COMP4900 report March 2

Professor Xu Dan’s seminar on visual scene understanding really improved my understanding on related research area such as machine learning and big data processing. He first introduced the necessity of the usage of visual scene, which is partially because that the price of human resource is rather high comparing to machine. He then described some scene of use of visual scene, for example, visual scene display and visual scene understanding. The definition of visual scene display is that it supports context and personalisation. Its goal is to create a more personalised and situational AAC system to reduce cognitive needs and make learning easier, and support social interaction and thought exchange. He also gave the point that the visual scene understanding system nowadays are still heavily rely on the traditional pipelines, which definitely slow the whole system and made the cost of building a visual scene understanding system rise extraordinarily. The scientist are now focus on developing a system called Deep learning based visual SLAM system, which can reduce this kind of disadvantage, but the challenges are mainly focus on the Key-frame detection, global pose optimization, 3D

reconstruction, which are all cutting-edge technology nowadays. I gained a lot from this seminar, and also got a lot more familiar with the visual scene understanding system.