III. Model documentation and write-up

Information included in this section may be shared publicly with challenge results. You can respond to these questions in an e-mail or as an attached file. Please number your responses.

1. Who are you (mini-bio) and what do you do professionally? If you are on a team, please complete this block for each member of the team.

    I am a bioinformatic engineer with a master's degree of Biochemistry. I also interested in take part in various data science competition. Have a competition grandmaster tie at Kaggle and also have some wins at Chinese domestic competitions.

2. What motivated you to compete in this challenge?

    I miss the competition hold in Kaggle last few months. This one is similar, thus I decided to join in.

3. High level summary of your approach: what did you do and why?

    Metric learning, for it is the most popular and maybe the most useful one for such problem.

    This is a typical database retrieval problem. Since the dataset is small and very unbalanced, I focus on improving the representation of the whale. The major parts of the network includes CNN and transformer backbone, a Arcface head and a focal loss. Swin transformer performs better on local validation and is similar to Efficientnet on the leaderboard. Arcface head is commonly used on ReID and also performs well at this dataset.

    I also tried some state of the art method like DOLG and adaptive Arcface, but the gap between the test set and the train set makes these methods merely the overfitting of cross validation. I also tried tricks like database augmentation and query extension, due to the size of the train set, these methods help merely a little.

4. Do you have any useful charts, graphs, or visualizations from the process?

    I focus the mAP metric, no other graphic.

5. Copy and paste the 3 most impactful parts of your code and explain what each does and how it helped your model.

    1. the focal loss
    2. the model structure
    3. the knn

6. Please provide the machine specs and time you used to run your model.

- CPU (model): 3960x
- GPU (model or N/A): 2x3090
- Memory (GB): 128GB
- OS: Ubuntu server
- Train duration: 24h
- Inference duration: ~3h

7. Anything we should watch out for or be aware of in using your model (e.g. code quirks, memory requirements, numerical stability issues, etc.)?

   a. Dataset is small, and result is different if seed changed, thus even I fixed all possible seed on my local machine, retrain may leads to very significant differences.
   b. I train this model on dual RTX3090 24G, any dual card V100 32G, A100 RTX A6000 and A5000 should be okay.

8. Did you use any tools for data preparation or exploratory data analysis that aren't listed in your code submission?

   No.

9. How did you evaluate performance of the model other than the provided metric, if at all?

   No.

10. What are some other things you tried that didn't necessarily make it into the final workflow (quick overview)?

    I have many other runs with adaptive margin, DOLG and transformer backbone with better CV (I have a CV setup similar to the test set, top vs top and left vs top). Sadly these never give me better LB, thus the submission workflow is the best LB.

11. If you were to continue working on this problem for the next year, what methods or techniques might you try in order to build on your work so far? Are there other fields or features you felt would have been very helpful to have?

    I want to try some reconstruction method similar to those used in the face recognition, generate side view based on given front view.