A Purely Visual Re-ID Approach for Bumblebees (*Bombus terrestris*)

Parzival Borlinghaus [a,*], Frederic Tausch [b], Luca Rettenberger [c]

[a] *Institute for Operations Research, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany*
[b] *apic.ai GmbH, Karlsruhe, Germany*
[c] *Institute for Automation and Applied Informatics (IAI), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany*

ARTICLE INFO

ABSTRACT

Entomologists have widely applied re-identification techniques to better understand insects and their interaction with the environment. While humans can re-identify other humans and some mammals quite well, entomologists rely on gluing markers on insects to perform this task. This paper presents an approach for purely visual re-identification of bumblebees (*Bombus terrestris*) without the need to use markers. Non-invasive identification methods offer the possibility to observe the interaction of bumblebees with their environment without disturbance. Both a CNN model and a simple body shape model were used to investigate how they can be re-identified within a colony. The best-performing model, BumbleNet, correctly identifies more than two-thirds (CMC-1 score) of the individuals. Bumblebees are known for their substantial variations in body shape. To understand whether other features can also play a role in re-identification, different augmentations are applied during the training of BumbleNet. It was found that non-body-shape features increased the performance of BumbleNet by 25 percentage points (CMC-1 score). This also explains the observed superiority of the CNN-based BumbleNet compared to the BumbleShape model, that is solely based on body size parameters.

1. Introduction

Animal re-identification (re-ID) is a laborious, albeit useful, task. Re-ID of single individuals allows to draw conclusions about the diversity, relative abundance and behavioral patterns of different animal species. Re-ID based on camera traps is particularly popular because it is less invasive and can reduce the amount of work required. In addition, camera-based methods also offer the possibility of machine-assisted analysis of the video data [1]. Machine-assisted animal re-ID has been successfully demonstrated for mammals in many cases [2–5], including apes, whales and elephants. However, re-ID is not limited to mammals; promising examples of insect re-ID already exist. Arbuckle et al. [6] developed the Automatic Bee Identification System (ABIS), which can reliably identify bee species based on wing patterns. From this, Francoy et al. [7] showed that these patterns also allow the identification of the Africanized bee. Furthermore, Kastberger et al. [8] have shown that wing patterns can be used not only to classify bee species but also to distinguish individuals of the giant honeybee. Murali et al. [9] succeeded in re-identifying fruit flies using a convolutional neural network (CNN). In a recently published honeybee re-ID dataset, Chan et al. [10] used the abdomen as a re-ID feature and highlighted the importance of time invariant features.

Surprisingly, to our knowledge, there are no attempts to transfer visual animal re-ID to the bumblebee domain, especially since it is known that bumblebees differ much more in body shape and live in much smaller colonies than honeybees [11]. This is despite the fact that bumblebees, as important pollinators, have been the subject of repeated research. Mola and Williams [12] review methods for studying the movements of bumblebees, including QR codes, paint stains or RFID chips, all of which being particulary important for ecotoxicological studies as required for pesticide testing [13]. However, tagging individuals is extremely laborious and only possible for a small number of individuals [14]. Furthermore, it cannot be ruled out that tagging individuals may affect the observed behaviour. For these reasons, it is desirable to replace the existing procedures with a purely visual re-identification.

This paper comprises two contributions. (1) We investigate to what extent bumblebees, within a colony, are re-identifiable on visual material and (2) which features are of decisive importance. To address these research questions, a Convolution Neuronal Network (CNN) based model, a simple body shape model and a random baseline model were benchmarked using the *Bumblebee Re-ID Dataset* [15].

2. Material and methods

For the purpose of determining behavioural patterns, entomologists are interested in identifying bumblebees as they re-enter the hive [11, 16]. The underlying problem can thus be characterised as a bumblebee retrieval problem across the same camera. In an actual use case, a bumblebee would be captured by the camera when leaving the hive and some images of this individual would be stored together with an ID in the so-called gallery for later retrieval. Images of returning bumblebees (probe images) are then compared against the gallery. There may be several images of different bumblebees in the gallery at this time. Since bumblebees are registered when they fly out, it can be assumed that in a real application all incoming bumblebees are already known to the system, which motivates the classification as closed re-ID [17].

The re-ID task is commonly divided into object detection, tracking and id retrieval [18]. While the computer vision community already provides solutions for the first two tasks [19–21], this work focuses on the id retrieval, i.e. algorithms that output a ranked candidate list of potential matches for any probe image. In the following, it is explained how re-ID models can be evaluated and how the corresponding bumblebee data set is composed.

2.1. Evaluation metrics

When evaluating re-ID algorithms, the Cumulative Matching Characteristics (CMC) curve is often used. CMC-k represents the probability that a query identity (probe) appears in the top-k ranked retrieved results (candidate list) [18,22]. The candidate list is generated by the re-ID algorithm and contains images from the gallery set ranked by the matching probability assigned by the algorithm. Regardless of how many true matches there are in the gallery, only the first match is taken into account in the CMC calculations, making it impossible to assess recall ability. This is acceptable in practice if the users' main concern is to find a match at the top of the ranking, but is otherwise undesirable [18]. This disadvantage does not apply to the mean Average Precision metric (mAP). In contrast to the CMC score, mAP is able to differentiate between algorithms that have the first correct match at the same position by also taking further matches and their positions into account [22]. Since CMC scores are easy to understand, they are reported together with the more precise mAP.

2.2. Description of dataset

The dataset was previously published by Tausch et al. [15]. The images were filtered and finally show 99 individuals from two commercial bumblebee hives with more than 22 images available per individual. The dataset was split randomly into a test and training dataset, keeping the ratio of male and female individuals constant.

In both the test and training datasets, $g = 17$ images were selected per individual as gallery images and $p = 5$ images were selected as probe images. To minimise temporal features in the image sequence of a bumblebee, gallery images were randomly selected from the first $(n - g - p)/2 + g$ images and probe images from the remaining ones, where n denotes the number of available images. Training was performed on the first $(n - g - p)/2 + g$ images of the training dataset. There are a minimum of 22 images per individual (M=69.6, STD=65.3).

2.3. Re-ID models

Three bumblebee re-ID models were developed and compared. While the BodyShape model receives a list of body features as input, the BumbleNet model has to learn suitable features from the image data itself. Furthermore, the two models are compared with a third model that serves as a random baseline.

2.3.1. BodyShape model

The BodyShape model consists of a feature extractor and a learned metric. After applying an image segmentation, a rotated bounding box and a rotated ellipse was fitted to the resulting segmentation map. This allowed the extraction of several features that form the input for the BodyShape model (segmentation area, segmentation perimeter, enclosing rotated bounding box width, enclosing rotated bounding box height, enclosing rotated ellipse width and enclosing rotated ellipse height). In order to perform a classification based on these 7 properties, a distance metric was learned that outputs small distances for equal IDs and large distances for different IDs. For this purpose, the "Keep It Simple and Straightforward Method" (KISSME) by [23] was applied.

However, attention must be paid to the extraction of the body characteristics. It is important to note here that the extremities (i.e. wings and legs) must not have any influence on the determined body size due to their potential for change. This means that it is not sufficient to calculate a binary mask by thresholding or background subtraction and then determine the stated characteristics. Instead of this error-prone procedure, a UNet architecture similar to the one presented by Ronneberger et al. [24] was trained to segment the bumblebee's body by removing legs and wings. Fig. 1 shows the input, label and output of the segmentation net. Considering the average body size of bumblebees of $95 mm^2$ [11], the image shows a very large specimen. The average deviation of the computed bumblebee size from the labelled size, converted into millimetres, is $18 mm^2$.

2.3.2. BumbleNet model

FaceNet by Schroff et al. [25] is a face recognition system which employs CNNs to map images of human faces to a dense features space that is descriptive for the person depicted. Such a feature vector for a particular image is called its embedding. The goal is to obtain a function that maps faces of the same person to feature vectors that are as similar as possible while being dissimilar to other individuals. Thus, the embedding can be used to re-identify a particular individual. Using FaceNet to re-identify bumblebees is a sensible choice, as CNN-based approaches have already produced promising results when applied to animals [1].

To obtain the desired embedding, the triplet loss is employed [25]. The triplet loss minimizes the distance between the embeddings of an anchor image and a positive image, while maximizing the distance to a negative image. The positive image shows the same bumblebee as the anchor image while the negative image is guaranteed to show a different individual. Fig. 2 illustrates the triplet loss in practice. To obtain the embeddings we use the ResNet-18 [26] architecture as backbone and an embedding size of 128, which both have proven to be a sensible choice in a conducted hyperparameters search. This neural network will be referenced as BumbleNet in the following.

In total, three variants of the BumbleNet were trained, which only differ in terms of pre-processing of the training data. In doing so, the network can be made more robust for real-world applications and provide insights into research question (2).

Similar to the BodyShape model, it makes sense to use the segmented bumblebee for training rather than the raw bumblebee image, except that the rationale is different: by having a large number of images of
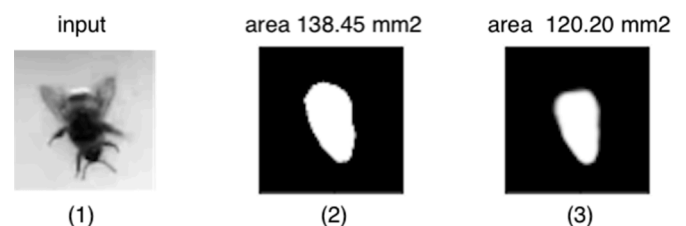

Fig. 1. Foreground segmentation using a UNet. (1) Input, (2) labelled contours, (3) UNet output. Note the desired absence of legs and wings.
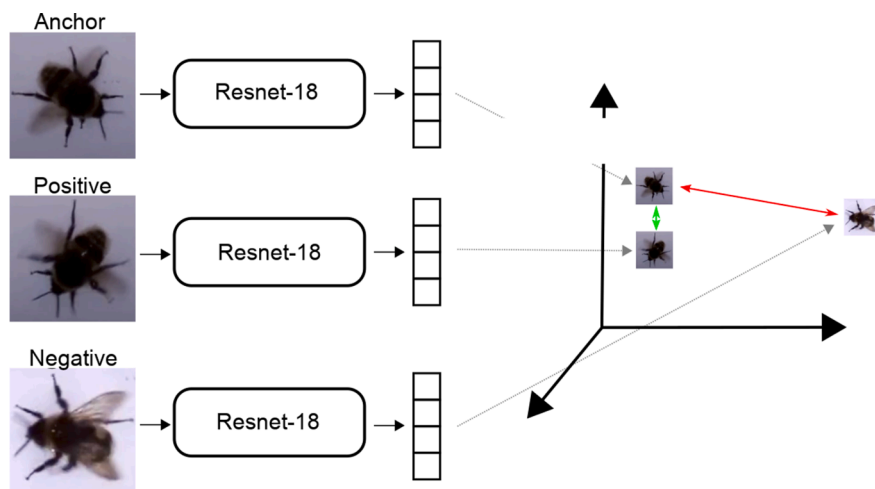
Fig. 2. Illustration of triplet loss. (Left) Two images of the same bumblebee (anchor and positive) are compared to an image of a different bumblebee (negative). (Center) The ResNet-18 provides the embeddings for each of the three images. (Right) In the embedding space, the anchor and the positive image should be as close to each other as possible while the negative image should be far away from both. During training, the model learns to find better embeddings so that the green distances are minimised and the red distances are maximised.

each individual, the network can learn that the position of the wings and legs is not a helpful cue for re-identification. What cannot be learned from limited training data, however, is that bumblebees in the real world could leave the hive without pollen and later return carrying pollen and thus significantly change their appearance. By segmenting the body, this pitfall can be eliminated at the cost of lost detail. In attempt to quantify the information loss of the segmentation, the performance is given with and without (see BumbleNetRaw and BumbleNetSeg).

Additionally, to obtain insights into which features are of decisive importance, a third variant of the BumbleNet model was trained. When training the BumbleNetSegScaled model, scaling was added to the list of possible input augmentations. Randomly scaling the images between 80% and 130% of the initial dimensions prevents the model to use the individual's size as a feature. A comparison with and without the scaled inputs allows statements about the relevance of this feature. Augmentations that do not affect the bumblebee shape are used in all three cases, including slight colour jittering, gaussian noise, contrast variations and rotation. A comparison of the models can be found in Table 1.

2.3.3. Random baseline model

To show that a model has learned important appearance descriptors, it must show better results than a random baseline model. This role is accomplished by a model that generates randomly ordered lists of candidates. The model's performance is then calculated as the average mAP or CMC score of all possible candidate lists.

Table 1
Comparison of the three BumbleNet variants.

BumbleNet variants	Model characteristics
BumbleNetRaw	The training data is neither scaled nor segmented. This model has the most information available to find potential re-ID features. However, the model could learn unwanted features that are not related to the bumblebee itself but are somehow present in the image series, like special lighting conditions or distinctive debris. BumbleNetRaw could be confused by bumblebees carrying or not carrying pollen in real applications.
BumbleNetSeg	The training data was preprocessed by segmenting the bumblebee's body. This removes noise (and possibly information) from the input image and makes the network more robust for real-world applications.
BumbleNetSegScaled	The training data was preprocessed as described above, but scaling was added to the list of input augmentation methods. This prevents the model from using body size as a feature and forces it to learn other characteristics. In this way, the influence of the body shape feature on the re-id task can be determined.

3. Results

Fig. 3 shows the CMC accuracy for all models. With a CMC-1 score of 0.6792 BumbleNetRaw is the model that gives the best results on the test data. However, segmenting the training data lowers the CMC-1 score by 42%. This shows that by segmenting, the model loses information that would otherwise improve re-identification in the test data. As expected, the CMC score improves for higher values of k. For CMC-10, the accuracy already exceeds 88% (BumbleNetRaw) (see Table 2). If the body size feature is omitted from BumbleNetSeg, which was implemented by randomly scaling the training data when training the BumbleNetSegScaled model, the CMC-1 score drops by another 37%. If visual features are not used at all (BodyShape model), the performance drops by almost 66% compared to the best model. A comparison with the random baseline, however, shows that all models succeeded in re-identifiying bumblebees based on distinct characteristics.

A look at the mAP results in Table 2 does not provide any new information about the choice of the best (or second best) model: The mAP scores of these models are broadly consistent with the CMC scores, so there is nothing to suggest that considerations of the model's recall capability should influence model choice. However, the BodyShape model may be preferred over BumbleNetSegScaled due to a better mAP score.

4. Implications for application

Real-world applications allow some simplifications that improve re-ID. These include (1) reducing the gallery set and (2) statistically modeling the probability of the presence of single bumblebees. Detecting an individual in a small (gallery) set is easier than in a large set. The data set used here includes two bumblebee colonies. In reality, however, it is only necessary to distinguish the individuals of one colony. The set can be reduced even further if a dynamic gallery set is used. The intuition then is that one does not compare a returning bumblebee with all known bumblebees, but only with the subset that previously left the hive. Modeling the presence of individuals helps to learn from contradictory observations, for example, that the same bumblebee leaves twice without returning in between. Since each re-ID involves a confidence, it might make sense to believe not the most likely candidate bumblebee, but the one that does not contradict previous observations. If particularly accurate data is required, a study can be limited to bumblebees that are particularly easy to identify. If the selection of bumblebees with extreme shapes does not affect the study, it may be possible to obtain even more accurate results.
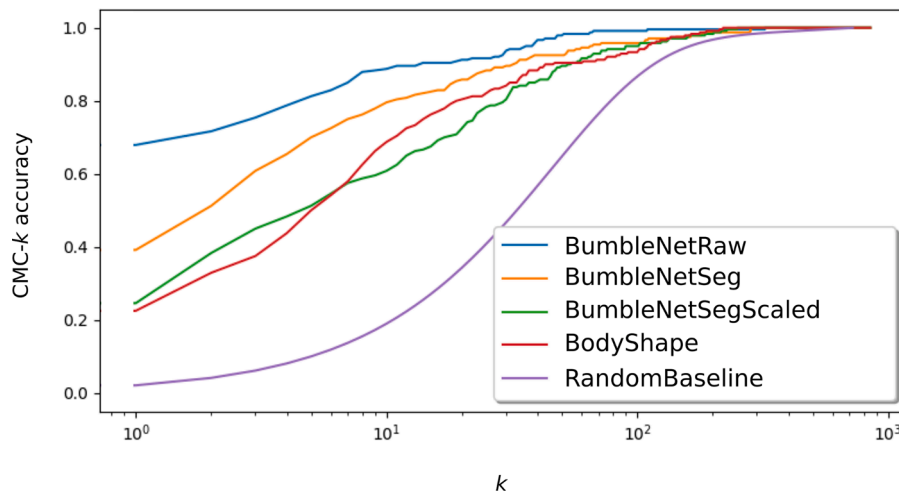
Fig. 3. For $k = 1, \cdots, 833$ (gallery size) the graph shows the probability that a probe image is among the top-k elements of the ranked retrieved results (CMC-*k*).

Table 2

Comparison of models performance with respect to CMC-1, CMC-5, CMC-10 and mAP score, ordered by CMC-1.

Models	CMC-1	CMC-5	CMC-10	mAP
BumbleNetRaw	**0.6792**	**0.8125**	**0.8875**	**0.3509**
BumbleNetSeg	0.3916	0.7000	0.7958	0.2446
BumbleNetSegScaled	0.2458	0.5125	0.6083	0.1295
BodyShape	0.2291	0.4958	0.6958	0.1876
RandomBaseline	0.0208	0.0999	0.1898	0.0292

5. Discussion & Conclusion

In relation to research question (1), to what extent bumblebees are re-identifiable on visual material, we have been able to show promising results. The best model is able to correctly assign two thirds of the individuals at first go (BumbleNetRaw). Since the model may have used some undesired information, the performance could be worse in reality. A safe lower bound for the model performance is the BumbleNetSeg with a CMC-1 score of 39%.

It was already known that bumblebees show a remarkable diversity in their body size. This finding was confirmed by comparing the BumbleNetSeg and BumbleNetSegScaled models, as omitting the size information led to a sharp drop in the CMC-1 score. However, it was also found that other features play a role, still allowing a fourth of the bees to be correctly identified on the first attempt (see BumbleNetSegScaled, Table 2). The non-BodyShape model (BumbleNetSegScaled) slightly outperforms the BodyShape model, indicating that the general appearance contains more information than the individual's shape parameters.

Due to the low contrast and resolution of the images, other features could become more important as image quality increases. Therefore, a dataset with more detailed and higher quality images, including pollen, would be desirable. Furthermore, a new dataset should integrate long-term and short-term re-ID data as proposed by Chan et al. [10]. Long-term data would then make it possible to transfer models without segmentation (e.g. BumbleNetRaw) into applications with confidence, as unwanted features are largely avoided.

Finally, a bumblebee re-ID is possible on the given dataset. A real-world application would benefit from clever boosting through dynamic galleries and statistical modelling of presence probabilities. Relying solely on variations in body shape misses the potential of other visual features.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

References

[1] S. Schneider, G.W. Taylor, S. Linquist, S.C. Kremer, Past, present and future approaches using computer vision for animal re-identification from camera trap data, Methods Ecol. Evol. 10 (4) (2019) 461–470, https://doi.org/10.1111/2041-210X.13133.
[2] C. Brust, T. Burghardt, M. Groenenberg, C. Käding, H. Kühl, M.L. Manguette, J. Denzler, Towards Automated Visual Monitoring of Individual Gorillas in the Wild, 2017 IEEE International Conference on Computer Vision Workshops (ICCVW) (2017), https://doi.org/10.1109/ICCVW.2017.333.
[3] A. Freytag, E. Rodner, M. Simon, A. Loos, H.S. Kühl, J. Denzler, Chimpanzee Faces in the Wild: Log-Euclidean CNNs for Predicting Identities and Attributes of Primates, in: B. Rosenhahn, B. Andres (Eds.), Pattern Recognition, Lecture Notes in Computer Science, Springer International Publishing, Cham, 2016, pp. 51–63, https://doi.org/10.1007/978-3-319-45886-1_5.
[4] M. Körschens, B. Barz, J. Denzler, Towards Automatic Identification of Elephants in the Wild, 2018.
[5] W. Wang, R.A. Solovyev, A.L. Stempkovsky, D.V. Telpukhov, A.A. Volkov, Method for Whale Re-identification Based on Siamese Nets and Adversarial Training, Opt. Mem. Neural Netw. 29 (2) (2020) 118–132, https://doi.org/10.3103/S1060992X20020058.
[6] T. Arbuckle, S. Schröder, V. Steinhage, D. Wittmann, Biodiversity Informatics in Action: Identification and Monitoring of Bee Species using ABIS. Proc. 15th Int. Symp. Informatics for Environmental Protection, 2002.
[7] T.M. Francoy, D. Wittmann, M. Drauschke, S. Müller, V. Steinhage, M.A.F. Bezerra-Laure, D. De Jong, L.S. Gonçalves, Identification of Africanized honey bees through wing morphometrics: two fast and efficient procedures, Apidologie 39 (5) (2008) 488–494, https://doi.org/10.1051/apido:2008028.
[8] G. Kastberger, S. Radloff, G. Kranner, Individuality of wing patterning in Giant honey bees (Apis laboriosa), Apidologie 34 (3) (2003) 311–318, https://doi.org/10.1051/apido:2003020.
[9] N. Murali, J. Schneider, J. Levine, G. Taylor, Classification and Re-Identification of Fruit Fly Individuals Across Days With Convolutional Neural Networks. 2019 IEEE Winter Conference on Applications of Computer Vision (WACV), IEEE, Waikoloa Village, HI, USA, 2019, pp. 570–578, https://doi.org/10.1109/WACV.2019.00066.
[10] J. Chan, H. Carrión, R. Mégret, J. Rivera, T. Giray, Honeybee Re-identification in Video: New Datasets and Impact of Self-supervision:. Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, SCITEPRESS - Science and Technology Publications, 2022, pp. 517–525, https://doi.org/10.5220/0010843100003124.

[11] D. Goulson, Bumblebees: Their Behaviour and Ecology, Oxford University Press, 2003.

[12] J.M. Mola, N.M. Williams, A review of methods for the study of bumble bee movement, Apidologie 50 (4) (2019) 497–514, https://doi.org/10.1007/s13592-019-00662-3.

[13] European Food Safety Authority, Guidance on the risk assessment of plant protection products on bees (Apis mellifera, Bombus spp. and solitary bees), EFSA Journal 11 (7) (2013), https://doi.org/10.2903/j.efsa.2013.3295.

[14] F. Tausch. Visual Re-Identification of Bumblebees Returning from Foraging Trips, Karlsruhe Institute of Technology, Karlsruhe, 2020. Bachelor's Thesis.

[15] F. Tausch, S. Stock, J. Fricke, O. Klein, Bumblebee Re-Identification Dataset. 2020 IEEE Winter Applications of Computer Vision Workshops (WACVW), IEEE, Snowmass Village, CO, USA, 2020, pp. 35–37, https://doi.org/10.1109/WACVW50321.2020.9096909.

[16] M.A. Becher, G. Twiston-Davies, T.D. Penny, D. Goulson, E.L. Rotheray, J.L. Osborne, Bumble-BEEHAVE: A systems model for exploring multifactorial causes of bumblebee decline at individual, colony, population and community level, J. Appl. Ecol. 55 (6) (2018) 2790–2801, https://doi.org/10.1111/1365-2664.13165.

[17] S. Chan-Lang, Q.-C. Pham, C. Achard, Closed and Open-World Person Re-Identification and Verification. 2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA), 2017, pp. 1–8, https://doi.org/10.1109/DICTA.2017.8227416.

[18] L. Zheng, L. Shen, L. Tian, S. Wang, J. Wang, Q. Tian, Scalable person re-identification: A benchmark. Proceedings of the IEEE international conference on computer vision, 2015, pp. 1116–1124.

[19] J. Marstaller, F. Tausch, S. Stock, DeepBees - Building and Scaling Convolutional Neuronal Nets For Fast and Large-Scale Visual Monitoring of Bee Hives. 2019

[20] F. Tausch, K. Schmidt, M. Diehl, Current achievements and future developments of a novel AI based visual monitoring of beehives in ecotoxicology and for the monitoring of landscape structures. preprint, Animal Behavior and Cognition, 2020, https://doi.org/10.1101/2020.02.04.933580.

[21] G. Van Horn, O. Mac Aodha, Y. Song, Y. Cui, C. Sun, A. Shepard, H. Adam, P. Perona, S. Belongie, The inaturalist species classification and detection dataset. Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 8769–8778.

[22] M. Ye, J. Shen, G. Lin, T. Xiang, L. Shao, S.C. Hoi, Deep learning for person re-identification: A survey and outlook, IEEE transactions on pattern analysis and machine intelligence 44 (6) (2021) 2872–2893.Publisher: IEEE

[23] M. Koestinger, M. Hirzer, P. Wohlhart, P.M. Roth, H. Bischof, Large scale metric learning from equivalence constraints. 2012 IEEE conference on computer vision and pattern recognition, IEEE, 2012, pp. 2288–2295.

[24] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention, Springer, 2015, pp. 234–241.

[25] F. Schroff, D. Kalenichenko, J. Philbin, Facenet: A unified embedding for face recognition and clustering. Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 815–823.

[26] K. He, X. Zhang, S. Ren, J. Sun, Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778, https://doi.org/10.1109/CVPR.2016.90.

IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), IEEE, Seoul, Korea (South), 2019, pp. 271–278, https://doi.org/10.1109/ICCVW.2019.00036.