400-Level Project Proposals FYDNI 2023

A list of possible 400-level projects is given below. You do not have to restrict yourself to this list, and can make up your own project topic. However, to do so you will have to arrange for a supervisor. You can find more information about the research interests of the staff members on the Computer Science Department Web pages.

You should submit your project selections via BlackBoard* by Friday 4th March. Before making these selections you should talk to the supervisors of the projects you are interested in. In most cases the best way is to email them to make a time to do so – supervisors' email addresses are given with each project.

In order to allocate the projects as fairly as possible we ask that you give first, second, and third choice projects, and that you choose projects from three different supervisors.

There are two different project papers - COSC480 and COSC490. If you are enrolled in the Computer Science Honours degree you take COSC490, students in the first year of a Masters take either paper, and most others take COSC480. Most projects are suitable for either COSC480 or COSC490, but a few might be only suitable for COSC480, which will be indicated after the project title.

*On BlackBoard select the COSC480_490_FYDNI_2023 paper, and then Project/Research Project/Assessment/Project Choice.

1. **Searching sparse learned document representation**

   Andrew Trotman (andrew.trotman@otago.ac.nz)

   Modern search engines apply transformers to documents to get a representation that, when searched, produce high-quality results. These representations are not well understood, and the effect of throughput in an impact-ordering search engine has not been studied sufficiently. Indeed, there are two such open source research search engines, implemented in different languages, and using different optimisations. In this project we will investigate the strengths and weaknesses of each of the differences in order to provide insights into how to best implement search over sparse learned document representations.

2. **Adding a poly emulator to MAME**

   Andrew Trotman (andrew.trotman@otago.ac.nz)
   David Eyers

   During the 1980s Polycorp (a New Zealand company) designed and manufactured computers – the Poly series – that were based on the 6809 CPU. We have some of those computers, some of the ROMs, some of the disks, and some of the Operating System source code. In this project we want to take the resources we have and add them to MAME, the Mega-Emulator. This will require extensions to MAME's existing 6809 source code to emulate parts of the CPU that only the Poly used, as well as implementing the machine as a subsystem in MAME. Once we have achieved this the Poly will be preserved (in software) and made as openly available as we can, for those with an interest in the history of New Zealand's own computing industry.

3. **Does stemming really work?**

   Andrew Trotman (andrew.trotman@otago.ac.nz)

   Search engines such as Google and Bing "stem" your search terms. That is, if you enter "running" it will also search for "run" and "ran". But does this really produce better results? There are several stemming algorithms already published, and some authors have implied that stemming can be learned using a deep learning. But when do stemmers work and why do stemmers work? In this project we will build a dictionary stemmer by scraping the Wictionary, and seeing which words are related and how. We will then use this stemmer to measure the performance of other stemmers – in particular as document collections grow in size and documents grow in lengths. Finally, we will try to learn a new stemmer using the data from the Wictionary.

4. **Real-time banjo music transcription**

   Andrew Trotman (andrew.trotman@otago.ac.nz)

   Some musical instruments, such as the Banjo, can either play chords or pure notes. Indeed, there is a style of Banjo (called Clawhammer) in which notes are typical and chords are rare. It is typical to play Old Time music in this style (Old Time is similar to Bluegrass). As the name suggests, these tunes are old, and they have been passed from generation to generation acoustically. In this project we will build a system that, in real time, will listen to a piece of music being played by a Banjoist and transcribe it to musical score (Tab) in real time. Ideally this will run on a mobile phone so that we can take it to skilled players and transcribe some tunes as they are being played.

5. **Augmenting Māori Rock Art**

   Steven Mills (steven.steven.mills@otago.ac.nz)
   Stefanie Zollmann

   There are hundreds of Māori rock art sites around the South Island, some of which are accessible to the public. Understanding these sites can be difficult, however, since the images can be weathered and hard to make out. This project will develop digital tools to provide augmentation and interaction to help people experience these culturally significant works. There are two main aspects to this work – an interactive display for an exhibition scheduled to open in late 2023, and an on-site Augmented Reality tool to help visitors find and understand rock art near Moeraki. This project is a collaboration with Otago Museum and Te Rūnanga o Moeraki, and will likely involve visiting the rock art sites.

6. **Voxel Video Completion**

   Steven Mills (steven.steven.mills@otago.ac.nz)
   Lech Szymanski

   The Ātea project, a collaboration between academics and Te Rūnanga o Awarua in Bluff has developed an experience where users can meet in a virtual replica of the marae and play back recordings of stories told by kaumatua. The people (both the participants and recorded storytellers) are rendered as 'voxel videos', using regular grids of blocks like in Minecraft or a Lego model. Voxel videos have many advantages, but our capture system often misses parts of the figures. This project will use deep learning to complete and smooth out the models, and potentially increase their resolution. You will get experience with the latest deep learning methods, and contribute to an exciting project with great cultural value.

7. **360° Video Transitions**

   Steven Mills (steven.steven.mills@otago.ac.nz)
   Holger Regenbrecht

   Seamless Integration of multiple 360° videos into one VR Scene From a given set of (loop-able) 360° videos (Insta360 X3, ambisonic sound Zoom H3 VR) we want to create an immersive user experience which allows users wearing a head-mounted display (e.g. Meta Quest2) to transition from one 360 video position to the next in a seamless as possible fashion. This project would involve Unity or Unreal programming to create the interactive experiences, and could be extended to integrate reconstructed 3D objects into the VR scene with a meaningful transition between 360° video and 3D VR

8. **Capturing 3D Shape**

   Steven Mills (steven.steven.mills@otago.ac.nz)

   Recovering 3D shape from multiple images is a common task in computer vision. While the fundamental algorithms are still many areas to investigate. This is a mix-and-match project with a lot of different things you might explore. I'm interested in learning to segment objects from the background in images for reconstruction; stitching together partial reconstructions; building a turntable-based device to automatically capture and segment images; evaluating and extending readily available 3D reconstruction packages; etc. Each of these could be a project in themselves, or they could be combined in interesting ways.

9. **Casual Collaboration**

   Geoff Wyvill (geoff@otago.ac.nz)
   David Eyers

   A 3D, collaborative sculpture application (Sculpture) has featured at many of the recent University Science Expos. After consulting a brief help-screen (and possibly not even doing that!) users of the provided computer terminals very rapidly begin to build (and sometimes destroy) a 3D sculpture collaboratively. The overall success of this exhibit stems in part from its ease of configuration and its usability.

   This project seeks to abstract and redesign the application platform underlying Sculpture, to help modernise the hosting of the application itself, and to support other similar collaborative systems to be built and demonstrated in future.

   Today's operating systems, web frameworks, game engines, and other forms of middleware provide tools for building cooperating software that were not available when Sculpture was first implemented. Further, high quality open source libraries are available that implement conflict-free replicated data types (CRDTs). CRDTs enable asynchronous collaborative editing such as within Google Docs. They specifically address how to avoid inconsistencies developing between distributed replicas of an artefact, as asynchronous editing occurs. The artefact in the case of Sculpture is the copy of the sculpture contained within the memory of each participating, connected computer.

   For a 490 project you would build a new application on top of the collaborative framework, or continue extension and evaluation of the framework itself.

10. **Tartini Reloaded**

    Geoff Wyvill (geoff@otago.ac.nz)
    David Eyers

    Tartini is an open source application that helps practical musicians and music students. Its main function is to display the pitch and loudness of each note as a passage of music is played. Unlike an ordinary commercial tuner, it does this continuously and responsively so the musician can discover technical faults such as pitch drop at the end of notes for wind players and poor bow changes for string players. It records the sound so a complete passage can be analysed note by note afterwards and imperfections identified.

    Tartini was created between 2003 and 2008 as a PhD project, and continues to run on old (quarantined) computers within the Department. Geoff and David have demonstrated Tartini at the University's Science Expo, including during the event last year.

    The goal of this project is to rework Tartini to run on a modern software platform. One possible target is a recent macOS version. A more ambitious target would be porting Tartini to the web. It is likely that JavaScript within a web browser now runs sufficiently quickly to handle Tartini's digital signal processing code.

    More information about Tartini can be found here.

11. **Finding EEG biomarkers for depression**

    Zhiyi Huang (zhiyi.huang@otago.ac.nz)
    Veronica Liesaputra

    Depression is a global problem, especially under the ongoing Covid and other global issues. Brain wave (EEG) is the most convenient way to investigate the status of the brain. It has rich information about the brain's mental status. In this project, we are going to use deep neural networks to classify the EEG data collected from depressed patients. It can help find the biomarkers for depression. We had a successful previous COSC490 project which uses Layer-wise Relevance Propagation (LRP) to find EEG biomarkers. In this project, we are going to use synthesised EEG data to study the capability of LRP in terms of retrieving the latent biomarkers embedded in EEG data. It is an extended study of the previous COSC490 project.

12. **Visualising EEG data with music and nature**

    Zhiyi Huang (zhiyi.huang@otago.ac.nz)
    Veronica Liesaputra

    EEG data has information of the brain's mental status, such as excitement and calmness. According to this information, it would be helpful for people to understand their mental status by visualising the mind state with music and natural scenes. Alternatively, it would be helpful to use music and natural scenes to alter the mental status, which is called neurofeedback treatment. In either way, we need to transform the EEG information to music and natural scenes such as flowing water or beautiful flowers. In this project, we will investigate the possible ways to visualise EEG data.

13. **Emotion recognition of legged robot**

    Zhiyi Huang (zhiyi.huang@otago.ac.nz)
    Lech Szymanski

    The department purchased a legged robot recently. It can walk under the guide of a remote controller. However, it is not autonomous enough or smart enough. For example, it cannot recognise the emotion of the human master. In this project, we are going to enable emotion recognition in the legged robot by acquiring facial images and EEG data from the human master. We will use affective facial images to train a deep learning network to recognise the different emotions. In addition, we will use EEG data for the network to understand fake facial emotion and thus enhance emotion recognition. Ideally the outcome of the project could be applied into the legged robot and enable it to dance when the master is happy or lie down when the master is sad. The SDK of the robot is available for us to control the robot through C++ or Python.

14. **Robot dog**

    Brendan McCane (brendan.mccane@otago.ac.nz)

    We have a robot dog. This project will involve implementing some cool demos on the dog. For example, some simple dog behaviours: heel, follow, sit etc. Then perhaps some more complicated behaviours: e.g. fetch. And finally, if all that goes well, then some learnt behaviours based on direct feedback from humans.

15. **Detecting cancer**

    Brendan McCane (brendan.mccane@otago.ac.nz)

    I have a dataset related to detection of cancer in cells. Some initial experiments indicate that deep networks are extremely good at distinguishing cancer cells from non-cancer cells. This project will involve more thorough experimentation on the data set including implementation/application of several different classification methods.

16. **Deep trouble**

    Lech Szymanski (lech.szymanski@otago.ac.nz)
    David Eyers

    Transfer learning facilitates convenient reuse of machine learning models. Rather than training from scratch, in the context of reuse, partial retraining is performed over an existing source model. For example, training can be applied over a neural network with weights copied from the source network that remain fixed but for within a few final layers. Is there, from a security perspective, a potential risk of original training process embedding malicious data? Consider an unrealistically oversimplified, hypothetical scenario, in which a target neural network is trained, by transfer learning from a source neural network, to recognise faces for the purpose of controlling access to a building. Is it possible to embedded a pattern in the source capable of 'surviving" target training and later causing a strong response that may act as a "master key"?

    This project will involve devising means of embedding source networks with "malicious" patterns. The aim is to determine the extent to which it is possible to have these patterns survive the transfer learning and whether they can affect a target network in any meaningful way.

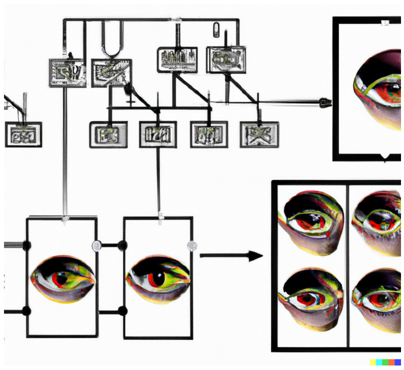
17. **Stochastic parrot meets University policy**

Lech Szymanski (lech.szymanski@otago.ac.nz)
David Eyers

The University has **lots** of policy documents. Can large language models reliably "make sense" of the University policy database? Can AI help us to navigate the essence of these policies, and what is ambiguous or inconsistent within them? Dialogue systems such as ChatGPT can accept a few thousand input symbols at present, which has to cover the entire dialogue to be processed. This is unlikely to be enough context to process the types of documents we are proposing. We will probably have to fine-tune the underlying statistical language model on our dataset to encode the University-specific knowledge into the network. The work on this project will involve creation and preparation of the training dataset (from various University documents), investigating creative ways of providing ChatGPT with task-specific context, and/or using transfer learning techniques to fine tune GPT for the specific task of advising on University policies.

18. **Deep recollections**

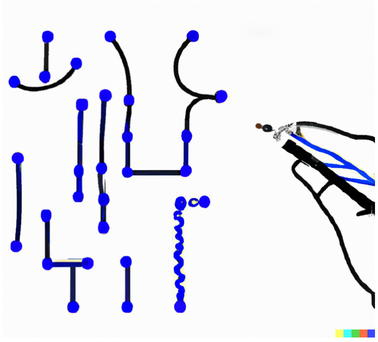
Lech Szymanski (lech.szymanski@otago.ac.nz)



Artificial neural networks and deep learning works, surprising well sometimes. The key aspects of this success are: availability of massive amounts of training data, unreasonably large and very deep network models (convolutional, if we are dealing with images), and optimisation methods that force the network to overtrain without sacrificing its generalisation capabilities. This is surprising because the latter of these two aspects are precisely what machine learning theory says not to do if you want a model to extract *useful* patterns from data. So, does a neural network learn *useful* patterns, or does it just memorise its input? The aim of this project is to investigate the degree of memorisation in neural networks. For example, if we repurpose a network pre-trained for classification to do auto-encoding, how much, and what type of information on the input gets passed through the network? Will this be a viable measure of memorisation? If so, can we use it to track memorisation over different layers of the network? Can get a glimpse of the patterns the network pays attention to?

19. **Teaching neural networks to write**

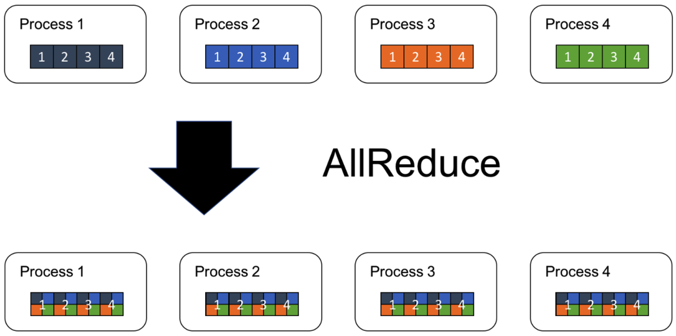
Lech Szymanski (lech.szymanski@otago.ac.nz)



Convolutional neural networks (CNNs) can outperform humans on visual tasks, such as object classification from images. However, these models often fail when presented with specially crafted adversarial examples that would not fool a human. What if, in addition to training the network to read, we were to teach it to write (by controlling a virtual pen). Would the CNN learn to see the writing more akin to the way we see it? For this project you will be developing a CNN- based agent and train it using supervised and reinforcement learning (RL) techniques to re-construct written text from images. This project will give you a background in machine learning (specifically CNNs and RL) and practical experience with TensorFlow.

20. **Optimising AllReduce in Distributed Deep Learning**

Yawen Chen (yawen.chen@otago.ac.nz)
Zhiyi Huang



Communication among GPUs is one of the many challenges when training distributed deep learning models in a large-scale environment. The latency of exchanging gradients over all GPUs is a severe bottleneck in data-parallel synchronised distributed deep learning. In synchronised data-parallel distributed deep learning, the major computation steps are: 1. Compute the gradient of the loss function using a minibatch on each GPU. 2. Compute the mean of the gradients by inter-GPU communication. 3. Update the model. To compute the mean, a collective communication operation called "AllReduce" is used. This project aim to investigate optimisation of AllReduce in Distributed Deep Learning under different type of interconnection systems (e.g. different topologies and scenarios).

21. **Understanding communication in federated learning**

Haibo Zhang (haibo.zhang@otago.ac.nz)

Federated learning has emerged as a new paradigm of distributed learning that trains a global prediction model across multiple devices without the need to share the training data. However, there is limited understanding of how communication protocols specifically contribute to the performance of federated learning. This project aims to set up a testbed to capture the traffic generated by federated learning and analyse its traffic pattern. If time allows, a communication protocol for federated learning will be designed and evaluated in experiments.

22. **Efficient wireless communication for UAV swarms to avoid obstacles**

    Haibo Zhang (haibo.zhang@otago.ac.nz)

    Unmanned Aerial Vehicle (UAV) swarm systems have enabled many long-haul, delay sensitive and hazardous applications. As UAVs are powered by batteries and the airspace they work in is usually highly dynamic, planning trajectories in an energy-efficient way to avoid both UAV-to-UAV and UAV- to-obstacle collisions has become a challenging problem. We have designed a path planning method for UAV swarms to avoid obstacles based on Particle Swarm Optimisation (PSO) and Artificial Potential Field (APF). However, our method assumes the wireless channel used for exchanging position and velocity information among UAVs is reliable. This project aims to study the impact of channel reliability on the robustness of collision avoidance, and design an efficient wireless communication scheme to improve the performance of collision avoidance in the presence of channel fading.

23. **Machine learning based link scheduling for wireless communication**

    Haibo Zhang (haibo.zhang@otago.ac.nz)

    A wireless channel is a shared channel. If multiple devices send data to the same receiving device over the same channel simultaneously, all transmitted data are typically corrupted and discarded due to collisions. Time Division Multiple Access (TDMA) is well-known medium access control protocol to schedule different packet transmissions in different time slots to avoid collisions. However, constructing the optimal transmission schedule to achieve high transmission reliability is a challenging problem since the channel quality sensed by each transmitting device can change over time. This project aims to investigate the feasibility of using federated learning to schedule wireless transmissions. It is expected that federated learning could enable multiple devices to collaboratively learn a shared prediction model of the wireless channel without the need of exchanging training samples among different devices.

24. **AI for code generation and explanation**

    Anthony Robins (anthony.robins@otago.ac.nz)

    Survey the ways in which AI systems (e.g. GitHub Copilot, ChatGPT) are being used for code generation. In particular, explore errors in generated code. Is there any pattern? Are there similarities or differences across languages? What is the best way to detect errors, and can the systems be used to repair errors? Explore the use of code explanation in similar ways. This is an open ended project, let's see where it takes us...

25. **Attractor spaces in Hopfield nets**

    Anthony Robins (anthony.robins@otago.ac.nz)

    Hopfield networks (a kind of neural network) have dynamic behaviour that can be characterised in terms of gradient descent in a multidimensional attractor space. Such spaces have some interesting properties – and some serious problems. This project will involve implementing and exploring a Hopfield type network to further develop our understanding of its dynamic behaviour.

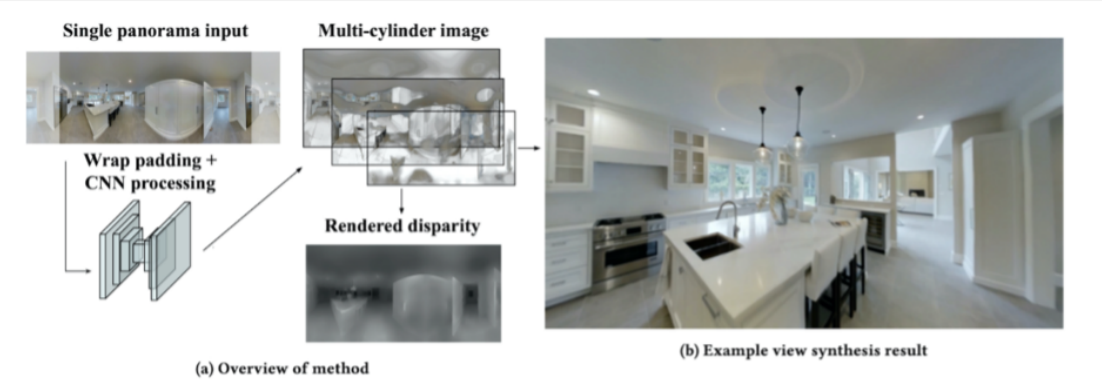
26. **Visualising deep networks**

Anthony Robins (anthony.robins@otago.ac.nz)
Lech Szymanski

Deep networks are powerful – but mysterious. This project will involve reviewing current methods for visualising the behaviour of deep networks, implementing an existing method or methods, and possibly developing a new method or methods. This project will involve the use of a standard deep learning toolkit such as TensorFlow.

27. **PanoSynthVRVideo**

Stefanie Zollmann (stefanie.zollmann@otago.ac.nz)
Jonathan Ventura (Assoc. Prof. visiting from Cal Poly, US)



PanoSynthVR is an approach that provides 360-degree view synthesis on current virtual reality hardware (such as an Oculus Quest) from a single panoramic image input. Our lightweight method automatically converts a single panoramic input into a multi-cylinder image representation that supports real-time, free-viewpoint view synthesis rendering for virtual reality. Currently, this approach is limited to single images.

In this project, we will work on extending this approach to videos considering a temporal component and making sure that the output is temporally consistent. You will work with an Oculus Quest and implement data processing in python, rendering will be done in WebXR.

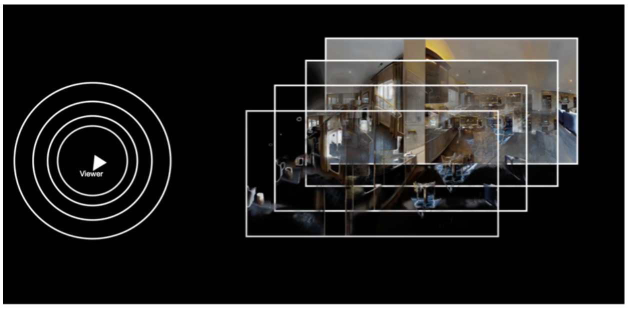
28. **Stable VR360 outpainting**

Stefanie Zollmann (stefanie.zollmann@otago.ac.nz)
Jonathan Ventura (Assoc. Prof. visiting from Cal Poly, US)

In this project we will work on expanding monocular videos so that they can be rendered in VR. Video captured on your smartphone, for example, have a limited field-of-view: when viewed in VR, most of the scene will appear black. We aim to "fill in" the missing parts of the scene using an AI technique called "outpainting." In the example above, we simply asked Stable Diffusion to fill in missing pixels outside of the image. We would like to further explore this technique to improve the image quality and ensure temporal coherence throughout the video. This project would suit a student that has an interest in VR and immersive technology as well as in image and video processing and AI.

29. **6DOF Video Benchmark**

Stefanie Zollmann (stefanie.zollmann@otago.ac.nz)
Jonathan Ventura (Assoc. Prof. visiting from Cal Poly, US)



In our previous work, we developed a system called PanoSynthVR that can produce a complete 3D surround-view scene from a single panoramic image input. To improve upon that work, we would like to build a large dataset suitable for developing and evaluating panoramic view synthesis methods. Existing datasets contain only low resolution panoramas, are synthetically generated, and are restricted to mostly indoor scenes. We propose to build a larger and more diverse dataset for learning high-resolution panoramic view synthesis. Using a pole or helmet-mounted camera, we will collect high-res 360-degree video both outdoors and indoors using a consumer panorama camera. We can also collect 360 videos from YouTube VR travel channels. We will reconstruct the camera trajectories using existing open source software. We will aim to match the scale of existing datasets and provide enough examples for neural network training. Once we have built the dataset, we can evaluate existing methods and explore the development of improved methods using the data.

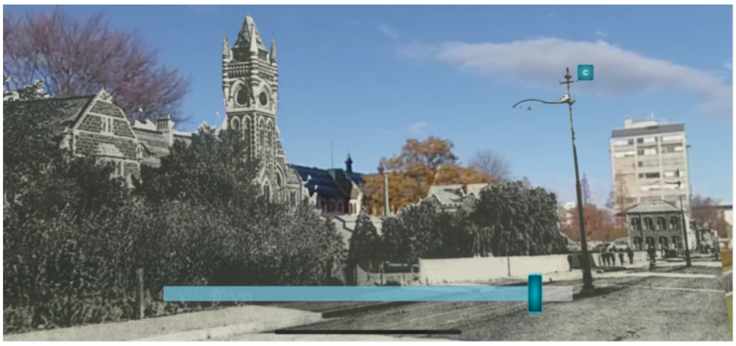
30. **Geolocated Augmented Reality**

Stefanie Zollmann (stefanie.zollmann@otago.ac.nz)
Jonathan Ventura (Assoc. Prof. visiting from Cal Poly, US)

Geo-referenced Augmented Reality (AR), such as provided by Google's Geospatial API, allows developers to create Augmented Reality (AR) applications by attaching content to outdoor locations. Thereby visual positioning services (VPS) allow to localise users and provide content right in place. Some of the application scenarios for this type of technology are navigation, tourism, sightseeing, gaming or even sports events. While this technology has great potential, the impact of different technical factors (such as how accurate users can be localised or how much latency is created) on the user experience within these systems is not well-researched yet. In this project, we will develop an AR application that can be used on mobile phones and allows us to further investigate these aspects. This project requires an interest in AR technology, openness to work with Unity or other AR content creation SDKs, as well as an interest in working on user experience.

31. **Augmented Reality Rephotography**

Stefanie Zollmann (stefanie.zollmann@otago.ac.nz)
Jonathan Ventura (Assoc. Prof. visiting from Cal Poly, US)
Tobias Langlotz (InfoScience)



In this project, we are interested in investigating novel ways to explore historical photographs in the present context. For many applications, a before-after view can be helpful to understand changes, e.g. to monitor the progress of a building project, historical applications, tourism or even environmental change analysis. We developed a prototype that aligns historical photographs and integrates them into the current view of the users. However, there are still challenges, such as ensuring that dynamic objects in the surrounding are well integrated and not just covered by the historic photograph and providing a suitable user interface that allows an easy transition between past and present.
This project will require a strong interest in Augmented Reality and an openness to learn new tools such as Snap Technology's Lensstudio. You will also have an opportunity to work with exciting smart glasses hardware such as the Snap Spectacles.

32. **Long-form Summarisation**

    Veronica Liesaputra (veronica.liesaputra@otago.ac.nz)
    Andrew Trotman

    With the advent of neural architectures, we see significant advances in automatic text summarisation systems. However, due to the input length limit, it is still not possible to summarise an entire book or movie. In this project, you will be creating/using a hierarchical deep learning model to summarise a long document.

33. **Have I heard this story before?**

    Veronica Liesaputra (veronica.liesaputra@otago.ac.nz)
    Andrew Trotman

    People can identify similar narratives in everyday life. For example, an analogy with the Cinderella story may be made in describing the unexpected success of an underdog in seemingly different stories. Currently, similar stories are clustered based on the list of words that are in the story, regardless of the words' order. Thus, search engines often return many irrelevant stories. In this project, we are trying to create a neural story embedding that identifies similar stories based on the interactions of the main characters in the story.

34. **Classifying Students' Emotions and Learning Experiences**

    Veronica Liesaputra (veronica.liesaputra@otago.ac.nz)
    Claudia Ott

    Multiple studies have shown that students' emotional states influence their general well-being and their learning success. With the advances of AI and sensors, there are many techniques we can use to automatically detect human emotions through various modalities such as facial expressions, tone of voice, text, posture, etc. However, the accuracy for detecting emotion and learning experiences is low. In this project, we will use an aspect-based extraction model to identify students' emotions and learning experiences based on their reflective writings.

35. **Few-shot object detections**

    Veronica Liesaputra (veronica.liesaputra@otago.ac.nz)

    Deep neural networks have transformed computer vision field, enabled in part by the development of large and diverse sets of curated training data. However, in many image classification tasks, labeled examples are not available. In this project, you will create/use a few-shot image classifier that could produce labels for a given set of unlabelled images based on few labelled images per class.