# COMP47250 Team Software Project
# Deep Metrics For Visual Retrieval
# Project Plan


Team Stay Safe

<Zhao Yuan zhao.yuan@ucdconnect.ie>

<Ning Tao ning.tao@ucdconnect.ie>

<Yili Lai yili.lai@ucdconnect.ie>

<Hanyuxi Zhou hanyuxi.zhou@ucdconnect.ie>

<Yikai Wang yikai.wang@ucdconnect.ie>

<Yuping Tian yuping.tian@ucdconnect.ie>



University College Dublin


June, 2020

## 1 Project Objectives

This project is dedicated to get a fair comparison between existing deep metric learning (DML) models and try to propose a new, improved metric with state of the art performance.

Basic goal**:**

- Implement and evaluate 12 models, conduct detailed studies and comparisons of their strengths and weaknesses.
- Develop an image retrieval system using 6 of these models.
- Build up a web page to display the achievements above.

Ultimate goal**:**

- Try to create an innovative model with better performance.
- Utilize the innovative model and achieve a best image retrieval system.

In the background learning phase, we will learn more about deep metric learning by understanding its principles, application scenarios and the training methods corresponding to different algorithms.

Task:

- Learn systematically about related terms, algorithm principles and related applications and write a literature report.

In models' reproduction phase, we will implement different metrics and understand their principles, compare the differences between them with the same parameters and conditions.

Task:

- Implement 12 evaluation algorithms.
- Develop a webpage to display our findings.

In the research and innovation phase,we will try to come up with a new metric with better performance.

Task:

- Generate a model that could be used as a test to compare with the benchmark.
- Write a report which could introduce the idea and code and display the result of our metric (with the experiment data).

In the summarize phase, we will summarize our work and deploy an accessible front-end web page, which can access the image retrieval model that we deployed on the back-end and provide services.

Task:

- Develop a user interface to apply on the web and make it available for users to access our model.
- Deploy a backend end server for our model.
- Write the final report.

## 2 Project Plan

<table>
<tr><td rowspan="5">Phase 1</td><td colspan="2">Background Learning Phase</td><td>Week</td><td>End date</td></tr>
<tr><td>Aim</td><td>Understanding the background of deep metric learning through papers</td><td rowspan="4">1 - 3</td><td rowspan="4">14/6/2020</td></tr>
<tr><td>Activities</td><td>Task 1. Learn the concept of deep metric learning<br>Task 2. Learn the processing of deep metric learning and the different methods of each process<br>Task 4. Learn the concept of losses<br>Task 5. Learn any relative knowledge in this area</td></tr>
<tr><td>Milestones</td><td>I. Understand relative knowledge of deep metric learning<br>II. Write a literature review<br>III. Build a simple model demo according to understanding knowledge</td></tr>
<tr><td>Deliverables</td><td>A literature review report</td></tr>
<tr><td rowspan="5">Phase 2</td><td colspan="2">Models Reproduction Phase</td><td>Week</td><td>End date</td></tr>
<tr><td>Aim</td><td>Reproducing some models from the paper of the past few years</td><td rowspan="4">4 - 6</td><td rowspan="4">5/7/2020</td></tr>
<tr><td>Activities</td><td>Task 1. Learn 12 algorithms in papers from 2006 to 2019<br>Task 2. Reproduce algorithms according to the papers<br>Task 3. Evaluate each models<br>Task 4. Analyse the performance and benchmark of each models</td></tr>
<tr><td>Milestones</td><td>I. Have all models constructed<br>II. Have all models evaluated<br>III. Analyse and benchmark all models</td></tr>
<tr><td>Deliverables</td><td>I. 12 existing deep metric learning models<br>II. Analysis and benchmark report of 12 models</td></tr>
<tr><td></td><td colspan="2">Innovation phase</td><td>Week</td><td>End date</td></tr>
</table>

<table>
<tr><td rowspan="4">Phase 3</td><td>Aim</td><td>Constructing a new model which could achieve better performance</td><td rowspan="4">6 - 10</td><td rowspan="4">2/8/2020</td></tr>
<tr><td>Activities</td><td>Task 1. Proposing a new deep metric learning algorithm<br>Task 2. Implementing proposed model<br>Task 3. Evaluating the new model<br>Task 4. Benchmark the new model with reproducing models</td></tr>
<tr><td>Milestones</td><td>I. Build the new model<br>II. Test the new model<br>III. Benchmark the new model with existing models</td></tr>
<tr><td>Deliverables</td><td>I. A innovative deep metric learning model<br>II. Benchmark report</td></tr>
<tr><td rowspan="5">Phase 4</td><td colspan="2">Document and Presentation Preparation Phase</td><td>Week</td><td>End date</td></tr>
<tr><td>Aim</td><td>Writing the report and preparing anything help to present the project</td><td rowspan="4">6 - 13</td><td rowspan="4">18/8/2020</td></tr>
<tr><td>Activities</td><td>Task 1. Visualize the outcome<br>Task 2. Write a project report<br>Task 3. Build a website<br>I. Website would present the final outcome and any relative data source of this project<br>II. Website would be able to access models<br>Task 4. Prepare the final presentation</td></tr>
<tr><td>Milestones</td><td>I. Write the final report<br>II. Build the outcome website<br>III. Prepare presentation</td></tr>
<tr><td>Deliverables</td><td>I. Report<br>II. Final presentation<br>III. A website</td></tr>
</table>

*Table 1: Specific Plan*

## 3 Roles

**Technical Roles:**

<u>Understanding Phase:</u> All members will work on learning related knowledge in this phase.

Implementation Phase: As shown in table below, every group member will take responsibility for 2 of the 12 algorithms including corresponding loss types suggested by Microsoft mentors.

| Team Member | Algorithm | Loss Type |
|---|---|---|
| **Zhao Yuan** | MultiSimilarity (MS) | Embedding |
| | ProxyNCA | Classification |
| **Yikai Wang** | Contrastive | Embedding |
| | CosFace | Classification |
| **Ning Tao** | Triplet | Embedding |
| | Normalized Softmax (N. Softmax) | Classification |
| **Yili Lai** | Margin | Embedding |
| | ArcFace | Classification |
| **Hanyuxi Zhou** | FastAP | Embedding |
| | SoftTriple | Classification |
| **Yuping Tian** | Signal to Noise Ratio Contrastive (SNR) | Embedding |
| | MS+Miner | Embedding |

*Table 2: Algorithm Allocation*

Innovation Phase: All team members will work on creating new evaluation methods and perform all procedures.

Wrap Off Phase: Every member will analyze their own algorithm performance and work out their own report. Zhao Yuan is responsible for the combination of the whole report and Yili Lai is responsible for expression optimization. Yuping Tian and Ning Tao are responsible for collecting visual results (figures and tables) and the website for presentation will be established by Yikai Wang and Hanyuxi Zhou.

**Non-Technical Roles:**

For Non-Technical Roles, different members will be responsible for different tasks during different phases, the table below shows the details.

| Roles | Member Name | Responsibility |
|---|---|---|

| | | |
|---|---|---|
| **Team Leader / Spokesman** | Zhao Yuan (Phase 1)<br>Yikai Wang (Phase 2)<br>Ning Tao (Phase 3)<br>Hanyuxi Zhou (Phase 4) | Organize and unite all team members |
| | | Contact with mentors and college |
| **Project Manager** | Yikai Wang (Phase 1)<br>Yili Lai (Phase 2)<br>Yuping Tian (Phase 3)<br>Ning Tao (Phase 4) | Control thinking and research direction |
| | | Control project schedule |
| | | Collect questions |
| **Meeting Recorder** | Hanyuxi Zhou (Phase 1)<br>Zhao Yuan (Phase 2)<br>Yuping Tian (Phase 3)<br>Yili Lai (Phase 4) | Record and summarize all points in meetings |
| | | All members may help with meeting records |
| **Document Manager** | Yili Lai (Phase 1)<br>Hanyuxi Zhou (Phase 2)<br>Zhao Yuan (Phase 3)<br>Yikai Wang (Phase 4) | Manage and modify formats and contents of documents |
| **Meeting Organizer** | Yuping Tian (Phase 1)<br>Ning Tao (Phase 2)<br>Yikai Wang (Phase 3)<br>Zhao Yuan (Phase 4) | Determine topics and times of group meetings and ensure every member takes part |
| **Github Repo Manager** | Ning Tao (Phase 1)<br>Yuping Tian (Phase 2)<br>Yili Lai (Phase 3)<br>Yikai Wang (Phase 4) | Organize all contents in Github repo and deal with conflicts |

*Table 3: Role and Responsibility*

## 4 Architecture

The project uses Microsoft Azure as the cloud computation environment. Every team member has a 100 USD student subscription plan (see figure 1 below), besides our mentor provided us with an additional subscription (see figure 2) with 1500 USD credit. Overall, the team has a 2100 USD Azure budget for the project.

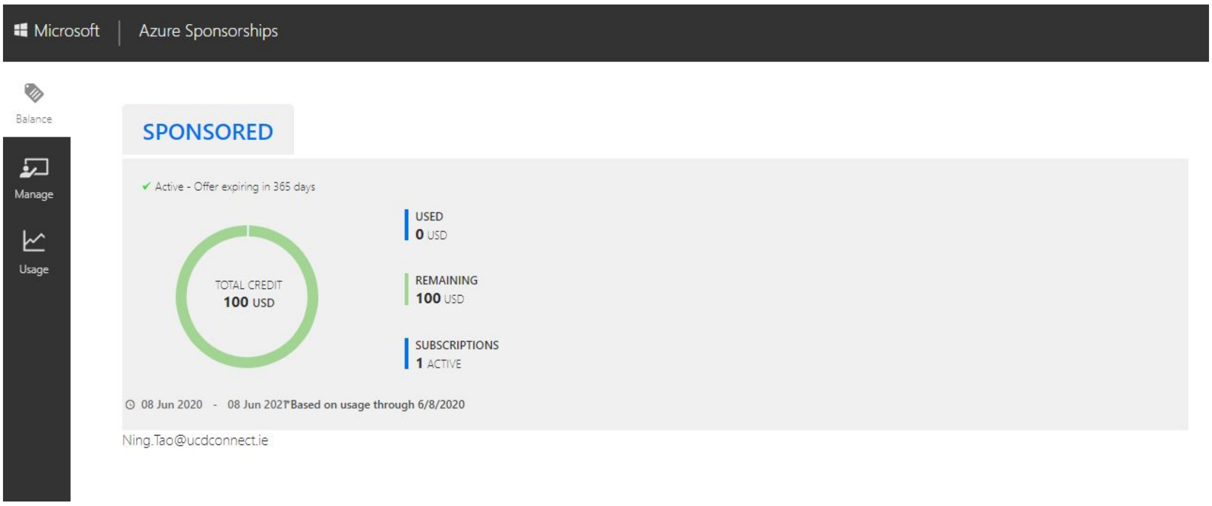


*Figure 1: Azure Student Subscription*


*Figure 2: Additional sponsorship*

## 5 Data Plan

Four popular datasets in the image retrieval area will be used to train and test our models, which are listed below. All images in these datasets are in JPG format and most dataset structures are stored as text files. And all evaluation results generated, such as loss, precision @1, MAP@R , R-precision and so on, will be stored in csv files and presented by graphs in our final website.

1. **Cars-196**

(From Stanford Artificial Intelligence Laboratory: [http://ai.stanford.edu/~jkrause/cars/car_dataset.html](http://ai.stanford.edu/~jkrause/cars/car_dataset.html))

Cars-196 dataset (1.8G) contains 16185 pictures of 196 classes of cars, in which 8,144 of them are training images and the other 8,041 are testing images. Images from each class are approximately equally divided into training and testing images.

Annotations of each picture including bounding box(measured in pixels), class and filename. Class names and annotations are stored in MAT files and a matlab script for evaluating training accuracy is provided.

2. **Caltech-UCSD Birds-200-2011**

(From Caltech and UCSD Vision: http://www.vision.caltech.edu/visipedia/CUB-200-2011.html)

CUB-200-2011 dataset (1.1G) is a new version of CUB-200 and is made up of 11,788 pictures of 200 bird species. Images in this dataset overlap with images in ImageNet. Training and testing split of this dataset is suggested in a txt file, dividing all pictures approximately equally into training and testing sets.

Each image is assigned with a class and a bounding box label measured in pixels. Each picture is also divided into 15 parts by pixel location of the center of the part and part locations also have another version perceived by multiple Mechanical Turk users. Moreover, this dataset contains 322 binary attribute labels (on a per-class level) of images obtained from MTurk workers.

3. **Stanford Online Products**

(From Computational Vision and Geometry Lab at Stanford University: https://cvgl.stanford.edu/projects/lifted_struct/)

Stanford Online Products dataset (2.9G) is made up of 120,053 images of 22,634 online products in 12 categories from eBay.com, with each product containing no more than 12 images. 59,551 of them are used for training and the other 60,502 images are for testing.

4. **In-shop Clothes Retrieval**

(From Multimedia Laboratory at The Chinese University of Hong Kong: http://mmlab.ie.cuhk.edu.hk/projects/DeepFashion/InShopRetrieval.html)

In-shop Clothes Retrieval dataset (10.2G) is a large subset of DeepFashion dataset, containing 52,712 in-shop clothes images of 7,982 clothing items (14 categories for women/9 categories for men) and ~200,000 cross-pose/scale pairs. Images are centered and resized to 256*256 and are annotated by bounding box, fashion landmark, clothing type, clothing segmentation mask and pose type.

The dataset is divided into a training set of 25,882 images from 3,997 items and a test set of 28,760 images from 3,985 items. But the test set is divided again into query set (14,218 images) and gallery set (12,612 images) so users use images in the query set to retrieve images in the gallery set. All dividing information is stored in json format.

## 6 Github

Up to now all members of the team have made regular non-arbitrary commits to the team Github repo, the following table shows what they have done:

| Member | Commits |
|---|---|

| | |
|---|---|
| Zhao Yuan | Established **Email Communication** file to record communications with college and Microsoft Mentors.<br>Update **Project Plan** file about Roles and Github section. |
| Hanyuxi Zhou | Established **Meeting Minutes** file to record all points and conclusions of every meeting.<br>Update **Project Plan** file about Data Plan section. |
| Ning Tao | Established **ReadMe** file which introduces the whole group.<br>Established **Project Plan** file to introduce the plan for the whole project.<br>Update **Project Plan** file about Architecture section. |
| Yuping Tian | Update **Project Plan** file about Team Management section. |
| Yikai Wang | Update **Project Plan** file about Project Objectives section. |
| Yili Lai | Update **Project Plan** file about Project Plan section. |

*Table 4: Github Contributions*

The following figure also justified that all six members are familiar with Github.

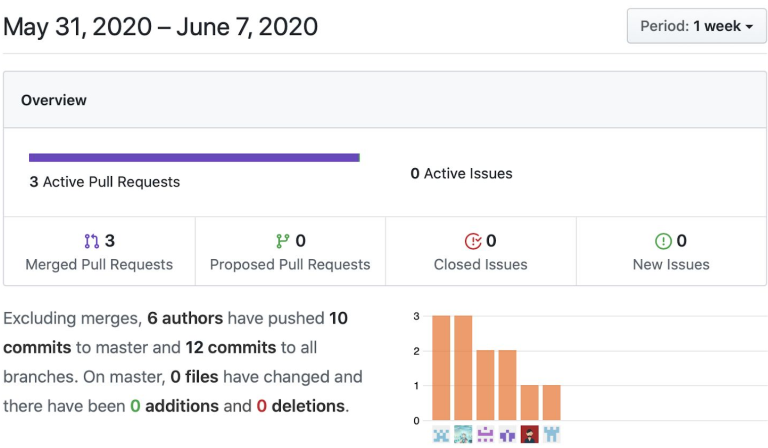


*Figure 3: Github Contribution Visualization*

# 7 Team Management

## Meetings

Due to the Covid-19, meetings are all taking place online, the Microsoft Teams is used as the meeting platform. There will be an one hour meeting and a 30 minutes office hour each week with mentors from Microsoft. Besides, there are daily meetings within the group members. The minutes of meetings will be uploaded to github after each meeting.

## Demonstration

Task assignments and progress tracking are done on Microsoft Teams. In addition to chatting and meeting, it can also provide a function called planner, which allows users to set targets (Figure 4) and date of completion (Figure 5).

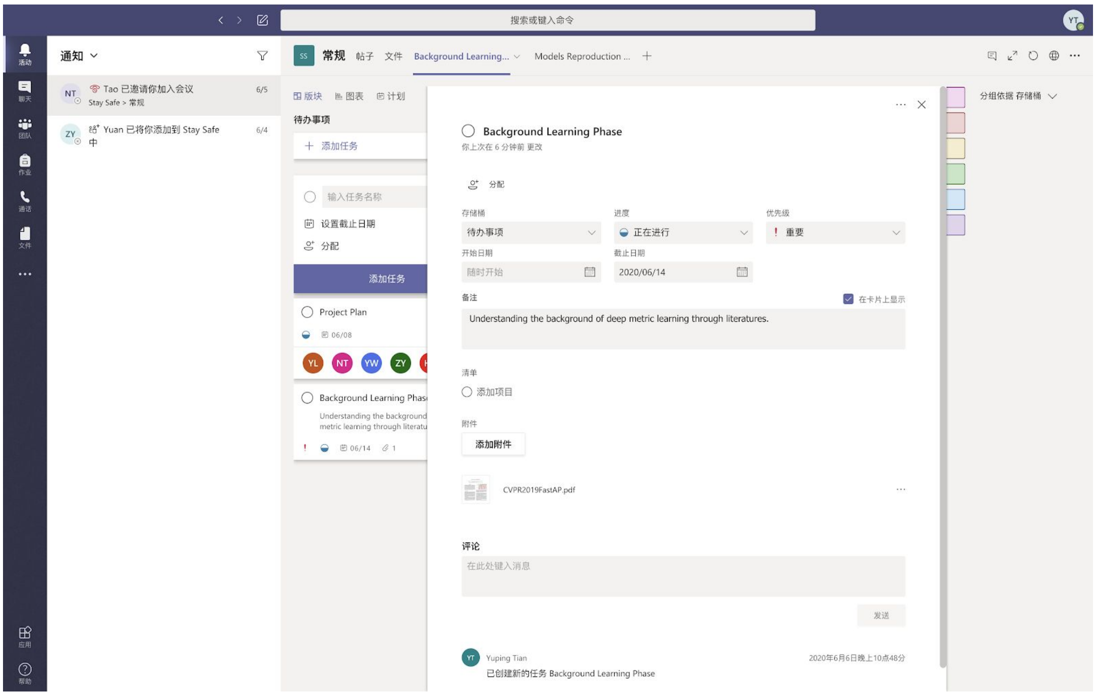


*Figure 4:Task Information*

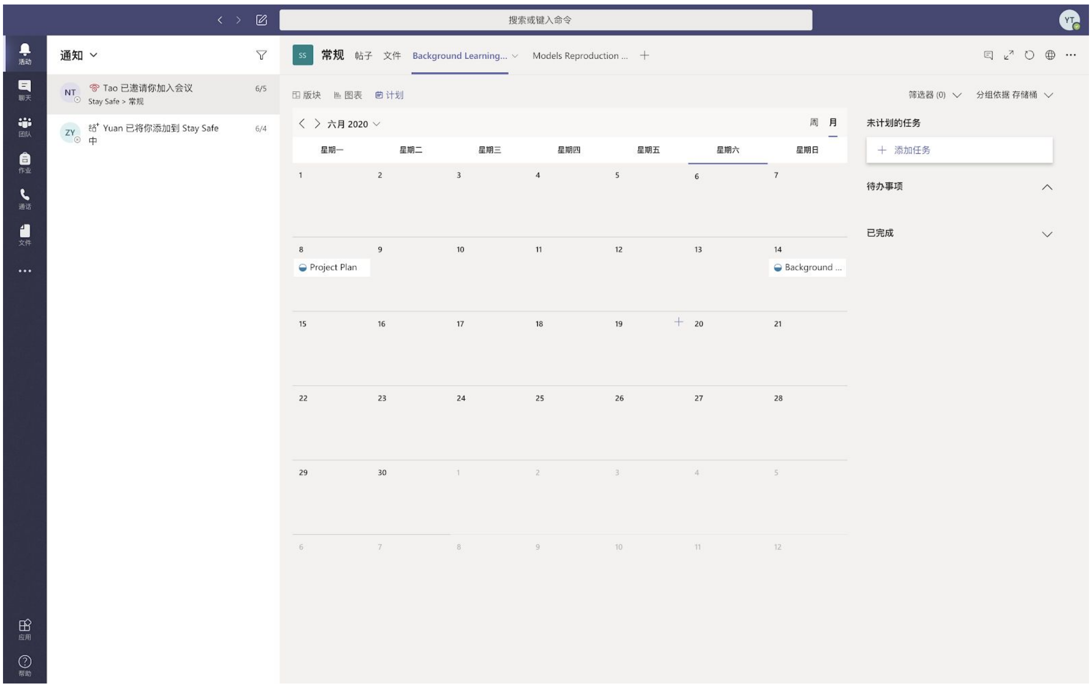


*Figure 5: Timetable*