|  |  |  |  |
| --- | --- | --- | --- |
| **First Supervisor** | Dr Daqi Liu | | |
| **School/Department** | School of Computing and Mathematical Sciences | | |
| **Email** | dl336@leicester.ac.uk | **Telephone Ext** |  |

**Section 2 – *Project Information***

|  |  |  |
| --- | --- | --- |
| **Project Title** | Multi-modal scene understanding | |
| **Project Highlights:** | 1. | Proposing balanced learning techniques to mitigate the generic biased prediction issue. |
| 2. | Improving the holistic inference capabilities of the multi-modal scene understanding systems. |
| 3. | Interpreting the ever-increasing visual data accurately and efficiently. |
| **Project Summary** | | |
| **In the current information age, interpreting the ever-increasing visual data generated by social media or other sources is urgently needed across many related sectors, e.g., robotics, security, or the creative industries. Multi-modal scene understanding research is the key within the above interpretation tasks, which is essentially a critical aspect of computer vision that involves not only identifying objects in a scene but also understanding their relationships. However, such research is still far from desired due to various challenges, such as biased prediction, inefficiency, and inferior inference capabilities.**  **One primary aim of this PhD project is to solve the above challenges by proposing advanced deep learning-based multi-modal scene understanding paradigms. These paradigms will improve the holistic inference capabilities as well as the overall efficiencies of the multi-modal scene understanding systems. Moreover, the generic biased prediction issues caused by the imbalanced distributions of the training datasets will also be mitigated.**  **The applications of this project are broad and varied, e.g., object detection, scene graph generation, image captioning, or visual question answering. The overarching aim is to interpret the ever-increasing visual data so that the tremendous semantic information embedded within the above visual data can be properly exploited to meet urgent needs across various related sectors.** | | |