

# 基于改进MobileNet V2的茶叶病害识别方法

严春雨, 李 飞

贵州大学大数据与信息工程学院, 贵州 贵阳

收稿日期: 2022年6月10日; 录用日期: 2022年7月27日; 发布日期: 2022年8月4日

## 摘 要

自然场景下采集的茶叶病害样本背景复杂, 且存在类别样本数量不平衡的现象。结合茶叶病害特征, 提出一种基于改进MobileNet V2的茶叶病害识别方法。在MobileNet V2倒残差结构中引入坐标注意力机制, 使网络将注意力定位于目标区域, 减少无关信息的干扰, 有效地学习茶叶病害特征。同时, 将交叉熵损失替换为焦点损失, 解决茶叶病害样本类别不平衡导致网络训练效果不佳的问题。在茶叶病害数据集上进行验证实验, 实验结果表明, 改进后的MobileNet V2网络识别率达96.31%, 参数量仅为2.27 MB, 对比其他模型具有较高性价比。改进后的MobileNet V2网络能高效地对自然环境中茶叶病害进行识别, 为茶叶病害识别提供了新思路。

## 关键词

病害识别, MobileNet V2, 注意力机制, 损失函数

# Tea Disease Identification Method Based on Improved MobileNet V2

Chunyu Yan, Fei Li

College of Big Data and Information Engineering, Guizhou University, Guiyang Guizhou

Received: Jun. 10<sup>th</sup>, 2022; accepted: Jul. 27<sup>th</sup>, 2022; published: Aug. 4<sup>th</sup>, 2022

## Abstract

The background of tea disease samples collected in natural scenes is complex, and there is an imbalance in the number of categories of samples. Combined with the characteristics of tea diseases, a tea disease identification method based on improved MobileNet V2 is proposed. The coordinate attention mechanism is introduced into the MobileNet V2 inverted residual structure, so that the network can focus on the target area, reduce the interference of irrelevant information, and effectively learn the characteristics of tea diseases. At the same time, the cross-entropy loss is replaced

by the focal loss to solve the problem of poor network training effect caused by the imbalance of tea disease sample categories. The verification experiment is carried out on the tea disease data set. The experimental results show that the improved MobileNet V2 network has a recognition rate of 96.31% and a parameter size of only 2.27 MB, which is more cost-effective than other models. The improved MobileNet V2 network can efficiently identify tea diseases in the natural environment, which provides a new idea for the identification of tea diseases.

## Keywords

Disease Recognition, MobileNet V2, Attention Mechanism, Loss Function

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

茶叶作为一种重要的经济作物,在当前精准扶贫形势下是茶农脱贫的重要产业[1]。然而,茶树在生长过程中容易产生病害,导致茶叶产量和质量下降,给茶农造成严重的经济损失。因此,能准确识别茶叶病害,并及时采取相应防护措施,对减少茶叶产量损失和提高茶叶品质具有重要意义。

传统的植物病害鉴定主要依靠人工识别,易受主观性影响。随着计算机技术的发展,图像处理和机器学习方法已广泛应用于植物病害识别[2][3]。孙瑜[4]等人使用粒子群算法优化支持向量机(SVM),对黄花菜叶部病害识别精度为81.67%。杨涛[5]等人利用K-means聚类对麦冬叶部病斑做分割处理,利用SVM对提取后的病害特征进行识别,识别精度为93.30%。黄太远[6]等人对茶叶病害进行特征提取后通过SVM进行分类,识别精度为93.75%。传统的机器学习方法在识别植物病害时必须预先提取病害特征,由于病斑大小不一,且具有复杂的纹理和颜色特征,人工提取的特征可能无法反映植物病害的基本属性。上述方法对植物病害的识别率较低。

深度学习方法不需要人工提取特征,目前已在目标识别和其他领域得到了广泛的应用[7][8]。徐岩[9]等人在ResNet50基础上融合通道注意力和空间注意力,对马铃薯病害识别精度为93.86%。鲍文霞[10]等人基于VGG16网络,引入迁移学习和SK模块,对苹果病害识别精度为94.70%。黄林生[11]等人在ResNet网络中引入SE注意力和Inception模块,对农作物病害识别精度为95.62%。上述研究虽然均取得较好的效果,但网络参数较大,不利于实现病害识别的移动端部署。本文基于自然环境下的3种茶叶病害和健康叶制作茶叶病害数据集,在MobileNet V2[12]中嵌入坐标注意力(Coordinate Attention,简称CA)[13],通过赋予目标区域高权重以增强网络对茶叶病害的特征提取能力,在达到高识别率的同时不增加网络参数。引入焦点损失(Focal Loss)函数[14]解决茶叶病害图像数量不均衡的现象,使网络训练过程中聚焦于难分类样本,提高网络识别性能。

## 2. 图像采集与预处理

### 2.1. 图像采集

实验中使用的图像数据均采于南京雨花台风景区和贵阳羊艾茶园,共采集图像1900张,其中茶藻斑病414张、茶轮斑病503张、茶云纹叶枯病371张以及正常叶612张,部分样本图像如图1所示。



**Figure 1.** Tea diseased leaf diagram and normal leaf diagram

**图 1.** 茶叶病害叶片图与正常叶片图

## 2.2. 图像预处理

深度学习依赖于大量数据集, 为避免出现模型训练过拟合现象, 本实验对已采集茶叶病害图像进行数据增强操作。通过随机旋转、调整亮度、调整对比度、添加椒盐噪声等方法对病害图像进行扩充, 图像数据增强情况如表 1 所示。对数据增强后共 7150 张图像。采集的原始图像尺寸为  $2736 \times 3648$ , 由于本模型要求输入固定维度, 图像中像素过多会导致冗余信息增加, 因此本试验将图像尺寸统一为  $224 \times 224$  像素, 按 7:3 的比例随机划分训练集和验证集。

**Table 1.** Sample status after data enhancement

**表 1.** 数据增强后样本情况

类别	原始图像数量/张	数据增强后图像数量/张
茶藻斑病	414	1656
茶轮斑病	503	1803
茶云纹叶枯病	371	1855
正常叶	612	1836
总计	1900	7150

## 3. 基于 MobileNet V2 的茶叶病害识别模型

### 3.1. MobileNet V2

MobileNet V2 是 2019 年提出的一种轻量级网络, 沿用了 MobileNet V1 中提出的深度可分离卷积, 并引入倒置残差(Inverted residuals)结构进行改进, 以提高网络性能。MobileNet V2 整体网络结构如图 2 所示, 主要由一个卷积核大小为  $3 \times 3$  的卷积层、17 个含有倒置残差结构的瓶颈层(Bottleneck)以及一层  $7 \times 7$  大小的平均池化层(Avgpool)构成。

残差结构有助于解决网络梯度消失和梯度爆炸现象, 使网络更深的同时又能保证良好的性能。传统残差结构对特征通道维度先压缩后扩展, 而倒置残差结构首先通过  $1 \times 1$  逐点卷积(Pointwise Convolution, PWConv)对输入特征维度进行压缩, 其次送入深度卷积(Depthwise Convolution, DWConv)提取各通道的特征, 最后使用  $1 \times 1$  逐点卷积对特征维度进行压缩, 避免了传统残差结构先压缩特征维度, 导致上层输入通道过少, 网络仅对空间中低维特提取效果较好的现象。另外为防止高维特征压缩至低维时出现特征信息丢失, MobileNet V2 将倒置残差结构中降维卷积后的非线性激活函数 ReLU6 换为线性函数(Linear)输

出, 从而保留了特征信息的多样性, 增强了网络的特征提取能力。MobileNet V2 的倒置残差结构如图 3 所示。

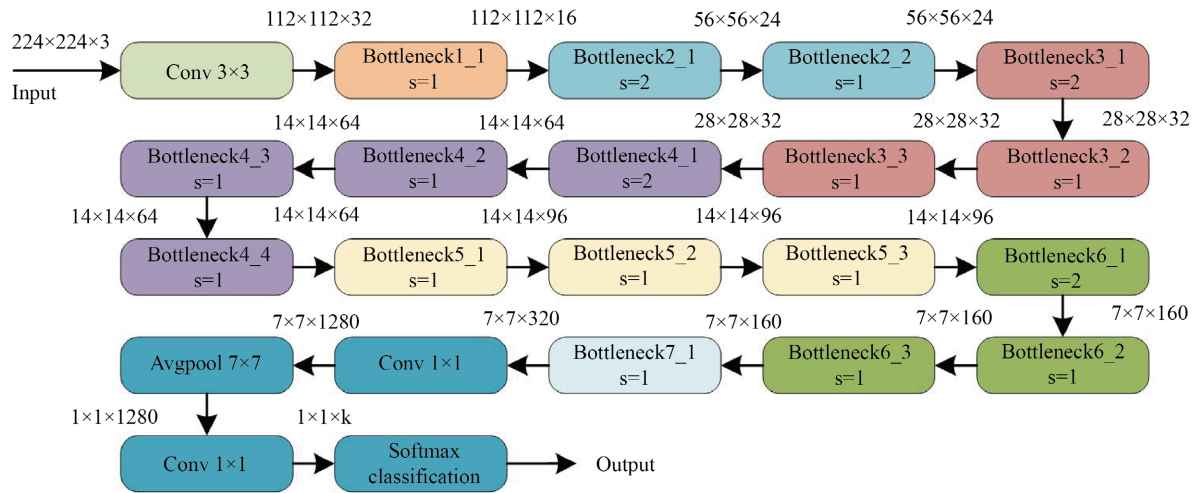


Figure 2. Structure of MobileNet V2 network  
图 2. MobileNet V2 网络结构

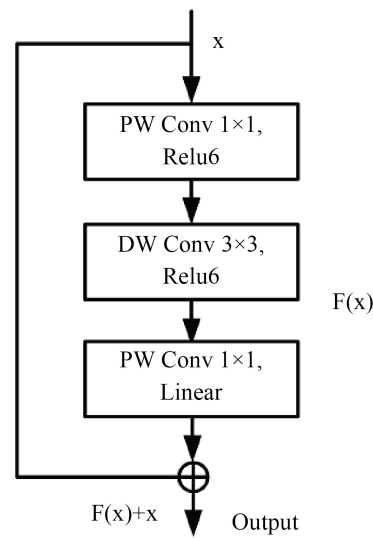


Figure 3. Structure of inverted residuals  
图 3. 倒置残差结构

### 3.2. 坐标注意力机制

注意力机制通过将网络注意力聚焦于目标区域, 减少无关信息的干扰, 提高视觉信息处理的性能。茶叶病害图像中含有丰富的目标与背景信息, 传统神经网络很难区分两者间的差异, 通过引入坐标注意力机制提高网络对茶叶病斑特征的敏感程度, 从注意力的分配方面改善网络性能, 坐标注意力机制结构如图 4 所示。

坐标注意力机制分为两部分, 第一部分为坐标信息嵌入: 将位置信息嵌入通道注意力, 对输入特征分别进行水平和垂直方向上的平均池化, 得到尺寸为  $C \times H \times 1$  和  $C \times 1 \times W$  的输出特征图,  $H$  和  $W$  分别

为输入特征图的高和宽,  $C$  为通道数。每个特征图都包括输入特征图沿着一个空间方向的远距离依存关系, 并保存沿着另一个空间方向的精确位置信息, 使网络能更准确地定位目标区域。

第二部分为坐标注意力生成: 首先, 将第一部分的两幅输出特征图通过 **Concat** 操作拼接为  $C \times 1 \times (H + W)$  的特征图。其次, 将特征图送入  $1 \times 1$  卷积层, 通道维度从  $C$  维降低为原来的  $C/r$ , 经过归一化处理和非线性激活后, 对获得的结果沿空间维度切分为两个张量, 尺寸分别为  $C/r \times H \times 1$  和  $C/r \times 1 \times W$ 。最后, 通过两个  $1 \times 1$  卷积层将特征图通道维度升至  $C$  维, 利用 **Sigmoid** 函数进行非线性激活。

最后通过乘法加权计算融合空间信息与通道特征, 完成坐标注意力的施加。

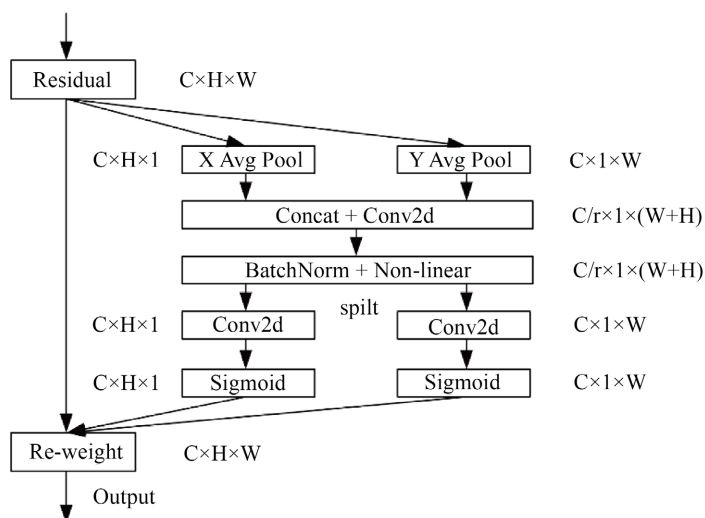


Figure 4. Structure of Coordinate Attention

图 4. 坐标注意力结构

引入坐标注意力的 MobileNet V2 倒置残差结构如图 5 所示, 通过在倒置残差结构模块中最后一层逐点卷积后嵌入坐标注意力机制, 使网络更关注目标区域, 有效区分目标与背景, 提高网络对病害的分类效果。

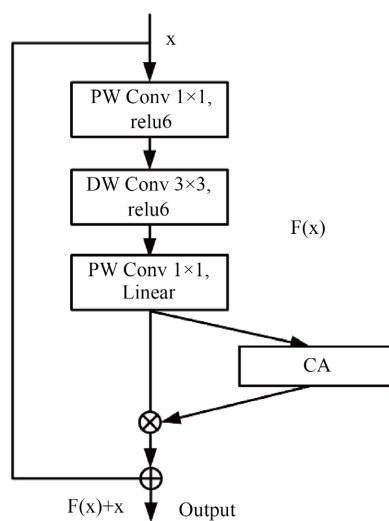


Figure 5. Inverted residuals of the embedded coordinate attention

图 5. 嵌入坐标注意力的倒置残差结构

### 3.3. 损失函数

交叉熵损失(CrossEntropy loss)常出现于分类问题, 为网络训练常用的损失函数, 公式为:

$$CE(p_t) = -\log(p_t) \quad (1)$$

式中,  $p_t$  为类别  $t$  的预测概率。由于交叉熵损失函数对所有样本损失的权值是相同的, 而本文实验使用的茶叶病害数据集存在类别样本不均衡的现象, 导致网络训练结果不佳, 因此本文实验利用焦点损失函数替换交叉熵损失函数。焦点损失是由交叉熵损失改进得到, 通过在交叉熵损失函数中引入  $\gamma$  和  $\alpha$  两个参数重新分配样本权重, 参数  $\alpha$  用于解决正负样本不均衡的问题, 参数  $\gamma$  用于解决难易样本不均衡的问题, 使网络在训练时更关注难分类的样本, 焦点损失函数公式为:

$$FL(p_t) = -\alpha_i(1 - p_t)^\gamma \log(p_t) \quad (2)$$

式中,  $p_t$  为类别  $t$  的预测概率,  $\gamma$  和  $\alpha$  为超参数, 如果  $\gamma = 0$ , 则  $FL(p_t)$  等价于  $CE(p_t)$ 。

## 4. 实验结果与分析

### 4.1. 实验环境

实验环境主要配置如表 2 所示。

**Table 2.** Experimental environment configuration  
**表 2.** 实验环境配置

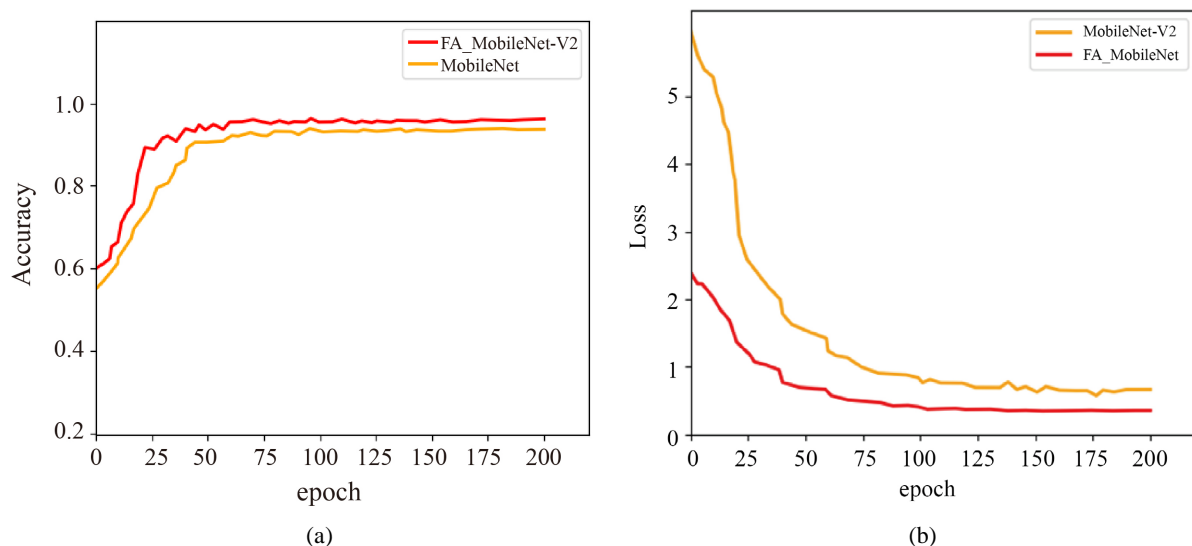
系统环境	Windows 10
GPU 版本	RTX3060
CUDA 版本	11.0
编程语言	Python
Pytorch 版本	1.7.1

### 4.2. 实验结果分析

本文实验使用尺寸为  $224 \times 224$  像素大小图像分别对 MobileNetV2 原网络和改进后的网络 FA\_MobileNet 进行训练和验证, 学习率和 batch\_size 分别设置为 0.001 和 16, 使用 Adam 优化器, 200 次迭代后两种网络验证集准确率和损失值变化曲线如图 6 所示, 可以明显看出改进后的网络在验证集上识别率高于 MobileNetV2, 收敛快的同时损失值更低, 验证了改进后网络的有效性。

为验证本文改进方法的有效性, 在 MobileNet V2 中使用不同改进策略组合, 实验结果如表 3 所示。可以看出, 在 MobileNet V2 网络中分别引入坐标注意力机制和焦点损失函数后对茶叶病害数据集的识别准确率较原始网络均有所提升, 分别提升了 1.55% 和 0.96%。融合两种改进策略的 FA\_MobileNet 网络性能优于单独使用一种策略的网络性能, 对茶叶病害的识别准确率提高了 2.39%, 网络参数量仅增加 0.04MB。验证了嵌入坐标注意力和焦点损失函数对网络性能提升的有效性。

从网络竞争力角度出发, 本文实验选取 AlexNet [15]、VGG16 [16] 和 GoogLeNet [17] 这三种卷积神经网络与 FA\_MobileNet 网络进行对比试验, 不同网络比较结果如表 4 所示。可以明显看出 FA\_MobileNet 网络对茶叶病害的识别效果最好, 识别率达 96.31%, 较 AlexNet、VGG16 和 GoogLeNet 分别高出 8.66%、3.67% 和 2.16% 个百分点, 网络参数量仅为 2.27 MB。实验表明, FA\_MobileNet 网络相较于其他网络具有更好的竞争优势, 能够满足移动端对复杂背景下茶叶病害识别的要求。



**Figure 6.** Variation curve of network validation accuracy and loss. (a) Variation curve of the accuracy; (b) Variation curve of loss

**图 6.** 网络验证集准确率和损失值变化曲线。(a) 准确率变化曲线; (b) 损失值变化曲线

**Table 3.** Ablation test

**表 3.** 消融试验

组 Group	坐标注意力 CA	焦点损失函数 Focal Loss	准确率 Accuracy/%	参数量 Parameters/MB
1	-	-	94.02	2.23
2	√	-	95.57	2.27
3	-	√	94.98	2.23
4	√	√	96.41	2.27

注: -表示未使用该方法, √表示使用该方法。

**Table 4.** Comparison of different network performance

**表 4.** 不同网络性能比较

Model	Accuracy/%	Parameters/MB
AlaxNet	87.65	57.03
VGG16	92.69	134.29
GoogLeNet	94.15	5.6
FA_MobileNet	96.31	2.27

## 5. 结论

本文针对复杂背景下茶叶病害识别困难的问题,提出一种轻量级茶叶病害识别网络。在 MobileNet V2 网络中引入坐标注意力,同时考虑通道关系和位置信息,达到减少背景信息干扰的效果并使网络更准确定位至目标区域。利用焦点损失函数替换交叉熵损失函数,有效解决茶叶病害样本不均衡的问题,提高网络识别性能。与其他网络进行对比,实验表明本文改进后的网络具有识别率高,网络参数量小的优势,识别率和参数量分别为 96.31% 和 2.27 MB,具有较强的实用性。

## 参考文献

- [1] 王娜. 乡村振兴背景下茶叶产业发展研究[J]. 福建茶叶, 2022, 44(7): 12-14.
- [2] 王聘, 柴秀娟. 机器学习在植物病害识别研究中的应用[J]. 中国农机化学报, 2019, 40(9): 171-180
- [3] 刘翠翠, 杨涛, 马京晶, 等. 基于 PCA-SVM 的麦冬叶部病害识别系统[J]. 中国农机化学报, 2019, 40(8): 132-136.
- [4] 孙瑜, 张永梅, 武玉军. 基于粒子群算法和支持向量机的黄花菜叶部病害识别[J]. 中国农学通报, 2022, 38(8): 135-140.
- [5] 杨涛, 雷进, 朱皓睿, 等. 基于图像特征融合的麦冬叶部病害识别[J]. 湖北农业科学, 2021, 60(7): 135-138+144.
- [6] 黄太远, 李旺, 邱亚西, 等. 基于形状特征和 SVM 多分类的铜仁地区茶叶病害识别研究[J]. 种子科技, 2020, 38(6): 7-9.
- [7] 刘钧文, 钟德星, 邵会凯, 等. 基于边缘计算的紧致化学纹识别系统[J]. 中国科学: 技术科学, 2022, 52(5): 704-712.
- [8] 丁永军, 张晶晶, 李民赞. 基于卷积胶囊网络的百合病害识别研究[J]. 农业机械学报, 2020, 51(12): 246-251+331.
- [9] 徐岩, 李晓振, 吴作宏, 等. 基于残差注意力网络的马铃薯叶部病害识别[J]. 山东科技大学学报(自然科学版), 2021, 40(2): 76-83.
- [10] 鲍文霞, 吴刚, 胡根生, 等. 基于改进卷积神经网络的苹果叶部病害识别[J]. 安徽大学学报(自然科学版), 2021, 45(1): 53-59.
- [11] 黄林生, 罗耀武, 杨小冬, 等. 基于注意力机制和多尺度残差网络的农作物病害识别[J]. 农业机械学报, 2021, 52(10): 264-271.
- [12] Sandler, M., Howard, A., Zhu, M., et al. (2018) MobileNet V2: Inverted Residuals and Linear Bottlenecks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 4510-4520. <https://doi.org/10.1109/CVPR.2018.00474>
- [13] Hou, Q., Zhou, D. and Feng, J. (2021) Coordinate Attention for Efficient Mobile Network Design. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Nashville, 20-25 June 2021, 13713-13722. <https://doi.org/10.1109/CVPR46437.2021.01350>
- [14] Lin, T.Y., Goyal, P., Girshick, R., et al. (2017) Focal Loss for Dense Object Detection. *Proceedings of the IEEE International Conference on Computer Vision*, Venice, 22-29 October 2017, 2980-2988. <https://doi.org/10.1109/ICCV.2017.324>
- [15] Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2012) Imagenet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, **60**, 84-90. <https://doi.org/10.1145/3065386>
- [16] Simonyan, K. and Zisserman, A. (2014) Very Deep Convolutional Networks for Large-Scale Image Recognition.
- [17] Szegedy, C., Liu, W., Jia, Y., et al. (2015) Going Deeper with Convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, 7-12 June 2015, 1-9. <https://doi.org/10.1109/CVPR.2015.7298594>