

重庆邮电大学

CHONGQING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS

# 专业学位硕士论文

MASTER THESIS FOR PROFESSIONAL DEGREE



## 论文题目

重庆邮电大学学位论文

## 格式模板

## 学科专业

# 电子科学与技术

学 号

**S20202222**

作者姓名

张三

## 指导教师

李四

学院

光电工程学院/国际半导体学院

学校代码	10617	UDC	xxxxxx
分 类 号	xxxxxx	密级	

学 位 论 文

重庆邮电大学学位论文格式模板

某 某

指导教师	某某某	教 授
	某 某	副教授

申请学位级别	博士	学科专业	XXXX
专业学位领域	XXXXXX		
答辩委员会主席	某某某	教授	论文答辩日期 2021 年 5 月 20 日
学位授予单位和日期	重庆邮电大学	2021 年 6 月	

**Dissertation Templeplate for Master Degree of  
Engineering in CHONGQING UNIVERSITY OF  
POSTS AND TELECOMMUNICATIONS**

A Doctoral Dissertation Submitted to  
Chongqing University of Posts and Telecommunications

Discipline	XXXX
Student ID	XXXX
Author	XXXX
Supervisor	XXXX
School	XXXX

# 重庆邮电大学

## 学位论文独创性声明

本人郑重声明：所呈交的学位论文，是本人在导师指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文中不包含其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在论文中以明确方式标明。本人完全知晓本声明的法律后果由本人承担。

学位论文作者签名：

日期： 年 月 日

# 重庆邮电大学

## 学位论文使用授权书

本人同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。

本学位论文属于：

☐ 公开论文

☐ 涉密论文，保密\_\_\_\_年，过保密期后适用本授权书。

（请在以上方框内选择打“√”）

作者签名：

导师签名：

日期： 年 月 日

## 摘 要

学位论文是研究生从事科研工作的成果的主要表现，集中表明了作者在研究工作中获得的新发明、新理论或新见解，是研究生申请硕士或博士学位的重要依据，也是科研领域中的重要文献资料和社会的宝贵财富。

为进一步规范我校研究生学位论文撰写格式，提高研究生学位论文质量，参照国家标准《学位论文编写规则》(GB/T 7713.1-2006)，结合我校实际，制定本模板。

**关键词：**学位论文，撰写规范，论文模板，重庆邮电大学

## ABSTRACT

Dissertation /Thesis is postgraduate ' s main academic performance to display her/his works of scientific research, which shows the author ' s new invention, new theory or new opinion in her/his research. It is the crucial document for the graduate students to apply for degree, and it is also the important scientific research literature and the valuable wealth of society.

In order to further standardize the format of dissertation/thesis writing and improve graduate dissertation/thesis quality, this temolate is formulated with reference to the national standard "Rules for Dissertation Writing" (GB/T 7713.1-2006) and the reality of CQUPT.

**Keywords:** Dissertation/Thesis, Writing Specification, Thesis Template, Chongqing University of Posts and Telecommunications

## 目 录

摘 要 .....	I
ABSTRACT .....	II
图目录 .....	V
表目录 .....	VI
主要符号表 .....	VII
缩略词表 .....	VIII
第 1 章 绪论 .....	1
1.1 研究背景及意义 .....	1
1.2 国内外研究现状 .....	2
1.2.1 基于反向传播的显著图解释技术 .....	2
1.2.2 基于类激活映射的显著图解释技术 .....	3
1.2.3 基于扰动的显著图解释方法 .....	3
1.3 论文研究的主要内容 .....	4
1.4 论文组织结构 .....	4
第 2 章 相关理论介绍 .....	5
2.1 引言 .....	5
2.2 .....	5
2.3 字数要求 .....	5
2.3.1 硕士论文要求 .....	5
2.3.2 博士论文要求 .....	5
2.4 字体和段落 .....	5
2.5 章节标题 .....	6
2.5.1 三级标题 .....	6
2.6 本章小结 .....	6
第 3 章 基于输入图片多尺寸放大的卷积神经网络显著图解释方法 .....	7
3.1 本章引言 .....	7
3.2 问题描述和研究思路 .....	7
3.3 基于输入图片多尺寸放大的卷积神经网络显著图解释方法 .....	8
3.3.1 生成高分辨率掩膜 .....	8
3.3.2 掩膜优化 .....	10
3.3.3 图 .....	10

3.3.4 表 .....	12
3.4 公式格式 .....	13
3.5 印制要求 .....	14
3.6 本章小结 .....	14
第 4 章 总结与展望 .....	15
4.1 主要结论 .....	15
4.2 研究展望 .....	15
参考文献 .....	17
附录 A 各学院中英文名称对照表 .....	21
附录 B 常见一级学科中英文名称对照表 .....	22
附录 C 常见专业学位类别中英文名称对照表 .....	23
作者简介 .....	24
1. 基本情况 .....	24
2. 教育和工作经历 .....	24
3. 攻读学位期间的研究成果 .....	24
3.1 发表的学术论文和著作 .....	24
3.2 申请（授权）专利 .....	24
3.3 参与的科研项目及获奖 .....	24
致 谢 .....	26



## 图目录

图 2-1 不同章节图片排版测试 .....	6
图 3-1 不同缩放系数 $\nu$ 的缩放效果 .....	10
图 3-2 理论效率与 $\gamma$ 和 $\varphi$ 的关系。 .....	12

## 表目录

表 3-1 电流类型对效率的影响 .....	13
表 3-2 高效率功放性能对比 .....	13

## 主要符号表

符号	说明	页码
$c$	电磁波的相平面速度	10

## 缩略词表

英文缩写	英文全称	中文全称
CQUPT	Chongqing University of Posts Telecommunications	重庆邮电大学

## 第1章 绪论

### 1.1 研究背景及意义

当今的社会正处于智能化趋势的浪潮中，由深度学习理论所衍生的相关技术被广泛的应用在人们所熟知的各个领域。在学术界和工业界的互相融合促进下，深度学习算法不断推陈出新，深度神经网络也不断进化并发展出适用与文字、图像、视频等信息介质的自动化识别和信息提取的高效的深度神经网络模型，许多相关领域深度神经网络模型已经是现代社会正常运转的不可或缺的一部分。

但是，当前主流的深度神经网络并不具备良好的可解释性，即便这些深度神经网络在各种测试任务下展现出了很高的准确率。例如在图像分类应用中，将待分类的图片送入训练好的深度神经网络会得到不同类别物体的置信概率分数，即便某一类别的置信概率分数是 99.99%，我们也无法得到深度神经网络做出这一决策的依据，即该深度神经网络的输出结果并不具备可解释性。并且随着应用场景的复杂化多样化，深度神经网络结构日趋复杂，参数数量日趋庞大，这使得深度神经网络的“黑盒”特点变得更为突出。这种难以解释的黑盒特性使得深度神经网络在可靠性要求极高的领域，诸如医疗影像、自动驾驶、航空航天等领域的应用就受到限制。

除此之外，上述的黑盒特性也会成为当前的深度神经网络的研究过程当中的“拦路虎”。研究者往往是将训练好的深度神经网络在既有的数据集上根据各种外部的量化指标评价训练效果，但是在实际应用过程当中，模型可能会对现实世界中某些特殊的数据或者人为恶意伪造的数据给出异常的结果。如果训练的深度神经网络不能对这些意外数据有着良好的鲁棒性，那么该神经网络就不能得到人们的信任。因此这使得相关研究者必须从更加全面的角度判断深度神经网络的性能表现并且解释其训练的神经网络的结果输出，而不只是依赖单一的评价指标来判断神经网络的性能表现。

因此，从深度神经网络实际应用和可靠性的角度出发，都需要有适用于深度神经网络黑盒特性的可解释方法，所以许多研究人员将目光转移到深度神经网络的可解释性上，他们或试图从既有的深度神经网络内外寻找其的决策依据，或试图从原理上构建可解释的深度神经网络，这两类研究路径分别就是后解释的人工智能和自解释的人工智能。

通过利用针对深度神经网络的可解释方法，不仅可以使得研究者和用户知道深度神经网络的决策依据，理解其中的决策机理，增强人对神经网络输出结果的信任程度，还可以使得研究者对神经网络的可靠性进行针对性的验证和测试。加

强对深度神经网络可解释方法的研究有助于研究人员用更加全面且灵活双向的方式和神经网络进行交互，能做到“知其然并知其所以然”。可解释性的研究赋予了深度神经网络更多的可能性并加强了其可靠性。

本文主要聚焦于图像分类神经网络的可解释性研究，更加具体的是利用图像即显著图的方式来提供图像分类神经网络的可解释性，根据显著图中给出权值数据来解释图像分类神经网络的输出结果并判断其内部决策过程是否合理可靠。同时本文的研究侧重点在于不改变神经网络内部参数，仅利用既有训练好的深度神经网络模型来进行可解释性研究。既有的图像分类神经网络模型在工业界和学术界已经得到了大量的应用，并且拥有很好的性能表现，因此在不改变模型的前提下，本文的研究可以提供简单明了能直接使用的可解释方法，对已经训练完成模型的输出结果进行显著图分析解释，将单一的，不可解释的输出结果转变为直观明了的，可解释的图像结果，提高了模型的可靠性和可信任性。此外，显著图给出的分析结果可以帮助研究者有效分析模型是否学习到正确的特征，提供神经网络决策的关键依据。

## 1.2 国内外研究现状

### 1.2.1 基于反向传播的显著图解释技术

利用反向传播的机制来实现可视化解释是较早的一些工作所采用的方法。Zeiler 等人<sup>[1]</sup>提出了一种基于反卷积的可视化方法，反卷积将特征值逆映射回了输入图片的像素空间，借此说明图片中的哪些像素参与激活了该特征值。在这项工作的基础上，导向反向传播方法<sup>[2]</sup>提出在反向传播时通过抑制输入和梯度小于 0 的值，从而突出可视化目标的重要特征。DeepLIFT<sup>[3]</sup>，LRP(Layer-wise Relevance Propagation)<sup>[4]</sup>方法通过修改反向传播的规则，将输出层的贡献逐渐向下分配直到输入层，以此来获得输入图片中每个像素对输出相关性分数。Simonyan 等人<sup>[5]</sup>提出使用输入的梯度作为可视化解释的一种手段，这种方法认为输入的某些像素对网络的预测结果起到了主要作用，它直接计算网络输出的特定类别分数对输入的梯度，但是输入的梯度中包含明显的噪声，导致显著图可视化十分模糊。SmoothGrad 方法<sup>[6]</sup>和 VarGrad 方法<sup>[7]</sup>的原理都是共同的，它们向输入图片中多次添加噪声生成一组包含噪声的图片，通过平均化结果使得生成的显著性图更加平滑。为了解决梯度饱和问题，Sundararajan 等人提出了一种积分梯度方法<sup>[8]</sup>，该方法结合了直接计算梯度和基于反向传播的归因技术 DeepLIFT 和 LRP 的分而治之的设计思想，满足敏感性和实现不变性的公理。这些研究虽然有坚实的理论基础，但是它们的可视化结果对于人类来说不容易理解，而且噪声较多。此外这些方法中许多是和具

体类别不相干的,无法对指定类别给出显著图可视化解释结果。另外有研究<sup>[9]</sup>指出其中有些方法的可靠性是值得怀疑的,它们对深度神经网络的参数不敏感,即使网络没有经过训练也能得到相似的结果。

### 1.2.2 基于类激活映射的显著图解释技术

基于类激活映射的方法是目前被大量研究和应用的一种流行的方法。这类方法利用了卷积神经网络中靠近输出端的信息,这些信息的包含着和预测结果相关的丰富的类别信息,所以这类方法能够给出类别相关的显著图可视化解释。**CAM**<sup>[10]</sup>方法首先提出了将卷积神经网络的最后一层全局平均池化后得到权重和该层提取的特征图线性相乘后累加从而生成显著图。**Grad-CAM**<sup>[11]</sup>对**CAM**方法进行了改进,无需修改网络结构,利用反向传播的梯度取均值作为权重。**Grad-CAM++**<sup>[12]</sup>进一步改进了**Grad-CAM**,它对不同像素的梯度进行加权,生成的显著图中能将同一类别物体出现多次的情况给较好的展示出来;**XGrad-CAM**<sup>[13]</sup>通过分析敏感性和实现不变性公理采用了另一种加权方法来获得特征图的权重,它引入了特征图中的像素权重为对应梯度进一步加权。为了减少噪声的影响,**Smooth Grad-CAM++**<sup>[14]</sup>,**SS-CAM**<sup>[15]</sup>也采用了**SmoothGrad**<sup>[6]2</sup>和**VarGrad**<sup>[7]2</sup>中的向输入图片中多次添加噪声的措施。**Score-CAM**<sup>[2]</sup>和**Ablation-CAM**<sup>[16]</sup>没有使用反向传播中的梯度作为特征图的权重,它们将前向传播中从最后一层卷积层获得的特征图作为掩膜来扰动输入图片,利用网络输出值或者下降值作为特征图的权重,这种方法有效避免了使用梯度而产生的噪声,取得了良好的效果。**Relevance-CAM**<sup>[17]</sup>将卷积神经网络中提取的特征图再分别输入到卷积神经网络中,对每张特征图的预测结果进行层间相关性传播(**Layer-wise Relevance Propagation**),得到对应的相关性图,将相关性图全局平均池化后作为特征的权重,该方法有效解决了之前的**CAM**方法对卷积神经网络中间层可视化解释不足的问题。**CAMERAS**<sup>[18]</sup>提出了将输入图片进行多尺度放大再输入到网络当中,将提取到的不同分辨率的特征图和梯度全部放大到和原图分辨率一样。然而,基于类激活映射的方法只针对卷积神经网络进行设计,无法便捷的迁移到其他网络当中,比如基于**Transformer**的图片分类神经网络模型。

### 1.2.3 基于扰动的显著图解释方法

基于扰动的显著图可视化方法最突出的优点便是这种方法基本只关心网络的输入和输出,可以在不同结构网络轻松的应用,即使这些网络中的细节千差万别。**Zeiler**等人<sup>[12]</sup>使用一个固定形状的方块来对输入图片进行扰动,观察输出变化从而找到对模型而已输入图片中最重要的部分。**SHAP**方法<sup>[19]</sup>使用了**Shapley**值,计算不同输入像素在是否存在的条件下对网络预测结果的影响,公平分配贡献度给

这些像素点。LIME<sup>[20]</sup>方法通过在较小的范围内扰动输入图片，得到输出结果，利用输入数据和输出数据重新训练一个可解释的代理模型去逼近黑盒模型在局部的决策边界，从而获得不同特征的重要性。RISE<sup>[21]</sup>利用蒙特卡洛方法随机生成的数量巨大的掩膜来扰动输入图片，将模型对特定类别输出概率作为的权重，再将多个所有掩膜加权相乘得到可视化结果。有研究提出了基于机器学习的方法，将掩膜作为优化对象，通过定义限制掩膜的损失函数，不断迭代从而找到最优的掩膜。在此基础上，Fong 等人<sup>[22]</sup>提出通过限制搜索区域，重新设计损失函数，达到了很好的效果。基于扰动的可视化算法因为能够对模型预测结果进行直接观察，所以可以较为真实的反映模型决策机制。但是，这种类型算法往往需要多次迭代，需要付出高额的计算成本

### 1.3 论文研究的主要内容

学位论文……

### 1.4 论文组织结构

本文……



## 第2章 相关理论介绍

### 2.1 引言

### 2.2

学位论文包括前置部分、主体部分和结尾部分共三大部分，各部分组成及顺序如所示。

学位论文各部分独立为一部分，每部分应从新的一页开始。

论文的正文（中间各章）是论文的核心部分，一般由标题、文字叙述、图、表格和公式等部分构成。由于涉及的学科、选题、研究方法等有很大的差异，可以有不同的写作表达方式，但应遵循本学科通行的学术规范，必须实事求是，客观真切，准确完备，合乎逻辑，层次分明，简练可读。引用他人研究成果时，应注明出处，不得将其与本人的工作混淆。

### 2.3 字数要求

字数要求

#### 2.3.1 硕士论文要求

各学科和学部自定。

#### 2.3.2 博士论文要求

各学科和学部自定。

### 2.4 字体和段落

学位论文中的中文统一用宋体，数字和英文统一用 Times New Roman 字体。从中文摘要开始，所有文字段落和标题行间距均取固定值 20 磅；所有段落按两端对齐、首行缩进 2 个全角字符方式书写内容。

中、英文混排时，除小数点以及引用的分图序号、公式序号等外，宜使用全角标点符号（逗号、冒号、括号、引号等）；英文段落中，符号使用应遵循英文书写习惯，统一使用半角符号，并规范使用空格。

其他要求：

- (1) 各级标题不得置于页面的最后一行，即须与下段同页；
- (2) 两个标题之间无正文时，第二个标题的段前距设置为 0 磅；

(3) 图、表、公式统一采用单倍行距；

(4) 只有一、两行文字的，不得单独作为一页内容；除各章最后一页外，中间页面不得出现较大空白；

(5) 必要时，可在规定的格式要求基础上适当微调，以利于排版，但显示效果不得与规定的格式要求存在明显差距。

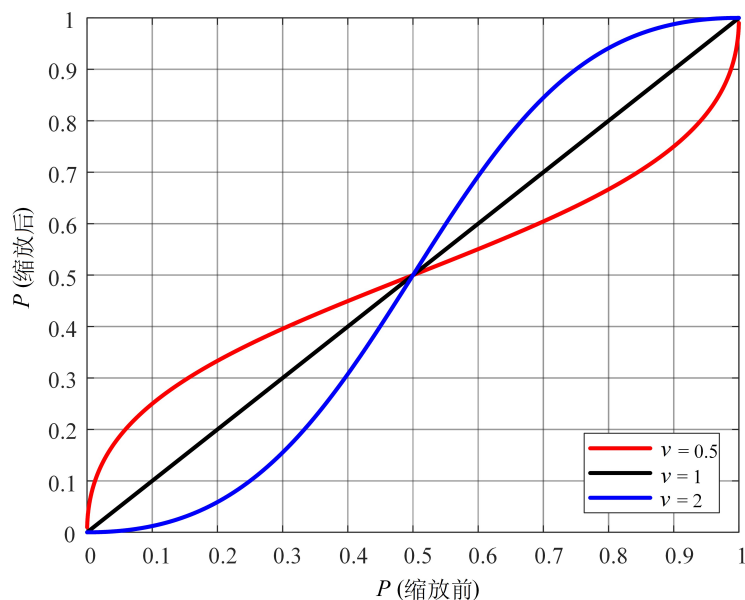


图 2-1 图片排版测试

Fig. 2-1 Scaling results with different scaling coefficients  $v$

## 2.5 章节标题

目录中章节标题只显示到 3 级标题，正文中最多显示到 4 级标题。

### 2.5.1 三级标题

#### 2.5.1.1 四级标题

## 2.6 本章小结

本章介绍了……

## 第3章 基于输入图片多尺寸放大的卷积神经网络显著图解释方法

### 3.1 本章引言

### 3.2 问题描述和研究思路

当前大多数基于类激活映射的显著图解释方法诸如 CAM、Grad-CAM、Grad-CAM++、XGrad-CAM、Score-CAM 等都存在一个共有问题，即最终生成的显著图的分辨率较低，这导致其只能在显著图中呈现一个模糊的解释效果，无法聚焦更加精细的特征。究其原因是这些基于类激活映射的方法一般提取的都是卷积神经网络的最后一层卷积层的输出作为特征图的来源，因该层包含较为丰富的类别特征信息，之后通过各种加权方法组合这些特征图来得到最终的显著图。因为卷积神经网络的结构特性，其最后一层卷积层的会输出多个通道的低分辨率特征图，因此无论如何组合这些特征图，最终得到的也只是一张分辨率和最后一层的卷积层输出的单一通道的特征图相当的初级显著图。当然一般情况下为了得到显著图和原始输入图片的特征对应关系，都会将这张原始输入图片进行分辨率层面的放大，使用的一般也是双线性插值算法，但这并不意味着原始显著图的有效信息的增多。

这里以 VGG19 这个基于卷积神经网络的图片分类模型举例，若输入图片的尺寸是  $3 \times 224 \times 224$ ，其中 3 表示 RGB 颜色通道， $224 \times 224$  表示图片分辨率，那么该模型最后一层卷积层的输出的特征图尺寸是  $512 \times 14 \times 14$ ，其中数字 512 是通道数量， $14 \times 14$  是特征图的分辨率。若我们使用双线性插值函数  $\phi(s, H, W)$  将原始显著图分辨率提升到  $224 \times 224$ ，那么意味着我们插入了额外 99.9% 的信息，而最终的显著图中的这些额外的像素信息并不能反映图片分类模型对原始输入图片的对应像素的感兴趣程度。即便有部分文献<sup>[9][11]</sup>试图改进这种插值函数，但是它们仍然引入了外部不相关的信息。

因此为了解决上述的问题，本文从特征图的提取这一关键点着手。考虑到若使用原始输入图片，那么卷积神经网络最后一层卷积层输出的单一特征图分辨率较低且包含的特征信息有限，因此本文提出基于输入图片多尺寸放大的卷积神经网络显著图解释方法，其核心就是将输入图片进行逐级多尺寸的双线性插值放大，例如将输入图片放大至  $334 \times 334$ 、 $434 \times 434$ 、 $534 \times 534$  等分辨率，这时能得到一组分辨率不同的输入图片，接着将这组输入图片分别送入基于卷积神经网络的图像分类模型中，再从其最后一层卷积层中分别提取特征图和梯度矩阵数据，基于卷积神经网络的采样特性，可以得到不同分辨率的特征图，更高分辨率的输入图片

即对应更高分辨率的特征图，也即拥有更丰富的特征信息。因此若将这些特征图的信息进行融合即能得到分辨率更高的掩膜，再用这些更高分辨率的掩膜对原始输入图片进行扰动，即可得到相应的掩膜权重，用权重和高分辨率掩膜进行融合即能得到包含更多特征信息并且分辨率更高的显著图。

### 3.3 基于输入图片多尺寸放大的卷积神经网络显著图解释方法

#### 3.3.1 生成高分辨率掩膜

令原始输入图片为  $I_0 \in \mathbb{R}^{3 \times H_0 \times W_0}$ ，其中 3 表示 RGB 颜色通道数量， $H_0$  和  $W_0$  分别原始输入图片的长和宽的像素个数，即原始输入图片的分辨率可以表示为  $\zeta_0 = (H_0, W_0)$ 。此外，我们用  $\mathcal{F}$  表示一个预训练好的基于卷积神经网络的图像分类模型，那么  $\mathcal{F}_c(I_0)$  则表示在输入图片为  $I_0$  的情况下，图像分类模型  $\mathcal{F}$  对于类别  $c$  的输出分数，值得注意的是，该分数是未经归一化指数函数（softmax 函数）之前的分数。将原始输入图片送入图像分类模型  $\mathcal{F}$ ，并从中最后一层卷积层  $l$  当中提取所有通道的特征图集合，该特征图集合表示为  $\mathbf{A}_0^*$ 。通过获得的类别  $c$  的输出分数  $\mathcal{F}_c(I_0)$ ，对该分数反向传播至最后一层卷积层特征图，则可以获得对应特征图关于类别  $c$  的梯度矩阵  $\mathbf{G}_0^*$ ，计算公式如下：

$$\mathbf{G}_0^* = \frac{\partial \mathcal{F}_c(I_0)}{\partial \mathbf{A}_0^*} \quad (3-1)$$

3-1公式中，梯度矩阵集合  $\mathbf{G}_0^*$  一共有  $K$  个通道，通道数量和特征图集合  $\mathbf{A}_0^*$  一致，并且每个通道的梯度矩阵和特征图一一对应。

接着将原始输入图片  $I_0$  经由双线性插值函数  $\varphi(I_0, \zeta_t)$  上采样至图片  $I_t$ ， $t$  表示第  $t$  次上采样结果，经由上采样后的  $I_t$  的分辨率可表示为  $\zeta_t$ ， $\zeta_t$  是由于原始输入图片的分辨率  $\zeta_0$  逐步递增得来的。为了控制计算的时间成本，考虑将  $\zeta_{max} = (H_{max}, W_{max})$  作为上采样分辨率的上限，同时设  $N$  为最大迭代次数即上采样次数，那么第  $t$  次上采样得到的图片分辨率可由以下公式计算得出：

$$\zeta_t = \zeta_0 + \lfloor \frac{\zeta_{max}}{N} \rfloor (t - 1) \quad (3-2)$$

在上述迭代过程当中，每次上采样获得的输入图片  $I_t$  的分辨率都不同，则将其输入到图像分类模型当中  $\mathcal{F}$  从最后一层卷积层提取的特征图集合  $\mathbf{A}_t^*$  的分辨率也不同，而且随着输入图片分辨率的提高， $\mathbf{A}_t^*$  的分辨率也会相应提高，相对应的  $\mathbf{A}_t^*$  也能包含更多和类别  $c$  相关的细节特征信息。因此若将这些不同的分辨率的特征图集合

进行融合则能够得到更多的特征信息。融合的方法可由以下的公式进行表示：

$$\bar{\mathbf{A}} = \frac{1}{t_{max}} \sum_{t=0}^{t_{max}} \varphi(\mathbf{A}_t^*, \zeta_0) \quad (3-3)$$

公式3-3中， $t_{max}$  表示最大的有效迭代次数，此处的“有效”指的是只有当图像分类模型的输出结果中概率分数最大的类别是  $c$  时才采用这次迭代的特征图集合  $\mathbf{A}_t^*$ ，其他情况下均舍弃，因此不难得出  $t_{max} \leq N$ 。同时  $\varphi(\mathbf{A}_t^*, \zeta_0)$  表示将特征图集合  $\mathbf{A}_t^*$  中的每一个通道上的特征图均上采样至原始输入图片的分辨率  $\zeta_0$ ，这样做的目的是方便将不同分辨率的特征图集合进行融合，也方便后续对原始输入图片进行扰动。

在每次迭代过程中，都会计算并提取保存图像分类模型关于类别  $c$  的输出分数  $\mathcal{F}_c(I_t)$  对于卷积层  $l$  的特征图集合  $\mathbf{A}_t^*$  的反向传播梯度  $\mathbf{G}_t^*$ ，和特征图集合一样，将不同迭代过程中保存的梯度矩阵集合  $\mathbf{G}_t^*$  也进行融合，可以得到  $\bar{\mathbf{A}}$  分辨率尺寸一致梯度矩阵集合，融合的公式如下所示：

$$\bar{\mathbf{G}} = \frac{1}{t_{max}} \sum_{t=0}^{t_{max}} \varphi(\mathbf{G}_t^*, \zeta_0) \quad (3-4)$$

同样，公式3-4中，同时  $\varphi(\mathbf{G}_t^*, \zeta_0)$  表示将梯度矩阵集合  $\mathbf{G}_t^*$  中的每一个通道上的梯度矩阵均上采样至原始输入图片的分辨率  $\zeta_0$ 。由于融合后的特征图集合中不仅包含类别  $c$  特征信息，也包含其他类别的特征信息，因此为了将类别  $c$  的特征信息进行凸显，其他类别的特征信息进行削弱，我们将利用融合后的梯度矩阵  $\bar{\mathbf{G}}$ ，它在一定程度上反映了特征图上不同像素对类别分数  $\mathcal{F}_c(I_0)$  的敏感程度，或者说是重要程度<sup>[11]3</sup>。仿照 Grad-CAM<sup>[11]4</sup> 中的方法，我们将  $\bar{\mathbf{G}}$  中每张梯度矩阵进行全局平均池化，在每个通道上得到一个权重值，每个通道下的权重值都和  $\bar{\mathbf{A}}$  中的每个通道下的特征图一一对应，所有通道下权重值的集合的计算公式如下：

$$\bar{\mathbf{W}} = \frac{1}{H_0 \times W_0} \sum_{i=1}^{H_0} \sum_{j=1}^{W_0} \bar{\mathbf{G}} \quad (3-5)$$

额外说明的是，融合后的特征图集合  $\bar{\mathbf{A}}$  的尺寸是  $[1 \times K \times H_0 \times W_0]$ ，它对应的权重集合  $\bar{\mathbf{W}}$  的尺寸是  $[1 \times K \times 1 \times 1]$ ，其中  $K$  就是第  $l$  层卷积层输出的特征图通道数量。

本小节的最后，初始掩膜集合  $\mathbf{M}$  可以通过以下公式计算得出：

$$\mathbf{M} = \bar{\mathbf{A}} \cdot \bar{\mathbf{W}} \quad (3-6)$$

其中，运算符  $\cdot$  表示  $\bar{\mathbf{A}}$  中的每个通道下的特征图中的每个像素值都和  $\bar{\mathbf{W}}$  中对应每个权值相乘。

### 3.3.2 掩膜优化

$\mathbf{M}$  是经由输入图片多尺寸上采样放大后从卷积层中提取的特征图和梯度矩阵结合而得到的，相比单一原始输入图片得到掩膜（比如 Score-CAM<sup>[23]2</sup>）， $\mathbf{M}$  包含更丰富的类别特征信息，但是目前得到  $\mathbf{M}$  仍然不能直接当作掩膜来扰动输入图片，它还有两个明显的缺点。第一点是  $\mathbf{M}$  的尺寸是  $[1 \times K \times H_0 \times W_0]$ ，其中  $K$  是卷积层  $l$  输出的特征图的通道数量，一般我们取最后一层卷积层作为  $l$ ，因此  $K$  的值将会是数百至上千，显然这样掩膜的数量太多了，若像 Score-CAM<sup>[23]3</sup> 一样逐个将掩膜去扰动原始输入图片得到权值，那样将会极为耗时。第二个缺点  $\mathbf{M}$  就是使用全局平均池化后的梯度作为特征图的权重，由于 ReLU 函数的零梯度问题<sup>[24]</sup>，这意味着  $\mathbf{M}$  显然会有噪声。因此，为了使  $\mathbf{M}$  成为合格的掩膜，需要让它变得更加纯净。

图、表在版面中居中放置，图编号和图题居中列在图下。编号采用阿拉伯数字分章连续编号，例如“图 3-1”，“表 3-1”以及“式 3-7”。

### 3.3.3 图

下面给出图片示例：

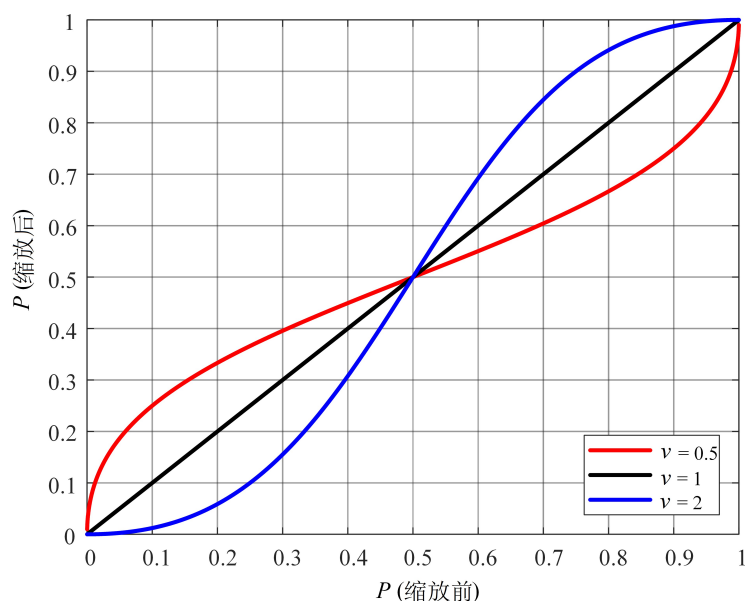


图 3-1 不同缩放系数  $v$  的缩放结果

Fig. 3-1 Scaling results with different scaling coefficients  $v$

图片标题与图片之间的间距使用默认设置即可，与上下文的间距由于 LATEX 动态排版特性，需要大家手动调整。

。

---

**Algorithm 1** MSG-CAM
 

---

**输入:** 基于卷积神经网络的图像分类模型  $\mathcal{F}$ , 原始输入图片  $I_0 \in \mathbb{R}^{3 \times H_0 \times W_0}$ , 上采样分辨率上限  $\zeta_{max} = (H_{max}, W_{max})$ , 卷积层  $l$ , 最大迭代次数  $N$ , 类别  $c$ , 每个分组中掩膜数量  $B$ , 高斯模糊参数:  $kernel\_size, sigma$ , interpolation function  $\phi(\cdot)$

**输出:** 显著图  $L_{MSG-CAM} \in \mathbb{R}^{3 \times H_0 \times W_0}$

```

1: 初始化  $L_{MSG-CAM} \leftarrow 0$ ,  $\mathcal{A}_0 \leftarrow 0$ ,  $\mathcal{G}_0 \leftarrow 0$ ,  $r \leftarrow 0$ ,  $t_{max} \leftarrow 0$ ,  $t_m \leftarrow 0$ ,  $kernel\_size = 51$ ,  $sigma = 50$ , 基准图片  $I'_0 = gaussian\_blur2d(I_0, kernel\_size, sigma)$ ,  $c = \mathcal{F}(I_0)$ 
2: while  $t \leq N$  do
3:    $t \leftarrow t + 1$ 
4:    $\zeta_t \leftarrow \zeta_0 + \lfloor \frac{\zeta_{max}}{N} \rfloor (t - 1)$ 
5:    $I_t \leftarrow \phi(I_0, \zeta_t)$ 
6:   if  $\mathcal{F}(I_t) \rightarrow c$  then
7:      $t_{max} \leftarrow t_{max} + 1$ 
8:      $\mathbf{A}_t^* \leftarrow \mathcal{A}(I_t, l)$ 
9:      $\mathbf{G}_t^* \leftarrow \nabla \mathcal{J}(l, c)$ 
10:     $\mathcal{A}_t \leftarrow \mathcal{A}_{t-1} + \phi(\mathbf{A}_t^*, \zeta_0)$ 
11:     $\mathcal{G}_t \leftarrow \mathcal{G}_{t-1} + \phi(\mathbf{G}_t^*, \zeta_0)$ 
12:   end if
13: end while
14:  $\bar{\mathbf{A}} \leftarrow \mathcal{A}_t / t_{max}$ 
15:  $\bar{\mathbf{G}} \leftarrow \mathcal{G}_t / t_{max}$ 
16:  $\bar{\mathbf{W}} \leftarrow \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n \bar{\mathbf{G}}$ 
17:  $\mathbf{M} = \bar{\mathbf{A}} \cdot \bar{\mathbf{W}}$ 
18:  $K \leftarrow \mathbf{M}$  的通道数量;
19:  $g \leftarrow$  每个分组中掩膜数量;
20: while  $r < B$  do
21:   在组内生成单一的掩膜
      $M_r = ReLU(\sum_{k=r \times g}^{(r+1) \times g - 1} M^k)$ 
22:   初始化  $M'_r \leftarrow$  将  $M^r$  进行去噪操作和归一化
23:   用掩膜扰动原始输入图片
      $I_r = I_0 \odot M'_r + I'_0 \odot (1 - M'_r)$ 
24:   计算每张掩膜的权重
      $\alpha_r = \mathcal{F}_c(I_r) - \mathcal{F}_c(I'_0)$ 
25:    $L_{MSG-CAM} \leftarrow L_{MSG-CAM} + \alpha_r M_r$ 
26:    $r \leftarrow r + 1$ 
27: end while
28: return  $ReLU(L_{MSG-CAM}^c)$ 

```

---





表 3-1 电流类型对效率的影响  
Table 3-1 Current type impact on efficiency

电流类型	A	B	C
aaa	aa1	bb1	cc1

参照写作指南。表格格式参照写作指南。表格格式参照写作指南。表格格式参照写作指南。表格格式参照写作指南。表格格式参照写作指南。表格格式参照写作指南。表格格式参照写作指南。表格格式参照写作指南。表格格式参照写作指南。

表 3-2 高效率功放性能对比  
Table 3-2 High-efficiency power amplifier performance comparison

带宽 (GHz)	功率 (dBm)	效率 (%)	线性度 (dBc)	信号带宽 (MHz)
1.4–2.6	32–34	30–40 (DE)	–25 – –30 (ACLR)	5
2.1–2.7	39	45 (DE) @ 2.14 GHz	–31 (ACLR)	5
	(average)	40 (DE) @ 2.655 GHz	–30 (ACLR)	
3.5	38.1	59 (PAE)	30 (C/I)	5
1.6–2.6	36.0–38.5	45–60 (PAE)	30 (C/I)	5
	35.3–37.5	40–55 (PAE)	–30 (ACLR)	20

3.4 公式格式

$$\left\{ \begin{array}{l} 0.794 \leq \zeta \leq 1 \\ 0.631 \leq \gamma = \frac{0.631}{\zeta^2} \leq 1 \\ -\frac{1}{2\gamma} \leq \delta \leq \frac{1}{2\gamma} \\ Z_{c,low,f} = 2R_{opt}(\gamma + j\delta) \\ Z_{c,2f} = Z_{c,low,2f} = -j\frac{3\pi}{4}\gamma\delta R_{opt} \end{array} \right. \quad (3-7)$$

$$v(\theta) = V_{DD} \cdot (1 - \alpha \cos(\theta + \varphi) + \beta \cos(3\theta + 3\varphi)) \cdot (1 - \gamma \sin(\theta + \varphi)) \quad -1 \leq \gamma \leq 1 \quad (3-8)$$

公式格式测试。  $\Theta = \{\theta_k(n), \forall k, n\}$

### 3.5 印制要求

涉密学位论文的印刷、制作、传递、存档等，须符合国家、学校相关保密要求。学位论文一律左侧装订。

中文摘要之前的前置部分（封面、中、英文题名页、独创性声明和使用授权书），采用单面印刷。

从中文摘要开始，采用双面印刷。

中文摘要及之后的前置部分，包括中文摘要、ABSTRACT、目录、图目录（如有）、表目录（如有）、主要符号表（如有）、缩略词表（如有），在双面印刷时，若某部分页数为奇数，则该部分最后一页单面印刷。例如：若“摘要”只有1页，则其页码是“I”，第“I”页纸的背面为空白（无页眉或页码）；“ABSTRACT”用新的一张纸印刷，页码从“II”开始。

从第1章第1页开始，至论文最后1页，所有页面均双面印刷。例如：若第1章的最后1页为第17页，则第2章的第1页在第17页的背面印刷，页码为“18”（页眉是“重庆邮电大学博士（硕士）学位论文”）。

一次性双面打印整本学位论文技巧：除用于打印的版本外，电子版论文中一律不得出现空白页。论文打印建议使用PDF格式。为方便一次性双面打印，有时可在单面印刷的部分（如封面、中、英文题名页、独创性声明和使用授权书），或者双面打印只有1页的某部分内容（如摘要、ABSTRACT等）后插入1页空白页，该空白页不编排页眉页码；论文中出现的页码应前后连续，不得中断。

### 3.6 本章小结

本章介绍了……

## 第4章 总结与展望

### 4.1 主要结论

本文主要……

### 4.2 研究展望

更深入的研究……

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

## 参考文献

- [1] Matthew D Zeiler and Rob Fergus. Visualizing and understanding convolutional networks. In *European conference on computer vision*, pages 818–833. Springer, 2014.
- [2] Jost Tobias Springenberg, Alexey Dosovitskiy, Thomas Brox, and Martin Riedmiller. Striving for simplicity: The all convolutional net. *arXiv preprint arXiv:1412.6806*, 2014.
- [3] Avanti Shrikumar, Peyton Greenside, and Anshul Kundaje. Learning important features through propagating activation differences. In *International conference on machine learning*, pages 3145–3153. PMLR, 2017.
- [4] Alexander Binder, Grégoire Montavon, Sebastian Lapuschkin, Klaus-Robert Müller, and Wojciech Samek. Layer-wise relevance propagation for neural networks with local renormalization layers. In *International Conference on Artificial Neural Networks*, pages 63–71. Springer, 2016.
- [5] Karen Simonyan, Andrea Vedaldi, and Andrew Zisserman. Deep inside convolutional networks: Visualising image classification models and saliency maps. In *In Workshop at International Conference on Learning Representations*. Citeseer, 2014.
- [6] Daniel Smilkov, Nikhil Thorat, Been Kim, Fernanda Viégas, and Martin Wattenberg. Smoothgrad: removing noise by adding noise. *ArXiv preprint*, abs/1706.03825, 2017.
- [7] Julius Adebayo, Justin Gilmer, Ian Goodfellow, and Been Kim. Local explanation methods for deep neural networks lack sensitivity to parameter values. *ArXiv preprint*, abs/1810.03307, 2018.
- [8] Mukund Sundararajan, Ankur Taly, and Qiqi Yan. Axiomatic attribution for deep networks. In *International conference on machine learning*, pages 3319–3328. PMLR, 2017.
- [9] Julius Adebayo, Justin Gilmer, Michael Muelly, Ian J. Goodfellow, Moritz Hardt, and Been Kim. Sanity checks for saliency maps. In Samy Bengio, Hanna M. Wallach,

- Hugo Larochelle, Kristen Grauman, Nicolò Cesa-Bianchi, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, December 3-8, 2018, Montréal, Canada*, pages 9525–9536, 2018.
- [10] Bolei Zhou, Aditya Khosla, Àgata Lapedriza, Aude Oliva, and Antonio Torralba. Learning deep features for discriminative localization. In *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016*, pages 2921–2929. IEEE Computer Society, 2016.
- [11] Ramprasaath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. Grad-cam: Visual explanations from deep networks via gradient-based localization. In *IEEE International Conference on Computer Vision, ICCV 2017, Venice, Italy, October 22-29, 2017*, pages 618–626. IEEE Computer Society, 2017.
- [12] Aditya Chattopadhyay, Anirban Sarkar, Prantik Howlader, and Vineeth N Balasubramanian. Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks. In *2018 IEEE winter conference on applications of computer vision (WACV)*, pages 839–847. IEEE, 2018.
- [13] Ruigang Fu, Qingyong Hu, Xiaohu Dong, Yulan Guo, Yinghui Gao, and Biao Li. Axiom-based grad-cam: Towards accurate visualization and explanation of cnns. In *31st British Machine Vision Conference 2020, BMVC 2020, Virtual Event, UK, September 7-10, 2020*. BMVA Press, 2020.
- [14] Daniel Omeiza, Skyler Speakman, Celia Cintas, and Komminist Weldermariam. Smooth grad-cam++: An enhanced inference level visualization technique for deep convolutional neural network models. *ArXiv preprint*, abs/1908.01224, 2019.
- [15] Haofan Wang, Rakshit Naidu, Joy Michael, and Soumya Snigdha Kundu. Ss-cam: Smoothed score-cam for sharper visual feature localization. *ArXiv preprint*, abs/2006.14255, 2020.
- [16] Harish Guruprasad Ramaswamy et al. Ablation-cam: Visual explanations for deep convolutional network via gradient-free localization. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 983–991, 2020.

- 
- [17] Jeong Ryong Lee, Sewon Kim, Inyong Park, Taejoon Eo, and Dosik Hwang. Relevance-cam: Your model already knows where to look. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 14944–14953, 2021.
- [18] Mohammad AAK Jalwana, Naveed Akhtar, Mohammed Bennamoun, and Ajmal Mian. Cameras: Enhanced resolution and sanity preserving class activation mapping for image saliency. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 16327–16336, 2021.
- [19] Scott M. Lundberg and Su-In Lee. A unified approach to interpreting model predictions. In Isabelle Guyon, Ulrike von Luxburg, Samy Bengio, Hanna M. Wallach, Rob Fergus, S. V. N. Vishwanathan, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA*, pages 4765–4774, 2017.
- [20] Marco Túlio Ribeiro, Sameer Singh, and Carlos Guestrin. ”why should I trust you?”: Explaining the predictions of any classifier. In Balaji Krishnapuram, Mohak Shah, Alexander J. Smola, Charu C. Aggarwal, Dou Shen, and Rajeev Rastogi, editors, *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, August 13-17, 2016*, pages 1135–1144. ACM, 2016.
- [21] Vitali Petsiuk, Abir Das, and Kate Saenko. RISE: randomized input sampling for explanation of black-box models. In *British Machine Vision Conference 2018, BMVC 2018, Newcastle, UK, September 3-6, 2018*, page 151. BMVA Press, 2018.
- [22] Ruth Fong, Mandela Patrick, and Andrea Vedaldi. Understanding deep networks via extremal perturbations and smooth masks. In *2019 IEEE/CVF International Conference on Computer Vision, ICCV 2019, Seoul, Korea (South), October 27 - November 2, 2019*, pages 2950–2958. IEEE, 2019.
- [23] Haofan Wang, Zifan Wang, Mengnan Du, Fan Yang, Zijian Zhang, Sirui Ding, Piotr Mardziel, and Xia Hu. Score-cam: Score-weighted visual explanations for convolutional neural networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, pages 24–25, 2020.

- [24] Qinglong Zhang, Lu Rao, and Yubin Yang. A novel visual interpretability for deep neural networks by optimizing activation maps with perturbation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 3377–3384, 2021.



## 附录 A 各学院中英文名称对照表

序号	中文名称	英文名称
01	通信工程学院	School of Communications and Information Engineering

## 附录 B 常见一级学科中英文名称对照表

代码	中文名称	英文名称
0810	信息与通信工程	Information and Communication Engineering

## 附录 C 常见专业学位类别中英文名称对照表

代码	中文名称	英文名称
1256	工程管理	Engineering Management

## 作者简介

### 1. 基本情况

张某某，男，重庆人，1993 年 8 月出生，重庆邮电大学 XX 学院 XX 专业 2018 级博士研究生。

### 2. 教育和工作经历

### 3. 攻读学位期间的研究成果

#### 3.1 发表的学术论文和著作

#### 3.2 申请（授权）专利

#### 3.3 参与的科研项目及获奖

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。

以下文字用于测试。



感谢老师、同学们的关心、支持和帮助！  
感谢老师、同学们的关心、支持和帮助！  
感谢老师、同学们的关心、支持和帮助！  
感谢老师、同学们的关心、支持和帮助！  
感谢老师、同学们的关心、支持和帮助！  
感谢老师、同学们的关心、支持和帮助！  
感谢老师、同学们的关心、支持和帮助！  
感谢老师、同学们的关心、支持和帮助！  
感谢老师、同学们的关心、支持和帮助！  
感谢老师、同学们的关心、支持和帮助！  
感谢老师、同学们的关心、支持和帮助！  
感谢老师、同学们的关心、支持和帮助！