

A Review of Computer Vision Approaches for Non-Invasive Animal Re-Identification in Large-Scale Camera Networks

¹Kingsley Chinazaekpere Ndupu, ²Serif Oyindamola Oyesiji, ³Stanley Nwakamma, ⁴Harouna Wendpanga Yann Christian Sankara.

¹Kennesaw State University, Georgia, USA; kingsleynd34@gmail.com

²Harrisburg University of Science and Technology, Harrisburg, PA; oyesijiserif@gmail.com

³Independent Researcher, USA; nwakammastanley@gmail.com

⁴Independent Researcher, Dallas, TX, USA; christiansankara@gmail.com

Corresponding Author; kingsleynd34@gmail.com

Abstract: The ability to recognize individual animals repeatedly and reliably underpin a wide range of scientific and operational activities, from population ecology and conservation management to precision livestock farming and biosecurity. Traditional approaches to individual identification, including physical tagging, branding, collaring, and manual photographic matching, are invasive, labor intensive, or both, and they scale poorly to the camera networks now deployed across protected areas, rangelands, and farms. This review synthesizes the computer vision literature on non-invasive animal re-identification, with particular attention to the pipeline that has become dominant in large-scale deployments: convolutional object detection, most prominently the YOLO family of single-stage detectors, followed by deep feature extraction into high-dimensional embedding spaces in which individual identity is resolved by similarity search. The review traces the development of the field from early pattern-matching systems built on hand-crafted features through the adoption of deep metric learning, examines species-specific applications spanning big cats, primates, cetaceans, pinnipeds, elephants, bears, and domestic cattle and pigs, and analyzes the systems-level challenges introduced by distributed multi-camera networks, including open-set recognition, cross-camera appearance variation, embedding indexing at scale, and edge deployment under constrained power and bandwidth budgets. Persistent obstacles are identified in dataset scarcity and imbalance, domain shift between capture locations, the absence of standardized evaluation protocols, and the integration of algorithmic outputs with statistical population models. The review closes with directions for future research, emphasizing self-supervised representation learning, cross-species transfer, and uncertainty-aware identification suitable for downstream ecological inference.

1. Introduction

Individual identity is the atomic unit of much of animal science. Estimates of abundance, survival, movement, social structure, disease transmission, and behavioral repertoire all depend, in one form or another, on the capacity to say that the animal observed today is, or is not, the animal observed last week at a different place (Kühl & Burghardt, 2013). For more than a century that capacity was purchased at the cost of intervention: animals were trapped, marked, tagged, banded, branded, or collared so that their identity could be read on re-encounter. The costs of such intervention are now well documented, encompassing capture stress and injury, altered behavior, tag loss that biases demographic estimates, and the sheer expense of fieldwork at scale. Non-invasive alternatives, chief among them the photographic identification of naturally occurring marks, have therefore held enduring appeal, and camera-based monitoring has grown into one of the standard instruments of wildlife science (Burton et al., 2015).

Two developments transformed the scale of the problem over the past decade. The first is the proliferation of the sensors themselves. Camera traps have become inexpensive, robust, and ubiquitous, and coordinated deployments now operate at regional, national, and continental scales, producing image volumes measured in the tens of millions of frames per project (Steenweg et al., 2017). Aerial platforms have widened the same trend, with unmanned aerial vehicles moving from experimental trials to routine survey of extended corridors and remote terrain, and with the operational literature on their deployment maturing alongside (Dagodzo, 2018b; Dagodzo & Ahiaekwe Patrick, 2020). Precision livestock operations have followed a parallel trajectory, instrumenting barns, milking parlors, and pastures with fixed and mobile cameras for continuous monitoring of health and welfare (Neethirajan, 2020). The measurement burden this instrumentation is meant to relieve is visible in the animal science literature, where growth, digestibility, and physiological response are recorded animal by animal across feeding trials and management systems (Aye & Tawose, 2016; Oluwadele, Tawose, & Adetumbi, 2024). The second development is the maturation of deep learning as a general-purpose visual recognition technology (LeCun et al., 2015). Convolutional neural networks displaced hand-engineered features across nearly every benchmark in computer vision, and the techniques developed for human-centric tasks, notably face verification and person re-identification, transferred with striking effectiveness to the animal domain (Schneider et al., 2019).

The convergence of these developments defines the subject of this review: automated, non-invasive re-identification of individual animals from images collected by large networks of cameras. The canonical modern pipeline decomposes the problem into stages. An object detector, very often a member of the YOLO family chosen for its favorable speed and accuracy trade-off,

localizes animals within cluttered natural scenes. Detected regions are cropped, optionally aligned or segmented, and passed to a feature extraction network trained with metric learning objectives to produce a compact high-dimensional embedding. Identity is then resolved not by classification over a fixed label set but by measuring distances between embeddings, so that new, previously uncatalogued individuals can be accommodated without retraining. Finally, systems deployed across many cameras must associate identities over space and time, maintain searchable galleries of known individuals, and often perform some or all of this computation on constrained hardware at the network edge.

Interest in this pipeline extends well beyond ecology. Conservation organizations use automated re-identification to maintain longitudinal registries of endangered populations and to support anti-poaching enforcement (Berger-Wolf et al., 2017). Livestock producers pursue the same capability for traceability, disease surveillance, and individualized management, where the replacement of ear tags with camera-based identification promises welfare and labor benefits (Awad, 2016). Operational safety domains draw on the same detection substrate, with data-driven wildlife hazard monitoring and incident reduction at airports an active applied line built on automated recognition of animal presence (Adeyelu & Dagodzo, 2024). Such systems are consumed by regulatory processes as much as by operations, a pattern already documented for evidence frameworks governing protected airport surfaces (Adeyelu, 2022). Methodologically, animal re-identification has also become a productive testbed for open problems in machine learning, because it combines fine-grained discrimination, severe data imbalance, open-set conditions, and distribution shift in a single application (Tuia et al., 2022).

Despite rapid growth, the literature remains fragmented across ecology, agricultural engineering, and computer science venues, with limited cross-citation and little standardization of tasks or metrics. The purpose of this review is to synthesize that literature into a coherent account of where the field stands. The remainder of the paper is organized as follows. Section 2 states the scope and approach of the review. Section 3 traces the historical development from manual photographic matching to learned recognition. Sections 4 and 5 examine the two core algorithmic stages, object detection and embedding-based identification. Sections 6 and 7 survey applications in wildlife and livestock respectively. Sections 8 through 11 address the systems dimension: multi-camera association, similarity search at scale, datasets and evaluation, and edge deployment. Section 12 considers integration with population modeling, Sections 13 and 14 discuss open challenges and future directions, and Section 15 concludes.

2. Scope and Approach of the Review

This review is concerned with visual, image-based re-identification of individual animals, and it treats the topic from the perspective of systems deployed across multiple cameras rather than single-sensor laboratory studies. Re-identification is understood in the sense inherited from the person re-identification literature: given a query image of an animal, retrieve all prior observations of the same individual from a gallery accumulated across cameras, times, and places (Ye et al., 2022). This formulation subsumes both closed-set identification, in which every query belongs to a known catalogue, and the more realistic open-set variant in which new individuals continually enter the population. Acoustic identification, genetic identification, and radio-frequency tagging fall outside the scope, as do studies concerned solely with species-level classification, although species classification is discussed where it functions as a stage within an individual-identification pipeline.

The literature base was assembled from the principal venues in which the field publishes: ecological methods journals, agricultural engineering journals, and the computer vision conference ecosystem including its workshop series, where much of the applied animal biometrics work first appears (Ravoor & Sudarshan, 2020). Survey and perspective articles at the intersection of biology and computer vision were used to seed backward and forward citation searches (Vidal et al., 2021; Weinstein, 2018). Emphasis was placed on work published from 2013 onward, the period in which deep learning entered the field, while foundational earlier work on photographic matching is retained where it establishes concepts still in use. Methodological papers from the broader vision literature are cited where the animal re-identification community has adopted them directly, since the field is to a large degree an importer of techniques developed for other domains (Christin et al., 2019). The same importing dynamic is visible in adjacent monitoring domains, where detection and classification architectures built for generic scene understanding are retargeted at vegetation, corridor, and asset inspection tasks (Dagodzo, Ahiaeké Patrick, & Aliliele, 2022).

The synthesis is organized around a reference architecture rather than around individual species or projects, because the striking feature of the past decade is precisely the convergence of previously disparate species-specific efforts onto a shared pipeline of detection, embedding, and retrieval. Species-specific material is therefore presented as a set of instantiations of that architecture, which allows the review to identify which design choices are genuinely species-dependent and which reflect accidents of disciplinary history. Throughout, attention is paid to the constraints that distinguish large-scale camera networks from curated datasets: unconstrained pose and illumination, partial occlusion, low and variable image quality, extreme class imbalance across individuals, and the operational need for real-time or near-real-time processing on limited hardware.

3. From Species Recognition to Individual Re-Identification

Photographic identification of individuals predates computer vision by decades, resting on the observation that many taxa carry stable, individually distinctive markings: pelage patterns in felids and zebras, fluke pigmentation and trailing-edge nicks in

whales, whisker spot constellations in lions and seals, and scarring in elephants and manta rays. Early computational assistance formalized what field biologists had done by eye. Kelly (2001) evaluated computer-aided matching of cheetah photographs in the Serengeti, demonstrating both the promise and the sensitivity of semi-automated matching to image quality. Arzoumanian et al. (2005) adapted an astronomical point-pattern matching algorithm to the spot constellations of whale sharks, an influential early demonstration that highly specialized algorithms from unrelated fields could solve individual identification. Hiby et al. (2009) fitted three-dimensional surface models to tiger flank images so that stripe patterns photographed at different camera angles could be compared, and showed that even images of confiscated skins could be matched to living animals, a capability with direct enforcement value.

General-purpose tools followed. Bolger et al. (2012) released a computer-assisted mark-recapture system built on SIFT keypoints, and Town et al. (2013) applied related keypoint methods to manta ray ventral patterns. The most widely adopted of these systems, HotSpotter, matched patterned species through viewpoint-invariant keypoint descriptors and demonstrated accurate instance recognition across zebras, giraffes, leopards, and lionfish from a single algorithmic core (Crall et al., 2013). These systems shared a common structure: hand-crafted local features in the tradition of SIFT, SURF, HOG, and related descriptors (Lowe, 2004; Dalal & Triggs, 2005; Bay et al., 2008), geometric verification, and a human in the loop to confirm proposed matches. The underlying principle, scoring a query region against stored exemplars, is the same one that supports template matching in medical image screening (Ejofodomi et al., 2014). They achieved real scientific impact, populating registries that supported capture-recapture analyses, but their reliance on visible, high-contrast patterning limited taxonomic reach, and their per-comparison cost limited gallery size.

In parallel, the ecological community was industrializing image collection without yet industrializing analysis. Camera trapping matured from a specialist technique into a standard survey method, with foundational work establishing its use for density estimation of individually identifiable species such as tigers, jaguars, and snow leopards (Karanth, 1995; Karanth & Nichols, 1998; Silver et al., 2004; Jackson et al., 2006). Methodological syntheses codified survey design and flagged the analytical hazards of misidentification (O'Connell et al., 2011; Foster & Harmsen, 2012), while spatial capture-recapture models raised the statistical value of knowing not merely that an individual was detected but where and when (Royle et al., 2013). Large collaborative deployments such as Snapshot Serengeti generated annotated archives of a scale that manual processing could not sustain (Swanson et al., 2015), creating the backlog that automated methods would later address.

Automation arrived first at the species level. Early classifiers using hand-crafted features achieved modest accuracy on camera trap imagery (Yu et al., 2013), and the first convolutional approaches improved on them substantially (Chen et al., 2014; Gomez Villa et al., 2017). The decisive demonstration came from Norouzzadeh et al. (2018), who trained deep networks on millions of Snapshot Serengeti images to identify, count, and describe animals at accuracy comparable to crowdsourced human annotation, and estimated savings of tens of thousands of hours of labeling effort. Subsequent work extended species classification across projects and continents, quantified its fragility under transfer to new sites, and embedded it in active learning loops to reduce annotation cost (Tabak et al., 2019; Willi et al., 2019; Beery et al., 2018; Norouzzadeh et al., 2021). Generic detection models trained across many projects, most prominently the MegaDetector line, decoupled animal localization from project-specific classification and became de facto infrastructure for the camera trap community (Beery et al., 2019).

Individual-level recognition followed the same trajectory a few years behind. Reviews at the transition point framed the opportunity explicitly: hand-crafted pipelines had proven the concept, deep feature learning promised species-agnostic generality, and the person re-identification literature offered a ready-made experimental paradigm (Schneider et al., 2019). Within a few years, deep metric learning systems had been demonstrated across a taxonomically diverse set of species, in several cases exceeding the matching performance achievable by trained human observers (Schneider et al., 2020), and the field's center of gravity shifted from bespoke pattern matchers to the detection-plus-embedding architecture that the remainder of this review examines.

4. Object Detection Foundations and the YOLO Family

Reliable localization is the entry point of every large-scale pipeline, because camera network imagery is dominated by empty frames, partial views, and scenes in which animals occupy a small fraction of the field of view. The deep detection literature that the animal domain draws on developed rapidly after convolutional networks demonstrated decisive advantages in image classification (Krizhevsky et al., 2012; Simonyan & Zisserman, 2015; Szegedy et al., 2015; He et al., 2016), supported by large annotated corpora such as ImageNet, PASCAL VOC, and Microsoft COCO that supplied both pre-training material and evaluation protocols (Deng et al., 2009; Everingham et al., 2010; Lin et al., 2014). Two-stage detectors formalized detection as region proposal followed by classification, progressing from R-CNN through Fast R-CNN to the fully learned proposal mechanism of Faster R-CNN (Girshick et al., 2014; Girshick, 2015; Ren et al., 2015), with feature pyramid networks extending accuracy across object scales (Lin et al., 2017a).

Single-stage detectors traded a measure of accuracy for dramatic gains in speed by predicting boxes and classes directly from dense feature maps. The original YOLO formulation cast detection as a single regression problem evaluated in one forward pass (Redmon et al., 2016), and successive revisions improved anchor design, multi-scale prediction, and backbone capacity through YOLO9000 and YOLOv3 (Redmon & Farhadi, 2017, 2018). SSD offered a contemporaneous single-stage alternative (Liu et al.,

2016), and focal loss addressed the foreground-background imbalance that had limited single-stage accuracy (Lin et al., 2017b). Later community-driven iterations, including YOLOv4 with its systematic study of training refinements, the anchor-free YOLOX, and YOLOv7, continued to advance the speed-accuracy frontier that makes the family attractive for high-throughput and embedded use (Bochkovskiy et al., 2020; Ge et al., 2021; Wang et al., 2023). Transformer-based designs subsequently reframed detection as set prediction and imported attention-based backbones (Carion et al., 2020; Dosovitskiy et al., 2021; Liu et al., 2021), while compound-scaled convolutional families provided strong accuracy at controlled computational budgets (Tan & Le, 2019; Tan et al., 2020).

The animal re-identification literature has adopted these detectors essentially wholesale, with the YOLO family the most common choice in deployed systems for three reasons. First, throughput: a network of hundreds of cameras generating motion-triggered bursts produces frame volumes for which two-stage inference budgets are uncomfortable, whereas single-stage detectors sustain real-time rates on modest GPUs and, in compressed form, on edge accelerators. Second, robustness under transfer: detection is a coarser task than identification, and detectors trained across heterogeneous animal imagery generalize to new sites far better than classifiers do, a property exploited by generic animal detectors that serve as shared front ends across projects (Beery et al., 2019). Third, engineering maturity: the family's tooling for training, augmentation, quantization, and export to embedded runtimes lowers the barrier for research groups outside computer science, which describes much of the ecological user base (Christin et al., 2019). Comparable adoption patterns appear in aerial inspection pipelines, where single-stage detectors are run over unmanned aerial vehicle imagery to flag encroachment and vegetation contact along extended rights of way (Dagodzo & Ahiaeké Patrick, 2022; Dagodzo et al., 2022).

Detection in natural scenes nonetheless presents failure modes that curated benchmarks underrepresent. Camera trap animals are frequently truncated at frame edges, motion-blurred by fast triggers at night, camouflaged against vegetation, or present only as a partial flank filling the frame at close range. Empty-frame suppression matters as much as localization accuracy, because false triggers from vegetation movement can constitute the majority of captures. Studies of cross-location generalization show that detectors, like classifiers, degrade when confronted with backgrounds, species, and camera placements absent from training data, motivating training corpora assembled deliberately across many projects and locations (Beery et al., 2018). Analogous degradation under changing background, altitude, and sensor configuration is reported in corridor inspection deployments that combine aerial imagery with laser scanning and geographic information systems (Dagodzo & Ahiaeké Patrick, 2021b). Within re-identification pipelines specifically, the detector also carries responsibilities beyond the bounding box: pose and viewpoint filtering, part localization such as heads, flanks, fins, or flukes, and quality gating that discards crops too degraded to embed reliably, functions demonstrated in integrated detection pipelines built for identification front ends (Parham et al., 2018).

The practical consequence for system designers is that detector choice is rarely the binding constraint on identification accuracy, but detector configuration often is. The recall of the detection stage bounds the recall of the entire pipeline, since an animal never detected can never be identified, and the cropping and alignment behavior of the detector determines the statistical distribution of inputs that the embedding network must tolerate. Pipelines that co-design the two stages, for example by training the embedding network on crops produced by the deployed detector rather than on hand-curated boxes, consistently report better field performance than those that treat the stages independently, a lesson inherited directly from the person re-identification literature (Zheng et al., 2016).

5. Deep Feature Extraction and Metric Learning for Re-Identification

The defining algorithmic commitment of modern animal re-identification is the replacement of classification over a fixed set of identities with the learning of an embedding function that maps images into a high-dimensional vector space in which distance encodes identity. The formulation originates in signature verification and face recognition, where Siamese networks trained with contrastive objectives were introduced to compare pairs of inputs (Chopra et al., 2005; Hadsell et al., 2006), and it reached maturity with FaceNet, which trained a single network with a triplet loss to produce compact embeddings supporting recognition, verification, and clustering in one representation (Schroff et al., 2015). The attraction for animal applications is structural. Populations are open: new individuals are born, immigrate, or are photographed for the first time continually, and a classifier with a fixed output layer must be retrained to accommodate them, whereas an embedding gallery accommodates them by insertion. Populations are also long-tailed, with a few well-photographed individuals and many singletons, a regime in which per-identity classifiers are starved of data but pairwise objectives can still learn.

The person re-identification literature supplied the experimental paradigm and much of the technical repertoire. Benchmark formulations based on query and gallery sets with rank-based evaluation were standardized on datasets such as Market-1501 (Zheng et al., 2015), and comprehensive surveys trace the field's progression through part-based models, attention mechanisms, and domain adaptation (Zheng et al., 2016; Ye et al., 2022). Specific techniques that transferred to animals include the batch-hard triplet mining strategy that made plain triplet losses competitive (Hermans et al., 2017), part-based descriptors that pool features over horizontal stripes or learned regions (Sun et al., 2018), and the collection of training refinements, including warmup schedules, label smoothing, and batch normalization necks, codified as strong baselines (Luo et al., 2019). From the face recognition line, margin-based softmax variants such as center loss, CosFace, and ArcFace, which shape the embedding space through angular margins rather than explicit

pair sampling, are now common choices for animal identification networks because they scale gracefully to many identities (Wen et al., 2016; Wang et al., 2018; Deng et al., 2019). Inheriting objectives from face recognition also means inheriting its vulnerabilities, including demographic bias and presentation attacks, which have prompted a dedicated countermeasure literature (Adeniyi, Odejebi, & Taiwo, 2025). Sobering methodological work has shown that reported gains among metric learning losses are smaller than the literature suggests once evaluation is controlled (Musgrave et al., 2020), a caution directly relevant to a field prone to comparing methods across incompatible protocols.

Transfer learning is the practical foundation of nearly all of this work, since few animal datasets can support training from scratch. Networks pre-trained on large generic corpora are fine-tuned on species-specific galleries, exploiting the general observation that learned visual representations transfer across domains (Pan & Yang, 2010; Zhuang et al., 2021). Curation of the fine-tuning corpus is itself a design variable, since context-aware selection of training material changes what the adapted network encodes (Oyesiji, Nwakamma, & Ojukwu, 2024), and parameter-efficient adaptation offers per-population specialization without full fine-tuning of the backbone (Oyesiji, Nwakamma, & Ojukwu, 2023). Aggressive data augmentation partially compensates for small galleries, with standard geometric and photometric transformations supplemented by domain-motivated corruptions such as synthetic motion blur, infrared conversion, and occlusion simulation (Shorten & Khoshgoftaar, 2019). More recently, self-supervised pre-training has entered the field, on the argument that unlabeled animal imagery is abundant even where identity labels are scarce. Contrastive frameworks and their negative-free and distillation-based successors learn representations from unlabeled images that rival supervised pre-training (Chen et al., 2020; He et al., 2020; Grill et al., 2020), and self-supervised vision transformers have been observed to develop attention patterns that localize salient object structure without supervision (Caron et al., 2021), a property of evident value for pattern-bearing animal surfaces.

Interpretability and diagnosis tools support the training loop. Low-dimensional projections of embedding spaces are routinely used to inspect whether individuals form separable clusters and whether nuisance factors such as camera identity or illumination dominate the learned geometry (van der Maaten & Hinton, 2008), and gradient-based attribution methods verify that networks attend to individually distinctive anatomy, stripes, spots, scars, and whisker fields, rather than to background context that will not survive a change of location (Selvaraju et al., 2017). These diagnostics matter more in animal applications than in many vision tasks because the most dangerous failure mode is silent: a network that identifies individuals through the camera site at which they were photographed can post excellent within-site validation scores and collapse entirely across sites.

Network analysis of pairwise similarity offers a complementary route from embeddings to identities. Rather than assigning each query independently to its nearest gallery neighbor, similarity graphs over entire photographic archives can be clustered so that connected components correspond to individuals, an approach demonstrated at scale on wild populations with deep metric embeddings feeding graph-based grouping (Miele et al., 2021). This perspective aligns identification with how the outputs are consumed downstream, since capture-recapture analysis requires a partition of observations into individuals rather than a ranked retrieval list, and it provides natural points at which human verification effort can be concentrated on uncertain graph edges rather than spread uniformly over all matches.

6. Wildlife Re-Identification Across Taxa

The taxonomic breadth of deep re-identification is now considerable, and the species-specific literature is best read as a set of variations on the shared architecture, with the interesting differences concentrated in which anatomical surface carries identity and how it is presented to the camera. Among patterned felids, tigers have served as the flagship. The ATRW benchmark introduced a substantial annotated dataset of Amur tigers with pose keypoints and evaluated detection, pose estimation, and re-identification jointly, establishing that pose-normalized embeddings materially outperform naive whole-body descriptors for laterally patterned species whose stripes deform with body articulation (Li et al., 2020). The benchmark formalized within-dataset and cross-dataset protocols that subsequent big cat work has adopted, and its lineage connects directly back to the three-dimensional surface matching insight of the pre-deep era (Hiby et al., 2009).

Primates recruited face recognition most directly, since primate faces are individually distinctive in ways homologous to human faces. Early systems adapted face detection and identification pipelines to chimpanzees in captive and wild settings (Loos & Ernst, 2013), log-Euclidean convolutional formulations improved identity and attribute prediction on wild chimpanzee archives (Freytag et al., 2016), and deep pipelines applied to fourteen years of village-scale video achieved frame-level face recognition robust enough to support longitudinal social analysis (Schofield et al., 2019). Purpose-built primate face recognition has also been directed at enforcement problems, with systems designed to identify trafficked golden monkeys and lemurs from photographs (Deb et al., 2018). Elephants, whose identity cues are distributed across ear contours, tusk configuration, and head morphology rather than surface pattern, have been addressed with detection followed by learned head descriptors (Körschens et al., 2018). For species and terrains that ground-level cameras serve poorly, the same architecture is being carried aloft, drawing on the platform, airspace, and regulatory experience accumulated in unmanned aerial survey programs (Dagodzo, 2018a; Dagodzo & Ahiaekwe Patrick, 2023b).

Marine mammals contributed some of the field's largest operational systems, reflecting decades of photographic identification practice in cetacean science. Curvature-based representations of fluke and fin trailing edges provided strong pre-deep baselines for dolphins and whales (Weideman et al., 2017). Deep systems then absorbed the task: right whale identification from

aerial and vessel imagery was demonstrated with convolutional pipelines in a widely noted data science collaboration (Bogucki et al., 2019), and a fully automated humpback fluke matching system now operates at global scale across an archive contributed by thousands of researchers and citizen scientists, reporting matching accuracy high enough to have effectively retired manual matching for that species (Cheeseman et al., 2022). Pinnipeds with pelage patterns have been served by Siamese matching of pattern crops, demonstrated on the small and endangered Saimaa ringed seal population (Nepovinnikh et al., 2020). Among bears, whose soft tissue and seasonal mass change defeat pattern-based approaches, facial landmark detection followed by deep embeddings achieved individual recognition adequate for monitoring applications (Clapham et al., 2020).

Two systems-level observations emerge from this species survey. The first is that platform projects have been as consequential as algorithms. Wildbook and its associated detection and curation pipeline assembled detection, matching algorithms, human verification interfaces, and data management into an infrastructure that hosts identification catalogues for dozens of species, and in doing so defined the operational pattern, algorithm-assisted human curation of a growing gallery, that most large deployments follow (Berger-Wolf et al., 2017; Parham et al., 2018). The second is that human-competitive accuracy is no longer exceptional. Across several taxa, learned embeddings match or exceed experienced observers on held-out matching tasks (Schneider et al., 2020), which shifts the binding constraints from matching accuracy per se to the surrounding system: detection recall, gallery hygiene, cross-site generalization, and the quantification of confidence for downstream analysis.

At the same time, the distribution of effort across taxa remains uneven in ways that reflect data availability rather than scientific priority. Species with conspicuous patterns, long-standing photo-identification traditions, and organized image archives are well served; cryptic, unpatterned, and rarely photographed species are not. Reviews of the multi-species literature note that most published systems are trained and evaluated on single species and single sites, with genuinely cross-species architectures and evaluations still uncommon (Ravoor & Sudarshan, 2020). This unevenness matters for the large-network setting, since regional camera networks observe whole communities, and per-species bespoke systems multiply engineering and maintenance costs in proportion to community richness.

7. Livestock Re-Identification in Precision Agriculture

Livestock applications share the wildlife pipeline while inverting several of its constraints, and the agricultural literature has developed largely in parallel with limited cross-citation. Identification of cattle in particular has a long pre-vision history, and reviews of the transition from mechanical and electronic identification, ear tags, brands, and RFID, to image-based biometrics catalogue both the welfare and reliability motivations for camera-based alternatives: tags are lost, damaged, duplicated, and fraudulently swapped, and their application is painful (Awad, 2016). Muzzle prints, the bovine analogue of fingerprints, supported some of the earliest biometric work, and deep networks over muzzle point imagery achieve high identification accuracy in controlled capture settings (Kumar et al., 2018). Identity is the key that binds every other per-animal record, and the nutrition literature makes the granularity of those records explicit, since intake, digestibility, growth, and carcass outcomes are estimated per animal and pooled across small experimental groups (Aye & Tawose, 2015; Oluwadele, Ekeocha, Aganga, et al., 2024; Tawose, Ekeocha, & Oluwadele, 2023).

For unconstrained settings, coat patterning carries identity in patterned breeds. Detection and identification of Holstein Friesian cattle from overhead and unmanned aerial viewpoints was demonstrated with convolutional pipelines that localize animals and read their dorsal patterns (Andrew et al., 2017), and subsequent work reformulated the problem in explicitly open-set terms, using deep metric learning so that animals never seen during training can be enrolled and re-identified from their embeddings alone, reporting strong accuracy on herds an order of magnitude larger than the training population (Andrew et al., 2021). Facial recognition has been pursued for species and breeds without distinctive coat patterns, with pig face recognition demonstrated under farm conditions and its specific difficulties, pose variation, dirt, and rapid growth-driven appearance change, documented (Hansen et al., 2018). In poultry systems the same identity problem arises at flock scale, where phenotypic characterization and physiological stress indexing of scavenging birds depend on reliable attribution of observations to individuals (Ekeocha et al., 2021; Oluwadele, Tawose, Ekeocha, et al., 2023). Framework-level work has combined detection backbones with recurrent components to identify individual cattle in walkway video (Qiao et al., 2019).

The farm setting is distinctive in three respects. First, the population is closed and enumerable: the set of legitimate identities is known, entry and exit are administratively recorded, and the open-set problem reduces to flagging anomalies rather than enrolling unknowns, which permits classification-style formulations that wildlife systems cannot use. Second, the environment is partially controllable: camera placement at races, feeders, and milking stations constrains pose and range, and lighting can be engineered, so livestock systems often operate closer to the curated end of the difficulty spectrum. Third, the tolerance for error is asymmetric and economically defined: misidentification propagates into feeding, medication, and breeding records, so systems are evaluated against traceability and management requirements rather than ecological estimator bias. Haematological and biochemical monitoring under alternative feeding regimes, and the feed quality surveillance that accompanies it, illustrate the record streams that identity must link correctly (Tawose, Ekeocha, & Oluwadele, 2024; Tawose & Oluwadele, 2025; Tawose, Oluwadele, et al., 2023). Systematic reviews of machine vision in animal farming situate identification within this broader monitoring agenda, alongside behavior recognition, body condition scoring, and welfare assessment, and note that individual identity is the key that links all per-animal analytics together

(Wurtz et al., 2019; Neethirajan, 2020). Vision-based identification also sits inside a wider mechanization agenda in which handling and processing equipment is instrumented and evaluated for throughput and product quality, so that per-animal records and per-machine records can be reconciled (Bello, Adama, Tawose, & Bolaji, 2022; Bello, Tawose, Adama, & Bolaji, 2022; Bello, Oyelaran, et al., 2024).

Convergence between the agricultural and wildlife literatures is nonetheless visible and productive. The open-set metric learning formulation developed for cattle is precisely the wildlife formulation; conversely, the engineering discipline of continuous barn-scale video processing, with its attention to throughput, hardware cost, and integration with management software, anticipates the operational maturity that large wildlife networks are only beginning to reach. A shared research agenda on lifelong appearance change, calves growing, coats weathering, wounds healing, is emerging in both communities, since neither farm nor field systems can afford galleries that silently expire as animals age.

8. Multi-Camera Tracking and Open-Set Identification in Large Networks

Scaling from single cameras to networks changes the problem qualitatively, because identity must now be maintained across sensors that never share a field of view, across revisit intervals ranging from seconds to years, and across appearance shifts induced by viewpoint, illumination, season, and animal condition. The computer vision community formalized the within-network portion of this problem as multi-target multi-camera tracking, contributing datasets and performance measures that quantify how well trajectories are associated across cameras (Ristani et al., 2016). Within a single camera, tracking-by-detection remains the dominant paradigm: detections are linked frame to frame by motion models and appearance descriptors, with the SORT and DeepSORT algorithms establishing the standard combination of Kalman filtering, assignment, and learned appearance features (Bewley et al., 2016; Wojke et al., 2017), and later work showing that associating even low-confidence detection boxes materially reduces identity fragmentation under occlusion (Zhang et al., 2022). Standardized benchmarks and metrics, from the CLEAR MOT measures through the MOT challenge series to the higher-order HOTA metric that balances detection and association quality, supplied the evaluation vocabulary that animal deployments now borrow (Bernardin & Stiefel, 2008; Milan et al., 2016; Luiten et al., 2021).

Animal camera networks stress these tools in specific ways. Wildlife cameras are sparse and non-overlapping by design, spaced across landscapes at densities set by ecology rather than coverage geometry, so cross-camera association cannot rely on trajectory continuity and rests almost entirely on appearance embeddings, with capture times and inter-camera distances contributing weak spatiotemporal priors. Holding those priors in a spatial database rather than in ad hoc spreadsheets is an integration pattern already established in utility asset management and enterprise drone operations (Dagodz & Ahiaek Patrick, 2021a; Dagodz & Ahiaek Patrick, 2023a), and geospatial predictive analytics developed for site selection suggests that camera placement can be treated as an optimization variable rather than a fixed survey constraint (Tonoyan, Dada, & Ayivi-Donkor, 2022; Akanbi & Sunday, 2024). Triggered still capture produces short bursts rather than continuous video, reducing within-camera tracking to burst grouping. Nocturnal captures arrive in near-infrared, in which color-based descriptors collapse and pelage patterns present differently than in daylight, effectively creating a cross-modality matching problem within a single gallery. Livestock networks sit at the opposite pole, with dense overlapping coverage of confined spaces, continuous video, and crowding-driven occlusion as the dominant difficulty. The shared requirement is an association layer that treats the embedding gallery, rather than the track, as the persistent representation of identity, with tracks serving as evidence aggregation windows that produce higher-quality gallery entries than single frames. Maintaining consistent state across many semi-autonomous sensing agents is studied directly in mobile robot fleets, where coordination policies rather than per-agent perception set system-level performance (Akanbi & Sunday, 2025; Akanbi, Ganiu, & Sunday, 2025).

Open-set recognition is the second defining property of the network setting. In field populations the gallery is always incomplete: births, immigration, and first captures introduce individuals that no threshold-free nearest-neighbor assignment can handle correctly. Practical systems therefore operate in a verification-plus-enrollment regime, accepting a match only when embedding distance passes a calibrated threshold and otherwise proposing a new identity for human confirmation. The difficulty is that the score distributions of genuine and impostor pairs shift with image quality, viewpoint, and site, so a single global threshold is rarely adequate, and the cost structure is asymmetric: a false merge of two individuals corrupts capture histories in ways that bias abundance estimates downward, while a false split inflates them. The graph-clustering perspective, in which the archive-wide similarity graph is partitioned and ambiguous edges are routed to human review, provides one principled response, concentrating scarce expert attention where it changes the partition (Miele et al., 2021). Platform deployments institutionalize the same logic through curation queues in which algorithmic candidates are verified before entering the canonical catalogue (Berger-Wolf et al., 2017).

Cross-site generalization deserves separate emphasis because it is the dominant failure mode reported when systems trained at one location are deployed at another. Recognition performance measured within a site systematically overstates network performance, since networks entangle identity with background, camera model, and site-specific capture statistics; the phenomenon is the individual-level analogue of the cross-location degradation documented for species classification (Beery et al., 2018). Defenses are those of domain generalization generally: training across many sites, augmentation toward the deployment distribution, background suppression through detection-stage segmentation, and diagnostic attribution to confirm that identity evidence is

anatomical rather than contextual (Selvaraju et al., 2017). The design lesson for network operators is that gallery construction should deliberately mix sites and seasons per individual wherever field protocols allow, because embedding networks trained on single-site sightings of each individual have never been shown what invariance across sites looks like.

Temporal drift, finally, is intrinsic rather than adversarial. Animals grow, scar, molt, and senesce; equipment weathers; vegetation regrows in front of fixed cameras. Long-horizon identity maintenance therefore requires galleries that are versioned in time, with per-individual exemplar sets refreshed as new verified sightings accumulate and confidence decaying for individuals unseen over long intervals. These requirements are operational rather than algorithmic, but they determine whether a network's identification layer remains trustworthy in year five, and they are conspicuously under-described in a literature whose evaluations are overwhelmingly cross-sectional.

9. Embedding Indexing, Retrieval, and Scalability

Once identity resides in embedding space, identification at network scale is a similarity search problem, and its computational character is set by gallery size, query rate, and dimensionality. Exhaustive comparison is linear in gallery size and rapidly becomes the bottleneck for archives accumulating millions of detections. The approximate nearest neighbor literature supplies the standard remedies. GPU-resident indexes support billion-scale search with product quantization and inverted file structures, trading controlled amounts of recall for orders-of-magnitude throughput gains (Johnson et al., 2021), while graph-based indexes such as hierarchical navigable small world structures provide state-of-the-art query complexity on CPU and support the incremental insertion that continuously enrolling systems require (Malkov & Yashunin, 2020). Analogous serving pressures in recommendation and ranking have produced a compression literature aimed at holding retrieval models within tight latency and memory budgets, which transfers directly to identification galleries queried at network scale (Oyesiji, Nwakamma, & Ojukwu, 2025).

System designers face three recurring choices. The first is exemplar policy: whether each individual is represented by all verified embeddings, a curated exemplar subset, or an aggregated prototype, a choice that trades gallery size against robustness to viewpoint and condition variation. The second is index freshness: incremental structures accommodate continuous enrollment, but periodic re-training of the embedding network invalidates the entire index and forces re-embedding of the archive, so network operators must version embeddings jointly with model checkpoints and budget for full re-indexing as a routine maintenance operation rather than an exceptional event. The third is the division of labor between retrieval and verification: approximate search returns candidate lists cheaply, after which exact distances, geometric verification on pattern keypoints for suitable species, or human review adjudicate the final match, an architecture inherited from the pre-deep systems in which fast candidate generation preceded careful pairwise scoring (Crall et al., 2013).

Scalability also has a human dimension that the retrieval literature does not capture. In operational catalogues, the scarce resource is verification labor, and the effective throughput of the identification layer is set by how efficiently candidate matches can be confirmed. Interfaces that present ranked candidates with aligned pattern regions, that batch decisions by individual, and that route only low-margin decisions to experts multiply the effective capacity of a fixed curation team, which is why platform projects report interface and workflow design as first-order determinants of catalogue quality (Berger-Wolf et al., 2017). Because the galleries themselves carry sensitive location data, sensitivity classification and traceability mechanisms from enterprise data governance supply the access control patterns that such catalogues require (Aliliele, Mbonu, & Iwuanyanwu, 2024). Retrieval quality and human efficiency compound: each improvement in embedding discrimination shortens candidate lists, and each improvement in interface design lowers the cost of the residual ambiguity.

10. Datasets, Benchmarks, and Evaluation Protocols

The dataset landscape explains much of both the field's progress and its remaining immaturity. On the species-classification side, large public archives exist and have been transformative: Snapshot Serengeti established the template of openly released, expert-and-crowd annotated camera trap corpora at the million-image scale (Swanson et al., 2015), subsequent multi-project releases enabled the cross-location generalization studies that revealed domain shift as the central obstacle (Beery et al., 2018), and platform initiatives now aggregate camera trap data globally with standardized metadata (Ahumada et al., 2020). Individual-level datasets are far scarcer, smaller, and more fragmented, because identity annotation requires either invasive ground truth or expert matching effort that only long-term studies can supply. ATRW remains among the largest purpose-built wildlife re-identification benchmarks with its several thousand images of a few hundred tigers (Li et al., 2020); most published studies otherwise train and evaluate on private or semi-private galleries of tens to a few hundreds of individuals, and livestock studies similarly report on single herds (Andrew et al., 2021; Hansen et al., 2018).

This scarcity has methodological consequences. Small closed galleries invite closed-set evaluation, and closed-set accuracy is the least informative number for network deployment, where the operative questions are open-set: how often is a new individual wrongly matched to the gallery, and how often is a known individual wrongly enrolled as new. The person re-identification community's protocol vocabulary, cumulative match characteristic curves, mean average precision over query-gallery splits, and cross-camera evaluation in which query and gallery are drawn from disjoint cameras (Zheng et al., 2015; Ye et al., 2022), transfers

directly and is increasingly used, but reporting remains heterogeneous, and cross-paper comparison is further undermined by the sensitivity of metric learning results to implementation detail documented in controlled re-evaluations (Musgrave et al., 2020). Reviews of the animal re-identification literature repeatedly identify the absence of shared benchmarks with standardized open-set splits as the field's principal reproducibility deficit (Schneider et al., 2019; Ravor & Sudarshan, 2020).

Evaluation design must also respect the downstream use of the outputs, which distinguishes animal re-identification from its person-centric parent. When identification histories feed capture-recapture estimators, the relevant error quantities are the rates of false merges and false splits and their correlation structure over individuals and occasions, because these translate directly into estimator bias; a system selected for rank-1 accuracy may be dominated, for ecological purposes, by one selected for calibrated match scores at a conservative operating point (Foster & Harmsen, 2012; Royle et al., 2013). Benchmark design that reports full score distributions and operating characteristics, rather than single summary ranks, would allow ecological users to select operating points against their own loss functions, and represents a low-cost convention the field could adopt immediately. Keeping a growing catalogue verifiable is a related problem, and continuous auditing methods developed for technology risk oversight offer one tested route to it (Aliliele, Mbonu, Uzoka, & Iwuanyanwu, 2025).

Data programs that would most accelerate progress are identifiable from the gaps. Longitudinal galleries with verified identities spanning years would support the study of appearance drift that cross-sectional datasets cannot touch. Multi-site galleries with per-individual sightings across cameras would make cross-site generalization measurable rather than anecdotal. Community-scale galleries covering many species at one site would support the multi-species architectures that regional networks actually need (Ravor & Sudarshan, 2020). Citizen science and platform archives suggest the supply side is available: the humpback system's global archive demonstrates that distributed contribution plus algorithmic matching plus expert curation can assemble identity datasets of a scale no single project could fund (Cheeseman et al., 2022), and crowd annotation pipelines validated for species labels extend naturally to match verification (Swanson et al., 2015; Willi et al., 2019).

11. Edge Deployment in Distributed Camera Networks

Large networks are bandwidth-constrained and power-constrained before they are compute-constrained, and this shapes where in the network each pipeline stage runs. Remote wildlife cameras typically operate on batteries or small solar panels with intermittent satellite or cellular connectivity, so transmitting full-resolution imagery to central servers is often infeasible, and the economically transmittable unit is the detection crop, the embedding vector, or merely the event record. The edge computing literature frames the general trade-offs of pushing inference toward the data source, latency, bandwidth, privacy, and autonomy against the constraints of embedded compute and energy budgets (Shi et al., 2016), and surveys of edge intelligence catalogue the model-partitioning and offloading strategies available between device, gateway, and cloud tiers (Zhou et al., 2019). Work on intermittently powered sensing makes those trade-offs concrete, since radio frequency and ambient harvesting budgets determine how much computation and communication a node can afford between energy arrivals, and reliability, latency, and energy cannot be tuned independently of one another (Asiedu & Quainoo, 2023a; Asiedu & Quainoo, 2023b; Asiedu & Quainoo, 2024; Asiedu, Quainoo, & Asiedu, 2025). Protocol-level results show that reliable delivery under intermittent power requires scheduling aligned to energy epochs rather than assuming continuous availability (Asiedu & Quainoo, 2025), and lightweight predictive models of harvest can smooth duty cycles when the energy supply is uncertain (Asiedu, Quainoo, & Asiedu, 2026). Achievable uplink throughput is equally consequential for a network that must move crops or embeddings off the device, and it is governed by impedance matching, channel conditions, and radio coexistence within the wireless system on chip (Quainoo, Ogundapo, & Asiedu, 2024; Quainoo, Ogundapo, & Asiedu, 2025a; Quainoo & Ogundapo, 2026a; Quainoo & Ogundapo, 2026b), which is why disciplined test automation is part of the deployment path rather than an afterthought (Quainoo, Ogundapo, & Asiedu, 2025b).

Model compression is the enabling toolset. Pruning and weight quantization compress trained networks by an order of magnitude with limited accuracy cost (Han et al., 2016), integer-only quantization schemes co-designed with inference runtimes make eight-bit deployment routine on mobile-class accelerators (Jacob et al., 2018), and knowledge distillation transfers the behavior of large teachers into deployable students (Hinton et al., 2015). Recent on-device multimodal work applies exactly this toolkit of quantization, pruning, and distillation to reach privacy-preserving inference inside embedded energy envelopes (Nwakamma, Ojukwu, & Oyesiji, 2025), and portable component-model runtimes address the language and sandboxing constraints of the heterogeneous device fleets that real networks accumulate (Ojukwu, Oyesiji, & Nwakamma, 2025). Architecture families designed for constrained budgets, the MobileNet line with its depthwise separable convolutions and inverted residuals and the compound-scaled EfficientNet and EfficientDet families, provide the standard backbones for on-device detection and embedding (Howard et al., 2017; Sandler et al., 2018; Tan & Le, 2019; Tan et al., 2020). The YOLO family's persistent popularity in the animal domain owes much to how readily its recent members compress into this regime while retaining useful accuracy (Wang et al., 2023).

The canonical partition that emerges in deployed animal networks runs detection at the device, embedding at the device or gateway depending on hardware, and gallery search plus association centrally. On-device detection converts continuous or triggered capture into a sparse event stream and discards empty frames at the source, which is the single largest bandwidth economy available given the false-trigger rates of motion-activated cameras (Beery et al., 2019). On-device embedding compresses each detection into a vector hundreds of times smaller than its source crop, at the cost of freezing the embedding model version at the edge; networks

that anticipate model revision transmit crops for their highest-value detections alongside embeddings so that the archive can be re-embedded centrally when models improve. Central association is retained because gallery search, enrollment decisions, and human verification are inherently global operations over the network's accumulated state. Unattended field nodes are also an attack surface, and the defensive literature on malware in cloud-native and edge microservices, on cyber-physical intrusion in operational technology, and on crypto-agile migration of transport security applies with little modification (Idika et al., 2021; Ojukwu, Oyesiji, & Nwakamma, 2023; Nwakamma, Ojukwu, & Oyesiji, 2026). As identification pipelines acquire autonomous orchestration layers, the threat model widens further to include prompt injection, tool poisoning, and excessive agency (Nwakamma, Ojukwu, & Oyesiji, 2024b).

Livestock deployments relax the energy constraint but tighten latency and reliability requirements, since identification gates real-time management actions at feeders, sorters, and milking systems. Continuous multi-camera video at barn scale makes per-frame cloud offload uneconomical, and the same compression toolset serves instead to fit detection, tracking, and embedding within on-premises hardware budgets (Neethirajan, 2020). In both settings, the systems literature remains thinner than the algorithms literature: published accounts of multi-year, multi-hundred-camera deployments with quantified energy, bandwidth, and maintenance costs are rare, and the field's engineering knowledge is disproportionately held in project teams rather than in the archival record (Glover-Kapfer et al., 2019). Published accounts of national-scale unmanned aerial programs in resource-constrained settings document precisely the logistics, regulation, and maintenance burdens that comparable camera networks will encounter (Dagodzo & Ahiaeké Patrick, 2025), and predictive maintenance models for distributed systems offer a template for anticipating node failure before it opens silent coverage gaps (Mayo et al., 2021).

12. Integration with Population Ecology and Management Decisions

The scientific value of automated re-identification is realized only when its outputs enter the statistical and operational frameworks that consume individual identity. In wildlife science the primary consumer is capture-recapture inference. Photographic capture histories have supported abundance and density estimation since the foundational tiger studies (Karanth, 1995; Karanth & Nichols, 1998), the approach generalized across cryptic carnivores (Silver et al., 2004; Jackson et al., 2006), and spatially explicit models now use the locations of identified encounters to estimate density directly (Royle et al., 2013). These estimators inherit the error structure of the identification layer. Misidentification analyses established well before automation showed that even modest false-split rates inflate abundance estimates substantially, while false merges deflate them, and that identification error is most dangerous when it is unmodeled (Foster & Harmsen, 2012). Automated systems change the terms of this problem favorably in one respect, their error rates are measurable on held-out data in a way that ad hoc manual matching never was, but exploiting that advantage requires estimators that accept match confidence as data, and the coupling of embedding-based identification with error-aware capture-recapture models remains an underdeveloped interface.

Survey design and analysis practice are also reshaped. Reviews of camera trapping as a survey method emphasize that design choices, camera density, placement, and duration, should follow from the target ecological process (Burton et al., 2015), and comparative work finds camera-based methods competitive with or superior to alternative survey instruments across many contexts (Wearn & Glover-Kapfer, 2019). Automated identification alters the economics underlying these designs: when the marginal cost of processing an image approaches zero, longer deployments and denser networks stop being analysis-limited, and continental-scale individual-level monitoring, of the kind envisioned when camera networks were first proposed as a biodiversity observatory, becomes a data management problem rather than a labor problem (Steenweg et al., 2017; Ahumada et al., 2020). The supporting analytical infrastructure is well described outside ecology, where lifecycle performance models and resilient logistics frameworks turn continuous sensor streams into operational decisions (Adelanwa, Basnet, & Anene, 2023; Anene & Clement, 2022), and where spatially explicit risk models built for data-limited environmental monitoring show how sparse observations can be propagated into continuous surfaces for early warning and regulatory reporting (Lamidi & Olamide, 2018; Azeez & Badmus, 2024; Badmus, 2025).

Management and enforcement applications impose their own requirements. Longitudinal registries of identified individuals support translocation planning, human-wildlife conflict attribution, and the forensic matching of confiscated animal products to known individuals, the capability first demonstrated for tiger skins (Hiby et al., 2009). In production agriculture, identity is the linchpin of traceability regimes and per-animal health analytics, and the acceptance criteria are set by regulatory and commercial standards rather than estimator bias (Awad, 2016; Wurtz et al., 2019). Where identity histories feed disease surveillance, the outputs enter One Health programs that join animal, environmental, and human health data, including work on environmental vectors of avian pathogens and on rapid biosensor-assisted detection of resistant organisms (Ayinde, 2025; Ayinde et al., 2024; Mensah et al., 2026; Olorunkosebi, Ayinde, & Mensah, 2023). Surveillance dashboard architectures from public health demonstrate how such detections are made actionable for decision makers rather than left in analyst notebooks (Ogunboye et al., 2023; Oparah et al., 2023). Across both domains, the common institutional requirement is auditability: decisions with legal or regulatory consequence require that a claimed identification be reviewable, which favors pipelines that retain evidence images, match scores, and verification records rather than emitting bare identity labels.

The interface between identification systems and their consumers is, finally, where uncertainty quantification becomes non-optional. Point identifications suffice for browsing and curation; inference and enforcement require calibrated confidence. The

elements exist separately in the literature, score calibration in recognition, measurement-error models in ecology, but their integration is occasional rather than systematic, and reviews from the ecological side consistently list uncertainty propagation from vision outputs into ecological models among the field's outstanding needs (Weinstein, 2018; Christin et al., 2019; Tuia et al., 2022).

13. Challenges and Open Problems

Data remains the first-order constraint. Individual-level labels are expensive by nature, galleries are small and long-tailed, and the resulting benchmarks are too weak to discriminate among methods or to certify deployment readiness. The imbalance is structural: a handful of well-studied populations contribute most labeled identities, while the taxa and regions where monitoring needs are greatest contribute least. Self-supervised pre-training on unlabeled animal imagery offers the most credible route around the labeling bottleneck (Chen et al., 2020; Caron et al., 2021), and platform-mediated crowd contribution has demonstrated that archives of sufficient scale can be assembled when incentives and infrastructure align (Cheeseman et al., 2022), but the field currently lacks the shared, openly licensed, open-set benchmark suite that would convert these resources into measurable progress (Schneider et al., 2019; Ravoor & Sudarshan, 2020).

Generalization failures constitute the second cluster. Cross-site domain shift, day-night modality shift, and long-horizon appearance drift all degrade deployed performance relative to published validation figures, and all three are underrepresented in evaluation practice because cross-sectional single-site datasets cannot express them (Beery et al., 2018). The silent failure mode, networks that exploit background and site context as identity shortcuts, is particularly corrosive because it inflates exactly the numbers on which deployment decisions are made; attribution-based auditing mitigates it diagnostically (Selvaraju et al., 2017), and detection-stage segmentation mitigates it architecturally, but a principled and routine defense is not yet standard practice. Related, the field has no settled approach to lifelong gallery maintenance: enrollment policy, exemplar refresh, embedding version management, and confidence decay for unseen individuals are handled by local convention in each deployment, without the shared vocabulary that would let practices be compared and improved.

Open-set behavior and calibration form the third cluster. The decisive events in network operation, enrolling a genuinely new individual versus matching a known one, occur precisely where nearest-neighbor confidence is least trustworthy, and score distributions shift with quality, viewpoint, and site in ways that defeat global thresholds. Verification-and-enrollment workflows with human adjudication are the current answer, and graph-based formulations rationalize where that human effort is spent (Miele et al., 2021), but principled open-set recognition with calibrated novelty detection, evaluated against ecologically weighted error costs, remains open. Its resolution matters doubly because downstream capture-recapture inference needs exactly the calibrated error characterization that current systems do not emit (Foster & Harmsen, 2012).

The fourth cluster is systems knowledge. The algorithmic literature has raced ahead of the deployment literature: multi-year accounts of operating costs, failure modes, maintenance burdens, and data governance for large identification networks are scarce (Glover-Kapfer et al., 2019), and the human dimension of curation, the workflows, interfaces, and quality assurance that platform projects identify as decisive (Berger-Wolf et al., 2017), is almost invisible in the peer-reviewed record. The consequence is that new deployments repeatedly rediscover operational lessons, and that the field's claims about scalability rest on extrapolation more than on documented practice. A modest reporting convention, publishing deployment postmortems with quantified error, cost, and labor figures, would compound quickly across a community whose installations now number in the thousands of cameras.

Ethical and dual-use considerations, finally, are inherited from the surveillance technologies the field adapts. Individual-level location histories of endangered animals are sensitive by nature, since the same registry that supports conservation supports poaching if exposed, and platform operators already treat location data governance as a security problem. Livestock identification raises conventional data protection questions around farm business data. The community's norms here are developing pragmatically, but the review literature has begun to note that the governance of animal biometric data deserves deliberate treatment rather than inheritance by default from open-data conventions designed for species-level records (Tuia et al., 2022). The closest available templates come from regulated industries, where responsible governance structures for automated decision systems, legal and ethical risk modeling for data protection, and algorithmic accountability requirements are already codified (Annan, 2024; Mbonu et al., 2018; Ominyi & Anichukwueze, 2024; Ominyi, Anichukwueze, & Uzougbo, 2026). Automated evaluation and release gating, including the use of model-based judges, is one emerging mechanism for keeping a deployed recognition model within its certified specification (Nwakamma, Ojukwu, & Oyesiji, 2024a).

14. Future Research Directions

Four directions appear most consequential for the next phase of the field. The first is foundation-style representation learning for animal identity. The ingredients, abundant unlabeled imagery, self-supervised objectives that discover object structure without labels, and transformer architectures whose attention maps localize distinctive anatomy, are all in place (Dosovitskiy et al., 2021; Caron et al., 2021; Grill et al., 2020), and a representation pre-trained across taxa and fine-tuned per population would directly attack both the data scarcity and the cross-species fragmentation documented above. Cross-species transfer is the natural evaluation: a field that can enroll a new species with tens rather than thousands of labeled identities will have changed its own economics.

The second is uncertainty-aware identification designed for its consumers. This entails calibrated match scores stable across sites and conditions, explicit novelty probabilities for enrollment decisions, and estimator interfaces through which identification uncertainty propagates into abundance, survival, and movement inference rather than being resolved prematurely to point labels (Royle et al., 2013; Foster & Harmsen, 2012). The graph formulation of archive-wide identification provides a congenial substrate, since posterior uncertainty over graph partitions maps naturally onto uncertainty over capture histories (Miele et al., 2021).

The third is the co-design of edge and network architecture with identification workloads. As camera networks densify and hardware accelerators reach the device tier, the partition of detection, embedding, and association across tiers becomes a design space to be optimized against energy, bandwidth, latency, and model-update constraints rather than a fixed convention (Zhou et al., 2019; Jacob et al., 2018). Concrete open problems include embedding schemes robust to model version skew across a heterogeneous device fleet, on-device continual adaptation within certified bounds, and event-triggered sensing policies in which identification confidence feeds back into capture behavior. Energy-aware scheduling of that feedback loop is the point at which identification, communication, and power management become a single design problem (Asiedu, Quainoo, & Asiedu, 2026). Fusion with other remote sensing modalities, whose interpretation traditions are long established in earth observation, is a natural extension once identity is treated as one layer in a georeferenced monitoring stack (Dagodzo, 2018c).

The fourth is consolidation of the field's empirical infrastructure: shared open-set benchmarks with longitudinal and multi-site structure, standardized reporting of operating characteristics rather than single summary ranks, and publication norms that surface deployment experience alongside algorithmic novelty (Schneider et al., 2019; Musgrave et al., 2020; Ravor & Sudarshan, 2020). None of this is glamorous, and all of it is rate-limiting. The trajectory of the neighboring species-classification problem, transformed within five years by open data, shared detectors, and platform infrastructure (Swanson et al., 2015; Beery et al., 2019; Ahumada et al., 2020), is the available proof that infrastructure investments of this kind repay themselves quickly in this domain.

15. Conclusion

Non-invasive animal re-identification has moved, within roughly a decade, from a collection of ingenious species-specific pattern matchers to a shared architecture in which single-stage detection, deep metric embedding, and similarity search jointly resolve individual identity across networks of cameras. The transition has been driven less by domain-specific invention than by the disciplined importation of methods matured elsewhere in computer vision, and its most visible results are operational: global-scale matching systems for flagship species, open-set identification of farm animals without tags, and platform infrastructures in which algorithms and human curators jointly maintain identity catalogues of scientific and regulatory standing.

The review's central finding is that the binding constraints have migrated outward from the matching algorithm to the surrounding system. Within curated evaluations, learned embeddings now rival expert human matchers across a widening set of taxa. In deployed networks, performance is governed instead by detection recall under field conditions, by generalization across sites, modalities, and years, by the calibration of open-set decisions, by the scalability of gallery search and human verification, and by the fidelity with which identification uncertainty reaches the ecological and operational models that consume identity. Each of these is tractable, and each is currently limited more by evaluation practice, shared data, and documented deployment experience than by algorithmic capability.

The field's near-term agenda follows directly. Self-supervised, cross-species representations promise to relax the labeled-data constraint that has shaped the literature's uneven taxonomic coverage. Uncertainty-aware identification promises to make automated outputs statistically consumable rather than merely convenient. Edge-aware system design promises networks whose sensing, computation, and communication are jointly organized around identity as the unit of information. And shared benchmarks with open-set, longitudinal, and multi-site structure promise to make progress on all of the above measurable. A field that completes this agenda will have delivered something of consequence well beyond computer vision: continuous, individual-resolution observation of animal populations at the scale of landscapes, obtained without touching a single animal.

References

- Adelanwa, A., Basnet, A., & Anene, U. N. (2023). Data driven digital transformation models for lifecycle performance management in infrastructure delivery. *International Journal of Advanced Multidisciplinary Research and Studies*, 3(6), 2646–2662.
- Adeniyi, A. I., Odejebi, O., & Taiwo, T. (2025). Countermeasures against bias and spoofing in modern facial recognition systems. *World Journal of Advanced Research and Reviews*, 25(1), 1914–1930.
- Adeyelu, O. O. (2022). A regulatory evidence framework for assessing unauthorized structure violations within protected airport obstacle limitation surfaces. *International Journal of Scientific Research in Civil Engineering*, 6(6), 248–283.
- Adeyelu, O. O., & Dagodzo, D. (2024). Advances in artificial intelligence and data-driven wildlife hazard monitoring and incident reduction at international airports in West Africa. *International Journal of Scientific Research in Science and Technology*, 11(5), 872–910.
- Ahumada, J. A., Fegraus, E., Birch, T., Flores, N., Kays, R., O'Brien, T. G., Palmer, J., Schuttler, S., Zhao, J. Y., Jetz, W., Kinnaird, M., Kulkarni, S., Lyet, A., Thau, D., Duong, M., Oliver, R., & Dancer, A. (2020). Wildlife Insights: A platform to maximize the potential of camera trap and other passive sensor wildlife data for the planet. *Environmental Conservation*, 47(1), 1–6.

- Akanbi, O., & Sunday, E. A. (2024). An integrated path-planning and slotting optimization model for AMR-enabled high-density warehousing. *International Journal of Advanced Multidisciplinary Research and Studies*, 4(6), 3226–3243.
- Akanbi, O., & Sunday, E. A. (2025). A systematic review of AI-driven autonomous mobile robots (AMR) in scalable micro-fulfillment centers. *International Journal of Advanced Multidisciplinary Research and Studies*, 5(6), 2363–2379.
- Akanbi, O., Ganiu, O. S., & Sunday, E. A. (2025). A multi-agent AI framework for swarm intelligence in autonomous mobile robot (AMR) fleet coordination. *International Journal of Scientific Research in Humanities and Social Sciences*, 2(1), 81–109.
- Aliliele, C., Mbonu, I. S., & Iwuanyanwu, U. (2024). A conceptual framework for enterprise data sensitivity classification and regulatory traceability mechanisms. *International Journal of Advanced Multidisciplinary Research and Studies*, 4(6), 3103–3124.
- Aliliele, C., Mbonu, I. S., Uzoka, E., & Iwuanyanwu, U. (2025). A review of AI assisted continuous auditing systems in technology risk and cybersecurity oversight. *Gyanshauryam, International Scientific Refereed Research Journal*, 8(4), 210–250.
- Andrew, W., Gao, J., Mullan, S., Campbell, N., Dowsey, A. W., & Burghardt, T. (2021). Visual identification of individual Holstein-Friesian cattle via deep metric learning. *Computers and Electronics in Agriculture*, 185, 106133.
- Andrew, W., Greatwood, C., & Burghardt, T. (2017). Visual localisation and individual identification of Holstein Friesian cattle via deep learning. *Proceedings of the IEEE International Conference on Computer Vision Workshops*, 2850–2859.
- Anene, U. N., & Clement, T. (2022). A resilient logistics framework for humanitarian supply chains: Integrating predictive analytics, IoT, and localized distribution to strengthen emergency response systems. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 8(5), 398–424.
- Annan, A. O. (2024). Algorithmic accountability and trade secret protection in artificial intelligence. *Shodhshauryam, International Scientific Refereed Research Journal*, 7(5), 315–347.
- Arzoumanian, Z., Holmberg, J., & Norman, B. (2005). An astronomical pattern-matching algorithm for computer-aided identification of whale sharks *Rhincodon typus*. *Journal of Applied Ecology*, 42(6), 999–1011.
- Asiedu, W., & Quainoo, R. (2023a). How far can energy harvesting take us? A systematic review of radio frequency strategies for energy autonomous sensing. *Shodhshauryam, International Scientific Refereed Research Journal*, 6(1), 448–466.
- Asiedu, W., & Quainoo, R. (2023b). Toward maintenance free wireless infrastructure: Simulating performance and reliability in large scale intermittently powered IoT networks. *Gyanshauryam, International Scientific Refereed Research Journal*, 6(1), 489–510.
- Asiedu, W., & Quainoo, R. (2024). Rethinking energy, reliability, and latency trade-offs in green communication for next generation IoT. *International Journal of Multidisciplinary Research and Growth Evaluation*, 5(6), 1987–1994.
- Asiedu, W., & Quainoo, R. (2025). GleanCast: Epoch-aligned reliable broadcast in batteryless intermittent sensor networks. *International Journal of Engineering and Modern Technology*, 11(12), 205–218.
- Asiedu, W., Quainoo, R., & Asiedu, A. (2025). A conceptual framework for characterizing fundamental energy, reliability, and latency trade-offs in green communication paradigms for next-generation IoT. *International Journal of Advanced Multidisciplinary Research and Studies*, 5(6), 2447–2453.
- Asiedu, W., Quainoo, R., & Asiedu, A. (2026). Predictive power management for intermittent IoT devices using lightweight machine learning under uncertain energy harvest. *Gulf Journal of Engineering and Technology*, 2(4), 108–118.
- Awad, A. I. (2016). From classical methods to animal biometrics: A review on cattle identification and tracking. *Computers and Electronics in Agriculture*, 123, 423–435.
- Aye, P. A., & Tawose, O. M. (2015). Acceptability and utilization of graded levels of *Gmelina arborea* leaves and cassava peels concentrate by West African dwarf sheep. *International Journal of Advances in Agriculture*, 4(2), 415–422.
- Aye, P. A., & Tawose, O. M. (2016). Physiological responses of West African dwarf sheep fed graded levels of *Gmelina arborea* leaf and cassava peel concentrates under different management systems. *Agriculture and Biology Journal of North America*, 7(4), 185–195.
- Ayinde, M. K. (2025). West Nile virus dynamics in American robins (*Turdus migratorius*; AMRO) of Illinois [Unpublished thesis]. Western Illinois University.
- Ayinde, M. K., Penddah, A. L., Mensah, G. K., & Olorunkosebi, M. T. (2024). Environmental microplastics as vectors for avian virus transmission and their implications for food safety within a one health framework. *Journal of Current Research in Food Science*, 5(2), 210–219.
- Azeez, L. O., & Badmus, O. (2024). Predictive analytical framework for identifying vapor intrusion risks across urban redevelopment zones. *International Journal of Scientific Research in Humanities and Social Sciences*, 1(1), 524–555.
- Badmus, O. (2025). Integrated predictive and remote-sensing framework for early warning and regulatory compliance in environmentally sensitive urban zones. *IIARD International Journal of Geography and Environmental Management*, 11(12), 212–239.
- Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. (2008). Speeded-up robust features (SURF). *Computer Vision and Image Understanding*, 110(3), 346–359.
- Beery, S., Morris, D., & Yang, S. (2019). Efficient pipeline for camera trap image review. *arXiv preprint arXiv:1907.06772*.
- Beery, S., Van Horn, G., & Perona, P. (2018). Recognition in terra incognita. *Proceedings of the European Conference on Computer Vision*, 456–473.
- Bello, K. A., Adama, A., Tawose, O. M., & Bolaji, B. O. (2022). Development and performance evaluation of a poultry bird de-feathering machine. *FUOYE Journal of Engineering and Technology*, 7(4).
- Bello, K. A., Oyelaran, O. A., Tawose, O. M., Omoyi, C. O., & Yussouff, A. A. (2024). Development of a gum production machine from *Manihot esculenta* waste peels. *FUOYE Journal of Innovation Science and Technology*, 2(1).

- Bello, K. A., Tawose, O. M., Adama, A., & Bolaji, B. O. (2022). Factor analysis of poultry birds de-feathering machines. *FUOYE Journal of Engineering and Technology*, 7(3).
- Berger-Wolf, T. Y., Rubenstein, D. I., Stewart, C. V., Holmberg, J. A., Parham, J., Menon, S., Crall, J., Van Oast, J., Kiciman, E., & Joppa, L. (2017). Wildbook: Crowdsourcing, computer vision, and data science for conservation. *arXiv preprint arXiv:1710.08880*.
- Bernardin, K., & Stiefelhagen, R. (2008). Evaluating multiple object tracking performance: The CLEAR MOT metrics. *EURASIP Journal on Image and Video Processing*, 2008, 246309.
- Bewley, A., Ge, Z., Ott, L., Ramos, F., & Upcroft, B. (2016). Simple online and realtime tracking. *Proceedings of the IEEE International Conference on Image Processing*, 3464–3468.
- Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). YOLOv4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*.
- Bogucki, R., Cygan, M., Khan, C. B., Klimek, M., Milczek, J. K., & Mucha, M. (2019). Applying deep learning to right whale photo identification. *Conservation Biology*, 33(3), 676–684.
- Bolger, D. T., Morrison, T. A., Vance, B., Lee, D., & Farid, H. (2012). A computer-assisted system for photographic mark-recapture analysis. *Methods in Ecology and Evolution*, 3(5), 813–822.
- Burton, A. C., Neilson, E., Moreira, D., Ladle, A., Steenweg, R., Fisher, J. T., Bayne, E., & Boutin, S. (2015). Wildlife camera trapping: A review and recommendations for linking surveys to ecological processes. *Journal of Applied Ecology*, 52(3), 675–685.
- Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., & Zagoruyko, S. (2020). End-to-end object detection with transformers. *Proceedings of the European Conference on Computer Vision*, 213–229.
- Caron, M., Touvron, H., Misra, I., Jégou, H., Mairal, J., Bojanowski, P., & Joulin, A. (2021). Emerging properties in self-supervised vision transformers. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 9650–9660.
- Cheeseman, T., Southerland, K., Park, J., et al. (2022). Advanced image recognition: A fully automated, high-accuracy photo-identification matching system for humpback whales. *Mammalian Biology*, 102, 915–929.
- Chen, G., Han, T. X., He, Z., Kays, R., & Forrester, T. (2014). Deep convolutional neural network based species recognition for wild animal monitoring. *Proceedings of the IEEE International Conference on Image Processing*, 858–862.
- Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020). A simple framework for contrastive learning of visual representations. *Proceedings of the International Conference on Machine Learning*, 1597–1607.
- Chopra, S., Hadsell, R., & LeCun, Y. (2005). Learning a similarity metric discriminatively, with application to face verification. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 539–546.
- Christin, S., Hervet, É., & Lecomte, N. (2019). Applications for deep learning in ecology. *Methods in Ecology and Evolution*, 10(10), 1632–1644.
- Clapham, M., Miller, E., Nguyen, M., & Darimont, C. T. (2020). Automated facial recognition for wild bears. *Ecology and Evolution*, 10(23), 12883–12892.
- Crall, J. P., Stewart, C. V., Berger-Wolf, T. Y., Rubenstein, D. I., & Sundaresan, S. R. (2013). HotSpotter: Patterned species instance recognition. *Proceedings of the IEEE Workshop on Applications of Computer Vision*, 230–237.
- Dagodzo, D. (2018a). A conceptual framework for UAV integration into national power grid inspection programs. *Iconic Research and Engineering Journals*, 2(5), 391–412.
- Dagodzo, D. (2018b). A review of UAV applications in electrical transmission line inspection: Methods, technologies, and challenges. *Iconic Research and Engineering Journals*, 2(6), 234–254.
- Dagodzo, D. (2018c). Hydrothermal alteration interpretation of LANDSAT and ASTER data. Lambert Academic Publishing.
- Dagodzo, D., & Ahiaeké Patrick, M. C. (2020). UAV-based pipeline and corridor monitoring: A review of current practices and emerging technologies. *Iconic Research and Engineering Journals*, 3(10), 574–597.
- Dagodzo, D., & Ahiaeké Patrick, M. C. (2021a). A review of GIS applications in utility asset management and infrastructure planning. *Iconic Research and Engineering Journals*, 5(3), 468–492.
- Dagodzo, D., & Ahiaeké Patrick, M. C. (2021b). An integrated framework for UAV, LiDAR, and GIS in infrastructure corridor management. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 7(5), 497–524.
- Dagodzo, D., & Ahiaeké Patrick, M. C. (2022). A review of right-of-way encroachment detection methods using geospatial technologies. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 8(1), 638–667.
- Dagodzo, D., & Ahiaeké Patrick, M. C. (2023a). A framework for integrating drone operations into enterprise GIS systems for utility companies. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 9(4), 924–957.
- Dagodzo, D., & Ahiaeké Patrick, M. C. (2023b). A review of UAV regulatory frameworks in developing economies: Progress, gaps, and recommendations. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 9(4), 958–994.
- Dagodzo, D., & Ahiaeké Patrick, M. C. (2025). A framework for national-scale UAV deployment in power infrastructure: Lessons from developing economies. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 11(4), 779–824.
- Dagodzo, D., Ahiaeké Patrick, M. C., & Aliliele, C. (2022). AI and deep learning for vegetation classification in power corridor management: A review. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 8(1), 668–698.

- Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 886–893.
- Deb, D., Wiper, S., Gong, S., Shi, Y., Tymoszek, C., Fletcher, A., & Jain, A. K. (2018). Face recognition: Primates in the wild. *Proceedings of the IEEE International Conference on Biometrics Theory, Applications and Systems*, 1–10.
- Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 248–255.
- Deng, J., Guo, J., Xue, N., & Zafeiriou, S. (2019). ArcFace: Additive angular margin loss for deep face recognition. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4690–4699.
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., et al. (2021). An image is worth 16x16 words: Transformers for image recognition at scale. *Proceedings of the International Conference on Learning Representations*.
- Ejofodomi, O. A., Gideon, E. N., Oladipo, G. O., & Oshomah, E. R. (2014). Automated detection of architectural distortion in mammograms using template matching. *International Journal of Biomedical Science and Engineering*, 2(1), 1–6.
- Ekeocha, A. H., Aganga, A. A., Adejoro, F. A., Oyebanji, A., Oluwadele, J. F., & Tawose, O. M. (2021). Phenotypic characteristics of indigenous chickens in selected regions of Nigeria. *Journal of World's Poultry Research*, 11(3), 352–358.
- Everingham, M., Van Gool, L., Williams, C. K. I., Winn, J., & Zisserman, A. (2010). The PASCAL visual object classes (VOC) challenge. *International Journal of Computer Vision*, 88(2), 303–338.
- Foster, R. J., & Harmsen, B. J. (2012). A critique of density estimation from camera-trap data. *Journal of Wildlife Management*, 76(2), 224–236.
- Freytag, A., Rodner, E., Simon, M., Loos, A., Köhl, H. S., & Denzler, J. (2016). Chimpanzee faces in the wild: Log-Euclidean CNNs for predicting identities and attributes of primates. *Proceedings of the German Conference on Pattern Recognition*, 51–63.
- Ge, Z., Liu, S., Wang, F., Li, Z., & Sun, J. (2021). YOLOX: Exceeding YOLO series in 2021. *arXiv preprint arXiv:2107.08430*.
- Girshick, R. (2015). Fast R-CNN. *Proceedings of the IEEE International Conference on Computer Vision*, 1440–1448.
- Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 580–587.
- Glover-Kapfer, P., Soto-Navarro, C. A., & Wearn, O. R. (2019). Camera-trapping version 3.0: Current constraints and future priorities for development. *Remote Sensing in Ecology and Conservation*, 5(3), 209–223.
- Gomez Villa, A., Salazar, A., & Vargas, F. (2017). Towards automatic wild animal monitoring: Identification of animal species in camera-trap images using very deep convolutional neural networks. *Ecological Informatics*, 41, 24–32.
- Grill, J.-B., Strub, F., Altché, F., et al. (2020). Bootstrap your own latent: A new approach to self-supervised learning. *Advances in Neural Information Processing Systems*, 33, 21271–21284.
- Hadsell, R., Chopra, S., & LeCun, Y. (2006). Dimensionality reduction by learning an invariant mapping. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1735–1742.
- Han, S., Mao, H., & Dally, W. J. (2016). Deep compression: Compressing deep neural networks with pruning, trained quantization and Huffman coding. *Proceedings of the International Conference on Learning Representations*.
- Hansen, M. F., Smith, M. L., Smith, L. N., Salter, M. G., Baxter, E. M., Farish, M., & Grieve, B. (2018). Towards on-farm pig face recognition using convolutional neural networks. *Computers in Industry*, 98, 145–152.
- He, K., Fan, H., Wu, Y., Xie, S., & Girshick, R. (2020). Momentum contrast for unsupervised visual representation learning. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 9729–9738.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778.
- Hermans, A., Beyer, L., & Leibe, B. (2017). In defense of the triplet loss for person re-identification. *arXiv preprint arXiv:1703.07737*.
- Hiby, L., Lovell, P., Patil, N., Kumar, N. S., Gopalaswamy, A. M., & Karanth, K. U. (2009). A tiger cannot change its stripes: Using a three-dimensional model to match images of living tigers and tiger skins. *Biology Letters*, 5(3), 383–386.
- Hinton, G., Vinyals, O., & Dean, J. (2015). Distilling the knowledge in a neural network. *arXiv preprint arXiv:1503.02531*.
- Howard, A. G., Zhu, M., Chen, B., et al. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*.
- Idika, C. N., Salami, E. O., Ijiga, O. M., & Enyejo, L. A. (2021). Deep learning driven malware classification for cloud-native microservices in edge computing architectures. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 7(4).
- Jackson, R. M., Roe, J. D., Wangchuk, R., & Hunter, D. O. (2006). Estimating snow leopard population abundance using photography and capture-recapture techniques. *Wildlife Society Bulletin*, 34(3), 772–781.
- Jacob, B., Kligys, S., Chen, B., et al. (2018). Quantization and training of neural networks for efficient integer-arithmetic-only inference. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2704–2713.
- Johnson, J., Douze, M., & Jégou, H. (2021). Billion-scale similarity search with GPUs. *IEEE Transactions on Big Data*, 7(3), 535–547.

- Karanth, K. U. (1995). Estimating tiger *Panthera tigris* populations from camera-trap data using capture-recapture models. *Biological Conservation*, 71(3), 333–338.
- Karanth, K. U., & Nichols, J. D. (1998). Estimation of tiger densities in India using photographic captures and recaptures. *Ecology*, 79(8), 2852–2862.
- Kelly, M. J. (2001). Computer-aided photograph matching in studies using individual identification: An example from Serengeti cheetahs. *Journal of Mammalogy*, 82(2), 440–449.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25, 1097–1105.
- Kumar, S., Singh, S. K., Singh, R., & Singh, A. K. (2018). Deep learning framework for recognition of cattle using muzzle point image pattern. *Measurement*, 116, 1–17.
- Körschens, M., Barz, B., & Denzler, J. (2018). Towards automatic identification of elephants in the wild. *arXiv preprint arXiv:1812.04418*.
- Kühl, H. S., & Burghardt, T. (2013). Animal biometrics: Quantifying and detecting phenotypic appearance. *Trends in Ecology & Evolution*, 28(7), 432–441.
- Lamidi, O. B. A., & Olamide, A. (2018). Spatially explicit risk modeling framework for tracking subsurface contaminant migration in data-limited remediation sites. *Iconic Research and Engineering Journals*, 2(6), 178–198.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444.
- Li, S., Li, J., Tang, H., Qian, R., & Lin, W. (2020). ATRW: A benchmark for Amur tiger re-identification in the wild. *Proceedings of the 28th ACM International Conference on Multimedia*, 2590–2598.
- Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017a). Feature pyramid networks for object detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2117–2125.
- Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017b). Focal loss for dense object detection. *Proceedings of the IEEE International Conference on Computer Vision*, 2980–2988.
- Lin, T.-Y., Maire, M., Belongie, S., et al. (2014). Microsoft COCO: Common objects in context. *Proceedings of the European Conference on Computer Vision*, 740–755.
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., & Berg, A. C. (2016). SSD: Single shot multibox detector. *Proceedings of the European Conference on Computer Vision*, 21–37.
- Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., & Guo, B. (2021). Swin transformer: Hierarchical vision transformer using shifted windows. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 10012–10022.
- Loos, A., & Ernst, A. (2013). An automated chimpanzee identification system using face detection and recognition. *EURASIP Journal on Image and Video Processing*, 2013, 49.
- Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2), 91–110.
- Luiten, J., Osep, A., Dendorfer, P., Torr, P., Geiger, A., Leal-Taixé, L., & Leibe, B. (2021). HOTA: A higher order metric for evaluating multi-object tracking. *International Journal of Computer Vision*, 129(2), 548–578.
- Luo, H., Gu, Y., Liao, X., Lai, S., & Jiang, W. (2019). Bag of tricks and a strong baseline for deep person re-identification. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 1487–1495.
- Malkov, Y. A., & Yashunin, D. A. (2020). Efficient and robust approximate nearest neighbor search using hierarchical navigable small world graphs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(4), 824–836.
- Mayo, W., Ogbale, J. I., Okoruwa, P. O., & Babatope, O. M. (2021). Designing an AI-predictive maintenance model for e-commerce systems using machine learning and cloud analytics. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 7(5), 416–440.
- Mbonu, I. S., Aliliele, C., Iwuanyanwu, U., & Oluoha, O. M. (2018). A conceptual framework for legal and ethical risk modeling in enterprise data protection governance systems. *Iconic Research and Engineering Journals*, 2(2), 207–226.
- Mensah, G. K., Olorunkosebi, M. T., Penddah, A. L., & Ayinde, M. K. (2026). Advanced nanomaterials for mitigating foodborne pathogens: Efficacy, mechanisms, and safety considerations. *International Journal of Multidisciplinary Evolutionary Research*, 7(1), 17–28.
- Miele, V., Dussert, G., Spataro, B., Chamaillé-Jammes, S., Allainé, D., & Bonenfant, C. (2021). Revisiting animal photo-identification using deep metric learning and network analysis. *Methods in Ecology and Evolution*, 12(4), 690–702.
- Milan, A., Leal-Taixé, L., Reid, I., Roth, S., & Schindler, K. (2016). MOT16: A benchmark for multi-object tracking. *arXiv preprint arXiv:1603.00831*.
- Musgrave, K., Belongie, S., & Lim, S.-N. (2020). A metric learning reality check. *Proceedings of the European Conference on Computer Vision*, 681–699.
- Neethirajan, S. (2020). The role of sensors, big data and machine learning in modern animal farming. *Sensing and Bio-Sensing Research*, 29, 100367.
- Nepovimnykh, E., Eerola, T., & Kälviäinen, H. (2020). Siamese network based pelage pattern matching for ringed seal re-identification. *Proceedings of the IEEE Winter Conference on Applications of Computer Vision Workshops*, 25–34.
- Norouzzadeh, M. S., Morris, D., Beery, S., Joshi, N., Jojic, N., & Clune, J. (2021). A deep active learning system for species identification and counting in camera trap images. *Methods in Ecology and Evolution*, 12(1), 150–161.

- Norouzzadeh, M. S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M. S., Packer, C., & Clune, J. (2018). Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proceedings of the National Academy of Sciences*, 115(25), E5716–E5725.
- Nwakamma, S., Ojukwu, J., & Oyesiji, S. O. (2024a). LLM-as-a-judge for automated model governance: A review of methods, biases, and release-gating practices. *World Journal of Innovation and Modern Technology*, 8(6), 185–216.
- Nwakamma, S., Ojukwu, J., & Oyesiji, S. O. (2024b). Securing agentic AI enterprise workflows against prompt injection, tool poisoning, memory manipulation, and excessive agency threats. *World Journal of Innovation and Modern Technology*, 8(6), 217–248.
- Nwakamma, S., Ojukwu, J., & Oyesiji, S. O. (2025). On-device multimodal small language models for privacy-preserving edge intelligence through quantization, pruning, distillation, and energy-efficient inference. *World Journal of Innovation and Modern Technology*, 9(12), 271–302.
- Nwakamma, S., Ojukwu, J., & Oyesiji, S. O. (2026). Crypto-agile migration to hybrid post-quantum cryptography in enterprise TLS, PKI, VPN, and cloud-native systems and services. *Gulf Journal of Computer Sciences*, 1(1).
- O'Connell, A. F., Nichols, J. D., & Karanth, K. U. (Eds.). (2011). *Camera traps in animal ecology: Methods and analyses*. Springer.
- Ogunboye, I., Adebayo, I. P. S., Anioke, S. C., Egwuatu, E. C., Ajala, C. F., & Awuah, S. B. (2023). Enhancing Nigeria's health surveillance system: A data-driven approach to epidemic preparedness and response. *World Journal of Advanced Research and Reviews*.
- Ojukwu, J., Oyesiji, S. O., & Nwakamma, S. (2023). Cyber-physical ransomware defense in operational technology using physics-informed detection, digital twins, network telemetry, and resilient recovery. *International Journal of Computer Science and Mathematical Theory*, 9(5), 193–228.
- Ojukwu, J., Oyesiji, S. O., & Nwakamma, S. (2025). WebAssembly component model and WASI for portable cloud-to-edge software: Language interoperability, security, performance, and container comparisons. *International Journal of Computer Science and Mathematical Theory*, 11(12), 232–263.
- Olorunkosebi, M. T., Ayinde, M. K., & Mensah, G. K. (2023). Integrating artificial intelligence and advanced biosensors for rapid and sensitive detection of antimicrobial resistant foodborne pathogens. *Journal of Advances in Microbiology Research*, 4(2), 175–182.
- Oluwadele, J. F., Ekeocha, A. H., Aganga, A. A., Tawose, O. M., Odumboni, A., Ofodome, C. I., & Akinlabi, E. Y. (2024). Effects of feeding *Pennisetum purpureum* silage supplemented with selected farm residues on growth performance and meat quality of West African dwarf rams. *SVU-International Journal of Agricultural Sciences*, 6(3), 190–196.
- Oluwadele, J. F., Tawose, O. M., & Adetumbi, T. (2024). Evaluation of feed intake, growth performance, and carcass characteristics of West African dwarf rams fed mango leaves, Napier grass, neem seed cake, and concentrate for fattening. *EUREKA: Life Sciences*, (4), 11–19.
- Oluwadele, J. F., Tawose, O. M., Ekeocha, A. H., Akinlabi, E. Y., Adeitan, O. O., & Bello, K. A. (2023). Physiological indicators and stress index of scavenging chickens at LAFARGE (Ewekoro) and DANGOTE (Ibese) cement factory areas of Ogun State, Nigeria. *Journal of Austrian Society of Agricultural Economics*, 19(4).
- Ominyi, M., & Anichukwueze, C. C. (2024). Conceptualizing responsible AI governance structures for automated decision systems in regulated industries. *International Journal of Social Sciences and Management Research*, 10(11), 403–430.
- Ominyi, M., Anichukwueze, C. C., & Uzougbo, N. S. (2026). Conceptualizing enterprise AI governance functions as institutional standards for auditable and defensible risk management. *Journal of Accounting and Financial Management*, 12(7), 326–356.
- Oparah, O. S., Ezeh, F. E., Olatunji, G. I., & Ajayi, O. O. (2023). Framework for designing national real-time disease surveillance dashboards for public health stakeholders. *Shodhshauryam, International Scientific Refereed Research Journal*, 6(1), 208–227.
- Oyesiji, S. O., Nwakamma, S., & Ojukwu, J. (2023). Parameter-efficient multilingual adaptation of on-device language models: A survey. *International Journal of Engineering and Modern Technology*, 9(3), 287–320.
- Oyesiji, S. O., Nwakamma, S., & Ojukwu, J. (2024). A conceptual framework for context-aware data curation in fine-tuning pipelines. *International Journal of Engineering and Modern Technology*, 10(11), 197–228.
- Oyesiji, S. O., Nwakamma, S., & Ojukwu, J. (2025). Compressing recommender and ranking models for edge deployment: A survey. *International Journal of Engineering and Modern Technology*, 11(12), 205–237.
- Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359.
- Parham, J., Stewart, C., Crall, J., Rubenstein, D., Holmberg, J., & Berger-Wolf, T. (2018). An animal detection pipeline for identification. *Proceedings of the IEEE Winter Conference on Applications of Computer Vision*, 1075–1083.
- Qiao, Y., Su, D., Kong, H., Sukkari, S., Lomax, S., & Clark, C. (2019). Individual cattle identification using a deep learning based framework. *IFAC-PapersOnLine*, 52(30), 318–323.
- Quainoo, R., & Ogundapo, O. (2026a). A system-level power behavior model for Bluetooth and Wi-Fi coexistence in dual-mode wireless devices. *International Journal of Computer Science and Mathematical Theory*, 12(2), 298–358.
- Quainoo, R., & Ogundapo, O. (2026b). Wireless system-on-chip performance in next-generation Internet of Things devices: A systematic review of integrated transceiver testing methodologies. *World Journal of Innovation and Modern Technology*, 10(5), 76–126.
- Quainoo, R., Ogundapo, O., & Asiedu, W. A. (2024). Modeling the impact of impedance mismatch on wireless link performance: A framework for prototype development and system optimization. *International Journal of Computer Science and Mathematical Theory*, 10(2), 62–116.
- Quainoo, R., Ogundapo, O., & Asiedu, W. A. (2025a). Predicting throughput degradation in Wi-Fi 6 networks under varying channel conditions: A physical layer performance model. *International Journal of Engineering and Modern Technology*, 11(12), 205–264.
- Quainoo, R., Ogundapo, O., & Asiedu, W. A. (2025b). Test automation in wireless hardware engineering: A comprehensive review of scripting frameworks and instrument control strategies. *International Journal of Engineering and Modern Technology*, 11(10), 405–464.

- Ravoor, P. C., & Sudarshan, T. S. B. (2020). Deep learning methods for multi-species animal re-identification and tracking: A survey. *Computer Science Review*, 38, 100289.
- Redmon, J., & Farhadi, A. (2017). YOLO9000: Better, faster, stronger. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 7263–7271.
- Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improvement. *arXiv preprint arXiv:1804.02767*.
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 779–788.
- Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. *Advances in Neural Information Processing Systems*, 28, 91–99.
- Ristani, E., Solera, F., Zou, R., Cucchiara, R., & Tomasi, C. (2016). Performance measures and a data set for multi-target, multi-camera tracking. *Proceedings of the European Conference on Computer Vision Workshops*, 17–35.
- Royle, J. A., Chandler, R. B., Sollmann, R., & Gardner, B. (2013). *Spatial capture-recapture*. Academic Press.
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). MobileNetV2: Inverted residuals and linear bottlenecks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4510–4520.
- Schneider, S., Taylor, G. W., & Kremer, S. C. (2020). Similarity learning networks for animal individual re-identification: Beyond the capabilities of a human observer. *Proceedings of the IEEE Winter Conference on Applications of Computer Vision Workshops*, 44–52.
- Schneider, S., Taylor, G. W., Linquist, S., & Kremer, S. C. (2019). 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.
- Schofield, D., Nagrani, A., Zisserman, A., Hayashi, M., Matsuzawa, T., Biro, D., & Carvalho, S. (2019). Chimpanzee face recognition from videos in the wild using deep learning. *Science Advances*, 5(9), eaaw0736.
- Schroff, F., Kalenichenko, D., & Philbin, J. (2015). FaceNet: A unified embedding for face recognition and clustering. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 815–823.
- Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. *Proceedings of the IEEE International Conference on Computer Vision*, 618–626.
- Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge computing: Vision and challenges. *IEEE Internet of Things Journal*, 3(5), 637–646.
- Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, 6, 60.
- Silver, S. C., Ostro, L. E. T., Marsh, L. K., et al. (2004). The use of camera traps for estimating jaguar *Panthera onca* abundance and density using capture/recapture analysis. *Oryx*, 38(2), 148–154.
- Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *Proceedings of the International Conference on Learning Representations*.
- Steenweg, R., Hebblewhite, M., Kays, R., et al. (2017). Scaling-up camera traps: Monitoring the planet's biodiversity with networks of remote sensors. *Frontiers in Ecology and the Environment*, 15(1), 26–34.
- Sun, Y., Zheng, L., Yang, Y., Tian, Q., & Wang, S. (2018). Beyond part models: Person retrieval with refined part pooling. *Proceedings of the European Conference on Computer Vision*, 480–496.
- Swanson, A., Kosmala, M., Lintott, C., Simpson, R., Smith, A., & Packer, C. (2015). Snapshot Serengeti, high-frequency annotated camera trap images of 40 mammalian species in an African savanna. *Scientific Data*, 2, 150026.
- Szegedy, C., Liu, W., Jia, Y., et al. (2015). Going deeper with convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1–9.
- Tabak, M. A., Norouzzadeh, M. S., Wolfson, D. W., et al. (2019). Machine learning to classify animal species in camera trap images: Applications in ecology. *Methods in Ecology and Evolution*, 10(4), 585–590.
- Tan, M., & Le, Q. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. *Proceedings of the International Conference on Machine Learning*, 6105–6114.
- Tan, M., Pang, R., & Le, Q. V. (2020). EfficientDet: Scalable and efficient object detection. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 10781–10790.
- Tawose, O. M., & Oluwadele, J. F. (2025). Comparative analysis of the nutritional potentials of selected tuber peel meals as feed supplements for ruminant animals. *Nigerian Journal of Agriculture and Agricultural Technology*, 5(3A).
- Tawose, O. M., Ekeocha, A. H., & Oluwadele, J. F. (2023). Nutritional quality and utilization of water hyacinth-cassava peels silage by West African dwarf (WAD) goats. *FUOYE Journal of Agriculture and Human Ecology*, 6(1), 35–45.
- Tawose, O. M., Ekeocha, A. H., & Oluwadele, J. F. (2024). Haematological and biochemical responses of West African dwarf goats fed water hyacinth-cassava peels silage. *Nigerian Journal of Animal Production*, 50(3).
- Tawose, O. M., Oluwadele, J. F., Ekeocha, A. H., Odumboni, A. A., & Egbeyemi, O. S. (2023). Aflatoxin detection and quantification in poultry feeds available in selected areas in Ekiti State. *FUOYE Journal of Agriculture and Human Ecology*, 7(2).

- Tonoyan, A., Dada, O., & Ayivi-Donkor, S. S. (2022). Advances in geospatial analysis: Predictive analytics methods for store location selection and market entry return on investment. *International Journal of Marketing and Communication Studies*, 6(2), 78–105.
- Town, C., Marshall, A., & Sethasathien, N. (2013). Manta Matcher: Automated photographic identification of manta rays using keypoint features. *Ecology and Evolution*, 3(7), 1902–1914.
- Tuia, D., Kellenberger, B., Beery, S., et al. (2022). Perspectives in machine learning for wildlife conservation. *Nature Communications*, 13, 792.
- van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9, 2579–2605.
- Vidal, M., Wolf, N., Rosenberg, B., Harris, B. P., & Mathis, A. (2021). Perspectives on individual animal identification from biology and computer vision. *Integrative and Comparative Biology*, 61(3), 900–916.
- Wang, C.-Y., Bochkovskiy, A., & Liao, H.-Y. M. (2023). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 7464–7475.
- Wang, H., Wang, Y., Zhou, Z., Ji, X., Gong, D., Zhou, J., Li, Z., & Liu, W. (2018). CosFace: Large margin cosine loss for deep face recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 5265–5274.
- Wearn, O. R., & Glover-Kapfer, P. (2019). Snap happy: Camera traps are an effective sampling tool when compared with alternative methods. *Royal Society Open Science*, 6(3), 181748.
- Weideman, H. J., Jablons, Z. M., Holmberg, J., et al. (2017). Integral curvature representation and matching algorithms for identification of dolphins and whales. *Proceedings of the IEEE International Conference on Computer Vision Workshops*, 2831–2839.
- Weinstein, B. G. (2018). A computer vision for animal ecology. *Journal of Animal Ecology*, 87(3), 533–545.
- Wen, Y., Zhang, K., Li, Z., & Qiao, Y. (2016). A discriminative feature learning approach for deep face recognition. *Proceedings of the European Conference on Computer Vision*, 499–515.
- Willi, M., Pitman, R. T., Cardoso, A. W., et al. (2019). Identifying animal species in camera trap images using deep learning and citizen science. *Methods in Ecology and Evolution*, 10(1), 80–91.
- Wojke, N., Bewley, A., & Paulus, D. (2017). Simple online and realtime tracking with a deep association metric. *Proceedings of the IEEE International Conference on Image Processing*, 3645–3649.
- Wurtz, K., Camerlink, I., D'Eath, R. B., et al. (2019). Recording behaviour of indoor-housed farm animals automatically using machine vision technology: A systematic review. *PLoS ONE*, 14(12), e0226669.
- Ye, M., Shen, J., Lin, G., Xiang, T., Shao, L., & Hoi, S. C. H. (2022). Deep learning for person re-identification: A survey and outlook. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(6), 2872–2893.
- Yu, X., Wang, J., Kays, R., Jansen, P. A., Wang, T., & Huang, T. (2013). Automated identification of animal species in camera trap images. *EURASIP Journal on Image and Video Processing*, 2013, 52.
- Zhang, Y., Sun, P., Jiang, Y., et al. (2022). ByteTrack: Multi-object tracking by associating every detection box. *Proceedings of the European Conference on Computer Vision*, 1–21.
- Zheng, L., Shen, L., Tian, L., Wang, S., Wang, J., & Tian, Q. (2015). Scalable person re-identification: A benchmark. *Proceedings of the IEEE International Conference on Computer Vision*, 1116–1124.
- Zheng, L., Yang, Y., & Hauptmann, A. G. (2016). Person re-identification: Past, present and future. *arXiv preprint arXiv:1610.02984*.
- Zhou, Z., Chen, X., Li, E., Zeng, L., Luo, K., & Zhang, J. (2019). Edge intelligence: Paving the last mile of artificial intelligence with edge computing. *Proceedings of the IEEE*, 107(8), 1738–1762.
- Zhuang, F., Qi, Z., Duan, K., et al. (2021). A comprehensive survey on transfer learning. *Proceedings of the IEEE*, 109(1), 43–76.