

# Team name: Joker

# Team members: Jiacheng Zhou, Ruitao Shen, Shihang Liu, Siwei Guo, Yuanchi Guo

# Date: 03/08/2024

This journal has to be uploaded to 1) Canvas (will create assignments). **AND** 2) posted to your project website every two weeks by 11:59 pm on the second Friday (except the spring break). When uploading to Canvas, change the file name to "Team\_name\_MM-DD-2023" where MM = month and DD= day of upload).

---

## Team roles for this report (write down name):

Facilitator(s): Siwei Guo

Recorder(s): Ruitao Shen, Shihang Liu

Deliverer(s): Yuanchi Guo

Planner(s): Jiacheng Zhou

See last page for description of roles. Obviously one person can take more than one role or there can be more than one person per role or make your own roles!

- 0. Describe briefly what the main goal of your team is (so the peer reviewer has some context). E.g. we are working on image classification for blah de blah. Our goal is blah de blah etc. In the initial part of the semester before your proposal it is ok to put down "we are still coming up with ideas on team project".**

Our project centers around image classification. Following a comparison of animal picture resources and fruit picture resources, we opted to pursue fruit and vegetable image classification due to the extensive variance within animal species, which could complicate the identification process. Our specific objective is to utilize a supervised method to train a model capable of identifying similar fruit categories, such as apple pink lady, apple golden, and apple crimson, based on subtle physical feature differences like colors and

shapes. Similar fruit classification always confuses costumers, sellers and market managers. It's worth mentioning that while fruit is our current focus, this may shift as we progress through the analysis process. Our current group plan involves preparing the images for future analysis, including performing tSNE and PCA for dimension reduction. Additionally, we are conducting a literature review to gather valuable suggestions on machine learning methods and models for our future analysis.

- I. What was done during the report period regarding the project:** If you want to include code include this in the Appendix. Describe what the group did (including contributions of individual team members) with regards to the group project during this report period. Give enough details so I understand what you folks have been doing over the week. Include dates of your meeting(s) and who met on these days.

Our group consistently meets every Sunday morning from 10 am to 11 am. In the meeting on 2/25, we deliberated on the quality of the current image dataset obtained from Kaggle, assessing its research potential and image qualities. Jiacheng conducted a literature review on general image classification, compiling notes with normal objects, impressive figures, and typical processes that inspired our research direction. Yuanchi delved into the implications of image classification in real-world scenarios from economic perspective. Shihang researched various models and general research methodologies relevant to image identification projects. Ruitao explored the differences between using Python and R for image processing in identification tasks. Siwei investigated PCA and tSNE methods for dimensionality reduction and their application in our project. After deliberation, we decided to prioritize image preparation tasks such as resizing and dimension reduction, followed by selecting the best model based on the results obtained.

In the meeting on 3/3, we discussed reduced-dimension image techniques and collaborated on the project proposal. Each group member was responsible for writing one part of the proposal. Jiacheng drafted a section on the general image classification process, drawing insights from established research and projects as our research guide. Yuanchi outlined the research implications from various perspectives, including costumers, sellers, and market managers. Siwei contributed code for PCA and tSNE to apply to our images and also wrote the image preparation section, which included figures and code snippets. Shihang and Ruitao identified the final research models we will use together with provided some code as examples, including CNNs, providing a specific direction for the model anlysis.

- II. What were obstacles faced if any in working on the project?** This could be technical (like not being able to implement or understand particular techniques) or time issues (midterms for other courses etc).

Currently we had a technical obstacle. The PCA works on a small set of test images and successfully provides us with reconstructed pictures. However, when applying the PCA code to the entire image folder, RStudio runs but never finishes. One possibility is that we need to wait longer until it finishes, but we don't know how long it will take or whether it will successfully apply to the entire folder. Moreover, our analysis

is more targeted and seems does not apply to all images. We need to find a way to add a more representative code.

**III. What is the plan for the next reporting period including what each team member is planning to work on. Describe goals and potential timelines (“ I plan to finish understanding x to see if it can be implemented for our project by Wednesday etc”. )**

Before the next reporting period, we will finalize the preparation of our image dataset and begin applying potential models to attempt identification. Siwei will endeavor to extend PCA and t-SNE to apply to all images rather than just a small subset. Yuanchi and Jiacheng plan to process image standardization to ensure all images undergo the necessary adjustments to conform to established standards we set. Shihang and Ruitao plan to further refine the final project model by using CNN to add more possibilities for images with different backgrounds.

While in the biweekly document above you will describe what your team did with regards to the team project (with proper attributions of who did what in the week) there are 4 pre-defined roles. I urge you to have different people do these jobs every week so that you gain experience in each of the jobs. There can also be more than one person per job for example 2 people recording the weekly journal.

**Facilitator:** Manages the group for this week including setting up times for group members to meet, making sure everyone has a say in the meetings etc.

**Recorder:** Person in charge of recording the meetings as well as the happenings of the past two weeks and describing what was accomplished in the meeting and writing up this report.

**Deliverer:** Person in charge of checking the entire report and uploading the file to dropbox folder and project website, as well as the representative of the group getting in touch with the instructor.

**Planner:** Person in charge of what will be happening next two weeks as well as thinking about longer term goals (what more needs to be done for the project).

**Team contact:** Person I can email if I see any issues in the biweekly report instead of mass spamming everyone in the team.