RMIT UNIVERSITY

Deep Learning
COSC 2779/2972 | Semester 2 2022
Assignment 1: Deep Neural Networks for 3D Shape 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	Week 7, Friday 9th September 2022, 05:00pm
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

Recognizing 3D shape is a fundamental task in computer vision, which is widely used in autonomous driving, robotics and scene understanding. Most autonomous vehicles use "*Light Detection and Ranging*" (Lidar) technology to sense the environment. Lidar creates precise, three-dimensional (3D) representation of the environment where the output is a set of 3 dimensional points in space (point-cloud). Here is a quick video introduction to what lidar is: `https://youtu.be/PRg5RNU_JLk`. Extracting high level information from point-clouds is vital for applications and one key task is to recognise 3D shapes from point-clouds. The question is how can we use deep learning to do this task?

Many techniques has been proposed for this task over the years. One of the first (and popular) techniques proposed for this was Multi-view CNN[1] in 2015. This technique was motivated by how humans identify 3D shape. Human recognizes 3D shapes by observing the projection from different viewpoints of the object. By adopting many viewpoints and mining the relationship between them, this intuitive idea, which infers the categories from 2D rendered images, has proven to be effective in computer vision. A modified version of the pipeline from Multi-view CNN for shape recognition is shown in 1.

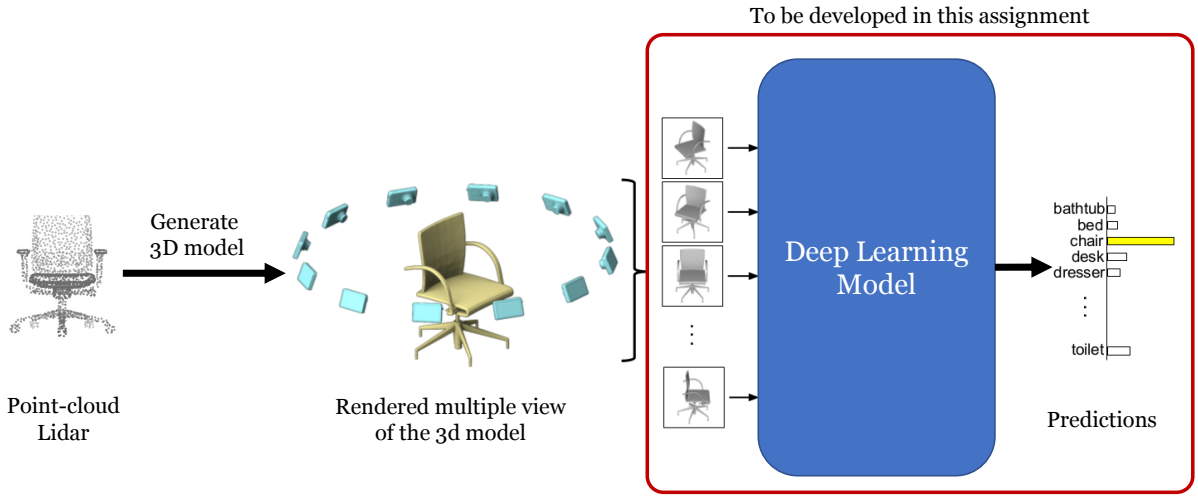

Figure 1: Pipeline of Multi-view CNN

[1]Hang Su et. al.Multi-view convolutional neural networks for 3d shape recognition ICCV 2015.

In the above pipeline, first the 3D point-cloud is converted to a 3D model. Then, multiple 2D views of the 3D model is rendered using a graphics package. These 2D views can now be classified into the relevant class using a deep network.

In this assignment, you will develop an **end-to-end trained deep convolutional neural network (CNN)** to identify the 3D shape from a set of 2D views. For each object you will get four 2D views where each 2D view is an image of the size $(64, 64, 3)$. Each 3D shape belongs to one of 20 classes.
Please read the following requirements and guidelines carefully.

- You need to design a network that takes in 2D image(s) as input, and predict the class (outputs). For higher grades (CR/DI/HD) you should develop a **single network** that uses all the four views to make a prediction.
- *It is not acceptable to develop multiple stand alone single network that output class and then do post-processing to derive the final prediction.* When using multiple views, they should be integrated to a single network.
- 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 real world and discuss limitation if there are any (**ultimate judgement**).
- Predict the result for the test set (the final test set will be released 1 day prior to the due date).

3.2 Dataset

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

- "`README.md`": Description of dataset.
- "`data.zip`": Contain all the images that are to be used for model development.
- "`data.csv`": Contain file names (for the 4 views) and the label of the 3D shape. This data is to be used in developing the models.

- "`test_data.csv`": **Will be released later (one day prior to the due date)**. Contain file names (for the 4 views). You need to predict the class label for this data and submit the prediction via canvas. The teaching team will use this data to evaluate the performance of the model you have developed.

- "`s1234567_predictions.csv`": **Will be released later (one day prior to the due date)**. Shows the expected format for your predictions on the unseen test data. You should organize your predictions in this format. Any deviation from this format will result on zero marks for the results part. Change the number in the filename to your student ID.

The original data is from ShapeNet published with *Chang, A.X., Funkhouser, T., Guibas, L., Hanrahan, P., Huang, Q., Li, Z., Savarese, S., Savva, M., Song, S., Su, H. and Xiao, J., 2015. Shapenet: An information-rich 3d model repository. arXiv preprint arXiv:1512.03012.*

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.*

- **Week4**: Read the specification and familiarize with the problem. Explore the data set.

- **Week5**: Design the experiments. Develop the data loading mechanism. Search relevant literature and read. Start writing the report.

- **Week6**: Develop the model - design and train. Do model analysis. Update the report.

- **Week7**: Do model analysis. Predict test set. Complete the report. Submission!

5 Submission

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

1. **A report** (of no more than 3 pages , plus up to 2 pages for appendices) critically analysing your approach and ultimate judgement. Should be in PDF format.

2. A **set of predictions** from your ultimate judgement for the test data. Should be in CSV 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 & Analysis 20%;
- Performance on test set (Unseen data) 30%;
- 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.

Performance on test set (Unseen data): You must use the model chosen in your ultimate judgement to predict the class on unseen testing data (provided in `test_data.csv`). Your ultimate prediction will be evaluated, and the performance of all of the ultimate judgements will be published.

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 3 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 3 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. A 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 . . . "*