

Species-Aware Metric Graph Clustering for Animal Re-Identification and Discovery at AnimalCLEF 2026

Biao Liu¹, Zhongyuan Han^{1,*} and Zihao Liang¹

¹Foshan University, Foshan, China

Abstract

AnimalCLEF 2026 asks participating systems to partition animal images into identity clusters in a mixed setting that combines re-identification of known individuals with the discovery of previously unseen ones. This paper presents Species-Aware Metric Graph Clustering (SAMGC), Team Bron's submission to the task. SAMGC uses MegaDescriptor and MiewID image descriptors and dispatches images from each data subset to a dedicated inference branch. Across these branches, SAMGC combines mutual-kNN graph construction, pseudo-label refinement, partition-level co-association, descriptor fusion, and multi-seed ensembling; the discovery-only subset is handled by density-based clustering. The final submission achieved public and private ARI scores of 0.35662 and 0.35682, ranking 37th and 33rd, respectively. Code is publicly available at <https://github.com/cssds/samgc-animalclef2026>.

Keywords

AnimalCLEF, animal re-identification, individual discovery, metric graph clustering, open-set recognition

1. Introduction

Individual animal re-identification aims to link observations of the same animal across time, locations, and imaging conditions. It provides visual evidence for long-term population monitoring, migration and home-range analysis, and individual-level behavioral studies [1]. Unlike species classification, individual re-identification depends on fine-grained appearance cues such as coat texture, spot configurations, scale or scute patterns, scars, and local body markings. These cues may be visible only under particular poses, scales, or body-region coverage. A useful representation must therefore distinguish individuals while producing consistent similarity estimates despite occlusion, illumination changes, background variation, and differences in acquisition devices.

Recent work has moved beyond single-species, closed-set identity classification toward multi-species, open-set, and discovery-oriented settings. Multi-species datasets and toolkits have supported the development of transferable visual representations [2, 3]. Nevertheless, identity evidence differs substantially across taxa: some species are characterized primarily by global texture, whereas others depend on local spots or specific anatomical regions. Real monitoring systems also cannot assume that every test individual is represented in a reference database. They must match confident observations to known individuals while organizing repeated observations of unseen animals into new identity clusters. The objective therefore extends from classifying individual images to partitioning an entire test collection by identity.

The AnimalCLEF 2026 task, part of the LifeCLEF 2026 biodiversity evaluation campaign, focuses on animal re-identification and individual discovery [4, 5]. Participants receive test images and species metadata and must assign a predicted identity cluster label to each image. The task comprises four data subsets: LynxID2025, which draws in part on CzechLynx [6]; SalamanderID2025, as defined in the AnimalCLEF 2026 task overview [5]; SeaTurtleID2022 [7]; and TexasHornedLizards [8]. The first three subsets provide identity-labeled reference images, whereas TexasHornedLizards has no reference database of known identities. Test images may therefore depict either database identities

CLEF 2026 Working Notes

*Corresponding author.

EMAIL: hyticen@gmail.com (B. Liu); hanzhongyuan@gmail.com (Z. Han); 3059421373@qq.com (Z. Liang)

ORCID: 0009-0000-3031-9758 (B. Liu); 0000-0001-8960-9872 (Z. Han); 0009-0008-9499-5710 (Z. Liang)



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

or previously unseen individuals. Submissions are evaluated by comparing the predicted and hidden ground-truth identity partitions using the Adjusted Rand Index (ARI).

The main difficulty is the substantial variation in visual evidence and supervision across the four subsets. Lynx identification relies mainly on global coat texture, salamander identification on local spot configurations, sea turtle identification on head and scute patterns, and Texas horned lizard identification on fine-grained ventral markings. These cues lead to different similarity distributions. Moreover, only the first three subsets can use known identities to adapt a metric space; the TexasHornedLizards branch must rely on the structure of the test set itself. A single feature combination, threshold, or self-training strategy is therefore unlikely to fit every subset. In graph-based partitioning, such mismatch is particularly costly, because a single erroneous edge can propagate through a connected component and merge distinct individuals.

To address this setting, we developed SAMGC, Team Bron’s submission to AnimalCLEF 2026. SAMGC uses MegaDescriptor and MiewID as base image descriptors and routes test images according to species metadata. Each branch then applies metric learning, pseudo-labeling, partition ensembling, or density-based clustering according to the supervision and visual evidence available in its corresponding data subset.

2. Related Work

2.1. Multi-Species Benchmarks and Representation Learning

Animal re-identification research is shifting from species-specific models trained under narrow acquisition conditions toward unified benchmarks and transferable visual representations [1]. WildlifeDatasets provides a common interface for multiple animal re-identification datasets and evaluation protocols [2], while WildlifeReID-10k extends the scale to more than ten thousand individual animals and approximately 140,000 images [3]. MegaDescriptor and MiewID are trained on such multi-species data and provide reusable global identity representations [2, 3, 9]. Species-specific datasets further show that discriminative evidence differs across taxa. SeaTurtleID2022 emphasizes long-term appearance changes, identity annotation, and segmentation [7]; CzechLynx covers individual identification, pose estimation, and segmentation for Eurasian lynx [6]; and Texas horned lizard identification relies on fine-grained ventral spot patterns [8]. These observations motivate the use of transferable representations as a common starting point while adapting the final decision rule to each data subset.

2.2. Fine-Grained Matching and Open-Set Re-Identification

Global descriptors support efficient large-scale retrieval. When identity evidence is concentrated in spots, scales, scars, or other local markings, however, local feature matching, segmentation-aware filtering, and calibrated multi-model scores can provide complementary information. WildFusion combines several recognition signals through calibrated similarity fusion [10]. AnimalCLEF 2025 systems also integrated global and local matching, re-ranking, segmentation-based filtering, and species-specific metric learning [11, 12, 13]. In open-set re-identification, a system must retrieve known identities while rejecting queries for which the reference database provides insufficient evidence. The challenge therefore involves not only closed-set classification accuracy but also similarity calibration and separation between known and unknown samples.

2.3. From Open-Set Retrieval to Individual Discovery

Individual discovery extends open-set recognition into a partitioning problem. Unknown samples must not only be rejected from known identities but also be organized into new identity clusters. A graph can encode images as nodes and pairwise same-identity evidence as edges; connected components or hierarchical clustering can then convert this evidence into predicted individuals. K-reciprocal re-ranking refines distances using reciprocal neighborhood structure [14]. Pseudo-label self-training updates

a representation from high-confidence clusters [15]. Co-association clustering combines multiple partitions according to how frequently pairs of images are assigned to the same cluster [16]. These techniques address neighborhood reliability, unlabeled adaptation, and partition stability, respectively, and form the basis of the subset-specific branches in SAMGC.

3. Method

3.1. Problem Formulation

Let $X = \{x_i\}_{i=1}^N$ denote a set of N test images, where x_i is the i th image and s_i identifies its data subset. The system assigns a cluster label $\hat{y}_i$ to every image. Together, these labels define a predicted identity partition $\hat{\mathcal{P}}$ of the test set. Images x_i and x_j are predicted to depict the same individual if and only if $\hat{y}_i = \hat{y}_j$. The official metric is the ARI [17] between $\hat{\mathcal{P}}$ and the hidden ground-truth identity partition $\mathcal{P}^*$.

The supplied metadata are first used to separate the test images into four disjoint data subsets. For subset s , let $X_s = \{x_i : s_i = s\}$. The corresponding branch clusters images only within X_s and returns a subset-level identity partition $\hat{\mathcal{P}}_s$. Here, an identity partition is the set of clusters corresponding to predicted individuals; it is distinct from the metadata-based division into input subsets. The four subset-level partitions are then assembled into the global prediction $\hat{\mathcal{P}}$ using disjoint cluster-label namespaces. No cross-subset cluster relations are created, which prevents invalid cross-species merges and allows each branch to use a decision rule suited to its visual evidence.

3.2. Representation and Graph Construction

Each image is represented by visual descriptors. We use MegaDescriptor and MiewID as the base representations. In branches with multiple descriptor views, each descriptor is first L2-normalized, concatenated along the feature dimension, and normalized again, thereby avoiding direct addition of raw descriptors with different dimensionalities. In the three database-assisted branches, the concatenated representation is mapped to a 512-dimensional embedding by a subset-specific projection head. The discovery-only Texas horned lizard branch instead uses normalized MiewID descriptors directly. When a branch trains several projection heads from different random seeds, embedding averaging is applied only to heads that share the same architecture, descriptor set, and identity-label space. Heterogeneous evidence, such as local keypoints or segmented regions, can be incorporated at the similarity-score level without requiring the same dimensionality as the global descriptors.

For a database-assisted branch, let $g_m(x_i) \in \mathbb{R}^{d_m}$ denote the raw feature produced for image x_i by descriptor view m . Let $h_{\theta_{s,q}}$ denote the projection head trained for subset s with random seed q . Descriptor fusion and multi-seed ensembling are defined as

$$\begin{aligned} u_i &= \text{norm}([\sqrt{w_1} \bar{g}_1(x_i); \dots; \sqrt{w_M} \bar{g}_M(x_i)]), \\ z_i &= \text{norm}\left(\frac{1}{S} \sum_{q=1}^S h_{\theta_{s,q}}(u_i)\right), \quad \bar{g}_m(x_i) = \frac{g_m(x_i)}{\|g_m(x_i)\|_2}. \end{aligned} \quad (1)$$

Here, M is the number of descriptor views, and the dimension d_m may differ across views. The nonnegative value w_m is the weight of view m , and semicolons inside the brackets denote vector concatenation. We define $\text{norm}(v) = v/\|v\|_2$. The variable S is the number of random seeds, u_i is the input to the projection head, and z_i is the final embedding. A branch with one projection head uses $S = 1$. Because z_i is normalized, the similarity between images x_i and x_j is the cosine similarity $s_{ij} = z_i^\top z_j$.

The graph-based branches convert pairwise similarities into clustering evidence using a mutual-kNN graph. For such a branch, let $G_s = (V_s, E_s)$, where V_s contains the test images in X_s and E_s contains accepted same-identity relations. Let $\mathcal{N}_{k_s}(i)$ be the k_s nearest neighbors of image i within subset

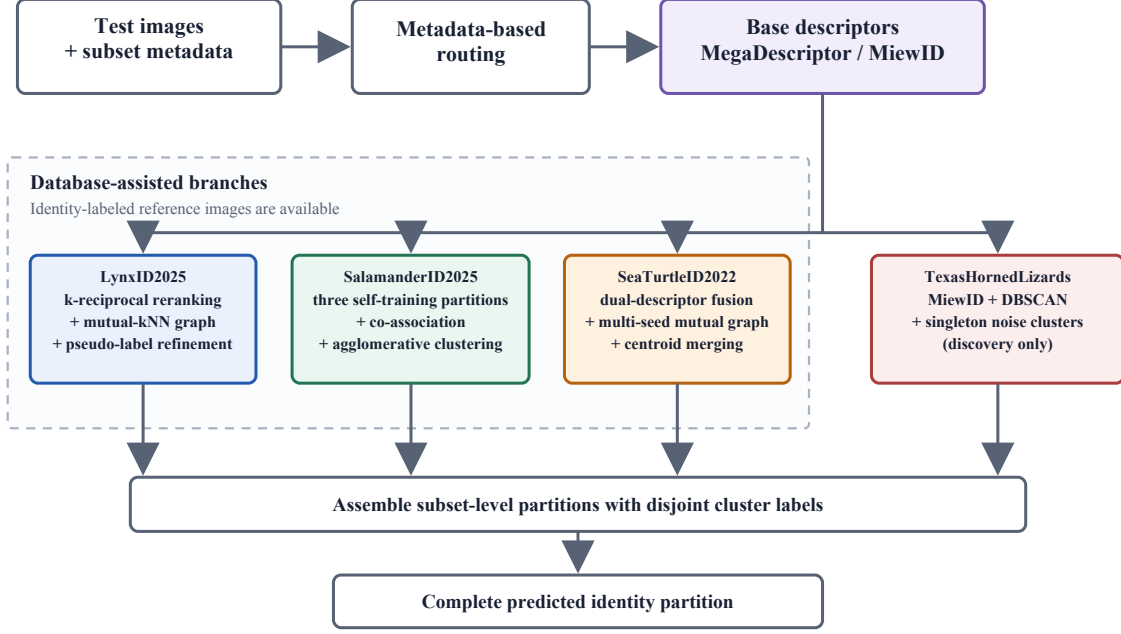


Figure 1: Overview of SAMGC. Test images are routed by metadata, represented with base image descriptors, and processed by subset-specific branches. The Lynx branch combines k-reciprocal re-ranking, mutual-kNN graph clustering, and pseudo-label refinement. The Salamander branch aggregates three self-training partitions through co-association and agglomerative clustering. The SeaTurtle branch performs centroid merging after extracting graph connected components. Without an identity-labeled reference database, the Texas horned lizard branch uses DBSCAN. The four subset-level partitions are assembled with disjoint cluster labels.

s . We define $b_i^{(s)}$ as the similarity to the first neighbor outside this list, namely the local boundary immediately after the top- k_s neighbors. An undirected edge is created only when the two images are mutual neighbors, their similarity exceeds a branch-specific threshold, and the pair is sufficiently separated from both local boundaries:

$$e_{ij}^{(s)} = \mathbf{1}[j \in \mathcal{N}_{k_s}(i)] \mathbf{1}[i \in \mathcal{N}_{k_s}(j)] \mathbf{1}[s_{ij} \geq \tau_s] \mathbf{1}[s_{ij} - b_i^{(s)} \geq \delta_s] \mathbf{1}[s_{ij} - b_j^{(s)} \geq \delta_s]. \quad (2)$$

In Eq. (2), $e_{ij}^{(s)} \in \{0, 1\}$ indicates whether an edge connects images i and j in subset s . The value τ_s is the similarity threshold, δ_s is the local margin threshold, and $\mathbf{1}[\cdot]$ is the indicator function. Setting $\delta_s = 0$ makes the margin conditions inactive for top- k_s neighbors. The connected components of this graph define

$$\mathcal{C}_s = \text{CC}(V_s, E_s), \quad (3)$$

where $\text{CC}(V_s, E_s)$ denotes the set of connected components over the test images in X_s .

The component partition $\mathcal{C}_s$ is not used as the final output in all branches. It is the final partition for Lynx and provides each of the three input partitions for Salamander. For SeaTurtle, it defines the initial clusters before centroid merging. The Texas branch skips graph construction and applies DBSCAN directly. The final subset-level identity partition $\hat{\mathcal{P}}_s$ is therefore determined by its branch-specific clustering operator.

3.3. Subset-Specific Branches

Branch selection in SAMGC is motivated by two factors. The first is the availability of identity-labeled reference images, which determines whether identity supervision can be used to adapt the metric

representation. The second is the stability of pairwise evidence among test images, which determines whether one metric graph is sufficient or whether several representations or self-training partitions should be combined.

Lynx. Lynx coat texture usually provides reliable global similarity cues, but similar poses and partial occlusion can still produce high-scoring false neighbors. In connected-component clustering, one erroneous edge may merge two distinct individuals. The Lynx branch therefore first applies k-reciprocal re-ranking to obtain a refined similarity matrix and then builds a mutual-kNN graph for a conservative initial identity partition. In the initial Lynx graph, s_{ij} in Eq. (2) denotes re-ranked similarity; in the subsequent graph, it denotes cosine similarity in the projected embedding space. Stable predicted clusters are treated as temporary identity classes. Their images are used together with identity-labeled database images to train a metric projection head. The mutual-kNN graph is then rebuilt in the adapted embedding space, and its connected components define the predicted individuals. This procedure uses both known-identity supervision and high-confidence pseudo-clusters while limiting the propagation of erroneous edges and unreliable pseudo-labels.

Salamander. Salamander identity cues are concentrated in local spot patterns whose visibility changes with pose, occlusion, and imaging conditions. A partition from one self-training run can consequently be sensitive to the pseudo-cluster selection rule. Rather than choosing one partition, the Salamander branch combines self-training partitions produced with different pseudo-cluster filtering strengths. Each partition is obtained by one round of pseudo-label training followed by re-clustering: selected pseudo-clusters and labeled reference identities are used to train a projection head, after which the test images are clustered again in the adapted metric space. To retain same-cluster relations that recur across these partitions, we compute the pairwise co-association score

$$a_{ij} = \frac{1}{P} \sum_{p=1}^P \mathbf{1}\{c_i^{(p)} = c_j^{(p)}\}. \quad (4)$$

Here, i and j index two Salamander test images, P is the number of input partitions, and p is the partition index. The value $c_i^{(p)}$ is the cluster label assigned to image i in partition p , and $\mathbf{1}\{c_i^{(p)} = c_j^{(p)}\}$ indicates whether the two images belong to the same cluster in that partition. In the final system, each pseudo-cluster filtering setting produces one partition after embeddings from different seeds are averaged, giving $P = 3$ partitions for co-association. The score $a_{ij} \in [0, 1]$ is the fraction of partitions in which the image pair is clustered together. The branch converts this score into a pairwise distance and applies agglomerative clustering, reducing the influence of split or merge errors from any single partition.

SeaTurtle. The visibility of sea turtle head and scute patterns varies with viewpoint, scale, and the local region captured in an image, making graph edges from a single representation unstable. Random initialization of the projection head can also change the connectivity of image pairs near a decision boundary. The SeaTurtle branch combines complementary evidence from MegaDescriptor and MiewID. Their normalized descriptors are concatenated and projected into a shared 512-dimensional metric space. Outputs from different random seeds are then averaged as defined in Eq. (1). A mutual-kNN graph is built from the fused similarities, connected components are extracted, and component centroids are merged conservatively. This design avoids adding raw descriptors of different dimensionalities and reduces dependence on a single representation or training run.

Texas horned lizard. This subset has no identity-labeled reference images. The metric space therefore cannot be calibrated with known identities, and pseudo-label training lacks reliable identity anchors. The Texas horned lizard branch instead applies conservative DBSCAN [18] directly to the MiewID representations of the test images. Samples labeled as noise are converted into singleton clusters rather than being forced into larger clusters without sufficient supporting evidence.

Table 1 summarizes the subset-specific branch designs in SAMGC.

Table 1
Subset-specific branch design in SAMGC.

Subset	Primary challenge	Branch design	Rationale
LynxID2025	Asymmetric nearest-neighbor matches can cause transitive merges	Mutual-kNN graph; pseudo-label refinement	Retain mutually supported edges and limit pseudo-label propagation
SalamanderID2025	Variable spot visibility makes single-run partitions unstable	Multiple self-training partitions; co-association clustering	Retain same-cluster evidence recurring across partitions
SeaTurtleID2022	Graph edges vary across descriptors and random seeds	Dual-descriptor fusion; multi-seed graph; centroid merging	Combine complementary cues, reduce variation across training runs, and conservatively merge component-level evidence
TexasHornedLizards	No labeled identity anchors; high risk from pseudo-labeling	Conservative DBSCAN; singleton noise clusters	Avoid forcing weakly supported samples into larger clusters

4. Experimental Setup and Results

4.1. Data and Evaluation

We used the official AnimalCLEF 2026 data, whose test metadata assigns each image to one of four data subsets. LynxID2025 concerns Eurasian lynx identification and draws on sources including CzechLynx [6]. The organization of SalamanderID2025 for this task is specified in the AnimalCLEF 2026 overview [5]. SeaTurtleID2022 is a long-term sea turtle re-identification dataset [7], while TexasHornedLizards concerns identification from ventral spot patterns [8]. The first three subsets provide identity-labeled database images that can be used to train or calibrate metric representations. We treat TexasHornedLizards as a discovery-only subset because it provides no labeled database identities as anchors for supervised adaptation.

The system assigns one predicted identity cluster label to each test image, and images with the same label are interpreted as the same individual. Predictions from the four data subsets are assembled into one submission file. Subset prefixes are added to cluster labels to prevent collisions between species. Each submission covers the complete test set, while the public and private ARI scores are computed by the organizers on different hidden evaluation subsets.

4.2. Implementation Details

The three subsets with identity databases use the same projection-head architecture, but their parameters are trained separately for each subset. The projection module is a two-layer MLP with dimensions $d \rightarrow 1024 \rightarrow 512$. Its hidden layer uses batch normalization, ReLU, and dropout with a rate of 0.2, and the output is L2-normalized. Training uses an ArcFace classification head [19] with a scale of 24.0 and an angular margin of 0.3. We optimize the model with AdamW [20], an initial learning rate of 10^{-3} , weight decay of 10^{-4} , and a batch size of 256. Graph parameters are selected on identity-disjoint validation splits. Approximately 20% of the database identities are held out in each split, corresponding to 20–120 held-out identities depending on the subset. Evaluation is repeated across multiple random seeds.

For LynxID2025, the input combines horizontal-flip-averaged MiewID features with MegaDescriptor features. Initial k-reciprocal re-ranking uses $k_1 = 10$, $k_2 = 1$, and $\lambda = 0.5$, followed by the mutual-kNN graph rule in Eq. (2). Only predicted clusters containing at least four images are retained as

pseudo-identities. For a retained pseudo-cluster C , the weight assigned to its pseudo-labeled samples is

$$w(C) = \min\left(1.5, 0.75\sqrt{\frac{|C|}{4}}\right), \quad (5)$$

where $|C|$ is the number of images in the pseudo-cluster. Projection heads are trained for 80 epochs with random seeds 42, 2024, and 3407, and their normalized outputs are averaged. The final mutual-kNN graph uses $k = 12$, a similarity threshold of 0.45, and a margin of 0. No additional centroid-based cluster merging is applied.

For SalamanderID2025, the input comprises a multi-crop-averaged MiewID view and a horizontal-flip-averaged MiewID view, weighted by 1.0 and 0.25, respectively. The three self-training partitions use minimum pseudo-cluster sizes of 2, 3, and 4. For each setting, projection heads are trained for 90 epochs with seeds 42, 2024, and 3407. Their embeddings are averaged before graph clustering, so each setting contributes one partition to the co-association matrix. The graph parameters for the three input partitions are $(k, \tau, \delta) = (5, 0.36, 0.13)$, $(5, 0.36, 0.125)$, and $(5, 0.34, 0.11)$. The final co-association distance is $d_{ij} = 1 - a_{ij}$. Average-linkage agglomerative clustering is applied to this precomputed distance matrix with `n_clusters=558`. This cluster count was selected with the same identity-disjoint validation protocol used for the graph parameters.

For SeaTurtleID2022, MegaDescriptor and MiewID descriptors are normalized with equal weights and concatenated. Projection heads are trained for 120 epochs with seeds 42, 2024, 3407, 777, and 1234. The five normalized 512-dimensional outputs are averaged and L2-normalized again. The final graph uses $k = 24$, a similarity threshold of 0.52, and a margin of 0.01. After connected components are extracted, the centroid of each component is computed as the mean of its member embeddings and then L2-normalized. Components with a centroid cosine similarity of at least 0.60 are merged in one union-find pass over the centroid-similarity graph.

For TexasHornedLizards, DBSCAN is applied to a precomputed cosine-distance matrix of MiewID descriptors with $\epsilon = 0.24$ and `min_samples=2`. Every noise sample is converted into a singleton cluster.

4.3. Results and Analysis

SAMGC achieved public and private ARI scores of 0.35662 and 0.35682, ranking 37th and 33rd, respectively. The close public and private scores suggest consistent behavior across the two hidden evaluation subsets, with no substantial public-to-private performance shift.

To analyze the effect of major design choices under the same task setting, we report several representative internal submissions from Team Bron. All rows in Table 2 correspond to methods developed and submitted by our team, and all reported ARI values are public-leaderboard scores returned for those submissions.

The results show the limitations of applying a single global clustering strategy to this task. Adaptive-density clustering obtained an ARI of 0.10274, indicating that local density structure alone did not adequately balance false splits and false merges. Unified metric-graph clustering increased the score to 0.20841. Applying the same supervised metric calibration and graph-construction rule to every data subset, however, remained insufficient.

Methods targeted at difficult individual subsets performed better. The Salamander pairwise GNN reached an ARI of 0.26771, while the SeaTurtle dual-descriptor multi-seed ensemble reached 0.31141. These results suggest that pairwise relation modeling and model ensembling are useful when applied to subsets whose visual evidence matches their assumptions. SAMGC obtained the highest ARI among the compared submissions, at 0.35662. This result supports organizing evidence at the subset level: Lynx uses mutual-neighbor constraints, Salamander combines several self-training partitions, SeaTurtle integrates two descriptors and multiple random seeds before centroid merging, and Texas horned lizard uses density-based clustering in the discovery-only setting.

Table 3 presents the progression of the main technical routes. The initial feature-fusion system obtained an ARI of 0.19190. With the other subsets fixed, replacing the SeaTurtle branch increased

Table 2

Public-leaderboard ARI of representative Team Bron methods.

Method	Main mechanism	ARI
Adaptive-density clustering	Local-density thresholds with hierarchical stability clustering	0.10274
Unified metric-graph clustering	A single metric-projection and mutual-kNN graph-clustering rule applied across all data subsets	0.20841
Salamander pairwise GNN	Pairwise relation modeling with GNN-based clustering	0.26771
SeaTurtle dual-descriptor ensemble	Dual-descriptor fusion with multi-seed averaging and centroid merging	0.31141
SAMGC (ours)	Mutual-kNN filtering, pseudo-label refinement, co-association clustering, descriptor fusion, and conservative DBSCAN	0.35662

Table 3

Progression of representative Team Bron submissions on the public leaderboard.

Technical route	Main change	ARI	Clusters
Initial feature-fusion system	MegaDescriptor and MiewlD representations with basic connected-component clustering	0.19190	1604
SeaTurtle metric-graph branch	SeaTurtle branch replaced; other subsets fixed	0.24384	1580
Salamander metric-graph branch	Salamander branch replaced in the preceding system	0.28115	1501
SeaTurtle dual-descriptor multi-seed ensemble	Dual-descriptor projection, five random seeds, and centroid merging	0.31141	1495
Salamander pseudo-label partition ensemble	Aggregation of self-training partitions with a fixed 558-cluster cut	0.33085	1496
Lynx reciprocal-neighborhood re-ranking	Lynx replaced by a reciprocal-neighborhood re-ranking branch	0.35366	1398
SAMGC (ours)	Conservative pseudo-label refinement added to the Lynx branch	0.35662	1399

the score to 0.24384, and subsequently replacing the Salamander branch raised it to 0.28115. The SeaTurtle dual-descriptor multi-seed ensemble and the Salamander pseudo-label partition ensemble then increased the public ARI to 0.31141 and 0.33085, respectively. Reciprocal-neighborhood re-ranking for Lynx produced the largest subsequent gain, and adding pseudo-label refinement to this branch yielded the final SAMGC submission.

This progression indicates that the gains arose from different sources across the data subsets. SeaTurtle benefited primarily from descriptor and multi-seed ensembling, Salamander from partition-level co-association, and Lynx from reduced over-splitting followed by conservative pseudo-label adaptation.

5. Conclusion

This paper presented SAMGC, the species-aware metric graph clustering method submitted by Team Bron to AnimalCLEF 2026. SAMGC uses a shared representation-and-clustering framework with subset-specific inference branches. Each branch organizes pairwise or partition-level evidence according to

the supervision and visual cues available in its data subset. The final system generated predictions for all test images and achieved public and private ARI scores of 0.35662 and 0.35682, ranking 37th and 33rd, respectively.

Comparisons with global density clustering, unified metric-graph clustering, and single-subset enhancements support the use of subset-specific metric and clustering evidence under heterogeneous supervision. Future work will examine more systematic subset-specific validation protocols and incorporate local keypoints, segmented regions, and anatomical cues as score-level evidence within SAMGC.

Code and Model Availability

The implementation, configuration files, and submission-generation scripts for SAMGC are publicly available at <https://github.com/cssds/samgc-animalclef2026>. We do not redistribute the AnimalCLEF challenge images, hidden labels, cached features, intermediate experiment outputs, or trained model checkpoints. External pretrained vision models must be obtained under the license terms of their original providers. The lightweight SAMGC projection heads can be retrained from the released scripts and configurations.

Acknowledgments

This work is supported by the Guangdong Province Basic and Applied Basic Research Fund (Guangdong-Foshan Joint Funds, 2024A1515140026).

Generative AI Use Statement

During the preparation of this manuscript, the authors used ChatGPT for grammar checking, language editing, draft translation, and organizing revisions. The tool was not used to generate experimental results, hidden labels, or leaderboard scores. The authors reviewed and edited all AI-assisted content and take full responsibility for the published version.

References

- [1] Stefan Schneider, Graham W. Taylor, Stefan S. Linquist, and Stefan C. Kremer. Past, Present and Future Approaches Using Computer Vision for Animal Re-Identification from Camera Trap Data. *Methods in Ecology and Evolution*, 10(4):461–470, 2019. doi:10.1111/2041-210X.13133.
- [2] Vojtěch Čermák, Lukas Pícek, Lukáš Adam, and Kostas Papafitsoros. WildlifeDatasets: An Open-Source Toolkit for Animal Re-Identification. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 5953–5963, 2024.
- [3] Lukáš Adam, Vojtěch Čermák, Kostas Papafitsoros, and Lukas Pícek. WildlifeReID-10k: Wildlife Re-Identification Dataset with 10k Individual Animals. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2025.
- [4] Lukas Pícek, Lukáš Adam, Stefan Kahl, Robert Bossy, Laura Chrobak, Hervé Goëau, Kostas Papafitsoros, Holger Klinck, Willem-Pier Vellinga, Robert Planqué, Tom Denton, Kevin Barnard, Claire Nédellec, Louise Deléger, Marine Courtin, Giulio Martellucci, Ilyass Moummad, Fabrice Vinatier, Pierre Bonnet, and Alexis 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.
- [5] Lukáš Adam, Kostas Papafitsoros, Dean A. Williams, Daniella Biffi, and Lukas Pícek. 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.

- [6] Lukas Pícek, Jakub Straka, Miroslav Jířík, Elisa Belotti, Martin Duřa, Josefa Krausová, Michal Bojda, Vojtěch Čermák, Luděk Bufka, Rostislav Dvořák, Luboslav Hrdý, Václav Kocourek, Jiří Labuda, Luděk Toman, Vlado Trulík, and Martin Váňa. CzechLynx: A Dataset for Individual Identification and Pose Estimation of the Eurasian Lynx. *Scientific Data*, 13:511, 2026. doi:10.1038/s41597-026-06853-9.
- [7] Lukáš Adam, Vojtěch Čermák, Kostas Papafitsoros, and Lukas Pícek. SeaTurtleID2022: A Long-Span Dataset for Reliable Sea Turtle Re-Identification. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 7146–7156, 2024.
- [8] Daniella Biffi, Mary R. Tucker, Alexis Ackel, and Dean A. Williams. Identification of Individual Texas Horned Lizards (*Phrynosoma cornutum*) Using Genotypes and Ventral Spot Patterns. *Ecology and Evolution*, 15(3):e71167, 2025. doi:10.1002/ece3.71167.
- [9] Lasha Otarashvili, Tamilselvan Subramanian, Jason Holmberg, J. J. Levenson, and Charles V. Stewart. Multispecies Animal Re-ID Using a Large Community-Curated Dataset. *arXiv preprint arXiv:2412.05602*, 2024.
- [10] Vojtěch Čermák, Lukas Pícek, Lukáš Adam, Lukas Neumann, and Jiří Matas. WildFusion: Individual Animal Identification with Calibrated Similarity Fusion. In *Computer Vision – ECCV 2024 Workshops*, Lecture Notes in Computer Science, pages 18–36, Springer, 2024. doi:10.1007/978-3-031-92387-6_2.
- [11] Lukáš Adam, Kostas Papafitsoros, Roman Kovář, Vojtěch Čermák, and Lukas Pícek. Overview of AnimalCLEF 2025: Recognizing Individual Animals in Images. In *Working Notes of CLEF 2025 - Conference and Labs of the Evaluation Forum*, CEUR Workshop Proceedings, volume 4038, pages 2898–2908, CEUR-WS.org, 2025.
- [12] Roman Pakhomov, Grigory Demidov, Kristian Bogdan, Svyatoslav Lanskiikh, Danis Dinmuhametov, and Andrey Khlopotnykh. Individual Wildlife Recognition via Hybrid Global-Local Matching and Segmentation-Aware Filtering. In *Working Notes of CLEF 2025 - Conference and Labs of the Evaluation Forum*, CEUR Workshop Proceedings, volume 4038, pages 3139–3146, CEUR-WS.org, 2025.
- [13] Nelly Semenova. Meta-Algorithm for Open-Set Animal Re-ID: WildFusion, XGBoost, and Dual-Backbone ArcFace. In *Working Notes of CLEF 2025 - Conference and Labs of the Evaluation Forum*, CEUR Workshop Proceedings, volume 4038, pages 3154–3165, CEUR-WS.org, 2025.
- [14] Zhun Zhong, Liang Zheng, Donglin Cao, and Shaozi Li. Re-ranking Person Re-identification with k-Reciprocal Encoding. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3652–3661, 2017. doi:10.1109/CVPR.2017.389.
- [15] Dong-Hyun Lee. Pseudo-Label: The Simple and Efficient Semi-Supervised Learning Method for Deep Neural Networks. In *ICML 2013 Workshop on Challenges in Representation Learning*, 2013.
- [16] Ana L. N. Fred and Anil K. Jain. Combining Multiple Clusterings Using Evidence Accumulation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(6):835–850, 2005. doi:10.1109/TPAMI.2005.113.
- [17] Lawrence Hubert and Phipps Arabie. Comparing Partitions. *Journal of Classification*, 2(1):193–218, 1985. doi:10.1007/BF01908075.
- [18] Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96)*, pages 226–231, AAAI Press, 1996.
- [19] Jiankang Deng, Jia Guo, Niannan Xue, and Stefanos Zafeiriou. ArcFace: Additive Angular Margin Loss for Deep Face Recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4690–4699, 2019. doi:10.1109/CVPR.2019.00482.
- [20] Ilya Loshchilov and Frank Hutter. Decoupled Weight Decay Regularization. In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2019.