

文章编号: 2095-2163(2022)01-0046-07

中图分类号: TP183

文献标志码: A

基于深度学习的驾驶员分心行为识别模型

成福朋, 赵芸

(浙江科技学院 信息与工程学院, 杭州 310023)

摘要: 针对 VGG16 网络识别驾驶员分心时参数量多、分类准确率低的问题, 提出了一种基于多尺度通道的分类模块 MCAM。分类模块 MCAM 包括 MCM 模块和 MSE 模块。MCM 模块将卷积特征图分离为 4 个子特征图, 使用不同的卷积核对 4 个子特征图进行空间信息提取, 以期提高分类准确率; MSE 模块为改进的通道注意力, 其使用一维卷积来改进原始通道注意力中的全连接层, 降低了全连接层大量参数问题。将 MCAM 嵌入 VGG16 中, 并使用非对称卷积辅助降低参数量。实验结果表明, 使用 MCAM 模块在 State Farm Distracted Driver 分心数据集上达到了 97.50% 的识别准确率, 并降低了 VGG16 网络的参数量。

关键词: 驾驶员分心; 注意力; 深度学习; 多尺度卷积

Driver distracted behavior recognition model based on deep learning

CHENG Fupeng, ZHAO Yun

(School of information and electronic engineering, Zhejiang University of Science and Technology, Hangzhou 310023, China)

[Abstract] In order to solve the problem of the large number of parameters and low classification accuracy when the VGG16 network recognizes the driver's distraction, a classification module MCAM based on multi-scale channel dependence is proposed. The classification module MCAM includes MCM module and MSE module. The MCM module separates the convolution feature map into four sub-feature maps, and uses different convolution kernels to extract spatial information from the four sub-feature maps in order to improve the classification accuracy; the MSE module is an improved channel attention, in which one-dimensional convolution is used to improve the fully connected layer in the original channel attention, which reduces a large number of parameter problems in the fully connected layer. Embed MCAM in VGG16, and use asymmetric convolution to help reduce the amount of parameters. The experimental results show that the use of the MCAM module on the State Farm Distracted Driver data set achieves a recognition accuracy of 97.50%, and reduces the number of parameters of the VGG16 network.

[Key words] Driver distraction; Attention; Deep learning; Multiscale convolution

0 引言

据统计数据显示, 大约 25% 的车祸是由于驾驶员分心造成的^[1]。目前, 驾驶员分心检测主要是利用深度学习方法进行识别。王冠^[2]等人根据 Hough 算法进行车道线检测和识别, 然后使用多点透视算法对驾驶员头部姿态进行估计, 分心行驶识别率达到了 80.8%。陈军^[3]等人设计了二级级联网络, 训练速度快, 结构特征冗余较少, 准确率达到 93.3%。王加^[4]等人采用车辆偏航角以及转向盘转速的标准差作为辨识特征量, 准确率达到 85%。Eraqi^[5]等人提出了一种基于深度学习的驾驶员分心检测解决方案, 该方案由卷积神经网络的遗传加权集成组成, 证明了使用遗传算法的加权集成分类器具有更好的分类可信度,

实现了 90% 的准确率。白中浩^[6]等人提出一种基于图卷积的多信息融合驾驶员分心行为检测方法, 准确率达到 93%。刘伟^[7]等人提出一种使用驾驶员局部定位信息, 帮助卷积神经网络识别驾驶员分心驾驶的方法, 准确率达到 96.10%。本文基于文献[20]中的 PSA 模型, 提出改进的 MCAM (Multi-scale Convolutional Attention Module) 分类模块, 通过加入改进的注意力机制来提高分类的准确率。

1 MCAM 分类模块

MCAM 分类模块分为 MCM 模块和 MSE 模块。设计 MCAM 分类模块的目的是解决 VGG16 网络中多层卷积造成网络参数量过大的问题, 并且 MCAM 分类模块中加入了改进的通道注意力, 能够让通道

作者简介: 成福朋(1995-), 男, 硕士研究生, 主要研究方向: 人工智能、深度学习; 赵芸(1982-), 女, 博士后, 教授, 主要研究方向: 人工智能、深度学习。

收稿日期: 2021-10-04

哈尔滨工业大学主办 ◆ 学术研究与应用

间进行通信,并且为每个通道分配不同的权重,使其网络能够关注重要的特征,抑制不重要的特征,以此提高分类的准确率^[20]。

原始 VGG16 的网络结构如图 1 所示^[11]。其中 5 个卷积模块的卷积核尺寸均为 3×3,卷积核使用的数量分别是 64、128、256、512、512。这样的卷积模块设置存在的问题:一是全部使用 3×3 卷积,没有考虑到其它尺寸的卷积核是否可以提取到更多的特征;二是卷积核数量太多,导致网络参数量增多。因此为了解决这两个问题,提出了 MCM 模块。

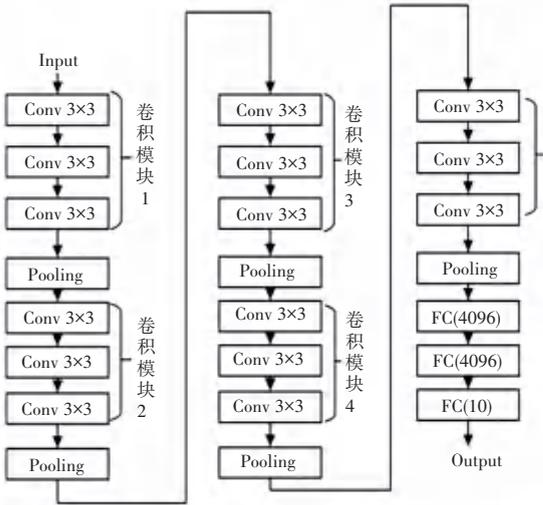


图 1 VGG16 网络结构

Fig. 1 VGG16 network structure

MCAM 分类模块如图 2 所示^[20]。该分类模块主要操作步骤为:

- (1) 上一层的特征图进入 MCM 模块,在通道维度上拆分为 4 个子特征图;
- (2) 4 个子特征图分别使用 3×3、5×5、7×7、9×9 进行卷积;
- (3) 卷积结果进入空间注意力,得到新的 4 个子特征图;
- (4) 使用 MSE 模块提取新的子特征图的注意力向量后,在通道维度上进行拼接。

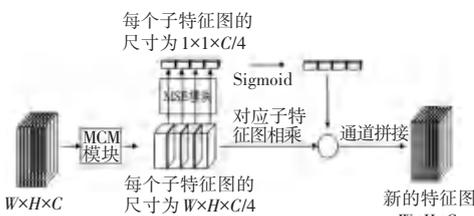


图 2 MCAM 分类模块图

Fig. 2 MCAM classification module diagram

在 VGG16 中插入 MCM 模块后的结构变化如图 3 所示,并将其命名为 MC-VGG16。

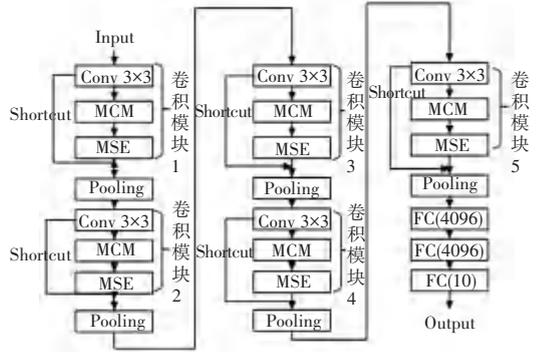


图 3 MC-VGG16 网络结构

Fig. 3 MC-VGG16 network structure

1.1 MCM 模块

MCM 模块^[20]主要是对特征图进行分离并提取不同尺度下的空间信息,解决 VGG16 网络中单尺寸卷积核的问题,并且将 VGG16 网络中每个卷积模块的卷积核数量减少一倍,使用空间注意力弥补特征的损失。空间注意力的优势是只有一个 7×7×1 的卷积,可以保证参数量不会增加太多,而且提取每个通道的特征图在 (x,y) 位置的特征,增强特征表达。因此,MCM 模块对输入特征图进行通道分离,然后进行多尺度下的特征提取。MCM 模块架构如图 4 所示^[20]。

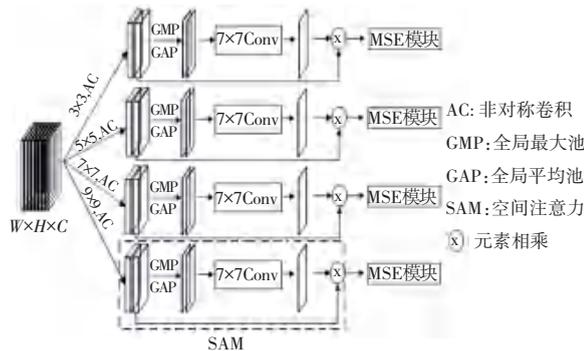


图 4 MCM 模块

Fig. 4 MCM module

图 4 中,对于输入特征图 $W \times H \times C$,将其在通道维度下分离成 4 个子特征图,每个子特征图尺寸为 $W \times H \times C/4$ 。在 4 个尺寸的卷积核进行卷积时,使用非对称卷积,将参数量进一步减少。图中 AC 表示非对称卷积^[19],具体过程如图 5 所示。

图 5 中,左边为 3×3 的卷积核,右边 1×3 和 3×1 是非对称卷积,中间 5×5 为特征图。非对称卷积的做法是将一个 3×3 的卷积核分解为两步:首先使用 1×3 来对特征图进行卷积,然后使用 3×1 对 1×3 的

结果卷积。这样得到的结果与一次 3×3 的卷积结果相同。

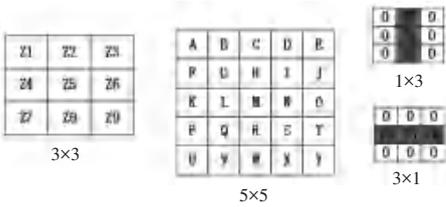


图5 非对称卷积过程

Fig. 5 Asymmetric convolution process

对于非对称卷积第一步得到的 3×3 特征图,首先是做填充,将 3×3 特征图填充为 5×5 ,最后使用 3×1 的卷积核去卷积计算,同样参与计算的参数是 3 个,则两次参与计算的参数是 6 个。而使用 3×3 的卷积核对 5×5 的特征图进行卷积的计算,参与计算的参数是 $3 \times 3 = 9$,则节约的参数量为: $(9-6)/9 = 0.33$ 。因此,非对称卷积减少了大约 33% 的参数量。

为了弥补减少卷积核数量时损失的特征,将 4 个子特征图分别进行空间注意力的提取。对于特征图 $W \times H$ 中的 (x, y) 位置,在通道维度上取平均值后得到最后特征图 $W' \times H' \times 1$,全局最大池化也是同样的操作,只是取的是最大值。将全局平均池化和全局最大池化的结果在通道维度上进行拼接,得到输出特征图 $W' \times H' \times 2$,然后进行 7×7 的卷积,输出特征图 $W' \times H' \times 1$,最后与输入特征图 $W' \times H' \times C$ 进行元素相乘。通过空间注意力,将原始图像中的空间信息变换到另一个空间中并保留了关键信息。在经过空间注意力后,4 个子特征图作为 MSE 模块的输入。因此, MCM 模块中对于输入特征图 $W \times H \times C$,首先进行通道维度下的分离,平均分离为 4 个子特征图,分离如公式(1)所示^[20]:

$$Result = split(F, count = 4, axis = 3) \quad (1)$$

式中, F 表示输入特征图 $W \times H \times C$; $split$ 表示分离操作; $count$ 表示将 F 分离为子特征图的个数; $axis$ 表示在通道维度上进行分离; $Result$ 表示分离结果。

分离操作完成后,分别对 4 个子特征图进行不同尺寸的卷积操作,卷积如公式(2)所示^[20]:

$$Conv_i = conv2d(Result_i, channel/4, size = y) \quad (2)$$

式中, $Result_i$ 表示保存在 $Result$ 中的第 i 个子特征图; $channel$ 表示输入子特征图的通道数; $channel/4$ 表示使用卷积核的数量; $size$ 表示使用的卷积核的尺寸为 $y \times y$; $Conv_i$ 表示第 i 个子特征图的卷积结果。 $Result_i$ 中的 i 变量取值为: $\{0, 1, 2, 3\}$, $size$ 的 y 取值为: $\{3, 5, 7, 9\}$ 。

在计算时,相同位置的数字要一一对应,第 0 个

子特征图对应 y 为 3,第 1 个子特征图对应 y 为 5,第 2 个子特征图对应 y 为 7,第 3 个子特征图对应 y 为 9。为了进一步降低参数量,使用图 5 所示的非对称卷积。

卷积操作完成后,需要对每个子特征图进行空间注意力的计算,计算公式为:

$$F' = F * Conv(GMP(F) + GAP(F)) \quad (3)$$

式中, F 表示输入的特征图 $W \times H \times C/4$; GMP 和 GAP 分别表示全局最大池化和全局平均池化; $+$ 表示在通道维度上进行特征图拼接; $*$ 表示 7×7 卷积结果与 F 进行元素相乘得到输出结果 F' 。

将 7×7 卷积后的特征图 $W' \times H' \times 1$ 标记为 F_{conv} 。则对于输入特征图 F 与 F_{conv} 进行元素相乘时,由于 F 的通道维度是 $C/4$, F_{conv} 的通道维度为 1。因此分别将 F 的每一通道维度的特征图与 F_{conv} 进行对应位置元素相乘,输出的特征图 F' 为 $W' \times H' \times C/4$ 。

1.2 MSE 模块

MSE 模块基于 SENet^[9] 的思想,主要作用是提升网络准确率,并在此基础上减少网络参数量。SENet 模块中使用全连接层来限制模型复杂度,增加泛化能力。然而,全连接层还是带来了大量的参数。因此,SENet 模块虽然使模型性能提升,但是也带来了参数量增加的问题。SENet 如图 6 所示^[9]。

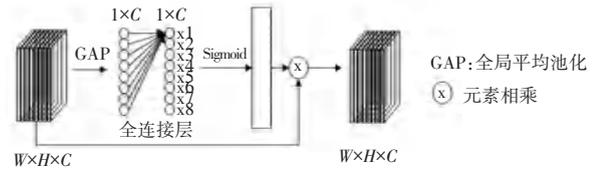


图6 SENet

Fig. 6 SENet

在 SENet 中,全连接层的计算采用全局卷积计算方式。但这样的卷积方式参数量较多,特别是当输入特征图 $W \times H \times C$ 的 C 增加时,全连接层的通道 C 也会增加,参数量就会成倍的增长。因此, MSE 模块沿用 SENet 中加权的方式,改用一维卷积而不是用全连接层。由于 MSE 模块中特征图通道之间的通信是在通道 K 邻域内^[10],这意味着不需要将输入特征图 $1 \times C$ 中的 C 个通道完全进行卷积,只在 C 通道的邻域内进行计算即可。MSE 模块如图 7 所示^[10]。

完整的 MSE 流程如图 8 所示。

其中,一维卷积过程如图 9 所示。

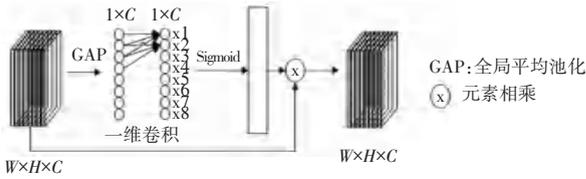


图 7 MSE 模块
Fig. 7 MSE module

2.2 评价指标

对于 MC-VGG16 网络, 使用准确率 $Accuracy(A)$ 、精确率 $Precision(P)$ 、召回率 $Recall(R)$ 进行模型性能评价, 其公式如公式(4) ~ 公式(6) 所示。

$$Accuracy(A) = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

$$Precision(P) = \frac{TP}{TP + FP} \quad (5)$$

$$Recall(R) = \frac{TP}{TP + FN} \quad (6)$$

式中, TP 表示模型预测为正例, 实际也为正例; FP 表示模型预测为正例, 实际为负例; TN 表示模型预测为负例, 实际上为正例; FN 表示模型预测为负例, 实际上为负例。

数据集中有十类驾驶员分心行为, 当前类别识别时标记为正例, 其他类别则为负例。例如对当前 c_0 类来说, 其是正例, 其他 9 类均是负例。

2.3 数据集

实验中使用 State Farm Distracted Driver 数据集^[8]。数据集由驾驶员 10 种行为的训练集图像 17 462 个, 测试集图像 4 961 个组成, 图片是 RGB 图像, 分辨率为 320×240 。10 种行为分别是: 安全驾驶、右手编辑手机、右手打电话、左手编辑手机、左手打电话、操作收音机、喝东西、向后找东西、整理发型或化妆、与乘客交谈。训练集有驾驶员行为的真实标签, 测试集不含有驾驶员行为的真实标签。将测试集中的 10 种驾驶员图像手动分类, 并将训练集和测试集中的每一类别分别编号为 $c_0, c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8, c_9$ 。训练集和测试集每个类别的数量见表 1、表 2。

表 1 训练集细节

Tab. 1 Training Set details

| 类别编号 | 类别标签名称 | 图像数量 |
|-------|---------|--------|
| c_0 | 安全驾驶 | 1 938 |
| c_1 | 右手编辑手机 | 1 765 |
| c_2 | 右手打电话 | 1 804 |
| c_3 | 左手编辑电话 | 1 826 |
| c_4 | 左手打电话 | 1 812 |
| c_5 | 操作收音机 | 1 801 |
| c_6 | 喝东西 | 1 811 |
| c_7 | 向后找东西 | 1 559 |
| c_8 | 整理发型或化妆 | 1 488 |
| c_9 | 与乘客交谈 | 1 658 |
| 总计 | | 17 462 |

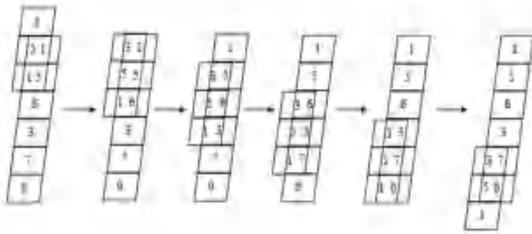


图 9 一维卷积过程

Fig. 9 One dimensional convolution process

图 9 中, $(3, 5, 1)$ 表示一维卷积核 3×1 , $(1, 5, 6, 3, 7, 0)$ 表示重组后的一维特征图 $1 \times C$ 。一维卷积核的尺寸为 3×1 , 其尺寸的确定是由于在 MSE 模块中, 特征图通道之间的通信是在通道 K 邻域内, 因此这里选取 $K = 3$ 进行示例说明。卷积核中的数据分别是 3、5、1。其中, 5 为锚点。在卷积时, 锚点依次移动到特征图的每一个位置处, 对应位置相乘再求和, 得到卷积后的特征图尺寸也为 $1 \times C$ 。

MSE 模块中, 由于卷积时很容易做 $padding$, K 邻域一般取奇数, 并且只有是奇数时, 一维卷积核才有锚点。

2 实验结果与分析

2.1 实验环境与参数

实验中, 集成开发环境为 JetBrains PyCharm 2018.1.4 x64, 操作系统为 win10, 内存 32 G, 8 核 CPU, 显卡为 NVIDIA GeForce RTX 2080 Ti, 开源框架为 Keras。训练网络时, $batch_size$ 设置为 16, $epoch$ 设置为 120, 优化器为 Adam, 学习率为 1×10^{-5} 。

表2 测试集细节
Tab. 2 Testing Set details

| 类别编号 | 类别标签名称 | 图像数量 |
|------|---------|-------|
| c0 | 安全驾驶 | 551 |
| c1 | 右手编辑手机 | 502 |
| c2 | 右手打电话 | 513 |
| c3 | 左手编辑电话 | 519 |
| c4 | 左手打电话 | 514 |
| c5 | 操作收音机 | 511 |
| c6 | 喝东西 | 514 |
| c7 | 向后找东西 | 443 |
| c8 | 整理发型或化妆 | 423 |
| c9 | 与乘客交谈 | 471 |
| 总计 | | 4 961 |

最后对数据集的处理是数据增强,这是为了在进行模型识别时丰富图像的多样性,数据增强的目标是减少网络的过拟合,增加模型的泛化性。实验中图形增强主要是将图像进行水平翻转和竖直翻转,如图10所示。



(a) 原始图像 (b) 水平翻转 (c) 竖直翻转

图10 数据增强

Fig. 10 Data Augmentation

数据集中有10类驾驶员分心行为,当前类别识别时标记为正例,其它类别则为负例。例如,对当前c0类来说,其是正例,其他9类均是负例。

2.4 消融实验

MC-VGG16网络是基于VGG16网络而来,因此在实验时需将两种网络进行对比。又由于MSE模块中特征图的通道通信是在 K 邻域内,为了找出效果更好的 K ,将 K 取值为 $\{3, 5, 7, 9\}$,依次实验。实验结果见表3。

由表中数据可见,与原始VGG16网络相比,MC-VGG16网络中 $K = 3, 5, 7, 9$ 时,都比原始网络VGG16的准确率高,且统计出的参数量也比VGG16网络的参数量要少。当 $K = 3$ 时,比原始VGG16的网络降低了13,356,088个参数。其中,当 $K = 9$ 时,分类效果比 $K = 3$ 时稍低,但也降低了VGG16网络

13,355,888个参数。能够降低参数数量的原因:一是所用模型中的卷积核数量较少,二是非对称卷积的使用。实验结果证明,MC-VGG16不仅可以提高分类准确率,更能明显的降低原始VGG16模型的参数量。

表3 不同模型准确率和参数减少量对比

Tab. 3 Comparison of accuracy and parameter reduction of different models

| 模型 | 测试准确率/% | 相比原模型减少的参数量 |
|--------------------------|---------|-------------|
| 原始 VGG16 ^[11] | 93.88 | - |
| MC-VGG16 MSE; $K = 3$ | 97.50 | 13,356,008 |
| MC-VGG16 MSE; $K = 5$ | 97.30 | 13,355,968 |
| MC-VGG16 MSE; $K = 7$ | 97.24 | 13,355,928 |
| MC-VGG16 MSE; $K = 9$ | 97.16 | 13,355,888 |
| VGG16+MCM+SENet | 97.20 | 13,345,724 |
| VGG16+MCM | 95.53 | 13,356,088 |

为了证明MCM模块中多尺寸卷积以及空间注意力的有效性,仅在VGG16模型中添加MCM模块后,与原始VGG16网络进行了对比。从表3中看出,VGG16+MCM模型的准确率达到95.53%,比原始VGG16网络的准确率高了1.65%,并且参数量减少了13,356,088个。因此,MCM模块的加入不仅降低了VGG16的参数量,并且提高了准确率。

由于MSE模块基于SENet模块进行改进,因此需要将MSE模块和SENet模块进行对比。由表3中可以看出,MC-VGG16+SENet模型虽然减少了原始VGG16模型13,345,724个参数,但是却不如当 $K = 3$ 时的MC-VGG16模型减少的参数量多,且MSE模块比SENet模块具有更高的准确率。结论是:当 $K = 3$ 时,模型准确率较高,且参数量最少。

2.5 损失和准确率结果

当 $K = 3$ 时,MC-VGG16模型的测试准确率达到97.50%,其主要原因是MCM模块和MSE模块改进了原始VGG16网络中的卷积层,解决了使用卷积层带来大量参数的问题;对子特征图进行加权,保留重要特征,删除不重要特征,不仅提升了准确率,而且辅助使用非对称卷积进一步来减少参数量,实现了更高的准确率和更低的参数量。 $K = 3$ 时模型的训练准确率、测试准确率、训练损失、测试损失如图11所示。

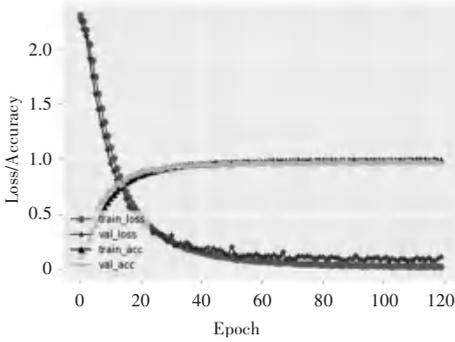


图 11 准确率和损失曲线

Fig. 11 Accuracy and loss curve

图中横轴代表迭代次数 Epoch, 纵轴代表准确率和损失。整体而言, 模型没有出现过拟合的现象, 证明当 $K = 3$ 时模型分类效果理想。

2.6 不同方法对比

为了证明 MC-VGG16 模型的有效性, 将其与其它分类模型进行对比。不同模型的准确率见表 4。对比结果证明: MC-VGG16 模型具有更高的分类准确率, 并且使用 MCAM 模块可以降低网络的参数量。

表 4 不同方法的准确率比较

Tab. 4 Comparison of accuracy of different methods

| 模型 | 准确率/% |
|----------|-------|
| MC-VGG16 | 97.50 |
| 文献[3] | 93.3 |
| 文献[6] | 90 |
| 文献[7] | 96.10 |
| 文献[12] | 96.74 |
| 文献[13] | 97 |
| 文献[14] | 96.23 |
| 文献[15] | 97 |

几种模型对每一类别的准确率的对比结果见表 5。

表 5 不同方法对每一类别的准确率比较

Tab. 5 Comparison of accuracy of different methods for each category

| 类别 | VGG19 ^[16] /% | PAV-Hint CNN ^[17] /% | Multi-stream CNN ^[18] | MC-VGG16/% |
|----|--------------------------|---------------------------------|----------------------------------|------------|
| c0 | 84 | 85 | 88 | 97 |
| c1 | 78 | 80 | 85 | 100 |
| c2 | 83 | 82 | 87 | 96 |
| c3 | 81 | 83 | 88 | 99 |
| c4 | 79 | 82 | 87 | 98 |
| c5 | 80 | 81 | 87 | 97 |
| c6 | 78 | 81 | 89 | 98 |
| c7 | 80 | 82 | 85 | 98 |
| c8 | 79 | 81 | 84 | 96 |
| c9 | 78 | 80 | 85 | 95 |
| 平均 | 80 | 82 | 87 | 97 |

表 5 中通过对比 VGG19、PAV-Hint CNN、Multi-stream CNN 3 种方法每一类别的准确率, 证明了 MC-VGG16 模型分类效果更好。

2.7 可视化结果

从数据集选取十类分心图像进行识别, 如图 12 所示。图像下方的文字表示该图像的实际行为类别, 在每个图像上显示的文字为模型预测的图像类别结果。由此可见, 模型可以正确预测将每个图像所属的类别, 证明 MC-VGG16 模型具有良好的识别效果。



图 12 可视化结果

Fig. 12 Visualization results

3 结束语

本文的 MCAM 模型的主要目的是降低 VGG16

网络中的参数量, 并且加入注意力机制来提高分类准确率, 实验表明该模块是有效的。在实验过程中使用非对称卷积来进一步减少网络的参数量, 并且

分析了通道间通信的 K 邻域的取值不同也会导致最后网络识别的效果不同。尽管模型取得了不错的识别效果,但是当两种分心行为相似时会出现识别错误,后期研究会结合更多的特征进行研究。此外,数据集中的图像都是清晰可见的,因此下一步的研究方向可以是如何对低分辨率的图像进行分类识别,以此保证驾驶员的生命安全。

参考文献

- [1] HATOYAMA K, NISHIOKA M, KITAJIMA M, et al. Perception of Time in Traffic Congestion and Drivers' Stress [C]// International Conference on Transportation and Development. IEEE, 2019:165.
- [2] 王冠,李振龙. 基于反向双目识别的驾驶员分心检测[J]. 科学技术与工程, 2018,18(17):87-93.
- [3] 陈军,张黎,周博,等. 基于级联卷积神经网络的驾驶员分心驾驶行为检测[J]. 科学技术与工程, 2020,20(14):265-271.
- [4] 王加,陈慧. 基于驾驶员操纵及车辆运动轨迹信息的驾驶分心辨识方法[J]. 汽车技术,2013(10):14-18.
- [5] ERAQI H M, ABOUELNAGA Y, SAAD M H, et al. Driver Distraction Identification with an Ensemble of Convolutional Neural Networks[J]. Journal of advanced transportation, 2019(1):1-12.
- [6] 白中浩,王韞宇,张林伟. 基于图卷积网络的多信息融合驾驶员分心行为检测[J]. 汽车工程, 2020,42(8):1027-1033.
- [7] 刘伟,周广平,杨春亭. 基于局部信息和卷积网络的分心行为识别[J]. 信息技术, 2020,44(7):12-16.
- [8] Kaggle: State farm distracted driver detection. Retrieved From <https://www.kaggle.com/rightway11/state-farm-distracted-driver-detection>. 2019.
- [9] JIE H, LI S, GANG S, et al. Squeeze-and-Excitation Networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017(99).
- [10] WANG Q, WU B, ZHU P, et al. ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks [C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2020:11531-11539.
- [11] SIMONYAN K, ZISSERMAN A. Very Deep Convolutional Networks for Large-Scale Image Recognition [J]. Computer Science, 2014. CarXiv: 1409.1556.
- [12] HUANG C, WANG X, CAO J, et al. HCF: A Hybrid CNN Framework for Behavior Detection of Distracted Drivers[J]. IEEE Access, 2020, 8:109335-109349.
- [13] DHAKATE K R, DASH R. Distracted Driver Detection using Stacking Ensemble [C]// 2020 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS). IEEE, 2020:1002.
- [14] ALOTAIBI M, ALOTAIBI B. Distracted driver classification using deep learning [J]. Signal, Image and Video Processing, 2020, 14(3):617-624.
- [15] ABDUL J V, JANET B, REDDY U S. Real Time Detection of driver distraction using CNN [C]// 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT). 2020:185-191.
- [16] KOESDWIADY A, BEDAWI S M, OU C, et al. End-to-End Deep Learning for Driver Distraction Recognition [J]. Springer, Cham, 2017:11-18.
- [17] QI T, XU Y, QUAN Y, et al. Image-based action recognition using hint-enhanced deep neural networks [J]. Neurocomputing, 2017, 267(dec.6):475-488.
- [18] HU Y, LU M, LU X. Driving behaviour recognition from still images by using multi-stream fusion CNN [J]. Machine Vision and Applications, 2019, 30(5):851-865.
- [19] SZEGEDY C, VANHOUCHE V, IOFFE S, et al. Rethinking the Inception Architecture for Computer Vision [C]// IEEE. IEEE, 2016:2818-2826.
- [20] ZHANG H, ZU K, LU J, et al. EPSANet: An Efficient Pyramid Split Attention Block on Convolutional Neural Network. arXiv preprintXiv:20105.14447,2021.

(上接第45页)

阶段,通过把图像分成两部分并分别分割,大大提高了分割精度,有效抑制了传统分水岭算法的过分割现象。实验证明,本文的算法满足 T 细胞菌斑图像的计数要求,对 T 细胞菌斑图像有良好的分割效果。由于本文算法只需稍作修改,因此也适用于其它具有类似特征的菌斑或细胞图像的分割,应用前景较广。

参考文献

- [1] 丁毅,穆银玉. 结核感染 T 细胞斑点试验检测的临床意义 [J]. 现代实用医学, 2018, 30(10):87-88.
- [2] 刘琳,王敬亚,祖蓓蓓,等. T-SPOT.TB 试验对 240 例疑似结核患者的诊断分析 [J]. 中华临床医师杂志(电子版), 2016, 10(3):440-442.
- [3] SHAH M I, MISHRA S, SARKAR M, et al. Automatic Detection and Classification of Tuberculosis Bacilli from ZN-stained Sputum Smear Images using Watershed Segmentation [C]// International Conference on Signal Processing. 2016.
- [4] 桑艳艳,李昕. 基于改进分水岭算法的菌落图像分割 [J]. 电子测量技术, 2019,42(6):87-93.
- [5] 张力新,张黎明,杜培培,等. 基于改进水平集的菌落图像智能计数算法 [J]. 天津大学学报(自然科学与工程技术版), 2019, 52(1):84-89.
- [6] 邱桑敏,夏雨人. 一种快速霍夫变换算法 [J]. 计算机工程, 2004(2):148-150.
- [7] 司秀娟,王红强,宋良图. 基于图像处理的菌落自动计数系统 [J]. 北京生物医学工程, 2017,36(6):633-638.
- [8] OTSU N. A Threshold Selection Method from Gray-Level Histogram [J]. Automatica, 1975, 11(23):285-296.
- [9] BIENIEK A, MOGA A. An efficient watershed algorithm based on connected components [J]. Pattern Recognition, 2000, 33(6):907-916.
- [10] 魏光杏,吴锡生. 克服分水岭算法过分割的方法 [J]. 服装学报, 2008(1):38-41.
- [11] 陈洁,胡永,刘泽国. 基于标记的分水岭图像分割算法研究 [J]. 软件, 2012(9):115-117.
- [12] 胡敏,蔡慧芬. 基于形态学标记连通的分水岭图像分割 [J]. 电子测量与仪器学报, 2011,25(10):864-869.