

ORIGINAL RESEARCH

Open access

Continual Multi-Domain Adaptation for Unsupervised Holstein Cow Re-Identification

Amira Khalil^{1*}, Sara Fathy¹, Mahmoud Adel², Hany Nabil³, Rania Mostafa⁴

Abstract

In the dairy industry, maintaining the welfare of every cow and curbing financial losses requires individual-level surveillance, which is achieved through cow Re-Identification (Re-ID). Computer vision-based Re-ID hinges on visually distinguishing characteristics, such as the specific coat patterns seen in breeds like Holstein. However, manually labelling every cow across all farms is financially unfeasible. We aim to formulate Re-ID strategies that work on both annotated and non-annotated farms, accommodating new individuals and varied settings. Unsupervised Domain Adaptation (UDA) approaches aim to close this divide by transferring knowledge from labelled source domains to unlabelled target domains. However, they have largely been tailored to person and vehicle Re-ID scenarios. Our research proposes Cumulative Unsupervised Multi-Domain Adaptation (CUMDA) to address the challenges of constrained identity variation and dissimilar farm appearances. CUMDA gathers knowledge across all domains, reinforcing specialization in familiar domains while strengthening generalization to unknown ones. Our offerings include a CUMDA technique that adapts to multiple unlabelled target domains while safeguarding source-domain efficacy, along with thorough cross-dataset testing across three cattle Re-ID datasets. These tests reveal substantial gains in source conservation, target-domain specialization, and adaptation to unseen domains.

Keywords Re-identification, Domain adaptation, Holstein cattle, Unsupervised learning, Monitoring

*Correspondence:

Amira Khalil

amira.khalil@alexu.edu.eg

¹ Department of Continual Domain Adaptation for Cow Re-ID, Faculty of Agriculture, Alexandria University, Alexandria, Egypt

² Department of Unsupervised Animal Re-Identification, Faculty of Agriculture, Cairo University, Cairo, Egypt

³ Department of Multi-Domain Adaptation for Livestock, Faculty of Agriculture, Mansoura University, Mansoura, Egypt

⁴ Department of Holstein Cow Identification, Faculty of Agriculture, Helwan University, Cairo, Egypt

Introduction

Historically, farmers have been responsible for overseeing their dairy cows' health and behavior. Spotting early cues of estrus or anomalous behavior in these creatures is vital, not only to ensure their overall welfare but also to reduce potential financial losses. Moving from hands-on monitoring to vision-based individual following creates a demand for exact cow identification within any farm. This elaborate operation is commonly referred to as cow Re-Identification (Re-ID). Here, we target cow Re-ID in the Holstein breed, which is distinguished by its visually conspicuous spot patterns. We also connect sightings of the same animal over a brief interval, meaning across a designated timeframe during which its outward appearance is expected to be stable.

Earlier, Re-ID represented a major theme in image understanding, particularly within pedestrian video observation contexts [1]. The core purpose of Re-ID here is to identify or recover individuals of interest across camera systems whose

viewing zones do not overlap. The rise of supervised learning, powered by deep Convolutional Neural Networks (CNNs) [2], has substantially elevated the results achievable in supervised person Re-ID [3].

This Re-ID model has moved beyond human tracking to include a wider set of practical uses, covering vehicle Re-ID for traffic observation [4] and animal Re-ID for livestock surveillance [5–8].

Supervised Re-ID calls for labor-intensive dataset tagging. In the particular case of cow Re-ID, the difficulty and expense of tagging each cow on every farm necessitate a Re-ID method that generalizes reliably across both labelled and unlabelled farms. Such versatility proves indispensable for introducing new cows into existing farms and for launching the technology in unfamiliar farm settings. Despite this, Re-ID suffers a sharp performance drop whenever the test image spread deviates from that encountered during training.

To meet these hurdles, UDA techniques have been formulated for Re-ID [9–11]. UDA adapts a model to a domain of study by using an annotated dataset from a separate domain (source) along with unlabelled data from the domain of interest (target). While studied in depth for person and vehicle Re-ID, UDA is especially applicable to cow Re-ID, given that annotating each cow across many farms is unworkable.

Nevertheless, the distinct constraints and requirements of cow Re-ID have driven further advances, which are the main focus of this article. In cow Re-ID, each farm constitutes a unique domain, defined by its specific collection of animals, camera setups, and surrounding conditions. Consequently, the problem of cross-domain performance decay becomes more relevant. Furthermore, each domain typically contains only a small number of animals, limiting the generalization power of Re-ID architectures trained on isolated datasets, as practiced in standard UDA.

To sharpen the model's ability to distinguish cows, we proposed an original architecture: Cumulative Unsupervised Multi-Domain Adaptation (CUMDA). CUMDA strives to learn Re-ID insights from data acquired from numerous farms (across varied domains), enabling finer specialization within known domains and stronger performance when encountering previously unseen domains.

To the best of our knowledge, current UDA Re-ID methods have mostly addressed multi-domain settings with multiple annotated source domains. No UDA Re-ID solution has been devised expressly for numerous unlabelled target domains. What is more, whereas the Domain Generalization construct [12] aims to create models effective on unseen target domains, it diverges from CUMDA. CUMDA presupposes familiarity with all target domains at training time, along with the availability of unlabelled data drawn from these domains, aiming to enhance specialization and cross-domain Re-ID performance by building knowledge. Even though generalization is not the chief goal, we foresee that a CUMDA model, through the accumulation of knowledge across varied target domains, can improve its generalization to unseen domains.

Also, customary UDA frameworks often emphasise maximising target-domain accuracy while losing or sidelining source-domain knowledge [11]. In real-world use, whenever new cameras are set up on a farm, adjusting the Re-ID model to those novel devices becomes desirable whilst maintaining strong results on those already in place—a feature we label “source conservation.”

To sum up, although present-day Re-ID approaches could be straightforwardly applied to cattle Re-ID, none of them entirely satisfy the peculiar limitations and practical requirements of cow Re-ID that justify opting for a CUMDA framework instead of a standard UDA route. Motivated by these practical factors and the need for cross-domain cattle Re-ID solutions, this paper sets out to engineer a cross-domain model specifically suited to the specific hurdles of cattle Re-ID across multiple farms. This topic has drawn little research. We intend to develop such an approach within the CUMDA framework, emphasizing source conservation alongside the accumulation of knowledge across multiple unsupervised target domains. The paper puts forward two principal contributions:

- A Source-Guided CUMDA method, capable of boosting cross-domain re-ID results on one or several target domains, while managing to accumulate knowledge from multiple domains and retain the source performance.
- Wide-ranging cross-dataset trials for CUMDA re-ID on 2 publicly available cow re-ID datasets and a private one.

Related Work

This study sits at the convergence of two research streams: cattle re-id and cross-domain re-id.

Cattle re-ID

The primary emphasis of Re-ID has been on pedestrians and vehicles [3]. A limited body of work addresses animal re-ID, concentrating, for instance, on Amur tiger re-ID [6, 8]. However, Amur tigers exhibit spot patterns on their fur that are visually quite distinct from those of cows, and the way these spots distort during motion is tied to their specific body structure. This observation leads us to consider that cows exhibit a sufficiently divergent appearance from Amur tigers to warrant treatment as a distinct category of concern for re-ID [8]. It is precisely this uniqueness of bovine appearance that has spurred dedicated investigations [13–17], exploring it through supervised learning or self-supervised learning paradigms [17] that presuppose the availability of tracklets. That said, neither supervised nor self-supervised learning lends itself well to training that integrates information from multiple farms, owing to the mounting cost of annotation under the supervised scheme or the inability to account for domain variations when handling multi-farm data within the self-supervised framework. To the best of our understanding, cattle re-ID has therefore never tackled the cross-domain setting, nor the multi-domain one.

Cross-domain person and vehicle re-ID via UDA paradigm

Cross-domain re-ID has been widely investigated for person and vehicle re-ID challenges. It can be categorized into two main families of techniques.

Domain-translation methods were among the earliest explored for cross-domain person re-ID. Among them, image-to-image translation approaches try to render source images in the visual style of the target domain [7, 18–23]. They build upon generative architectures such as CycleGAN [24]. During training, they are required to restyle images while preserving the identity class. Doing so enables carrying over the identity label and performing supervised re-ID training on source images that now bear a target-like appearance. Domain-invariant feature learning methods [25–30] impose direct constraints on the feature space of the trained model, for instance by aligning feature distributions of different domains, so that the re-ID model, shaped to be discriminative on the source domain, retains that discriminative power on the target domain within this domain-agnostic space.

Given the modest cross-domain gains delivered by domain-translation approaches, research interest has shifted toward pseudo-labeling approaches [31–37]. These involve assigning identity labels to target domain images by clustering features extracted via an initial encoder, which is generally trained in a supervised fashion for re-ID on the source domain. Pseudo-labeling strategies have made significant strides in cross-domain re-ID results [9–11]. In pursuit of still better outcomes, further mechanisms have been crafted to either produce higher-quality pseudo-labels or fortify the overall framework against label noise [11, 38–45].

The method introduced in the present paper likewise falls within the family of pseudo-labeling methods. In contrast to prior art, the focus here is on a previously unexamined object class for cross-domain re-ID—specifically, cross-domain cow re-ID—for which no UDA solution has ever been contemplated. Beyond that, whereas earlier pseudo-labeling studies for cross-domain person re-ID address a single target domain scenario, this work confronts a more arduous yet more practical problem, one that is more closely aligned with real-world cattle re-ID requirements: Cumulative Unsupervised Multi-Domain Adaptation (CUMDA). This cross-domain paradigm seeks to leverage knowledge from one or more

domains of interest (e.g., different cattle farms) to improve cross-domain re-ID performance across all domains (farms), including the model's source-domain re-ID capacity.

Materials and Methods

Lower part: proposed pseudo-labeling scheme for multi-target CUMDA re-ID. Black arrows trace the pseudo-labeling and training loop; gray arrows trace the clustering hyperparameter optimization steps. The setup assumes a collection of n target domains $T_1, \dots, T_n$ together with a source domain S . For each target, SOURCE CALIBRATION derives an associated labelled source validation set $\{x_{val}^{S1}\}, \dots, \{x_{val}^{Sn}\}$. Following FEATURE EXTRACTION across all source validation and target sets, SOURCE-GUIDED AUTO HP TUNING determines target-specific optimal hyperparameter (HP) values $\lambda^*1, \dots, \lambda^*n$ from the calibrated source validation sets by maximizing a clustering quality measure Q . Target-specific PSEUDO LABELING BY CLUSTERING is subsequently executed. TRAINING is then performed jointly on all pseudo-labelled target sets as well as on the labelled source domain, through minimization of LSG-CUMDA.

Our objective is to engineer a CUMDA re-ID technique that improves re-ID accuracy across all observed domains. In more concrete terms, the method is expected to:

- Specialize toward one or several target domains to improve performance on them
- Maintain solid performance on the source domain

We also expect such a model to transfer well to a novel, previously unseen target domain, as it should progressively build knowledge across multiple domains.

As depicted in the upper section of **Figure 1**, present-day pseudo-labeling methods are designed for UDA re-ID, aiming to enhance re-ID performance in a single target domain by relying solely on data from that domain. Accordingly, they must be redesigned to fulfill the objectives of CUMDA re-ID outlined above and to accommodate data from multiple domains. This section lays out the central components of our CUMDA re-ID method.

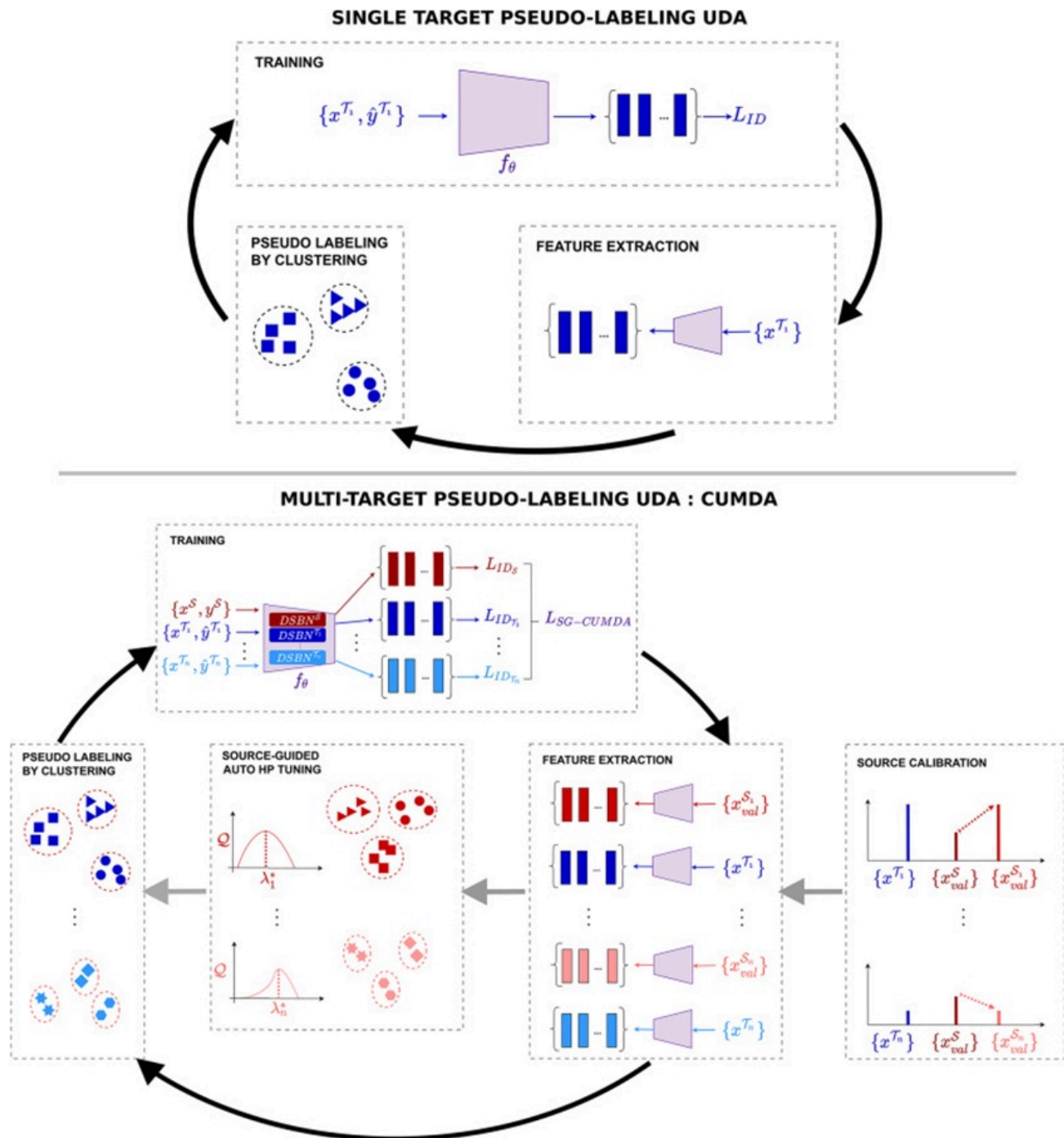


Figure 1. Upper part: pseudo-labeling paradigm for single-target UDA re-ID. Black arrows trace the pseudo-labeling and training loop. FEATURE EXTRACTION is applied to images $\{x^{T_1}\}$ belonging to the target domain T1 by means of a feature encoder f_θ . PSEUDO LABELING BY CLUSTERING produces pseudo-labels $\{\hat{y}^{T_1}\}$ over the clustered features. TRAINING proceeds on the pseudo-labelled target set $\{x^{T_1}, \hat{y}^{T_1}\}$ by minimizing L_{ID} .

The lower part introduces our CUMDA re-ID method, whose elements will be justified throughout this section.

General notations. We let $T_1, \dots, T_n, n \in \mathbb{N}$ be a collection of n target domains of interest, and S a source domain, from which a pool of labelled data S (drawn from S) and unlabelled data $T_1, \dots, T_n$ (drawn from $T_1, \dots, T_n$, the target domains) are at our disposal.

Pseudo-labeling by clustering

A feature encoder $f_\theta, \theta \in \mathbb{R}^p, p \in \mathbb{N}$ (commonly a CNN) undergoes training on the labelled source dataset T , through the minimisation of a re-ID loss function (e.g.: Classification Loss, Triplet Loss as found in [46], a blend of both as described in [47] ...) $L_{ID}(\theta, T)$, with respect to θ . Afterwards, a clustering function C_λ characterised by hyperparameters (HP) $\lambda \in \mathbb{R}^m, m \in \mathbb{N}$ is applied to forecast pseudo-labels for data samples inside each target set, drawing on the feature representation of their respective data. This yields pseudo-labelled target sets $\hat{T}_1, \dots, \hat{T}_n$:

$$\forall k \in [1, n], \hat{T}_k = C_\lambda(f_\theta, T_k) \quad (1)$$

We refer to this procedure, captured by Eq. 1, as Pseudo-Labeling by Clustering (PLC). The resulting pseudo-labels will subsequently define the loss function that steers learning toward the targets.

Source-guided CUMDA re-ID learning

What CUMDA re-ID aims to achieve is the dual capability to improve cross-domain performance on one or more target domains while preserving source re-ID performance. Motivated by the source-guided loss function introduced by [11] for single-target domain UDA re-ID, we formulate a fresh Source-Guided loss function tailored for CUMDA re-ID. Thus, f_θ undergoes fine-tuning via the minimisation of a Source-Guided CUMDA (SG-CUMDA) loss function LSG-CUMDA that bundles together all individual re-ID loss terms tied to each domain, as expressed below:

$$L_{SG-CUMDA}(\theta, (S, \hat{T}_1, \dots, \hat{T}_n)) = L_{ID}(\theta, S) + \sum_{k=1}^n L_{ID}(\theta, \hat{T}_k) \quad (2)$$

Alleviating the domain gap with domain-specific batch normalization

A common assertion is that the domain gap can significantly undercut performance [11]. The methodology we propose aims to counteract this at the level of the batch normalization layers.

Batch Normalization (batchnorm) is a widely used technique for accelerating and stabilising the training of deep neural networks by suppressing internal covariate shift [48]. It operates by normalising the batch after each convolutional or linear layer, using the batch-wise mean and variance of the activations, along with learnable affine parameters that rescale the features into batch-normalised outputs.

Prior investigations have highlighted the harmful consequences of calculating batchnorm statistics from data spanning multiple domains during training [49]. Domain-Specific Batch Normalization (DSBN) layers have proven valuable across a range of domain adaptation challenges, including UDA classification [50] and UDA re-ID [11]. The approach boils down to using domain-specific batch norm affine parameters and estimating domain-specific means and variances. All other network parameters remain shared and are applied uniformly across domains. Since f_θ is instantiated as a CNN, DSBN layers are inserted after each convolutional and fully-connected layer.

Improving pseudo-labels with multi-target automatic source-guided selection of Pseudo-labeling Hyperparameters

The effectiveness of pseudo-labeling UDA approaches hinges heavily on the quality of the assigned labels, which, in turn, depends on the careful adjustment of the clustering hyperparameter λ . In pedestrian and vehicle re-ID contexts, it has been established that the optimal λ value, λ^* , varies with both the distribution of the target dataset in feature space and its inherent statistics (e.g., the average number of shots per identity) [38]. A majority of studies centred on pedestrian and

vehicle UDA re-ID use a single set of values, tuned empirically on one particular cross-dataset pairing, for every other cross-dataset combination evaluated thereafter. It has been demonstrated that this practice can result in substantially inferior outcomes compared to those achieved by a properly fitted value [38].

In real-world cow re-ID, the unlabelled nature of the target makes it impossible to assemble a labelled validation set for tuning this parameter. What is more, the conventional λ value used in person re-ID may not transfer effectively to the cow re-ID setting, given the distinct traits of cow datasets (e.g., colour distributions, viewpoints). We therefore put forward an automated strategy for tuning λ directly from the labelled source data. To do so, we adopt the HyperParameters Automated by Source & Similarities (HyPASS) algorithm [38]. HyPASS was originally conceived and validated for single-target domain pedestrian and vehicle UDA re-ID. It optimises clustering hyperparameters using labelled data from the source validation split. In practice, HyPASS performs model selection over $C\lambda$ to identify the value λ^* defined as $\lambda^* = \operatorname{argmax}_{\lambda} \lambda Q(C\lambda, S^{val})$, where Q represents a clustering quality metric, and S^{val} stands for a labelled validation subset drawn from the source data. This step is illustrated in **Figure 1** under the label source-guided auto HP tuning.

HyPASS-SC for CUMDA: improving robustness to domain gap. In this paper, we adapt HyPASS to the specifics of our cow re-ID CUMDA setting by selecting a dedicated value λ_k for each target dataset T_k . The PLC originally defined by **Eq. 1** is accordingly reformulated as a Domain-specific PLC written as:

$$\forall k \in [1, n], \hat{T}_k = C_{\lambda_k}(f_{\theta}, T_k). \quad (3)$$

The inner workings of HyPASS rest on reducing the domain gap. To make this feasible in the multi-target scenario, we pursue reduction along two axes:

- At the feature level, we exploit target-specific DSBN to shrink the domain gap within the feature space
- At the dataset statistics level, we introduce a novel Source validation Calibration (SC) technique

If the cross-dataset statistics gap may reasonably be neglected in academic person and vehicle re-ID benchmarks, as done in [38], it becomes a decisive factor in the cow re-ID problem at hand. Indeed, within established person re-ID collections, typical statistics vary only modestly, staying between 21 and 36 shots per ID (consider, for example, Market1501 [51], DukeMTMC [52], personX [53], and MSMT17 [18]). This uniformity does not always extend to other domains. The situation is especially acute in animal-focused re-ID, where collecting data is often more difficult, resulting in greater discrepancies in shots per ID. Furthermore, unlike person or vehicle re-ID operating in open-world scenarios, the total number of cows on a farm of interest is usually known or can be approximated with little difficulty in cross-domain re-ID applications. This knowledge enables us to devise SC for cattle re-ID, a technique that adjusts the number of shots per identity in the source partition to match the target's corresponding figure. SC produces target-specific source validation subsets S^{val}_k starting from S^{val} , thereby narrowing the cross-dataset statistics gap with the associated target training set T_k . SC is marked as Source Calibration in **Figure 1**. HyPASS is subsequently applied on S^{val}_k to determine λ^*_k , the optimal hyperparameter setting for clustering applied to T_k . The combined application of HyPASS and SC will be referred to as HyPASS-SC throughout the remainder of the paper.

Experiments

Datasets

Three separate datasets are employed throughout this paper: Cows2021 [17], HolsteinCattleRecognition [53], and the privately held CowFisheye. A visual overview of each dataset's content is offered in **Figure 2**.



Figure 2. Visual overview of the content of each dataset. From left to right: Cows2021 [17], HolsteinCattleRecognition [53], CowFisheye (private).

Cows2021

Cows2021 [17] supplies RGB images and video sequences covering 186 individuals. Recording was conducted at an elevation of 4 m using a downward-pointing pinhole camera.

Image extraction. The frames were sampled over a one-month recording period. An oriented bounding-box detector paired with a tracker isolates cow images from the video feed. The resulting boxes are centred on the torsos, with heads excluded, and every individual faces to the right. Readers are directed to [17] for a fuller description of the acquisition protocol.

IDs & samples. Since this study requires labelled annotations for initial supervised training to ensure fair performance comparisons against unsupervised UDA, only the labelled portion of the data is utilised. In total, 8670 images covering 181 distinct individuals were obtained, yielding a mean of 48 shots per identity. Table 1 provides a more granular breakdown of the image distribution.

Table 1. Dataset statistics. For Cows2021 [17] and HolsteinCattle [53], * denotes that the dataset is derived from the RGB-annotated segment of the full dataset, under a 50/50 ID split for Train/Test. Train and test IDs do not overlap.

Dataset	# gallery images	# query images	# test IDs	# train images	# train IDs
Cows2021*	3213	855	91	4602	90
HolsteinCattle*	414	204	68	609	68
CowFisheye	2224	151	16	6334	62

Complexity. Despite the month-long capture window, illumination conditions and cow viewpoints remain largely consistent across acquisitions. Occlusion within the images is negligible or non-existent.

HolsteinCattleRecognition

HolsteinCattleRecognition [53] comprises RGB and infrared imagery of 1237 individuals. Recording relied on a pinhole camera installed at ground level, 5 m distant from the milking equipment being filmed. For compactness, the remainder of the paper refers to it as HolsteinCattle.

Image extraction. Imagery was gathered across nine consecutive days. Every frame contains exactly one cow situated inside the milking machine. Readers are directed to [53] for a fuller description of the acquisition protocol.

IDs & samples. For this study, only the RGB portion is considered. A total of 1227 images portraying 136 distinct individuals were obtained, averaging 9 shots per identity. Table 1 provides a more granular breakdown of the image distribution.

Complexity. This dataset presents cattle that are partially hidden and placed at different positions within the milking apparatus.

CowFisheye

CowFisheye is a privately sourced collection of RGB images depicting 78 individuals. Data were collected on a single farm using 4 downward-facing fisheye cameras mounted 6 m above ground level. Identity labels were assigned by hand. This dataset is representative of the inherently difficult data encountered in scenarios that demand around-the-clock cow identification at any spot on the farm.

Image extraction. Images were sampled from six days of recording, spanning both daylight and darkness. A detector outputs frames that contain the entire cow body plus head, horizontally aligned and oriented to the right.

IDs & samples. Images were manually picked and annotated to ensure robust variation in viewpoint and lighting for each animal. The extraction process yielded 8709 images covering 78 distinct individuals, which gives an average of 112 instances per identity. **Table 1** and **Figure 3** provide a more granular breakdown of the image distribution.

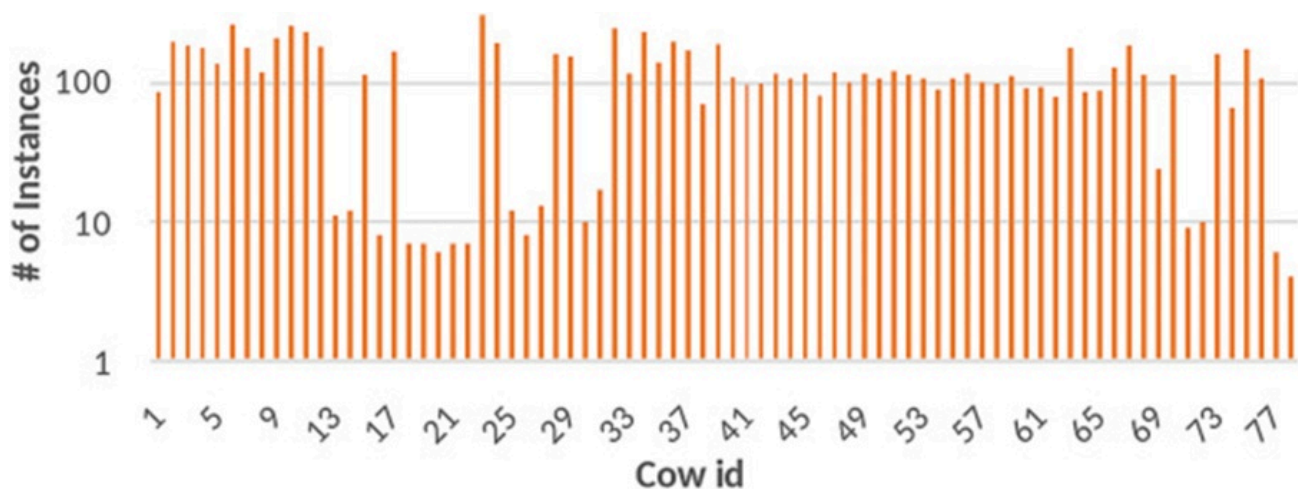


Figure 3. Image count per ID within the CowFisheye dataset.

Complexity. The challenging nature of CowFisheye mirrors the intended real-world task: continuous 24/7 re-ID of cows spanning the entire camera network. In particular, the low-angle camera arrangement necessary for full barn coverage can result in heavy occlusion of cows by equipment or herdmates. Fisheye lenses introduce characteristic distortions. On top of that, capture occurs under dramatically shifting illumination, which can cause a single cow to look markedly different from one image to the next. During night hours specifically, a near-infrared mode is triggered, producing monochrome photographs. **Figure 4** offers a glimpse into the complexity embodied by the CowFisheye dataset.



Figure 4. A glimpse into the complexity of the CowFisheye dataset. Left: changing illumination, Center: occlusion by equipment/cows, Right: diverse viewpoints. All frames show the very same individual.

Experimental settings

Use case

Our scenario is specific to animal re-ID across multiple farms, assuming the availability of labelled images from one farm together with one or several sets of unlabelled images originating from multiple farms. The CUMDA re-ID approach pursues two concurrent goals:

- Conservation of performance on the source (labelled) domain
- Specialization to the target (unlabelled) domains

We also count on stronger generalization to a novel, unseen target domain—the equivalent of a newly added farm in a practical deployment.

Accordingly, results are reported for each training configuration across all three datasets. Throughout the experimental section, we examine the separate building blocks of our CUMDA method: source guidance, DSBN, and HyPASS-SC. It is worth noting that this paper does not set out to fine-tune the network architecture or manually search for hyperparameters; instead, it advances a principled approach for performing effective CUMDA re-ID through pseudo-labeling.

Framework

To confer some resilience against pseudo-label noise, we use the state-of-the-art Mutual Mean Teaching (MMT) framework [9] combined with a Resnet-18 backbone pre-trained on ImageNet. The Resnet-18's last stride is reduced to 1, thereby increasing the spatial resolution of the feature maps. The DBSCAN clustering routine is applied to k-reciprocal encoded features with $k = 30$. DBSCAN's n_{min} and eps parameters are held at $n_{min} = 0.4$ and $eps = 0.6$, their commonly adopted values [38]. These remain fixed, except for eps , which undergo tuning via HyPASS-SC. Any other unspecified settings follow the original MMT publication [9] exactly. The contributions of this paper, however, are not tied to MMT and could be integrated into any UDA framework.

Data preprocessing

Mini-batches of size 16 are composed per domain, each with $P = 4$ identities and $K = 4$ shots per identity. Because CUMDA batches are assembled from one mini-batch in the source domain and one in each target domain, their overall size fluctuates. Specifically, it depends on n_{domain} , which counts how many domains participate in training. Here, the

batch size is $16 \times \text{ndomain}$. All images are resized to 128×128 pixels. Standard re-ID data augmentations, including cropping and flipping, are activated during training. Random erasing was omitted after experimental observations showed it reduced cow re-ID accuracy.

Initial supervised pre-training

The network is trained for 20 epochs on the selected dataset. The learning rate is maintained at $\text{lr} = 3.5 \times 10^{-4}$. Training on source images draws on both triplet and cross-entropy losses [9].

Domain adaptation

Training proceeds for 15 epochs, each composed of 200 iterations. This duration is deliberately chosen to guard against overfitting on the smaller-scale datasets. The learning rate is kept at $\text{lr} = 3.5 \times 10^{-4}$. The training phase employs both triplet and cross-entropy losses. The source and target domains make use of a single shared fully connected layer for classification. Whenever MMT is used, evaluation is invariably carried out on model 1, mirroring real-world constraints in which one cannot determine the best-performing model on the target because it lacks annotations. For adaptation across multiple target datasets, DSBN is extended to include a dedicated Batch Normalization (BN) module for each domain. During inference, the BN belonging to the domain that most closely resembles the test domain in visual terms is selected. Concretely, if the test domain's BN was estimated during training, it is applied during testing. When it was not, the BN from CowFisheye is used for testing on Cows2021 or HolsteinCattle, and the BN from Cows2021 is used for testing on CowFisheye.

Testing

Given that most datasets originate from a single camera viewpoint, evaluation is performed without excluding images captured by that device. The metric adopted for reporting is mean Average Precision (mAP). This measure captures the network's capacity to successfully retrieve, from a gallery, the multiple shots belonging to a queried individual, with higher values indicating superior performance.

This metric represents the proportion of queries for which the topmost gallery result is correct, thereby reflecting retrieval precision. Nevertheless, rank-1 is sensitive to flaws in the dataset, particularly pronounced in the cattle re-id context examined in this work (constrained image diversity, label noise, and similar issues). Accordingly, all commentary within the main text will focus on mAP, as it demonstrates greater resilience and provides a more faithful picture of overall dataset-wide performance.

No re-ranking procedures are invoked during evaluation. The identity split for Cows2021 and HolsteinCattle adheres to the original ascending order: the first half of the identities constitutes the training partition. In contrast, the second half serves as the test partition.

HyPASS-SC

HyPASS-SC searches for the optimal DBSCAN hyperparameter eps within the interval $[0.35, 0.65]$, which corresponds to the empirically viable range documented for applications involving human datasets [38]. A stark contrast of roughly one order of magnitude exists between the average shots per identity in Holstein Cattle (9) and the other two datasets (48 and 112). Consequently, in configurations where HolsteinCattle serves as either target or source, HyPASS-SC constructs, respectively, a subsampled or oversampled version of the source validation partition. The aim is to align the shots-per-identity count of the source validation split with that of the target. The random subsampling of Cows2021 or CowFisheye presents no particular difficulty. At the same time, oversampling for Holstein cattle is performed by replicating the identical set of data augmentations already applied during the training phase. Crucially, the source data actually used for model training is not altered by this calibration process in any way. In a practical on-farm deployment, the total cow count is known, and a rough estimate of instances per identity can be derived by dividing the total number of captured images by the known or estimated herd size.

Results and Discussion

Effectiveness of our CUMDA, single target

Supervised training & direct transfer. The outcomes of supervised training are assembled in [Table 2](#). When training and testing are performed on the same dataset, the resulting scores provide a window into the inherent difficulty of each dataset—arranged from most to least challenging: CowFisheye, HolsteinCattle, and Cows2021.

Table 2. Performance (mAP in % and accuracy in % of the rank-1 closest element within the gallery) of models trained with full supervision on a single source dataset, together with direct transfer results on every target dataset in the absence of adaptation.

Method	CowFisheye		Test				Train	
	rank-1	mAP	HolsteinCattle		Cows2021		Target	Source
			rank-1	mAP	rank-1	mAP		
Supervised training	45.7	16.8	13.7	8.0	98.2	95.3	None	Cows2021
Supervised training	25.2	12.7	91.7	81.2	73.1	29.1	None	HolsteinCattle
Supervised training	75.5	50.5	24.5	13.6	95.1	71.0	None	CowFisheye

The table also presents cross-domain outcomes for models trained under supervision on a source dataset and then directly evaluated on the remaining datasets, without any adaptation step. The conspicuously low scores recorded on these alternate datasets underscore just how indispensable domain adaptation is for this problem. Within [Table 2](#), a pair of cross-domain configurations stands apart. To begin with, the direct transfer setting Cows2021 → HolsteinCattle records the single weakest target-domain result, a mere 8.0% mAP, against a ceiling of 81.2% that could theoretically be attained were ground-truth labels available for that domain. In contrast, the direct transfer setting CowFisheye → Cows2021 achieves the strongest result among all cross-domain pairings evaluated, hitting 71.0% mAP on the target side. These contrasting outcomes signal that the magnitude of the domain shift between farms plays a decisive role in shaping the network's aptitude for handling novel image styles, and that this shift is considerably more pronounced between HolsteinCattle and the other two datasets than it is between CowFisheye and Cows2021. Moreover, the richness and diversity encapsulated in CowFisheye—a product of its wide-ranging viewpoints and frequent occlusions—appear to facilitate bridging domain gaps, an interpretation supported by the noticeably higher direct transfer figures obtained when pre-training is anchored in CowFisheye.

UDA baseline. From this point forward in the paper, the label UDA will denote domain adaptation conducted without source guidance. The mechanics of this approach are visually summarized in the upper half of [Figure 1](#). It bears repeating that the MMT framework was adopted as the baseline for this study [\[9\]](#). It remains interchangeable with any alternative baseline, a flexibility already demonstrated in prior work [\[38, 54\]](#).

As evidenced by [Table 3](#), target-dataset performance sees an uplift, with Δ mAP readings falling across the interval [+1.7 p.p., +55.5 p.p.]; that is to say, the lower limit of the observed performance swing equals +1.7 percentage points (p.p.) while the upper limit reaches +55.5 p.p. These gains confirm that domain adaptation delivers meaningful improvements on the target dataset across all configurations tested. At the same time, the systematic erosion across all diagonal entries in [Table 3](#) reveals that the network suffers a substantial loss of what it originally learned from the source dataset. The magnitude of this forgetting is considerable, with Δ mAP values lying within [−61.6 p.p., −11.2 p.p.].

Table 3. Performance of MMT [\[9\]](#). mAP (in %), Δ mAP (in p.p.) denotes the deviation measured against the initial supervised models (cf. [Table 2](#)).

Method	Test						Train	
	CowFisheye		HolsteinCattle		Cows2021		Target	Source
	ΔmAP	mAP	ΔmAP	mAP	ΔmAP	mAP		
UDA	-7.7	9.1	+4.6	12.6	-56.1	39.2	HolsteinCattle	Cows2021
UDA	+7.5	24.3	+2.8	10.8	-11.2	84.1	CowFisheye	Cows2021
UDA	+3.3	16.0	-55.7	25.5	+55.5	84.6	Cows2021	HolsteinCattle
UDA	+1.7	14.4	-61.6	19.6	+25.8	54.9	CowFisheye	HolsteinCattle
UDA	-30.6	19.9	-5.1	8.5	+17.9	88.9	Cows2021	CowFisheye
UDA	-36.2	14.3	+9.1	22.7	-40.2	30.8	HolsteinCattle	CowFisheye

The capacity to generalize to a dataset not seen during training is inconsistent and appears to converge toward the performance baseline of a network trained exclusively on the target dataset. As a case in point, consider the CowFisheye → HolsteinCattle setup: the network's mAP on Cows2021 falls sharply from its pre-UDA level of 71.0% (cf. [Table 2](#)) to a post-UDA figure of just 30.8% (cf. [Table 3](#)). This diminished value sits remarkably close to the 29.1% mAP registered by a network whose supervision was restricted entirely to Holstein cattle (cf. [Table 2](#)). The overarching conclusion is that the UDA paradigm reshapes the network to suit a single dataset, without instituting any mechanism to protect performance on other datasets.

CUMDA. For the remainder of this paper, the designation CUMDA will be applied to domain adaptation that jointly incorporates source guidance, DSBN, and HyPASS-SC. The mechanics of this approach are visually summarized in the lower half of [Figure 1](#). The results achieved through CUMDA are tabulated in [Table 4](#). Relative to UDA (cf. [Table 3](#)), the improvements are wide-ranging.

Table 4. Performance of our CUMDA method (source guidance, DSBN, and HyPASS-SC). mAP (in %), ΔmAP (in p.p.) denotes the deviation measured against the initial supervised models (cf. [Table 2](#)).

Method	CowFisheye ΔmAP	CowFisheye mAP	Holstein Cattle ΔmAP	Holstein Cattle mAP	Cows2021 ΔmAP	Cows2021 mAP	Target Domain	Source Domain
CUMDA	+3.3	20.1	+7.9	15.9	+0.1	95.4	HolsteinCattle	Cows2021
CUMDA	+11.4	28.2	+5.7	13.7	-0.3	95.0	CowFisheye	Cows2021
CUMDA	+2.6	15.3	-0.4	80.8	+58.1	87.2	Cows2021	HolsteinCattle
CUMDA	+4.1	16.8	-0.1	81.1	+28.8	57.9	CowFisheye	HolsteinCattle
CUMDA	+9.6	60.1	-0.3	13.3	+19.6	90.6	Cows2021	CowFisheye
CUMDA	+9.6	60.1	+25.0	38.6	-0.1	70.9	HolsteinCattle	CowFisheye

To begin, CUMDA delivers results that uniformly exceed those of the corresponding supervised network across all three evaluated domains: the source, the target, and a third domain that remained entirely unseen. Looking at the source domain in particular, performance either holds steady or advances, with a maximum observed lift of +9.6 p.p. in mAP recorded for CowFisheye. On the target side, the performance jump is marked, with mean ΔmAP values of +38.9 p.p., +16.5 p.p., and +7.8 p.p. for Cows2021, HolsteinCattle, and CowFisheye, respectively, in configurations where each takes its turn as the adaptation target. As for generalization to the unseen domain, average gains of +14.4 p.p., +2.7 p.p., and +3.0 p.p. are observed on those same three datasets.

Furthermore, when evaluated specifically in the target domain, our proposed CUMDA method outperforms UDA. To render this comparison concrete, we examine the differences between the numbers in [Table 4](#) and those in [Table 3](#). On the source domain, the uplift is striking, climbing as high as $81.1 - 19.6 = +61.5$ p.p. for the cross-domain configuration HolsteinCattle $\rightarrow$ CowFisheye. On the target domain, performance improves in every case, with the magnitude of improvement ranging from +1.7 p.p. to +15.9 p.p. Generalization to the unseen domain also improves, with shifts ranging from -0.7 p.p. to +40.1 p.p.

When viewed collectively, these findings attest to the network's twofold capability: safeguarding what was originally learned from the source dataset, and simultaneously extracting and combining information from every available domain to achieve steady, across-the-board performance gains. Our CUMDA method thus guarantees source-domain conservation, fosters deeper specialization in each domain encountered during training, and improves generalization performance on domains never observed during training.

Ablation study, single target

Having established the performance advantages of our CUMDA approach over both direct transfer and conventional UDA, we now turn to a detailed examination of the contribution of each constituent component. This ablation study reports the averaged changes in performance, always referenced against a network trained in a fully supervised manner on the source. The corresponding figures are organized across [Tables 5-8](#).

Table 5. Performance of UDA expressed relative to direct transfer. ΔmAP (in p.p.) denotes the mAP (in %) gap compared with direct transfer (cf. [Table 2](#)).

	Test		
	CowFisheye	HolsteinCattle	Cows2021
Test set as	ΔmAP	ΔmAP	ΔmAP
Source	-33.4	-58.7	-33.7
Target	+4.6	+6.9	+36.7
Unseen	-2.2	-1.2	-7.2

Table 6. Performance of source-guided UDA expressed relative to direct transfer. ΔmAP (in p.p.) denotes the mAP (in %) gap compared with direct transfer (cf. [Table 2](#)).

	Test		
	CowFisheye	HolsteinCattle	Cows2021
Test set as	ΔmAP	ΔmAP	ΔmAP
Source	+9.2	-0.6	-0.4
Target	+5.7	+6.0	+36.6
Unseen	+3.3	+1.1	+14.2

Table 7. Performance of source-guided + DSBN UDA expressed relative to direct transfer. ΔmAP (in p.p.) denotes the mAP (in %) gap compared with direct transfer (cf. [Table 2](#)).

	Test		

	CowFisheye	HolsteinCattle	Cows2021
Test set as	ΔmAP	ΔmAP	ΔmAP
Source	+7.5	-2.4	+0.3
Target	+8.2	+10.0	+40.2
Unseen	+2.5	+3.1	+13.0

Table 8. Performance of our single-target CUMDA expressed relative to direct transfer. ΔmAP (in p.p.) denotes the mAP (in %) gap compared with direct transfer (cf. Table 2).

	Test		
	CowFisheye	HolsteinCattle	Cows2021
Test set as	ΔmAP	ΔmAP	ΔmAP
Source	+9.6	-0.3	-0.1
Target	+7.8	+16.5	+38.9
Unseen	+3.0	+2.7	+14.4

Source guidance. The averaged outcomes when source guidance is activated appear in Table 6. We contrast these directly against the UDA-only results collected in Table 5.

Retaining the source as a labelled stream throughout the training cycle yields a pronounced upswing in accuracy on that very same source dataset. When set beside standard UDA, the mean improvement reaches $33.7 - 0.4 = +33.3$ p.p., $+58.1$ p.p., and $+42.6$ p.p. for cases where Cows2021, HolsteinCattle, and CowFisheye, respectively, serve as the source. Meanwhile, target-dataset results stay largely flat, with average drifts of merely -0.1 p.p., -0.9 p.p., and $+1.1$ p.p. Because the model now absorbs information from both the source and the target simultaneously, its ability to extend to unfamiliar data improves by margins of $+21.4$ p.p., $+2.3$ p.p., and $+5.5$ p.p. on Cows2021, HolsteinCattle, and CowFisheye, respectively.

The takeaway is clear: source guidance lets the model hold its ground on the target while reliably safeguarding what was learned on the source, something plain UDA fails to do. Joint exposure to both domains also translates into more robust generalization to data from a domain never seen during training.

DSBN. Our strategy for narrowing the domain gap is to adopt DSBN. Averaged measurements for the tandem of source guidance and DSBN are collated in Table 7, which we then measure against the source-guidance-only configuration from Table 6.

Contrasted with source-guided UDA, the addition of DSBN delivers a meaningful target-side lift: $40.2 - 36.6 = +3.6$ p.p., $+4.0$ p.p., and $+2.5$ p.p. for Cows2021, HolsteinCattle, and CowFisheye, respectively. The flip side is a slight softening of source-dataset scores, with shifts of $+0.7$ p.p., -1.8 p.p., and -1.7 p.p. Generalization to an unseen domain, viewed in aggregate, stays essentially where it was.

In short, across the experiments reported here, DSBN alone does not consistently translate into better cow re-identification. Its value, however, becomes apparent further down the line: it is what makes HyPASS viable.

HyPASS-SC.

It is precisely the presence of DSBN that enables source-informed automatic tuning of pseudo-labeling hyperparameters, as demonstrated here with HyPASS-SC. The full combination—source guidance + DSBN + HyPASS-SC, i.e., our complete CUMDA—has its averaged results recorded in **Table 8**. These are juxtaposed with the source-guided + DSBN figures found in **Table 7**.

Relative to source-guided + DSBN UDA, target-domain performance on Holstein Cattle jumps by $16.5 - 10.0 = +6.5$ p.p. One plausible reason is the marked dissimilarity between Holstein Cattle and the other domains, which likely shifts the ideal clustering parameter configuration considerably away from default values. On Cows2021, HyPASS-SC appears marginally weaker than source-guided + DSBN, trailing by -1.3 p.p. This outcome suggests that HyPASS-SC is not a universal silver bullet, particularly when the source validation split is small. Nonetheless, its practical usefulness endures because it eliminates the burden of manual parameter tuning. On the source side, performance edges up by deltas of $-0.1 - 0.3 = -0.4$ p.p., $+2.1$ p.p., and $+2.1$ p.p., even overtaking the numbers posted by source-guided UDA (cf. **Table 6**). Generalization to unseen data improves by $+1.4$ p.p., -0.4 p.p., and $+0.5$ p.p.

We conclude that HyPASS-SC exerts a net positive influence on results across the board. It essentially guarantees well-calibrated clustering for every target, which, in turn, supports stronger source retention, finer target specialization, and better generalization to novel domains compared with the stripped-down variants evaluated.

Focusing on the cross-domain pair HolsteinCattle → CowFisheye, **Figure 5** presents a t-SNE projection of the embedding space that traces how the representations evolve as domain adaptation, source guidance, and the remaining CUMDA components are successively engaged. The figure lays bare: the collapse of source performance under UDA, the recovery and broad enhancement brought by source guidance, and the best overall cluster structure achieved only when every piece of our CUMDA method is active together.

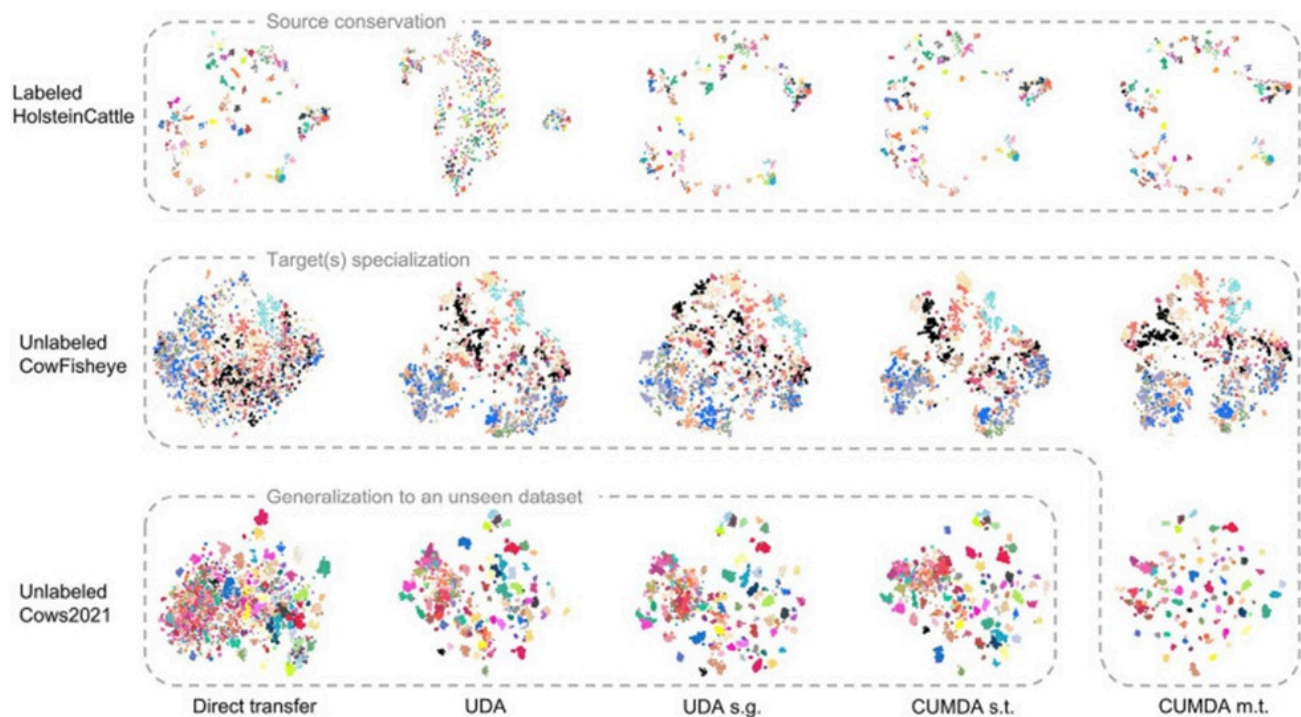


Figure 5. Progression of the embedding space on all validation splits for the cross-domain scenario HolsteinCattle → CowFisheye. t-SNE visualization where every identity is assigned a random colour and size. Each point represents a single image projected into the embedding space. Rows correspond to datasets; columns correspond to methods. Acronyms: UDA: UDA baseline with MMT; s.g.: source-guided; CUMDA: Cumulative Unsupervised Multi-Domain

Adaptation; s.t.: single-target; m.t.: multi-target (with Cows2021). Ideally, points of identical colour should group into compact, well-separated clusters. Best viewed in colour.

CUMDA re-ID operating over multiple targets simultaneously (CowFisheye and Cows2021) is likewise depicted, with unmistakable improvements visible on all three datasets.

Benefit of the source calibration in HyPASS-SC

HyPASS exhibits a strong dependence on the statistical properties of the dataset at hand—in our particular context, this translates to sensitivity to the shots-per-identity counts of Cows2021 (48), HolsteinCattle (9), and CowFisheye (112).

To confirm that our newly introduced source validation set calibration genuinely matters, we pit HyPASS-SC against uncalibrated HyPASS on two representative cross-domain transfers: HolsteinCattle → Cows2021 (a scenario that demands oversampling) and CowFisheye → HolsteinCattle (a scenario that demands subsampling).

We quantify clustering quality using the Adjusted Rand Index (ARI), a standard metric for assessing the agreement between two partitions. Here, it is computed between the ground-truth labels of the target training portion and the cluster assignments output by the algorithm, using the implementation provided by scikit-learn². **Figure 6** plots the trajectory of the ARI as we progressively align the statistical profiles of source and target. A higher ARI reading denotes a more faithful clustering. All comparisons are drawn at the 3000th training iteration.

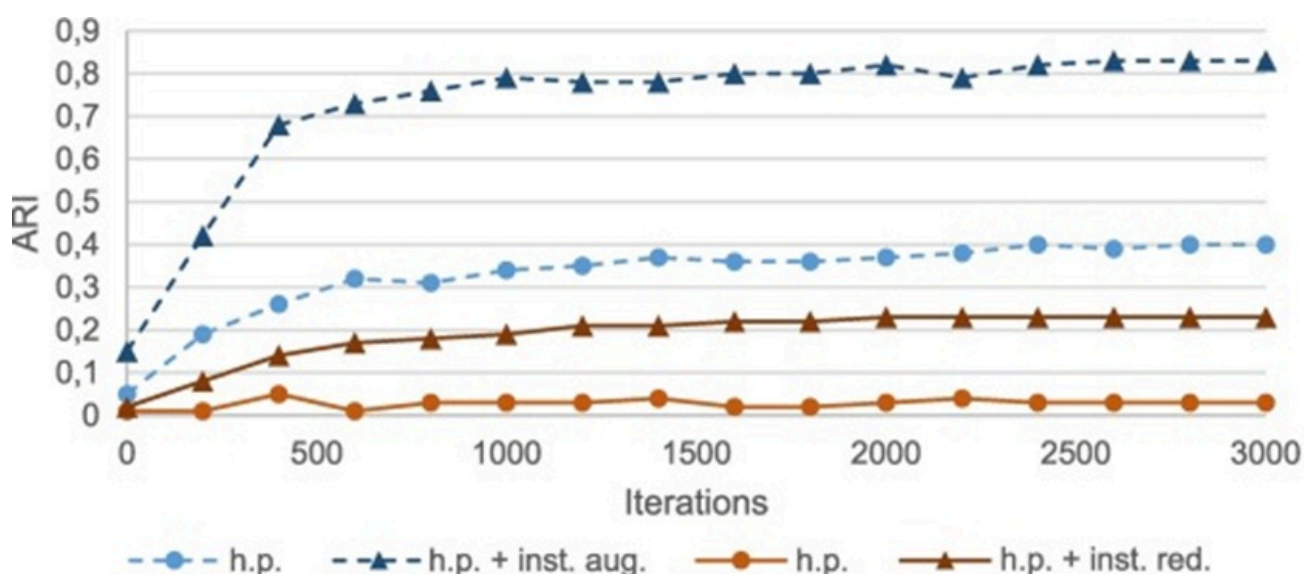


Figure 6. Trajectory of the Adjusted Random Index (ARI) on target clustering. How calibrating the source validation set shapes HyPASS behaviour. Blue dashed curves: HolsteinCattle → Cows2021; red curves: CowFisheye → HolsteinCattle.

Acronyms: h.p.: HyPASS [38]; inst. Red.: source instance reduction; inst. Aug.: source instance augmentation. (For interpretation of the colour references in this figure, the reader is referred to the web version of this article.)

In the HolsteinCattle → Cows2021 direction, we oversample HolsteinCattle from its native ~9 shots/ID up to roughly 90 shots/ID. The effect is striking: the ARI more than doubles, rising from 0.40 in the uncalibrated condition to 0.83 after calibration. Translating this into mAP, target-side Cows2021 performance moves from 77.4% (uncalibrated) to 87.2% (calibrated).

In the CowFisheye → HolsteinCattle direction, we subsample CowFisheye from ~112 shots/ID down to approximately 9 shots/ID. Here again, the ARI experiences a pronounced boost, rising from 0.03 without calibration to 0.23 with it. In mAP terms, the target HolsteinCattle score rises from 29.3% (uncalibrated) to 38.6% (calibrated).

These observations drive home the point that calibrating the source validation set is paramount when working with datasets whose shots-per-identity figures differ widely—a recurring reality in animal-oriented data collections. We take this opportunity to reiterate that source calibration has been uniformly applied across all experimental results reported earlier in this paper.

Effectiveness of our CUMDA, multiple targets

A core objective of this work is to broaden domain adaptation to accommodate multiple target domains within a single training process. This setup mirrors authentic field requirements, where a model initially trained on one labelled dataset must subsequently adapt to numerous distinct farm environments. Averaged outcomes for our complete CUMDA formulation (source guidance, DSBN, and HyPASS-SC) are collated in [Table 10](#). We contrast these against the source-guided variant, whose figures appear in [Table 9](#).

Table 9. Performance of source-guided + multi-target UDA expressed relative to direct transfer. ΔmAP (in p.p.) denotes the mAP (in %) gap compared with direct transfer (cf. [Table 2](#)).

	Test		
	CowFisheye	HolsteinCattle	Cows2021
Test set as	ΔmAP	ΔmAP	ΔmAP
Source	+10.5	−1.5	−0.3
Target	+3.8	+7.4	+36.0

Table 10. Performance of our CUMDA method expressed relative to direct transfer. ΔmAP (in p.p.) denotes the mAP (in %) gap compared with direct transfer (cf. [Table 2](#)).

	Test		
	CowFisheye	HolsteinCattle	Cows2021
Test set as	ΔmAP	ΔmAP	ΔmAP
Source	+10.4	−0.9	0.0
Target	+7.5	+13.4	+38.8

Focusing on the target datasets, our approach outperforms source-guided UDA by margins of +2.8 p.p., +6.0 p.p., and +3.7 p.p. for Cows2021, HolsteinCattle, and CowFisheye, respectively, while maintaining source-dataset performance. These numbers testify to the tangible practical importance of the CUMDA method in multi-dataset deployment scenarios. We attribute this performance edge to the synergistic pairing of DSBN and HyPASS-SC.

DSBN partially closes the domain gap by performing domain-aware normalization and simultaneously enabling multiple domains to share a single backbone network, a property that actively encourages generalization. HyPASS-SC complements this by delivering clustering parameters that are individually optimized for the statistical peculiarities of each target dataset. This dual mechanism ensures that clustering remains reliable across all target domains, thereby improving network performance across all datasets.

Limitations

The preceding sections have established that our approach yields marked improvements in unsupervised domain adaptation performance compared with traditional pseudo-labeling-based UDA pipelines. That said, one cannot overlook

that a domain's performance when treated as an unlabelled target remains substantially below what can be achieved when the same domain is fully labelled and used as a source.

This performance shortfall is a defining characteristic of the cow re-ID challenge. In the person re-ID domain, colour cues routinely ease the pseudo-labeling process to the point where unsupervised UDA often narrows the gap to fully supervised results [9]. Cow re-ID offers no such shortcut: colour conveys negligible discriminative signal for pseudo-labeling, obliging the system to latch onto finer-grained cues such as silhouette and spot morphology. Moreover, both in this investigation and in cow re-ID research at large, the domain shift between the chosen datasets can be substantially wider than the inter-dataset gaps typically observed in person re-ID benchmarks.

Stated differently, even with automatic tuning of the pseudo-labeling hyperparameters, the ceiling of the pseudo-labeling paradigm itself ultimately constrains the entire pipeline. Pseudo-label fidelity is bounded by two factors: the inherent capabilities of the clustering algorithm in use and the degree to which the network can extract genuinely discriminative representations. An extended treatment of how different pseudo-labeling strategies shape UDA outcomes is available in [38].

Conclusion

This paper has introduced a novel CUMDA re-ID method, purpose-built for effective cumulative multi-domain adaptation and grounded in the concrete constraints and operational requirements of cattle re-ID. The contribution directly addresses the realities of practical deployment, where a model must continually improve its performance across a growing collection of farms for which manual annotation is seldom feasible. Our work extends domain adaptation beyond the conventional single-target scenario to settings involving multiple target domains that exhibit wide divergences in dataset statistics and visual appearance—precisely the kinds of obstacles that surface routinely in real-world applications. The CUMDA method we propose is built upon three foundational elements: source guidance to prevent catastrophic forgetting, domain gap reduction through domain-specific normalization, and an automated, source-informed hyperparameter selection procedure for clustering that builds on HyPASS. Recognizing the acute statistical diversity in agricultural datasets, we have also contributed a source calibration technique that substantially improves HyPASS performance under such conditions. We have evaluated our approach against direct transfer, standard domain adaptation, and source-guided domain adaptation. The experimental record demonstrates consistent and significant performance gains, realized as stronger retention of source knowledge, deeper specialization within each target domain, and more robust generalization to domains that were never observed during training, compared with classical pseudo-labeling-based UDA methods.

Acknowledgements

Partial financial support was received from the AIHerd company.

This publication was made possible by the use of the FactoryIA supercomputer, financially supported by the Ile-de-France Regional Council.

The CowFisheye dataset was kindly provided by the AIHerd company.

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 26 Aug 2024

Revised: 23 Sep 2024

Accepted: 06 Nov 2024

Published online: 15 March 2025

Rights and permissions

Open Access The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at <https://creativecommons.org/licenses/by-nc-sa/4.0/>.

References

Zheng L, Yang Y, Hauptmann AG. Person re-identification: past, present and future. arXiv [Preprint]. 2016. arXiv:1610.02984.

Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst*. 2012;25:1097-105.

Ye M, Shen J, Lin G, Xiang T, Shao L, Hoi SC. Deep learning for person re-identification: a survey and outlook. *IEEE Trans Pattern Anal Mach Intell*. 2021;44(6):2872-93.

Khan SD, Ullah H. A survey of advances in vision-based vehicle re-identification. *Comput Vis Image Underst*. 2019;182:50-63.

Schneider S, Taylor GW, Kremer SC. Similarity learning networks for animal individual re-identification-beyond the capabilities of a human observer. In: *IEEE Winter Conf Appl Comput Vis Workshops*. 2020; p. 44-52.

Liu R, Zhang L, Guo L. Part-pose guided Amur tiger re-identification. In: *IEEE/CVF Int Conf Comput Vis Workshops*. 2019a.

Liu J, Zha ZJ, Chen D, Hong R, Wang M. Adaptive transfer network for cross-domain person re-identification. In: *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*. 2019b; p. 7202-11.

Liu N, Zhao Q, Zhang N, Cheng X, Zhu J. Pose-guided complementary features learning for Amur tiger re-identification. In: *IEEE/CVF Int Conf Comput Vis Workshops*. 2019c.

Ge Y, Chen D, Li H. Mutual mean-teaching: pseudo label refinery for unsupervised domain adaptation on person re-identification. In: *Int Conf Learn Represent*. 2019.

Ge Y, Zhu F, Chen D, Zhao R, Li H. Self-paced contrastive learning with hybrid memory for domain adaptive object re-id. In: *Adv Neural Inf Process Syst*. 2020; p. 11309-21.

Dubourgieux F, Audigier R, Loesch A, Ainouz S, Canu S. Unsupervised domain adaptation for person re-identification through source-guided pseudo-labeling. In: *IEEE Proc Int Conf Pattern Recogn*. 2021; p. 4957-64.

Wang J, Lan C, Liu C, Ouyang Y, Qin T. Generalizing to unseen domains: a survey on domain generalization. In: *Int Joint Conf Artif Intell*. 2021; p. 4627-35.

Bergamini L, Porrello A, Dondona AC, Del Negro E, Mattioli M, D'alterio N, et al. Multi-views embedding for cattle re-identification. In: *IEEE Int Conf Signal-Image Technol Internet-Based Syst*. 2018; p. 184-91.

Andrew W, Gao J, Mullan S, Campbell N, Dowsey AW, Burghardt T. Visual identification of individual Holstein-Friesian cattle via deep metric learning. *Comput Electron Agric*. 2021;185:106133.

Andrew W, Hannuna S, Campbell N, Burghardt T. Automatic individual Holstein Friesian cattle identification via selective local coat pattern matching in RGB-D imagery. In: *IEEE Proc Int Conf Image Process*. 2016; p. 484-8.

Andrew W, Greatwood C, Burghardt T. Visual localisation and individual identification of Holstein Friesian cattle via deep learning. In: *Proc IEEE Int Conf Comput Vis Workshop*. 2017; p. 2850-9.

Gao J, Burghardt T, Andrew W, Dowsey AW, Campbell NW. Towards self-supervision for video identification of individual Holstein-Friesian cattle: the Cows2021 dataset. *arXiv [Preprint]*. 2021. arXiv:2105.01938.

Wei L, Zhang S, Gao W, Tian Q. Person transfer GAN to bridge domain gap for person re-identification. In: *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*. 2018; p. 79-88.

Bak S, Carr P, Lalonde JF. Domain adaptation through synthesis for unsupervised person re-identification. In: *Comput Vis ECCV*. 2018; p. 189-205.

Deng W, Zheng L, Ye Q, Kang G, Yang Y, Jiao J. Image-image domain adaptation with preserved self-similarity and domain-dissimilarity for person re-identification. In: *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*. 2018; p. 994-1003.

Zhong Z, Zheng L, Li S, Yang Y. Generalizing a person retrieval model hetero-and homogeneously. In: *Comput Vis ECCV*. 2018; p. 172-88.

Huang Y, Wu Q, Xu J, Zhong Y. Sbsgan: suppression of inter-domain background shift for person re-identification. In: *Proc IEEE Int Conf Comput Vis*. 2019; p. 9527-36.

Chen Y, Zhu X, Gong S. Instance-guided context rendering for cross-domain person re-identification. In: *Proc IEEE Int Conf Comput Vis*. 2019; p. 232-42.

Zhu JY, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: *Proc IEEE Int Conf Comput Vis*. 2017; p. 2223-32.

Wang J, Zhu X, Gong S, Li W. Transferable joint attribute-identity deep learning for unsupervised person re-identification. In: *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*. 2018; p. 2275-84.

Lin S, Li H, Li CT, Kot AC. Multi-task mid-level feature alignment network for unsupervised cross-dataset person re-identification. In: *29th Br Mach Vis Conf (BMVC)*. 2018.

Chang X, Yang Y, Xiang T, Hospedales TM. Disjoint label space transfer learning with common factorised space. In: *Proc AAAI Conf Artif Intell*. 2019.

Li YJ, Yang FE, Liu YC, Yeh YY, Du X, Frank Wang YC. Adaptation and re-identification network: an unsupervised deep transfer learning approach to person re-identification. In: *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit Workshops*. 2018; p. 172-8.

Li YJ, Lin CS, Lin YB, Wang YCF. Cross-dataset person re-identification via unsupervised pose disentanglement and adaptation. In: *Proc IEEE Int Conf Comput Vis*. 2019; p. 7919-29.

Qi L, Wang L, Huo J, Zhou L, Shi Y, Gao Y. A novel unsupervised camera-aware domain adaptation framework for person re-identification. In: *Proc IEEE Int Conf Comput Vis*. 2019; p. 8080-9.

Song L, Wang C, Zhang L, Du B, Zhang Q, Huang C, et al. Unsupervised domain adaptive re-identification: theory and practice. *Pattern Recogn.* 2020;102:107173.

Zhang X, Cao J, Shen C, You M. Self-training with progressive augmentation for unsupervised cross-domain person re-identification. In: *Proc IEEE Int Conf Comput Vis.* 2019; p. 8222-31.

Jin X, Lan C, Zeng W, Chen Z. Global distance-distributions separation for unsupervised person re-identification. In: *Comput Vis ECCV.* 2020; p. 735-51.

Tang H, Zhao Y, Lu H. Unsupervised person re-identification with iterative self-supervised domain adaptation. In: *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit Workshops.* 2019.

Zhai Y, Lu S, Ye Q, Shan X, Chen J, Ji R, et al. Ad-cluster: augmented discriminative clustering for domain adaptive person re-identification. In: *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit.* 2020a; p. 9021-30.

Zou Y, Yang X, Yu Z, Kumar BV, Kautz J. Joint disentangling and adaptation for cross-domain person re-identification. In: *Comput Vis ECCV.* 2020; p. 87-104.

Fu Y, Wei Y, Wang G, Zhou Y, Shi H, Huang TS. Self-similarity grouping: a simple unsupervised cross domain adaptation approach for person re-identification. In: *Proc IEEE Int Conf Comput Vis.* 2019; p. 6112-21.

Dubourvieux F, Loesch A, Audigier R, Ainouz S, Canu S. Improving unsupervised domain adaptive re-identification via source-guided selection of pseudo-labeling hyperparameters. *IEEE Access.* 2021;9:149780-95.

Chen G, Lu Y, Lu J, Zhou J. Deep credible metric learning for unsupervised domain adaptation person re-identification. In: *Comput Vis ECCV.* 2020; p. 643-59.

Zhai Y, Ye Q, Lu S, Jia M, Ji R, Tian Y. Multiple expert brainstorming for domain adaptive person re-identification. In: *Comput Vis ECCV.* 2020b; p. 594-611.

Zhao F, Liao S, Xie GS, Zhao J, Zhang K, Shao L. Unsupervised domain adaptation with noise resistible mutual-training for person re-identification. In: *Comput Vis ECCV.* 2020; p. 526-44.

Peng J, Wang Y, Wang H, Zhang Z, Fu X, Wang M. Unsupervised vehicle re-identification with progressive adaptation. *arXiv [Preprint].* 2020. arXiv:2006.11486.

Yu HX, Zheng WS, Wu A, Guo X, Gong S, Lai JH. Unsupervised person re-identification by soft multilabel learning. In: *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit.* 2019; p. 2148-57.

Zhong Z, Zheng L, Luo Z, Li S, Yang Y. Invariance matters: exemplar memory for domain adaptive person re-identification. In: *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit.* 2019; p. 598-607.

Luo C, Song C, Zhang Z. Generalizing person re-identification by camera-aware invariance learning and cross-domain mixup. In: *Comput Vis ECCV.* 2020; p. 224-41.

Hermans A, Beyer L, Leibe B. In defense of the triplet loss for person re-identification. *arXiv [Preprint].* 2017. arXiv:1703.07737.

Luo H, Gu Y, Liao X, Lai S, Jiang W. Bag of tricks and a strong baseline for deep person re-identification. In: *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit Workshops.* 2019.

Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. In: *Proc Int Conf Mach Learn.* 2015; p. 448-56.

Zajkac M, Zolna K, Jastrzebski S. Split batch normalization: improving semi-supervised learning under domain shift. arXiv [Preprint]. 2019. arXiv:1904.03515.

Chang WG, You T, Seo S, Kwak S, Han B. Domain-specific batch normalization for unsupervised domain adaptation. In: Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit. 2019; p. 7354-62.

Zheng L, Shen L, Tian L, Wang S, Wang J, Tian Q. Scalable person re-identification: a benchmark. In: Proc IEEE Int Conf Comput Vis. 2015.

Sun X, Zheng L. Dissecting person re-identification from the viewpoint of viewpoint. In: Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit. 2019; p. 608-17.

Bhole A, Falzon O, Biehl M, Azzopardi G. A computer vision pipeline that uses thermal and RGB images for the recognition of Holstein cattle. In: Comput Anal Images Patterns. 2019;11679:108-19.

Dubourvieux F, Audigier R, Loesch A, Ainouz S, Canu S. A formal approach to good practices in pseudo-labeling for unsupervised domain adaptive re-identification. Comput Vis Image Underst. 2022;223:103527.