MSc Seminar

Wednesday August 23, 2023 at 11am in Reynolds 1101

Kristina Kacmarova

*Generation of Synthetic Datasets for
Vision-Based Animal Monitoring*

Advisor: Dr. Stacey Scott
Advisory: Dr. Neil Bruce

Abstract:

Precision Livestock Farming (PLF) is a rapidly growing field that aims to use technology to improve the health and welfare of farm animals. This includes the development of innovative monitoring techniques through the use of machine vision. While some machine vision tasks, such as animal identification are fairly reliable, more complex tasks such as the detection of health issues are still in their early stages due to being unable to generate enough data for rare cases. Ensuring data quality can be challenging, as the existence of imbalanced classes within datasets can lead to overfitting. Creating relevant datasets for PLF research is not an easy task, as it requires significant time and financial investments.

The goal of this project is to address the challenges associated with real-world data collection for vision-based deep learning models. The solution proposed is the creation of synthetic datasets. Synthetic data is artificially created data that mimics the characteristics of real-world data but can be generated at a much larger scale and with more diversity than what is possible with real-world data alone. This allows for more efficient and effective training of machine learning models, especially when there are sparse or imbalanced datasets. Synthetic datasets offer greater control over data variability (e.g., rare conditions) and can be annotated automatically. This approach can save time and resources while still providing reliable data for training computer vision models. By evaluating the quality, diversity, and realism of the generated data, this study seeks to provide valuable insights into the potential of different solutions for overcoming imbalanced datasets, data scarcity and for improving the overall performance of animal-related machine learning tasks.