Statement of Purpose

Adam Kosiorek

My journey with computer science began at the best tech university in the country, where I enrolled in the BSc in Robotics. I learned a lot about physics and mechanics, material science and manufacturing technologies; not until my first programming class, however, have I realized how a truly intelligent entity could be controlled. Paired with my passion for vision and perception, which I have developed as a photojournalist, it steered me towards my first computer vision course. I immediately knew that this intersection of image processing and machine learning was the area I wanted to focus on. Without any systems good enough for real world deployment, it was challanging as well as aligned with my original passion for intelligent machines.

To learn from the best, I went for an internship to the Computer Vision Lab at Samsung R&D in Warsaw. My project was to design and implement a Bag of Words-based pipeline for object classification. Consequently, I learned about the state-of-the-art in keypoint detection and feature description, which later helped me to appreciate the importance of automatically learned low level features in convolutional neural networks (CNNs). I focused my research mainly on the visual codebook generation within Bag of Words. The final step was to optimize it and port to the Android platform. In my bachelor thesis I investigated how spatial information affects classification performance by implementing a similar system for Kinect-gathered point cloud classification. Afterwards, I continued my work at Samsung, where I turned to deep learning techniques for object classification and image duplicate detection.

My observation was that deep learning, although often superior to older approaches, required huge computational resources and datasets. To mitigate this problem I went to study computational science at the Technical University in Munich. With cutting-edge computer vision research groups it seemed a great choice. Currently, I am working with Caner Hazirbas and Rudolph Triebel of the Computer Vision group headed by Prof. Cremers. In our recent project we investigate the introspective capacity of neural networks. It is understood as the ability of a classifier to assess uncertainty of its prediction given the data [ref]. It is a paramount problem in mobile robotics or medicine where a wrong classification might lead to loss of life. While high classification accuracy is desired, it is impossible to assess whether a given prediction is accurate in a test setting, where no labels are available. It is, therefore, vital to assess uncertainty of predictions. Our intermediate results show that neural networks, augmented with additional layers and a novel cost function, can be jointly trained for classification and uncertainty estimation. The topic might expand into my master thesis.

This summer I did an internship at Bloomberg in London, where I worked on fraud detection in financial transactions. The problem can be cast as unsupervised anomaly detection with further verification in a supervised settings. I learned how difficult it can be to introduce an innovative approach in a corporate setting. This, together with my earlier

industrial experience, convinced me that I do want to pursue a PhD. I love solving scientific problems that do not have "the best" solution or a reference specification by going into the deepest details.

In my doctoral research project I would like to focus on deep neural networks for reinforcement learning and recurrent neural networks. The former fascinates me since it is similar to how humans learn. Specifying a reward function or designating an expert in the inverse problem is far easier than providing supervised training data. Recurrent neural networks, on the other hand, are well suited to sequential information processing. If beliefs about content could be introduced in a form of priors inferred from previous elements of the sequence, it might be possible to increase object classification accuracy in videos and to shrink the network size. It would reduce the amount of computation needed, possibly leading to an energy efficient system for portable devices. Another interesting problem is an end-to-end RNN for optical flow computation. The only end-to-end neural network, while efficient and accurate, does not use temporal information. It computes optical flow for any two possibly unrelated images [ref]. I would like to investigate RNN designs capable of computing optical flow while presented with a single image at a time. It might lead to a better accuracy with fewer computations, thus better energy efficiency.

Passion for solving technical problems coupled with my demonstrated skills show that I am ready and extremely well motivated to carry out research as a PhD student at the University of Oxford. I am confident that I will be able to contribute significantly to and benefit immensely from my stay in Oxford. Thank you for considering my application.