PRESIDENT UNIVERSITY

Where tomorrow's leaders come together

Name : Sarah Ruen Soekarno (MM Tech 2019)

ID Number : 023201905035

Subject : Digital Literacy (Final Term Project)

# EXPLORING THE POTENTIAL OF MACHINE LEARNING AND COMPUTER VISION IN THE RETAIL INDUSTRY

## ABSTRACT

*Initially, this Artificial Intelligence technology was introduced to simplify work or solve problems more quickly. But along with its increased ability, some companies see its potential even more than that. The concept of machine learning is the development of systems that can learn "alone" without the need to be programmed by humans repeatedly. Computer vision itself is a process of how a computer or machine can see. Merging these two concepts is used in various industries, including players in retail industry. One of the examples that using these two concepts at the same time is Amazon Go. Amazon Go is a prototype grocery store that is operated under the auspices of the e-commerce company Amazon.com. Amazon applies a technology called "just walk out technology" which use Computer Vision based Machine Learning for tracking and estimating the intention of everyone in the store.*

***Keywords: Machine Learning, Computer Vision, Retail***

# INTRODUCTION

Initially, this Artificial Intelligence technology was introduced to simplify work or solve problems more quickly. But along with its increased ability, some companies see its potential even more than that. AI continues to evolve to benefit many different industries such as the advent of machine learning and computer vision. Since computer vision systems are difficult to be fully programmed, because of both the complexity of real-world modeling and the large amount of required knowledge, we can greatly benefit from the application of ML techniques. Machine learning is currently the only known way to develop computer vision systems that are robust and easily reusable in different environments. Although the two fields have a long-standing tradition and can be considered technologically mature, past research in applying ML techniques to CV problems has been limited.

Merging these two concepts is used in various industries, including players in retail industry. According to Weinswig in a panel discussion at the National Retail Federation (NRF) who is the CEO and founder of US retail and technology consultancy Coresight Research, even the retail industry will be the most popular among other industries in terms of being promoted, even higher than adoption in banking , manufacturing, and health care have largely utilized more advanced technology in recent years. The challenge of the retail industry continues to make activists have to work extra hard to maintain existence. Not that the retail industry is extinct. However, the development of the digital world and all that was given was fully approved by their shops. This challenge looks very heavy for large retail companies that are famous for closing their shops. Call it Lotus and 7-Eleven. Retail Businessmen are aware that good customer service is vital to determining business success. AI has the ability to revolutionize the online retail industry by making the shopping experience on the web easier. It also encourages players in offline retailing to take advantage of AI because of the rise of the emergence of online retailing that supports the continuity of their business is very important to improve competitiveness and increase customer convenience in purchasing their products so they are not inferior to offline promotions. In their research, Felix and Reinhard have also implemented a variety of artificial intelligence applications that can be applied to the retail industry.

One example of the application of artificial intelligence specifically for machine learning and computer vision that is applied to the offline retail business is Amazon Go. Amazon Go is a prototype grocery store that is operated under the auspices of the e-commerce company Amazon.com (Wingfield, 2016). Amazon itself offers players in the largest online retail field in the world that offers various types of products. But Amazon began to penetrate into the offline world offline by utilizing AI in its daily operations. In this research, the writer will find out about the use of special machines and computer vision in the retail world. It is hoped that this research can inspire players in online retailing to take advantage of AI entering its business.

# LITERATURE REVIEW

## 1. Artificial Intelligence

All efforts within the domain of information technologies, independently of an academically motivated and enforced separation of the research fields, have followed the assumption and goal of the transfer of task to be overtaken by machines in the last decades. Artificial intelligence was completely replace humans in the performance of tasks. Artificial intelligence tries to figure out how to get machines to use language, to form abstractions and concepts, to solve types of problems that are currently reserved for humans, and to improve themselves (Dahria, 2008).

In addition, it is pointed out here that the ideas of are pursued in order to evaluate the use of in retail: AI is the science that enables machines problem types and tasks that cannot yet be performed by computers and in which people are currently better. This means that the examples are not simply emulated, but patterns and laws from the data are recognized (James, 2018).

Artificial intelligence (AI) is an area of computer science that emphasizes the creation of intelligent machines that work and react like humans. Some of the activities computers with artificial intelligence are designed for include: Speech recognition, Learning, Planning, Problem solving. In this topic we shall discus the following subjects; Deep learning, Machine learning, Computer Programming, Medical field. Deep Learning has enabled many practical applications of Machine Learning and by extension the overall field of AI. Deep Learning breaks down tasks in ways that makes all kinds of machine assists seem possible, even likely. Driverless cars, better preventive healthcare, even better movie recommendations, are all here today or on the horizon. AI i s the present and the future. With Deep Learning’s help, AI may even get to that science fiction state we’ve so long imagined.

Machine Learning at its most basic isthe practice of using algorithms to parse data, learn from it, and then make a determination or prediction about something in the world. So rather than hand-coding software routines with a specific set of instructions to accomplish a particular task, the machine is “trained” using large amounts of data and algorithmsthat give it the ability to learn how to perform the task

## 2. Machine Learning

Machine learning is a set of techniques that supports complex data prediction by presenting the data itself with learning algorithm.

The term machine learning was defined for the first time by Arthur Samuel in 1959. According to Arthur Samuel, machine learning is a field of computer science that provides learning skills for computers to learn something without detail programming (O'Hagan, 2019).

According to (Mohri, Rostamizadeh, & Talwalkar, 2012) machine learning can be determined as a method of computing based on experience to improve performance or make accurate predictions. The definition of experience is information that has been previously available and can be used for learning data.

Machine Learning is a method used for creating program by the process of data learning. The difference with other computer programs is that Machine Learning has an ability to learn by itself while other computer programs are usually static (W. Yuciana, 2014) The Learning method of Machine Learning is following the the way of human learning, which is learning from examples. Machine learning will learn the pattern of the examples that has been analyzed in order to determine the answers to the next questions. In fact, all problems cannot be solved by machine learning programs. However, it happens very often that complex algorithms also can be solved easily by Machine Learning.

Some examples of machine learning programs that have been used in everyday life: Spam Detection, Face Detection, Product Recommendations, Virtual Assistant, Medical Diagnosis, Credit Card Fraud Detection, Digit Identification, Stock Trading and Customer Segmentation.

The process carried out in the machine learning workflow is to collect the dataset you want to test, then choose the method you want to use for example linear regression, logistic regression, neural network, svm and others, provide training on the selected method, evaluate the method chosen and who last did Predictions (Ahmad, 2017)

The initial accuracy of machine learning programs is usually very poor. However, over time, the more we practice the program, the more examples the program will learn, the better and more accurate the program will be (Stepan, 2020).

**Gathering Data**

The process of gathering data depends on the type of project we desire to make, if we want to make an ML project that uses real-time data, then we can build an IoT system that using different sensors data. The data set can be collected from various sources such as a file, database, sensor and many other such sources but the collected data cannot be used directly for performing the analysis process as there might be a lot of missing data, extremely large values, unorganized text data or noisy data. Therefore, to solve this problem Data Preparation is done.

**Data pre-processing**

Data pre-processing is a process of cleaning the raw data i.e. the data is collected in the real world and is converted to a clean data set. In other words, whenever the data is gathered from different sources it is collected in a raw format and this data isn't feasible for the analysis.

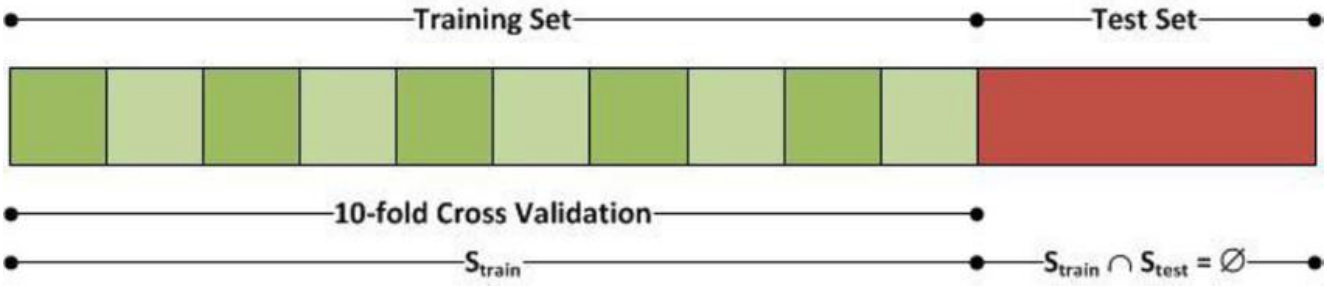
Therefore, certain steps are executed to convert the data into a small clean data set, this part of the process is called as data pre-processing.

These are some of the basic pre — processing techniques that can be used to convert raw data.

**Training and testing the model on data**

For training a model we initially split the model into 3 three sections which are 'Training data' ,'Validation data' and 'Testing data'.

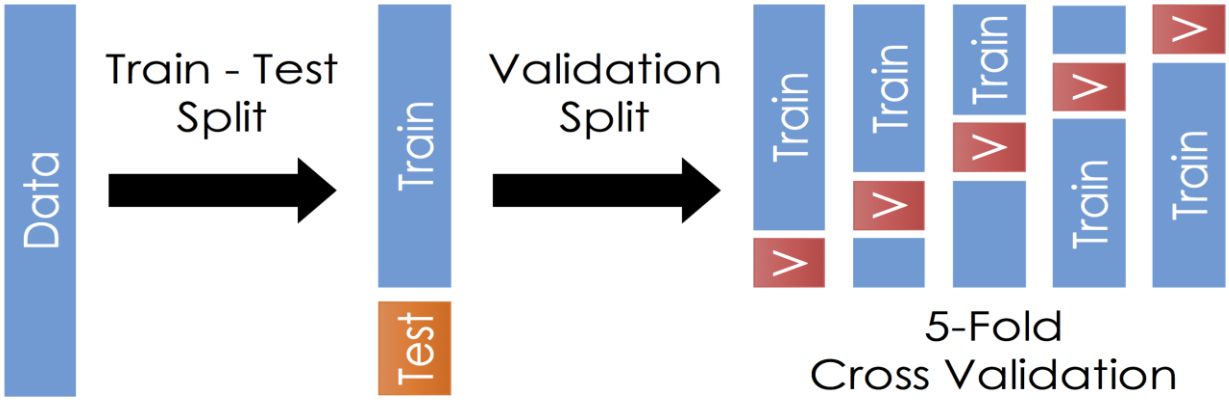
You train the classifier using 'training data set', tune the parameters using 'validation set' and then test the performance of your classifier on unseen 'test data set'. An important point to note is that during training the classifier only the training and/or validation set is available. The test data set must not be used during training the classifier. The test set will only be available during testing the classifier.




Training set: The training set is the material through which the computer learns how to process information. Machine learning uses algorithms to perform the training part. A set of data used for learning, that is to fit the parameters of the classifier.

Validation set: Cross-validation is primarily used in applied machine learning to estimate the skill of a machine learning model on unseen data. A set of unseen data is used from the training data to tune the parameters of a classifier.

Test set: A set of unseen data used only to assess the performance of a fully-specified classifier.




Once the data is divided into the 3 given segments we can start the training process.

In a data set, a training set is implemented to build up a model, while a test (or validation) set is to validate the model built. Data points in the training set are excluded from the test (validation) set. Usually, a data set is divided into a training set, a validation set (some people use 'test set' instead) in each iteration, or divided into a training set, a validation set and a test set in each iteration.

## 3. Computer Vision

Computer vision is a discipline that studies how to reconstruct, interpret, and understand a 3-dimensional view of its 2-dimensional display in terms of the nature of the display structure. Computer Vision is concerned with modeling and imitating human vision using software and hardware on a computer. Computer Vision combines science in the fields of computer science, electrical engineering, mathematics, physiology, biology, and cognitive science. Knowledge from all these fields is needed to understand and simulate the operation of human vision. The supporting functions are between other.

The process of capturing images or image acquisition in humans starts from the eye, then visual information is translated into a format that can then be manipulated by the brain. Like the process, Computer Vision needs an eye to capture a visual signal.

The camera will translate a scene or image. The camera observes an event in one lane at a time, scans it and divides it into hundreds of similar horizontal lines. Each line creates an analog signal whose amplitude explains the change in brightness along the signal line.

Computers do not work with analog signals, therefore an Analog to Digital Converter is needed, it is needed to process all these signals by the computer. This ADC will convert the analog signal represented in the form of a single signal information into a stream of a number of binary numbers. These binary numbers are then stored in memory and will become raw data to be processed.

**Digital Image Representation**

Digital image consists of a finite set of values called picture elements or pixels for short. These pixels are arranged in a regular grid (or raster) of rows and columns, and so it can be useful to think of an image as a matrix. Every pixel in a greyscale image (also called an intensity image) is an 8-bit unsigned integer, meaning that it can have an integer value between 0 and 255. A value of 0 corresponds to pitch black, a value of 255 to pure white, and values between these extremes produce various grey levels between black and white.

A color image is also stored as a raster of pixels. Every pixel is now represented by three integer values between 0 and 255: one for red, one for green and one for blue. These three primary intensities are added to reproduce a certain color on the screen, and this commonly used way of representing color is called the RGB color scheme see Fig. 1.0.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| R: 000<br>G: 000<br>B: 000 | R: 036<br>G: 036<br>B: 036 | R: 073<br>G: 073<br>B: 073 | R: 109<br>G: 109<br>B: 109 | R: 146<br>G: 146<br>B: 146 | R: 182<br>G: 182<br>B: 182 | R: 219<br>G: 219<br>B: 219 | R: 255<br>G: 255<br>B: 255 |
| R: 255<br>G: 000<br>B: 000 | R: 255<br>G: 192<br>B: 000 | R: 255<br>G: 255<br>B: 000 | R: 146<br>G: 208<br>B: 080 | R: 000<br>G: 176<br>B: 080 | R: 000<br>G: 176<br>B: 240 | R: 000<br>G: 112<br>B: 192 | R: 112<br>G: 048<br>B: 160 |

Fig. 1.0 Examples of RGB color values [0:255].

An image may be defined as a two-dimensional function, f(x, y), where (x) and (y) are spatial (plane) coordinates, and the amplitude of at any pair of coordinates (x, y) is called the Intensity or gray level of the image at that point. When (x), (y), and the amplitude values of (f) are all finite, discrete quantities, we call the image a digital image. The field of digital image processing refers to processing digital images by means of a digital computer. Note that a digital image is composed of a finite number of elements, each of which has a particular location and value. These elements are referred to as picture elements, image elements, and pixels. Pixel is the term most widely used to denote the elements of a digital image.

**Camera Model**

Here we develop a basic camera model. The model performs well as long as the lens is thin and no wide-angle lens is used. In practice the image plane is located behind the lens, but to simplify calculations and relations between the coordinate systems, the image plane can be put in front of the lens. Fig. 1.7 illustrates the camera model with the image plane located in front of the lens.

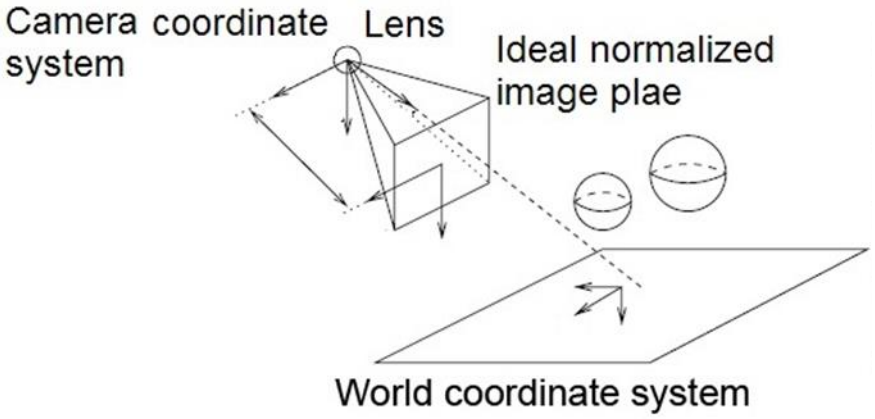



Fig. 1.2 Illustration of the camera model, image in front of the lens to simplify calculations [Courtesy of Maria Magnusson Seger].

The hardest part of this model is keeping track of the different coordinate systems. Before we continue, we have to define an image-coordinate system in the image plane of the camera. In most electro-optical cameras, this image plane is defined by the sensor plane.

The image plane is also provided with a coordinate system to record the position of features on the image. In practice of course, positions will be measured in pixel coordinates, so ultimately we will have to make provision to measure in pixel coordinates. The object or scene to be captured is described in terms of a world coordinate system. It is therefore often convenient to fix the world coordinate system to the object or scene.

This section is about deriving the relationships between these coordinate systems. Written in homogeneous coordinates it is a linear relationship expressed in terms of a matrix called the camera matrix.

**Image Processing Algorithms**

Image processing algorithms start with a given picture or video and ends up with useful information about a certain object like its position and orientation or object number and so on. In the present study the pencil position and orientation and the end effector position are considered the data to be determined by image processing algorithms.

The proposed algorithm should guide the gripper to grasp the objects from its centroid, so the centroid of objects should be obtained. The blob analysis block in the Simulink software is very similar to the "region props" function in MATLAB. They both measure a set of properties for each connected object in an image file. The properties include area, centroid, bounding box, major and minor axis, orientation and so on. The details of the proposed Simulink models will be explained in the next section. In the following sub-section three different image processing algorithms are discussed.

## 4. Retailing

Retail business can be understood as all activities related to sales activities and adding value of goods or services directly to the consumers.

First, the management of goods is located in the scientific field of trade marketing, which is defined as analytical processes, target formulation, strategy selection and the composition and control of the marketing mix in a trading company. The four areas of the marketing mix are central to the decisions to be taken within the scope of trade marketing. The basic concept of the 4Ps, introduced by McCarthy, structures the marketing into four separate components: product, price, location and promotion.

The ordering of goods includes all activities relating to the replenishment of the shops, the filling of the shelves and the reaction to customer requirements. This includes the processes between central warehouses and stores, between suppliers and warehouses, and between suppliers and stores depending on the type of trading company. Serving customers includes tasks that are intuitively attributed to trade, such as sales advice and the actual core activities at the cash desks in the store. The operational tasks include the initiation, execution and downstream processes of a transaction.

The transport of goods and logistics includes all tasks related to the storage of goods. This includes any functions related to the creation or management of the warehouse structure for the transfer and management of the warehouse in general like the management of storage locations and optimization of shelf space. It also covers the operational tasks between central warehouses in the individual stores, from the manufacturer to the direct goods stores, and the coordination of these activities.

The retail sector is characterized by an oligopolistic market with strong intra-competition between existing retailers and rising inter-competition between traditional and new digital players in many countries around the world. With Amazon Fresh about to enter the grocery market, this competition is intensifying. Due to the nature of stationary trade in particular, the work areas can be described as focused on manual human activities (Howe, 1992).

Lots of changes are happening in the retail industry. Of the many changes, there are two things that are very rapid changes, namely new technology and customer communication. These two things are related to one another. Utilizing technology in the retail business will make it between customers and sellers as if without distance, whenever wherever communication can still be done until the goods are received by consumers. Changes in consumer behavior in shopping from online to online are increasingly apparent. This is supported by the development of e-commerce and online transaction security systems that make consumers more comfortable doing online shopping. Developments that have made retail businesses in the world of ecommerce increasingly good on a small scale or even established retail companies have begun to see this potential.

The Amazon Go concept is an evolution from the self-checkout operational model that is used at Walmart and other grocery stores. However, self-service in the retail arena has been evolving for

over a hundred years. In 1916, Piggly-Wiggly made the innovative decision to allow customers to explore the aisles of products without the assistance of an employee. Then, in 1992, some of the first self-checkouts debuted (NCR, 2014).

Since then, the self-checkout model has been widely accepted and immensely popular. According to a study conducted by computer and point-of-sale terminal company, NCR, those who use the self-checkout indicated that they appreciate the convenience and ease of use that it provides. One interesting note of those surveyed, was that they indicated that they still liked knowing an attendant was nearby to provide help if needed (NCR, 2014).

Paul Forsell, an Assistant Manager with a major retail and grocery store, estimated that roughly 20 to 30 percent of customers prefer self-checkout. The rest give preference to the traditional cashier check-out method. He indicated that while selfcheckouts are popular, some people prefer to not have to deal with the technology themselves. In fact, Walmart's attempt to have customers in 200 stores pre-scan their purchases using their "Scan & Go" app failed when customers could not figure how to use it. Walmart discontinued the experiment (Anderson, 2014)

## 5. About Amazon Go

Amazon Go is a prototype grocery store that is operated under the auspices of the e-commerce company Amazon.com (Ives & Addams, 2019). Before shopping on Amazon Go, shoppers only need to download the application on their mobile phones. The application is also connected to the user's credit card so that every item to be purchased is billed directly to their application account.

If the application is not installed on their cellphones, prospective buyers on Amazon Go will not be able to enter the store. This shop installs a scanning machine as we usually encounter it at the commuter train station as the entrance. By attaching a cellphone that has the Amazon Go app installed, one can enter it. Once inside, buyers are free to take the goods into their shopping bags. The Amazon machine knows when an item is taken off its shelf and identifies it as an item to be purchased. Conversely, when the item is returned to the shelf, the machine will realize it as an item that is not purchased. When the customer finishes shopping and leaves the store, the items in the virtual basket will immediately be paid for with electric money (balance) on the Amazon Go app and will show the receipt for the purchase. When a customer leaves, the sensor will

automatically detect the customer's ID who is leaving the store and will give an order to cut the customer's balance on the Amazon Go application and show the purchase receipt (Grewal, Roggeveen, & Nordfalt, 2017).

All processes just happen without human intervention. The process at Amazon Go relies on hundreds of sensors, cameras and machine learning that are scattered throughout the store. These sensors and cameras witness the activities of shoppers in the shop (Sujata & Hasandeep, 2019).

**Shopping Process**

In order to shop at an Amazon Go store, the consumer must first create an Amazon account, have a smart phone, and download the Amazon Go app. Once these 82 Polacco, Backes / Journal of Business and Management, 24 (1), March 2018, 79-92 items are in place, the customer scans their Amazon Go app, located on their smartphone, upon entering the retail store. At this point, the customer is free to walk around the store, shop, and add and replace items to or from their virtual cart (Amazon Go Editorial Staff, 2017). For Amazon Go to accomplish this paradigmatic shift in the operational model of retail shopping, it relies heavily on technological innovation. The technology Amazon Go uses automatically senses when an item is picked up, put back on the shelf, and tracks who committed the action. Once the customer is satisfied with the items they have chosen, they simply walk out the door - no lines, no checkouts, and no waiting. The purchased items are charged to the customer's Amazon account and a receipt is sent to the Amazon Go app (Amazon Go Editorial Staff, 2017).

**Technology Used**

Amazon Go uses what they call, "Just Walk Out Technology". This technology is responsible for keeping track of items taken from, and in some cases, returned to, the store's shelves. It also keeps track of the individual's virtual cart (Amazon Go Editorial Staff, 2017).

For this to be possible, Amazon uses technology similar to that of self-driving cars. The system relies heavily on sensor fusion, computer vision, and deep learning algorithms (Amazon Go Editorial Staff, 2017). While they have not revealed too much about their technological advancement, it appears that Amazon plans on making cameras central to their operational strategy. These cameras will track not only the products and their placement, but also the individuals who do the shopping (Swanson, 2016).

# CONCLUSION & RECOMMENDATION

Artificial intelligence tries to figure out how to get machines to use language, to form abstractions and concepts, to solve types of problems that are currently reserved for humans, and to improve themselves. It contains Computer vision which is a discipline that studies how to reconstruct, interpret, and understand and machine learning which is a field of computer science that provides learning skills for computers to learn something without detail programming. Merging these two concepts is used in various industries, including players in retail industry. In Amazon Go, all processes just happen without human intervention. The process at Amazon Go relies on hundreds of sensors, cameras and machine learning that are scattered throughout the store.

Artificial intelligence continues to grow and develop and will not slow down for a while. Applying Artificial Intelligence Especially Machine Learning and Computer Vision to the strategies of the players in the retail industry has been proven to have a good impact, especially in helping customers have a better experience and provide the insights needed to be successful.

## REFERENCES


Dahria, M. (2008). Kecerdasan Buatan. Jurnal Saintikom, 185.

Grewal, D., Roggeveen, A. L., & Nordfalt, J. (2017). The Future of Retailing. Journal of Retailing, 1-6.

Howe, W. S. (1992). Retailing Management. Hampshire: Macmillan.

Ives, B., Cossick, K., & Addams, K. 2019. Amazon Go: Disrupting retail?. Journal of Information Technology Teaching Casenotes, 5-8. DOI: 10. 1177/2043886918819092.

James, S. (2018). Artificial Intelligence and Machine Learning and Marketing Management. Hampshire: Southampton University.

Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2012). Foundations of Machine Learning. London: MIT Press.

O'Hagan, I. (2019). The Routledge Handbook of Translation and Technology. Oxfordshire: Routledge.

Sujata, J., Mukul, P., & Hasandeep, K. 2019. Role of Smart Communication Technologies for Smart Retailing. International Journal of Innovative Technology and Exploring Engineering, 8(6), 1015-1019. DOI: 10.35940/ijitee.F1209.0486S419.

Stepan, A. (2020, April 9). Computer vision powered by machine learning will change the way we see the world. Qualcomm. https://www.qualcomm.com/news/onq/2018/04/24/computer-vision-powered-machine-learning-will-change-way-we-see-world.

Weber, Felix Dominix, & Schutte, Reinhard. 2019. State-of-the-art and adoption of artificial intelligence in retailing. Digital Policy, Regulation and Governance, Vol. 21 No. 3, pp. 264-279. https://doi.org/10.1108/DPRG-09-2018-0050

Wingfield, N. (2016, December 5). The New York Times Company. Retrieved from The New York Times: https://www.nytimes.com/2016/12/05/technology/amazon-moves-to-cut-checkout-line-promoting-a-grab-and-go-experience.html