

基于传统方法与深度学习方法的图片相似度算法比较^①



王华溢, 黄要诚, 蔡波

(武汉大学 国家网络安全学院, 武汉 430072)

通信作者: 蔡波, E-mail: caib@whu.edu.cn

摘要: 图片相似度比对作为计算机视觉的一个研究方向, 具有广泛的应用前景, 例如人脸识别、行人重识别和目标跟踪等。然而, 目前有关图片相似度算法的总结和归纳相对较少, 并且将其应用在实际工业生产中也存在挑战。本文总结了传统图像处理算法和深度学习图像处理算法在图片相似度比对方面的原理与表现, 旨在选取最佳的算法用于药品图片相似度比对的场景中。在传统图像处理算法中, ORB 算法在测试集上表现最佳, 准确率为 93.09%; 在深度学习算法中, 采用改进的孪生网络结构、发明了一种标签生成法、设置特定的数据增强策略并增加一个特征面分类网络, 从而提高了训练效率和准确率。最终的测试结果显示, 改进的孪生网络表现最佳, 可以实现 98.56% 的准确率和 27.80 次/s 的推理速度。综上所述, 采用改进的孪生网络算法更适用于药品图片的快速比对, 并且有望在未来的医药行业中得到广泛应用。

关键词: 图片相似度算法; 传统图像处理; 深度学习; 孪生网络; 分类网络

引用格式: 王华溢, 黄要诚, 蔡波. 基于传统方法与深度学习方法的图片相似度算法比较. 计算机系统应用, 2024, 33(2): 253-264. <http://www.c-s-a.org.cn/1003-3254/9413.html>

Comparison of Image Similarity Algorithms Based on Traditional Methods and Deep Learning Methods

WANG Hua-Yi, HUANG Yao-Cheng, CAI Bo

(School of Cyber Science and Engineering, Wuhan University, Wuhan 430072, China)

Abstract: As a research direction of computer vision, image similarity comparison has a wide range of applications, such as face recognition, person re-identification, and target tracking. However, the summary and induction of image similarity algorithms are relatively few, and there are challenges in applying them to actual industrial production. This study summarizes the principle and performance of traditional image processing algorithms and deep learning image processing algorithms in image similarity comparison, aiming to select the best algorithm for the scene of drug image similarity comparison. Among the traditional image processing algorithms, the ORB algorithm performs best on the test set, with an accuracy of 93.09%. In the deep learning algorithm, the study adopts an improved Siamese network structure, invents a label generation method, sets a specific data augmentation strategy, and adds a feature surface classification network to improve the training efficiency and accuracy. The final test results show that the improved Siamese network performs best and can achieve an accuracy of 98.56% and an inference speed of 27.80 times/s. In summary, the improved Siamese network algorithm is more suitable for the fast comparison of drug images and is expected to be widely used in the future pharmaceutical industry.

Key words: image similarity algorithm; traditional image processing; deep learning; Siamese network; classification network

① 收稿时间: 2023-07-01; 修改时间: 2023-08-08, 2023-10-08; 采用时间: 2023-10-16; csa 在线出版时间: 2023-12-26
CNKI 网络首发时间: 2023-12-28

随着数字图像处理技术的不断发展和应用, 图片相似度算法在计算机视觉、图像检索等领域中扮演着重要的角色. 截至目前, 许多相关文献将一些图片相似度算法用于实际应用中, 例如人脸识别^[1]、签名比对^[2]、车牌比对^[3]、目标跟踪^[4]、行人重识别^[5]等, 然而对于传统图片相似度比对算法与基于深度学习的图片相似度比对算法的概括与总结, 很少有文献提及. 为此, 本文归纳与总结了一些传统图片相似度比对算法与基于深度学习的图片相似度比对算法, 并将这些算法用于药品图片相似度比对这一实际应用中. 表1列举了本文要对比的一些图片相似度比对算法.

表1 常用图片相似度比对算法

图像比对算法	速度	复杂度	泛化能力
直方图法 ^[6]	中	低	低
余弦相似度法 ^[7]	低	中	低
全局哈希法 ^[8]	高	低	低
局部哈希法 ^[9,10]	中	中	中
孪生网络 ^[2]	低	高	高

表1中, 直方图法、余弦相似度法、全局哈希法、SIFT与ORB算法属于传统图像处理范畴, 孪生网络是基于深度学习的图像处理算法.

直方图法通过统计图像中不同像素点的个数, 得出像素点分布直方图, 进而使用几种相关性计算方法计算两张图片像素点分布直方图之间的相关性, 最后根据计算结果判断两张图片是否相似; 余弦相似度法常用于文本相似度或图片相似度的计算. 在应用余弦相似度比较两张图片是否相似之前, 需先将图片转换为向量形式, 之后, 通过计算两向量的余弦相似度来判断两张图片是否相似; 哈希法分为全局哈希法和局部哈希法. 其中, 全局哈希法包括均值哈希(AHash)、感知(PHash)和差值哈希(DHash)算法, 基本原理是将图像转换为一个固定长度的二进制编码, 通过计算两编码的距离判断两张图片是否相似; 局部哈希法包括SIFT(scale-invariant feature transform)、ORB(oriented FAST and rotated BRIEF)算法, 基本原理是找到并量化图像中的特征点, 通过计算两张图片量化特征点间的距离判断是否相似; 孪生网络是一种特殊的神经网络结构, 其中包含两个完全相同的子网络, 也被称为“孪生子网络”. 这两个子网络具有相同的结构和参数, 并且共享权重, 但是在训练过程中它们被独立地处理不同的输入数据. 其常被用于比较两个输入数据之间

的相似度或距离, 在图像比对、文本匹配、人脸识别等领域被广泛应用. 表2列出了传统图像处理算法与深度学习图像处理方法的优缺点.

表2 传统方法与深度学习方法的优缺点

图像处理	优点	缺点
传统方法	精度高、可解释性强、资源消耗低	特征提取效果不佳、泛化能力不足
深度学习方法	特征提取能力强、泛化能力强	模型可解释性差、需要大量训练数据

表2中, 传统图像处理方法经过长时间的发展和研究, 其算法和模型已经非常成熟, 具有相对较高的精度; 此外, 设计出的算法会利用专家的先验知识, 使算法的结果更加可解释, 并有利于深入研究算法的性能. 相较于深度学习方法, 传统方法的算法和模型规模较小, 需要的计算资源和时间较少. 然而, 对于数据量较大、复杂度高的任务, 传统方法的效率和精度会有所下降, 其泛化能力不足, 而且对于一些复杂的图像特征提取效果也不佳.

深度学习方法能够从原始数据中自动学习并提取出高阶特征, 从而处理更加复杂、抽象的问题, 减少了人工特征提取的繁琐过程. 同时, 其还具有良好的可扩展性和可伸缩性, 也可以实现模型的并行化和加速, 具有较强的泛化能力. 然而, 深度学习模型的结构和参数规模较大, 其内部的处理机制比较复杂, 导致其结果往往难以被解释和理解. 此外, 对于数据量较小的任务, 深度学习算法往往需要更多的数据才能达到较好的效果, 且存在过拟合的风险.

综上所述, 表1中的传统处理方法普遍存在计算速度偏高、复杂度与泛化能力偏低的特点, 基于深度学习的图像处理算法普遍存在计算速度偏低、复杂度与泛化能力偏高的情况. 然而, 在实际应用中, 通常利用GPU加速神经网络的训练和推理, 因此使用孪生网络进行推理的速度与传统方法进行推理的速度旗鼓相当.

本文主要贡献包括以下4个方面.

(1) 我们首先比较了一系列图片相似度比对算法在药品图片相似度比对任务中的表现. 其次, 我们改进了效果最好的孪生网络算法, 进一步提升了其性能. 实验结果显示, 该算法的推理速度约为38次/s, 准确率达到了98.5%. 最终, 该算法成功应用于该任务.

(2) 为解决药品图片特征点少所造成相似度匹配

失效的情况,本文在图片相似度比对前先判断图片中特征点数量是否达到比对标准.我们设计了一套准则来制作标签,使用 ResNet 分类网络,最终在测试集上的分类准确率达 95.51%.

(3) 鉴于孪生网络需要成对的输入,本文设计出一种标签制作方法,在训练期间既能保证数据在 GPU 与 CPU 之间的传输速率,又能保证正负样本均衡.

(4) 鉴于样本图像受拍摄现场光照以及药盒颜色的影响,本文设计了一系列数据增强策略来模拟现实情况,增强孪生网络的泛化能力.

1 相关工作

直方图法的起源可以追溯到 19 世纪初,当时英国的数学家皮尔逊提出了统计学中的频数分布表,为后来直方图法的发展提供了基础.在图像相似度比对领域,直方图法也得到了广泛的应用.王宁邦等人^[11]利用直方图匹配法,使用巴氏系数 (Bhattacharyya) 判断两张资产图片的直方图是否相似,进而实现资产图片的核对;薛向阳等人^[12]提出了一种基于聚类的颜色直方图匹配法,用于检索自然风格相近的图像;芦俊等人^[13]使用直方图法并结合差分进化算法进行目标图片与模板图片的匹配;邹承明等人^[14]使用直方图法与灰度共生矩阵法结合的方式进行图片匹配;陈会鸽等人^[15]使用目标图片直方图与模板图片直方图作差的方式进行中央空调风管附尘度检测;丁建平等^[16]使用相关系数计算不同视频帧中目标图像灰度直方图的相关性,从而进行铁路目标追踪.

在 18 世纪,瑞士数学家欧拉在研究数学向量时首次提出了余弦相似度的概念.后来余弦相似度法也被广泛应用于图像处理、文本处理等领域.李阳齐等人^[17]将音乐属性转化为特征向量,使用余弦相似度计算不同音乐间的相似性;刘冰等人^[18]使用余弦相似度进行当前位置强度向量与定位服务器数据库中的位置强度向量比对,从而定位当前位置;魏玮等人^[19]利用余弦相似度更正 SIFT 算法误匹配的人脸特征,进而提高人脸匹配的准确度.

哈希法起源于 1992 年,当时 NASA 的天文学家发明了一种叫做“星形哈希 (star-hash)”算法^[20],用于识别并匹配天文图像.徐贤局等人^[21]使用 AHash 得到无疵点织物和有疵点织物的图片向量,使用汉明距离计算差异,定位疵点位置;赵杰等人^[22]利用 PHash 得到不同

图片区域的向量,通过计算相邻图片块向量的相似性及位移矢量检测图像篡改区域;郭玮^[23]利用分块图像的 PHash 进行不同区域类型遥感图像的匹配;吴子岳等人^[24]使用拉普拉斯变换与 DHash 结合的方式得到不同视频帧目标区域的向量,利用汉明距离计算其差异,从而实现目标跟踪.

图像间的相似度比对可以间接转化为特征点间的相似度比对.特征检测算法起源于 20 世纪六七十年代,当时的研究主要是基于手工设计的滤波器和边缘检测算子,例如 Sobel 算子^[25]、Laplacian 算子^[26]等,用于寻找图像中的边缘和角点等特征.1996 年, Lowe 等人^[9]提出 SIFT 算法.于明等人^[27]使用 SIFT 算法进行充电桩的识别与定位;韩宇等人^[28]使用改进的 SIFT 算法实现无人机航拍图片的快速匹配;任晶秋等人^[29]使用 Harris 与 SIFT 结合的方式进行图片的快速匹配.

ORB 是另一种特征点检测与描述算法,由 Rublee 等人^[10]于 2011 年提出.其设计在保持速度优势的同时提高 SIFT 特征算法的精度,以适应嵌入式设备等计算能力有限的场景.杨金玲等人^[30]使用改进的 ORB 算法进行近红外和可见光图像的特征点检测与匹配;姚三坤等人^[31]使用改进的 ORB 算法进行视频不同帧场景的匹配.

孪生网络最初应用于签名识别和文本相似度计算.该模型是由 Bromley 等人^[2]在 1993 年提出,它包含两个完全相同的子网络,子网络间共享权重.训练时并行处理两个输入序列,对输出进行比较并计算损失.2006 年, Hadsell 等人^[32]提出了一种计算孪生网络损失的方法: Contrastive Loss, 该方法利用欧氏距离来度量两个样本之间的相似度;火善栋^[33]使用孪生网络进行人脸识别,训练时使用 Contrastive Loss;吴正文^[34]使用 U-Net 类型的孪生网络,并结合注意力机制进行遥感建筑物变化检测.

以上文献只是针对某个任务,选取一种或两种图片相似度比对算法进行分析,为此,本文将使用上述提及的所有算法用于药品图片相似度比对,并从中选取最佳算法用于实际工业生产中.

2 图片相似度算法对比研究

2.1 直方图法

直方图法统计图片的颜色和灰度分布情况.计算直方图时,将像素值范围划分成若干区间,每个区间内

像素点数量即为该区间的统计量。

本文实验中, 区间范围为 0-255, 每个像素值代表一个区间. 图 1 展示了两张样例图片, 图 2 展示了图 1 中两张图片的像素分布直方图。



图 1 两张样例图片

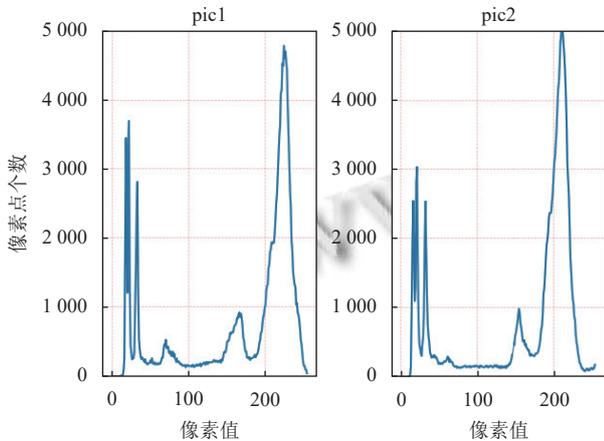


图 2 两张图片的像素分布直方图

计算像素分布直方图的相关性方法有 Correlation、Chi-Square、Intersection、Bhattacharyya distance.

Correlation 计算公式见式 (1):

$$d(H_1, H_2) = \frac{\sum_I (H_1(I) - \bar{H}_1)(H_2(I) - \bar{H}_2)}{\sqrt{\sum_I (H_1(I) - \bar{H}_1)^2 \sum_I (H_2(I) - \bar{H}_2)^2}} \quad (1)$$

其中, $\bar{H}_k = \frac{1}{N} \sum_j H_k(j)$, N 为直方图中 bin 的个数 (即组数). 计算结果最大值为 1, 最小值为 -1, 取绝对值后, 值越大, 相关度越高。

Chi-Square 计算公式见式 (2):

$$d(H_1, H_2) = \sum_I \frac{(H_1(I) - H_2(I))^2}{H_1(I)} \quad (2)$$

计算结果越小, 相关度越高, 最小值为 0, 最大值无上界。

Intersection 计算公式见式 (3):

$$d(H_1, H_2) = \sum_I \min(H_1(I), H_2(I)) \quad (3)$$

计算结果越大, 相关度越高, 最大值无上界, 最小

值为 0.

Bhattacharyya distance 计算公式见式 (4):

$$d(H_1, H_2) = \sqrt{1 - \frac{1}{\sqrt{H_1 H_2 N^2}} \sum_I \sqrt{H_1(I) \cdot H_2(I)}} \quad (4)$$

计算结果越小, 相关度越高, 最小值为 0, 最大值为 1. 表 3 列出了图 1 的相关性计算结果。

表 3 4 种度量计算结果

度量方法	计算结果	值域
Correlation	0.0078	[-1, 1]
Chi-Square	59.0565	[0, +∞]
Intersection	0.0769	[0, +∞]
Bhattacharyya distance	0.9993	[0, 1]

表 3 中, 4 种度量计算结果与实际几乎相反, 且易受外界因素影响, 例如亮度、药盒颜色等, 因此不能作为图片相似度比对的最佳算法, 在实验部分会进一步说明。

2.2 余弦相似度法

余弦相似度是一种常用的图片相似度比对法. 在预处理阶段, 将两张图片调整为同样大小、灰度化图片. 此步的目的是获得相同规模的图片向量, 方便计算. 利用得到的两个图片向量 A, B , 代入式 (5) 中, 得到结果。

$$cosine_similarity = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}} \quad (5)$$

结果在 [-1, 1] 之间, 值接近 1 表示两个向量相似, 接近 -1 表示两个向量不相似. 本文为余弦相似度设置一个合适的阈值, 例如 0.8, 表示当两张图片的余弦相似度大于等于 0.8 时, 判定两张图片相似. 以图 1 为例, 两张图片的余弦相似度为 0.801 83, 可以判定其相似。

2.3 全局哈希法

哈希法通过将图片转换为一系列数字表示, 从而快速比较两张图片的相似度. 全局哈希法将整张图片作为一个整体进行哈希处理, 得到一组唯一的哈希码. 全局哈希法包含: AHash、PHash、DHash.

AHash 计算过程如下.

(1) 缩小图片至 8×8 大小.

(2) 灰度化缩小后的图片, 计算出所有像素的平均灰度值 \hat{P} .

(3) 根据式 (6) 比较每个像素的灰度值 $P(i, j)$ 与平

均灰度值 \hat{P} , 得到由 0、1 组成的 64 位哈希码。

$$P'(i, j) = \begin{cases} 0, & \hat{P} \geq P(i, j) \\ 1, & \hat{P} < P(i, j) \end{cases} \quad (6)$$

PHash 在 AHash 的基础上增加 DCT (discrete cosine transform), 进而获取图片低频信息, 计算过程如下。

- (1) 将图片缩小至 32×32 大小并进行灰度化处理。
- (2) 进行 DCT 变换, 取前 8×8 个 DCT 系数, 计算其平均值 \widehat{dct} 。

(3) 根据式 (7) 比较每个 DCT 系数 $dct(i, j)$ 与 DCT 系数的平均值 \widehat{dct} , 得到由 0、1 组成的 64 位哈希码。

$$dct'(i, j) = \begin{cases} 0, & \widehat{dct} \geq dct(i, j) \\ 1, & \widehat{dct} < dct(i, j) \end{cases} \quad (7)$$

DHash 计算过程如下。

- (1) 将图片缩小到 9×8 并灰度化, 得到 $P(i, j)$ 。
- (2) 根据式 (8) 比较相邻像素之间的差异, 得到由 0、1 组成的 64 位哈希码。

$$P'(i, j) = \begin{cases} 1, & P(i, j-1) \geq P(i, j) \\ 0, & P(i, j-1) < P(i, j) \end{cases} \quad (8)$$

得到两张图片的哈希码后, 统计对应位置元素相同的个数并设定阈值, 如果相同的个数大于等于阈值, 则判定两张图片相似。以图 1 为例, 表 4 展示了 3 种全局哈希法的计算结果。

表 4 3 种哈希计算结果

度量方法	计算结果	哈希码长度
AHash	32	64
PHash	23	64
DHash	26	64

表 4 中, 3 种方法的计算结果约占哈希码总长度的一半, 后续的实验会说明这 3 种算法并不适用。

2.4 SIFT 算法

SIFT 算法是一种尺度不变的特征转换算法, 能够定位图像中的局部性特征, 在尺度空间中寻找极值点, 并提取出其位置、尺度、旋转不变量。

SIFT 算法流程分为 3 步: 尺度空间检测和极值点定位、特征点方向赋值、特征点描述子生成。

(1) 尺度空间检测和极值点定位

为使特征点具备尺度不变性, SIFT 算法使用高斯核函数 $G(x, y, \sigma)$ 对图像 $I(x, y)$ 进行不同尺度的滤波 (卷积运算), 得到一系列不同尺度的尺度空间 $L(x, y, \sigma)$, 见式 (9)、式 (10)。

$$L(x, y, \sigma) = G(x, y, \sigma) \times I(x, y) \quad (9)$$

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (10)$$

之后使用高斯差分对其中的尺度空间进行极值点定位。高斯差分尺度空间的计算见式 (11)。

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) \times I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \quad (11)$$

图 3 展示了高斯差分过程, SIFT 算法取 4 层差分空间的中间两层做极值点检测, 见图 4。每一个采样点与其相邻的 26 个点进行比较, 如果采样点的值是相邻点中的最大值或最小值, 则被视为极值点。

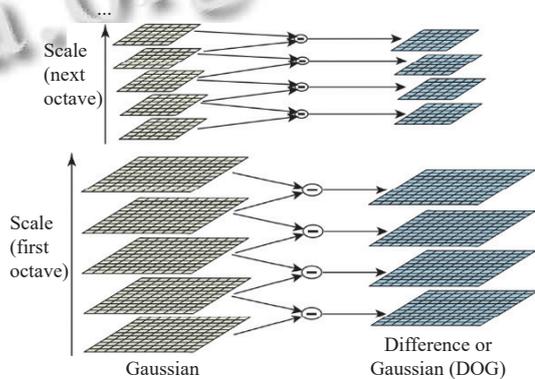


图 3 高斯差分

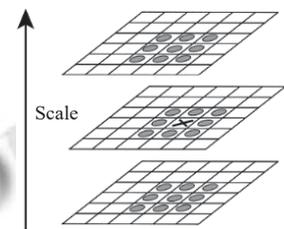


图 4 极值点检测

(2) 特征点方向赋值

为使特征点具备旋转不变性, 需要为每一个特征点分配一个基准方向。对于某尺度下的图像 $L(x, y)$, 用像素插值估计图像梯度幅值 $m(x, y)$ 与方向 $\theta(x, y)$, 见式 (12)。之后统计以特征点 $3\sigma' = 3 \times 1.5\sigma$ 为中心范围内的像素梯度幅值及方向。

$$\begin{cases} m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \\ \theta(x, y) = \tan^{-1} \frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \end{cases} \quad (12)$$

利用梯度直方图统计法, 将得到的梯度方向划分成 36 个区域 (每区域 10°), 取直方图峰值的方向定为

主方向.

(3) 特征点描述子生成

此时每一个特征点都包含位置、尺度与方向信息, 根据这些信息计算特征点的描述子. 式 (13) 划定以关键点为中心的距离范围, 默认 $d=4$.

$$r = \frac{3\sqrt{2}\sigma(d+1)}{2} \quad (13)$$

图 5 展示 $d=2$ 时, 关键点梯度直方图统计结果, 每个区域 45° , 一个关键点可由 $d \times d \times 8$ 位向量表示. 实际每个关键点会产生 128 位特征描述子.

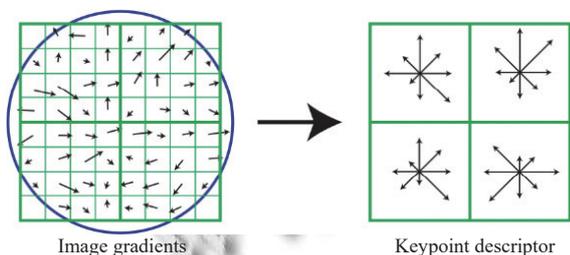


图 5 关键点描述子的生成

特征匹配时, 使用曼哈顿距离度量两张图片特征点描述子之间的距离, 并采用 Brute-Force 算法找到最接近的特征点. 设定阈值, 当成功匹配的特征点对数量超过阈值时, 认定两张图片相似. 在图 6 中, 使用 SIFT 算法对图 1 中的特征点进行匹配, 总共得到 240 对相似关键点 (两张图片分别检测到 391、368 个关键点).

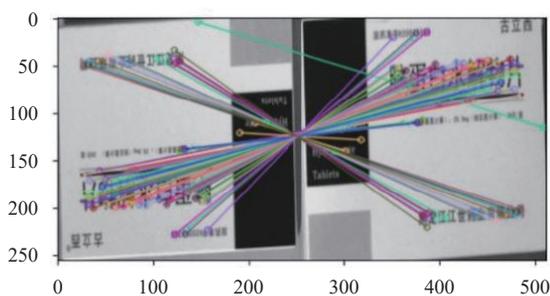


图 6 SIFT 算法匹配结果

2.5 ORB 算法

ORB 算法是一种快速检测图像上的关键点并创建其特征向量的算法, 由 oriented FAST 特征检测^[35]与 rotated BRIEF 特征描述算法^[36]组成. 为满足缩放不变性, 首先构建图像金字塔, 见式 (14). 之后对每一层图像进行关键点检测与描述.

$$\begin{cases} \sigma_s = \sigma_0^s \\ Size = \left(H \times \frac{1}{\sigma_s}\right) \times \left(W \times \frac{1}{\sigma_s}\right) \end{cases} \quad (14)$$

其中, 第 1 行表示第 s 层的缩放尺度, 第 2 行表示第 s 层的图像尺寸, 默认 $\sigma_0 = 1.2$.

(1) Oriented FAST 特征检测算法步骤

1) 对每一个像素点, 确定其周围 $R \times R$ 范围内的边缘像素点, 若其中有连续 K 个像素值与中心点像素值相差大于阈值 T , 则该中心点视为关键点.

2) 使用 Harris 角点^[37]响应函数解决特征点分布密集的情况: 计算每个特征点的响应值 V , 选取前 N 个 V 值较大的关键点作为最终的关键点.

3) 使用强度质心法 (intensity centroid)^[38]计算特征点的方向, 强度质心法假设角的强度偏离其中心.

$$m_{pq} = \sum_{x,y} x^p y^q I(x,y) \quad (15)$$

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}}\right) \quad (16)$$

$$\theta = \arctan(m_{01}, m_{10}) \quad (17)$$

式 (15) 计算特征点所在图像块的矩量, $I(x, y)$ 为图像灰度值, $x, y \in [-r, r]$, r 为以特征点为中心的图像块半径, 通过式 (16) 计算出灰度值质心, 通过式 (17) 计算特征点与灰度值质心之间的方向 θ , 并将其作为主方向.

(2) Rotated BRIEF 特征描述算法

1) 在原始 BRIEF 特征描述算法中, 以特征点为中心, $S \times S$ 的范围内, 随机选取 N 组点对, 由式 (18)、式 (19) 计算出关键点的二进制特征描述.

$$\tau(p; x, y) = \begin{cases} 1: p(x) < p(y) \\ 0: p(x) \geq p(y) \end{cases} \quad (18)$$

$$f_n(p) = \sum_{1 \leq i \leq n} 2^{i-1} \tau(p; x_i, y_i) \quad (19)$$

其中, $p(x)$ 表示图像区块 p 在点 x 的灰度值, $f_n(p)$ 表示特征点的 n 位向量表示. 由于原始 BRIEF 对噪声敏感, 在计算式 (19) 前需进行平滑处理: 在 31×31 的图像块上使用 5×5 的子窗口平滑图像.

2) 鉴于图像平面旋转后造成匹配效果差的现象, 提出 steered BRIEF: 根据特征点的旋转方向调整 BRIEF 的方向. 具体地, 在特征点 $S \times S$ (默认 $S=31$) 范围内选取 n 对点集, 根据角度 θ 与旋转矩阵 R_θ , 得到旋转后的坐标, 见式 (20), 之后利用式 (21) 得到旋转后的特征点

描述,最后使用贪婪搜索选取 256 个低相关性的像素块对,作为最终的特征点描述子。

$$D_\theta = R_\theta D, D = \begin{pmatrix} x_1, x_2, \dots, x_n \\ y_1, y_2, \dots, y_n \end{pmatrix} \quad (20)$$

$$g_n(p, \theta) = f_n(p) \quad (21)$$

特征匹配时,以两幅图片特征点描述子间的汉明距离为度量,使用 Brute-Force 算法匹配距离两幅图最近的特征点. 设定阈值,当成功匹配的特征点对数量超过阈值时,认定两张图片相似. 在图 7 中,使用 ORB 算法对图 1 中的特征点进行匹配,总共得到 277 对相似关键点(两张图片分别检测到 410、387 个关键点)。

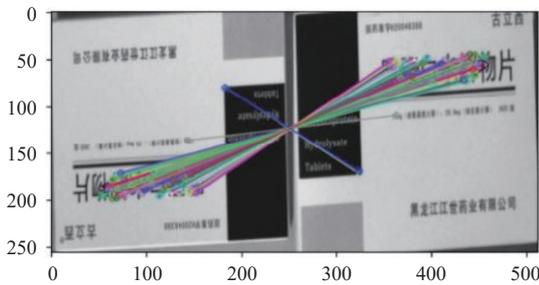


图 7 ORB 算法匹配结果

2.6 孪生网络

孪生网络是一种深度学习模型,该模型结构由两个相同的子网络组成,每个子网络都有相同的层次结构和参数.这两个子网络并行地接受输入,并通过共享权重来提取输入样本的特征表示.这种共享权重的设计使得网络能够学习到一种通用的特征表示,从而对输入样本进行比较和匹配.经典孪生网络结构见图 8。

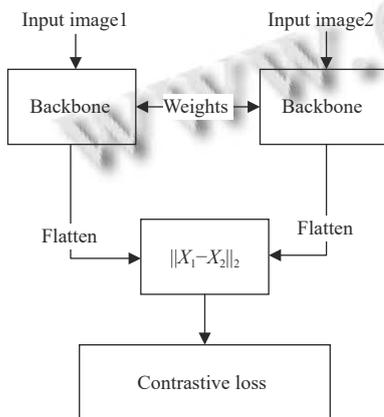


图 8 经典孪生网络结构

图 8 中,对于成对的输入图片,孪生网络将每个样本通过共享权重的 Backbone (子网络) 进行前向传播,

提取得到对应的图片特征向量.然后,将两特征向量维度从多维变换成一维 (Flatten 操作),在训练时使用对比损失 (式 (22)),测试时根据计算得到的欧氏距离 (式 (23)),设定阈值判定两张图片是否相似。

$$L = \frac{1}{2} (Y) D_w^2 + \frac{1}{2} (1 - Y) \{ \max(0, m - D_w)^2 \}^2 \quad (22)$$

$$D_w(X_1, X_2) = \|X_1 - X_2\|_2 = \sqrt{\sum_{i=1}^P (X_1^i - X_2^i)^2} \quad (23)$$

式 (22) 中,两样本相同则 $Y=1$, 否则 $Y=0$, 默认 $m=2$. 表示如果两样本的距离达到临界值,则断定两样本不同,网络区分正确,如果两样本相似,则只考虑样本间的距离作为损失。

另外,本文提出了一种改进的孪生网络结构 (见图 9),用来进一步提升识别准确率。

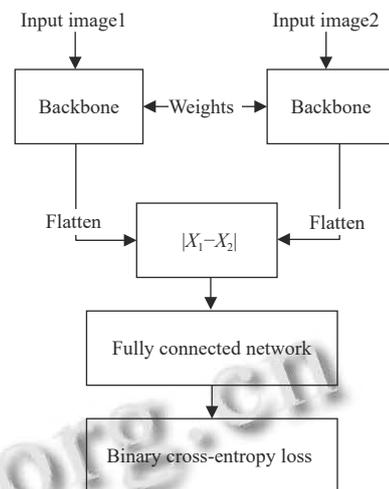


图 9 改进的孪生网络结构

类似地,图片经过 Backbone 提取成对的特征向量后,对成对向量作差并取绝对值,然后使用全连接网络进一步学习差异向量,训练时损失计算使用二分类交叉熵损失,见式 (24):

$$L = \frac{1}{N} \sum_i -[y_i \log(p_i) + (1 - y_i) \log(1 - p_i)] \quad (24)$$

其中, N 为样本数, y_i 表示样本 i 的标签, p_i 表示样本 i 预测为正类的概率. 测试时全连接网络的输出经过 Sigmoid 函数 (将结果约束在 $[0, 1]$ 范围内), 并设定阈值判定两张图片是否相似。

相比经典孪生网络结构,改进的孪生网络具有以下优势。

- (1) 充分捕捉输入样本间的差异信息. 通过将向量

作差(距离度量)得到差异向量,其中每个维度表示对应特征的差异.差异向量包含丰富的信息,因为它们反映了输入样本在各个特征上的差异程度.传入全连接层后,网络可以通过学习适当的权重和偏差来进一步提取和组合这些差异特征,从而计算出更具区分性的 Loss.然而传统方案中仅仅使用距离度量作为 Loss 计算,无法充分利用向量中的细微差异,从而限制了模型的表达能力.

(2) 训练稳定性高.在传统方案中,对比损失在训练过程中易出现 Loss=0 的情况($Y=0, m-D_w^2 \leq 0$).在训练初期,这会导致 Loss 值波动明显(时而为 0,时而很大).相反,改进的孪生网络结构使用交叉熵损失函数,在计算 Loss 前先计算 *Softmax*,为模型的预测结果提供概率分布,不会出现某类别概率完全为 0 的情况,这表明 Loss $\neq 0$,也意味着在训练初期 Loss 波动幅度相对较小,训练过程更稳定.

鉴于实时性的要求,Backbone 选用轻量级的 VGG16 和 ResNet18 卷积网络.VGG16 包含 5 个卷积模块,1 个全连接模块.每个卷积模块由多层 3 \times 3 卷积、ReLU 激活函数、最大池化层组成,全连接模块由 3 个全连接层组成;ResNet18 包含 5 个卷积模块、1 个均值池化层和 1 个全连接层.第 1 个卷积模块由 7 \times 7 卷积组成,其余 4 个卷积模块由多层 3 \times 3 卷积组成,每个模块中包含残差连接.

3 特征面分类策略

3.1 存在问题

如图 10,数据集中存在药盒面是完全空白、大部分空白或因光照及药盒颜色导致图片模糊不清的情况,这些情况会使特征检测与提取算法失效.



图 10 特征少的药品图片样例

3.2 问题解决

为解决上述问题,在进行特征提取前,设计一个分

类网络判断药盒面是否有较多的特征.选用 ResNet18 作为二分类网络,训练时采用二分类交叉熵损失(式(24)),最终分类网络在测试集的准确率达 95.51%,推理速度达 77.10 张/s.

4 实验

由于传统的特征检测方法是根据图像本身固有属性进行的(例如像素值的比较、检测图像的点),不需要利用大量数据去学习、提取图像的特征,所以对于传统方法的图片相似度比对只在测试集中展示其准确率.

4.1 数据集与标签

数据集含有 153 种药品,每种药品拍摄多张药盒 4-6 个面的图片并放入对应类别的文件夹中,每类药品图片数量从 59-752 张不等,平均 299 张/类别,数据集总共 45 826 张图片.表 5 展示了数据集所有药品中,每个药盒面图片数量的最小值、最大值以及平均值.

表 5 所有药品类别中药盒每面图片数量统计

药盒面	数量最小值	数量最大值	数量平均值
正面	9	160	59
反面	12	110	49
侧面1	10	200	55
侧面2	10	110	52
侧面3	10	265	54
侧面4	7	269	53

我们将 1-133 类中的图片划分为训练集,总计 42 437 张,134-153 类的图片划分为测试集,总计 3 389 张.图 11 展示了从训练集中随机挑选的 5 种药品图片,每种药品随机选取 4 个面.

4.1.1 特征面分类的标签制作

根据第 3.2 节提出的解决方案,通过设置图片的名称来制作标签.具体的,将图片名称设置为“种类标识_图片序号_特征面种类(1 表示满足特征匹配的图片)”格式,写入到.txt 文件内,方便训练时读取.

4.1.2 图片对比标签的制作

我们提出了一种标签制作的方法,能够在不影响 CPU 与 GPU 之间传输速度的情况下,保持正负样本的平衡.常规图片对的采集见算法 1.

算法 1. 常规训练数据采集算法

输入: 训练数据集 *ImageDataset*.

输出: *img1*、*img2*、*Label*. //图片对、标签

```

1)  $Img1 = Select(ImageDataset);$  /*从训练数据集中随机抽取一张图片*/
2)  $T = Select(0, 1)$  //从集合{0, 1}中随机取一个数
3) If  $T == 1$  //选取与  $Img1$  相同的图片//
4) While True
5)  $Img2 = Select(ImageDataset);$ 
6) If  $Img1 == Img2$ : Break;
7) End While
8) End If
9) Else //选取与  $Img1$  不同的图片
10) While True
11)  $Img2 = Select(ImageDataset);$ 
12) If  $Img1 != Img2$ : Break;
13) End While
14) End Else
15) Return  $Img1, Img2, Label$ 

```

但算法 1 存在以下缺点.

(1) 在训练的过程中, 随机的选择图片非常影响数据在 CPU 与 GPU 的传输速率.

(2) 由于随机性, 最终只会产生 4 万多对训练数据, 而且仅考虑了类间非相似图片对, 没有考虑类内非相似图片对.

为解决上述问题, 本文设计了一种高效的方法, 在训练前生成图片名称对, 在训练的时候仅仅读取文本中的内容即可, 图片命名格式同第 4.1.1 节的命名格式. 见算法 2.

算法 2. 高效训练数据采集算法

输入: 训练数据集 $ImageDataset$.

输出: Cls_Img_K 、 Cls_Img_{K+1} 、 $Label1$ 、 Cls_Img_K 、 $Img2$ 、 $Label2$.

```

1) For Each  $Cls\_Imgs$  in  $ImageDataset$ : /*获取每一类药品所有图片的路径与名称*/
2) For Each  $Img_K$  in  $Cls\_Imgs$ :
3) For Each  $Img_{K+1}$  in  $Cls\_Imgs$ :
4) If  $Cls\_Img_K == Cls\_Img_{K+1}$ :
5)  $T = Select(0, 1);$ 
6) If  $T == 1$ : //从同类中选非同面当负样本
7) While True
8)  $Img2 = Select(Cls\_Imgs);$ 
9) If  $Cls\_Img_K != Img2$ : Break;
10) End While
11) End If
12) Else //从非同类中选某面当负样本
13) While True
14)  $Img2 = Select(ImageDataset);$ 
15) If  $Cls\_Img_K != Img2$ : Break;
16) End While
17) End Else
18)  $Write(Cls\_Img_K, Cls\_Img_{K+1}, Label1)$  /*正样本对图片信息及标签写入文本*/
19)  $Write(Cls\_Img_K, Img2, Label2)$  /*负样本对图片信息及标签写入文本*/

```

算法 2 中, 正样本的获取充分考虑每一张图片, 负样本的获取不仅考虑了类间非相似图片, 而且考虑类内非相似图片, 最后训练集总共获得约 310 万对图片, 测试集总共获得约 12 万对图片, 在训练与测试的时候仅读取.txt 文件中的样本名称即可.



图 11 数据集图片示例

4.1.3 数据增强

在训练神经网络模型时, 为增强其泛化能力, 通常对训练数据进行数据增强. 针对药盒图片, 设置了图片上下翻转、水平翻转、旋转以及色彩变换. 其中上下翻转、水平翻转的概率设置为 0.5; 旋转角度设置为

0 或 180°; 为减轻光照及药盒颜色对模型的影响, 制定了色彩变换策略, 主要从亮度、对比度、饱和度以及色调 4 个方面进行调整. 由于数据集中存在很多忽明忽暗的图片, 所以在亮度方面调整力度较大, 参数范围在[0.5, 1.5]、对比度与饱和度参数范围在[0.9, 1.1], 色

调参数范围在[-0.1, 0.1].

4.2 训练与测试细节

本文的实验配置环境如下: 对于特征面分类网络与孪生网络的训练, 在武汉大学超算平台上完成, 操作系统为 CentOS 7.5 Linux, 显卡型号为 Nvidia Tesla V100 16 GB, 采用 PyTorch 深度学习框架; 对于所有图片相似度对比算法的测试, 在本地运行, 操作系统为 Windows 10, 显卡型号为 Nvidia RTX 3050Ti, 采用 PyTorch 深度学习框架及 OpenCV 算法包.

4.2.1 特征面分类网络

在训练与测试前, 将图片调整到 224×224 大小. 训练时, 设置 Epoch=30, Batchsize=64, 初始学习率 Lr=0.001, 采用 Adam 优化器并设置学习率每隔 7 epoch 削减 1/10.

4.2.2 孪生网络

在训练与测试前, 将图片调整到 224×224 大小. 训练时, 设置 Epoch=5, Batchsize=64, 初始学习率 Lr=0.0005, 采用 Adam 优化器并设置学习率每隔 2 Epoch 削减 1/2, 根据经验, 图 8 的两种方案在训练与测试时设置阈值为 0.8. 为增强孪生网络的泛化能力, 训练时不考虑特征面是否有效 (见第 3 节).

4.3 结果与时间对比

4.3.1 特征面分类结果

训练完成后, 使用准确率最高的权重进行测试, 结果见表 6.

表 6 特征面分类网络准确率与推理速度

数据集	准确率 (%)	推理速度 (次/s)
训练集	99.39	—
测试集	95.51	77.10

模型在测试集上效果表现良好, 且推理速度快, 满足实时性的要求.

4.3.2 传统图像算法比对结果

由于传统图像算法不需要大量训练数据, 在计算准确率与推理速度时, 仅利用测试集中的数据.

对于直方图法、余弦相似度法、全局哈希法, 由于方法相较于 ORB 与 SIFT 更不稳定, 我们从测试集抽取 1000 对图片进行正负样本对输出值分析, 见表 7.

表 7 中, 直方图法下标 1、2、3、4 分别代表 Correlation、Chi-Square、Intersection、Bhattacharyya distance 这 4 种度量方法. 对于直方图的 4 种距离计算方法, 正样本对的距离输出均值与负样本对的距离输

出均值基本一致, 所以很难设置一个阈值将其准确区分; 对于余弦相似度比较法, 情况与直方图法类似, 而且由于其中涉及矩阵运算, 且没有 GPU 加速, 速度会略慢; 同理, 3 种全局哈希法出现同样的情况; 对于 SIFT 与 ORB 算法, 正样本对的距离输出均值与负样本对的距离输出均值差别较大, 可以设定一个阈值进行准确判断: 取正样本输出均值与负样本输出均值的平均值, 计算得 SIFT 算法阈值约为 160, ORB 算法阈值约为 200. 之后使用计算得到的阈值测试 SIFT 与 ORB 算法在完整测试集中的表现, 见表 8.

表 7 正负样本对输出值分析

图像算法	正样本对输出均值	负样本对输出均值	推理速度 (次/s)
直方图法 ₁	0.9689	0.9771	11.78
直方图法 ₂	0.0521	0.0360	11.77
直方图法 ₃	197.2986	195.4865	12.42
直方图法 ₄	4.1986	3.2433	12.09
余弦相似度法	0.9337	0.9184	1.83
AHash	25.0540	28.6500	106.55
PHash	12.5000	16.8120	107.60
DHash	29.2080	31.4540	108.03
SIFT	227.8600	92.8060	26.18
ORB	285.7280	114.6220	30.87

表 8 SIFT 与 ORB 算法的表现

图像算法	阈值	准确率 (%)	推理速度 (次/s)
SIFT	160	82.38	22.69
ORB	200	93.09	31.88

由于 ORB 用到的算子运行速度快、准确率高, 所以 ORB 算法表现较好, 准确率达 93.09%, 推理速度达 31.88 次/s.

4.3.3 孪生网络比对结果

根据训练结果, 选择使网络损失最低、准确率最高的权重文件进行测试, 训练与测试结果见表 9.

表 9 孪生网络准确率与推理速度

Backbone模型	数据集	准确率 (%)	推理速度 (次/s)
VGG16 ₁	训练集	99.33	—
	测试集	97.37	31.64
VGG16 ₂	训练集	99.16	—
	测试集	96.57	30.62
ResNet18 ₁	训练集	99.84	—
	测试集	98.50	38.45
ResNet18 ₂	训练集	99.99	—
	测试集	95.51	38.04

表 9 中, 下标 1 代表使用改进的孪生网络结构, 下标 2 代表使用经典孪生网络结构. 可以看到使用改

进的孪生网络准确率更高,正是由于其使用表达能力强的全连接网络能够充分捕捉输入样本间的差异信息;而经典的孪生网络结构,对于两特征向量的关系学习较弱,仅仅在计算损失的时候进行关系学习.根据表 8、表 9,选取准确率 Top2 的模型进行最终的测试.

4.3.4 总结结果

本节将特征面分类网络与孪生网络结合,采用以下策略进行测试:对于测试过程中的某一对图片,用特征面分类网络判断这两张图片是否有一张满足特征面匹配的条件,如果是则进行图片比对,否则跳过(不参与最终准确率的计算),在测试集中的表现见表 10.

表 10 综合结果

网络模型	准确率 (%)	推理速度 (次/s)
ResNet18+VGG16 ₁	97.77	22.72
ResNet18+ResNet18 ₁	98.56	27.80

根据表 10,最终选择 ResNet18+ResNet18₁ 模型,其准确率高达 98.56%,推理速度达约 28 次/s,既满足了准确率的要求,又满足了实时性的要求.

5 结论

本文首先阐述了一系列图片相似度比对算法及相关工作,并着重改进孪生网络算法.之后设置实验,从中选取准确率最高、实时性较好算法用于实际应用中.

根据实验结果发现传统图像处理算法虽然在实时性方面表现优秀,且不需要用大量数据去训练,但其泛化能力较弱,并不满足实际要求;在深度学习图像处理算法上:为提高神经网络的泛化性能,设计一种高效的标签生成方法,在保证正负样本平衡的情况下,加速数据在 CPU 与 GPU 之间的传输速度,使训练过程更高效;在数据预处理方面,考虑到实际场景中光照、药盒颜色等影响因素,我们从亮度、对比度、饱和度与色调等方面按照一定概率进行数据增强;在对数据集进行检查时,发现并不是每个面都具有较强的特征,所以进行图片相似度比对前,先使用特征面分类网络过滤特征少的面;在网络结构方面,我们提出了一种改进的孪生网络,在 Backbone 为 ResNet18₁ 的情况下,测试集上准确率最高达到 98.50%,比同类型经典孪生网络结构准确率高出 2.99%;将特征面分类网络与改进的孪生网络结合,在测试集上最高准确率达 98.56%,推理速度达到约 28 次/s,同时满足准确性和实时性要求.

目前工作仍存在不足之处:在阈值的选择上仍有缺陷,现在是根据正负样本对产生的相似度值取平均来设定阈值或根据数据的分布情况设定阈值,但这种做法合理性较弱、主观性较强,为此如何科学且有效地选择阈值是未来要研究的方面.

参考文献

- 1 Taigman Y, Yang M, Ranzato M, *et al.* DeepFace: Closing the gap to human-level performance in face verification. Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE, 2014. 1701–1708.
- 2 Bromley J, Bentz JW, Bottou L, *et al.* Signature verification using a “Siamese” time delay neural network. International Journal of Pattern Recognition and Artificial Intelligence, 1993, 7(4): 669–688. [doi: 10.1142/S0218001493000339]
- 3 Eltoun IS, Xue ZJ. Automatic gate control system based on vehicle license plate recognition. ERSRA Publications, 2014, 3(8): 79–83.
- 4 Cen MB, Jung C. Fully convolutional siamese fusion networks for object tracking. Proceedings of the 25th IEEE International Conference on Image Processing (ICIP). Athens: IEEE, 2018. 3718–3722.
- 5 Ye M, Shen JB, Lin GJ, *et al.* Deep learning for person re-identification: A survey and outlook. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 2022, 44(6): 2872–2893. [doi: 10.1109/TPAMI.2021.3054775]
- 6 Rother C, Minka T, Blake A, *et al.* Cosegmentation of image pairs by histogram matching-incorporating a global constraint into MRFs. Proceedings of the 2016 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE, 2006. 993–1000.
- 7 Nguyen HV, Bai L. Cosine similarity metric learning for face verification. Proceedings of the 10th Asian Conference on Computer Vision. Queenstown: Springer, 2011. 709–720.
- 8 黄嘉恒, 李晓伟, 陈本辉, 等. 基于哈希的图像相似度算法比较研究. 大理大学学报, 2017, 2(12): 32–37. [doi: 10.3969/j.issn.2096-2266.2017.12.007]
- 9 Lowe DG. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 2004, 60(2): 91–110. [doi: 10.1023/B:VISI.0000029664.99615.94]
- 10 Rublee E, Rabaud V, Konolige K, *et al.* ORB: An efficient alternative to SIFT or SURF. Proceedings of the 2011 IEEE International Conference on Computer Vision. Barcelona:

- IEEE, 2011. 2564–2571.
- 11 王宁邦, 徐博, 夏百川, 等. 一种颜色直方图计算相似度的资产图片核对算法. 智能计算机与应用, 2018, 8(2): 59–62, 67. [doi: 10.3969/j.issn.2095-2163.2018.02.013]
 - 12 薛向阳, 罗航哉, 朱兴全, 等. 一种新的颜色相似度定义及其计算方法. 计算机学报, 1999, 22(9): 918–922. [doi: 10.3321/j.issn:0254-4164.1999.09.004]
 - 13 芦俊, 丛卫华. 一种基于直方图相似度的快速图像匹配算法. 浙江省信号处理学会 2013 学术年会论文集——信号处理在海洋. 杭州: 浙江省信号处理学会, 2013. 222–225.
 - 14 邹承明, 薛栋, 郭双双, 等. 一种改进的图像相似度算法. 计算机科学, 2016, 43(6): 72–76. [doi: 10.11896/j.issn.1002-137X.2016.06.015]
 - 15 陈会鸽, 曹毅. 基于直方图相似度度量的中央空调风管附尘度检测. 自动化与信息工程, 2009, 30(3): 22–24, 34. [doi: 10.3969/j.issn.1674-2605.2009.03.005]
 - 16 丁建平, 余学才, 黄凯, 等. 基于红外图像灰度直方图相似性的铁路跟踪. 激光与红外, 2012, 42(12): 1385–1388. [doi: 10.3969/j.issn.1001-5078.2012.12.015]
 - 17 李阳齐, 常佳宁, 潘琳, 等. 余弦相似度算法在音乐识别中的研究. 中国科技信息, 2022(12): 110–111. [doi: 10.3969/j.issn.1001-8972.2022.12.034]
 - 18 刘冰, 李文书. 基于余弦相似度的指纹匹配算法的室内定位方法. 科技通报, 2017, 33(3): 198–202. [doi: 10.13774/j.cnki.kjtb.2017.03.041]
 - 19 魏玮, 张芯月, 朱叶. 改进的 SIFT 结合余弦相似度的人脸匹配算法. 计算机工程与应用, 2020, 56(6): 207–212. [doi: 10.3778/j.issn.1002-8331.1812-0346]
 - 20 Charikar MS. Similarity estimation techniques from rounding algorithms. Proceedings of the 34th Annual ACM Symposium on Theory of Computing. Montreal: ACM, 2002. 380–388.
 - 21 徐贤局, 顾敏明, 潘海鹏. 基于均值 Hash 的图案织物疵点检测. 光电子·激光, 2018, 29(12): 1305–1311. [doi: 10.16136/j.joel.2018.12.0130]
 - 22 赵杰, 毕秀丽. 基于局部感知 Hash 的图像篡改取证算法. 计算机与数字工程, 2019, 47(11): 2880–2883. [doi: 10.3969/j.issn.1672-9722.2019.11.047]
 - 23 郭玮. 基于感知哈希的遥感图像检索方法研究[硕士学位论文]. 武汉: 华中科技大学, 2015.
 - 24 吴子岳, 杨帅, 王董测, 等. 自适应图像增强的差值哈希算法对运动目标的跟踪定位研究. 全球定位系统, 2018, 43(5): 98–104. [doi: 10.13442/j.gnss.1008-9268.2018.05.018]
 - 25 Sobel I, Feldman G. An isotropic 3×3 image gradient operator. In: Freeman H, ed. Machine Vision for Three-dimensional Scenes. Sea Harbor Drive: Academic Press, 1990, doi: 10.13140/RG.2.1.1912.4965.
 - 26 Marr D, Hildreth E. Theory of edge detection. Proceedings of the Royal Society B: Biological Sciences, 1980, 207(1167): 187–217.
 - 27 于明, 李雪薇. 基于 SIFT 算法的电动车充电孔双目视觉识别技术. 化工自动化及仪表, 2023, 50(3): 316–322. [doi: 10.20030/j.cnki.1000-3932.202303008]
 - 28 韩宇, 宗群, 邢娜. 基于改进 SIFT 的无人机航拍图像快速匹配. 南开大学学报(自然科学版), 2019, 52(1): 5–9.
 - 29 任晶秋, 武小凯. 改进 SIFT 的图像匹配算法. 吉林大学学报(信息科学版), 2022, 40(4): 672–676. [doi: 10.19292/j.cnki.jdxxp.2022.04.014]
 - 30 杨金玲, 曹先革, 王霞迎, 等. 基于改进 ORB 算法的近红外和可见光图像特征匹配方法. 测绘工程, 2022, 31(6): 7–12. [doi: 10.19349/j.cnki.issn1006-7949.2022.06.002]
 - 31 姚三坤, 刘明. ORB 特征点提取和匹配的研究. 电子设计工程, 2023, 31(2): 43–47. [doi: 10.14022/j.issn1674-6236.2023.02.010]
 - 32 Hadsell R, Chopra S, LeCun Y. Dimensionality reduction by learning an invariant mapping. Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2006. 1735–1742.
 - 33 火善栋. 用孪生网络进行人脸识别与效果分析. 现代计算机, 2022, 28(20): 71–74, 83. [doi: 10.3969/j.issn.1007-1423.2022.20.015]
 - 34 吴正文. 基于改进 UNet 孪生网络的高清遥感建筑物变化检测. 信息技术与信息化, 2022(10): 176–179.
 - 35 Rosten E, Drummond T. Machine learning for high-speed corner detection. Proceedings of the 9th European Conference on Computer Vision. Graz: Springer, 2006. 430–443.
 - 36 Calonder M, Lepetit V, Strecha C, et al. BRIEF: Binary robust independent elementary features. Proceedings of the 11th European Conference on Computer Vision. Heraklion: Springer, 2010. 778–792.
 - 37 Harris CG, Stephens MJ. A combined corner and edge detector. Proceedings of the 1988 Alvey Vision Conference. Manchester: Alvey Vision Club, 1988. 23.1–23.6.
 - 38 Rosin PL. Measuring corner properties. Computer Vision and Image Understanding, 1999, 73(2): 291–307. [doi: 10.1006/cviu.1998.0719]

(校对责编: 孙君艳)