UNIVERSITY OF MINES AND TECHNOLOGY

TARKWA

FACULTY OF ENGINEERING

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

A PROJECT REPORT ENTITLED

RECYCLERAI: A CONVOLUTIONAL NEURAL NETWORK FOR DETECTING AND CLASSIFYING RECYCLABLE GARBAGE

BY

KILU ENOCK

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF A BACHELOR OF SCIENCE IN COMPUTER SCIENCE AND ENGINEERING

PROJECT SUPERVISOR

……………………………………….

MRS MILLICENT AGANGIBA

TARKWA, GHANA

DECLARATION

I declare that this project work is my own work. It is being submitted for the degree of Bachelor of Science in Computer Science and Engineering in the University of Mines and Technology (UMaT), Tarkwa. It has not been submitted for any degree or examination in any other University.

……………………………………..

(Signature of Candidate)

…………………. Day of June, 2020.

ABSTRACT

Recycling is important for sustainable development. The recycling process involves the sorting of recyclables into various classes. There are six major types of recyclables. Sorting of the recyclables into their various classes is a very time consuming and less productive activity. Consumers are also often confused about how to determine the correct way to dispose of garbage. This project aims to leverage the use of machine learning, particularly deep learning to simply the detection and classification of solid waste or garbage. Leveraging on Convolutional Neural Network(CNN), a single label image classification model is developed to accurately detect and classify solid wastes into the six groups of recyclables. After conducting experiments with real-world datasets, the model achieved an accuracy of 90.82%.

ACKNOWLEDGMENTS

My sincere gratitude goes to the Almighty God for His mighty hand that is at work in my life. I am greatly indebted to my supervisor, Mrs. Millicent Agangiba for her valuable insights and contribution without which this project would not have come to fruition. Much gratitude also goes to Mr. William Agangiba for his support throughout my studies and in this project as well. I would also like to acknowledge all other lecturers for being a constant source of inspiration.

Finally, I am grateful for my roommates and friends who were always there to support me both mentally and physically.

TABLE OF CONTENTS

Content **Page**

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Statement Of Problem

Waste management is a global issue that requires immediate global attention. In the 1940s, the first industrial production of synthetic plastics started and has increased considerably. In 1990, each individual in the world produced approximately 250kg of municipal solid waste resulting in a total of 1.3 x 10^9 tonnes of municipal solid waste (Beede and Bloom, 1995). A decade later it was found out that this amount has doubled to 2.3 x 10^9 tonnes(Beede and Bloom, 1995). One of the main reasons for this waste generation is due to urbanization. It is projected that by 2050, 68% of the world population will be moving to urban cities and that will cause a considerable increase in the amount of municipal waste generated globally (Hannah and Max, 2018). Around half of the world's municipal waste generated is from developed countries (Canada, Japan, Australia, Western Europe, New Zealand) and the remaining from Africa and Asia (Hannah and Max, 2018). Improper disposal of municipal solid waste mostly causes the release of greenhouse gases that account for global warming and leachate which cause surface water and groundwater pollution which causes various water-borne diseases (Aljaradin and Persson, 2012).

Ghana also faces a lot of problems associated with waste management. According to Miezah *et al,* 2015, Ghana generates about 12,710 tonnes of municipal solid waste out of which 14% is plastic. Poor management of these plastic waste generated has resulted in choked gutters, poor drainage system and polluted water bodies (such as the Korle Lagoon in Accra which has large amount of untreated industrial waste). These clogged gutters and polluted water bodies serve as breeding grounds for diseases and pollution. In Accra, it is estimated that only 60% of the waste generated is being collected(Kwasi Owusu Boadi, 2002). Government efforts made to aid in the collection of waste by engaging in house-to-house garbage collection has proven not to be successful at handling such large volumes of waste generated in the country.

One of the global solutions for waste management is recycling. Recycling is the process of breaking down and reusing materials that would have otherwise been disposed of. Not only does recycling helps reduce the amount of waste sent to landfills and incinerators, but it also in reducing the amount of pollution, creating green and safer environment(Lienig and Bruemmer, 2017). There are some private as well as government institutions that have recently been engaged in recycling such as the Sewage Systems Ghana Limited and Gamihot enterprise. One of the main problems that these recycling agencies face is the automatic segregation of garbage into recyclables. A lot of time and manpower is spent trying to manually segregate these garbages into respective recyclable classes. Consumers are also most times confused about how to determine the correct way to dispose of a large variety of materials used in packaging.

The world is moving towards automation. Without automation, the collection of waste continues to be a labor-intensive activity and very time-consuming. This project, therefore, proposes an automated solution system; RecylerAI. It is envisaged that RecylerAI, would be a classification model that can classify recyclable garbage into their respective classes. Using machine learning, specifically deep learning, the model would be accurately trained to classify recyclable garbage. Deep learning is an artificial intelligence function that imitates the workings of the human brain in processing data and creating patterns for use in decision making. The deep learning algorithm used is the Convolutional Neural Networks (CNN). The model would detect and classify garbage into six recyclable groups; Cardboard, Glass, Metal, Paper, Plastic, Trash.

1.2 Project Objectives

The objective of this project is to;

i. Detect and classify garbage into the six recyclable groups; and

ii. Build an interfacing client on top of the model (Mobile/Web application).

1.3 Methods Used

The following methods was employed in this project;

i. Review of relevant literature;

ii. Convolutional Neural Networks;

iii. Deep Learning; and

iv. Transfer Learning.

1.4 Tools And Facilities.

The following tools and facilities was used for the project;

i. Internet;

ii. GitHub;

iii. PyTorch; and

iv. Google Cloud Platform.

1.5 Scope Of Work

The project is limited to the detection and classification (sorting) of garbage into six recyclable groups;

i. Cardboard;

ii. Glass;

iii. Metal;

iv. Paper;

v. Plastic; and

vi. Trash.

1.6 Organization Of Work

This project is organised into five chapters as follows: Chapter 1 outlines the problem statement, project objectives, methods and facilities used as well as the scope and organisation of the work. Chapter 2 presents a literature review of this work. It begins with a review of the state of recycling. It then proceeds to discuss general literature on machine learning, deep learning and focuses more on CNN. Further, the works of other researchers that are relevant to the problem statement are also reviewed. Chapter 3 encompasses discussions on the design of the deep learning model. This includes the reason for the choice of Convolutional Neural Network (CNN). Chapter 4 covers experiments to determine the effectiveness of the proposed model. Chapter 5 presents the conclusion and recommendations and provides directions for future studies.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

The main aim of this project is to develop a model using deep learning(CNN) to accurately identify and classify garbage into their recyclable groups. This chapter examines the state of recycling, the various types of materials that are known to be recyclable. This chapter further discusses machine learning and its role in deep learning and concludes with a review of some related works.

2.2 State of Recycling

Due to the increasing population and manufacturing processes, there was an enormous amount of solid waste generated during the middle age periods (Williams, 2005). Authorities moved overhauled garbage bins which were breading an increasing number of rats and spreading diseases from the street. This was considered the first garbage collection attempt in history. The first form of recycling was recorded in the 18th century and involves the recycling of metal scraps and paper (Palmer,1995). Recycling became very widespread and much adopted during the periods of world war 1 and world war 2. The main reason behind the recycling was the shortage of materials and economic crisis. A lot of countries and individuals throughout the 20th century adopted the idea of recycling because of the major economic and industrial benefits. After 2010, many recycling firms were established, with the European Union owning up to 50% of the world's share of the recycling industries and had since employed over half a million people. Products to recycle are changing as technology advances and overexposure to certain waste products are known to be very harmful and even worse carcinogenic (Álvarez-Chávez *et al.*,2012). The efficiency of recycling depends mainly on the stable supply of recyclable materials. Three major supply chains are used for the collection of these recyclables. They are mandatory recycling collection, container deposit legislation and refuse bins. Container deposit legislation typically involves a small amount of money for the return of recyclables typically plastic, bottles and metal. In recent years, many countries have banned the use and disposal of certain solid wastes. Despite the increasing effort of the government and individuals in this field, most recycling processes still rely heavily on human efforts and capabilities.

2.3 Recyclable Materials

A recyclable (recyclate) is a raw material that is sent to and processed in a waste recycling plant or materials recovery facility which will be used to form new products. A material is considered recyclable when the ratio of the target material compared to the non-target material and non-recyclable material is considerably high (Ferrari *et al.*, 2019). The reason being that only the target material will be recycled and hence a high amount of non-recyclable material reduces the efficiency of the recycling process. Plastics are one of the most common recyclables. All plastics are not recyclables. Most plastics are assigned resin numbers which may determine their rate of being recyclable. Generally, plastics with resin numbers from 1-3 are known to be commonly recyclable. Over the years there have been some commonly recyclable municipal solid wastes. These are;

i. Paper;

ii. Glass;

iii. Plastic;

iv. Metal;

v. Cardboard; and

vi. Trash.

2.4 Machine learning

The intelligence of computer systems depends greatly on Artificial Intelligence (AI), specifically Machine learning. Machine learning is the process of teaching a computer how to make accurate predictions based on past studies. According to Mitchell (1997), a computer program is said to learn from experience (E) with respect to some class of tasks (T) and performance measure (P), if its performance at tasks in (T), as measured by (P), improves with experience (E). The experience(E) in this case will be recognizing the trend of prices over some time frame. The success of prediction will be the performance (P). The task in this situation will be the computer predicting stock for a future time frame. The fundamental difference between machine learning and traditional computer software is that with machine learning, the computer has not been explicitly programmed to instruct the system to make predictions, rather the computer learns to make accurate predictions from previously learned data. Machine learning has become one of the most successful subsets of Artificial intelligence. Machine learning is highly related to statistics. The major difference is that machine learning deals with large, complex data sets for which traditional statistics will be impractical (Chollet, 2017).

2.5 Main Types of Machine Learning

Machine learning is a broad field with various sub-branches which is dedicated to the study of a specific problem. However, the two main categories are unsupervised learning and supervised learning. The various types of machine learning are discussed below.

2.5.1 Supervised Learning

Supervised learning is teaching a computer by example. It can formally be defined as the process of determining a mathematical function that maps an input to an output based on examples of input-output pairs (Russel, 2012). It infers this function from training data consisting of training examples which are usually normalized to a vector (Mohri,2012). An example of supervised learning is a system exposed to a large data of preprocessed handwritten digits and overtime being able to accurately recognize these handwritten digits. There are two major classes of supervised learning. They are;

i. Regression predicts a quantity. The output of regression is continuous. A basic example is the prediction of stock prices and house prices; and

ii. Classification, however, predicts quantity. The output of classification is categorical. An example is predicting whether a mail is a spam or not.

2.5.2 Unsupervised Learning

Unlike supervised learning, unsupervised learning finds previous unknown patterns in large, complex data sets without any existing labels. This branch of machine learning consists of identifying interesting transformations of the input data without the help of any targets, for data visualization, data compression, or data denoising, or to better understand the correlations present in the data at hand (Chollet, 2017). Unsupervised learning algorithms are not designed to target a specific type of data, but rather to group similar data or find interesting patterns. One major application of unsupervised learning is in the field of density estimation in statistics (Jordan, 2004). The main types of unsupervised learning are dimensionality reduction and clustering.

i. Clustering: is a type of unsupervised learning with the goal of creating groups of data points such that points in the same group are highly correlated or similar to each other and those in other cluster are dissimilar.

ii. Dimensionality reduction: refers to the process of reducing the dimensions of a large feature set. Dimensionality reduction helps reduce the complexity of models and therefore prevent model overfitting, ensuring greater accuracy. Less dimensions also mean less computing power, lower storage and removal of redundant data.

2.5.3 Semi-supervised Learning

This approach falls between supervised learning and unsupervised learning. During training, semi-supervised learning combines a large amount of unlabeled data with a small amount of labeled data. One major benefit of semi-supervised learning is that it can greatly improve the accuracy of supervised learning by using readily available unlabeled data when the labeled data are scarce or expensive. Semi-supervised learning gives greater accuracy and requires less human effort.

2.5.4 Reinforcement Learning

Reinforcement learning is concerned with how agents take certain action in an environment to maximize some notion of cumulative reward (Tanner, 2018). Reinforcement learning allows machines and software agents to automatically determine the ideal behaviour within a specific context to maximize its performance. One typical example of reinforcement learning is Google's DeepMind's Deep Q-Network which has outperformed humans in a variety of vintage computer games. The model is given various pixels of the game with additional information such as distance between objects on screen and state of the game. The agent then determines how the state of the game and the action it takes relates to the score it achieves. After playing the game several times, the system builds a model of which actions will maximize its score and eventually becomes very good at playing the game.

2.6 Machine Learning Algorithms

The are various machine learning algorithms that can be used to solve a variety of real-world problems. Each algorithm is suited for a particular task. The most common machine learning algorithms are described in the following sub-sections.

2.6.1 Linear Regression

Linear regression expresses a relationship between input variables(x) and output(y) as an equation in the form $y = a + bx$. The goal of linear regression is to find the values of the coefficients of a and b. a is the intercept and b is the slope of the equation of the line. The main objective is to find the values of these coefficients that fit a straight line nearest to most of the points. This line is called the best line of fit or the regression line. A good regression model (a model that can make accurate predictions) is one in which the distance between the data points and the regression line is very minimal.

2.6.2 Logistic Regression

Unlike linear regression predictions that are continuous values, logistic regression predictions are discrete. The output of logistic regression is a probability that a given input belongs to a particular class. An example of logistic regression is predicting whether it will rain today or not. The output of such a prediction will be a probability indicating whether it will rain or not. Logistic regression is called so because it makes use of the logistic function $\mathrm{h}(\mathrm{x}) = 1/(1+\mathrm{ex})$. This is an s-sharped curve.

2.6.3 Classification and Regression Trees (C.A.R.T)

CART stands for Classification and Regression Trees. It consists of a number of nodes. The root node is the parent node, internal nodes and leaf nodes(nodes without any children or descendants). Each internal node represents a single input variable(x) and the leaf node represents an output variable(y). CART algorithms generally follow the following steps; Walk the internal nodes to arrive at a leaf node and then output the value at that node. CART, particularly decision trees are very useful because of their explanatory power. They provide a straightforward and intuitive explanation of how a decision was made. However, their main disadvantage is that they can be subject to overfitting and underfitting, particularly when using a small dataset. This can limit the generalisability and robustness of the resultant models (Song and Lu, 2015).

2.6.4 Artificial Neural Networks(ANNs)

Artificial Neural Networks were obtained from biological neural networks. A biological neural network is simply an interconnected network or group of neurons. A neuron is an electrically excited cell that forms the basis of the biological nervous system. The neurons consist of four main parts which are dendrites, nucleus, soma, and axon. An artificial neural network simulates the functions of biological neurons. The concept of Artificial Neural Networks was first proposed by McCulloch and Pitts in 1943. The original goal of ANNs was to enable computer systems to solve problems in a way similar to how the human brain solves problems. In recent times, ANNs have been used to solve a lot of tasks such as speech recognition, computer vision, speech translation, video games, computer-generated painting(Bethge et al, 2015). ANNs form the foundation of deep learning, another field that has proven to best at computer vision-related and speech recognition tasks. The most basic artificial network is the perceptron which is illustrated below.

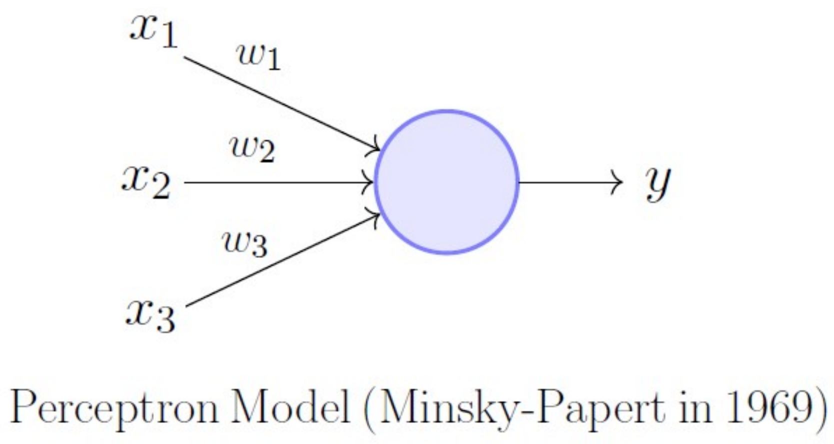

Fig 2.1 Perceptron Model (Source: Medium, 2019).

2.6.5 Deep Learning

The particular set of algorithms that will be used in this project will be deep learning algorithms. Deep learning is based on artificial neural networks. Deep learning involves a lot of useful layers of meaningful representation usually called the hidden layers. Each layer performs a particular transformation to the input data. The specification of each layer's transformation of the input depends on the layer's weights. The process of deep learning algorithms are similar to each other; the inputs data passes through various layers of varying weights and transformations. A loss function, either a cross-entropy loss function or quadratic loss function is then applied to the output. This loss function measures how well the neural network is performing. After calculating the loss, adjustments are made to the parameters of the model to reduce the loss. The adjustment makes use of an optimizer which uses a backpropagation algorithm (Rumelhart et al., 1986). At each layer, a dropout function may be applied to reduce model overfitting. Deep learning algorithms have proven to be best suited for recent state-of-the-art computer vision and speech recognition. The two broad deep learning algorithms are;

i. Convolutional Neural Networks (CNNs) which are mostly used for computer vision-related tasks such as object detection and classification; and
ii. Recurrent Neural Networks (RNNs) which are predominantly used in text and speech-related tasks such as chat-bots, speech recognition and translation.

2.7 Convolutional Neural Networks(CNNs)

Convolutional Neural Network is a feed forward artificial neural network. The difference between CNN and other deep learning neural networks is that the layer of the neurons that comprise the CNN is organized into three dimensions; the spatial dimensionality of the input(the width and height) and the depth. The neurons at any layer of the network is only connected to a small region before it. CNNs primarily assumes images as its inputs.

2.7.1 CNN Architecture

CNNs consist of three main layers. These are the convolutional layers, pooling layers and the fully connected layers. A simple CNN for detecting hand written digits is show below.

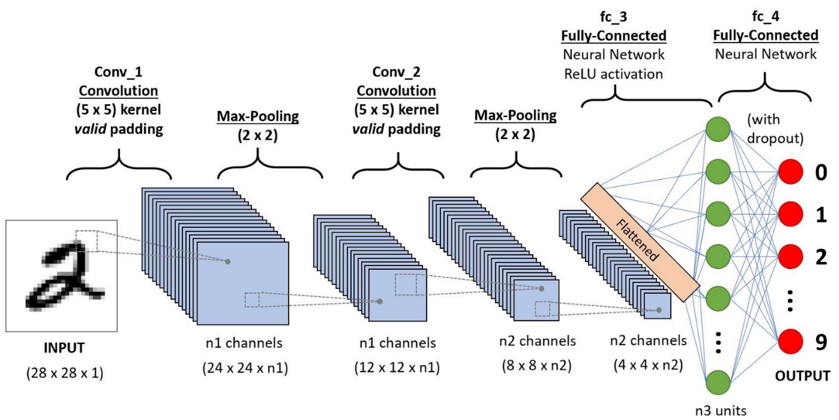

Fig 2.2 CNN Architecture for MNIST (Source: ***towardsdatascience.com*****, 2019)**

Convolutional Layer

The convolutional layer is an important layer in the convolutional neural network. This layer consists of primarily learnable kernels or filters. These kernels are usually small in dimension, however they cover the entire depth of the input. When the inputs data(images) passes through the network and encounters the convolutional layer, the layer of the inputs convolves with the filter to produce an activation 2D map. As the filter strides through the input image, the matrix multiplication of the filter and the portion of the input image is calculated. This will enable the network to learn kernels that get activated when they encounter a specific feature at a particular position of the input image. The convolutional layer consists of an important layer known as the receptive field. This receptive field is a small local area of a neuron that connects to the previous layer. This layer is very important in reducing the size of the model to make training of the model faster. The Convolutional layers helps to greatly reduce the complexity of the model by using optimization techniques. These optimization techniques depend on the depth, the stride and the zero-padding.

The depth of the output data produced can be reduced so as to minimize the total number of neurons in the network and help make training very fast. Another optimization technique employed in the convolutional layer is setting the stride number. The greater the stride number, the lower the amount of overlapping and hence the faster and less computationally intensive the model will be. Zero padding gives further control over the output volumes by padding the border of the input.

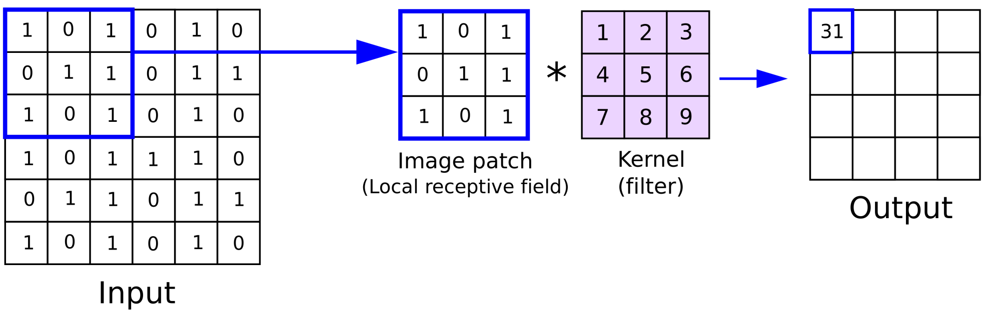

Fig 2.3 Convolutional Layer (Source: anhreynolds.com, 2019)

Pooling Layer

The pooling layer is most often placed after the convolutional layer. This layer does not add any learnable parameters to the model. It also consist of the stride and receptive field.

The main objective of the pooling layer is to greatly reduce the dimensionality and complexity of the model. The most used pooling layer is the max-pooling layer. The are two main methods of max-pooling. This is the overlapping pooling and the general pooling. The overlapping pooling has the stride set to 2 and a kernel size of 3. The general pooling is responsible for performing data normalization.

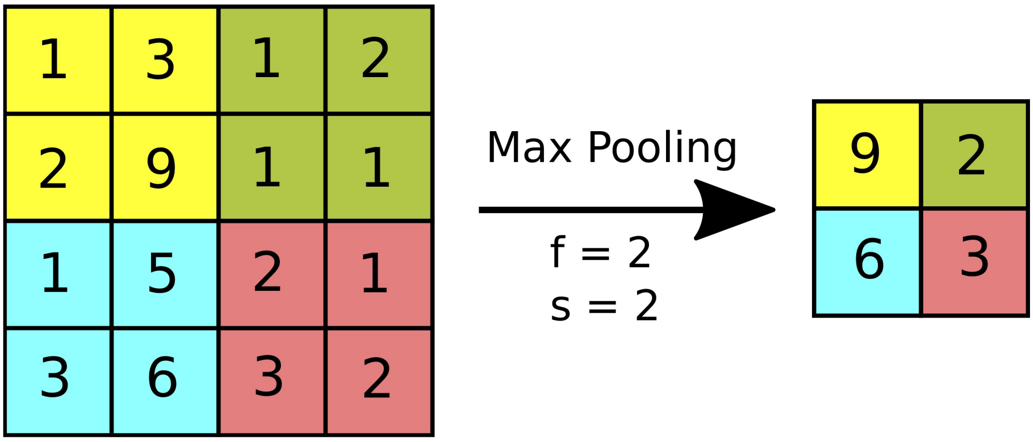

Fig 2.4 Pooling layer, (Source: *anhreynolds.com*, 2019)

Full Connected Layer

The fully connected layers consists of neurons that are connected to all adjacent neurons to form a complete connected layer hence the name, fully connected layer.

2.8 Related Works

One of the early works done to illustrate intelligent waste classification was based on a classical pattern recognition system called the Bayesian framework (Liu, 2010). This methodology was not very useful because it required a lot of manual or hand-extracted features.

One methodology that has proven to be very accurate in image detection and classification is the Convolutional Neural Network (CNN). AlexNet is an example of a highly capable CNN architecture (Krizhevsky, 2012).

In another study (Donovan,2016) autoTrash was developed using raspberry pi with a camera module to automatically sort waste as either recyclable or compost. AutoTrash was built on top of Google's TensorFlow framework. The autoTrash classified solid waste as either recyclable or compost.

Similarly, a study by Mittal et al., (2016), built a smartphone application that helped local citizens to track and report solid wastes in their communities. Their study employed transfer learning by utilizing a pre-trained AlexNet with images extracted from Bing image search result. The model attained an overall accuracy of 87.69%.

CHAPTER 3

MODEL DESIGN

Design of deep learning models involves many related steps. The accuracy and performance of a model strongly depend on the decisions taken whiles designing the model. The model considered for the trash classification and sorting is the Convolutional Neural Network(CNN). The general steps or procedures involved in a CNN are;

i. Data gathering and pre-processing;
ii. Division of dataset into training and validation sets;
iii. Actual model construction;
iv. Training of the model;
v. Hyper-parameter tuning; and
vi. Making final predictions with the model.

This chapter focuses and elaborates more on the procedures enumerated above to train, classify and sort recyclable trash.

3.1 Data Gathering and Pre-processing

A model accuracy depends mostly on the data that is supplied to the input. Data gathering is therefore an essential procedure that cannot be overlooked. The dataset collected was a mixture of dataset that was handpicked and those that were obtained from the internet. The dataset consists primarily of the six groups of recyclable garbage. The dataset collected from the internet was from a publicly released dataset that was made available by a group of students at Stanford University that performed a similar project. The dataset collected from the public/internet consisted of 400-500 images each of the six groups of recyclable objects making a total of about 2,400 images. The collected data consisted of varying sizes and orientations as well as varying light and pose. Handpicked images were taken with a personal mobile camera also at varying angles and also had different sizes. In order for the model to both accurately predict and also not over-fit(can only work for dataset seen before), data pre-processing has to be done. Fig 3.1 shows a sample of the dataset collected.

Fig 3.1 Sample Datasets

A pre-processing script was written in python using the standard package known as the PIL or pillow, to automatically resize the images into 224 x 244.

These data(images) were then augmented. Image augmentation is usually done to artificially increase the number of images that can be seen by the network during training. Image augmentation means randomly resizing and cropping the images and also flipping the images horizontally. Different random transformations are also applied to images during each epoch or step of the training so that the neural network can identify and learn various versions of the image. Data augmentation is done before any normalization takes place. Test sets and validation sets are not augmented, however, these data sets are also resized and normalized. Fig 3.2 below shows an example of the data augmentation technique.

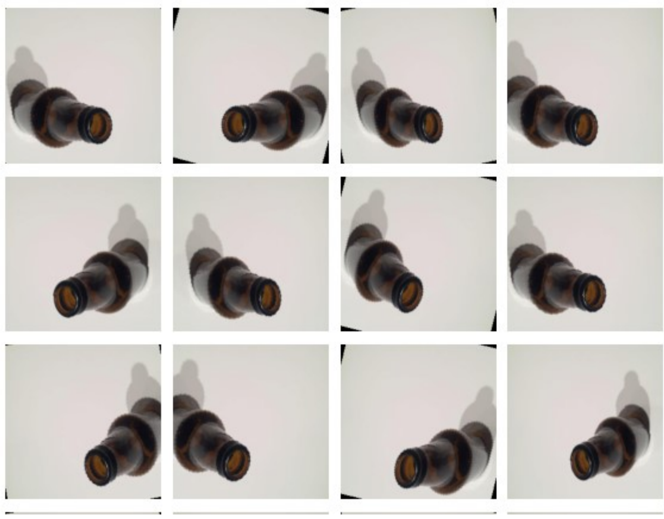

Fig 3.2 Data augmentation applied on a single glass image

3.2 Division of dataset into training and validation sets

After data selection and pre-processing, another important step is the division of the processed data set into training, testing and validation sets. The training set consists of a set of data sets that will be used to fit or train the model. The validation/training sets used refers to the remaining dataset that is not used for training but rather used to estimate the performance of the model to tune the models hyper-parameter. This provides an unbiased sense of model effectiveness (Khun and Johnson, 2013). A python script was used to automatically divide and group the dataset into training, validation, and test sets in the ratio of 50% (training sets), 30%(validation sets) and 20%(testing sets). Each sets consists of several images. Below is the distribution graph of the images in the training sets. It can be seen that paper has the largest number of images in the training sets.

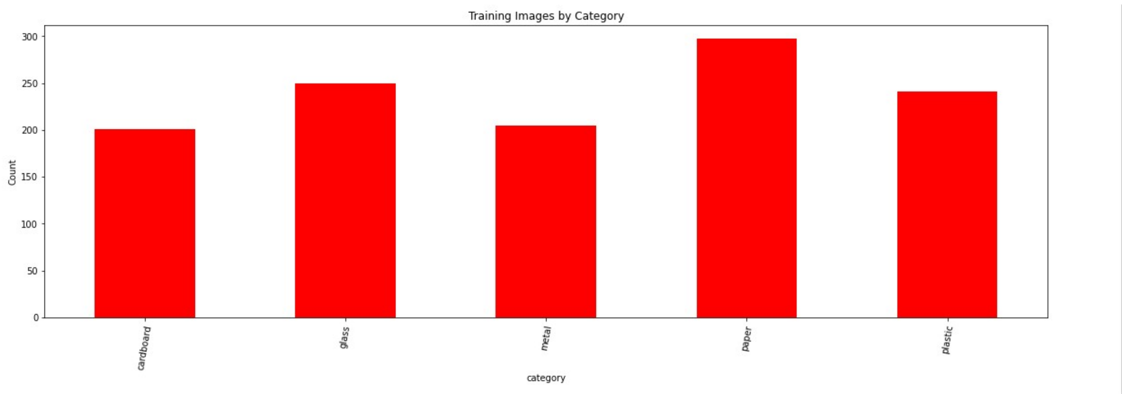

Fig 3.3 Distribution of images in training sets

3.3 Model Construction

Training a model completely from scratch is very time consuming and also tedious(Canziani *et al., 2016.*). The model used for this project was therefore constructed from a pre-trained deep learning model. Canziani *et al.* make an exhaustive list of pre-trained models that can be used for practical deep learning projects as well as their trade-off , such as memory consumption, time used during processing, etc. Some of these pre-trained models include AlexNet, ResNet-152, GoogleNet, ResNet-50, VGG-16, etc that can be used and built upon. Considering the various analysis and deductions he made, I decided to use ResNet-50 since it has a good trade-off between accuracy and inference time.The Selection of a pre-trained model(in this case ResNet-50) is from a concept called transfer learning. Transfer learning is a machine learning concept that focuses on knowledge gained whiles solving a problem(training) and applying it to a related problem (West *et al.*, 2007).

ResNet-50, as the name suggests, is a convolutional neural network with 50 deep layers. It is capable of handling more than a million training dataset, which makes it very performant. It has rich feature representation for a wide range of images(Kaiming et al, 2016). It receives an input of 224x224 image size. This is the main reason the images were augmented to that size during the dataset pre-processing. ResNet-50 is performant because it solves a common deep learning problem known as the vanishing gradient problem(Huang et al., 2017). The vanishing gradient problem occurs when more and more layers are added to deep neural networks. This occurs as a result of the gradients becoming extremely small as they are backpropagated or continuously multiplied through the network. Vanishing gradient problem causes the network performance to become stagnant or degrade. ResNet solves this vanishing gradient problem by using an identity shortcut connection that skips one or more layers.

To make use of the ResNet-50 pre-trained model, the last layer of the ResNet50 was replaced with a custom classifier. The custom classifier consisted of fully connected linear layers with ReLu activation, a dropout with a probability of 40%, and a log softmax output. The linear layers also known as the transition layers are used for the actual convolution, the ReLU activations (Nair and Hinton, 2010) are used throughout the network except for the final fully connected layer which uses a softmax non non-linearity. The softmax function computes final class scores that will be fed into the loss function or output.

3.4 Model Training

Model training involves continuously adjusting the models parameters in order to attain a desired result or output. The model training involves a series of steps that start with choosing a training loss, an optimizer and continuously calculating the gradient of the loss with respect to the model's parameter in order to update these parameters with the optimizer. The process of continuously calculating the gradients in order to adjust the model's parameter is know as backpropagation (Rumelhart *et al.*, 1986).

3.4.1 Loss Function

The loss function is the objective of CNN. The loss function evaluates how well a learning algorithm models the given data or inputs. If the predictions of a model deviate from the actual results, the loss function produces a large number indicating that the algorithm used does not best model the input data. There are a variety of loss functions that can be used for various machine learning tasks. They are mostly grouped into two classes; regression loss functions and classification loss functions. The Negative Loss Likelihood (NLL) function which is a classification loss function was used. This is because the model classifies multiple classes. The NLL function is given by the mathematical equation illustrated below.

$$CrossEntropyLoss =- \left(y_i log\left(\hat{y}_i\right)\right) + (1 - y_i) log\left(1 - \hat{y}_i\right) \tag{3.1}$$

3.4.2 Optimizer

The are various optimizers that are used for various learning problems. Optimizers use the gradients of the loss function to reduce the error of the model's prediction. The optimizer used for the training of this model is Adam optimizer. The Adam optimizer is an optimization algorithm that is used in place of the classical stochastic gradient descent optimization algorithm (Dierderik P. *et al.*, 2014). Adam is a very performant optimization algorithm. According to the authors or creators of Adam optimizer, Adam combines two main optimization techniques of the classical stochastic gradient descent.

i. It makes use of the Adaptive Gradient Algorithm. This algorithm helps improve performance on problems with sparse gradients. This makes Adam useful in natural language processing and computer vision problems; and
ii. Uses Root Mean Square Propagation. This algorithm enables Adam to perform well on non-stationary or online problems.

3.4.3 Choosing A Learning Rate

The learning rate controls how much the weights of a network are adjusted with to the loss gradient. If the learning rate is too low, training will progress very slowly. However, if the learning rate is too high, it can cause the loss to diverge from the minima. Choosing a good learning rate is therefore important to the accuracy of the model. A technique proposed by Smith(2018) was used for choosing the learning rate. A simple experiment is performed where the learning rate is gradually increased after each mini batch. The loss at each increment is recorded and plotted. Analyzing the slope of the plot illustrated in fig3.4 indicates that 1e-03 is the optimum learning rate.

```
In [22]: learn.lr_find(start_lr=1e-6,end_lr=1e1)
         learn.recorder.plot()

LR Finder is complete, type {learner_name}.recorder.plot() to see the graph.
Min numerical gradient: 5.13E-03
```

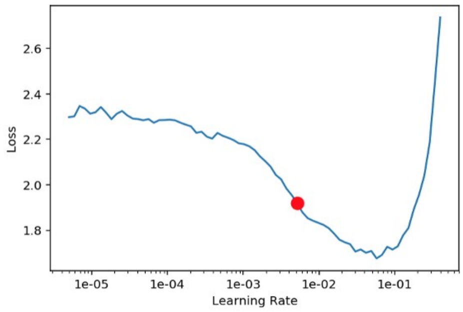

Fig 3.4 Learning Rate

3.4.4 Actual Model Training

The training of the model is done using Pytorch. Training of a model using Pytorch is an explicit process that grants you full control over what happens during the training process. The training is carried for a fixed number of times, called the epochs. For training the model, a total of 20 epochs was used. The training data consisted of a loader function that loads image datasets with a batch size of 32. This means that each data batch consist of 32 images. For each pass in the epoch, the images in a single batch will be passed through our model(resNet-50) to get the outputs. This process is known as forward passing. The loss in each pass is computed using the loss or cost function. The loss is calculated against the output from the model at each epoch. The gradient of the loss is then calculated to the model's trainable parameter. This process is known as backpropagation. The gradients were computed only for a small set of parameters that belong to the new or custom layer added to the ResNet-50 model. This helps in faster training of the model. Total loss and accuracy are then calculated for the whole epochs, which will be used later for validating the model's accuracy.

3.4.5 Early Stopping

Early stopping is a very important step in the model training process. It halts the training process when the validation has not decreased for several epochs. The main benefit of the early stopping is to prevent model overfitting (Burnharm, 2002). Model overfitting occurs when training loss continues to decrease but the validation loss stays same and hence the model starts to memorize the training data. Early stopping also helps so that we can save the best model when the validation loss is not decreasing again. Early stopping is implemented by iterating though the validation data, calculating the loss and comparing it to the previous loss.

CHAPTER 4

RESULTS AND DISCUSSION

This chapter presents experimental results from the testing of the model on the test set and verification of the model's performance of the Convolutional Neural Network that was used in this work for the identification and classification of recyclable garbage.

4.1 Making Inference

After training the model, there is the need to make inference, that is, test the model on data that it has not seen before. Inference is done on single images as well as the whole test data sets. Inference was done on the single images using the top N accuracy metrics. The top N accuracy is used to measure how often a class falls in the top N values of the softmax distribution. Below are some sample test images and their associated probability graph.

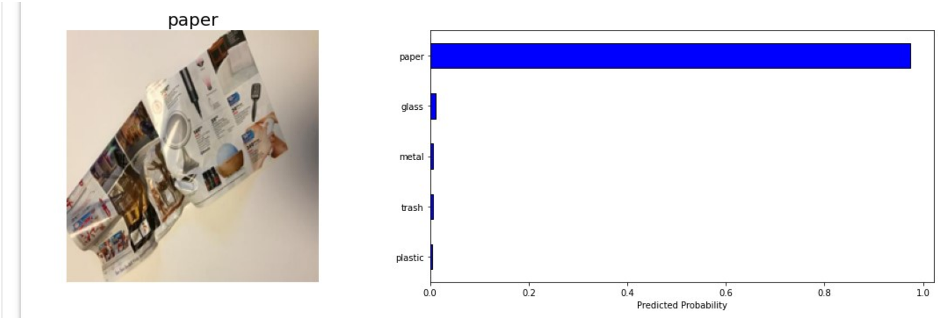

Fig 4.1 Inference on Model Using Top 5 Accuracy Metric

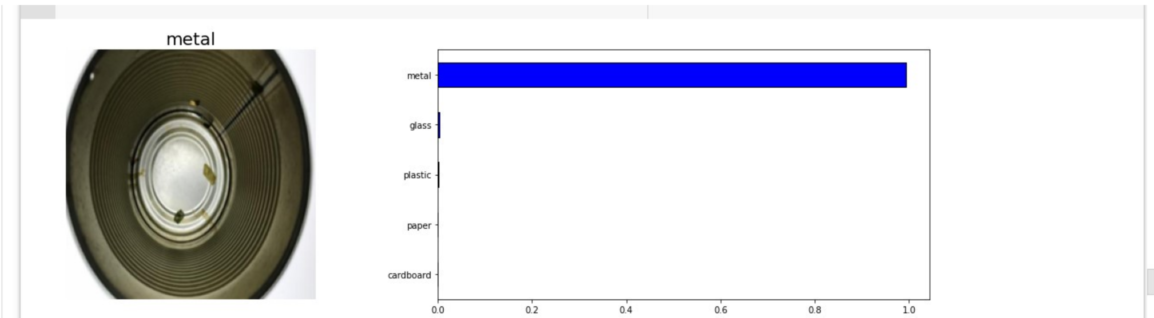

Fig 4.2 Inference on model using top 5 accuracy metric

4.2 Confusion Matrix

A confusion matrix is used in Figure 4.1 to present the results attained by the learner visually. The accuracy of the model can be calculated by taking the average of the values lying across the main diagonal of the matrix.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4.1)$$

where FP = False Positives, FN = False Negatives, TP = True Positives and TN = True Negatives. The accuracy can also be calculated as $1-e$ where e is the error rate. The accuracy of the model is found to be 90.82%. The accuracy of the individual classes is illustrated in table 4.1

Class	Accuracy(%)
Cardboard	93.069%
Glass	89.189%
Metal	95.833%
Paper	87.755%
Plastic	93.458%
Trash	86.42%

Table 4.1 Accuracy on model classes

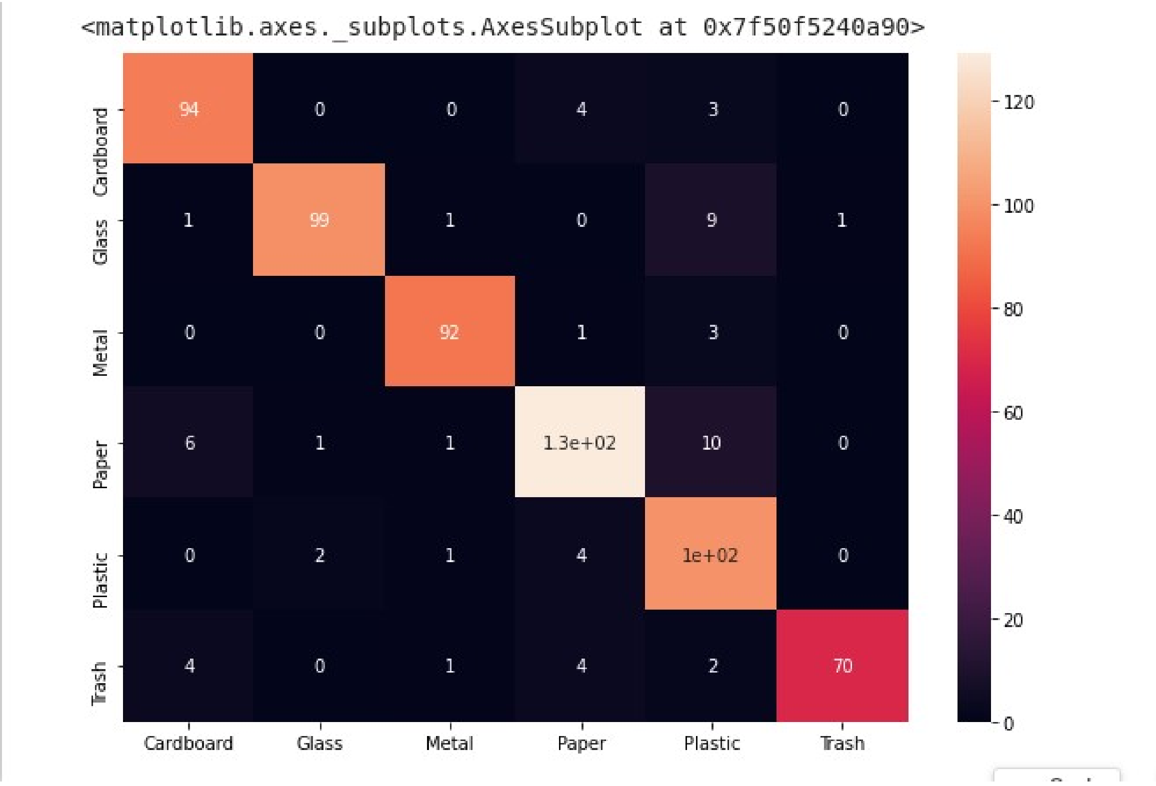

Fig 4.3 Confusion Matrix of Results

4.3 Mobile Application Results

In order to access and make use of the model, the model was deployed and a mobile client was built on top of it. The mobile application provides provides a simpler means for a user to take a picture of garbage with their mobile phone and camera and instantly determine the class of recyclable that the garbage belongs to.

DASHBOARD SCREEN

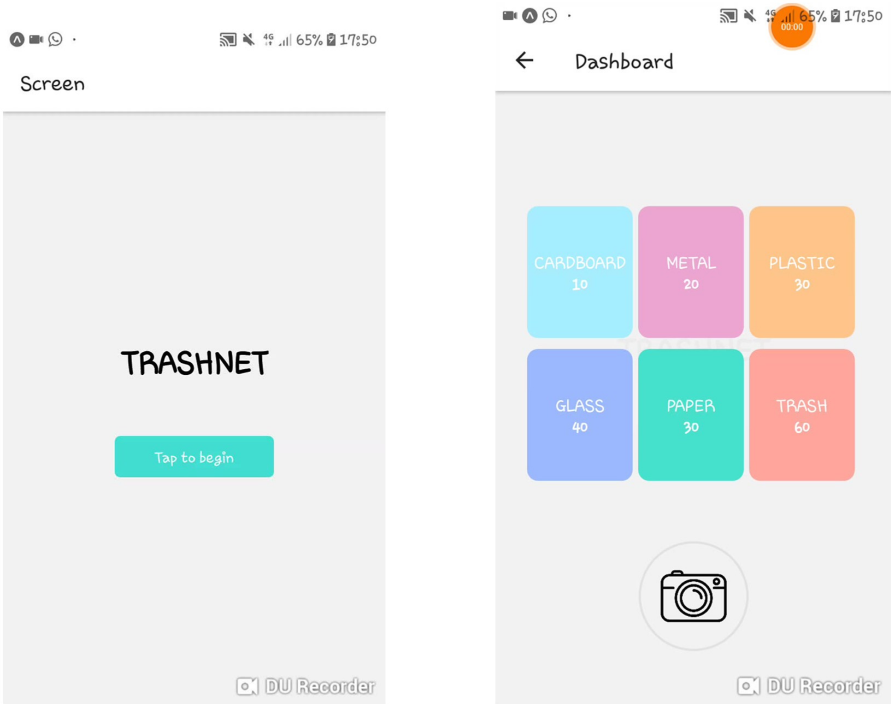

Fig 4.4 Dashboard Screen

MATERIAL DETECTION / CLASSIFICATION

Fig 4.5 Material Detection and Classification

CHAPTER 5

CONCLUSION AND RECOMMENDATION

5.1 Conclusion

This project developed RecyclerAI- a single-label image classification model to identify and categorize six classes. Using a Convolutional Neural Network, the model achieved an accuracy of 90.824%. The approaches employed in this project helped in accurately classifying the recyclable materials. Transfer learning was used to speed up the training process. The model developed in this project can be incorporated or serve as the basis of an autonomous recycling robot.

5.2 Recommendation

After the completion of this project, it is recommended that future projects should address the problem of identifying recyclable garbage as an object detection problem. It is recommended that future models should be focused more on obtaining datasets for the trash category so as to increase the model's accuracy.

REFERENCES

A. Krizhevsky, I. A. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in Advances in Neural Information Processing Systems 25, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105. Available: http://papers.nips.ccaccessed day, month year .

Aljaradin, M. and Persson, K.M., 2012. Environmental impact of municipal solid waste landfills in semi-arid climates-case study–Jordan. The open waste management journal,5(1), pp.28-39.

Bethge, Matthias; Ecker, Alexander S.; Gatys, Leon A. (26 August 2015). "A Neural Algorithm of Artistic Style"

C. Liu, L. Sharan, E. H. Adelson, and R. Rosenholtz, “Exploring features in a bayesian framework for material recognition,” in Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on. IEEE,2010, pp. 239–246.

C. R. Álvarez-Chávez, S. Edwards, R. Moure-Eraso, and K. Geiser, “Sustainability of bio-based plastics: general comparative analysis and recommendations for improvement,” Journal of Cleaner Production,vol. 23, no. 1, pp. 47–56, 2012.

Chollet, F. (2017), Deep Learning with Python, Manning Publications, New York, 384pp

Dummies.com, “Which Materials Can You Recycle?”.Viewed: November 12, 2019, <https://www.dummies.com/home-garden/green-living/recycling/which-materials-can-you-recycle/>

García T. Andres, Aragón R. Oscar , 2015, Intelligent Waste Separator, https://www.researchgate.net/publication/285611174_Intelligent_Waste_Separator. Accessed: November 12, 2019.

Global Waste Management Organization, 2015, “Waste management – still a global challenge in the 21st century: An evidence-based call for action”, http://journals.sagepub.com/doi/pdf/10.1177/0734242X15616055. Viewed: November 14, 2019.

Gripe Ghana (2019), “Recycling starts with segregation of your waste - Gripe Ghana”. Viewed: November 12, 2019, <https://www.ghanaweb.com/GhanaHomePage/NewsArchive/Recycling-starts-with-segregation-of-your-waste-Gripe-Ghana-731355#>

Hunt, V. Daniel (1985)."Smart Robots".Smart Robots: A Handbook of Intelligent Robotic Systems. Chapman and Hall. p.141.

J. A. Palmer, “Environmental thinking in the early years: Understanding and misunderstanding of concepts related to waste management,” Environmental Education Research, vol. 1, no. 1, pp. 35–45, 1995.

J. Donovan,“Auto-trash sorts garbage automatically at the techcrunch disrupt hackathon.” Available: https://techcrunch.com/2016/09/13/auto-trash-sorts-garbage-automatically-at-the-techcrunch-disrupt-hackathon/.

Konstantinos Sirlantzis, Paul Oprea(2019) “Handbook Of Electronic Assistive Technology”, Vol 1, pp 5-8.

Kwasi O. Boadi, 2002, Urban Waste Pollution in the Korle Lagoon, Accra, Ghana, https://www.researchgate.net/publication/226671931_Urban_Waste_Pollution_in_the_Korle_Lagoon_Accra_Ghana. Accessed: November 12, 2019.

Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar (2012) Foundations of Machine Learning.

P. T. Williams, Waste treatment and disposal. John Wiley & Sons,2005.

Rumelhart, D. E., Hinton, G. E. and Williams, R. J. (1986), “Learning representations by back-propagating errors”, Nature, Vol. 323, pp. 533 – 536.

Stuart J. Russell, Peter Norvig (2010)Artificial Intelligence: A Modern Approach, Third Edition.

Tanner, G. (2018), “What is Machine Learning and why is it important”, https://medium.com/datadriveninvestor. Accessed: January 29, 2020.

Ferrari, M., Benedetti, A. and Cirisano, F., 2019. Superhydrophobic coatings from recyclable materials for protection in a real sea environment. Coatings, 9(5), pp.303-317, Khun, M. and Johnson, K. (2013), Applied Predictive Modeling, Springer, Basel, 600pp.

Canziani, A., Paszke, A. and Culurciello, E., 2016. An analysis of deep neural network models for practical applications. arXiv preprint arXiv:1605.07678.

West, J., Ventura, D. and Warnick, S., 2007. Spring research presentation: A theoretical foundation for inductive transfer. Brigham Young University, College of Physical and Mathematical Sciences, 1(08).

He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778. 2016.

Nair, V. and Hinton, G.E., 2009. 3D object recognition with deep belief nets. In Advances in neural information processing systems (pp. 1339-1347).

Kingma, D.P. and Ba, J., 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

Burnham, K.P. and Anderson, D.R., 2002. A practical information-theoretic approach. Model selection and multimodel inference, 2nd ed. Springer, New York, 2.

O'Shea, K. and Nash, R., 2015. An introduction to convolutional neural networks. arXiv preprint arXiv:1511.08458.