

北京信息科技大学

# 毕业设计（论文）

题    目：\_\_\_\_\_

学    院：\_\_\_\_\_

专    业：\_\_\_\_\_

学生姓名：\_\_\_\_\_班级/学号\_\_\_\_\_

指导老师/督导老师：\_\_\_\_\_

起止时间：\_\_\_\_\_年  月  日至\_\_\_\_\_年  月  日

## 目录

第一章 绪论 .....	4
1.1 研究背景 .....	4
1.2 研究目的与意义 .....	4
1.3 国内外研究现状 .....	5
1.4 论文结构与组织 .....	6
第二章 相关研究与技术概述 .....	7
2.1 情感分析研究现状 .....	7
2.2 多模态情感分析方法 .....	7
2.3 图像特征提取 .....	7
2.4 文本特征提取 .....	8
2.5 多模态特征融合 .....	8
2.6 小结 .....	8
第三章 多模态情感分析系统的设计 .....	10
3.1 系统架构 .....	10
3.2 微信小程序 .....	11
3.2.1 UI 设计 .....	11
3.2.2 视图层设计 .....	12
3.2.3 业务逻辑层设计 .....	12
3.2.4 数据层设计 .....	12
3.3 数据库设计与实现 .....	12
3.4 小结 .....	13
第四章 数据收集与处理 .....	15
4.1 新浪微博爬虫实现 .....	15
4.2 数据处理 .....	16
4.2.1 数据清洗 .....	16
4.2.2 数据预处理 .....	16
4.3 数据标注 .....	17
4.4 小结 .....	18
第五章 图文融合多模态情感分析模型 .....	19
5.1 模型架构 .....	19
5.2 图像特征提取 .....	19
5.2.1 预训练的 VGG19 模型 .....	19
5.2.2 Flatten 层和全连接层 .....	19
5.3 文本特征提取 .....	19
5.3.1 单词向量输入 .....	19
5.3.2 自定义 Transformer 编码器 .....	20
5.3.3 全局平均池化和全连接层 .....	20
5.4 图文多模态特征融合 .....	20
5.4.1 注意力层 .....	20
5.4.2 全连接层及分类输出 .....	20
5.5 模型训练与优化 .....	20
5.5.1 损失函数与优化器 .....	20
5.5.2 回调函数 .....	20

5.6 小结 .....	20
第六章 实验与结果分析 .....	22
6.1 实验环境与数据集 .....	22
6.1.1 实验环境 .....	22
6.1.2 数据集 .....	22
6.2 实验指标 .....	22
6.3 模型性能评估 .....	22
6.3.1 模型最终性能 .....	22
6.3.2 对比实验 .....	23
6.4 小结 .....	24
第七章 结束语 .....	25
7.1 研究总结 .....	25
7.2 创新与贡献 .....	25
7.3 局限性与未来工作 .....	25
致谢 .....	27
参考文献 .....	28

# 第一章 绪论

## 1.1 研究背景

随着互联网的高速发展和智能手机的普及，社交媒体平台已迅速成为人们获取信息、表达观点和情感的主要渠道，越来越多的社交媒体平台应运而生。新浪微博、微信等平台深入人们的日常生活，在很多人的生活中都成为了不可或缺的一环。在这些平台上，人们以图片、文字、视频等形式展示生活、表达情感，极大地丰富了网络信息的多样性。

根据微博 2022 第三季度财报显示，在 2022 年的九月份，微博拥有 5.84 亿的月度活跃用户。这庞大的用户群所产生的海量数据，隐藏着多领域的巨大商业价值和社会价值<sup>[1]</sup>。

情感分析是自然语言处理和人工智能领域的研究热点，其主要目的是识别和提取文本中的情感信息。然而，在当下，社交媒体数据不再局限于单一的文本模态，而是呈现为更多样化的模态信息，仅依赖文本进行情感分析不足以准确识别人类表达出的复杂情感<sup>[2]</sup>。因此，需要研究多模态情感分析方法，以更好地捕捉和理解社交媒体中的情感信息。

多模态情感分析应运而生，它以单模态情感分析为基础，有效融合文字、图像或音频等数据信息，帮助模型做出更精确的判断。然而，现有的情感分析方法主要关注单一模态的信息处理，忽略了文本和图像等多模态信息的综合分析。因此，本研究旨在开发一种基于多模态信息的情感分析方法，提高新浪微博数据的分析效果，以满足社交媒体时代多样化信息处理的需求。这将有助于更好地挖掘社交媒体数据中的价值，并为相关领域提供更准确的支持。

## 1.2 研究目的与意义

本研究的主要目的是开发一种基于多模态信息的情感分析方法，通过融合文本和图像特征，提高社交媒体数据的情感分析准确性。此外，本研究还将探讨不同融合策略对情感分析性能的影响，为多模态情感分析方法的优化提供理论依据。本研究的意义主要体现在如下几个方面：

(1) 提高情感分析的准确性：多模态的情感分析方法可以充分利用文字和图像信息，提高情感识别的准确性。这有助于在社交媒体数据中对情感信息的挖掘更加精准，对企业和政府部门的数据支持也更加精准。同时，提高情感分析的准确性也有助于进一步了解用户在社交媒体上的行为和需求，为用户提供更好的服务和体验。

(2) 丰富情绪分析应用场景：舆情监测、产品点评分析、金融市场预测等领域可以广泛使用多模态情绪分析方法。在舆情监控方面，多模态情感分析方法有助于及时发现网络舆情热点，为企业和政府部门提供有力的决策支持。在产品评论分析方面，多模态情感分析方法可以帮助企业更好地了解消费者的需求和反馈，为产品的优化和改进提供依据。在金融市场预测方面，多模态情感分析方法可以有效分析市场上的情感信息，为投资者提供更为精确的市场分析结果，降低投资风险。

(3) 促进多模态情感分析方法的研究：本研究的理论和实践成果可以为多模态情感分析方法的进一步发展提供借鉴和启迪。通过探讨不同融合策略对情感分析性能的影响，有助于研究者在设计多模态情感分析方法时，选择更为合适的融合策略。此外，这项研究还有助于拓展其他领域的多模态情感分析方法的应用。

(4) 促进人工智能与社会科学的交叉融合：本研究将人工智能领域的技术应用于情感分析问题，有助于促进人工智能与社会科学的交叉融合。通过多模式的情感分析方法，可以更好地了解人们在社交媒体上的行为方式、情感方式和沟通方式，提供有价值的研究数据和洞见，应用于社会科学领域，如心理学，社会学，传播学等。同时，这种交叉融合也有助于开拓人工智能领域的新应用场景，为人工智能技术的进一步发展提供更多可能性。

(5) 增强社交媒体平台的功能和价值：通过本研究开发的多模态情感分析方法，可以帮助社交媒体平台更好地分析和利用用户生成的内容，为用户提供更加丰富和精准的信息推荐、

广告定向和内容策划等服务。这将进一步提升社交媒体平台的功能和价值,吸引更多的用户和合作伙伴,促进社交媒体行业的持续发展。

(6) 为信息传播和舆论引导提供支持:多模态情感分析方法可以帮助更准确地识别和提取社交媒体数据中的情感信息,为信息传播和舆论引导提供有力支持。政府部门和企业可以利用多模态情感分析方法及时发现和应对网络舆论热点,有效地引导网络舆论走向,维护网络空间的秩序和安全。

### 1.3 国内外研究现状

多模态情感分析较早的研究着重于融合策略、特征提取和模型建立等方面。在情感分析研究的初期,采用的方法是词频加机器学习方法,代表工作为Turney等人<sup>[3]</sup>的研究。后来发现单纯用bag of words无法很好解决情感词组合、否定/程度词等问题,一些工作开始搭建情感词典、情感规则,例如SentiWordNet<sup>[4]</sup>、SO-CAL<sup>[5]</sup>。随后,深度学习在NLP流行后,研究者尝试了多种提取特征的模型结构,包括基于卷积神经网络(CNN)<sup>[6]</sup>、循环神经网络(RNN)<sup>[7]</sup>、长短期记忆网络(LSTM)<sup>[8]</sup>等方法。

在多模态情感分析方面,国内学者近年来也有一系列成功的成果。例如,清华大学研究团队基于注意力机制,开发了一种多模态情感分析方法<sup>[9]</sup>,通过动态权重分配,提高情感分析的准确性,实现文字与图像特征的自适应融合。此外,国内研究者还针对中文社交媒体数据的特点,提出了一些特定于中文环境的多模态情感分析方法,如考虑汉字的字形特征、利用预训练中文词向量<sup>[10]</sup>等。

文本和图像作为网络媒体和社交平台中最常见的两种数据源,探究两种模态信息之间的语义相关性和情感联系受到了更多的关注。Schmidt S, Stock W G.等人的研究<sup>[11]</sup>使用机器学习方法研究了文本与图像情感之间的紧密联系性。Cao等人<sup>[12]</sup>对微博文本和图像进行情感分析,从文本中提取有用的单词、标签和标点符号,从图像中获得低层次的特征如颜色直方图、GIST描述符等,然后构建情感概念词典和情感本体,并通过归一化和线性插值法进行有效融合。近几年,随着硬件设备和方法的进步,科研工作者多侧重于通过深度学习的方法进行跨模态的情感分析。Yu等人的研究<sup>[13]</sup>中利用两个CNN独立模型,分别对文字、图像特征进行提取,然后通过后期融合和Logistic回归进行情感预测。Chen等人的研究<sup>[14]</sup>与之相类似,将提取到的两个模态信息融合在一个池化层中,再输送到全连接层进行情感分析。Yu和Chen的研究<sup>[13][14]</sup>,只是将通过CNN提取到的特征向量进行简单的拼接操作,不能很好的建模模态间复杂的关系。You等人<sup>[15]</sup>首先通过图像情感数据集对预训练的CNN模型进行微调,再将倒数第二层提取出来作为图像特征,用无监督的学习方式对Word2VEC模型进行训练,对文档或句子进行分布式表示提取,最后分别对特征级融合和决策级融合方法进行实验,两者的效果都优于单一模式。之后又在文献[15]的基础上建立了跨模态一致回归模型<sup>[16]</sup>,模型考虑为有关联的不同模态之间添加一致约束,具体通过在损失函数中为每个模态特征和连接特征增加惩罚项KL散度实现。

综上所述,多模态情感分析在国内外均取得了较为丰富的研究成果,但仍存在一定的技术挑战和应用空间。有关的模态情绪分析研究,现在看来还是有一些不足:

(1) 数据集的文本信息都是以英语为主,对于中文内容的研究有限,并不能很好的分析。

(2) 应用市场上并没有具备情感分析功能的产品,用户无法使用 and 了解。

因此,本研究依托于微信小程序来开发,设计了一个基于微博社交媒体数据的多模态情感分析系统,功能包含图像和文本的上传、情感倾向分析、对情感数据的持久化。

本研究旨在充分借鉴国内外研究成果,探索一种高效的多模态情感分析方法,并结合新浪微博数据的特点,为相关领域的研究与应用提供新的理论与实践支撑。

## 1.4 论文结构与组织

本论文共分 7 章，各章节内容安排如下：

第一章为绪论，主要介绍研究背景、研究目的与意义、国内外研究现状以及论文结构与组织。

第二章为相关研究与技术概述，介绍情感分析研究现状、多模态情感分析方法、图像特征提取、文本特征提取、多模态特征融合的内容。

第三章为多模态情感分析系统的设计：阐述了系统整体架构与微信小程序的设计与实现，包括 UI 设计、视图层设计、业务逻辑层设计与数据层设计，以及数据库设计。

第四章为数据收集与处理，描述了新浪微博资料的爬取工作，包括微博爬虫的设计，以及数据预处理过程中的文本清洗、分词、停用词过滤等操作。

第五章为图文融合多模态情感分析模型：详细介绍了模型的架构、图像特征提取、文本特征提取、特征融合与情感分类，以及模型的训练与优化。

第六章为实验与结果分析：描述了实验环境、数据集、实验方法与指标，对模型性能进行了评估，并对实验结果进行了分析与讨论。

第七章为结束语：总结了本研究的内容和成果，阐述了创新与贡献，以及局限性与未来工作方向。

## 第二章 相关研究与技术概述

### 2.1 情感分析研究现状

情感分析 (Emotional Analysis) 是自然语言处理领域的一项重要工作, 其目标是挖掘和提取文本资料中的情感信息, 包括情绪、观点、态度等。近年来, 无论是学术界还是工业界, 情感分析技术都受到了广泛的关注和研究。许多研究者提出了各种不同角度、不同领域的解决方法, 以情感分析任务为目标, 主要包括基于词汇的方法、机器学习方法和深度学习方法。这些方法都在各自的应用场景中做出了自己的贡献。然而, 随着多模态数据的快速发展, 研究者们越来越关注如何有效地融合多种信息源 (如文本、图像和声音等) 以进行更为准确的情感分析。

### 2.2 多模态情感分析方法

多模态情感分析方法为提高情感分析的精准度, 整合了不同模态的信息 (例如: 文字, 图像, 音频等)。此类方法独立处理各个模态的信息, 再通过特定的融合策略将抽取的特征结合, 从而达成情感预测。

本研究通过以下实例揭示单模情感分析的局限性, 以突出多模情感分析方法的优势。设想一个在线产品评论场景, 评论者在文本中写道“这个产品质量不错”, 然后附上了一张产品损坏的图片。如果仅仅依赖于文本的情感分析, 会错误地推断出评论者对该产品的评价是积极的, 忽视了图片传递的负面信息。在此例中, 文本和图像所传递的情感信息存在冲突, 这种情况在实际生活中并不罕见。

因此, 在这项研究中, 引入了多模态情感分析的方法。这种方法通过深度学习模型对文本和图像数据的特征进行提取和融合, 实现了对社交媒体数据的多模态情感分析。不仅考虑了文本信息, 而且还包含了图像信息, 因此能更全面、准确地捕捉和理解用户的情绪表达。这种方法更适应于处理现实世界中复杂、多元的情感表达, 从而提高情感分析的准确度和可靠性。

### 2.3 图像特征提取

图像特征提取是多模态情感分析中必不可少的一步。有多种方法可用于提取图像特征, 如手动提取颜色、纹理和形状等低级特征, 或利用深度学习模型自动提取高级特征。其中, 深度学习模型已经成为当前图像特征提取的主流方法, 其可以有效地捕获图像的高级抽象特征。

常见的深度学习模型有 AlexNet、VGGNet、GoogLeNet、ResNet 等。这些模型都在图像识别任务中取得了优异的性能, 具有强大的特征提取能力。然而, 它们在特性和性能上有所不同。比如, AlexNet 是深度学习中的开创者, 但其模型结构相对较简单; VGGNet 包括 VGG16 和 VGG19, 它们的模型结构更深, 能够捕获更复杂的特征; GoogLeNet 引入了 Inception 模块, 使模型在大幅减少运算量的同时保持性能; ResNet 通过引入残差结构, 成功训练了超过 1000 层的深度网络。

在这些模型中, 本研究选择使用 VGG19 进行图像特征提取。VGG19 模型包含 19 个卷积层、全连接层和池化层, 其架构深度适中, 能够有效捕捉图像的局部和全局特征。具体来说, VGG19 模型的较低层次主要捕获图像的局部特征, 如边缘、颜色和纹理等; 而较高层次则能够捕捉更复杂的抽象特征, 如物体的形状和结构等。因此, VGG19 模型可以从不同层次对图像进行深度理解, 能够从多个角度理解图像的情感表达。

使用预训练的 VGG19 模型, 只需要将图像数据输入到 VGG19 模型中, 通过前向传播的方式获取模型的中间层输出, 作为图像的特征表示。这种方法的优点在于, 不需要从头开始训练模型, 而是利用在大量图像数据上预训练的 VGG19 模型的权重, 可以大大节省计算资源和时间。

## 2.4 文本特征提取

在多模态情感分析中，文本特征提取同样占据了关键地位。针对文本数据，也存在多种特征提取方法，如基于统计的词频-逆文本频率（TF-IDF）方法、基于词袋模型的方法，以及基于深度学习的词嵌入和自注意力机制等方法。

在本研究首先使用 jieba 分词库对文本进行处理。jieba 是最出名的一款中文分词工具，它是基于前缀词典，来对输入文本进行切分，通过动态规划算法，计算得到最大概率路径，得到最终切分形式。它能有效地将原始文本切分成更易于处理的词汇单元。

接下来，本研究利用预训练的 Word2Vec 模型将分词结果转换为词向量。Word2Vec 是一种能够将单词映射到高维向量空间中的基于神经网络的单词嵌入方法，它可以将词语映射到一个高维的向量空间中，在向量空间中，语义相近的单词距离也会比较接近。Word2Vec 模型的训练方式主要有两种，即连续词袋模型 (CBOW) 和 Skip-GRAM 模型，它们分别从不同的角度捕捉单词的上下文信息。

最后，本研究利用 Transformer 模型对词向量进行进一步的特征提取。Transformer 模型是一种基于自注意力机制的深度学习模型，它通过自注意力机制，可以关注到文本中所有位置的信息，捕获到文本中的长距离依赖关系和复杂的语义信息。Transformer 模型的一个主要优点是其并行计算能力，相比于传统的循环神经网络 (RNN) 和长短期记忆网络 (LSTM)，Transformer 模型在处理长文本时更具优势。

在具体的实现中，本研究使用预训练的 Word2Vec 和 Transformer 模型对文本数据进行特征提取。这种方法的优点在于，利用预训练的模型，能够直接获取高质量的文本特征，大大节省了计算资源和时间。

## 2.5 多模态特征融合

多模态情感分析中的关键问题是如何有效融合来自不同模态的特征。目前，存在多种多模态特征融合的方法，如简单的特征拼接、特征级联、特征交叉和特征选择等。然而，这些传统的特性整合方法往往忽略了不同的模式之间的交互关系，不能充分利用不同模态的信息。

为解决这一问题，本研究采用了注意力机制来实现自适应的特征融合。注意力机制最早来源于人类的视觉观察习惯，即在观察大量信息时，人类会自动关注到那些重要的信息，而忽视那些不重要的信息。在多模态特征融合中，注意力机制可以自动学习到不同模态特征的重要性，从而对不同模态的特征赋予不同的权重。

具体来说，本研究先将来自图像和文本的特征通过一个全连接层进行初步融合，然后通过一个注意力层计算出不同模态特征的权重。这个权重是基于数据自动学习得到的，可以反映出不同模态特征在情感分析中的重要性。最后，根据这个权重，本研究可以得到一个融合了不同模态特征的情感表示。

这种以注意力机制为基础的多模态特征融合法的好处是可以根据不同模态的特征自动适应不同的数据并自动调整权重，达到更好的融合性。而且该方法相对来说运算复杂度较低，适用于对大规模的数据的处理。

## 2.6 小结

本章主要围绕情感分析的研究现状、多模态情感分析方法、图像特征提取、文本特征提取以及多模态特征融合进行了详细的概述。

首先，我们介绍了情感分析的研究现状，讨论了从基于词汇的方法到深度学习方法的各種情感分析技术，以及当前研究者如何关注融合多模态数据进行更准确的情感分析。接着，我们对多模态情感分析方法进行了深入讨论，解释了如何结合文本、图像和音频等多模态信息以提高情感分析的精准度，并通过实例揭示了多模态情感分析的优势。

其次，我们分别探讨了图像特征提取和文本特征提取的方法和技术。在图像特征提取部



分，我们讨论了使用深度学习模型（特别是 VGG19 模型）进行图像特征提取的优点和方法。在文本特征提取部分，我们介绍了 jieba 分词库的使用，预训练的 Word2Vec 模型以及基于自注意力机制的 Transformer 模型在文本特征提取中的应用。

最后，我们讨论了多模态特征融合的方法，特别是使用注意力机制来实现自适应的特征融合的方法，包括如何通过注意力机制自动学习不同模态特征的重要性，以及如何利用这些权重得到一个融合了不同模态特征的情感表示。

本章为理解情感分析的各种技术和方法提供了全面的概述，为下一步的研究和实验工作奠定了基础。

## 第三章 多模态情感分析系统的设计

### 3.1 系统架构

系统的架构主要包括三个核心部分：一个部署在阿里云服务器上的后端，负责处理小程序的请求并与数据库进行交互；一个本地运行的 API，使用 ngrok 进行内网穿透，负责处理图像和文本数据，并返回情感分析结果；以及用户端的小程序，负责与用户进行交互、向后端发送请求和展示情感分析结果。这三个部分协同工作，共同实现了用户上传文本和图片数据的情感分析功能。

阿里云服务器上的后端扮演了核心角色，主要负责处理用户端小程序发出的请求。这部分后端代码采用 Flask 框架编写，用于接收用户提交的文本和图片数据，并将这些数据转发给本地运行的 API。此外，用户历史记录的存储、查询等与数据库的交互，也由后端负责。为了保证系统的安全性和稳定性，阿里云服务器上的后端设置了相应的权限和安全策略。

本地运行的 API 主要负责对接收到的文本和图片数据进行处理。使用 ngrok 进行内网穿透，使得阿里云服务器上的后端可以访问到本地 API。该 API 采用 Flask 框架编写，利用预训练的深度学习模型对文本和图片数据进行特征提取，并结合多模态方法对这些特征进行融合，最后通过情感分类模型返回情感分析结果。

用户端小程序则是系统与用户交互的界面，主要负责收集用户输入的文本和上传的图片数据，同时向后端发送请求，并在获取情感分析结果后，以友好的方式展示给用户。用户端小程序具备良好的用户体验设计，用户能够轻松地使用本系统进行情感分析。

通过这三个部分的紧密合作，本系统有效地实现基于微博社交媒体数据的多模态情感分析功能，为用户提供准确、及时的情感识别服务。

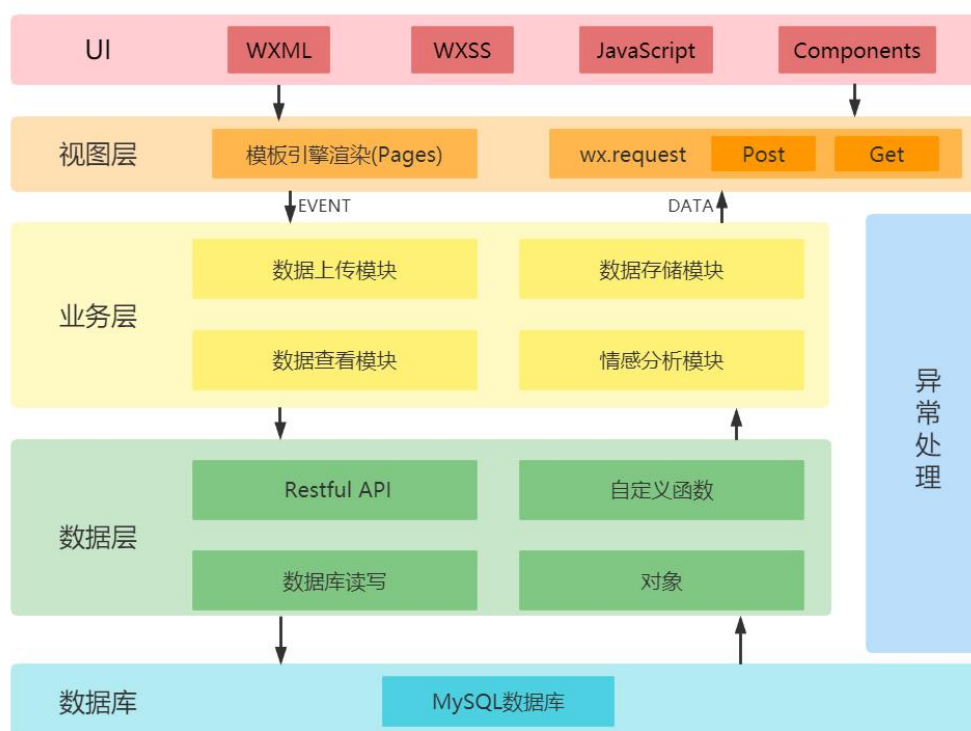


图 1 系统架构图

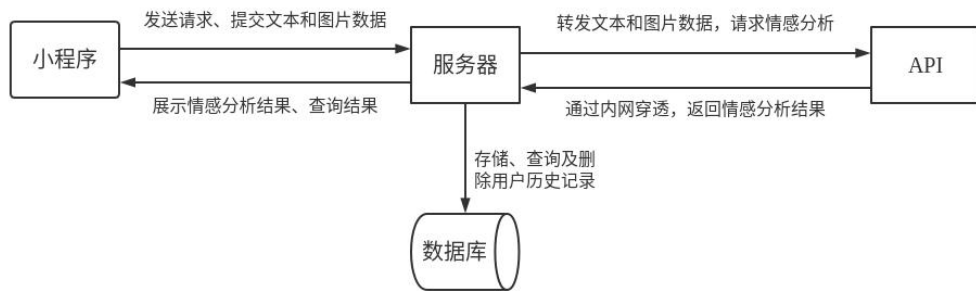


图 2 系统流程图

## 3.2 微信小程序

本节详细介绍了微信小程序的设计，包括 UI 设计、视图层设计、业务逻辑层设计以及数据层设计。

### 3.2.1 UI 设计

微信小程序的 UI 设计追求简洁易用，主要包括以下几个界面，以使用户能够轻松地使用本系统进行情感分析：

(1) 登录界面：用户可以使用微信账号进行登录，以便将分析结果与个人信息关联。登录界面的设计简洁明了，突出登录按钮，避免了冗余元素，确保用户能够快速进入小程序。

(2) 主界面：用户可以在此界面上传文本和图片数据，点击“提交”按钮进行情感分析。为了确保用户能够轻松上手，主界面的设计风格统一且清晰，采用醒目的颜色搭配与字体设计。同时，主界面的布局注重易用性，输入框、上传图片按钮和分析按钮都易于找到和操作。

(3) 用户分析记录界面：用于显示该用户之前分析的结果。此界面列出用户过去的情感分析记录，按照时间顺序排列，方便用户快速查看和回顾。每条记录展示简要的情感类型（积极、消极或中性）。



图 3 小程序预测界面



图 4 小程序记录界面

### 3.2.2 视图层设计

视图层主要负责界面的渲染，包括以下几个部分，以实现系统功能的高效呈现：

(1) 登陆界面：使用微信提供的登录组件，用户可以快速登录小程序。为了确保用户能够顺利登录，视图层需要正确处理微信登录组件的回调，将用户信息传递给业务逻辑层。

(2) 主界面：包含用于输入文本和上传图片的表单，以及用于触发情感分析的按钮。视图层实现了表单的响应式设计，以适应不同尺寸的屏幕。上传图片功能能够处理不同格式和尺寸的图片，并在上传前进行压缩，以优化传输速度。

(3) 用户分析记录界面：展示用户过去的情感分析结果。当用户按下记录按键时便会访问服务端获取用户历史记录。

### 3.2.3 业务逻辑层设计

业务逻辑层是整个系统的核心部分，负责处理用户输入的数据、与后端进行交互以及协调视图层的显示。业务逻辑层主要包括以下功能模块，确保用户可以顺利完成情感分析任务：

(1) 登录模块：用户使用微信账号登录，业务逻辑层需要处理微信登录组件的回调，并将用户信息发送给后端。后端通过微信 API 获取用户的 openid 作为唯一标识，确保每位用户的分析记录都能准确关联到相应的账号。同时，业务逻辑层可以处理登录异常，例如网络问题或者用户拒绝授权，以优化用户体验。

(2) 上传数据模块：用户在主界面上传文本和图片数据，业务逻辑层将数据发送至后端进行情感分析。在分析完成后，后端将分析结果返回给业务逻辑层，然后业务逻辑层将情感类型显示在主界面上。

(3) 获取历史分析记录模块：用户可以在用户分析记录界面查看历史分析结果。业务逻辑层需要从后端获取分析结果数据，并按照记录的 id 排序，即按照时间先后顺序展示。

(4) 删除记录模块：用户可以在用户分析记录界面选择对记录进行删除。业务逻辑层会将 record\_id 发送到后端，在数据库中进行删除。之后会重新加载历史分析记录界面，以显示最新内容。

### 3.2.4 数据层设计

数据层作为整个系统的基础，主要负责与后端进行数据交互和管理本地数据存储，确保数据在各个模块之间顺畅传递。数据层主要包含以下功能：

(1) 登录模块：数据层负责调用微信提供的登录 API，获取用户的 code，并向后端发送 code。后端根据 code 获取用户的 openid 作为唯一标识，以便在之后的数据交互中识别用户身份。

(2) 上传数据模块：数据层对用户输入的文本和图片数据进行预处理，包括对文本进行编码以及对图片进行压缩。预处理完成后，将数据封装为 FormData 对象，通过 POST 请求发送至后端。同时，数据层需要处理网络异常和数据上传失败的情况，确保用户能够顺利完成数据上传操作。

(3) 获取历史分析记录模块：数据层负责从后端获取用户的历史分析记录，通过 GET 请求与后端交互。接收到后端返回的数据后，数据层对数据进行解析和排序，按照时间先后顺序传递给业务逻辑层进行展示。同时，数据层还会处理网络异常和数据获取失败的情况，确保用户能够顺利查看历史分析记录。

(4) 删除记录模块：数据层负责获得记录对应的 record\_id，通过 POST 方式向后端请求，告知数据库应该对这条记录进行删除。如果未收到 record\_id，则会返回报错信息以及 400，如果成功删除则会返回删除成功，以便用户知道删除任务已完成。

## 3.3 数据库设计与实现

本系统使用 MySQL 数据库对用户数据和分析结果进行存储。数据库设计包括以下几个

表:

users 表: 存储用户的唯一标识 (openid)、昵称等信息。

user\_records 表: 存储用户提交的文本、图片文件名、情感分析结果等信息。

具体表结构如下:

表 1 数据库 user 表结构

字段名	类型	描述
openid	varchar(64)	用户唯一标识
nickname	varchar(64)	用户昵称

表 2 数据库 user\_records 表结构

字段名	类型	描述
id	int	记录编号
user_id	varchar(255)	用户唯一标识
text	text	用户输入文本
image	varchar(255)	图片文件名
api_result	varchar(32)	情感分析结果

后端采用了轻量级的 Flask 框架和 pymysql 库来实现与数据库的交互,确保数据传输的安全性和高效性。后端主要提供以下功能:

(1) 用户身份识别与管理: 后端接收来自数据层的用户 code, 通过调用微信 API 获取用户的 openid, 并将其作为唯一标识。在获取 openid 后, 后端检查 users 表中是否已存在该用户。用户已经存在的, 会对用户资料进行更新; 如果用户不存在, 在 users 表中加入新用户。这一过程确保了用户信息的准确性和完整性。

(2) 数据存储与情感分析结果记录: 后端接收来自数据层的文本和图片数据, 并将其存储到 user\_records 表中。在数据存储的过程中, 后端还需要调用情感分析模块, 对用户上传的数据进行情感分析。分析完成后, 后端将情感分析结果一同存储到 user\_records 表中, 便于后续的数据查询和展示。

(3) 历史分析记录查询与提取: 后端根据用户的 openid 查询 user\_records 表, 获取用户的历史分析记录。在查询过程中, 后端需要按照时间顺序对记录进行排序, 确保最新的分析记录排在最前面。查询完成后, 后端将查询结果返回给数据层, 以供前端展示。

(4) 历史分析记录删除: 后端根据用户发送的 record\_id 查询 user\_records 表, 查找对应的记录。如果存在就删除这一条记录, 最后将删除结果返回给数据层。

### 3.4 小结

本章介绍了一种基于微博社交媒体数据的多模态情感分析系统的设计。首先, 概述了整个系统的架构, 然后详细介绍了微信小程序的设计, 包括 UI 设计、视图层设计、业务逻辑层设计以及数据层设计。最后, 介绍了数据库的设计。

本系统采用微信小程序作为前端界面, 搭建了简洁易用的用户界面, 用户可以通过上传

文本和图片进行情感分析。后端采用 Flask 框架和 MySQL 数据库，实现了用户数据和分析结果的存储和查询功能。本地 API 通过内网穿透的方式提供情感分析服务，处理用户上传的文本和图片数据，并返回分析结果。

## 第四章 数据收集与处理

### 4.1 新浪微博爬虫实现

为了从新浪微博上采集数据，本研究首先实现了一个功能完善且稳定的爬虫。该爬虫的核心任务是自动访问新浪微博页面，抓取指定网页下的微博数据，并提取所需要的信息。爬虫实现过程中加入了代理池、模拟浏览器请求、设置随机访问区间等技术，保证了数据采集的稳定与完整，以确保数据采集的稳定性和完整性。

爬虫实现过程如下：

(1) 利用 Python 编写爬虫程序，选用了成熟的爬虫库 requests 来执行网络请求和解析页面。

(2) 通过正则表达式解析页面源代码，提取所需的微博内容、发布时间、用户信息等。

(3) 设定爬虫的访问频率限制，以防止因过于频繁访问而遭到封禁，同时保证数据采集的效率。

(4) 使用代理池技术，收集大量可用代理 IP，并定期检测代理的有效性，确保爬虫在发送请求时可以随机选择一个可用代理，从而降低被封禁的风险。

(5) 模拟浏览器请求，自动切换不同浏览器的 User-Agent，使得每次请求都具有独特的特征，降低被识别为爬虫的概率。

(6) 设置随机访问间隔，根据设定的时间范围内生成随机的等待时间，避免过于规律的访问模式被识别为爬虫行为。

(7) 将抓取到的数据以 CSV 格式进行存储，方便后续数据分析和处理。

整个爬虫的流程如下：

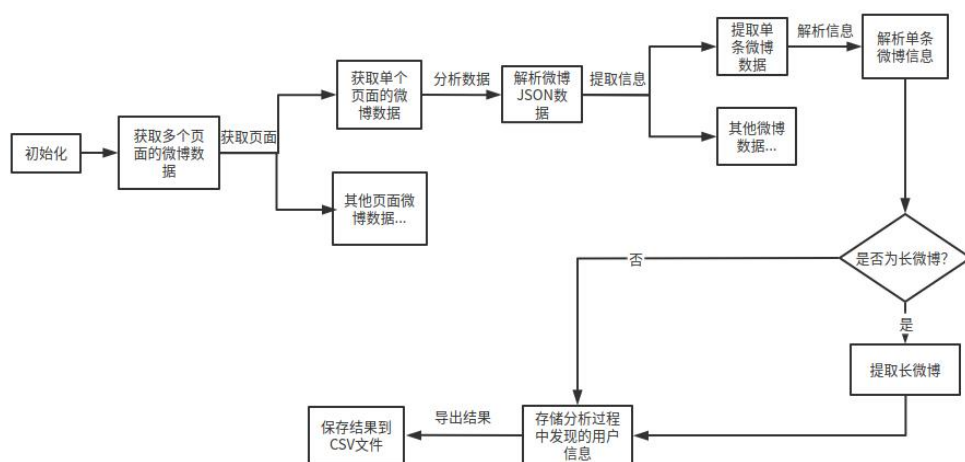


图 5 新浪微博爬虫流程图

(1) 初始化：在 Spider 类的 `__init__` 方法中，本研究设置了请求头 (headers)、Cookie 信息以及代理池的初始化。同时，创建了一个空列表用于存储爬取到的所有微博信息。

(2) 获取多个页面的微博数据：在 `get_pages` 方法中，本研究遍历指定的页数范围，分析每个页面的 URL 结构，并调用 `get_one_page` 方法获取每一页的微博数据。最后将所有结果保存到 CSV 文件中。

(3) 获取单个页面的微博数据：在 `get_one_page` 方法中，本研究首先调用 `get_weibo_json` 方法获取页面中的微博 JSON 数据。然后解析 JSON 数据，提取微博信息，并将结果添加到 `page_content` 列表中。在发送请求时，从代理池中随机选取一个可用代理 IP，并设置随机访问间隔，以降低被封禁的风险。

(4) 解析微博 JSON 数据：在 `get_weibo_json` 方法中，本研究构造请求参数，并结合代理池与模拟浏览器请求技术，使用 `requests` 库发送 GET 请求，获取微博数据并将其解析为 JSON 格式。

(5) 提取单条微博数据：在 `get_one_weibo` 方法中，本研究从 JSON 数据中提取微博信息，包括微博正文、图片 URL、点赞数、转发数、评论数等。对于长微博，本研究调用 `get_long_weibo` 方法获取完整的微博内容。

(6) 解析单条微博信息：本研究在 `parse_weibo` 方法中，从微博信息中提取用户 ID、昵称、微博 ID、正文、图片 URL、点赞数、转发数、评论数等所需字段，并进行数据清洗，去除乱码、特殊字符等，确保数据准确无误。

(7) 获取长微博内容：在 `get_long_weibo` 方法中，本研究发送 HTTP 请求获取长微博完整内容，同时将长微博中的图片网址、点赞、转发、评论等信息进行解析，并将完整的长微博内容与其他信息一起存储。

(8) 保存结果到 CSV 文件：在 `SaveCSV` 方法中，本研究将提取到的微博信息和用户信息分别保存到不同的 CSV 文件中，例如：`Spider_weibo.csv` 和 `Spider_user.csv`。通过设置 CSV 文件的编码格式和字段分隔符，确保数据的可读性和通用性。

(9) 异常处理和日志记录：在整个爬虫过程中，本研究对可能出现的异常情况进行了处理，并将异常信息和运行日志记录到本地文件中，便于后续分析、调试和优化爬虫性能。

通过以上实现过程和流程，本研究构建了一个功能完善、稳定且高效的新浪微博爬虫，可以有效应对新浪微博的反爬虫机制，从而确保数据采集的稳定性和完整性，为后续的数据分析和研究提供有力支持。

## 4.2 数据处理

为了提高模型的准确性和可靠性，本研究对收集到的新浪微博数据进行了一系列的清洗与预处理工作。由于原始数据中存在许多不符合研究需求或包含噪声的数据，本研究通过以下方法对数据进行优化：

### 4.2.1 数据清洗

(1) 去除重复数据：本研究对收集到的数据进行去重处理，确保数据集中每条记录都是独一无二的，以降低数据冗余度和提高数据质量。

(2) 删除无关信息：本研究删除与研究无关的内容，例如广告、推销信息、无意义的符号等。这一步骤有助于提高数据的实际有效性，使得后续分析更加精确。

(3) 剔除异常数据：针对包含异常数据的记录，如图片无法正常加载、文本内容为空等，本研究将其移除，以减少异常数据对模型的影响。

(4) 表情符号处理：由于微博文本中可能包含大量表情符号，但本研究中不使用这部分信息，因此本研究删除标签符号，使得文本更加简洁明了。

(5) 去除 URL、@用户和#话题#：本研究使用正则表达式和 `re` 库中的字符串处理函数去除文本中的 URL、@用户和#话题#标签，进一步提高文本的简洁性和可读性。

(6) 仅保留一对一的数据：在本研究中，本研究仅探讨一段文本对应一个图片的情况，因此需要删除多张图片、没有图片、没有文本的情况。这一步骤有助于保证数据集的一致性。

(7) 图片文本匹配判定：在数据清洗过程中，本研究还需删除因网络原因导致文本或者图片有一方丢失的数据，以确保数据集中的图片和文本完整匹配。

### 4.2.2 数据预处理

(1) 文本进行分词：本研究采用中文分词工具 `jieba` 对微博文本进行分词处理，然后去除停用词（如“的”、“和”等常用词），提取关键词和短语。这一步骤有助于提高文本分析的准确性。

(2) 图片处理：对于收集到的图片数据，本研究进行预处理操作，包括缩放、裁剪等，



以保持图片的统一尺寸。这有助于提高后续图片分析和处理的效率。

(3) 标签处理：针对图片或文本的标签，本研究进行处理以适应模型的需求。具体来说，这项研究对标签进行了独热编码（One-Hot Encoding），这样可以将类别型标签转换为二进制向量表示。独热编码有助于模型更好地理解 and 处理类别数据，进而提高模型的准确性和效率。

(4) 多模态数据合成：这种方法是一种结合了文本和图像数据的生成技术。对于每个类别，从原始样本中随机选择一组文本和图像，然后将它们组合成新的多模态样本。例如，将一个正面情感的文本与一个正面情感的图像组合，以生成一个新的正面情感样本。这样在保持原始样本语义信息的同时，可以创建新的合成样本，以平衡不同情感倾向数据的数量。

(5) 划分数据集：根据情感标签和数据类型将数据分组，然后在每个组中进行训练、验证和测试集的划分。确保每个数据集具有相似的情感分布和原始/合成数据比例。

### 4.3 数据标注

为了对数据集进行更为精确的情感分析，本研究需要对数据集中的文本和图片内容进行情感标注。而多模态情感标注因为包含图片这种媒体资源，无法在 csv 中直接直观的进行标注。因此，在这一过程中，本研究基于 Kivy 库编写了一个自定义的数据标注程序，用以对数据集中的文本和图片数据进行可视化的情感标注。用户可以选择 photo、text 以及 total 对应的三分类情感，点击 next 即可进行下一组数据的标注。此程序旨在帮助用户更方便地为文本和图片数据分别标注负面、中性或正面的情感。



图 6 数据标注程序界面

在标注过程中，用户可以针对每个数据样本的文本和图片内容分别进行情感标注，同时为整个数据样本标注一个综合性情感评价。这样的操作不仅可以提高标注的精确度，还能够确保在多模态情感分析中，各种情感因素得到充分的考虑。为了便于后续处理和分析，将这些标注的情感数据保存到名为“tag.csv”的文件中。

通过这个自定义的数据标注程序，本研究我们最终成功标注了一份规模为 1500 组的新浪微博数据集，名为 Weibo2023。为了方便后续处理和分析，我们将这些标注的情感数据保存到名为“tag.csv”的文件中。这些数据将为训练多模态情感分析的深度学习模型提供有力

支持，提高模型的预测准确性。

#### 4.4 小结

本部分的研究主要聚焦于新浪微博数据的获取、处理和标注过程。首先，通过构建高效稳定的爬虫，实现了对新浪微博上的数据自动采集爬虫。使用了代理池、模拟浏览器请求和随机访问间隔等技术，以保证数据采集的稳定性和完整性。

其次，对收集到的数据进行了一系列的清洗与预处理工作，包括去除重复数据、删除无关信息、剔除异常数据、处理表情符号、删除 URL 和@用户，以及保留一对一的数据，从而提高数据质量和实际有效性。并且对文本进行了分词处理，图片进行了预处理，标签进行了独热编码，以提高模型的准确性和效率。

最后，利用基于 Kivy 库的自定义数据标注程序，对数据集中的文本和图片数据进行了情感标注，提高了标注效率和准确度。还成功标注了一份规模为 1500 组的新浪微博数据集。此过程为后续数据分析和研究提供了有力的支持，确保了后续分析的准确性和可靠性。

## 第五章 图文融合多模态情感分析模型

### 5.1 模型架构

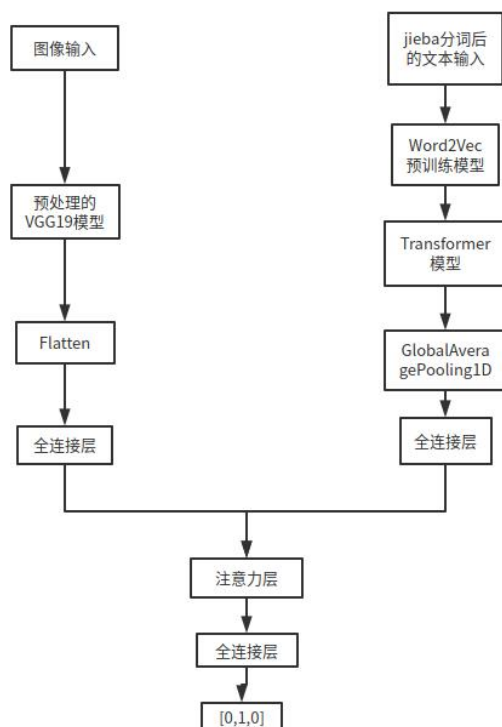


图 7 本研究模型架构图

本项目中的图文融合多模态情感分析模型通过结合文本和图像特征进行情感分类任务。

图像特征模型以图像作为输入，然后通过预训练的 VGG19 模型提取图像特征。接着，使用 Flatten 层将特征展平，然后通过全连接层（Dense）得到图像特征输出。

文本特征模型以 jieba 分词后的文本经过 Word2Vec 预训练模型得到的单词向量为输入，使用基于 Transformer 的文本特征模型来获得文本特征输出。

在特征融合上，采用注意力层来自适应权重融合特征，最后经过四个全连接层（Dense）来获得最终的 3 个类别的概率输出。

### 5.2 图像特征提取

#### 5.2.1 预训练的 VGG19 模型

本项目的图像特征模型以（224，224，3）的图像作为输入，并通过预训练的 VGG19 模型提取图像特征。VGG19 是一个经典的卷积神经网络（CNN），在许多图像分类任务中表现优异。

#### 5.2.2 Flatten 层和全连接层

在通过预训练的 VGG19 模型提取图像特征后，本项目使用 Flatten 层将特征展平，以便将其输入到后续的全连接层（Dense）。展平操作是将多维特征映射为一维向量，以便于进一步处理。接着，通过全连接层（Dense）对展平后的图像特征进行处理，以获得最终的图像特征表示。

### 5.3 文本特征提取

#### 5.3.1 单词向量输入

在本研究中，输入的单词向量的维度为  $\text{MaxWordNum} * \text{WORD\_SHAPE}$ 。这些单词向量是通过预训练的中文维基 word2vec 模型生成的。在保留语义关系的同时，这种模式能够对高维空间中的汉语词汇进行有效映射。

### 5.3.2 自定义 Transformer 编码器

本研究使用了一个自定义的 Transformer 编码器来处理输入序列。这个编码器的关键部分包括位置编码（PositionalEncoding）层，多头自注意力（MultiHeadAttention）层，以及层归一化（LayerNormalization）层。

位置编码层用于在模型中注入词序列的位置信息。在没有明确顺序依赖性的模型中，为了使模型能够捕捉到文本中的顺序信息，本研究在输入单词向量后加入了位置编码层。该层通过在每个单词向量上添加一个由正弦和余弦函数生成的位置向量。

多头自注意力层是 Transformer 编码器的核心，它允许模型在处理输入序列时，能够考虑到输入中的不同特征。而层归一化则用于规范化多头自注意力层的输出，使其保持适当的尺度和分布。通过这三个关键组件的结合，自定义的 Transformer 编码器能有效地处理并理解输入序列。

### 5.3.3 全局平均池化和全连接层

在获取自定义 Transformer 编码器的输出后，本研究采用了全局平均池化（GlobalAveragePooling1D）层对特征向量进行池化。该层能够降低模型复杂性，提高计算效率，也有助于缓解过拟合。经过全局平均池化后，本研究使用全连接层（Dense）对池化后的输出向量进行处理，以得到最终的文本特征表示。

## 5.4 图文多模态特征融合

### 5.4.1 注意力层

在融合文本和图像特征时，本研究使用自定义的加性注意力层（AdditiveAttention）来实现注意力机制。该层能够帮助模型学习在整合特征时如何分配不同权重，从而让模型能够关注更为重要的信息。

在加性注意力层中，首先计算文本和图像特征之间的相关性。然后，通过 Softmax 激活函数将相关性转换为注意力权重。利用这个注意力权重，将文本和图像特征加权相加，生成一个上下文向量。这个向量包含了模型认为对情感分类最重要的信息。最后将这个向量与文本特征和图像特征连接起来，形成一个融合特征向量。

### 5.4.2 全连接层及分类输出

使用三个全连接层（Dense）对融合特征向量进行处理，激活函数均为 relu。在最后一个全连接层中，使用 Softmax 激活函数得到 3 个类别的概率输出，进行情感三分类。

## 5.5 模型训练与优化

### 5.5.1 损失函数与优化器

整个模型使用类别交叉熵（categorical\_crossentropy）作为损失函数。这个损失函数在多分类任务中表现优异，因为它可以衡量模型预测的概率分布与真实标签的概率分布之间的差异。在优化器方面，本模型选择了表现稳定的 Adam 优化器。这种优化器可以自适应地调整学习率，从而加速训练过程并提高模型的收敛速度。

### 5.5.2 回调函数

为了防止模型过拟合，在训练过程中采用了早停（EarlyStopping）策略来作为回调函数。早停策略可以在验证集上的性能不再提升时提前终止训练过程。这种策略能够避免模型在训练数据上过度拟合，从而提高模型在未知数据上的泛化能力。

## 5.6 小结

本章详细讲解了图文融合多模态情感分析模型的设计。首先，介绍了模型的总体架构，该模型通过融合从图像和文本中提取的特征来进行情感分类。对于图像特征，本研究利用预训练的 VGG19 模型进行提取，然后通过 Flatten 层和全连接层进行处理。对于文本特征，本研究采用基于 Transformer 的模型，首先将 jieba 分词后的文本通过 Word2Vec 预训练模型得到单词向量，然后进行特征提取。

在特征融合阶段，本研究采用了注意力机制，通过加性注意力层来自适应地调整文本和图像特征的权重，以生成一个融合特征向量。接着，该向量经过多个全连接层进行处理，最终得到三个类别的概率输出，完成情感分类任务。

在模型训练和优化部分，本研究选择了类别交叉熵作为损失函数，并使用了 Adam 优化器来提高模型的收敛速度。为了防止过拟合，本研究还引入了早停策略，通过在验证集上的性能不再提升时提前终止训练，以提高模型在未知数据上的泛化能力。

## 第六章 实验与结果分析

### 6.1 实验环境与数据集

#### 6.1.1 实验环境

本实验在 AutoDL 平台进行，使用的 GPU 型号为 RTX A5000。开发环境在 Ubuntu18.04 上进行，下面是本实验环境的关键版本配置：

```
Python == 3.8,  
Tensorflow == 1.15.5,  
Cuda ==11.4,  
numpy == 1.19.2,  
pandas == 1.3.3
```

#### 6.1.2 数据集

本实验训练使用的数据集为从 kaggle 平台获取的 Weibo nCoV Data 数据集。这个数据集的原始规模为 100000 组，每一组数据包括微博 id、微博发布时间、发布人账号、微博中文内容、微博图片、微博视频、情感倾向(使用-1、0、1 分别表示消极、中性、积极)。经过数据清洗后得到符合本实验需求的数据共有 44177 组，每一组数据包括微博中文内容、微博图片以及情感倾向。最后经过多模态数据合成，得到的数据共有 81198 组。

本实验还使用了自己标注的一个数据集作为验证集，名为 Weibo 2023 Data 数据集。这个数据集的规模为 1500 组。每一组数据包括微博中文内容、微博图片、以及情感倾向(使用-1、0、1 分别表示消极、中性、积极)。

### 6.2 实验指标

本实验采用以下指标来评估模型的性能：

(1)准确率 (Accuracy, 简称 ACC) :正确分类的样本数占总样本数的比例。用于衡量模型在所有类别上的整体性能。对于分类问题，如本实验的情感分析任务，准确率可以直观地反映模型的分类能力。

(2)损失 (Loss) : 模型预测结果与真实标签之间的差异。损失函数用于衡量模型的预测误差，它是优化算法在训练过程中试图最小化的目标。在本实验的情感分析任务中，采用交叉熵损失作为损失函数。较低的损失值通常表示模型的预测效果更好。

### 6.3 模型性能评估

#### 6.3.1 模型最终性能

(1)训练过程表现：

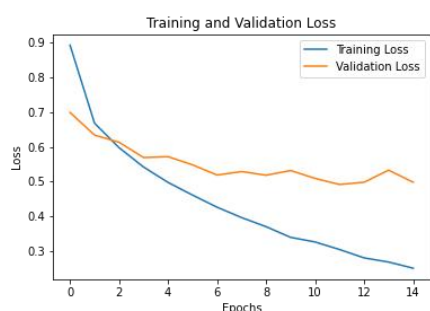


图 8 本模型 Loss 对比

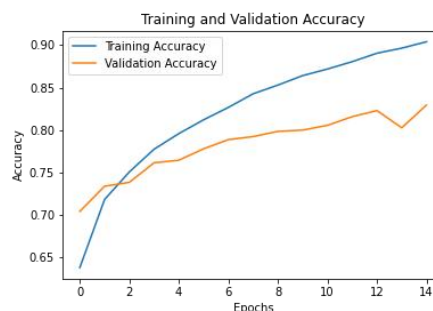


图 9 本模型 Acc 对比

训练集准确率从 0.6382 增加到 0.9040，而验证集准确率从 0.7043 增加到 0.8295。这表明模型在训练过程中逐渐提高了分类性能。训练集的精确度达到了 90.4%，但验证集的准确率最终为 82.95%。

训练集损失从 0.8924 降低到 0.2516，验证集损失在 0.4921 至 0.6991 之间波动，最终

为 0.4987。训练集合损失的减少，表明模型逐步拟合了训练数据。

模型在训练过程中的准确率和损失表现良好，模型出现了一定程度的过拟合，验证集损失的波动表明模型的泛化性能可能会有一些影响。

(2) 在 Weibo nCoV Data 数据集上表现：

使用 Weibo nCoV Data 数据集上训练之前划分的验证集来进行预测，这部分验证集的规模为 12179 组，Acc 准确率约等于 0.823。这表明模型在这个数据集上具有较好的分类性能。然而，这个验证集是从训练过程中划分出来的，可能与训练数据具有相似的分布特征。

(3) 在 Weibo 2023 Data 数据集上表现：

用自己标注的 Weibo 2023 Data 数据集来进行预测，这部分数据集的规模为 1500 组，Acc 准确率约等于 0.484。模型在这个新的数据集上的泛化性能较差，训练数据和测试数据之间的分布差异较大，导致模型无法很好地泛化到新的数据集。

6.3.2 对比实验

表 3 对比试验结果

准确率对比	Acc(Weibo nCoV Data)	Acc(Weibo 2023 Data)
VGG19	0.573	0.382
Transformer	0.789	0.468
VGG19+Transformer+Concatenate	0.802	0.474
VGG19+Transformer+Attention	0.823	0.484

为了验证本研究提出的 VGG19+Transformer+Attention 模型在处理微博数据上的性能，本研究设计了一系列对比实验。

(1) VGG19 模型依赖于图像特征进行分类。在 Weibo nCoV 数据集上，其准确率为 0.573，而在 Weibo 2023 数据集上为 0.382。这可能是因为微博数据中的图片和情感之间的关联性不如文本和情感之间的关联性明显。此外，微博中的图片可能存在很大的多样性和噪声，这也可能影响到了模型的性能。

(2) Transformer 模型仅依赖文本特征进行分类。相较于 VGG19 模型，Transformer 在这类任务中表现更好。在 Weibo nCoV 数据集上，准确率达到 0.789，而在 Weibo 2023 数据集上为 0.468。Transformer 模型主要依赖于文本特征进行分类。从实验结果可以看出，Transformer 模型在这个任务上的表现要优于 VGG19 模型。这说明了文本特征在情感分类任务中的重要性，因为文本通常能够更直接地反映出作者的情感倾向。

(3) VGG19+Transformer+Concatenate 模型将 VGG19 和 Transformer 的输出进行拼接，充分利用图像和文本特征。在 Weibo nCoV 数据集上，该模型的准确率为 0.802，在 Weibo 2023 数据集上为 0.474。此模型尝试结合图像特征和文本特征进行分类，其性能优于仅使用单一特征的模型。这表明，尽管图像特征在这项任务中的作用可能不及文字特征，但是当它与文本特征结合时，仍然可以提供有价值的补充信息，帮助提升模型的性能。

(4) VGG19+Transformer+Attention 模型采用注意力机制融合 VGG19 和 Transformer 的输出。在 Weibo nCoV 数据集上,该模型的准确率为 0.823,在 Weibo 2023 数据集上为 0.484。在所有实验中,这个模型的表现最好。这主要归功于 Attention 机制的作用,它可以自动学习到不同特征之间的关联性,根据这些关联性分配不同的权重,这样模型就可以更加聚焦于对分类结果影响最大的特征,从而提升性能。从实验结果来看,尽管该模型在 Weibo2023 数据集上的表现并没有达到很高的准确率,但是相比其他模型仍然有明显的优势,这表明 Attention 机制在处理此类复杂任务时的有效性。

## 6.4 小结

本章主要详细介绍了实验的环境,数据集,指标,模型性能评估和结果分析。首先,本研究描述了实验的环境和使用的数据集,包括从 Kaggle 平台获取的 Weibo nCoV Data 数据集以及自己标注的 Weibo 2023 Data 数据集。然后,本研究选择了准确率和损失两种指标来评估模型的性能。

本研究评估了 VGG19 模型、Transformer 模型、VGG19+Transformer+Concatenate 模型以及 VGG19+Transformer+Attention 模型在这两个数据集上的性能。通过比较,本研究发现,虽然单一特征的模型(如 VGG19 模型和 Transformer 模型)在任务上有一定的效果,但结合多种特征的模型(如 VGG19+Transformer+Concatenate 模型和 VGG19+Transformer+Attention 模型)在任务上的性能更好。尤其是 VGG19+Transformer+Attention 模型,凭借 Attention 机制的优势,其在处理微博数据的性能最优。

尽管本研究的模型在 Weibo nCoV Data 数据集上的性能较好,但在新的 Weibo 2023 Data 数据集上的泛化性能较差。这可能是因为训练数据和测试数据之间的分布差异较大,导致模型无法很好地泛化到新的数据集。

总的来说,本章通过详细的实验和结果分析,展示了本研究的 VGG19+Transformer+Attention 模型在微博情感分类任务上的优越性,同时也指出了其在处理新的数据集时可能存在的问题。这些发现对于本研究进一步改进模型和优化任务策略有着重要的指导意义。



## 第七章 结束语

### 7.1 研究总结

本论文主要研究了基于微博社交媒体数据的多模态情感分析,旨在设计并实现一个高效准确的情感分析系统,以识别与分析社交媒体中的文本和图像内容。为实现这一目标,论文从理论和实践两个方面展开研究。

首先,本研究通过回顾国内外情感分析研究现状,对多模态情感分析方法、图像特征提取、文本特征提取、深度学习模型等方面进行了系统梳理。其次,论文详细介绍了新浪微博数据的爬取和预处理过程,为后续分析奠定了基础。接着,论文设计并实现了一个基于微博社交媒体数据的多模态情感分析系统,包括了微信小程序的设计与实现。此外,本研究还详细阐述了图文融合多模态情感分析模型的架构、特征提取、特征融合与情感分类等方面,并对模型进行了训练与评估。

### 7.2 创新与贡献

(1) 将多模态情感分析应用于微博社交媒体数据,实现了对文本和图像信息的综合分析,提高了情感分析的准确性和效果。

(2) 设计并实现了一个基于微博社交媒体数据的多模态情感分析系统,通过微信小程序为用户提供了一种便捷的情感分析工具。

通过实验与结果分析,证明了本研究提出的多模态情感分析模型在微博社交媒体数据中的有效性和可行性,为后续研究提供了有益的参考。

### 7.3 局限性与未来工作

本研究在数据集泛化性能方面存在一定局限性,主要表现在以下几个方面:

(1) 模型调参问题:模型在训练过程中出现一定程度的过拟合现象。

(2) 使用多模态数据合成平衡数据集:为了平衡数据集中不同情感特征数据的数量,采用了一种结合文本和图像数据的生成技术。这种方法虽然可以创建新的合成样本以平衡不同情感倾向数据的数量,但可能导致生成的样本与真实场景存在差异,从而影响模型的泛化性能。

(3) 数据时效性:nCoV数据集采集自2019年新冠肺炎疫情期间,与当前(2023年)微博热门话题和关注点存在较大差异。这导致训练数据和测试数据分布不匹配,影响了模型在新数据集上的泛化性能。

(4) 数据主题差异:2019年的微博内容主要与疫情相关,而2023年的微博内容几乎不再与疫情密切相关。这种主题差异进一步加大了训练数据和测试数据之间的分布差异。

(5) 社交媒体数据多样性:由于社交媒体上话题和表达方式的多样性,模型在应对新数据集时可能难以捕捉到潜在规律,从而影响泛化性能。

针对上述问题,未来会尝试以下方法以提高模型在新数据集上的泛化性能:

(1) 调整模型参数:针对过拟合问题,尝试继续调整模型参数,提高模型泛化性能。

(2) 优化多模态数据合成方法:在使用多模态数据合成技术时,可以尝试改进生成策略,以生成更接近真实场景的样本。此外,可以探索其他数据平衡方法,如重新采样或使用类别权重,以减轻合成数据可能对模型泛化性能的影响。

(3) 更新训练数据:收集更多近期的微博数据作为训练数据,以便模型更好地适应新的数据分布,提高泛化性能。

(4) 数据增强:采用数据增强技术,在现有数据集基础上生成具有更高多样性的训练样本,提高模型泛化能力,使其能适应更多不同类型的数据。

(5) 结合其他信息来源:除微博文本和图像数据外,尝试引入其他信息来源,如用户画像、话题关联度等,以提供更丰富的上下文信息,帮助模型更好地捕捉潜在规律,提高泛化

性能。

## 致谢

在这几个月完成毕业设计的过程中,我深感到学术研究的严谨性与探索的挑战性。在此,我要对所有给予我帮助与支持的人表示衷心的感谢。

首先,我要由衷地感谢我的指导老师乔秀明老师,她在我的研究过程中给予了无微不至的关怀和耐心指导。从研究方向的选择、实验设计到论文撰写,导师始终以严谨的治学态度和高度的责任心为我树立了学术榜样。在我遇到困难时,无论有多么晚,导师也总是为我提供思路,帮助我解决难题,让我在这个毕设项目中学到了很多知识。

其次,我要感谢实习期间的同事们还有宿舍的舍友和同学们,在我面临学术困境时,他们总是愿意与我一起探讨研究内容,分享新的思路和心得体会。

此外,我要感谢我的家人对我学业的支持和理解。他们始终关心我的成长,为我提供了坚实的后盾,让我在学术道路上勇往直前。他们的爱和鼓励使我更有信心面对未来的挑战。

再次向所有给予我帮助的人表示衷心的感谢!

## 参考文献

- [1] Nran MT, But INA. Azzal M T. Open source software adoption evaluation through feature level sentiment analysis using Twitter data[J]. Turkish Journal of Electrical Engineering & Computer Sciences, 2015, 24(5): 4481-4496.
- [2] ABDU S A, YOUSEF A H, SALEM A. Multimodal video sentiment analysis using deep learning approaches, a survey[J]. Information Fusion, 2021, 76(2021): 204-226.
- [3] Turney, P. D. Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews. In Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics, July 6-12, 2002, Philadelphia, PA, USA, pp. 417-424. ACL, 2002.
- [4] Andrea Esuli and Fabrizio Sebastiani. 2006. SENTIWORDNET: A publicly available lexical resource for opinion mining. In Proceedings of the Fifth International Conference on Language Resources and Evaluation, pages 417-422.
- [5] Taboada, Maite, Julian Brooke, Milan Tofiloski, Kimberly Voll, and Manfred Stede. Lexicon-Based Methods for Sentiment Analysis. Computational Linguistics, 2011. 37(2): 267-307.
- [6] Cambria E, Hazarika D, Poria S, et al. Benchmarking Multimodal Sentiment Analysis[C]. In: Proceedings of International Conference on Computational Linguistics and Intelligent Text Processing. Springer, Cham, 2017: 166-179.
- [7] Bahdanau D, Cho K, Bengio Y. Neural Machine Translation by Jointly Learning to Align and Translate[C]. In: Proceedings of International Conference on Learning Representations ICLR 2015. San Diego: CA, 2015.
- [8] Zadeh A, Chen M, Poria S, et al. Tensor Fusion Network for Multimodal Sentiment Analysis[OL]. arXiv Preprint, arXiv: 1707.07250, 2017.
- [9] Guo M, Xu T, Liu J, et al. Attention mechanisms in computer vision: A survey[J]. Computational Visual Media, 2022, 8(03): 331-368.
- [10] Zijun Sun, Xiaoya Li, Xiaofei Sun, Yuxian Meng, Xiang Ao, Qing He, Fei Wu, and Jiwei Li. 2021. ChineseBERT: Chinese Pretraining Enhanced by Glyph and Pinyin Information.
- [11] Schmidt S, Stock W G. Collective indexing of emotions in [gnages. A study in emotional information retrieval[J]. Journal of the American Society for Information Science & Technology, 2009, 60(5): 863-876.
- [12] Cao D, Ji R, Lin D, et al. A cross-media public sentiment analysis system for microblog[J]. Multimedia Systems, 2016, 22(4): 479-486.
- [13] Yu Y, Lin H, Meng J, et al. Visual and Textual Sentiment Analysis of a Microblog Using Deep Convolutional Neural Networks[J]. Algorithms, 2016, 9(2): 41.
- [14] Chen X, Wang Y, Liu Q. Visual and textual sentiment analysis using deep fusion convolutional neural networks[C]. 2017 IEEE International Conference on Image Processing (ICIP). IEEE, 2017: 1557-1561.
- [15] You Q, Luo J, Jin H, et al. Joint visual-textual sentiment analysis with deep neural networks[C]. Proceedings of the 23rd ACM international conference on Multimedia. 2015: 1071-1074.
- [16] You Q, Luo J, Jin H, et al. Cross-modality consistent regression for joint visual-textual sentiment analysis of social multimedia[C]. Proceedings of the Ninth ACM international conference on Web search and data mining. 2016: 13-22.

