**MSc. in Computing**

**Practicum Approval Form**

# Section 1: Student Details

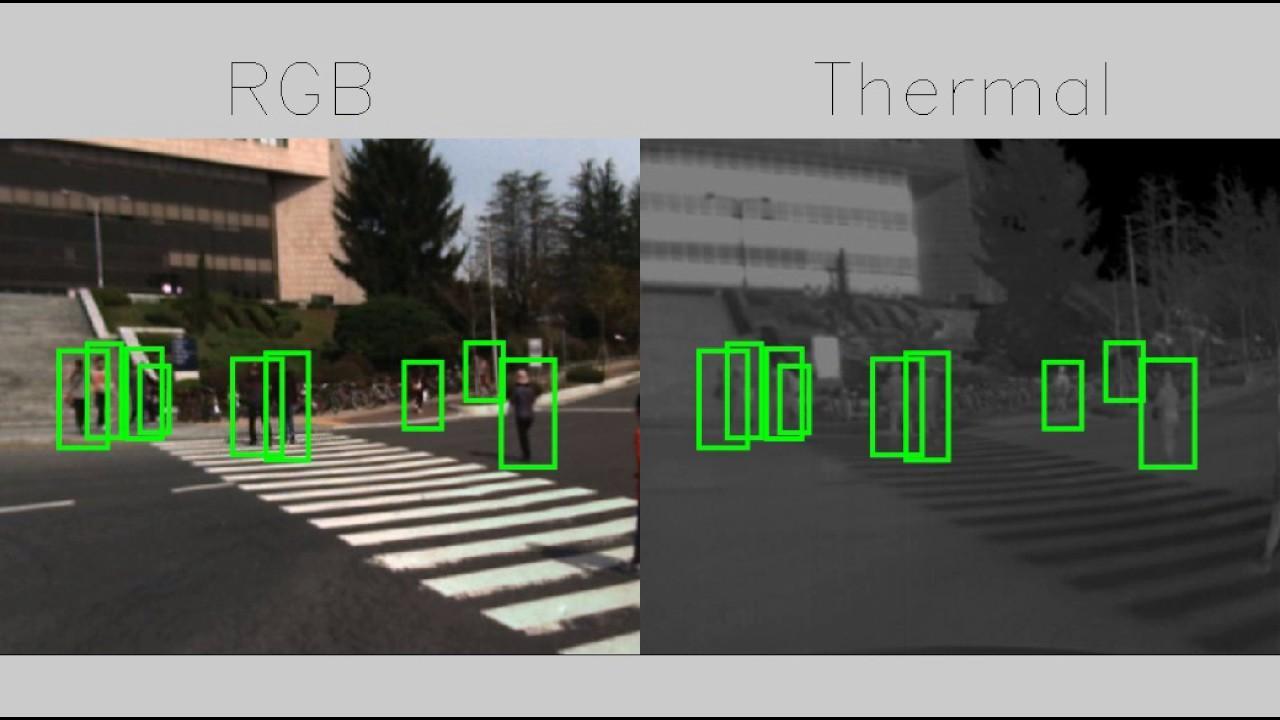
|  |  |
| --- | --- |
| Project Title: | Analysis of multimodal imagery for autonomous driving |
| Student ID: | 19211077 |
| Student name: | Shivani Firke |
| Student email | shivani.firke2@mail.dcu.ie |
| Chosen major: | Data Analytics |
| Supervisor | Suzanne Little |
| Date of Submission | 8th November 2019 |

**What is the topic of your proposed practicum? (100 words)**

Project Proposal: Analysis of multimodal imagery for autonomous driving.

Autonomous Driving is an effort to ease the commute, save lives, and to make efficient utilization of time. Object detection plays a key role in understanding the surroundings and helps autonomous vehicles to make appropriate decisions. We would implement a benchmark algorithm that would compare the images of a thermal camera and RGB camera to check the performance of the algorithm on the dataset so that there can be a possibility of implementing infrared cameras in autonomous vehicles.

The limitation with RGB cameras is such that the quality of the images captured depends on lighting conditions such as low light conditions or glare from headlights of oncoming vehicles. However, thermal imaging can overcome this limitation as thermal cameras work independently of lighting conditions and capture images based on heat signatures. But thermal imaging has a shortcoming of its own, that is, thermal images would be less informative for objects that share temperature in a similar range. We plan to develop and illumination-aware system by assigning weights to the features and training the neural network to detect and eliminate less relevant information.



**Please provide details of the papers you have read on this topic (details of 5 papers** expected).

1. Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi, “You Only Look Once: Uniﬁed, Real-Time Object Detection”, *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779 - 788, June 2016*
2. Y. Choi et al., "KAIST multi-spectral day/night data set for autonomous and assisted driving", *IEEE Transactions on Intelligent Transportation Systems, vol. 19, no. 3, pp. 934-948, Mar. 2018.*
3. A. González, Z. Fang, Y. Socarras, J. Serrat, D. Vázquez, Xu J., A.M. López, “Pedestrian detection at day/night time with visible and fir cameras: A comparison”, *Sensors, Multidisciplinary Digital Publishing Institute (MDPI), 16 (6) (2016), p. 820*
4. Yuxiang Sun, Weixun Zuo and Ming Liu, “RTFNet: RGB-Thermal Fusion Network for Semantic Segmentation of Urban Scenes”*, IEEE Robotics and Automation Letters, Volume 4, July 2019, pp. 2576-2583*
5. C. Li, D. Song, R. Tong, M. Tang, “Illumination-aware Faster R-CNN for robust multi-spectral pedestrian detection,'' Pattern *Recognition, vol. 85C, (2019), pp. 161–171.*

**How does your proposal relate to existing work on this topic described in these papers?** (200 words)

Computer vision in autonomous vehicles is still a challenging problem despite significant advances. One of the major obstacles in this is varying light conditions during daytime and night as well as glare from different light sources. From the referred papers, it can be inferred that the combination of visible and non-visible imaging techniques, namely RGB and infrared, can increase the detection accuracy. Our aim is to assess and compare the accuracy of a benchmark algorithm on different imaging modalities. Specifically, we want to compare detection accuracy on test images recorded at day and night time if trained using (a) plain color images; (b) just infrared images; and (c) both of them. We plan to evaluate the algorithm on the publicly available KAIST multispectral dataset and the FLIR dataset.

From the referred literature, we have inferred that both RGB and thermal imaging have their own limitations and the current work in this field is being carried out to develop a fusion RGB-T architecture, so as to eliminate the individual limitations. But nevertheless, this architecture still suffers from the drawback that in certain situations, the RGB or thermal image may be more informative than the other. For example, in the case of objects with similar temperatures would pose a problem for thermal imaging and would be less informative. To overcome this problem, the proposed solution would be assigning weights to the features based on its contribution to the information or discarding it completely in case it contributes negligibly.

**What are the research questions that you will attempt to answer? (200 words)**

* Can an RGB camera and a thermal camera be used in tandem in autonomous vehicles for better results in object detection?
* Is the performance of the benchmark algorithm independent of the type of image?
* Is the neural network capable of detecting and eliminating redundant information to reduce the problem of false detection?

**How will you explore these questions? (Please address the following points. Note that three or four sentences on each will suffice.)**  
- *What software and programming environment will you use?*

We will be using Python for data processing and training the neural network, and libraries such as Keras, TensorFlow, Pytorch, etc.

*- What coding/development will you do?*

We plan to implement a benchmark algorithm to compare the performance of that algorithm on RGB and thermal images and develop a neural network capable of assigning weights to the features in thermal images in order to detect and use only contributive information.

*- What data will be used for your investigations?*

We plan to use the KAIST dataset for RGB and Thermal images and FLIR dataset for Thermal images if required.

KAIST Dataset: <https://soonminhwang.github.io/rgbt-ped-detection/>

FLIR Dataset: <https://www.flir.com/oem/adas/adas-dataset-form/>

- *Is this data currently available, if not, where will it come from?*

Yes, both of the above-mentioned datasets are publicly available.

- *What experiments do you expect to run?*

We plan to run the algorithm on RGB and thermal images in the KAIST dataset to evaluate its performance on both. In addition to this, we plan to use a weighting mechanism based on illumination for the features in images and train the network to detect information based on illumination.

*- What output do you expect to gather?*

We expect to gather performance results for a certain benchmark algorithm on both RGB and thermal images for autonomous vehicles. Furthermore, we expect our neural network to be able to give lesser weight to less relevant information and eliminate it.

*- How will the results be evaluated?*

The results will be evaluated based on performance metrics such as miss rates and accuracy.