**Research Proposal Form**

|  |
| --- |
| **Name and Surname:**  Keane Schembri |
| **MCAST Email:** keane.schembri.a106267@mcast.edu.mt |
| **Group:**  IT-SWD-6.2A |
| **Research Title:** Traditional Maltese Food and Snacks Classifier |
| **Hypothesis and/or Research Questions:**  Using computer vision as one of deep learning methods, it is possible to classify various traditional Maltese food.  The significance of this research is to provide a tool that assists people who are not familiar with traditional Maltese food such as tourists to distinguish between various food easily.  This study will help answer several research questions including;   * *How can a reliable dataset be obtained or constructed?* * *Which algorithm is best used to classify food?* * *Will image processing be able to reach a high level of accuracy?* |
| **Outline of Key Literature:**  *Computer vision* is a study of developing techniques that help machines to visually understand the content of an image. Such techniques include *Image Processing* which is the process of digitally creating a new image from an existing raw image, this will allow enhancing the content in some way. (Szeliski, 2010) (Gerónimo, Serrat, López, & Baldrich, 2013)  Through computer vision, machines can be trained to recognize pixels and patterns by providing a set of relevant images with the respective label and making use of the correct tools. (Alpaydin, 2020) In order to train machines, one can make use of algorithms such as YOLOv3 which is an extremely fast and accurate algorithm. YOLOv3 applies a single neural network to the full image and divides this image into regions and bounding boxes. These bounding boxes will predict probabilities for each region. (Redmon & Farhadi, 2018) Since YOLOv3 makes use of a single neural network it can execute instructions much faster in comparison to other algorithms. (Zheng, et al., 2018)  Other algorithm includes R-CNN or Fast R-CNN which are region-based rather than object detection algorithms this means that region-based draws a bounding box around the object of interest instead. |
| **Overview of Methodology:**  To successfully make use of computer vision, one needs to follow a pipeline that will help understand a workflow. Such pipeline includes the following;  Problem Statement > Data Collection > Data Preparation > Model Training > Model Inference  This pipeline starts with *Problem Statement* where one can identify which type of computer vision technique to use; in this case, one can follow a classification type which is the ability to understand what’s inside an image according to this visual content. (Le, 2018)  To create such dataset, images are scraped from google and sorted depending on their relevance towards the project, this will cover the *Data Collection* stage.  *Data Preparation* involves preparing the dataset by annotating and labelling the previously scraped images with their corresponding label. This process will prepare the dataset for training.  In the *Model Training s*tage*,* Google Colabs environment will allow a model to be trained using Google's powerful GPU’s. This will reduce time and resources to execute such tasks. Training the model will make use of the YOLOv3 algorithm due to its many benefits including its best used to classify food and is extremely fast.  In the final stage *Model Inference,* is to be discussed and evaluate various aspects that determine the efficiency of the research. Such as the level of accuracy this model has and the ability to classifying between different food. |
| **Ethical Considerations:**  To train and test such model, various images where used however none of these images included the use of anyone's personal information or breaches any data protection procedures, hence following legal principles. |
| **References:**  Alpaydin, E. (2020). *Introduction to Machine Learning.* Cambridge, Massachusetts: The MIT Press.  Gerónimo, D., Serrat, J., López, A. M., & Baldrich, R. (2013). *Traffic Sign Recognition for Computer VisionProject-Based Learning.* IEEE Transactions on Education.  Le, J. (2018, April 12). *The 5 Computer Vision Techniques That Will Change How You See The World.* Retrieved from Heartbeat: https://heartbeat.fritz.ai/the-5-computer-vision-techniques-that-will-change-how-you-see-the-world-1ee19334354b  Nelson, J. (2020, January 9). *Training a YOLOv3 Object Detection Model with a Custom Dataset.* Retrieved from Roboflow: https://blog.roboflow.ai/training-a-yolov3-object-detection-model-with-a-custom-dataset/  Redmon, J., & Farhadi, A. (2018). *YOLOv3: An Incremental Improvement.* University of Washington.  Szeliski, R. (2010). *Computer vision: algorithms and applications.* Springer Science & Business Media.  Zheng, Y., Kong, J., Jin, X., Su, T., Nie, M., & Bai, Y. (2018). *Real-Time Vegtable Recognition System based on Deep Learning Network for Agricultural Robots.* China: IEE; Chinese Automation Congress . |