7 presents the Confusion Matrix for the YOLO model, revealing its performance on a classification task. The values in the table represent the count of true positives, true negatives, false positives, and false negatives. Specifically, the YOLO model achieved 257 true positives, 455 true negatives, 88 false positives, and 56 false negatives, indicating its ability to predict positive and negative instances accurately and where it may have made classification errors.

Evaluation of YOLO model in terms of Accuracy Performance

Table 8. Results of Accuracy Evaluation

Model	Precision	Recall	F1-score	mAP	Accuracy
YOLOv8	0.8211	0.7449	0.7812	84.50%	83.18%

Table 8 presents the evaluation results for the YOLO model, showcasing various performance metrics. The precision of the model is 0.8211, indicating the accuracy of optimistic predictions, while the recall is 0.7449, representing the model's ability to capture all positive instances. The F1-score, a harmonic mean of precision and recall, is 0.7812, and the model achieved an accuracy of 83.18%, with a mAP of 84.50%, providing a comprehensive overview of its effectiveness in object detection.

Evaluation of YOLO model in terms of Speed Performance

Table 9. Results of Speed Evaluation

Model	Inference Time	Frames per second
YOLOv8	90 ms	30 fps

Table 9 presents the speed evaluation results for the YOLO model, which details its inference time and corresponding frames per second (fps) performance. The model demonstrates an inference time of 90 milliseconds, indicating the average it takes to process a single image. With a commendable frame-per-second rate of 30, the YOLO model showcases its real-time processing capability. It is processing 30 frames in one second, making it suitable for applications requiring quick and efficient object detection.

Maximum Distance for Detecting Humans

Table 10. Maximum Distance for Detecting Humans in Meters

DISTANCE	ACCURACY
4.0 m	85.8%
5.0 m	84.6%
7.0 m	80.7%
8.0 m	78.6%
18.0 m	75.3%

Table 10 outlines the maximum distances for detecting humans, measured in meters, along with corresponding accuracy percentages. At a distance of 4.0 meters, the system achieves an accuracy of 85.8%. The accuracy remains at 84.6% when the distance increases to 5.0 meters. The accuracy gradually decreases as the model is tested at greater distances, reaching 80.7% at 7.0 meters, 78.6% at 8.0 meters, and 75.3% at 18.0 meters.

Maximum Distance for Detecting Trash

Table 11. Maximum Distance for Detecting Trash in Meters

DISTANCE	ACCURACY
4.0 m	83.7%
4.5 m	81.9%
5.0 m	75.4%
6.0 m	73.2%
6.6 m	72.6 %

Table 11 presents the maximum distance for detecting trash in meters and corresponding accuracy percentages. At a distance of 4.0 meters, the model demonstrates an accuracy of 83.7%. The accuracy slightly decreases to

81.9% at 4.5 meters and 75.4% at 5.0 meters. As the distance increases to 6.0 meters, the accuracy drops to 73.2%, and at 6.6 meters, it is 72.6%.

CHAPTER V

SUMMARY, CONCLUSION AND RECOMMENDATION

Summary

The research on litter monitoring in a complicated setting follows an experimental design to develop a litter surveillance model that can assist sectors. The study's primary goal is to develop a model capable of recognizing and storing instances of littering by combining three distinct types of models: You Only Look Once version 8 (YOLOv8), ResNet-50, and Single Shot Detection-MobileNet version 2 (SSD-MobileNetv2).

Roboflow, an open-source image repository, obtains photos by taking photographs outside to construct a dataset. The dataset is preprocessed by annotating, scaling, and blurring the images. It is then divided into three sections: train, valid, and test sets. The study split the data into train, valid, and test sets. The train set constitute 70% of the data, followed by the valid set at 20% and the test set at 10%. The study trained the models with various hyperparameters that yielded favorable outcomes based on their previous research. The study subsequently uses the test set to evaluate the accuracy and speed of the model to determine the best-performing model. In three scenarios that were evaluated, the model was tested with one person and rubbish, multiple humans and trash, and trash concealed by a human.

The study utilized the Euclidean distance of bounding boxes, poly-fitted to measure the real-time distance of items in the camera to determine if someone littered. This approach is augmented with the Law of Cosines to calculate the distance between humans and rubbish. To minimize the chance of incorrect predictions, computer vision models employed a detection threshold of two meters. Used the date and time function and OpenCV to create time logs and picture outputs of identified littering. YOLO, the best-performing model, successfully identified trash in three different environmental situations.

Conclusion

This section presents the conclusions drawn from the objectives through a thorough examination of results and discussions, covering diverse testing methodologies, assessments, and evaluations. It is written as follows:

1. To integrate visual data from cameras and apply Law of Cosines to calculate the distance between trash and humans to create a comprehensive surveillance model capable of robust detection of littering in complex environments.

The trained model could distinguish humans and trash at first, but it had troubled identifying instances of littering. Adding the Law of Cosine improved the model, increasing its knowledge of the distance between people and trash. These improvements added to the overall effectiveness of the surveillance model by enabling accurate identification and efficient monitoring of the spatial interactions between the two classes.

2. To develop a model that integrates time monitoring and recording for securing instances of littering for documentation and analysis.

The incorporation of the datetime function was crucial in accomplishing the goal of creating a model for tracking and documenting occurrences of littering. This feature improved the model’s effectiveness as it is essential in identifying littering incidents. The model’s capacity to review instances using screenshots, videos, and Excel files guarantees the preservation of any occurrences found. It is a valuable tool for examining and evaluating possible littering incidents caught on camera.

3. To evaluate the YOLO model in terms of Accuracy and Speed in detecting littering instances in real-time.

Of all the algorithms, the YOLOv8 method is the best option because it performs exceptionally well on test like the confusion matrix, accuracy, mAP, and others. Impressing with its high efficiency, the YOLO model beats rival not only in terms of accuracy but also in terms of inference speed. YOLOv8 is the best option for real-time instances because it processes frames half as quickly as SSD-MobileNetv2 and ResNet-50. A thorough literature review confirming the algorithm's remarkable speed and accuracy in real-time applications adds credence to confirm these results. In terms of distance, YOLOv8 could detect humans and trash at a maximum distance of 6-7 meters from the camera.

Recommendation

As the study reaches its conclusion, valuable insights are extended to future researchers aiming to advance the field of littering detection through sophisticated surveillance models. Thus, the following recommendations can be made:

1. Prioritizing investment in advanced camera technology is critical for improving the capabilities of littering detection models. A high-quality camera with improved resolution and image clarity can significantly

contribute to accurately identifying and differentiating humans and various types of litter.

2. Use GAN as a preprocessing tool for enhance images of humans and trash to generate realistic and refined images.

3. To enhance the diversity and robustness of the model, future researchers should focus on expanding the dataset used for training.

4. To increase the model's accuracy, researchers must investigate various training parameters such as learning rate, batch size, and optimization techniques, allowing for fine-tuning that can considerably improve overall performance.

REFERENCES

N -Dimensional Approximation of Euclidean Distance. (2020, March 1). https://ieeexplore.ieee.org/abstract/document/8723511

2020 Index IEEE Transactions on Pattern Analysis and Machine Intelligence Vol. 42. (2021). *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *43*(1), 1–34. https://doi.org/10.1109/tpami.2020.3036557

Abdel-Shafy, H. I., & Mansour, M. S. M. (2018). Solid waste issue: Sources, composition, disposal, recycling, and valorization. *Egyptian Journal of Petroleum*, *27*(4), 1275–1290. https://doi.org/10.1016/j.ejpe.2018.07.003

Ali, M. (2022, August 22). *Supervised Machine learning*. https://www.datacamp.com/blog/supervised-machine-learning?fbclid=IwAR0ep598kbI8DTSI4XbSgqh8S49FvJD2u6-MYSGvq-Qz3iupmcN6Ieda7nQ

Ahmed, I., Ahmad, A., Rodrigues, J. J. P. C., Jeon, G., & Din, S. (2021). A deep learning-based social distance monitoring framework for COVID-19. Sustainable Cities and Society, 65, 102571. https://doi.org/10.1016/j.scs.2020.102571

Anwar, A. (2022, July 13). What is Average Precision in Object Detection & Localization Algorithms and how to calculate it? *Medium*. https://towardsdatascience.com/what-is-average-precision-in-object-detection-localization-algorithms-and-how-to-calculate-it-3f330efe697b

Behrens, T., Schmidt, K., Rossel, R. a. V., Gries, P., Scholten, T., & MacMillan, R. (2018). Spatial modelling with Euclidean distance fields and machine learning. *European Journal of Soil Science*, *69*(5), 757–770. https://doi.org/10.1111/ejss.12687

Cheung, K., Lee, C., Jun, W. W., & Lightsey, G. (2019). Single-Satellite Doppler Localization with Law of Cosines (LOC). https://ieeexplore.ieee.org/. https://doi.org/10.1109/aero.2019.8742181

Civancik-Uslu, D., Puig, R., Hauschild, M. Z., & Fullana–i–Palmer, P. (2019). Life cycle assessment of carrier bags and development of a littering indicator. *Science of the Total Environment*, *685*, 621–630. https://doi.org/10.1016/j.scitotenv.2019.05.372

Córdova, M., Pinto, A., Hellevik, C. C., Alaliyat, S., Hameed, I. A., Pedrini, H., & Da S Torres, R. (2022). Litter Detection with Deep Learning: A Comparative Study. *Sensors*, *22*(2), 548. https://doi.org/10.3390/s22020548

Dobrilović, D., Stojanov, Ž., Stojanov, J., & Malić, M. (2020). Tools for modelling distance estimation based on RSSI. *Tools for Modelling Distance Estimation Based on RSSI*. https://doi.org/10.47350/iccs-de.2020.04

Dridi, S. (2022). Supervised Learning - A Systematic Literature Review. *Supervised Learning - a Systematic Literature Review*. https://doi.org/10.31219/osf.io/tysr4

Gharib. (n.d.). Emulating the Doppler-Shift for LoRa based Low Earth Orbit Satellite Communication. *Emulating the Doppler-Shift for LoRa Based Low Earth Orbit Satellite Communication*. https://digitalcommons.calpoly.edu/eesp/605/

Gollapudi, S. (2019). Learn Computer Vision using OpenCV. In *Apress eBooks*. https://doi.org/10.1007/978-1-4842-4261-2

GPU-accelerated Faster Mean Shift with euclidean distance metrics. (2022, June 1). IEEE Conference Publication | IEEE Xplore. https://ieeexplore.ieee.org/abstract/document/9842579

Han, L., Gao, F., Zhou, B., & Shen, S. (2019). FIESTA: Fast Incremental Euclidean Distance Fields for Online Motion Planning of Aerial Robots. FIESTA: Fast Incremental Euclidean Distance Fields for Online Motion Planning of Aerial Robots. https://doi.org/10.1109/iros40897.2019.8968199

Hoffmann, M. (2019, October 7). *Generating valid Euclidean distance matrices*. arXiv.org. https://arxiv.org/abs/1910.03131

Hohmann, M., Devriendt, K., & Coscia, M. (2023). Quantifying ideological polarization on a network using generalized Euclidean distance. *Science Advances*, *9*(9). https://doi.org/10.1126/sciadv.abq2044

How do you feel about the overall looks of the app? (n.d.). https://azpdf.org/article/literature-review-on-mobilenet-literaturereview.z1dlexxv?fbclid=IwAR1CxKurubswAPNZYPgGhp4BFGm80UJSpLTgf3XeshV-KQMYV1R-HTBBmUM

Jia, T., Kapelan, Z., De Vries, R., Vriend, P., Peereboom, E. C., Okkerman, I., & Taormina, R. (2023). Deep learning for detecting macroplastic litter in water bodies: A review. *Water Research*, *231*, 119632. https://doi.org/10.1016/j.watres.2023.119632

Jiang, P., Ergu, D., Liu, F., Cai, Y., & Ma, B. (2022). A review of Yolo algorithm developments. *Procedia Computer Science*, *199*, 1066–1073. https://doi.org/10.1016/j.procs.2022.01.135

Ju, R., & Cai, W. (2023). Fracture detection in pediatric wrist trauma x-ray images using YOLOV8 algorithm. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2304.05071

Jun, W. W., Cheung, K., Lightsey, E. G., & Lee, C. (2019). Localizing in Urban Canyons using Joint Doppler and Ranging and the Law of Cosines

Method. *Proceedings of the Satellite Division's International Technical Meeting*. https://doi.org/10.33012/2019.16897

Kukil, & Kukil. (2023, August 4). *Intersection over union IOU in object detection segmentation*. LearnOpenCV – Learn OpenCV, PyTorch, Keras, Tensorflow with Examples and Tutorials. https://learnopencv.com/intersection-over-union-iou-in-object-detection-and-segmentation/

Liu, J., Pan, C., & Yan, W. Q. (2022). Litter Detection from Digital Images Using Deep Learning. *SN Computer Science*, *4*(2). https://doi.org/10.1007/s42979-022-01568-1

Lou, H., Duan, X., Guo, J., Liu, H., Gu, J., Bi, L., & Chen, H. (2023). DC-YOLOV8: Small-Size Object Detection Algorithm based on Camera Sensor. *Electronics*, *12*(10), 2323. https://doi.org/10.3390/electronics12102323

Mariyam, S., Cochrane, L., Zuhara, S., & McKay, G. (2022a). Waste Management in Qatar: A Systematic Literature review and recommendations for system strengthening. *Sustainability*, *14*(15), 8991. https://doi.org/10.3390/su14158991

Mariyam, S., Cochrane, L., Zuhara, S., & McKay, G. (2022b). Waste Management in Qatar: A Systematic Literature review and recommendations for system strengthening. *Sustainability*, *14*(15), 8991. https://doi.org/10.3390/su14158991

Ogundokun, R. O., Misra, S., Akinrotimi, A. O., & Oğul, H. (2023). MobileNet-SVM: a lightweight deep transfer learning model to diagnose BCH scans for IOMT-Based imaging sensors. *Sensors*, *23*(2), 656. https://doi.org/10.3390/s23020656

Rangel-Buitrago, N., Williams, A. T., Costa, M. F., & De Jonge, V. N. (2020). Curbing the inexorable rising in marine litter: An overview. *Ocean & Coastal Management*, *188*, 105133. https://doi.org/10.1016/j.ocecoaman.2020.105133

Real-Time Littering Detection for Smart City using Deep Learning Algorithm. (2021, March 16). IEEE Conference Publication | IEEE Xplore. https://ieeexplore.ieee.org/document/9385721

Rezatofighi, H., Tsoi, N., Gwak, J., Sadeghian, A., Reid, I. R., & Savarese, S. (2019). Generalized Intersection over Union: A Metric and A Loss for Bounding Box Regression. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1902.09630

Sasman, M., Dolan, C. B., Villegas, D. C., Eyob, E., & Barrett, C. (2021). The Influence of Marginalization on Cultural Attitudes and Trash Disposal Practices in Esfuerzo de Paraíso of the Dominican Republic: A Qualitative Interview Study. *International Journal of Environmental*

Research and Public Health, *18*(6), 2872. https://doi.org/10.3390/ijerph18062872

Shah, D. (2023, April 24). Mean Average Precision (MAP) explained: Everything you need to know. *V7*. https://www.v7labs.com/blog/mean-average-precision

Sharma, V. (2021, December 29). ResNets: Why Do They Perform Better than Classic ConvNets? (Conceptual Analysis). *Medium*. https://towardsdatascience.com/resnets-why-do-they-perform-better-than-classic-convnets-conceptual-analysis-6a9c82e06e53

Singh, J., & Kaur, R. (2021). Influencing the intention to adopt Anti-Littering Behavior: an approach with modified TPB model. *Social Marketing Quarterly*, *27*(2), 117–132. https://doi.org/10.1177/15245004211013333

Suhaimi, N. a. D., Suhaimi, D., & Abas, H. (2020). A SYSTEMATIC LITERATURE REVIEW ON SUPERVISED MACHINE LEARNING ALGORITHMS. *ResearchGate*. https://www.researchgate.net/publication/344869267_A_SYSTEMATIC_LITERATURE_REVIEW_ON_SUPERVISED_MACHINE_LEARNING_ALGORITHMS

Suharto, E., Suhartono, S., Widodo, A. P., & Sarwoko, E. A. (2020). The use of mobilenet v1 for identifying various types of freshwater fish. *Journal of Physics*, *1524*(1), 012105. https://doi.org/10.1088/1742-6596/1524/1/012105

Suwanda, R., Syahputra, Z., & Zamzami, E. M. (2020). Analysis of Euclidean distance and Manhattan distance in the K-Means Algorithm for variations number of Centroid K. Journal of Physics: Conference Series, 1566(1), 012058. https://doi.org/10.1088/1742-6596/1566/1/012058

Terven, J. R., Córdova-Esparza, D., & Romero-González, J. (2023). A comprehensive review of YOLO architectures in computer vision: from YOLOV1 to YOLOV8 and YOLO-NAS. *Machine Learning and Knowledge Extraction*, *5*(4), 1680–1716. https://doi.org/10.3390/make5040083

Tsang, S. (2022, April 2). Review — ResNet-D: Bag of Tricks for Image Classification with Convolutional Neural Networks. *Medium*. https://sh-tsang.medium.com/review-resnet-d-bag-of-tricks-for-image-classification-with-convolutional-neural-networks-7ee77669851b

Ulloa-Torrealba, Y. Z., Schmitt, A., Wurm, M., & Taubenböck, H. (2023). Litter on the streets - solid waste detection using VHR images. *European Journal of Remote Sensing*, *56*(1). https://doi.org/10.1080/22797254.2023.2176006

Wang, H., Li, K., & Xu, C. (2022). A new generation of ResNet model based on artificial intelligence and few data driven and its construction in image recognition model. *Computational Intelligence and Neuroscience*, *2022*, 1–10. https://doi.org/10.1155/2022/5976155

Wang, W., Hu, Y., Zou, T., Li, H., Wang, J., & Wang, X. (2020). A New Image Classification Approach via Improved MobileNet Models with Local Receptive Field Expansion in Shallow Layers. *Computational Intelligence and Neuroscience*, *2020*, 1–10. https://doi.org/10.1155/2020/8817849

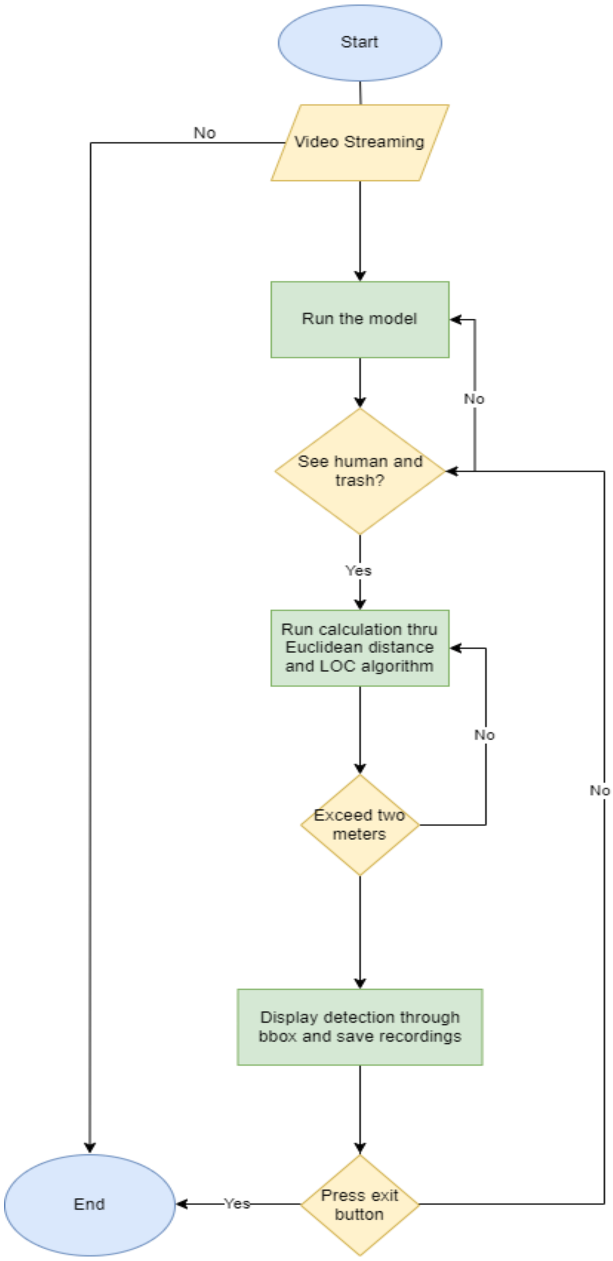
Wu, Z., Shen, C., & Van Den Hengel, A. (2019). Wider or deeper: Revisiting the RESNET model for visual recognition. *Pattern Recognition*, *90*, 119–133. https://doi.org/10.1016/j.patcog.2019.01.006

APPENDIX

A

Specification and Design Phase

Model Flowchart



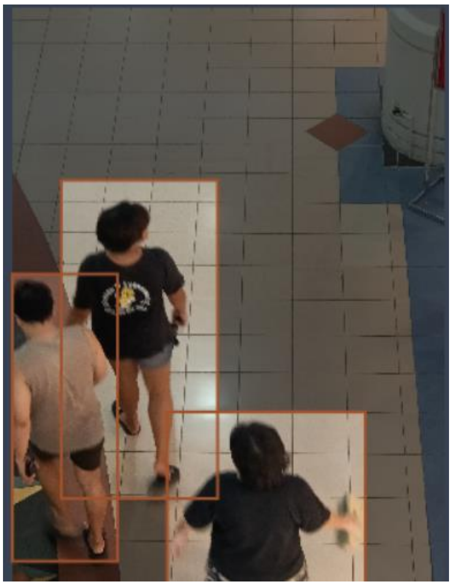
The diagram exhibits the process of the algorithm for littering detection. It starts with checking if the algorithm is connected to a video stream. It would run the model if it were connected. The model would loop until it sees a human and trash in the stream. After locating the classes, it would calculate and verify if it exceeds the distance threshold and loop until it finally reaches the condition. It would then display the boxes and save the recording after reaching the threshold. The algorithm would run until we press the exit button.

APPENDIX

B

Hardware and Software Resources

Sample Image of Human dataset



The sample image of Human dataset has a bounding box around a detected human, annotated using Roboflow. This annotation signifies the presence and location of humans, aiding in training machine learning models to recognize and analyze human behavior relevant to littering incidents.

Sample of Trash dataset

In the sample image of the Trash dataset used for littering detection, a bounding box is annotated around detected trash items, including plastic bags, trash bags, plastic bottles, and more. This annotation in Roboflow indicates the specific locations of these trash items within the image. This dataset is valuable for training litter detection models, allowing them to identify and analyze various types of litter, aiding in developing effective litter monitoring systems.

Source Code

```
import os
from ultralytics import YOLO
# from cvzone.HandTrackingModule import HandDetector
import cv2
from scipy.spatial import distance as dist
import numpy as np
import cvzone
import math
from openpyxl import load_workbook
from datetime import datetime

filepathexcel =
"C:/Users/ASUS/Desktop/detectionalgo/runs/detect/train2/litter
detect.v9i.yolov8/Book1.xlsx"
wb = load_workbook ( filepathexcel )
output_directory =
"C:/Users/ASUS/Desktop/detectionalgo/runs/detect/train2/frames"
os.makedirs(output_directory, exist_ok=True)
model = YOLO (
"C:/Users/ASUS/Desktop/detectionalgo/runs/detect/train2/weights/18k_o
penvino_model" )

y = [20, 40, 60, 80, 100, 120, 140, 160, 180, 200, 220, 240]
x1 = [245, 215, 200, 180, 160, 145, 138, 122, 114, 110, 98, 80]
x2 = [940, 933, 877, 804, 738, 678, 621, 568, 522, 479, 444, 406]

coffh = np.polyfit ( x2, y, 2 )
cofft = np.polyfit ( x1, y, 2 )

# date and time & excel sheet
sheet = wb.active
time = datetime.now ()
arraypointstime = [time]
timeanddate = arraypointstime

# frame count
frame_count =0
desired_frame_size = (1920, 1080)
# Initialize the webcam
cap = cv2.VideoCapture ( "C:/Users/ASUS/Downloads/Vid of littering
2.mp4" )
fps = int ( cap.get ( cv2.CAP_PROP_FPS ) )

fourcc = cv2.VideoWriter_fourcc ( 'm', 'p', '4', 'v' )
frame_width = int ( cap.get ( cv2.CAP_PROP_FRAME_WIDTH ) )
frame_height = int ( cap.get ( cv2.CAP_PROP_FRAME_HEIGHT ) )

out = cv2.VideoWriter ( 'runs/detect/train2/output.mp4', fourcc, fps,
```

```
(frame_width, frame_height), True )
# Main loop to continuously get frames from the webcam

while True:
    ret, frame = cap.read ()

    if not ret:
        print ( "Failed to grab frame" )
        break

    # Predict object bounding boxes using YOLO
    results = model.predict ( frame, stream = True, conf = 0.15,
max_det = 10 )
    arraypoints = [0]
    arraypoints2 = [0]
    minus = []

    for result in results:
        boxes = result.boxes.cpu ().numpy ()
        for box in boxes:
            humanbox = box.xyxy[0].astype ( int )

            b = str ( result.names[int ( box.cls[0] )] )
            cv2.rectangle ( frame, tuple ( humanbox[:2] ), tuple (
humanbox[2:] ), (0, 0, 0), 5 )
            font = cv2.FONT_HERSHEY_PLAIN
            cv2.putText ( frame, str ( datetime.now () ), (20, 40),
                         font, 2, (0, 255, 0), 2, cv2.LINE_AA )
            cls = int ( box.cls[0] )
            # Calculate the center x and y coordinates for the
bounding box
            if b == "Human":
                hpoint1 = (humanbox[0], humanbox[1])
                hpoint2 = (humanbox[2], humanbox[3])
                euclidean_distancehuman = dist.euclidean ( hpoint1,
hpoint2 )
                A, B, C = coffh
                humandistance = A * euclidean_distancehuman ** +B *
euclidean_distancehuman + C
                arraypoints.append ( humandistance )

            if b == "Trash":
                print ( "trash" )
                tpoint1 = (humanbox[0], humanbox[1])
                tpoint2 = (humanbox[2], humanbox[3])
                euclidean_distancetrash = dist.euclidean ( tpoint1,
tpoint2 )
                X, Y, Z = coffh
                trashdistance = X * euclidean_distancetrash ** +Y *
euclidean_distancetrash + Z
                arraypoints2.append ( trashdistance )
                print ( trashdistance, "trashdis" )

        euclihuman = int ( np.mean ( arraypoints ) )
        eudlitrash = int ( np.mean ( arraypoints2 ) )
        print ( "trash", eudlitrash )
        print ( "human", euclihuman )
        if eudlitrash < euclihuman and euclihuman != 0 and eudlitrash
!= 0:
            cosineangle = (eudlitrash / euclihuman)
            distance_squared = (eudlitrash ** 2 + euclihuman ** 2) -
```

```
(2 * eudlitrash * euclihuman) * cosineangle

            # Ensure that the value inside math.sqrt is non-negative
            distance = math.sqrt ( max ( distance_squared, 0 ) )

            if distance >= 39:
                (cv2.rectangle ( frame, tuple ( humanbox[:2] ), tuple
( humanbox[2:] ), (0, 0, 255), 5 ))
                #out.write ( cv2.rectangle ( frame, tuple (
humanbox[:2] ), tuple ( humanbox[2:] ), (0, 0, 0), 5 ) )
                frame_count +=1
                resized_frame = cv2.resize ( frame,
desired_frame_size )
                cv2.rectangle ( resized_frame, tuple ( humanbox[:2]
), tuple ( humanbox[2:] ), (0, 0, 255),5 )  # Red rectangle
                frame_path = os.path.join ( output_directory,
f"frame_{frame_count:04d}.jpg" )
                cv2.imwrite ( frame_path, resized_frame )
                # saving time
                data = sheet.append ( timeanddate )
                print ( "littering" )

        if eudlitrash < euclihuman and euclihuman != 0 and eudlitrash
!= 0:
            cosineangle = (eudlitrash / euclihuman)
            distance_squared = (eudlitrash ** 2 + euclihuman ** 2) -
(2 * eudlitrash * euclihuman) * cosineangle

            # Ensure that the value inside math.sqrt is non-negative
            distance = math.sqrt ( max ( distance_squared, 0 ) )
            if distance >= 39:  # equals to 2m
                # detecting
                (cv2.rectangle ( frame, tuple ( humanbox[:2] ), tuple
( humanbox[2:] ), (0, 0, 255), 5 ))
                #out.write ( cv2.rectangle ( frame, tuple (
humanbox[:2] ), tuple ( humanbox[2:] ), (0, 0, 0), 5 ) )
                frame_count += 1
                resized_frame = cv2.resize ( frame,
desired_frame_size )
                cv2.rectangle ( resized_frame, tuple ( humanbox[:2]
), tuple ( humanbox[2:] ), (0, 0, 255),5 )  # Red rectangle
                frame_path = os.path.join ( output_directory,
f"frame_{frame_count:04d}.jpg" )
                cv2.imwrite ( frame_path, resized_frame )
                out.write ( frame )
                # saving time
                data = sheet.append ( timeanddate )
                print ( "littering" )

    out.write ( frame )
    cv2.imshow ( "Detecting", frame )
    wb.save ( filepathexcel )
    if cv2.waitKey ( 1 ) & 0xFF == ord ( 'x' ):
        wb.save ( filepathexcel )
        out.write ( frame )
        break

# Release the webcam and destroy OpenCV windows
cap.release ()
cv2.destroyAllWindows ()
```

APPENDIX

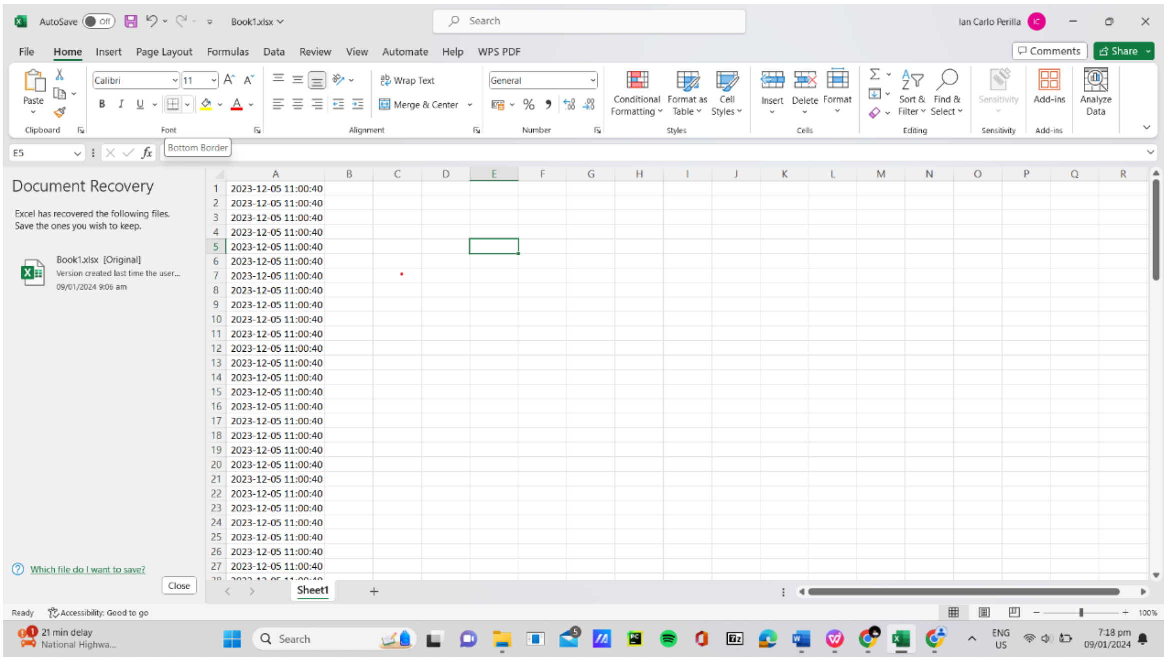
C

Input/Output Screenshots

Raw Data

The raw data used in the study comprises 3,000 images of humans and 6,000 images of trash, captured in diverse angles and environmental settings. This raw data includes a variety of perspective and scenarios, providing a comprehensive representation of humans and different types of litter in various real-world conditions.

Timestamp



The timestamp feature used in the study involves recording incidents of littering in an Excel file, capturing both the data and time of each occurrence. This temporal data provides a chronological record of when littering incidents occurred, offering valuable insights into patterns, trends, and potential factors influencing littering behavior. Analyzing the timestamped data can contribute to a better understanding of temporal aspects of littering, aiding in developing more effective litter prevention and management strategies.

APPENDIX

D

Testing and Evaluation of the Algorithms

Testing of YOLO Model

In the study using YOLO model, the testing process involves assessing the model's performance on a separate dataset not used during training. The model makes predictions on litter presence and location, its accuracy is evaluated against ground truth annotations using metrics like precision and recall. Findings from this testing phase inform potential model adjustment and optimizations for effective litter detection.

Confusion Matrices Scores

YOLOv8		ResNet- 50		SSD-MobileNetv2	
257*	88	59	286	80	264
56	455	152	490	131	512*

The superscript * indicates that the model exhibits high True Positive and Negative detections, coupled with low False Negative and Positive detections.

The confusion matrix score for YOLOv8, ResNet-50, and SSD-MobileNetv2 show their respective True Positive and Negative detections, highlighted with superscript * to indicate strong performance. In contrast, low

False Negative and Positive detections suggest these models accurately identify and exclude litter in the study.

Evaluation Results of Accuracy

Model	Precision	Recall	F1-score	mAP	Accuracy
YOLOv8	0.8211	0.7449	0.7812	84.50%	83.18%
ResNet- 50	0.2796	0.1710	0.2122	23.09%	55.62%
SSD-MobileNetv2	0.3791	0. 2326	0.2882	29.46%	60.00%

The evaluation results indicate that the YOLO model outperforms ResNet-50 and SSD-MobileNetv2 in precision, recall, F1-score, mean Average Precision (mAP), and overall accuracy detecting litter in the study. ResNet-50 and SSD-MobileNetv2 show lower performance across these metrics, with YOLO achieving the highest accuracy at 83.18%.

Evaluation Results of Speed

Model	Inference Time	Frames per second (fps)
YOLOv8	90 ms	30
ResNet- 50	250 ms	15
SSD-MobileNetv2	200 ms	15

The evaluation results indicate that YOLO has the fastest inference time 90 ms, achieving 30 frames per second (fps), outperforming ResNet-50 and SSD-MobileNetv2, which have slower inference times of 250 ms and 200 ms, respectively, resulting in 15 fps for each.

APPENDIX E

Publishable Paper

UTILIZATION AND OPTIMIZATION OF DEEP LEARNING ALGORITHMS IN CLASSIFYING LITTERING IN A COMPLEX ENVIRONMENT

Ian Carlo R. Perilla, Jerika G. Mana-ay, Ericson Z. Recto, Jonardo R. Asor
College of Computer Studies (CCS), Laguna State Polytechnic University, Los Baños, Laguna, Philippines

ARTICLE INFO

Keywords:
Deep Learning
Real-time Littering Detection
ResNet-50
SSD-MobileNetv2
YOLOv8

ABSTRACT

The study explores the use of deep learning algorithms and the challenges related to technology adoption in littering detection. By effectively employing a camera as a surveillance tool to monitor specific littering activities from a distance, the collected diverse datasets encompassing plastic bags, trash bags, plastic bottles, and more underwent preprocessing using Roboflow to ensure the quality of the images. The dataset comprises 3,000 images of Humans and 6,000 images of Trash. Divide it into three sets: allocate 70% of the data to the training set, assign 20% to the test set, and reserve the remaining 10% to validate the model's predictions. A comprehensive analysis comparing three leading deep learning algorithms—YOLOv8, ResNet-50, and SSD-MobileNetv2—covers two critical dimensions. The first dimension involves assessing their performance in accurately classifying littering incidents, while the second focuses on evaluating their computational efficiency for real-time applications. The findings indicate YOLO as the optimal choice for real-time littering detection, boasting an 82% accuracy in classifying Trash and Humans, with an impressive 90ms inference speed during real-time applications. On the other hand, SSD MobileNetv2 is the second-best model to perform real-time litter detection with an accuracy of 60% and 200ms inference time.

Corresponding Author:
Mana-ay, Jerika G.
College of Computer Studies,
Laguna State Polytechnic University,
Malinta, Los Baños, Laguna 4030.
Email: jerikamanaay01@gmail.com

1. INTRODUCTION

Waste pollution remains a significant issue in the Philippines, made worse by rapid industrialization and urbanization. Ineffective policies have caused mass pollution, and the issue is getting worse due to poor waste management facilities and illicit dumping. Littering, fueled by negligence and inadequate infrastructure, endangers the environment and public health in addition to ruining cityscapes. Although there have been attempts to reduce littering, such as enacting Republic Act No. 9003, difficulties still exist.

Littering is a significant environmental issue in the Philippines despite efforts to reduce it. According to studies, humans consume significant amounts of plastic waste, exacerbating litter on beaches and oceans. To preserve a sustainable and healthy environment, law enforcement, public education efforts, and appropriate waste disposal techniques must be used to combat illegal littering.

The study intends to innovatively monitor and identify litterers using computer vision techniques such as YOLO, ResNet-50, and SSD-MobileNetv2. Authorities developed and test algorithms to effectively control littering by distinguishing between different types of waste and providing real-time results.

Statement of Objectives

This study aims to provide a robust model for detecting instances of littering within intricate and challenging environments. Thus, this study specifically aims to:

1. To integrate visual data from cameras and apply the Law of Cosine to calculate the distance between trash and humans to create a comprehensive surveillance model capable of robust detection of littering in complex environments.
2. To develop a model that integrates time monitoring and recording for securing instances of littering for documentation and analysis.
3. To evaluate the YOLO model in terms of Accuracy and Speed in detecting littering instances in real-time.

2. RELATED LITERATURE

YOLO

YOLO (You Only Look Once) is a family of fast object detection neural networks capable of achieving real-time performance by dividing an input image into grid cells. In each cell, YOLO predicts a bounding box and object classification [1]. YOLO-v5 stands out as the most promising approach to date for object detection due to its fast-training stage and superior performance compared to previous versions. This achievement is realized through genetic programming to adjust the set of prior boxes. YOLO operates as a one-stage detector [2], simultaneously training bounding box coordinates and corresponding class label probabilities. This design choice makes YOLO faster than two-stage detectors like R-CNN, necessitating an external region proposal algorithm. YOLO finds wide applications in various fields, including autonomous driving, surveillance, and robotics.

In their study, researchers defined YOLO as a real-time object detection system based on deep neural networks [3]. Scientists extensively use YOLO for detecting microplastics in water bodies and studies related to detection systems based on deep neural networks. Furthermore, as described [4], YOLO is an object detection algorithm that simultaneously performs object detection and classification using a convolutional neural network (CNN) architecture. Being a single-stage detector, it divides an image into a grid to predict bounding boxes and class probabilities for each grid cell, earning YOLO its reputation for speed and accuracy in object detection tasks.

ResNet

ResNet is an abbreviation for Residual Neural Network. This deep learning architecture addresses the vanishing gradient problem by introducing shortcut connections between layers, enabling the training of intense neural networks [5]. To enhance ResNet's performance with limited data, [6] modified the architecture, addressing a common challenge in the practical applications of deep learning.

ResNet-50 and ResNet-101 are backbone models in the proposed technique for litter identification from digital images using machine learning in the paper by [7]. The study by [6] presented several modifications to the original ResNet architecture, including a new bottleneck structure and dense connections, aimed at reducing the number of parameters and improving the model's efficiency. They evaluated the performance of their modified ResNet models on several benchmark datasets for image classification tasks and compared them with other state-of-the-art models.

MobileNet

In the study by [8], MobileNet is introduced as a convolutional neural network architecture that utilizes depth-wise separable convolution, except for its initial layer. Unlike traditional convolutional networks, depth-wise separable convolution reduces the computational requirements for image convolution. Andrew G. Howard and his team created MobileNet v1, which enables users to perform object detection and classification tasks.

The architecture of MobileNet incorporates depth-wise separable convolution, comprising two components: depth-wise convolution and 1x1 pointwise convolution [9]. Inception-ResNet-v2, another significant design, combines the benefits of residual and inception networks, showcasing superior feature extraction capabilities for BCH images using convolutional neural networks (CNNs). [10]

Euclidean Distance Algorithm

To calculate the distance between two points, one needs to square the x and y coordinates of both points, then add the results together and find the square root of the sum. This formula is also like the Euclidean distance algorithm. The formula is also like the Euclidean distance algorithm. The algorithm finds the shortest path between the two points in a cartesian coordinate plane. The formula can be written as:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Law of Cosines

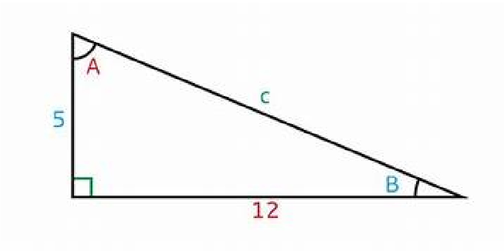


Figure 1. Law of Cosines

The diagram in Figure 1 illustrates the use of the Law of Cosines (LOC) to determine the length of the third side ‘c’ or distance in a triangle. The algorithms use all the information, employing the sides ‘a’ and ‘b’ and the angle ‘c’. The algorithm calculates side ‘c’ length by applying the LOC. The approach also uses the Side-Angle-Side (SAS) theorem to calculate the length of the remaining side of the triangle using the unknown values of side ‘b’, angle ‘c’, and side ‘a’. With this all-encompassing method, the algorithm can effectively extract the geometric relationships inside the triangle, making it easier to determine distance accurately by combining LOC and SAS concepts.

Conceptual Framework

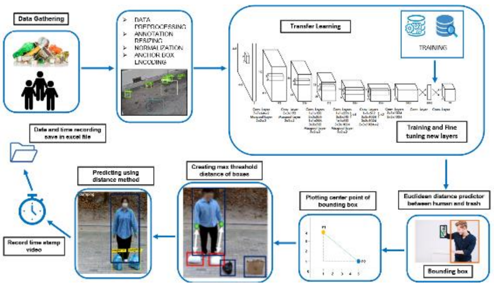


Figure 2. Conceptual Framework

The model used two kinds of data, primary and secondary data, which would process. The images annotated and given bounding boxes. The data set to numerical values through annotation by determining the point of bounding boxes. First, the study used the images to customize pre-trained models by adding extra layers to the frozen model weights. After detecting the boxes, the model used Euclidean distance to determine the distance between the human and the trash. The box’s center would be acquired and converted to numerical values. Then, the system used the numerical values’ distance to create a certain threshold, determining whether someone is littering by detecting if humans and trash detected reached the maximum distance. The time module record the time of littering occurrence and save it in an xlsx file, which anyone can use to review whether someone is littering.

3. METHODOLOGY

3.1 Research Design

This study employed an experimental research design, manipulating one or more variables to observe their effect on another variable [11]. The researchers manipulated the input and measured the system’s output accuracy and speed to determine the YOLOv8 algorithm.

3.2 Data

The study utilized the dataset sourced from the open-source website Roboflow. Roboflow provides diverse datasets of annotated trash, convertible to various formats. Additionally, the platform enables the uploading and annotation of datasets. The dataset includes images from Roboflow and images captured in different locations, featuring trash and humans—the two classes under detection in the study.

3.3 Model Development

Model development is a crucial aspect when creating an algorithm for the study. To ensure the selection of appropriate models, various algorithms suitable for image recognition of trash were filtered and narrowed down.

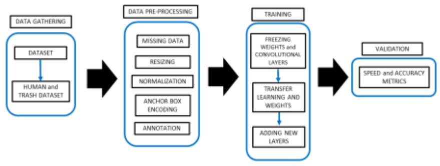


Figure 3. Model Development

Figure 3 illustrates the baseline for the algorithms the study used in this investigation. The study employ various strategies to preprocess the dataset using different designs and pipelines to fine-tune the dataset incorporated within the model.

Solution

Figure 4. Solution for Littering

Following the training phase on the dataset containing images of humans and trash, the researchers plan to enhance the model by incorporating advanced algorithms for spatial awareness. This integration enabled the model to identify individual instances of litter and understand the spatial relationships between humans and trash in a given environment. By leveraging this spatial awareness, the model aims to provide a more nuanced and context-aware litter detection system capable of discerning intentional littering behaviors and improving overall accuracy in varied scenarios.

Experiment

Table 1. List of Hyperparameter Controlled Variables

HYPERPARAMETER	YOLO
Learning rate	0.001
Batch size	16
Number of epochs	100
Confidence threshold	0.25
Weight decay	0.0005
Optimizer	SGD
IOU	0.7

The experiment in this study involved the utilization of various hyperparameters. Table 1 displays the hyperparameters intended for use in the study, sourced from different studies. The hyperparameters utilized in YOLO are based on the study of [11]. This study underscores the significance of YOLO in the realm of real-time applications, attributed to its of optimal trade-off between model accuracy and inference speed.

Model Evaluation

The accuracy matrix uses two tests, Intersection over Union (IOU) and Mean Average Precision (mAP), to determine the model's accuracy by obtaining the actual positive and negative values for litter detection. The mAP utilizes the data from IOU to compute the average precision of models. It uses inference time and frames per second (fps) for speed. Inference time refers to the time the model needs to predict, while fps represents the number of frames the model can process.

Mean Average Precision (mAP)

An evaluation metric commonly used for object detection. The mAP takes into account the trade-off between precision and recall, considering both false positives (FP) and false negatives (FN). This characteristic makes mAP a suitable metric for most detection application [12].

$$mAP = \frac{1}{k}\sum_{i}^{k} AP_i$$

The mean average precision (mAP) is the mean of all precision values across all classes. AP stand for Average Precision, which is determined separately for every class, where k is the total amount of classes.

$$Precision = \frac{TP}{TP + FP}$$

The precision is computed by dividing the number of correct predictions by the total number of correct and incorrect forecasts. In the context of the study, precision represents the accuracy of the model in detecting objects, specifically littering and other objects. A higher precision rate indicates a higher level of accuracy for the model, and vice versa.

$$Recall\ (TPR) = \frac{TP}{TP + FN}$$

In contrast to accuracy, the proportion of actual positive predictions to the total of true positive and false negative guesses determines recall. The model's ability to properly detect all instances of genuine litter is measured by recall.

F1-score

$$f1 = \frac{2*(Precision\ Recall)}{Precision+Recall}$$

It evaluates the performance of two classifiers using precision and recall, with the F1 score reaching a minimum of 1 and a minimum of 0. Precision and recall both contribute equally to the F1-score as a percentage, which is calculated using the specified procedure.

Accuracy

$$Accuracy = \frac{(TP + TN)}{TP + FP + TN + FN}$$

The study by [13] accuracy as the measure of how well three different machine learning techniques can predict clinical outcomes after shoulder arthroplasty. It refers to the ability of the machine learning models to correctly predict the outcome of the surgery based on the input data.

Confusion Matrix

This matrix compiles all the performance of the model based on the test data. In the study, it used a multiclass confusion matrix which helps in the classification of trash that is evident in the study. According to [14] research, the confusion matrix is a matrix that provides the mix of expected vs real class occurrences, and it plays an important part in evaluating algorithm performance.

Speed Matrix

Inference time. It refers to time needed for the model to predict or infer the occurrence of littering. The study utilized the model's results, as demonstrated in YOLOv8 where it was used to analyze the inference time by processing video frames and predicting the objects within each frame.

Frames per second (fps). The number of frames a model can process in real-time or within a video. The study utilized this metric to evaluate the model's ability to efficiently handle a large volume of images in a matter of seconds.

4. RESULTS AND DISCUSSION

Objective 1. Integrating camera visual data with the Law of Cosine to compute distances between trash and humas results in a viable surveillance model for litter detection in complex environment.

Integration of Visual Data from Cameras

Figure 5. Visual Data from Camera

The researchers used OpenCV to capture from the camera and create a visual representation of the surroundings. The system converts the raw data into frames, each comprising pixels. In this case, the video pixel dimensions are 1152 x 640 pixels, as shown in Figure 5.

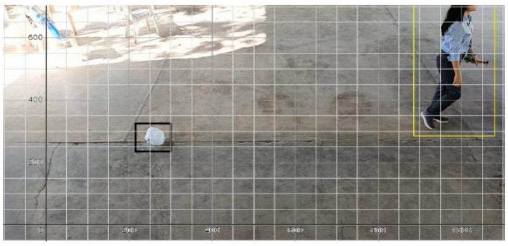


Figure 6. Cartesian Coordinates of Pixels

Figure 6 shows that the model used the pixels to convert coordinates represented as numerical data similar to Cartesian coordinates by obtaining the width and height of the pixels. In the study [15] pixels from OpenCV can be used to convert Cartesian coordinates with a 1:1 ratio.

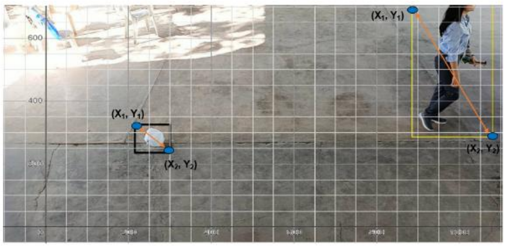


Figure 7. Computed Euclidean Distance using dist. function in Python

As shown in Figure 7, the dist. function provides the coordinates of the two points of the bounding boxes. The bounding box coordinates are in the format '(x1, y1, x2, y2) where (x1, y1) is the top-left corner, and (x2, y2) is the bottom-right corner. $d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$is the formula used by dist. function to calculate the pixel distance between objects. The distance obtained would be poly-fitted to convert the pixel size to the approximate distance between human and trash. The model utilized these coordinates to determine instance of littering.

Application of Law of Cosine to Calculate the Distance Between Trash and Humans

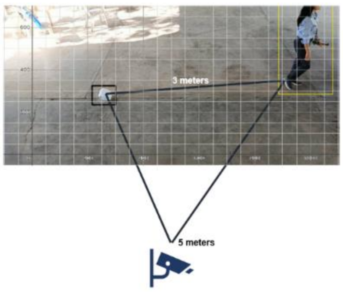
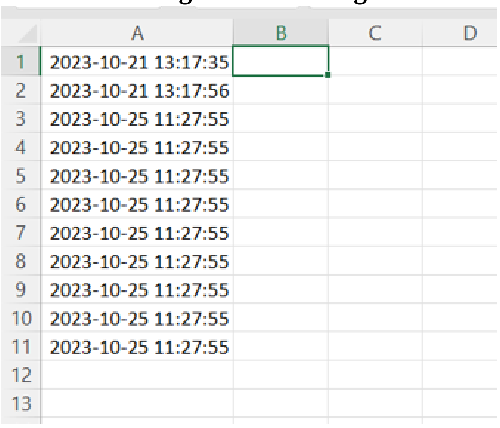


Figure 8. Law of Cosine Application

Figure 8 illustrates the application of the LOC to calculate the distance between Trash and Humans in the study. The model utilized the distance between the camera and the two detected objects and inputted these distances into $c = \sqrt{a^2 + b^2 + 2\,ab\cos y}$ the LOC formula. Wherein ‘c’ is the opposite side, which in the study is the distance between the two classes is equal to the square root of the sum of squared distance between camera and trash (a), camera and human (b) subtracted by the twice of ab multiplied to cosine γ. The model compares which length is more significant than between a and b, which switches the values to prevent miscalculations.

Objective 2. To develop a model that integrates the monitoring and recording to secure instances of littering for documentation and analysis.

Time Monitoring and Recording



	A	B	C	D
1	2023-10-21 13:17:35			
2	2023-10-21 13:17:56			
3	2023-10-25 11:27:55			
4	2023-10-25 11:27:55			
5	2023-10-25 11:27:55			
6	2023-10-25 11:27:55			
7	2023-10-25 11:27:55			
8	2023-10-25 11:27:55			
9	2023-10-25 11:27:55			
10	2023-10-25 11:27:55			
11	2023-10-25 11:27:55			
12				
13				

Figure 9. Date and time Recording saves in xlsx file

The model utilized the 'open workbook' function to open an Excel file and save the time and date of the littering instances, as shown in Figure 9. Furthermore, OpenCV integrates timestamping into the output video, enabling one to rewind the video to the exact date and time of the littering occurrence.

Image and Video Recording

Figure 10. Snapshot of Littering being recorded by the Model

The model identifies the two classes, and if a person departs or throws garbage in the frame, the color of the trash bounding box would change to red. This process generates photos that can help identify littering. The mentioned timestamp can also serve as reference for video replay to identify the person responsible for the trash in the areas, as illustrated in Figure 10.

Objective 3. To evaluate the YOLO model in terms of Accuracy and Speed in detecting littering instances in real-time.

Confusion Matrices Scores

Table 2. Confusion Matrix of YOLO model

YOLO	
257	88
56	455

Table 2 presents the Confusion Matrix for the YOLO model, revealing its performance on a classification task. Specifically, the YOLO model achieved 257 true positives, 455 true negatives, 88 false positives, and 56 false negatives, indicating its ability to predict positive and negative instances accurately and where it may have made classification errors.

Evaluation of YOLO model in terms of Accuracy Performance

Table 3. Results of Evaluation

Model	Precision	Recall	F1-score	mAP	Accuracy
YOLO	0.8211	0.7449	0.7812	84.50%	83.18%

Table 3 presents the evaluation results for the YOLO model, showcasing various performance metrics. The precision of the model is 0.8211, indicating the accuracy of optimistic predictions, while the recall is 0.7449, representing the model’s ability to capture all positive instances. The F1-score, a harmonic mean of precision and recall, is 0.7812, and the model achieved an accuracy of 83.18%, with a mAP of 84.50%, providing a comprehensive overview of its effectiveness in object detection.

Evaluation of YOLO in terms of Speed Performance

Table 4. Results of Speed Performance

Model	Inference Time	Frames per second (fps)
YOLO	90 ms	30

Table 4 presents the speed evaluation results for the YOLO model, which details its inference time and corresponding frames per second (fps) performance. The model demonstrates an inference time of 90 milliseconds, indicating the average it takes to process a single image. With a commendable frame-per-second rate of 30, the YOLO model showcases its real-time processing capability. It is processing 30 frames in one second, making it suitable for applications requiring quick and efficient object detection.

Maximum Distance for Detecting Humans

Table 5. Maximum Distance for Detecting Humans in Meters

DISTANCE	ACCURACY
4.0 m	85.8%
5.0 m	84.6%
7.0 m	80.7%
8.0 m	78.6%
18.0 m	75.3%

Table 5 outlines the maximum distances for detecting humans, measured in meters, along with corresponding accuracy percentages. At a distance of 4.0 meters, the system achieves an accuracy of 85.8%. The accuracy remains at 84.6% when the distance increases to 5.0 meters. The accuracy gradually decreases as the model is tested at greater distances, reaching 80.7% at 7.0 meters, 78.6% at 8.0 meters, and 75.3% at 18.0 meters.

Maximum Distance for Detecting Trash

Table 6. Maximum Distance for Detecting Trash in Meters

DISTANCE	ACCURACY
4.0 m	83.7%
4.5 m	81.9%
5.0 m	75.4%
6.0 m	73.2%
6.6 m	72.6 %

Table 6 presents the maximum distance for detecting trash in meters and corresponding accuracy percentages. At a distance of 4.0 meters, the model demonstrates an accuracy of 83.7%. The accuracy slightly decreases to 81.9% at 4.5 meters and 75.4% at 5.0 meters. As the distance increases to 6.0 meters, the accuracy drops to 73.2%, and at 6.6 meters, it is 72.6%.

5. SUMMARY, CONCLUSION AND RECOMMENDATION

5.1 Summary

The research on litter monitoring in a complicated setting follows an experimental design to develop a litter surveillance model that can assist sectors. The study's primary goal is to develop a model capable of recognizing and storing instances of littering by combining three distinct types of models: You Only Look Once version 8 (YOLOv8), ResNet-50, and Single Shot Detection-MobileNet version 2 (SSD-MobileNetv2).

Roboflow, an open-source image repository, obtains photos by taking photographs outside to construct a dataset. The dataset is preprocessed by annotating, scaling, and blurring the images. It is then divided into three sections: train, valid, and test sets. The study split the data into train, valid, and test sets. The train set constitute 70% of the data, followed by the valid set at 20% and the test set at 10%. The study trained the models with various hyperparameters that yielded favorable outcomes based on their previous research. The study subsequently uses the test set to evaluate the accuracy and speed of the model to determine the best-performing model. The researchers evaluate the models in three scenarios: one person and rubbish, multiple humans and trash, and trash concealed by a human.

The study utilized the Euclidean distance of bounding boxes, poly-fitted to measure the real-time distance of items in the camera to determine if someone littered. This approach is augmented with the Law of Cosines to calculate the distance between humans and rubbish. To minimize the chance of incorrect predictions, computer vision models employed a detection threshold of two meters. The researchers used the date and time function and OpenCV to create time logs and picture outputs of identified littering. YOLO, the best-performing model, successfully identified trash in three different environmental situations.

5.2 Conclusion

This section presents the conclusions drawn from the objectives through a thorough examination of results and discussions, covering diverse testing methodologies, assessments, and evaluations. It is written as follows:

1. To integrate visual data from cameras and apply Law of Cosines to calculate the distance between trash and humans to create a comprehensive surveillance model capable of robust detection of littering in complex environments.

The trained model could distinguish humans and trash at first, but it had troubled identifying instances of littering. Adding the Law of Cosine improved the model, increasing its knowledge of the distance between people and trash. These improvements added to the overall effectiveness of the surveillance model by enabling accurate identification and efficient monitoring of the spatial interactions between the two classes.

2. To develop a model that integrates time monitoring and recording for securing instances of littering for documentation and analysis.

The incorporation of the datetime function was crucial in accomplishing the goal of creating a model for tracking and documenting occurrences of littering. This feature improved the model's effectiveness as it is essential in identifying littering incidents. The model's capacity to review instances using screenshots, videos, and Excel files guarantees the preservation of any occurrences found. It is a valuable tool for examining and evaluating possible littering incidents caught on camera.

3. To evaluate the YOLO model in terms of Accuracy and Speed in detecting littering instances in real-time.

Of all the algorithms, the YOLOv8 method is the best option because it performs exceptionally well on test like the confusion matrix, accuracy, mAP, and others. Impressing with its high efficiency, the YOLO model beats rival not only in terms of accuracy but also in terms of inference speed. YOLOv8 is the best option for real-time instances because it processes frames half as quickly as SSD-MobileNetv2 and ResNet-50. A thorough literature review confirming the algorithm's remarkable speed and accuracy in real-time applications adds credence to confirm these results. In terms of distance, YOLOv8 could detect humans and trash at a maximum distance of 6-7 meters from the camera.

5.3 Recommendation

As the study reaches its conclusion, the researchers extend valuable insights to future researchers aiming to advance the field of littering detection through sophisticated surveillance models. Thus, the researchers recommend the following:

1. Prioritizing investment in advanced camera technology is critical for improving the capabilities of littering detection models. A high-quality camera with improved resolution and image clarity can significantly contribute to accurately identifying and differentiating humans and various types of litter.

2. Use GAN as a preprocessing tool for enhance images of humans and trash to generate realistic and refined images.

3. To enhance the diversity and robustness of the model, future researchers should focus on expanding the dataset used for training.

4. To increase the model's accuracy, researchers must investigate various training parameters such as learning rate, batch size, and optimization techniques, allowing for fine-tuning that can considerably improve overall performance.

ACKNOWLEDGEMENTS

The Researchers would like to express their deepest gratitude to the Almighty God for providing them with the strength, wisdom, and guidance throughout this research journey.

Special thanks to their dedicated adviser, Mr. Jonardo R. Asor, whose valuable insights and unwavering support played a crucial role in the success of this thesis.

They extend their appreciation to the esteemed faculty members, Mr. Gene Marck B. Catedrilla, Ms. Carolina R. Joval, and Ms. Jocelyn O. Padallan, for their continuous encouragement and scholarly contributions that enriched their academic experience.

Heartfelt thanks to their families for their understanding, patience, and constant encouragement. Their support has been a pillar of strength, fueling their perseverance.

Lastly, the Researchers acknowledge themselves for the determination, hard work, and collaborative spirit invested in this endeavor. Together, they have navigated challenges and celebrated achievements, making this research a collective triumph.

REFERENCES

[1] M. Córdova *et al.*, "Litter Detection with Deep Learning: A Comparative Study," *Sensors*, vol. 22, no. 2, p. 548, Jan. 2022, doi: 10.3390/s22020548.

[2] B. Korchani and K. Sethom, "Real-Time Littering Detection for Smart City using Deep Learning Algorithm," *IEEE*, Mar. 2021, doi: 10.1109/iccspa49915.2021.9385721.

[3] T. Jia *et al.*, "Deep learning for detecting macroplastic litter in water bodies: A review," *Water Research*, vol. 231, p. 119632, Mar. 2023, doi: 10.1016/j.watres.2023.119632.

[4] Y. Z. Ulloa-Torrealba, A. Schmitt, M. Wurm, and H. Taubenböck, "Litter on the streets - solid waste detection using VHR images," *European Journal of Remote Sensing*, vol. 56, no. 1, Feb. 2023, doi: 10.1080/22797254.2023.2176006.

[5] F. Hussain, S. G. Abbas, M. Husnain, U. U. Fayyaz, F. Shahzad, and G. A. Shah, "IoT

DoS and DDoS Attack Detection using ResNet," *IEEE*, Nov. 2020, doi: 10.1109/inmic50486.2020.9318216.

[6] H. Wang, K. Li, and C. Xu, "A new generation of ResNet model based on artificial intelligence and few data driven and its construction in image recognition model," *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1–10, Mar. 2022, doi: 10.1155/2022/5976155.

[7] P. Jiang, D. Ergu, F. Liu, Y. Cai, and B. Ma, "A review of Yolo algorithm developments," *Procedia Computer Science*, vol. 199, pp. 1066–1073, Jan. 2022, doi: 10.1016/j.procs.2022.01.135.

[8] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," *arXiv (Cornell University)*, Jan. 2018, doi: 10.48550/arxiv.1801.04381.

[9] E. Suharto, S. Suhartono, A. P. Widodo, and E. A. Sarwoko, "The use of mobilenet v1 for identifying various types of freshwater fish," *Journal of Physics: Conference Series*, vol. 1524, no. 1, p. 012105, Apr. 2020, doi: 10.1088/1742-6596/1524/1/012105.

[10] R. O. Ogundokun, S. Misra, A. O. Akinrotimi, and H. Oğul, "MobileNet-SVM: a lightweight deep transfer learning model to diagnose BCH scans for IOMT-Based imaging sensors," *Sensors*, vol. 23, no. 2, p. 656, Jan. 2023, doi: 10.3390/s23020656.

[11] R.-Y. Ju and W. Cai, "Fracture detection in pediatric wrist trauma x-ray images using YOLOV8 algorithm," *arXiv (Cornell University)*, Apr. 2023, doi: 10.48550/arxiv.2304.05071.

[12] D. Shah, "Mean Average Precision (MAP) explained: Everything you need to know," *V7*, Apr. 24, 2023. https://www.v7labs.com/blog/mean-average-precision

[13] C. K. B, R. Punitha, and Mohana, "YOLOV3 and YOLOV4: Multiple object detection for surveillance applications," *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*, Aug. 2020, doi: 10.1109/icssit48917.2020.9214094.

[14] I. Markoulidakis, G. Kopsiaftis, I. Rallis, and I. Georgoulas, "Multi-Class Confusion Matrix Reduction method and its application on Net Promoter Score classification problem," *ACM Digital Library*, Jun. 2021, doi: 10.1145/3453892.3461323.

[15] S. Gollapudi, *Learn Computer Vision using OpenCV*. 2019. doi: 10.1007/978-1-4842-4261-2.

APPENDIX

F

Grammar Rate

grammarly Report: THESIS_CHAPTER 1-5_part1

THESIS_CHAPTER 1-5_part1

by B

General metrics

66,466	9,869	760	39 min 28 sec	1 hr 15 min
characters	words	sentences	reading time	speaking time

Score

98

This text scores better than 98% of all texts checked by Grammarly

Writing Issues

100	58	42
Issues left	Critical	Advanced

Plagiarism

This text hasn't been checked for plagiarism

grammarly Report: THESIS_CHAPTER 1-5_part2

THESIS_CHAPTER 1-5_part2

by B

General metrics

48,358	6,878	809	27 min 30 sec	52 min 54 sec
characters	words	sentences	reading time	speaking time

Score

96

This text scores better than 96% of all texts checked by Grammarly

Writing Issues

100	66	34
Issues left	Critical	Advanced

Plagiarism

This text hasn't been checked for plagiarism

APPENDIX

G

Plagiarism Checker

UTILIZATION AND OPTIMIZATION OF DEEP LEARNING ALGORITHMS IN CLASSIFYING LITTERING IN A COMPLEX ENVIRONMENT(Perilla,Recto,Manaay)

ORIGINALITY REPORT

12%	10%	7%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.coursehero.com Internet Source	2%
2	www.mdpi.com Internet Source	1%
3	sbsapin.webs.com Internet Source	1%
4	www.papercamp.com Internet Source	<1%
5	link.springer.com Internet Source	<1%
6	proceeding.unnes.ac.id Internet Source	<1%
7	docshare.tips Internet Source	<1%
8	export.arxiv.org Internet Source	<1%
9	www.fuse.org.ph Internet Source	<1%
10	ebin.pub Internet Source	<1%
11	mb.com.ph Internet Source	<1%
12	dspace.lib.uom.gr Internet Source	<1%
13	arxiv.org Internet Source	<1%
14	www.jetro.go.jp Internet Source	<1%
15	www.theseus.fi Internet Source	<1%
16	"Poster Presentations Part II (pp. 627–727)", Annals of Nutrition and Metabolism, 2009 Publication	<1%
17	www.thefreelibrary.com Internet Source	<1%
18	naga.gov.ph Internet Source	<1%
19	"Computational Science – ICCS 2023", Springer Science and Business Media LLC, 2023 Publication	<1%

20	Basma Korchani, Kaouthar Sethom. "Real-Time Littering Detection for Smart City using Deep Learning Algorithm", 2020 International Conference on Communications, Signal Processing, and their Applications (ICCSPA), 2021 Publication	<1 %
21	Grover, Utkarsh. "Investigating the Effectiveness of Fisheye Camera for Depth Estimation and Object Detection", New York University Tandon School of Engineering, 2023 Publication	<1 %
22	www.batangascity.gov.ph Internet Source	<1 %
23	Laurenz Luigi B. Cruz, Takenori Shimozono. "Transport of floating litter within Manila Bay, Philippines", Marine Pollution Bulletin, 2021 Publication	<1 %
24	journalofbigdata.springeropen.com Internet Source	<1 %
25	repository.nida.ac.th Internet Source	<1 %
26	Farrukh Hassan, Lukman AB Rahim, Muhammad Irfan Khan, Muhamed Hariz Mohamed Adnan, Shahid Ali, Faiza Qayyum. "Unlocking Corrosion Secrets: A Comprehensive Analysis of Acoustic Emission Data Using Continuous Wavelet Transform and Deep Learning for Steel Pipeline Monitoring", Research Square Platform LLC, 2023 Publication	<1 %
27	Garfo, Selorm. "Machine Vision for Autonomous Systems with Application to Stormwater Pipeline", North Carolina Agricultural and Technical State University, 2023 Publication	<1 %
28	www.medrxiv.org Internet Source	<1 %
29	Raimund Schnürer, René Sieber, Jost Schmid-Lanter, A. Cengiz Öztireli, Lorenz Hurni. "Detection of Pictorial Map Objects with Convolutional Neural Networks", The Cartographic Journal, 2020 Publication	<1 %
30	en.wikipedia.org Internet Source	<1 %
31	digilib.isi.ac.id Internet Source	<1 %
32	news-aws.abs-cbn.com Internet Source	<1 %
	www.publishing.globalcsrc.org	

33 Internet Source <1%

34 Juan Terven, Diana-Margarita Córdova-Esparza, Julio-Alejandro Romero-González. "A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS", Machine Learning and Knowledge Extraction, 2023
Publication <1%

35 ouci.dntb.gov.ua
Internet Source <1%

36 "Intelligent Systems and Data Science", Springer Science and Business Media LLC, 2024
Publication <1%

37 Dimitris V. Politikos, Argyro Adamopoulou, George Petasis, Francois Galgani. "Using artificial intelligence to support marine macrolitter research: A content analysis and an online database", Ocean & Coastal Management, 2023
Publication <1%

38 Qinghe Zheng, Sergio Saponara, Xinyu Tian, Zhiguo Yu, Abdussalam Elhanashi, Rui Yu. "A real-time constellation image classification method of wireless communication signals based on the lightweight network MobileViT", Cognitive Neurodynamics, 2023
Publication <1%

39 Wei Qi Yan. "Computational Methods for Deep Learning", Springer Science and Business Media LLC, 2023
Publication <1%

40 www.kdnuggets.com
Internet Source <1%

41 E Suharto, Suhartono, A P Widodo, E A Sarwoko. "The use of mobilenet v1 for identifying various types of freshwater fish", Journal of Physics: Conference Series, 2020
Publication <1%

42 U. Sirisha, S. Phani Praveen, Parvathaneni Naga Srinivasu, Paolo Barsocchi, Akash Kumar Bhoi. "Statistical Analysis of Design Aspects of Various YOLO-Based Deep Learning Models for Object Detection", International Journal of Computational Intelligence Systems, 2023
Publication <1%

43 dspace.spbu.ru
Internet Source <1%

44 www.ijsrcseit.com
Internet Source <1%

einj.org

45	Internet Source	<1 %
46	journals.lww.com Internet Source	<1 %
47	journals.plos.org Internet Source	<1 %
48	mdpi-res.com Internet Source	<1 %
49	repository.umsu.ac.id Internet Source	<1 %
50	scholar.ppu.edu Internet Source	<1 %
51	ssdl.gatech.edu Internet Source	<1 %
52	theses.hal.science Internet Source	<1 %
53	Haozhe Cheng, Xu Han, Pengcheng Shi, Jihua Zhu, Zhongyu Li. "Multi-Trusted Cross-Modal Information Bottleneck for 3D self-supervised representation learning", Knowledge-Based Systems, 2023 Publication	<1 %
54	jurnal.polgan.ac.id Internet Source	<1 %
	www.diva-portal.org	

55	Internet Source	<1 %
56	www.nature.com Internet Source	<1 %
57	www.scribd.com Internet Source	<1 %
58	Lin, Wu, Fu, Wang, Zhang, Kong. "Dual-NMS: A Method for Autonomously Removing False Detection Boxes from Aerial Image Object Detection Results", Sensors, 2019 Publication	<1 %
59	bmcmedimaging.biomedcentral.com Internet Source	<1 %
60	coek.info Internet Source	<1 %
61	thesai.org Internet Source	<1 %
62	"Advances in Brain Inspired Cognitive Systems", Springer Science and Business Media LLC, 2020 Publication	<1 %
63	"Biometric Recognition", Springer Science and Business Media LLC, 2017 Publication	<1 %

64	"IEEE MI-STA2023 Conference Proceeding", 2023 IEEE 3rd International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering (MI-STA), 2023 Publication	<1 %
65	"Modelling and Development of Intelligent Systems", Springer Science and Business Media LLC, 2023 Publication	<1 %
66	9pdf.net Internet Source	<1 %
67	Alghamdi, Thoria. "Facial Expressions-Based Pain Assessment System Using Deep Learning Techniques.", University of Colorado at Denver, 2023 Publication	<1 %
68	Gene Marck Bañares Catedrilla, Jefferson Llobit Lerios, Sherwin Banaag Sapin, Manuel C Lanuang, Chester Alexis C Buama. "An android-based mobile educational game for disaster preparedness: an input to risk reduction management", Indonesian Journal of Electrical Engineering and Computer Science, 2021 Publication	<1 %

69	Lawrence, Tom. "Deep Neural Network Generation for Image Classification Within Resource-Constrained Environments Using Evolutionary and Hand-Crafted Processes", University of Northumbria at Newcastle (United Kingdom), 2023 Publication	<1 %
70	Md Mijanur Rahman, Fatema Khatun, Ashik Uzzaman, Sadia Islam Sami, Md Al-Amin Bhuiyan, Tiong Sieh Kiong. "A Comprehensive Study of Artificial Intelligence and Machine Learning Approaches in Confronting the Coronavirus (COVID-19) Pandemic", International Journal of Health Services, 2021 Publication	<1 %
71	R. Barth, J. IJsselmuiden, J. Hemming, E.J. Van Henten. "Synthetic bootstrapping of convolutional neural networks for semantic plant part segmentation", Computers and Electronics in Agriculture, 2017 Publication	<1 %
72	Shilian Zheng, Jiachen Hu, Luxin Zhang, Kunfeng Qiu, Jie Chen, Peihan Qi, Zhijin Zhao, Xiaoniu Yang. "FM-Based Positioning via Deep Learning", Institute of Electrical and Electronics Engineers (IEEE), 2023 Publication	<1 %

73	Shubham Mahajan, Akshay Raina, Xiao-Zhi Gao, Amit Kant Pandit. " -19 detection using hybrid deep learning model in chest x-rays images ", Concurrency and Computation: Practice and Experience, 2021 Publication	<1 %
74	aran.library.nuigalway.ie Internet Source	<1 %
75	edepot.wur.nl Internet Source	<1 %
76	eprints.gla.ac.uk Internet Source	<1 %
77	iieta.org Internet Source	<1 %
78	ijosi.org Internet Source	<1 %
79	repository.tudelft.nl Internet Source	<1 %
80	repository.uwl.ac.uk Internet Source	<1 %
81	www.arxiv-vanity.com Internet Source	<1 %
82	www.frontiersin.org Internet Source	<1 %

83	www.globalscientificjournal.com Internet Source	<1 %
84	www.researchgate.net Internet Source	<1 %
85	www.researchsquare.com Internet Source	<1 %
86	www.semanticscholar.org Internet Source	<1 %
87	www.uwstout.edu Internet Source	<1 %
88	www.wne.uw.edu.pl Internet Source	<1 %
89	Yassine, Barhoumi. "Efficient Scopeformer: Towards Scalable and Rich Feature Extraction for Intracranial Hemorrhage Detection Using Hybrid Convolution and Vision Transformer Networks", Rowan University, 2023 Publication	<1 %
90	hdl.handle.net Internet Source	<1 %
91	strathprints.strath.ac.uk Internet Source	<1 %
92	El-Naggar, Susan Awm. "Ear Biometrics: A Comprehensive Study of Taxonomy,	<1 %

Detection, and Recognition Methods", West Virginia University, 2023
Publication

93	Roseline Oluwaseun Ogundokun, Sanjay Misra, Akinyemi Omololu Akinrotimi, Hasan Ogul. "MobileNet-SVM: A Lightweight Deep Transfer Learning Model to Diagnose BCH Scans for IoMT-Based Imaging Sensors", Sensors, 2023 Publication	<1 %

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	On		

APPENDIX

H

Curriculum Vitae