程序报告

姓名：叶睿宸　　**学号：**22551338　　**学院（系）专业：**软件学院人工智能

1 算法描述

1.1 算法历史演变

MobileNet 是一种轻量级的深度学习网络架构，专为移动设备和嵌入式设备设计，具有高效的计算性能和较低的资源消耗。

MobileNet 有如下特点：

1.轻量化设计：MobileNet 使用深度可分离卷积（Depthwise Separable Convolution）来减少计算量和参数量。并且深度可分离卷积将标准卷积分解为两个步骤（Depthwise Convolution 和 Pointwise Convolution），这种分解大幅减少了计算复杂度：

2.可调的宽度和分辨率：MobileNet 引入了两个超参数：宽度乘子（Width Multiplier）和分辨率乘子（Resolution Multiplier），通过调整这两个参数，可以在精度和效率之间进行权衡。

3.适用于移动设备：MobileNet 的设计目标是减少模型大小和计算量，使其能够在资源受限的设备（如手机、嵌入式设备）上运行。

MobileNet 的经历了三个版本的演进，演变过程如下

1.MobileNetV1：最早的版本，提出了深度可分离卷积的概念。通过减少计算量和参数量，显著提高了效率。

2.MobileNetV2：引入了倒残差结构（Inverted Residuals）和线性瓶颈（Linear Bottleneck）。倒残差结构通过跳跃连接减少信息丢失。线性瓶颈通过限制激活函数的使用，减少特征信息的损失。在相同计算量下，精度显著提升，且更适合迁移学习任务

3.MobileNetV3：结合了神经架构搜索（NAS）和网络优化技术。引入了注意力机制（SE 模块）以进一步提升性能。并且提供了两种版本：MobileNetV3-Small 和 MobileNetV3-Large，分别适用于不同的计算资源需求。

MobileNet 主要应用在图像分类、目标检测（如 SSD-MobileNet）、语义分割（如 DeepLab-MobileNet）、人脸识别、手势识别、嵌入式设备上的实时推理任务。

MobileNet 的优缺点：

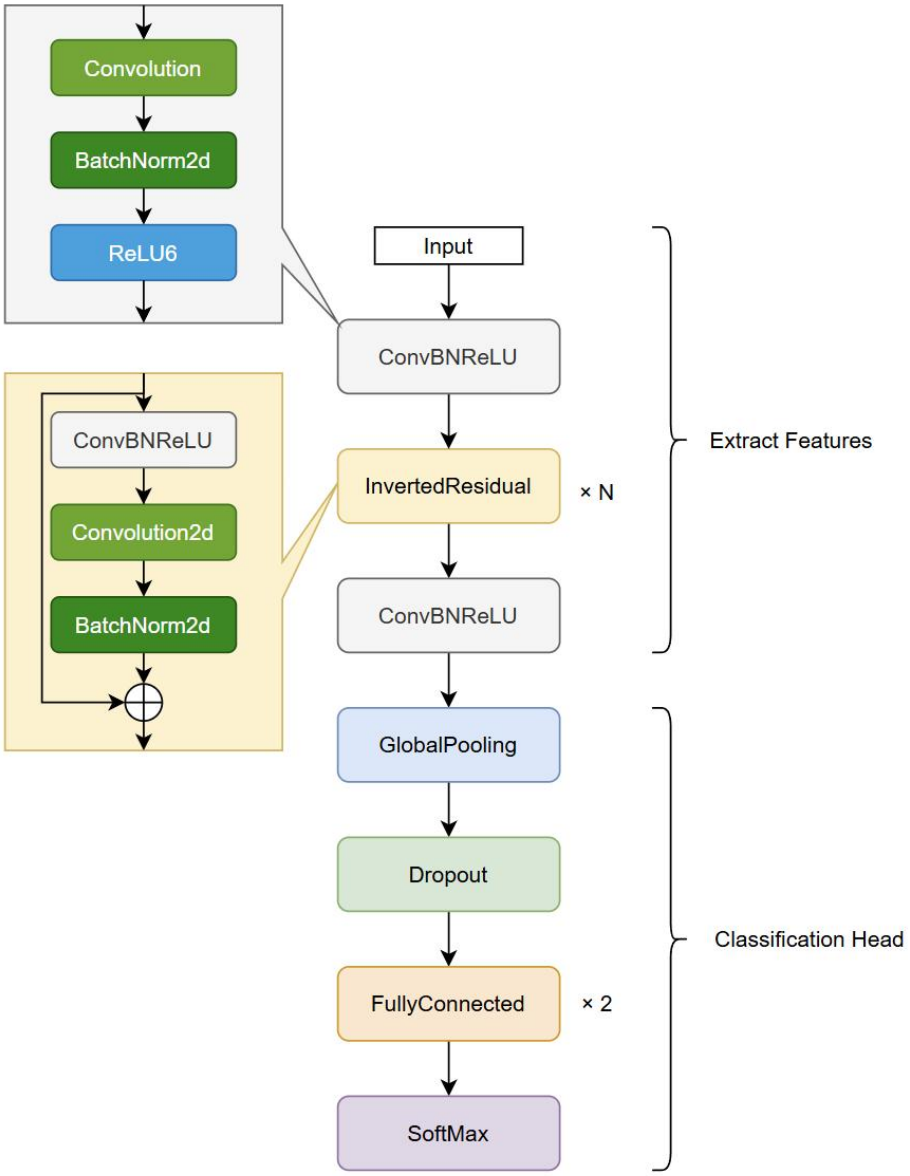
优点：参数量小，计算效率高。适合移动设备和嵌入式设备。易于迁移到其他任务（如目标检测、分割）。

缺点：相较于更大的网络（如 ResNet、EfficientNet），精度可能略低。对高分辨率图像或复杂任务的表现有限。

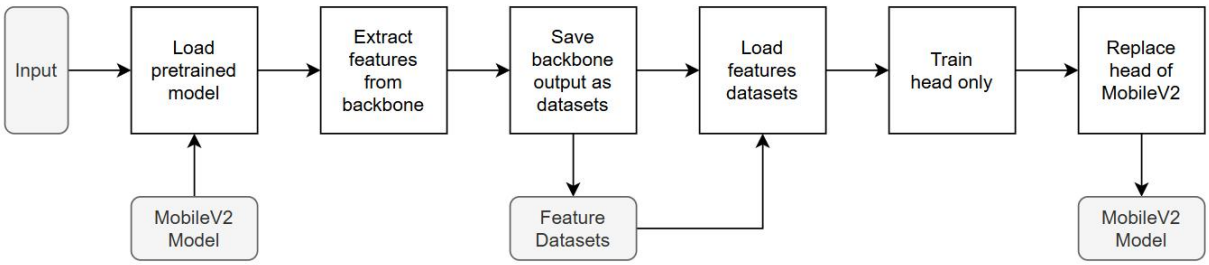
MobileNet 是一种高效的深度学习网络架构，专为资源受限的设备设计。其轻量化和高效性使其成为移动端和嵌入式设备深度学习任务的首选。

1.1 算法流程图

如下展示的是 MobileNetV2 的模型架构图：



如下展示的是模型训练流程图：



2 算法性能分析

2.1 计算效率

MobileNetV2 使用深度可分离卷积和倒残差结构，显著减少了参数量和计算量。

训练方式，通过加载预训练权重（ImageNet 数据上预训练的 MobileNetv2），利用预训练权重的 backbone 部分得到输入训练集的 features。这种方式可以避免，每次训练输入都需要重复的经过相同的 backbone 得到相同的 features，大大减少了无用的计算耗费。因此这种方式下，原先的训练集仅使用到一次，后续都使用 backbone 的输出 features 作为训练集。

由于是采用分离训练的方式，因此需要冻结 MobileNet 除 classification head 之外的部分，只训练 classification head。本身是在预训练模型 MobileNetv2 上实现的迁移学习，再加上只对头部一小部分参数进行微调，因此这个训练过程是非常轻量级和迅速的，因此我这边选择在 CPU 上运行，而参数或数据集直接缓存到内存中。

这种方式非常适合在资源受限的场景上进行训练，这边在训练集上训练 100 个 epoch 只需要耗时 10 分钟左右。采用这种轻量级的微调方式，我可以在有限的时间内尝试更多的超参数，使得模型的泛化能力能够进一步提升，找到最佳的训练超参。

在模型的推理阶段同样可以采用两阶段的方式，测试集预先输入网络的 backbone 得到 feature 数据集，后续再将 feature 数据集作为 head 的输入，得到最终的分类结果。由于推理过程只需要进行一轮，因此采用两阶段和一阶段的方式在计算效率上没有较大的差异。

但是如果是在训练过程中，想要验证下当前模型在测试集上的泛化能力，这个过程会涉及多轮相同测试集的推理，因此在训练过程中的验证阶段采用两阶段的推理方式可以大大降低模型的计算效率上。

2.2 计算精度

由于通过加载预训练权重（ImageNet 数据上预训练的 MobileNetv2）并进行微调，模型在垃圾分类任务中可以达到较高的分类精度。因为模型已经具备了较好的语义理解和特征捕获能力，因此在下游任务上只需要进行 few-shot 的训练就可以达到较好的精度。可以发现训练集的样本数相较于 Imagenet-1k 而言是非常小的。

为了防止模型出现过拟合的情况，这边通过逐步提高正则化参数的大小，限制模型的复杂度，从而使得模型在训练集和测试集上的提升进程相当。并且采用较小的学习率配合余弦退火的优化器算法，能让模型在训练过程中稳步提升，有效的避免了模型在梯度下降过程中，出现震荡的情况。

在训练过程中，我选择模型在测试集上的准确度作为最后模型权重抉择的依据，因为在测试集上较高的准确度就意味着模型会有更好的泛化能力。原代码采用保留所有 ckpt 的方式，并且记录训练过程中每个 ckpt 的训练损失、验证损失等参数，结合两者去决定模型权重的保留，这种方式不仅会消耗大量的存储空间，且需要人工的查看日志筛选最优 ckpt，这种方式是比较费精力的。我这边采用每轮训练记录最优准确度，通过每轮验证精度与最优精度的对比，选择是否覆盖之前保存的 ckpt，这种方式仅需保留一个模型权重文件，且无需人工的参与。

原模型架构不包含 dropout 层，在训练过程中，发现对模型泛化能力的提升有限，最终做多种训练超参数的组合下在测试集上的最高精度只能达到 91.9 %。为了进一步提升模型的泛化能力，我在原有的 head 模块中，增加了一个 dropout 层，并且将原有的 fully connected 的层拆成了两部分。两种变革的目的都是为了进一步提升模型的泛化能力，最终采用优化后

的模型架构进行训练，最高测试集精度可以达到 93.5 %。

3 研究展望

3.1 模型优化

采用混合精度训练：使用 float16 和 float32 的混合精度训练，可以进一步提升训练效率并减少显存占用。

模型压缩：结合剪枝（Pruning）和量化（Quantization）技术，进一步减少模型大小和推理时间。适合在超低功耗设备上部署。

改进网络结构：引入注意力机制（如 SE 模块或 CBAM）增强特征提取能力。借鉴 MobileNetV3 的改进（如 NAS 和 h-swish 激活函数）进一步提升性能。

3.2 数据增强与迁移学习

数据增强：使用更丰富的数据增强技术（如随机裁剪、颜色抖动、混合增强）提升模型的鲁棒性。

迁移学习：在更大规模的垃圾分类数据集上预训练模型，然后迁移到特定任务中。结合领域自适应技术，提升模型在不同场景下的泛化能力。

3.3 多任务学习

联合学习：将垃圾分类与其他相关任务（如垃圾检测、垃圾分割）联合训练，提升模型的多任务能力。

多模态学习：融合图像和其他模态（如文本描述、传感器数据）进行分类，提升模型的表现。

3.4 部署与应用

边缘设备部署：将模型部署到边缘设备（如树莓派、智能垃圾桶）上，实现实时垃圾分类。

云端与边缘协同：结合云端计算和边缘设备，利用云端的强大计算能力进行模型更新和优化，边缘设备负责实时推理。

3.5 可解释性与鲁棒性

模型可解释性：使用可视化技术（如 Grad-CAM）分析模型的决策过程，提升模型的透明性。

鲁棒性研究：研究模型对噪声、模糊图像或对抗样本的鲁棒性，提升模型在真实场景中的可靠性。

参考文献

[1] Howard, A. G., Zhu, M., Chen, B., et al. (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. arXiv preprint arXiv:1704.04861.

[2] Sandler, M., Howard, A., Zhu, M., et al. (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 4510-4520.

[3] Howard, A., Pang, R., Adam, H., et al. (2019). Searching for MobileNetV3. Proceedings of the IEEE International Conference on Computer Vision (ICCV), 1314-1324.

[4] Loshchilov, I., & Hutter, F. (2017). SGDR: Stochastic Gradient Descent with Warm Restarts. arXiv preprint arXiv:1608.03983.

[5] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770-778.

[6] Kingma, D. P., & Ba, J. (2015). Adam: A Method for Stochastic Optimization. arXiv preprint arXiv:1412.6980.

[7] Hinton, G., Srivastava, N., & Krizhevsky, A. (2012). Improving Neural Networks by Preventing Co-adaptation of Feature Detectors. arXiv preprint arXiv:1207.0580.

[8] Shorten, C., & Khoshgoftaar, T. M. (2019). A Survey on Image Data Augmentation for Deep Learning. Journal of Big Data, 6(1), 1-48.

[9] Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). How Transferable Are Features in Deep Neural Networks?. Advances in Neural Information Processing Systems (NeurIPS), 27, 3320-3328.

[10] Perez, L., & Wang, J. (2017). The Effectiveness of Data Augmentation in Image Classification Using Deep Learning. arXiv preprint arXiv:1712.04621.

[11] Huawei Technologies Co., Ltd. (2020). MindSpore: An AI Computing Framework. Online Documentation.

[12] Zhang, Y., Wang, X., & Li, J. (2021). MindSpore: A Unified AI Framework for Model Development and Deployment. Proceedings of the International Conference on Artificial Intelligence and Big Data (ICAIBD), 1-6.

[13] Han, S., Pool, J., Tran, J., & Dally, W. (2015). Learning Both Weights and Connections for Efficient Neural Networks. Advances in Neural Information Processing Systems (NeurIPS), 28, 1135-1143.

[14] Jacob, B., Kligys, S., Chen, B., et al. (2018). Quantization and Training of Neural Networks for Efficient Integer-Arithmetic-Only Inference. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2704-2713.

[15] Frankle, J., & Carbin, M. (2019). The Lottery Ticket Hypothesis: Finding Sparse, Trainable Neural Networks. Proceedings of the International Conference on Learning Representations (ICLR).

[16] SHI, Cuiping, et al. A waste classification method based on a multilayer hybrid convolution neural network. Applied Sciences, 2021, 11.18: 8572.

[17] MALIK, Meena, et al. Waste classification for sustainable development using image recognition with deep learning neural network models. Sustainability, 2022, 14.12: 7222.