

[文章编号] 1003—4684(2021)05-0042-05

一种 EfficientNet 网络下雾天道路能见度估测方法

王年涛, 王淑青, 张鹏飞, 顿伟超, 黄剑锋

(湖北工业大学电气与电子工程学院, 湖北 武汉 430068)

[摘 要] 针对传统的激光雷达能见度仪对雾天能见度检测存在成本高和覆盖率低的问题,提出一种 EfficientNet 网络模型对道路摄像头拍摄雾天图像进行能见度实时估测的方法。该方法首次将 EfficientNet 网络应用于雾天道路能见度的估测。首先对道路摄像头拍摄的雾天视频进行定时抽帧采集图像,然后在深度学习框架下使用优化的 EfficientNet 深度学习网络进行训练分类,最后用训练好的网络对验证集进行检验。实验结果显示, EfficientNet 分类网络对能见度分类的平均准确率达 96%,平均准确率高于 VGG16、Darknet53、ResNet50 对照网络,能够实现大雾天气下道路能见度的估测。

[关键词] EfficientNet; 雾天图像; 图像分类; 能见度

[中图分类号] TP391.41 **[文献标识码]** A

气象能见度是视力正常的人将目标从背景中识别出来的最大水平距离^[1],以气象光学视程(MOR, Meteorological Optical Range)作为能见度的特征量来表示能见度。能见度是气象、公路行车、飞机飞行中常见指标,能见度估测对道路行车及机场航班起降安全非常重要,对雾天能见度估测的研究,减少了交通事故的发生率。对道路交通、自动驾驶、气象预报领域有一定的应用价值。

传统的雾天图像能见度检测算法采用人工提取图像特征边缘、颜色、对比度等研究能见度的检测方法,如文献[2]采用小波变换提取图像边缘特征点,根据可见目标的匹配结果估计能见度,对 500m 内的能见度检测有效果,但是超出了该范围就无法检测。对于通过已有观测点的气象环境数据,推断环境因素与能见度的关系问题,已经得到许多专家学者的关注研究。文献[3]提出了长短期记忆短临预测模型以实现大雾的临期预测,最终验证模型预报效果良好。文献[4]提出了通过语义模糊场景理解图像去雾算法,其用于图像去雾的理论对于雾天能见度的估计有一定参考价值。深度学习算法在环境能见度估测及预报问题的应用并不鲜见,文献[5]使用 AlexNet 网络训练道路摄像头采集的图像,模型得到的平均识别率达 78.02%,精度一般;文献[6]采用一种端到端的深度学习算法,利用序数关系使网络提取的能见度属性在空间上按规律分布,根据图像属性之间的相对位置关系估测能见度,是一种较

为新颖的能见度估测方法。

本文使用现有的视频及图像数据训练网络模型,提取重要特征信息,建立特征与环境的联系,根据特征的变化实现能见度数值的估测。前期通过在一个固定的地点对激光雷达能见度仪和摄像机采集的图像数据进行训练,后期就可以直接利用训练好的网络对该地点摄像头拍摄的图像进行估测,大大减少了能见度仪的使用,该方法后期维护成本低,可以将该方法应用于城市道路、机场、自动驾驶等领域。常见的图像分类网络有 VGGNet^[7]、Darknet^[8]、ResNet^[9]、EfficientNet^[10]等,本文采用优化的 EfficientNet 网络对雾天图像进行训练分类,传统的图像分类网络只对网络深度或宽度或输入图像分辨率某个方面进行改变, EfficientNet 网络的优势在于同时改变网络的深度、宽度和图像分辨率,通过多次训练,找到一组分类最优的复合参数,经过优化训练好的 EfficientNet 网络直接用于该地点雾天图像估测,测试结果显示,估测效果优于 VGGNet、Darknet、ResNet 对照网络。

1 数据预处理

1.1 数据采集

雾天道路能见度估测数据集是一个自制数据集,选择一个有雾的时间段,并选择合适的道路,调整好激光雷达能见度仪和摄像头角度,同时开启能见度仪和摄像头,将视频尺寸设置为 1280×720,视

[收稿日期] 2021—05—06

[第一作者] 王年涛(1995—),男,湖北黄冈人,湖北工业大学硕士研究生,研究方向为图像处理

[通信作者] 王淑青(1969—),女,河北衡水人,湖北工业大学教授,研究方向为智能检测与控制

频的帧速率默认为 30 f/s。激光雷达能见度仪记录了拍摄视频每一帧图像的能见度真实数据,由于测量的能见度范围从 0~1200 m,故将能见度的测量精度设置为 50 m,采集的视频通过程序每 5 s 进行一次取图,筛选出 10230 张图片。图 1 是从采集的数据集中随机抽取的雾天图像。



图 1 随机选取部分雾天图像

1.2 数据的分类标注

在深度学习网络执行分类任务之前,所待分类的数据均需要根据图片的时间信息与激光雷达能见度仪的测量数据信息根据程序进行匹配与分类,即采用监督学习。将每张获得的图片根据能见度数值进行分类,方便模型的训练与验证,根据激光雷达能见度仪检测的能见度真实数据和对应时间点的照片,将数据集分为 18 个类别,再将分类的图片随机选取 72% 的标记图像作为训练集,18% 作为测试集,10% 作为验证集。

1.3 深度学习网络配置

训练深度学习网络对平台的软硬件配置要求较高,由于 GPU 对图像的计算能力远高于 CPU,所以本次训练将在单 GPU 上进行,网络训练平台的配置:操作平台 Windows 10,深度学习框架 Tensorflow, GPU 型号 NVIDIA GeForce GTX 1080 Ti,显存 11G,内存 16G, GPU 加速库 CUDA 10.1 和 CUDNN 7.5.0。

同时设置网络模型的训练参数,训练参数的配置信息:输入图像尺寸采用默认采集的图像尺寸 1280×720;经过多次迭代实验,发现网络迭代达到 500 次之后,平均准确率和损失值满足分类要求,故将网络迭代次数确定为 500 次,同时将批大小设为 16;初始学习率设置过大损失值会偏大,初始学习率过小训练的时间会相对较长,设置初始学习率为 0.001;优化器的任务是在全部的数据集训练一次后计算损失函数的梯度,根据实验对比,选择 Adam

优化器。

2 能见度估测方法

2.1 雾天道路能见度估测原理

道路能见度估测流程如图 2 所示,选择有雾的时间段和合适的道路,通过激光雷达能见度仪和摄像头同步采集有雾天气下的道路视频,再对拍摄的视频进行定时抽帧,将每张图片 and 能见度仪测量的能见度数据进行一一对应,即得到了带有能见度值的雾天图像,根据图片能见度对数据集进行分类标注,使用 VGG16、Darknet53、ResNet50、EfficientNet 深度学习网络进行训练分类,对分类结果通过多项指标进行对比分析,将测试的结果与激光雷达能见度仪实测的结果对照,判断能见度估测准确度是否达到标准,经过多次迭代训练,估测的准确度满足要求后,即可对该道路监控视频进行能见度估测。

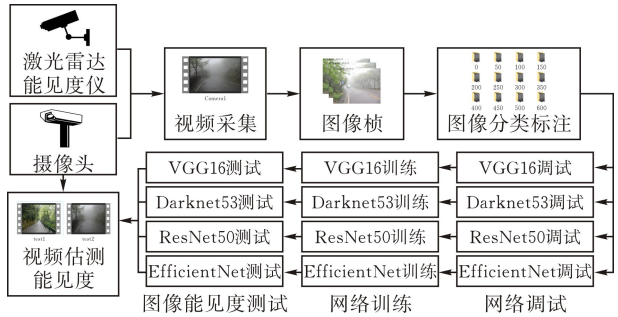


图 2 雾天道路能见度估测流程图

2.2 EfficientNet 网络算法原理

通常大多数深度学习网络是改变网络深度、宽度和图像分辨率中某个方面, EfficientNet 网络采用复合模型缩放思想,寻找一组合适缩放系数,对以上三个维度同时改变,使缩放后的网络达到最优的分类效果^[11]。复合缩放网络算法的优点是:不仅分类精度高于单改变某一个参数,而且占用物理内存较小。 EfficientNet 网络复合参数调节方法原理见图 3,其中图 3a 有固定的深度、宽度和分辨率,图 3b 对网络宽度进行变换,图 3c 对网络深度进行变换,图 3d 对输入图像的分辨率变换,图 3e 同时对网络深度、宽度和分辨率变换。

EfficientNet 网络的关键是找到最合适的一组网络深度、网络宽度和图像分辨率的复合参数。 EfficientNet 网络的算法原理如式(1):

$$N(d, w, r) = \bigodot_{i=1 \dots s} f_i^{L_i} (X_{(H_i, W_i, C_i)}) \tag{1}$$

其中 d, w, r 是缩放网络分别对宽度、深度、分辨率的缩放系数, $N(d, w, r)$ 是分类网络, \bigodot 表示连乘运算, i 表示阶段, X 表示输入, f_i 表示基础网络层, L_i 表示基础网络层 f_i 在第 i 个阶段中的重复次数, 共有 s 个阶段, H_i 表示网络的深度, W_i 表

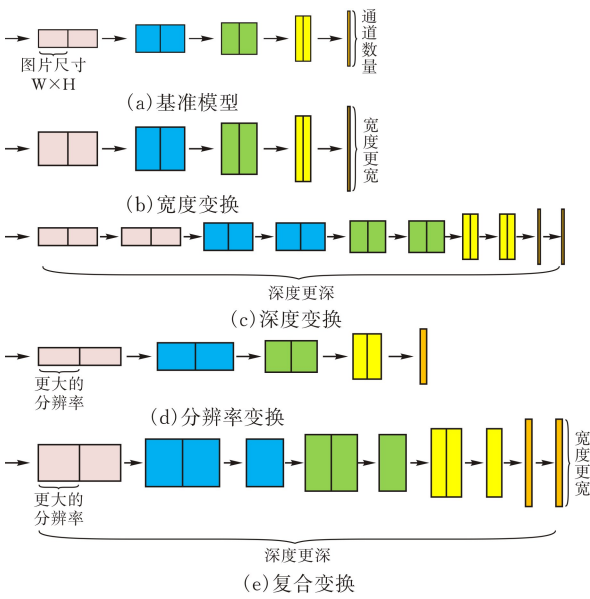


图3 EfficientNet网络复合参数调节方法原理图

示网络的宽度, C_i 表示网络的通道数。

为进一步研究 d, w, r 对网络的影响, 建立三者之间的联系, 在网络参数和计算量满足一定的条件下, 有如式(2)关系:

$$\begin{aligned} & \max_{d, w, r} A_{\text{accuracy}}(N(d, w, r)) \\ & \text{s.t. } N(d, w, r) = \bigodot_{i=1, \dots, s} \hat{f}_i^{d \cdot \hat{L}_i}(X_{(r, \hat{H}_i, r, \hat{W}_i, \hat{C}_i)}) \quad (2) \\ & \text{Memory}(N) \leq \text{target_memory} \\ & \text{FLOPS}(N) \leq \text{target_flops} \end{aligned}$$

其中 $\hat{f}, \hat{L}_i, \hat{H}_i, \hat{W}_i, \hat{C}_i$ 是基础网络中预先设置的参数, \hat{f} 是预设网络结构, \hat{L}_i 是预设层数, \hat{H}_i, \hat{W}_i 是预设分辨率, \hat{C}_i 是预设通道数, $\text{Memory}(N)$ 为网络参数量, target_memory 为参数量阈值, $\text{FLOPS}(N)$ 为每秒浮点运算次数, target_flops 为浮点运算量的阈值, $\max_{d, w, r} A_{\text{accuracy}}$ 为模型的最大精度, 通过此算法找到最大的模型精度。为了将 d, w, r 三个参数统一为一个参数, 引入参数 φ , 如式(3)所示:

$$\begin{aligned} & \text{depth: } d = \alpha^\varphi \\ & \text{width: } w = \beta^\varphi \\ & \text{resolution: } r = \gamma^\varphi \quad (3) \\ & \text{s.t. } \alpha \cdot \beta^2 \cdot \gamma^2 \approx 2 \\ & \alpha \geq 1, \beta \geq 1, \gamma \geq 1 \end{aligned}$$

其中 α, β, γ 是对应维度的资源分配参数, 参数 φ 是控制系数, 参数 φ 取 0 到 7 之间不同的整数, 对应了网络的 8 个不同的模型, 即 EfficientNetB0, EfficientNetB2, ..., EfficientNetB7。

2.3 EfficientNet 网络结构和参数

EfficientNet 网络共有 9 个阶段, 基线网络 EfficientNetB0 网络结构见表 1。

第一阶段是通过一个 3×3 的卷积核进行步长为 2 的普通卷积, 其中包含批量归一化 BN(Batch

Normalization)和激活函数 Swish 操作; 接下来经过第二到第八阶段重复堆叠移动翻转瓶颈卷积(Mobile inverted Bottleneck Convolution, MBConv)结构, MBConv 卷积的具体过程如图 4 所示, 先通过 1×1 的卷积核升维, 再通过 Depwise Conv 卷积和注意力机制(Squeeze-and-excitation Nets, SENets)模块操作^[12], 即在深度可分离卷积前加了一层逐点卷积; 第九阶段是通过 1×1 的卷积层卷积, 再通过池化和全连接处理。

表 1 EfficientNet-B0 的网络结构

阶段 (i)	操作 (f_i)	分辨率 ($H_i \times W_i$)	通道 (C_i)	层数 (L_i)	步长 (S)
1	Conv 3×3	224×224	32	1	2
2	MBconv1, $k3 \times 3$	112×112	16	1	1
3	MBconv6, $k3 \times 3$	112×112	24	2	2
4	MBconv6, $k5 \times 5$	56×56	40	2	2
5	MBconv6, $k3 \times 3$	28×28	80	3	2
6	MBconv6, $k5 \times 5$	14×14	112	3	1
7	MBconv6, $k5 \times 5$	14×14	192	4	2
8	MBconv6, $k3 \times 3$	7×7	320	1	1
9	Conv 1×1 & pooling & FC	7×7	1280	1	—

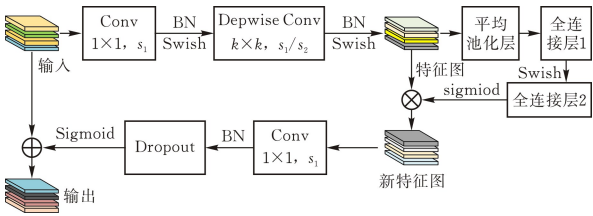


图4 MBConv 结构

EfficientNet 网络通过以上 9 个阶段操作提高了分类的平均准确率, 经多次训练发现 EfficientNetB7 的分类准确度高于其它基线网络, 在 α, β, γ 满足一定条件下, 通过训练, 找到最适合该雾天图像能见度分类的网络参数, 经过反复调参并对分类准确度结果对比, 最终选取了一组分类精度相对较高的复合参数: $\alpha = 1.18, \beta = 1.09, \gamma = 1.30$ 。

3 实验结果分析

本次实验采用 EfficientNetB7 网络对数据集进行训练, 并设置 VGG16、Darknet53、ResNet50 为对照网络, 通过反复调参找到最适合于本模型的网络复合参数。统计的实验数据见表 2。由数据看出, EfficientNetB7 网络的平均准确率最高, 比 ResNet50 平均准确率高 5%, 比 Darknet53 平均准确率高 9%, 比 VGG16 网络的平均准确率高 11%。

网络损失函数采用交叉熵损失函数, 计算雾天图片分类结果中能能见度实测结果与网络预测结果之间的误差, 算法如式(4):

$$H(p,q)=-\sum_{i=1}^np(x_i)\ln q(x_i) \tag{4}$$

其中, p 是实测数据分布,训练之前输入数据已经按照能见度分类,即 p 已经确定; q 为网络预测的概

率。损失函数反映了能见度仪实测概率分布与预测概率分布之间的误差程度,交叉熵损失值越小,表示网络效果越好。

表 2 能见度估测准确率统计

能见度等级/m	数据集/张	训练集/张	测试集/张	验证集/张	准确率/%			
					VGG16	Darknet53	ResNet50	Efficient-NetB7
0	230	165	41	24	100.00	100.00	100.00	100.00
50	505	363	90	52	85.73	86.00	90.32	98.23
100	420	302	75	43	80.23	90.21	92.00	95.00
150	400	288	72	40	75.36	87.03	91.06	96.32
200	415	298	74	43	78.65	79.20	89.29	95.33
250	448	322	80	46	89.32	85.38	90.00	97.10
300	523	376	94	53	80.10	82.01	89.12	96.72
350	789	568	142	79	76.07	81.06	87.32	96.30
400	750	540	135	75	68.02	77.60	89.60	85.98
450	800	576	144	80	85.32	88.38	88.17	89.59
500	509	366	91	52	87.00	86.23	89.35	95.28
600	500	360	90	50	89.10	86.00	87.20	96.00
700	550	396	99	55	86.98	85.69	80.01	87.63
800	600	432	108	60	85.09	87.00	89.80	98.00
900	621	447	111	63	87.32	90.10	96.00	98.63
1000	700	504	126	70	89.93	91.01	96.00	100.00
1100	720	518	129	73	91.20	90.63	95.69	100.00
1200	750	540	135	75	90.44	92.96	97.21	100.00
合计/平均	10230	7361	1836	1033	85.00	87.00	91.00	96.00

图 5 是四种网络的损失值曲线图,其中蓝色线是 Darknet53 网络的损失值曲线,橙色线是 VGG16 的损失值曲线,黄色线是 ResNet50 的损失值曲线,紫色线是 EfficientNetB7 的损失值曲线,由图 5 可以看出,EfficientNetB7 模型的损失值最小,Darknet53 模型损失值最大。迭代至 300 次以后,四种网络的损失值趋于稳定。EfficientNetB7 网络模型的损失值在 0.1 附近区间内波动,模型的损失值较低,训练及测试结果较好。

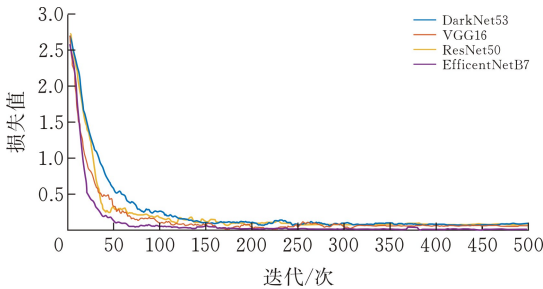


图 5 损失值曲线

将验证集图像输入训练好的 EfficientNetB7 网络,得到部分分类效果图(图 6),结果显示,网络分类的效果和能见度仪测试的实际结果一致。验证结果表明训练好的 EfficientNet 网络可以对该路段进行雾天能见度估测。从训练结果和验证结果上可以看出能见度估测的变化趋势,在视频开始拍摄和拍

摄结束时间段,雾天图像分类的平均准确率较高,中间时间段的平均准确率相对稍低,根据特征提取的算法原理分析,产生这种结果的原因是:在视频拍摄开始阶段雾天图像的特征较少,易于分类,卷积计算的速度也较快;视频拍摄结束阶段,雾天图像的清晰度较高,图像特征突出,故分类的平均准确率较高。



图 6 部分分类结果图

4 结论

基于传统的能见度测量仪器成本高和覆盖率低的问题,本文借助激光雷达能见度仪,提出一种 Ef-

ficientNet 分类网络对雾天图片能见度进行分类的方法,通过优化算法调参,找到最适用于雾天环境下道路图像分类网络的最优复合参数进行训练,通过四种常用分类网络对比,EfficientNet 分类网络的平均准确率达 96.00%,模型的损失值为 0.1,训练好的网络能在没有激光雷达能见度仪的情况下,在该拍摄路段估测出雾天道路视频的能见度值,降低了道路能见度检测成本,同时提高了道路能见度检测覆盖率。

[参 考 文 献]

[1] 黑龙江省气象局,中国气象局气象探测中心,山西省气象局,等. GB/T 35223—2017 地面气象观测规范气象能见度[S]. 北京:中国标准出版社,2017.

[2] 陈钊正,周庆逵,陈启美. 基于小波变换的视频能见度检测算法研究与实现[J]. 仪器仪表学报,2010,31(1):92-98.

[3] 韩婷婷. 基于深度学习的大雾短临预报研究[D].合肥:安徽大学,2019.

[4] SAKARIDIS C, DAI D, VAN GOOL L. Semantic foggy scene understanding with synthetic data[J]. International Journal of Computer Vision, 2018, 126(9):973-992.

[5] 苗开超,王传辉,张亚力,等. 基于 AlexNet 算法的道

路能见度估测方法[J]. 计算机与现代化,2019,15(6):87-91+103.

[6] 吴琪,汪小武,章军,等. 利用序数关系实现相对学习的雾图像能见度检测算法[J]. 计算机辅助设计与图形学学报,2020,32(12):1938-1947.

[7] RUSSAKOVSKY O, DENG J, SU H, et al. Imagenet large scale visual recognition challenge[J]. International journal of computer vision, 2015, 115(3):211-252.

[8] 刘博,王胜正,赵建森,等. 基于 Darknet 网络和 YOLOv3 算法的船舶跟踪识别[J]. 计算机应用,2019,39(06):1663-1668.

[9] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016:770-778.

[10] TAN M, LE Q. Efficientnet: Rethinking model scaling for convolutional neural networks[C]//International Conference on Machine Learning. PMLR, 2019:6105-6114.

[11] 赵凯旋,刘晓航,姬江涛.基于 EfficientNet 与点云凸包特征的奶牛体况自动评分方法[J].农业机械学报,2021,52(5):73,192-201.

[12] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8):2011-2023.

A Road Visibility Estimation Method Based on Efficientnet Network in Foggy Weather

WANG Niantao,WANG Shuqing,ZHANG Pengfei,DUN Weichao,HUANG Jianfeng

(School of Electrical and Electronic Engin.,Hubei Univ. of Tech.,Wuhan 430068,China)

Abstract: Aiming at the problems of high cost and low coverage in foggy visibility detection with traditional lidar visibility meters, an EfficientNet network model is proposed to estimate the visibility of foggy weather images taken by road cameras in real time. This method applies the EfficientNet network to the visibility estimation of foggy roads for the first time. First, the foggy video taken by the road camera is timed to collect images, and then the optimized EfficientNet deep learning network is used for training and classification under the deep learning framework. The experimental results show that the EfficientNet classification network has an average accuracy rate of 96% for visibility classification, which is higher than the VGG16, Darknet53, and ResNet50 comparison networks, which can realize the estimation of road visibility in foggy weather.

Keywords: Efficientnet;fog image;image classification;visibility

[责任编辑:张岩芳]