Journal of Universal Computer Science, vol. 30, no. 1 (2024), 49-92
submitted: 21/2/2023, accepted: 12/9/2023, appeared: 28/1/2024 CC BY-ND 4.0

Image Filtering Techniques for Object Recognition in Autonomous Vehicles

Ngo Le Huy Hien
(Leeds Beckett University, Leeds, United Kingdom
https://orcid.org/0000-0002-1439-8289, ngolehuyhien@gmail.com)

Ah-Lian Kor
(Leeds Beckett University, Leeds, United Kingdom
https://orcid.org/0000-0001-9514-3773, a.kor@leedsbeckett.ac.uk)

Mei Choo Ang
(Universiti Kebangsaan Malaysia, Bangi, Malaysia
https://orcid.org/0000-0001-5316-3936, amc@ukm.edu.my)

Eric Rondeau
(University of Lorraine, Nancy, France
https://orcid.org/0000-0002-9265-1673, eric.rondeau@univ-lorraine.fr)

Jean-Philippe Georges
(University of Lorraine, Nancy, France
https://orcid.org/0000-0003-1011-8312, jean-philippe.georges@univ-lorraine.fr)

Abstract: The deployment of autonomous vehicles has the potential to significantly lessen the variety of current harmful externalities, (such as accidents, traffic congestion, security, and environmental degradation), making autonomous vehicles an emerging topic of research. In this paper, a literature review of autonomous vehicle development has been conducted with a notable finding that autonomous vehicles will inevitably become an indispensable future greener solution. Subsequently, 5 different deep learning models, YOLOv5s, EfficientNet-B7, Xception, MobilenetV3, and InceptionV4, have been built and analyzed for 2-D object recognition in the navigation system. While testing on the BDD100K dataset, YOLOv5s and EfficientNet-B7 appear to be the two best models. Finally, this study has proposed Hessian, Laplacian, and Hessian-based Ridge Detection filtering techniques to optimize the performance and sustainability of those 2 models. The results demonstrate that these filters could increase the mean average precision by up to 11.81%, reduce detection time by up to 43.98%, and significantly reduce energy consumption by up to 50.69% when applied to YOLOv5s and EfficientNet-B7 models. Overall, all the experiment results are promising and could be extended to other domains for semantic understanding of the environment. Additionally, various filtering algorithms for multiple object detection and classification could be applied to other areas. Different recommendations and future work have been clearly defined in this study

Keywords: Autonomous Vehicle, Deep Learning, Object Recognition, Image Filtering Techniques, Navigation System
Categories: I.0, I.2, I.3, I.4, I.5, I.6, J.6
DOI: 10.3897/jucs.102428

1 Introduction

Unlike traditional vehicles, autonomous vehicles [Yair, 2022] are land-based vehicles integrating vehicular automation [Francis and Anoop, 2014] that can sense their surroundings and operate securely [Hu et al., 2020a] with minimal or no human intervention [Brummelen et al., 2018]. To achieve this, autonomous vehicles employ a fully or partially automated driving system that enables them to adapt to external situations similar to a skilled driver [Yair, 2022]. The autonomous driving system incorporates a range of sensors to sense its environment, including GPS (Global Positioning System), thermographic cameras, radar, sonar, odometry, LiDAR (light detection and ranging), and IMU (Inertial Measurement Units) [Hu et al., 2020a, Marchegiani and Fafoutis, 2022, Thrun, 2010, Feng and Liu, 2019, Jedrasiak et al., 2013]. Traffic signals are normally detected by video cameras, which could also recognize other vehicles, street signs, and pedestrians. Radar sensors focus on the location and movement of surrounding objects. Meanwhile, LiDAR sensors assess distances, detect road boundaries, and recognize lane markers by bouncing light pulses off the car's environment. Automobiles' wheels are also equipped with ultrasonic sensors to detect obstacles and other vehicles while parking [Schwarting et al., 2018].

Those sensory data are analyzed and interpreted by cutting-edge control systems to build a map of the surrounding environment for planning/executing relevant navigation routes, avoiding obstacles, and recognizing essential signs [Hu et al., 2021]. The systems then issue instructions to the car's actuators, which regulate steering, acceleration, and braking [Yair, 2022]. The program could comply with traffic regulations and avoid obstacles due to its object identification capabilities, obstacle avoidance algorithms, hard-coded rules, and predictive modeling [Friedrich, 2016]. Notably, precise scene perception and accurate localization are critical prerequisites for effective and safe driverless car navigation systems. These 2 functions must accurately obtain and evaluate data gathered about real surroundings [Nguyen and Le, 2013]. To achieve this, multiple primary sensors, such as LiDAR and digital cameras, are installed on autonomous vehicles to collect precise contextual data.

In general, Deep Learning models have been widely applied in autonomous cars [Uçar et al., 2017] with 3 major tasks: segmentation, detection, and classification [Guan et al., 2016, Hien, 2022]. These models incur a variety of data source inputs (e.g., 1-D speech, 2-D pictures or videos, 3-D CAD (Computer-Aided Design), and 3-D LiDAR point cloud) [Wiseman, 2022] which is the primary focus of most research [Bagloee et al., 2016]. However, technological limitations and the high cost of proprioceptive sensors have a significant impact on the implementation of fully self-driving systems [Wiseman, 2022].

- Firstly, one of the most serious concerns is incorrect object detection and classification [Bagloee et al., 2016]. The data generated by various sensors is accumulated and processed by the autonomous system. However, with just a few pixels of variation in an image created by a digital camera, a vehicle may mistake a stop sign for a different sign such as a speed restriction sign, which could be potentially dangerous. If the algorithm similarly wrongly identifies a pedestrian as a traffic light, it would not expect it to move [Pozna and Antonya, 2016]. Therefore, object recognition and classification are critical for research relating to autonomous vehicles and this is the driver of this research.
- Secondly, a majority of objects, including vehicles, street signs, and pedestrians, are identified and classified from image data by digital cameras due to their low cost

and high efficiency [Vivacqua et al., 2017]. Camera-based vision is comparable to that of a human driver and has become the most widely utilized data in perception tasks using 2-D appearance-based representation [Paya et al., 2009]. Nevertheless, poor image quality acquired under extreme conditions (e.g., low illumination, rain, fog/mist/haze, snowstorms, and sandstorms) would adversely affect the performance of deep learning models [Ondrus et al., 2020]. Consequently, there is a need to employ different filtering techniques to enhance the quality of the images to improve the deep learning models.

– Thirdly, it is essential to monitor the energy consumption of ICT programs and applications in order to explore energy efficiency and its sustainable co-existence with the hardware [Fettweis and Zimmermann, 2008]. Deep learning technologies have been employed extensively across different areas and have emerged as the leading technology in the ICT sector [Haldorai et al., 2021]. However, most deep learning studies continue to primarily focus on precision without any consideration of computing resources or energy consumption [Hilty et al., 2009]. The possible main reason is that data scientists or AI experts do not prioritize sustainability because they are yet to acquire software/application-related energy audit skills [García-Martín et al., 2019]. To address this issue, this research discusses energy monitoring and analysis for deep learning applications. Additionally, this research also aims to conduct a further carbon footprint analysis, after optimizing deep learning models.

Those identified gaps have prompted this study to construct and evaluate different Deep Learning models for 2-D Object Recognition in autonomous vehicles, with the following set of Research Objectives (RO).

RO1: To conduct a critical autonomous vehicle-related literature review by addressing the following Research Questions (RQ).

– **RQ1.1** What are the potentials in the development context of autonomous vehicles?

– **RQ1.2** What are the challenges of autonomous vehicles?

– **RQ1.3** Will the environmental impact of autonomous vehicles be better compared to the current traditional transportation mode?

– **RQ1.4** What deep learning applications and frameworks are currently applied in autonomous vehicles?

– **RQ1.5** How are current deep learning models employed in 2-D object recognition for autonomous vehicles?

RO2: To identify a suitable dataset for preprocessing and analysis.

RO3: To propose a range of filtering techniques for deep learning models in object recognition and classification in 2-D image data, optimizing their performance and sustainability, with the following RQs:

– **RQ3.1** What are the performance and sustainability metrics for current state-of-the-art deep learning models for 2-D object recognition?

– **RQ3.2** Is the mean average precision, detection time, and sustainability of object recognition by deep learning models with proposed filtering techniques better compared to the models without them?

RO4: To make evidence-based recommendations and inform future relevant research work.

To address the above research objectives and questions, firstly, a literature review is undertaken to furnish an overview of autonomous vehicles, their limitations, and opportunities, as outlined in Section 2. Relevant research methodologies are discussed in Section 3, followed by an investigation of the performance and sustainability of current state-of-the-art deep learning models for 2-D image data object recognition in Section 4. The novelty of this research includes proposing varying filtering techniques to enhance digital camera images to achieve a better level of performance and sustainability (Section 5). Finally, pertinent discussion and evidence-based recommendations are presented in Sections 6 and 7. The findings are promising, and they will lead to more studies in the future, relating to various filtering algorithms for object detection and classification in autonomous cars.

2 Literature Review

2.1 Overview of Autonomous Vehicles

2.1.1 Autonomous Vehicles

In contrast to traditional vehicles, autonomous vehicles [Yair, 2022] (also known as driverless vehicles [Marchegiani and Fafoutis, 2022], self-driving cars [Mike et al., 2017], robocars [Feng and Liu, 2019], or robotic cars [Thrun, 2010]) are land-based vehicles that integrate vehicular automation [Francis and Anoop, 2014], sense their environment and operate safely [Hu et al., 2020a] with no or minimal human intervention [Brummelen et al., 2018]. Autonomous vehicles have been assessed in terms of their environmental [Kopelias et al., 2020], practical [Fan and Xu, 2019], and human lifestyle impacts [Pozna and Antonya, 2016]. It has been deployed in a variety of forms, including shared self-driving taxis [Martinez and Viegas, 2017], and connected vehicle platoons [Hu et al., 2020b].

A car navigation package incorporates a GPS [Rahiman and Zainal, 2013] and a geographic information system (GIS) [Quan, 2019] to collect location data including latitude and longitude coordinates. The positioning system determines the relative vehicle location using an inertial navigation system (INS) [Liu et al., 2020]. Furthermore, an autonomous car is equipped with an electronic map (EM) that saves data on road and traffic infrastructure [Yuan et al., 2018]. Subsequently, the vehicle's position is estimated followed by relevant path planning [Chen et al., 2014].

The three basic channels for environment perception are visual perception, radar perception, and laser perception [Zhu et al., 2017]. Radar perception involves the calculation of the distance via time estimation from which a radar sensor's wave is transmitted and reflected [Dickmann et al., 2016]. Reflection time and reflection signal strength are exploited in laser perception for the creation of cloud data for target points such as position, shapes, and states. Light Detection and Ranging (LiDAR) is used to avert accidents and in emergency braking circumstances. LiDAR systems produce a high number of laser pulses per second. After interacting with nearby objects, these pulses are reflected. Calculations based on the speed of light and the distance traveled by the pulse aid in the creation of a three-dimensional representation [Li and Ibanez-Guzman, 2020].

Self-driving cars normally use short-range vehicle-to-vehicle communication to connect with their environment and other cars [Shieh et al., 2018]. This form of communication requires a real-time and redundant infrastructure [Hien et al., 2021a]. The concepts of MANETs (mobile ad hoc networks) – the spontaneous establishment of a wireless network of mobile devices – are applied to the transportation domain to produce vehicular Ad hoc networks (VANETs). VANETs are used by connected automobiles to communicate with one another [Sood and Kanwar, 2014]. To facilitate the integration of autonomous vehicles with existing connected vehicle technology, the autonomous vehicle must employ the same communication standard that is currently used for connected vehicle technology [Jiang et al., 2011].

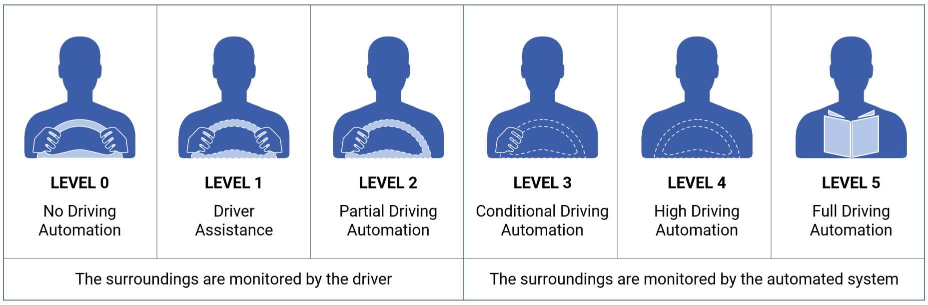

Figure 1: Society of Automotive Engineers (SAE) 6 levels of driving automation.

The Society of Automotive Engineers (SAE) outlines a system named SAE J3016 (revised periodically) for 6 levels of driving automation, ranging from Level 0 (completely manual) to Level 5 (completely autonomous), as depicted in Figure 1 [Automotive Engineers, 2021]. At the beginning of Level 3, the autonomous driving system can monitor the driving surroundings. Autonomous cars would first advance through these 6 levels of driver-aid technologies before being able to operate on public roads [Cui and Sabaliauskaite, 2017].

Numerous efforts to produce a fully autonomous commercial vehicle are in various development phases but self-driving cars are yet to be offered to the general public [Litman, 2017]. Level 3 and above vehicles continue to represent a minuscule fraction of the market [Markets, 2021]. As of May 2022, Tesla is currently at Level 2 because the autonomous system operates without driver supervision only under specific, and constrained situations [Morris, 2021]. However, according to Elon Musk, Tesla is expected to reach Level 4 autonomy in 2022, when critical development milestones of full self-driving (FSD) are accomplished to perform better than a human driver in averting accidents [Fox, 2021]. Moreover, Ultra Cruise and Super Cruise, which are currently offered on Cadillac and Chevrolet automobiles, are classified as Level 2 since the human driver must maintain complete attention anytime [Ulitskaya, 2021]. Audi AI Traffic Jam Pilot, first announced in 2017, is supposed to be the new A8's highlight feature, enabling the driver to assume control of the vehicle driving on the highway [Hartmann, 2017]. However, Audi has scrapped plans to update its premium A8 vehicle with Level 3 autonomous driving in April 2020, citing the need for governments to build a legislative framework before making the technology available broadly [Davies, 2020]. Honda is the first business to offer a legally authorized Level 3 vehicle in March 2021 [Honda, 2021]. Mercedes-Benz is the second company to receive legal approval for a Level 3 that exceeded regulatory conditions in December 2021 [Group, 2021].

Due to current technological infrastructures and regulations, most Level 4 cars are designed for ridesharing within a specific region [DeKort, 2019]. In 2020, Waymo, a subsidiary of Alphabet, becomes the first manufacturer to offer an autonomous taxi trip to the general public within the Phoenix region in Arizona. It is currently categorized as a Level 4 since the driver is not needed in the vehicle despite a remote observer is still required [Lee, 2020]. Magna, a Canadian automaker, has developed MAX4 - a technology that allows Level 4 autonomy for both urban and motorway environments. They are collaborating with Lyft to equip high-tech kits that transform cars into self-driving vehicles [Magna, 2022]. Around the Tokyo 2020 Olympic Village, Toyota has provided a potential Level 4 solution [Davis, 2021]. NAVYA, a French company, has been manufacturing and marketing Level 4 taxis and shuttles in the United States, which operate fully on electrical power with a peak speed of 55 mph [Vijayenthiran, 2018]. The remaining manufacturers are still in the developing phase to launch a Level 4 autonomous vehicle. For example, Volvo and Baidu form a strategic cooperation in 2018 to jointly develop Level 4 electric vehicles for the Chinese robo-taxi sector [Bolduc, 2018]. In terms of Level 5, fully autonomous vehicles are being tested in various parts of the world, but none is currently accessible to the general public [Mane, 2021].

2.1.2 Potentials of Autonomous Vehicles

Driverless cars are currently being integrated into the product catalogs of all major automotive original equipment manufacturers (OEMs) [Kang et al., 2019]. Furthermore, recent advancements in VANET (Vehicular Ad hoc NETwork) applications and services [Maalej et al., 2018] as well as connected automobile technology [Contreras-Castillo et al., 2017], have prompted companies such as Google [Lee, 2020], car manufacturers (e.g. Tesla [Fox, 2021] and Audi [Group, 2021] to research on self-driving automobile technologies. Other leading manufacturers, including BMW, Ford, Mercedes-Benz, Toyota, Hyundai, Kia, Honda, Nissan, Volvo, Volkswagen, and General Motors, have undertaken programs such as semi-automatic pilot driving, emergency braking, smart parking, and accident alert [Hussain and Zeadally, 2018]. Volkswagen and Microsoft have broadened their cloud-based technology partnership to include autonomous vehicle development [Lawrence, 2021]. This transition will significantly transform how transportation solutions are accessed and utilized in the future, as well as how metropolitan regions are planned, designed, and built to facilitate a cleaner and greener coexistence of different transportation solutions such as automobiles, electric motorcycles, trucks, public transportation, supply vehicles, and bicycles [Kang et al., 2019].

With continuous improvement, autonomous cars are steadily gaining acceptance [Wiseman, 2022]. According to ABI Research, by 2025, there will be around 8 million semi-autonomous or autonomous vehicles in operation [Research, 2018]. The Victoria Transport Policy Institute anticipates that half of the new vehicles will be autonomous by 2045, with half of the whole fleet operating autonomously by 2060 [Litman, 2017]. Meanwhile, the COVID-19 pandemic is depleting manufacturers' financial reserves and threatens to hinder progress in self-driving car research, a sector that Bain predicts will reach a critical point in 2028 [Klaus et al., 2020].

Figure 2 depicts the forecast trend of market penetration and benefits for autonomous vehicles. It is predicted that by 2045, half of the new cars will be autonomous. Level 4 autonomy is expected to alleviate driver stress and boost productivity, but the most advantages associated with Level 5 autonomy would be autonomous pickup and drop-off [Litman, 2017].

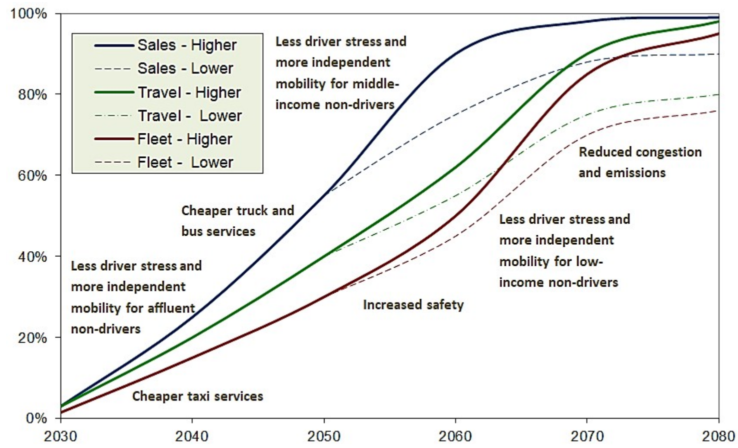

Figure 2: Predictions of Fleet, Sales, Travel, and Benefits of Autonomous Vehicles [Litman, 2017].

Autonomous vehicles would help ensure road safety, reduce traffic congestion, enhance security, and mitigate environmental degradation [Saeed et al., 2016]. Cities may reduce carbon dioxide (CO_2) emissions by up to 80% if they embrace three revolutions (3R) within vehicle technology: electrification, automation, and ride-sharing. By 2050, these three urban transportation revolutions might decrease the worldwide cost of automobiles, infrastructure, and transportation system operations by 40%, depending on the successful deployment of ride-sharing and sustainable energy sources consumption [Lew et al., 2018]. Moreover, autonomous vehicles can increase mobility and accessibility [William and Pande, 2021], as well as optimize land use [Hawkins and Habib, 2019].

On the consumer side, autonomous vehicles have the potential to relieve vehicle occupants of driving and navigation tasks, lowering stress, and enhancing productivity [Ahangar et al., 2021]. This enables individuals to perform other duties, such as playing and working while commuting or resting during their lengthy and stressful traffic travels [Andrew, 2015]. Furthermore, self-driving automobiles eliminate limitations imposed by driver ability or condition. It provides non-drivers with more autonomous mobility and has the potential to minimize drivers' chauffeuring duties and transit subsidy demands, particularly for the blind, intoxicated, elderly, underage, and impaired people [Hussain and Zeadally, 2018]. A driving test or driving license is no longer necessary since humans do not need to drive [Ilková and Ilka, 2017]. With fewer traffic accidents and collisions, autonomous vehicles can improve safety while also lowering crash risks and insurance costs. Customers can also benefit from lower costs of taxi and commercial transportation [Saeed et al., 2016].

On the social side, self-driving cars might become a solution to the problem of depleted fuel economy, lessening the environmental burden of existing transportation modes [Litman, 2017]. Moreover, autonomous vehicles may promote commercial car-sharing and ride-sharing, lowering overall vehicle ownership and travel expenses [Mike et al., 2017]. Better control of traffic flow will enhance road capacity and minimize traffic congestion [Grigorescu et al., 2020]. Furthermore, self-driving vehicles may reduce the demand for parking at destinations, leading to less parking costs [Saeed et al., 2016].

2.1.3 Challenges of Autonomous Vehicles

There are numerous technological challenges associated with autonomous cars. The distance traveled determines the system's reliability. If no accident happens, a self-driving car would need to travel approximately 291 million miles to achieve a 95 percent equivalent to a certified human driver [Hars, 2016]. Systems programming will require an in-depth assessment of product development and supply chain [Nastjuk et al., 2020]. Furthermore, AI is now incapable of successfully operating in densely populated metropolitan regions [Ondrus et al., 2020]. Security and data privacy of autonomous cars might be challenging due to mandatory location monitoring and data sharing [Bagloee et al., 2016]. The car's computer, as well as the communication system between vehicles, might be compromised [Andrew, 2015]. Extreme weather, such as rain and snow, as well as intentional interference (e.g., spoofing and jamming), may have an impact on the car's navigation and sensory systems [Bagloee et al., 2016]. It is crucial to determine the severity of driving lane impediments, such as safely straddling a pothole or debris or avoiding moving animals [Hars, 2016]. The car may be involved in an accident if the main sensor and backup sensors break. High-definition maps may be required for autonomous vehicles to correctly function. Where these maps may be out of date, they must be able to assume sensible decisions. Additionally, some temporary construction zones may not be updated on the maps and databases [Faisal et al., 2019]. Most importantly, for autonomous vehicles to correctly function, existing road infrastructure may need to be upgraded [Gill, 2021].

Furthermore, the development of self-driving automobiles encounters a variety of social obstacles. The high cost of creating and deploying self-driving cars poses as a barrier, raising concerns about end-user affordability [Singh and Saini, 2021]. Furthermore, uncertainty about future regulations and policies may defer the deployment of driverless cars on the road [Saeed et al., 2016]. Transparent as well as unambiguous regulations and procedures that address potential consumers' concerns are essential [Ilková and Ilka, 2017]. Another important concern is how self-driving cars make a judgment with justifiable behaviors in potentially challenging emergencies [Hussain and Zeadally, 2018]. The development of self-driving automobiles may reduce the need for cheap mobility choices such as walking, biking, and public transportation [Gill, 2021]. Another area of concern is the health consequences preferred vehicle usage over public transportation and walking since less exercise leads to a rise in obesity [Bagloee et al., 2016]. Reliance on autonomous driving results in less experienced drivers when manual driving is required [Andrew, 2015]. Terrorists or other criminals might theoretically employ driverless automobiles as autonomous explosives. As a result, there is a need to build a legal framework and regulatory laws for self-driving automobiles [Andrew, 2015].

Though autonomous cars have made significant advances, there are still constraints that must be appropriately addressed. Despite the potential difficulties of self-driving cars, the common view is that the benefits outweigh the drawbacks [Ahangar et al., 2021]. While autonomous car companies have made great technological advances, people must be prepared to adopt them. To ensure that technology continues to progress, the above-mentioned technological and societal concerns must be addressed.

2.1.4 Carbon Footprint

Undeniably, environmental challenges caused by transportation becomes more severe as urbanization advances [Vos et al., 2021]. Carbon dioxide (CO_2) emissions (the highest contributor to climate change) have continued to rapidly rise, according to statistics from Our World in Data. Humans currently release around 38.02 billion tons of CO_2 each year (data from 2020) [Data, 2022]. Notably, passenger and freight transportation contribute 20% to 30% of world greenhouse gases (GHG) [Straka et al., 2021], while passenger vehicles account for 75% of overall CO_2 emissions [Climate Change, 2019, Hien and Kor, 2021]. Consequently, to address this global challenge, electric cars (EVs) will ultimately replace fuel-powered cars.

a) *The calculation of CO_2 emissions of EVs*

It takes a lot of energy to produce batteries made from materials such as lithium, copper, and refined silicon. Thus, energy consumed for autonomous cars' production, and its associated CO_2 emissions amount to half of its lifecycle emissions (i.e., 6.3 tons of CO_2). For fossil-fuel vehicles, the amount represents only 17% of their lifecycle emissions (i.e., 5.6 tons of CO_2) [Bjørn and Hauschild, 2012]. Additionally, many electric cars are likely to require battery replacement after several years. When emissions of the replaced battery are included, the total CO_2 produced by an electric vehicle jumps to 12.6 tons, compared to 5.6 tons for a gas vehicle. Furthermore, due to the energy required for collecting and recycling metals in the battery, disposal would double the emissions [Verma et al., 2021].

Considering charging, if the energy required to recharge the electric car comes mostly from coal-fired power plants, it will emit about 15 ounces of CO_2 for every 50,000 miles traveled, which is 3 ounces more than a comparable gas-powered car [Bjørn and Hauschild, 2012]. If electric cars are driven for over 67,000 miles with green electricity, they would be greener than gasoline cars. For instance, for a covered distance of 90,000 miles, the EV emits 24% less CO_2 than its gas-powered counterpart, which is equivalent to 8.7 tons of CO_2 [Bjørn and Hauschild, 2012]. Research has shown that EVs driven by the current European energy mix would have a 10% to 24% lower global warming potential than traditional diesel or gasoline vehicles with lifespans of 150,000 km [Hawkins et al., 2012].

b) *The actual efficiency of EVs*

Though Nissan claims that its EV efficiency is about 100 miles, the actual figure is approximately 58 miles [Woodhill, 2011]. With cold weather or fast driving, this is further reduced to 30 miles. On the other hand, a traditional vehicle could travel 300 to 400 miles. A Nissan's Leaf takes 20 hours to fully recharge on 110V household power. This duration can be reduced to 8 hours by using a 240V charger which costs more. Although there are pricey 480V chargers that can reduce this duration to 4 hours, Nissan has warned against it because of the negative impact of frequent use on the battery's life [Woodhill, 2011]. However, autonomous vehicles and their supply chain have the potential to significantly increase toxicity, freshwater eutrophication, eco-toxicity, and metal depletion [Hawkins et al., 2012].

c) *Comparison of the environmental impact of an electric car and a gas-fueled car*

When considering the mitigation of natural resource consumption, EVs will be the preferred option [Swedish Environmental Research Institute, 2019]. By 2061, the world will have to transition away from gas-powered vehicles, and coal-fueled power plant industries to renewable energy sources [Jacobson, 2017]. The environmental impact of

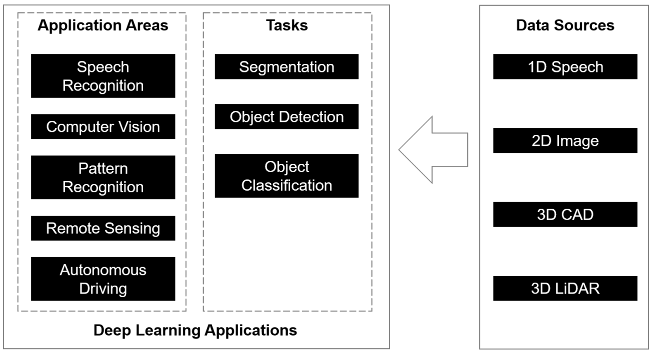

Figure 3: Deep Learning Applications and Different Data Sources for Autonomous Vehicles.

autonomous cars will be minimal if they are charged with renewable solar and wind energy [Roth, 2021].

The environmental impacts of autonomous vehicles depend on how electricity is generated. If autonomous cars are charged with renewable solar and wind energy [Roth, 2021], they will be fairly clean. The environmental impacts of autonomous vehicles also depend on the lifetime of the vehicle that customers use. The calculation of EVs' CO_2 emissions above estimated that when the electric car is driven for over 67,000 miles with green power, such as renewable solar and wind energy, EVs are greener than gasoline cars.

Overall, with the current context of major electricity production by coal-fired power plants, electric cars may not seem to be a good solution. However, with a long-term vision of using green power, EVs are proven to have real value in terms of being environmentally friendly in the long run. Conditions for doing so include a substantial supply of green energy, usage phase energy consumption, vehicle lifetime, battery replacement schedules, and the transition from private vehicle use to shared and community use. Therefore, improving the sustainability level of autonomous vehicles necessitates collaboration on minimizing vehicle production supply chain impacts and encouraging clean power sources in electrical infrastructure decision-making.

2.2 Deep Learning Applications in Different Data Sources for Autonomous Vehicles

While self-driving cars are mainly still in their early stage, Deep Learning has been exploited for Advanced Driver-Assistance Systems (ADAS) [Kukkala et al., 2018], particularly for knowledge discovery and contextual awareness [Mike et al., 2017]. Recent developments in deep learning and sensor technologies in autonomous driving have the potential to improve traffic efficiency, and road safety [Yair, 2022]. Deep Learning models have been deployed for: segmentation, detection, and classification [Guan et al., 2016]. These models utilize a diverse range of data sources, such as 1-D speech, 2-D images or videos, 3-D CAD (Computer-Aided Design), and 3-D LiDAR point cloud [Che et al., 2019] (Figure 3).

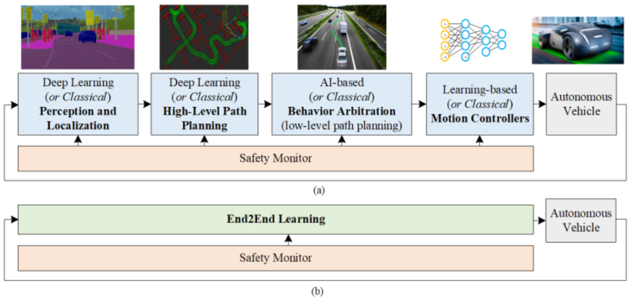

Figure 4: Deep Learning-based Decision-Making Architectures: (a) sequential perception-planning-action pipeline, (b) End2End system [Yan et al., 2017].

Precise scene perception and localization are important prerequisites for decision-making, and the safety of autonomous vehicles. These two functionalities involve the collection and analysis of detailed real-world environments-related data [Nguyen and Le, 2013] (e.g., LiDAR, radar, or digital camera data). Current 3-D deep learning frameworks (i.e., extensions of 2-D structures) have been deployed for object recognition, scene interpretation, classification, and semantic segmentation [Yan et al., 2017].

A Deep Learning-based decision-making architecture design for autonomous vehicles might assume the form of a sequential perception-planning-action pipeline or an End2End system. The sequential pipeline system is separated into 4 components that are hierarchically organized: perception and localization, high-level path planning, low-level path planning, and motion controllers. The core of the End2End system comprises deep learning algorithms, as indicated in Figure 4. A safety monitor routinely ensures the safety of each module [Grigorescu et al., 2020].

Google's TensorFlow and Caffe frameworks employ deep learning algorithms to train a neural network for object/ scene detection and learning so that autonomous vehicles can operate appropriately [Gallardo et al., 2017]. They may also be utilized for lane identification, obstacle avoidance, and trajectory forecast [Grigorescu et al., 2020]. Deep Learning is also widely used in autonomous cars for object categorization (e.g., vehicles, pedestrians, animals, or static structures). They can also detect clear space around the car before switching lanes, etc. [Muhammad et al., 2020].

2.3 2-D Deep Learning Object Recognition Models for Autonomous Vehicles

Object detection and classification are the initial steps (after sensor steps) in the perception phase of autonomous vehicles. The data collected from these steps should be merged for data fusion. Next, the data fusion output becomes the input for the decision and planning phase, before yielding output to the control system and invoking relevant actions [Behere and Torngren, 2015]. Therefore, object recognition is the initial and critical step that tremendously impacts subsequent actions in autonomous vehicle systems. This is also the reason why object recognition is the main focus of this study.

Object detection involves the use of bounding boxes for producing object localizations [Wang et al., 2022] and recognizing objects based on their salient attributes [Che et al., 2019]. Single-stage and double-stage detectors are the most often used structures for 2-D object detection in pictures. ”You Only Look Once” (YOLO) [Jiang et al., 2022], the Single Shot multi-box Detector (SSD) [Kumar and Srivastava, 2020], CornerNet [Law et al., 2019], and RefineNet [Zhou et al., 2022] are all renowned single-stage detectors. RCNN, Faster-RCNN [Cheng et al., 2018], and R-FCN [Zhang and Chi, 2020] are two-stage detectors that divide the object detection process into two parts: region of interest for object candidates’ generation and bounding box classification. In general, single-stage detectors do not perform as well as double-stage detectors, but they are substantially quicker.

Figure 5 shows a comparison of the object recognition algorithms using the Pascal VOC 2012 dataset and their measured mean Average Precision (mAP) with an Intersection over Union (IoU) value of 50 and 75, respectively. The first four techniques on the right are single-stage detectors, whereas the rest are double-stage detectors. Due to their greater complexity, two-stage detectors have a reduced runtime performance in frames per second (FPS) [Grigorescu et al., 2020].

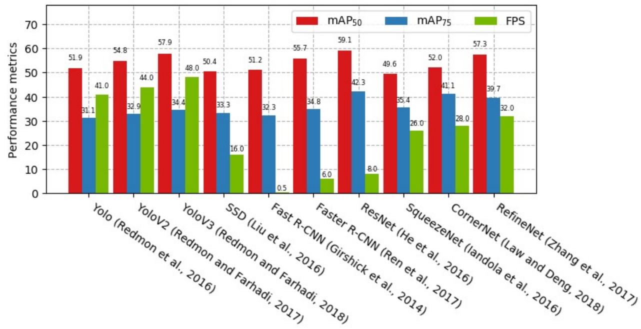

Figure 5: Comparison of object detection and recognition performance of different deep learning models on Pascal VOC 2012 benchmarking dataset [Grigorescu et al., 2020].

It can be seen from Figure 5 that the ResNet model seems to perform the best with the Pascal VOC 2012 benchmarking dataset, compared to other models. Meanwhile, SqueezeNet, SSD, and Fast R-CNN are models that need to be optimized in terms of mean Average Precision. Another comparative analysis has been conducted on the VOC 2012 dataset. The results have been tabulated in Table 1 and depicted in Figure 6.

From Table 1 and Figure 6, R-FCN (ResNet101) appears to be the best model when compared to YOLOv1 and R-CNN(Alex). By experimenting on another dataset – COCO with the latest deep learning models, the comparison results are shown in Figure 7. It evidences that the DyHead model (Swin-L, multi-scale, self-training) is currently the best in terms of Box Average Precision, compared to state-of-the-art models on the COCO test-dev dataset [Code, 2022].

Model	mAP (mean Average Precision)
R-CNN(Alex)	53.3
R-CNN(VGG16)	62.4
Bayes	66.4
Fast R-CNN	68.4
SutffNet30	70.0
NOC	68.8
MR-CNN & S-CNN	73.9
HyperNet	71.4
OHEM + Fast R-CNN	80.1
ION	76.4
Faster R-CNN	70.4
YOLO	57.9
YOLO + Fast R-CNN	70.7
YOLOv2	78.2
SSD300	79.3
SSD512	82.2
R-FCN (ResNet101)	85.0

Table 1: Comparison of object detection performance of different deep learning models on VOC 2012 Dataset [Zhao et al., 2019].

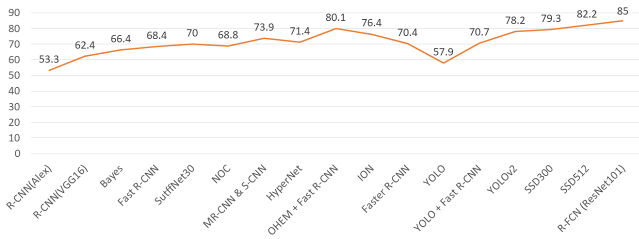

Figure 6: Comparison of object detection performance of different deep learning models on VOC 2012 Dataset [Zhao et al., 2019].

3 Methodology

3.1 Macro Methodology

In this study, the chosen dataset for the deep learning experiments on object recognition and classification of image data is Berkeley Deep Drive Dataset (BDD100K). This research follows the data analytics lifecycle (see Figure 8) which is adopted from EMC Education Services [Ultralytics, 2022].

The data analytics lifecycle comprises 6 phases. The initial stage is the Discovery phase which involves the identification of the problem, context, hypothesis, and goals for which the data will be used. The datasets are then processed and compacted into a single

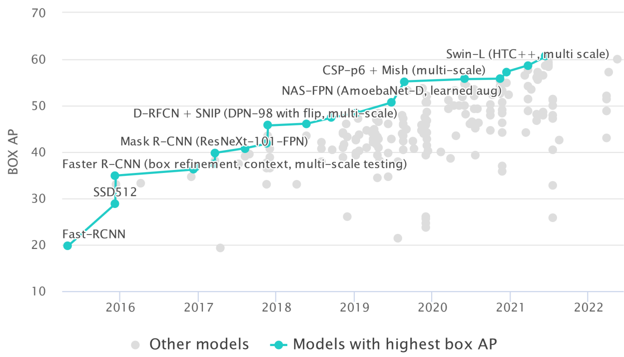

Figure 7: Comparison of object detection performance of different deep learning models on COCO test-dev Dataset [Code, 2022].

spreadsheet in Phase 2 - Data Preparation. The dataset is explored and visualized in Phases 3 and 4 - Model Planning and Building - utilizing 2 levels of data analytics: (i.e., descriptive and inferential statistical analysis). Finally, in Phases 5 and 6, key analytics and prediction outcomes are communicated to inform relevant operations.

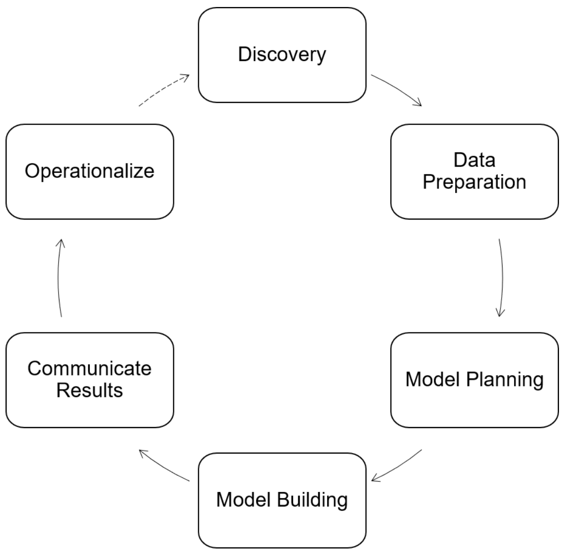

Figure 8: The Data Analytics Lifecycle.

Overall, BDD100K is a large-scale driving dataset with diverse types of annotations provided by UC Berkeley, USA. This dataset can be used to recognize 2-D and 3-D objects, segment instances, mark lanes, and determine drivable regions. It is presently the largest dataset for self-driving AI, with over 100,000 videos of over 1,100-hour driving experiences at various times of the day and weather conditions. The BDD100K dataset collects video data using cameras (in collaboration with Nexar). The footage is captured with a Nexar dashcam at 720p and 30 frames per second. The collection includes videos from several cities in the USA, including New York, Berkeley, San Francisco, and the Bay Area.

Frame extraction is performed on 100,000 video clips to provide images with respective bounding box coordinates of the objects, considerably lowering the pre-processing time for this dataset. The frame of each video clip is extracted from the 10^{th} second of the video, allowing viewers to distinguish items. All photographs have a resolution of 1280 x 720 pixels in RGB. A JSON file containing the properties of each video, including the time of day, objects, scene, weather, and their respective bounding boxes, has been given (Table 2).

Dataset	BerkeleyDeepDrive BDD100k
Time and location	2018, NewYork, Berkeley, San Franciso, Bay Area (USA)
Traffic condition	Different weather, lighting, and road conditions
Sensor configuration	Monocular color camera, gyroscope, GPS, IMU, smartphone sensors, magnetometer
Data type and size	100000 videos, 40s each (1.8 TB) Video (mov) Image (jpg) Label (json)
Supplied resources	Raw data for training, validation, and testing; annotations, bounding box, lane marking, drivable area, pixel/instance-level, segmentation
Usage scenarios	Object detection, lane marking detection, drivable area detection, semantics egmentation

Table 2: Berkeley Deep Drive Dataset Details.

To reiterate, in the Discovery stage, the problem, context, and objectives for using this dataset are defined. The main aims of the study are to propose different filtering techniques for deep learning models in object detection and classification in 2-D image data. Therefore, only images of the dataset are being used for this purpose. The dataset consists of 100,000 images that are split into 3 subsets: 70,000 images in the training; 10,000 images in the validation; and 20,000 images in the testing.

Figure 9 shows that most photos have 9 to 21 different objects in each frame. However, some images contain 91 objects in each frame. There are multiple pedestrians on the sidewalk that the driverless car will need to identify in a picture with 91 bounding boxes.

Figure 10 depicts that the Berkeley dataset has seven distinct settings, including residential, city, and highway environments. The Berkeley dataset contains photos and videos from four separate locations in the United States: Berkeley, New York, The Bay Area, and San Francisco. This provides a sufficient distinction between the photos and ought to show enhanced object recognition results.

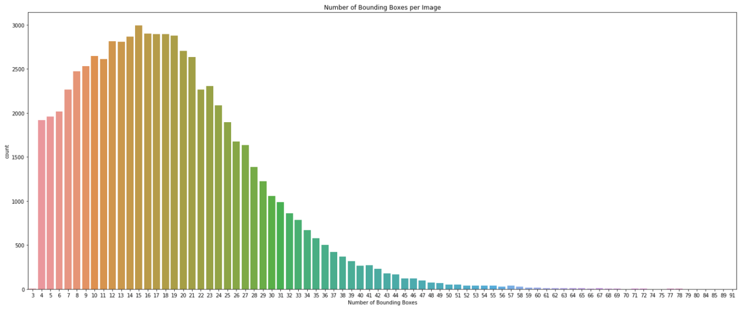

Figure 9: The number of Bounding Boxes per Image in the BDD100K Dataset [Hisho, 2022].

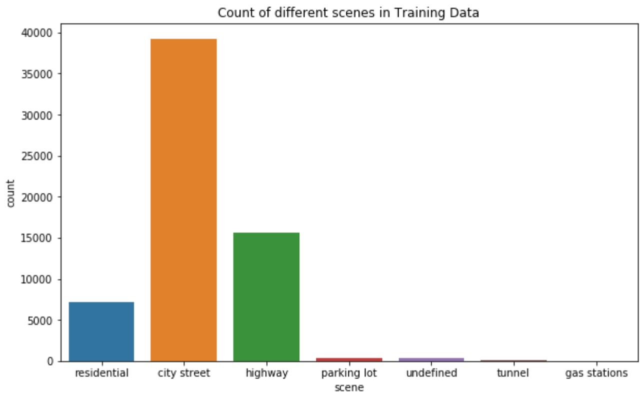

Figure 10: The number of images in different scenes in the BDD100K Dataset [Hisho, 2022].

As can be seen from Figure 11, the number of items in this dataset makes it a preferred option. Though the dataset is skewed (due to more 'vehicle' items), there is still a wide range of objects. The class imbalance may be 'corrected' by duplicating the photographs to match the number of 'car' objects. However, all images in the training dataset contain the label 'car.' The animal category (which would need to be spotted by driverless cars in the actual world) is missing in the dataset [Hisho, 2022].

In Phase 2 (Data Preparation), all the labels are in JSON format. They are converted to COCO JSON format followed by the YOLO format. To ensure that each image matches its label, a search of on-labeled data is conducted for removing images without labels. In the end, there are a total of 80821 remaining images, with 69,863 images in the testing; 10,000 images in the validation; and 958 images in the testing subset.

In Phases 3 and 4, the analyzed and visualized results of different models are presented in Section 5. Finally, in Phases 5 and 6, relevant findings and discussion are presented in Sections 6 and 7.

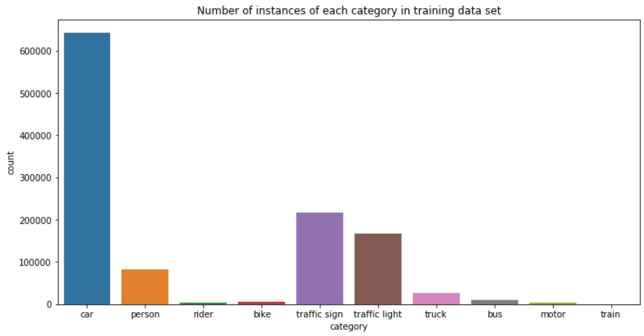

Figure 11: The number of images in different object categories in the BDD100K Dataset [Hisho, 2022].

3.2 Micro Methodology

3.2.1 Deep Learning Models

This study has selected 5 different deep learning models, which are commonly used for object detection, which are YOLOv5s, EfficientNet-B7, InceptionV4, Xception, and MobilenetV3.

a) *YOLOv5s*

You Only Look Once (YOLO) is one of the most common model architectures and object detection methods. It employs a neural network architecture to achieve high accuracy and overall processing speed [Jiang et al., 2022]. A YOLO model's first premise is residual blocks. They utilize 77 remaining blocks in the initial building design to form grids in the picture, as presented in Figure 12 [Wang and He, 2021].

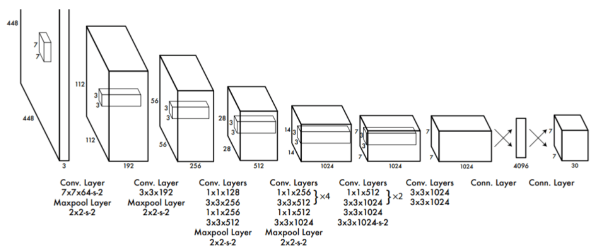

Figure 12: YOLO Architecture [Jiang et al., 2022].

Each grid serves as a center point, and a specific forecast is created for each grid. In the second grid, the bounding boxes are created by considering the center point for a particular forecast. While classification tasks perform well for each grid, separating the bounding boxes for each prediction is more challenging. The third and last grid is to utilize the intersection of union (IOU) to obtain the optimal bounding boxes for a specific object identification task [Alves-Oliveira et al., 2020].

Figure 13 presents the performance of different versions of YOLOv5 and EfficientDet in terms of Average Precision in the validation set and GPU Speed on the COCO Dataset. Results reveal that YOLO models have faster GPU speed for object detection, compared to different EfficientDet models. YOLOv5s is the fastest model amongst the 4 versions of YOLO but with the lowest average precision. In contrast, the YOLOv5x is the slowest model but with the highest average precision among the 4 versions [Ultralytics, 2022].

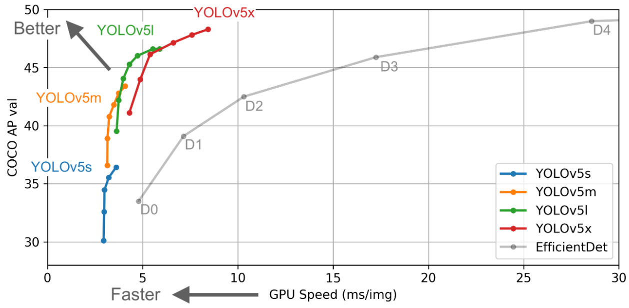

Figure 13: Performance of different versions of YOLO on the COCO Dataset [Ultralytics, 2022].

b) *EfficientNet-B7*

EfficientNet is an architecture and scaling approach to a convolutional neural network [Hien et al., 2021b] that uses a compound coefficient to consistently scale all depth/width/resolution dimensions (Figure 14). This scaling approach evenly scales network depth, width, and resolution using a predetermined set of scaling coefficients [Hoang and Jo, 2021].

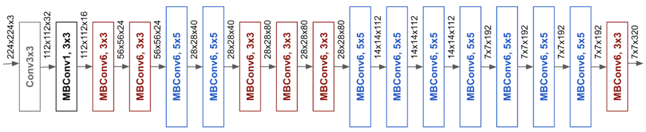

Figure 14: The architecture of EfficientNet [Hoang and Jo, 2021].

As can be seen from Figure 15, EfficientNet models outperform earlier CNNs in terms of accuracy and efficiency [Baheti et al., 2020].

c) *Xception*

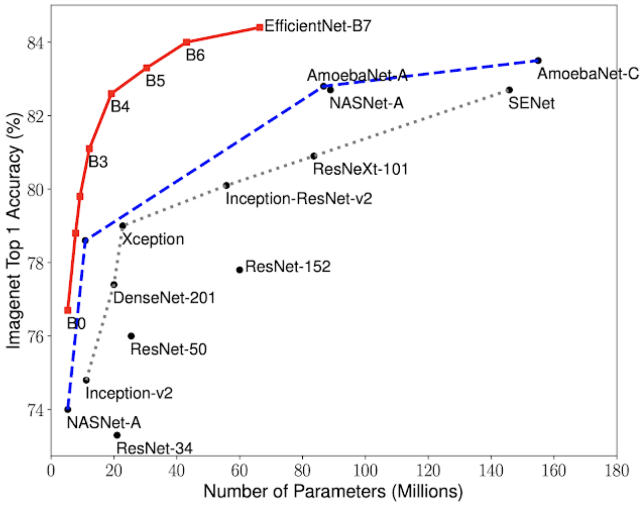

Figure 15: Performance of different versions of EfficientNet [Baheti et al., 2020].

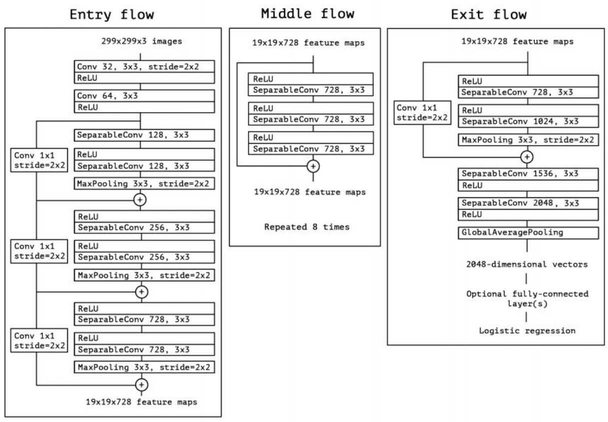

Figure 16: The architecture of Xception [Chollet, 2017].

Xception is a variation of the deep CNN architecture with Depthwise Separable Convolutions. Inception modules in CNN are introduced by Google as an intermediary step between normal convolution and the depthwise separable convolution process. In this context, a depthwise separable convolution may be viewed as an Inception module with an infinite number of towers. However, in a new deep CNN architecture (motivated by Inception), the Inception modules are substituted with depthwise separable convolutions [Chollet, 2017].

As presented in Figure 16, data is routed repeatedly 8 times through the entry flow and the middle flow, and ultimately, through the exit flow. In most conventional classification problems, the Xception algorithm outperforms VGG-16, ResNet, and Inception V3.

d) *MobilenetV3*

MobileNetV3 is a tuned CNN for mobile phone CPUs by using a mix of hardware-aware NAS (network architecture search) augmented by the NetAdapt algorithm. The network is a new efficient version of nonlinearities suitable for the mobile environment. It complements search approaches with an innovative powerful network design. The network architecture is depicted in Figure 17, which includes the use of hard swish activation and squeeze-and-excitation modules in the MBConv blocks [Kavyashree and El-Sharkawy, 2021].

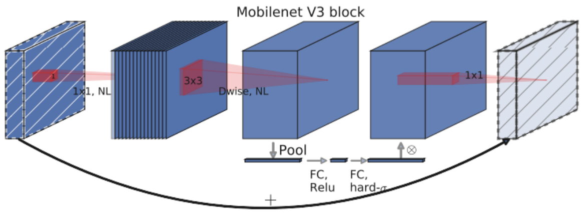

Figure 17: MobilenetV3 Architecture [Kavyashree and El-Sharkawy, 2021].

a) *InceptionV4*

Inception is an architectural design of a deep neural network that comprises replicated components known as Inception modules. CNN employs these Inception Modules to provide more efficient computation and deeper networks by reducing dimensionality using stacked 1×1 convolutions. The modules are created to address challenges such as computational cost and overfitting. Figure 18 depicts the architecture of Inception-v4 [Szegedy et al., 2017].

3.2.2 Filtering Techniques

a) *Hessian Filter*

A hybrid Hessian filtering technique can be used to filter 2-D and 3-D images. In 2-D images, the Hessian matrix can be defined as

$$H = \begin{bmatrix} H_{\mathrm{xx}} & H_{\mathrm{xy}} \\ H_{\mathrm{xy}} & H_{\mathrm{yy}} \end{bmatrix}$$

which is obtained by convolving the image with the Gaussian kernel's second derivatives in the x- and y-directions respectively. This filter is suitable for detecting continuous edges, such as wrinkles, vessels, and rivers. It may be utilized to compute a fraction of the entire image that contains such objects. The technique of Hessian is nearly identical to the Frangi filter, but it uses an alternative smoothing method.

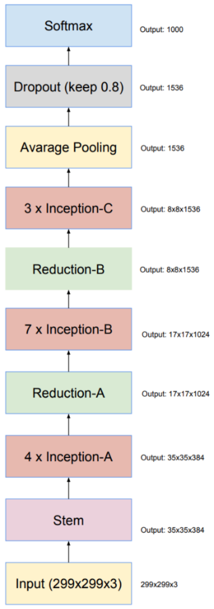

Figure 18: The overall schema of InceptionV4 [Szegedy et al., 2017].

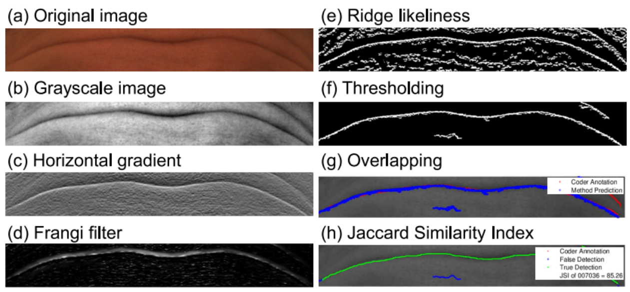

Figure 19: Application of Hybrid Hessian filter for winkle detection.

Figure 19 presents an application of how a Hybrid Hessian Filter detects the edges of human wrinkles. [Ng et al., 2014]

In image processing and computer vision, the Hessian matrix is often used to describe image processing operators. In normal mode analysis, the Hessian matrix can compute the various molecular frequencies in infrared spectroscopy.

In Figure 20, a driving scene image (on the left) has been applied with a Hessian filter (on the right).

b) Laplacian Filter

In the edge area of an image, the pixel intensity shows a ”jump” or a high variation

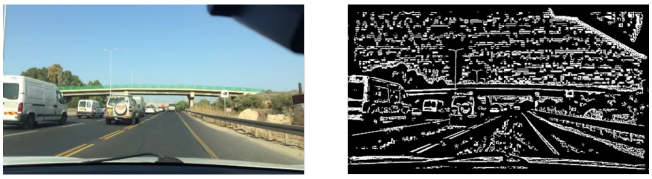

Figure 20: Application of Hessian filter to a driving image.

of intensity. Using Sobel Operator to get the first derivative of the intensity, an edge is characterized by a maximum, as illustrated in Figure 21 [OpenCV, 2022].

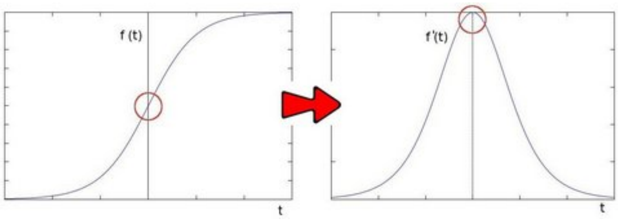

Figure 21: The first derivative of the intensity [OpenCV, 2022].

And while obtaining the second derivate, the value becomes zero, as shown in Figure 22.

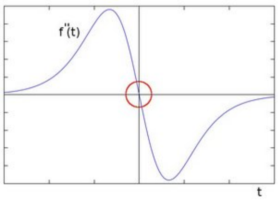

Figure 22: The second derivative of the intensity [OpenCV, 2022].

Therefore, this criterion can be used to detect edges in an image. Since images are 2-D, it is necessary to obtain the derivative in both dimensions.

The Laplacian operator is defined by:

$$\text{Laplace}(f) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

The Laplacian operator is implemented in OpenCV using the function Laplacian. Since the Laplacian utilizes the images' gradient, it employs the Sobel operator internally for its calculation.

To apply a Laplacian filter to an image, first, noise needs to be removed from the image by applying a Gaussian blur followed by converting the original image to a grayscale. A Laplacian operator can be applied to the grayscale image and store the output image, as illustrated in Figure 23.

Figure 23: Application of Laplacian filter to a driving image in this research.

c) *Hessian-based Ridge Detection Filter*

Ridges are eigenvalues of the matrix of second-order derivate of an image, also known as the Hessian matrix. Therefore, it is possible to create a ridge detector function using scikit-image to extract local minima ridges. First, noise needs to be removed from the image by applying a Gaussian blur followed by converting the original image to a grayscale. Then the grayscale image can be put into a ridge detector function to output the local minima ridges and save them as a new image presented in Figure 24.

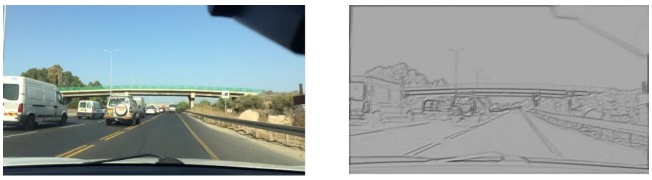

Figure 24: Application of Ridge Detection filter to a driving image in this research.

d) *Proposed Filtering Techniques*

In this research, a summary of the proposed filtering techniques for object detection in images (shown in Figure 25) is implemented as follows:

1. One of the 3 proposed filters, including Hessian, Laplacian, and Ridge Detection filters is applied to an input image;
2. The output of (1) is converted to grayscale, retaining only the edges of the image.
3. The grayscale image is then overlayed with the original input image with varying overlaying levels of 5% to 20%. This task includes alpha blending and masking with NumPy.

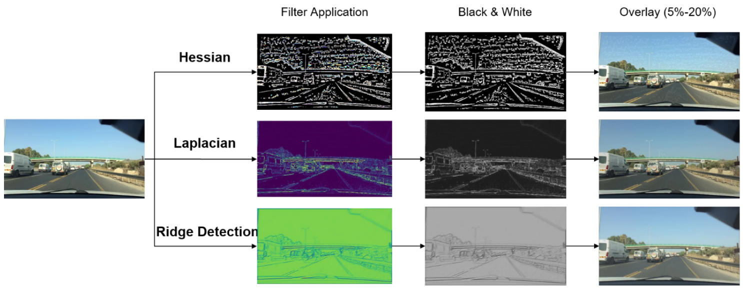

Figure 25: Proposed filtering techniques for object recognition in image data for autonomous vehicles.

4. The final image is fed into the deep learning model for object detection.

The critical reason behind this approach is that the final image will have highly sharpened edges.

3.2.3 Monitoring Energy Consumption

Type	Specification
Model	HP Pavilion 15
Operating System	Windows 10.0, 64 bits
Processor	Intel® CoreTM i7-8550U CPU @1.80GHz 1.99GHz
RAM	12.0GB
HDD Storage	1TB

Table 3: Specifications of the Computer used for Experiments.

Joulemeter power measuring software has been chosen for this study due to its simple and user-friendly interface. It can track and monitor the power consumption of individual apps. The PC specifications for all the experiments have been tabulated in Table 3.

Joulemeter does not allow automated calibration in Windows 10. As a result, manual calibrations are carried out using the setups shown in Figure 26. Furthermore, for the calibration, the battery must be at least 50% charged without a background program or process [Kothari and Bhattacharya, 2009].

3.2.4 Green House Gas Footprint Audit

Conversion factors are commonly used by the United Kingdom and international organizations to report greenhouse gas (GHG) emissions [Business et al., 2021]. The types of emissions are as follows:

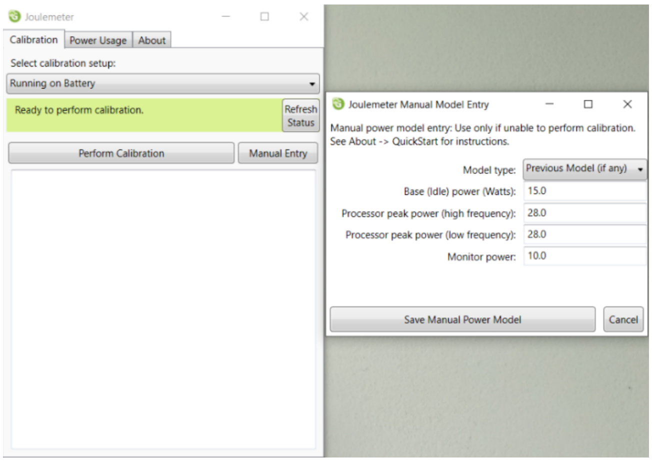

Figure 26: Joulemeter Manual Calibration.

- Scope 1 (direct emissions): Emissions from operations that the organization maintains or generates.
- Scope 2 (energy indirect): Emissions from using electricity, heat, steam, and cooling of the environment
- Scope 3 (other indirect): Emissions created by activities not defined as Scope 2.

The conversions for Levels 2 and 3 emissions are used for this research. This study focuses on the UK electricity, transmission, and distribution (T&D) variables, which are shown in Table 4.

Scope	**Activity**	**Type**	**Unit**	**kg CO2e**	**kg CO2**	**kg CH4**	**kg N2O**
2	Electricity generated	Electric-ity: UK	kWh	0.21233	0.21016	0.0008	0.00137
3	T&D-UK electricity	Electric-ity: UK	kWh	0.01879	0.01860	0.00007	0.00012
3	Distribution district heat & steam	5% loss	kWh	0.00899	0.00890	0.00006	0.00003

Table 4: GHG Conversion Factors from the UK Government Data.

4 Preliminary Research

In this preliminary research, 5 different deep learning models (YOLOv5s, EfficientNet-B7, Xception, MobilenetV3, and InceptionV4) have been built for object detection in autonomous vehicles with the BDD100K dataset. The results are depicted in Table 5 and Figures 27 to 29.

Model	Training time per epoch (s)	Inference time per epoch (s)	Log loss
Xception	3548.529	0.061	0.431
InceptionV4	3704.766	0.074	0.468
MobileNetV3	761.87	0.050	0.538
EfficientNet-B7	6224.268	0.112	0.425
YOLOv5s	2299.266	0.044	0.403

Table 5: Performance of Deep Learning models for object detection.

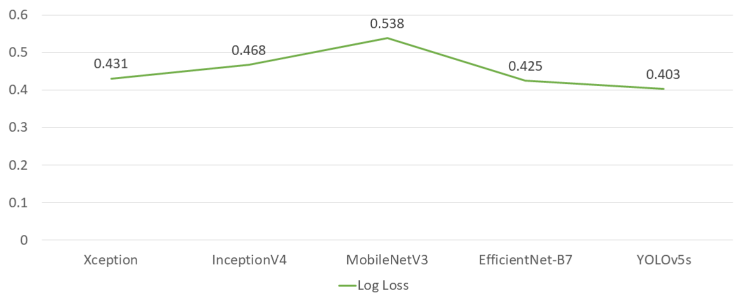

Figure 27: Log losses of the models.

Table 5 and Figure 27 evidence that YOLOv5s and EfficientNet-B7 models perform the best with the lowest log losses. This result informs a decision for the implementation of the 2 models in the next section.

Regarding the training and inference times, MobileNetV3 has the lowest performance. And in contrast, despite generating one of the best accuracies, EfficientNet-B7 takes the longest time for training and inference. Therefore, EfficientNet-B7 should be taken into consideration for optimization in Section 5. Meanwhile, YOLOv5s has become the best model in terms of accuracy, training, and inference time. This is a promising result and could act as a reference in model decision-making for researchers and manufacturers.

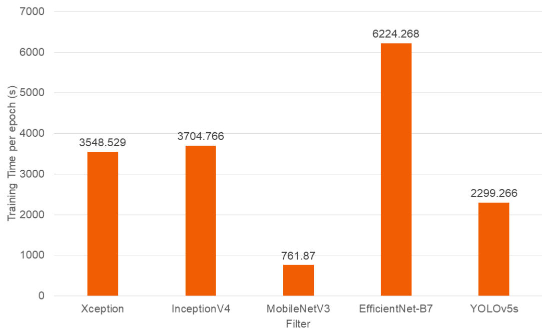

Figure 28: Training time per epoch.

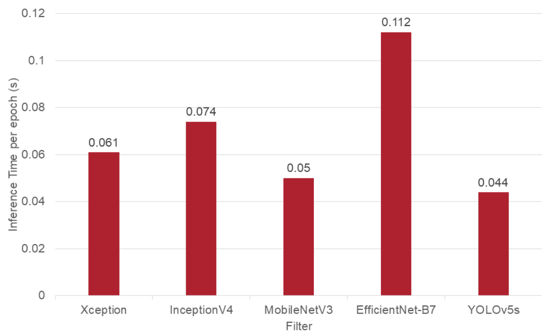

Figure 29: Inference time per epoch.

5 Findings and Discussion

5.1 Comparison of Filtering Techniques applying to the YOLOv5s model

Table 6 and Figure 30 have indicated that by using different filtering techniques, the mean Average Precision (mAP) of YOLOv5s models increased remarkably. Particularly, the Ridge Detection filter with 10% overlay, yields the highest mAP boost of 6.69%, from 56.67% to 63.36% (an 11.81% increase of the original mAP). This is followed by the Laplacian filter with a 5% overlay and the Hessian filter with a 10% overlay.

Figure 31 plots the mAP values for the Ridge Detection filter with a 10% overlay applied on 1000 test images. The graph shows that the mAP of YOLOv5s with a Ridge Detection filter is higher than the one without Ridge Detection Filter for most of the tested images.

T-test:

A series of 2-tailed t-tests are implemented to evaluate if there is a significant difference between the mean Average Precision (mAP) of object detection using YOLOv5s with and without different filters.

– Null Hypothesis (H0): mAP of model A = mAP of model B

Filter	Mean Average Precision (%)	Pre-process Time (ms)	Inference Time (ms)	Non-maximum Suppres-sion Time (ms)	Total Time per Image (ms)
No Filter	56.67	36.83	326.70	3.15	363.53
Hessian 20%	57.10	102.78	215.38	3.98	318.16
Hessian 10%	60.58	38.25	168.05	1.75	206.30
Hessian 5%	59.65	48.30	298.70	3.53	347.00
Laplacian 20%	57.56	37.20	167.63	1.78	204.83
Laplacian 10%	59.51	73.25	161.65	1.95	234.90
Laplacian 5%	62.43	112.45	158.78	1.53	271.23
Ridge Detection 20%	60.32	37.90	163.80	1.55	201.70
Ridge Detection 10%	63.36	39.23	167.85	2.85	207.08
Ridge Detection 5%	59.74	39.65	164.00	1.75	203.65

Table 6: Comparison of Filtering Techniques applying to the YOLOv5s model.

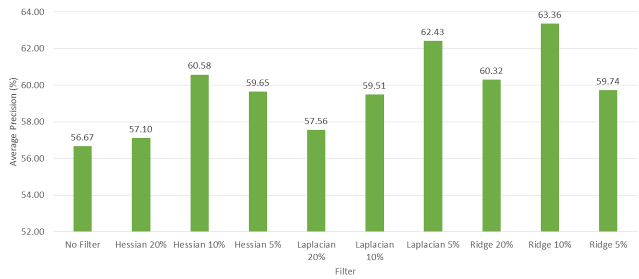

Figure 30: Mean Average Precision of Object Detection in YOLOv5s with different filtering techniques.

- Alternative Hypothesis (Ha): mAP of model A $\neq$ mAP of model B
- Chosen Confidence Level: 99%, which means the significance level $\alpha = 0.01$.

A series of paired 2-tail t-tests are conducted for the filtering techniques listed in Table 6 and the results are plotted into a heatmap in Figure 32. The value **1** indicates that there is a significant difference in the mAP (i.e., reject the Null Hypothesis) and **0** means there is no significant difference (i.e., accept the null hypothesis).

Figure 31 shows that mAP of most of the filtering techniques is significantly different from each other. Most importantly, the mAP values of the YOLOv5s model with all filtering techniques are significantly different from those that are without any filter.

Figures 31 and 32 evidence that the mAP of object detection for deep learning models with filtering techniques is better compared to the models without them. Remarkably, the total object detection time could be reduced up to 43.98% by using the Ridge Detection filter with a 5% overlay.

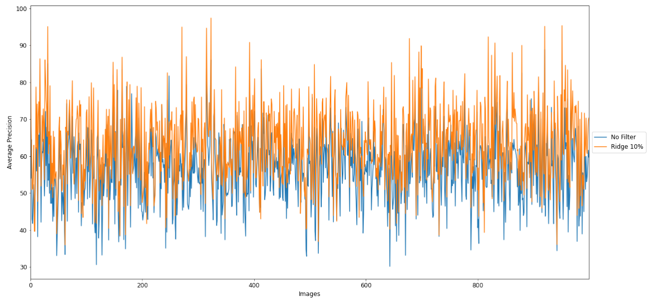

Figure 31: Means of Average Precision of YOLOv5s model with and without Ridge Detection Filter (10% overlay).

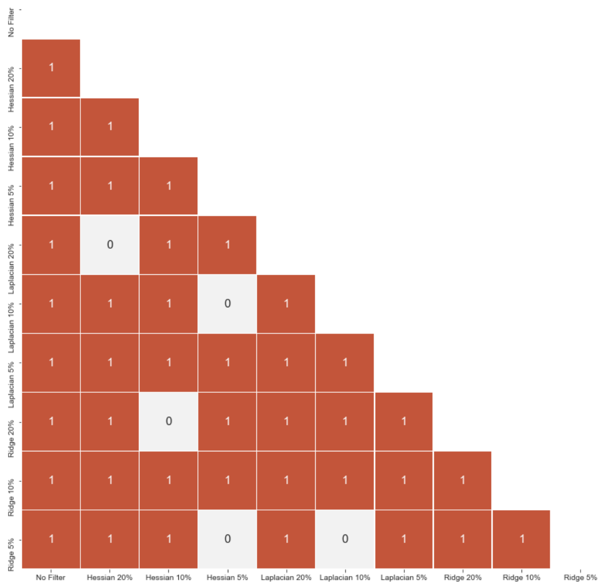

Figure 32: T-Test on mean Average Precision for the different filtering techniques on the YOLOv5s model.

Overall, object detection in the YOLOv5s models with 3 proposed filters (i.e., Hessian, Laplacian, and Hessian-based Ridge Detection) outperforms those without any filter. In this experiment, the Ridge Detection Filter appears to be the best-performing filter.

5.2 Comparison of Filtering Techniques applied to the EfficientNet-B7 model

Using the same methodology of Section 5.1, the results when comparing different filtering techniques applied to the EfficientNet-B7 model are presented below.

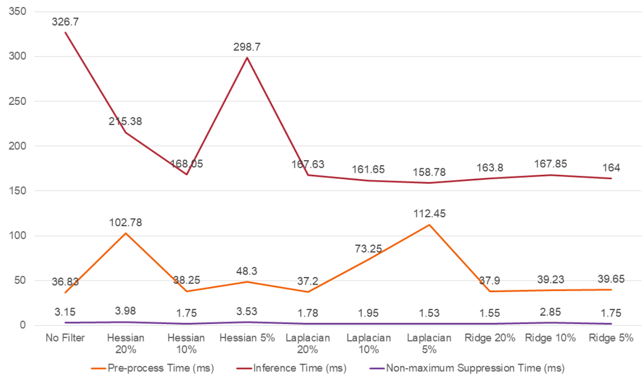

Figure 33: The pre-process, inference, and non-maximum suppression times of object detection in YOLOv5s with different filtering techniques.

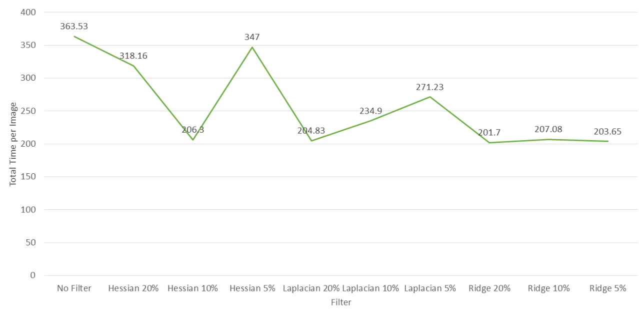

Figure 34: The total object detection time in YOLOv5s with different filtering techniques.

Table 7 and Figure 35 prove that with different filtering techniques, the mAP of EfficientNet-B7 models has increased considerably. Especially, the Laplacian filter with a 5% overlay has the highest mAP increase of 1.05%, from 42.46% to 44.76% (an increase of 5.42% from the original mAP), followed by the Ridge Detection filter with a 10% overlay and the Hessian filter with a 5% overlay.

Figure 36 also shows the mAP of EfficientNet-B7 with Laplacian filter with a 5% overlay is higher than without the filter for most of the 1000 tested images.

Figure 37 indicates that the mAP of the EfficientNet-B7 models with all filtering techniques is significantly better than the ones without filters (except for the Laplacian filter with 20% overlay).

Figure 38 also proves that the mean processing time for object detection by deep learning models with proposed filtering techniques is better compared to the ones without them. Notably, the total detection time could be reduced by up to 21.02% by using the Ridge Detection filter with a 5% overlay, shown in Figure 39.

In general, object detection in the EfficientNet-B7 models with 3 proposed filters (i.e., Hessian, Laplacian, and Hessian-based Ridge Detection) outperforms those without

Filter	Mean Average Precision (%)	Pre-process Time (ms)	Inference Time (ms)	Total Time per Image (ms)
No Filter	42.46	50.27	1417.64	1467.91
Hessian 20%	43.68	50.21	1140.07	1190.28
Hessian 10%	44.54	52.16	1129.73	1181.89
Hessian 5%	44.55	50.09	1178.97	1229.06
Laplacian 20%	42.98	48.40	1132.60	1181.00
Laplacian 10%	44.05	50.99	1263.43	1314.42
Laplacian 5%	44.76	56.88	1451.26	1508.14
Ridge 20%	44.32	53.61	1541.04	1594.65
Ridge 10%	44.72	49.22	1325.21	1374.43
Ridge 5%	44.54	47.17	1119.69	1166.86

Table 7: Comparison of Filtering Techniques for the EfficientNet-B7 model.

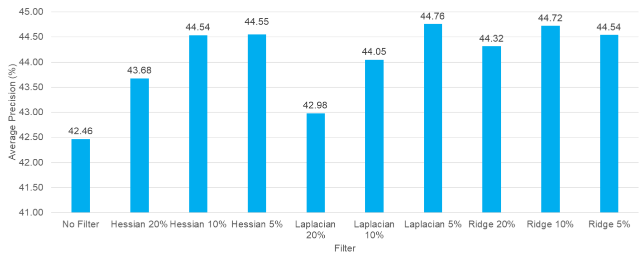

Figure 35: Mean Average Precision of object detection in EfficientNet-B7 with different filtering techniques.

any filter. And in this experiment, the Laplacian Filter becomes the best-performing filter among the rest.

5.3 Green House Gas Footprint Audit

a) Comparison of Energy Consumption of 5 Deep Learning Models

Aggregated energy consumption of many deep learning models with different filtering techniques is measured 10 times followed by an average, as depicted in Table 8. The total hardware energy is the sum of the energy consumed by the CPU, Disk, Monitor, and Base. The aggregated total energy consumption is the sum of total hardware energy and aggregated application energy. Lastly, the aggregated Average Energy Consumption per second (J/s) is calculated by dividing Aggregated Total Energy Consumption (KJ) by Aggregated Total Time (s) and then multiplying by 1000.

Among the 5 deep learning models, MobileNet appears to be the greenest model, while EfficienNet-B7 would be the least green. The Aggregated Total Energy Consumption (KJ)

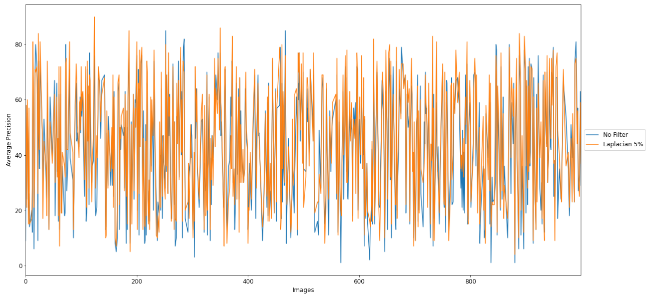

Figure 36: Compare the means of Average Precision of the EfficientNet-B7 model without and with Laplacian Filter with 5% overlay.

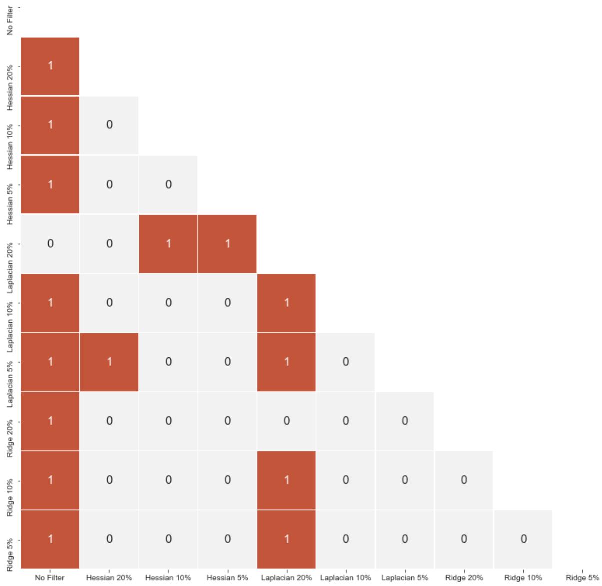

Figure 37: T-Test on Average Precision among different filtering techniques on the EfficientNet-B7 model.

of the 5 deep learning models is tabulated in Table 8. Subsequently, they are converted to GHG footprint (see Figure 40) using the greenhouse gas conversion factors in Table 4.

b) Comparison of YOLOv5s Models with and without filters

Similar to previous section a, energy consumption and emissions data are presented in Table 9 and Figure 41. From Table 9, it is clear that the energy consumption and emission of object recognition for YOLOv5s with proposed filtering techniques is less compared to the models without them. Moreover, using a Hessian filter with a 20% overlay and a Ridge Detection filter with a 5% overlay can significantly reduce energy

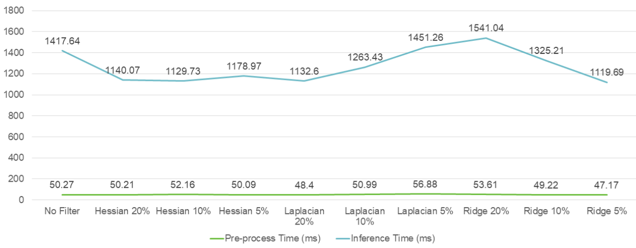

Figure 38: The pre-process, and inference time for object detection in EfficientNet-B7 with different filtering techniques.

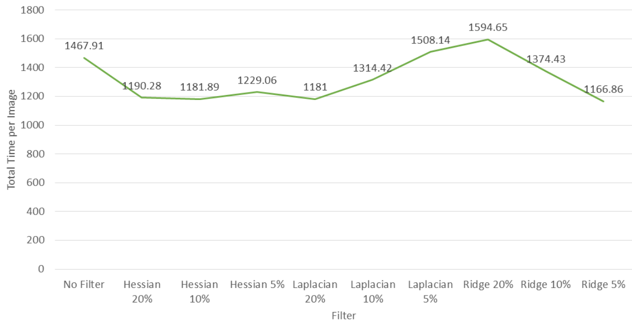

Figure 39: Total detection time for object detection in EfficientNet-B7 with different filtering techniques.

consumption and total emission, by 50.69% and 50.22% respectively.

c) Comparison between the EfficientNet-B7 Model with and without using filters

Similar to section a and b, Table 10 and Figure 42 indicates that the energy consumption and emission of object recognition for EfficientNet-B7 with proposed filtering techniques are better compared to the models without them. Especially, using a Hessian filter with 5% can significantly reduce energy consumption and total emission of the EfficientNet-B7 model by 33.38%.

6 Conclusion

The preliminary research of this study aims to build 5 different Deep Learning models, YOLOv5s, EfficientNet-B7, Xception, MobilenetV3, and InceptionV4, which are commonly used for 2-D object recognition. The findings reveal that YOLOv5s and EfficientNet-B7 models perform the best among the 5, with the lowest log losses. While MobileNetV3 has the lowest performance, it delivers relatively satisfactory training and inference time, as well as the lowest energy consumption. On the other hand, despite holding one of the best performances, EfficientNet-B7 consumes the longest time for training and inference; and incurs the highest energy consumption. Meanwhile, YOLOv5s has become the best model in terms of performance and sustainability. This result could act as a

Model	Aggregated Average Hardware Energy Consumption (KJ)					Aggregated Application Energy (KJ)	Aggregated Total Time (s)	Aggregated Total Energy Consumption (KJ)
	CPU (KJ)	Monitor (KJ)	Disk (KJ)	Base (KJ)	Total Hardware Energy (KJ)			
Xception	4.517	0.909	0.007	4.545	9.978	0.223	354859	10.201
InceptionV4	4.630	0.912	0.005	4.560	10.107	0.143	370484	10.249
MobilenetV3	0.823	0.216	0.001	1.080	2.119	0.076	76192	2.195
EfficientNet-B7	8.444	1.515	0.007	7.575	17.542	0.322	622438	17.863
YOLOv5s	2.184	0.609	0.003	3.045	5.840	0.137	229931	5.978

Table 8: Comparative Analysis of Energy Consumption of 5 Deep Learning Models for Object Recognition.

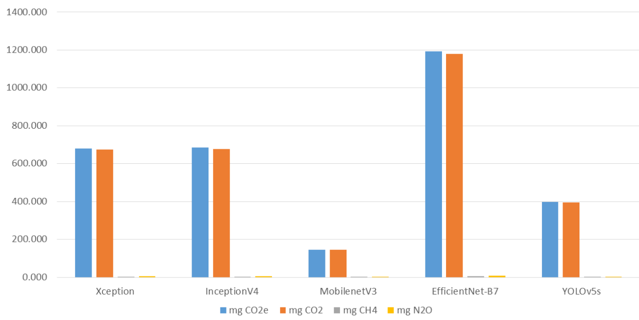

Figure 40: GHG Emissions for 5 different deep learning models.

good benchmarking, leading to a decision for optimizing YOLOv5s and EfficientNet-B7 models in this study.

This research primarily proposes 3 filtering techniques, Hessian, Laplacian, and Hessian-based Ridge Detection filters, for 2 deep learning models (YOLOv5s and EfficientNet-B7) in 2-D object recognition, optimizing their performance and sustainability. Applying to the BDD100K image dataset, those filtering techniques boosted the mean Average Precision (mAP) of the YOLOv5s and EfficientNet-B7 models remarkably.

- In the YOLOv5s model, the Ridge Detection filter with a 10% overlay brings an increase of the mAP to 11.81%, from 56.67% to 63.36%. Slightly lower than that, the Laplacian filter with a 5% overlay and the Hessian filter with a 10% overlay also

Model	Aggregated Average Hardware Energy Consumption (KJ)					Aggregated Application Energy (KJ)	Aggregated Total Time (s)	Aggregated Total Energy Consumption (KJ)
	CPU (KJ)	Monitor (KJ)	Disk (KJ)	Base (KJ)	Total Hardware Energy (KJ)			
No Filter	2.184	0.609	0.003	3.045	5.840	0.137	229.931	5.978
Hessian 5%	1.374	0.273	0.003	1.365	3.014	0.026	95.956	3.040
Hessian 10%	1.358	0.282	0.000	1.410	3.050	0.024	98.162	3.075
Hessian 20%	1.283	0.273	0.000	1.365	2.921	0.026	94.228	2.948
Laplacian 5%	1.328	0.297	0.000	1.485	3.110	0.025	101.022	3.135
Laplacian 10%	1.283	0.330	0.001	1.650	3.264	0.026	96.842	3.289
Laplacian 20%	1.347	0.288	0.000	1.440	3.075	0.022	97.439	3.098
Ridge 5%	1.317	0.273	0.000	1.365	2.955	0.022	94.668	2.976
Ridge 10%	1.357	0.285	0.000	1.425	3.067	0.025	97.321	3.091
Ridge 20%	1.390	0.288	0.001	1.440	3.119	0.037	99.877	3.156

Table 9: Comparison of Energy Consumption for YOLOv5s Models with and without using filters.

reached satisfactory results. A series of 2-tail t-tests are also implemented and has confirmed that all filtering techniques have significantly different mAP compared to the model without filters. Additionally, the mAP of most of the filtering techniques is also significantly different from each other. The mean detection time of object recognition with proposed filtering techniques is better compared to the models without them. Remarkably, the total object detection time could be reduced up to 43.98% by using the Ridge Detection filter with a 5% overlay. Using a Hessian filter with a 20% overlay and a Ridge Detection filter with a 5% overlay can also significantly reduce energy consumption and emission, by 50.69% and 50.22% respectively. Therefore, the 3 proposed filters have enhanced the YOLOv5s model resulting in better performance and sustainability in object recognition compared to ones without filters.

– In the EfficientNet-B7 model, the Laplacian filter with a 5% overlay increased the mAP to 5.42%, from 42.46% to 44.76%. Moreover, the Ridge Detection filter with a 10% overlay and Hessian filter with a 5% overlay also delivered promising results. A set of 2-tail t-tests has indicated that all filtering techniques have a significantly different mAP compared to the model without filters, except for the Laplacian filter with a 20% overlay. Furthermore, the proposed filters can also help reduce the mean

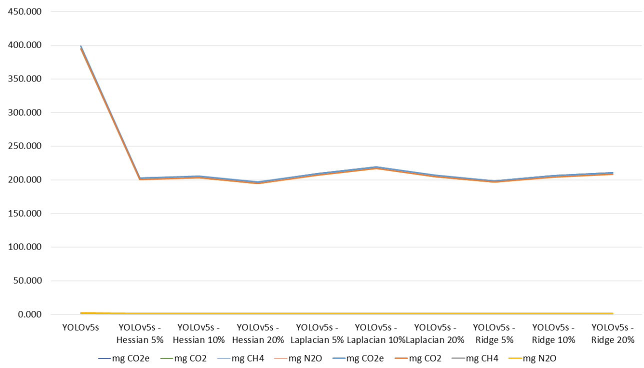

Figure 41: Emission Data of YOLOv5s with and without using different filters.

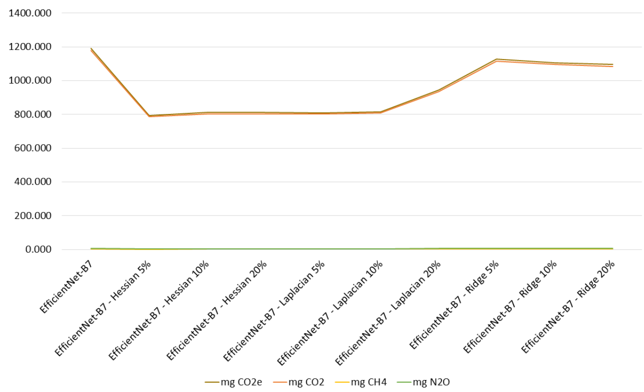

Figure 42: Emissions Data of EfficienNet-B7 with and without using filters.

detection time. Particularly, the total detection time could be reduced up to 21.02% by using the Ridge Detection filter with a 5% overlay. And using the Hessian filter with 5% can significantly reduce energy consumption and emission by 33.38%. Again, the 3 proposed filters have enabled the EfficientNet-B7 model to outperform its performance and sustainability when not using filters for object recognition.

Despite substantial efforts that have been conducted in the past [He et al., 2022, Wang et al., 2021, Cai and Vasconcelos, 1483], this study has proposed high-efficiency filtering techniques for object recognition and classification for autonomous vehicles. These results may lead to more future studies relating to filtering algorithms for object detection and classification in autonomous cars.

7 Recommendations and Future Work

One of the notable findings from the Literature Review section is that autonomous vehicles are currently yet to be an environmental-friendly solution in transportation,

Model	Aggregated Average Hardware Energy Consumption (KJ)					Aggregated Application Energy (KJ)	Aggregated Total Time (s)	Aggregated Total Energy Consumption (KJ)
	CPU (KJ)	Monitor (KJ)	Disk (KJ)	Base (KJ)	Total Hardware Energy (KJ)			
No Filter	8.444	1.515	0.007	7.575	17.542	0.322	622.438	17.863
Hessian 5%	5.657	1.026	0.000	5.130	11.813	0.087	396.411	11.900
Hessian 10%	5.837	1.038	0.001	5.190	12.066	0.089	394.872	12.155
Hessian 20%	5.799	1.050	0.000	5.250	12.099	0.079	395.932	12.178
Laplacian 5%	5.712	1.059	0.000	5.295	12.066	0.072	393.187	12.137
Laplacian 10%	5.852	1.053	0.000	5.265	12.170	0.071	386.141	12.241
Laplacian 20%	6.892	1.200	0.001	6.000	14.093	0.084	454.514	14.177
Ridge 5%	8.293	1.413	0.008	7.065	16.778	0.134	563.343	16.912
Ridge 10%	8.087	1.398	0.001	6.990	16.475	0.124	560.665	16.599
Ridge 20%	7.990	1.383	0.000	6.915	16.288	0.127	556.514	16.415

Table 10: Comparative Energy Consumption of EfficientNet-B7 Model with and without using filters.

however, they will become an indispensable and greener solution in the long-term vision. Scientists and manufacturers are recommended to use green power for manufacturing and operating autonomous vehicles, optimize the vehicle lifetime, schedule battery replacements, and conduct the transition from private vehicle use to shared and community use. Improving the sustainability level of autonomous vehicles necessitates collaboration on minimizing vehicle production supply chain impacts and encouraging clean power sources in electrical infrastructure decision-making.

Based on the findings of the Preliminary Research, the YOLOv5s and EfficientNet-B7 models are recommended for object detection and classification in autonomous vehicles instead of the Xception, MobilenetV3, and InceptionV4 models. Particularly, the YOLOv5s model is highly recommended for mobile applications since it holds a good balance of precision, detection time, and sustainability.

In the main research section, the proposed Hessian, Laplacian, and Hessian-based Ridge Detection filters have significantly improved the performance and sustainability of YOLOv5s and EfficientNet-B7 models. Therefore, these image-filtering techniques can be applied in further research to reduce motion blurs and mitigate some of the limitations of image data in autonomous vehicles. Furthermore, while constructing massive deep learning systems, developers should employ energy-tracking tools to reduce the ICT sector's carbon impact. This approach also provides a better consumer experience in terms of how long the batteries in technical products last. This is since the impact of deep learning would be considerably bigger than that demonstrated in this study when a program of a few thousand lines of code is produced.

Future research may gear towards evaluating more deep learning models, and not only in object detection and classification models in autonomous vehicles. Further development of different filtering techniques should be implemented such as optimizing the filters, using different overlaying techniques, or proposing more filtering approaches. Future work also tends to propose different filtering techniques in 3D object detection in LiDAR point cloud data. To enhance the analytical report, a full evaluation of how deep learning affects power use and emissions in a program should be undertaken using more parameters, such as memory utilization. More profiling tools as well as footprint audit tools can be used in later studies to extract more accurate auditing data.

Acknowledgements

This study is supported and implemented under the framework of the Erasmus Mundus Joint Masters Degree (EMJMD) in Green Networking and Cloud Computing (GENIAL) program (610619-EPP-1-2019-1-FR-EPPKA1-JMD-MOB).

References

[Ahangar et al., 2021] Ahangar, M., Ahmed, Q., Khan, F., and Hafeez, M. (2021). A survey of autonomous vehicles: Enabling communication technologies and challenges. *Sensors*, 21(3):706.

[Alves-Oliveira et al., 2020] Alves-Oliveira, P., Gomes, S., Chandak, A., Arriaga, P., Hoffman, G., and Paiva, A. (2020). Software architecture for yolo, a creativity-stimulating robot. *SoftwareX*, 11:100461.

[Andrew, 2015] Andrew, M. (2015). *The road to autonomy: driverless cars and the implications for insurance*.

[Automotive Engineers, 2021] Automotive Engineers, S. (2021). Sae levels of driving automationtm refined for clarity and international audience. *Society of Automotive Engineers*, 3 5. Accessed 10 3 2022.

[Bagloee et al., 2016] Bagloee, S., Tavana, M., Asadi, M., and Oliver, T. (2016). Autonomous vehicles: challenges, opportunities, and future implications for transportation policies. *Journal of modern transportation*, 24(4):284–303.

[Baheti et al., 2020] Baheti, B., Innani, S., Gajre, S., and Talbar, S. (2020). *Eff-unet: A novel architecture for semantic segmentation in unstructured environment*. IEEE.

[Behere and Torngren, 2015] Behere, S. and Torngren, M. (2015). *A functional architecture for autonomous driving*.

[Bjørn and Hauschild, 2012] Bjørn, A. and Hauschild, M. (2012). Absolute versus relative environmental sustainability: What can the cradle-to-cradle and eco-efficiency concepts learn from each other? *Journal of Industrial Ecology*, 18(2):321–332.

[Bolduc, 2018] Bolduc, D. (2018). Volvo, baidu to co-develop evs with level 4 autonomy for china. *Automotive News Europe*, 1 11. Accessed 11 3 2022.

[Brummelen et al., 2018] Brummelen, J., O'Brien, M., Gruyer, D., and Najjaran, H. (2018). Autonomous vehicle perception: The technology of today and tomorrow. *Transportation research part C: emerging technologies*, 89:384–406.

[Business et al., 2021] Business, U., Energy, and Strategy, I. (2021). *Greenhouse gas reporting: conversion factors 2021*.

[Cai and Vasconcelos, 1483] Cai, Z. and Vasconcelos, N. (1483). Cascade r-cnn: high quality object detection and instance segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 5:43.

[Che et al., 2019] Che, E., Jung, J., and Olsen, M. (2019). Object recognition, segmentation, and classification of mobile laser scanning point clouds: A state of the art review. *Sensors*, 19(4):810.

[Chen et al., 2014] Chen, B., Yuan, H., Li, Q., Lam, W., Shaw, S., and Yan, K. (2014). Map-matching algorithm for large-scale low-frequency floating car data. *International Journal of Geographical Information Science*, 28(1):22–38.

[Cheng et al., 2018] Cheng, B., Wei, Y., Shi, H., Feris, R., Xiong, J., and Huang, T. (2018). *Revisiting rcnn: On awakening the classification power of faster rcnn*.

[Chollet, 2017] Chollet, F. (2017). *Xception: Deep learning with depthwise separable convolutions*.

[Climate Change, 2019] Climate Change, I. P. (2019). *The Fifth Assessment Report of IPCC*. IPCC, Geneva, Switzerland.

[Code, 2022] Code, P. W. (2022). Object detection on coco test-dev. Accessed 30 05 2022.

[Contreras-Castillo et al., 2017] Contreras-Castillo, J., Zeadally, S., and Ibáñez, J. (2017). A seven-layered model architecture for internet of vehicles. *Journal of Information and Telecommunication*, 1(1):4–22.

[Cui and Sabaliauskaite, 2017] Cui, J. and Sabaliauskaite, G. (2017). On the alignment of safety and security for autonomous vehicles. In *IARIA CYBER*. Barcelona, Spain.

[Data, 2022] Data, O. W. (2022). Co2 emissions. Accessed 20 01 2022.

[Davies, 2020] Davies, C. (2020). Audi abandons self-driving plans for current flagship. *Slash Gear*, 28 4. Accessed 11 3 2022.

[Davis, 2021] Davis, R. (2021). Hyperdrive daily: The driverless shuttle helping toyota win gold. *Bloomberg*, 2 8. Accessed 11 3 2022.

[DeKort, 2019] DeKort, M. (2019). The hype of geofencing for autonomous vehicles. *Medium*, 10 3. Accessed 11 3 2022.

[Dickmann et al., 2016] Dickmann, J., Klappstein, J., Hahn, M., Appenrodt, N., Bloecher, H., Werber, K., and Sailer, A. (2016). *Automotive radar the key technology for autonomous driving: From detection and ranging to environmental understanding*. IEEE.

[Faisal et al., 2019] Faisal, A., Kamruzzaman, M., Yigitcanlar, T., and Currie, G. (2019). Understanding autonomous vehicles. *Journal of transport and land use*, 12(1):45–72.

[Fan and Xu, 2019] Fan, C. and Xu, X. (2019). Influences of autonomous cars on the insurance market from the perspectives of insurance companies and auto insurance agencies. *Journal of Applied Finance & Banking*, 9(4):11–35.

[Feng and Liu, 2019] Feng, Y. and Liu, H. (2019). Augmented reality for robocars. *IEEE Spectrum*, 56(12):22–27.

[Fettweis and Zimmermann, 2008] Fettweis, G. and Zimmermann, E. (2008). Ict energy consumption-trends and challenges.

[Fox, 2021] Fox, E. (2021). Tesla likely to achieve level 4 autonomy in 2022, says elon musk. *Tesmanian Blog*, 29 12. Accessed 10 3 2022.

[Francis and Anoop, 2014] Francis, J. and Anoop, B. (2014). Significance of hand gesture recognition systems in vehicular automation-a survey. *International Journal of Computer Applications*, 99(7):50–55.

[Friedrich, 2016] Friedrich, B. (2016). The effect of autonomous vehicles on traffic.

[Gallardo et al., 2017] Gallardo, N., Gamez, N., Rad, P., and Jamshidi, M. (2017). Autonomous decision making for a driver-less car.

[García-Martín et al., 2019] García-Martín, E., Lavesson, N., Grahn, H., Casalicchio, E., and Boeva, V. (2019). How to measure energy consumption in machine learning algorithms.

[Gill, 2021] Gill, T. (2021). Ethical dilemmas are really important to potential adopters of autonomous vehicles. *Ethics and information technology*, 23(4):657–673.

[Grigorescu et al., 2020] Grigorescu, S., Trasnea, B., Cocias, T., and Macesanu, G. (2020). A survey of deep learning techniques for autonomous driving. *Journal of Field Robotics*, 37(3):362–386.

[Group, 2021] Group, M.-B. (2021). First internationally valid system approval for conditionally automated driving. *Mercedes-Benz Group*, 17 12.

[Guan et al., 2016] Guan, H., Li, J., Cao, S., and Yu, Y. (2016). Use of mobile lidar in road information inventory: A review. *International Journal of Image and Data Fusion*, 7(3):219–242.

[Haldorai et al., 2021] Haldorai, A., Murugan, S., and Ramu, A. (2021). Evolution, challenges, and application of intelligent ict education: An overview. *Computer Applications in Engineering Education*, 29(3):562–571.

[Hars, 2016] Hars, A. (2016). *Misconception 7: To Convince us That They are Safe, Self-Driving Cars Must Drive Hundreds of.* Millions of Miles.

[Hartmann, 2017] Hartmann, C. (2017). *TechDay piloted driving – The traffic jam pilot in the new Audi A8*. Audi Media Center.

[Hawkins and Habib, 2019] Hawkins, J. and Habib, K. (2019). Integrated models of land use and transportation for the autonomous vehicle revolution. *Transport reviews*, 39(1):66–83.

[Hawkins et al., 2012] Hawkins, T., Singh, B., Majeau-Bettez, G., and Strømman, A. (2012). Comparative environmental life cycle assessment of conventional and electric vehicles. *Journal of Industrial Ecology*, 17(1):53–64.

[He et al., 2022] He, M., Wang, Y., Wu, J., Wang, Y., Li, H., Li, B., Gan, W., Wu, W., and Qiao, Y. (2022). Cross domain object detection by target-perceived dual branch distillation.

[Hien, 2022] Hien, N. (2022). Deep learning analysis and optimization for different data sources in autonomous vehicles. *Luleå University of Technology*, 2022. Accessed 11 23 2022.

[Hien et al., 2021a] Hien, N., Hoang, H., Hieu, N., and Tien, N. (2021a). Keyphrase extraction model: A new design and application on tourism information. *Informatica*, 45(4).

[Hien et al., 2021b] Hien, N., Huy, L., and Hieu, N. (2021b). Artwork style transfer model using deep learning approach. *Cybernetics and Physics*, 10(3):127–137.

[Hien and Kor, 2021] Hien, N. and Kor, A.-L. (2021). Analysis and prediction model of fuel consumption and carbon dioxide emissions of light-duty vehicles. *Applied Sciences*, 12(2):803.

[Hilty et al., 2009] Hilty, L., Coroama, V., Eicker, M., Ruddy, T., and Müller, E. (2009). The role of ict in energy consumption and energy efficiency. In *Report to the European Commission, DG INFSO, Project ICT ENSURE*. European ICT Sustainability Research, Graz University.

[Hisho, 2022] Hisho, R. (2022). Berkeley deep drive dataset — bdd100k. Accessed 25 05 2022.

[Hoang and Jo, 2021] Hoang, V. and Jo, K. (2021). *Practical Analysis on Architecture of EfficientNet*. IEEE.

[Honda, 2021] Honda (2021). Honda to begin sales of legend with new honda sensing elite. *Honda*, 4 3.

[Hu et al., 2020a] Hu, J., Bhowmick, P., Arvin, F., Lanzon, A., and Lennox, B. (2020a). Cooperative control of heterogeneous connected vehicle platoons: An adaptive leader-following approach. *IEEE Robotics and Automation Letters*, 5(2):977–984.

[Hu et al., 2020b] Hu, J., Bhowmick, P., Arvin, F., Lanzon, A., and Lennox, B. (2020b). Cooperative control of heterogeneous connected vehicle platoons: An adaptive leader-following approach. *IEEE Robotics and Automation Letters*, 5(2):977–984.

[Hu et al., 2021] Hu, J., Bhowmick, P., Jang, I., Arvin, F., and Lanzon, A. (2021). A decentralized cluster formation containment framework for multirobot systems. *IEEE Transactions on Robotics*, 37(6):1936–1955.

[Hussain and Zeadally, 2018] Hussain, R. and Zeadally, S. (2018). Autonomous cars: Research results, issues, and future challenges. *IEEE Communications Surveys & Tutorials*, 21(2):1275–1313.

[Ilková and Ilka, 2017] Ilková, V. and Ilka, A. (2017). *Legal aspects of autonomous vehicles—an overview*. IEEE.

[Jacobson, 2017] Jacobson (2017). 100roadmaps for 139 countries of the world. *Joule*, 1(1):108–121.

[Jedrasiak et al., 2013] Jedrasiak, K., Daniec, K., and Nawrat, A. (2013). The low cost micro inertial measurement unit. In *IEEE 8th Conference on Industrial Electronics and Applications (ICIEA), Melborne*, Australia.

[Jiang et al., 2022] Jiang, P., Ergu, D., Liu, F., Cai, Y., and Ma, B. (2022). A review of yolo algorithm developments. *Procedia Computer Science*, 199:1066–1073.

[Jiang et al., 2011] Jiang, T., Alfadhl, Y., and Chai, K. (2011). *Efficient dynamic scheduling scheme between vehicles and roadside units based on IEEE 802.11 p/WAVE communication standard*. IEEE.

[Kang et al., 2019] Kang, Y., Yin, H., and Berger, C. (2019). Test your self-driving algorithm: An overview of publicly available driving datasets and virtual testing environments. *IEEE Transactions on Intelligent Vehicles*, 4(2):171–185.

[Kavyashree and El-Sharkawy, 2021] Kavyashree, P. and El-Sharkawy, M. (2021). Compressed mobilenet v3: a light weight variant for resource-constrained platforms.

[Klaus et al., 2020] Klaus, S., Wendt, T., Stark, W., Gottfredson, M., Tsang, R., and Schallehn, M. (2020). Electric and autonomous vehicles: The future is now. Accessed 28 05 2022.

[Kopelias et al., 2020] Kopelias, P., Demiridi, E., Vogiatzis, K., Skabardonis, A., and Zafiropoulou, V. (2020). Connected & autonomous vehicles–environmental impacts–a review. *Science of the total environment*, 712:135237.

[Kothari and Bhattacharya, 2009] Kothari, N. and Bhattacharya, A. (2009). Joulemeter: Virtual machine power measurement and management. Msr tech report,.

[Kukkala et al., 2018] Kukkala, V., Tunnell, J., Pasricha, S., and Bradley, T. (2018). Advanced driver-assistance systems: A path toward autonomous vehicle. *IEEE Consumer Electronics Magazine*, 7(5):18–25.

[Kumar and Srivastava, 2020] Kumar, A. and Srivastava, S. (2020). Object detection system based on convolution neural networks using single shot multi-box detector. *Procedia Computer Science*, 171:2610–2617.

[Law et al., 2019] Law, H., Teng, Y., Russakovsky, O., and Deng, J. (2019). Cornernet-lite: Efficient keypoint based object detection. arXiv:1904.08900.

[Lawrence, 2021] Lawrence, A. (2021). Volkswagen and microsoft collaborate on autonomous driving tech. Accessed 29 05 2022.

[Lee, 2020] Lee, T. (2020). Waymo finally launches an actual public, driverless taxi service. *ARS Technica*, 10 8. Accessed 11 3 2022.

[Lew et al., 2018] Lew, F., Mason, J., and Meroux, D. (2018). *Three Revolutions in Urban Transportation*. ITDP, UC Davis.

[Li and Ibanez-Guzman, 2020] Li, Y. and Ibanez-Guzman, J. (2020). Lidar for autonomous driving: The principles, challenges, and trends for automotive lidar and perception systems. *IEEE Signal Processing Magazine*, 37(4):50–61.

[Litman, 2017] Litman, T. (2017). *Autonomous vehicle implementation predictions*. Victoria Transport Policy Institute, Victoria, Canada.

[Liu et al., 2020] Liu, B., He, S., Xing, Z., and Kang, N. (2020). Study on the inertial navigation system for autonomous driving.

[Maalej et al., 2018] Maalej, Y., Sorour, S., Abdel-Rahim, A., and Guizani, M. (2018). Vanets meet autonomous vehicles: Multimodal surrounding recognition using manifold alignment. *IEEE Access*, 6:29026–29040.

[Magna, 2022] Magna (2022). Max4: Magna's formula for winning the self-driving car race. Accessed 11 3 2022.

[Mane, 2021] Mane, C. (2021). Here's when we'll see driverless cars with level 5 autonomy. *Hot Cars*, 31 8. Accessed 11 3 2022.

[Marchegiani and Fafoutis, 2022] Marchegiani, L. and Fafoutis, X. (2022). How well can driverless vehicles hear? a gentle introduction to auditory perception for autonomous and smart vehicles.

[Markets, 2021] Markets, R. A. (2021). Autonomous commercial vehicle global market report 2021: Covid-19 growth and change to 2030. Accessed 2022 3 10].

[Martinez and Viegas, 2017] Martinez, L. and Viegas, J. (2017). Assessing the impacts of deploying a shared self-driving urban mobility system: An agent-based model applied to the city of lisbon, portugal. *International Journal of Transportation Science and Technology*, 6(1):13–27.

[Mike et al., 2017] Mike, D., Medasani, S., Behringer, R., and Trivedi, M. (2017). Self-driving cars. *Computer*, 50(12):18–23.

[Morris, 2021] Morris, J. (2021). Why is tesla's full self-driving only level 2 autonomous? *Forbes*, 12 3. Accessed 10 3 2022.

[Muhammad et al., 2020] Muhammad, K., Ullah, A., Lloret, J., Ser, J., and Albuquerque, V. (2020). Deep learning for safe autonomous driving: Current challenges and future directions. *IEEE Transactions on Intelligent Transportation Systems*, 22(7):4316–4336.

[Nastjuk et al., 2020] Nastjuk, I., Herrenkind, B., Marrone, M., Brendel, A., and Kolbe, L. (2020). What drives the acceptance of autonomous driving? an investigation of acceptance factors from an end-user's perspective. *Technological Forecasting and Social Change*, 161:120319.

[Ng et al., 2014] Ng, C.-C., Yap, M., Costen, N., and Li, B. (2014). Automatic wrinkle detection using hybrid hessian filter. In *Asian Conference on Computer Vision*, page 609–622.

[Nguyen and Le, 2013] Nguyen, A. and Le, B. (2013). 3d point cloud segmentation: A survey.

[Ondrus et al., 2020] Ondrus, J., Kolla, E., Vertal, P., and Saric, Z. (2020). How do autonomous cars work? *Transportation Research Procedia*, 44:226–233.

[OpenCV, 2022] OpenCV (2022). Laplace operator. *OpenCV*, 29 10. Accessed 30 10 2022.

[Paya et al., 2009] Paya, L., Fernandez, L., Reinoso, O., Gil, A., and Ubeda, D. (2009). Appearance-based dense maps creation. In *6th Int Conf on Informatics in Control*.

[Pozna and Antonya, 2016] Pozna, C. and Antonya, C. (2016). Issues about autonomous cars. In *IEEE 11th International Symposium on Applied Computational Intelligence and Informatics (SACI*, Timisoara, Romania.

[Quan, 2019] Quan, J. (2019). Enhanced geographic information system-based mapping of local climate zones in beijing, china. *Science China Technological Sciences*, 62(12):2243–2260.

[Rahiman and Zainal, 2013] Rahiman, W. and Zainal, Z. (2013). *An overview of development GPS navigation for autonomous car*. IEEE.

[Research, 2018] Research, A. (2018). Abi research forecasts 8 million vehicles to ship with sae level 3, 4 and 5 autonomous technology in 2025. Accessed 28 05 2022.

[Roth, 2021] Roth, S. (2021). California just hit 95along for the ride? *Los Angeles Times*, 29 04. Accessed 01 01 2022.

[Saeed et al., 2016] Saeed, A., Madjid, T., Mohsen, A., and Tracey, O. (2016). Autonomous vehicles: challenges, opportunities, and future implications for transportation policies. *Journal of Modern Transportation*, 24:284–303.

[Schwarting et al., 2018] Schwarting, W., Alonso-Mora, J., and Rus, D. (2018). Planning and decision-making for autonomous vehicles. *Annual Review of Control, Robotics, and Autonomous Systems*, 1:187–210.

[Shieh et al., 2018] Shieh, W., Hsu, C., Lin, C., and Wang, T. (2018). Investigation of vehicle positioning by infrared signal-direction discrimination for short-range vehicle-to-vehicle communications. *IEEE Transactions on Vehicular Technology*, 67(12):11563–11574.

[Singh and Saini, 2021] Singh, S. and Saini, B. (2021). *Autonomous cars: Recent developments, challenges, and possible solutions*. IOP Publishing.

[Sood and Kanwar, 2014] Sood, M. and Kanwar, S. (2014). *Clustering in MANET and VANET: A survey*. IEEE.

[Straka et al., 2021] Straka, W., Kondragunta, S., Wei, Z., Zhang, H., Miller, S., and Watts, A. (2021). Examining the economic and environmental impacts of covid-19 using earth observation data. *Remote Sensing*, 15(5).

[Swedish Environmental Research Institute, 2019] Swedish Environmental Research Institute (2019). ”new report on climate impact of electric car batteries. *ivl*, 04 12. Accessed 01 01 2022.

[Szegedy et al., 2017] Szegedy, C., Ioffe, S., Vanhoucke, V., and Alemi, A. (2017). Inception-v4, inception-resnet and the impact of residual connections on learning.

[Thrun, 2010] Thrun, S. (2010). Toward robotic cars. *Communications of the ACM*, 53(4):99–106.

[Ulitskaya, 2021] Ulitskaya, J. (2021). ‘super-duper cruise’ didn’t cut it: Gm’s ultra cruise aims to elevate hands-free driving. *Cars.com*, 7 10. Accessed 11 3 2022.

[Ultralytics, 2022] Ultralytics (2022). Yolov5 documentation. Accessed 31 05 2022.

[Uçar et al., 2017] Uçar, A., Demir, Y., and Güzeliş, C. (2017). Object recognition and detection with deep learning for autonomous driving applications. *Simulation*, 93(9):759–769.

[Verma et al., 2021] Verma, S., Dwivedi, G., and Verma, P. (2021). *Life cycle assessment of electric vehicles in comparison to combustion engine vehicles: A review*. Proceedings, Materials Today.

[Vijayenthiran, 2018] Vijayenthiran, V. (2018). Navya already sells fully self-driving cars, including in us. *Motor Authority*, 17 9. Accessed 11 3 2022.

[Vivacqua et al., 2017] Vivacqua, R., Bertozzi, M., Cerri, P., Martins, F., and Vassallo, R. (2017). Self-localization based on visual lane marking maps: An accurate low-cost approach for autonomous driving. *IEEE Transactions on Intelligent Transportation Systems*, 19(2):582–597.

[Vos et al., 2021] Vos, J., Cheng, L., Kamruzzaman, M., and Witlox, F. (2021). *The indirect effect of the built environment on travel mode choice: A focus on recent movers*.

[Wang and He, 2021] Wang, D. and He, D. (2021). Channel pruned yolo v5s-based deep learning approach for rapid and accurate apple fruitlet detection before fruit thinning. *Biosystems Engineering*, 210:271–281.

[Wang et al., 2022] Wang, N., Wang, Y., and Er, M. (2022). Review on deep learning techniques for marine object recognition: Architectures and algorithms. *Control Engineering Practice*, 118:104458.

[Wang et al., 2021] Wang, X., Huang, T., Liu, B., Yu, F., Wang, X., Gonzalez, J., and Darrell, T. (2021). Robust object detection via instance-level temporal cycle confusion.

[William and Pande, 2021] William, R. and Pande, A. (2021). *Gaps and Opportunities in Accessibility Policy*. San Jose State University, California.

[Wiseman, 2022] Wiseman, Y. (2022). Autonomous vehicles. In *Research Anthology on Cross-Disciplinary Designs and Applications of Automation*, page 878–889.

[Woodhill, 2011] Woodhill, L. (2011). Electric cars are an extraordinarily bad idea. *Forbes*, 14. Accessed 05 01 2022.

[Yair, 2022] Yair, W. (2022). Autonomous vehicles. In *Research Anthology on Cross-Disciplinary Designs and Applications of Automation*, page 878–889.

[Yan et al., 2017] Yan, L., Li, Z., Liu, H., Tan, J., Zhao, S., and Chen, C. (2017). Detection and classification of pole-like road objects from mobile lidar data in motorway environment. *Optics & Laser Technology*, 97:272–283.

[Yuan et al., 2018] Yuan, L., Li, D., and Hu, S. (2018). A map-matching algorithm with low-frequency floating car data based on matching path. *EURASIP Journal on Wireless Communications and Networking*, 2018(1):1–14.

[Zhang and Chi, 2020] Zhang, Y. and Chi, M. (2020). Mask-r-fcn: A deep fusion network for semantic segmentation. *IEEE Access*, 8:155753–155765.

[Zhao et al., 2019] Zhao, Z., Zheng, P., Xu, S., and Wu, X. (2019). Object detection with deep learning: A review. *IEEE transactions on neural networks and learning systems*, 30(11):3212–3232.

[Zhou et al., 2022] Zhou, H., Chen, H., Zhang, Y., Wei, M., Xie, H., Wang, J., Lu, T., Qin, J., and Zhang, X. (2022). *Refine-Net: Normal Refinement Neural Network for Noisy Point Clouds*.

[Zhu et al., 2017] Zhu, H., Yuen, K., Mihaylova, L., and Leung, H. (2017). Overview of environment perception for intelligent vehicles. *IEEE Transactions on Intelligent Transportation Systems*, 18(10):2584–2601.