MSc. in Computing
Practicum Approval Form

Section 1: Student Details

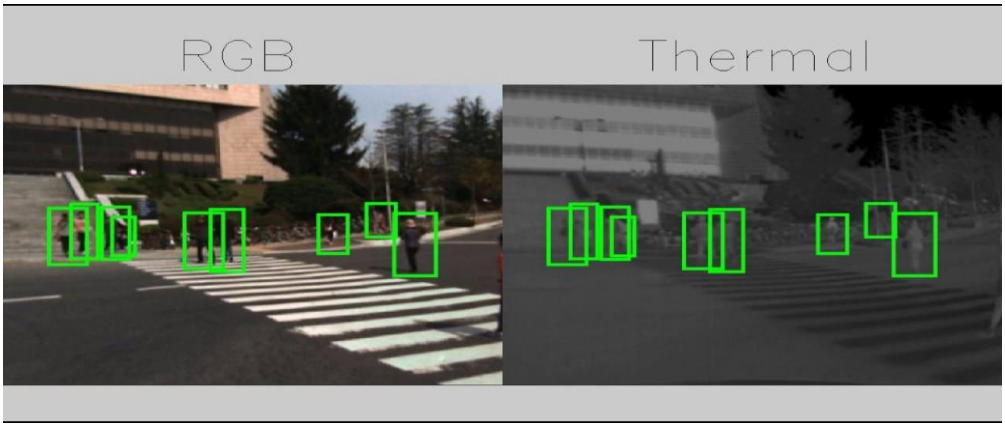
Project Title:	Analysis of Multimodal Imagery for Autonomous Driving
Student ID:	19211077, 19210993
Student name:	Shivani Firke, Anshika Sharma
Student email	shivani.firke2@mail.dcu.ie, anshika.sharma27@mail.dcu.ie
Chosen major:	Data Analytics
Supervisor	Suzanne Little
Date of Submission	29th November 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, N. Kim, S. Hwang 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, 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)

1. This paper introduces a single network, object detection algorithm, You Only Look Once. As the name suggests, YOLO frames the bounding boxes by assigning each object a confidence score and classify the object on the basis of how confident it is about the categorization of the object. The paper compares the performance of YOLO to other object detection algorithms like, DPM and R-CNN. As a result, YOLO out performs both the algorithm as it is a one-time evaluation, although it has its own drawbacks.
2. This paper provides a detailed study of multi-spectral KAIST Dataset. This dataset comprises of a wide collection of images for autonomous driving. From urban to residential, day and night, including sunrise, sunset, afternoon, morning, night and dawn. Both RGB and thermal camera are coaligned to attain the same orientation in both images. A detailed examination of RGB, Thermal, fusion of RGB and Thermal and LiDAR images along with the appropriate camera specifications is presented.
3. This paper examines the accuracy gain of different pedestrian models (holistic, part-based, patch-based) when trained with images in the far infrared spectrum. Comparison of detection accuracy on test images recorded at day and night-time is done, if trained and tested using (a) plain color images; (b) infrared images; and (c) both of them. It proposes an early fusion approach to combine features from both modalities, i.e., color and infrared images.
4. This paper points out that the primary drawback of RGB semantic segmentation networks is that RGB images are susceptible to degradation with inadequate lighting conditions while thermal cameras generate images using thermal radiation emitted by objects and are independent of lighting conditions. So, this paper proposes to fuse

both the RGB and thermal information in a novel deep neural network. An Encoder-Decoder design concept and ResNet is employed for feature extraction.

5. This paper compares six convolutional network fusion architectures and analyses their adaptations, enabling a vanilla architecture to obtain detection performances comparable to state-of-the-art results. Based on the discovery that pedestrian detection confidences from color to thermal images are correlated with illumination conditions, it proposes an illumination-aware Faster R-CNN to give an illumination measure of the input image

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 a neural network, and libraries such as Keras, TensorFlow, Pytorch etc as required.

- What coding/development will you do?

We plan to implement the YOLO algorithm to compare its performance 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 relevant 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 YOLO 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 miss rates and accuracy.

Division of Work:

TASKS	Anshika Sharma	Shivani Firke
Dataset Collection	Multi-Spectral KAIST Dataset	FLIR Thermal Dataset
Algorithm Implementation	Implement YOLO algorithm for RGB	Implement YOLO algorithm for thermal
Evaluation of Algorithm	Evaluate the algorithm on training RGB dataset	Evaluate the algorithm on training Thermal dataset
Comparison of Evaluation Metrics	Comparison with other RGB detectors	Comparison with other thermal detectors
Illumination-aware Architecture	Create illumination information weights for input	Use RGB architecture or RGB-T fusion architecture to test the illumination information input
Visualization	Create and analyse visualizations	Create and analyse visualizations