

AnimalCLEF25

ANIMAL IDENTIFICATION
DEEP NEURAL NETWORK

@CVPR-FGVC & LifeCLEF

BAD GENIUS



Team members

- 1. MUHAMMAD MUHAIRIS BIN AZMAN (2114599)**
- 2. MUHAMMAD HAFIDZUDDIN HANIF DANIAL BIN NORIZAL (2123651)**
- 3. SHAREEN ARAWIE BIN HISHAM (2116943)**
- 4. IBRAHIM BIN NASRUM (2116467)**

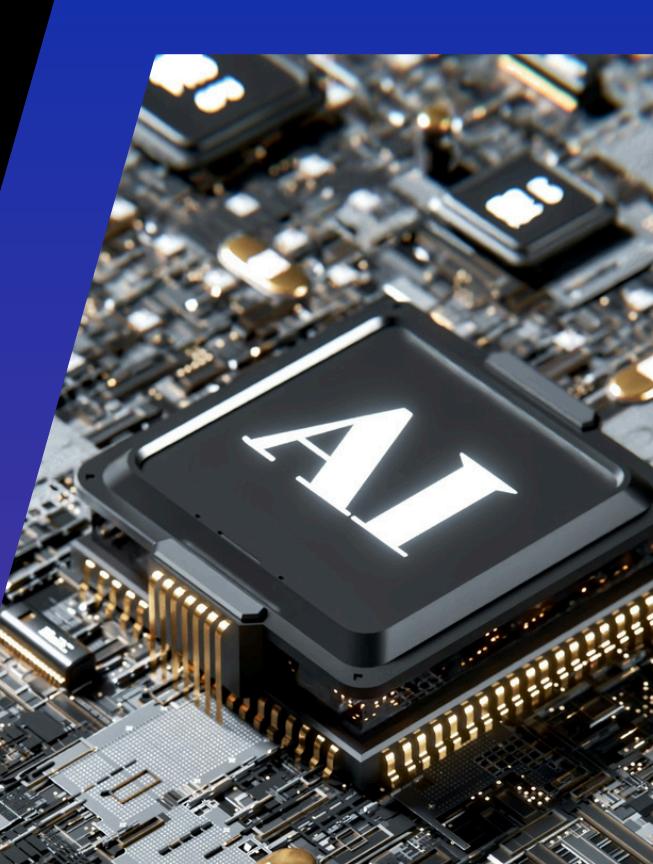


TABLE OF CONTENTS

1. COMPETITION OVERVIEW

2. PROBLEM STATEMENT

3. MODEL DESIGN

4. TRAINING & EVALUATION

5. RESULT & DISCUSSION

6. CONCLUSION



Competition Background

Objective

Build a model to **identify and classify** individual animals from image

Species

Comprises of **Loggerhead Sea Turtles** ,
Salamanders and **Eurasian Lynx**

Datasets

Provided pretraining set (**AnimalCLEF2015**) and a test set split into database (labeled) and query (to identify). Use only provided datasets external data requires approval.

Evaluation

Based on **classification performance** under real-world challenging conditions.



COMPETITION OVERVIEW

Available Tools

WildlifeReID-10k Dataset Paper

1

This dataset helps improve AI models for identifying animals across different species, ensuring fair and reliable comparisons.

2

WildlifeDatasets ✓

A Jupyter notebook tutorial is included to help users get started. The package works well with Wildlife tools, which has a pre-trained AI model (MegaDescriptor) and training tools for better wildlife recognition.

3

WildlifeTools ✓

The wildlife-tools library offers a simple interface for various tasks in the Wildlife Re-Identification domain. It covers use cases such as training, feature extraction, similarity calculation, image retrieval, and classification. It complements the wildlife-datasets library, which acts as dataset repository.

4

MegaDescriptor ✓

The first foundation model for recognizing individual animals across many species. It beats other popular models (like CLIP and DINOv2) in accuracy and comes in three sizes (Small, Medium, Large) for different needs.

5

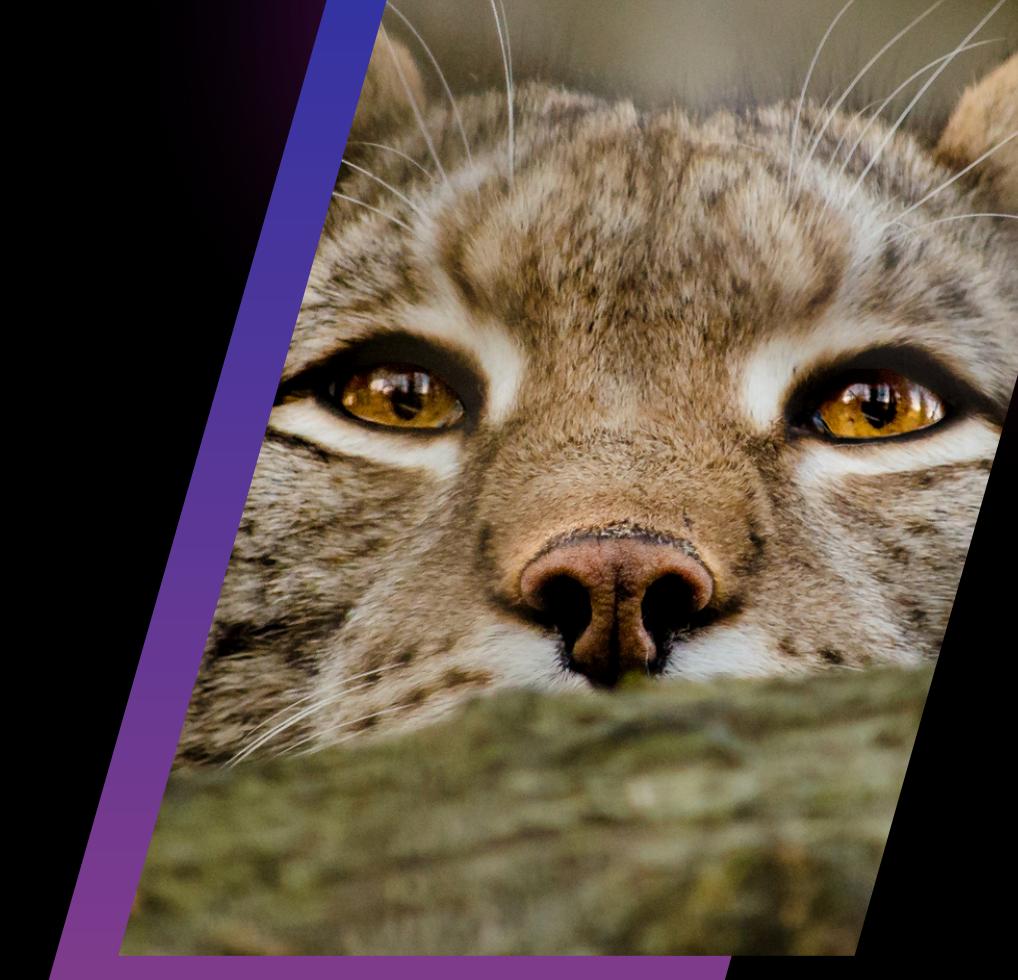
WildFusion

WildFusion is a new AI method that enhances animal identification by combining global features (which recognize an animal's overall appearance using models like MegaDescriptor or DINOv2) with local features (which match unique details like spots or stripes using tools like LoFTR and LightGlue), resulting in more accurate and reliable wildlife recognition.

Note:



Tools and libraries used in this project



PROBLEM STATEMENT

Problem statement: Identifying individual animals is important for tracking and protecting wildlife. Traditional ways like tagging or using GPS collars can harm the animals and cost a lot, especially in remote places. Camera traps are a better, non-invasive option, but they take thousands of pictures, which are hard and slow to check by hand. It's even harder for animals that look very similar, like tigers or leopards, where small details like fur patterns or scars are the only way to tell them apart.

MODEL DESIGN

Model specification

Model name: MegaDescriptor-L-384

A **Swin-L** image feature model

It is a pre-trained **transformer-based deep neural network** model

In this project, we will do **self-supervised feature extractor**

Model Type: Animal re-identification / feature backbone

Params (M): 228.8 million parameters

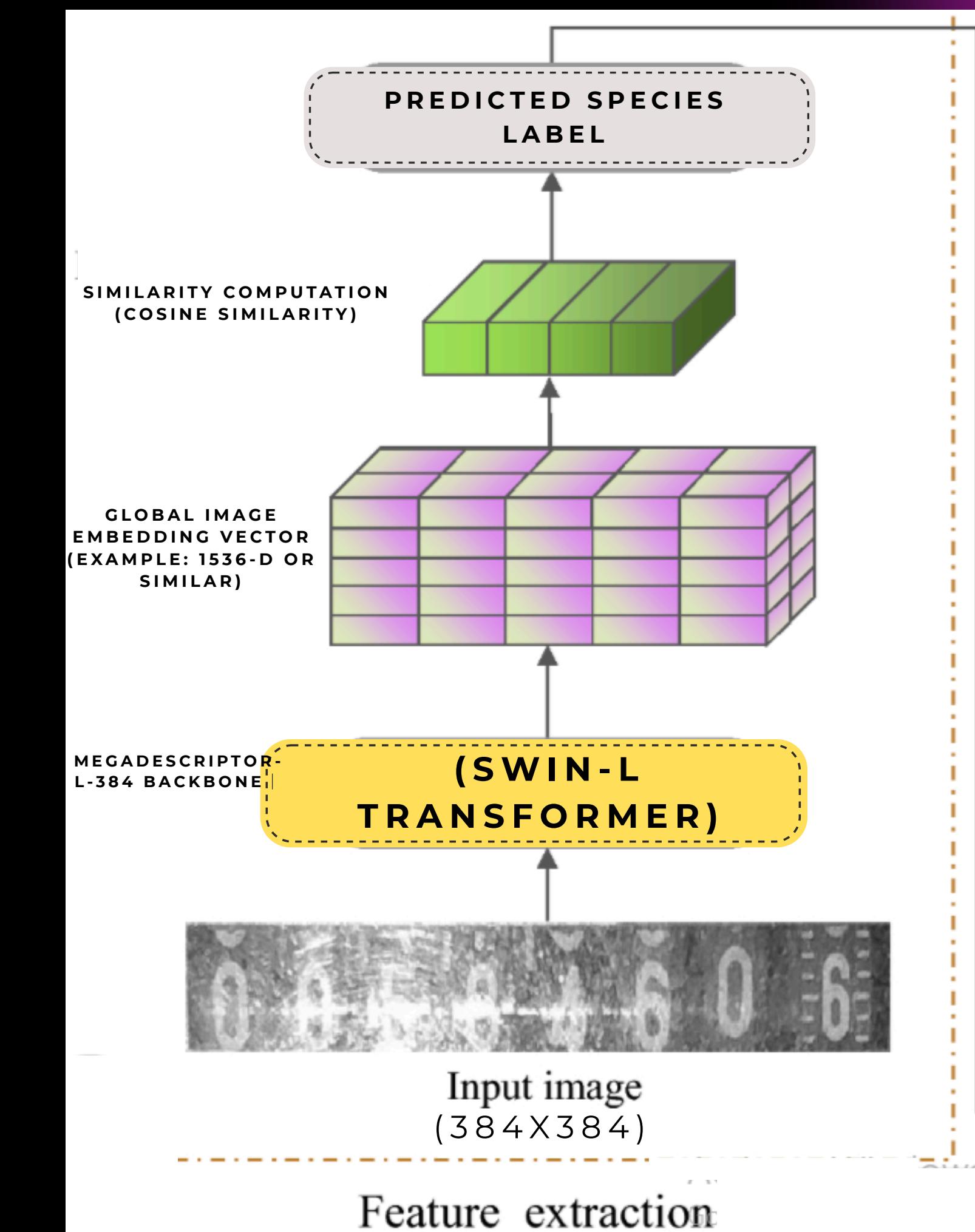
Image size: 384 x 384 pixels

Architecture: swin_large_patch4_window12_384



MODEL DESIGN

Architecture Diagram



MODEL DESIGN

Dataset Overview

Dataset specification:

Total images = 15,209

Unlabeled images = 2,135

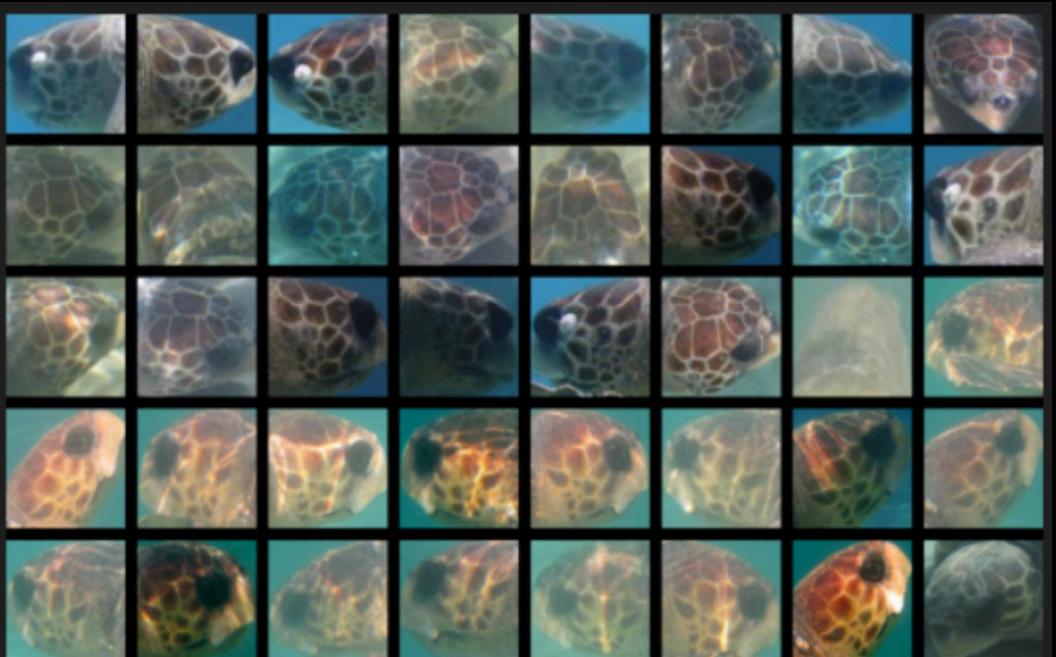
Number of classes = 3

Image format = jpg

Dataset classes:

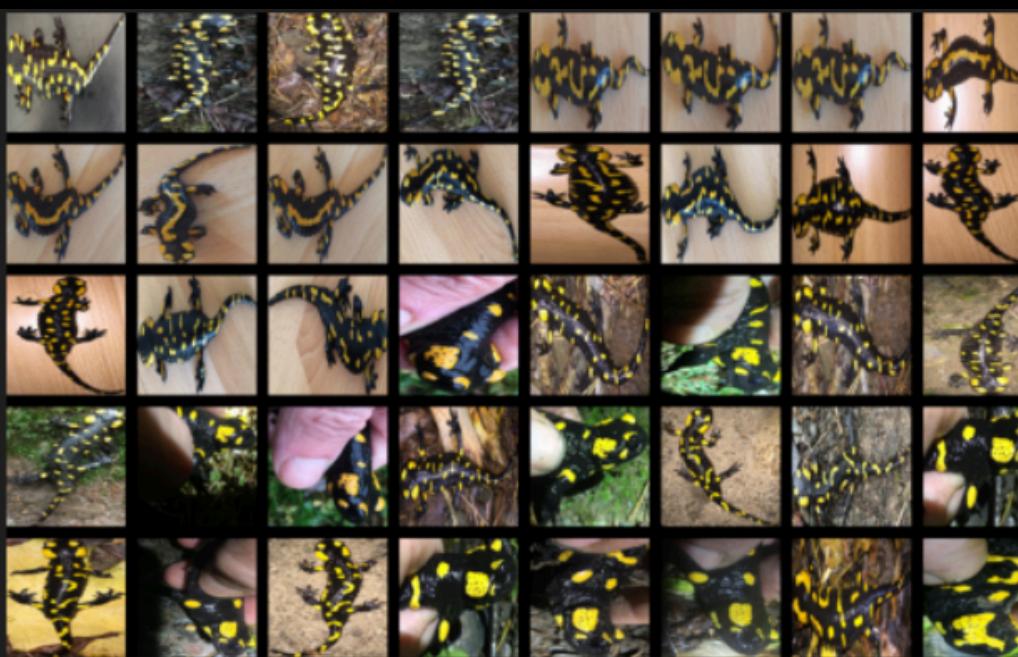
1 LOGGERHEAD SEA TURTLE

9,230 IMAGES



2 SALAMENDERS

2,077 IMAGES



3 EURASIAN LYNXES

3,902 IMAGES



Justification

- **Why use Transformer-based architecture?**

- i) Global context modeling (robust to pose/background variation)
- ii) Fine-grained local details (essential for distinguishing visually similar species)

- **Why pretrain on animal re-identification datasets?**

- i) Features highly relevant for recognizing individual animals and fine-grained species differences
- ii) Strong inductive bias tailored for this competition

Data Preprocessing

1. Proper folder structure

**2. Transform the input
image into 384x384 pixels**

3. Convert image to tensor

**4. Normalize the image
matching the model
requirement**

RESULT & DISCUSSION

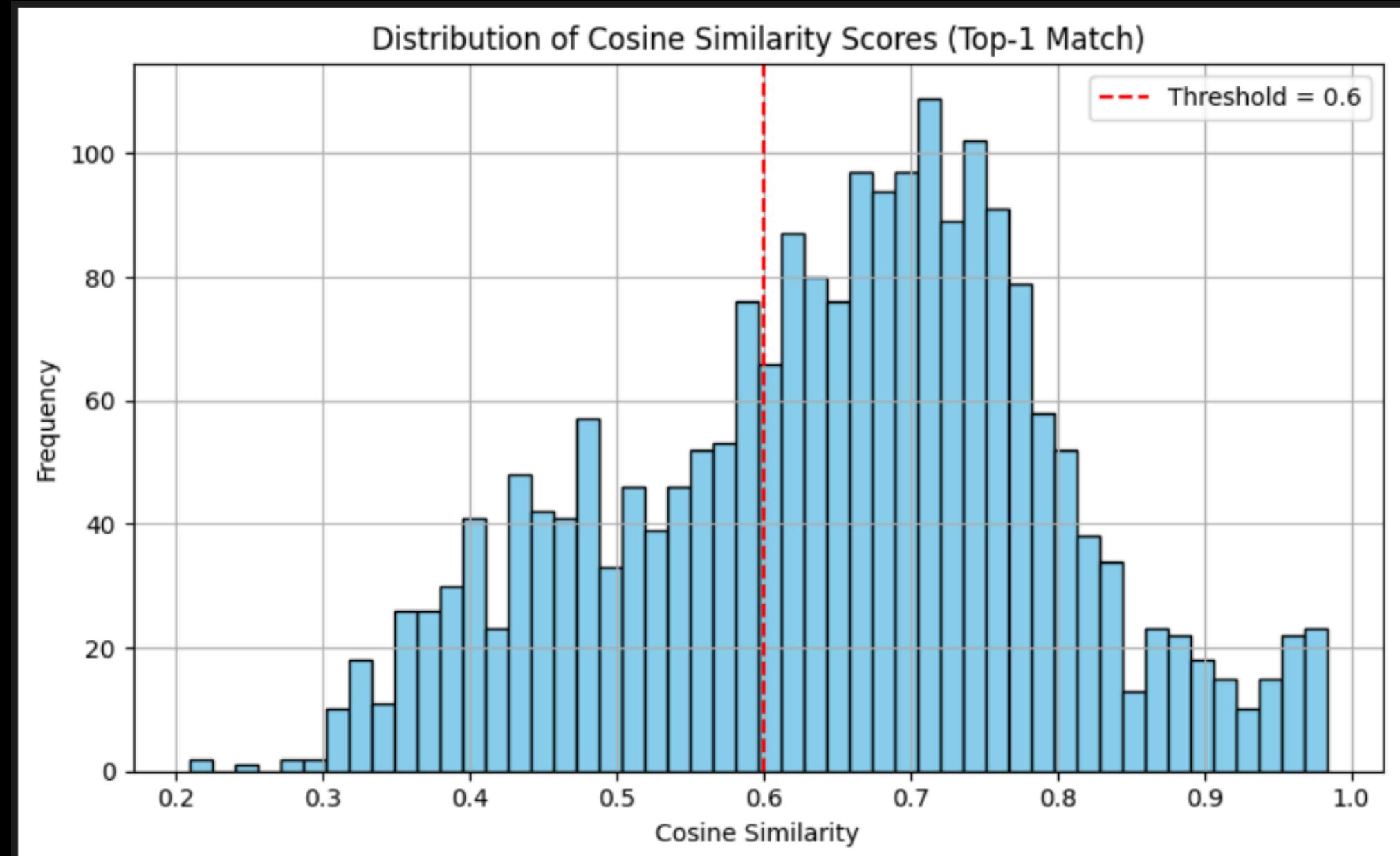
- **Cosine similarity** is used to compare feature vectors extracted from images to determine the similarity between images.
- **Matches/ Threshold:**
 - $> 0.6 \rightarrow$ good result
 - $< 0.6 \rightarrow$ too ambiguous

Inference time : 47 minutes

Batch amount : 476 batches

Running on VS Code with GPU

Result



RESULT & DISCUSSION

Result

Demonstration of image classification
by **MegaDescriptor-L-384**

image_id	identity set by MegaDescriptor	original identity
3	LynxID2025_lynx_75	LynxID2025
5	new_individual	LynxID2025
12	LynxID2025_lynx_95	LynxID2025
13	LynxID2025_lynx_28	LynxID2025
18	LynxID2025_lynx_05	LynxID2025
19	new_individual	LynxID2025
27	new_individual	LynxID2025
33	LynxID2025_lynx_62	LynxID2025
36	LynxID2025_lynx_65	LynxID2025
45	new_individual	LynxID2025
46	LynxID2025_lynx_28	LynxID2025
47	new_individual	LynxID2025
48	LynxID2025_lynx_31	LynxID2025
49	LynxID2025_lynx_32	LynxID2025
55	LynxID2025_lynx_32	LynxID2025
58	LynxID2025_lynx_40	LynxID2025
61	LynxID2025_lynx_60	LynxID2025
62	new_individual	LynxID2025
63	new_individual	LynxID2025
64	new_individual	LynxID2025
65	new_individual	LynxID2025
67	new_individual	LynxID2025
68	LynxID2025_lynx_62	LynxID2025
74	LynxID2025_lynx_93	LynxID2025

CONCLUSION

LIMITATION

- **Large Dataset Size**, means high memory and storage demands making training and inference slower.
- **Big Model Architecture**, The model has a large number of parameters which increases GPU/CPU usage
- **High Computational Time**
Training, validation and even inference became time-consuming due to the dataset and model size.



CONCLUSION

ADVANTAGE

- **High Descriptive Power**

Though it is a big model it can captures details in images enabling distinction between very similar inputs.

- Can perform well with various datasets due to its **robust and large feature space**.

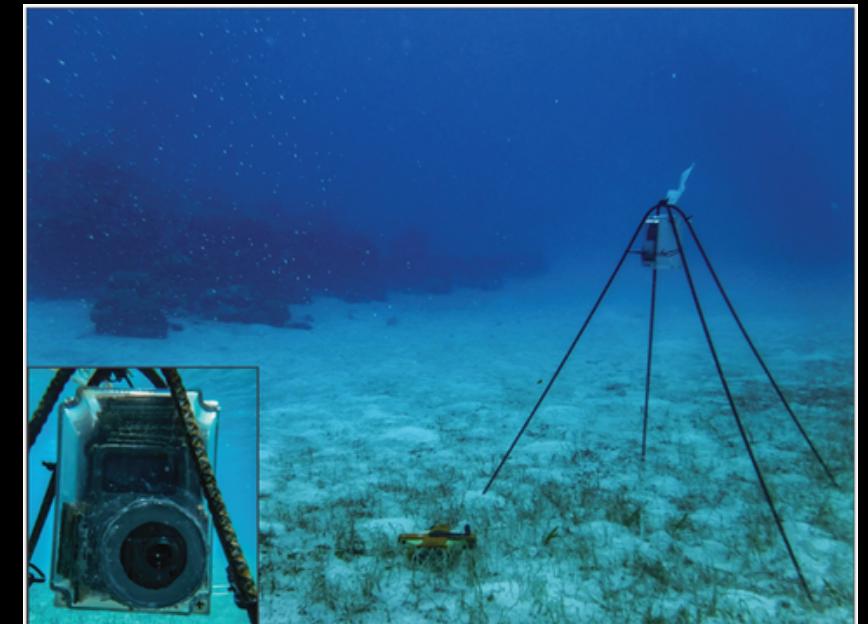
- The model is **versatile** it's suitable for face recognition, surveillance, visual search and object verification



CONCLUSION

APPLICATION

- **Wildlife monitoring apps** for conservationists and citizen scientist , Explore self-supervised pretraining or domain-specific pretraining
- **Automated species identification tools** for camera trap data
- Real-time biodiversity assessment on drones or **field cameras**



CONCLUSION

Future Planning

- **Improving the model**
 - Better fine tuning
 - More data augmentation
 - More variety of animals/ rare species
- **Apply the MegaDescriptor-L-384 into projects**
 - Compact animal identification device

Summary

This project successfully demonstrated the capability of MegaDescriptor-L-384 together with its advantages, limitations and applications in the industry especially in **Marine and Wildlife Ecology**.



**Thank You
For Your Attention !**



References

1. AnimalCLEF25 @ CVPR-FGVC & LifeCLEF. (2024).
Kaggle; Fine-Grained Visual Categorization
Community Prediction Competition.
<https://www.kaggle.com/competitions/animal-clef-2025/overview>
2. Group Github Link: https://github.com/HN-admin/BadGeniusAnimal_Identification

The user of this dataset must follow the provided license of SeaTurtleID2022 and acknowledge the accompanying paper. In particular, it prohibits commercial applications and being re-uploaded. If WildlifeReID-10k was used as a training database, its license file must be followed.

