**RESEARCH PROPOSAL FORM**

*(also referred to as the ‘Statement of Intent Form’, or SOI)*

***To be submitted by the researcher to the Institute Research Sub-Committee (IRC)***

|  |  |
| --- | --- |
| **Research Title:**  Multi-Label Classification of Motorized Vehicles | |
| **Institute Name:**  Institute of Information & Communication Technology | |
| **Course / Programme:**  B.Sc. Software Development (Hons.) | |
| **Level and year of study**  Level 6 Year 3 | |
| **Main area of study being proposed:**  Object classification is a very common concept used to detect and classify objects in target of a digital image or a real-time video footage, this process involves a set of two classes that are trained on a larger dataset of images in order to develop a model used to classify whether other images fall under the labels used to train with. However, this process called binary classification, does not extract most characteristics present in an image. Thus, this study would explore another mean of classification; multi-label classification.  The intention of this study is to analyze, detect and classify multiple characteristic of a motor vehicle (including cars, motorcycles, trucks and busses) thus not discriminating an object under only one particular label. The reason behind this study is to challenge pre-existing studies to extract as much information from an image as possible, information that is greatly needed throughout real-life scenarios. Such use case would include tracking undocumented motor vehicles using snapshot from CCTV footage, hence aiding local authorities.  By using such technique of multi-label classification on motor vehicles, vital information such as brand of vehicle (i.e. Volkswagen, Toyota etc.), type of vehicle (i.e. motorcycle, truck, car), Color, Model (i.e. Golf Mk1, Golf Mk2 etc.). This relevant information is much more valuable rather than labeling vehicle detected with their main category (i.e. car).  Many present studies of motor vehicles classification include using similar techniques that cover basic classification models including labeling an object in an image with either a ‘car’ label or a ‘truck’, and other similar projects follow. By reevaluating those projects this study would classify more labels that are able to differentiate and filter through more vehicles under the same category i.e. instead of classifying an object as a ‘car’, such study would classify a ‘Car, Red, Gold Mk4, Volkswagen’.  This study focuses on computer vision by detecting objects in target of an image and classifying the respective object with the appropriate labels. Computer vision technologies that this study will use mainly include convolution neural network for image analyzation. Such neural network would make use of the appropriate algorithm able to be trained and classify with multiple labels. | |
| **Name of Researcher:**  Keane Schembri | **Researcher’s I.D. Number:**  216799(M) |
| **Signature of Researcher:** | **Date of submission of Form:**  5th June 2020 |
| **Name of Tutor (or Recommended Tutor):** | |

|  |
| --- |
| **Personal Motivation for the Choice of Research Theme.** |
| A common concept of classification is binary classification, where an object being detected is labelled using one from two classes. This area has been explored and exhausted enough with very low significance since limited information is classified using this approach. Thus, I propose a more complex classification where a model labels all characteristics of the object in target. This implementation will challenge already researched studies by including a more diverse dataset and a wider array of labels. This study will also challenge future studies to implement similar models on a larger scale. Computer vision is a very popular concept and I believe this topic is greatly needed to solve real-life problems. |

|  |
| --- |
| **Outline of Key Literature and Theoretical Framework or Propositions.** |
| Computer vision is a technology able to transform data from digital images into a decision or a new representation (Bradski & Kaehler, 2008). Data transformation intent is to achieve any particular goal; such goals include solving tasks, perform alteration to an existing image or to identify an object in an image. Indeed, computer vision is a study of developing techniques that help machines to visually understand the content of an image.  One technique of computer vision is 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).  Other techniques include training a model to recognize and understand the content of an image. Understanding pixels and patterns are achievable by providing a machine with a set of relevant images with their respective label and training under the correct environment (Ethem, 2010). This technique led to multiple studies and research in order to detect and classify object in an image, valuable studies include classification of human long bone localization in ultrasound imaging (Lazuardi, et al., 2019), real-time vegetable recognition based for agricultural robots (Zheng, et al., 2018) and military ship detection (Cui, et al., 2019). However, these researches have only been capable of making use of a binary classification method which is the ability to classify objects in target under two groups; with two labels.  Although this is a very impressive model to be able to detect and classify the said object, very limited information is gathered from an image. An improved method is a multi-label classification, where an objected is labelled with as many labels required that fit its characteristics. Hence, rather than only limiting an object with 2 labels under the same category, we can classify multiple other characteristics and label them accordingly.  Multi-label classification is categorized into two different subdomains; problem transformation methods and algorithm adaptation methods. The first group include algorithm independent methods which transform multi-label classification task into one or more single-label classification, regression or ranking tasks. Whilst the second group contains methods that extend existent learning algorithms capable of handling multi-label data directly (Trohidis, et al., 2008). |

|  |
| --- |
| **Significance of the Study.** |
| The significance of this study is to challenge common classification models with real-life use case scenarios. This study will improve pre-existing binary classification with multiple labels to identify various other characteristics of the object in target, in this case, motor vehicles which are a very common object to target in real-life use case scenarios. This implementation could serve as a base study to identify stolen or illegal motorized vehicles on public roads trough real-time CCTV footage. |

|  |
| --- |
| **Hypotheses and/or Research Question/s** |
| Using object detection and classification as one of computer vision subdomains, it is able to detect, classify and multi-label characteristics of a motorized vehicle through digital images.  This study will address the following research questions:   * What are the viable labels that should be considered to classify different characteristics of an object in target? * How is a reliable dataset obtained, constructed, annotated and what should it consist in order to result in the best accuracy rates? * Which algorithm would best suit for multi-label classification? |

|  |
| --- |
| **Target Participants and Research Methods for Data Collection and Analysis** |
| Multiple datasets are present for use online without cost, combined with proper labelling is required which is a process expected to be executed manually trough annotation tools which best fit the algorithm of choice. In view of correctly labelling motorized vehicles accordingly, proper information should be extracted trough each vehicle brands website thus not misinforming a model to label and image with the incorrect class. |

|  |
| --- |
| **Anticipated Contributions of the Study.** |
| It is anticipated that the relevance of this study will serve as a base implementation for future studies to implement similar models on a larger scale in real-life scenarios, this includes being a funded research project. Contribute to local criminal units in order to trace undocumented / wanted motorized vehicles. Contribute to starting a discussion regarding the implementing multi-label classifications on transportation. |

|  |
| --- |
| **Dissertation Project Plan.** |
|  |

|  |
| --- |
| **Ethical Considerations.**  **Refer to *guidance points below. You are also additionally required to read MCAST Document 074 ‘Research Ethics Policy and Procedure’ that is available on the College website via link*** [***http://www.mcast.edu.mt/179***](http://www.mcast.edu.mt/179)  *Research shall be conducted in such a manner so as to avoid any psychological and physical harm to humans and animals and financial damage to organizations:*   1. *Only the supervisor and examiners will have access to any data gathered.* 2. *Participants will remain free to withdraw from the study at any time without having to provide any reason. In the case of withdrawal, all the records and information collection will be deleted.* 3. *The participant, who is the sole proprietor of the data provided, is granting that such data would be processed for this study purposes only.* 4. *The data collection process will be a transparent process.* 5. *All transcriptions and/or electronic recordings reflecting the data collected, once exhausted, are to be deleted* 6. *Confidentiality, anonymity and data protection procedures are to be ethically abided by.* 7. *The researcher would provide a soft copy of the study to the participant, if required.* |
| *Enter details here regarding possibility of issues regarding confidential personal data:*  Dataset used during this study will be of motorized vehicles thus not violating laws regarding confidential personal data. To prevent from breaking this legislation, any confidential information that could be present in said dataset i.e. where drivers identify is visible, should be blurred accordingly and not violated.  Any information related to this study where participants identity is used should not be divulged. |
| *Enter details here regarding possibility of physical harm:*  This study is to be researched under a computer-related form thus not present to any form of physical harm. |
| *Enter details here regarding possibility of moral harm:*  Due to this study being a computer-related task without the need of participants involvement, this study would not be exposed to any form of moral harm. |
| *Enter details here regarding possibility of business harm:*  Although brand, make and type of vehicle will be classified, this study will discuss neither advantages nor disadvantages of the classified object. Information used to label objects in target such as make and type has been made public by the respective company on the release of a vehicle. |

**List of Key References:**

|  |
| --- |
| References  Bradski, G. & Kaehler, A., 2008. *Learning OpenCV,* California: O’Reilly Media, Inc.  Cui, H. et al., 2019. *Ship Detection: An Improved YOLOv3 Method,* France: OCEANS 2019 - Marseille.  Ethem, A., 2010. *Introduction to machone learning,* s.l.: MIT Press.  Lazuardi, R. A. F. et al., 2019. *Human Bone Localization in Ultrasound Image Using YOLOv3 CNN Architecture,* Indonesia: CENIM.  Szeliski, R., 2010. *Computer vision: algorithms and applications,* s.l.: Springer Science & Business Media.  Trohidis, K., Tsoumakas, G., Kalliris, G. & Vlahavas, I., 2008. *Multi-label classification of music into emotions.,* s.l.: ISMIR.  Zheng, Y.-Y.et al., 2018. *Real-Time Vegetables Recognition System based on Deep Learning Network for Agricultural Robots,* China: CAC. |

***This section is to be filled in by the representative of the Institute Research Sub-Committee (IRC) prior to forwarding of this Form to the ‘MCAST Research Ethics Committee’ for final ethics approval:***

|  |  |  |
| --- | --- | --- |
| ***Nature of Ethical Consideration*** | ***Outcome (Tick)*** | ***Comments/Advice*** |
| All ethical issues have been adequately tackled. |  |  |
| Possibility of issues regarding misuse of data or some form of harm. |  |  |

|  |  |
| --- | --- |
| **Details of Representative to the Institute Research Sub-Committee.** | |
| Name | Signature |
| Designation | Date |

**Annex 1: Participant Information Letter**

***Sample: ***

**Title of Research: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

You are being invited to take part in a research study. Before you decide to participate, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information.

**What is the purpose of the study?**

This research is being undertaken on…

**Why have I been chosen?**

You have been chosen because…

**Do I have to take part?**

It is up to you to decide whether or not your take part. If you decide to take part you will be given this information sheet to keep and be asked to sign a corresponding consent form.

**What will happen to me if I take part?**

You will then be given a questionnaire on.../your data will be used…/your image will be used…

**What are the possible disadvantages and risks of taking part?**

There are no disadvantages or risks foreseen in taking part in the study.

**What are the possible benefits of taking part?**

By taking part you will be contributing to the development of a set of recommendations for…

**What if something goes wrong?**

If you wish to complain or have any concerns about any aspect of the way in which you have been approached or treated during the course of this study, please contact…(researcher is to give his/her MCAST email as a contact)

**Will my details be kept confidential?**

All information which is collected about you during the course of the research will be kept strictly confidential so that only the researcher carrying out the research will have access to such information and will not be shared with any other individuals. Participants should note that data/images collected from this project may be retained and published in an anonymized form. By agreeing to participate in this project, you are consenting to the retention and publication of data.

**What will happen to the results of the research study?**

The results will be written up into a dissertation for my final project of my Bachelor…

**Who is organizing the research?**

The research is conducted as part of a degree in …

**Who may I contact for further information?**

If you would like more information about the research before you decide…(researcher is to give his/her MCAST email as a contact)

*Thank you for your interest in this research…*

**Annex 2: Participant (or Guardian) Consent Form**

***Sample: ***

**Title of Research: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Name of Researcher: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

# Please initial box

1. I confirm that I have read and understand the Information Letter

for the above study and have had the opportunity to ask questions.

1. I understand that my/my charge’s participation is voluntary and that I/my charge am/are free to withdraw at any time without giving any reason.

3. I agree to allow my daughter/son/charge to take part in the above study.

*(Statement 3 is to be included only when guardians/parents are involved in giving consent)*

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_

Name of Participant/

Guardian Date Signature

Researcher Date Signature

*1 for participant; 1 for researcher*