Driverless Detection: Speed vs. Accuracy

Matthew Kyea

Binghamton University

Abstract

Currently, two types of object-detection systems are in use: ones that are accurate but slow, and ones that are less accurate but run in real-time. Through the evidence of various sources, this research paper proves that real-time systems, including the YOLO system, are the only practical choice for object detection in driverless vehicles due to their speed. Only real-time systems will allow driverless cars to reach the speeds that they must move at to positively affect society. Systems like YOLO have fast detection due to their use of a single neural network which quickly examines all parts of a photo or live feed once, while slower systems such as Faster R-CNN strive for increased accuracy by utilizing multiple networks to closely examine all parts of a picture. Studies by computer vision researchers show the difference in accuracy rates between these two types of systems is currently minor, and this gap is shrinking as real-time systems grow more precise and preserve their speed. In addition, by incorporating the philosophies of John Stuart Mill and John Rawls, this paper examines ethical reasons supporting a real-time system. This paper serves as a moral and technical guide to professionals who may work on driverless systems by describing why autonomous technologies must use real-time systems. Without speed in autonomous driving, these technologies will remain unsafe for a society with so much to gain from them.

Driverless Detection: Speed vs. Accuracy

Introduction

To successfully implement driverless cars, the underlying technologies which keep them moving safely must perform better than humans in critical situations. One of these technologies is the object recognition system, which a driverless car uses to identify its surroundings. The You Only Look Once (YOLO) object recognition system's use of a single neural network to examine a photo, video, or live feed allows for faster identification than systems like faster region convolutional neural networks (Faster R-CNN) which use multiple neural networks to examine portions of an image individually, proving the ability of YOLO and other real-time systems to reduce accidents in driverless vehicles. Researchers around the world are currently showing that YOLO is one of the most precise systems with the ability to classify objects from a video feed or in real-time, demonstrating YOLO's effectiveness as the object recognition system for time-sensitive applications. However, there are systems that are more accurate than YOLO: other tools, including Faster R-CNN, have higher accuracy rates but cannot run in real-time. Between these systems and YOLO there exists a trade-off between accuracy and speed, and driverless cars must use YOLO because driving without the ability to detect in real-time is a fruitless effort in which a vehicle remains one step behind its surroundings.

Background

The YOLO System

The YOLO object recognition system is a recently developed real-time detection method. Object recognition systems have the goal of detecting objects from pictures or videos and correctly classifying them like a human can, allowing computers to respond to changes in an environment or detect patterns and trends in images. In their 2016 conference paper, Redmon, Divvala, Girshick, and Farhadi (2016) introduce the YOLO system, which differs from other methods by using only one neural network which looks at the entire image at once. The authors describe a neural network as a system made of multiple layers, each with its own task. Some layers pull information from the image, then send the data to another layer to find the bounds of an object, until the network has found, labeled, and bound all objects in the image. With YOLO, the authors began by training their network with images and classifications for objects in these pictures. From these training cases, the network can look at a photo and compare the new image to earlier pictures, trying to detect objects the network knows. Finally, the object finally becomes

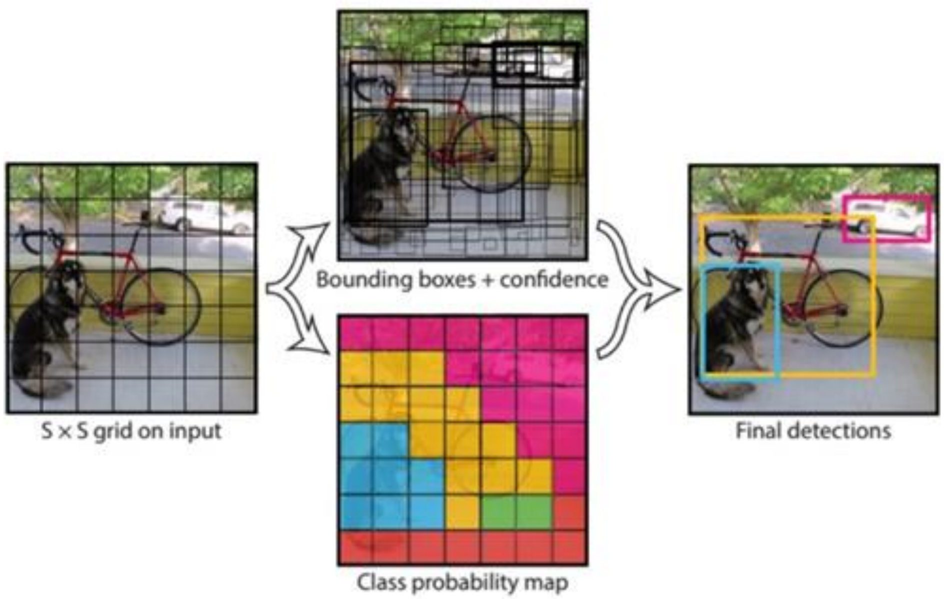

Figure 1. Model of the YOLO system.

bounded and labeled once YOLO recognizes a section of the photo as a known item from probabilities for sections of the image and bounding boxes for suspected objects (Redmon, Divvala, Girshick, & Farhadi, 2016, p.1-2). Figure 1 shows a high-level representation of this process, going from bounding boxes and probabilities to actual detections by the system (Towards Data Science, 2017, para.10). This revolutionary approach of a single network which looks once and handles all steps in classification has improved detection speeds greatly, allowing YOLO to run in real-time on common hardware such as smartphones (Redmon, 2017, 6:55). The system is also “general purpose” (Redmon, 2017, 5:50), meaning any data set with enough examples to learn from can train YOLO to detect objects important to various applications, including driverless vehicles.

Mean Average Precision

Many object detection systems are in use today, and researchers must have a way of comparing the accuracy rates between different implementations. The most widely used method of comparison is to train both systems with the same data sets, then test both implementations with the same sample images, and finally compute a mean average precision (mAP) representative of how well each object detection method recognized items in a group of photos. As explained in detail by Oksuz, Cam, Akbas, and Kalkan (2018), to compute the mAP of an object detection system for a group of test cases, a computer program must first calculate the average precision (AP) for each test image in the set. First, the tested system tries to detect objects in an image and produces a score for each item it has found. The computer compares each of these scores to a predetermined threshold number, and if the score is higher than this number, the program labels the associated bounding box as a prediction. Then, the computer compares each prediction to the set of ground truth boxes - a collection, provided by the tested data set, of boxes surrounding each object in the image. If a prediction box and its associated ground truth box for the same object intersect at a high enough rate, the prediction is a true positive, and a false positive otherwise. Next, the computer determines recall, the percentage of objects detected, and precision, the percentage of correct predictions, at the current threshold number. Finally, the computer changes the threshold number, and repeats all of the steps to create a recall-precision curve for the tested system and image. The AP is the area under this curve, and the mAP of a system on a particular data set of images is the mean of all average precisions computed for each photo (Oksuz, Cam, Akbas, & Kalkan, 2018, p.539). With mAP, researchers have a common ground in comparing different object detection systems, making it possible to tell when one implementation is more accurate then another.

Driverless Vehicles

Driverless vehicles can change transportation worldwide by removing human error from driving, but innovative solutions need to first overcome technological obstacles which prevent these cars from functioning. As described by David Silver (2017), a driverless car has five collaborating systems which make autonomous motion possible: computer vision, sensor fusion, localization, path planning, and control (Silver, 2017, 1:33). All of these groups have their own challenges and various implementations, with computer vision helping a car detect pedestrians, vehicles, and street signs by classifying them at a rate acceptable enough to allow for safe driving. A computer vision implementation for driverless cars must emphasize fast detection by using systems like YOLO; if detection is too slow, driving will be dangerous for a car which always stays a step behind, trying to detect objects it has already passed.

Precedents and Related Work

YOLO is not the only type of object recognition system in use today; earlier technologies, including Fast R-CNN, use a slightly different approach and still have relevant applications. These two object detection systems are similar at their core: a trained neural network looks for objects in new images based on training sets. However, Redmon, Divvala, Girshick, and Farhadi (2016) describe how Fast R-CNN systems take more steps than YOLO due to a region-based approach. The authors describe how Fast R-CNN splits an image into small sections, which a neural network examines closely rather than looking at the entire picture. Then, the system gives each region a prediction rating based on the training cases and adjusts the bounds on the region to find where an object begins or ends. Finally, the system compares detections to ensure the network detects each object only once (Redmon, Divvala, Girshick, & Farhadi, 2016, p.5). In summary, multiple independent components and networks compose the different steps Fast R-CNN takes to classify an object, requiring these subsystems to run hundreds of times on the many regions of a photo; the effects of using such methods in driverless cars would be disastrous. Although many studies have shown Fast R-CNN systems to be consistently more precise than YOLO (Park, Yoon, & Park, 2019, p.22) Fast R-CNN has a fatal flaw: the system cannot run in real-time. High detection rates and speed have an inverse relationship, as Fast R-CNN thoroughly examines every section of an image while YOLO takes a quick glance. With a car's object detection system always lagging behind, it becomes possible for the car to run a red light without noticing. Using such systems may prove useful in time-insensitive applications due to their high accuracy rate but are impractical for autonomous vehicles. Even though YOLO has a lower accuracy rate, YOLO makes object detection possible for driverless cars – if a car does not have the chance to react in real-time, it cannot drive.

Support

Technical Details

The dangers of driverless cars. On March 12, 2018, an Uber autonomous vehicle struck and killed Elaine Hertzberg as she crossed the street. This incident marked the first case of a pedestrian killed by a self-driving car and the beginning of valid concerns about the safety of autonomous vehicles (Zhou & Sun, 2019, p.61-67). In their article in *Communications of the ACM*, Zhi Quan Zhou and Liqun Sun (2019) examine the cause of this tragic accident and search for safer solutions. The authors have knowledge on the topic, as Zhou holds a bachelor's degree in computer science from Peking University as well as a doctorate in software engineering from the University of Hong Kong while Sun holds a master's degree in plasma physics from Donghua University. Currently, Zhou and Sun conduct research at the School of Computing and Information Technology in the University of Wollongong, Australia. In the authors' investigation of the incident, they believe a miscommunication between sensors and onboard computer systems caused Elaine's death, and that similar miscommunications could lead to future dangers involving driverless vehicles. As explained by Zhou and Sun, current technologies cannot manually check the millions of data points sent by sensors to the car's computer system, causing humans to miss unexpected behavior and uncommon bugs. The authors advocate for a new system, called metamorphic testing, which compares the inputs and outputs of a system over thousands of executions to find discrepancies from bugs. This comparison establishes a "metamorphic relation" (Zhou & Sun, 2019, p.63), which the system compares against future tests to ensure this association stays intact. If a bug results in outliers to this relationship, the system would know about the possibility of an error. In future scenarios, combining fast detection speeds and systems such as metamorphic testing will prove necessary to making driverless cars safe (Zhou & Sun, 2019, p.61-67). To allow metamorphic testing, driverless vehicles must use fast object detection to provide as much information about the environment as possible.

YOLO: a fast detection system. To achieve fast detection speeds and increase the amount of data provided to a vehicle via cameras and sensors, computer vision implementations in driverless cars must emphasize real-time detection methods such as YOLO, that recognize objects fractions of a second after a camera captures the surroundings. Redmon (2017) introduces YOLO in a TED talk by first comparing this system to earlier object detection methods. When Redmon started working with such systems, they took twenty seconds to classify an object from an image. According to Redmon, over the course of a few years detection speeds increased from twenty seconds per picture to two seconds per picture, and then systems gained the ability to analyze video at five frames per second. Despite these exponential advances, five frames per second is too slow for practical applications, and Redmon says he “would not want a system like this driving my car.” Redmon is currently pursuing his PhD in Computer Science from the University of Washington and has his master’s degree in computer science from Middlebury College. Despite his credibility, Redmon has a bias as one of the system’s creators (Redmon, 2017, 2:20-3:31). As a result, it is necessary to examine the results of independent studies that have tested YOLO.

Studies concerning YOLO. Computer vision researchers around the world are testing the results of object detection systems in various applications to discover the best methods for completing specific tasks. Yang, Zhang, Bo, Wang, & Chen (2019) have recently studied several of the object detection systems in use today and compared the strengths and weaknesses of each. The authors research object detection at the Minzu University of China, with “Fast vehicle logo detection in complex scenes” being the first contribution of Yang, Wang, and Chen to the field. The authors use the most recent version of YOLO, YOLOv3, and have trained their system on a data set of vehicle logos to find an acceptable solution to object detection of small objects. The authors have also modified YOLOv3’s last neural network layer to make small object detection easier. SSD512 and YOLOv2 beat the authors’ system in a few cases, but the authors state that their system is usually the most exact and the fastest. Table 1 shows that in these infrequent defeats, the average precision of the system which beat the authors' system was usually .02 points higher, and the overall mAP of the authors' system is .899 while Faster R-CNN's mAP is

Table 1
Detection results of different algorithms.

Number	**Name**	**Faster-RCNN**	**SSD-512**	**YOLOv2**	**Our Method**
0001	Beijing Automotive	0.852	0.877	0.863	**0.881**
0002	Ford	**0.774**	0.745	0.733	0.767
0003	SKODA	0.871	**0.893**	0.854	0.860
0004	Venucia	0.817	0.792	0.824	**0.835**
0005	HONDA	0.982	0.965	0.922	**0.988**
0006	NISSAN	0.969	0.974	0.952	**0.978**
0007	Cadillac	0.765	**0.845**	0.742	0.824
0008	SUZUKI	0.874	0.892	0.883	**0.914**
0009	GEELY	0.824	0.863	0.854	**0.870**
0010	Porsche	0.854	**0.863**	0.812	0.855
0011	JEEP	0.642	**0.751**	0.602	0.745
0012	BAOJUN	0.977	0.982	0.963	**0.984**
0013	ROEWE	0.865	0.915	0.878	**0.933**
0014	LINCOLN	0.834	**0.866**	0.786	0.842
0015	TOYOTA	**0.997**	0.983	0.988	0.992
0016	BUICK	0.873	**0.902**	0.857	0.884
0017	CHERY	0.923	0.947	0.904	**0.953**
0018	KIA	**0.872**	0.864	0.822	0.858
0019	HAVAL	0.781	**0.796**	0.653	0.776
0020	Audi	0.977	0.986	0.982	**0.989**
0021	LAND ROVER	0.879	**0.893**	0.859	0.882
0022	Volkswagen	0.917	0.931	0.908	**0.943**
0023	Trumpchi	0.924	0.944	0.915	**0.961**
0024	CHANGAN	0.803	**0.865**	0.786	0.847

Number	Name	Faster-RCNN	SSD-512	YOLOv2	Our Method
0025	Morris Garages	0.832	0.854	0.820	**0.876**
0026	Renault S.A	0.920	0.903	0.896	**0.963**
0027	LEXUS	0.867	0.902	0.869	**0.921**
0028	BMW	0.934	0.928	0.912	**0.987**
0029	MAZDA	0.899	0.911	0.883	**0.924**
0030	Mercedes-Benz	**0.953**	0.927	0.879	0.933
	MAP	0.875	0.892	0.853	0.899
	Time(s)	1.5	0.05	0.02	0.03

.875 (Yang, Zhang, Bo, Wang, & Chen, 2019, p.4). These occasional defeats by SSD512 and YOLOv2 were due to complex patterns in images and the authors' method of using the last network layer to reach a prediction, two issues the authors plan to fix (Yang, Zhang, Bo, Wang, & Chen, 2019, p.1-5). This study shows the strengths of YOLO in a time-sensitive scenario related to driverless vehicles, but the system may have weaknesses for dissimilar situations.

Nguyen, Nguyen, Le, Duong, & Nguyen (2019) have also compared some detection systems, but for a different application where fast detection is unnecessary. All of the authors conduct research at the Ho Chi Minh City University of Information Technology and Le, Duong, and Nguyen each have a PhD in computer science. The authors set out to create a new detection system which solves a problem that others cannot: recognition of objects that all systems miss or have difficulty detecting. The new system, called You Always Look Again (YALA), arose from a modification of Faster R-CNN which the authors trained and altered specifically for their application. In a comparison of the accuracy rates of various systems, including YOLO, SSD, Fast R-CNN, and YALA, the authors find that their system is the most precise system for detecting these missed objects. Table 2 shows that YALA has a mAP of 79.0, while YOLO and

SSD512 have a mAP of 66.4 and 76.8, respectively (Nguyen, Nguyen, Le, Duong, & Nguyen, 2019, p.10). However, YALA lacks the ability to run in real-time. Out of the systems with the

Table 2

Runtime comparison of different methods on VOC 2007 test.

Method	**Fast RCNN**	**Faster RCNN**	**YOLO**	**SSD300**	**SSD512**	**YALA-SS**	**YALA-MS**
mAP	70.0	73.2	66.4	74.3	76.8	77.5	79.0
Testing time (fps)	0.5	7	21	46	19	4	1

ability to run in real-time, SSD512 beat its predecessor SSD300 by two points in accuracy, but SSD300 ran faster than YOLO and SSD512 (Nguyen, Nguyen, Le, Duong, & Nguyen, 2019, p.1-10). The authors of this study find that YALA is the best choice for their application; however, the accuracy tests used an older version of YOLO (Redmon & Farhadi, 2018, p.1).

As described by Redmon and Farhadi (2018), YOLO has become more accurate over time while retaining its speed due to a few small improvements suggested by fellow researchers in the object detection field. Firstly, the developers of YOLO have now allowed the system to classify an object by multiple names. Now if the system detects a dog, it can classify it as both a dog and the specific breed, assuming the system has enough information to do so. This is called a multilabel classification and using this method of detection increases accuracy when the system is trained with advanced datasets containing several classifications for the same object. In addition, the developers also changed YOLO's neural network. The authors have redesigned their network by adding more layers in a different order and "shortcut connections" (Redmon and Farhadi, 2018, p.3), which the network can use to skip to certain layers from a related layer in specific circumstances.YOLOv3 has fifty-three layers, which is significantly less than some systems mentioned like ResNet-152, but is still able to run in real time and achieve accuracy rates close to these related systems that take twice as long to detect objects (Redmon & Farhadi, 2018, p.1-3). YOLOv3 has small improvements that increase accuracy while retaining speed, and studies must examine the most recent version of such systems to choose the one best for their application.

One final study by Park, Yoon, & Park (2019) also examines a time-insensitive scenario where speed is optional while testing the most recent version of YOLO. In their study, the authors try to create the best object detection method for tracking a driver's eyes and telling when the driver seems too tired or distracted to drive. In their search for the best method of detection, the authors find that their system, which uses Faster R-CNN, has a higher precision than other systems like YOLO. Despite YOLO's ability to run in real-time, the authors conclude by stating that their system is the best for this application. In this study and the work of Nguyen et al. (2019), greater accuracy outweighs the ability to run in real-time - these systems run in a time-insensitive application, so the authors can use a slower system to improve accuracy with no ill effects (Park, Yoon, & Park, 2019, p.1-26; Nguyen, Nguyen, Le, Duong, & Nguyen, 2019, p.1-10). This is the opposite of the first study described, where a system requires fast detection due to the speed at which the environment is seen by an autonomous vehicle. All of these unique studies by diverse groups of researchers describe a trade-off between accuracy and speed in current object detection methods that needs examination to determine the type of system that will allow society to benefit the most from safe driverless cars.

Social Impact

Driverless detection must be fast. The development of YOLO allows for a shift in how societies understand transportation: driverless technology becomes possible, safe, and ethical with a real-time system. However, for people to trust driverless cars and allow them to drive, these cars must drive better than humans and create a safer environment with fewer accidents. When humans drive, the driver's eyes feed information to the brain as events happen without a delay needed to classify the environment. If manufacturers of a time-sensitive application are considering using a specific computer vision method, they must ensure the detection system can recognize objects in real-time like humans. Systems like Faster R-CNN lack this ability, and although they have their uses in situations which value accuracy over speed, they have no place in an autonomous vehicle's detection system (Park et al., 2019, p.26).

Tragic cases, including the death of Elaine Hertzberg, also show why these systems need fast detection. Although Elaine's death resulted from a miscommunication in the car's systems rather than slow detection, a car will need all the data possible from its environment to respond accurately to situations such as a pedestrian appearing from the dark (Zhou & Sun, 2019, p.61-67). With a real-time system, the car has a chance to stop, assuming the communication between sensors and controller has no bugs like in Hertzberg's case. Without a real-time system, a car can only respond to a situation after the moment has already passed, braking for a pedestrian in the street after hitting them.

A final reason why these systems need fast detection relates to the speed society expects such vehicles to drive at. In theory, driverless cars should increase convenience and decrease travel time by driving faster and safer than humans. With a slow object recognition system, autonomous driving at speeds humans safely drive at remains impossible. Autonomous vehicles that drive at a speed slow enough for more exact detection systems to classify everything in the environment have little use for people trying to save time. Thus, real-time systems describe the safest and only practical choice, assuming a high enough accuracy.

Technical explanations. As shown by Nguyen et al. (2019, p.10), Yang et al. (2019, p.5), and Park et al. (2019, p.26), real-time systems are often less accurate than ones without the ability to detect in real-time. However, car manufacturers and developers should consider two technical reasons why these lower accuracy systems still have valid applications.

Firstly, the difference in accuracy rates between fast and slow systems is small. Although the slow systems have higher precision, newer real-time detection systems can eventually reach these higher accuracy rates, and these faster systems currently have mAP scores close to those of slower methods. The study by Nguyen et al. (2019, p.10) shows this closeness in mAP, as SSD512 scored only three points lower than the proposed system that lacks real time detection, YALA. Furthermore, in the study by Yang et al. (2019, p.5), when the author's modified YOLOv3 system occasionally lost to Faster R-CNN, the average precision difference was often less than .02. In addition, the author's system had an average mAP of .024 points higher than the mAP of Faster-R-CNN. Thus, although real-time systems usually lack some accuracy, these studies have shown the differences between them to be too small to cause serious concern.

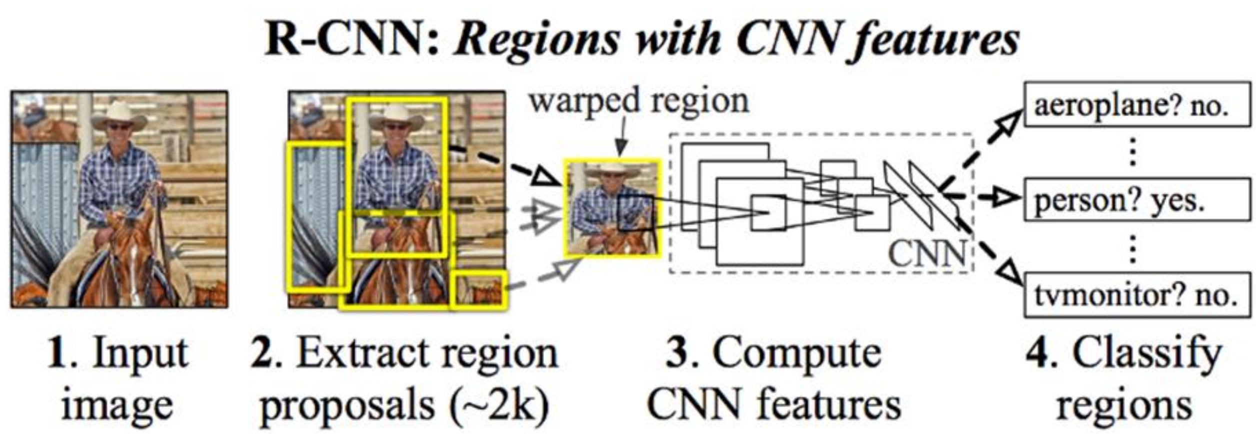

Figure 2. Model of the R-CNN system.

Secondly, the accuracy gap between these systems is shrinking due a difference in how each system works. Figure 2 illustrates this difference: in slower systems such as Fast R-CNN, multiple neural networks closely examine all parts of a photo one portion at a time, while in real-time systems a single neural network examines the entire image (Redmon et al., 2016, p.5; Athelas, 2017, para.14). This difference is important: Fast R-CNN has worked to improve in accuracy over time but cannot improve in speed without changing how the underlying network works. On the other hand, YOLO has increased in accuracy over three versions, shown by the improvements between version one in the study by Nguyen et al. (2019) to version two and three in the study by Yang et al. (2019), as well as a description of improvements provided by Redmon and Farhadi (2018) (Nguyen et al., 2019, p.10; Yang et al., 2019, p.5; Redmon & Farhadi, 2018, p.3). Therefore, YOLO has become more precise over time due to further improvements to the system while retaining its speed, but Fast R-CNN systems cannot become any faster without switching to an approach similar to YOLO's. Real-time systems have the potential to become both fast and exact, while the path of slow systems only allows for more accurate detection. One day, systems with perfect accuracy may exist using a Fast R-CNN type architecture and a YOLO type approach; however, only the YOLO system will have the added benefit of running in real-time. Although real-time systems tend to have less accuracy, the gap in precision is small and shrinking due to how the underlying network of each method works.

Ethical explanations. In addition to technical explanations in favor of YOLO, the philosophies of John Stuart Mill and John Rawls offer ethical explanations as to why driverless cars should use fast object detection. Mill's utilitarianism describes how the morality of an action lies in the eventual consequences for humanity, and that the results of a decision should increase total human happiness. With this philosophy, the importance of driverless vehicles helping humanity becomes clear. The amount of human happiness from driverless vehicles can only increase from cars that drive faster than humans. If a driverless vehicle uses a real-time system with less accuracy, cars will be able to drive faster due to their ability to detect at faster speeds. However, a lack of accuracy means a system could incorrectly classify a pedestrian and kill them. So, if driverless cars instead opted for a slower detector with higher accuracy, these systems could prevent deaths but would cause vehicles to move slowly. Thus, Mill would choose a real-time but less accurate system. Total human happiness will increase the most from cars that make driving easier and quicker, by driving at faster speeds. There may be individuals killed by driverless cars due to lessened accuracy, but these people would be in the minority, as autonomous systems would be worthless and illegal if they injured more people than they helped. Utilitarianism only focuses on the majority of individuals with increased happiness from real-time detection and disregards the minority that may suffer from lessened accuracy.

Rawls, using his veil of ignorance, would look at this least advantaged minority passed over by Mill: those who suffer from the lack of accuracy with a real-time detection system. Using the veil of ignorance encourages one to imagine if they became the person who suffers the most from a choice and decide on the morality of said choice from outside their own perspective. In extreme cases, people could die from this system’s lack of precision, incorrectly classifying a pedestrian as part of the road. This stance employs a convincing argument against such systems, but the same argument also applies to systems without the ability to run in real-time. Using the veil of ignorance results in the same effect: people will die from a system’s lack of speed, as the vehicle cannot identify a person until the car is too close to slow down. The ideal system should have both real-time detection and perfect accuracy, something impossible with current technology. So, Rawls would be against driverless cars in general, due to the death they may cause. However, a future with driverless cars means increased access to all: people with disabilities could travel in cars without hiring a driver. Due to this increase in access, detection systems may one day become ethical under Rawls's philosophy, and these systems can only become more precise via a real-time implementation. Mill's utilitarianism and Rawls's view on the least advantaged support the argument in favor of real-time detection.

YOLO vs SSD: focusing on real-time methods rather than specific systems. Proven through this research of object detection systems, driverless vehicles must use real-time detection. Ruling out slower systems leaves only a few choices: systems like YOLO or SSD which can run in real-time. The study by Nguyen et al. (2019) shows that YOLO is slower and less precise than SSD. However, the authors compared SSD to the original version of YOLO, which lacks improvements - such as multilabel classifications and more layers - present in the most recent third version (Nguyen et al., 2019, p.10; Redmon & Farhadi, 2018, p.3). Another study by Yang et al. (2019) uses the most recent version of YOLO and shows how it was the best system for the recognition of small objects, specifically vehicle logos, despite some troubles in creating the system (Yang et al., 2019, p.5). Although SSD may work just as well in driverless systems, this study has shown explicit evidence supporting the possibility of modifying YOLOv3 to work in a specific application related to driverless vehicles while outperforming SSD512. Both systems should be able to successfully drive an autonomous vehicle due to their real-time detection, but this study supplies a clear argument in favor of YOLO which is simple to apply to driverless vehicle research. By focusing on YOLO, researchers can examine a general trade-off between accuracy and speed in two diverse types of systems rather than dissect the smaller differences between various real-time systems.

Driverless cars for society. Assuming the reasoning behind the necessity of real-time detection is well-understood by car manufacturers and developers, autonomous cars will change transportation. Accidents previously caused by intoxication, sleep deprivation, distractions, and road rage cannot occur while a computer drives. In addition, if these vehicles were to eventually obtain perfect accuracy rates and can communicate with each other, the number of deaths from car accidents should drop close to zero. With perfect systems, few chances will exist for an autonomous vehicle to put someone in danger. In addition to reducing accidents, driverless cars will make human lives less stressful. Without needing to drive, individuals will have more free time and spend less time worrying about collisions driving home in traffic. However, this innovative technology has downsides. There exists the possibility of hacking driverless vehicles, and job loss for truckers, taxi drivers, and others working in the field of transportation. As with any groundbreaking technology, driverless cars will cause benefits and problems for society, and humanity must work to maximize the positives by counteracting the negatives to increase safety and ensure the technology is worthwhile.

Conclusion

Real-time object detection systems like YOLO must be the choice for computer vision implementations in driverless vehicles. Several studies have found systems similar to Faster R-CNN to have higher accuracy rates, but the differences in these rates are small enough to conclude that fast detection systems are still safe even with less precision. In addition, systems related to YOLO are improving in detection rates and speed, while the methods Faster R-CNN systems use only allow for an increase in accuracy. Thirdly and most importantly, driverless technology cannot have a useful or safe effect on society if a system takes multiple seconds to classify objects while trying to drive at useful speeds. YOLO's use of a single neural network rather than Faster R-CNN’s use of several networks allows for such an increase in detection speeds to the point that driverless technology will have less accidents and truly be possible. Automobile developers working on driverless technology should examine this research to understand why real-time detection is crucial for building both a useful and safe driverless vehicle. From this research, professionals should learn that real-time object detection is necessary and further examine the other computer vision technologies in use by driverless vehicles to look for solutions to any similar trade-offs in choosing an implementation that is accurate or one that is fast. In addition, these experts should also find and test the best real-time implementation for their application, as this research has found that YOLO and SSD can both safely drive an autonomous car. In conclusion, the society which has so much to gain from driverless technology must pay attention to how manufacturers implement these systems. Real-time systems are the only safe choice for object-detection in driverless cars, and the public must demand that professionals only use such systems in this revolutionary technology if autonomous vehicles are to help society rather than harm it.

References

Athelas. (2017). *Model of the R-CNN system*. Retrieved from https://cdn-images-1.medium.com/max/1600/0*Sdj6sKDRQyZpO6oH.

Nguyen, K., Nguyen, K., Le, D., Duong, D. A., & Nguyen, T. V. (2019). You always look again: Learning to detect the unseen objects. *Journal of Visual Communication and Image Representation, 60*, 206-216. doi:10.1016/j.jvcir.2019.02.020

Oksuz, K., Cam, B. C., Akbas, E., & Kalkan, S. (2018). Localization recall precision (LRP): a new performance metric for object detection. *Computer Vision – ECCV 2018 Lecture Notes in Computer Science*, 521-537. doi:10.1007/978-3-030-01234-2_31

Park, S. H., Yoon, H. S., & Park, K. R. (2019). Faster R-CNN and geometric transformation-based detection of driver's eyes using multiple near-infrared camera sensors. *Sensors, 19*(1), 197. doi:10.3390/s19010197

Redmon, J, & Farhadi, A. (2018, April). YOLOv3: an incremental improvement. *Computing Resource Repository*. Retrieved from https://pjreddie.com/media/files/papers/YOLOv3.pdf

Redmon, J. (2017, April). *How computers learn to recognize objects instantly* [Video file]. TED. Retrieved from https://www.ted.com/talks/joseph_redmon_how_a_computer_learns_to_recognize_objects_instantly

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016, June). *You only look once: unified, real-time object detection.* Paper presented at the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, Nevada, USA. doi:10.1109/cvpr.2016.91

Silver, D. (2017, November). *How self-driving cars work* [Video file]. TED. Retrieved from https://www.youtube.com/watch?v=Ly92UcnoEMY&t=2s

Towards Data Science. (2017). *Model of the YOLO system.* Retrieved from https://cdn-images-1.medium.com/max/1200/1*m8p5lhWdFDdapEFa2zUtIA.jpeg

Yang, S., Zhang, J., Bo, C., Wang, M., & Chen, L. (2019). Fast vehicle logo detection in complex scenes. *Optics & Laser Technology,110,* 196-201. doi:10.1016/j.optlastec.2018.08.007

Zhou, Z. Q., & Sun, L. (2019, March). Metamorphic testing of driverless cars. *Communications of the ACM*, *62*(3), 61-67. doi:10.1145/3241979