2020

An intelligent image processing model for context-aware digital signage: case of apparel advertisement.

Omogo, Thomas Omondi
Faculty of Information Technology
Strathmore University

An Intelligent Image Processing Model for Context-aware Digital Signage: Case of Apparel Advertisement

By

Thomas Omondi Omogo

111343

Submitted in partial fulfillment of the requirements for the award of the Degree of Master of Science in Information Technology at Strathmore University.

Faculty of Information Technology
Strathmore University
Nairobi, Kenya

June 2020

Declaration

I declare that this work has not been previously submitted and approved for the award of a degree by this or any other University. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

Thomas Omogo
Signature: ………………….
Date: ……………………….

Approval

The thesis of Thomas Omogo was reviewed and approved by the following:

Professor Ismail Ateya Lukandu
Associate Professor, Faculty of Information Technology
Strathmore University

Dr. Joseph Orero (PhD)
Dean, Faculty of Information Technology
Strathmore University

Dr. Bernard Shibwabo,
Director of Graduate Studies,
Strathmore University

Abstract

Digital Signage is a way of presenting content such as advertisements, news, menus, and directions on electronic displays on places with high human traffic, Stadiums, Transport hubs, Malls, Retail Stores, and Notice boards. People have a variety of preferences when it comes to style and design of apparel. In developing advertising content, people feel more engaged when viewing tailor-made content fashioned toward a set of characteristics such are skin tone, age, and gender. The ability of digital signs to detect its contextual surrounding and intelligently display an apparel advertisement has been on-demand to create an autonomous content generator for digital signage. Currently, digital advertising content is developed by a designer who has experience both in computer skills, digital design with creativity and innovation. Content management software preloaded onto the digital screens makes it simple to load and display content. However, small scale retailers who invest in digital signage fail to achieve its full potential due to limited knowledge in creating a context-aware advertisement. The proposed research applies machine learning technique to build a model which captures a potential customer's apparel features using image processing then display a recommended outfit based on the input features. The result is a digital signage that autonomously creates a context-aware apparel advertisement based on what a potential customer is wearing. The model was evaluated based on accuracy, precision, and recall. Anthropometric measurements coupled with apparel segmentation fed into R-CNN gave the best apparel classification. The model's yielded an accuracy of 94.39%, with a precision of 0.72 and 0.84 recall, respectively.

Keywords**:** Digital Signage; apparel advertising; image processing; Convolution neural network; support vector machine.

Table of Contents

List of Figures

List of Equations

Code Listings

List of Abbreviations/Acronyms

4K	-	Horizontal Display Television resolution of approximately 4000 pixels
CMS	-	Content Management System
CNN	-	Convolution Neural Network
DVD	-	Digital Versatile Disc
HoG	-	Histogram of Oriented gradients
IoT	-	Internet of Things
LAN	-	Local Area Network
LCD	-	Liquid Crystal Display
LED	-	Light-Emitting Diode
QoS	-	Quality of Service
SVM	-	Support Vector Machine
TV	-	Television
UHD	-	Ultra-High Definition
VHS	-	Video Home System

Chapter 1: Introduction

1.1. Background

The product life cycle is defined as the progressive stages a product goes through from when it is first introduced to the market until it is finally retired, however, some products never reach the final stage as they continue to serve a need in the market (Kurkin, Januska & Marlin, 2010). The process is referred to as product lifecycle management (PLM) that integrates data, people and businesses to create a framework for an enterprise to extend its initiative in product or service distribution (CIMdata, 2018). PLM is segmented into four phases: introduction, growth, maturity, and decline (Wong & Kenneth, 2009).

During the introduction, the product is unknown to the market and prices tend to be high. The product placement is also selective, and promotion is generally personalized and informative to attract a target niche in the market (Pearce, et al., 2017). Campaigns in the form of advertising take center stage in this phase with the focus to inform and persuade potential customers, content is geared towards educating the people on the attributes and benefits of the new product. After a successful introduction, the product moves into the growth stage. The product's attributes become familiar with the market and price begins to decline with the increase in volumes while placement becomes more widespread. Promotion and advertisement are focused on the product's need to satisfy their properties (Teresko, 2004).

In both Introduction and growth stages, advertising plays a vital part in product success by providing a direct line of communication to present and prospective customers. Advertising is a paid nonpersonal communication to influence people's interest in a product by focusing on customer's needs to buy the product (Kotler, 1984). Advertising is classified based on the media used, geographic area, target audience and the purpose (Abey, 2012). Broadcast advertising is a form of media advertising that constitutes branches like television, radio, and the internet. Furthermore, broadcast media is grouped as traditional and online advertising wherein online advertising, the primary transmission mode is using the internet and other forms of digital media to reach a wider range of potential customers (Digital News, 2015). Advertisers are putting more resources in research and development of advertising methods that are effective and efficient in determining a potential customer's purchase choice.

According to Kolowich (2015), the human brain takes 13 milliseconds for images to elicit emotion compared to 250 milliseconds on visual cues therefore effective advertising has strongly shifted to develop content better suited to elicit emotion. Digital media in advertising have provided content developers with the platform to achieve better results. Digital advertising is used interchangeably with online advertising which refers to the use of internet technology to deliver marketing messages to consumers (NSF, 2013). It includes the use of tools such as social media, email marketing, and digital banners on websites to display advertising content to customers based on customer historical data.

Electric signs refer to the use of an illuminant as an advertising medium. Neon signs are the earliest form of electronic signage recorded, it refers to the use of lighted luminous gas tubes that contain neon gas to display signs in various colors used mostly at night (Dulken, 2002). Electronic signage now involves the use of LED signs and HID which refer to digital signage. In digital signage, illuminated panels such as LCD, LED and projectors are used to display images, videos or text for the purpose of advertisement.

During the development of digital signages, retailers used analog video streams on cathode-ray-tube TV monitors to run VHS and DVDs with pre-recorded advertisements. However, due to price drops and advancement in technology, flat panel displays became affordable and less bulky. Megatrends such as technology miniaturization; trends to develop ever-smaller mechanical, optical and electronic devices, led to the integration of flat panel display with a microprocessor. This integration developed computer-driven signage, with dynamic content management and central control (Schaeffler, 2013). The digital signage foundation takes advantage of the technology development in digital media and content management systems running on-premise or on the cloud. Content is developed depending on the target audience coupled with hardware capabilities. In numerous architecture designs, designers must maintain a regular update to the aired content to relevant content is displayed (Satoh, 2011).

In the fashion industry, digital signage captures 80% of the static signs used to create an immersive experience of potential customers to interact with their favorite designs. Digital signage provides a method for customers to try various apparel without reducing the value of the product (Warrington, 2015). The ability of UHD screen to display high-resolution images mixed with the visual artistic nature of fashion holds great potential in creating an impression on every customer's subconscious. In the fashion industry, brand awareness is the primary asset a retailer intends to achieve because customers are willing to pay extra for a brand name (Dulken, 2002). The fashion industry, therefore, fits digital signage like a glove since, in fashion, visual cues play a key role in capturing and reminding the customers of a specific brand. In order to enhance brand penetration and awareness in the market, stylists use colors and designs that can only achieve their intended effect if displayed on a high definition crystal clear screen offered by digital signage which is the most effective attribute.

1.2. Problem Statement

Advertising effectiveness is defined as the degree to which advertising generates the desired effect on a potential customer and how well a company accomplishes this intent (Starch, 2009). Effectiveness is generally measured using words and image flags; these are simple phrases or images that a customer can recognize and can mention when inquiring about a product. In 2010, 70% of normal American citizens had seen a digital sign with apparel advertisement (Navori Labs, 2019), this is a clear indication that digital signage can expand consumer knowledge about a specific brand and create a vivid image of the product.

Entrepreneurs invest in digital signage with an aim to improve consumer reach and market their products using creative multimedia tools displayed on digital signage. Multimedia types such as animation, video, sound, and graphics inspire creativity in creating apparel advertisement, however, such content require creativity and analysis of contextual information from the prevailing environment. According to Retail Academics (2018), entrepreneurs who invest in digital signage unknowingly create themselves an additional role of developing, designing and continuous publishing of relevant digital signage content. Therefore, digital signage elicits excitement and attracts the attention of potential customers but declines over time due to the ability of entrepreneurs to continuously create context-aware advertisements to run on the digital display.

The content of your advertisement is what makes consumers decide to buy or not to buy your product (Ogilvy, 2011). The digital signage runs a regular advertisement in a loop but intelligently switches to a customized advertisement when a person crosses the line of sight of a high-resolution camera. The camera detects the contextual features of a potential customer and applies machine learning techniques to recommend an attractive and customized advertisement appealing to the customer.

1.3. Aim

The research aims to develop a digital signage model with the ability to intelligently display an apparel advertisement based on its contextual environment.

1.4. Research objectives

(i) To investigate the basic human anthropometric features in the selection of apparel design.
(ii) To review existing apparel detection models based on neural network for digital signage
(iii) To build a context-aware digital signage model for apparel advertising.
(iv) To test the model in apparel advertising on digital signage

1.5. Research Questions

(i) What are the fundamental anthropometric characteristics in the selection of apparel design?
(ii) What are the neural network models for developing context-aware digital signage?
(iii) How effective is the model in generating adaptive content for digital signage?
(iv) How efficient is the model proposed?

1.6. Justification

Digital signage in clothing advertisement has presented content creators with a dynamic platform able to accommodate a plethora of designs and eye-catching displays for entrepreneurs. However, there is a tight bond between the effectiveness of digital signage and content creators. The research should assist entrepreneurs by automating content creation and advertising clothing on digital displays. This recommendation is based on automating the input factors that a designer uses to create an appealing advertisement. Secondly, autonomous content generation in clothing advertising keeps digital signage dynamic and appealing, the research should tend towards improving the return on investment for entrepreneurs. Lastly, the results of this research should also be useful to researchers and future scientists by contributing to the body of knowledge in intelligent advertising. It provides innovative ways using artificial intelligence to detect clothing features and customize adverts based on them

1.7. Scope

The research focuses on the detecting clothing of the upper torso of the human body through image processing and intelligently creating a customized advertisement based on these features. When advertising clothing, the lifecycle involves determining the advertising objective. A Budget that determines how much should be spent in each section while the research stage involves studying the general characteristics of the target market. The media plan highlights the medium and evaluation focus on effectiveness. This research proposes to cover the target audience research and media plan steps in digital advertisement design.

1.8. Limitation

The research is limited by resource constraints as per academic requirements; therefore, the scope is narrowed to focusing on clothing adverts of the human upper torso. Computing resources may also limit the number of images used to train the model. The research, therefore, proposes to limit the number of images used to train the model. The image dataset to use in modeling is accessible to the research, hence this delimitation is expected to contribute to a reasonable degree of accuracy of the proposed research

Chapter 2: Literature Review

2.1. Introduction

Machine apparel detection depends on several inputs to correctly detect and classify apparel according to its features, size, and shape. In order to facilitate learning, the machine algorithm extracts feature based on the following: category, texture, fabric, shape, and style. Some of these algorithms use similar features in the detection of apparel while others work a combination of one or two algorithms. Advertising is shifting from the designer desk to artificial intelligence (Wired, 2015) therefore there is a need for research to develop an algorithm with the ability to integrate with low-cost advertising models like digital signage. This chapter reviews relevant literature to understand the concept and investigate the research problem. The anthropometric features in relation to the human body, the architecture of digital signage and the current process of running advertisements on retail stores are reviewed. Previous technology evolution of digital signs is also presented with attempts on how to intelligently generate persuasive and context-aware advertisement to improve effectiveness.

2.2. Human Anthropometry

The phrase 'Anthropometry' is stemmed from a blend of two words ***anthrop*** (human) and *metricos* (measurement). It refers to the scientific measurement and collection of data about human physical characteristics such as body dimensions, body volumes, masses of body segments, the center of gravity, and inertial properties. According to Hrdlicka (1920), one of the initial anthropologists, anthropometric measurements are used in apparel design, art and sculpture, military designs, medical, surgical preparations and dental research and techniques, detection of body defects and their correction and forensic identification. Apparels differ from other consumer products in a number of ways. Most products, for instance workspaces, workstations, switches, and automobile interiors are based on the combined anthropometric data of men and women, this approach cannot be applied to the design of clothing. The body shape is different for men and women, boys and girls and each data set have to be treated independently. Age is often not as important a consideration in general product design as it is in the design of clothing. Change in body shape, size, posture, strength, and mobility of joints with age leads to a change in clothing preferences. As clothing follows the complex double curvature of the body forming an intimate covering for the same, the measures must correspond to the specific group for which it is being designed. Additionally, the accuracy of fit desired by the users in their garments is much higher than the fit needed in. Figure 2.1 summaries the anthropometric measurements for men's apparel body scanning technologies.

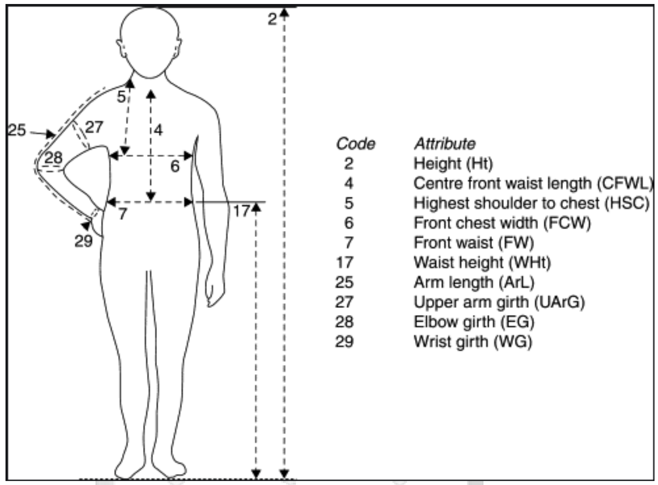


Figure 2. 1 Men Anthropometric Measurement

(Gill, Simeon & Paker, 2014)

2.3. Advertising using Digital Signage

In the 80s and 90s retails stores relied on the VHS tapes with pre-recorded content and cathode ray tube TVs to construct a video wall. When DVDs were invented in 1961, it offered digital content and therefore had improved benefits over VHS tapes. Fast forward to the early 2000s when flat-panel displays became an affordable and slim design, retailers replaced VHS and DVD media with computer-based signage players. The introduction of computer-based signage opened a world of possibilities for dynamic content design with centralized access to multiple video walls. However, the initial cost for setting up and digital signage was too expensive for retail stores (Samsung Display, 2011). The total cost of ownership was at $8,500 in 2004 and slowly decreased to $3,720 in 2010 (Wirespring, 2011) as illustrated in Figure 2.2.

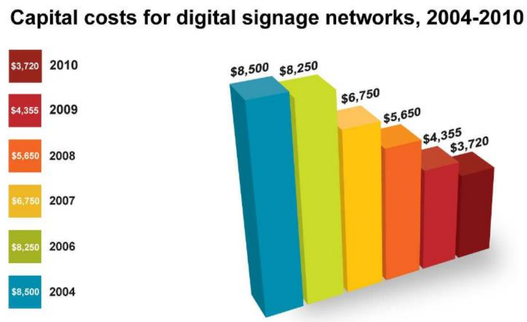


Figure 2. 2: Cost of Digital Signage Solutions

(Wire spring, 2011)

Acquisition cost coupled with operational expenses made digital signage inaccessible to retail business. This majorly due to the high cost of components that make up a digital sign. The rapid evolution of television hardware, the development of new management software and access to skilled labor caused a sharp decline in the cost of acquiring digital signage.

2.4. Digital Signage Design

There are four main components that merge to complete the digital signage system. The system components remain the same but are scaled accordingly depending on the intended installation target which ranges from menu boards on a single screen to a video wall that occupies a large stadium depicted in Figure 2.3.

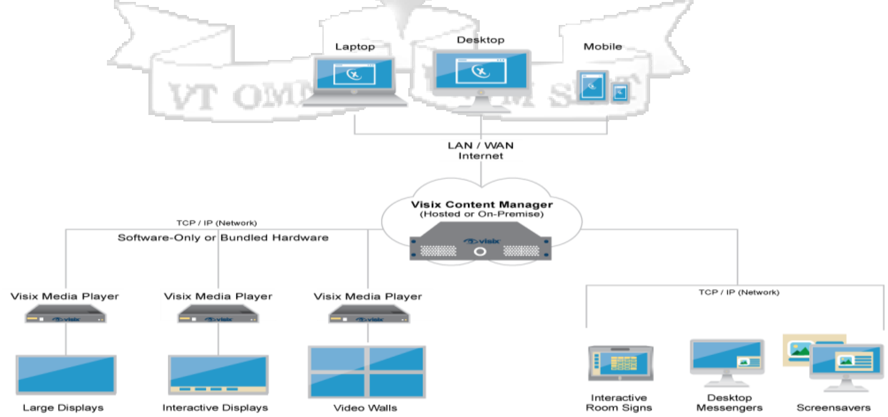


Figure 2. 3: Enterprise Cloud-based Digital Signage Solution

(Samsung, 2019).

2.4.1. **Software**

Software is the main engine driver for digital signage. It is composed of media players, device management software and finally the content management software.

(i) **Media player** – Make sure that the creators pick a player that can support various types and media formats (e.g. photos, music, videos). Manufacturers such as Samsung, Sony and Philips have introduced digital signs with design with onboard storage in case of lack of connection

(ii) **Content Management Software** – A content management software (CMS) manages the content to be played on the Display. There are many features available, which all come in a range of prices. Content can be uploaded and organized creating a playlist or distribute the same content to multiple players. Other CMS can only support a limited type of content, while there are those that can let you manage text, images, and videos

(iii) **Device management software -** Multiple Screens in different locations require central management. Digital Signage with cloud architecture can have screens on multiple continents. The right device management software ensures ability playback data from media players, check on players' health status, such as temperature, disk space, and performance, capture screenshots, update the software, or reboot the device remotely.

2.4.2. **Hardware**

Hardware is the physical part of the digital signage system. According to Farnham (2017), digital signage hardware forms a critical part of the system since it determines the content capability of the system. Hardware also informs the intended use of the overall system such as outdoor, indoor, crowds or single users. The hardware includes screens, media players, connectivity and wall mounts. Cloud option introduces internet connectivity on Wi-Fi or Local area Network (LAN)

(i) **Digital displays** – Digital displays for digital signage look very similar to standard consumer flat-screen TVs but are more robust and can be used 24/7 without screen burn. They are also often supplied brighter to suit ambient light and can be used in both portrait and landscape formats. Digital displays are LCD screen, (Liquid Crystal Display) and like modern televisions, are usually lit with LEDs from the edge. Most digital display companies (we included) do not manufacture digital signage displays but will source them from industry players like Samsung, LG, Philips, Panasonic, etc.

(ii) **Digital signage player** – Think of Blu-ray players. This hardware connects to another device and delivers the content to it. Make sure, however, that there are HDMI ports on both devices (and VGA, DVI, RCA, depending on the device) for compatibility.

2.5. Significance of smart Digital Signage in advertising.

Building smart digital signage will enable the ability to easily integrate various back-end systems, this allows retailers to fuse different concepts of advertising into one signage hardware (Haynes, 2013). According to Bauer (2018), digital signs presents a gateway to a single point of contact with all organizational systems, therefore a smart digital sign located in a storefront is able to detect customer attention and alert the sales representative of particular advertising that customer viewed. The sales representative can use this as an initial baseline to start a customer conversation. Research on audience measurement of digital signage shows a correlation of 52% in the display advertisement viewed and final purchase made (Ravnik, 2013).

There are 2.71 billion smart mobile phones around the world (Statista, 2019) with increased ability to facilitate near-field communications (NFC) and other mobile computing technologies that provide a connection between digital signage and smartphones to convey targeted advertisements. For large screens, digital signs, to retain its relevance, there is a need to build a close relationship with the small screen, smartphone (Haynes, 2013). Ultra-high-definition (UHD) displays also referred to as 4K, are changing the advertising atmosphere (Frost, 2009). The cost of ultra-high-definition has drastically dropped as depicted in Figure 2.4, and with the cost drop in production in UHD TVs, Digital signage flourished in creating 4K resolution content that gives an advertisement a realist feel.

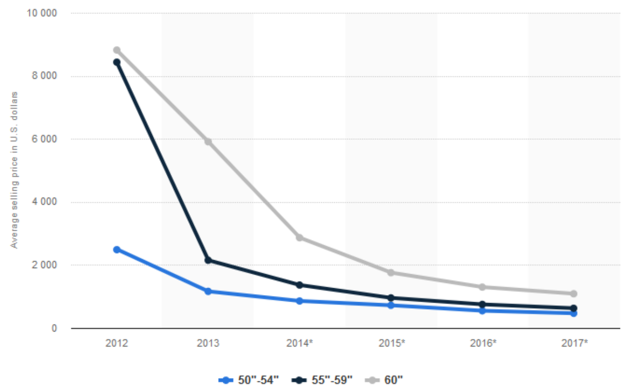


Figure 2. 4: Price Graph of UHD Display Screens

(Satista, 2019).

Digital signage content is quickly being developed on Ultra-high-definition which will swamp an organization bandwidth. The network infrastructure will be tested in terms of capacity to handle high bandwidth-consuming content, particularly if this content is pulled from the cloud (Khatri, 2013). Smart digital sign requires the ability to be able to maintain quality of service (QoS) despite a turbulent network stream.

2.6. Technological Challenges in Implementing Digital Signage

Researchers and technologists are currently testing smart signage to increase more human interaction. Machine-based content is already bringing generated on digital signs, this includes feeds from traffic, weather, stock exchange, sports scores, and social media trends. The content is driven by the dynamic changes in real human life powered by the internet of things (IoT) devices (Mackenzie, 2019).

Machine learning integration into digital signage focuses on the identification of human behavior to better personalize content based on the current situation. According to SightCorp (2019), machine learning has enabled face detection, people counting, demographic analysis, emotion recognition and face recognition and verification on digital signage. In addition, machine learning focuses on the following areas.

2.6.1. **Lack of a Personalized Experience**

Customer experience is important in providing a personalized experience. The machine-based system can learn and interact with people in a smart and unique way to every individual. Unlike consumer, local stores, intelligent digital signage can learn customers' names, say hello, recall customer preference, number of visits and the preferred section of a store (Wertime & Fenwick, 2017). When technology reaches astounding accomplishments, organizations should be able to inspire personalization, therefore, becoming beneficial in customer satisfaction rather than a privacy concern. A high processing camera can determine the personal feelings and emotions of a potential customer. Figure 2.5 depicts image processing on digital signage for context-aware emotion sensors. Intelligent digital signage can detect a customer's feature, emotions and age, then display an advertisement based on these features to attract attention towards the retail store (Ichiro et al, 2011).

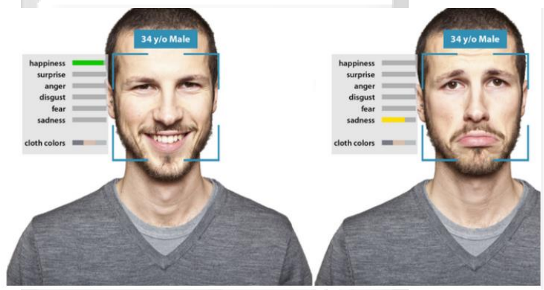


Figure 2. 5: Human Emotion Detection

(SightCorp, 2019).

2.6.2. **Content Generation Challenges**

The retail apocalypse refers to the migration of customers from a brick-and-mortar store to online stores, this implies that the retail market has failed to continuously create experiences for shoppers. Machine learning can create personalized attention to meet a large array of customers. Machine-based system achieves this feature in the following ways

2.6.3. **Lack of Context-Aware Digital signage**

Digital signage lacks the agility that enables intelligent content to be generated and scheduled, displayed at different times. Intelligent content dynamically changes to reflect the current surrounding environment of the digital sign (Euiseok & Nahm, 2015). According to the Indian Institute of Technology (2015), context-aware digital signage enables self-contained programmable entities that can travel from computer to computer and provide multimedia content for advertising or user-assistant services by using artificial intelligence.

2.7. Apparel detection models and Algorithms

For the purposes of this research, the approaches reviewed in the apparel detection model are classified into two: empirical models and machine learning algorithms

2.7.1. Empirical Models.

Empirical models have largely been used by researchers and scientists to detect and currently classify images from still photos and video streams. Computer vision problems like image classification and color detection have traditionally been designed using hand-engineered features such as SIFT (Lowe, 2004) and HoG (Dalal & Triggs, 2005). The descriptions based on bag-of-visual words descriptors (Yang & Jiang, 2007) that focused on classifying an image from the features extracted in comparison with labeled key words. Generally, most empirical models work by following the progression of a problem definition, a collection of features or data, model a formula, model verification and implementation (Arcilla, 2012). The traditional models-based image detection on features like edges and corners to approximate the representation of the image and determine the scale-invariant, in effect this method tries to standardize all images to corresponding features using some square-window dimension. In traditional computer vision, two techniques were majorly used.

2.7.1.1. SIFT (Scale-Invariant Feature Transform)

The SIFT algorithm deals with image detection based on features like edges and corners are not scale-invariant. There are periods when a corner looks like a corner but looks like a completely different item when the image is scaled up by a few factors. The SIFT algorithm uses a sequence of mathematical approximations to learn a depiction of an image that is scale-invariant. In effect, it attempts to standardize all images (if the image is scaled up, SIFT shrinks it; if the image is shrunk, SIFT enlarges it). This corresponds to the idea that if some feature can be detected in an image using some square window of dimension $\boldsymbol{\sigma}$ across the pixels, then if the image was scaled to be larger, a larger dimension $\boldsymbol{k\sigma}$ is required to capture the same corner.

2.7.1.2. **Speeded Up Robust Feature – SURF**

The drawback of SIFT is that the algorithm itself uses a sequence of approximations using a difference of Gaussians for normalizing the scale. Unfortunately, this approximation scheme is slow. SURF is simply a speeded-up version of SIFT. SURF is designed to find a quick and dirty approximation to the transformation of Gaussians using a technique called box blur. A box blur is the average value of all the image values in each rectangle and it can be computed efficiently. Figure 2.6 illustrates the earlier image processing methods.

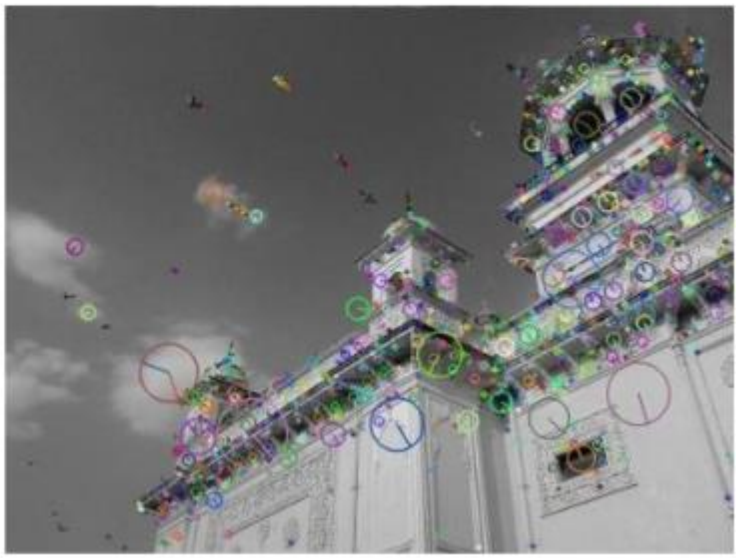


Figure 2. 6: Image Detection using SIFT Key points

(Bay & Tuytelaars, 2006)

2.8. Machine Learning algorithm used in Apparel Detection

In a supervised learning classification problem, a classifier uses labeled data to predict new or future events by mapping input to an output based on a training model (Russell & Norvig, 2010). When using machine learning in a classification problem, there are four issues that are considered: Categories used to classify class instances, training data, features of data to be represented and categorization algorithm (Feldman & Sanger, 2007). The research review two machine learning algorithms in apparel detection and recommendation.

2.8.1. **Artificial Neural Networks – ANN**

ANN is a parallel distributed processor that has a natural tendency for storing experimental knowledge (Chen & Yen, 2012) used in data characterized by non-linear ties, high dimensionality noisy, complex, imprecise and error-prone data. A key benefit of neural networks is that a model of the system can be built from the available data. Image classification using neural networks is done by texture feature extraction and then applying the back-propagation algorithm. The texture is characterized by the spatial distribution of gray levels in a neighborhood. In texture classification, the aim is to assign an unknown sample image to one of a set of known texture classes. Textural features are scalar numbers, discrete histograms or empirical distributions. In the design, four textural features namely the angular second moment, contrast, correlation, and variance are considered. Texture and tone have an inexpressible relationship to one another. They are always present in an image, although on occasion one property can overlook the other. In order to capture the spatial dependence of gray-level values, which contribute to the perception of texture, a two-dimensional dependence, and the texture analysis matrix is considered. Since texture shows its characteristics by both pixel and pixel values, there are many approaches used for texture classification. Figure 2.7 is an architecture of NN with combined gray value and textural features. In this, four layers consisting of three inputs, seven first layer hidden nodes, eleven-second layer hidden nodes and five output nodes are considered.

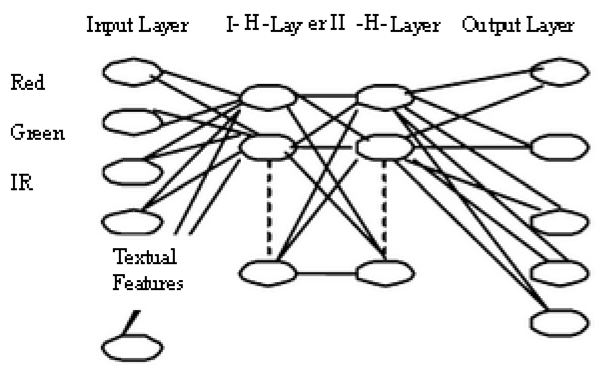


Figure 2. 7: Architecture of ANN combined with textual features

(Jazouli, 2017)

2.8.2. **Convolution Neural Network – CNN**

Convolution neural network (CNN) is a feedforward class of neural networks that processes data in a grid or matrix topology such as images. Figure 2.8 shows a binary image contains pixel arranged in a matrix with values to denote how bright and color each pixel should be (Brownlee, 2019).

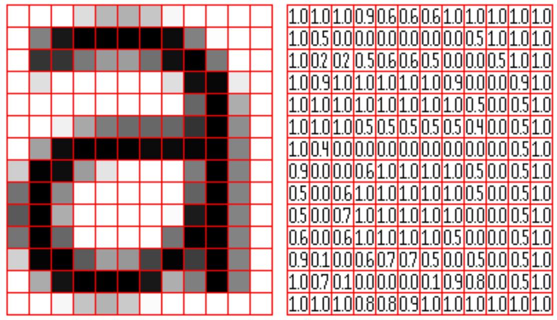


Figure 2. 8: Image representation as to the pixel grid

(Kolas, 2007)

CNN consists of neurons, weights, and biases. The neurons receive multiple inputs, generate a weighted sum of the weights, pass it via an activation function and reply with an output. This process is repeated over several layers to achieve higher accuracy. CNN operations are categorized into two phases: Learning and classification of objects.

During learning, an input with correct classification, these values are presented as a label then it is transformed depending on the weights of each connection as it approaches the output layer. Each unit from the output layer belongs to a different category and so now the output of the network is compared to the ideal output that would have been, has the pattern been correctly classified. In classification, the output unit with the highest value is selected as the best fitting category and passed forward to the next network is summarized in Figure 2.9.

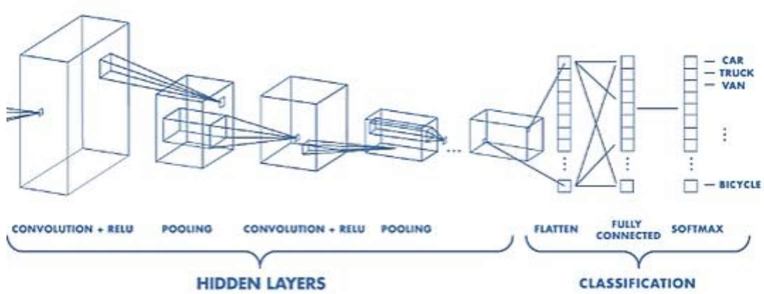


Figure 2. 9: CNN architecture

(Dielema & Sanders, 2013)

2.8.2.1. **Convolution Layer**

Convolution layer performs a product of two matrices where one set of matrixes is learnable parameters and the other is a receptive field. The previous layers featured are convolved with learnable parameters and relayed over an activation function to form an output feature map. Each output is combined with convolutions with multiple input maps. This is summarized in equation 2.1

$$\mathbf{x}_j^{\ell} = f\left(\sum_{i \in M_j} \mathbf{x}_i^{\ell-1} * \mathbf{k}_{ij}^{\ell} + b_j^{\ell}\right),$$

Equation 2. 1: Convolution Summary

M_j – denotes a selection of inputs and the convolution is of the valid border handling.

B – Output map additive bias, therefore if output *j* and map *k* both sum over input map *i* then the kernels applied to map *i* are different for out maps *j* and *k*.

2.8.2.2. **Pooling Layer**

In this layer, a downscaling of obtained layers is carried out, it reduces the image to its pixel density. A common pooling method id max-pooling (Hervatte, 2018) that is pooling by a ratio of 2 means the height and width of the image are halved. This helps in reducing the spatial size of the representation, which decreases the required amount of computation and weights. Figure 2.10 is a pooling operation that is processed on every slice of the representation.

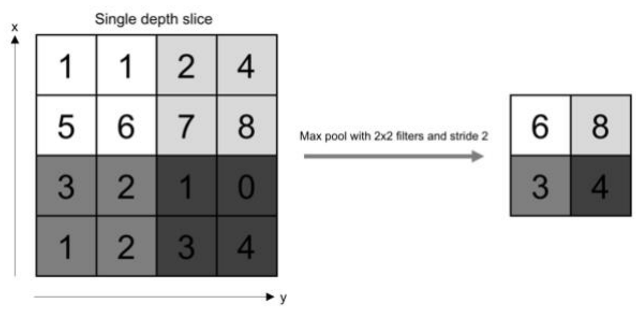


Figure 2. 10: Pooling Operation

(Kathy, 2015)

An activation map of six $W \ x \ W \ x \ F$, a pooling kernel of size E_I and stride S_i then the size of the output can be determined equation 2.2

$$W_{out} = \frac{W - E}{S} + 1$$ Equation 2. 2: Pixel activation

The output volume of size $W_{out} \ X \ W_{out} \ X \ F$. Pooling provides some translation invariance which means that an object would be recognizable regardless of where it appears on the frame

2.8.2.3. **Output layer**

After multiple layers, the output is required to form a class. The convolutions and pooling layers would only extract features and reduce parameters of the original images; therefore, we apply a forward pass is required to begin updating the weight and biases for error and loss reduction. Figure 2.11 summaries the layer's output.

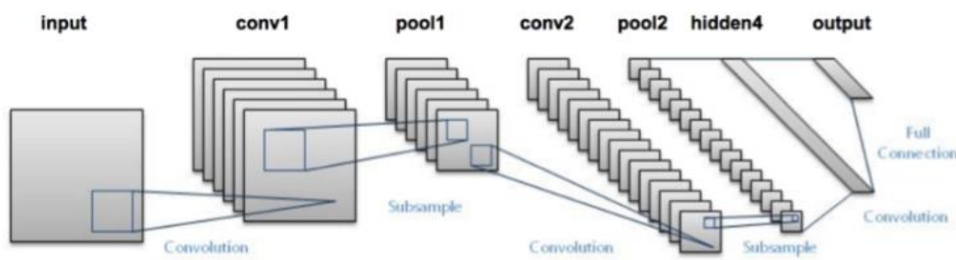


Figure 2. 11: Architecture of convolution neural network

(Dielema & Sanders, 2013)

2.8.3. **Support Vector Machine (SVM)**

The support vector machine was introduced by (Cortes & Vapnik, 1995) to determine an object in a multiple dimension hyperplane. Support vector finds a hyperplane $f(w, x) = w \ \ x \, X + b$ to separate two classes given features: x $\in R^m$.

SVM learns the features w by solving an optimization problem by equation 2.3

$$min \frac{1}{p} \mathrm{w}^T \mathrm{w} + C \sum_{i=1}^{p} max\left(0, 1 - y_i'(\mathrm{w}^T \mathrm{x}_i + b)\right)$$ Equation 2. 3:Learning Features

Where $W^T W$ is the L1 norm (also known as manhattan norm), C is the penalty restriction, y' is the label and $W^T X + b$ is the predictor function. This completes L1-SVM with a loss, however, it is corresponding, L2-SVM provides more stable results depicted on equation 2.4

$$min\frac{1}{p}\|\mathrm{w}\|_2^2 + C\sum_{i=1}^{p} max\left(0, 1 - y_i'(\mathrm{w}^T\mathrm{x}_i + b)\right)^2$$

Equation 2. 4: L1-SVM

Where $||W||_2$ is known as the L2 norm (as defined as *Euclidean norm*) with the squared hinges loss.

2.9. The architecture of Machine apparel Detection

In this section we explore the building blocks that effectively enable machine image detection, it features a combination of a sequential algorithm that optimizes the image quality at each step

2.9.1. **Conversion to grayscale**

Grayscale in computer imaging, refers to an image with two colors; black and white. Grayscale is a result of measuring the intensity of light in each pixel according to a weighted combination of frequencies (Wilhelm & Burge, 2010). The intensity of a pixel is expressed within a given range between a minimum and maximum. Figure 2.12 is the grayscale image of a T-shirt, this is the first process in machine image detection.

Figure 2. 12: T-shirt conversion to grayscale

2.9.2. **Background Elimination**

Foreground detection is used to eliminate the image background. Two images are captured separately, and the difference is one image has the background and the second image does not. The absolute value is achieved by subtracting the pixel value which results in the same location in the two images and the value is out range is considered as the apparel pixel location. The difference in pixel of both images, D can be defined as illustrated in equation 2.5.

$$D = |Q_{x,y} - P_{x,y}|$$

Equation 2. 5: Image Difference Function

Where $P_{x,y}$ represents the pixel of the image with background and $Q_{x,y}$ represents a pixel of image with clothes including the background. The final operation to eliminate the background is written as equation 2.6;

$$Q_{x,y} = \begin{cases} Q_{x,y} & D > R \\ 1(white & D < R \end{cases}$$

Equation 2. 6: Background Elimination

Where R represents the threshold value for background elimination.

Figure 2.13 is an image with an eliminated background. The background is a source of noise which reduces the accuracy of image detection.

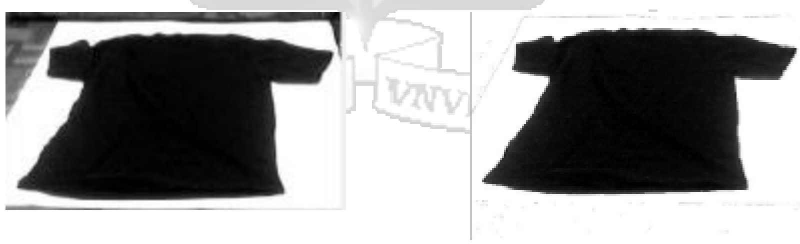


Figure 2. 13: Background elimination process

2.9.3. **Conversion to Binary image**

The algorithm takes an input of grayscale with only the foreground to improve the accuracy of detection. In order to achieve higher accuracy, a threshold operation is performed on a grayscale image to set minimum value. The grayscale (black-and-white) is converted to a binary image defined as shown in equation 2.7

$$p(q) = \begin{cases} 0 & P < T \\ 1 & q < T \end{cases}$$ Equation 2. 7: Gamma Encoding

Where T represents the binary threshold value and q the grayscale image value. The resulting image is Figure 2.14.

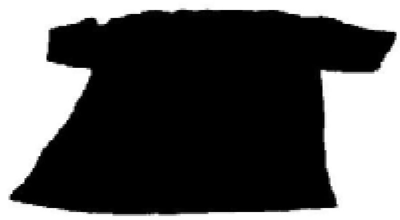


Figure 2. 14: Binary image of a T-shirt

2.9.4. **Image Noise removal and Enhancement**

According to Gonzalez (2002), noise reduction and enhancement are referred to as erosion (represented as ⊖) and dilation (represented as ⊕). The aim of the operation on a binary image is to erode the boundaries of the foreground pixels and shrink them size, the areas become clearer as in Figure 2.14.

2.9.4.1. **Erosion (⊖)**

A structuring element is used to shape, and probe a given image with the objective of finding if a shape fits or misses the outline of an image (Matheron, 2000). The knowledge of an image depends on the way we observe it. In the structuring the focus is on the image size and shape shown in equation 2.8 and 2.9

$$A \ominus B = \{z \in E | B_z \subseteq A\},$$ Equation 2. : Structring Element Function

Where B_z refer to the translation of an image B by the vector Z i.e.

$$B_z = \{b + z | b \in B\}, \forall z \in E.$$

Equation 2. 8: White Pixel Elimination

In structuring, element B has a center (i.e. a disk or a square), and this center is located on the origin of E, then the erosion of A by B can be understood as the locus of points reached by the center of B when B moves inside A.

2.9.4.2. **Dilation (⊕)**

In Dilation, the translation symmetry is used. The translational symmetry of an object refers to a translation that does not alter the image.

Translational invariance implies that, at least in one direction, the object is infinite: for any given point p, the set of points with the same properties due to the translational symmetry form the infinite discrete set depicted in equation 2.10

$$\{p + na | n \in Z\} = p + Z\ a.$$

Equation 2. 9: Scaling Factor

The dilation of an image A by B is defined in equation 2.11

$$A \oplus B = \bigcup_{b \in B} A_b,$$

Equation 2. 10: Translation Function

Where A_b is the translation of A by b; the dilation is dependent and given by equation 2.12

$$A \oplus B = B \oplus A = \bigcup_{a \in A} B_a.$$

Equation 2. 11: Dilation Function

If *B* has a center on the origin, then the dilation of *A* by *B* can be understood as the locus of the points covered by *B* when the center of *B* moves inside *A*.

Figure 2. 15: Image Enhancement

2.9.5. **Edge Detection of apparel**

The center of a binary image of apparel is determined by applying a matrix algorithm to a grayscale image. The algorithm of detecting apparel edges is applied to the pixel image of the apparel with the help of a threshold method around the pixels to change the center of the pixel. The algorithm is illustrated using matrices as follows.

The matrix used to represent the pixel in a 3x3 plane from a binary image represented in the following equation 2.13.

Equation 2. 12: Pixel Matrix

$$\text{Plane, p}=\begin{bmatrix} p1 & p2 & p3 \\ p4 & p5 & p6 \\ p7 & p8 & p9 \end{bmatrix}$$

P_5 represents the pixel of clothes and the sum of square windows excluding the center from the equation which is represented in equation 2.14.

Equation 2. 13: Matrix Sum

$$Sum,\ S = \sum_{i=1}^{9} I_i \ \ , i \neq 5$$

The center pixel of the square window is the *S* calculated from illustrated in equation 2.15

$$C = \begin{cases} 0(black) & 2 \leq S \leq 6 \\ 1(white) & S < 2 \ or \ S > 6 \end{cases}$$

Equation 2. 14: Pixel Center

Where the pixel will be indicated as the edge of apparel if S value is between 2 & 6

2.9.5.1. **Corner Identification**

The corner points are referenced as the junction of two or multiple edge lines (Zhen & Teoh, 1999). The method uses matrices with center detection from equation 2.13. Initially, the algorithm identifies the quantity of black and white pixel of an apparel image after erosion and dilation in order to detect the edges of an image object

$$\text{SUM} = \sum_{i=1}^{25} A_i \quad \text{where } A_i = 1(background) \text{ or } 0(clothes)$$

Equation 2. 15:Center Detection

In a 5x5 pixel where B13 is the center, the algorithm detection is summarized in table 2.1

B1	B2	B3	B4	B5
B6	B7	B8	B9	B10
B11	B12	C(B13)	B14	B15
B16	B17	B18	B19	B20
B21	B22	B23	B24	B25

B1	B2	B3		
B4	B5	B6		
B7	B8	B9/C1	C2	C3
		C4	C5	C6
		C7	C8	C9

M1

		C1	C2	C3
		C4	C5	C6
B1	B2	B3/C7	C8	C9
B4	B5	B6		
B7	B8	B9		

M2

B1				C1
B2	B6		C6	C2
B3	B7	B9/C9	C6	C3
B4	B8		C8	C4
B5				C5

M3

B1	B2	B3	B4	B5
	B6	B7	B8	
		B9/C9		
	C6	C7	C8	
C1	C2	C3	C4	C5

M4

Table 2. 1: Measuring intensity variation

diagonal; (M1-M2), horizontal (M3) and Vertical (M4) position

The corner algorithm utilizes the intensity measurements to determine the edges of apparel. The steps are highlighted in an iterative process with reference to the center of a matrix of the pixel. In a 5 X 5 matrix, the input of the equation becomes M1, M2, M3, and M4 passed through an algorithm in Figure 2.17 and the resulting edge of the T-shirt displayed on equation 2.17.

$$M1, M2, M3, M4 = \sum_{i=1}^{9} (A_i - B_i)^2 \qquad \text{Equation 2. 16: Edge Detection}$$

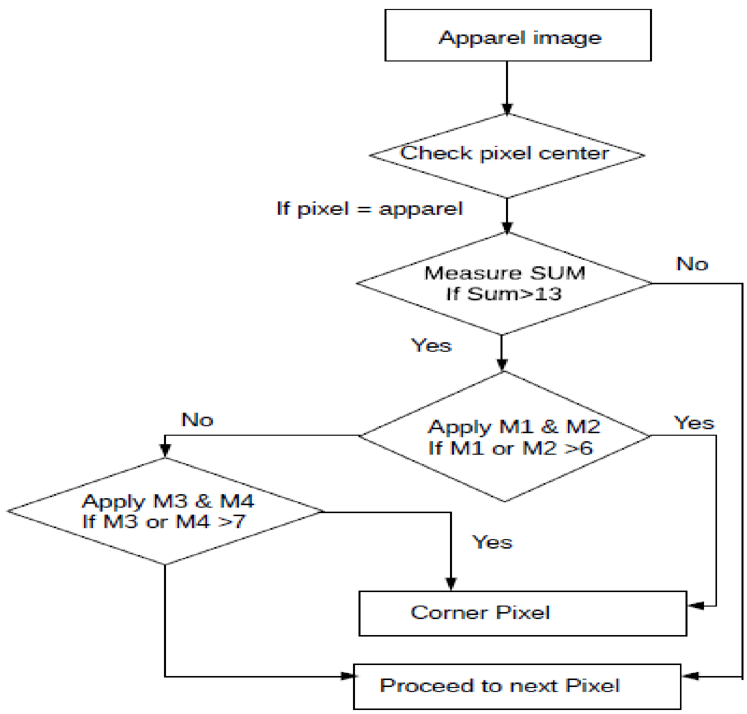


Figure 2. 16: Apparel Edge Detection algorithm

(Zhen, 2015)

The machine is able to outline the with high precision the edges on apparel. The edges enable a machine recommendation algorithm to recommend an appropriate clothing size to a customer in the event of a purchase, Figure 2.16 highlights the end detection of a T-shirt for size determination.

Figure 2. 17: Edge detection algorithm

2.10. Conceptual Framework

Founded on the literature studied and numerous gaps identified, this work proposes the following conceptual frame to detect apparel from a live stream on digital signage and recommend a targeted clothing advertisement. Data images will be collected from Google and open-source repository then goes through the preprocessing phase including pose estimation and torso segmentation, then edge estimation and background removal.

The solution is of two sections, for high accuracy rates in image detection CNN algorithm will use to train the model and second part SVM algorithm used to classify and recommend applicable contextual advertisement. The performance will be evaluated based on the metrics: accuracy, precision, views, and F-Score. The model will be trained until it reaches acceptable levels of performance where it can detect apparel fashion and determine a new similar fashion based on a live feed from a digital signage camera. The advertisement will then be presented on the digital signage screen to the potential customer. Figure 2.17 presents a visual summary of the framework

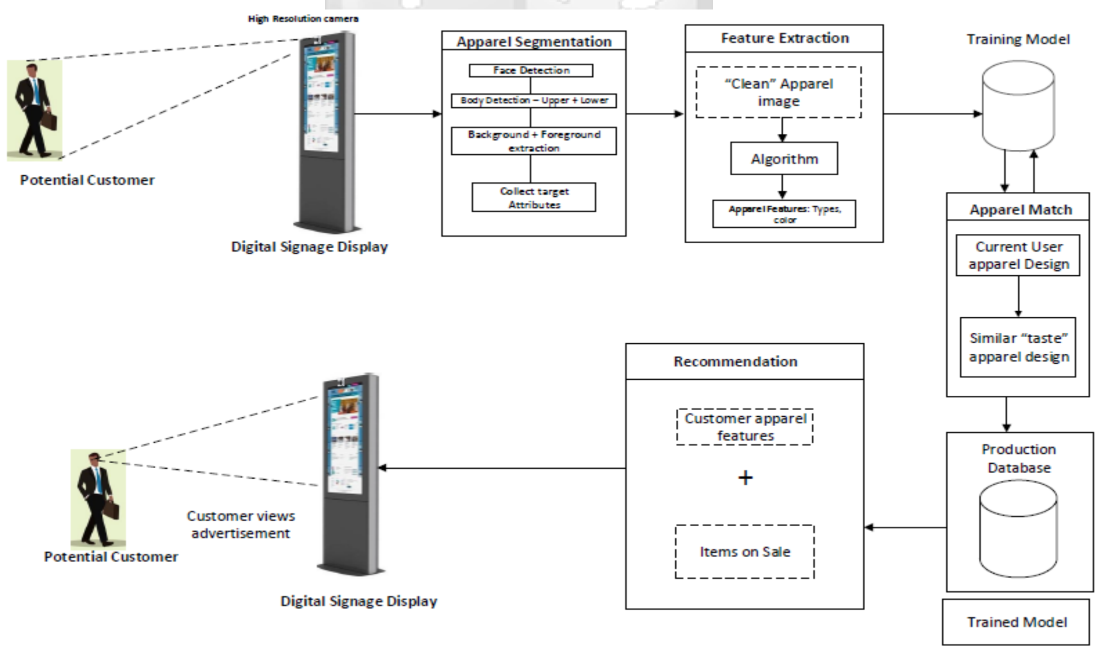


Figure 2. 18: Conceptual Framework

Chapter 3: Research Methodology

3.1. Introduction

Research can be defined as a systematic and objective process of solving problems with the aim of determining facts (Howard & Laird, 2013). The research and procedures that were adopted in carrying out the research are presented. Research was guided by the objectives of the study informed by the lucrative market space in contextual advertisements. The research exploited the advancement in artificial intelligence in machine learning and image detection to structure apparel advertisements in a targeted fashion. Experiments were designed to determine the best of features and conditions to present an appealing advertisement on digital signage.

3.2. Research Design

This research design involved appraising the tools and measurement methods used in collecting and analyzing variables outlined in the objectives (Creswell, 2014). Research design ensured an adequate sample size is used in carrying out the study (Thyer, 1993). An experimental approach was used to develop a model to detect apparel from a camera integrated with a digital signage display. Machine classes labeled with a predefined preference of apparel was used to train the model. The model was then tested using apparel images as a validation test to ensure an optimized performance was presented.

3.3. Data Collection

The research used secondary data from an open-source repository. ImageNet by Google currently has 14 million images of various categories. The research utilized a specific dataset of business class apparel. The images used were 150 by 150 pixels to improve the accuracy of the research since any measurements larger would require higher computational power. Figure 3.1 is an ImageNet dataset.

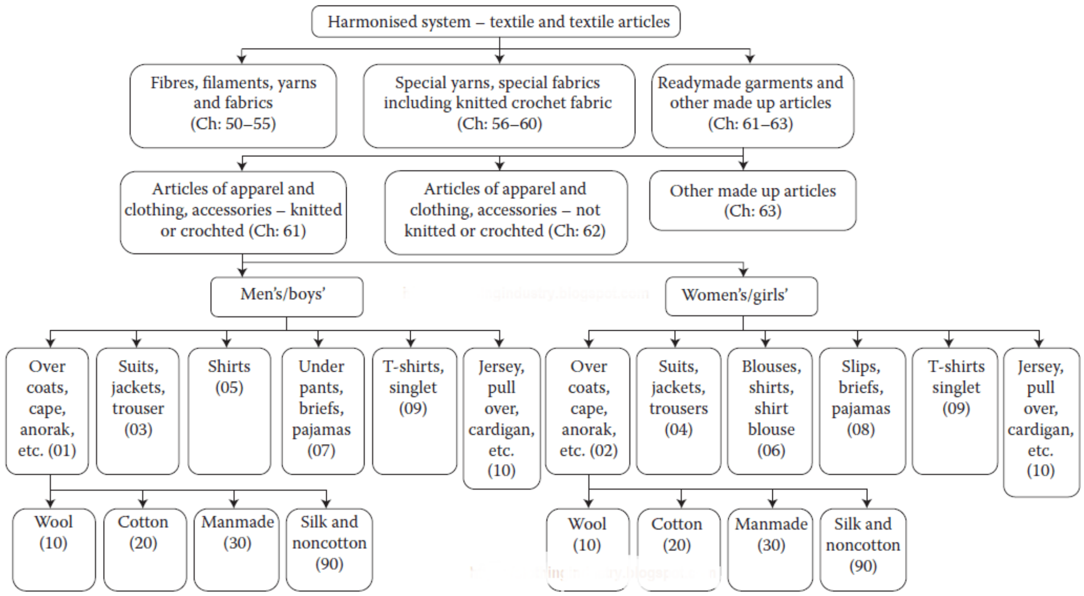


Figure 3. 1: Apparel database architecture

(ImageNet, 2015)

3.4. Data Preprocessing

The image labels were tagged with fashion recommendation to train the system. The learning processing takes an apparel category with a similar recommendation of the same category in Figure 3.2.

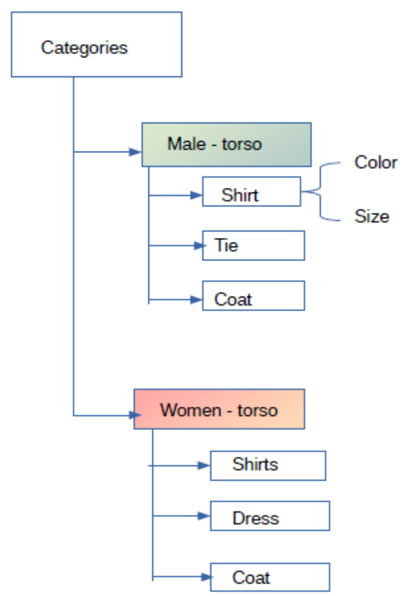


Figure 3. 2: Data Preprocessing

3.5. Model Training

The image samples were presented by labels to enable the application of the CNN machine learning algorithm. The images were split into training sets and testing set, these are used to train and validate the model respectively. Eight percent of the labeled data will be used to train the data and twenty percent used to evaluate the model.

3.6. System Development Methodology

Rapid Application Development (RAD) system development methodology which emphasizes quick short development time, with minimal functionality to test the concept present for review (Kerr & Hunter, 1993).

3.7. Overview of RAD Architecture

RAD is divided into four distinct processes shown in Figure 3.3.

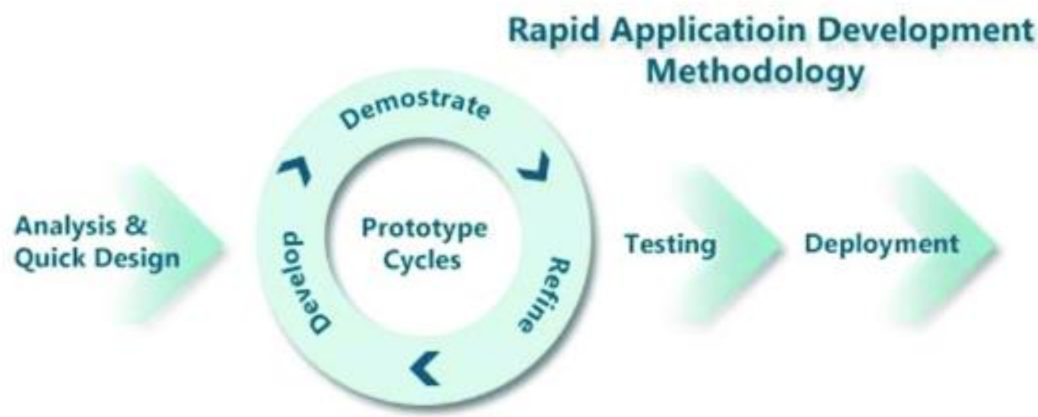


Figure 3. 3: Rapid Application Design

(Martin, 2001)

3.7.1. Analysis and Quick Design

In this phase, we appraised the data images to be collected and develop a skeleton design of the model. The intention is to design the general framework after which all the modules will be updated and tested. Challenges are expected in capturing the fashion requirements in the apparel of various categories of potential customers.

3.7.2. Prototype and Development Cycle

The prototype cycles included three major stages since majority problems and changes arise and are addressed through this stage. In development, a rapid coding of the model was done with backbone units of training and reviews considered A quick review of isolated components to improve on the weakness of previous code was done. In this phase, RAD enabled a broken-down and itemization of tasks which enabled quick error resolution.

3.7.3. Refine and Testing

The model training dataset is reserved at 80%, which ensured optimization is iterative until acceptable results are achieved. Unit testing was also carried iteratively. Labels define units of the model; this improved the ability to improve and tweak components at each level of the model.

3.8. System Implementation Tools

3.8.1. TensorFlow

TensorFlow is an open-source library for dataflow with mathematics libraries in machine learning and precisely python libraries in neural networks. TensorFlow was developed by (Google Brain, 2015) and composed of open source licenses. Tensor Processing Unit (TPU) provided an AI accelerator in quick learning throughput.

3.8.2. Keras

The cross is a python library that provided a high-level neural network API for running TensorFlow. Keras was developed considering requirements to quickly experiments that are developing an idea to a result in the least possible delay. Keras provides fast and easy prototyping and supports convolution networks and recurrent networks.

3.9. Model Evaluation and Validation

System evaluation and validation involved testing of the accuracy of the developed model. As a baseline, the evaluation of the model was on the ability of the model to correctly detect apparel and correctly display an apparel advertisement to the customer. Improved accuracy and display of apparel advertisement as compared to a similar model will serve as the validation method.

3.10. Ethical considerations

This research intended to adhere to the County outdoor Advertising Control Bill, 2018 by obtaining the necessary license from the Nairobi County during testing. The research intends to maintain advertising with extra care taken to protect customers. Digital Signage advertising model intends to use data ethical in advertising and avoid any methods that exploit emotional appeals, taking advantage of less-educated individuals and political associations. The research also will require the consent of a user to design and present an alternative apparel. Data is a critical part of advertising; the research intends to keep any demographic features collection confidential and safely destroyed after generation and use of such data.

Chapter Four: System Design and Architecture

4.1. Introduction

System analysis and design play a critical role in examining various component parts and their interactions, information flow, and transformation and overall architecture of the system that satisfy specific requirements. It enables the system analyst to have an aerial view of the system structure and giving a diagrammatic view of the system architecture to its stakeholders. The analysis and design of the systems conform to world-wide accepted UML diagram notation clearly depicting the environment of the system with all the entities interacting within the system. To accomplish this, design diagrams were developed including Use Case diagrams, sequence diagrams, and data flow diagrams.

4.2. Requirements Analysis

This research aims to develop a model to observe users' apparel and display a recommended apparel advertisement on the digital signage platform. Based on this objective, this section outlines the various requirements to be fulfilled in the proposed model.

4.2.1. Functional Requirements

(i) The system should accept a series of images from the camera

(ii) The system should transform each of the images and extract relevant features

(iii) The system should identify the apparel of the people in the image

(iv) The system should classify the apparel from the image accordingly

(v) The system should provide an output recommending similar or the combination of apparel currently on sale

(vi) The system should display the apparel on sale on the screen of the digital signage

4.2.2. **Nonfunctional Requirements**

4.2.2.1. Usability Requirement

The system is intended for entrepreneurs in the fashion industry with a basic level of technology configuration. The intended interaction between the user and the model shall be straightforward to allow entrepreneurs to easily set up the digital signage and project apparel in stock for potential customers.

4.2.2.2. Supportability Requirement

The main input channel for this model is a camera, therefore the system should be able to support a wide range of cameras that support various natural environmental conditions such as night, rain, fog and high light intensity among others. Essentially, this focuses on the implementation within a real-world environment.

4.2.2.3. Reliability Requirement

The system should consistently provide accurate results provided with the same set of input. The systems should accurately detect and classify input apparel and display matching apparel on the digital signage, additionally system needs to reduce false classifications as highlighted as follows:

(i) A regular snapshot of an instance of images captured at different times and natural conditions on local digital signage and remote processing server. This also provides a contingency for back up in case of disaster.

(ii) In case the model fails to run after reconfiguration or upgrade/update, the administrator should be able to revert to the last stable version that contains the prediction model and proceed to troubleshoot the model for occurring issues.

(iii) The model should be able to work in various naturally occurring conditions such as night, day, rain, fog and fog.

(iv) The model should be able to predict potential customer apparel taste with a satisfactory degree of accuracy.

.

4.2.2.4. Ease of reconfiguration

The system implements image processing using a convolution neural network. It may require reconfiguration of the weights to attain higher accuracy levels in the classification of the image apparels. Therefore, the system should have an easier method to enable quick reconfiguration and achieve better results.

4.2.2.5. Security Requirements

The system constantly captures images of people in a specific location which intrudes personal privacy. Information captured needs to be stored behind a system with strong security measures to maintain the privacy of such data. In addition, reconfiguration of the system requires authorized access to increase the security of the system.

4.3. System Design and Architecture

The system architecture depicts the general layout of the Intelligent apparel advertising prototype on digital signage and the main components it is made up of as illustrated in Figure 4.1. The system uses a camera as a source of input images. A camera built-in a digital signage display will capture a series of images of a person, these images will then go through a generated model to determine the apparel classification. The classified apparel will then pass through a preprocessor to determine fashionable combinations currently existing in the store database and return a recommendation to be displayed.

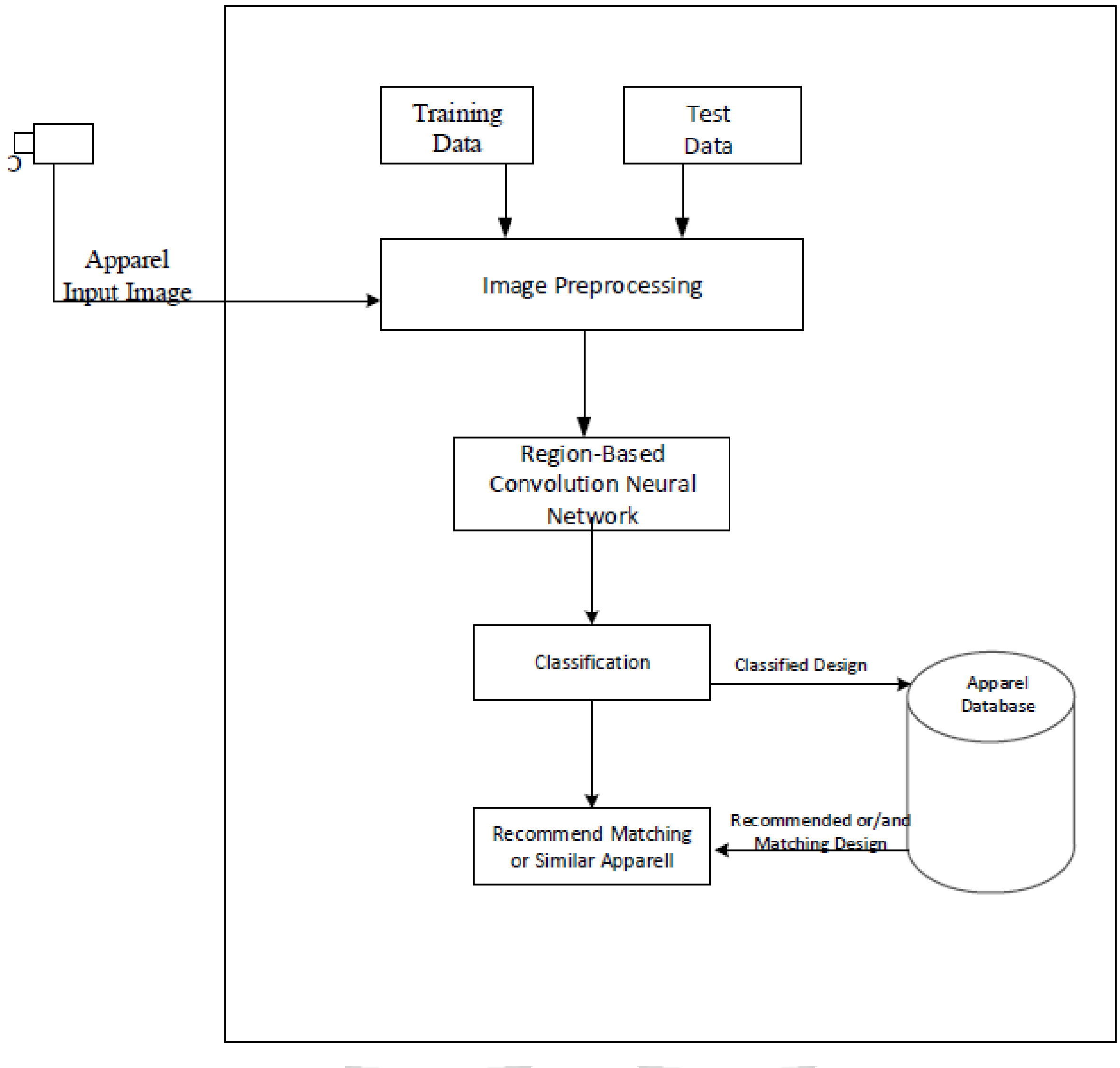


Figure 4. 1: System Architecture

4.4. Use Case Diagram

A use case diagram in the UML diagram represents the users interacting with the system. The system proposed consists of three main actors: Developer whose main role is to train the model, test the model and carry out maintenance. The developer also can include other apparel classification in the form of a system update. The system administrator is responsible for managing the users and apparel design to be advertised in the system. Finally, the merchant can receive information about trends analyzed by the system in form text or email reports. Figure 4.2 summarizes the concept.

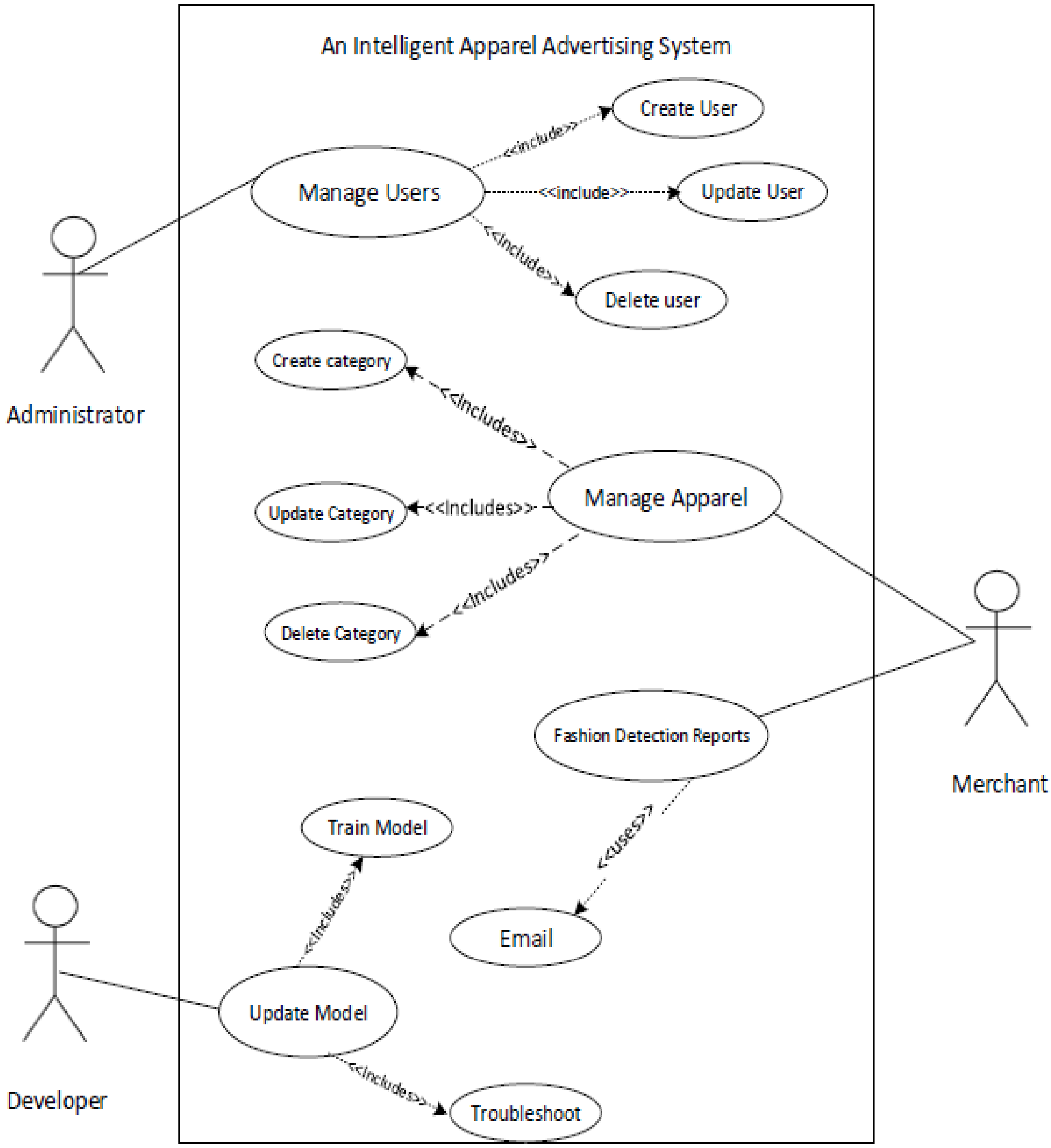


Figure 4. 2: Use Case Diagram

Use Case: Merchant Registration	
Primary Actor: Administrator	
Precondition: Administrator Panel Activated - Login	
Postcondition: Administrator Panel Deactivated -Log off	
Main Success Scenario	
Actor responsibility	**System responsibility**
1. Administrator enters the details of the Merchant	
	2. The system captures the information entered
	3. The system validates if the user information is entered correctly
	4. The system saves the information in the database
5. The administrator exits the panel.	

4.5. Data Flow Diagram

Dataflow is an illustration of representing the flow of data in a system. The DFD provides information in terms of inputs and inputs of each entity in the system, additionally, it captures the storage of data from the process. In this system, the diagram will outline entities such as the user of the system

4.5.1. Context Diagram

The context diagram represents the overview of the data flow at a high level between processes and entities in the system. This gives a diagrammatic overview of data flow in the system. Figure 4.3 depicts the context diagram of the desired system.

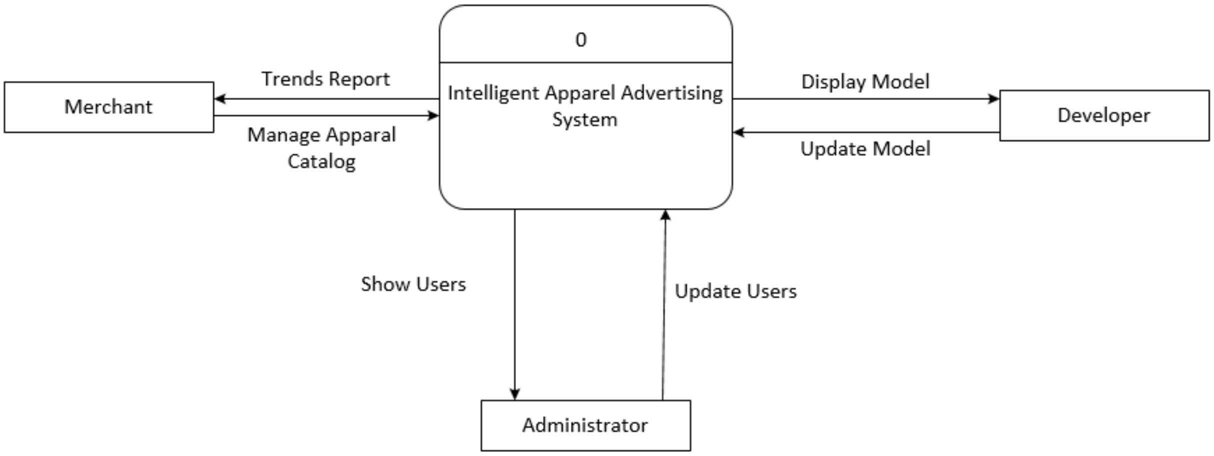


Figure 4. 3: Context Diagram

4.5.1.1. Level 0 Diagram

Generally, entities, data stores, and processes are clearly illustrated using a context diagram. Level 0 goes a notch further to dissect the flow of information to support the context diagram which further provides a clear understanding of the data flow process.

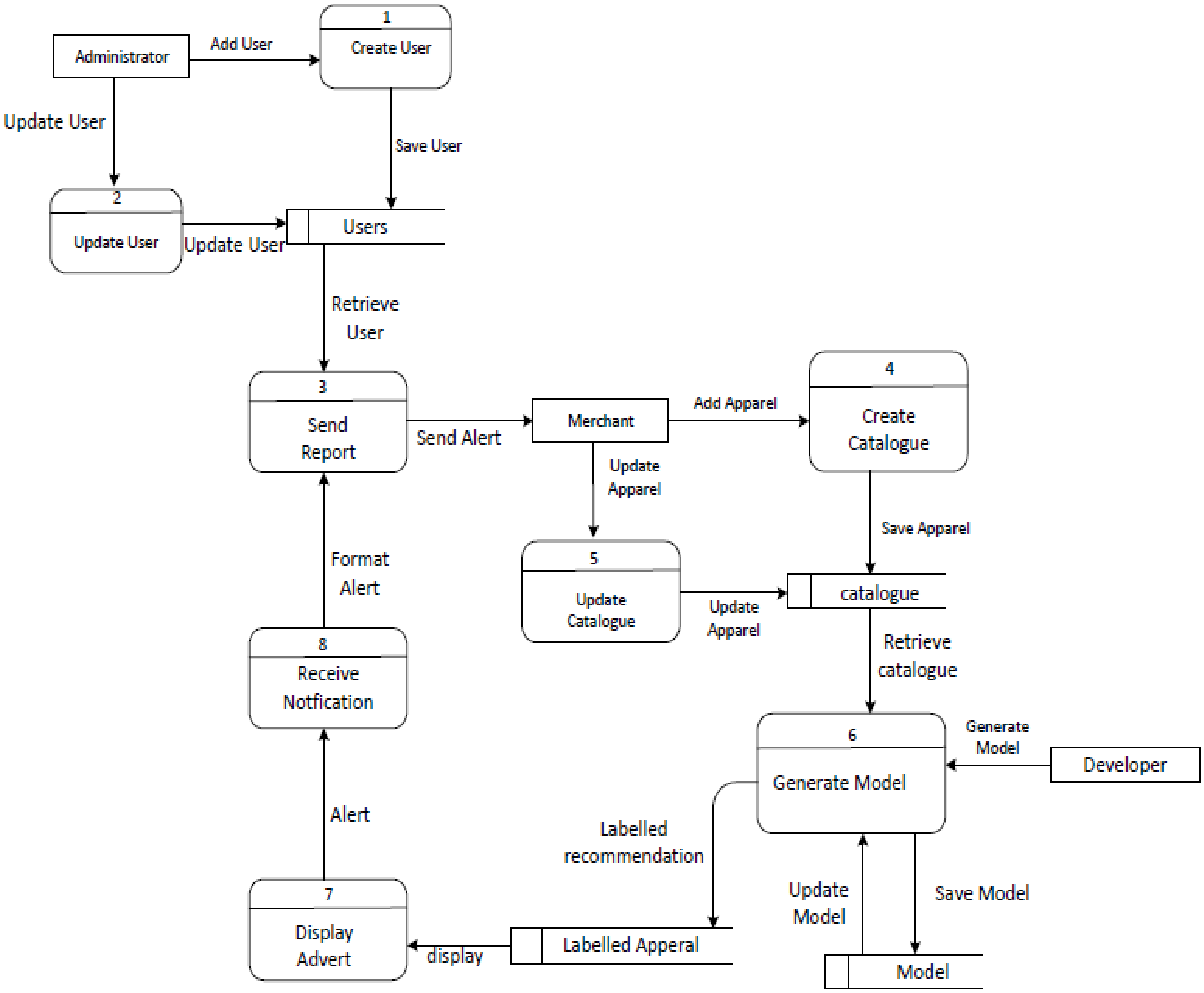


Figure 4. 4: Level 0 Diagram

4.5.1.2. Sequence Diagram

Sequence diagrams illustrate an exhaustive way in which messages are exchanged between objects over time in the system. The diagrams breakdowns a series of steps to show how a system achieves an objective. Figure 4.5 illustrated the steps taken by an intelligent image processing model for context-aware digital signage in creating an apparel advertisement. The model is initialized by the developer who trains it using a series of training dataset images of men's apparel limited to the upper torso. In Figure 4.5, the images are obtained from a live feed camera mounted on digital signage and stored in a defined frame format. The images are then preprocessed to conform to required specifications, after which features are extracted from the image in a process known as feature extraction. The model classifies the apparel image into predefined labels. The model then recommends an apparel design, color, and size based on the input features from the camera live feed.

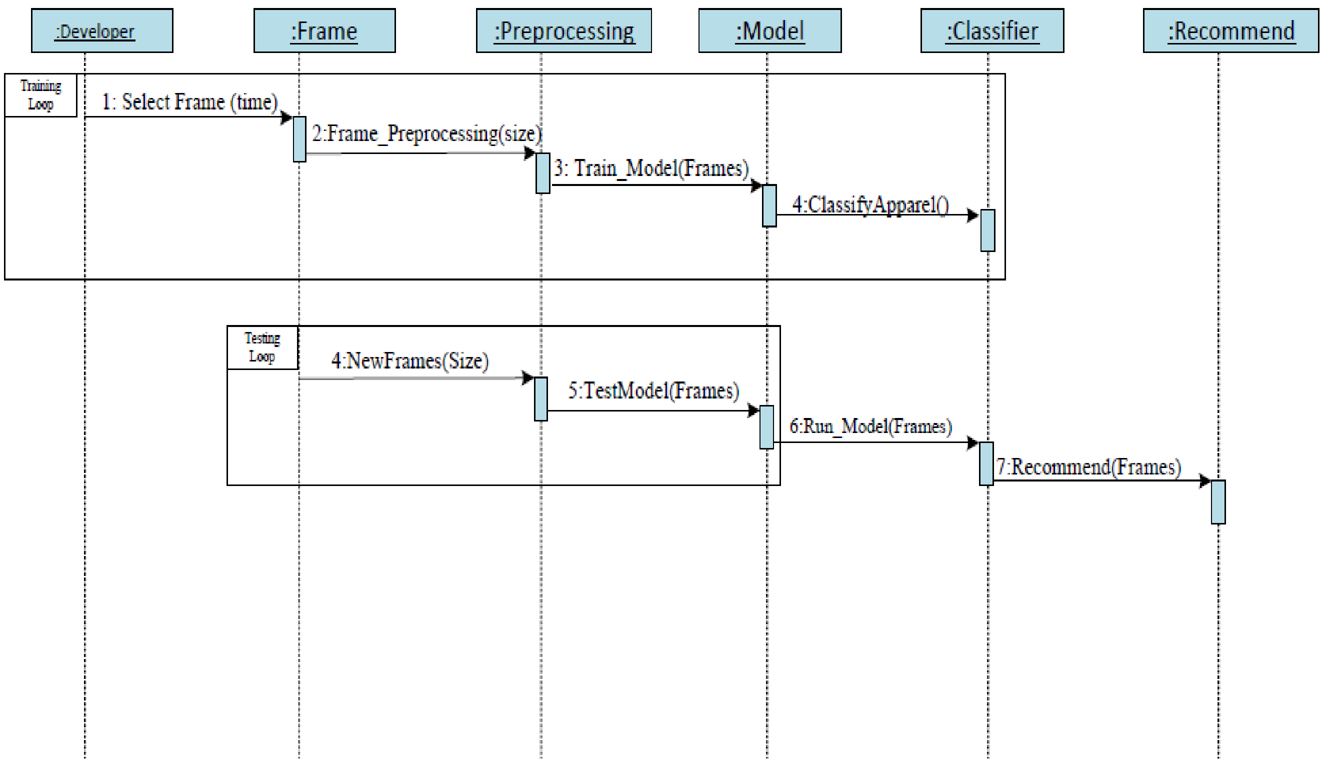


Figure 4. 5: Sequence Diagram

4.6. Class Diagram

Class diagrams deliver an overview perspective of the structure of the model, that is, the classes in the model and the interactions between these components. Figure 4.6 exemplifies the structure for developers' better understanding of the system and better knowledge transfer.

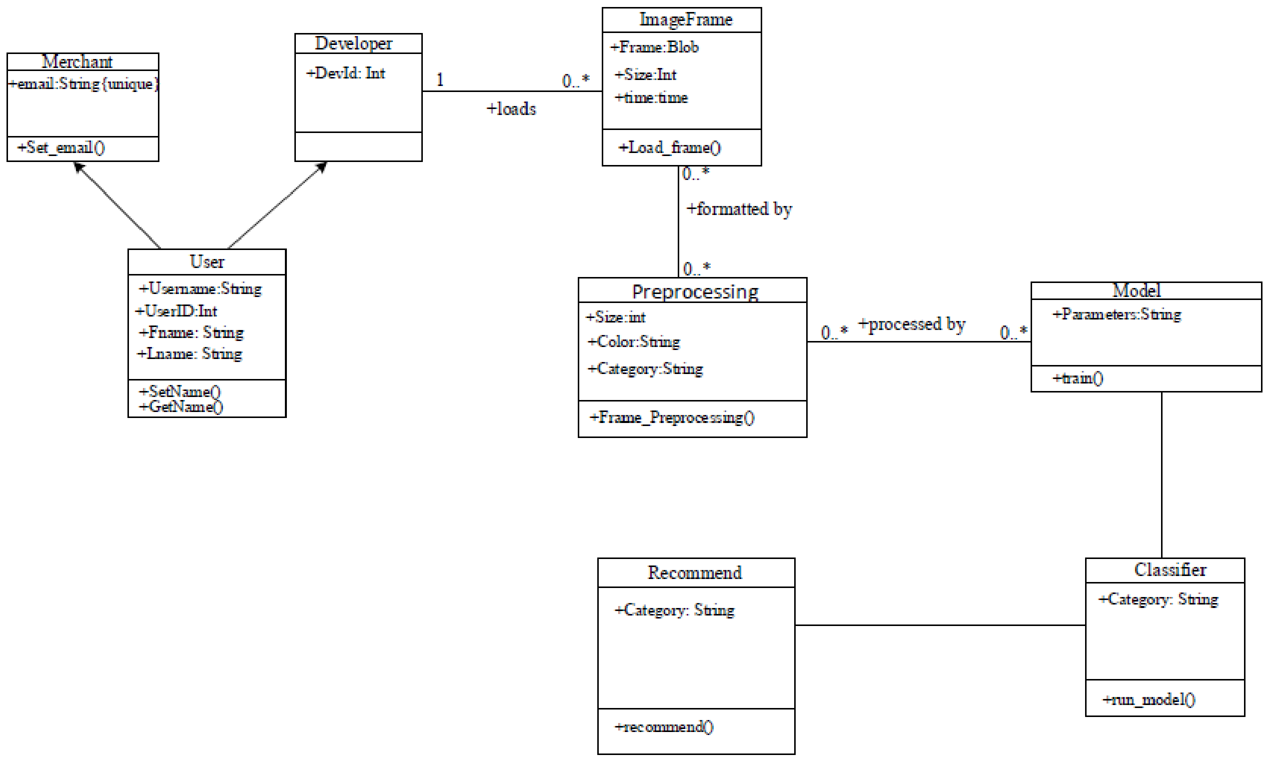


Figure 4. 6: Class Diagram

4.7. Database Schema

A database schema shows the various entities in a system and the relationships that exist among them, it also highlights the associations between these entities. The schema shows the constraints integrated into the system. Figure 4.7 reveals the database schema for the Intelligent image processing system for digital signage.

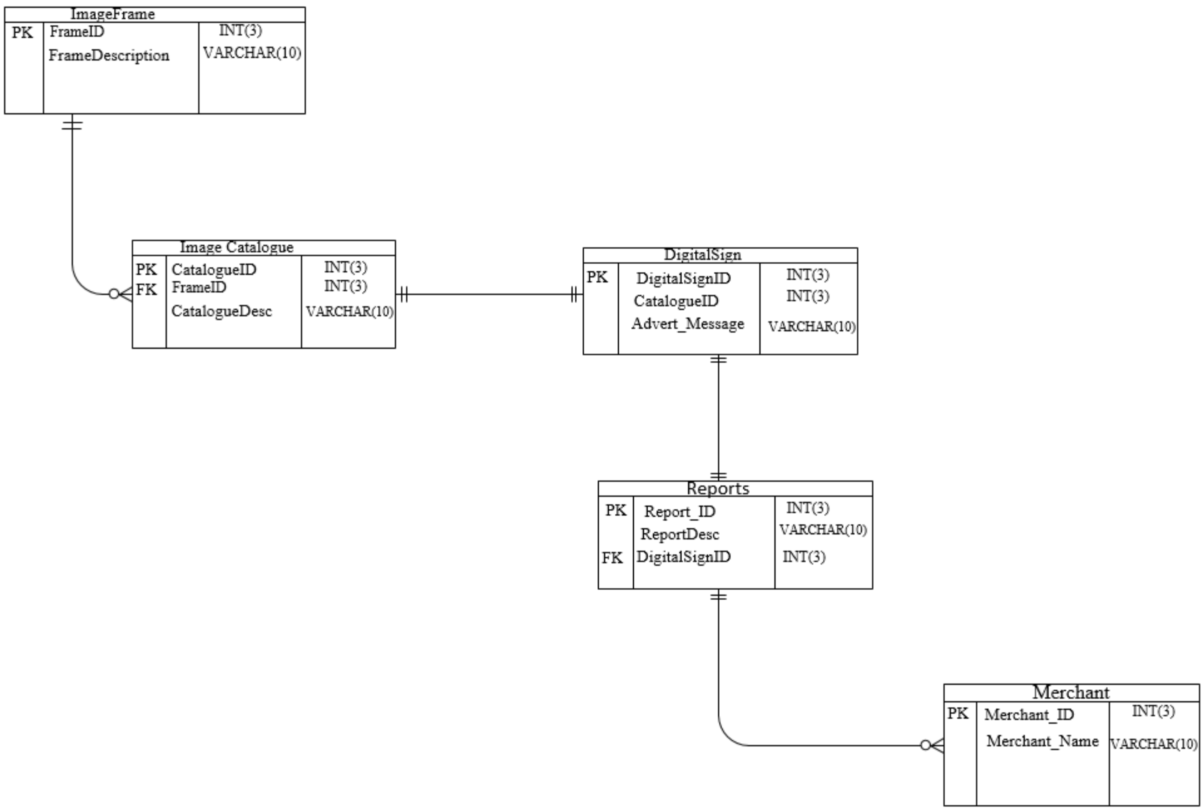


Figure 4. 7: Database Schema

Chapter Five: Implementation and Testing

5.1. Introduction

The implementation and testing chapter incorporates a series of steps that highlight how the model is delivered and tested. The implementation of the model was achieved through the following activities. The initial step, before developing the model was to determine the feature selection to determine the attributes required for the study target class, which is a recommendation of apparel for advertising. The next step was image preprocessing which entailed the conversion of the apparel image to the desired predefined format required for the model and successive annotation of the images into specific classes. The transformation involves converting the image from its original size into a 28 by 28-pixel image thereafter converting it into grayscale. The grayscale image is passed into a function that eliminates the foreground and the background of the apparel image. After image preprocessing, the image is passed to R-CNN as input and returns a label with a classified image. Based on the apparel classification class, the model then provides a recommendation. The intelligent image processing model for context-aware digital signage in cooperates detection of apparel images, correctly classifying the image and recommending another apparel on the screen of a digital signage. The study takes note of all past research on apparel image classification. The data from Kaggle, Fashion Product dataset, was normalized and split into a training set, and validation set. The training dataset was fed into a neural network which is elaborated in the subsequent sections.

5.2. Hardware and Software Environment

The model development and experiments were carried out in the Cloud platform. However, in a real-world environment, latencies in uploading the image to a cloud platform and obtaining a recommendation would take a longer time. Increased response time reduces the effectiveness of the model in terms of advertising. A could environment is preferred due to the high demand for computing resources, CPU, GPU, and RAM. Python was the underlying development language with extensions machine learning libraries such as TensorFlow (Abadi et al., 2016), Scikit-learn (Pedregosa et al., 2011), Numpy (Colbert, Varoquaux, & Walt, 2011), Keras (Chollet et al., 2015). The hardware and software environment components are summarized in Table 5.1

Table 5. 1 Hardware and Software configuration

Software	Specific Library	Version
Python (3.7)	TensorFlow	2.1.0
	Keras	2.3.0
	Numpy	1.16.3
	OpenCV	4.0.0
	Description	**Version**
Hardware	GPU	AMD Ryzen 7 M265 8GB
	CPU	Intel Broadwell
	RAM	16GB

5.3. Model Components

The model proposed is composed of two major components: the apparel image classification component and the recommendation part. These are discussed in the sub-sections.

5.3.1. Dataset Preprocessing

A total dataset from Kaggle contained 6206 images of different fashion component scrapped from an e-commerce site. The study focuses on men apparel, specifically the upper torso that is Shirt, T-shirt. The study further narrows down to Shirts and T-shirts which are commonly worn by men on the upper torso. Therefore, the images were reduced to 1210 images that focused on the men's upper torso to be used for the study. Additionally, the images were divided into the training and validation set as 840 and 370 respectively. The Logitech C920 captures images in high resolution which are converted to reduce the size. This process was achieved using the python function on Listing 5.1.

```
// Convert apparel images to 150 by 150 pixels
model = tf.keras.Sequential([
    tf.keras.layers.Flatten(input_shape=(150, 150, 1)),
    tf.keras.layers.Dense(128, activation=tf.nn.relu),
    tf.keras.layers.Dense(10)
])
```

Listing 5. 1 Convert to 150 by 150 pixel

5.3.2. **Conversion from RGB to Grayscale**

The captured images are in full RGB color, the images were converted to grayscale with a python function that accepts the image name and path as input and outputs a converted image. OpenCV library was used to achieve this in code listing 5.2

```
// Convert image to grayscale
import cv2
def image_source(name=none,  path=none):
   image = cv2.imread('name+path')
         gray = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
return gray
```

Listing 5. 2 Conversion from RGB to Grayscale

5.4. Model Configuration

A subsequent set of labels were generated for each set where labels 1 assigned to a shirt and 2 assigned to a T-shirt for classification. Additionally, from class generation, the apparel images were enumerated with categorical classes to encode and retain the meaning of classification of an image of apparel. The correctly classified were then encoded with an apparel label recommendation class that will then be displayed to the customer. The images were split into two different clusters. One set was marked for one class use, which is to identify only one apparel. In the second class, the second cluster of images was marked with two classes, which is a shirt and a T-shirt in the same image frame. This aimed to provide the comparative in the two categories of upper torso apparel scenarios. The training feature set, training dataset, and testing labels dimension have been highlighted in Table 5.2.

Table 5. 2 One class Configuration

Dataset	Dimensions	Class Configuration
Training	(840, 150,150)	1
Validation	(370, 150,150)	1

Table 5. 3 Two class configuration

Dataset	Dimensions	Class Configuration
Training	(840, 150,150)	2
Validation	(370, 150,150)	2

The image annotation was done using an opensource tool, VGG image annotator tool version 2.0.5 which generated a JSON file with the desired class configuration. The image in that dataset was annotated to illustrate the different classes in the study. Figures 5.1 and 5.2 represents images used in the different annotations used for the one and two class configurations.

Figure 5. 1 Single apparel image per frame

The one class configuration represents a single apparel detected in a single image frame. The image frame contains a single frame with only one apparel detected in this image. Partial frames, in this case, are eliminated and the focus is on the apparel image that occupies most of the image frame

Figure 5. 2 Multiple apparels image per frame

In the two-class scenario, more than one apparel image appears on the image frame, this represents a variation in the study where two people walking together past the digital signage. In these cases, the model can detect both images and recommend and advertisement to both parties. The two-class configuration would be the default working configuration. The resulting JSON file contains a unique identifier to each of the images with the details respectively. Listing 5.3 shows the JSON file for the two-class configuration.


```
{"shirt.jpg24562":{"filename":"shirt.jpg","size":24562,"regions":
[{"shape_attributes":{"name":"rect","x":89,"y":111,"width":281,"height":304},
"region_attributes":{"Apparel":"Shirt"}}],"file_attributes":{}}
,"T-shirt.jpg26889":{"filename":"T-shirt.jpg","size":26889,"regions":
[{"shape_attributes":{"name":"rect","x":35,"y":145,"width":289,"height":337},
"region_attributes":{"Apparel":"T-shirt"}}],"file_attributes":{}},
"2 Tshirt.jpg21310":{"filename":"2 Tshirt.jpg","size":21310,"regions":
[{"shape_attributes":{"name":"rect","x":519,"y":100,
"width":501,"height":660},"region_attributes":{"Apparel":"T-shirt"}},
{"shape_attributes":{"name":"rect","x":5,"y":87,"width":501,"height":666},
"region_attributes":{"Apparel":"T-shirt"}}],"file_attributes":{}}}
```


Listing 5. 3 Two Class JSON Configuration Output

5.5. R-CNN Model

R-CNN offers improved image classification reducing the reading of blank spaces and focus on the region with value. The region proposal in this study represent areas within which apparel appears in the frame. The study did not re-configure the hidden layers of the model or the activation function used in the R-CNN algorithm. The configuration was to maintain a baseline in the different configured classes of the study. The developed model was trained under these two configurations.

5.5.1. One class configuration

In this class, the labeled dataset was annotated with a label that represents one apparel within the image frame. The image was then resized, converted to grayscale and eliminate the background as a separate entity. The aim was to identify one apparel in one picture. The code configuration for defining different classed used for different parameters. The code Listing 5.4 is summarized as; define the name of the class, the type of processor and steps per epoch for a training class and finally the detection minimum confidence.

```
class Baseline(Config):
name = "singleFrame"
image_GPU = 1
num_class = 1 + 1
steps_epoch = 100
detection_conf = 0.9
```

Listing 5. 4 Baseline instance

The images were loaded using a function that derived the annotation of the x and y coordinates derived from the VGG tool. These annotations focused on the upper torso area covered by apparel, in this case, a shirt or t-shirt. The function takes the JSON file as input to access the different annotation of the regions with specific labels

5.5.2. Two Class configuration

The class configuration represents the practical application of the model where multiple apparels are captured in one single video frame. The images contain different people with different poses and walking style. The most distinct position is an upright walking position visible to the camera mounted on the digital signage. This serves as the code configuration for two people in a frame. The people may be wearing two distinct apparel or similar apparel. The code for two-class configurations is the same as a single class, the only difference in the total number of classes represented since the background is considered as a class on its own.

In Listing 5.5, the images are then passed on a model that can detect and out a labeled item. Training is done bey calling a function that will feed the training dataset into the model, the model learns to associate the images with labels and epoch define the number of full iterations of the training.

```
#Training the whole network for 5 epochs first
checkpoint_callback = ModelCheckpoint('./models/vgg_weights_best_pattern.hdf5', monitor='val_acc', verbose=0, save_best_only=True, save_weights_only=False, mode='auto', period=1)
tf_model.fit_generator(
        train_generator,
        samples_per_epoch = nb_train_samples,
        nb_epoch = 10,
        validation_data = validation_generator,
        nb_val_samples = nb_validation_samples,
        verbose = 1,
        initial_epoch = 5,
        callbacks=[checkpoint_callback]
)
```

Listing 5. 5 Training set

The recommendation module is based on four classes. These classes are comprised of where the apparel covers the body, the apparel pattern, color, and style. The recommendation function is based on the output of the correctly classified apparel image. Essentially, the image classified from either one class or the two-class configuration returns a label with attributes such as color and design. Based on the attributes label the model can return an image of apparel in stock for advertising. As we train the classification module, we freeze the recommendation module then unfreeze it after the validation accuracy stabilizes. The code Listing 5.6 illustrates the apparel recommendation for advertisement display

```
#Initialization for recommendation
class_part = {
    0: 'FullSleeve',
    1: 'HalfSleeve',}
class_pattern = {
    0: 'Plain',
```

```
    1: 'Striped'}
class_color = {
    0: 'Black',
    1: 'Blue',
    2: 'Green',
    3: 'Red',
    4: 'White'}
class_style = {
    0: 'Casual',
    1: 'Formal',}
```

Listing 5. 6: Recommendation instance

The apparel currently in stock is loaded into the model. The images are coded under these four classes to enable co-relation between the classified images and recommendation. The model was trained with 10 epochs with 100 steps per epoch.

5.6. System implementation

5.6.1. Trends Notification

The model works by capturing images of apparel as potential customers walk past the digital signage and display a corresponding advertisement of similar or different apparel matching the apparel of the customer. The result is for a merchant to be able to review all the trends captured within a specific period. A merchant is the only contact with the system which is set to autogenerate or promptly generate on command.

5.6.2. System Testing

The model objective was providing context-aware digital signage for advertising in a multi-object environment. The object, in this case, is apparels that appear on an image frame using a real-time camera. These images are captured analyzed for apparel classified correctly and an apparel recommended. The recommended apparel is displayed on the screen of a digital signage as an advertisement. The test case was generated to depict the performance of the system, these tests seek to validate the system’s function. Table 5.3 summaries the tests

Table 5. 4: Test cases

Test case	Importance Level	Results
Does the model capture image?	High	The model captures images using a camera
Does the model classify the apparel images?	High	The model analyzes the image captured and classifies them
Does the model display an apparel recommendation?	High	A recommended apparel is printed
Does the model generate a trends report?	High	A report is generated

Chapter Six: Discussions

6.1. Introduction

The study seeks to develop an intelligent image processing system for apparel detection and recommendation on digital signage. This assumes correctly capturing apparel images from a camera, classifying the apparel image and recommending an appropriate apparel image as an advertisement on digital signage. The last decade has seen the increase of internet usage (Satisa, 2019) with products advertising growing at 16% per quarter (PWC, 2018). The research reviews the works of (Subbarayudu & Vardhan, 2015) on clothing pattern recognition using SVM and compare the features of SVM to our study. Additionally, the research reviews SVM efficiency in apparel detection in terms of accuracy and precision. According to Schindler (2018), convolution neural network (CNN) presents promising results in apparel classification and recommendation in meta-data enrichment for e-commerce applications and therefore we review their contribution in this study. The accuracy of the model developed is evaluated based on how it handles natural environment factors such as light, and human-based factors such as height, body poster, and gait (Russakovsky & Deng, 2015). This chapter highlights the different evaluation metrics used to select the best model suited for context-aware digital sign in the case of apparel advertisement. Finally based on the metrics we conclude on the model performance against previous studies.

6.2. Model Validation

Model Validation involves the process of accessing the effectiveness of a model with reference to measurement metrics (Philip & Hardin, 2012). The validation aims to enable researches differentiated between the different model performance outputs and accuracy (Barlas, 2016). The research compares various image classifiers to determine the best training model for apparel classification

6.2.1. Support Vector Machine (SVM)

According to Subbarayudu (2015), SVM provided an overall 91% accuracy in apparel pattern recognition for visual impaired person and therefore the study reviews the inner workings of SVM to achieve such results. Scikit-Learn API supports two support vector models that are linear and nonlinear. The study achieves great insight in apparel pattern classification in a controlled setting with three classes: apparel, color and pattern. The study utilizes both linear and nonlinear for training and testing feature set of dimensions [batch * length* width] that is reshaped to [1410,1500]. The model findings are presented in Table 6.1 and 6.2

Table 6. 1 Linear SVM Outcomes Summary

Metric	**Value (%)**
Accuracy	90.47
Precision	90.08
Recall	91.00

Table 6. 2 Non-Linear SVM Outcome Summary

Metric	**Value (%)**
Accuracy	87.35
Precision	99.08
Recall	69.96

6.2.2. **Convolution Neural Network (CNN)**

In their study on Fashion and apparel classification, (Schindler, Lidy & Kerner, 2018) focused on classification accuracy in the task of differentiating products based on gender and category for e-commerce advertising. CNN pioneered the research in computer visual imagery (Tsantekidis & Avraam, 2017) but has seen improvement in terms of processing efficiency. The major drawback in this study was on the algorithm, that is, CNN iterative process that covers white spaces instead on focus on image region in detection. The Convolution Neural Network (CNN) API was modeled using Keras API with the following specifications: convolution layer 150 by 150 input stream image passed onto the network layer with a kernel of size 6 by 6 and 2 by 2 stride size. The final pooling layer set to 2 by 2 max-pooling layer. The ensuing convolution layer, a kernel of size 5 by 5 was used and active the rectified linear units (RELU's) Two connected layers then succeeded, first one activated by RELU and second by SoftMax functions. The code Listing 6.0 illustrates the process for 5 epochs. The subsequent results for the CNN model based on the Keras API build are outlined in Table 6.3

Table 6. 3 CNN Outcome Summary

Metrics	**Value (%)**
Accuracy	92.96
Precision	94.08
Recall	54.26

6.2.3. Region Convolution Neural Network (R-CNN)

R-CNN is geared towards detecting semantic objects of a class based on certain domain in each image (Shaoqing & Ren, 2015). The study utilizes R-CNN as opposed to CNN because R-CNN addresses the shortcomings of CNN in object detection which is vital to the study. Apparel classifiers are optimized by region-based object detection as opposed to CNN which rests towards class object detection (Guan, Yifeng & Yuan, 2012). The model loads a parsed file and determines a bounding box for each object in the image, the model then extracts image dimensions and extract details from the annotation file. The accuracy of detection is improved by generating a region box around the apparel area then passed for detection. The model was trained for 5 epochs and the outcomes summarized in Table 6.4

Table 6. 4 R-CNN Outcome Summary

Metric	**Value (%)**
Accuracy	98.46
Precision	97.08
Recall	40.26

6.3. R-CNN Cost Function

The cost function serves as a marker that seeks to evaluate the model by measuring the inconsistency between predicted values and real values (Christian & Cramer, 2009). The cost function must accurately extract all features of the model down to a single digit in such a manner that an improvement in that digit is a sign of a better model (Feng & Song, 2012). The scalar value allows the model to be ranked and compared. The study uses cross-entropy to measure R-CNN in the difference between correctly classified apparel of dataset provided to the model. Figure 6.1

shows a training cost function and testing function against the number of epochs used in the study of one class configuration model.

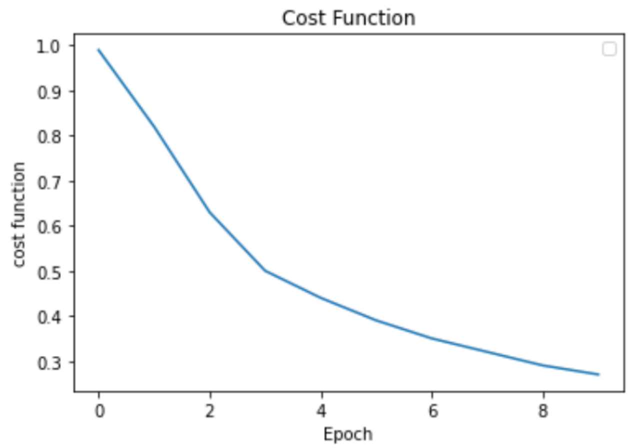


Figure 6. 1 R-CNN Cost function against Epoch

The results from Figure 6.2 seek to check for overfitting, the assessment of the training loss towards the testing loss function implemented to evaluate the model. Additionally, it was noted that the model precision improved with the number of iterations. The accuracy figures show that the model had not yet over-learned the training dataset hence comparison shows increased performance in the validation as referenced in Figure 6.2

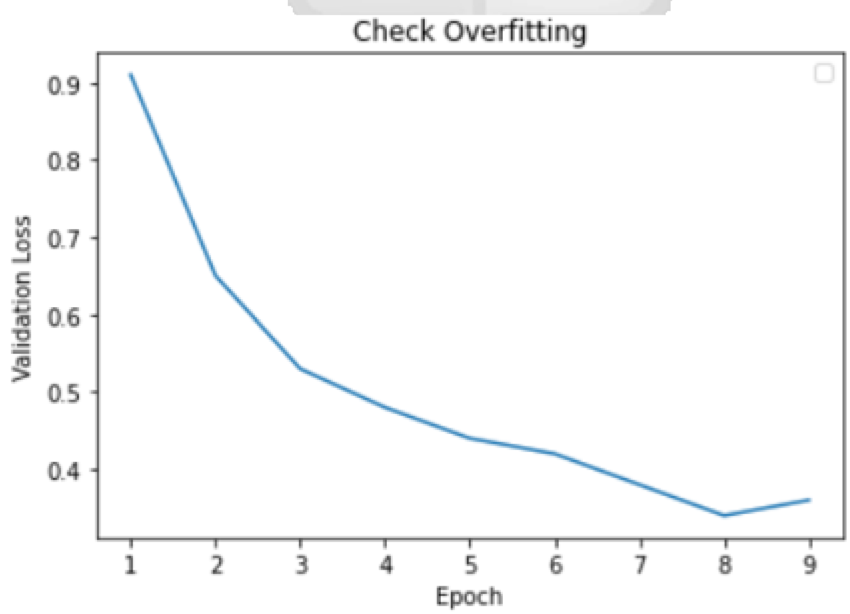


Figure 6. 2 Overfitting

6.4. Confusion Matrix

Validation of the Intelligent image processing system for context-aware digital was evaluated after the training process. The execution metrics of apparel detection, classification and recommendation were based on accuracy, precision, and recall. In the study, a function was created to implement this feature. The input is predicted and actual labels and computes the metrics based on equation 6.1

$$\sum_{i=1}^{m} \left(\hat{y}^{(i)} - y^{(i)}\right)$$

Equation 6.1: Confusion Matrix

The formula computes the sum difference between the predicted $(\check{Y}^{(i)})$ and actual values $(Y^{(i)})$. The subsequent error is backpropagated and network weights and bias updated using gradient descent rule.

The code Listing 6.1 for the subject formula is as follows

```
def confusion_matrix(actual,predicted):
    tn,fp,fn,tp = 0,0,0,0
    for a,p in zip(actual,predicted):
    if np.argmax(a) == 1:
    if np.argmax(a) == np.argmax(p):
    tp+=1 # truly positive actual-predicted
    else:
    fp+=1
    else:
    if np.argmax(a) == np.argmax(p):
    tn+=1
    else:
    fn+=1 #falsely negative
return tn,fp,fn,tp
```

Listing 6. 1 Confusion Matrix

6.5. Model Selection

In the subsequent experiments carried out during the study, the best performing classification model was chosen to fit real-world environment implementation. The main feature of the best model was based on correctly classifying and recommending apparel with the highest precision, accuracy. In addition, the model would need to exhibit pixel-level stability since a camera is the main source of input data to the model. The two best models which exhibited these properties were: CNN and R-CNN. With reference to these, the latter had a precision of 98.04% while the former 97.08% and accuracy of 92.96% and 98.46% respectively. R-CNN was selected as the model algorithm for the intelligent image processing model for context-aware digital signage.

6.6. Real-time Classification

Fundamentally, for a precise and responsive context-aware digital signage, the model ought to support real-time analysis of captured images from a video stream.

The model was appraised on its capability to provide real-time advertisement of apparel based on a live video stream. This was achieved with the use of sample videos with potential customers wearing two sets of different apparel walking past the digital sign. Table 6.6 shows the performance of the model in different video timeframes.

Table 6. 5 Real-time analysis 2 Second video stream

Video Duration (Seconds)	**Time to Classify (Seconds)**	**Generation Time (Seconds)**
25	320	2
12	270	2
60	1357	2
97	**1947**	
Average analysis time	**20 Seconds**	

Table 6.5 shows the model that captures an image from the video stream every two seconds and analyses each image for apparel then provides an advertisement as a recommendation. When an image is generated every two seconds, the model takes 20 seconds to detect apparel and display an advertisement to a potential customer.

Conspicuously, this is entirely slow for a digital sign since an average person with a height of 5ft 7 inches and 28 inches stride takes 6.75 seconds to walk 5 meters. Five meters is the average distance a person would take to walk past a digital sign. As a result of this challenge, the model was optimized to generate images for analysis after 5 seconds rather than 2 seconds. Notably, this adjustment would reduce the number of displayed advertisements. Table 6.7 illustrated a 5-second apparel detection against time.

Table 6. 6 Real-time analysis 5 Second video stream

Video Duration (Seconds)	Time to Classify (Seconds)	Generation Time (Seconds)
25	96	5
12	64	5
60	298	5
97	**458**	
Average analysis time	**4.7 Seconds**	

The additional time in the generation of images from the digital sign video stream leads to a decrease in the number of images classified but also reduces the time taken for analysis and recommending an advertisement. The reduction in time improves the likelihood of a customer viewing the advertisement. This is an acceptable loss since there are minimal changes in a real-time environment in a time frame of between two and five seconds.

6.7. The validity of the Proposed Model

Digital Signage research has mainly focused on converting noticeboards from a physical static platform to online dynamic boards with enhanced capabilities that include cameras, microphones, and high-resolution display screens. Studies in the digital signage space focused on improving the content displayed and how effectively and easily capture the attention of a customer. The proposed model aims to improve the agility of digital signage by incorporating a context-aware module to proactively create targeted advertisements.

In our studies, we focus on the fundamental human need which is clothing. The model receives captured apparel images from a video stream mounted on top of a digital sign, classifies the image and recommends a matching or similar apparel to the customer within 4.7 seconds. This considers the minimum amount of time required for an average customer to walk past the digital sign with ample time to view the display.

Additionally, the model can summarize the captured trends which provide an entrepreneur with an advantage in stocking apparel designs and understanding the target population. The model is aimed to capture the knowledge and experience of creative digital marketing with context awareness to ensure it is dynamic. Lastly, the system does not invade the personal space of a customer to collect personal data in a bid to offer personalized advertising but utilizes the current context to provide the same service.

Chapter Seven: Conclusions and Recommendations

7.1. Conclusions

The primary objective of this study was to develop a machine learning model to improve the content displayed in digital signs. The decision by an entrepreneur to invest in a digital signage platform is based on the ability to provide relevant content. An entrepreneur needs to constantly develop relevant content to keep up with emerging trends. The study mainly focused on apparel advertisement based digital signage current context. The study focuses on intelligently displaying advertisements based on observed apparel features from a potential customer walking past the digital signage.

The researcher sub-divided the main goal into sub-goals by examining existing literature on digital signage architecture, apparel image processing methodologies, application of machine learning tools on apparel advertisement and recommendation and finally a set of learning algorithms techniques that are optimistic in the research field apparel image processing.

Several early image processing algorithms were developed, and their outcomes used as a benchmark measurement of precision, accuracy, and recall in the current study. The aim was to develop a validation method for the proposed model and evaluate how various algorithms faired against it. In the end, several tests were carried out to determine the best model for apparel feature extraction and recommendation.

After all experiments, Region-Based Convolution Neural Network (R-CNN), a region-based type of CNN outperformed most of the classical models in image detection. The main advantage of R-CNN was based on the algorithms' ability to create boundaries within different areas of the human torso for better analysis. Considering this, we can, therefore, segment various parts of the body covered in different types of apparel with the aim of maximizing accuracy. A good accuracy score of feature extraction from an apparel image will lead to an improved recommendation score.

7.2. Challenges

The researcher was able to achieve the outline objectives, however, quite a significant amount of challenges was encountered during the study. It should be noted that the challenges did not affect the outcome of the study to a greater extent. First, there is not locally data repository for local apparel. Most fashion databases collected by big cooperation such as Google cannot be easily customized to fit the local setting, that is African apparel such as kitenge. Most data, in this case, are from developed countries which might not accurately fit the local population without retraining.

Secondly, many apparel stores do not have data on apparel sales or unwilling to share such information for the purpose of this research due to privacy concerns. The data would have been vital as a starting point to under the type and volumes of apparel largely consumed in Kenya. Thirdly, it is quite a challenge to replicate an apparel advertising method; most apparel advertisements are done using a prominent personality who wears the apparel and display its features. In this study, a display of just apparel without a human or humanoid was not appealing to potential customers. Most customers are used to see fashion advertisements of a humanoid with different sizes such as small, medium and various types of large. The design presents a level of association that was difficult to achieve in the recommendation phase of the model.

Lastly, image processing in machine learning is carried out mostly by the graphics processing unit (GPU). Camera feeds the input images into the machines; the GPU acts the computer optical nerve that processes the input from the camera. In contrast to Moore's law of double processing power every two years, GPU specification has remained expensive and limited to desktop and high-end laptop machines. Additionally, cloud-based resources could not be used for this study due to the input stream for camera feed which would increase the latency in image frame uploads.

7.3. Recommendations

The proposed model is essentially recommended for shopping malls with dynamic naturally occurring elements. The model accuracy value of 0.9, it may give false positives leading to misleading detection and recommendation. Also, the model requires locally captured images to better train the model to achieve better context-awareness. This would result in a more intelligent model with local accepted apparel designs. Finally, the research requires high computing resources in terms of CPU and GPU. The research proposed the following recommendations.

(i) Train the model with more dynamic images obtained locally to reflect the people's culture in apparel designs, this will ensure the model can correctly detect and classify apparel images on local context and in turn provide an appealing recommendation in form of an advertisement.

(ii) Train the model to segment the human body into various parts covered by apparel. Thus, an entrepreneur can focus on their area of specialty in apparel sale. Consequently, the model can be trained to only detect and advertise shirts or trousers depending on the preference of the entrepreneur.

(iii) Digital signage architecture to include GPU processing abilities. The main hardware required for this model is computing resources in the form of a CPU. Future digital signage should in cooperate GPU capabilities to catapult it into the next generation of intelligent signage.

(iv) To facilitate a high rate of adoption, the project should be implemented liaising with entrepreneurs and fashion designers. This will help a higher rate of production knowledge and usage

7.4. Suggestion and Future works

The fashion industry is a multibillion-dollar enterprise with numerous unexplored components in design, detection, classification, and recommendation. The study instigated manual approach image labeling in different classes. This led to a limited number of images and classes used in the study. The researcher recommends an automated image annotation process to increase the scope.

Secondly, the study focused on male upper torso apparel as part of a proof of concept for this study. In the future, expanding the scope to cover the upper torso for males and both upper and lower torso for female. The model could also intelligently detect its environment such as weather, time and customer age group among others. Finally, for the purpose of future image processing studies in the fashion industry, a local data repository for apparel is required. This includes sales data, a digital catalog of such images and labeling. This will spur research in this field due to availability and access to such data.

References

AI, machine learning and digital signage (DSConnection) – Digital Signage Pulse. (2019, January 4). Retrieved from https://digitalsignagepulse.com/news/ai-machine-learning-and-digital-signage/

Alamanos, E., Brakus, J. J., & Dennis, C. (2014). How Digital Signage Affects Shoppers' In-Store Behavior: The Role of the Evoked Experience. *Ideas in Marketing: Finding the New and Polishing the Old*, 748-751. 10.1007/978-3-319-10951-0_276

Arsan, T., Parkan, A., & Konu, H. (2014). Design and Implementation of Remotely Managed Embedded Digital Signage System. *International Journal of Computer Science, Engineering and Applications*, *4*(3), 1-15.

Attitudes to Sponsored and Branded Content (Native Advertising). (n.d.). Retrieved from http://www.digitalnewsreport.org/essays/2015/attitudes-to-advertising/

Bauer, C., Garaus, M., Strauss, C., & Wagner, U. (2018). Research Directions for Digital Signage Systems in Retail. *Procedia Computer Science*, *141*, 503-506.

Bauer, C., & Spiekermann, S. (2011). Conceptualizing Context for Pervasive Advertising. *Pervasive Advertising*, 159-183.

Burger, W., & Burge, M. J. (2013). SIFT—Scale-Invariant Local Features. *Principles of Digital Image Processing*, 229-296.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273-297.

Deng, J., Dong, W., Socher, R., Li, L., Kai Li, & Li Fei-Fei. (2009). ImageNet: A large-scale hierarchical image database. *2009 IEEE Conference on Computer Vision and Pattern Recognition*.

Guan, L., Kung, S., & Hwang, J. (2000). Intelligent multimedia processing. *Multimedia Image and Video Processing*.

Feedforward neural networks 1. What is a feedforward neural network? (n.d.). Retrieved from http://www.fon.hum.uva.nl/praat/manual/Feedforward_neural_networks_1__What_is_a_feedforward_ne.html

Kauffman, G. B. (1999). Inorganic Chemistry, 2nd Ed. By Gary L. Miessler and Donald A. Tarr. Prentice-Hall: Upper Saddle River, NJ, 1999. Figures, tables, charts. xiv + 642 pp, 21.0 26.0 cm. $92.00. ISBN 0-13-841891-8 (Photo reproduced with permission from Prentice-Hall, Inc). *The Chemical Educator*, *4*(5), 200-200.

Ling Guan; Yifeng He; Sun-Yuan Kung (1 March 2012). Multimedia Image and Video Processing. CRC Press. pp. 331–. ISBN 978-1-4398-3087-1.

Luo, B., Cross, A., & Hancock, E. (1999). Corner detection via topographic analysis of vector-potential. *Pattern Recognition Letters*, *20*(6), 635-650.

Nodari, A., Ghiringhelli, M., Zamberletti, A., Vanetti, M., Albertini, S., & Gallo, I. (2012). A mobile visual search application for content-based image retrieval in the fashion domain. *2012 10th International Workshop on Content-Based Multimedia Indexing (CBMI).*

Number of smartphone users worldwide 2014-2020. (N.d.). Retrieved from https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/

Number of smartphone users worldwide 2014-2020. (N.d.). Retrieved from https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/

Ravnik, R., & Solina, F. (2013). Audience Measurement of Digital Signage: Quantitative Study in Real-World Environment Using Computer Vision. *Interacting with Computers*, *25* (3), 218-228.

Schaeffler, J. (2008). Who…s Right for Digital Signage? *Digital Signage*, 137-156.

Schaeffler, J. (2008). When and Where to Use Digital Signage? *Digital Signage*, 101-111.

Shaoqing, Ren (2015). "Faster R-CNN". Advances in Neural Information Processing Systems. arXiv:1506.01497. Liu, Wei (October 2016). SSD: Single shot

O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg.

L. Fei-Fei, "ImageNet Large Scale Visual Recognition Challenge," International Journal of Computer Vision (IJCV), 2015.

The Top 3 Free Digital Signage Software Options for Your Next Event. (2019, June 21). Retrieved from https://blog.capterra.com/7-free-and-open-source-digital-signage-software-options-for-your-next-event/

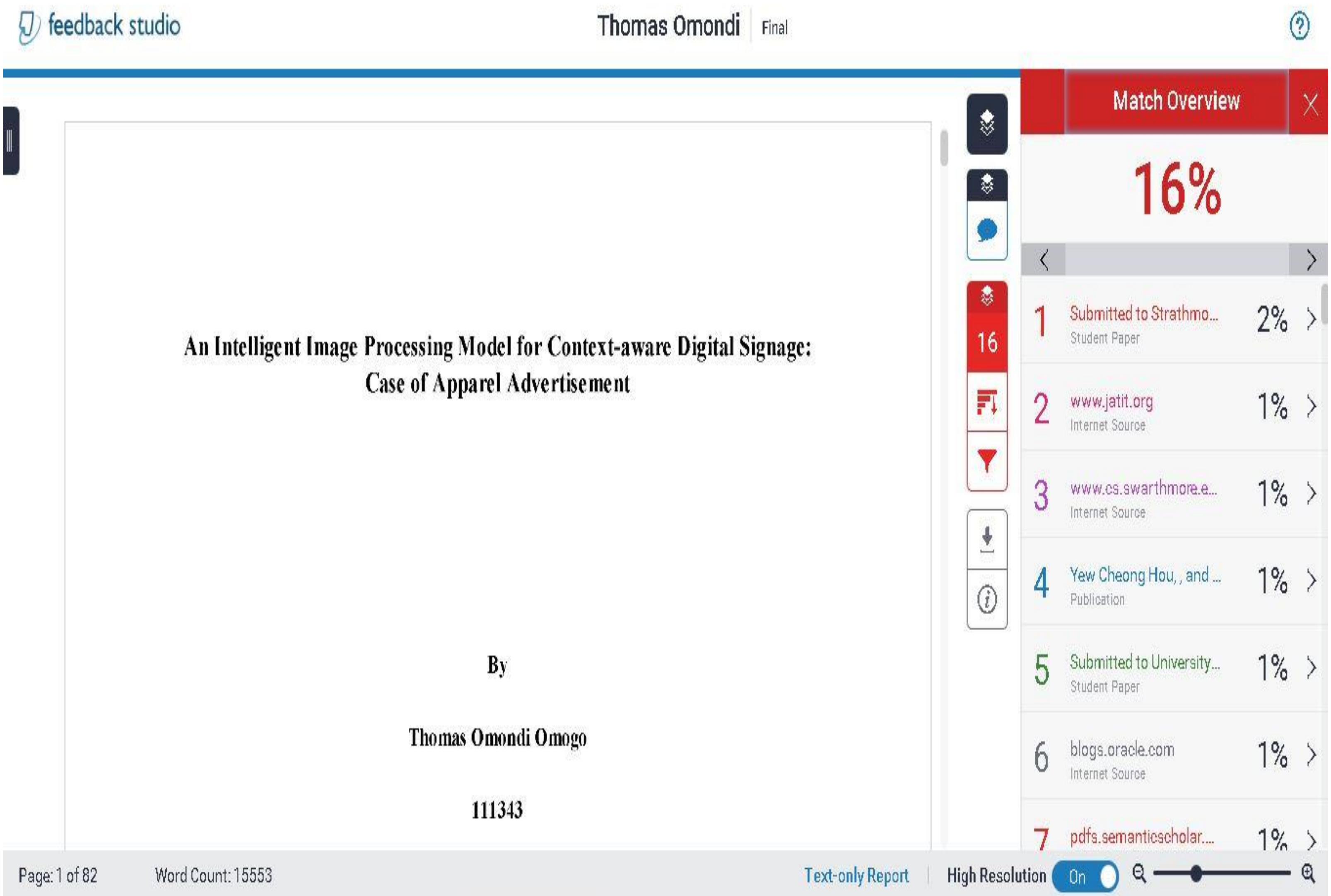
Wertime, K., & Fenwick, I. (2011). *DigiMarketing: The Essential Guide to New Media and Digital Marketing*. Hoboken, NJ: John Wiley & Sons.

Yu, C. H. (2008). Book Review: Creswell, J., & Plano Clark, V. (2007). Designing and Conducting Mixed Methods Research. Thousand Oaks, CA: Sage. *Organizational Research Methods*, *12*(4), 801-804.

Zhang, T., Xu, C., Zhu, G., Liu, S. and Lu, H., 2012. A Generic Framework for Video Annotation via Semi-Supervised Learning. IEEE Transactions on Multimedia, 14(4), pp.1206-1219.

Appendices

8.1. Appendix A: Plagiarism Report



8.2. Appendix B : Ethical Approval Letter

Strathmore
UNIVERSITY

3rd December 2019

Mr Omondi, Thomas
thomas.omondi@strathmore.edu

Dear Mr Omondi,

RE: An Intelligent Image Processing Model for Context-aware Digital Signage Case of Apparel Advertisement

This is to inform you that SU-IERC has reviewed and **approved** your above research proposal. Your application approval number is **SU-IERC0573/19**. The approval period is 3rd **December, 2019 to 2nd December, 2020.**

This approval is subject to compliance with the following requirements:

i. Only approved documents including (informed consents, study instruments, MTA) will be used
ii. All changes including (amendments, deviations, and violations) are submitted for review and approval by SU-IERC.
iii. Death and life threatening problems and serious adverse events or unexpected adverse events whether related or unrelated to the study must be reported to SU-IERC within 72 hours of notification
iv. Any changes, anticipated or otherwise that may increase the risks or affected safety or welfare of study participants and others or affect the integrity of the research must be reported to SU-IERC within 72 hours
v. Clearance for export of biological specimens must be obtained from relevant institutions.
vi. Submission of a request for renewal of approval at least 60 days prior to expiry of the approval period. Attach a comprehensive progress report to support the renewal.
vii. Submission of an executive summary report within 90 days upon completion of the study to SU-IERC.

Prior to commencing your study, you will be expected to obtain a research license from National Commission for Science, Technology and Innovation (NACOSTI) https://oris.nacosti.go.ke and also obtain other clearances needed.

Yours sincerely,

for: Dr Virginia Gichuru,
Secretary; SU-IERC

STRATHMORE UNIVERSITY INSTITUTIONAL ETHICS REVIEW COMMITTEE (SU-IERC)
03 DEC 2019
TEL: +254 (0)703 034 000
P. O. Box 59857 - 00200
NAIROBI - KENYA

Cc: Prof Fred Were,
Chairperson; SU-IERC

Ole Sangale Rd, Madaraka Estate. PO Box 59857-00200, Nairobi, Kenya. Tel +254 (0)703 034000
Email info@strathmore.edu www.strathmore.edu

8.3. Appendix C: Use Cases

Table 8.1 Merchant Registration : Use for Merchant registration

Use Case: Merchant Registration	
Primary Actor: Administrator	
Precondition: Administrator Panel Activated - Login	
Postcondition: Administrator Panel Deactivated -Log off	
Main Success Scenario	
Actor responsibility	**System responsibility**
1. Administrator enters the details of the Merchant	
	2. The system captures the information entered
	3. The system validates if the user information is entered correctly
	4. The system saves the information in the database
5. The administrator exits the panel.	

Table 8.2 Apparel Management.

This use case illustrates apparel stock management for effective advertising.

Use Case: Apparel Management	
Primary Actor: Merchant	
Precondition: Merchant Credentials are Active; Merchant Panel is Active	
Postcondition: Merchant panel is deactivated	
Main Success Scenario	
Actor responsibility	**System responsibility**
1. The merchant enters the details of the apparel on sale	
	2. The system captures the information entered
	3. The system validates if the apparel information is entered correctly
	4. The system saves the information in the database
5. Merchant exits the panel.	

Table 8.3 Update Model

The Update model use case review the process in making necessary technical changes to the model.

Use Case: Update Model	
Primary Actor: Developer	
Precondition: The model should have existing weights	
Postcondition: The model has existing weights	
Main Success Scenario	
Actor responsibility	**System responsibility**
1. The developer loads the system into memory.	

2. The developer loads new series of apparel images to the system.	
3. The developer initializes the system with the new series of images.	
	4. The system updates its features from the images.
	5. The features enable predictions while new weights are updated in the model.
	6. The system stores the new weights and saves the generated model.
7. The developer exits the system.	
Alternative 5 fails	
1. The developer initializes with new adjusted weights	
	2. The system undertakes steps 5 and 6.
3. The developer undertakes step 7.	

8.4. Appendix D: Code Snippets

Figure 8.1 illustrates import libraries that aid in apparel image detection and classification

```
import os
import h5py
import image
import operator

import matplotlib.pyplot as plt
import time, pickle, pandas

import numpy as np

import keras
from PIL import Image
import glob

from keras.preprocessing.image import ImageDataGenerator
from keras.models import Sequential, load_model
from keras.layers import Convolution2D, MaxPooling2D, ZeroPadding2D ,Conv2D
from keras.layers import Activation, Dropout, Flatten, Dense
from keras.callbacks import TensorBoard, ModelCheckpoint
from keras import backend
from keras import optimizers
from keras import applications
```

Figure 8. 1 Import Function

Figure 8.2: Code listing highlights the feature extraction of color from apparel images.

```
from sklearn.cluster import KMeans
import matplotlib.pyplot as plt
import numpy as np
import cv2
from collections import Counter
from skimage.color import rgb2lab, deltaE_cie76
import os


def RGB2HEX(color):
    return "#{:02x}{:02x}{:02x}".format(int(color[0]), int(color[1]), int(color[2]))

def get_image(image_path):
    image = cv2.imread(image_path)
    image = cv2.cvtColor(image, cv2.COLOR_BGR2RGB)
    return image


def get_colors(image, number_of_colors, show_chart):
    modified_image = cv2.resize(image, (600, 400), interpolation=cv2.INTER_AREA)
    modified_image = modified_image.reshape(modified_image.shape[0] * modified_image.shape[1], 3)

    clf = KMeans(n_clusters=number_of_colors)
    labels = clf.fit_predict(modified_image)

    counts = Counter(labels)

    center_colors = clf.cluster_centers_
    # We get ordered colors by iterating through the keys
    ordered_colors = [center_colors[i] for i in counts.keys()]
    hex_colors = [RGB2HEX(ordered_colors[i]) for i in counts.keys()]
    rgb_colors = [ordered_colors[i] for i in counts.keys()]

    return rgb_colors
```

Figure 8. 2 Feature Extraction

Figure 8.3: Illustrate the code listing for combination of features extracted from apparel ready for recommendation.

```
def squeeze_h():
    d8 = data.H
    df8 = d8.transform(score)
    dff8 = df8.groupby(d8.index // N).mean()
    return dff8

def df_combine():
    dfc = pd.concat([squeeze_a(), squeeze_b()], axis=1)
    dfc1 = pd.concat([squeeze_c(), squeeze_d()], axis=1)
    dfc2 = pd.concat([squeeze_e(), squeeze_f()], axis=1)
    dfc3 = pd.concat([squeeze_g(), squeeze_h()], axis=1)
    dfc_fin = dfc = pd.concat([dfc, dfc1,dfc2,dfc3], axis=1)
    return dfc_fin

def df_write():
    df = df_combine()

    export_csv = df.to_csv ('C:/Model/Combine/shirtscombine.csv', index = None, header=True)
    return export_csv

def main():
    print(df_combine())
    df_write()
    predict_csv(df_write())
    pass

def predict_csv(df_combined):
    with open('saved_model.pkl', 'rb') as fid:
        classifier = pickle.load(fid)
    fid.close()
    data = pd.read_csv(df_combined)
    Ypredict = classifier.predict(data)
    print(Ypredict)
```

Figure 8. 3 Feature Combination

Figure 8.4 Illustrates the code listing for weighing the apparel features and submitting a recommendation to be displayed to the Digital Signage.

```
def similarity(feature_data,inp_feature_data):
    num_samp=inp_feature_data.size
#     print (num_samp)
    sim_score={}
    for i in range(len(feature_data)):
        score=0
#         show_sample(data_images[i])
#         print(feature_data[i])
        score_m= inp_feature_data - feature_data[i]
#         print (score_m)
        score = num_samp-np.count_nonzero(score_m)
        sim_score[i]=score
#         print (score)

    return sim_score

similarities=similarity(feature_data,inp_feature_data)
sorted_similarities = sorted(similarities.items(), key=operator.itemgetter(1),reverse=True)
#print (sorted_similarities)
num_reco=30
num_data=feature_data.size
for i in range(num_reco):
    ind = sorted_similarities[i][0]
    print("Score : ", sorted_similarities[i][1])
    show_sample(data_images[ind])

    return display(data_images)
```

Figure 8. 4:Apparel Display

Figure 8.5 illustrates code listing for generating an email report of recommended apparel

```
import smtplib
import ssl
smtp_server = "smtp.gmail.com"
port = 587
sender_email = "noreply@gmail.com"

# Create a secure SSL context
context = ssl.create_default_context()

try:
    server = smtplib.SMTP(smtp_server,port)
    server.ehlo()
    server.starttls(context=context)
    server.ehlo()
    server.login(sender_email, password)

except Exception as e:

    print(e)
finally:
    server.quit()
```

Figure 8. 5 : Email notification