PhD Project Proposal

ECIT Interdisciplinary PhD Programme

Proposed Project Title: Towards Safe AI: Improve Adversarial Robustness for General Computer Vision Tasks

Principal Supervisor(s): Dr Yang Hua

Project Description:
Thanks to the resurgence of deep learning, in recent years, computer vision has achieved stunning progress and been applied to many applications, e.g., face recognition and autonomous driving. However, despite the near, and sometimes beyond, human-like results on several benchmarks, we still do not understand the true reliability and robustness of this approach, which will cause potential safety issues if we deploy these algorithms directly to real-world scenarios. For example, so-called adversarial examples, almost indistinguishable from natural data to the human eye, can be generated based on adversarial training and cause state-of-the-art classifiers to make incorrect predictions with high confidence. Therefore, it is urgent to improve adversarial robustness for general computer vision tasks, including classification, detection and recognition, by considering both the theoretical and practical challenges for AI safety.

Objectives:

The goal of this Ph.D. project is to develop novel algorithms to improve adversarial robustness for general computer vision tasks. The main work directions as below:
- Investigate different approaches to generate adversarial samples to attack/fool the existing state-of-the-art framework on general computer vision tasks, including classification, detection and recognition
- Develop novel algorithms that defend against adversarial-example attacks on general computer vision tasks.
- Utilize general adversarial approach to mutually improve the power of adversarial attack and defence.

Academic Requirements:

Students entering the programme will normally be required to have a 2.1 BSc/BEng in Computer Science, Electrical and Electronic Engineering, or a maths based engineering or physical science degree, or equivalent qualification recognised by the University. Students holding an appropriate MEng or MSc (Software conversion) will normally be required to have a 2.1 or commendation (distinction) respectively. Furthermore, additional criteria may be applied. All applicants must have significant mathematical and programming experience.

GENERAL INFORMATION:
This 4 year PhD studentship, potentially funded by the Department for Employment and Learning (DEL), commences on 1 October 2019.

Eligibility for both fees and maintenance depends on the applicants being either an ordinary UK resident or those EU residents who have lived permanently in the UK for the 3 years immediately preceding the start of the studentship. Non UK residents who hold EU residency may also apply but if successful may receive fees only.

Applicants should apply electronically through the Queen's online application portal at: https://dap.qub.ac.uk/portal/

Deadline for applications: Friday 1 March 2019

Contact details:

Supervisor Name: Yang Hua
QUB Address: ECIT, Queens Road, Queen's Island,
Belfast BT3 9DT

Tel: +44 (0)28 9097 1816
Email: Y.Hua@qub.ac.uk