DS@GT ARC at AnimalCLEF 2026: Species-Aware Graph Construction for Multi-Species Animal Re-Identification

Evan Sinclair Smith[1,*], Anthony Miyaguchi[1,*], Snigdha Palamari[1,*] and Danté Evangelista[1,*]

[1]Georgia Institute of Technology, North Ave NW, Atlanta, GA 30332

Abstract

Automated individual animal re-identification is essential for large-scale biodiversity monitoring; however, field imagery complicates separating identity cues from nuisance variation in pose, illumination, background, resolution, and species-specific morphology. The DS@GT ARC submission to AnimalCLEF 2026 introduces a multi-species image-clustering system for re-identifying Eurasian lynx, fire salamanders, loggerhead sea turtles, and Texas horned lizards. Instead of relying on a single descriptor or nearest-neighbor retrieval, this approach formulates re-identification as species-aware graph construction over candidate image pairs. The pipeline integrates tailored preprocessing, global candidate retrieval, LightGlue-based local verification with multiple keypoint families, LightGBM pair scoring, conservative edge admission, and Leiden community detection. This design directly addresses a primary failure mode of clustering-based re-identification: high-scoring false pairs that act as bridge edges and merge distinct individuals through transitive closure. Across species, ablation studies demonstrate that local feature support, foreground-aware preprocessing, and species-specific backbone selection enhance pair evidence, while graph operating points determine the trade-off between fragmentation and over-merging. The selected submission achieved a public ARI of 0.733 and a private ARI of 0.674, ranking fifth among 230 teams. These results indicate that robust wildlife re-identification requires not only strong visual representations but also calibrated integration of global similarity, local identity markings, neighborhood context, and graph-level constraints. The code can be found at https://github.com/dsgt-arc/animalclef-2026.

Keywords

Re-identification, Species Identification, Global Descriptors, Local Feature Matching, Graph Clustering, Fine-Grained Visual Classification, AnimalCLEF 2026, CEUR-WS

1. Introduction

Individual animal identification is a central task in biodiversity monitoring because repeated sightings of the same animal support estimates of population size, movement, and behavior. In this study, we developed a multi-stage computer vision system for image-based individual re-identification in AnimalCLEF 2026, a field-style clustering challenge hosted through Kaggle and described in the official AnimalCLEF overview [1, 2]. The challenge provides images from four species with scientific names: Lynx lynx (Eurasian lynx) [3], Salamandra salamandra (fire salamander), Caretta caretta (loggerhead sea turtle) [4], and Phrynosoma cornutum (Texas horned lizard) [5]. The competition evaluated submissions with the Adjusted Rand Index (ARI), which measures agreement between predicted clusters and ground-truth identities while penalizing both over-splitting one individual into multiple clusters and over-merging different individuals into the same cluster [1, 6].

AnimalCLEF is part of the LifeCLEF lab, which aims to automate identification and understanding of life forms in the field of biodiversity informatics, implementing machine learning and computer vision methods [7]. LifeCLEF is part of the Conference and Labs of the Evaluation Forum (CLEF), which is structured into two parts: evaluation labs like LifeCLEF and the peer-reviewed Conference [8]. AnimalCLEF 2026 changed the task from the 2025 open-set assignment setting to a clustering setting across four species, including an unlabeled Texas horned lizard discovery set. In AnimalCLEF

CLEF 2026: Conference and Labs of the Evaluation Forum, September 21-24, 2026, Jena, Germany

*Corresponding author.

esmith446@gatech.edu (E. S. Smith); acmiyaguchi@gatech.edu (A. Miyaguchi); spalamari3@gatech.edu (S. Palamari); devangelista8@gatech.edu (D. Evangelista)

0009-0007-9097-275X (E. S. Smith); 0000-0002-9165-8718 (A. Miyaguchi); 0009-0001-1729-0396 (S. Palamari); 0009-0005-7587-6271 (D. Evangelista)

Figure 1: Sample images from each of the animals in the dataset: Lynx, Salamander, Turtle, Lizard.

2025, competitors were challenged with designing a model that would determine whether the depicted animal was new (not present in the training set) or known (where its identity must be provided) [9]. AnimalCLEF 2026 expanded the competition by focusing on individual-animal clustering across multiple species, including a discovery-only Texas horned lizard subset, and by evaluating submissions with the Adjusted Rand Index (ARI) [1, 2]. Our system placed 5th out of 230 teams in AnimalCLEF 2026 by combining global descriptors, local feature matching, learned pair scoring, and conservative graph construction across the four species. This approach demonstrates that calibrated graph construction, not just stronger descriptors, is central to reliable clustering-based re-identification.

2. Related Work

Animal re-identification aims to recognize unique animals across images by using traits such as coat patterns, spots, scale patterns, body texture, or other stable natural markings [10, 11]. Animal re-identification enables estimates of population size, migration, and behavior by linking sightings of the same individual across time and sites. Field images vary in pose, lighting, background, and resolution. Prior AnimalCLEF work treats background, pose, lighting, and segmentation as robustness concerns because models can rely on capture conditions instead of identity markings when transferred to new images [12, 13].

More recent AnimalCLEF approaches start with global image descriptors that convert each image into a feature vector and place visually similar animals close together. Domain-specific models such as MegaDescriptor are trained on animal re-identification datasets and have outperformed general-purpose vision models in individual-animal matching studies [10, 14, 15]. The DS@GT 2025 pipeline compared DINOv2 and MegaDescriptor and found that triplet learning was more effective when applied to MegaDescriptor, a model already trained for animal re-identification [16, 14]. Other work has explored MiewID as a multi-species embedding model, where training across several species improved candidate retrieval for previously unseen animals [17, 18, 15].

Because global descriptors summarize the entire image, they may miss small but important local identity cues. AnimalCLEF 2025 participants combined global retrieval with local feature matching. WildFusion calibrates and fuses global similarity scores, such as MegaDescriptor or DINOv2, with local matching scores from methods such as LoFTR and LightGlue [19, 20, 21]. The first-place hybrid global-local system first used global similarity to find likely matches, then used a weighted fusion of local matching methods to compare fine details such as markings and body patterns [12]. Similarly, multilevel feature fusion methods combine deep image features with keypoint-based features, then use a threshold to decide whether an animal is known or unknown [22]. Fusion pipelines using MegaDescriptor, ALIKED, EVA02, WildFusion calibration, and test-time augmentation reported gains from combining

global, local, and semantic features across species [13].

Prior systems also learned calibrated pairwise similarity scores. Gradient-boosting approaches combined pretrained embeddings, pairwise comparison features, and supervised decision boundaries to determine whether two images depicted the same animal [23]. Other pipelines used WildFusion, XGBoost, and ArcFace-based fine-tuning to combine different types of information, including global descriptors, local matches, neighbor context, and species-specific models [15].

The 2026 clustering setting makes the task much more challenging than in the previous open-set assignment task. Prior systems relied on global retrieval, local verification, and learned pair scoring; the clustering setting also requires deciding which pair signals to treat as identity edges. A high-scoring false pair can be harmless if rejected, but damaging if it becomes a bridge edge inside a transitive graph.

3. Data Overview

The four animals featured in this year's competition have been selected to demonstrate the diversity of animal photography and the associated challenges of re-identification. There are images on land and at sea, in various lighting conditions and perspectives, with and without labels for training. Figure 1 presents just a few of the images within the dataset to give a flavor of the unique traits that the animals possess.

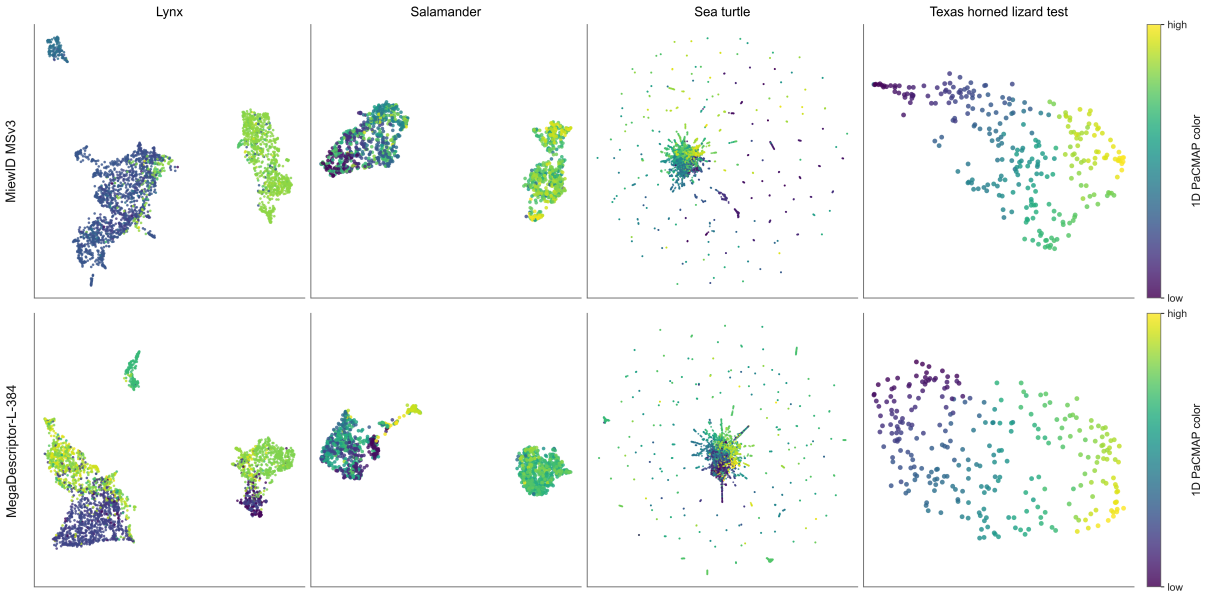

Figure 2: Visualization of clustering behavior for different animal species per MegaDescriptor and MiewID. The clustering behavior observed in Lynxes is likely due to a shift between day and night imaging. The clustering behavior of salamanders likely stems from images taken from a natural top-down perspective rather than those captured and held in hand. Sea turtles form a large, connected component near the center, with many surrounding clusters, likely due to the nature of thumbnail-sized images in the dataset. The distribution over horned lizards demonstrates a smooth topology in the localized embedding space which reflects the consistent photography framing.

The low-dimensional embedding structure in Figure 2 mirrors these per-species traits and challenges. Lynxes are pre-segmented images from a camera-trap dataset and include stills extracted from videos taken in daylight, alongside low-light camera footage from nights. The main identifier is the spotting and striping pattern on the fur, which is visible in many of the images. The primary challenge is a significant domain shift in lighting, as many low-light images are difficult to differentiate. Salamanders are photographed in a natural forest environment, with some photos of the backs showing unique yellow spots and glossy skin, and a few images taken with hands in the frame. The turtle datasets are taken underwater, primarily capturing the head and scale-like patterns that uniquely identify individuals [4].

Table 1
Per-dataset statistics for the AnimalCLEF 2026 corpus. Identity statistics are computed on the training (database) split only. TexasHornedLizards has no training data (discovery-only task).

Dataset	Train	Test	IDs	Median	Max	Singletons
LynxID2025	2,957	946	77	17	353	7.8 %
SalamanderID2025	1,388	689	587	1	12	52.8 %
SeaTurtleID2022	8,729	500	438	13	190	0.2 %
TexasHornedLizards	—	274	—	—	—	—
Total	13,074	2,409	1,102	—	—	—

This dataset contains many thumbnails and other low-quality images that are unlikely to be informative for re-identification. The horned lizard's images are framed, showing the animal's belly scales held up by hand. There are a few dozen clusters of darkened scales along the lizard's belly that appear to be its primary identifying markers. The challenge for the lizards is that there are no labeled individuals in this set, and thus it is a pure clustering challenge.

Alongside these visual differences, the species also differ in their statistical structure (Table 1). Some species have fewer examples than others, and salamanders in particular contain many singleton identities with only one labeled training image. At roughly 13,000 training and 2,400 test images, the corpus is small enough to run on Kaggle notebooks or a personal machine.

4. Methodology

We model individual animal re-identification as species-specific graph construction over candidate image pairs. For each species, the system treats test images as nodes, proposes likely same-individual pairs, scores them using global, local, and neighborhood evidence, and converts retained edges into predicted identity clusters. Cross-species edges are not considered. The pipeline consists of species-specific preprocessing, candidate retrieval, pair scoring, edge admission, and graph clustering (Figure 3).

4.1. Preprocessing and Global Representations

Image preprocessing was selected separately for each species before retrieval and local verification. We used square padding, cropped black borders around Lynx images to keep the animal in view, applied vertical-flip test-time augmentation, and used SAM3 masks to isolate salamanders from noisy backgrounds [25].

Because the four species differ in pose, background structure, marking scale, and supervision, representation choices are evaluated by species rather than through a single universal backbone. The global backbones are MegaDescriptor-L-384 [10] and MiewID-MSv3 [17, 18]. We also tested the fusion of edge-level evidence from both backbones.

4.2. Candidate Pair Retrieval and Neighborhood Context

Candidate pair retrieval reduces the otherwise quadratic set of possible image pairs to a smaller shortlist. For each species, the nearest-neighbor search first keeps a neighborhood around each image. A second, usually smaller budget decides which candidate pairs receive expensive local verification and pair scoring. During training, a separate budget controls how many labeled neighbor pairs are mined to provide positive examples and hard negatives for the pair scorer. Test-time budgets instead determine which unlabeled pairs are eligible to become graph edges.

The retrieval stage also produces neighborhood-context features used downstream. For a candidate pair, these include retrieval rank, reciprocal-neighbor indicators, shared-neighbor counts, and Jaccard overlap between local neighborhoods. These features encode whether direct visual similarity agrees with the surrounding neighborhood structure.

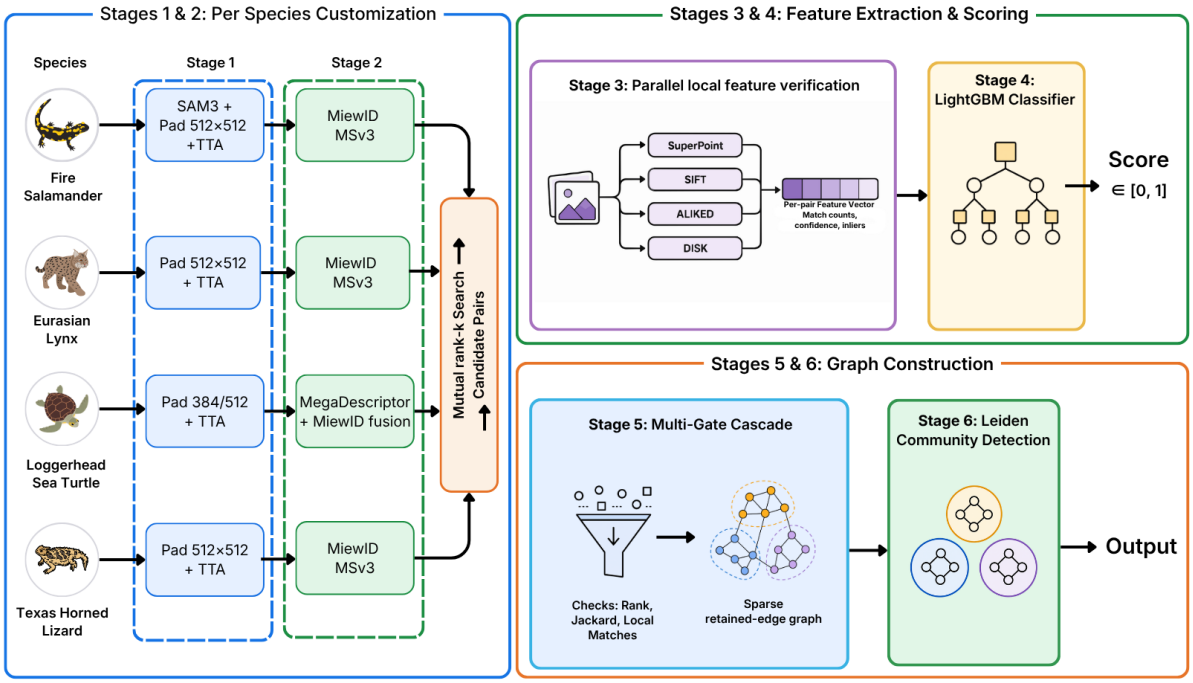

Figure 3: Species-aware submission pipeline. The system follows a fixed sequence: species-specific preprocessing and global representation, mutual-rank candidate retrieval, LightGlue local verification [20], LightGBM pair scoring [24], edge-admission gates, and graph clustering.

4.3. Local Verification and Pair Features

Global descriptors summarize the full image, but individual identity often depends on local markings, spots, texture, or scale patterns. Shortlisted pairs receive local-verification features from LightGlue matching with SuperPoint, SIFT, ALIKED, and DISK keypoints [20, 26, 27, 28, 29]. Each branch compares the two images at a fixed working resolution and returns match-count and match-confidence summaries of the two images, including mean confidence, total support, confidence ratios, and thresholded support counts.

Local matching does not directly assign identities. Instead, it adds pair-level evidence about whether two shortlisted images share the same markings, spots, or texture. A pair can look similar under the global embedding model but still be rejected if local matching and neighborhood context are weak. Conversely, a lower-ranked retrieval pair can still be kept if local matching provides strong support and the surrounding graph evidence is consistent with a repeated individual.

4.4. Learned Pair Scoring

The pair scorer estimates whether a candidate image pair depicts the same individual. We train a LightGBM gradient-boosted decision-tree model [24] on labeled candidate pairs from the training identities, using global similarity features, retrieval-rank features, neighborhood-context features, and local-verification summaries. Positive examples share a training identity, while negative examples come from different identities. Per-image caps and rank-band sampling are used to reduce class imbalance and expose the model to hard negatives rather than only visually unrelated pairs.

Pair scores are not final identity predictions. They are edge-strength proposals for graph construction. One false-positive edge can merge two otherwise separate identity components, whereas one false-negative edge may split a repeated individual into smaller clusters. The graph stage applies additional gates to convert pair scores into retained identity edges.

4.5. Graph Construction and Clustering

Graph construction converts scored candidate pairs into a sparse graph. Images are nodes, candidate same-individual relationships are proposed edges, and retained edges are selected using pair score, retrieval rank, local support, and neighborhood consistency. The edge filters include high-confidence core thresholds, lower-expansion thresholds, rank caps, local-support gates, common-neighbor and Jaccard gates, and component-shape controls, such as degree or component caps, when configured.

After edge filtering, the retained graph is clustered to produce predicted identity labels. Connected components treat every retained edge as a hard union. Leiden clustering can instead split a weighted graph into communities when edge strengths vary within a connected component [30]. Backend comparisons are meaningful only when the edge set and weights being clustered are comparable.

4.6. Final Competition Configuration

The final competition system used species-specific operating points rather than a single global setting. This reflected the structure of the four datasets: LynxID2025 required conservative graph growth across camera-trap domains, SalamanderID2025 contained many singleton identities where false bridge edges were costly, SeaTurtleID2022 favored stable global retrieval under low-resolution imagery, and TexasHornedLizards lacked labeled identities for supervised pair mining. Across species with local verification, we summarized four LightGlue branches, SuperPoint, SIFT, ALIKED, and DISK, using match-count and confidence features in a LightGBM pair scorer. The scorer used 250 trees, learning rate 0.06, 63 leaves, row and column subsampling of 0.8, and ℓ_2 regularization of 0.5; its output was treated as graph-edge evidence rather than as a direct identity classifier.

Table 2
Species-specific settings in the selected competition-period system. Pair budgets are shown as retrieved/scored candidate counts.

Species	Representation and preparation	Candidate and training budgets	Graph setting and final role
LynxID2025	MiewID, 512 px; vertical flip	test 40/20; train 30/20, top-48	core/expand 0.93/0.88; Leiden 0.02, edge floor 0.82; at least two shared neighbors; known-identity attachment threshold 0.94, margin 0.04, support 2; retained baseline
SalamanderID2025	MiewID, 512 px; SAM3 foreground mask and square pad; vertical flip	test 224/128 with 50 context neighbors; train 80/40, top-48; graph 60/30 with 60 context neighbors and top-72 tuning	core/expand/very-core 0.9618/0.9255/0.9811; singleton/merge 0.9829/0.9546; Leiden 0.03, edge floor 0.80, seed 7, eight iterations; promoted replacement
SeaTurtleID2022	MegaDescriptor, 384 px; vertical flip	test 50/25; train 30/30, top-48	core/expand 0.92/0.85; Leiden 0.02, edge floor 0.81; at least two shared neighbors; retained baseline
TexasHornedLizards	MiewID, 512 px	test 60/30; no supervised train-pair mining	core/expand 0.95/0.90; at least three shared neighbors; known-identity attachment disabled; retained baseline

Table 2 summarizes the selected per-species operating points. The selected full system retained the baseline configurations for LynxID2025, SeaTurtleID2022, and TexasHornedLizards, and replaced only the SalamanderID2025 predictions with the more selective SAM3, vertical-flip, MiewID, and Leiden safe-graph configuration. This was the main empirical lesson from our ablations: improvements did not transfer uniformly across species, so the final configuration selected the species-specific operating point with the strongest full-system behavior rather than promoting every exploratory improvement. The code can be found at https://github.com/dsgt-arc/animalclef-2026.

4.7. Evaluation Protocol

Public and private leaderboard scores are aggregate ARI values for complete submissions. The public score is computed on the public portion of the hidden test set, while the private score is computed on the remaining hidden test data; the private leaderboard uses approximately 52% of the test images. The Adjusted Rand Index used for scoring is defined in Eq. 1.

Local validation used identity-grouped cross-validation on the labeled training images. Training identities, not individual images, were assigned to folds, so images of the same known animal could not appear in both the training and validation portions of a fold. For each validation fold, the system predicted clusters for the held-out images and compared them with the known identity labels using ARI. These local CV scores were used to compare candidate settings before leaderboard submission.

Ablation and sensitivity studies differ in what they hold fixed. Some studies change one component, while others remove a component without retraining the pair scorer or retuning graph thresholds. We therefore interpret these studies as sensitivity evidence unless the table or text states that the scorer was retrained and graph settings were retuned.

5. Results

5.1. Main Competition Results

AnimalCLEF 2026 evaluates submissions with the Adjusted Rand Index (ARI), a chance-corrected agreement measure between the hidden identity partition U and the predicted cluster partition V [1, 2, 6]. Let $n_{ij} = |U_i \cap V_j|$, $a_i = \sum_j n_{ij}$, $b_j = \sum_i n_{ij}$, $n = \sum_{ij} n_{ij}$, and $\binom{x}{2} = x(x-1)/2$. The ARI is

$$\mathrm{ARI} = \frac{\sum_{ij} \binom{n_{ij}}{2} - \frac{\sum_i \binom{a_i}{2} \sum_j \binom{b_j}{2}}{\binom{n}{2}}}{\frac{1}{2}\left[\sum_i \binom{a_i}{2} + \sum_j \binom{b_j}{2}\right] - \frac{\sum_i \binom{a_i}{2} \sum_j \binom{b_j}{2}}{\binom{n}{2}}}. \tag{1}$$

A score of 1 indicates identical partitions, values near 0 indicate chance-level agreement, and negative values indicate worse-than-chance agreement. In this challenge, false splits and false merges both reduce ARI.

The strongest competition-period DS@GT result scored 0.73275 public and 0.67424 private ARI, placing 5th on the final private leaderboard among 230 teams [31]. The LynxID2025, SeaTurtleID2022, and TexasHornedLizards blocks were retained from the frozen baseline, while SalamanderID2025 was the only promoted replacement, using MiewID retrieval, SAM3 mask-square-pad preprocessing, vertical-flip augmentation, local verification, raw LightGBM pair scoring, and evidence-gated Leiden clustering. Later, SeaTurtle and Lynx variants showed that stronger local or post-deadline operating points could trade off public and private performance, but they did not yield a better official-period all-species submission. All reported public/private values are full-submission ARI, not hidden per-species ARI.

Table 3 anchors this result against the rest of the competition. The five highest-ranked teams on the private leaderboard fell within a narrow band of about 0.030 ARI (0.674–0.704), with DS@GT trailing fourth place by roughly 0.001 ARI, while all five scored approximately three times higher than the organizer WildFusion and MegaDescriptor baselines (0.206–0.212) [31]. The challenge was therefore both difficult, with baselines clustered near 0.21, and closely contested among the leading systems.

Public/private scores progressed from early MegaDescriptor without trained pair scorers to the DS@GT all-species system and the selected Salamander replacement, which produced the highest competition-period public score (Figure 4). Post-deadline SeaTurtle improved public score without improving private score, while wider Lynx graphs improved private score at public-score cost.

Table 3
AnimalCLEF 2026 final (private) leaderboard: top five teams and organizer baselines. The private leaderboard determines the official ranking. DS@GT (team *DS@GT LifeCLEF*) placed 5th of 230 teams on the private split and 3rd on the public split (public ARI 0.73275). Organizer baseline entries are listed for reference and are not ranked. Source: Kaggle leaderboard [31].

Rank	Team	Private ARI	Entries
1	M.I.A	0.70393	270
2	Ram Lab	0.70050	154
3	VIPL-VSU	0.69632	88
4	HSU Lab	0.67537	174
5	**DS@GT LifeCLEF (ours)**	**0.67424**	**37**
Organizer baselines			
—	WildFusion + HDBSCAN	0.21221	1
—	MegaDescriptor + MiewID + DBSCAN	0.20646	1

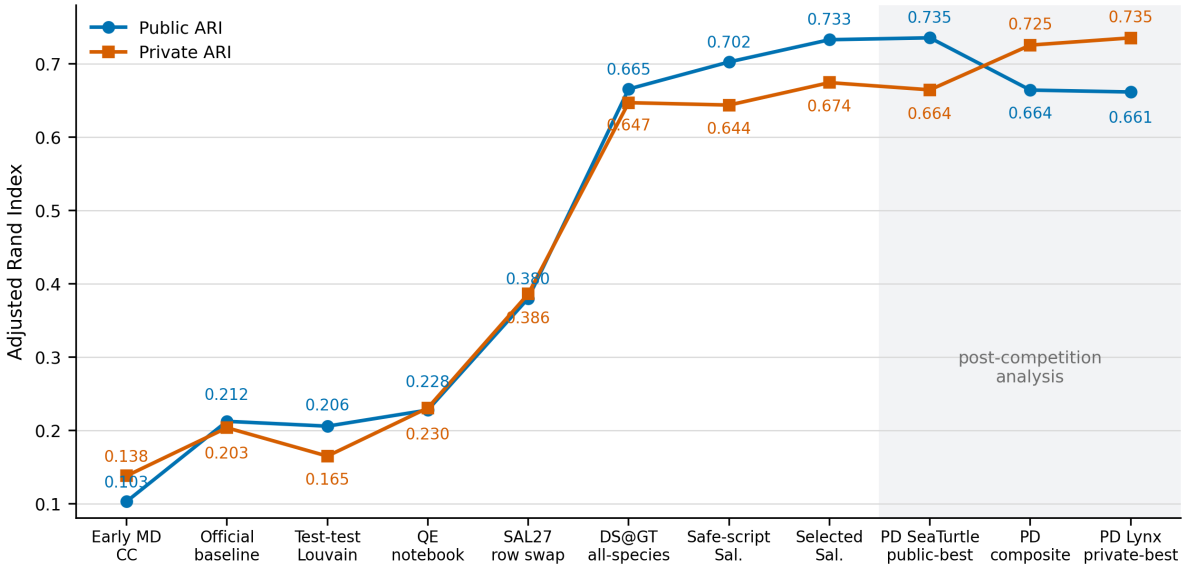

Figure 4: Public/private submission score trajectory. Scores include early MegaDescriptor and test-to-test graph submissions, the DS@GT all-species system, the selected Salamander replacement, and post-deadline analyses. All scores are full-submission ARI. Shaded points indicate submissions scored after the official competition closed.

5.2. Global Retrieval and Backbone Selection

Retrieval metrics differed by species: MiewID MSv3 [18, 17] produced stronger curves for Lynx and Salamander than MegaDescriptor [10] in the evaluated budgets, while SeaTurtle retrieval was strongest when MegaDescriptor and MiewID evidence were fused (Figure 5; Tables 4–5). The final ARI still relied on local verification, edge admission, and clustering.

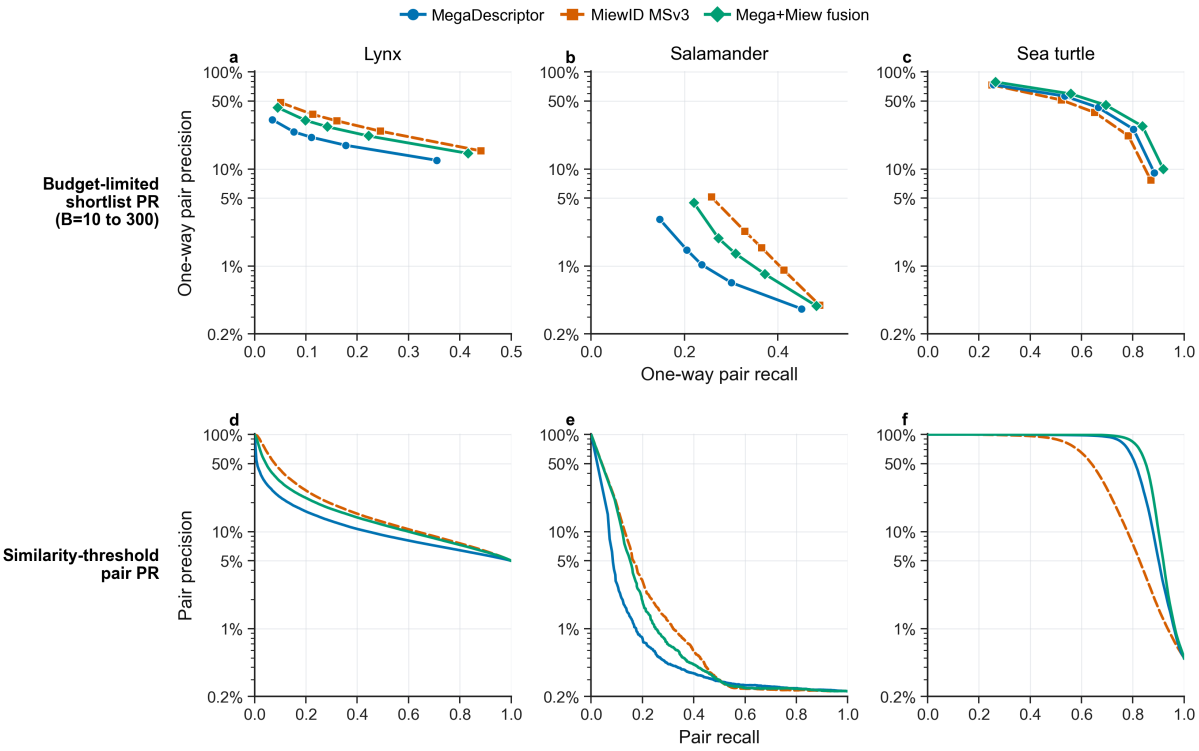

Figure 5: Global backbone precision-recall tradeoffs by species. Top panels show fixed-budget shortlist retrieval; bottom panels show threshold-sweep retrieval before local verification or graph filtering.

Table 4
Fixed-budget global shortlist precision-recall by species. Cells give one-way pair recall/precision as percentages rounded to one decimal. Bold model names mark the strongest displayed retrieval model for each species across the shown budgets.

Species	Model	Retrieval budget B; recall / precision (%)				
		10	30	50	100	300
Lynx	MegaDescriptor	3.5% / 32.3%	7.7% / 24.1%	11.1% / 21.2%	17.8% / 17.5%	35.5% / 12.3%
	MiewID MSv3	5.1% / 48.3%	11.3% / 36.5%	16.0% / 31.3%	24.5% / 24.5%	44.1% / 15.4%
	Mega+Miew fusion	4.5% / 43.0%	10.0% / 31.7%	14.2% / 27.2%	22.2% / 21.9%	41.6% / 14.5%
Salamander	MegaDescriptor	14.7% / 3.0%	20.5% / 1.5%	23.7% / 1.0%	30.1% / 0.7%	45.1% / 0.4%
	MiewID MSv3	25.8% / 5.1%	33.0% / 2.3%	36.6% / 1.6%	41.3% / 0.9%	49.0% / 0.4%
	Mega+Miew fusion	22.0% / 4.5%	27.3% / 1.9%	31.0% / 1.3%	37.2% / 0.8%	48.3% / 0.4%
SeaTurtle	MegaDescriptor	25.6% / 73.7%	53.4% / 56.4%	66.5% / 43.0%	80.2% / 25.7%	88.4% / 9.2%
	MiewID MSv3	25.0% / 73.5%	52.2% / 51.4%	65.2% / 38.2%	78.3% / 22.0%	87.0% / 7.7%
	Mega+Miew fusion	26.5% / 78.7%	55.8% / 59.4%	69.5% / 45.4%	83.7% / 27.6%	91.9% / 9.9%

Table 5
Score-threshold global pair-retrieval summary. Pair prior and candidate counts are species-level. AP values are rounded to three decimals; bold marks the strongest threshold curve within each species.

Species	Pair prior	Positive / candidate pairs	Average precision		
			MegaDescriptor	MiewID MSv3	Fusion
Lynx	0.050	218,226 / 4,370,446	0.123	**0.198**	0.169
Salamander	0.002	2,183 / 962,578	0.055	**0.082**	0.080
SeaTurtle	0.005	186,777 / 38,093,356	0.804	0.636	**0.844**

5.3. Component Evidence

Table 6 shows which kinds of evidence the LightGBM pair scorer relied on most when judging candidate image pairs [24, 32]. Across species, retrieval-rank features carried the highest per-feature attribution by a wide margin, followed by the DISK LightGlue branch. The SuperPoint branch was most prominent for Salamander, while for SeaTurtle the Global cosine features (driven primarily by MiewID similarity) rose to the second strongest per-feature family. Neighborhood-context and SIFT features contributed the least per feature, consistent with a scorer that combines several signals rather than ranking pairs by a single embedding score.

Table 6
Feature-family attribution in the LightGBM pair scorer. Each cell shows the average attribution of a single feature in that family when the scorer judges a candidate same-individual image pair, measured by mean|SHAP| [32] and divided by the family's feature count so that wide families (the four LightGlue branches, WildFusion summaries) are directly comparable with narrow ones (Global cosine, Neighborhood context). Values are in the model's raw-output (log-odds) units, so a value of 0.10 means each feature in that family on average shifts the pair-decision logit by about ± 0.10 per pair — larger is more influential. *Mean* averages the three species. Bold marks the most-attributed family within each species column; rows are ordered by *Mean*.

Feature family	Salamander	SeaTurtle	Lynx	Mean
Retrieval rank	**0.222**	**0.277**	**0.259**	0.253
LightGlue DISK	0.160	0.124	0.160	0.148
Global cosine	0.097	0.143	0.074	0.105
LightGlue SuperPoint	0.140	0.085	0.085	0.103
WildFusion summaries	0.062	0.067	0.056	0.062
LightGlue ALIKED	0.053	0.046	0.048	0.049
Neighborhood context	0.043	0.034	0.041	0.039
LightGlue SIFT	0.038	0.033	0.041	0.037

Table 7
Ablations grouped by species. Public/private scores are full-submission ARI, rounded to three decimals. Each species block is headed by its selected submission (reference); subsequent rows change or remove a single component.

Reference or change	Public	Private
SalamanderID2025		
Best submission (reference)	**0.733**	**0.674**
Remove SAM3 mask-square-pad	0.660	0.629
MegaDescriptor instead of MiewID	0.650	0.637
Remove vertical-flip TTA	0.714	0.663
SeaTurtleID2022		
Best submission (reference)	**0.688**	**0.630**
MiewID only	0.684	0.629
MegaDescriptor+MiewID edge mean	0.679	0.616
Remove vertical-flip TTA	0.670	0.618
LynxID2025		
Best submission (reference)	**0.702**	**0.703**
Remove vertical-flip TTA	0.656	0.616

Salamander was most sensitive to the global representation and foreground pre-processing: replacing MiewID with MegaDescriptor and removing the SAM3 mask-square-pad pre-processing both reduced the public and private scores, while removing vertical-flip augmentation caused a smaller but consistent loss. SeaTurtle showed the opposite pattern: although retrieval evidence favored MegaDescrip-

tor+MiewID fusion, the downstream replacement controls favored MegaDescriptor alone, suggesting that stronger retrieval did not automatically translate into better graph clustering. Removing TTA from SeaTurtle lowered the public score with little change to the private score. For Lynx, removing vertical-flip TTA reduced the reference from $0.702/0.703$ to $0.656/0.616$. Across species, the LightGlue match families also carried substantial attribution in the learned pair scorer (Table 6), consistent with local evidence being a useful graph-construction signal.

5.4. Edge Admission and Cluster Shape

Graph construction changed ARI by controlling retained edge sets and component sizes. Plain threshold graphs and support-gate removals produced graph shapes consistent with bridge-edge over-merging, creating large transitive components and reducing ARI. A post-deadline backend-isolation check that ran connected components and Leiden on matched retained edge sets found lower transitive over-merging with Leiden, most clearly for SeaTurtle, but leaderboard scores remained mixed. The selected Salamander edge filter retained a small graph, whereas ungated threshold and support-gate removal variants produced larger components associated with lower private ARI (Figure 6).

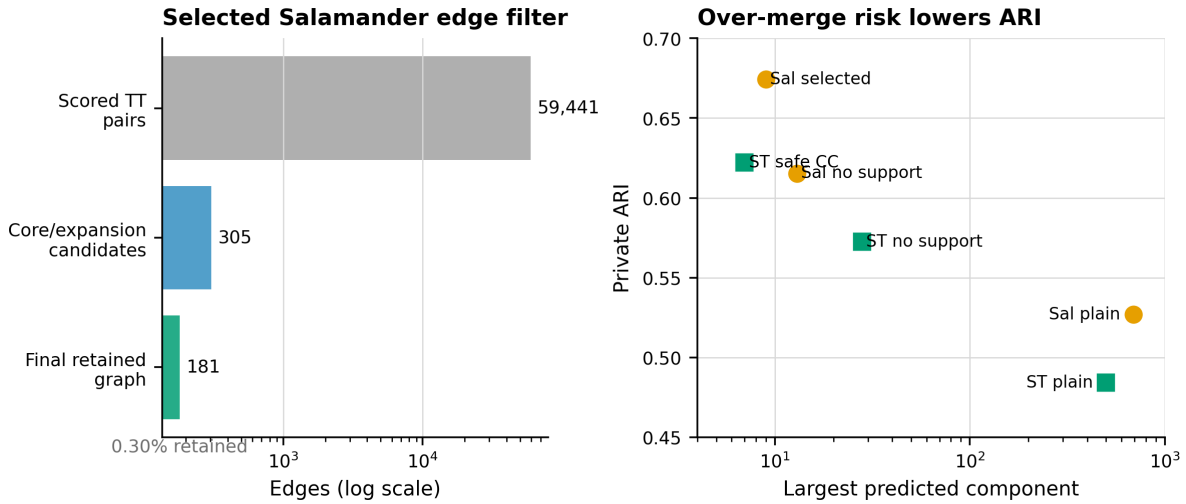

Figure 6: Edge admission and over-merge risk. The selected Salamander edge-admission filter retains a small subset of the candidate-pair universe, while ungated threshold graphs and local-support gate removals produce larger components and lower private ARI. The largest failures occur when threshold graphs collapse many images into one dominant component.

The selected Salamander submission reduced fragmentation relative to the frozen baseline while keeping the largest component small: clusters fell from 599 to 574, singletons from 549 to 523, and same-cluster pairs rose from 158 to 260. The private-strong widened-CC Lynx analysis instead used a 311-image largest component and 55,447 same-cluster pairs, coinciding with higher private ARI and lower public ARI.

6. Discussion

Animal re-identification in AnimalCLEF 2026 depended on edge admission as much as descriptor quality. Stronger retrieval increased both true same-individual candidates and plausible false candidates; once accepted as graph edges, false pairs could propagate through transitive closure or community detection. The selected submission combined broad candidate generation with conservative graph construction, producing more non-singleton structure while keeping the largest predicted component small (Figure 6). K-reciprocal reranking [33] was also tested as a post-retrieval refinement, but did not improve ARI for any species.

The strongest competition-period result came from the selected Salamander submission. It scored 0.733 public / 0.674 private ARI and placed DS@GT 5th out of 230 teams. Relative to the frozen baseline, salamander reduced fragmentation while preserving a conservative graph shape: clusters fell from 599 to 574, singletons fell from 549 to 523, and the largest predicted component increased only from 7 to 9. The gain, therefore, came from adding supported same-individual links, not from letting a few large components absorb uncertain images.

Salamander also had an exploratory spot-based branch. We tested HSV color thresholding, SAM3 prompts for yellow spots, spot masks, centroid geometry, constellation and Delaunay features, and Kornia-based keypoint matching. These probes were motivated by the species' distinctive markings, but they remained exploratory. Spot-level cues may be useful, but only when they improve calibrated pair evidence and edge admission.

SeaTurtle retrieval metrics did not translate reliably into final clustering performance. The fused MegaDescriptor and MiewID retrieval curve was strongest for candidate generation, but downstream graph controls did not consistently convert that advantage into private-score gains. Local support still mattered: disabling the local-support gate reduced the SeaTurtle graph-support candidate from 0.622 to 0.572 private ARI.. The final graph operating point, rather than retrieval alone, shaped transfer.

Lynx had the strongest public/private graph-shape divergence. In post-deadline Lynx analyses, using aggregate public/private scores, larger connected structures improved private ARI and reduced public ARI. Our Lynx graph-support replays were mostly neutral around the low accepted-edge connected-components reference, so no single local branch was globally preferable. The main Lynx uncertainty was the graph operating point: the public and private splits rewarded different component-size and edge-admission tradeoffs.

6.1. Lynx 3D

One exploratory direction used pose estimators to re-project images and reduce the effects of occlusion and soft-body deformation. The hypothesis was that orientation information is implicitly captured in images, and that projecting animals into a canonical pose could help local feature matchers align individual-specific patterns in pixel space.

The first direction followed pelage-unwrapping [34], in which fur texture is flattened onto a surface. The approach used monocular depth and surface-normal estimates, then optimized a projection that made the normals more consistent. We tried to reproduce this direction and used Lynx as the main testbed because Lynx have fur and are often less occluded than the other species. In practice, the resulting unwraps were not usable for downstream matching. The distortion was severe, and each image produced its own UV space rather than a shared space across images. The underlying model also relied on foreground-background contrast, whereas the Lynx images in the competition were already pre-segmented and many were captured at night, creating a likely domain shift.

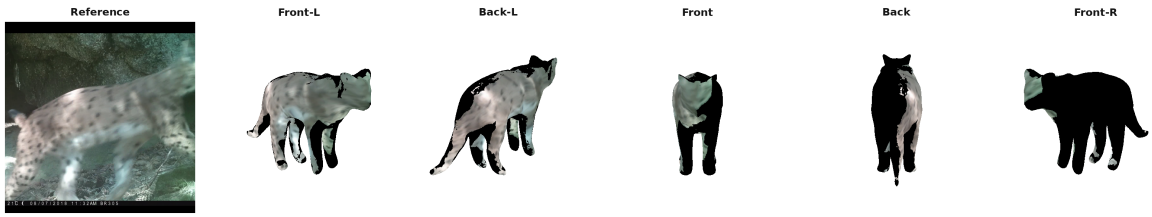

Figure 7: Lynx 3D and UV-projection exploration. 3D-Fauna-style reconstruction produced visually coherent single-image projections in some cases, but multi-image synthesis and downstream matching were limited by pose error, sparse surface coverage, and domain shift.

The second direction used 3D-Fauna, a feed-forward model that reconstructs posed and textured 3D meshes from quadruped images [35]. Given a stack of Lynx image embeddings, we averaged the embeddings to obtain a single shared UV space. The intended benefit was that pixels from the original

image could be projected onto a shared surface where the same identifying markings would occupy similar locations across pose and viewpoint changes.

In practice, pose and orientation were often plausible but difficult to validate quantitatively. The generated latent texture was not useful for re-identification because it mostly contained blurry color regions rather than high-frequency individual markings. Gaussian point-cloud projection and related rendering attempts produced either high-frequency speckling from sparse surface coverage or blurry texture from uncertain correspondence. A single-image UV render was more visually coherent (Figure 7), but the resulting image was outside the training distribution of global embedding models.

We therefore treated UV projections as possible local-matching inputs rather than as replacement global embeddings. Local matchers can be more tolerant of unusual image geometry when the feature extractor captures repeatable visual structure; SIFT, for example, is not learned from the same wildlife-image distribution as the global embedding models. For Lynx, however, adding 3D-Fauna UV projections degraded overall model performance.

6.2. SeaTurtle Upscaling

SeaTurtle also received an exploratory 3D treatment, but a quadruped model is poorly matched to turtle heads. The goal was to model the head, where large-scale shaped markings and beak geometry carry much of the identity evidence. A separate shape-embedding parameterization had limited success. Even with a general-purpose feature extractor such as DINOv3 [36], the model often failed to recover a useful orientation or pose when the head was represented as a semi-ellipsoid or related geometric surface. Background water was also projected onto the surface in unnatural ways, so this SeaTurtle 3D direction was not used in the final pipeline.

The 3D exploration still exposed useful preprocessing issues. We needed a segmentation model that could separate the turtle head from the background water. SAM3 center-object segmentation worked well for many images, but small thumbnail-sized images and full-body turtle poses led to degenerate cases. We also tried upscaling very small images, often less than 50×50 pixels, because simple interpolation produced blurry results. Neural upscaling and JPEG artifact removal were considered as ways to recover line and scale-boundary evidence that might help turtle re-identification, but these steps were not promoted into the final system.

6.3. Texas Horned Lizard Detections

Texas horned lizards posed challenges due to the lack of labeled training identities, which prevented the use of the same validation loop as for Lynx, Salamander, and SeaTurtle. The lizards were photographed with consistent framing and pose, and individual identification relied on their unique belly scaling patterns. This pattern forms a regular lattice, allowing for quantification of total area in units of scales and mapping of binary coloring on a shared template. The labeling strategy involved normalizing image sizes to ensure a consistent pixel-to-unit-distance ratio, followed by cropping each image to a square grid ten units wide and manually labeling the visible scales. A U-Net model was employed to predict scale locations, with its output used to pre-annotate subsequent annotation rounds. This semi-supervised process produced a model capable of identifying individual scales (Figure 8).

Although this model was not deployed in production, future work could leverage it to determine the diameter of the belly section and the relative positions of dark spots from the lizard's center. Subsequent methods could include point matching and custom fingerprinting. The first method would use the detected points as features for algorithms such as LightGlue, aiming to find a bipartite matching between two point sets that preserves spatial locality between images. The second approach would generate features from the images, such as U-Net embeddings, and directly cluster them, assuming that scale detection naturally clusters individuals. This process may be further refined by generating a regular grid, normalizing dot locations, and comparing direct matches within this simplified geometry to enhance interpretability.

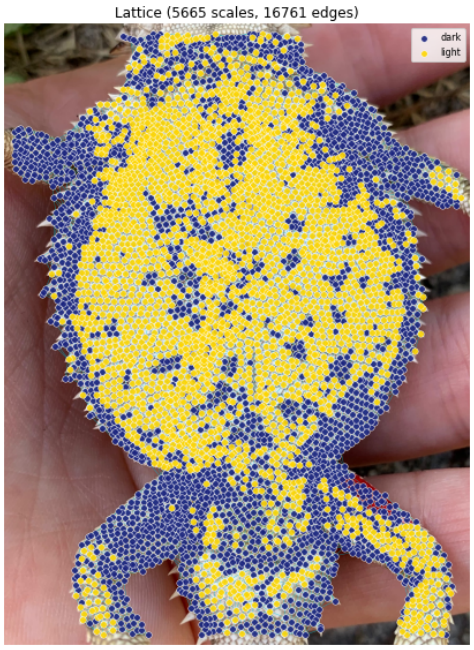

Figure 8: Texas horned lizard scale-detection exploration. Scale-level segmentation and Delaunay-style geometry could provide species-specific cues, but the current evidence is exploratory until connected to identity-clustering validation.

7. Future Work

The highest-priority future work is controlled operating-point validation. Small changes in retrieval budget, edge-admission thresholds, and clustering backend changed whether candidate pairs became useful identity links or false bridge edges, so future studies should sweep threshold values under fixed retrieval caches and report ARI together with graph-shape diagnostics such as cluster count, singleton count, largest component, and same-cluster pairs. This validation protocol is especially important for SeaTurtle and Lynx, where retrieval quality and final clustering performance were not always aligned.

Once the operating point is controlled, two additions are natural to test. Texas horned lizards should be evaluated with a species-specific ventral-spot branch when usable ventral crops are available; Biffi et al. [5] report that HotSpotter [37] matching on ventral spot crops reaches 94% match accuracy against PIT-tag and multilocus-genotype ground truth, with near-100% accuracy when top-10 candidate review is allowed. A DINOv3 backbone fine-tuned on WildlifeReID-10k [36, 11] could also refresh the global descriptor, but it should be judged through the full retrieval-to-graph pipeline rather than retrieval metrics alone. Longer term, a Generalized Category Discovery-inspired formulation [38, 39] is worth exploring as an alternative to pair scoring followed by graph construction.

8. Conclusions

The DS@GT AnimalCLEF 2026 system treated clustering-based animal re-identification as a species-aware graph construction problem rather than a single-descriptor ranking problem. The selected submission achieved 0.733 public ARI and 0.674 private ARI, placing 5th out of 230 teams. Post-deadline analyses linked SeaTurtle and Lynx transfer differences to local support, graph shape, and sensitivity to public/private split. Edge admission determined whether retrieved pairs formed a predicted identity structure or incurred over-merge risk.

Acknowledgements

We thank the Data Science at Georgia Tech (DS@GT) CLEF competition group for their support. This research was supported in part by research cyberinfrastructure resources and services provided by the Partnership for an Advanced Computing Environment (PACE) at the Georgia Institute of Technology, Atlanta, Georgia, USA [40].

Declaration on Generative AI

During the preparation of this manuscript, the authors used Grammarly to assist with grammar and spelling checks and ChatGPT to help improve clarity and revise sentence wording. The authors reviewed and edited all AI-assisted text as appropriate and take full responsibility for the final content of the publication.

References

[1] Kaggle, LifeCLEF, AnimalCLEF26 @ CVPR & CLEF, Kaggle competition, 2026. URL: https://www.kaggle.com/competitions/animal-clef-2026, accessed: 2026-05-15; companion challenge page: https://www.imageclef.org/AnimalCLEF2026.

[2] L. Adam, K. Papafitsoros, D. A. Williams, D. Biffi, L. Picek, Overview of AnimalCLEF 2026: Discovery and re-identification of individual animals, in: Working Notes of CLEF 2026 - Conference and Labs of the Evaluation Forum, 2026.

[3] L. Picek, J. Straka, M. Jirik, E. Belotti, M. Dul'a, J. Krausová, M. Bojda, V. Cermak, L. Bufka, R. Dvořák, et al., Czechlynx: A dataset for individual identification and pose estimation of the eurasian lynx, Scientific Data 13 (2026) 511.

[4] L. Adam, V. Čermák, K. Papafitsoros, L. Picek, SeaTurtleID2022: A long-span dataset for reliable sea turtle re-identification, in: 2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2024, pp. 7146–7156.

[5] D. Biffi, M. R. Tucker, A. Ackel, D. A. Williams, Identification of individual Texas horned lizards (Phrynosoma cornutum) using genotypes and ventral spot patterns, Ecology and Evolution 15 (2025) e71167. URL: https://onlinelibrary.wiley.com/doi/full/10.1002/ece3.71167. doi:10.1002/ece3.71167.

[6] L. Hubert, P. Arabie, Comparing partitions, Journal of Classification 2 (1985) 193–218. doi:10.1007/BF01908075.

[7] L. Picek, L. Adam, S. Kahl, R. Bossy, L. Chrobak, H. Goëau, K. Papafitsoros, H. Klinck, W.-P. Vellinga, R. Planqué, T. Denton, K. Barnard, C. Nédellec, L. Deléger, M. Courtin, G. Martellucci, I. Moummad, F. Vinatier, P. Bonnet, A. Joly, Overview of LifeCLEF 2026: AI Challenges for Biodiversity Understanding and Ecosystem Management, in: International Conference of the Cross-Language Evaluation Forum for European Languages (CLEF), Springer, 2026.

[8] CLEF Initiative, CLEF 2026: Conference and labs of the evaluation forum, Conference website, 2026. URL: https://clef2026.clef-initiative.eu/, accessed: 2026-05-15.

[9] LifeCLEF, AnimalCLEF 2025: Individual animal identification, LifeCLEF challenge page, 2025. URL: https://www.imageclef.org/AnimalCLEF2025, accessed: 2026-05-15.

[10] V. Čermák, L. Picek, L. Adam, K. Papafitsoros, WildlifeDatasets: An open-source toolkit for animal re-identification, in: 2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2024, pp. 5941–5951. URL: https://doi.org/10.1109/WACV57701.2024.00585. doi:10.1109/WACV57701.2024.00585.

[11] L. Adam, V. Čermák, K. Papafitsoros, L. Picek, WildlifeReID-10k: Wildlife re-identification dataset with 10k individual animals, in: 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2025, pp. 2090–2100. URL: https://doi.org/10.1109/CVPRW67362.2025.00197. doi:10.1109/CVPRW67362.2025.00197.

[12] R. Pakhomov, G. Demidov, K. Bogdan, S. Lanskikh, D. Dinmuhametov, A. Khlopotnykh, Individual wildlife recognition via hybrid global-local matching and segmentation-aware filtering, in: Working Notes of the Conference and Labs of the Evaluation Forum (CLEF 2025), volume 4038 of *CEUR Workshop Proceedings*, Madrid, Spain, 2025, pp. 3139–3146. URL: https://ceur-ws.org/Vol-4038/paper_251.pdf.

[13] D. Kim, B. Park, H. Bae, S. Lee, C. Lee, Fusion of global and local descriptors with feature calibration for multi-species animal re-identification: AnimalCLEF 2025, in: Working Notes of the Conference and Labs of the Evaluation Forum (CLEF 2025), volume 4038 of *CEUR Workshop Proceedings*, Madrid, Spain, 2025, pp. 3069–3079. URL: https://ceur-ws.org/Vol-4038/paper_245.pdf.

[14] A. Miyaguchi, C. Maruthaiyannan, C. R. Clark, DS@GT AnimalCLEF: Triplet learning over ViT manifolds with nearest neighbor classification for animal re-identification, in: Working Notes of the Conference and Labs of the Evaluation Forum (CLEF 2025), volume 4038 of *CEUR Workshop Proceedings*, Madrid, Spain, 2025, pp. 3128–3138. URL: https://ceur-ws.org/Vol-4038/paper_250.pdf.

[15] N. Semenova, Meta-algorithm for open-set animal re-ID: WildFusion, XGBoost, and dual-backbone ArcFace, in: Working Notes of the Conference and Labs of the Evaluation Forum (CLEF 2025), volume 4038 of *CEUR Workshop Proceedings*, Madrid, Spain, 2025, pp. 3154–3165. URL: https://ceur-ws.org/Vol-4038/paper_253.pdf.

[16] M. Oquab, T. Darcet, T. Moutakanni, H. V. Vo, M. Szafraniec, V. Khalidov, P. Fernandez, D. Haziza, F. Massa, A. El-Nouby, M. Assran, N. Ballas, W. Galuba, R. Howes, P.-Y. Huang, S.-W. Li, I. Misra, M. Rabbat, V. Sharma, G. Synnaeve, H. Xu, H. Jégou, J. Mairal, P. Labatut, A. Joulin, P. Bojanowski, DINOv2: Learning robust visual features without supervision, Transactions on Machine Learning Research (2024). URL: https://openreview.net/forum?id=a68SUt6zFt.

[17] L. Otarashvili, T. Subramanian, J. Holmberg, J. J. Levenson, C. V. Stewart, Multispecies animal re-ID using a large community-curated dataset, 2024. URL: https://arxiv.org/abs/2412.05602. doi:10.48550/arXiv.2412.05602. arXiv:2412.05602.

[18] Conservation X Labs, MiewID-msv3, Hugging Face model card, 2024. URL: https://huggingface.co/conservationxlabs/miewid-msv3, wildlife re-identification feature extractor; model created 2024-10-15 and last modified 2025-02-13; accessed: 2026-05-15.

[19] V. Čermák, L. Picek, L. Adam, L. Neumann, J. Matas, WildFusion: Individual animal identification with calibrated similarity fusion, 2024. URL: https://arxiv.org/abs/2408.12934. doi:10.48550/arXiv.2408.12934. arXiv:2408.12934.

[20] P. Lindenberger, P.-E. Sarlin, M. Pollefeys, LightGlue: Local feature matching at light speed, in: 2023 IEEE/CVF International Conference on Computer Vision (ICCV), 2023, pp. 17581–17592. URL: https://doi.org/10.1109/ICCV51070.2023.01616. doi:10.1109/ICCV51070.2023.01616.

[21] J. Sun, Z. Shen, Y. Wang, H. Bao, X. Zhou, LoFTR: Detector-free local feature matching with transformers, in: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021, pp. 8918–8927. URL: https://doi.org/10.1109/CVPR46437.2021.00881. doi:10.1109/CVPR46437.2021.00881.

[22] J. Tan, A. Wang, Open-set animal re-identification via multilevel feature fusion, in: Working Notes of the Conference and Labs of the Evaluation Forum (CLEF 2025), volume 4038 of *CEUR Workshop Proceedings*, Madrid, Spain, 2025, pp. 3242–3247. URL: https://ceur-ws.org/Vol-4038/paper_258.pdf.

[23] S. Fedorchenko, S. Arefiev, Gradient boosting similarity in entity matching, in: Working Notes of the Conference and Labs of the Evaluation Forum (CLEF 2025), volume 4038 of *CEUR Workshop Proceedings*, Madrid, Spain, 2025, pp. 3011–3018. URL: https://ceur-ws.org/Vol-4038/paper_240.pdf.

[24] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, T.-Y. Liu, LightGBM: A highly efficient gradient boosting decision tree, in: Advances in Neural Information Processing Systems, volume 30, Curran Associates, Inc., 2017, pp. 3146–3154. URL: https://proceedings.neurips.cc/paper_files/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf.

[25] N. Carion, L. Gustafson, Y.-T. Hu, S. Debnath, R. Hu, D. Suris, C. Ryali, K. V. Alwala, H. Khedr, A. Huang, J. Lei, T. Ma, B. Guo, A. Kalla, M. Marks, J. Greer, M. Wang, P. Sun, R. Rädle, T. Afouras, E. Mavroudi, K. Xu, T.-H. Wu, Y. Zhou, L. Momeni, R. Hazra, S. Ding, S. Vaze, F. Porcher, F. Li, S. Li, A. Kamath, H. K. Cheng, P. Dollár, N. Ravi, K. Saenko, P. Zhang, C. Feichtenhofer, SAM

3: Segment anything with concepts, 2025. URL: https://arxiv.org/abs/2511.16719. doi:10.48550/arXiv.2511.16719. arXiv:2511.16719, version 2 revised 2026-03-28.

[26] D. DeTone, T. Malisiewicz, A. Rabinovich, SuperPoint: Self-supervised interest point detection and description, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2018, pp. 224–236. URL: https://openaccess.thecvf.com/content_cvpr_2018_workshops/w9/html/DeTone_SuperPoint_Self-Supervised_Interest_CVPR_2018_paper.html.

[27] D. G. Lowe, Distinctive image features from scale-invariant keypoints, International Journal of Computer Vision 60 (2004) 91–110. doi:10.1023/B:VISI.0000029664.99615.94.

[28] X. Zhao, X. Wu, W. Chen, P. C. Y. Chen, Q. Xu, Z. Li, ALIKED: A lighter keypoint and descriptor extraction network via deformable transformation, IEEE Transactions on Instrumentation and Measurement 72 (2023) 1–16. doi:10.1109/TIM.2023.3271000.

[29] M. Tyszkiewicz, P. Fua, E. Trulls, DISK: Learning local features with policy gradient, in: Advances in Neural Information Processing Systems, volume 33, Curran Associates, Inc., 2020, pp. 14254–14265. URL: https://proceedings.neurips.cc/paper_files/paper/2020/file/a42a596fc71e17828440030074d15e74-Paper.pdf.

[30] V. A. Traag, L. Waltman, N. J. van Eck, From louvain to leiden: Guaranteeing well-connected communities, Scientific Reports 9 (2019) 5233. doi:10.1038/s41598-019-41695-z.

[31] Kaggle, AnimalCLEF26 @ CVPR & CLEF: Leaderboard, Kaggle leaderboard export and private leaderboard API output, 2026. URL: https://www.kaggle.com/competitions/animal-clef-2026/leaderboard, accessed with Kaggle API on 2026-05-15; public export showed DS@GT LifeCLEF score 0.73275 and public rank 3; private leaderboard output showed private score 0.67424 and rank 5.

[32] S. M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, in: Advances in Neural Information Processing Systems, volume 30, Curran Associates, Inc., 2017, pp. 4765–4774. URL: https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf.

[33] Z. Zhong, L. Zheng, D. Cao, S. Li, Re-ranking person re-identification with k-reciprocal encoding, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 3652–3661. URL: https://openaccess.thecvf.com/content_cvpr_2017/html/Zhong_Re-Ranking_Person_Re-Identification_CVPR_2017_paper.html. doi:10.1109/CVPR.2017.389.

[34] A. Algasov, E. Nepovinnykh, F. Zolotarev, T. Eerola, H. Kälviäinen, P. Zemčík, C. V. Stewart, Unsupervised pelage pattern unwrapping for animal re-identification, 2025. URL: https://arxiv.org/abs/2506.15369. doi:10.48550/arXiv.2506.15369. arXiv:2506.15369.

[35] Z. Li, D. Litvak, R. Li, Y. Zhang, T. Jakab, C. Rupprecht, S. Wu, A. Vedaldi, J. Wu, Learning the 3D fauna of the web, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024, pp. 9752–9762. URL: https://openaccess.thecvf.com/content/CVPR2024/html/Li_Learning_the_3D_Fauna_of_the_Web_CVPR_2024_paper.html. doi:10.1109/CVPR52733.2024.00931.

[36] O. Siméoni, H. V. Vo, M. Seitzer, F. Baldassarre, M. Oquab, C. Jose, V. Khalidov, M. Szafraniec, S. Yi, M. Ramamonjisoa, F. Massa, D. Haziza, L. Wehrstedt, J. Wang, T. Darcet, T. Moutakanni, L. Sentana, C. Roberts, A. Vedaldi, J. Tolan, J. Brandt, C. Couprie, J. Mairal, H. Jégou, P. Labatut, P. Bojanowski, DINOv3, 2025. URL: https://arxiv.org/abs/2508.10104. doi:10.48550/arXiv.2508.10104. arXiv:2508.10104.

[37] J. P. Crall, C. V. Stewart, T. Y. Berger-Wolf, D. I. Rubenstein, S. R. Sundaresan, HotSpotter — patterned species instance recognition, in: 2013 IEEE Workshop on Applications of Computer Vision (WACV), 2013, pp. 230–237. doi:10.1109/WACV.2013.6475023.

[38] X. Wen, B. Zhao, X. Qi, Parametric classification for generalized category discovery: A baseline study, in: 2023 IEEE/CVF International Conference on Computer Vision (ICCV), 2023, pp. 16544–16554. URL: https://doi.org/10.1109/ICCV51070.2023.01521. doi:10.1109/ICCV51070.2023.01521.

[39] W. Zhang, B. Zhang, Z. Teng, W. Luo, J. Zou, J. Fan, Less attention is more: Prompt transformer for generalized category discovery, in: Proceedings of the IEEE/CVF Con-

ference on Computer Vision and Pattern Recognition (CVPR), 2025, pp. 30322–30331. URL: https://openaccess.thecvf.com/content/CVPR2025/html/Zhang_Less_Attention_is_More_Prompt_Transformer_for_Generalized_Category_Discovery_CVPR_2025_paper.html. doi:10.1109/CVPR52734.2025.02823.

[40] PACE, Partnership for an Advanced Computing Environment (PACE), 2017. URL: http://www.pace.gatech.edu.