

# 基于高斯过程分类的小样本图像识别

林 宽<sup>1</sup>, 李卓雅<sup>2</sup>, 李新民<sup>1\*</sup>

<sup>1</sup>青岛大学数学与统计学院, 山东 青岛

<sup>2</sup>青岛海尔智能家电科技有限公司, 山东 青岛

收稿日期: 2022年9月8日; 录用日期: 2022年10月19日; 发布日期: 2022年10月28日

## 摘 要

当今图片分类课题应用广泛, 各种神经网络模型层出不穷。计算机性能和图片样本量对模型的训练和预测结果均有影响。针对于图片分类可能遇到的小样本问题, 本文提出使用高斯过程分类模型和HSV图片分解的组合方式, 与传统机器学习模型和RGB分解方式进行比较, 发现该组合方式的准确率最高, 效果最好。

## 关键词

图片分类, 小样本, 高斯过程分类, HSV

# Image Recognition of Small Sample Based on Gaussian Process Classification

Kuan Lin<sup>1</sup>, Zhuoya Li<sup>2</sup>, Xinmin Li<sup>1\*</sup>

<sup>1</sup>School of Mathematics and Statistics, Qingdao University, Qingdao Shandong

<sup>2</sup>Qingdao Haier Intelligent Home Appliance Technology Co., Ltd., Qingdao Shandong

Received: Sep. 8<sup>th</sup>, 2022; accepted: Oct. 19<sup>th</sup>, 2022; published: Oct. 28<sup>th</sup>, 2022

## Abstract

Nowadays, image classification is widely used, and various neural network models emerge in endlessly. The performance of computer and the sample size of pictures have an influence on the training and prediction results of the model. Aiming at the problem of small samples that may be encountered in image classification, this paper proposes a combination method of the Gaussian process classification model and HSV image decomposition. Compared with the traditional ma-

\*通讯作者。

chine learning model and RGB decomposition method, it is found that this combination method has the highest accuracy and the best effect.

## Keywords

Image Classification, Small Sample, Gaussian Process Classification, HSV

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

深度学习因提取特征能力强, 识别精度高, 实时性快等优点被广泛应用[1]。科研工作者开始将深度学习应用到图像识别中, 并且在最近的几年中取得了巨大的进步, 图像的识别效率及识别效果比传统的识别手段有了巨大的进步。最近几年在人工神经网络的大发展背景下, 基于卷积神经网络[2] (Convolutional Neural Network, CNN)的图像分类技术在智能数据采集和高效处理中也取得了较好的效果[3]。以 CNN 模型为基础, 将多层卷积和多层池化结合产生新的网络模型, 可提高网络结构的准确度[4]。经典的卷积神经网络模型有 LeNet-5 [5]、Alex Net [6]、Google Net [7]、VGGNet [8]、ResNet [9]、MobileNet [10]以及 DenseNet 等。很多网络的出现是对已有网络的排列组合, 加深了层次或者加多了节点数目。

近年来现有文献中基于小样本学习的图像分类算法都是采用深度学习[11]。将深度学习中的技术用于小样本学习中, 比如使用数据增强技术来增加样本的数量。迁移学习、元学习、对偶学习、贝叶斯学习以及图神经网络方法也被用于小样本图像分类的任务。然而, 深度神经网络的训练质量受训练集中有标记样本数目的强烈影响[12]。在实际的问题中, 数据获取往往是困难的, 这之中既有个人隐私的问题, 比如人脸数据, 也有问题对象本身就很少的问题, 比如识别珍稀保护动物的问题[13]。其次是算力问题, 深度学习方法在提高算法性能的同时, 往往伴随着庞大的网络运算, 这也就使得深度学习的方法很难部署在计算资源受限的设备上。那么此时如何用小样本的图像数据进行分类、同时节省算力, 达到很好的分类效果便成了研究问题的重点。

机器学习中的机器视觉利用机器模拟人脑, 通过设备获取图像, 然后经过一定的处理和特征提取, 最后识别出相关信息[14]。机器学习的核心是分类器的训练[15]。通过不断的训练和调整, 最终得到一个适用的分类器模型, 实现对事物的准确识别[16]。

高斯过程观测变量空间是连续域, 时间或空间。高斯过程观测变量空间是实数域的时候, 我们就可以进行回归而实现预测, 即高斯过程回归(Gaussian Process Regression, GPR) [17]。高斯过程观测变量空间是整数域的时候(观测点是离散的), 我们就可以进行分类, 即高斯过程分类(Gaussian Process Classification, GPC)。由于模型特性, 高斯过程在小样本数据集上效果很好。于是我们便可以将高斯过程分类这一方法, 应用于小样本的图像分类中, 避免了深度学习模型的庞大架构和数据需求, 也能达到不俗的分类效果。

本论文从图片分解方式出发, 结合分类算法研究了小样本图像识别工作。主要内容包括: 本文的内容结构安排如下, 首先在第 2 节介绍高斯过程回归和高斯过程分类模型, 然后在第 3 节介绍图片的分解方式和原理, 在第 4 节给出了不同情况下小样本的图片分类的实验结论。最后在第 5 节给出结论。

## 2. 高斯过程分类模型

高斯过程的应用包括回归、分类两种问题。输入与输出变量均为连续变量的预测被称为回归问题，而输出变量为有限个离散变量的预测被称为分类问题[18]。高斯过程分类其实就是在高斯过程回归的基础上，将高斯过程回归的输出结果通过一个连接函数获得相应的标签概率结果，进而依据一个合适的划分阈值便可输出不同的类别结果[19]。

假设有数据集  $D$  包含  $n$  个观测值  $D = \{(x_i, y_i) | i = 1, \dots, n\}$ ，其中  $x$  表示  $D$  维的输入向量(协变量)， $y$  表示输出向量(因变量)，所有  $n$  种情况的列向量输入可以集合在一个  $D \times n$  的设计矩阵  $X$  中。我们希望通过训练集的数据构建模型  $y = f(x)$ ，其中  $f(\cdot)$  是协变量与因变量的映射关系，即潜变量函数。而且可以将这个模型应用于新的数据集  $\{x^*, y^*\}$  中，即测试集数据。在已知  $x^*$  时用于预测得到输出的  $y^*$ 。对于更加真实的建模情况，我们通常无法只获得函数  $f(x)$  本身，而是有噪声的情况

$$y = f(x) + \varepsilon, \varepsilon \sim N(0, \delta_n^2 \mathbf{I}_n).$$

假设  $f(x)$  服从高斯过程  $f(x) \sim GP(m(x), K(x, x'))$ ，可以得到

$$y \sim GP(m(x), K(x, x') + \delta_n^2 \mathbf{I}_n),$$

其中  $m(x)$  为均值函数， $K(x, x')$  为核函数。由高斯过程性质可以得到  $y^*$  和  $y$  的联合分布

$$\begin{pmatrix} y \\ y^* \end{pmatrix} \sim GP \left( m(X, X^*), \begin{bmatrix} K(X, X) + \delta_n^2 \mathbf{I}_n & K(X, X^*) \\ K(X^*, X) & K(X^*, X^*) \end{bmatrix} \right).$$

从而可以导出条件分布，也就是高斯过程回归的关键预测方程[17]

$$y^* | y \sim GP \left( K(X^*, X) [K(X, X) + \delta_n^2 \mathbf{I}_n]^{-1} y, \right. \\ \left. K(X^*, X^*) - K(X^*, X) [K(X, X) + \delta_n^2 \mathbf{I}_n]^{-1} K(X, X^*) \right).$$

当因变量不是连续值，而是标签时，即  $y \in \{-1, 1\}$ ，便出现了分类问题。在二元分类问题中，用高斯过程进行预测背后的思想是很简单的，即在函数  $f(x)$  上放置一个高斯过程先验，然后通过某个连接函数  $\sigma(\cdot)$  将结果挤压在  $[0, 1]$  区间内：

$$p(y = 1 | x) = \sigma(f(x)).$$

接下来进行贝叶斯推断，首先计算应用于测试集数据的潜变量函数的分布

$$p(f^* | X, y, X^*) = \int p(f^* | X, X^*, f) p(f | X, y) df,$$

其中  $p(f | X, y) = \frac{p(y | f) p(f | X)}{p(y | X)}$  是潜变量函数的后验分布，利用这个分布可以产生一个基于数据集的概率预测

$$p(y^* = 1 | X, y, X^*) = \int \sigma(f^*) p(f^* | X, y, X^*) df^*.$$

在回归的情况下，预测的计算是直接可以进行解析计算的，因为相关积分都是高斯分布的。但在分类问题中，非高斯似然使得结果难以计算，后验分布和概率的预测都是无法进行的，所以需要解析逼近，例如拉普拉斯近似算法和期望传播算法。

本文实验中的高斯过程分类，是采用拉普拉斯近似算法。拉普拉斯算法在积分中利用高斯函数  $q(f | X, y)$  近似后验  $p(f | X, y)$ 。对  $\log p(f | X, y)$  最大后验进行二阶泰勒展开，可以得到高斯近似[17]

$$q(f|X, y) = N(f|\hat{f}, A^{-1}) \propto \exp\left(-\frac{1}{2}(f - \hat{f})^T A(f - \hat{f})\right),$$

其中  $\hat{f} = \arg \max_f p(f|X, y)$ ,  $A = -\nabla\nabla \log p(f|X, y)|_{f=\hat{f}}$  是该点后验负对数的 Hessian 矩阵。

根据贝叶斯定理,  $p(f|X, y) = \frac{p(y|f)p(f|X)}{p(y|X)}$  是潜变量函数的后验分布, 其中  $p(y|X)$  与  $f$  无关, 那么只对分子研究可以得到边际似然函数:

$$\begin{aligned} \Psi(f) &= \log p(y|f) + \log p(f|X) \\ &= \log p(y|f) - \frac{1}{2}f^T K^{-1}f - \frac{1}{2}\log|K| - \frac{n}{2}\log 2\pi. \end{aligned}$$

对上式求微分可以得到:

$$\begin{aligned} \nabla\Psi(f) &= \nabla \log p(y|f) - K^{-1}f, \\ \nabla\nabla\Psi(f) &= \nabla\nabla \log p(y|f) - K^{-1} = -W - K^{-1}, \end{aligned}$$

其中  $W = -\nabla\nabla \log p(y|f)$  是对角阵。因为似然函数中的因子( $y_i$  分布只依赖于  $f_i$ , 而不是  $f_{j \neq i}$ )。如果似然函数  $p(y|f)$  是对数凹函数,  $W$  的对角元素是非负的, Hessian 矩阵是负定的, 因此  $\Psi(f)$  是凹的且具有唯一最大值。

在找到  $\hat{f}$  的最大后验之后, 可以指定后验的后验的拉普拉斯近似为一个高斯函数, 其均值和协方差矩阵由  $\Psi$  的负逆 Hessian 给出:

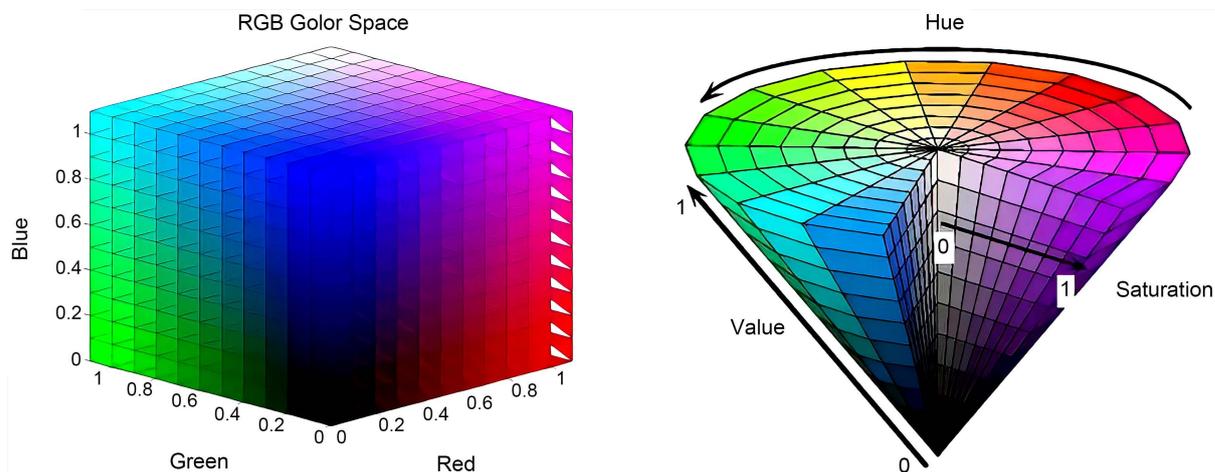
$$q(f|X, y) = N(\hat{f}, (K^{-1} + W)^{-1}).$$

### 3. 图片分解方式

在计算机图像处理中, 常采用 RGB 模型和 HSV 模型[20]。RGB 模型是基于人类视觉三原色——红(Red)、绿(Green)、蓝(Blue)理论建立的色彩空间, 即认为用红、绿、蓝 3 种色光作适当混合, 可以引起电磁光谱上所有任何颜色的感知。RGB 空间的颜色是一种加色, 对于三种基色的光, 当它们按不同比例混合时, 会形成不同颜色。三个成分的范围都是 0 到 255, 当它们以不同的比例混合时, 就可以组合出 2563 种颜色[21]。RGB 颜色模型是图像处理中最常用的颜色模型, 现有的图像采集设备最初采集到的颜色信息是 RGB 值, 图像处理中使用的其他颜色空间也是从 RGB 颜色空间转换来的[22]。但 RGB 颜色空间不直观, 从 RGB 值中很难判断该值所表示的颜色, 因此, RGB 颜色空间不符合人对颜色的感知心理[23]。

由于 RGB 模型中这三色分量高度相关, 且是一种不均匀的颜色空间, 即两种颜色之间的知觉差异(色差)不能表示为该颜色空间中两点间的距离, 所以, 它主要是作为一种面向硬件设备(如彩色监视器、彩色摄像机等)的色彩空间模型。

HSV (Hue-Saturation-Value)颜色空间是一种面向视觉感知的颜色模型, 人眼的色彩知觉主要包括 3 个要素: 色调、饱和度和亮度。HSV 模型是基于人的视觉感知特性建立的色彩空间, 其中色度(Hue)表示不同的颜色, 如红绿蓝; 饱和度(Saturation)表示颜色的深浅, 如深绿浅绿; 亮度(Value)表示颜色的明暗程度, 如很亮(亮白)很暗(黑暗)。它具有两个重要的特点: 其一, 亮度分量(V)与图像的彩色信息无关; 其二, 色度分量(H)、饱和度分量(S)与人感受颜色的方式是紧密相连的[24]。三个分量的值都在 0~1 之间。HSV 颜色空间反映了人观察彩色的方式, 同时也有利于图像处理。RGB 模型和 HSV 模型对颜色空间的分解如图 1 所示。



**Figure 1.** RGB (left) decomposition model and HSV (right) decomposition model  
**图 1.** RGB (左)分解模型和 HSV (右)分解模型

## 4. 小样本下的图片分类

### 4.1. 二分类

本文图片来源为 kaggle 网站上的一个数据集,其中包含了 10 种动物的图片,抽取其中猫狗的部分图片构建二分类问题。实验模拟用 Python 进行,用到的传统机器学习模型和高斯过程分类模型均可从 sklearn 库中调用。对于图片读入后的分解方式,应用 opencv 编写自定义函数,实现图片应用 HSV 分解和 RGB 分解后得到不同的特征数据。

每张图片本质上是  $256 \times 256$  的矩阵,经过 HSV 和 RGB 分解函数后,每张图片会得到一个  $1 \times 768$  的向量,即按照 H、S、V 方向和 R、G、B 三个方向分解,每个方向会得到 256 个特征向量。再整合  $n$  张图片进行训练,训练数据便是  $n \times 768$  的矩阵。由于特征变量个数太多,会成为高维问题使模型无法训练或者效果不好,加入主成分分析对数据进行降维,保留 85% 的主成分,既保证了不损失太多信息,又达到了降维和减少变量的目的。采用传统的机器学习分类模型(随机森林、支持向量机、单隐层 BP 神经网络)与高斯过程分类模型一起训练和测试。最后对每张图片所属类别加上标签,将训练数据和对应标签按照 8:2 分割成训练集和测试集,保证所有模型训练集数据相同情况下,分别对 HSV 和 RGB 分解方式下计算各个模型在测试集数据的准确率,用于区分不同模型的优劣。

首先确定每个类别样本量是均衡的,然后在猫狗各 100 张图片、总样本量  $N = 200$  的基础上,不断缩减样本量为  $N = 160$  和  $N = 120$ 。在  $N = 200$  时,按照 8:2 分割数据集测试集图片是 40 张,所以保持样本量缩减时测试集图片仍为 40 张,这样避免因测试集图片数量不同对准确率的影响。实验结果见表 1 和表 2。

**Table 1.** Accuracy of each model under HSV decomposition

**表 1.** HSV 分解方式下各模型准确率

	随机森林	支持向量机	高斯过程分类	神经网络
$N = 200$	0.700	0.700	0.820	0.744
$N = 160$	0.750	0.800	0.820	0.738
$N = 120$	0.650	0.700	0.750	0.650

**Table 2.** Accuracy of each model under RGB decomposition**表 2.** RGB 分解方式下各模型准确率

	随机森林	支持向量机	高斯过程分类	神经网络
$N = 200$	0.600	0.730	0.780	0.669
$N = 160$	0.680	0.680	0.820	0.638
$N = 120$	0.780	0.700	0.800	0.633

从表 1 和表 2 的结果可以发现：同种分解方式下，随机森林、支持向量机和神经网络模型的效果受样本量的影响很大，其中神经网络模型明显准确率随着样本量增大而提高。高斯过程分类模型的效果一直保持很高的准确率，优于其他模型。而且在同一样本量下，使用 HSV 分解方式的准确率大部分情况下高于 RGB 分解方式。

## 4.2. 多分类

在二分类问题的基础上，引入羊的图片，与猫和狗的图片一起构成多分类问题。研究模式和训练模型仍与二分类时相同。

依然确定每个类别样本量是均衡的，然后以猫、狗、羊各 100 张图片，总样本量  $N = 300$  的基础上，不断缩减样本量  $N = 240$  和  $N = 180$ 。在  $N = 300$  时，按照 8:2 分割数据集测试集图片是 60 张，保持样本量缩减时测试集图片仍为 60 张，避免因测试集图片数量不同对准确率的影响。实验结果见表 3 和表 4。

**Table 3.** Accuracy of each model under HSV decomposition**表 3.** HSV 分解方式下各模型准确率

	随机森林	支持向量机	高斯过程分类	神经网络
$N = 300$	0.550	0.550	0.600	0.625
$N = 240$	0.530	0.600	0.620	0.617
$N = 180$	0.550	0.550	0.600	0.608

**Table 4.** Accuracy of each model under RGB decomposition**表 4.** RGB 分解方式下各模型准确率

	随机森林	支持向量机	高斯过程分类	神经网络
$N = 300$	0.420	0.500	0.450	0.567
$N = 240$	0.380	0.400	0.420	0.528
$N = 180$	0.420	0.420	0.430	0.517

从表 3 和表 4 的结果可以发现：随着第三类图片的加入，分类问题变得复杂，图片样本也多种多样，各模型准确率相比于二分类均有所下降。HSV 分解方式下高斯过程分类模型和神经网络模型效果大致相同，且优于其他模型。RGB 分解方式下，各模型均受分解方式的很大影响，准确率明显下降，神经网络模型的准确率仍高于其他模型。固定样本量时，相同模型使用 HSV 分解方式的准确率高于 RGB 分解方式。

## 5. 结论

根据本文的实验结果，当我们再遇到小样本图片分类问题时，可以采用高斯过程分类模型和 HSV 图

片分解方式这样的组合，以便达到较高的模型准确率。我们可以将这样的模式应用于实际生活场景中，例如医学的临床实验、病理推断、稀有动植物、文物鉴定等。根据已有的图片样本进行模型训练和调参，训练出效果较好的模型，再用于鉴别新样本，进行分类。

## 参考文献

- [1] 郑远攀, 李广阳, 李晔. 深度学习在图像识别中的应用研究综述[J]. 计算机工程与应用, 2019, 55(12): 20-36.
- [2] He, K., Zhang, X., Ren, S. and Sun, J. (2015) Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **37**, 1904-1916. <https://doi.org/10.1109/TPAMI.2015.2389824>
- [3] 季长清, 高志勇, 秦静, 汪祖民. 基于卷积神经网络的图像分类算法综述[J]. 计算机应用, 2022, 42(4): 1044-1049.
- [4] 盖荣丽, 蔡建荣, 王诗宇, 仓艳, 陈娜. 卷积神经网络在图像识别中的应用研究综述[J]. 小型微型计算机系统, 2021, 42(9): 1980-1984.
- [5] LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P. (1998) Gradient-Based Learn in Applied to Document Recognition. *Proceedings of the IEEE*, **86**, 2278-2324. <https://doi.org/10.1109/5.726791>
- [6] Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2012) Imagenet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems 25: 26th Annual Conference on Neural Information Processing Systems 2012, Lake Tahoe*, 3-6 December 2012, 1097-1105.
- [7] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Anguelov, D., Erhan, D., Vanhoucke, V. and Rabinovich, A. (2015) Going Deeper with Convolutions. *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition*, Boston, 7-12 June 2015, 1-9. <https://doi.org/10.1109/CVPR.2015.7298594>
- [8] Simonyan, K. and Zisserman, A. (2021) Very Deep Convolutional Networks for Large-Scale Image Recognition. *3rd International Conference on Learning Representations, ICLR 2015, San Diego*, 7-9 May 2015, 1-14. <https://arxiv.org/pdf/1409.1556.pdf>
- [9] He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, 27-30 June 2016, 770-778. <https://doi.org/10.1109/CVPR.2016.90>
- [10] Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T. and Andreetto, M. and Adam, H. (2021) Mobilenets: Efficient Convolutional Neural Networks for Mobile Vision Applications. <https://arxiv.org/pdf/1704.04861.pdf>
- [11] 刘颖, 雷研博, 范九伦, 王富平, 公衍超, 田奇. 基于小样本学习的图像分类技术综述[J]. 自动化学报, 2021, 47(2): 297-315.
- [12] 吕昊远, 俞璐, 周星宇, 邓祥. 半监督深度学习图像分类方法研究综述[J]. 计算机科学与探索, 2021, 15(6): 1038-1048.
- [13] 葛轶洲, 刘恒, 王言, 徐百乐, 周青, 申富饶. 小样本困境下的深度学习图像识别综述[J]. 软件学报, 2022, 33(1): 193-210. <https://doi.org/10.13328/j.cnki.jos.006342>
- [14] 王海霞, 李凯勇. 基于机器学习的图像敏感信息识别方法[J]. 现代电子技术, 2021, 44(19): 66-70.
- [15] 安强强, 张峰, 李赵兴, 张雅琼. 基于机器学习的图像分割研究[J]. 自动化与仪器仪表, 2018(6): 29-31.
- [16] 田朝辉, 金鑫, 赵耿, 李晓东. 基于深度学习的物体识别验证码破解方法[J]. 计算机仿真, 2018, 35(3): 216-220.
- [17] Williams, C.K. and Rasmussen, C.E. (2006) *Gaussian Processes for Machine Learning*. MIT Press, Cambridge. <https://doi.org/10.7551/mitpress/3206.001.0001>
- [18] 叶婧. 基于高斯过程回归的锂电池数据处理[D]: [硕士学位论文]. 北京: 北京交通大学, 2016.
- [19] 何亚东. 云南大理州林火预测的高斯过程分类模型研究[D]: [硕士学位论文]. 昆明: 云南大学, 2020.
- [20] 石美红, 申亮, 龙世忠, 胡西民. 从 RGB 到 HSV 色彩空间转换公式的修正[J]. 纺织高校基础科学学报, 2008(3): 351-356.
- [21] 刘国宏, 于竞翔, 任丽君. RGB 检测方法研究进展[J]. 分析科学学报, 2020, 36(4): 591-596.
- [22] 姜兰池, 沈国强, 张国焯. 基于 HSV 分块颜色直方图的图像检索算法[J]. 机电工程, 2009, 26(11): 54-57.
- [23] 魏宝刚, 李向阳, 鲁东明, 潘云鹤. 彩色图像分割研究进展[J]. 计算机科学, 1999(4): 59-62.
- [24] 张国权, 李战明, 李向伟, 魏伟一. HSV 空间中彩色图像分割研究[J]. 计算机工程与应用, 2010, 46(26): 179-181.