

## **A1. Background on you/your team**

- Competition Name: Google Universal Image Embedding
- Team Name: IRonCLIP
- Private Leaderboard Score: 0.685
- Private Leaderboard Place: 6

[Member 1]

- Name: Socratis Gkelios
- Location: Thessaloniki, Greece
- Email: socratis.gkelios@ymail.com

[Member 2]

- Name: Anestis Kastellos
- Location: Thessaloniki, Greece
- Email: kastellosa@nup.ac.cy

[Member 3]

- Name: Savvas A. Chatzichristofis
- Location: Pafos, Cyprus
- Email: s.chatzichristofis@nup.ac.cy

## A2. Background on you/your team

If part of a team, please answer these questions for each team member. For larger teams (3+), please give shorter responses.

- **What is your academic/professional background?**

[Member 1]

Socratis Gkelios received his diploma degree from the Department of Electrical and Computer Engineering, Democritus University of Thrace, in 2019. He is currently a PhD candidate in Computer Science at the same department and a Research Assistant at the Centre for Research and Technology Hellas (CE.R.T.H.), Information Technologies Institute (I.T.I.). His research is focused on computer vision, image retrieval, deep learning, and autonomous driving.

[Member 2]

Anestis Kastellos received his BSc degree from the Department of Mathematics, University of Ioannina, in 2021. He is currently a MSc student in Artificial Intelligence at the European University of Cyprus and a Research Associate at the Centre for Research and Technology Hellas (CE.R.T.H.), Information Technologies Institute (I.T.I.) at Visual Computing Lab (V.C.L.). His research is focused on computer vision, robotic learning, and deep learning.

[Member 3]

Savvas A Chatzichristofis is the Vice-Rector for Research & Innovation and a Professor of AI at the Department of Computer Science of the Neapolis University Pafos (Cyprus). He also serves as director of the Intelligent Systems Laboratory.

His research focuses on the intersection of Artificial Intelligence, Computer Vision, and Robotics. He has been involved in numerous R&D projects funded by European or National funds. His contribution received several distinctions, grants, scholarships, and awards.

- **Did you have any prior experience that helped you succeed in this competition?**

[Member 1]

Experience with Kaggle competitions, deep learning, and image retrieval algorithms.

[Member 2]

Experience with computer vision and deep learning

[Member 3]

Long experience on image retrieval and computer vision tasks

- **What made you decide to enter this competition?**

[Member 1]

Universal image embedding is a topic closely related to my PhD.

[Member 2]

For educational purposes

[Member 3]

Universal image embedding is a topic closely related to my research interests

- **How much time did you spend on the competition?**

[Member 1]

Approximately 4 hours per day.

[Member 2]

Approximately 2 hours per day

[Member 3]

Approximately 1 hour per day.

- **If part of a team, how did you decide to team up?**

[Member 1]

Savvas Chatzichristofis is my supervisor on my PhD and he was interested in exploring together the competition. Anestis Kastellos is a colleague that was briefly in touch with the competition and wanted to help the team.

- **If you competed as part of a team, who did what?**

Socratis and Anestis were responsible for the code implementation and researching relevant past competitions and papers. Savvas was in charge of directing and developing our submission approach.

### A3. Summary

4-6 sentences summarizing the most important aspects of your model and analysis, such as:

- **The training method(s) you used (Convolutional Neural Network, XGBoost)**

*Please refer to the submitted paper*

- **The most important features**

*N/A*

- **The tool(s) you used**

Pytorch, OpenClip, Sklearn, and many publicly available datasets.

- **How long it takes to train your model**

On our hardware, it takes approximately 13 hours to generate the final solution.

### A4. Features Selection / Engineering

- **Did you use external data? (if permitted)**

*Please refer to the submitted paper*

### A5. Training Method(s)

- **What training methods did you use?**

*Please refer to the submitted paper*

- **Did you ensemble the models?**

Unfortunately, we did not have time to explore further model ensembles.

- **If you did ensemble, how did you weight the different models?**

-

### A6. Interesting findings

- **What was the most important trick you used?**

The pretrained CLIP model on LAION-2B dataset and constructing a big relevant training dataset.

- **What do you think set you apart from others in the competition?**

The main difference with many existing solutions was the fact that we achieved to finetune successfully the model's backbone without a massive distribution shift.

- **Did you find any interesting relationships in the data that don't fit in the sections above?**

By increasing the dataset size, the finetuning of the backbone was easier.

## **A8. Model Execution Time**

Many customers care about how long the winning models take to train and generate predictions:

- **How long does it take to train your model?**

~13 hours

- **How long does it take to generate predictions using your model?**

Assuming a batch size of 32, an RTX 3090 GPU and a Ryzen 5900X CPU (num\_workers = 4), we could process 120 images/second with FP16.

## **A9. References**

*Please refer to the submitted paper*