Contrastive learning with multi-centroid proxy loss for domain adaptive cattle identification

Jianmin Zhao[1*], Qiusheng Lian[2], Yufei Zhao[3]

(1. *School of Automation and Electrical Engineering, Inner Mongolia University of Science and Technology, Baotou 014010, Inner Mongolia, China*;
2. *School of Information Science and Technology, Yanshan University, Qinhuangdao 066004, Hebei, China*;
3. *Vocational and Technical College, Inner Mongolia Agricultural University, Baotou 014109, Inner Mongolia, China*)

Abstract: Visual identification of cattle in the wild is essential to provide continuous individual monitoring applicable to precision livestock farming. Supervised learning heavily relies on annotation, which is really a time-consuming work for cattle labeling. Domain adaptive cattle identification aims at transferring knowledge learned from the source domain with rich annotations to the unlabeled target domain. Pseudo-label-based contrastive learning with a unique centroid inevitably incorporates information from different identities due to imperfect clustering. Thus, a contrastive learning with multi-centroid proxy (CL-MCP) for domain adaptive cattle identification was proposed with more local centroids to enhance the reliability of pseudo-labels for a more compact cluster in the target feature space. Firstly, a target domain feature storage module and a momentum update strategy were proposed to progressively update the target domain features for effective training with a stable clustering space. Secondly, a multi-centroid storage module and a proxy representation method were proposed to learn more informative local clusters and provide a representative proxy for each class to efficiently form correct clusters in the feature space. Finally, the domain-specific proxy-level contrastive loss was presented to enlarge the similarity between a query and its positive proxy while reducing the similarities among the query and its negative proxies for more compact clustering. It is encouraging to find that our CL-MCP mechanism performs better than Deep Metric Learning (DML) approaches for identifying individuals from different farms or unseen viewpoints or of a new breed. The datasets, MVCAID100, CNSID100, and Cattle-2022, are available on https://pan.baidu.com/s/19hoWd__7NMLvdLNp-otDSRg (code: d3oa). The results of this study can provide an effective cattle identification method applicable to automated production monitoring, behavioral and physiological observation, health and welfare supervision in precision livestock farming, and animal science research.

Keywords: cattle identification, precision livestock farming, unsupervised domain adaptation (UDA), contrastive learning, unsupervised learning

DOI: 10.25165/j.ijabe.20261901.9242

Citation: Zhao J M, Lian Q S, Zhao Y F. Contrastive learning with multi-centroid proxy loss for domain adaptive cattle identification. Int J Agric & Biol Eng, 2026; 19(1): 187–196.

1 Introduction

Continuous identification of cattle in the wild with an image of any view provides an essential way for real-time cattle monitoring that is applicable for precision livestock farming and animal science research[1,2]. Deep metric learning approaches have achieved success in cattle identification with dorsal images[3,4] and multi-view images[5-8]. However, as is known to all, supervised learning needs a large number of well-labeled images, and cattle ID annotation is labor-intensive and requires the assistance of professional breeders. Moreover, in real livestock farming scenarios, there are many more unlabeled individuals for cattle identification on new farms. Thus, how to transfer the learning ability from the source domain with rich labels to the unlabeled target domain is essential for solving the dependence of annotation in model training.

To address this problem, Unsupervised Domain Adaptation (UDA) offers effective approaches. Existing methods based on pseudo-label have achieved great success in human and vehicle recognition tasks. A two-stage training scheme is used to solve the problem, as 1) generating pseudo-labels for the training dataset on the target domain by clustering, and 2) optimizing the model on either the source domain, the target domain, or both of them under the supervision of contrastive loss with the class representation.

Traditionally, elements in contrastive loss include an anchor and its related positive and negative samples. Although the instance-level contrastive loss performs well in deep metric learning tasks, it performs poorly on the domain adaptation due to noises caused by the imperfect performance of clustering on the target domain. Thus, proxy-level contrastive loss, where the proxy is the average of the features or a learnable weight in the model, achieves success in the domain adaptive object re-identification (re-ID)[9].

As shown in Figure 1a, the cluster using only one centroid for a class tends to merge the information from different individuals due to the limitation of a sample to choose a more proper center. When it is used as the proxy to calculate contrastive loss, the supervision would probably mislead the training. As shown in Figure 1b, the multi-centroid mechanism provides more local clusters that would possibly reduce the clustering difficulty and enhance the reliability of pseudo-labels[10]. It is very helpful to learn representative proxies for contrastive loss that give effective optimization direction for

Received date: 2024-07-22 **Accepted date:** 2025-12-01

Biographies: Qiusheng Lian, PhD, Professor, research interest: computer vision, Email: lianqs@ysu.edu.cn; **Yufei Zhao**, MS, Associate Professor, research interest: herbivore production, Email: zhaoyfly@126.com.

***Corresponding author: Jianmin Zhao**, PhD, Professor, research interest: computer vision and precision livestock farming. Inner Mongolia University of Science and Technology, Arding Street No.7, Baotou, Inner Mongolia, China. Tel: +8613451326535, Email: zhao_jm@imust.edu.cn.

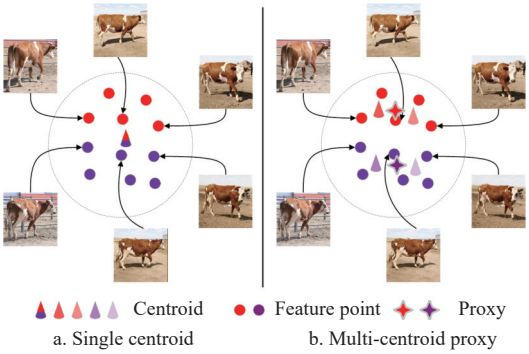

Figure 1 Comparison between the unique centroid and the multi-centroid proxy under imperfect clustering

model training.

The target of our work is to transfer the model's feature learning ability from cattle with rich annotation to the unlabeled newborn individuals or those on new farms. A mechanism of contrastive learning with multi-centroid proxy (CL-MCP) was proposed in this study, which mainly consists of a target domain feature storage module (TDFSM), a multi-centroid storage module (MCSM), and domain-specific proxy-level contrastive loss. To demonstrate the effect of our proposed CL-MCP mechanism on the identification of different unlabeled individuals in different environments, domain adaptive experiments were conducted among OpenCows2020[3], MVCAID100[5], and CNSID100[6] datasets. Moreover, unsupervised experiments were also conducted on them, as well as on the Cattle-2022[7] dataset. It is encouraging to find that our CL-MCP algorithm has good performance on both unsupervised domain adaptation and unsupervised learning tasks.

2 Related works

2.1 Cattle identification based on computer vision

Visual biometrics of cattle are unique and well-operated to identify an individual according to some observable biological characteristics[1]. With the development of computer vision, biological metrics, including nearly all the visual parts from muzzle to coat pattern, have been gradually becoming the main features in cattle identification research, as shown in Figure 2.

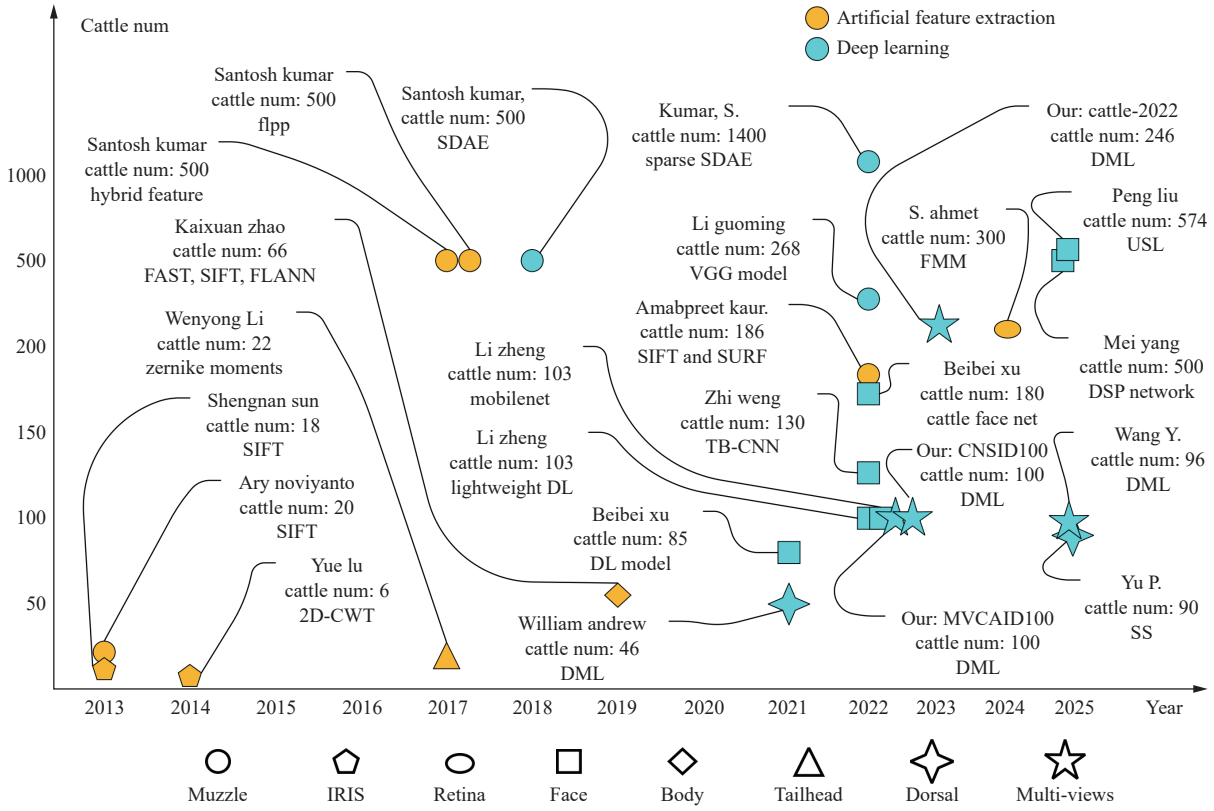

Figure 2 Development of cattle identification approaches, biometrics, and dataset scale

Muzzle image[11-19], iris image[20,21], and retina vascular pattern (RVP) image[22] all have been applied to cattle identification, and have greatly promoted the research and application of biometrics in cattle identification. Compared with the difficulties for sampling the above images, the images of parts such as cattle face[23-29], tailhead[30], profile[31], or dorsal part[3,4,32] can be easily obtained in the real farming environment and have been studied extensively in recent years.

Chinese Simmental and Holstein Cow's coat patterns are a brown coat with intrinsic white stripes, and a white coat with black stripes, respectively, and are visually akin to those generated from Turing's reaction-diffusion systems. The coat patterns above make it possible to identify individuals from any view. Our previous works give several approaches to cattle identification with the coat pattern and have made some progress in cattle recognition in the wild[5-7].

From Figure 2, it can also be seen that the datasets for cattle identification are becoming larger and larger in both the number of images and the scale of individuals. For biometrics used in precision livestock farming, it also gradually tends to use images from any view that are easily accessible in a real farming environment. The details of the used datasets are shown in Table 1.

Along with the development of biometrics, the algorithms for cattle identification based on computer vision have been studied from artificial feature extraction to deep learning with deep convolutional neural networks (DCNNs), as shown in Figure 2.

Particularly, approaches based on DML (Deep Metric Learning) have achieved great success in recent years. In the cattle face recognition task, TB-CNN[25], RetinaFace reciprocal ArcFace loss[26], lightweight deep learning model[24], and MobileNet[23] have been proposed. For the identification of cattle in the wild, SoftMax-based reciprocal triplet loss[3] is proposed to identify with dorsal images from a fixed top-view camera or unmanned aerial vehicle (UAV). In our previous works, compact loss[5], MCA loss[6], and mcSAP loss[7] based on DML are presented to recognize cattle individuals from any view in the wild.

Table 1 Details of datasets for cattle identification

Authors/Datasets	Year	Identities	Images	Details
Sun et al.[21]	2013	18	90	Iris
Lu et al.[20]	2014	6	60	Iris
Kumar et al.[16,17]	2017	500	5000	Muzzle
Li et al.[30]	2017	22	1965	Tailhead
Zhao et al.[31]	2019	66	528(video)	Body side
OpenCows2020[3]	2021	46	4376	Dorsal
Xu et al.[27]	2021	85	3000	Cattle face
Kaur et al.[12]	2022	186	930	Muzzle
Kumar et al.[14]	2022	1400	14 000	Muzzle
Li et al.[18]	2022	286	4923	Muzzle
Weng et al.[25]	2022	130	-	Cattle face
Xu et al.[26]	2022	180	4636	Cattle face
Zheng et al.[24]	2022	103	10 239	Cattle face
MVCAID100[5]	2022	100	4073	Multi-view
CNSID100[6]	2022	100	11 635	Multi-view
Cattle-2022[7]	2023	246	10 076	Multi-view
Ahmet et al.[22]	2024	300	2430	Retina
CattleFace2025[28]	2025	574	11 880	Cattle face
CattleFace-500[29]	2025	500	25 430	Cattle face
MultiCamCows2024[32]	2025	90	101 329	Dorsal
BeefCattle2024[8]	2025	96	3646	Multi-view

2.2 Unsupervised domain adaptation

Compared with domain translation-based methods[33-35], pseudo-label-based UDA methods[9,10,36-39] directly learn from the target domain, which is beneficial for target information capturing. Pseudo-label-based approaches take advantage of clustering labels generated from the target domain to continuously improve the DCNNs to learn compact clusters in the target space.

Therefore, how to improve the reliability of the clustering pseudo-labels and alleviate the impact of noise samples is of vital importance during the training stage. Self-similarity grouping (SSG) is proposed to specify multi-scale pseudo-labels with artificial local features to modify cluster centers for label generation reliability[36]. A mutual mean-teaching method is presented to fine-tune the soft pseudo-center to enhance clustering ability in the target domain with effective resistance to noise[37]. Hybrid memory is presented in self-paced contrastive learning (SpCL) to dynamically generate supervisory samples for contrastive loss, and particularly, clustering metric benchmarks for independence and compactness are given in it to measure the clustering reliability quantitatively[9]. In real-world scenarios, images of the same class always have several local clusters rather than one, e.g., cars from different perspectives and cattle of different postures. A multi-centroid representation network (MCRN) is proposed to provide more inner-class local centers to alleviate the difficulty of clustering, and it provides competitive performance in the domain adaptive person re-ID task[10].

2.3 Contrastive learning

In recent years, contrastive learning (CL) has provided an effective way for unsupervised visual representation learning[40,41] and is an intuitive approach for loss function construction in unsupervised domain adaptation[9,10].

In contrastive learning, the objective is to increase the inter-class similarity and meanwhile push the samples of different identities away by effective contrastive loss. Normalized temperature-scaled cross-entropy loss is presented in SimCLR by using random cropping, color distortion, and Gaussian blur for positive and negative samples argumentation[42]. A dynamic dictionary with a queue and a moving-averaged encoder is proposed to facilitate contrastive unsupervised learning in momentum contrastive (MoCo) that achieves competitive results on ImageNet classification[43]. In unsupervised domain adaptation, SpCL[9] provides a uniform contrastive loss that calculates the loss across all the source and target domains uniformly. Meanwhile, MCRN[10] represents independent contrastive learning to optimize the weights on the source and target domains separately.

3 Materials and methods

Given the source domain data $\mathbb{D}^s$ and the target domain data $\mathbb{D}^t$, that $\mathbb{D}^s = \left\{(x_i^s, y_i^s)|_{i=1}^{N^s}\right\}$ consists of N^s labeled samples (x_i^s, y_i^s) and $\mathbb{D}^t = \left\{(x_i^t)|_{i=1}^{N^t}\right\}$ of N^t unlabeled samples (x_i^t). A contrastive learning with multi-centroid proxy (CL-MCP) was proposed to transfer the learning ability from $\mathbb{D}^s$ to $\mathbb{D}^t$ for compact clustering in the target feature space.

In the CL-MCP mechanism, target domain feature storage module (TDFSM) was proposed to store the features of the target domain training samples and gradually optimize them to provide a more stable clustering structure, which is essential for consistent training. Multi-centroid storage module (MCSM) is presented to store and update the inner-class local centers in both the source and target domain. It provides multiple local centers for feature clustering, which is beneficial for enhancing the reliability of pseudo-labels. Particularly, the domain-specific proxy-level contrastive loss is proposed based on the representation of multi-centroid proxy, and it achieves competitive results in domain adaptive cattle identification. The details are shown in Figure 3.

3.1 Target domain feature storage module

Inspired by hybrid memory in SpCL[9], target domain feature storage module (TDFSM) is presented for memory and update of features of the target domain training samples. The features in it are gradually updated at a certain pace controlled by momentum during the training stage. The pace-adjusted update strategy is essential to ensure the stability of the target domain clustering structure, which is crucial to stabilize the training process and improve the performance of the model on the target domain.

The momentum updating strategy for features in TDFSM is represented as follows:

$$\tilde{f}_i^{ep} \leftarrow \nu_t f_i^{ep-1} + (1-\nu_t) f_i^{ep} \tag{1}$$

where, ep is the recent epoch, f_i^{ep} is the feature of the recent i-th epoch, and ν_t is the momentum for pacing the update rate of the features in TDFSM. The elements in TDFSM are updated right after each epoch with the previous features and the newly extracted ones. At the beginning of each training iteration, all the features in TDFSM are clustered using DBSCAN to generate pseudo-labels for the current training stage.

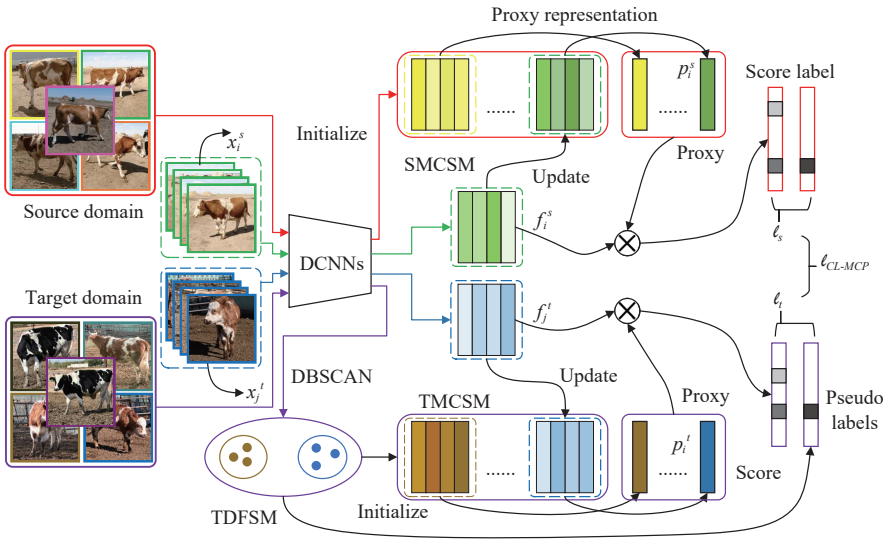

Figure 3 Structure of CL-MCP algorithm

The momentum strategy helps to obtain a more continuous and stable feature clustering structure. It is of vital importance for avoiding the training oscillation due to the poor pseudo-label stability, which in turn affects the new cluster generated by the inferior model. In Section 4, quantitative validation for the effect of the update rate of features in TDFSM is given with experiments.

3.2 Multi-centroid storage module

Due to the fact that pseudo-label changes continuously with training, it is impossible to construct a parameterizable framework for multiple centroids as SoftTriple does[44]. A multi-centroid storage module (MCSM) was proposed in this study, including SMCSM and TMCSM for the source and target domain respectively, to memorize and update inner-class local centers during training.

For the source domain, SMCSM is represented as $[...,c_i^1,...,c_i^m,...,c_i^M,...,c_{N^s}^1,...,c_{N^s}^m,...,c_{N^s}^M]^s$, where c_i^m is the m-th sub-center for the i-th class. The size of SMCSM is $M \times N^s \times d$, where M, N^s, and d are the number of inner-class local centroids, the number of classes in the source domain, and the feature dimension, respectively. SMCSM is only initialized once at the beginning of training and then gradually updated to provide nonparametric dynamic local centers for proxy representation in our proposed domain-specific contrastive loss.

Unlike the fixed structure of SMCSM, TMCSM changes after each iteration as the pseudo-labels are dynamically generated when a new iteration starts. TDFSM is as $[...,\tilde{c}_i^1,...,\tilde{c}_i^m,...,\tilde{c}_i^M,...,\tilde{c}_{N^t}^1,...,\tilde{c}_{N^t}^m,...,\tilde{c}_{N^t}^M]^t$, where, N^t is the number of classes in the target domain that changes with the clustering results during training. Thus, TDFSM is re-initialized at the beginning of each iteration and updated within it.

The initialization method of the m-th centroid for the class i is represented as

$$c_i^m = \frac{1}{|\mathbb{E}_i|}\sum_{f_j \in \mathbb{E}_i} f_j \tag{2}$$

where, $\mathbb{E}_i$ is the set of the i-th class, f_j is the feature of the j-th sample in class i. For initialization, the features are extracted from the specific domain to initialize the centroids in MCSM with Equation (2) and gradually updated during training.

In order to update all the centroids in a class synchronously, each batch samples P classes and M instances per class. The Hungarian algorithm is used to search for the optimal permutation for M instances that has the best bipartite matching for sample-centroid pairs with the highest similarity. The best bipartite matching $\tilde{\rho}$ is represented as,

$$\tilde{\rho} = \arg\max_{\rho \in \mathbb{P}_M} \sum_{i=1}^{M} \left\langle f_i, c_{\rho[i]} \right\rangle \tag{3}$$

where, $\mathbb{P}_M$ is the permutation set, $\langle , \rangle$ denotes the inner product, $c_{\rho[i]}$ is the i-th sub-center in the sub-center permutation ρ, and f_i is the feature of the i-th sample in the class. Under the constraint of Equation (3), the best permutation of samples with the highest matching degree would be arranged to update the corresponding inner-class centroid $\tilde{c}_{\tilde{\rho}[i]}$ in MCSM with

$$\tilde{c}_{\tilde{\rho}[i]} \leftarrow \nu_c c_{\tilde{\rho}[i]} + (1-\nu_c) f_i \tag{4}$$

where, ν_c is the momentum for controlling the rate of update, $c_{\tilde{\rho}[i]}$ is the i-th sub-center in the best permutation, and f_i is the best matched feature to it. Specifically, if the inner centroids are to be the same during training, such as right after the initialization, the feature vectors are randomly arranged around the corresponding centers for updating.

3.3 Domain-specific multi-centroid proxy contrastive loss

With multiple centroids clustering, it is beneficial to enhance the reliability of pseudo-labels and alleviate the impact of noise in the target domain. However, the sub-center lacks global information, and also the local centroid-based contrastive learning has difficulty in conveniently selecting representative positive and hard negatives[10]. Thus, we propose a domain-specific multi-centroid proxy-based contrastive loss to learn more informative and representative proxies to enlarge the similarities among the instances in the same class and push away the different ones for more compact clustering in the target domain.

It is assumed that there are M inner-class centroids $[c_{y_i}^1,...,c_{y_i}^m,...,c_{y_i}^M]$, and the proxy p_{y_i} of class y_i can be represented as,

$$p_{y_i} = \sum_{m=1}^{M} \frac{\exp\left(\frac{1}{\tau}\left\langle f_i, c_{y_i}^m \right\rangle\right)}{\sum_{m=1}^{M}\exp\left(\frac{1}{\tau}\left\langle f_i, c_{y_i}^m \right\rangle\right)} c_{y_i}^m \tag{5}$$

where, $c_{y_i}^m$ is the m-th local centroid of class y_i, and τ controls the Entropy regulation of multi-centroid as analyzed in SoftTriple[44] and is set to 0.05. From the representation, it can be seen that the proxy p_{y_i} would capture more global information with multiple local

centers of class y_i.

The distribution characteristics of the source domain and the target domain are different, and the cross-domain negatives have little valid information. Thus, the domain-specific contrastive loss was employed for training on the source domain and the clustered target domain with pseudo-labels separately. And we simply discard the unclustered target domain samples in the related iteration.

Given a sample $x_i^s \in \mathbb{X}^s$ from the source domain training set and its features extracted with the DCNN model, the objective of the domain-specific proxy contrastive learning on the source is represented as,

$$\ell_s = -\log \frac{\exp\left(r\left\langle f_i^s, p_{y_i}\right\rangle\right)}{\sum_{j=1}^{N^S} \exp\left(r\left\langle f_i^s, p_j\right\rangle\right)} \tag{6}$$

where, p_j is the source domain proxy of the relevant class, and N^s is the number of identities in the source training set. r is the radius of the projection hypersphere of the feature points and is set to 24.

Similarly, given $x_i^t \in \mathbb{X}^t$ from the clustered target domain training set and f_i^t as its feature, the objective on the target domain is represented as

$$\ell_t = -\log \frac{\exp\left(r\left\langle f_i^t, \tilde{p}_{\tilde{y}_i}^t\right\rangle\right)}{\sum_{j=1}^{\tilde{N}^t} \exp\left(r\left\langle f_i^t, \tilde{p}_j^t\right\rangle\right)} \tag{7}$$

where, $\tilde{p}_j$ is the proxy, and $\tilde{N}^t$ is the number of the relevant pseudo-classes from the clustered target training set. r is set to 24, the same for the source domain.

The total objective of the proposed domain-specific proxy-level contrastive learning is as,

$$\ell_{CP-MCP} = \ell_s + \ell_t \tag{8}$$

From the above, the training stage of multi-centroid proxy contrastive learning consists of several procedures, including pseudo-label generation, MCSM, TDFSM initialization, and updating. The processes for training are listed in Algorithm 1.

Algorithm 1 Training procedure of CL-CMP on unsupervised domain adaptation

Require: Source-domain training set $\mathbb{X}^s$ and target-domain training set $\mathbb{X}^t$;
Require: Initialize the backbone with ImageNet-pretrained ResNet50;
Require: Extract the features f_i^s of x_i^s in $\mathbb{X}^s$;
Require: Initialize SMCSM with f_i^s using Equation (2);
Require: Extract the features f_i^t of x_i^t in $\mathbb{X}^t$;
Require: Initialize TDFSM with f_i^t;
1. **for** $ep = 1$ **to** num_epochs **do**
2. Generate pseudo-labels in TDFSM with DBSCAN;
3. Initialize TMCSM with clustered features in TDFSM using Equation (2);
4. **for** $\left\{x_i^s\right\} \subset \mathbb{X}^s$, $\{x_i^t\} \subset \mathbb{X}^t$ **do**
5. Extract features $f_i^s = F(x_i^s;\theta)$, $f_i^t = F(x_i^t;\theta)$;
6. Calculate ℓ_{CP-MCP} with Equation (8);
7. Update TDFSM with Equation (1);
8. Update MCSM with Equation (4);
9. Optimize the network with back propagation;
10. **end for**
11. **end for**

4 Experiment results

4.1 Datasets and evaluation protocols

To show the performance of our proposed CL-MCP algorithm, a series of domain-adaptive cattle identification experiments was conducted on three datasets, including MVCAID100[5], CNSID100[6], and OpenCows2020[3]. Unsupervised experiments were also conducted on them, as well as on Cattle-2022[7]. The details of the datasets above are shown in Table 2.

Table 2 Details of the datasets for evaluation

Datasets	Train	Gallery	Query	Details	Breed
OpenCows2020[3]	2365	2122	249	Dorsal	Holstein
MVCAID100[5]	1991	1539	523	Multi-view	C. S. & Holstein
CNSID100[6]	5822	4158	1655	Multi-view	C. S.
Cattle-2022[7]	5026	3586	1464	Multi-view	C. S. & Holstein

Note: C. S. denotes Chinese Simmental.

OpenCows2020 Dataset. OpenCows2020 is a Holstein-Friesian cattle identification dataset with dorsal images from a fixed top-view camera or UAV[3]. It includes 4367 dorsal images from 46 identities. It contains only single-view dorsal images and, therefore, is only used as the target domain for domain adaptation evaluation.

Multi-view Cattle Identification Datasets. MVCAID100[5] and Cattle-2022[7], proposed in our previous works for visual identification of Chinese Simmental and Holstein in the wild, are used for the evaluation of domain adaptation and unsupervised learning. MVCAID100[5] has two breeds, and CNSID100[6] only includes identities of Chinese Simmental. Neither the identities nor the farming environments in MVCAID100 are the same as those in CNSID100. Cattle-2022 has some identities that are the same as the other two datasets, and thus is only used for unsupervised learning evaluation. All of them are available on https://pan.baidu.com/s/19hoWd__7NMLvdLNp-otDSRg (code: d3oa). Some of the samples in the datasets above are shown in Figure 4.

Open cows2020

MVCAID100

CNSID100

Cattle-2022

Figure 4 Samples in multi-view cattle identification datasets and OpenCows2020

Evaluation Protocols. In the training stage, only labeled data in the source-domain training set are provided. For domain adaptive cattle identification, cumulative matching characteristic (CMC) and mean average precision (mAP) were adopted to evaluate the performance on the target domain query set. The accuracy of the *k*-NN classifier for the query set in the target domain was also used for direct verification of the performance of cattle identification.

4.2 Implementation details

In the experiments, two NVIDIA RTX2080Ti GPUs were used as the main hardware for training and testing, and the algorithm platform was built on UBUNTU 18.04 and Pytorch 1.7.1.

The DCNNs backbone architecture used in this work was ResNet50, with weights pretrained on ImageNet. The input images were resized to 224×224 and flipped horizontally, and erased with a probability of 0.5 for data augmentation. The results in our work are reported with the 2048-dimensional feature using the GAP layer output by removing the classifier layer in ResNet50.

The model was optimized by Adam with a weight decay of 5e-4. The initial learning rate for the backbone was set to be 3.5e-4, and decays by 0.1 per 10.0 steps with a total of 30 epochs. The DBSCAN algorithm is adopted for pseudo-labels generation, and the maximum nearest neighbor distance in it is set to 0.6.

The evaluation is conducted on the query set in the target domain, and the performance is evaluated with Recall@k (k=1, 2, 4), mAP, and k-NN (k=5) accuracy.

4.3 UDA performance on the target domain

The comparison experiments were conducted with current deep metric learning approaches on cattle identification in the wild, including compact loss[5] and MCA loss[6], for evaluation of the transferring learning ability to new breeds, unseen viewpoints, or different farming environments. For evaluation of the UDA performance of compact loss and MCA loss, the features were directly extracted using ResNet50 and weights with the accuracy of 93.31%[5] and 98.13%[6], respectively. After a series of experiments, the CL-MCP method performs much better than the deep metric learning approaches for the identification of cattle of a new breed, from a different environment, or from an unseen viewpoint.

MVCAID100→CNSID100. In this experiment, MVCAID100[5] is the source domain, and CNSID100[6] is the target for evaluating adaptive ability from the source to the unseen individuals in different farm environments. The number of inner-class centroids is set to be two, and then, 32 identities and two instances per individual were sampled for a mini-batch. The momentum for MCSM v_c is set to be 0.2, and the momentum for TDFSM v_t is set to be 0.5. The results are listed in Table 3.

Table 3 Results of UDA evaluation on MVCAID100→CNSID100.

Methods	R@1	R@2	R@4	mAP/%	*k*-NN accuracy/%
Compact loss[5]	85.1	90.5	95.0	22.9	81.0
Our CL-MCP	96.1	98.4	99.5	27.0	93.1

It is encouraging to find that our proposed CL-MCP method outperforms the approach with compact loss trained on MVCAID100[5] in domain adaptation to different farming scenarios.

CNSID100→MVCAID100. In this experiment, CNSID100[6] is the source domain, and MVCAID100[5] is the target for validation of the adaptive performance from one breed to more breeds. Sixteen identities and four instances for each one are sampled in the mini-batch, and the number of inner-class centroids is four. The momentum for MCSM v_c and for TDFSM v_t is set to be 0.3 and 0.6, respectively. The results are listed in Table 4.

Our CL-MCP method greatly outperforms the approach with MCA loss trained on CNSID100. The MVCAID100 dataset contains cattle from two breeds, including Chinese Simmental and Holstein. Thus, it shows that our proposed CL-MCP approach has relatively good adaptability even with a larger gap between the different breeds in the source and target domains.

Table 4 Results of UDA evaluation on CNSID100→MVCAID100

Methods	R@1	R@2	R@4	mAP/%	*k*-NN accuracy/%
MCA loss[6]	86.4	93.1	96.9	36.7	84.3
Our CL-MCP	94.5	97.1	98.5	45.9	92.0

MVCAID100→OpenCows2020. In the experiment, MVCAID 100[5] is used as the source and OpenCows2020[3] as the target for domain adaptation evaluation from multi-views to a different single view. We set the number of inner-class centroids to be two, and then in each mini-batch we sample 32 identities and two instances for each one. The momentum for MCSM v_c is set to be 0.2, and the momentum for TDFSM v_t is set to be 0.5. The results are listed in Table 5.

Table 5 Results of UDA evaluation on MVCAID100→OpenCows2020

Methods	R@1	R@2	R@4	mAP/%	*k*-NN accuracy/%
Compact loss[5]	88.8	94.0	96.8	43.8	88.0
Our CL-MCP	97.6	98.4	99.2	51.7	97.2

It can be seen that our CL-MCP algorithm performs better than the approach with compact loss trained on MVCAID100[5] on adaptation from multi-view to an unseen view. We can also find that the proposed CL-MCP algorithm can identify the individual from any viewpoint in the natural state.

CNSID100→OpenCows2020. In this experiment, CNSID100[6] is used as the source and OpenCows2020[3] as the target to evaluate the adaptive identification performance for unseen breeds and unseen views. The number of inner-class centroids is four. Batch size is 64 with 16 classes and four images per each one. The momentum for MCSM v_c and for TDFSM v_t is set to be 0.2 and 0.6, respectively. The results are listed in Table 6.

Table 6 Results of UDA evaluation on CNSID100→OpenCows2020

Methods	R@1	R@2	R@4	mAP/%	*k*-NN accuracy/%
MCA loss[6]	90.4	93.2	95.6	42.4	87.1
Our CL-MCP	96.4	99.2	99.6	55.9	97.6

The CL-MCP proposed in this paper outperforms the approach with MCA loss[6] trained on CNSID100. It can be seen that CL-MCP has better domain adaptation ability to identify individuals of different breeds from unseen viewpoints, which is of vital importance for identifying a large number of unlabeled individuals of different breeds.

4.4 Ablation study

The local centers in MCSM can only be updated discontinuously in each mini-batch. Thus, the update range and speed are of vital importance for effective learning. Moreover, the update speed of the target-domain features in TDFSM is crucial for stable clustering for pseudo-label generation. Thus, complex experiments were conducted to analyze the impact of the update speed of target domain features, the update speed, and the range of

multi-centers on the performance of unsupervised domain adaptation.

4.4.1 Momentum v_t for TDFSM

The momentum v_t controls the speed of updating the target domain features in the TDFSM, which has a direct impact on the stability of the clustering structure. The performance was evaluated first with a change of the update momentum v_t on the MVCAID100→CNSID100 adaptive task. In the experiments, the momentum v_c for MCSM was fixed at 0.2, and v_t was adjusted in the range of 0.2 to 0.8. The results are listed in Table 7.

Table 7 Results on different settings of v_t on MVCAID100→CNSID100 (v_c = 0.2)

v_t	0.2	0.3	0.4	0.5	0.6	0.7	0.8
Top1	94.4	94.1	94.4	96.1	95.3	95.8	95.2
Top2	97.8	97.2	97.2	98.4	98.1	98.1	97.2
Top4	98.9	98.9	99.0	99.5	99.2	99.2	99.3
mAP	22.5	23.1	23.5	27.0	26.6	26.0	25.3
k-NN accuracy	92.1	91.8	91.4	93.1	92.6	91.7	91.8

It is consistent with the analysis that the update speed of the features in TDFSM directly affects the stability of the target clustering structure during training. If the target features update too slowly with a small v_t, the pseudo-labels generation would lag behind the performance. Contrary to fast updating, it would lead to turbulence of the target clustering structure that negatively affects the learning process in the target domain.

4.4.2 Momentum v_c for MCSM

The momentum v_c determines the update speed of the multi-centroid in MCSM, which has a direct impact on the learning of local centers for alleviating the clustering difficulty and generating informative and representative proxies.

The performance was evaluated with a change of the momentum v_c on the MVCAID100→CNSID100 adaptive task. In the experiments, the momentum v_t was fixed to be 0.5 based on the analysis above, and v_c was adjusted in the range from 0.1 to 0.6. The results are listed in Table 8.

Table 8 Results on different settings of v_c on MVCAID100→CNSID100 (v_t=0.5)

v_c	0.1	0.2	0.3	0.4	0.5	0.6
Top1	95.5	96.1	95.7	95.4	95.8	93.4
Top2	97.6	98.4	97.9	97.2	98.0	96.6
Top4	98.6	99.5	99.2	99.2	99.3	98.4
mAP	25.9	27.0	24.4	24.1	23.5	22.3
k-NN accuracy	92.0	93.1	91.2	91.1	90.6	89.0

As can be seen from the analysis in Table 8, the proper setting of the update speed of the multi-centroid in MCSM was beneficial for effective inner local centers learning, and is good for clustering and proxy representation. Along with the analysis of multi-centroid learning, the experiments were continued, and the effects of the multi-centroid number M and the update scale within a mini-batch were discussed.

4.4.3 Multi-centroid number M

Multiple local centroids provide more local clusters, which is helpful to reduce the clustering difficulty and improve the reliability of the pseudo-labels. M was set the same in the source and target domain, and experiments were conducted to evaluate the effects of M on the MVCAID100→CNSID100 adaptive task. v_c and v_t were set to 0.2 and 0.5, respectively. The results are listed in Table 9.

Table 9 Results on different multi-centroids number M on MVCAID100→CNSID100

M	1	2	3	4	5	6
Top1	93.8	96.1	94.9	95.9	94.6	94.7
Top2	97.3	98.4	97.1	98.0	97.2	97.5
Top4	98.9	99.5	98.6	99.2	98.9	99.0
mAP	23.6	27.0	23.7	21.5	20.6	19.0
k-NN accuracy	91.7	93.1	90.6	91.5	90.5	89.1

From the results in Table 9, it can be seen that more centers with large M help to capture more inner-class local centers for more reliable generation of pseudo-labels. However, too many of them would form many local clusters for the samples to jump among them, which eventually affects the stability of the cluster structure. Moreover, in order to update all the centroids in a class synchronously, each batch samples P classes and M instances per class, where M is equal to the number of local centers. Thus, due to the fixed batch size $P\times M$, when the number of inner-class centers M is enlarged, P has to be set smaller accordingly, which directly reduces the update range of the proxies each time. Then we continued to do experiments on the analysis of the balance between the sampling number of identity number P and the instance number M within the mini-batch.

4.4.4 Effect of M and P in a mini-batch

With the above discussions about the sampling balance between P and M in a mini-batch, a series of experiments are conducted to quantitatively analyze the effect of the update scale of the centroid in the MCSM. In the experiment, the update scale is fixed with P=8 to evaluate the effect of the multi-centroid number M by continuously increasing it from 1 to 8. The results are shown in Table 10.

Table 10 Effects of M with fixed update scale as P=8 on MVCAID100→CNSID100

M	1	2	3	4	5	6	7	8
Top1	89.4	92.1	93.4	93.2	92.9	93.5	95.1	92.9
Top2	93.2	96.0	96.0	96.6	96.7	96.3	96.7	95.8
Top4	95.8	98.4	98.4	98.2	98.4	98.1	98.5	97.9
mAP	16.2	17.5	17.3	17.4	17.6	18.4	18.6	17.1
k-NN accuracy	80.6	85.8	87.7	89.0	88.8	88.4	89.1	87.6

From Table 10, it can be seen that the performance improves with the increasing number of local centroids. It shows that our multiple centroid learning is essential for more reliable cluster and pseudo-labels generation, which are much more important for the performance.

However, due to the fixed batch size constrained by the hardware, larger M has to reduce the P, which makes the proxy update scale decrease. This would result in inferior training in which only a small number of P proxies are updated in each mini-batch, and the others cannot get updated in time. We experimentally confirm this with a series of experiments by changing the scale controller P under a fixed M=2. The details are listed in Table 11.

Table 11 Effects of P with fixed centroid number as M equals 2 on MVCAID100→CNSID100

M	8	16	24	32
Top1	91.5	94.9	94.3	96.1
Top2	96.0	97.5	96.9	98.4
Top4	98.4	98.9	98.5	99.5
mAP	17.5	21.0	22.0	27.0
k-NN accuracy	85.8	89.9	90.3	93.1

It can be seen that the performance improves a lot with a larger scale of proxies being updated each time. Then it effectively provides more reliable and timely updated proxies for our multi-centroid proxy-level contrastive loss to guide the optimization. Thus, the balance between the number of multi-centroid M and the scale controller P in a mini-batch is helpful to get good performance with a certain hardware. In particular, if the number of samples in a class under a pseudo-label is less than M, the samples in this ID will be abandoned in this iteration.

4.5 Reliability of pseudo-labels generation

Pseudo-labels of the target domain participate in the training as the exact supervision criteria. Therefore, the reliability of the pseudo-labels is one of the key factors affecting the performance. Our proposed CL-MCP method provides more local clusters to lower the difficulty of clustering and enhance the reliability of the pseudo-labels for fewer noisy samples. We conducted experiments to compare our CL-MCP method with the independence and compactness metrics proposed in SpCL[9] for the evaluation of the reliability of pseudo-labels.

DBSCAN is used to generate a standard cluster set, a loose cluster, and a tight cluster with different neighbor-distance parameters of 0.60, 0.62, and 0.58, respectively, as in SpCL[9]. Then we set up the independence and compactness thresholds, the same as those in SpCL, for determining reliable clusters. We conducted pseudo-label reliability comparing experiments on the MVCAID100→CNSID100 adaptive task, and the results are listed in Table 12.

Table 12 Results on pseudo-labels reliability enhancement methods on MVCAID100→CNSID100

Methods	R@1	R@2	R@4	mAP/%	k-NN accuracy/%
SpCL[9]	90.6	98.1	99.2	23.6	92.2
CL-MCP*(uni-centroid)	93.8	97.3	98.9	23.6	91.7
CL-MCP**(uni-centroid+metrics)	95.2	97.8	99.2	25.7	91.7
CL-MCP*** (multi-centroid+metrics)	94.6	97.4	98.8	24.9	91.5
CL-MCP	96.1	98.4	99.5	27.0	93.1

It is clearly indicated in Table 12 that CL-MCP* with a single centroid but without independence and compactness metrics[9] has the poorest performance due to a lack of reliability cluster enhancement. The performance of CL-MCP** improves with the help of reliability metrics, and it shows the importance of generating dependable pseudo-labels. However, the performance of CL-MCP*** with both multi-centroid learning and clustering metrics in SpCL[9] slides down. The reason for this is that the informative positives, located among the local centroids, tend to jump back and forth, and then they would be simply discarded as unstable samples by the cluster metrics in SpCL[9]. Therefore, the performance of CL-MCP*** falls compared with that of CL-MCP**. It is encouraging to find that our CL-MCP method has the best performance among them and gives a simple way to improve the reliability of the pseudo-labels.

4.6 Unsupervised cattle identification

To validate the performance of our proposed CL-MCP method on unsupervised cattle identification, we conduct experiments on OpenCows2020[3], MVCAID100[5], CNSID100[6], and Cattle-2022[7], respectively.

The objective of CL-MCP for unsupervised identification is presented as,

$$\ell_{CP-MCP-USL} = \ell_t \tag{9}$$

The detailed processes for training with CL-MCP on unsupervised learning are listed in Algorithm 2.

Algorithm 2 Training procedure of CL-CMP on unsupervised learning

Require: Target-domain training set;
Require: Initialize the backbone with ImageNet-pretrained ResNet50;
Require: Extract the features f_i^t of x_i^t in $\mathbb{X}^t$;
Require: Initialize TDFSM with f_i^t;
1. **for** $ep = 1$ **to** num_epochs **do**
2. Generate pseudo-labels in TDFSM with DBSCAN;
3. Initialize TMCSM with clustered features in TDFSM using Equ.2;
4. **for** $\{x_i^t\} \subset \mathbb{X}^t$ **do**
5. Extract features $f_i^t = F(x_i^t;\theta)$;
6. Calculate $\ell_{CP-MCP-USL}$ with Equation (9);
7. Update TDFSM with Equation (1);
8. Update TMCSM with Equation (4);
9. Optimize the network with back propagation;
10. **end for**
11. **end for**

For training on unsupervised learning, the backbone model, optimizer, model weight initialization, and learning rate settings are the same as those on domain adaptation experiments. The updated momentum v_t and v_c for TDFSM and MCSM, and the number of multi-centroid M, are as shown in Table 13.

Table 13 Parameters of CL-MCP under unsupervised learning

Methods	v_c	v_t	M
OpenCows2020[3]	0.3	0.8	2
MVCAID100[5]	0.2	0.6	4
CNSID100[6]	0.2	0.9	2
Cattle-2022[7]	0.2	0.9	2

The evaluation of unsupervised identification is conducted on the query set in the target domain, and the performance is evaluated with Recall@k (k=1, 2, 4), mAP, and k-NN accuracy. The results on the datasets are listed in Table 14.

Table 14 Results of CL-MCP under unsupervised learning

Methods	R@1	R@2	R@4	mAP/%	k-NN accuracy/%
OpenCows2020[3]	97.2	99.2	99.6	52.8	98.4
MVCAID100[5]	87.2	92.4	95.4	30.7	80.1
CNSID100[6]	94.9	97.2	98.9	21.5	90.3
Cattle-2022[7]	77.9	85.1	90.8	24.8	70.2

It is encouraging to find out that our proposed CL-MCP method not only has better performance in unsupervised domain adaptation, but also in unsupervised cattle identification. Particularly, our CL-MCP performs even better than SoftMax-based reciprocal triplet loss with the k-NN accuracy of 98.19% under supervised learning[3] on OpenCows2020. However, the accuracy is relatively low on the MVCAID100 and Cattle-2022 datasets. That is because the images in them are from two breeds of cattle, including Chinese Simmental and Holstein. This also increases the challenge of clustering in the target domain and brings more noise into the pseudo-labels. Thus, learning with less reliable labels results in relatively poor performance on them.

5 Conclusions and discussion

In this study, a multi-centroid proxy contrastive learning was introduced for unsupervised domain adaptive cattle identification. The CL-MCP method provides more local centers for alleviating the clustering difficulties and enhancing the reliability of the pseudo-labels generation. Meanwhile, the TDFSM is presented to offer a consistently updated and stable clustering structure, and the MCSM to provide effective multiple centers for clustering and proxy representation. Finally, the domain-specific proxy-level contrastive loss gives sufficient supervision for more compact clustering. A series of experiments were conducted among OpenCows2020, MVCAID100, and CNSID100 datasets for domain adaptive evaluation and on OpenCows2020, MVCAID100, CNSID100, and Cattle-2022 for unsupervised learning. The results convincingly demonstrate the effect of the proposed approach for the identification of unlabeled cattle from any viewpoint in the wild.

The population and the image number in the datasets for evaluation for our unsupervised domain adaptation meet the needs for cattle identification in small and medium-sized pastures. However, comparing with other UDA datasets for person or vehicle re-ID, the scale of cattle identification datasets is still too small, which limits the ability to investigate the performance. Therefore, we will continuously increase the number of individuals and images of cattle in real farming scenarios to support the related research.

And still, the cluster reliability metrics remain limited for the pseudo-labels generation to quantitatively measure the dependence of the cluster. In our future works, we will work on the representation for the reliability metrics of the cluster and continue to study the methods for alleviating the impact of noise.

Moreover, in UDA methods, there is still a lack of gap metrics for measuring the difference between the source and target domains. The relationship between the domain gap and the adaptation ability is still not clear. Therefore, we will be theoretically studying the issues of the gap measurement and the adaptation ability related to it.

Acknowledgements

This work was financially supported in part by the National Natural Science Foundation of China (Grant No. 32460858), the Natural Science Foundation of Inner Mongolia Autonomous Region in China (Grant No. 2023MS06014 and Grant No. 2019LH06006), Inner Mongolia Autonomous Region Science and Technology Plan Project in China (Grant No. 2021GG0224), Fundamental Research Funds for Inner Mongolia University of Science & Technology[088], and Inner Mongolia Autonomous Region Science and Technology Special Project in China (Grant No. 2019ZD025).

[References]

[1] Awad A I. From classical methods to animal biometrics: A review on cattle identification and tracking. Computers and Electronics in Agriculture, 2016; 123: 423–435.

[2] Qiao Y L, Kong H, Clark C, Lomax S, Su D, Eiffert S, et al. Intelligent perception for cattle monitoring: A review for cattle identification, body condition score evaluation, and weight estimation. Computers and Electronics in Agriculture, 2021; 185: 106143.

[3] Andrew W, Gao J, Mullan S, Campbell N, Dowsey A W, Burghardt T. Visual identification of individual Holstein-Friesian cattle via deep metric learning. Computers and Electronics in Agriculture, 2021; 185: 106133.

[4] Andrew W, Greatwood C, Burghardt T. Visual localization and individual identification of Holstein-Friesian cattle via deep learning. In: 2017 IEEE International Conference on Computer Vision Workshops (ICCVW), Venice: IEEE, 2017; pp.2850–2859. doi: 10.1109/ICCVW.2017.336.

[5] Zhao J M, Lian Q S. Compact loss for visual identification of cattle in the wild. Computers and Electronics in Agriculture, 2022; 195: 106784.

[6] Zhao J M, Lian Q S, Xiong N N. Multi-center agent loss for visual identification of Chinese Simmental in the wild. Animals, 2022; 12(4): 459.

[7] Zhao J M, Lian Q S. Multi-centers softmax reciprocal average precision loss for deep metric learning. Neural Computing and Application, 2023; 35(16): 11989–11999.

[8] Wang Y W, Mücher S, Wang K W, Wang W S, Kooistra L. Deep metric learning for individual cattle identification using coat patterns: Proposal for a best practice. Computers and Electronics in Agriculture, 2025; 238: 110754.

[9] Ge Y X, Zhu F, Chen D P, Zhao R, Li H S. Self-paced contrastive learning with hybrid memory for domain adaptive object re-id. arXiv, 2014; arXiv: 2006.02713. doi: 10.48550/arXiv.2006.02713.

[10] Wu Y H, Huang T T, Yao H T, Zhang C, Shao Y J, Han C C, et al. Multi-centroid representation network for domain adaptive person re-ID. In: Proceedings of the 36th AAAI Conference on Artificial Intelligence, 2022; pp.2750-2758. doi: 10.1609/aaai.v36i3.20178.

[11] Kaur A, Kumar M, Jindal M K. Cattle identification with muzzle pattern using computer vision technology: a critical review and prospective. Soft Computing, 2022; 26: 4771–4795.

[12] Kaur A, Kumar M, Jindal M K. Shi-Tomasi corner detector for cattle identification from muzzle print image pattern. Ecological Informatics, 2022; 68: 101549.

[13] Kumar S, Chaube M K, Kumar S. Secure and sustainable framework for cattle recognition using wireless multimedia networks and machine learning techniques. IEEE Transactions on Sustainable Computing, 2022; 7(3): 696–708.

[14] Kumar S, Kumar S, Shafi M, Chaube M K. A novel multimodal framework for automatic recognition of individual cattle based on hybrid features using sparse stacked denoising autoencoder and group sparse representation techniques. Multimedia Tools and Applications, 2022; 81(21): 31075–31106.

[15] Kumar S, Pandey A, Sai Ram Satwik K, Kumar S, Singh S K, Singh A K, Mohan A. Deep learning framework for recognition of cattle using muzzle point image pattern. Measurement, 2018; 116: 1–17.

[16] Kumar S, Singh S K. Automatic identification of cattle using muzzle point pattern: A hybrid feature extraction and classification paradigm. Multimedia Tools and Applications, 2017; 76: 26551–26580.

[17] Kumar S, Singh S K, Singh R S, Singh A K, Tiwari S. Real-time recognition of cattle using animal biometrics. Journal of Real-Time Image Processing, 2017; 13(3): 505–526.

[18] Li G M, Erickson G E, Xiong Y J. Individual beef cattle identification using muzzle images and deep learning techniques. Animals, 2022; 12(11): 1453.

[19] Noviyanto A, Arymurthy A M. Beef cattle identification based on muzzle pattern using a matching refinement technique in the SIFT method. Computers and Electronics in Agriculture, 2013; 99: 77–84.

[20] Lu Y, He X F, Wen Y, Wang P. A new cow identification system based on iris analysis and recognition. International Journal of Biometrics, 2014; 6: 18–32.

[21] Sun S N, Yang S C, Zhao L D. Noncooperative bovine iris recognition via SIFT. Neurocomputing, 2013; 120: 310–317.

[22] Saygili A, Cihan P, Ermutlu C Ş, Aydin U, Aydın U. CattNIS: Novel identification system of cattle with retinal images based on feature matching method. Computers and Electronics in Agriculture, 2024; 221: 108963.

[23] Li Z, Lei X M. Cattle face recognition under partial occlusion. Journal of Intelligent and Fuzzy Systems, 2022; 43: 66–77.

[24] Li Z, Lei X M, Liu H. A lightweight deep learning model for cattle face recognition. Computers and Electronics in Agriculture, 2022; 195: 106848.

[25] Weng Z, Meng E S, Liu S Q, Zhang Y, Zheng Z Q, Gong C L. Cattle face recognition based on a two-branch convolutional neural network. Computers and Electronics in Agriculture, 2022; 196: 106871.

[26] Xu B B, Wang W S, Guo L F, Chen G P, Li Y F, Cao Z, et al. CattleFaceNet: A cattle face identification approach based on Retinaface and Arcface loss. Computers and Electronics in Agriculture, 2022; 193: 106675.

[27] Xu B B, Wang W S, Guo L F, Chen G P, Wang Y W, Zhang W J, et al. Evaluation of deep learning for automatic multi-view face detection in

cattle. Agriculture, 2021; 11(11): 1062.

[28] Liu P, Zhao J M. Part-attention-based pseudo-label refinement reciprocal compact loss for unsupervised cattle face recognition. Electronics, 2025; 14(12): 2343.

[29] Yang M, Li Q, Zhao J M, Wang Y M, Gao J F, Xue F F. Dual-stream modality prompt network for individual cattle identification. International Journal of Machine Learning and Cybernetics, 2025; 16: 10925–10938.

[30] Li W Y, Ji Z T, Wang L, Sun C H, Yang X T. Automatic individual identification of Holstein dairy cows using tailhead images. Computers and Electronics in Agriculture, 2017; 142: 622–631.

[31] Zhao K X, Jin X, Ji J T, Wang J, Ma H, Zhu X F. Individual identification of Holstein dairy cows based on detecting and matching feature points in body images. Biosystems Engineering, 2019; 181: 128–139.

[32] Yu P, Burghardt T, Dowsey A W, Campbell N W. Holstein-Friesian re-identification using multiple cameras and self-supervision on a working farm. Computers and Electronics in Agriculture, 2025; 237(PartB): 110568.

[33] Chen Y B, Zhu X T, Gong S G. Instance-guided context rendering for cross-domain person re-identification. In: 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul: IEEE, 2019; pp.232–242. doi: 10.1109/ICCV.2019.00032.

[34] Wei L H, Zhang S L, Gao W, Tian Q. Person transfer GAN to bridge domain gap for person re-identification. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City: IEEE, 2018; pp.79–88. doi: 10.1109/CVPR.2018.00016.

[35] Zou Y, Yang X D, Yu Z D, Vijaya Kumar B V K, Kautz J. Joint disentangling and adaptation for cross-domain person re-identification. In: European Conference on Computer Vision (ECCV), 2020; pp.87–104. doi: 10.1007/978-3-030-58536-5_6.

[36] Fu Y, Wei Y C, Wang G S, Zhou Y Q, Shi H H, Uiuc U, et al. Self-similarity grouping: A simple unsupervised cross domain adaptation approach for person re-identification. In: 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019; pp.6111–6120. doi: 10.1109/ICCV.2019.00621.

[37] Ge Y X, Chen D P, Li H S. Mutual mean-teaching: Pseudo label refinery for unsupervised domain adaptation on person re-identification. International Conference on Learning Representations (ICLR), 2020. doi: 10.48550/arXiv.2001.01526.

[38] Song L C, Wang C, Zhang L F, Du B, Zhang Q, Huang C, et al. Unsupervised domain adaptive re-identification: Theory and practice. Pattern Recognition, 2020; 102: 107173.

[39] Zhang X Y, Cao J W, Shen C H, You M Y. Self-training with progressive augmentation for unsupervised cross-domain person re-identification. In: 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul: IEEE, 2019; pp.8221–8230. doi: 10.1109/ICCV.2019.00831.

[40] Chen X L, He K M. Exploring simple Siamese representation learning. In: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville: IEEE, 2021; pp.15745–15753. doi: 10.1109/CVPR46437.2021.01549.

[41] Tian Y L, Krishnan D, Isola P. Contrastive multiview coding. In: European Conference on Computer Vision (ECCV), 2020; pp.776–794. doi: 10.1007/978-3-030-58621-8_45.

[42] Chen T, Kornblith S, Norouzi M, Hinton G. A simple framework for contrastive learning of visual representations. In: International Conference on Machine Learning, 2020; arXiv: 2002.05709. doi: 10.48550/arXiv.2002.05709.

[43] He K M, Fan H Q, Wu Y X, Xie S N, Girshick R. Momentum contrast for unsupervised visual representation learning. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle: IEEE, 2020; pp.9726–9735. doi: 10.1109/CVPR42600.2020.00975.

[44] Qian Q, Shang L, Sun B G, Hu J H, Tacoma T, Li H. Softtriple loss: Deep metric learning without triplet sampling. In: 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul: IEEE, 2019; pp.6449–6457. doi: 10.1109/ICCV.2019.00655.