

ADVERTISING
GENERATION
SYSTEM

基于LVG-DFLA 算法的情感广告 生成系统

汇报人：
沈雅卉

日期：
2025.4.19



目录 CONTENTS



01 PART ONE
研究背景与内容

02 PART TWO
LVG-DFLA模型架构

03 PART THREE
技术创新点

04 PART FOUR
实验结果与分析



研究背景与内容



研究背景

- 全球社交媒体活跃用户达50.4亿，占世界人口62.3%
- 多模态数据量呈爆炸式增长
- 对情感图像信息生成技术需求迫切

研究意义

- 推动情感计算、自然语言处理技术进步
- 降低艺术创作门槛，实现创作自动化
- 助力数字营销转型，增强品牌情感连接

研究目标

构建一个基于情感可控算法的广告作品生成系统，实现从文本到情感图像的智能生成。

系统核心功能

- 输入处理
 - 接收用户文本输入
- 输出结果
 - 生成具有对应情感色彩的不同场景广告图像作品
 - 确保画面真实性和情感表现力

研究背景与内容



技术背景

情感图像技术主要沿着三条技术路线发展：基于GAN的方法、基于扩散模型的方法和基于多模态预训练模型的方法。

深度学习中的生成对抗网络（GANs）是图像生成技术的一个重要里程碑。GANs由生成器和判别器组成，生成器尝试生成逼真的图像来欺骗判别器，而判别器则努力区分真实图像和生成的图像。通过这种相互对抗的训练方式，生成器不断学习和改进，能够生成越来越逼真的图像。

相比于传统的生成对抗网络，扩散模型具有以下优势：能够生成细节丰富、逼真度高的图像；训练过程更加稳定，避免了GANs中常见的模式崩溃问题；可以生成多样化的图像，同时保持与文本描述的一致性。

多模态预训练模型（CLIP为例），它可以理解文本与图像之间的关系。CLIP通过对比学习，在大规模无标注数据集上进行了预训练，可以高效地计算文本描述与图像的相似度，为指导图像生成提供了强大工具。

研究背景与内容

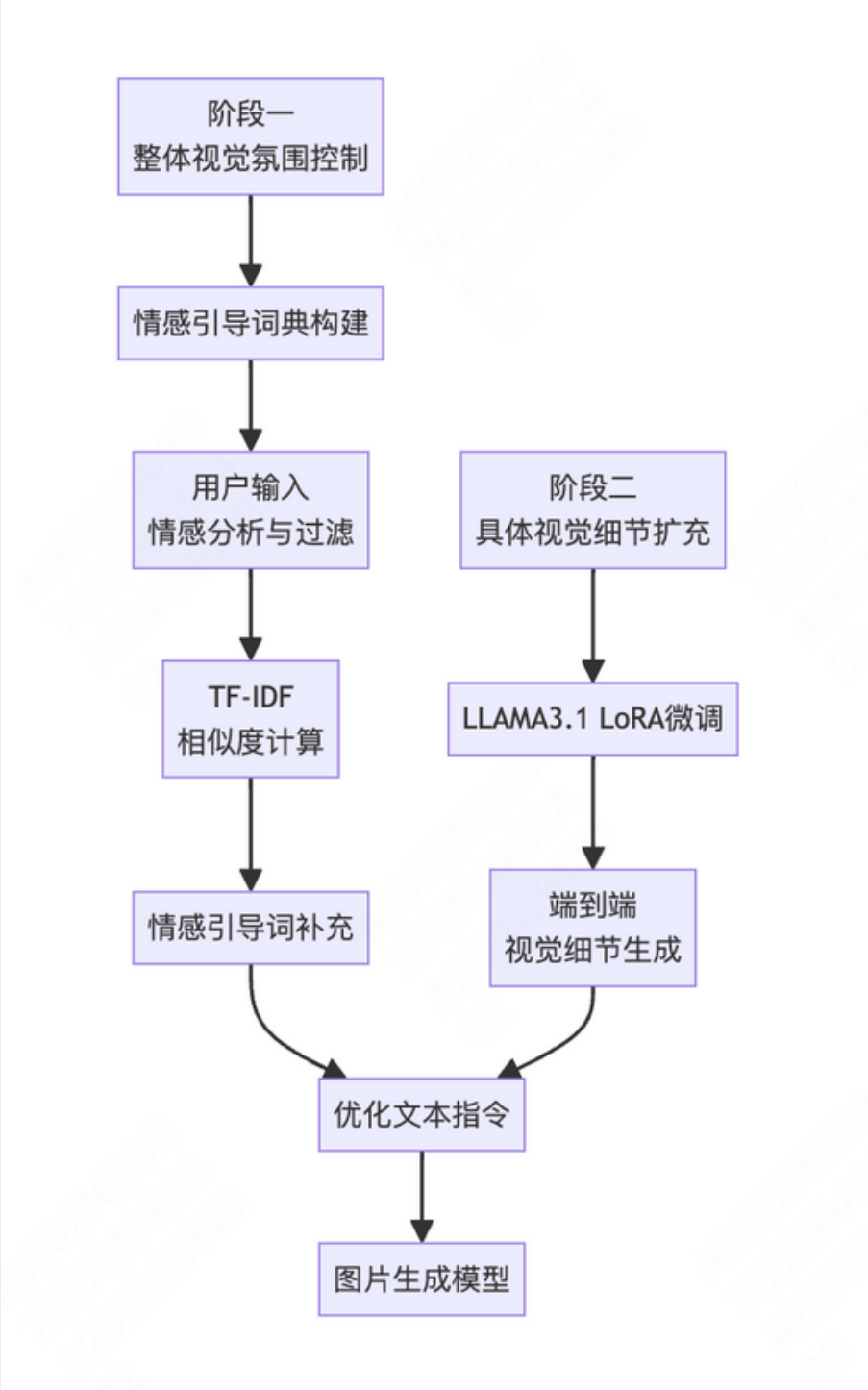


研究内容

Baseline：视频生成算法 FG-ECVG可以解决大模型生成的视觉内容情感属性模糊、交互性弱的问题,可实现文本指令到视频内容的高情感可控、强交互性自动化生成。该 算法包括两部分:一是基于效价-唤醒-控制(VAD)情感模型构建一个引导词典,并基于该词典对输入文本做情感极性 分析和情感引导词匹配,实现了整体视觉氛围的情感控制;二是基于检索-增强-生成(RAG)算法构建一个视觉细节扩 写框架,为用户输入的文本指令添加结构化的类人情感视觉元素,提升了生成内容的情感颗粒度。

为了得到情感表达更加准确的高质量图片内容,本研究提出基于文本指令优化的图片生成算法LVG-DFLA,在原模型基础上,采用LoRA监督微调Llama3.1-8B-Instruct模型代替RAG算法进行视觉细节扩充,进一步提升画面的真实性和情感表现力, 实现了从用户输入文本和指定情感类别到具有特定情感的图片内容的高情感可控、强交互性自动化生成,为情感图片内容创作提供了高效可靠的技术支持。

LVG-DFLA模型架构



- 基于VAD情感模型构建情感引导词典
- 利用NRC VAD词典和DiffusionDB数据集
- 通过TF-IDF算法实现文本相似度匹配

- 采用Llama3.1-8B-Instruct模型
- 使用LoRA技术进行高效微调
- 实现端到端的视觉细节描述生成

● 整体视觉氛围控制阶段

● 具体视觉细节扩充阶段

整体视觉氛围控制阶段



构建情感引导词典



计算TF-IDF向量



获取相似度前3的情感引导词



输出prompt+引导词

```
affective_dictionary.csv

1 emotion,affective_word,related_prompt
2 anger,destructive,['smashed glass mirror window, glass flying and moving, face in reflection of glass flying towards camera, destructive
3 anger,sinister,['a dark sinister creature with a noose for one hand and a hook for another realistic cinematic 3 5 mm ', 'a monster huge
4 anger,drastic,['the water levels along the coast are seen to be drastically receded, with ships stuck on the sand and docks a good dista
5 anger,bestial,['the strange mystical scene of human transformation into a perfect bestial cyborg 35mm doubleexposure photo mixed with ex
6 anger,brutal,['photograph, futuristic scene, interstellar, brutalist style architecture, straight edges, space, finely detailed oil pair
7 anger,eccentric,['kawaii pastel goth guy with dreads and eccentric clothing by jamie hewlett and artgerm, cel shading, toon shading, def
8 anger,bloody,['8 k made in abyss chubby, bear, wearing military uniform, bloody sharp teeth, devilish grin, red sky, halo reissuance sil
9 anger,dangerous,['beautiful dangerous female japanese cyberpunk cyborg hacker, emo, gothic, tattoos, grunge, noir, moody cinematic, futu
10 anger,outrageous,['a silver gelatin photo portrait of a royal cat, outrageously fluffy, on an embroidered velvet cushion on a neo rococo
11 anger,despotic,['prays to some despotic hooded God art by Stefan Koidl, Trending on artstation, artstationHD, artstationHQ, 4k, 8k', 'c
12 anger,ruthless,['hyper detailed ultra realistic portrait of clancy brown as cruel, ruthless, megalomaniacal, brutal, remorseless savage
13 anger,moody,['a massive gaping canyon on the moon, digital art, illustration, highly detailed, cinematic lighting, moody, 4 k, 8 k ', 'f
14 anger,barbaric,['detailed image of barbaric by Ayami Kojima, Amano, Karol Bak, Greg Hildebrandt, and Mark Brooks, rich deep colors Beksj
15 anger,fierce,['doom eternal, game concept art, muscular, exoskeleton, chiroptera head, chiroptera ears, shark teeth, ferocious, fierce,
16 anger,violent,['star wars clone troopers offer combat to soldiers in vietnam, photo, old picture, lush landscape, field, firearms, explo
17 anger,angry,['oil painting portrait of anthropomorphic female green fox animal, dressed in black hoodie fox animal, glowing eyes, oil pa
18 anger,aggressive,['photoreal portrait intricate darkwar robomosquito wearing a dark hoody, apocalyptic scenario, intricate malfunctionin
19 anger,hostile,['experts declare that there is no effective defence against air raids, but if the approach of a hostile air force could b
20 anger,furious,['movie poster paul walker, fast and furious, sexy, the crew drives great white sharks in the desert, climate change ', 'p
21 anger,ferocious,['doom eternal, game concept art, muscular, exoskeleton, chiroptera head, chiroptera ears, shark teeth, ferocious, hyper
22 disgust,grotesque,['a drawing of grand teton national park, a woodcut by martin schongauer, tumblr, heidelberg school, woodcut, grotesqu
23 disgust,invidious,['an invidious shadow crawling out of a black hole in the decadent court of the artdeco inner city of yaSattrra as the
24 disgust,loathsome,['the most loathsome purification lubricated by dripping viscous disgust', 'The 1080p photo of the despicably beautifu
25 disgust,horrid,['horrid, abominable, disgusting, vile, revolting, fanged creature, rembrandt ', 'horrid, abominable, disgusting, vile, r
26 disgust,eerie,['a painting by edward hopper of an eerie scene on a rainy day, illuminated by the cityscape ', 'portrait painting of a we
```

(情感引导词典)

具体视觉细节扩充阶段



自建数据集

- 498条原始场景文本
- 扩充至2988条情感样本
- 6种基本情绪类别覆盖

表 4-1 数据集示例

Scene	Emotion	Adjective-Noun Pairs
Beach	Anger	ferocious waves, brutal winds, angry skies, burning sand
	Disgust	ghastly seaweed, weird debris, eerie fog, rotting shells
	Fear	horrible storms, fearful darkness, desperate tides, crashing waves
	Joy	refreshing breeze, splendid sunshine, energetic waves, vibrant sand
	Sadness	desolate shore, stale air, alone footprints, fading light
	Surprise	crisp waves, stunning sunrise, playful dolphins, sparkling water

模型选择与优化

- 基础模型：Llama3.1-8B-Instruct
- 优化方法：LoRA监督微调技术
- 目标：提升视觉细节表现力

LoRA技术应用

- 通过低秩分解减少参数量
- 只更新部分参数，保持原始预训练权重
- 提高计算效率和资源利用率

细节扩充框架

- 输入：用户文本
- 输出：类人化的形容词-名词对 (ANPs)结构化指令
- 特点：端到端处理，无需额外检索步骤

前端展示



情感文本生成

beach

生成情感文本

生成视觉描述

生成图像

情感文本结果：

愤怒: beach, ferocious, brutal, angry

厌恶: beach, ghastly, weird, eerie

恐惧: beach, horrible, fearful, desperate

快乐: beach, refreshing, splendid, energetic

悲伤: beach, desolate, stale, alone

惊喜: beach, crisp, stunning, playful

视觉描述结果：

愤怒
ferocious wavesbrutal windsangry skiesraging fires

厌恶
ghastly trashweird odorseerie shadowsrotting food

恐惧
horrible darknessfearful silencedesperate escapesominous clouds

快乐
refreshing breeze splendid sunsetsenergetic wavescheerful seagulls

悲伤
desolate shores stale airalone figuresweeping willows

惊喜
crisp morningsstunning vistasplayful dolphinsunexpected rainbows

生成的图像：



愤怒: beach, ferocious, brutal, angry
视觉元素: ferocious waves, brutal winds, angry skies, raging fires



厌恶: beach, ghastly, weird, eerie
视觉元素: ghastly trash, weird odors, eerie shadows, rotting food



恐惧: beach, horrible, fearful, desperate
视觉元素: horrible darkness, fearful silence, desperate escapes, ominous clouds



快乐: beach, refreshing, splendid, energetic
视觉元素: refreshing breeze, splendid sunsets, energetic waves, cheerful seagulls

生成情感文本(氛围) ➡ 生成视觉描述(细节) ➡ 生成图像

实验结果与分析

表 4-4 定性消融实验结果

	愤怒 anger	厌恶 disgust	恐惧 fear	悲伤 sadness	喜悦 joy	惊喜 surprise
仅生成模型						
+氛围控制						
+细节控制						

表 4-4 定量实验结果

方法	ACC-2	ACC-6	FID	LPLPS
仅生成模型	66.67%	33.33%	16845.90%	70.43%
+氛围控制	83.33%	36.67%	19850.73%	73.74%
+细节控制 (FV-ECVG)	100.00%	46.67%	18846.00%	76.30%
+细节控制 (LVG-DFLA)	100.00%	83.33%	18112.32%	77.48%

定性实验

- 在引入情感引导词典后，模型的表现得到显著改善，明显提升了积极情感类别与消极情感类别之间的视觉差异性。
- 通过应用扩充框架，模型在细节层面的控制能力得到进一步提升。生成的视觉内容不仅能为观者带来细粒度的情感体验，同时在视觉丰富性和真实感方面也取得了显著进展。

定量实验

- 情感2分类和6分类任务准确率显著提升
- 生成内容具有更强的情感表现力
- 视觉细节描述更加丰富准确

T HANKS!

感谢您的观看

汇报人：
沈雅卉

日期：
2025.4.19

