Supervised Contrastive Vehicle Quantization for Efficient Vehicle Retrieval

Yongbiao Chen
chenyongbiao0319@sjtu.edu.cn
Shanghai Jiao Tong University
Shanghai, China

Kaicheng Guo
guokaicheng@sjtu.edu.cn
Shanghai Jiao Tong University
Shanghai, China

Fangxin Liu
liufangxin@sjtu.edu.cn
Shanghai Jiao Tong University
Shanghai, China

Yusheng Huang
huangyusheng@sjtu.edu.cn
Shanghai Jiao Tong University
Shanghai, China

Zhengwei Qi*
qizhwei@sjtu.edu.cn
Shanghai Jiao Tong University
Shanghai, China

ABSTRACT

This paper considers large-scale efficient vehicle re-identification (*Vehicle ReID*). Existing works adopting deep hashing techniques function by projecting vehicle images into compact binary codes in the *Hamming* space. Since *Hamming* distance is less distinct, a considerable amount of discriminative information will be lost, leading to degraded retrieval performances. Inspired by the recent advancements in contrastive learning, we put forward the very first product quantization based framework for large-scale efficient vehicle re-identification: **S**upervised **C**ontrastive **V**ehicle **Q**uantization (**SCVQ**). Specifically, we integrate the product quantization process into deep supervised learning by designing a differentiable quantization network. In addition, we propose a novel supervised cross-quantized contrastive quantization (**SCQC**) loss for similarity-preserving learning, which is tailored for the asymmetric retrieval in the product quantization process. Comprehensive experiments on two public benchmarks have evidenced the superiority of our framework against the state-of-the-arts. Our work is open-sourced at: https://github.com/chrisbyd/ContrastiveVehicleQuant.

CCS CONCEPTS

• **Information systems** → **Expert search**; • **Computing methodologies** → **Visual content-based indexing and retrieval**.

KEYWORDS

vehicle re-identification, deep learning, contrastive learning, product quantization, large-scale image retrieval

ACM Reference Format:
Yongbiao Chen, Kaicheng Guo, Fangxin Liu, Yusheng Huang, and Zhengwei Qi. 2022. Supervised Contrastive Vehicle Quantization for Efficient Vehicle Retrieval. In *Proceedings of the 2022 International Conference on Multimedia Retrieval (ICMR '22), June 27–30, 2022, Newark, NJ, USA.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3512527.3531432

Corresponding author is Zhengwei Qi.

ICMR '22, June 27–30, 2022, Newark, NJ, USA

ACM ISBN 978-1-4503-9238-9/22/06…$15.00
https://doi.org/10.1145/3512527.3531432

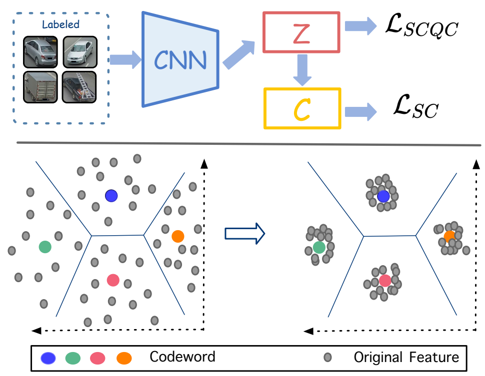


Figure 1: Above: an illustration of the overall framework. Z is the quantization layer, and C is the classifier. Below: a 2D conceptual Voronoi diagram showing one of the codebook training process in SCVQ. After training, the features are clustered around the codewords.

1 INTRODUCTION

Vehicle re-identification (*Vehicle ReID*) aims at accurately matching vehicles captured by disjoint surveillance cameras across different viewpoints. It has gained growing traction among computer vision research communities for its diverse applications in intelligent transportation, public security, and etc [3, 8, 14, 21, 24, 26, 30, 31]. The majority of works adopts real-valued features vectors that require high memory consumption and more computation, thus, is not applicable in a large-scale retrieval setting [8, 17, 18].

To resolve this challenge, researchers have investigated the application of deep Hamming hashing [1, 2, 9, 22, 32] into efficient large-scale vehicle re-identification, considering its low memory consumption and high retrieval speed. Despite the apparent successes they achieved, pioneering deep *Hamming* hashing-based **Vehicle ReID** methods suffer from two major weaknesses, further limiting their retrieval accuracy. ❶: Since the *Sign* function is not differentiable, these works generally adopt continuous optimization to relax the discrete optimization objective in the training stage, which might deviate from the original optimization target markedly. ❷: As pointed out by [29], *Hamming* distance is less distinct, which means a significant amount of fine-grained discriminative information will be sacrificed.

In light of these deficiencies, we investigate the adoption of product quantization [10, 19, 20] for deep hashing learning. Early works applying product quantization [16, 29] for deep hashing learning generally adopt a two-stage scheme. In the first stage, they learn deep discriminative features in a supervised fashion. Then, unsupervised k-means clustering is utilized to quantize the feature vectors into a *Cartesian* product of multiple codewords. [13, 28] further proposes to integrate the quantization process into the supervised deep feature learning process. Nonetheless, these methods are not tailed for the challenging *Vehicle ReID* scenario.

In this paper, we propose a supervised contrastive vehicle quantization (**SCVQ**) framework: the first product quantization based framework for large-scale, efficient vehicle re-identification. As illustrated in Fig. 1, it embeds the product quantization learning into the supervised learning as a trainable quantization layer. Inspired by the advancements of contrastive learning [4–6, 15], which adopts a contrastive loss called **InfoNCE** to pull close the similar pairs in the embedding space and push away the dissimilar ones, we innovate a novel quantization loss dubbed supervised cross-quantized contrastive (**SCQC**) loss for similarity-preserving learning. With this loss, the quantization layer and the backbone network can be trained simultaneously, and the asymmetric retrieval nature of product quantization is embedded into the training process. In addition to **SCQC**, to learn more discriminative features, we adopt a subspace classification loss on the quantized features **q** to learn vehicle identity discriminative features.

To sum up, we make the following contributions:

(1) We propose the first product quantization framework (**SCVQ**) for large-scale efficient vehicle re-identification. We innovate a novel supervised cross-quantized contrastive loss for similarity-preserving learning specially tailored for product quantization learning.
(2) We conducted comprehensive experiments across two widely-adopted public vehicle re-identification datasets across four varied hashing bits, and the results demonstrated the effectiveness and superiority of our method.

2 IMPLEMENTATION

2.1 Vehicle Product Quantization Network

The network architecture is illustrated in Fig. 1. We design a trainable product quantization network after the *Convolutional Neural Network* backbone, containing a randomly initialized trainable *codebook* $\mathcal{Z} = [\mathcal{Z}_1^s, \ldots, \mathcal{Z}_m^s, \ldots, \mathcal{Z}_M^s]$ which consists of M *sub-codebooks*. Each *sub-codebook* z_m^s contains K trainable *sub-codewords*: $\mathcal{Z}_m^s = [(z_m^s)_1, \cdots, (z_m^s)_k, \cdots, (z_M^s)_K]$. The conventional hard quantization is implemented with arg min function, shown in Eq. 1, which is not differentiable. We observe the equivalence when substituting the arg min with a softmax function as illustrated in Eq. 1 [28]. In such a way, we design a soft quantization function to approximate the original arg min as follows:

Hard Quantization to Soft Quantization. For an input feature $\mathbf{x} \in \mathbb{R}^L$, we obtain its m-th sub-vector $x_m^s \in \mathbb{R}^{L/M}$. The soft quantization function is formulated as:

$$\begin{aligned} Q_{soft}^m\left(\mathbf{x}_m^s\right) &= \lim_{\gamma\to+\infty} \sum_k \frac{e^{-\gamma\|\mathbf{x}_m^s-(\mathbf{z}_m^s)_k\|_2}}{\sum_k e^{-\gamma\|\mathbf{x}_m-(\mathbf{z}_m^s)_k\|_2}} (\mathbf{z}_m^s)_k \\ &= \arg\min_{(z_m^s)_k} \|\mathbf{x}_m^s - (z_m^s)_k\|_2^2 \\ &\text{s.t.} \quad k \in \{i\}_{i=1}^{i=K} \end{aligned} \tag{1}$$

when $\gamma \to +\infty$, the soft quantizer Q_{soft} equals the hard quantizer Q_{hard}. In this way, the quantized feature **q** for the original input feature **x** could be obtained through the following equation:

$$\mathbf{q} = [Q_{soft}^1(x_1^s), \cdots, Q_{soft}^m(x_m^s), \cdots, Q_{soft}^M(x_M^s)] \tag{2}$$

2.2 Supervised Cross Quantized Contrastive Learning.

As introduced in Sec. 2.1, the *PQ codebook* $\mathcal{Z}$ consists of M *sub-codebooks* $\mathcal{Z} = [\mathcal{Z}_1^s, \cdots, \mathcal{Z}_m^s, \cdots, \mathcal{Z}_M^s]$ while each *sub-codebook* $\mathcal{Z}_m$ contains K *sub-codewords* $\mathcal{Z}_m^s = [(z_m^s)_1, ..., (z_m^s)_k, ..., (z_m^s)_K]$. $(z_m^s) \in \mathbb{R}^d, d = \frac{L}{M}$. For an input feature vector **x**, we obtain its quantized counterpart **q** with the soft quantizer Q_{soft} as illustrated above. Then, both the original feature **x** and the quantized feature **q** are l2-normalized to produce $\hat{\mathbf{x}}$ and $\hat{\mathbf{q}}$.

Conventional scheme generates every sub-codebook $\mathcal{Z}_m^s$ through unsupervised k-means clustering on all the sub-vectors $\{(x_m^s)_i\}_{i=0}^{i=N_D-1}$ from the dataset where N_D is the number of training images. Since the product quantization learning is separated from the supervised learning process, the supervisory signals cannot contribute to the learning of the *codewords*, resulting in sub-optimal performances.

To simultaneously learn discriminative features and all the *codewords* in the *PQ codebook* $\mathcal{Z}$, we propose the *Supervised Cross Quantized Contrastive Quantization* (**SCQC**) Loss. Specifically, inspired by supervised contrastive learning [6, 15], we attempt to compare the original feature vector with the quantized features. For a batch of N_B randomly-sampled images with supervisory labels $\{(X_1, Y_1), \cdots, (X_{N_B}, Y_{N_B})\}$, we first define the cross-quantized pair as a pair of a original feature and a quantized feature $(\hat{\mathbf{x}}_i, \hat{\mathbf{q}}_j)$, where $\hat{\mathbf{x}}_i$ is the normalized feature for image X_i and $\hat{\mathbf{q}}_j$ is the normalized quantized one for image X_j. Then, we define a correlation function for cross-quantized pairs as $\mathcal{S}_C : (\hat{\mathbf{x}_i}, \hat{\mathbf{q}_j}) \mapsto \{0, 1\}$. $\mathcal{S}_C(\hat{\mathbf{x}_i}, \hat{\mathbf{q}_j}) = 1$ if $Y_i Y_j^T > 0$ and $\mathcal{S}_C(\hat{\mathbf{x}}_i, \ \hat{\mathbf{q}_j}) = 0$, otherwise.

Then, the loss objective can be formulated as:

$$L_{CQC} = \sum_{i\in I} \frac{-1}{|P(i)|} \sum_{p\in P(i)} \log \frac{\exp\left(\mathcal{S}(\hat{\mathbf{x}}_i, \hat{\mathbf{q}}_p)/\tau\right)}{\sum_{j\in A(i)} \exp\left(\mathcal{S}(\hat{\mathbf{x}}_i, \hat{\mathbf{q}}_j)/\tau\right)} \tag{3}$$

where $I = \{1, \cdots, N_B\}$ is the set of indexes of the input mini-batch and $A(i) = I \backslash \{i\}$. $P(i) = \left\{p \in A(i) : \mathcal{S}_C(\hat{\mathbf{x}}_i, \ \hat{\mathbf{q}_p}) = 1\right\}$ is the set of indices of all correlated images for X_i in the batch. Meanwhile, $\hat{\mathbf{x}}_i$ is the normalized descriptor for image X_i and $\hat{q}_p$ is the quantized feature for X_p. $\mathcal{S}(\hat{x}, \hat{q})$ is the cosine similarity function for **x** and **q**. $\tau \in \mathcal{R}^+$ is the temperature parameter. L_{SCQC} is asymmetric since all the pairs are cross-quantized pairs which contains an original feature and a quantized feature. It learns to pull close positive cross-quantized pairs and push away negative pairs. In such a way, the framework learns jointly the deep discriminative features and all the *codewords* in the *codebook* $\mathcal{Z}$.

Table 1: Vehicle Re-identification Results of State-of-The-Art Deep Hashing Methods on VehicleID and VeRi

Datasets	VehicleID				VeRi			
Methods	256 bits	512 bits	1024 bits	2048 bits	256 bits	512 bits	1024 bits	2048 bits
SH	23.98	23.64	21.09	17.71	3.33	2.75	2.08	1.57
ITQ	23.84	25.39	25.93	26.26	4.80	4.95	5.04	5.14
DSH	30.56	32.53	34.22	33.62	8.93	9.77	10.24	10.47
DHN	43.57	45.42	46.70	46.39	13.85	13.93	14.31	13.71
HashNet	62.95	64.52	65.96	66.61	27.13	28.40	29.94	29.30
DCH	37.08	34.55	26.04	24.13	19.75	19.23	12.15	14.96
DPN	57.16	61.92	63.55	64.65	10.03	11.26	13.52	13.58
EH	69.03	71.07	72.28	-	-	-	-	-
DVHN	**71.61**	**73.69**	**75.86**	**76.82**	**54.61**	**58.38**	**59.27**	**62.02**
SCVQ	**76.22**	**80.18**	**82.11**	**83.60**	**57.22**	**60.89**	**63.72**	**63.59**

Table 2: The Results of Variants on VehicleID and VeRi

Datasets	Variants	2080 bits		
		Rank@1	Rank@5	mAP %
VehicleID	**W/O SC**	72.42	94.31	81.88
	W/O SCQC	73.02	91.23	80.94
	Full Model	74.90	95.08	83.60
VeRi	**W/O SC**	84.62	92.85	61.78
	W/O SCQC	84.21	94.04	61.19
	Full Model	86.00	95.23	63.59

2.3 Subspace-wise Classification

To learn vehicle identity discriminative information and embed them into the *codewords* $\mathcal{Z}$, we further propose to learn multiple classifiers atop the quantized feature $\mathbf{q}$ which contains M weight matrices $[\mathbf{W}_1, \cdots, \mathbf{W}_M]$, where M is the number of subspaces. Each of the weight matrix $\mathbf{W}_m \in \mathbb{R}^{d \times N_C}$ consists of N_C vectors: $\mathbf{W}_m = [(w_m)_1, \cdots, (w_m)_{N_C}]$, where N_C is the number of vehicle identities. Take the m-th sub-vector q^s_m and the m-th weight matrix $\mathbf{W}_m$ for example, we could compute the class prediction logits as $\mathbf{W}^T_m (q_m)^T$. The class prediction logits for other sub-vectors could be computed similarly. In this way, we could derive the subspace-wise classification loss as:

$$\begin{aligned}\mathcal{L}_{\text{SC}} &= \frac{1}{M} \sum_{m=1}^{M} \sum_{i=1}^{N_C} L_{CE}(\mathbf{W}^T_m (q^s_m)^T_i, y_i) \\ &= -\frac{1}{M} \sum_{m=1}^{M} \sum_{i=1}^{N_C} \log \frac{e^{(w^T_m)_{y_i} (q^s_m)^T_i}}{\sum_{k=1}^{N_C} e^{(w^T_m)_k (q^s_m)^T_i}}\end{aligned} \tag{4}$$

where y_i is the labels for image I_i and $(q^s_m)_i$ is its m-th quantized sub-vector representation.

3 EXPERIMENTATION

We conduct extensive experiments to demonstrate the effectiveness of **SCVQ** against several state-of-the arts on two standard benchmarks. Datasets and implementations are available at GitHub [1]

3.1 Experimental Setup

The evaluation is conducted on two benchmarks: **VeRi-776** and **VehicleID**.

VeRi-776 [25] is the first vehicle re-identification dataset consisting of over 50,000 vehicle of 776 vehicles in total, which is capture by 20 cameras.

VehicleID [21] is a large-scale surveillance dataset comprised of 26,267 vehicle identities and 22,1763 vehicle images in total.

We compare the re-identification performance of **SCVQ** with the state-of-the-art methods, including 2 shallow method (**SH** [27], **ITQ** [11]), 5 deep hashing methods (**DSH** [22], **DHN** [32], **HashNet** [2], **DCH** [1], **DPN** [9]) and 2 deep hashing methods tailored for vehicle re-identification (**EH** [23], **DVHN** [7]).

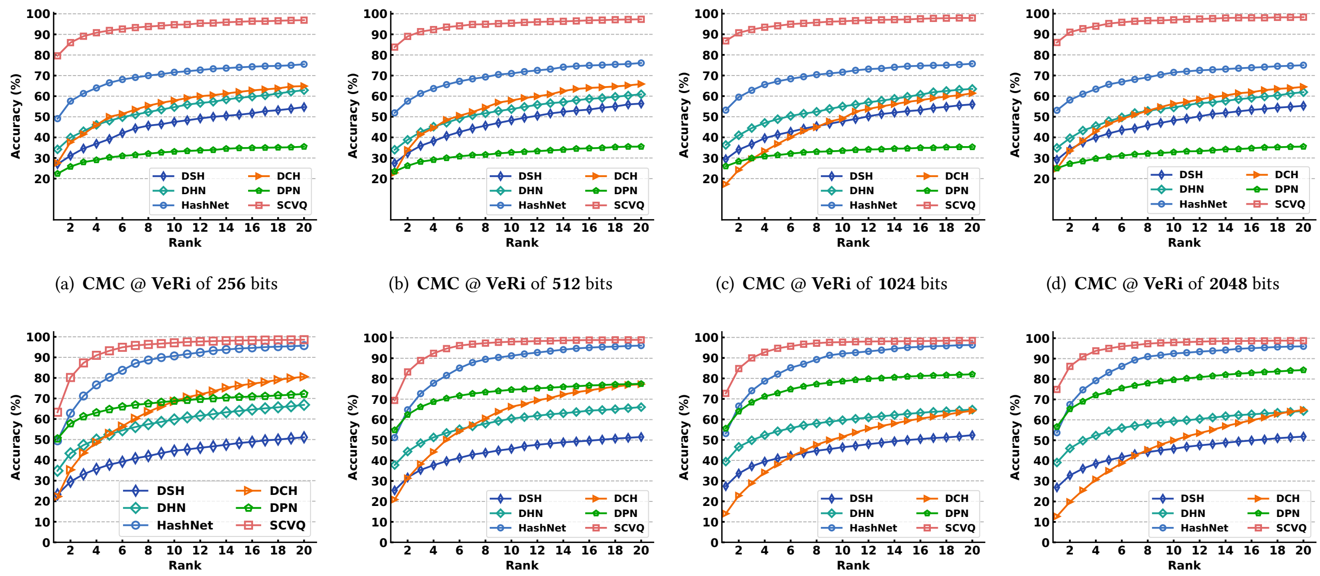
For fair comparisons, we adopt mean average precision (**mAP**) and cumulative matching curve (**CMC**) to measure the retrieval accuracy of all the competing methods.

ImageNet-pre-trained ResNet-50 [12] is adopted as the backbone for all the deep hashing methods. The input images are resized to 128×256. In the training stage, the *Adam* optimizer is used with an initial learning rate of $1e-4$ and weight decay of $5e-4$. Meanwhile, widely-adopted standard data augmentation techniques, including random cropping, random erasing, and random horizontal flipping are adopted. In terms of hyper-parameters, we set γ to 20.0 and α to 0.3. To generate binary codes of varied lengths, we fix the number of *sub-codewords* K to 256 and set the number of *sub-codebooks* M to 32, 64, 128, 256 to generate binary codes of 256, 512, 1024, 2046 bits. Our framework is implemented with *PyTorch*. The experiments are conducted on a server equipped with 80 Core Intel Xeon Gold 5218R CPU and two NVIDIA A100 GPUs.

3.2 Results

The **mAP**s of all the competing methods across four hashing bit lengths are presented in Tab. 1. It is easy to note that our proposed **SCVQ** is an absolute winner against all the competing methods across two benchmarks. Shallow hashing methods do not exhibit satisfactory results mainly because they could not utilize the supervisory information in the hashing encoding process. Deep hashing methods generally achieve better results. Compared with traditional deep hashing methods, we obtain significant performance improvements with an average increase in **mAP** of 15.52% and 32.66% on **VehicleID** and **VeRi**, respectively. The pronounced results could be ascribed to the combined effect of the **SCQC** for dealing with the large intra-class variation and the **SC** for inter-class discriminative learning. Compared with the state-of-the-art vehicle hashing methods, we outperform the best method **DVHN** by 6.03% and 2.78% in average **mAP** across all the hashing bits. The **CMC** curves of all the competing methods across four hashing bits

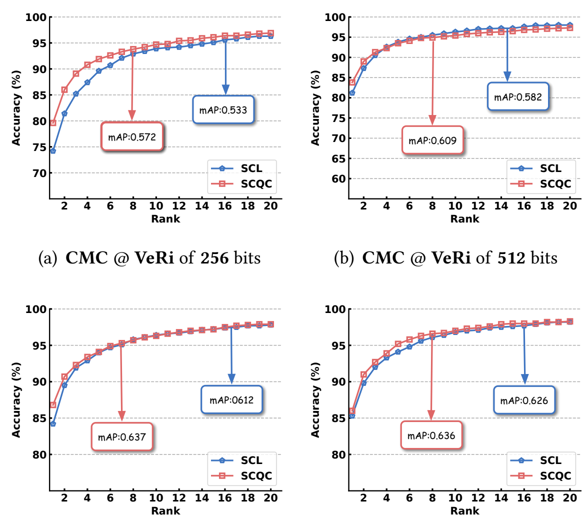
[1] https://github.com/chrisbyd/ContrastiveVehicleQuant



(a) **CMC @ VeRi** of **256** bits (b) **CMC @ VeRi** of **512** bits (c) **CMC @ VeRi** of **1024** bits (d) **CMC @ VeRi** of **2048** bits

(e) **CMC @ VehicleID** of **256** bits (f) **CMC @ VehicleID** of **512** bits (g) **CMC @ VehicleID** of **1024** bits (h) **CMC @ VehicleID** of **2048** bits

Figure 2: The CMC results against the State-of-the-art methods on two datasets across 256, 512, 1024, 2048 bits



(a) **CMC @ VeRi** of **256** bits (b) **CMC @ VeRi** of **512** bits

(c) **CMC @ VeRi** of **1024** bits (d) **CMC @ VeRi** of **2048** bits

Figure 3: The comparison of SCQC and the original supervised contrastive learning (SCL) loss

on **VehicleID** and **VeRi** are illustrated in Fig. 2. Note that **DVHN** is not open-sourced yet, so we do not include it in the **CMC** curve. It is rather evident that the curve of **SCVQ** consistently levitates beyond the competing counterparts.

3.3 Empirical Analysis

In this subsection, we conduct comprehensive experiments to evaluate the effectiveness of each component in **SCVQ** and the validity of the proposed quantization loss **SCQC**.

Analysis of the Components. To validate the contribution of $\mathcal{L}_{SC}$ and $\mathcal{L}_{SCQC}$ in our **SCVQ** framework, we test the performances of **SCVQ** when stripped of each component. The model without $\mathcal{L}_{SC}$ is denoted as **W/O SC**, the model without $\mathcal{L}_{SCQC}$ as **W/O SCQC** and the full model as **Full Model**.

From Tab. 2, it is evident that the best performance is achieved by integrating both the subspace-wise classification loss and the supervised cross-quantized contrastive loss. On **VehicleID**, **Full Model** surpasses **W/O SC** by **1.72**% and **W/O SCQC** by **2.66**%. On **VeRi**, it beats them by **1.81**% and **2.40**, respectively. We could also note that $\mathbf{L}_{SCQC}$ contributes slightly more to the overall performance. All these combined have evidenced the effectiveness of each component to the final performance of **SCVQ**.

Analysis of CQHT. To validate our quantization loss: "supervised cross-quantized contrastive loss", we conduct experiments comparing it to the traditional supervised contrastive loss (**SCL**) in the re-identification area. As demonstrated in Fig. 3, our loss consistently surpasses the original one in the end-to-end quantization setting. This is largely because our loss jointly considers the original feature descriptor and the quantized feature and embeds the asymmetric retrieval in product quantization into the metric learning process.

4 CONCLUSION

In this work, we propose a novel deep product quantization based framework (**SCVQ**) for large-scale vehicle re-identification. It integrates the product quantization process into the supervised learning and propose a joint supervised cross-quantized contrastive learning and subspace classification learning scheme. Our work is the very first supervised product quantization scheme for efficient large-scale vehicle re-identification.

ACKNOWLEDGMENTS

This work was supported in part by National NSF of China (Grant NO. 62141218), and Shanghai Key Laboratory of Scalable Computing and Systems. We thank Wu Wen Jun Honorary Doctoral Scholarship, AI Institute, Shanghai Jiao Tong University.

REFERENCES

[1] Yue Cao, Mingsheng Long, et al. 2018. Deep cauchy hashing for hamming space retrieval. In *CVPR*. 1229–1237.

[2] Zhangjie Cao, Mingsheng Long, et al. 2017. Hashnet: Deep learning to hash by continuation. In *ICCV*. 5608–5617.

[3] Zhigang Chang, Zhao Yang, Yongbiao Chen, Qin Zhou, and Shibao Zheng. 2021. Seq-Masks: Bridging the gap between appearance and gait modeling for video-based person re-identification. In *2021 International Conference on Visual Communications and Image Processing (VCIP)*. IEEE, 1–5.

[4] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. 2020. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*. PMLR, 1597–1607.

[5] Xinlei Chen, Haoqi Fan, Ross Girshick, and Kaiming He. 2020. Improved baselines with momentum contrastive learning. *arXiv preprint arXiv:2003.04297* (2020).

[6] Xinlei Chen and Kaiming He. 2021. Exploring simple siamese representation learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 15750–15758.

[7] Yongbiao Chen, Sheng Zhang, et al. 2021. DVHN: A Deep Hashing Framework for Large-scale Vehicle Re-identification. *arXiv preprint arXiv:2112.04937* (2021).

[8] Yongbiao Chen, Sheng Zhang, and Zhengwei Qi. 2020. MAENet: Boosting Feature Representation for Cross-Modal Person Re-Identification with Pairwise Supervision. In *Proceedings of the 2020 International Conference on Multimedia Retrieval*. 442–449.

[9] Lixin Fan, KamWoh Ng, Ce Ju, Tianyu Zhang, and Chee Seng Chan. 2020. Deep Polarized Network for Supervised Learning of Accurate Binary Hashing Codes.. In *IJCAI*. 825–831.

[10] Tiezheng Ge, Kaiming He, et al. 2013. Optimized product quantization. *TPAMI* 36, 4 (2013), 744–755.

[11] Yunchao Gong, Svetlana Lazebnik, et al. 2012. Iterative quantization: A procrustean approach to learning binary codes for large-scale image retrieval. *TPAMI* 35, 12 (2012), 2916–2929.

[12] Kaiming He, Xiangyu Zhang, et al. 2016. Deep residual learning for image recognition. In *CVPR*. 770–778.

[13] Young Kyun Jang and Nam Ik Cho. 2020. Generalized product quantization network for semi-supervised image retrieval. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 3420–3429.

[14] Pirazh Khorramshahi, Peri, et al. 2020. The devil is in the details: Self-supervised attention for vehicle re-identification. In *ECCV*. Springer, 369–386.

[15] Prannay Khosla, Piotr Teterwak, Chen Wang, Aaron Sarna, Yonglong Tian, Phillip Isola, Aaron Maschinot, Ce Liu, and Dilip Krishnan. 2020. Supervised contrastive learning. *arXiv preprint arXiv:2004.11362* (2020).

[16] Bin Liu, Yue Cao, et al. 2018. Deep triplet quantization. In *ACM MM*. 755–763.

[17] Fangxin Liu, Wenbo Zhao, Yongbiao Chen, Zongwu Wang, Tao Yang, and Li Jiang. 2021. SSTDP: Supervised Spike Timing Dependent Plasticity for Efficient Spiking Neural Network Training. *Frontiers in Neuroscience* 15 (2021).

[18] Fangxin Liu, Wenbo Zhao, Zhezhi He, et al. 2021. Bit-Transformer: Transforming Bit-level Sparsity into Higher Preformance in ReRAM-based Accelerator. In *Proceedings of the 40th International Conference on Computer-Aided Design*.

[19] Fangxin Liu, Wenbo Zhao, Zhezhi He, Yanzhi Wang, Zongwu Wang, Changzhi Dai, Xiaoyao Liang, and Li Jiang. 2021. Improving Neural Network Efficiency via Post-Training Quantization With Adaptive Floating-Point. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 5281–5290.

[20] Fangxin Liu, Wenbo Zhao, Zongwu Wang, Yilong Zhao, Tao Yang, Yiran Chen, and Li Jiang. 2022. IVQ: In-Memory Acceleration of DNN Inference Exploiting Varied Quantization. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* (2022).

[21] Hongye Liu, Tian, et al. 2016. Deep relative distance learning: Tell the difference between similar vehicles. In *CVPR*. 2167–2175.

[22] Haomiao Liu, Wang, et al. 2016. Deep supervised hashing for fast image retrieval. In *CVPR*. 2064–2072.

[23] Meihan Liu, Yongxing Dai, et al. 2020. Extending Hashing Towards Fast Re-Identification. In *ICIP*. IEEE, 1551–1555.

[24] Xinchen Liu, Liu, et al. 2020. Beyond the parts: Learning multi-view cross-part correlation for vehicle re-identification. In *ACM MM*. 907–915.

[25] Xinchen Liu, Wu Liu, et al. 2016. A deep learning-based approach to progressive vehicle re-identification for urban surveillance. In *ECCV*. Springer, 869–884.

[26] Shangzhi Teng, Shiliang Zhang, Qingming Huang, and Nicu Sebe. 2021. Viewpoint and scale consistency reinforcement for UAV vehicle re-identification. *International Journal of Computer Vision* 129, 3 (2021), 719–735.

[27] Yair Weiss, Antonio Torralba, Robert Fergus, et al. 2008. Spectral hashing.. In *Nips*, Vol. 1. Citeseer, 4.

[28] Tan Yu, Junsong Yuan, et al. 2018. Product quantization network for fast image retrieval. In *ECCV*. 186–201.

[29] C Yue, M Long, et al. 2016. Deep quantization network for efficient image retrieval. In *AAAI*. 3457–3463.

[30] Jianan Zhao, Fengliang Qi, Guangyu Ren, and Lin Xu. 2021. PhD Learning: Learning With Pompeiu-Hausdorff Distances for Video-Based Vehicle Re-Identification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2225–2235.

[31] Yi Zhou and Ling Shao. 2018. Aware attentive multi-view inference for vehicle re-identification. In *CVPR*. 6489–6498.

[32] Han Zhu, Mingsheng Long, et al. 2016. Deep hashing network for efficient similarity retrieval. In *AAAI*, Vol. 30.