

# Face Recognition method survey:4

首先对GhostNet进行一个介绍：

GhostNet: More Features from Cheap Operations

## 论文出发点或背景

因为有限的内存和计算资源，在嵌入式设备中部署卷积神经网络(CNNs)是困难的。特征图的冗余是那些成功CNNs的一个重要特征。这篇论文提出了一个新颖的Ghost模型，从廉价的运算中得到更多的特征图。

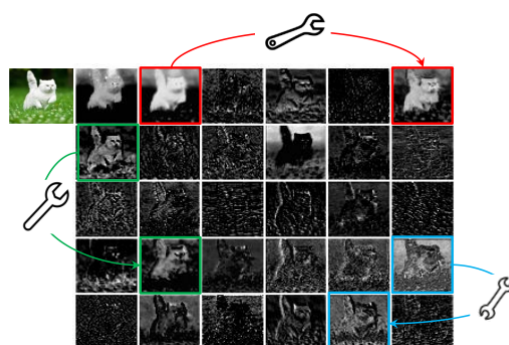


Figure 1. Visualization of some feature maps generated by the first residual group in ResNet-50, where three similar feature map pair examples are annotated with boxes of the same color. One feature map in the pair can be approximately obtained by transforming the other one through cheap operations (denoted by spanners).

## 论文创新思路

在训练有素的深度神经网络的特征图中，丰富甚至是冗余的信息常常保证了对输入数据的全面理解

我们不是避免冗余的特征图，我们倾向于用一个低成本的方法来使用这些冗余的特征图。

在深度神经网络中，一个普通的卷积层会被分成两个部分。第一个部分涉及普通的卷积，但是它们总的数量会得到严格地控制。根据从第一个部分得到的本征特征图，之后使用一系列简单的线性运算去生成更多的信息。

## 论文方法介绍

任意卷积层生成n个特征图的操作可以表示为

$$Y = X * f + b$$

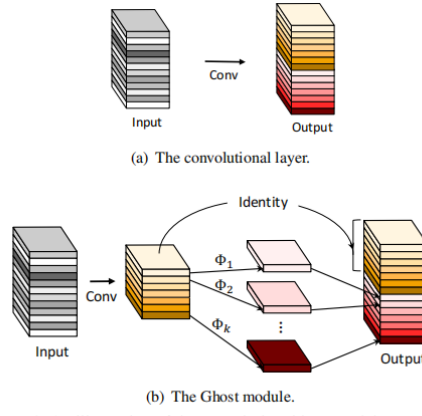


Figure 2. An illustration of the convolutional layer and the proposed Ghost module for outputting the same number of feature maps.  $\Phi$  represents the cheap operation.

我们指出，使用大量的FLOPs和参数来一个个的生成冗余的特征图是不必要的。假设用一些廉价的变换产生的输出特征图是一些本征特征图的“重影”，这些本征特征图尺寸通常很小而且由原始卷积核产生。特别地， $m$ 个本征映射图 $Y' \in R^{h' \times w' \times h}$ 使用基本的卷积产生：

$$Y' = X * f'$$

我们提出在本征特征图上 $Y'$ 上，根据下面的函数，使用一系列廉价的线性运算来生成 $s$ 个重影特征：

$$y_{ij} = \Phi_{i,j}(y'_i), \quad \forall i = 1, \dots, m, j = 1, \dots, s$$

这个提出来的Ghost模块和已有的高效的卷积方案有着很大的不同。i) 和这些广泛使用 $1 \times 1$ pointwise的卷积相比，Ghost模块中的基本的运算可以有自定义的内核模块。ii) 已有的方法采用pointwise卷积去跨通道处理特征，并且之后采用depthwise卷积去处理空间信息。相反，Ghost模块采用原始的卷积首先生成少量的本征特征图，然后利用廉价的线性运算去扩充特征和增加通道。iii) 这些去处理每个特征图的运算受之前高效架构里的depthwise运算或者shift运算限制，然而Ghost模型中的线性运算有这个很大的多样性。iv) 此外，恒等映射映射等价于Ghost模块中的线性变化，去保存本征特征图。

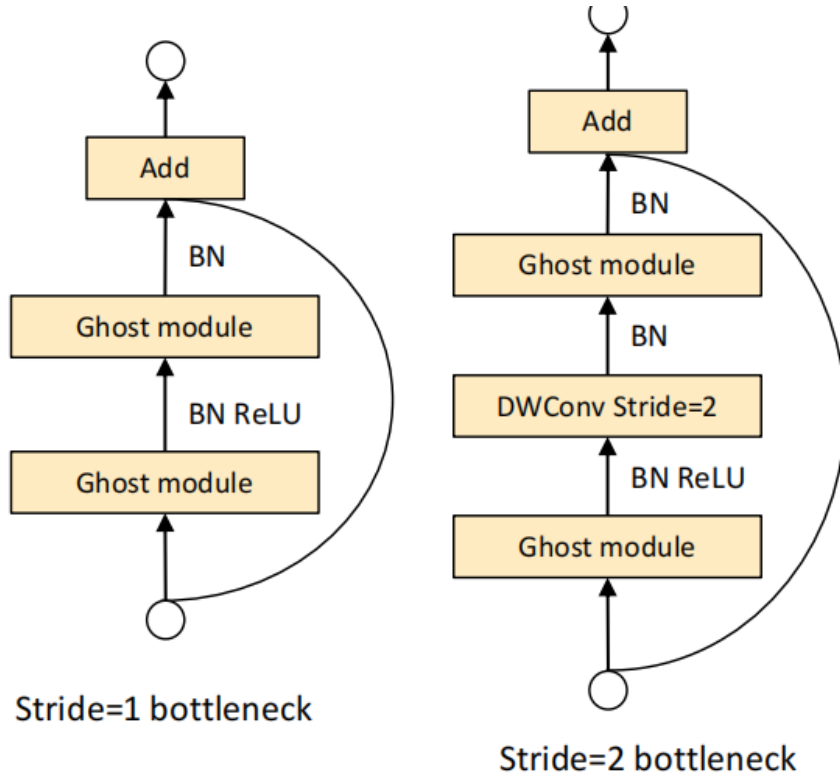


Figure 3. Ghost bottleneck. Left: Ghost bottleneck with stride=1; right: Ghost bottleneck with stride=2.

第一个Ghost模块充当着增加通道数和膨胀层的功能，我们指定输出和输入通道数之间的比例为膨胀比。第二个Ghost模块减少通道数量来匹配捷径通道，这个捷径连接了两个Ghost模块的输入和输出。

Table 1. Overall architecture of GhostNet. G-bneck denotes Ghost bottleneck. #exp means expansion size. #out means the number of output channels. SE denotes whether using SE module.

Input	Operator	#exp	#out	SE	Stride
$224^2 \times 3$	Conv2d $3 \times 3$	-	16	-	2
$112^2 \times 16$	G-bneck	16	16	-	1
$112^2 \times 16$	G-bneck	48	24	-	2
$56^2 \times 24$	G-bneck	72	24	-	1
$56^2 \times 24$	G-bneck	72	40	1	2
$28^2 \times 40$	G-bneck	120	40	1	1
$28^2 \times 40$	G-bneck	240	80	-	2
$14^2 \times 80$	G-bneck	200	80	-	1
$14^2 \times 80$	G-bneck	184	80	-	1
$14^2 \times 80$	G-bneck	184	80	-	1
$14^2 \times 80$	G-bneck	480	112	1	1
$14^2 \times 112$	G-bneck	672	112	1	1
$14^2 \times 112$	G-bneck	672	160	1	2
$7^2 \times 160$	G-bneck	960	160	-	1
$7^2 \times 160$	G-bneck	960	160	1	1
$7^2 \times 160$	G-bneck	960	160	-	1
$7^2 \times 160$	G-bneck	960	160	1	1
$7^2 \times 160$	Conv2d $1 \times 1$	-	960	-	1
$7^2 \times 960$	AvgPool $7 \times 7$	-	-	-	-
$1^2 \times 960$	Conv2d $1 \times 1$	-	1280	-	1
$1^2 \times 1280$	FC	-	1000	-	-

## 实际效果

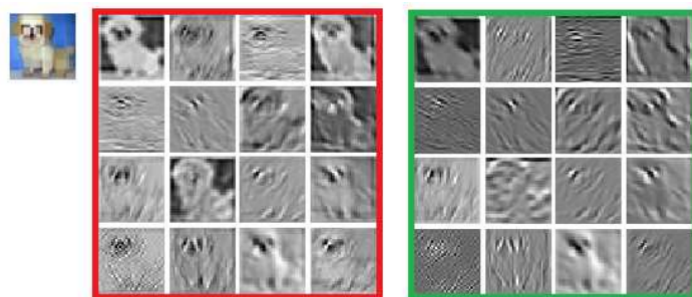


Figure 4. The feature maps in the 2nd layer of Ghost-VGG-16. The left-top image is the input, the feature maps in the left red box are from the primary convolution, and the feature maps in the right green box are after the depthwise transformation.

Table 8. Results on MS COCO dataset.

Backbone	Detection Framework	Backbone FLOPs	mAP
MobileNetV2 1.0× [44]	RetinaNet	300M	26.7%
MobileNetV3 1.0× [20]		219M	26.4%
GhostNet 1.1×		164M	26.6%
MobileNetV2 1.0× [44]	Faster R-CNN	300M	27.5%
MobileNetV3 1.0× [20]		219M	26.9%
GhostNet 1.1×		164M	26.9%

## 个人理解

GhostNet 给我最大的收获就是将突破的点放在了可视化得到的特征图上，之前的模型修改要是基于数学理论进行低秩近似，要么是通过视觉原理归纳出来一些宏观上的设计理念，然后去比较一些指标。GhostNet的话就是先去看看卷积得到的到底是些什么，然后得到的特征图有什么特点，然后发现有冗余的特征，然后通过一些更为cheap的操作去代替原有的标准卷积。

## GhostFaceNet

作者将Ghost操作融入了人脸识别，做的工作量不是很多，主要是将轻量化分类网络拿过来给face recognition任务使用，在原有的Ghost模块作为baseline的前提下，进行了一些修改，但是创新点都不是很多，比如给Ghost加了激活函数，SE模块之前在很多任务中都表现不错，所以给Ghost加了SE。其次就是改进了GDC部分

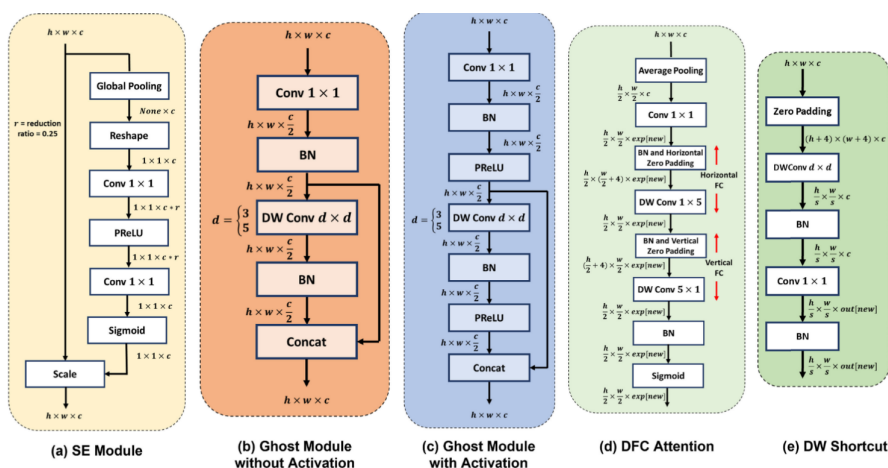


FIGURE 2. (a) SE Module. (b) Ghost module structure without PReLU activation function. (c) Ghost module structure with PReLU activation function. (d) DFC attention module. (e) Depthwise shortcut.

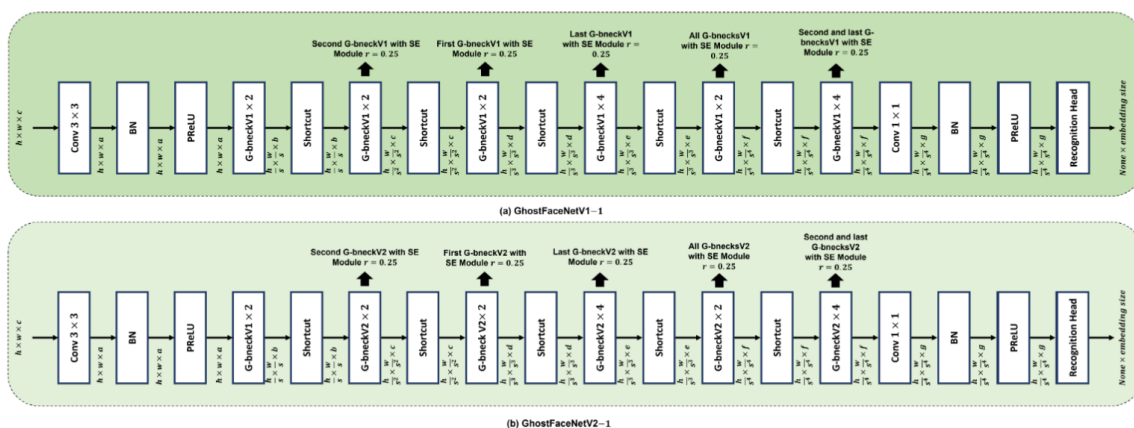
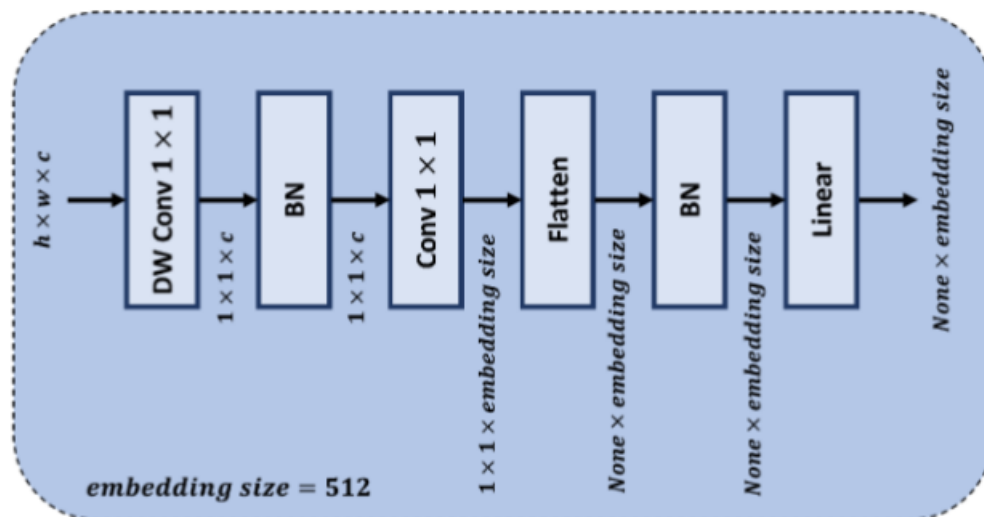


FIGURE 4. (a) Proposed GhostFaceNetV1-1 architecture. (b) Proposed GhostFaceNetV2-1 architecture.

文章改进了GDC模块，将GDC模块作为GhostNet中全连接层的替代。GDC 由  $7 \times 7$  内核大小的 GDC 层以及批量归一化层组成。然后使用大小为  $1 \times 1 \times$  嵌入大小的卷积层来创建所需的嵌入向量（大小）。然后将展平层、批量归一化层和线性激活函数添加到顶部以产生紧凑的 512 维嵌入向量（大小）。



**FIGURE 5.** Modified GDC, Recognition Head.