

# 国家级大学生创新创业训练计划 项目申报表 (创新训练项目)

推 荐 学 校 :	杭州电子科技大学
项 目 编 号 :	
项 目 名 称 :	国画影像智能分析与 可交互式生成算法
所属一级学科名称:	计算机科学与技术
项 目 负 责 人 :	叶涵
联 系 电 话 :	19818505821
指 导 教 师 :	高飞
联 系 电 话 :	13588033141
申 报 日 期 :	2022 年 3 月 25 日

浙江省教育厅 制  
二〇一三年五月

项目名称		国画影像智能分析与可交互式生成算法					
项目所属一级学科		计算机科学与技术					
项目类型		(√) 创新训练项目    ( ) 创业训练项目    ( ) 创业实践项目					
项目实施时间		起始时间： 2022 年 9 月    完成时间： 2023 年 12 月					
项目简介 (100字以内)	本项目拟基于深度学习技术，研究国画的智能分析算法，以及图像、文本等多模态交互驱动的国画风格迁移与动画合成方法。该研究在文博、教育、娱乐等领域具有广泛的应用前景，并将积极推进国画等传统文化的传承和发展。						
申请人或申请团队		姓名	年级	学号	所在院系/专业	联系电话	E-mail
	主持人	叶涵	2020级	20184501	卓越学院/智能计算与数据科学	19818505821	1315926935@qq.com
		钟鑫	2020级	20184106	卓越学院/智能计算与数据科学	15706740996	1140091006@qq.com
	成员	张孜远	2020级	20151521	卓越学院/智能计算与数据科学	19857150410	1872401992@qq.com
		陈俊鑫	2020级	20184102	卓越学院/计算机英才班	13575327781	841887695@qq.com
		郑宣博	2020级	20063114	卓越学院/智能计算与数据科学	18858662319	1015076779@qq.com

指导教师	第一指导教师	姓名	高飞	单位	杭州电子科技大学
		年龄	35	专业技术职务	副教授
	主要成果		<p>高飞，主要研究兴趣为计算机视觉与机器学习，涉及智能视觉艺术、医学影像分析、智能机器人等课题。现已在主流国际期刊、会议上发表论文 20 余篇，达到了国际领先水平。主持和参与多项国家自然科学基金、浙江省自然科学基金项目。以主要参与人员身份获得 2016 年陕西省科学技术奖一等奖及 2011 年陕西高等学校科学技术一等奖。同时担任多个国际顶级期刊和会议的审稿人。在学科竞赛方面，曾指导学生获得全国大学生创新创业项目国家级 1 项，中国大学生服务外包创新创业大赛国家级二等奖 1 项，“互联网+”创新创业大赛浙江省铜奖 1 项，浙江省大学生新苗人才计划 3 项等。</p>		
	第二指导教师	姓名	许金兰	单位	杭州电子科技大学
		年龄	35	专业技术职务	副教授
	主要成果		<p>许金兰，研究兴趣为等几何分析、等几何造型、图像处理等。已在国际国内重要学术期刊上共发表论文 10 余篇，主持在研国家自然科学基金面上项目一项，主持完成浙江大学 CAD&amp;CG 国家重点实验室开放课题、浙江省青年基金项目、国家自然科学基金青年项目各一项，参与多项国家级省部级科研项目。担任中国工业与应用数学学会几何设计与计算专委会委员、CCF 会员。</p>		
<p>一、申请理由（包括自身具备的知识条件、自己的特长、兴趣、已有的实践创新成果等）</p> <p><b>1. 团队背景</b></p> <p>该项目团队是由一群有理想，有规划，敢于创新，热爱科研的同学组成。团队拥有计算机视觉和深度学习方面学习的相对超前的深度和广度，参与项目所积累的丰富的代码量和随之形成独特而成熟的编程风格和水准，充沛的竞赛经验，项目经验，除技术实现外的经管方面专业人才。团队每位成员对于该项目都有着独特而统一的见解，每位成员有自己最擅长的部分，对技术方面熟练掌握，有着灵活大胆的思维能力，研发氛围良</p>					

好，共同为该项目的建成添砖加瓦。以此有信心达成项目的顺利展开。

该团队注重学科交叉融合。在沉浸本专业学科知识，紧随信息前沿，打造核心竞争力，提升专业硬实力的同时，有意识地积极探索多学科交叉融合的有效途径，以激发创新活力，力图实现更好的合作与分工，发挥跨学科优势，以此实现项目从 0 到 1 的蜕变。

## **2. 项目已有基础**

### **2.1 设施基础**

在实验设施基础方面，团队依托杭州电子科技大学计算机学院智能可视建模与仿真实验室。实验室负责人为徐岗教授，现拥有教师 6 名。团队主要研究方向包括计算机辅助设计与仿真、等几何分析、计算机视觉、机器学习等。已在相关领域的国内外主流期刊/会议上发表学术论文 80 余篇，其中在国际 TOP 期刊发表论文 30 余篇，多篇论文入选 ESI 热点论文和高被引论文。团队成员以负责人身份主持承担或完成多项国家自然科学基金项目（包括 1 项中德合作重点项目），并与多家企业开展联合研发和提供技术服务。拥有近 40 余块高性能计算显卡 GPU，具备实施本课题所有研发内容的硬件基础。

### **2.2 研究基础**

指导教师高飞和许金兰的主要研究兴趣为计算机视觉与机器学习，在智能视觉艺术、艺术机器人等领域达到了国际领先水平。现已在主流国际期刊、会议上发表论文 20 余篇，其中 ESI 高被引论文 1 篇。主持国家自然科学基金及浙江省自然科学基金共 3 项，并参与多项研究课题。研究团队已具备该项目所需的计算资源和相关研究基础。

## **3. 研究积累和已取得成绩**

现在，团队主持人叶涵和钟鑫已经参与艺术影像生成及风格化算法的研究超过 1 年时间，阅读了大量的相关论文。两位同学对于深度学习以及艺术影像生成领域的进展具有较为全面的了解，且具备丰富的算法编程实践经验。此外，叶涵和钟鑫还对国画影像的生成算法，进行了初步尝试和实现，取得了较好的效果。后续将做进一步研究和改进。团队成员张孜远、陈俊鑫和郑宣博，也已经积累丰富的编程实践经验，对于前后端软件系统的开发具有丰富的经验。现在，团队对于本项目的前后端系统研发已经具备了初步设计方案。两位同学后续将开展进一步细化和实践。

整体而言，我们有精良的研发团队、成熟的计算机技术和热忱的研究兴趣。本课题拟研发的技术及产品，旨在将科技与艺术进行完美融合，实现国画风格影像的自动生成，且具备便捷的可交互性。该课题在文博、教育、动画、游戏等领域具有广阔的前景，必将对国画等传统文化的传承起到积极的推进作用。

## 二、项目方案

### 1. 项目研究背景

#### 1.1 项目背景

##### 1.1.1 国画发展及应用场景

中华文化源远流长、灿烂辉煌。在 5000 多年文明发展中孕育的中华优秀传统文化，积淀着中华民族最深沉的精神追求，代表着中华民族独特的精神标识，是中华民族生生不息、发展壮大的丰厚滋养，是中国特色社会主义植根的文化沃土，是当代中国发展的突出优势，对延续和发展中华文明、促进人类文明进步，发挥着重要作用。而国画则是中国传统文化中最具特色和代表性的艺术媒介，彰显中华文化的精神内涵和审美风范。

国画一词起源于汉代，汉朝人认为中国是居天地之中者，所以称为中国，将中国的绘画称为“中国画”，简称“国画”。中国画在内容和艺术创作上，体现了古人对自然、社会及与之相关联的政治、哲学、宗教、道德、文艺等方面的认知；同时，作为一种“精神文化”，其魅力和意义就在于传承画中内容与其所独有的美学价值。例如，2022 年北京冬奥会的开闭幕式，大量采用了国画元素，向世界成果展示了中国传统文化的丰富内涵和独特审美。因此，对于国画的保护及研究，不仅是对中华优秀传统文化的保护与传承，也是对现代审美及艺术的发展壮大。

我们的项目主要拟研究基于照片、草图、文本的国画影像生成算法，以及视频、文本、语音等多模态交互驱动的国画视频合成技术。该研究在文博、教育、娱乐等领域具有广泛的应用前景，并将积极推进国画等传统文化的传承和发展。图 1 举例说明了几种可能的应用场景。下面，我们将对各个应用领域的背景信息进行介绍。

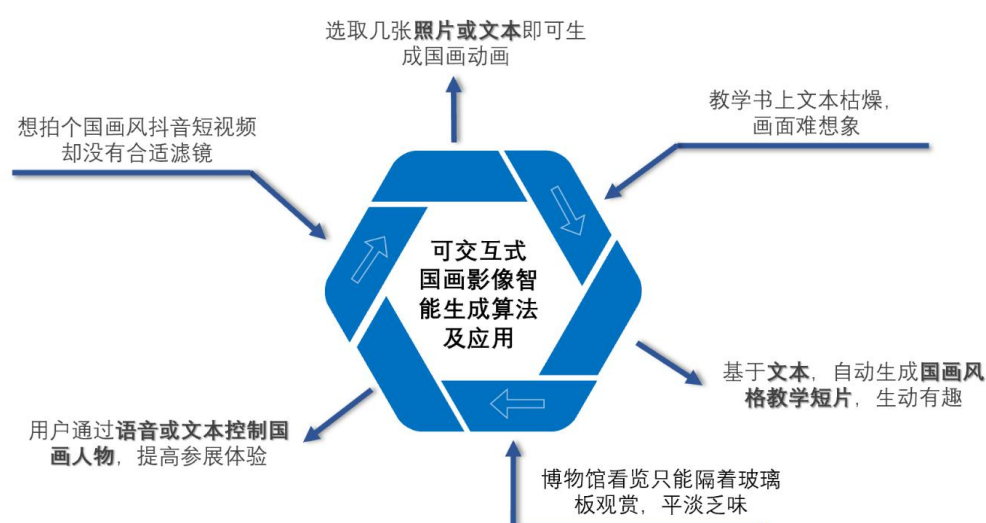


图 1 应用场景

### 1.1.2 教育行业亟需活力

我国近年来在传统文化的教育及传承方面越来越重视。例如，中共中央办公厅国务院办公厅印发的《关于实施中华优秀传统文化传承发展工程的意见》就指出，要把中国传统文化教育贯穿国民教育始终。国家教育机构也积极推进以幼儿、小学、中学教材为重点，构建中华文化课程和教材体系。编写中华文化幼儿读物，开展“少年传承中华传统美德”系列教育活动，创作系列绘本、童谣、儿歌、动画等。例如，在古诗文教学中，要让学生根据诗句去想象画面，充分发挥学生的想象力，最为直观的方法便是运用多媒体精心创设与教学内容相吻合的情境。形象、直观的效果可以增强对学生的感官刺激，调动学生的情感体验，激发学生的求知欲，让他们如临其境、感同身受，从而达到以情激情的目的。

然而，利用现有的多媒体及动画生成技术，其人力、时间、财力等成本非常高昂，而且需要大量专业的技术人员。这大幅度降低了传统文化教育走入校园的进度。我们的项目可以选取视频、图片甚至文本，更为便捷地生成国风动画，为教学提供便利，为现在的教育行业涌入一股“艺术的暖流”，去体验、去拥抱中华民族文化的温暖，陶冶情操。



图2 古诗国风动画

此外，2014 年左右，STEAM 教育理念开始引入我国，并在随后的几年内快速推广开来。家长的教育理念也不断升级，80 后、90 后家长开始注重兴趣培养，逐渐成为教育消费的主力军。国家也在积极推进素质教育，并由中共中央办公厅国务院办公厅于 2020 年印发了《关于全面加强和改进新时代学校美育工作的意见》，要求丰富艺术实践活动，加快艺术学科创新发展，营造社会氛围。艾瑞咨询 2021 年的调查报告（图 3）也显示，在不同阶段，家长对子女的艺术类课程教育都相当重视。国画又是中国艺术的瑰宝，必定有着更广阔的发展前景。

不同阶段家长的子女课程参与情况

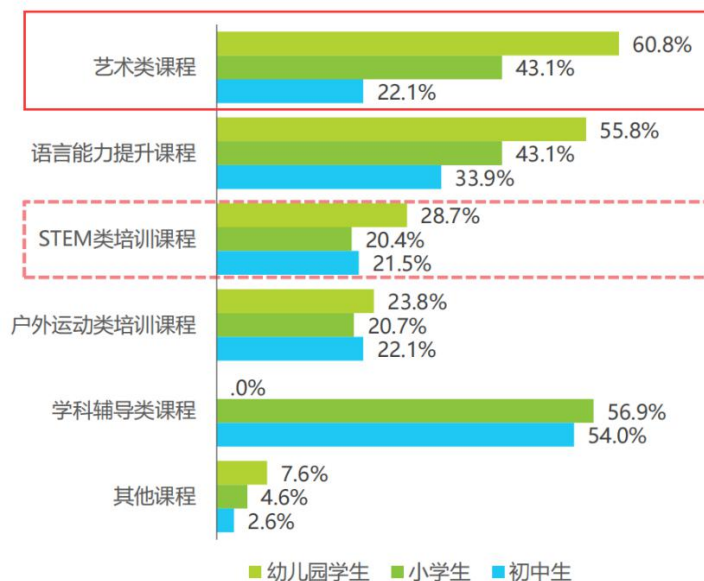


图 3 家长对子女课程参与情况

然而，现在国画教师、互动形式、公共资源、社会意识等都还存在不足。特别是受疫情影响，相关实践活动更为受限。这些困难也阻碍了国画等美学教育的推进。本项目的研究内容，可以基于多模态输入（如照片、草图等）生成国画图像，这可以应用于国画宣传及教育的多个阶段。此外，本项目还提供基于视频、语音、文本等方式的国画视频生成，可以与学生进行便捷、丰富的交互。基于本项目的美学教育方式，将注重寓教于乐，与游戏高度结合、应用导向、兴趣驱动。此外，我们的项目还可以将国风动画进行交互性升级，开发相应游戏，与教学相结合，做到寓教于乐，助推教育行业发展。

### 1.1.3 国风游戏成为趋势

中国移动游戏市场行业买入平稳增长阶段，业态创新给市场带来新的想象空间。2021 年中国移动游戏市场规模约 3078 亿元，同比增长 9.8%。在游戏行业政策监管趋严、行业规范化程度持续加深的背景下，预计在 2022 年移动游戏行业、仍将保持 10%以上的增速。行业未来的增长机会，将会来源于 5G、云游戏、元宇宙等新技术及新业态和移动游戏结合带来的想象空间，这也是移动游戏行业持续创新和变革的重要驱动因素。

在游戏内容创新上，单一玩法已经无法满足用户的需求，游戏研发将不同的内容题材及多种玩法机制进行有机融合，推动移动游戏产品的创新发展。其中，国风题材备受关注，中国传统文化的广度和深度成为游戏开发的无尽的灵感源泉。而便捷的国风动画生成为该类游戏的开发提供了便利。本项目的技术，可以用于基于用户照片自动生成个



性化的游戏角色形象，并且可以由用户通过视频、文本、语音等驱动游戏形象进行动作。这将大大推进游戏的个性化发展，并且大幅提升对游戏用户的吸引力。



图 4 游戏发展趋势

1.1.4 文博产业兴起

近几年，无论民办还是公立，博物馆像雨后春笋拔地而起，不仅为城市带来了新地标，更为普通人带来了更多与珍贵文化遗产亲密接触的机会。大型的综合博物馆在资金、藏品方面已经有一个较为充足的资源，能够采取多种形式如：讲座、流动展览、沙龙、体验展、音乐会等等。随之而来的是国潮风格的兴起，传统工艺、经典人物形象及博物馆 IP 受到消费者的喜爱与追捧。各大品牌纷纷与博物馆及国风综艺合作推出国潮联名。

“中国风”设计元素：综艺/博物馆IP

以建筑形象、馆藏文物及关联元素为设计理念融入国潮

综艺及博物馆IP已在多品类进行国潮跨界联名。代表性的配色、图案与纹理是主要设计方向。

iResearch 艾瑞咨询

综艺IP	博物馆IP
<p>——★ 国家宝藏 ★——</p> <p>➢ 服装联名：锦护膊（出土于尼雅遗址的汉代“五星出东方利中国”锦护膊，主要材质为织锦）</p> <p>➢ 白酒联名：三星堆纪念版</p> <p>➢ 外设：杜虎符鼠标、千里江山键盘等</p>	<p>——★ 敦煌博物馆 ★——</p> <p>➢ 服装联名：以敦煌地貌及色彩为设计理念</p> <p>➢ 彩妆联名：以瑞兽图案及矿石颜色为设计理念</p> <p>➢ 数码产品联名：孔雀、撒花飞天等限量版外观</p>
<p>——★ 上新了故宫 ★——</p> <p>➢ 帆布鞋联名：鞋头部分借鉴了太和殿的金色琉璃瓦当，鞋身装饰线条提取自太和殿和奉先殿的金钱纹带图样。</p> <p>➢ 坚果礼盒联名：设计灵感来源于皇极殿、养心殿宝匣以及剔红海水龙纹天球瓶内含的紫禁城色彩美学和图案样式。</p> <p>➢ 艾灸盒联名：包装设计灵感来自临溪亭；图案选择立体视角中的龙亭斗拱与临溪亭的彩绘天花藻。</p>	<p>——★ 河南博物院 ★——</p> <p>➢ 轮胎厂商推出联名款考古盲盒</p> <p>➢ 服装联名：以杜岭方鼎、武则天金简和汝窑天蓝釉鹅颈瓶图案为灵感</p>
	<p>——★ 陕西历史博物馆 ★——</p> <p>➢ 鞋服联名：以《阙楼仪仗图》图案和唐代马球的纹理为设计理念</p>

图 5 国风设计元素受到热捧



仅 2020 年，全国博物馆系统推出了 2000 多个线上展览，总浏览量超过 50 亿人次。另在国家相关政策的激励下，博物馆希望借助互联网和新技术的平台，将其所拥有的优秀文化资源与公众分享，把传统文化的内涵用富有创意的方式普及并传递下去，从而培养更多爱传统文化的年轻人。博物馆在相关国家政策的刺激之下，借助互联网和新技术的平台，将其所拥有的优秀文化资源与公众分享，把传统文化的内涵用富有创意的方式普及并传递下去，从而培养更多爱传统文化的年轻人。2016 年以来，故宫先后与阿里、腾讯、凤凰网等建立合作伙伴关系，一方面在天猫开设了官方旗舰店，主要有门票、文创、出版三个板块；另一方面和腾讯游戏推出基于故宫文化的大型游戏，从 2017 年下半年开始，双方合作更是介入到动漫和文学的创作，包括微电影的创作中；另外，还和凤凰网合作，利用 AR、MR、3D 等手段传播故宫文化。通过一系列的战略合作，故宫有望成为真正的“世界级的超级 IP”。



图 6 博物馆参观人次变化

此外，在多媒体传播的今天，人们对视听有了越来越高的要求。因此，很多博物馆推出了 AR、VR 等多种形式的展览。其中，最为普遍的就是让国画动起来。然而，现有技术主要依赖于大量人工手段，其成本非常高昂，且与用户的交互方式较为有限。本项目的研究将自动实现国画动画化，且提供丰富的用户交互方式。这可以大幅提升文博展览的形式多样性和用户体验，从而吸引更多用户，进一步促进传统文化的传承和保护。

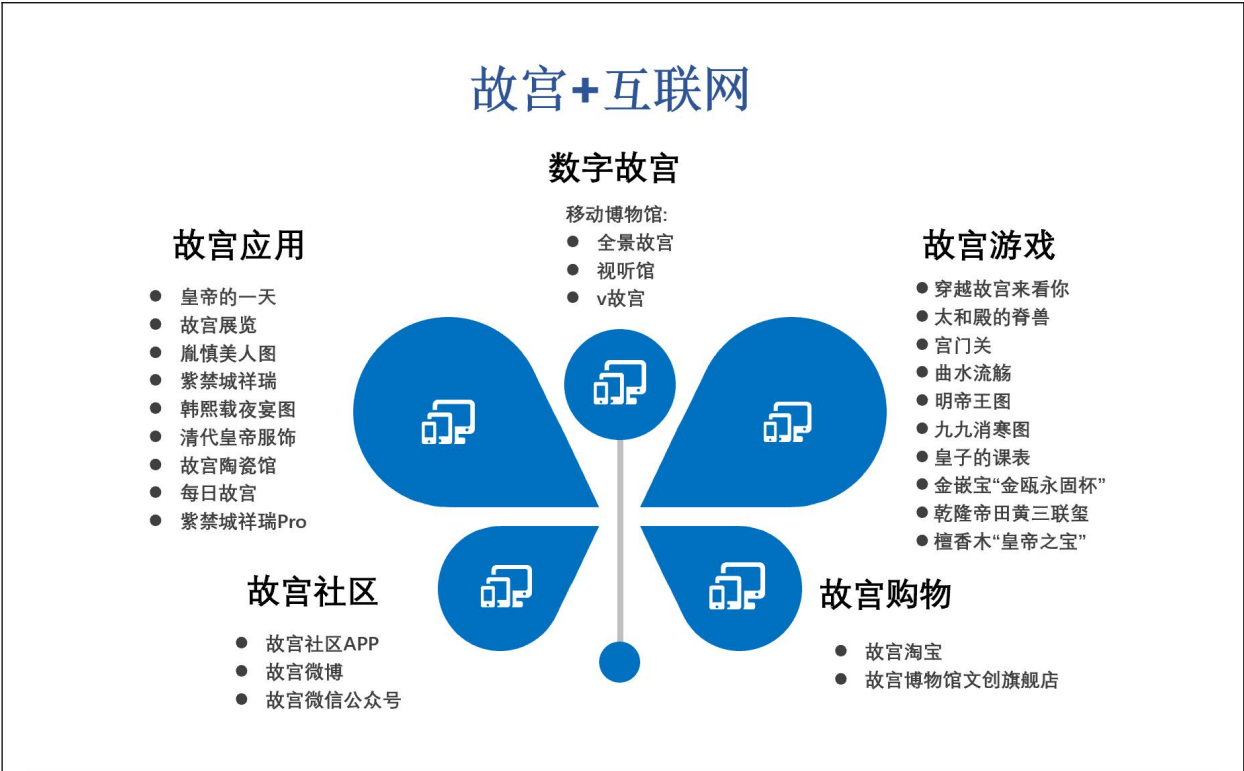


图 7 故宫与互联网应用

1.2 研究现状

本项目拟基于深度学习技术，研究国画的智能分析算法，以及图像、文本等多模态交互驱动的国画风格迁移与动画合成方法。该研究在文博、教育、娱乐等领域具有广泛的应用前景，并将积极推进国画等传统文化的传承和发展。相关领域尽管已经取得了一些进展，但仍然存在一些不足。具体分析如下。

如何建立和管理中国画数字图书馆和数字博物馆成为当前的一个热点研究问题，对中国画图像的处理技术是该研究问题的关键。如何发现这些画作背后的数字规律，如何有效地分析、识别和分类中国画的作者和艺术风格日趋成为研究热点。目前大量文献利用计算机技术对艺术画进行分析和研究，但针对中国画的研究仍存在如下问题：现有艺术画方面的成果主要是将自然图像进行油画、粉笔画等艺术风格模拟，或者是针对油画等西方画进行艺术风格分析。国画作为一种独特的艺术形式，与其他艺术风格（如油画、卡通、粉笔画等）相比有着截然不同的意境和韵味，因此一些已有的针对西方画等的研究不能直接用于中国画的分析中。

目前针对中国画的科研工作较少，主要研究工作如下：盛家川和江健民提取了整体和局部特征，并提出熵平衡(Fusion)算法对国画作者进行分类研究。盛家川研究了在小波域内不同尺度和频带上画作所体现出的不同深度信息特征，以对画作进行分类。Li J，

Wang J Z 等人设计了相关算法分类沈周、唐寅、张大千等中国画家的画作，该算法首先对画作提取小波特征，并设计了 MHMM 隐马尔可夫模型对画家进行分类研究。Jiang 等人提出了一种把中国画分为写意和工笔两类的算法，首先提取颜色、纹理等底层特征，进而采用支持向量机进行分类。王征和孙美君提出了复合特征的自适应遴选及对水墨画风格描述的优化，其还通过提取多种底层异构视觉特征对国画风格进行描述，对国画家进行分类预测。Fan 等人提出了一种色彩转换算法将照片转换成中国水墨画。

目前，针对中国画的工作大多集中在基于内容的分析和检索(content-based image analysis)，但存在以下问题：任何图像内容特征的可适应性都存在一定的局限性。例如，画马时的笔锋与画树叶时的笔锋肯定是不同的，如果不考虑每个笔锋所产生的条件而一味地分析笔锋的走向和力度分布等将会是很盲目的。此外，现有基于内容的研究会考虑画作中所有的信息，这样会使得研究重点较为分散，容易受到噪声信息的干扰。

## 1.4 可行性分析

### 1.4.1 国画现状、社会现状与软件现状概述

现今的国画传承主要面临以下问题：

（1）传统文化受众小，阳春白雪如何让大众都喜爱，是值得探讨和研究的课题，制作爆款需要有合适的机遇和成熟的营销方案。

（2）短视频文化崛起一方面体现多元文化的发展，另一方面也体现人们快节奏的生活，快餐文化势必不利于传统文化的发扬和沉淀，保持传统文化新发展的活力，延长产品在人们认知中存在的生命周期。

（3）软件技术实现难度较大，需要有完整的理论体系。另外，研究项目牵扯到其他学科领域，成果的展现依托于对本专业和其他学科专业的认知程度。

基于以上三点考虑，我们将国风、动画（短视频）与相关风格迁移技术结合起来，使得国画不但拥有其民族自身的文化内涵，也可以通过动画的方式，将其传播到大家的视野中。这种新的呈现方式一定会使大家对于国画本身以及其背后的文化底蕴产生浓厚的兴趣；同时，现存的成熟技术也给我们“自动生成国风动画”项目提供了保障。

### 1.4.2 政策支持

在物质文明飞速发展的同时，国家不忘精神文明建设，陆续出台一些政策扶持文化产业建设发展。

- 政策一：为贯彻落实《国务院关于印发“十三五”国家战略性新兴产业发展规划的通知》，进一步加快培育发展我省战略性新兴产业，浙江省政府制定《浙江省培育发

展战略性新兴产业行动计划》。计划中指出，要加快培育壮大数字创意产业，形成文化引领、技术先进、链条完整的数字创意产业发展格局。

- 政策二：深化文化领域的改革开放。十七届六中全会通过的《中共中央关于深化文化体制改革推动社会主义文化大发展大繁荣若干重大问题的决定》从法律、政策以及财政等三方面提出了具体的要求：一是强调法律保障；二是强调政策保障；三是强调财政保障。
- 政策三：2020 年 11 月文化和旅游部发布《关于推动数字文化产业高质量发展的意见》。在夯实数字文化产业发展基础、培育数字文化产业新型业态、构建数字文化产业生态等方面提出多项意见：
  - （1）夯实数字文化产业发展基础。
  - （2）培育数字文化产业新型业态。
  - （3）构建数字文化产业生态。
- 政策四：2020-2021 年，在新冠疫情的冲击下，以国务院、工信部等国家部门相继出台政策措施，引导数字文化产业发展风向，加强数字文化产品开发，建设数字化平台新经济，并计划在未来五年内引领出 20 家领军企业，打造出 5 个引领数字文化产业的产业集群，开发出一批具有示范性的数字文化产业项目约 200 个。

表 2 2025 年国家关于数字文化产业发展的战略目标

名称	2025 年目标
领军企业数	20 家
重要产业集群	5 个
示范性数字文化产业	200 个
支持技术创新方向	5G，大数据，云计算，人工智能，物联网，区块链

1.4.3 社会文化

近代以来，科技创新成为时代主旋律，传统文化在西方科技文明的围城下，可谓是举步维艰，一边是对科技的趋之若鹜，一边是对传统的弃若敝屣，一时间染上“文化乡愁”。然而，科技与传统文化并非二元对立，新与旧的博弈，从未非此即彼。

近年来，多个优秀传统文化 IP 活跃在人们视野：孙悟空、哪吒、姜子牙等一些列神话人物活跃在荧屏；《国家宝藏》一改往日枯燥严肃的风格，采用情景剧的方式，演绎文物背后的故事；《如果国宝会说话》通过 100 件国宝展现国人的创造智慧，吸引国人的同时也收获一大批外国的粉丝；《大学》有言：“苟日新，日日新，又日新”。己亥新春，上元之夜，紫禁城万盏灯火长明；敦煌经年，飞天再现，奇景不息。在 2021 年，

河南春晚节目《唐宫夜宴》从一众歌舞节目的混战中异军突起，成为今年爆款，为人们交口称赞，喜提热搜。



图 8 “国潮风”来袭

这些 IP 的爆火，无一不反映如今市场的走向：真正吸引观众，引发共鸣的关键不是大杂烩的歌舞，不是千篇一律、不走心的舞台套路，而是浓郁的国风国潮，是传统文化，华夏文明对骨血的召唤。泱泱文化似千里沃土，滋养科技之树，而科技之亭亭华盖盎然绿意，勃勃生机不绝。

## 2.项目研究目标及主要内容

### 2.1 研究目标

出于响应国家对传统文化艺术复兴的号召，本项目拟基于原有的国风动画自动生成等技术，进一步实现国画动画化，让画中的人物“活起来”、景物动起来。借乘当下短视频盛行之风，让我国优秀的国画艺术从纸上走出来，以动图或视频的形式展现在大众面前，进一步提高大众对国画艺术的关注度和喜爱度，增强文化自信。同时借由本次项目研究，提升我们的算法性能和算法适用度，在技术能力进一步提升后，本项目拟进行 APP、小程序及 API 开发，让中国传统文化的传播在互联网上更加便利有效。

### 2.2 主要内容

该项目涉及到的有如下几项关键技术：

- ✓ **属性分析：**通过对国画进行目标检测获取物体的信息，提取图像特征，建立图结构，构建图神经网络来进行国画属性分类。
- ✓ **风格迁移：**将输入图片转化为国画的风格，其中包含山水图水墨画、仕女图人像画、



水墨漫画、敦煌壁画以及其他种类的国画。

- ✓ **视频生成**：让静态的国画中的景物动起来；在获得对应的国画后，利用风格迁移类似技术将人脸转换为国画风格，而后利用 AI 技术生成相对应的视频。
- ✓ **APP 开发**：本项目拟开发基于 ios 的 ipad 应用程序，主要提供国画赏析、国画讲坛、国画体验、国画文创四个功能。

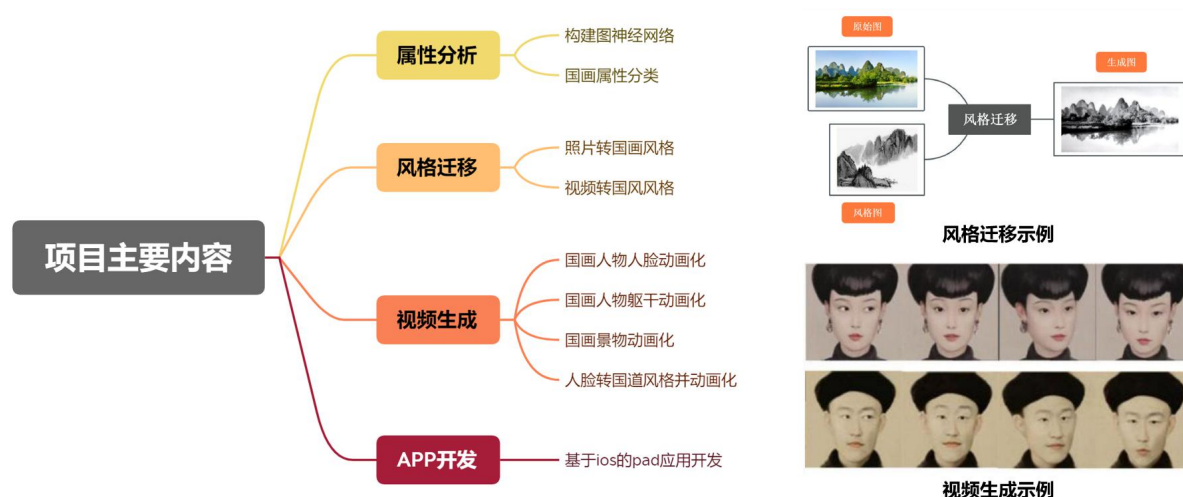


图 9 项目主要内容展示

### 3.项目创新特色概述

本项目的创新点，主要包含以下 4 方面：

(1) **基于专业构图理论的国画智能分析**：本项目拟基于国画专业领域中的构图理论，设计深度学习框架，自动实现国画的构图建模，并用于国画多种属性的自动预测。这为国画智能分析算法的设计提供一种新的路线。

(2) **基于语义指导的精细国画风格迁移**：本项目拟利用前景目标的分割掩膜、关键点位置等多种语义信息，进行国画风格过程中深度特征的精准自动调制，从而合成高质量逼真国画。这为国画风格迁移提供了一种新的算法框架。

(2) **基于多模态智能交互的国画动画化**：本项目拟利用视频、语音、文本等多种模态的信息，驱动国画从静态影像到动态影像的自动转换，从而实现丰富、便捷的交互方式。这为国画动画化任务提供了一种新的自动化解决方案。

(4) **“科技+艺术”助力传统文化的传承与发展**。本项目将科技与艺术进行完美结合，在教育、文博、游戏、娱乐等领域具有广泛的应用，让中华优秀传统文化以一个新的面貌展现在大众面前，必将对中国传统文化的传承与发展起到积极的推进作用。

## 4.项目研究技术路线

### 4.1 产品技术概况

该项目涉及到的有如下几项关键技术：（1）**国画属性分析**：将输入图片进行目标检测，提取特征并分析，对国画的属性进行分类。（2）**国画风格迁移**：将输入图片转化为国画的风格，其中包含山水图水墨画、仕女图人像画、水墨漫画、敦煌壁画以及其他种类的国画。（3）**国画动画化**：让静态的国画中的人物或其他目标物动起来；在用户开启摄像头后，利用风格迁移类似技术将人脸转换为国画风格，而后利用 AI 技术让转化后的人脸跟随用户的表情变化。（4）**APP 与 API 开发**：本项目拟开发 APP 和 API 便于用户调用，主要提供国画赏析、国画讲坛、国画体验、国画文创四个功能。

### 4.2 产品技术实现原理

该项目主要涉及绘画艺术专业知识和人工智能技术。在绘画艺术领域，主要需要利用其专业知识体系，为智能解析与交互提供专业指导。在人工智能方面，主要利用深度学习、计算机视觉、计算机图形学和自然语言处理，实现国画属性分析、国画风格迁移以及自动视频生成。

#### 4.2.1 属性分析

为实现国画的属性分析，本项目首先使用二维目标检测算法 YOLOv3 对国画进行目标检测。YOLO 算法在目前的目标检测中使用甚广，而且具有通用性强，运行速度快，识别精度高的一系列优势。此算法可以基于整张图片信息进行预测，其创新点在于借鉴了 ResNet 短接层的思想，使用了大量的残差跳层连接，残差跳层可以在加深网络的同时，极大地避免反向传播过程中梯度消失的问题。而且它提出了 one-stage 的检测思路，即目标定位和目标识别在一个步骤中完成，不需要网络划定边界框后再进行预测。

通过目标检测获取到国画中物体的信息后，我们利用特征编码器对所框选的物体提取图像特征，并将物体的图像特征作为节点、物体的位置作为边建立图结构。传统的神经网络比较适合用于欧式空间的数据，而图神经网络 GNN 可以把神经网络用在图结构 (Graph) 中，因此我们可以利用先前建立的国画信息图结构以构建图神经网络来进行国画属性分类。

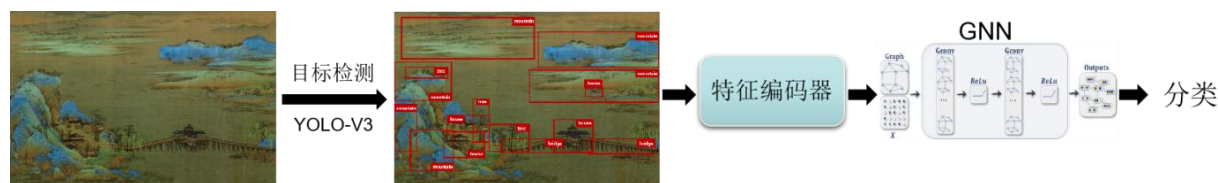


图 10 基于结构理论的国画属性分析算法流程



### 4.2.2 风格迁移

现有的风格迁移图像生成框架基于用户在单一模式下的输入生成图像，例如文本、分割图、草图或风格参考。在大多数情况下，往往无法利用用户的多模式输入，这降低了它们的实用性。为了解决这一局限性，我们拟使用 PoE-GAN 框架，基于多种输入模式或其任何子模式，甚至空集的图像去生成图像。其中，不同的输入模式适合传递不同类型的条件信息：

- (1) 分割图可以很容易地定义图像中语义类的粗略布局，即天空、云层、山脉和水域的相对位置和大小。在国画风格迁移中，我们可以在国画的基础上标出物体的关键点信息，基于物体各区域距离这些关键点的远近，把整个区域划分为不同的部位，最后将处理后的图作为分割编码器的输入。
- (2) 草图能够帮助我们在同一语义区域内指定结构和细节，例如单独的山脊。在我们等项目中，正值成长的儿童可在我们研发的 APP 上通过绘画出草图即刻生成出对应风格的国画，对于中国传统国画艺术的传播和教育起到极大积极作用。
- (3) 文本信息非常适合修改和描述图像中的对象或区域，这是使用分割或草图无法实现的。在我们研发的 APP 中，用户可以通过选择“写意”、“工笔”等选项来约束生成图像的技法，通过选择“山水画”、“花鸟画”等选项来约束图像的题材。
- (4) 对于风格参考图，用户可以选择我们以已经预先准备好的多种不同类型的风格参考图，也可由用户自由选择自己所想要的风格样式作为风格编码器的输入。

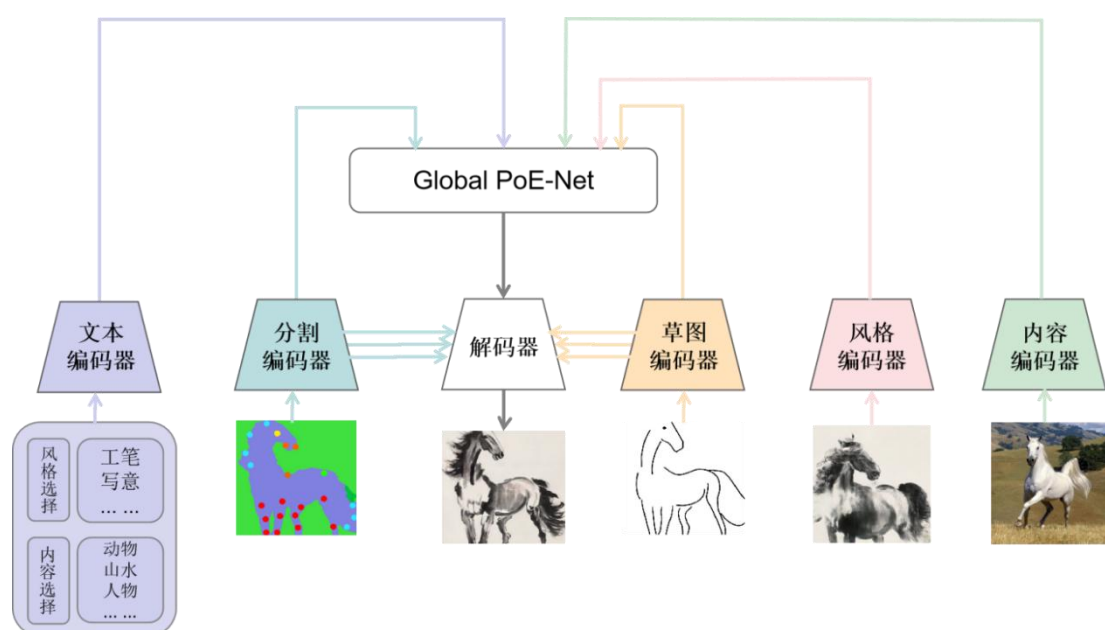


图 11 基于语义控制的国画生成器示意图

在此基础上，我们拟设计一个带有不同分辨率下的潜变量的潜在层次空间，这允许我们直接将信息从空间编码器的每个分辨率传递到潜在空间的相应分辨率，以便更好地保存高分辨率的控制信号。除了使用 **Product-of-experts** 提供多模态输入以外，我们拟采用多模态投影判别器，将投影判别器推广到处理多个条件输入，以获得最终损失。

图 11 为生成器体系结构的概述。我们将每个模态编码成一个特征向量，然后在 **Global PoE-Net** 中进行聚合。使用带有跳跃连接的卷积网络对分割图、图像内容和草图进行编码，使用残差网络对风格图进行编码以及使用 **CLIP** 去对文本进行编码。

### 4.2.3 视频自动生成

图像动画包括生成视频序列，根据驱动视频的运动为源图像中的对象设置动画。我们基于一阶运动模型（**First Order Motion Model**）框架来实现这一点，一阶运动模型通过自监督公式将外观和运动信息解耦。为了支持复杂的运动，我们使用由一组学习到的关键点及其局部仿射变换组成的表示。生成器网络对目标运动过程中产生的遮挡进行建模，并将从源图像中提取的外观与从驱动视频中提取的运动相结合以实现时间上的连续性以及空间上的一致性。一阶运动模型的算法框架如图所示，主要包括三个部分的网络：关键点检测器、密集移动网络和解码器。在框架的最后，我们使用编码后的源特征做变换，再进行解码生成出最终的结果。

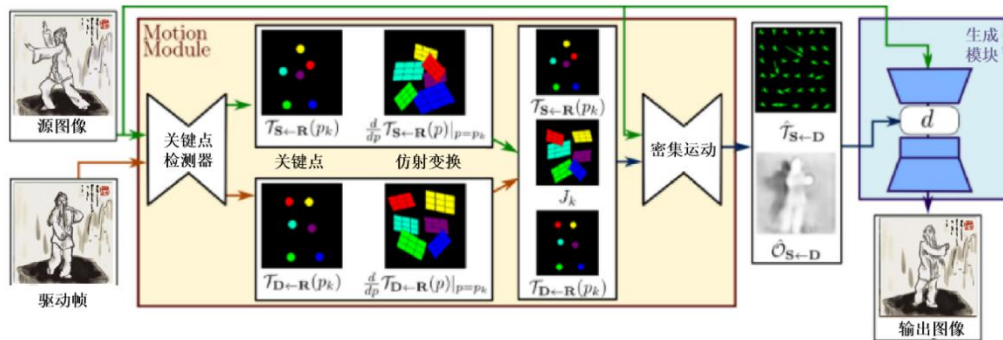


图 12 基于一阶运动模型的人物动画化方法流程

为了实现文本生成视频，我们基于文本驱动的视频生成（**Text2Video**）方法。根据训练数据（语音的音频和视频）构建声学-姿态字典（**phoneme poses dictionary**）。给定输入文本，可以包括英文，中文，数字和标点符号，使用 **TTS** 从文本生成语音。然后，应用强制对齐以获得音素时间戳，并在字典中查找对应元子。接下来，应用关键姿势插值和平滑模块来生成姿态序列。最后，使用 **vid2vid GAN** 生成视频。

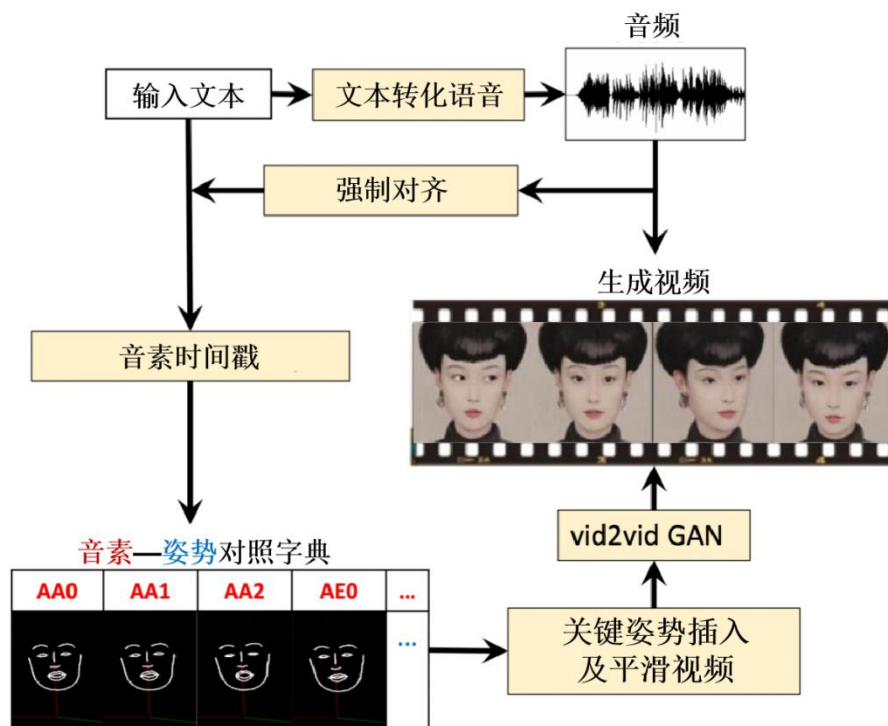


图 13 基于 Text2Video 的国画动画化框架图

## 4.3 APP 和 API 开发

### 4.3.1 APP 开发

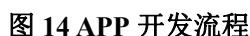
本项目拟针对平板用户，开发相关的应用。目前拟进行 ios 开发，可以应用在博物馆、教育应用中。开发时，首先需要确定产品方向。在生成、修订需求文档 后，根据需求文档画出 APP 线框图。设计师根据线框图设计视觉稿，开发人员根据视觉 稿搭建 UI 框架，主要是对 APP 图标、大概界面和功能进行设计。设计好 APP 的整体框架之后，开发人员根据需求文档开发功能，先进行大功能模块代码和大概的界面模块编写，将大概的界面和功能连接后，形成 APP 软件开发的大致 demo。测试人员再对 demo 进行测试，修复 Bug，完善 APP。APP 编程开发完毕之后，就可以生成安装包。在 IDE 中选择打包，在弹框中填写签名的 Key。技术手段我们目前打算将部分功能放到服务器上实现，即本地 APP 调用远程服务 器的 API。这样可以简化开发流程也可以减少终端性能开支，增强用户体验。

APP 提供如下几个功能可供用户使用：

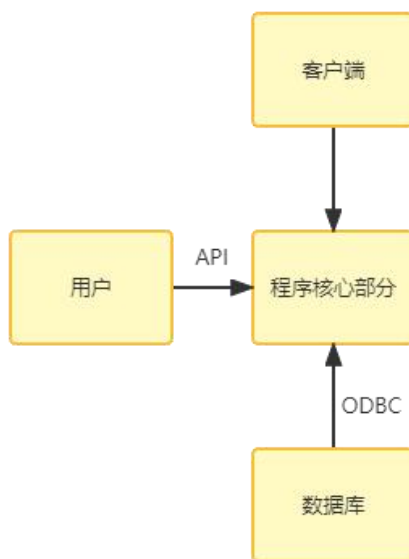
（1）国画赏析：提供各类国画的详细资料、尽量高清的图片以及名家评价赏析， 以供用户欣赏查阅。

(3) 国画体验：用户使用软件提供的画布绘制心仪的草图，或使用相册中的图片，甚至一段文本，便可以生成相对应的国画。软件能基于生成的国画，自动的为国画动画化，使国画有剧情化地动起来，让动画更加流畅，也给观者提供最佳的观看体验。此外，用户可以打开摄像头即可在手机中看见国画风格的自己，手机中的面部表情随着使用者的表情变化。拟提供道释画、仕女画、肖像画、历史故事画等多种风格。

API 开发部分拟使用 Python 实现，Python 相较于其他语言开发效率更高。将算法放到服务器上，用户可以通过网络接口调用该算法，之后服务器可以返回相关数据。API 所提供的功能与 APP 类似。



API 开发部分拟使用 Markdown 编写接口文档，开发语言拟使用 Python，框架拟使用 Flask。Python 语法有大量的相关库可以使用，因此使用 Python 是开发效率最高的。将相应算法放到服务器上，用户可以通过类似 `http://www.example.com/interface` 之类的网络接口调用该算法，通过 `get`、`post` 方法请求该接口，得到响应数据。当后端接口数据还未产生时，前端可以运用 `mock` 数据进行模拟开发。



**图 15 API 运行逻辑**

本项目拟结合所有研究内容，开发以下 5 项 API：

- 1) 国画详细资料及名家赏析推荐
- 2) 国画内人物动画化转换
- 3) 图像国画风格转换
- 4) 视频国画风格转换
- 5) 国画相关文创推荐

## 5. 研究进度安排

起止时间	开展内容
2022 年 4 月到 5 月	资料收集，继续完善和精化设计方案
2022 年 6 月到 9 月	收集所需样本数据集，对算法方案进行综合论证。
2022 年 9 月 到 2023 年 2 月	软件程序编写和调试。完善动画化和国画风格迁移算法设计，并实验验证其有效性。
2023 年 3 月到 5 月	进一步完善整个系统，改进算法性能。
2023 年 5 月到 9 月	系统整体模型改良，提升用户体验，做好最后产品提交的准备。
2023 年 9 月到 12 月	课题总结阶段：深入的总结和分析，完成有关报告和论文。

## 6. 项目组成员分工

成员姓名	成员分工
叶涵	资料收集，动画化效果改善，技术支持
钟鑫	动画化算法设计、风格迁移实现
张孜远	动画化效果测试，音频搭配，实验数据采集
郑宣博	风格迁移效果测试与完善，实验数据采集
陈俊鑫	资料收集，风格迁移效果改善，技术支持

## 三、学校提供条件（包括项目开展所需的实验实训情况、配套经费、相关扶持政策等）

### 1. 技术支持

本项目的开展受到杭州电子科技大学计算机学院智能可视建模与仿真实验室的支持。团队主要研究方向包括计算机辅助设计与仿真、等几何分析、计算机视觉、机器学习等。已在相关领域的国内外主流期刊/会议上发表学术论文 80 余篇，其中在 CMAME、CAD、CAGD、Computers & Structures 、IEEE TCYB、IEEE TIP 等国际权威 SCI 期刊发表论文 30 余篇，多篇论文入选 ESI 热点论文和高被引论文。团队成员以负责人身份主持承担或完成多项国家自然科学基金项目（包括 1 项中德合作重点项目）。并与多名国内外知名学者、出色研究小组保持着广泛和深入的学术交流与合作。实验室团队的研究方向为多媒体与机器学习、计算机视觉等人工智能相关前沿领域，与本项目所在领域相契合，能给予本项目组多方面的技术支持和指导。

### 2. 硬件支持

实验室的硬件资源（GPU 服务器等）充足，完全能够满足本项目在设计和训练深度学习模型时的计算需求。

### 3. 配套经费

正式立项后将有教育厅资金来支持完成本项目的设计实现。

#### 四、预期成果

本项目成果必须符合下列情况之一，且至少有一名项目组成员为主要参与人：

- 1) 项目结题报告 1 篇。
- 2) 计划开发 APP 软件 1 项。
- 3) 在国内核心期刊或以上投稿或发表 1-2 篇论文。
- 4) 申请或授权国家发明专利 1-2 项。
- 5) 申请或授权软件著作权 1-2 项。
- 6) 参加学科竞赛获得省级及以上荣誉。

#### 五、经费预算

总经费（元）	5000	财政拨款（元）	5000	学校拨款（元）	0
--------	------	---------	------	---------	---

注：总经费、财政拨款、学校拨款由学校按照有关规定核定数目进行填写

具体包括：

- 1) 打印，复印，印刷，PCB 制版等易耗材料费用 1000
- 2) 专业书籍购买费用 500
- 3) 调研等所需市内差旅费 500
- 4) 论文版面费，申请专利费 3000

#### 六、导师推荐意见

签名：

年 月 日



七、院系推荐意见

院系负责人签名：

学院盖章：

年 月 日

八、评审专家组意见：

负责人签名：

年 月 日

九、学校推荐意见：

学校负责人签名：

学校公章

年 月 日

注：表格栏高不够可增加。