**Project Documentation Draft**

Notes highlight in purple  
To Paraphrase highlight in green

Traditional Maltese Food and Snacks Classiﬁer

LITERATURE REVIEW Computer vision has become a very demanding technology that contributed to various sectors to identify various aspects; for instance, computer vision contributes a lot in the medical ﬁeld and helps practitioners such as nurses and doctors in the process of identifying, study or treat diseases. This technology is often found in agriculture, the tourism sector, sport, manufacturing amongst others. This technology works by taking an image, understand what the image is and interpret its content with a view to solve tasks and resolve queries.

Study relevant to pedestrian recognition discusses how computer vision includes two operations; to ﬁlter the image

by reducing noise that could hinder the object segmentation algorithms performance, and the second operation is subtracting the background from the region of interest (ROI). The main concern about this method is that the captured image conditions may have different lighting, shadows and level of reﬂections that might inﬂuence the overall performance of subtracting the background. This step could be quite challenging and plays a very important role in object detection systems, a solution is to use a sliding window that moves across an image and collect window view to the histogram of oriented gradients (HOG) however, this is an expensive operation. Therefore a less expensive solution are algorithms based on deep neural networks that are proven to have a better success rate.

Other relevant studies take into variable shape, size and the colour of the object to be detected. During a study regard vehicle intelligence and understanding of its surroundings, the researcher starts to question various aspects in relation to a typical driving scenario. Other variables discussed where the dynamic environment a vehicle is found in, thus the researcher takes the lighting of the environment and the background into consideration.

Throughout, various researchers have done intense study in the ﬁeld of computer vision, some of which have solved major tasks. One vital study in the medical sector has researched the use of computer vision in medical imaging to localize and examine human long bones. Images also come in the form of sound waves which are emitted by a B-scan sensor and are translated into a digital image more understandable to the human eye. However, a B-scan sensor is not able to distinguish between long bones and other specimens such as muscles, soft tissue, veins and other internal organs under the skin tissue. Hence, this study implements a deep learning system to be abletorecognizebone specimens.Thisstudy discussesvarious other algorithms but makes use of YOLOv3 convolution neural network algorithm to differentiate the said specimens in ultrasound images. Similarly, my research makes use of the same concepts this study does which classiﬁes long bones, or in my case traditional Maltese food, in digital images.

YOLOv3 Algorithm

YOLOv3 is a regression-based one-stage target detection algorithm just like SSD which can detect and classify

objects in an image at the same time. YOLOv3 divides a digital image into S x S non-overlapped grids and determines whether each grid has a target to predict bounding boxes around. Considering that YOLOv3 is a single end-toend network that performs feature extraction, location and classiﬁcation in a network, makes the algorithm extremely fast i.e. at 320x320 YOLOv3 runs at 22ms at 28.2mAP.

A convolution neural network (CNN) is a deep learning network that assigns importance to an image such as weights and can differentiate one aspect of an image from another. YOLOv3 is one of the fastest object detection algorithms that makes use of a CNN in real-time detection without loss of accuracy.