

AnimalCLEF25 @ CVPR-FGVC

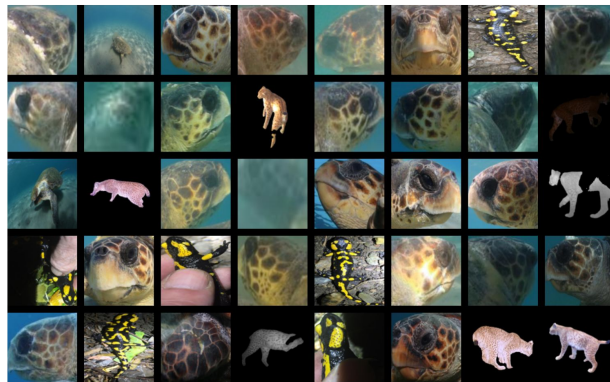
Animal Identification

Team number: 28

Team members: Hanxiao Wang, Yi Yu, Junyao Liu

Introduction

- Open-set individual animal identification task proposed in the CVPR 2025 Animal-CLEF challenge.

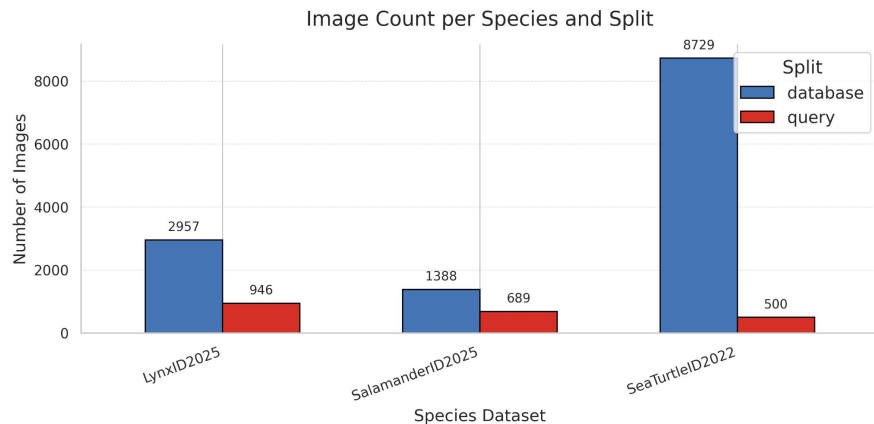


Motivation

- Track animal movement and population dynamics
- Support ecological monitoring in real-world environments
- Benchmark open-set recognition and fine-grained classification methods

Data

Primary Dataset: Animal CLEF 25



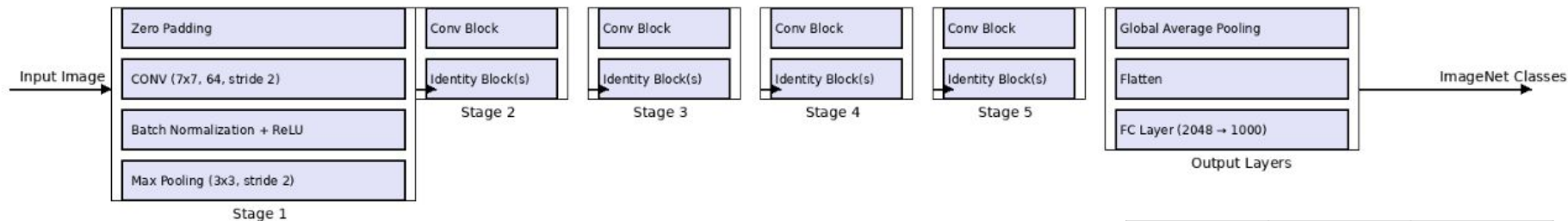
Known vs Unknown



Evaluation method

- Overall performance of the classification model:
 - Accuracy
 - Precision
 - Recall
 - F-1
- Specific to the open-set scenario:
 - BAKS: Balanced accuracy on known individuals
 - BAUS: Balanced accuracy on unknown individuals
 - Final Accuracy (geometric mean): combined with two indicators

Baseline approach — ResNet-50 (Pre-trained Only)



Setup

- **ResNet-50:** 50-layer CNN pretrained on **ImageNet-1k**

Deficiency

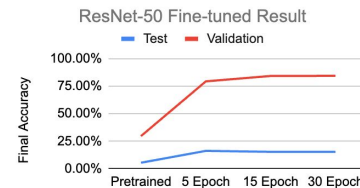
- **Domain mismatch:** Pretrained on ImageNet, not tailored to our dataset
- **General features** Fail to capture **fine-grained animal differences**
- **Resolution constraints:** Optimized for 224×224 images

		Pretrained
Validation	Accuracy	81.74%
	Precision	87.79%
	Recall	81.74%
	F1	76.00%
	BAKS	100.00%
	BAUS	8.81%
	Final Accuracy	29.68%
Test	Final Accuracy	5.50%

Advanced approach — ResNet-50 (Fine-tuned)

Setup

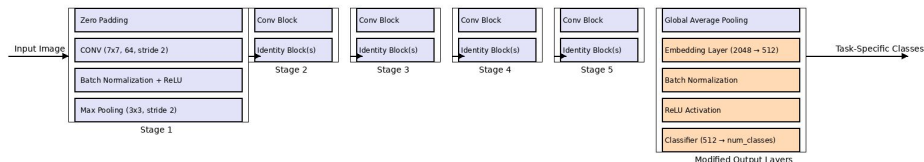
- **Frozen ResNet-50 backbone**
- **Custom classification head**
 - 2048 → 512 embedding layer
 - BatchNorm + ReLU
 - Final classifier tailored to our dataset



Training Strategy

- **Data Augmentation**
 - Resize (256×256) → Random Crop (224×224)
 - Horizontal flips, rotation, color jitter
- **Loss & Optimization:**
 - Cross Entropy Loss
 - Adam optimizer
 - StepLR scheduler

		Pretrained	5 Epoch	15 Epoch	30 Epoch
Validation	Accuracy	81.74%	92.59%	94.20%	94.24%
	Precision	87.79%	94.67%	95.81%	95.87%
	Recall	81.74%	92.59%	94.20%	94.24%
	F1	76.00%	92.41%	94.22%	94.28%
	BAKS	100.00%	100.00%	100.00%	100.00%
	BAUS	8.81%	63.01%	71.04%	71.23%
	Final Accuracy	29.68%	79.38%	84.29%	84.40%
Test	Final Accuracy	5.50%	16.28%	15.41%	15.41%



Advanced approach — Ensemble (ResNet & Pretrained Dino-V2)

Setup

- **Combine embeddings** from fine-tuned ResNet-50 and pretrained DINOv2
- **Normalize** each embedding to unit length
- **Concatenate** normalized vectors for final representation
- **Prediction** via nearest neighbor search using **cosine similarity**

Future Work

- **Alternative fine-tuning**
- **Threshold Optimization**
- **Advanced Ensembling**
- **Generalization**

		Pretrained	5 Epoch	15 Epoch	30 Epoch
Validation	Accuracy	99.53%	98.00%	99.61%	99.57%
	Precision	99.61%	98.62%	99.70%	99.66%
	Recall	99.53%	98.00%	99.61%	99.57%
	F1	99.55%	98.15%	99.64%	99.60%
	BAKS	100.00%	100.00%	100.00%	100.00%
	BAUS	97.65%	90.02%	98.04%	97.85%
	Final Accuracy	98.82%	94.88%	99.02%	98.92%
Test	Final Accuracy	12.68%	9.00%	10.83%	10.35%

Advanced approach — MegaDescriptor (Pretrained)

Global Embeddings (High-quality cosine similarity-ready embeddings)

(1)Image Retrieval (KNN)

(2)ViT-based, for Retrieval and Re-identification (ReID)

Future Work

- Continuously finetune
- Automatic threshold

	Baseline	Resnet-50	MegaDescriptor
accuracy	81.74%	94.23%	92.43%
precision	87.79%	95.87%	100%
recall	81.74%	94.24%	92.43%
F-1	76%	94.28%	96.07%
BAKS	100%	100%	58.25%
BAUS	8.81%	71.23%	0%
Final Accuracy	29.68%	84.40%	71.10%
Kaggle Score	5.50%	15.41%	23.05%

Workload

- Literature review (ResNet-50, DINO-V2, Mega Descriptor) (15%)
- Data preprocessing (20%)
- Baseline + advanced model implementation (35%)
- Performance evaluation (15%)
- Result analysis (15%)

Challenges

- Class imbalance, some identities have very few sample
- Computation of GPU limitation

To Do

- WildlifeReID-10k integration partially done (only 3 species)
- Explore possible reasons why the dino-v2 model performance deteriorates after training and adjust it if possible