

基于跨域联合空间注意网络的草图图像检索

于凌志*, 张熙凡

天津大学电气自动化与信息工程学院, 天津 300072

摘要 基于草图的图像检索以手绘草图为输入检索相应的自然图像, 让使用者即使没有精准的相似自然图像也可以自行绘制后检索。边缘图通常作为中间模态引入, 以减小草图和自然图像之间的域差异, 但现有方法忽视了边缘图与自然图像之间的内在联系。基于自然图像及其边缘图应有相近重点区域的设想, 提出了基于跨域联合空间注意网络的深度学习模型。该模型从边缘图与自然图像的融合特征中获取两者共用的空间注意掩模, 结合损失函数和辅助分类器进行端到端的训练。与现有代表性的基于草图的图像检索方法相比, 所提方法能有效地提取草图和自然图像的特征, 在 Sketchy 和 TU-Berlin 数据集上的平均精度均值(mAP)分别达 0.933 和 0.799, 优于大部分代表性方法。

关键词 深度学习; 图像检索; 注意力机制; 跨域检索

中图分类号 TP391.4

文献标志码 A

DOI: 10.3788/LOP202259.2215009

Cross-Domain Spatial Co-Attention Network for Sketch-Based Image Retrieval

Yu Lingzhi*, Zhang Xifan

School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China

Abstract Sketch-based image retrieval uses hand-drawn sketches as input to retrieve corresponding natural images, allowing users to draw and find desired natural images when no accurate query images are available. Edge maps are commonly used as an intermediate modality to bridge the domain gap between sketches and natural images. However, existing methods ignore the inherent relationship between edge maps and natural images. Based on the assumption that natural images and their corresponding edge maps have similar key regions, this paper proposes a deep learning model based on a cross-domain spatial co-attention network. The proposed model derives the shared spatial attention mask from the fused feature of the edge map and natural image, and it combines the loss function and auxiliary classifier for end-to-end training. When compared with existing representative sketch-based image retrieval methods, the proposed method can effectively extract the features of sketches and natural images, with mean average precision (mAP) values of 0.933 and 0.799 on the Sketchy and TU-Berlin datasets, respectively, outperforming most representative methods.

Key words deep learning; image retrieval; attention mechanism; cross-domain retrieval

1 引言

互联网的发展使得图像数据量也快速增长, 要充分利用信息技术带来的便利, 在海量的图像数据中如何进行有效的检索是一个重要的课题。主流的图像检索方法有基于文本的图像检索(TBIR)^[1]和基于内容的图像检索(CBIR)。然而 TBIR 依赖大量的人工标注且易受到主观描述的影响; 传统的 CBIR 需要用户提供精准的相关图片, 限制了应用场景。手绘草图是一种比文本标签更加直观的描述方式, 而且即使没有

受过绘画训练的人也能够通过手绘表达出脑海中的检索目标。由于草图的上述特点能够解决 TBIR 和 CBIR 方法存在的问题, 基于草图的图像检索(SBIR)成为了图像检索领域中的一个热点。

传统 SBIR 算法通常提取图像轮廓以减小图像和草图之间天然存在的域差异, 以特征描述符进行图像匹配完成检索。传统特征往往经过微调后即可用于 SBIR^[2-6], 但这些方法仍不能很好地解决巨大域差异带来的检索难题。得益于算力提升, 深度学习方法在草图相关图像检索领域的应用开始飞速发展, Yu 等^[7]受

收稿日期: 2021-08-30; 修回日期: 2021-09-28; 录用日期: 2021-10-27

通信作者: *yulingzhi777@sina.com

AlexNet启发,针对草图特点,采用大卷积核和多尺度网络融合,提出用于草图分类的深度学习模型(Skech-a-Net)。Yu等^[8]沿用基于Sketch-a-Net的卷积神经网络(CNN)来提取特征,将三元组损失^[9]用于SBIR,证明这一方法可以有效地取得跨域表示。Zhang等^[10]使用循环一致监督的生成对抗网络(GAN)实现草图和图像域的转换,利用生成器将草图和图像统一到图像域并获取哈希编码。Lei等^[11]在自然图像和边缘图上应用通道注意,并研究了草图与边缘图的参数共享,结合对比损失和分类损失训练网络。该方法引入了边缘图作为中间模态缩小草图与自然图像之间的域差异,并在特征融合前通过通道加权的方式筛选边缘图和自然图像特征中的公共部分,从而使图像特征更接近草图特征,但并未考虑自然图像及其边缘图的内在联系。

为解决上述问题,受注意力机制在视觉问题中应用的启发^[12-14],本文在引入边缘图作为缩小自然图像与草图域差异的中间模态数据基础上,基于对边缘图与自然图像应关注相似区域的考量,设计了跨域联合空间注意机制,以得到更好的空间注意和特征融合

结果;针对跨域检索需要紧凑的类内分布的特点,在三元组损失基础上添加了类内距离约束损失并构建网络;此外还设计了合理的辅助分类任务,使编码结果中保留更多语义信息,并提供适用于大规模检索的离散编码方法。

2 跨域联合空间注意的图像检索模型

本文目的是设计模型 $F(\cdot)$,得到图像的 M 位编码 $B \in \{0, 1\}^M$,使草图的编码与相似的自然图像编码接近,从而通过草图检索相应自然图像。草图与自然图像之间不仅有纹理色彩有无的区别,绘图的轮廓等也由于绘画者的认知和技巧与真实物体不同,为了缩小两者的域差异,具有真实物体轮廓和无纹理色彩等草图特点的边缘图成为理想的中间模态数据,因此将自然图像经Canny边缘检测和轮廓检测的结果融合得到用于训练的边缘图。

基于度量学习中三元组损失的思想设计了跨域检索网络,网络结构如图1所示。网络的输入为三元组形式,草图 S 作为锚点,与锚点同类的自然图像 I^+ 及其边缘图 E^+ 作为正样本,不同类自然图像 I^- 及其边缘图

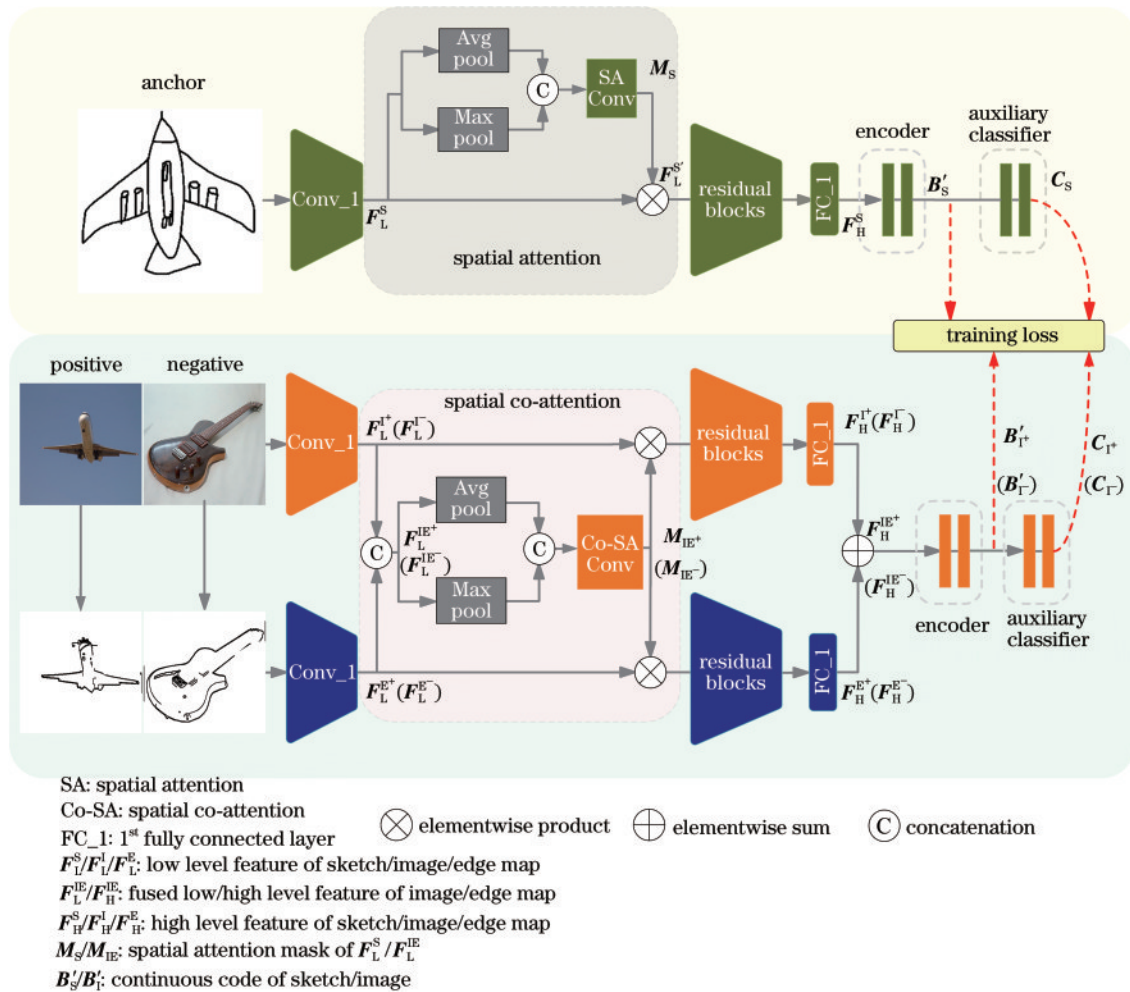


图1 所提方法的网络结构
 Fig. 1 Network architecture of the proposed method

E^- 作为负样本。为了将图像转换为编码,首先通过骨干网络提取图像特征,在骨干网络的第一个卷积层中加入空间注意机制,以实现更好的特征提取效果,并为自然图像及其边缘图设计了联合空间注意,以利用双模态的信息得到公共的注意区域;得到高层特征后,经编码器得到图像编码,结合辅助分类任务和损失函数训练模型。

2.1 网络结构

以 ResNet34 作为提取特征的骨干网络,记图像经 ResNet34-Conv_1 得到的低层特征为 F_L ,在草图分支第一组卷积后加入空间注意模块,从而使网络更关注关键区域。基础的空间注意实现与 Convolutional Block Attention Module^[15] 中相同,取特征图通道方向上最大池化和平均池化的结果,拼接后经卷积得到注意力掩模 M_S :

$$M_S = \sigma \left\{ \theta \left(\left[F_{L, Avg}^S; F_{L, Max}^S \right] \right) \right\}, \quad (1)$$

式中: σ 为 Sigmoid 激活函数; θ 为卷积核大小为 7×7 的卷积层。空间注意处理后的草图特征为

$$F_L^S = F_L^S \otimes M_S. \quad (2)$$

对于自然图像,认为自然图像及其边缘图应关注相似的空间区域,因此提出联合空间注意(Co-SA),将低层图像特征 F_L^I 和边缘图特征 F_L^E 在通道方向上拼接,得到低层融合特征 F_L^{IE} ,并以该融合特征计算其公共的注意掩模 M_{IE} :

$$F_L^{IE} = [F_L^I; F_L^E], \quad (3)$$

$$M_{IE} = \sigma \left\{ \theta \left(\left[F_{L, Avg}^{IE}; F_{L, Max}^{IE} \right] \right) \right\}. \quad (4)$$

这一注意掩模将同时应用于自然图像及其边缘图的低层特征:

$$F_L^I = F_L^I \otimes M_{IE}, \quad (5)$$

$$F_L^E = F_L^E \otimes M_{IE}. \quad (6)$$

经注意力模块处理的特征送入骨干网络的后续卷积部分,网络结构中 residual blocks 为 ResNet34 中的 Layer 2~Layer 33 及平均池化层,在骨干网络第一层全连接层后接编码器,得到编码结果。编码器以 $\psi(\cdot)$ 表示,由两层激活函数分别为 LeakyReLU 和 Sigmoid 的全连接层实现,其中自然图像及其边缘图的高层特征逐元素相加,得到融合的高层特征 F_H^{IE} ,草图和自然图像的高层特征经编码器得到连续编码 $B'_s \in (0, 1)^M$ 和 $B'_i \in (0, 1)^M$:

$$F_H^{IE} = F_H^I \oplus F_H^E, \quad (7)$$

$$B'_s = \psi_s(F_H^{IE}), \quad (8)$$

$$B'_i = \psi_i(F_H^{IE}). \quad (9)$$

由于草图与边缘图仍有较大的轮廓风格差异,将其视为两个域的数据,因而三种模态输入分支的参数相互独立。辅助分类器(AC)由两层全连接层实现,为使编码结果中保留更多语义信息,和以往的草图检索工作^[16]以编码器之前的高层特征作为分类任务的输入

不同,所提方法将编码结果 B' 作为分类器 $\varphi(\cdot)$ 的输入:

$$C_s = \varphi_s(B'_s), \quad (10)$$

$$C_i = \varphi_i(B'_i), \quad (11)$$

式中: C_s 和 C_i 分别为草图和自然图像经辅助分类器的输出。

2.2 针对跨域检索数据的损失函数

三元组损失^[9]最早在人脸识别任务中被提出,将训练样本分为锚点(anchor)、正样本(positive)和负样本(negative),旨在使正样本比负样本更加靠近锚点。损失函数的表达式为

$$L_{OriginTri} = \max \left(0, \|F_s(S) - F_i(I^+)\|_2 - \|F_s(S) - F_i(I^-)\|_2 + m \right), \quad (12)$$

式中: $F_s(S)$ 、 $F_i(I^+)$ 、 $F_i(I^-)$ 分别表示草图、正样本自然图像、负样本自然图像经网络模型 $F(\cdot)$ 得到的特征编码。原始的三元组损失通过参数 m 控制不同类间的间隔,当样本对满足 $\|F_s(S) - F_i(I^-)\|_2 \geq \|F_s(S) - F_i(I^+)\|_2 + m$ 时就不再惩罚其距离关系。然而在跨域检索任务中同类的跨域数据分布应尽可能紧凑,为了进一步减小类内样本间的距离,添加了要求正样本与锚点编码结果一致的约束:

$$L_{Intra-class} = \|F_s(S) - F_i(I^+)\|_2. \quad (13)$$

综上,用于调整图像编码间距离的损失函数为

$$L_{Triplet} = L_{OriginTri} + L_{Intra-class}. \quad (14)$$

采用交叉熵损失作为辅助分类任务的损失函数:

$$L_{cls} = \text{CrossEntropy}(C_s, Y_s) + \text{CrossEntropy}(C_i, Y_i), \quad (15)$$

式中: Y_s 和 Y_i 分别为草图和自然图像独热码形式的类别标签。

2.3 量化损失

大规模检索时连续编码的存储和查找计算成本较高,为了高效检索,采用量化函数 $f_q(\cdot)$ 将深度特征 $B' \in (0, 1)^M$ 转换为哈希编码 $B \in \{0, 1\}^M$:

$$f_q(x) = \begin{cases} 1, & x \geq 0.5 \\ 0, & x < 0.5 \end{cases}, \quad (16)$$

$$B = f_q(B'). \quad (17)$$

为了减小二值化时的量化误差,通过损失项约束端到端模型训练时的编码,使编码器产生草图和图像编码 B'_s 和 B'_i 时更接近离散值,损失项的表达式为

$$L_q = \|B'_s - f_q(B'_s)\|_1 + \|B'_i - f_q(B'_i)\|_1. \quad (18)$$

综上所述,结合量化损失,本文训练网络的损失函数为

$$L = L_{Triplet} + \alpha L_{cls} + \beta L_q. \quad (19)$$

3 实验与结果分析

3.1 数据集及评价指标

在以下两个数据集下进行实验和性能评估与对比: 扩展的 Sketchy 数据集^[17-18]包含 125 类物体, 共计 73002 张图像和 75471 张草图; 扩展的 TU-Berlin 数据集^[19-20]包含 250 类物体, 每一类有 80 张草图共计 20000 张, 相对应的自然图像共 204489 张。两数据集均通过众包平台收集非绘画专业工作者的草图作品得到, 因此所提方法并不要求使用者受过专业绘画训练。

为了保证公平的实验对比, 采取与文献^[10-11, 18]中相同的测试数据划分方法, 在 Sketchy 数据集和 TU-Berlin 数据集上分别从每一类抽取 50 张和 10 张草图作为各自的测试集, 并抽取了与测试集相同数量的草图作为验证集, 其余的草图与自然图像全部用于训练。

基于草图的图像检索常用评价标准为平均精度均值 (mAP) 和 Precision@ k 。设有 N_q 张查询草图, 图库中有 N_g 张自然图像, Precision@ k 为检索前 k 个结果的准确率, 即前 k 个返回结果中与查询草图相关的图像结果的占比。mAP 为更加精确的检索指标, 正确检索结果越靠前, 得到的 mAP 指标越高。

$$P_A = \frac{1}{N_+} \sum_{k=1}^{N_+} \frac{N_+^k}{k} \times p_{\text{position}}(k), \quad (20)$$

$$p_{\text{position}}(k) = \begin{cases} 1, & \text{if returned image at position } k \text{ is positive} \\ 0, & \text{otherwise} \end{cases}, \quad (21)$$

$$P_{m\Lambda} = \frac{1}{N_q} \sum_{i=1}^{N_q} P_{Ai}, \quad (22)$$

式中: N_+^k 表示前 k 个检索结果中正确的数量。

3.2 实验环境与设置

在 GPU 为 2080Ti 的平台上实现实验, 在 Pytorch 框架下实现代码。超参数设置: $\alpha = 2, \beta = 10^{-4}, m = 0.5\sqrt{M}$, 其中 M 为编码长度, 取码长为 128。训练时初始学习率为 10^{-4} , 训练 40 代后, 以在验证集上的测试结果为准调整学习率。每 3 代在验证集上计算一次 mAP, 当连续 4 次验证集测试指标无提升时学习率下降为原来的 1/10, 当学习率下降至 10^{-7} 且连续 4 次验证集指标无提升时结束训练。测试阶段, 按照图库中自然图像与查询草图的编码距离排序, 得到检索结果。

3.3 与代表性方法对比及检索样例

为了公平地与离散哈希方法^[10,18]以及未对连续特征离散化的 Semi3-Net^[11]进行对比, 分别测试了网络输出离散编码和连续编码的性能指标, SBIR 代表性方法在 Sketchy 和 TU-Berlin 数据集上的对比如表 1 所示。

与使用边缘图代替自然图像检索的方法^[21,25]相比, 所提方法获得了更好的检索精度, 这主要是因为从自然图像提取的边缘图像虽然减小域间差异, 但舍弃

表 1 不同方法在 Sketchy 和 TU-Berlin 数据集上的 mAP 指标
Table 1 mAP of different methods on Sketchy and TU-Berlin datasets

Method	Sketchy	TU-Berlin
3D shape ^[21]	0.084	0.054
HOG ^[22]	0.115	0.091
GF-HOG ^[6]	0.157	0.119
SHELO ^[23]	0.161	0.123
LKS ^[24]	0.190	0.157
SaN ^[7]	0.208	0.154
Siamese CNN ^[25]	0.481	0.322
Siamese-AlexNet ^[18]	0.518	0.367
GN Triplet ^[17]	0.529	0.187
Triplet-AlexNet ^[18]	0.573	0.448
DSH ^[18]	0.783	0.570
GDH ^[10]	0.810	0.690
Semi3-Net ^[11]	0.916	0.800
Proposed method (discrete)	0.932	0.797
Proposed method (continuous)	0.933	0.799

自然图像只用边缘图来检索会损失一部分图像信息。与仅使用草图和自然图像对的 SBIR 方法^[17]相比, 所提方法通过将边缘图作为中间模态有效地缩小了自然图像和草图之间的域差异。与新近的 Semi3-Net^[11]方法相比, 所提方法通过联合空间注意有效利用了自然图像与边缘图在空间上关键区域的相似性, 在 Sketchy 数据集上的 mAP 指标提高 1.7 个百分点, 在 TU-Berlin 数据集上性能相仿, 说明在草图数据充足的情况下所提方法有更好的特征提取和编码性能。

图 2 是测试集中部分图像返回的检索结果, 共列出 7 张查询草图, 每张返回检索结果的前 8 个, 方框标注的自然图像为不同类的错误返回结果, 无方框标注的自然图像为正确的检索结果。尽管分别以城堡和鸭子草图为查询时有教堂和海鸥类别的自然图被作为结

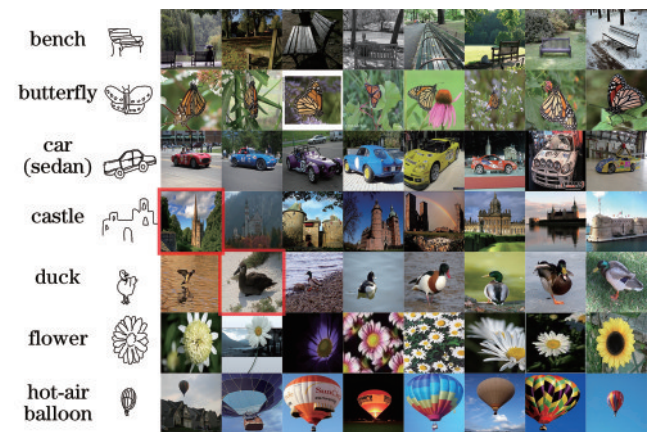


图 2 所提方法在 Sketchy 数据集上的检索样例
Fig. 2 Some retrieval examples of proposed method on Sketchy dataset

果返回,但这些错误结果仍与检索目标具有相似的语义和图形特征。检索结果表明,所提检索方法能够同时兼顾语义与图像的相似性。

3.4 消融实验

为验证所提方法的有效性,设计了一系列分别排除一部分网络结构或损失函数项的消融实验,以分别验证引入边缘图、联合空间注意、针对跨域检索的函数、辅助分类任务的有效性,并在 Sketchy 数据集上进行测试,消融实验部分的指标均为输出连续编码测得,结果如表 2 所示。

表 2 Sketchy 数据集上的消融实验
Table 2 Ablation study on Sketchy dataset

Method	mAP
w/o edge map branch	0.899
w/o spatial co-attention	0.923
w/o intra-class loss	0.928
w/o quantization loss	0.931
Full model	0.933

为了验证中间模态边缘图引入的效果,设计了去除边缘图分支的简化网络,与完整模型相比,去除边缘图分支的简化网络的 mAP 下降了 3.4 个百分点,表明中间模态的引入有效地缩小了草图与自然图像的域差异。在维持其他基础网络结构不变的前提下,将联合空间注意拆分为边缘图和自然图像,分别计算空间注意掩模并加权,在测试集上所提跨域联合空间注意能带来 1 个百分点的 mAP 性能提升,这表明所提联合空间注意方法能有效地使边缘图和自然图像关注同一区域,产生的空间注意掩模比两者各自计算得到的结果更有效。使用原始三元组损失和去除量化损失的损失函数分别训练模型,以评估所提损失函数的有效性。实验结果表明:针对跨域检索任务提出的类内距离约束损失能将两个不同域的输入更有效地对齐到公共特征空间,得到了类内分布更紧凑的图像编码,量化损失通过迫使码值接近离散值进一步提高了模型的检索性能。

为验证将编码结果 B' 作为辅助分类任务输入的有效性,将辅助分类器的输入改为编码之前的高层特征 F_H^S 和 F_H^E 作为对照,实验结果如表 3 所示。实验结果表明:相比直接以高层特征作为分类器的输入,使用图像编码 B' 作为输入可以使编码保留更多语义信息。

表 3 Sketchy 数据集上的辅助分类器实验
Table 3 Evaluation of auxiliary classifier on Sketchy dataset

Input of auxiliary classifier	Precision	mAP
High-level feature	0.996	0.911
Code	0.999	0.933

4 结 论

提出了基于跨域联合空间注意和三元组网络的草

图图像检索算法,通过端到端训练模型实现了很好的检索性能。研究表明:在基于草图的图像检索任务中,引入边缘图作为中间模态有助于减小域差异,在此基础上边缘图与自然图像的联合空间注意比单独实现的注意力机制更有效;在三元组网络下,以编码结果作为辅助分类器的输入可以使编码保留更多的语义信息;加大对类内距离约束的损失函数能进一步提升跨域检索模型的性能。从检索结果的样例可以看出,所提算法能有效检索到与查询草图语义相关及图形相似的自然图像,在 Sketchy 和 TU-Berlin 数据集上的性能指标也优于大多数代表性方法,实验证明所提算法是一种有效的基于草图的图像检索方法。受限于 Canny 边缘检测方法,本文从具有非边缘高频背景(如草地)的图像中提取的边缘图噪声较多,一定程度上影响了相应类别图像的特征提取,后续的研究可以尝试通过风格迁移等生成方法获取边缘图从而进一步提高检索性能。

参 考 文 献

- [1] Chang S K, Hsu A. Image information systems: where do we go from here? [J]. IEEE Transactions on Knowledge and Data Engineering, 1992, 4(5): 431-442.
- [2] Saavedra J M, Bustos B. Sketch-based image retrieval using keyshapes[J]. Multimedia Tools and Applications, 2014, 73(3): 2033-2062.
- [3] Hu R, Collomosse J. A performance evaluation of gradient field HOG descriptor for sketch based image retrieval[J]. Computer Vision and Image Understanding, 2013, 117(7): 790-806.
- [4] Hu R, Wang T H, Collomosse J. A bag-of-regions approach to sketch-based image retrieval[C]//2011 18th IEEE International Conference on Image Processing, September 11-14, 2011, Brussels, Belgium. New York: IEEE Press, 2011: 3661-3664.
- [5] Saavedra J M, Bustos B. An improved histogram of edge local orientations for sketch-based image retrieval[M]//Goesele M, Roth S, Kuijper A, et al. Pattern recognition. Lecture notes in computer science. Heidelberg: Springer, 2010, 6376: 432-441.
- [6] Hu R, Barnard M, Collomosse J. Gradient field descriptor for sketch based retrieval and localization[C]//2010 IEEE International Conference on Image Processing, September 26-29, 2010, Hong Kong, China. New York: IEEE Press, 2010: 1025-1028.
- [7] Yu Q, Yang Y X, Song Y Z, et al. Sketch-a-net that beats humans[C]//Proceedings of the British Machine Vision Conference 2015, September 7-10, 2015, Swansea. Durham: BMVA Press, 2015: 1-12.
- [8] Yu Q, Liu F, Song Y Z, et al. Sketch me that shoe[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 799-807.
- [9] Schroff F, Kalenichenko D, Philbin J. FaceNet: a unified embedding for face recognition and clustering[C]//2015

- IEEE Conference on Computer Vision and Pattern Recognition, June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 815-823.
- [10] Zhang J Y, Shen F M, Liu L, et al. Generative domain-migration hashing for sketch-to-image retrieval[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11206: 304-321.
- [11] Lei J J, Song Y X, Peng B, et al. Semi-heterogeneous three-way joint embedding network for sketch-based image retrieval[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2020, 30(9): 3226-3237.
- [12] 张文秀, 朱振才, 张永合, 等. 基于残差块和注意力机制的细胞图像分割方法[J]. 光学学报, 2020, 40(17): 1710001.
Zhang W X, Zhu Z C, Zhang Y H, et al. Cell image segmentation method based on residual block and attention mechanism[J]. Acta Optica Sinica, 2020, 40(17): 1710001.
- [13] 徐志刚, 闫娟娟, 朱红蕾. 基于多尺度残差注意力网络的壁画图像超分辨率重建算法[J]. 激光与光电子学进展, 2020, 57(16): 161012.
Xu Z G, Yan J J, Zhu H L. Mural image super resolution reconstruction based on multi-scale residual attention network[J]. Laser & Optoelectronics Progress, 2020, 57(16): 161012.
- [14] 鞠默然, 罗江宁, 王仲博, 等. 融合注意力机制的多尺度目标检测算法[J]. 光学学报, 2020, 40(13): 1315002.
Ju M R, Luo J N, Wang Z B, et al. Multi-scale target detection algorithm based on attention mechanism[J]. Acta Optica Sinica, 2020, 40(13): 1315002.
- [15] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11211: 3-19.
- [16] Song Y X, Lei J J, Peng B, et al. Edge-guided cross-domain learning with shape regression for sketch-based image retrieval[J]. IEEE Access, 2019, 7: 32393-32399.
- [17] Sangkloy P, Burnell N, Ham C, et al. The sketchy database: learning to retrieve badly drawn bunnies[J]. ACM Transactions on Graphics, 2016, 35(4): 1-12.
- [18] Liu L, Shen F M, Shen Y M, et al. Deep sketch hashing: fast free-hand sketch-based image retrieval[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 2298-2307.
- [19] Eitz M, Hays J, Alexa M. How do humans sketch objects? [J]. ACM Transactions on Graphics, 2012, 31(4): 1-10.
- [20] Zhang H, Liu S, Zhang C Q, et al. SketchNet: sketch classification with web images[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 1105-1113.
- [21] Wang F, Kang L, Li Y. Sketch-based 3D shape retrieval using Convolutional Neural Networks[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition, June 7-12, 2015, Boston, MA. New York: IEEE Press, 2015: 1875-1883.
- [22] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]//2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 20-25, 2005, San Diego, CA, USA. New York: IEEE Press, 2005: 886-893.
- [23] Saavedra J M. Sketch based image retrieval using a soft computation of the histogram of edge local orientations (S-HELO)[C]//2014 IEEE International Conference on Image Processing, October 27-30, 2014, Paris, France. New York: IEEE Press, 2014: 2998-3002.
- [24] Saavedra J M, Barrios J M. Sketch based Image Retrieval using Learned KeyShapes (LKS)[C]//Proceedings of the British Machine Vision Conference 2015, September 7-10, 2015, Swansea, UK. Durham: BMVA Press, 2015: 7.
- [25] Qi Y G, Song Y Z, Zhang H G, et al. Sketch-based image retrieval via Siamese convolutional neural network [C]//2016 IEEE International Conference on Image Processing, September 25-28, 2016, Phoenix, AZ, USA. New York: IEEE Press, 2016: 2460-2464.