



多感知兴趣区域特征融合的图像识别方法

闫涵, 张旭秀, 张净丹

引用本文:

闫涵, 张旭秀, 张净丹. 多感知兴趣区域特征融合的图像识别方法[J]. *智能系统学报*, 2021, 16(2): 263–270.

YAN Han, ZHANG Xuxiu, ZHANG Jingdan. Image recognition method based on multi-perceptual interest region feature fusion[J]. *CAAI Transactions on Intelligent Systems*, 2021, 16(2): 263–270.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.201906032>

您可能感兴趣的其他文章

基于深度学习的空间非合作目标特征检测与识别

Feature detection and recognition of spatial noncooperative objects based on deep learning
智能系统学报. 2020, 15(6): 1154–1162 <https://dx.doi.org/10.11992/tis.202006011>

基于增强AlexNet的音乐流派识别研究

Music genre recognition research based on enhanced AlexNet
智能系统学报. 2020, 15(4): 750–757 <https://dx.doi.org/10.11992/tis.201909032>

基于小样本学习的LCD产品缺陷自动检测方法

An automatic small sample learning-based detection method for LCD product defects
智能系统学报. 2020, 15(3): 560–567 <https://dx.doi.org/10.11992/tis.201904020>

深度学习的双人交互行为识别与预测算法研究

Human interaction recognition and prediction algorithm based on deep learning
智能系统学报. 2020, 15(3): 484–490 <https://dx.doi.org/10.11992/tis.201812029>

基于生成对抗网络的机载遥感图像超分辨率重建

Super-resolution reconstruction of airborne remote sensing images based on the generative adversarial networks
智能系统学报. 2020, 15(1): 74–83 <https://dx.doi.org/10.11992/tis.202002002>

REM记忆模型在图像分类识别中的应用

Application of REM memory model in image recognition and classification
智能系统学报. 2017, 12(3): 310–317 <https://dx.doi.org/10.11992/tis.201605010>

微信公众平台



关注微信公众号, 获取更多资讯信息

DOI: 10.11992/tis.201906032

网络出版地址: <https://kns.cnki.net/kcms/detail/23.1538.TP.20200630.1143.008.html>

多感知兴趣区域特征融合的图像识别方法

闫涵, 张旭秀, 张净丹

(大连交通大学 电气信息工程学院, 辽宁 大连 116028)

摘要: 针对自然图像识别过程中不同深度学习模型关注兴趣区域不同的现象, 本文引入深度卷积神经网络融合机制, 结合深度迁移学习方法, 给出了一种基于多感知兴趣区域特征融合的图像识别方法。本文将迁移学习方法引入牛津大学视觉组网络模型 (visual geometry group network, VGGNet) 和残差网络模型 (residual network, ResNet), 通过对单个分类模型进行热力图可视化及特征可视化, 得到了不同网络模型关联的特征区域不一样的结论。然后在此基础上分别设计特征拼接、特征融合加特征拼接及融合投票方法将不同模型特征进行融合, 得到 3 种新的融合模型。实验结果表明, 本文方法在 Kaggle 数据集上的识别准确率高于 VGG-16、VGG-19、ResNet-50、DenseNet-201 模型。

关键词: 深度学习; 图像识别; 迁移学习; 特征融合; 集成学习; 特征提取; CAM 可视化; 视觉组网络模型; 残差网络模型

中图分类号: TP311 文献标志码: A 文章编号: 1673-4785(2021)02-0263-08

中文引用格式: 闫涵, 张旭秀, 张净丹. 多感知兴趣区域特征融合的图像识别方法 [J]. 智能系统学报, 2021, 16(2): 263-270.

英文引用格式: YAN Han, ZHANG Xuxiu, ZHANG Jingdan. Image recognition method based on multi-perceptual interest region feature fusion[J]. CAAI transactions on intelligent systems, 2021, 16(2): 263-270.

Image recognition method based on multi-perceptual interest region feature fusion

YAN Han, ZHANG Xuxiu, ZHANG Jingdan

(School of Electrical Information Engineering, Dalian Jiaotong University, Dalian 116028, China)

Abstract: This paper presents the deep convolution neural network fusion mechanism and proposes an image recognition method based on multi-perceptual interest region feature fusion in combination with the deep-migration learning method. This is to solve the problem of different deep-learning models used on different interest regions when they recognize a natural image. The migration learning method is applied to the convolution neural net architectures, namely VGG and ResNet networks. Then, through the visualization of the heat map and the features of single classification model, a conclusion is drawn that the characteristic regions associated with different network models are different. Based on this, the methods of feature splicing, feature fusion and splicing, and fusion voting systems are designed to fuse different model features, obtaining three new fusion models. The experimental results show that the recognition accuracy of this method on Kaggle dataset is higher than that of VGG-16, VGG-19, ResNet-50, and DenseNet-201 models.

Keywords: deep learning; image recognition; migration learning; feature fusion; integrated learning; feature extraction; CAM visualization; VGGNet; ResNet

图像识别是计算机视觉领域的一个分支^[1-2], 图像识别是指利用计算机对图像进行特征提取,

收稿日期: 2019-06-18. 网络出版日期: 2020-06-30.

基金项目: 国家自然科学基金项目 (61471080/F010408); 国家支撑计划 (2015BAF20B02); 国家留学基金委资助计划 (201608210308); 辽宁省自然科学基金指导计划 (2019-ZD-0108).

通信作者: 闫涵. E-mail: 1346917459@qq.com.

从而识别各种不同模式下的目标和对象。图像识别在生活中的应用非常广泛, 如人脸识别、指纹识别、安防监控等各个领域^[3]。

图像识别方法可分为传统的图像识别方法与基于深度学习的图像识别方法。传统的图像识别手段主要建立在特征提取的基础上^[4], 首先提取

图像的底层像素信息或图像的特征,如图像边缘特征、纹理特征以及颜色特征等^[5-9]。也有学者采用小波变换、傅里叶变换、EMD(empirical mode decomp-osition)变换等方法进行特征提取^[10-12]。然后将提取的特征送入机器学习模型中进行分类识别。传统的图像识别方法都需要一定的先验性知识,而且对于一些相似的图像依靠人为寻找特征仍较为困难,算法实现的精度较低且难以转换。近年来,随着计算机性能的提升以及数据量的爆炸式增长,深度学习(deep learning)作为一种新兴的机器学习手段取得了飞速的发展^[13]。采用深度学习方法进行图像识别的优势在于图像识别精度高,无需特征工程,适应性强易于转换,可使用数据进行有效缩放,方便实际应用与部署^[14-15]。现有的深度学习网络浅层主要通过输入图像进行特征提取与降采样获得图像低阶特征,深层网络则通过更深层变换获得图像特征的语义信息与更高分辨率特征图^[16]。近年来,针对自然图像识别任务,一系列深度学习算法被提出,AlexNet是最早的深度学习网络模型^[17]。AlexNet有八层网络结构,采用新的relu激活函数和dropout机制抵抗过拟合以提高特征提取的有效性。VGGNet采用多个尺寸更小的卷积核模拟大尺寸卷积核^[18],在保证感受野的同时减少参数数量,同时引入 1×1 卷积核增加网络的表征能力、降低计算量。ResNet的提出主要通过跨层连接和拟合残差项的方法解决了随着网络层数的加深^[19],网络模型难以训练的问题。DenseNet是一种密集连接卷积神经网络,该模型设计了dense block模块,使得网络结构更窄,参数更少^[20]。同时该连接方式也使得特征和梯度的传递更加有效,网络更易训练。上述深度学习网络模型在图像识别任务中均有较好表现,但现有深度学习方法的图像识别任务仍有如下难点^[21]:1) AlexNet作为最早的深度学习模型虽然取得较好的图像识别效果,但是由于其层数只有八层,导致其在特征的表征能力上仍有上升空间。2) VGGNet、ResNet以及DenseNet网络模型各自采用创新技术达到模型层数加深的目的,提升了网络的学习能力。但是对于特定任务,从头训练上述模型成本代价过高。3) 随着人工智能技术的发展,深度学习图像识别网络层出不穷,但深度学习网络需要大量的数据进行训练及拟合,现实情况难以有足够的支撑网络进行训练。4) 在研究过程中,通过可视化数据特征发现,不同的网络模型关联和提取到的特征也是不同的,关注的图像特征区域也有所差异。如何在机器学习集成架构下,将不同的网络特征相融合

及集成,进一步提升网络的分类准确率,需要进行研究和讨论。

针对上述问题,本文提出了一种基于多感知兴趣区域特征融合的图像识别方法,对深度卷积模型VGG网络与残差网络引入融合机制与微调机制,结合深度迁移学习方法,将网络模型的识别能力进行增强。本文的主要工作如下:1) 利用特征可视化方法对模型的兴趣区域特征进行热力图标注,获得的结论为不同模型在识别相同图像时所关注兴趣区域有所差异;2) 根据特征可视化结论,引入模型多尺度性质和投票融合机制,设计3种网络融合方法;3) 在Kaggle数据集上进行对比实验以验证本文方法的有效性。实验结果表明本文方法比融合前方法具有更高的识别准确率。

1 相关理论

本文将迁移学习方法引入到VGG网络模型和ResNet网络模型作为算法的模型基础。通过CAM可视化机制对模型的兴趣特征进行可视化分析。最后通过集成学习方法将含有不同兴趣特征的网络模型进行融合得到本文算法。

1.1 迁移学习

在机器学习、深度学习和数据挖掘的大多数任务中都会假设训练和测试时采用的数据服从相同的分布并且来源于相同的特征空间^[17]。然而在深度学习实现的过程中,由于网络模型的深度较大且参数过多,在训练网络模型时往往会遇到以下问题:

1) 带标记的训练样本数量不够。当处理A领域任务时,缺少足够数量训练样本和标签。与A领域具有一定关联的B领域,却拥有数量巨大的训练样本。但由于B领域与A领域具有一定差异,因而导致其特征处于不同的空间分布,或样本服从不同分布。

2) 数据分布会发生变化。对于相同的任务,在不同时期、地点或受其他动态因素影响,数据分布会发生一定变化,因此需要重新收集数据并建立模型。

迁移学习作为一种机器学习的方法,可以把为任务A开发的模型作为初始点,重新应用在任务B的模型开发过程中。即把B领域中的知识迁移到A领域中来,提高A领域分类效果,不需要花大量时间去标注A领域数据。对于深度学习目标识别任务,在预训练模型中存在各种特征数据与权重信息。其中一些是与分类识别对象本身关联比较紧密的特征数据与权重信息,还有一些比较

共性的特征数据与信息,是可以被不同任务或者对象之间共享的。迁移学习是要迁移那些共性特征数据与信息,从而避免再次学习这些知识,实现

快速学习^[18]。迁移学习微调示意图如图 1 所示,为了方便图像进行卷积操作,首先使用函数 $F(0)$ 对图像进行预处理,去掉边缘像素值,缩小图像尺寸。

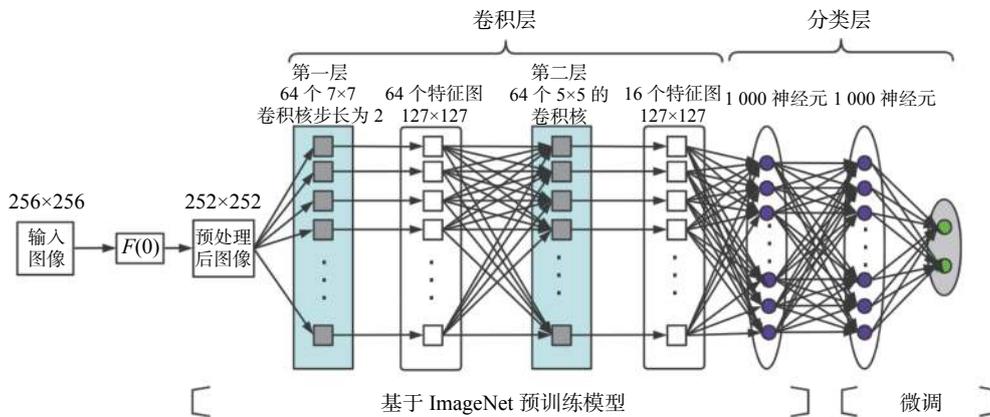


图 1 迁移学习微调示意

Fig. 1 Schematic diagram of migration learning fine-tuning

以上模型分为两部分,一部分是基于 ImageNet 训练的深度学习模型,另一部分为根据自身任务微调的网络模型。在训练时可针对特定任务,固定原始网络的相关层,修改网络的输出层,以使结果更符合任务需要。

1.2 集成学习

集成学习 (ensemble learning) 是使用一系列学习器进行学习,并使用某种规则把各个学习器结果进行整合,从而获得比单个学习器更加优越的泛化性能^[19-20]。集成学习算法策略主要有 Bagging 算法、Boosting 算法以及 Stacking 算法。本文涉及到的算法为 Bagging 算法,算法示意图如图 2 所示。

集成学习优势在于:

1) 使用相同数据对不同的个体学习器进行训练会得到不同结果,进而导致分类器的分类边界不同,还可能影响最终分类结果,产生错误输出。将多个个体学习器合并后,可获得更为合理的分类边界,提升分类准确率。

2) 对于数据集包含样本数量差异较大的情况,可以对样本数量较大的数据集进行划分操作并对样本数量较小的数据集进行放回操作,进而产生不同数据子集,再使用数据子集训练不同学习器,最终合并成为一个强学习器。

3) 当数据划分边界过于复杂时,单一使用线性分类器很难获得较好结果。可以考虑训练多个模型,再进行模型融合。

4) 对于多个难以直接融合的异构特征集,可以考虑使用每个数据集单独构建分类模型,再将多个模型进行融合。

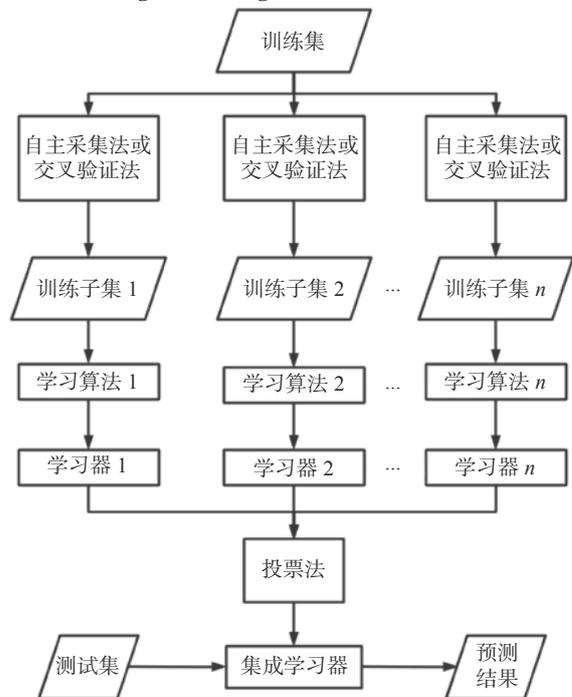


图 2 集成学习 Bagging 算法示意

Fig. 2 Schematic diagram of integrated learning Bagging algorithm

1.3 CAM 可视化

CAM 是深度学习中一种常用的网络可视化方法。利用 GAP(global average pooling) 替换掉全连接层,以把 GAP 视为一个特殊的平均值池化层,其池化核尺寸和整个特征图一样大,即求每张特征图所有像素的均值。GAP 的优点在于由于没有了全连接层,输入不需固定大小,因此可支持任意大小的输入。此外,引入 GAP 更充分地利用了空间信息,且没有了全连接层的各种参数,避免过拟合。同时在最后的卷积层生成了和目标类别数量一致的特征图,经过 GAP 以后再通

过 softmax 层得到结果,这样就给每个特征图赋予了很明确的意义,也就是类别置信度特征图。CAM 模型示意图如图 3 所示。

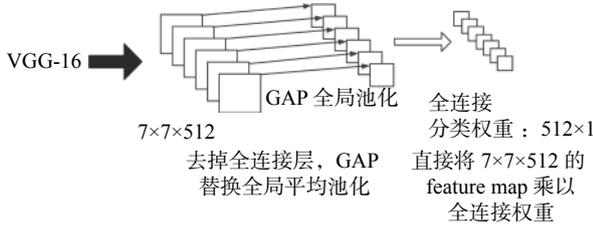


图 3 CAM 模型示意

Fig. 3 Schematic diagram of the CAM model

经过 GAP 得到最后一个卷积层每个特征图的均值,通过加权和得到输出。对每一个类别 C , 每个特征图 k 的均值对应 w 。当解释模型分类依据时,可将该类别对应的所有 w 求取出,并求出它们与自身对应的特征图的加权和。由于输出的大小和特征图是一致的,需要对它进行上采样并与原图进行叠加,得到 CAM 可视化图。CAM 以热力图的形式表明模型对目标进行分类的特征。图 4 是在 kaggle 数据集上 VGG-16 分类 CAM 可视化图。如图 4 所示,对于猫或者狗的分类判断区域全部集中在头部,头部特征的确能够区分猫和狗。

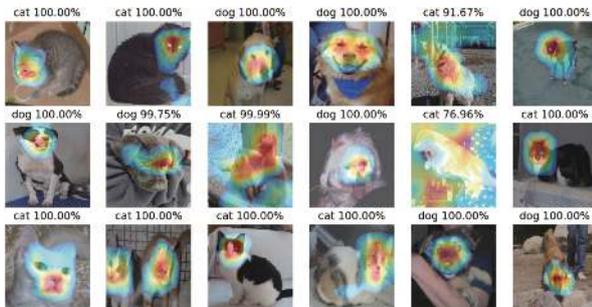


图 4 基于 VGG-16 猫狗大战数据 CAM 模型可视化图
Fig. 4 Visualization of CAM model based on VGG-16 cat and dog data figure

2 本文方法及实验

本文实验所采用的实验数据为自然图像猫狗分类数据集。该数据集主要分为两类:猫、狗。训练集为 20832 张,测试集为 4168 张。为了验证本文算法的有效性,在猫狗数据集上进行对比实验,同样的训练数据集和同样的测试集。实验在 2.2 GHz Intel Core i7 处理器 (4 核), macOS Mojave 10.14.4 系统上进行。编程环境为 python2 版本,编程语言为 python 语言,编译器为 jupyter notebook。仿真软件的深度学习框架为基于 TensorFlow 后端的 keras,显卡为 GTX 1070TI,显存大小为 8 GB。

2.1 VGG-16、VGG-19、ResNet 的 CAM 可视化实验

为了对比 3 种模型在图像识别时所关注的特征点,选取 16 张猫狗单独存在的图像和 4 张猫狗共存的模糊语义图像对 3 种模型进行测试。在实验前首先通过迁移学习与微调机制,固定网络模型的浅层权重参数,设置分类层与深层部分权重参数可优化。以 VGG-16 网络模型为例,其迁移学习微调示意图如图 5 所示。

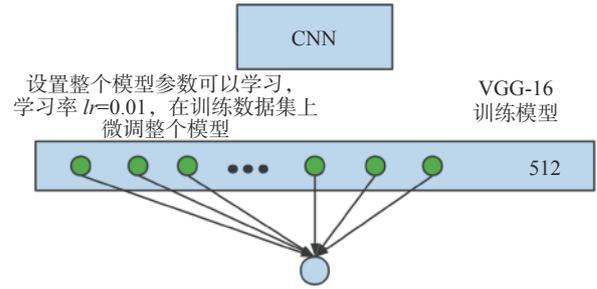


图 5 迁移学习微调结构示意

Fig. 5 Schematic diagram of the migration learning fine-tuning structure

图 6、图 7 为 VGG 网络模型和 ResNet 网络模型识别目标单独存在时的 CAM 可视化图。

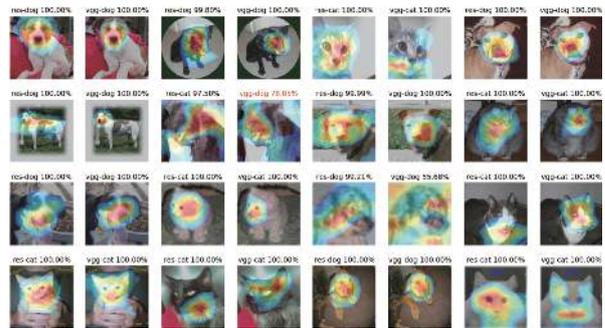


图 6 单目标 CAM 可视化结果
Fig. 6 Single target CAM visualization results



图 7 模糊语义下 CAM 可视化结果
Fig. 7 CAM visualization results under fuzzy semantics

通过模型识别单独目标的 CAM 可视化图可以看出,VGG 网络模型更容易将猫分类成狗,而 Resnet 识别结果与之相反,说明两种不同的网络模型在进行目标识别时关注的兴趣区域不同。通

过模型识别模糊语义的 CAM 可视化图可以看出, VGG 网络模型将 4 张测试图像分类成猫, 而 ResNet 将其分类成狗, 该现象也可验证上述结论。

2.2 本文方法

方案 1 特征拼接 (Multi view Fusion 方案)

不同的模型在同一张图片关注的兴趣区域不一, 可考虑将多个模型特征进行融合, 将会使得融合后的模型具有关注多视角兴趣特征的功能。本方案运用 DenseNet 的思想, 将不同尺度特征相拼接, 引入模型多尺度性质。设计的 Multi view Fusion 模型结构如图 8 所示。

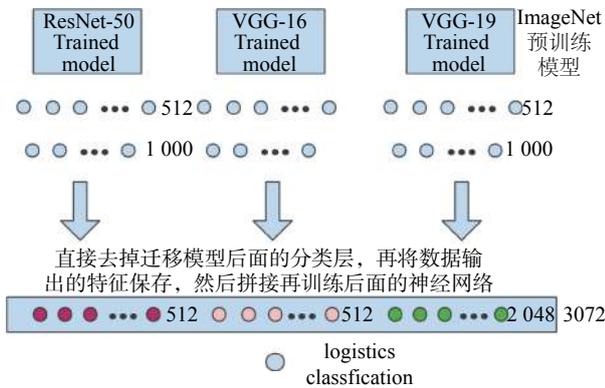


图 8 多视图融合模型融合示意

Fig. 8 Multi view Fusion model fusion diagram

将 3 个预训练模型提取特征之后, 选择预训练模型参数不进行更新, ResNet-50 提取得到 2048 个特征标量, 而 VGG-16 和 VGG-19 分别提取 512 个特征标量, 将 2048+512+512=3072 个特征合并成一个新的特征向量。然后加入 Dropout 层, 设置 drop 概率为 0.5。最后加入 1 个神经元, 指定激活函数为 sigmoid。对于二分类问题, 可以通过 logistic 二分类实现。分类损失设置为二进制交叉熵损失函数。二进制交叉熵损失函数公式如下:

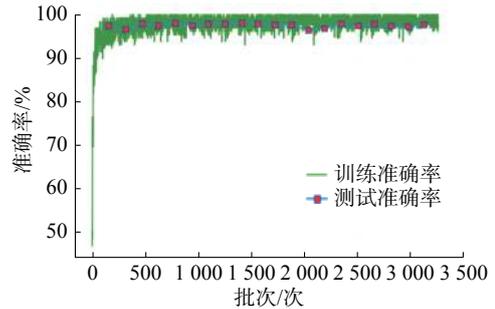
$$\text{sigmoid}(y) = \frac{1}{1 + e^{-y}} \quad (1)$$

$$L(y, \hat{y}) = -\frac{1}{2} [\hat{y} \log(\text{sigmoid}(y)) + (1 - \hat{y}) \log(\text{sigmoid}(1 - y))] \quad (2)$$

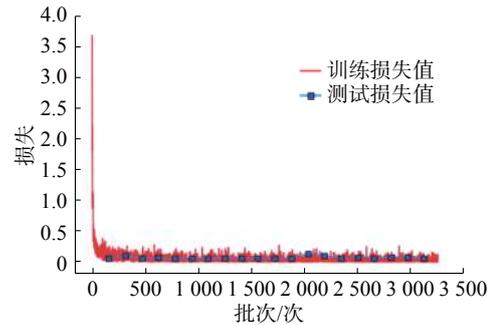
式中: y 为实际神经元输出; \hat{y} 为真实输出标签, 0 表示第一类, 1 表示第二类。

优化器选择 Adadelta 优化方式, 基准学习率设置为 1.0。模型训练迭代次数为 3500 次, batch size 设为 16。

Multi view Fusion 模型训练和测试的准确率和损失示意图如图 9 所示, 图 9(a) 为训练准确率的变化, 图 9(b) 为损失值变化。可以看出随着迭代次数的增加, 该算法准确率逐步上升, 误差值减小。



(a) 训练集准确率变化曲线



(b) 测试集准确率变化曲线

图 9 多视图融合模型训练和测试的准确率和误差
Fig. 9 Accuracy and error of training and testing of the Multi view Fusion model

方案 2 特征平均+特征拼接 (Multi view Fusion-tiny 方案)

方案 1 将 3 个不同的模型特征进行融合, 但是对于 VGG-16 和 VGG-19 来说, 这两个模型构造非常相似, 即提取的特征也会非常类似。两个类似的模型提取的特征之间融合, 造成特征的冗余, 并干扰特征提取。针对上述问题, 设计特征平均后再进行特征拼接的网络模型。采用 ResNet 和 DenseNet 的思想, 将 VGG-16 和 VGG-19 的特征进行特征相加, 然后与 ResNet-50 模型特征进行不同尺度特征拼接, 引入模型多尺度性质。该方法会使模型特征维度减少, 减少网络运行的测试时间。设计的 Multi view Fusion-tiny 模型如图 10 所示。

将 3 个预训练模型提取特征之后, 选择预训练模型参数不进行更新, ResNet-50 提取得到 2048 个特征标量, 而 VGG-16 和 VGG-19 分别提取 512 个特征标量, 首先将 VGG-16 和 VGG-19 两个模型特征进行平均, 得到一个新的 512 维度特征, 与 ResNet-50 的 2048 特征合并成一个新的特征向量, 一共为 2560 维度特征。相比方案一减少了特征维度, 以及模型训练和测试的时间。加入 Dropout 层, 设置 drop 概率为 0.5。加入 1 个神经元, 指定激活函数为 sigmoid。

Multi view Fusion-tiny 模型在每个批次上训练和

测试的准确率及误差示意图如图 11 所示, 图 11(a) 为训练准确率的变化, 即验证数据在每个迭代次数上准确率的变化。图 11(b) 为训练误差的变化。同 Multi view Fusion 模型训练时的准确率及误差图相比, Multi view Fusion-tiny 模型准确率曲线震荡减少, 稳定上升, 同时误差曲线下降。

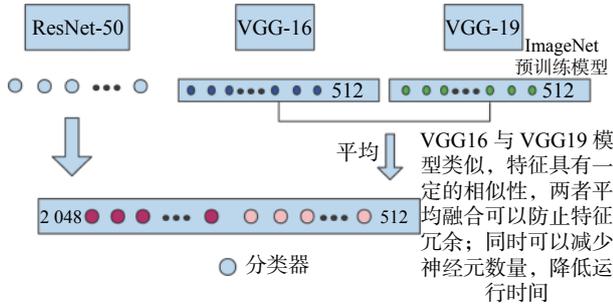
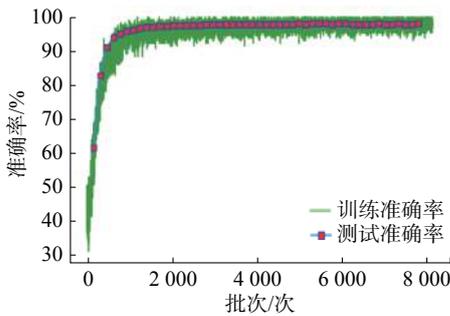
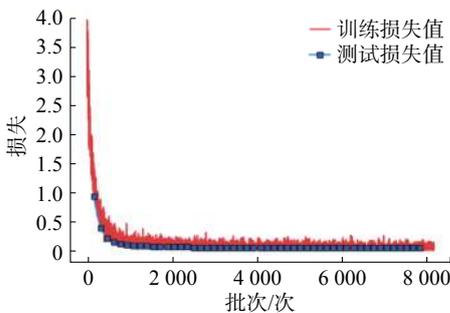


图 10 多视图融合微调模型融合示意
Fig. 10 Multi view Fusion-tiny model fusion diagram



(a) 训练集准确率变化曲线



(b) 测试集准确率变化曲线

图 11 多视图融合微调模型训练和测试准确率及误差
Fig. 11 Accuracy and error of training and testing of the Multi view Fusion-tiny model

方案 3 结果融合投票 (Voted Model)

以上两种方案, 融合之后还是单个模型的决策, 但对于 VGG-16 和 VGG-19, ResNet-50 3 个模型来说, 可以通过大多数分类的结果来决定最终模型的分类结果。即 3 个模型决策投票来决定分类。设计的模型投票决策如图 12 所示。

针对猫狗分类数据训练好的 3 个模型为 VGG-16、VGG-19、ResNet-50, 将这 3 个模型对同一张数据做测试, 会分别得到一个测试结果。然

后将这 3 个结果投票, 投票策略采取投票结果最多的情况作为最后的分类结果。

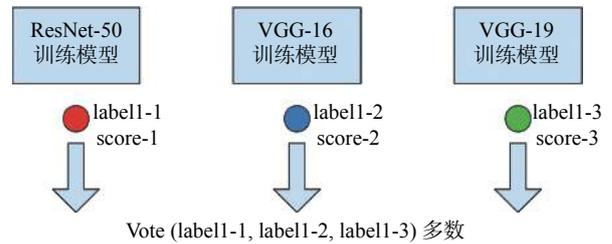


图 12 投票决策模型融合示意
Fig. 12 Voted Model fusion diagram

为定量分析本文基于多感知兴趣区域图像识别方法 3 种模型性能, 对本文算法与 VGG-16、VGG-19、ResNet 模型及 3 种模型的微调模型, DenseNet 模型在 kaggle 数据集上进行实验对比。识别准确率实验结果如表 1 所示。

表 1 模型准确率对比

对比模型	准确率(Acc)/%
VGG-16-No-trainable	98.00
VGG-16-trainable	98.02
VGG-19-No-trainable	97.72
VGG-19-trainable	97.84
ResNet-50-No-trainable	96.97
ResNet-50-trainable	97.05
DenseNet-trainable	98.03
多视图融合模型	98.03
多视图融合微调模型	98.035
投票决策模型	98.52

从表 1 可以看出, 本文设计的模型 Multi view Fusion、Multi view Fusion-tiny 和 Voted Model 模型比单个模型的准确率高, 说明了基于多感知兴趣区域特征融合的图像识别方法更为有效。同时也验证了不同的模型关注的图像兴趣区域不一, 学习到的特征不一, 多兴趣区域特征融合有利于提高分类精度的结论。同时 VGG-16、VGG-19 关联的特征兴趣区域极其相似, 为了防止特征冗余, 将这两个模型特征平均, 以降低特征维度和分类时间。

3 结束语

本文给出了一种基于多感知兴趣区域特征融合的图像识别方法, 通过对单独分类网络的复现与可视化分析, 发现不同的模型关注的图像兴趣

区域及学习到的特征有所差异。针对上述问题构建了多个模型融合机制,通过借鉴 DenseNet 模型的多尺度特征拼接及 ResNet 特征相加机制,设计了 Multi view Fusion 模型、Multi view Fusion tiny 模型和 Voted Model。实验结果表明本文算法在相似目标的二分类问题上具有更高的识别准确率。下一步的研究方向可放在网络结构轻量化与模型的加速上。

参考文献:

- [1] SPENCER JR B F, HOSKERE V, NARAZAKI Y. Advances in computer vision-based civil infrastructure inspection and monitoring[J]. *Engineering*, 2019, 5(2): 199–222.
- [2] CHELLAPPA R. The changing fortunes of pattern recognition and computer vision[J]. *Image and vision computing*, 2016, 55: 3–5.
- [3] 雷明. 机器学习与应用 [M]. 北京: 清华大学出版社, 2019: 26–33.
- [4] MARY N A B, DHARMA D. Coral reef image classification employing improved LDP for feature extraction[J]. *Journal of visual communication and image representation*, 2017, 49: 225–242.
- [5] YU Hua, YANG Jie. A direct LDA algorithm for high-dimensional data—with application to face recognition[J]. *Pattern recognition*, 2001, 34(10): 2067–2070.
- [6] 刘丽, 匡纲要. 图像纹理特征提取方法综述 [J]. *中国图象图形学报*, 2009, 14(4): 622–635.
LIU Li, KUANG Gangyao. Overview of image textural feature extraction methods[J]. *Journal of image and graphics*, 2009, 14(4): 622–635.
- [7] 李磊, 董卓丽. 利用改进图割的彩色图像分割算法 [J]. *武汉大学学报·信息科学版*, 2014, 39(12): 1504–1508.
LI Lei, DONG Zhuoli. Color image segmentation using improved graph cuts[J]. *Geomatics and Information Science of Wuhan University*, 2014, 39(12): 1504–1508.
- [8] 弓晓虹, 郑音飞, 秦佳乐, 等. 基于乘性梯度的医学超声图像边缘检测算法 [J]. *浙江大学学报(工学版)*, 2014, 48(10): 1871–1878.
GONG Xiaohong, ZHENG Yinfei, QIN Jiale, et al. Multiplicative gradient based edge detection method for medical ultrasound image[J]. *Journal of Zhejiang University (Engineering Science)*, 2014, 48(10): 1871–1878.
- [9] 张桂梅, 张松, 储珺. 一种新的基于局部轮廓特征的目标检测方法 [J]. *自动化学报*, 2014, 40(10): 2346–2355.
ZHANG Guimei, ZHANG Song, CHU Jun. A new object detection algorithm using local contour features[J]. *Acta automatica sinica*, 2014, 40(10): 2346–2355.
- [10] YE Qihong, XIANG Ming, CUI Zhendong. Fingerprint image enhancement algorithm based on two dimension EMD and Gabor filter[J]. *Procedia engineering*, 2012, 29: 1840–1844.
- [11] 张丽琼, 王炳和. 基于小波变换的脉象信号特征提取方法 [J]. *数据采集与处理*, 2004, 19(3): 323–328.
ZHANG Liqiong, WANG Binghe. Feature extraction methods for pulse signal based on wavelet transform[J]. *Journal of data acquisition & processing*, 2004, 19(3): 323–328.
- [12] YAGHOobi H, MANSOURI H, FARSANGI M A E, et al. Determining the fragmented rock size distribution using textural feature extraction of images[J]. *Powder technology*, 2019, 342: 630–641.
- [13] MESNIL G, BORDES A, WESTON J, et al. Learning semantic representations of objects and their parts[J]. *Machine learning*, 2014, 94(2): 281–301.
- [14] 胡越, 罗东阳, 花奎, 等. 关于深度学习的综述与讨论 [J]. *智能系统学报*, 2019, 14(1): 1–19.
HU Yue, LUO Dongyang, HUA Kui, et al. Overview on deep learning[J]. *CAAI transactions on intelligent systems*, 2019, 14(1): 1–19.
- [15] SERMANET P, EIGEN D, ZHANG Xiang, et al. Overfeat: integrated recognition, localization and detection using convolutional networks[J]. *Computer science*, 2013.
- [16] ZEILER M D, KRISHNAN D, TAYLOR G W, et al. Deconvolutional networks[C]//Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Francisco, USA, 2010: 2528–2535.
- [17] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. *Communications of the ACM*, 2017, 60(6): 84–90.
- [18] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[EB/OL]. <https://arxiv.org/abs/1409.1556>, 2014.
- [19] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA, 2016: 770–778.
- [20] HUANG Gao, LIU Zhuang, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA, 2017: 2261–2269.

- [21] 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.
ZHENG Yuanpan, LI Guangyang, LI Ye. Survey of application of deep learning in image recognition[J]. *Computer engineering and applications*, 2019, 55(12): 20–36.

作者简介:



闫涵, 硕士研究生, 主要研究方向为图像处理与模式识别。



张旭秀, 教授, 博士后, 主要研究方向为图像处理与模式识别、智能控制、信号处理。主持和参与完成国家自然科学基金项目、辽宁省自然科学基金项目 9 项。发表学术论文 40 余篇。



张净丹, 硕士研究生, 主要研究方向为图像处理与模式识别。

第二十三届中国机器人及人工智能大赛正式启动

中国机器人及人工智能大赛是一项历史悠久、影响广泛的全国性学科竞赛。作为中国人工智能学会最早主办的竞赛之一, 大赛已为我国培养了大量“能动手”、“敢创新”、“可协同”的复合型人才。目前, 大赛已列入中国高等教育学会发布的《2020 年全国普通高等学校学科竞赛排行榜》。

组织架构

主办单位:

中国人工智能学会

承办单位:

CAAI 机器人文化艺术专业委员会

乐聚(深圳)机器人技术有限公司

时间安排

大赛报名系统开放: 2021 年 5 月—6 月;

省赛、区域赛、全国初赛完成时间: 2021 年 7 月底;

全国决赛时间: 2021 年 8 月下旬;

注: 视疫情防控情况, 决定竞赛形式(线上或线下)

参赛详情

参赛对象:

普通高等教育组: 研究生、本科生;

职业教育组: 中职、高职(高专)、职业本科;

联系方式

大赛组委会联系人: 谭庆吉

电话: 15146409277

大赛组委会邮箱: craaic@126.com