RMIT UNIVERSITY

Deep Learning
COSC 2779/2972 | Semester 2 2023
Assignment 1: Introduction to Deep Convolutional Neural Networks
(Action Unit and Emotion Recognition)

Assessment Type	Individual assignment. Submit online via Canvas → Assignments → Assignment 1. Marks awarded for meeting requirements as closely as possible. Clarifications/updates may be made via announcements/relevant discussion forums.
Due Date	5.00pm, Friday 25 August 2023 (Week 6)
Marks	30%

1 Overview

In this assignment you will explore a real dataset to practice the typical deep learning process. The assignment is designed to help you become more confident in applying deep learning approaches. In this assignment you will:

- Develop a deep learning system to solve a real-world problem.
- Analyse the output of the algorithm(s).
- Research how to extend the DL techniques that are taught in class.
- Provide an ultimate judgement of the final trained model that you would use in a real-world setting.

To complete this assignment, you will require skills and knowledge from lecture and lab material for Weeks 1 to 6 (inclusive). You may find that you will be unable to complete some of the activities until you have completed the relevant lab work. However, you will be able to commence work on some sections. Thus, do the work you can initially, and continue to build in new features as you learn the relevant skills. *A deep learning model cannot be developed within a day or two. Therefore, start early.*

2 Learning Outcomes

This assessment relates to all of the learning outcomes of the course which are:

- Discuss and critically analyse a variety of neural network architectures; Evaluate and Compare approaches and algorithms on the basis of the nature of the problem/task being addressed.
- Synthesise suitable solutions to address particular machine learning problems based on analysis of the problem and characteristics of the data involved.

- Communicate effectively with a variety of audiences through a range of modes and media, in particular to: interpret abstract theoretical propositions, choose methodologies, justify conclusions and defend professional decisions to both IT and non-IT personnel via technical reports of professional standard and technical presentations.

- Develop skills for further self-directed learning in the general context of neural networks and machine learning; Research, Discuss, and Use new and novel algorithms for solving problems; Adapt experience and knowledge to and from other computer sciences contexts such as artificial intelligence, machine learning, and software design.

3 Assessment details

3.1 Task

Automatically detecting facial expressions has become an increasingly important task. It involves computer vision, machine learning and behavioral sciences, and can be used for many applications such as security, human-computer-interaction, driver safety, and health-care. Some interesting articles on this topic include (*"take the predictions and conclusions below with a pinch of salt"*): Businesswire, Financial Times, morphcast, The Atlantic, Washington Post, While there are many practical utilities of emotion recognition technology, there are many ethically challenging scenarios raised by such technology as well. A good understanding of the underlying technology will allow us to navigate such challenging topics.

In this assignment, you will develop an **end-to-end trained deep convolutional neural network (CNN)** to identify facial emotions and FACS (Facial Action Coding System) codes present in an image. For this we will use a subset of the The Extended Cohn-Kanade Dataset (CK+) provided on canvas Assignment 1 page. Read the Linked paper to get a good understanding of the data set. The CK+ dataset consists of 560 labeled images from 123 subjects. The task is to design a thorough experiment and determine how well the following two attributes can be predicted:

- **High Level Emotion:** Does the image show "Positive", "Negative" or "surpriced" subject?

- **FACS codes**: Does the image show a particular facial action (FACS code)? This is a yes(1)/no(0) output.

Figure 1 shows a couple of examples. The descriptions of the FACS codes can be found in Table 1 of the paper: Patrick Lucey et. al, The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression, CVPR 2010.

Please read the following **requirements and guidelines** carefully.

- You need to design a network that takes in an image as input and predicts the emotion class and the FACS codes. For higher grades (CR/DI/HD) you should develop a **single network** that makes all predictions.

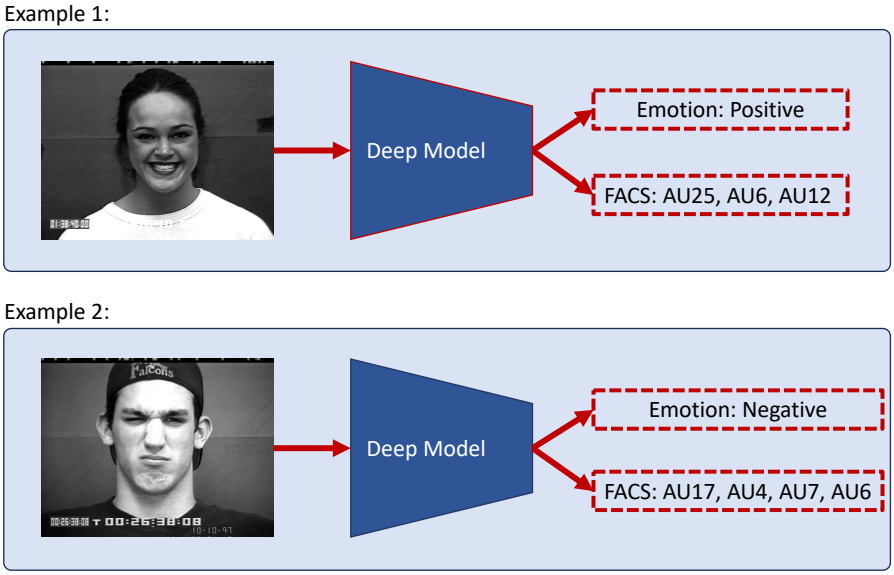

Figure 1: Examples of Emotion and facial action recognition pipelien.

- You may use pre-trained networks as part of your solution. However, there need to be "clearly identifiable" network segment(s) that is designed and trained by you. You should show how this segment is developed (tuned) in you code.
- For higher grades (DI/HD) you should write your own data loader.
- Only neural network based techniques can be used in the assignment. Other ML techniques such as SVM, RF cannot be used.
- (Report) You need to come up with a deep learning system, where each element of the system is *justified* using data analysis, performance analysis and/or knowledge from relevant literature.
- (Report) You should clearly explain your evaluation framework, including how you selected appropriate performance measures, and how you determined the data splits.
- (Report) Finally you need to analyse the results from your model using appropriate techniques and establish how adequate your model is to perform the task in the real world and discuss limitations if there are any (**ultimate judgment**).
- (Report) While there are many practical utilities of emotion recognition technology, there are many ethically challenging scenarios raised by such technology as well. discuss such challenges in general and biases in the dataset. **Don't need lengthy paragraphs. A point-form answer is acceptable**.

3.2 Dataset

The data set for this assignment is available on Canvas. There are the following files:

- "`cohn-kanade-images`": Contain all the images that are to be used for model development.

- "`data_labels.csv`": Contain file names and the label of the images. This data is to be used in developing & evaluating the models.
- "`CK+AgreementForm.pdf`": The agreement to use data. Need to fill in, sign and upload before the end of week 3.

The original data is from CK+ dataset published with *Lucey, Patrick, Jeffrey F. Cohn, Takeo Kanade, Jason Saragih, Zara Ambadar, and Iain Matthews. The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In 2010 IEEE computer society conference on computer vision and pattern recognition-workshops, pp. 94-101. IEEE, 2010*

Licence agreement: The dataset can only be used for the purpose of this assignment. Sharing or distributing this data or using this data for any other commercial or non-commercial purposes is prohibited.

4 Suggested Schedule

We expect that you will start the assignment immediately and follow a schedule similar to the one shown below. Do not fall behind, *A deep learning model cannot be developed within a day or two.*

- **Week 2-3**: Read the specification and familiarize yourself with the problem. Explore the data set and the task. Identify biases and ethical issues and start writing the report.
- **Week4**: Design the experiments. Develop the data loading mechanism. Search relevant literature and read. Start writing the report.
- **Week5**: Develop the model - design, and train. Do model analysis. Update the report.
- **Week6**: Do model analysis. Update the report. Submission!

5 Submission

You have to submit all the relevant material as listed below via Canvas.

1. Signed "CK+AgreementForm". Need to be submitted by Week 3.
2. **A report** (of no more than 4 pages , plus up to 2 pages for appendices) critically analysing your approach and ultimate judgement. Should be in PDF format.
3. Your **code** (Jupyter notebooks) used to perform your analysis. Should be a ZIP file containing all the support files. The Jupyter notebook(s) should be clearly commented in markdown format (see labs and lectorial exercises). **The final outputs should be visible**.

The submission portal on canvas consists of two sub-pages. page one for report submission and the second page for code and other file submission. More information is provided on canvas. Include only source code in a zip file containing your name. We

strongly recommend you to attach a README file with instructions on how to run your application. Make sure that your assignment can run only with the code included in your zip file! Include a PDF version of your report.

After the due date, you will have 5 business days to submit your assignment as a late submission. Late submissions will incur a penalty of 10% per day. After these five days, Canvas will be closed and you will lose ALL the assignment marks.

Assessment declaration:
When you submit work electronically, you agree to the assessment declaration - `https://www.rmit.edu.au/students/student-essentials/assessment-and-exams/assessment/assessment-declaration`

6 Teams

Not relevant. This is an individual assignment.

7 Academic integrity and plagiarism (standard warning)

Academic integrity is about honest presentation of your academic work. It means acknowledging the work of others while developing your own insights, knowledge and ideas. You should take extreme care that you have:

- Acknowledged words, data, diagrams, models, frameworks and/or ideas of others you have quoted (i.e. directly copied), summarised, paraphrased, discussed or mentioned in your assessment through the appropriate referencing methods
- Provided a reference list of the publication details so your reader can locate the source if necessary. This includes material taken from Internet sites. If you do not acknowledge the sources of your material, you may be accused of plagiarism because you have passed off the work and ideas of another person without appropriate referencing, as if they were your own.

RMIT University treats plagiarism as a very serious offence constituting misconduct. Plagiarism covers a variety of inappropriate behaviours, including:

- Failure to properly document a source
- Copyright material from the internet or databases
- Collusion between students

For further information on our policies and procedures, please refer to the following: `https://www.rmit.edu.au/students/student-essentials/rights-and-responsibilities/academic-integrity`.

8 Marking guidelines

A detailed rubric is attached on canvas. In summary:

- Approach 40%;

- Ultimate Judgment, performance & Analysis 30%;
- Discussion on Ethical issues and biases 20%;
- Implementation & Report Presentation 10%;

Approach: You are required to use a suitable approach to find a predictive model. Each element of the approach need to be *justified* using data analysis, performance analysis and/or published work in literature. *This assignment isn't just about your code or model, but the thought process behind your work.* The elements of your approach may include:

- Setting up the evaluation framework
- Selecting CNN architecture, loss function and optimization procedure.
- Hyper-parameter setting and tuning
- Identify problem specific issues/properties and solutions

Ultimate Judgement: You must make an *ultimate judgement* of the "best" model that you would use and recommend in a real-world setting for this problem. It is up to you to determine the criteria by which you evaluate your model and determine what is means to be "the best model". You need to provide evidence to support your ultimate judgement and discuss limitation of your approach/ultimate model if there are any.

Discussion on Ethical issues and biases: While there are many practical utilities of emotion recognition technology, there are many ethically challenging scenarios raised by such technology as well. discuss such challenges in general and biases in the dataset. **Don't need lengthy paragraphs. a point-form answer is acceptable**.

Critical Analysis & Report: Finally, you must compile a report describing and analysing the approach that you have taken to find a suitable model and make your ultimate judgement. Your report ***must*** be no longer that 4 pages , plus an additional 2 pages for appendices. The appendices must only contain references, figure, diagram, or data tables that provide evidence to support the conclusions and statements in your report.

Any over length content, or content outside of these requirements will not be marked. For example, if you report is too long, ONLY the first 4 pages pages of text will be read and marked.

In this report you should describe elements such as:

- Your final selected approach
- Why you selected this approach
- Parameter settings and other approaches you have tried.
- Limitation and improvements that are required for real-world implantation.

This will allow us to understand your rationale. We encourage you to explore this problem and not just focus on maximising a single performance metric. By the end of your report, we should be convinced that of your ultimate judgement and that you have considered all reasonable aspects in investigating this problem.

Remember that good analysis provides *factual statements, evidence and justifications for conclusions* that you draw. Statements such as:

"I did xyz because I felt that it was good"

is not analysis. This is an unjustified opinion. Instead, you should aim for statements such as:

"I did xyz because it is more efficient. It is more efficient because . . . "