

# A Smart Vehicular Forward Collision Alert System using Neural Network

Felix Larbi Aryeh  
Computer Sci. and Eng. Dept  
University of Mines and  
Technology  
Tarkwa, Ghana

Emmanuel Kwaku Eshun  
ICT Unit  
University of Mines and  
Technology  
Tarkwa, Ghana

Abdul-Harisu Inusah  
Computer Sci. and Eng. Dept  
University of Mines and  
Technology  
Tarkwa, Ghana

## ABSTRACT

Road accidents over the years have led to an increased mortality rate in Ghana. Most accidents are caused by stress, tiredness, and loss of concentration of drivers while driving. This paper sought to design and build an automatic detection forward collision alert system for already existing vehicles in Ghana. Specifically, a deep learning model was built using a convolutional neural network for detection of stationary/moving objects (person, cars, bicycle, truck, bus and others). The object's relative speed to that of the subject vehicle is used to determine the probability of a collision. A buzzing alarm sound occurs when the possibility of collision is higher. The proposed system is an economical technique of alerting drivers of any impending danger on our roads.

## Keywords

You Only Look Once (YOLO), Visual Geometry Group (VGG), Single Shot Multibox Detector (SSD), Advanced Driver Assistance System (ADAS).

## 1. INTRODUCTION

Ghana's road transport is the most prevalent form of public transport in Ghana. It represents 94% of freight and 97% of passenger traffic [2]. With a growing number of registered cars on our streets year after year, road usage in Ghana has had both positive and negative impact on the local economy. Roads have been crucial in combating poverty by stimulating rural and urban economic and social growth. Road traffic accident accounts for about 1.35 million deaths worldwide. Despite the number of resources invested, developing countries are the hardest whipped by road traffic accidents with a whopping 93% of world fatalities [8]. In Ghana, the total revenue from the road fund amounted to GHC 2.35 billion from January to December 2016, GHC 1.15 billion from traditional sources and GHC 2 billion from the loan agreement [6]. These are meant to ensure safe and good roads are provided, but the funds used for these undertakings do not correlate with the increasing number of road accidents. The National Road and Safety Commission's official figures indicate that fatalities and serious injuries increased by 15.6% and 6.77% respectively in 2015 and 2016 [9].

Road traffic accidents are caused by a variety of factors such as cell phone use during driving, over speeding, poor driving skills, failure to use seat belt and distracted driving. According to [15], performing in-vehicle tasks while driving has detrimental effects on driving performance and these effects in different environments are relatively stable across different driver age groups. [27] observed in their study that out of the 1700 observed motorists using a cell phone were more than ten times more likely to involve violations at the intersection in statutory stopping compared to those not

engaged. Statistics have shown that 20% of all accidents are related to fatigue. The American Automobile Association associates 17% of all fatal accidents in the United States with driver fatigue [17].

Over the years, there has been ongoing research on how best to prevent road accidents. Some of these efforts yielded into Advanced Driver Assistance Systems (ADAS). ADAS has several components such as Pedestrian Collision Warning System (PCW), Lane Departure Warning Systems (LDW) and Forward Collision Warning Systems (FCW) to reduce the risk of road accidents to the barest minimum. ADAS is currently available in new expensive cars [17].

## 1.1 Deep Learning

Deep learning is a machine learning technique used by computers to learn. It is used in computer tasks that are considered to be difficult. Deep learning models can achieve state-of-the-art outcomes that go beyond human capabilities [1]. To achieve this feat deep learning models, require large quantities of data. Deep learning helps computers perform classification tasks such as identifying or not identifying a mail as spam, determining an object's location, and many others. The AlphaGo program, which learned the go game using deep learning and won the 2016 world champion, is a stunning example of what deep learning can do [25]. Deep learning models use various architectures of the neural network, such as Convolution Neural Network (CNN), Recurrent Neural Network (RNN), Multilayer Perceptron (MLP), etc.

### 1.1.1 Neural Networks

Neural networks are basically a collection of neurons with input, hidden and output layers. Neurons and activation functions are the fundamental building blocks of a neural network. Activation functions help to activate a linear function which is a one-degree polynomial to a nonlinear function. It would be a linear model that cannot learn complicated data such as video, audio, images, speech without activating a neural network. There are various variants of activation function such as tanh, relu, softmax, sigmoid, leaky relu which enables complicated data to be learned by a neural network. Neural networks are modelled loosely after how the human brain is designed to recognize patterns. Why neurons? Neurons in the human brain are capable of receiving signals from its dendrites. A neuron sends out spikes of electrical activity through a long, thin strand called the axon which splits into thousands of branches. A structure called a synapse converts the activity from the axon into electrical effects at the end of each branch that inhibit or excite activity from the axon into electrical effects to the connected neurons. When a neuron receives enough exciting input compared to its inhibitory input, it sends a spike of electrical activity down its

axon. Learning occurs by changing the effectiveness of the synapse to change the influence of one neuron on other changes. Fig. 1 shows a diagram of the neuron.

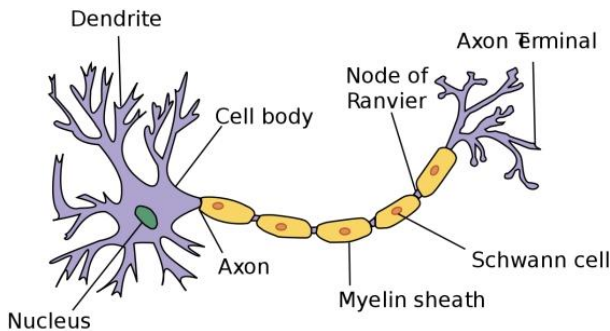


Fig 1: Diagram of a Neuron [3]

## 1.1.2 Types of Neural Networks

### 1.1.2.1 Multilayer Perceptron

It is the oldest deep learning type and simplest algorithm in deep learning. They are generally regarded as the "hello world" of deep learning. They are composed of an input layer for receiving the signal, an output layer that makes a decision or prediction about the input, and there are usually two or three hidden layers in between. They are widely used to solve the problems of supervised learning [11].

### 1.1.2.2 Convolutional Neural Networks

MLP's are not well suited for data such as images due to the spatial nature of images. MLP's work with vectors so transforming images into vectors will mean they will lose their spatial information such as form, edges, and shapes. Computer vision professionals used to manually extract interest variables called features before Convolution Neural Networks. But with the advent of CNN, a task that includes segmentation of images, classification of images, recognition of objects, and recognition of face is easy and performs better than conventional techniques [20]. The input features are taken in batch wise as a filter in a Convolutional Neural Network (CNN). This helps the model to remember image parts and do computations such as converting the image to grayscale from RGB or HSL. This helps to identify edges and shapes that can assist in classifying into various categories.

### 1.1.2.3 Recurrent Neural Networks

Recurrent Neural Network is the best neural network suited for sequential data. It is the first algorithm to remember its inputs because of its internal memory. Why do you need RNN? RNN's are very well suited for problems which have a lot of sequential data and that temporal dynamics that connect the data is more important than the spatial content of each spatial frame. The simplest Recurrent Neural Network can be seen as multiple copies of the same network passing information to its successor. A spectacular example is Deep Voice a text-to-speech conversion model developed by Baidu Artificial Intelligence Lab in California [10].

### 1.1.2.4 Kohonen Self Organizing Neural Networks

As the name suggests, a neural network that organizes itself based on the training data is a Kohonen Self Organizing Neural Networks. Its main goal is to input arbitrary dimension vectors into a discrete neuron map. The neurons remain constant during training, but the weights vary. Its self-organizing nature comes in phases, the first phase being all the neurons are initialised with a small weight and the input

vector. The neuron closest to the output moves in its direction in the second phase, followed by the other neurons. Using the Euclidean distance, the distance between the output and the neuron is calculated and the neuron with the least distance is the winning neuron.

### 1.1.2.5 Radial Basis Function Neural Networks

Radial basis function is an artificial neural network that takes the point distance from the center into consideration. The network measures the distance that can be used with Euclidean or any other measurement of distance. In classifying the models into different categories, the model depends on the maximum reach or radius of the circles. If a point is in or around the radius, it is highly likely to be classified in that class.

### 1.1.2.6 Modular Neural Networks

A modular Neural Network is a collection of independent computing networks for an output. Each network has its own unique set of inputs. Modular neural networks are very useful in breaking up large computational tasks into smaller components that reduce complexity. This reduction helps to increase computational speed by decreasing the number of connections.

## 2. RELATED WORKS

Traditionally, Haar Cascade was used to determine the existence of an object and the location of the object before deep neural networks. The first step in Haar Cascade is to accumulate Haar features that include a collection of rectangular regions in a detection window at a specific location, summarized in pixel intensities at each region, and calculate the difference between these amounts. The best features are then selected using a technique called AdaBoost, and each calculated feature is compared to a learned threshold to separate non-object from objects. They are then organized in a cascade to form a stronger classifier [29]. Convolution Neural Networks (CNN's) are very good for image classification but when it comes to object localization. They fall short. Unlike simple CNN classification algorithms, object detection algorithms do not only predict the type of object but also predict the object's boundary box. That is, bounding coordinates for the center of the object and the bounding width and height of the object. In self-driving cars, algorithms for object detection are very prominent to detect other cars, trucks, objects around them. They also assist in estimating pose, monitoring, etc.

## 2.1 Variants of Object Detection Algorithms

### 2.1.1 Region Convolutional Neural Network (R-CNN)

In their paper, [14], proposed a method in which 2000 regions were proposed on the basis of a selective algorithm. The proposed regions are then transformed into a square and fed into a Convolution Neural Network that produces as output a feature vector of 4096-dimensional. To classify the object within the proposed region, CNN is used to extract features such as shapes, edges before being fed to a Support Vector Machine (SVM). Besides predicting an object's presence, the algorithm also predicts four offset values used to increase the bounding box precision. The downside of using R-CNN is that it takes longer to train. Also, because of the calculation required to propose 2000 regions, it cannot be implemented in real time.

### 2.1.2 Fast Region Convolutional Neural Network (Fast R-CNN)

[13] proposed a better approach to reduce the computational speed required to propose regions. [13] proposed the use of convolutional feature maps to identify region proposal and warp them into squares using a Region of Interest (RoI) pooling layer to reshape them into a fixed size so they can be fed into the fully connected layer. Using the RoI feature vector, a softmax layer is used to predict the class of the proposed region and also the offset values for the bounding box.

### 2.1.3 Faster Region Convolutional Neural Network (Faster R-CNN)

[24] in their paper proposed a technique to improve the speed of the Fast R-CNN. It comprised of two modules, a first module for the region proposals and the Fast R-CNN detector. The first module eliminates the need for the selective search algorithm used in R-CNN which was the main bottleneck affecting the performance of the algorithm. The predicted region proposals are then reshaped using the RoI pooling layer which is then used to determine the object and its offset.

### 2.1.4 You Only Look Once (YOLO)

YOLO is an algorithm for object detection designed to target object detection in real time. In contrast to previous algorithms, YOLO does not propose an arbitrary number of regions, but rather looks into smaller parts of an image that is likely to contain an object. As a matter of fact, how it works is by splitting an image into grid of  $m$  bounding boxes within each grid. For each bounding box, the model predicts a class object and offset values. The bounding boxes above the threshold value are selected to locate the object in the image [23].

### 2.1.5 Single Shot Multibox Detector (SSD)

The Single Shot Multibox Detector (SSD) is an object detection algorithm which has a fine balance between accuracy and computation. SSD algorithm achieves this feat by running an input image into a convolutional network only once. The network then calculates a feature map. How is this done? The input image is passed through a Visual Geometry Group-16 (VGG-16) convolution layer for extracting the shapes, edges that are beneficial in identifying the category of image. They are then run through a  $3 \times 3$  sized convolutional kernel to predict the bounding boxes and classification probability [19].

## 2.2 Inverse Perspective Mapping

Inverse Perspective Mapping is a technique used to transform an image from one perspective to another. The transformed perspective is usually the eye perspective of the bird, but based on the angle of rotation ( $\Theta$ ) it can be changed. This method uses four points from the original image to project the transformed image.

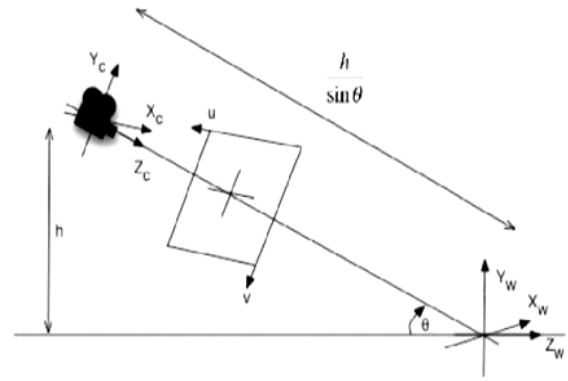


Fig. 2 Inverse Perspective Mapping [28]

From Fig. 2, the mapping of a surface ( $X_w, Y_w, Z_w$ ) is projected on the surface ( $u, v$ ) through a rotation ( $\Theta$ ). The transformation is done along with the optical axis camera scaled by the camera parameter matrix. The mapping can be expressed as:

$$(u, v, 1)^T = KTR(x, y, z, 1)^T$$

where  $R$  = angle of rotation

$T$  = translation matrix

$K$  = camera parameter matrix

## 2.3 Other Related Works

[18] worked on pre-collisional system algorithms ' potential effectiveness. These include only Forward Collision Warning (FCW), FCW and Pre-crash Brake Assist (PBA) and FCW, PBA and Pre-crash Brake Autonomous (PB). Based on the study, a sample of 1396 real-world collisions were simulated as if they occurred, corresponding to 1.1 million crashes. A probability-based framework was developed to consider the driver response to the warning. Conclusions made were, FCW prevents 3.2% of all crashes, FCW + PBA prevents 3.6% of all crashes, FCW + PBA + PB prevents 7.7% of all crashes. Thus, the use of pre-collision algorithms is very effective in preventing serious and fatal injuries caused by road accidents. CNN-based Pedestrian Collision Warning (PCW) system is accurate in pedestrian detection compared to traditional PCW systems as false alarms are reduced [17].

In their research, [28] proposed a solution for determining a vehicle's presence using the road surface subtraction algorithm and IPM to determine the vehicle's distance. In the study, [28] found metal parts of vehicles to reflect Radio Detection and Ranging (RADAR) effectively than human tissue. Therefore, it can be very difficult to detect pedestrians with RADAR especially at long distances. Power requirements and interference problems also make it a better choice to use optical camera sensors.

In their study, [22] explored how cameras can be used as Front Collision Alert (FCA) instead of using RADAR and LIDAR. The technique used explores the use of principles of image scale change for estimating the Time To Collision (TTC). Results using the proposed system on a test track and public road testing support a camera-based FCA deployment indicate that the system would meet the U.S. Department of Transportation New Car Assessment Program (NCAP) test requirement for forward collision warning.

In their research, [21] proposed an algorithm that provides a

more accurate solution to determine the forward object's distance. The solution proposed determines the virtual horizon from vehicle size and position. The virtual horizon is estimated if a vehicle's real width is known and the vehicle's position, vertical coordinate, is given in the picture. Due to pitch movement of subject vehicles such as vibration and acceleration, horizon positions can fluctuate leading to false distance calculation.

[26] used image sequences recorded by a monocular camera to work on a distance calculation algorithm. At its field of view, the camera captures real-time images in sequences. Subsequent images are processed to produce an accurate estimation of distance.

### **3. SYSTEM DESIGN**

#### **3.1 Feasibility Studies**

Arguably, in every paper, there is a need to ascertain what requirements are feasible and what isn't. Typically, there are cases where the cost of establishing the requirements exceeds the benefits gained in satisfying the problem which underpins them. There may also be cases where fulfilling the requirement is *prima facie*, illegal, unethical or contrary to scientific laws. Understanding feasibility studies expose the development team to the primary risk the team may come across. This help prevents dead ends possible to occur after months of development. In this paper, feasibility helped to prioritize efforts across all aspects of development. Below are some of the feasibility analysis done for this paper.

##### *3.1.1 Financial Feasibility*

To validate that, based on the finances at hand, the goals of this paper are achievable. Considerations were made to ascertain the cost of each component required for the successful completion of this paper. Questions such as, how much money could be raised for the paper? Are there organizations that want to fund the paper? What would the overall expenditure be? In a bid to answer these questions, funding was received from parents and others to purchase some of the components required for the paper based on priority. Unfortunately, there was no support from any organization to fund the paper.

##### *3.1.2 Technical Feasibility*

Here, the assessment of skills required for the completion of the project was taken into consideration. During the initial stages, the following are questions that came to mind. What are the skills required for this paper? Is there enough time to get trained within the shortest time frame for the successful completion of the paper? Enough time was invested to understand how image processing and machine learning work.

##### *3.1.3 Resource Feasibility*

Evaluating the resources available is important to ensure a successful project. Here the resources available to me was taken to consideration. Some questions that needed answering were: What resources are needed? What facilities are needed? Are resources available sufficient? To answer, the microcontroller used for this project was found to be sufficient but not fast and efficient enough. Preferably, a better fit would be Nvidia Jetson TX2 microcontroller. Also, there is no dataset on road users in Ghana that would have helped better finetune the model.

##### *3.1.4 Operational Feasibility*

The feasibility of deploying and operating the project was considered. Is there an existing system in place? If yes, how does it interface with its working environment? In an attempt

to answer the above questions, the system will interface with already existing vehicles providing an alarm in the near case of collision.

##### *3.1.5 Schedule Feasibility*

After completing the feasibility studies, the report was positive for undertaking the project. The next step is gathering requirement from the users and stakeholders involved. Techniques used was to ensure users will enjoy the visual appeal, outlook, and usability of the project. Their contributions and views were solicited through observation, prototyping, brainstorming sessions, and interviews.

#### **3.2 Requirement Gathering**

After completing the feasibility studies, the report was positive for undertaking the project. The next step is gathering requirement from the users and stakeholders involved. Techniques used was to ensure users will enjoy the visual appeal, outlook, and usability of the project. Their contributions and views were solicited through observation, prototyping, brainstorming sessions, and interviews.

##### *3.2.1 Interview*

Interviewing stakeholders is very important in creating a usable application. Without knowing the expectation of users, you are unlikely to satiate them. Understanding the perspective of every interviewee was crucial to properly address and weigh their inputs. Drivers of both commercial and private vehicles were interviewed on how such a system can help them.

##### *3.2.2 Prototyping*

Prototypes can be a very useful tool to communicate ideas effectively. Using prototypes with low fidelity makes a good listening tool by articulating specific needs which would have been difficult to explain in the abstract to the layman. The use of sketches, storyboards made it possible for stakeholders to provide quick feedback to meet their needs.

##### *3.2.3 Brainstorming*

It is generally used to collect a good number of ideas from people in requirement elicitation. Generally speaking, it helps to identify different discrete possibilities. Brainstorming sessions were held with friends, supervisors, lecturers, and drivers to ensure all possible options were catered.

##### *3.2.4 Observation*

To complete this project, observing stakeholders and users in their natural workflow was very important. Watching how users drive, their process workflow, pain points, awkward steps opened opportunities for improvement. Using this technique helped to uncover implicit requirements that were overlooked.

#### **3.3 Software Development Life Cycle (SDLC) Model**

To produce a high-quality system that exceeds user expectations works in its planned working environment efficiently and effectively. A suitable model of SDLC was carefully analysed. The waterfall model was removed early on in the decision-making process because of inflexibility. The big bang model was also eliminated because of its high-risk nature. In case the requirements were misunderstood in the beginning, there is a chance one can get to the end of the project and realize the project needs to be restarted. The spiral model was eliminated for its complexity. Looking at the remaining choices, the prototyping model, the V-shaped model, the iterative and incremental model a decision was

made with the prototyping model. This model fits their use case because it gives you the chance to build a prototype that can help you understand the user's needs. This project also has various moving parts that can be constructed separately before being combined to produce a final product. Fig. 3 illustrate the prototyping model.

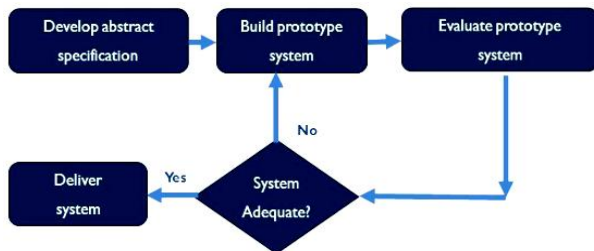


Fig 3: Diagram of the Prototyping Model

### 3.4 Functional Requirement

In the development of any system, its requirements need to be thought out clearly by all parties involved to prevent abandoning the project halfway through. Functional requirements of a system specify what the system does. The functional requirements for this project include:

- The system should be able to identify the presence of an object in front of it.
- The system should be able to determine the relative speed between the subject vehicle (S.V) and the stationary or moving object.
- Determine the distance between S.V and object.

The functional requirements can be best appreciated using the Unified Modelling Language (UML). UML is a modelling language used to visualize the design of the project. The UML's used in this project are a flowchart and an activity diagram.

#### 3.4.1 Flowchart

A flowchart is a visual representation of the sequences and decisions necessary to accomplish a task. It uses connecting lines and directional arrows and others to show the logical flow of the system from start to finish as shown in Fig 4.

#### 3.4.2 Activity Diagram

It is used to portray the control from the start point to the finish point showing the various decision paths that exist while the activity is executed. The detection of both sequential and concurrent processing of activities is possible. In this project, Fig. 5 shows the control flow drawn from one operation to another.

### 3.5 Non - Functional Requirement

The non-functional requirements describe how the system works. Generally, it specifies how the system should behave and constraints under which it works. The qualities which embody the non-functional requirements are scalability, performance, availability, security, manageability, data integrity interoperability capacity. The following are some of the non-functional requirements.

- Will the system suffer downtime?
- How efficient can the system predict forward collision?
- Can the system be hacked?

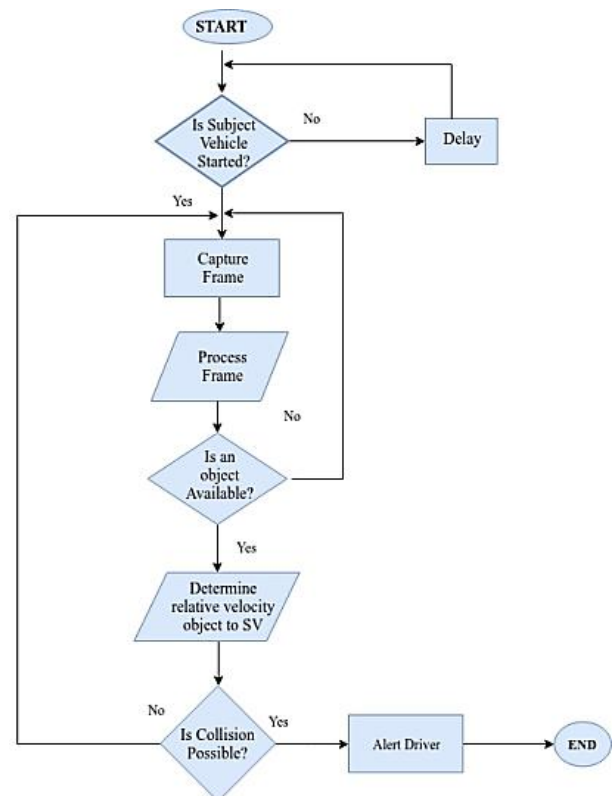


Fig 4: Flowchart of the System

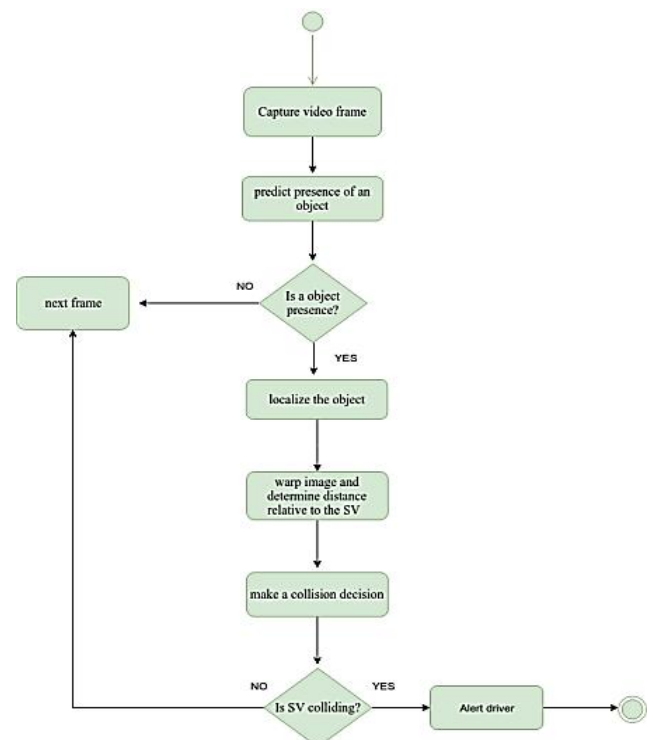


Fig 5: Activity Diagram of the System

### 3.6 Object Detection Model (SSD)

Single Shot Multibox Detector algorithm was selected after careful evaluation of the various variants of object detection algorithms. The algorithm achieves a 74.3% mAP on VOC2007 test at 59 fps on Nvidia Titan X for 300 x 300 input and 76.9% mAP on VOC2007 input [19]. It consists of a module for the extraction of features and a filter for the location of objects.

### 3.6.1 SSD Architecture

SSD is built on a Visual Geometry Group-16 (VGG-16) architecture without a fully connected layer. VGG-16 serves as the base network that allows feature extraction. It also leverages on transfer learning to get state of the art results. A set of other convolution layers are added to allow extraction at multiple scales and gradually decrease the size of the input to each next layer. Fig. 6 shows the architecture of a Single Shot Multibox Detector (SSD).

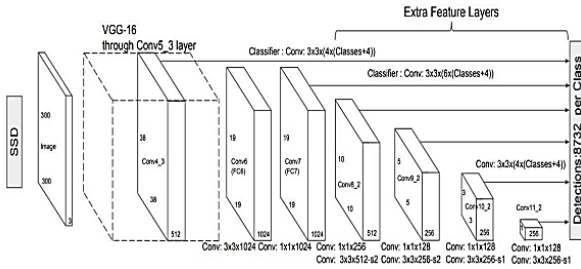


Fig 6: Architecture of Single Shot Multibox Detector [16]

### 3.6.2 Loss Function

There needs to be a way to check how accurate it is when it makes a prediction for a network to learn. The SSD network uses the location and confidence loss to determine how accurate its prediction is. The confidence loss measures how confident the network is of the object's computed bounding box. The location loss measures how far away the network forecasted the bounding boxes from the ground truth ones from the training set. The loss function contains two components, the loss confidence (Lconf) and the loss localization (Lloc).

$$L(x, c, lg) = \frac{1}{N} (L_{conf}(x, c) + \alpha L_{loc}(x, l, g))$$

Where N is the number of matched default boxes. If N = 0, the loss is set to 0. The predicted box is (l) and the ground truth box (g) parameters.

$$L_{loc}(x, l, g) = \sum_{i \in Pos} \sum_{m \in \{cx, mcy, w, h\}} x_{ij}^k smooth_{L1}(l_i^m - g_j^m)$$

The loss confidence equation below is a softmax loss over multiple classes confidences (c).

$$L_{conf}(x, c) = - \sum_{i \in Pos} x_{ij}^p \log(c_i^p) - \sum_{i \in -log} \log(c_i^o)$$

$$\text{Where } c_i^p = \frac{\exp(c_i^p)}{\sum_p \exp(c_i^p)}$$

### 3.6.3 Data Augmentation

To make the model robust in different environments. There's the need to teach it to be prepared for various object sizes, aspect ratio in the input. During training, the images are horizontally flipped to enable the model better.

### 3.6.4 Non-Maximum Suppression

A large number of predicted boxes are developed during the forward pass in an SSD for a single object. In other to find the predicted box closest to the ground truth object, a technique called Non-Maximum Suppression is used to prune boxes that have Intersection of Union (IoU) less than the accepted threshold usually 0.5. Thus, the predicted box closest to the ground truth is used.

### 3.6.5 Data Selection

For a neural network to perform well on a task, it requires large sets of data. For this paper, the Pattern Analysis, Statistical Modelling and Computational Learning Visual Object Classes 2012 dataset (PASCAL VOC2012), the widely known dataset for object detection, image segmentation, action classification, and large-scale recognition were employed. The dataset contains 27450 images of which 3422 are segmented distributed among 20 objects [12]. Table 1 shows the distribution of the images among the different objects.

Table 1: Distribution of objects in segmented datasets

Objects	Train		Validate	
	Image	Object	Image	Object
Aeroplane	88	108	90	190
Bicycle	65	94	79	103
Bird	105	137	103	140
Boat	78	124	72	108
Bottle	87	195	96	162
Bus	78	121	74	116
Car	128	209	127	249
Cat	131	154	119	132
Chair	148	303	123	245
Cow	64	152	71	132
Diningtable	82	86	75	82
Dog	121	149	128	150
Horse	68	100	79	104
Motorbike	81	101	76	103
Person	442	868	445	865
Pottedplant	82	151	85	171
Sheep	63	155	57	153
Sofa	93	103	90	106
Train	83	96	84	93
Tvmonitor	84	101	74	98
<b>Total</b>	<b>1464</b>	<b>3507</b>	<b>1449</b>	<b>3422</b>

(Source: [12])

Fig. 7 and 8 shown are an example of a sample and its segmented counterpart containing a person and a bicycle.





Fig 7: Segmented Image [12]



Fig 8: Original Image (Source: [12])

### 3.7 Hardware Design

The raspberry pi B+ microcontroller is a key component in the design process and is the main control unit that controls the operation of the other components. It requires about 2.5A of 5V Direct Current (D.C) for its operation. The Raspberry Pi Camera V2 is the main sensor used in this design. It is a custom designed add-on board for Raspberry Pi with an 8-megapixel native resolution featuring a fixed focus lens. The microcontroller is programmed to control the piezo buzzer operation. The piezo buzzer is an electronic component that is designed in the system to produce sound. Fig. 9 shows a block of how the various components interface with each other. Fig. 10 shows the schematic diagram of the system.

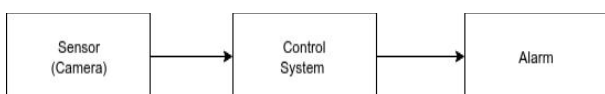


Fig 9: Block diagram of the Proposed Design

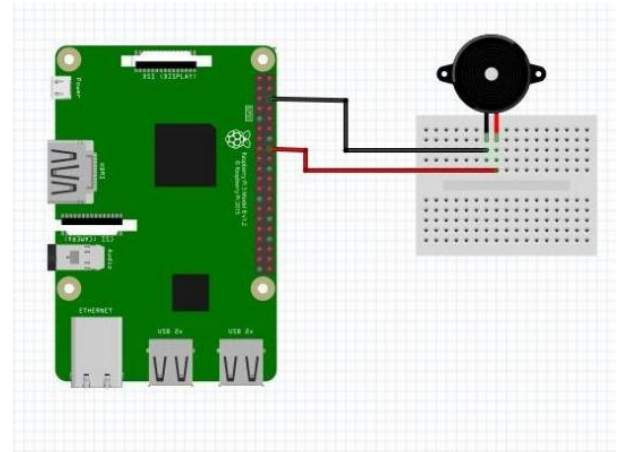


Fig. 10 Schematic Diagram of the System

#### 3.7.1 Components of the proposed design

- a. Control System (Microcontroller)
- b. Sensor (Camera)
- c. Alarm (Piezo Buzzer)

#### 3.7.2 Functions of the Components

- a. The Sensor converts images of the working environment into digital signals to be sent to the control system.
- b. The control system receives signals (static images) from the camera, processes it into a digital signal and sends the output to the alarm to operate.
- c. The alarm converts signal from control system into sound.

#### 3.7.3 System Hardware Components

##### 3.7.3.1 Piezo buzzer

Piezo buzzer is a sound-producing electronic component. Features such as light, simple construction and low price make it suitable for various applications such as car/truck reversal indicator, car bells, etc. Piezo buzzer is based on Jacques and Pierre Curie's inverse principle of piezoelectricity discovered in 1880.



Fig. 11 Piezo Buzzer ([4])

It is the phenomena of generating electricity when mechanical pressure is applied to certain materials and the vice versa is also true. Such materials are called piezoelectric materials.

When exposed to an alternating electrical field, there is the tendency of stretching or compressing according to the frequency of the signal producing the sound. Fig. 11 shows an example of a Piezo Buzzer.

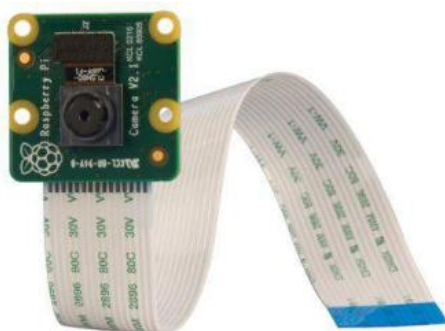
### 3.7.3.2 Raspberry Pi V2 Camera

The Raspberry Pi V2 Module is a Sony IMX219 8-megapixel sensor capable of taking high-definition still photographs and video. The module supports 1080p30, 720p60, and Video Graphic Array 90 (VGA90) video modes and is used in home security and wildlife camera traps. It attaches to Pi by means of one of the small sockets on the upper surface of the board and uses the Camera Serial interface (CSI) a specially designed interface for cameras. The board itself is small, at approximately 25mm x 23mm x 9mm. It also weighs just over 3g, making it ideal for mobile or other applications where size and weight are important. It is connected to the Raspberry Pi through a short ribbon cable. Fig. 12 shows an example of the Raspberry Pi V2 camera module. Table 2 shows the description and specification of the Raspberry Pi B+ Board.

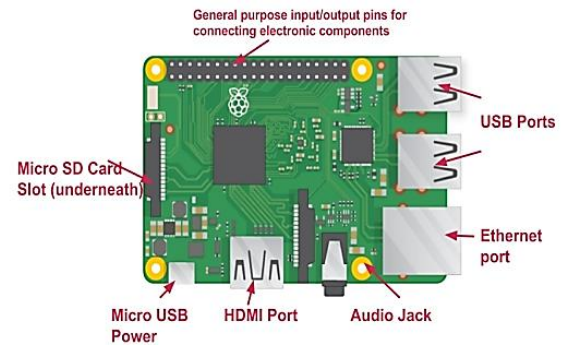
**Table 2: Description and Specification of the Raspberry Pi B+ Board**

Description	Specification
Processor	Broadcom BCM2837B0, Cortex-A53 64-bit SoC @ 1.4GHz
Memory	1GB LPDDR2 SDR
Connectivity	<ul style="list-style-type: none"> <li>2.4GHz and 5GHz IEEE 802.11.b/g/n/ac wireless LAN, Bluetooth 4.2, BLE</li> <li>Gigabit Ethernet over USB 2.0 (maximum throughput 300Mbps)</li> <li>4 × USB 2.0 ports</li> </ul>
Access	Extended 40-pin GPIO header
Video & sound	<ul style="list-style-type: none"> <li>MIPI DSI display port</li> <li>MIPI CSI camera port</li> <li>4 pole stereo output and composite video port</li> <li>1 × full size HDMI</li> </ul>
Multimedia	H.264, MPEG-4 decode (1080p30); H.264 encode (1080p30); OpenGL ES 1.1, 2.0 graphics
SD card support	Micro SD format for loading operating system and data storage
Input power	<ul style="list-style-type: none"> <li>5V/2.5A DC via micro USB connector</li> <li>5V DC via GPIO header</li> <li>Power over Ethernet (PoE)–enabled (requires separate PoE HAT)</li> </ul>
Environment	Operating temperature, 0–50°C

(Source: [5])



**Fig 12: Raspberry Pi V2 Camera Module (Source: [7])**



**Fig 13: Raspberry Pi B+ (Source: [5])**

## 4. RESULTS AND DISCUSSION

### 4.1 Overview

The implementations detail of the proposed are discussed in this chapter. Testing is done to evaluate the system and compared with currently existing systems. Also, in this chapter, a detailed cost analysis of the proposed system design is presented.

### 4.2 Operation of the Prototype

Following training the model, it was deployed on a raspberry pi 3 B+. The raspberry pi 3 B+ together with the picamera v2 successfully was able to take video feed (frames to the model) for the model to detect the presence of an object and localise the object. The frame is then warped to determine the distance between the object the subject vehicle. Snapshots below show the system and how it works. Fig. 14 shows a prototype of the proposed system. Fig. 15 shows how the system is able to predict the presence of a vehicle and draws a boundary box around it. Fig. 16 shows how the system estimates the likelihood of collision of a person not too far away. Fig. 17 shows how the system estimates the likelihood of collision of a person close.



**Fig 14: System Prototype**



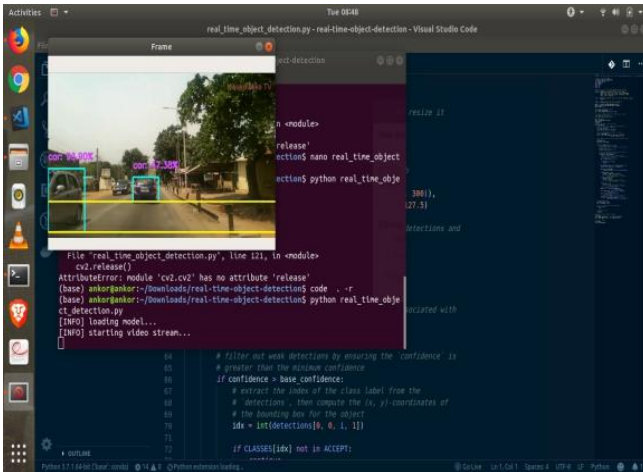


Fig 15: Detection of Cars Using the System

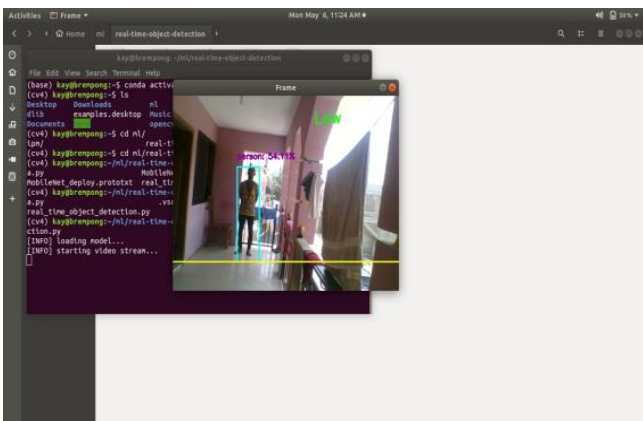


Fig. 16: Person Far Away from the System

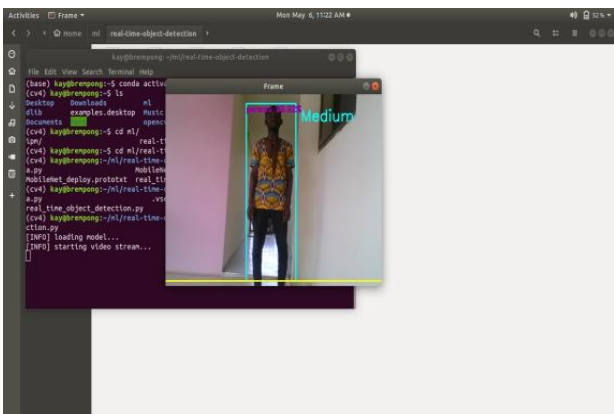


Fig 17: Person Within Better Distance of the System

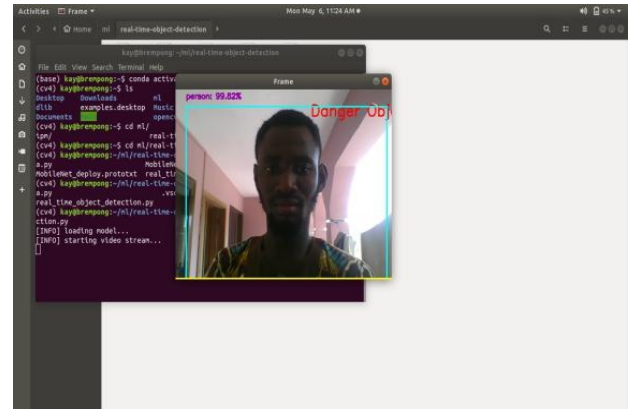


Fig 18: Person Too Close to The System

### 4.3 Cost Analysis

A detailed cost analysis of the proposed design was done to determine the true cost of this project. It showed that the total cost of GHC 347.50 was needed. Table 3 presents a summary of the detailed cost of the components used.

Table 3: Detailed Cost Analysis of the Proposed Design

Name	Quantity	Unit Cost (Ghc)	Total Cost (Ghc)
Raspberry Pi B+	1	255	255
Pi Camera V2 Module	1	90	90
Piezo Buzzer	1	2.5	2.5
<b>TOTAL COST</b>			<b>347.5</b>

## 5. CONCLUSION AND FUTURE WORK

### 5.1 Conclusion

Considering the technique employed and the gains made. It can be concluded that:

- The proposed design provides an economical way of preventing forward collision; and
- The technique used is very efficient at determining the presence of an object stationary/moving) and its relative velocity.

### 5.2 Recommendation

Future iterations of the system should consider implementing the following:

- A collision avoidance mechanism such as lane changing or brake application;
- Distance determination can be made more accurate on curved roads; and
- Object detection can also consider large potholes and also rocks on roads.

## 6. REFERENCES

[1] Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., Hasan, M., Van Essen, B. C., Awwal, A. A. and Asari, V. K. (2019), "A State-of-the-Art Survey on Deep Learning Theory and Architectures". Electronics, Vol. 8, Issue 3, p. 292.

[2] Anon. (2009), "Transport: Bedrock for Economic Development", <https://www.natcomreport.com/ghana/livre/transport.pdf>. Accessed: January 25,

- 2019.
- [3] Anon. (2012), “Neuron”, <https://www.simple.wikipedia/wiki/Neuron>. Accessed: February 2, 2019.
- [4] Anon. (2015), “Piezo Buzzer (Through Hole)”, <https://www.inventelectronics.com/product/piezo-buzzer-through-hole>. Accessed: February 8, 2019.
- [5] Anon. (2016), “Raspberry Pi 3 Model B+”, <http://www.raspberrypi.org/products/raspberry-pi-3-model-b+>. Accessed: February 3, 2019.
- [6] Anon. (2017a), “2017 State Ownership Report”, <http://www.mofep.gov.gh/sites/default/files/reports/economic/2017-State-Ownership-Report.pdf>. Accessed: February 3, 2019.
- [7] Anon. (2017b), “Getting Started with Picamera Module Your Raspberry Pi”, <https://www.electrone.com/getting-started-picamera-module>. Accessed: February 24, 2019.
- [8] Anon. (2018a), “Statistics: Road Safety Overview”, <http://www.nrsc.gov.gh/index.php/statistics>. Accessed: February 8, 2019.
- [9] Anon. (2018b), “Global Status Report on Road Safety 2018”, [https://www.who.int/violence\\_injury\\_prevention/road\\_safety\\_status/2018/en/](https://www.who.int/violence_injury_prevention/road_safety_status/2018/en/). Accessed: January 5, 2019.
- [10] Arik, S. Ö., Chrzanowski, M., Coates, A., Diamos, G., Gibiansky, A., Kang, Y., Li, X., Miller, J., Ng, A., Raiman, J. and Sengupta, S. (2017), “Deep voice: Real-Time Neural Text-to-Speech”, Proceedings of the 34th International Conference on Machine Learning, Vol. 70, pp. 195 - 204.
- [11] Du, K. L. and Swamy, M. N. S. (2014), “Multilayer Perceptrons: Architecture and Error Backpropagation”, Neural Networks and Statistical Learning, pp. 83 - 126.
- [12] Everingham, M. and Winn, J. (2011), “The PASCAL Visual Object Classes Challenge 2012 (VOC2012) Development Kit”, Pattern Analysis, Statistical Modelling and Computational Learning, Tech. Rep, pp. 1 - 32.
- [13] Girshick, R. (2015), “Fast R-CNN”, IEEE International Conference on Computer Vision, pp. 1440 - 1448.
- [14] Girshick, R., Donahue, J., Darrell, T. and Malik, J. (2014), “Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation”, IEEE Conference on Computer Vision and Pattern Recognition, pp. 580 - 587.
- [15] Horberry, T., Anderson, J., Regan, M. A., Triggs, T. J. and Brown, J. (2006), “Driver Distraction: The Effects of Concurrent in-Vehicle Tasks, Road Environment Complexity and Age on Driving Performance”, Accident Analysis & Prevention, Vol. 38, Issue 1, pp. 185 - 191.
- [16] Hulstaert, L. (2019), “A Beginner’s Guide to Object Detection”, Datacamp, p. 1.
- [17] Jung, H., Choi, M. K., Soon, K. and Jung, W. Y. (2018), “End-to-End Pedestrian Collision Warning System Based on a Convolutional Neural Network with Semantic Segmentation”, International Conference on Consumer Electronics (ICCE), pp. 1 - 3.
- [18] Kusano, K. D. and Gabler, H. C. (2012), “Safety Benefits of Forward Collision Warning, Brake Assist, and Autonomous Braking Systems in Rear-End Collisions”, IEEE Transactions on Intelligent Transportation Systems, Vol. 13, Issue 4, pp.1546 – 1555.
- [19] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y. and Berg, A. C. (2016), “SSD: Single Shot Multibox Detector”, European Conference on Computer Vision, pp. 21 - 37.
- [20] O’Shea, K. and Nash, R. (2015), “An Introduction to Convolutional Neural Networks”, pp. 1 – 11.
- [21] Park, K. Y. and Hwang, S. Y. (2014), “Robust Range Estimation with a Monocular Camera for Vision-based Forward Collision Warning System”, The Scientific World Journal, pp. 1 – 9.
- [22] Raphael, E., Kiefer, R., Reisman, P. and Hayon, G. (2011), “Development of a Camera-Based Forward Collision Alert System, SAE International Journal of Passenger Cars-Mechanical Systems, Vol. 4, pp. 467 - 478.
- [23] Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. (2016), “You Only look Once: Unified, Real-Time Object Detection”, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779 - 788.
- [24] Ren, S., He, K., Girshick, R. and Sun, J. (2015), “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, Advances in Neural Information Processing Systems, pp. 91 - 99.
- [25] Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A. and Chen, Y. (2017), “Mastering the Game of Go without Human Knowledge”. Nature, Vol. 550, Issue 7676, p. 354.
- [26] Stein, G. P., Ferencz, A. D. and Avni, O. (2012). “Estimating Distance to an Object using a Sequence of Images Recorded by a Monocular Camera”, United States Patent, pp. 1 – 15.
- [27] Strayer, D. L., Watson, J. M. and Drews, F. A. (2011), “Cognitive Distraction while Multitasking in the Automobile”, Psychology of Learning and Motivation, Vol. 54, pp. 29 - 58.
- [28] Tuohy, S., O’Cualain, D., Jones, E. and Glavin, M. (2010), “Distance Determination for an automobile Environment using Inverse Perspective Mapping in OpenCV”, Irish Signals and Systems Conference, Galway – Ireland, pp. 1 - 8.
- [29] Viola, P. and Jones, M. (2001), “Rapid Object Detection using a Boosted Cascade of Simple Features”, Accepted Conference on Computer Vision and Pattern Recognition, Vol. 1, Issue 1, pp. 511 - 518.