

学校代码 ： 10286

分 类 号 ： TP393

密 级 ： 公开

UDC ： 004.7

学 号 ： 123456

标题要长

才会有人注意

研究生姓名**： Josan**

导师姓名： **教授**

|  |  |
| --- | --- |
| 申请学位类别 工 学 硕 士 | 学位授予单位 东 南 大 学 |
| 一级学科名称 网络空间安全 | 论文答辩日期 2018 年 6 月 1 日 |
| 二级学科名称 网络空间安全 | 学位授予日期 20 年 月 日 |
| 答辩委员会主席 | 评 阅 人 |
|  |  |
|  |  |

2018年 5 月 30日

.

**硕士学位论文**

**这是一个很长的标题**

**假装还有字**

**专 业 名 称** ：  **网络空间安全**

**研究生姓名** ：  **Josan**

**导 师 姓 名** ：

The long title and the long title this is a test this is a test a test test test test

A Thesis Submitted to

Southeast University

For the Academic Degree of Master of

Engineering

BY

ZHANG Sansi

Supervised by

Prof. LI Sisan

School of Computer Science and Engineering

Southeast University

May 30, 2018

**东南大学学位论文独创性声明**



本人声明所呈交的学位论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得东南大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

研究生签名： 日期：

**东南大学学位论文使用授权声明**

东南大学、中国科学技术信息研究所（含万方数据）、国家图书馆、《中国学术期刊（光盘版）》电子杂志社有限公司有权保留本人所送交学位论文的复印件和电子文档，可以采用影印、缩印或其他复制手段保存论文。本人电子文档的内容和纸质论文的内容相一致。除在保密期内的保密论文外，允许论文被查阅和借阅，可以公布（包括以电子信息形式刊登）论文的全部内容或中、英文摘要等部分内容。论文的公布（包括以电子信息形式刊登）授权东南大学研究生院办理。

研究生签名： 导师签名： 日期：

摘 要

网络教学作为一种新的教学形态正逐渐引导着教育领域的变革。发展至关重要。

本文面向下一代网络教学，描述。可行性和合理性。

**关键词：**模式知识，网络教学

Abstract

Therefore, how to depict the ability of individuals to acquire knowledge and effectively evaluate the ability level of individuals to acquire knowledge is very important to the development of new generation of network teaching.

Finally, from the perspective of both individuals and groups, two teaching case recommendation experiments are given. The final result of the experiment verifies the feasibility and rationality of the evaluation model based on individual cognitive structure.

**Key words**: Pattern Knowledge, Network Teaching

目 录

[摘 要 I](#_Toc513538620)

[Abstract II](#_Toc513538621)

[目 录 III](#_Toc513538622)

[插图目录 V](#_Toc513538623)

[表格目录 VI](#_Toc513538624)

[第一章 引言 1](#_Toc513538625)

[1.1 研究背景 1](#_Toc513538626)

[1.2 研究目标 2](#_Toc513538627)

[1.3 研究内容 3](#_Toc513538628)

[1.4 论文结构 3](#_Toc513538629)

[第二章 相关研究 5](#_Toc513538630)

[2.1 认知相关理论研究 5](#_Toc513538631)

[2.1.1 认知理论的发展 5](#_Toc513538632)

[2.1.2 认知风格研究现状 6](#_Toc513538633)

[2.1.3 认知能力研究现状 7](#_Toc513538634)

[2.2 网络教学评测方法研究 8](#_Toc513538635)

[2.2.1 数据挖掘技术 8](#_Toc513538636)

[2.2.2 网络教学评测研究现状 11](#_Toc513538637)

[2.3 本章小结 12](#_Toc513538638)

[第三章 个体认知结构的表达模型 13](#_Toc513538639)

[3.1 建构个体认知结构模型的重要性和必然性 13](#_Toc513538640)

[3.1.1 重要性 13](#_Toc513538641)

[3.1.2 必然性 13](#_Toc513538642)

[3.2 支持隐性知识的知识表达模型 13](#_Toc513538643)

[3.2.1 模式知识（](#_Toc513538644)*[Pattern Knowledge, PK](#_Toc513538644)*[）形式化描述 14](#_Toc513538644)

[3.2.2 模式知识实例 14](#_Toc513538645)

[3.2.3 支持个体认知结构分析的知识表达模型 15](#_Toc513538646)

[3.3 个体认知结构模型及其相关概念 16](#_Toc513538647)

[3.3.1 个体认知结构模型总述 16](#_Toc513538648)

[3.3.2 个体认知结构模型概念描述 17](#_Toc513538649)

[3.3.3 个体认知结构模型数学描述 19](#_Toc513538650)

[3.4 本章小结 20](#_Toc513538651)

[第四章 个体认知结构发展水平的评测模型及其实现 21](#_Toc513538652)

[4.1 对个体认知结构发展水平进行评测的必要性和重要性 21](#_Toc513538653)

[4.2 个体认知结构发展水平的评测模型及其相关概念 21](#_Toc513538654)

[4.2.1 个体认知结构发展水平的评测模型总述 21](#_Toc513538655)

[4.2.2 个体认知结构发展水平的评测模型概念描述 22](#_Toc513538656)

[4.2.3 个体认知结构发展水平的评测模型数学描述 25](#_Toc513538657)

[4.3 评测模型的应用——教学案例推荐 27](#_Toc513538658)

[4.3.1 根据个体认知风格推荐教学案例 27](#_Toc513538659)

[4.3.2 根据群体信息推荐教学案例 27](#_Toc513538660)

[4.4 支持个体认知结构发展水平的评测模型的实现 27](#_Toc513538661)

[4.5 本章小结 28](#_Toc513538662)

[第五章 实验与结果分析 29](#_Toc513538663)

[5.1 实验概述 29](#_Toc513538664)

[5.1.1 实验数据准备 29](#_Toc513538665)

[5.1.2 支持评测个体认知结构的网络教学系统介绍 29](#_Toc513538666)

[5.2 实验一：基于个体认知风格的教学案例推荐 30](#_Toc513538667)

[5.2.1 实验方案 30](#_Toc513538668)

[5.2.2 实验过程 30](#_Toc513538669)

[5.2.3 实验结果分析 32](#_Toc513538670)

[5.3 实验二：基于协同过滤技术的教学案例推荐 32](#_Toc513538671)

[5.3.1 实验方案 33](#_Toc513538672)

[5.3.2 实验过程 33](#_Toc513538673)

[5.3.3 实验结果分析 34](#_Toc513538674)

[5.4 本章小结 35](#_Toc513538675)

[第六章 总结与展望 37](#_Toc513538676)

[6.1 本文的主要工作 37](#_Toc513538677)

[6.2 本文的不足与进一步研究 37](#_Toc513538678)

[致 谢 39](#_Toc513538679)

[参考文献 41](#_Toc513538680)

[附 录 43](#_Toc513538681)

[A. 所罗门学习风格自测问卷表 43](#_Toc513538682)

插图目录

[图 1-1 论文组织结构图 3](#_Toc512524804)

[图 2-1 知识挖掘过程 9](#_Toc512524855)

[图 3-1 C++数据表示知识平面 14](#_Toc512524909)

[图 3-2 基本知识 15](#_Toc512524910)

[图 3-3 模式知识 16](#_Toc512524911)

[图 3-4 个体认知结构模型 17](#_Toc512524912)

[图 4-2 教学用例集到认知风格的映射关系 22](#_Toc512524942)

[图 4-3 影响子图样例 24](#_Toc512524943)

[图 4-4 学习特征曲线样例 25](#_Toc512524944)

[图 4-5 评测模型实现的整体架构 28](#_Toc512524945)

[图 5-1 登录界面 30](#_Toc512524950)

[图 5-8 首次教学案例推荐 32](#_Toc512524957)

[图 5-9 选择群体推荐方式 33](#_Toc512524958)

[图 5-10 学习轨迹 34](#_Toc512524959)

[图 5-11 认知经验 34](#_Toc512524960)

[图 5-12 元认知能力 34](#_Toc512524961)

[图 5-13 默认教学案例 34](#_Toc512524962)

[图 5-14 基于群体推荐的教学案例 34](#_Toc512524963)

表格目录

[表 1-1 国内外网络教学平台属性信息 2](#_Toc512525073)

[表 5-1 15名大学生起始学习情况 29](#_Toc512525101)

[表 5-2 群体一和群体二学习整体情况比较表 32](#_Toc512525102)

[表 5-3 群体一和群体三学习整体情况比较表 35](#_Toc512525103)

# 引言

## 研究背景

随着大数据时代的到来，人们每天会从网络上接触到大量的数据信息。图像作为人们获取信息最直观、最简便的载体方式，具有重要的战略意义。智能手机以其移动、便捷、及时的特点改变了人们的信息接收习惯，刷微博、逛朋友圈已经取代阅读报纸成为了人们首选的获取资讯的方式。每个人既是信息的生产者又是传播者，人们随时记录、上传身边发生的事情，直观易懂的图像成为比文字更受青睐的传播符号。从图片、视频再到直播软件的推出是视觉传播技术的进步，同时也带来了新闻伦理方面的问题。

随着图像在网络上的流通和传播，由于图像大规模复制所带来的问题已经成为了当前社交媒体上的重点问题。因此，能够有效检测图像复制的技术显得十分必要，而图像复制相似度检测也正是是计算机视觉和图像处理领域中的一个研究热点。社交媒体的主要目的是让用户可以方便地与朋友和家人分享他们的生活和见解，但是，随着图像处理和制作软件的普及，社交媒体也变成了非法图像的传播平台。

因此，社交媒体平台需要使用图像复制相似度检测技术来确保用户上传的内容不包含暴力、色情等违法内容，该技术比使用手动内容审核更快速并且更大规模地审核内容，可以快速响应病毒式或突发攻击性内容。例如在恐怖袭击事件报道中，图像无疑是最有效率的传播手段，任何人只要下载了社交媒体软件，都可以第一时间将现场情况的图像传播出去。例如2015 年 8 月 13 日晚，法国巴黎 7 个地区先后遭遇了袭击，此后 ISIS 恐怖分子、 社交媒体上的个人用户和主流媒体的新媒体平台都发布了大量图像资料，视频中包括大量爆炸枪击等血腥 图像，该视频在 Youtube 上的点击量仅一天时间就达到了上百万次，并在Facebook 等社交媒体上得到多次 转发，触目惊心的画面通过网络给人们带来了极大的恐慌，甚至引发了“模仿犯罪”。如何快速响应具有负面影响的图像大规模传播，是互联网发展中要面对的重要问题之一。

图像复制相似度检测技术在社交媒体上的应用还不仅限于防止用户上传暴力或色情图片，它还可以用于监测图像的真实性。随着社交媒体的普及，人们越来越倾向于在线分享自己的生活照片，但也有一些人利用照片冒充他人或散布虚假信息。因此，社交媒体平台可以使用图像复制相似度检测技术来识别用户上传的照片是否已经存在，从而识别是否为虚假信息。

此外，图像复制相似度检测技术还可以用于防止图像的盗用。社交媒体平台上的图像通常是用户原创的，但在当前的流量时代，经常有其他用户将别人的原创图像直接复制并在其它平台上使用，那么原创者就会受到侵犯版权的威胁。因此，社交媒体平台如何识别和防止这种情况的发生就十分重要。图像复制相似度检测技术通过对图像进行分析和比对，可以有效地识别图像中的复制和篡改。例如，该技术可以检测图像中是否有重复的图像区域，以及这些重复的图像区域是否与原图像相似。此外，图像复制相似度检测技术还可以识别图像的旋转、缩放和剪切等变形，以判断图像是否被非法篡改。这将帮助版权所有者对图像使用情况进行监控。例如，版权所有者可以在网络上把他们的图像标记为受版权保护的图像，并使用图像复制相似度检测技术来检测这些图像是否被非法复制和使用。公司、组织和政府部门也可以通过这种方式保护它们的商标和标识。例如，公司可以使用图像复制相似度检测技术来检测是否有人非法使用他们的商标或标识，政府部门也可以检测非法使用国旗、国徽和其他重要标志的行为。

图像复制相似度检测技术的重要性不仅体现在互联网行业，它在多个领域也具有广泛的应用。在医学领域，图像复制相似度检测技术可以帮助医生快速检测出疾病的特征，在数字出版领域，图像复制相似度检测技术可以帮助出版商对图书中的图像进行版权检测，在安全领域，图像复制相似度检测技术可以帮助公安机关快速识别犯罪嫌疑人。

因此，图像复制相似度检测技术具有极高的现实价值和社会意义，它不仅能够提高工作效率，还能帮助我们解决许多实际问题，具有重要的应用前景。

## 研究目标

原始图像经过网络的传播会产生大量的副本，这些副本图像往往经过某些变形，比如分辨率变化，尺寸变化，色彩变化，添加数字水印等，如何快速的检测某张图像是否为另外一张图像的副本，比如一些包含负面内容的图像，这就是本课题要研究的内容：图像复制相似度检测技术。平台的审核者可以使用图像复制相似度检测技术来对违规内容图像的副本做出判断，使图像内容能够被更快的审核，并且这种方式比使用人工手动内容审核更加迅速，并且实现的规模更大。图像复制相似度检测技术能够快速检测出包含不健康内容图像的副本，尤其是针对网络突发事件，当突发事件产生出，其相关的图像是最新产生的，此前并未收录到社交媒体的数据库中，因此大规模不健康图像的副本在网络上进行传播时，如果对这些新产生的图像副本重新设计识别算法并部署检测模型，从时间上来说不能满足实时的要求，从资源成本的角度上来说需要大量资源支持，而直接通过图像复制相似度检测技术，根据某张原始图像，去检测该图像的复制图像，能够允许平台快速响应并对这些图像做下一步处理，比如在 2019 年美国发生的 Christchurch 枪击事件的现场直播，短时间在网络上出现了快速传播，此类事件要求非常快地采取行动，进行图像复制版本的检测和删除[2] 。

图像复制相似度检测技术长期以来受到关注，不断有新的方法被提出用于解决这个问题，然而由于互联网发展迅速，随之产生两个新问题难以解决：数据规模和新的图像变化。

1. 据统计，Google 的图片可检索量已达到百亿量级，百度的图片可检索量也达 50 亿，社交网站新浪微博的活跃用户 4.3 亿，图片日均发布量超过 1.2 亿张。消费类社区应用淘宝的后台服务器的图片存储量多达四百亿张，并且这个数量还在以每年十亿张的速度在增长。为图像复制相似度检测技术要处理的数据量极大，所以性能变得极其关键。一些低效的图像复制相似度检测技术在面对大规模数据时可能会变得极其缓慢，这将影响用户的体验。在浩如烟海的候选图像中，能够快速匹配对于大规模图像复制相似度检测技术的实用性至关重要。在应用图像复制相似度检测技术时，大多数查询图像在候选索引中并没有相应的匹配项。因此，检测系统的整体效率主要取决于其处理非匹配查询项的效率，因为查询图像的检测匹配对象是在检索候选对象中所占比例是非常低的，所以大规模数据会导致结果误报泛滥，这为图像复制相似度检测技术带来了额外的规模挑战。快速增长的数据规模是图像复制相似度检测技术要面临的一大难点，如何利用庞大的图像样本解决图像复制检索问题值得关注。
2. 另一方面为了满足更多的用户需求，诸如 TikTok，Insgram微博等等流行的社交媒体应用程序提供了丰富的图像和视频编辑选项，使得由普通用户产生的图像副本类型极大地被扩展。随着这些社交应用中的图像编辑功能的不断普及，用户可以对图像进行大量的修改，例如缩放、旋转、滤镜、去除背景等。这对图像复制相似度检测技术来说是一个巨大的挑战。此外，图像复制相似度检测技术也难以识别图像中隐藏的信息。例如，用户可以在图像中放置水印、文本等信息，而这样的信息很难被图像复制相似度检测技术识别，因为它们往往是隐藏在图像中的。并且，随着人们将更多的社交活动转移到移动设备上，手机屏幕截图和手机拍照已成为共享媒体的常见方式，通常用户希望分享的原始图像只会在手机屏幕截图中占据一小部分，屏幕截图会捕获很多额外的不相关内容（例如 app 界面或评论文本）。这种截图嵌套性质的图像，即将一张图像嵌入到另一张图像中，以及另一种局部图像，即截取一张图像的局部形成新的图像，很难通过图像复制相似度检测技术识别出图像中的信息，这些复杂图像也是图像复制相似度检测技术面对的另一大难题。

此外，用户在上传内容时可以从平台获得有关内容是否被阻止的即时反馈，如果上传被阻止，那么用户可以稍微修改它并重新尝试上传，这种行为相当于是一种复制检测攻击，因此社交平台上的图像复制相似度检测任务本质上是一种对抗性的任务。上面的变化要求图像复制相似度检测技术对于常见的图像变化必须具有较强的鲁棒性。

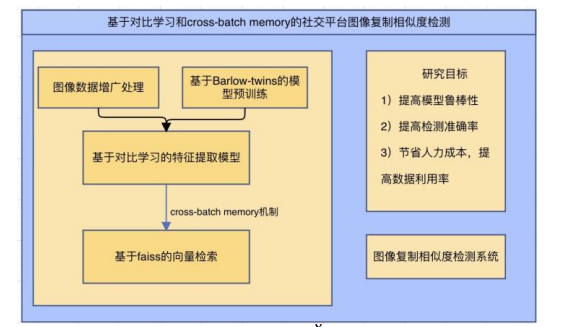
目前已有的图像相似度检索技术在解决随着网络发展出现的图像数据规模大，图像数据变化的问题上存在局限性，对于社交媒体中存在的图像变换鲁棒性不够强，难以应对目前复杂的数据规模和多变的场景，本课题通过对数据进行预处理，基于对比学习机制，使用深度学习作为特征提取模型的 encoder，引入 cross-batch memory 增强模型的鲁棒性，对于社交平台实时监控不良图像传播，打造清朗，文明，和谐的网络空间，维护网民身心健康有着重要作用。

## 研究内容

结合研究背景和研究目标，本课题中提出了一个用于图像复制相似度检测的方案，主要的内容可以概

括为以下三点：

1. 基于自监督学习的模型预训练和数据增广预处理：为了更好的利用社交平台中的海量数据，本课题引入了基于Barlow-Twins作为模型预训练的方法， Barlow-Twins 将神经科学中的冗余减少概念应用于自监督学习，训练的时候不需要构造大量的负样本，很适合用于模型预训练上获得合适的模型参数，为后面引入对比学习的模型训练做好铺垫。同时本课题使用Augly 库对数据集进行图像数据的增广处理，通过对图像的各类变形获得更多的图像变形副本，用来增加正样本的多样化，提高训练模型的准确率和鲁棒性。
2. 引入 cross-batch memory（XBM）的对比学习特征提取模型：在对社交平台中的图像应用图像复制相似度检测的过程中，大多数查询图像都不是某一图像的副本，在图像数据库中没有匹配对象。因此，为了利用大量不匹配的图像作为信息来训练模型，本课题计划 使用对比学习的方法，对比学习是一种自监督学习方法，用于在数据没有标签或只有部分数据有标签的情况下，通过让模型学习哪些数据点相似或不同来学习数据集的一般特征。对比学习的特性使得它非常适合图像进行复制相似度检索技术，少量的图像和其图像副本作为正样本对，不匹配的样本对作为负样本对，这样就能有效利用复制相似度检测在社交网络总体数据中的“低召回率”特征。社交平台中的数据存在着图像数量多，图像变化复杂的情况，因此在模型训练时应尽量扩大训练 样本的多样性来模拟现实社交平台中的数据。而在训练模型时往往会受到硬件资源的限制，需要采取mini-batch（小批量）训练的方式来训练模型，为了解决小批量训练带来的图像样本种类受限制的问题，本课题计划引入 cross-batch memory 方法，存储多个 mini-batch 的负样本，运用在当前 mini-batch 的训练中，增强训练样本的多样性。
3. 端到端的社交平台图像相似度检测系统：基于以上理论研究和构思，本课题设计并实现一种端到端的社交平台图像复制相似度检测系统， 该系统提供一个接口上传待检测图像，对于该图像，系统将使用上述经过训练的特征提取模型计算出该图像的特征向量，获得的向量会经由相似性搜索在给定的数据集中进行搜索，获得的图像即为图像复制相似度检测结果列表，并展示到图形界面中。该系统模型如下：



在本文中，我们提供了一系列技巧和强大的基线来竞争图像相似性挑战：NeurIPS'21 (ISC2021) [8] 的描述符跟踪。本次比赛建立了一个基准，以各种图像转换为特征，以模仿社交媒体中的真实案例。为了模仿大海捞针设置，查询和参考集都包含大多数不匹配的“干扰”图像。采用的评估指标是微平均精度，它会惩罚任何检测到的干扰查询对。该方法由三部分组成，即预训练、训练和测试。在预训练中，在ImageNet [6]上进行无监督预训练，而不是常用的有监督预训练我们凭经验发现 Barlow-Twins [36] 预训练优于其他一些无监督的预训练方法。在训练中，通过结合一袋技巧设计了一个强大的深度度量学习基线。此外，为了使样本对提供更多信息，使用了一些图像增强来生成训练图像。增强的多样性促进了学习鲁棒表示。在测试中，Descriptor Track设置直接使用欧几里得距离作为最终的匹配分数，从而禁止在Matching Track中进行分数归一化操作。针对这一特点，我们提出了一种新的描述符拉伸策略。它重新缩放（拉伸）描述符以稳定不同查询的分数。换句话说，它通过直接拉伸描述符来近似得分归一化效果。根据经验，这种拉伸策略显着提高了检索精度。所提出方法的图示如图 1 所示。总而言之，本文的主要贡献是： 1. 本文为图像副本检测提供了一系列技巧和强大的基线。 2. 提议的基线很好地处理了社交媒体中现实生活中的复制检测场景。 3. 我们的基线在 1 arXiv:2111.08004v2 [cs.CV] 2021 年 12 月 4 日的 526 名参与者中排名第三

## 论文结构

论文共分为6章，论文结构如图 1-1所示，每章的主要内容具体如下：



图 1-1 论文组织结构图

第一章为引言。

第二章介绍本文的相关研究工作。

第六章为结束语，总结了本文的研究工作，分析了本文工作中存在的不足，并在此基础上做进一步的展望。

# 相关研究

完整的教学评测不仅要有科学的认知理论作为思想指导，还需要有合理的方法论的指导。

## 图像复制相似度检测相关理论研究

图像复制相似度检测技术，可以看成是一种特殊的图像相似度检测技术，两者的不同之处主要在于前者更偏向检测图像是否为某图像的副本，因此要求对于图像变换具有更强的鲁棒性。但两者的主要实现方法是一致的，因此本课题将对图像相似度检索技术相关研究进行阐述。

传统的图像相似度检索技术是基于图像内容进行检索，提取出的图像特征的好坏直接决定了图像检索的效果。传统的描述图像特征的描述子有很多，基于全局的描述子有：表征图像颜色特征的颜色直方图[2]和颜色矩[3]、表征图像纹理特征的灰度共生矩阵（GLDM）[4]和 Gabor 描述子[5]等；基于局域的描述子有：局部二值模式（Local Binary Pattern，LBP）[6]、局部尺度不变特征（Scale Invariant Feature Transform， SIFT）[7]、方向梯度直方图（Histogram of Oriented Gradient, HOG）[8]等。传统的图像特征描述子特别是基于局域的局部特征描述子具有良好的稳定性，能够对图像的亮度、平移、旋转变化等各种畸变保持不变性。但是这类图像特征还是具有局限性的，它的精确率却难以达到高水平，并一直难以突破瓶颈，因为人工设计的特征描述符仅仅表示图像的表层特征，形状、背景、光线问题会大大影响图像的匹配精度。 所以无法表示深层图像特征，造成语义鸿沟。而深度神经网络 （Deep Neural Network, DNN）的出现改变了这一点。

传统的图像相似度检索技术是基于图像内容进行检索，图像特征描述子对图像检索的效果有着重要的影响。图像特征描述子可以分为基于全局的描述子和基于局域的描述子。基于全局的描述子包括颜色直方图、颜色矩、灰度共生矩阵（GLDM）、Gabor描述子等。基于局域的描述子包括局部二值模式（Local Binary Pattern，LBP）、局部尺度不变特征（Scale Invariant Feature Transform， SIFT）、方向梯度直方图（Histogram of Oriented Gradient, HOG）等。这些图像特征描述子，特别是基于局域的局部特征描述子，具有良好的稳定性，能够对图像的各种变化，诸如亮度、平移、旋转变化等变形保持较强的鲁棒性。然而，传统的图像特征描述子仍然存在局限性，由于人工设计的特征描述符仅仅表示图像的表层特征，形状、背景、光线问题会对图像的匹配精度产生很大的影响，这就造成了语义鸿沟，即无法表示深层图像特征，导致基于传统的图像特征描述子的精确率难以达到高水平。但是，随着深度神经网络的出现，这一局面得到了改变。

基于深度神经网络提取图像特征的基于深度学习的图像检索方法的出现，使得图像检索开始向着更快速更准确的方向发展。在文献[10]中，作者提出了一种名为REMAP的新型全局描述符，该方法基于卷积神经网络，旨在通过从多个卷积层学习和聚合深度特征的层次结构来实现对图像的更好的全局描述。其特点在于使用三元组损失进行端到端训练，可以更加准确的捕捉图像的特征。在文献[11]中，作者提出了一种计算掩码的策略，包括SIFT-mask、SUM-mask和MAX-mask，以选择具有代表性的局部卷积特征子集并去除冗余特征。这些策略的目的在于更好的捕捉图像的特征，以便对图像进行更准确的描述。文献[12]中，作者提出了一种高效的、自适应的编码驱动图，它通过在自动编码器的上下文中进行解码来更新，在框架中引入了双瓶颈，可以协同交换关键信息。这种方法可以更加高效地捕捉图像的特征，以便对图像进行更准确的描述。在文献[13]中，作者设计了两个二阶模块，其中一个专注于二阶空间信息，以提高局部和全局图像特征描述符的性能，另一个模块将二阶相似性 (SOS) 损失扩展到图像检索的全局特征描述符。文献[14]通过结合全局特征的广义均值池化和局部特通过高效的特征提取实现准确的检索。此外，文献[15]提出了一种新的基于深度学习的图像检索方法，该方法结合了预训练的卷积神经网络和弱监督学习，可以从图像中提取更加精细和鲜明的特征，从而提高图像检索的准确性。

文献[16]提出了一种新的多尺度注意力网络（MSA-Net），该网络在图像检索中同时利用多个注意力级别进行特征提取。文献[17]提出了一种新的长短时记忆卷积神经网络（LSTM-CNN），它可以通过结合长短时记忆网络的优点和卷积神经网络的优点来提高图像检索的性能。文献[18] 提出一个统一的上下文感知注意力网络（CAAN），它通过聚合全局上下文选择性地关注关键的局部片段（区域和单词），同时利用全局模态间对齐和模态内相关性来发现潜在的语义关系。

相比于传统特征描述子，深度神经网络对图像所包含的内容具有很强的抽象化表征能力，可以将图像特征全面、有效地组织起来。

深度度量学习是指利用深度神经网络学习输入数据之间的相似度或距离度量的一种方法。在深度度量学习中，常见的方法之一是pair-based方法，其基本思想是通过对比学习来学习样本之间的距离或相似度。

以下是几种常见的pair-based深度度量学习方法：

Contrastive Loss

Contrastive Loss是一种简单但有效的pair-based深度度量学习方法。它通过学习将同类样本对（positive pairs）的距离拉近，将异类样本对（negative pairs）的距离推远来学习样本之间的距离度量。具体来说，对于每个样本，选择一个与之同类的样本和一个与之异类的样本，然后使用Contrastive Loss计算它们之间的距离。该方法可以通过随机选择positive pairs和negative pairs来生成训练数据。

Triplet Loss

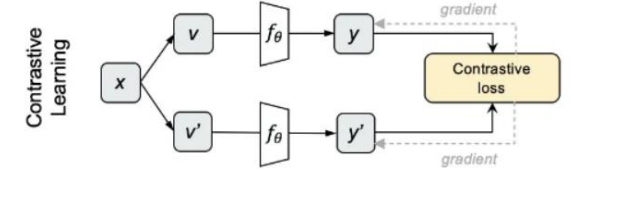
Triplet Loss是一种pair-based深度度量学习方法，它比Contrastive Loss更为常用和有效。该方法通过比较同类样本之间的距离和异类样本之间的距离，学习如何将同类样本间的距离缩小，将异类样本间的距离扩大。Triplet Loss的训练数据是由三元组（anchor, positive, negative）组成的，其中anchor和positive是同类样本，negative是异类样本。通过学习让anchor和positive之间的距离小于anchor和negative之间的距离，来学习距离度量。

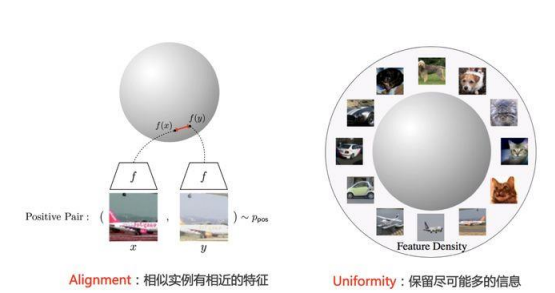
Quadruplet Loss

Quadruplet Loss是Triplet Loss的一种扩展。在Triplet Loss中，anchor和positive之间的距离被要求小于anchor和negative之间的距离，但没有要求anchor和negative之间的距离大于positive和negative之间的距离。Quadruplet Loss通过引入第二个negative样本，要求anchor和negative之间的距离大于positive和negative之间的距离，从而更有效地学习距离度量。

N-Pair Loss

N-Pair Loss是一种适用于多类分类的pair-based深度度量学习方法。该方法使用N个样本组成的N-Pair作为训练样本，其中包括一个anchor和N-1个与之同类的positive样本，以及M个与之不同类的negative样本。N-Pair Loss要求anchor与positive之间的距离小于anchor与任意一个negative之间的距离，从而学习样本之间的距离度量。N-Pair Loss的优点在于可以有效地处理多类分类问题。

对比学习本质上是深度度量学习的一种，深度度量学习作为一种结合了深度神经网络与度量学习的算法，不仅实现了利用深度神经网络提取更好的图像特征，还实现了利用度量学习提高图像特征的区分性。基于深度度量学习的图像检索算法的目标是学习将图像的原始特征映射到一个低维稠密的向量空间，这个向量空间也被称为嵌入空间（embedding space），而映射则通过深度神经网络来实现，在嵌入空间上的图像特征也可以称为图像特征的嵌入表达，在嵌入空间上使用距离函数计算各个图像特征之间的距离，相同类别的图像之间的距离比较近，不同类别的图像之间的距离比较远。常见的距离函数包括欧氏距离、曼哈顿距离、余弦距离等。在嵌入空间中，距离函数可以用内积来计算，即样本的相似性与它们的内积成正比。对比学习的目标是学习到一个映射函数，使得同类样本的内积尽可能大，异类样本的内积尽可能小。

对比学习系统一般具备两个属性：Alignment 和 Uniformity（参考下图）。所谓“Alignment”，指的是相似的例子，也就是正例，映射到单位超球面后，应该有接近的特征，也即是说，在超球面上距离比较近； 所谓“Uniformity”，指的是系统应该倾向在特征里保留尽可能多的信息，这等价于使得映射到单位超球面 的特征，尽可能均匀地分布在球面上，分布得越均匀，意味着保留的信息越充分。“分布均匀和保留信息” 两者之间的关联是：分布均匀意味着两两有差异，也意味着各自保有独有信息，这代表信息保留充分。

监督对比学习需要标注数据，通常使用softmax交叉熵等分类损失函数来对模型进行监督训练，因此训练模型的准确性取决于标注数据的质量和数量，同时还面临标签不准确或者标注不完整等问题。相比于有监督对比学习，无监督和自监督对比学习更加具有灵活性和可扩展性。监督学习近些年获得了巨大的成功，但是有如下的缺点：1）人工标签相对数据来说本身是稀疏的， 蕴含的信息不如数据内容丰富；2）监督学习只能学到特定任务的知识，不是通用知识，一般难以直接迁移到其他任务中。

与之相对的自监督学习不需要人工标注的类别标签信息，直接利用数据本身作为监督信息，学习样本数据的特征表达，应用于下游的任务。自监督学习又可以分为对比学习[24] (contrastive learning) 和生成学习(generative learning) 两条主要的技术路线。对比学习的核心思想是讲正样本和负样本在特征空间[25] 对比，学习样本的特征表示，难点在于如何构造正负样本。 对比学习首先学习未标记数据集上图像的通用表示形式，然后可以使用少量标记图像对其进行微调， 以提升在给定任务(例如分类)的性能。简单地说，对比表示学习可以被认为是通过比较学习。相对来说，生成学习(generative learning)是学习某些（伪）标签的映射的判别模型然后重构输入样本。在对比学习中， 通过在输入样本之间进行比较来学习表示。对比学习不是一次从单个数据样本中学习信号，而是通过在不 同样本之间进行比较来学习。可以在“相似”输入的正对和“不同”输入的负对之间进行比较[26] 。下图为对比学习训练过程。

自监督对比学习则需要一些辅助任务作为标签。这些任务需要使用少量的人工标注数据来生成，例如颜色匹配、空间变换、图像旋转等。自监督对比学习通过将图像分割成几个不同的部分，将一些部分作为“正”示例，另一些部分作为“负”示例，并使用对比损失来学习数据表示。常见的自监督对比学习方法包括MoCo（Momentum Contrast）、SimCLR（Simple Framework for Contrastive Learning of Visual Representations）、BYOL（Bootstrap Your Own Latent）、SwAV（Unsupervised Learning of Visual Features by Contrasting Cluster Assignments）等。

无监督对比学习不需要标注数据，而是通过自监督任务，如图像重建、自编码器、解码预测等，利用数据本身的自然结构来学习表示。无监督对比学习方法包括自编码器、变分自编码器、生成对抗网络（GAN）、自监督的预测等。其中，自编码器和变分自编码器的任务是重建原始数据，而GAN的任务是生成与原始数据相似的新数据。自监督预测的方法则是通过将原始数据中的一部分作为输入，另一部分作为输出，并利用对比损失来学习数据表示。

文献[]提出来Siamese Networks（孪生网络），这是一种特殊的神经网络结构，可以处理成对图像的检测和相似度对比任务。该方法训练两个网络，每个网络分别处理一个输入。这两个网络共享相同的参数，并且它们的输出向量被传递到距离函数中进行比较。Triplet Networks（三元组网络）在文献[]中被提出，该神经网络结构同样可以用于处理图像相似度比较任务。不同于Siamese Networks中处理一对数据，Triplet Networks处理三元组数据（anchor，positive和negative）。其中，anchor表示一个样本，positive表示与anchor相似的样本，negative表示与anchor不相似的样本。该方法旨在将anchor与positive的距离缩小，同时将anchor与negative的距离扩大。而文献[]中的Contrastive Multiview Coding（CMC）是一种无监督学习方法，它利用多个视角的信息来学习图像表示。CMC包括两个阶段：第一个阶段是视角编码，通过编码来自多个视角的数据来学习图像表示；第二个阶段是对比学习，通过比较同一样本的不同视角之间的距离来训练模型。在文献[]中提出一种用于学习图像表示的对比学习方法SimCLR，它利用数据增强技术来生成不同视角的图像，并使用大量的对比样本来训练模型。在训练过程中，模型通过最大化同一样本的不同视角之间的相似度，最小化不同样本之间的相似度来学习图像表示。以及最近提出的对比学习方法SwAV[]，它利用聚类技术来学习图像表示。SwAV使用数据增强技术来生成不同视角的图像，并通过最大化同一聚类中的样本之间的相似度，最小化不同聚类之间的相似度来学习图像表示。

对比学习在计算机视觉领域中的应用非常重要。它可以提高模型的泛化能力，解决数据利用率低的问题，提高模型的性能和准确性。而本文中提出的基于对比学习和cross-batch memory的社交平台图像复制相似度检测技术正是应用对比学习的一个案例。对比学习在图像复制检测技术中应用非常广泛，并且具有很多优点，它可以处理大规模数据，从而提高了检测的速度和效率。通过对比学习，可以学习到图像中的一些潜在特征，从而可以使用这些特征来快速地匹配和检测图像中的重复部分。这种方法比传统的方法更加快速和高效。此外，对比学习可以处理多种不同的噪声和干扰，从而提高了检测的准确性和鲁棒性。例如，当检测到复制图像时，对比学习可以通过学习图像中的噪声和干扰来排除这些噪声和干扰，从而提高检测的准确性。

综上所述，对比学习在图像复制检测技术中具有广泛的应用和很多优点，包括能够学习图像中的相似性和差异性、处理大规模数据、处理多种不同类型的图像和噪声、提高检测的准确性和鲁棒性等。随着对比学习技术的不断发展，它将在图像复制检测技术中发挥越来越重要的作用。

## 本章小结

本章围绕图像复制检测技术的研究现状。为接下来的基于对比学习和cross-batch memory的社交平台图像复制相似度检测技术模型提供了理论基础，为模型的评测提供了技术支撑。

# 数据增广与模型预训练

本章基于建构主义学习理论、认知风格理论和认知能力研究，寻求一种支持表述个体认知结构的表达模型。

### 基于Augly的数据增广

本文使用ISC2021数据集作为实验数据集，其中，query数据集中部分图像由reference数据集中部分图像经过各种变形获得，该数据集为了增加鲁棒性，在query数据集中混入了大量干扰项，也就是与变形图像和原始图像都毫无关系的图像，变形获得的副本图像在数据集中仅占据10%，剩下部分的图像都是无法用于匹配的“干扰”图像。

为了提高训练效果，获得更多的正样本用于训练，同时为了更加接近现实社交媒体中产生的多种多样的数据变换，本课题对数据集做了预增广处理。本课题通过Augly[22]构建了多种图像变化以扩充数据集的内容。AugLy是Facebook AI Research团队开源的一个强大的数据增强库，它提供了各种图像、文本和音频增强方法，可以帮助深度学习模型更好地学习数据的特征，从而提高模型的准确性。在本文中，我们结合Augly库构建了多种图像变换的种类。

基本增强集包括随机调整大小裁剪、随机旋转、随机像素化、随机像素洗牌、随机透视变换、随机填充、随机图像底层、随机颜色抖动、随机模糊、随机灰度、随机水平翻转、随机表情符号叠加、随机文本叠加、随机图像叠加和调整大小。在一个图像上使用所有基本增强的显示如图 3 所示。第一个图像是调整大小的原始图像

  transforms.ColorJitter(0.7, 0.7, 0.7, 0.2),

        RandomPixelization(p=0.25),

        ShufflePixels(factor=0.1, p=0.25),

        OneOf([EncodingQuality(quality=q) for q in [10, 20, 30, 50]], p=0.25),

        transforms.RandomGrayscale(p=0.25),

        RandomBlur(p=0.25),

        transforms.RandomPerspective(p=0.25),

        transforms.RandomHorizontalFlip(p=0.5),

        transforms.RandomVerticalFlip(p=0.1),

        RandomOverlayText(p=0.25),

        RandomEmojiOverlay(p=0.25),

RandomEdgeEnhance

随机调整大小裁剪：随机裁剪图像并调整其大小，可以增加模型对不同尺寸图像的鲁棒性。

随机旋转：随机旋转图像一定角度，可以增加模型对旋转变换的鲁棒性。

随机像素化：随机将图像像素化，即用相同颜色的像素块替换相邻像素，可以增加模型对类似像素化的图像处理的鲁棒性。

随机像素洗牌：随机将图像像素按一定规律洗牌，可以增加模型对类似像素洗牌的图像处理的鲁棒性。

随机透视变换：随机对图像进行透视变换，可以增加模型对透视变换的鲁棒性。

随机填充：在图像周围随机填充一定颜色的像素，可以增加模型对填充边界的图像处理的鲁棒性。

随机图像底层：随机选择一张图像，将当前图像的像素和该图像的像素按一定权重混合，可以增加模型对图像混合的鲁棒性。

随机颜色抖动：随机改变图像颜色通道的亮度、饱和度和对比度等属性，可以增加模型对颜色变换的鲁棒性。

随机模糊：随机对图像进行模糊处理，可以增加模型对模糊图像的鲁棒性。

随机灰度：随机将图像转换为灰度图像，可以增加模型对灰度图像的鲁棒性。

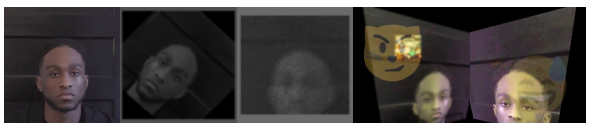
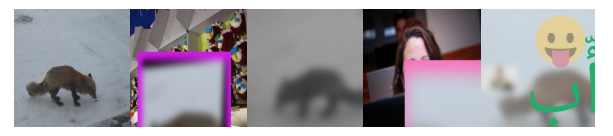
随机水平翻转：随机对图像进行水平翻转，可以增加模型对水平翻转的鲁棒性。

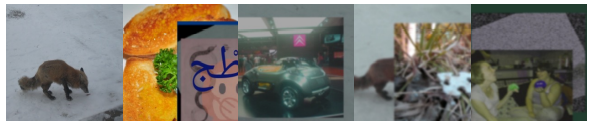
随机表情符号叠加：随机将一些表情符号叠加在图像上，可以增加模型对叠加物体的鲁棒性。

随机文本叠加：随机将一些文本叠加在图像上，可以增加模型对叠加物体的鲁棒性。

随机图像叠加：随机将一张图像与当前图像按一定权重混合，可以增加模型对图像混合的鲁棒性。

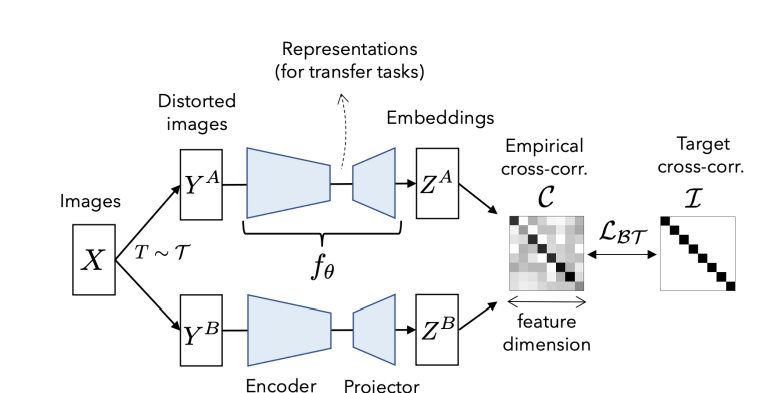
调整大小：调整图像的大小，可以增加模型对不同尺寸图像的鲁棒性。

总的来说，我们构建了较为丰富的图像增强方法，可以帮助图像复制检测模型更好地学习图像的各种特征，从而提高模型的准确性。这些增强方法不仅可以用于图像分类和目标检测等任务，还可以用于图像生成和图像重建等任务。为了更清晰地展示在本文中所使用的数据增强的种类，下面列出几种图像转换示例：



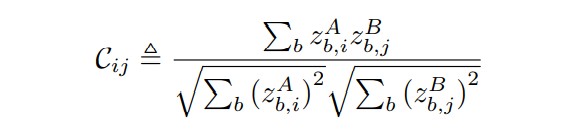
基于Barlow-Twins的模型预训练

在 ISC2021 中，一个类别被定义在一个非常“严格”的级别，即由精确复制、近精确复制和编辑副本生成的图像被认为在同一类别中，而属于同一实例或对象的两幅图像不属于同一类别。然而，当在 ImageNet 上使用完全监督的预训练时，两个图像共享相同的实例或对象属于一个类。因此，类别的定义在 ISC2021 和 ImageNet 之间是矛盾的。幸运的是，最近关于自我监督学习的研究提供了一个新的方向。它将每个图像定义为一个类别，并使用数据增强的不变性作为其主要训练标准 [8]。虽然直接采用自监督学习方法来训练副本检测模型似乎很自然，但我们选择放弃这种解决方案有两个原因。首先，自监督学习方法通​​常要花费很多天才能获得收敛结果。例如，在 ImageNet 上，BYOL [13] 或其 Momentum2 Teacher 版本 [19] 需要大约两周时间使用 8 个 NVIDIA Tesla V100 32GB GPU 使用 ResNet-152 [10] 主干预训练 300 个时期。此外，当在线使用复杂的转换时，时间变得更长，这是无法承受的。其次，即使资源是无限的，我们也无法仅通过自监督学习方法训练得到令人满意的表现，这可能是因为没有仔细调整超参数。因此，我们选择使用 BYOL [13]（其 Momentum2 Teacher 版本 [19]）和 Barlow-Twins [37] 在 ImageNet 上获得预训练模型。使用的增强遵循其原始实现中的默认设置。

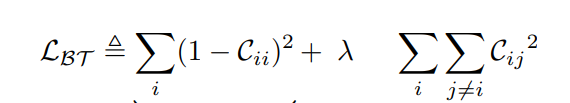
本文中选择使用 Barlow-Twins [23] 来进行本课题中的模型预训练。Barlow Twins由Jeffrey Barlow和Jared Kaplan在2021年提出，是最近提出的一种自监督学习方法Barlow-Twins算法的主要目的是利用无标签数据进行自监督学习，以帮助机器学习模型更好地理解数据分布。Barlow-Twins模型不需要依赖于大量的有标签数据来进行训练，从而可以在数据有限的情况下获得很好的表现。相比于其他方法使用大批量样本数据，该方法不需要构建过多负样本，因此很适合用于训练集数量较小的模型预训练。它的主要架构见图。

Barlow-Twins模型的基本结构是双胞胎网络。它由两个相同的神经网络组成，分别被称为“胞妹”和“胞姐”。双胞胎网络的设计灵感来自于人类双胞胎的相似性。与人类双胞胎一样，胞妹和胞姐也有着很高的相似性，但并不完全相同。这种设计可以使得网络学习到更加有意义的特征表示，从而提高模型的性能。

经过训练的双胞胎网络可以被用来提取有意义的特征表示。这些特征表示可以被用于下游任务，例如图像分类、目标检测、语义分割等。这两个网络使用相同的权重和结构，在训练过程中双胞胎网络对由同一个数据样本进行旋转、翻转、缩放等操作，生成的不同增广样本，将两个样本送入上下两个网络（上下网络相同），具体来说，一个网络接收原始图像，另一个网络接收通过变换得到的图像。这两张图像经过encoder提取特征后得到两个不同的embedding，这两个embedding被用来计算它们的cross-correlation matrix（互相关矩阵），其中互相关矩阵的元素是由下面公式计算得到：

其中 Z A 和 ZB本次batch中第b个图像的embedding，，i, j是embdding的第i, j维。 C 是一个方阵，其大小与Z A 和 ZB的维数相同，Ci j 是方阵C的元素，值介于 -1（即完全反相关）和 1（即完全相关）之间。

Barlow-Twins模型利用上面的互相关矩阵来鼓励网络学习到更加相似的特征表示。具体来说，模型通过最小化互相关矩阵来使得互相关矩阵的元素之间的相似度更加接近。这样，模型就可以学习到更加有意义的特征表示，从而提高其在计算机视觉任务中的表现。下面的式子为Barlow-Twins的损失函数：

其中 λ 是一个正常数，权衡损失的第一项和第二项的重要性，Ci i 是互相关矩阵C的对角线元素，Ci j是互相关矩阵C的非对角线元素。Barlow-Twins的主要思想就是利用损失函数来降低冗余，使上述互相关矩阵尽可能接近identity matrix（单位矩阵），即对角线元素接近1，非对角线元素接近0，因此就代表着同一样本不同增广版本经过网络后提取的两个特征向量的相同维度分量非常相似，不同维度分量的冗余应该最小化（意味着表示不同的信息）。

Barlow-Twins算法已经在许多计算机视觉领域的应用中取得了显著的成果。例如，Barlow-Twins算法可以用于图像分类、目标检测、语义分割等任务。此外，Barlow-Twins算法也可以与其他神经网络模型结合使用，例如Transformer等。

本章小结

## 模型

### Resnet50

ResNet-50是深度学习中最为经典的卷积神经网络之一，它在许多计算机视觉任务中表现出了出色的性能，已经被广泛应用于许多实际应用中。其中，最为典型的应用之一是图像分类（image classification），即将输入的图像分为不同的类别。在ImageNet数据集上，ResNet-50的分类准确性可以达到约76％。此外，ResNet-50还可以应用于目标检测（object detection）、语义分割（semantic segmentation）等其他计算机视觉任务中。

ResNet-50的结构是由50个卷积层、池化层、全连接层以及一些辅助层组成的。它采用了残差块（residual block）的设计思路，这种设计可以帮助模型更好地训练深度网络，避免了梯度消失和梯度爆炸等问题。具体来说，ResNet-50中的残差块包括了两个3×3大小的卷积层，以及一个跨层连接（shortcut connection），用于直接将输入的特征图传递到输出层，从而保证信息不会丢失。此外，为了减小特征图的大小，ResNet-50中还采用了池化层来对特征图进行下采样操作。

本文选择使用ResNet-50作为图像复制相似度检测模型的backbone，ResNet-50作为经典模型，在对比学习任务中有着很强优势，它可以提高模型的性能、加快训练速度、提高泛化能力，并且可以通过预训练模型来更快地构建和优化，由于ResNet-50在大规模图像数据集（如ImageNet）上进行了预训练，因此可以将该模型作为预训练模型，在其它图像相关任务中进行微调，从而大大减少了模型训练时间和计算成本。

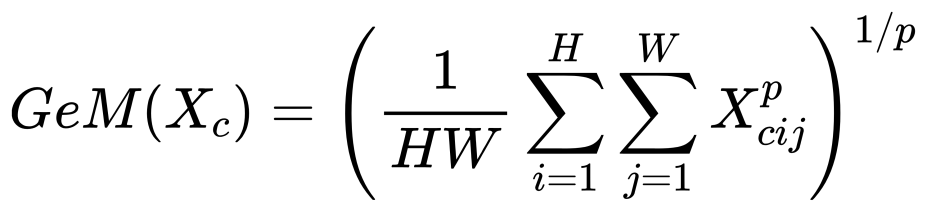
### Gem

在深度学习中，池化操作通常被用于减小特征图的空间大小，以便于降低计算复杂度和减小过拟合。常见的池化操作包括最大池化（Max Pooling）、平均池化（Average Pooling）等。相对于这些传统的池化操作，Generalized Mean (GeM) Pooling是在2018年由Zhou等人提出的一种池化方法，旨在改进平均池化和最大池化的缺点。GeM池化基于平均池化，但是使用了一种可调节的参数$p$，以便对特征图中不同位置的信息进行自适应的加权。

在这项工作之前，已经有许多工作致力于改进传统的池化方法。其中，最大池化和平均池化是最常用的方法，它们在许多深度学习任务中都取得了不错的效果。然而，这些池化方法存在一些缺陷。最大池化可能会丢失特征图中的有用信息，因为它只选择特征图中最大的值作为池化结果；而平均池化会平均特征图中的所有值，这可能会导致一些噪声的影响。因此，GeM池化提供了一种基于平均池化的改进方法，旨在同时避免上述两种缺陷。

自从GeM池化的提出，它已经在许多计算机视觉任务中得到了广泛的应用。例如，它已被成功地用于图像分类、目标检测、行人重识别等任务中。同时，GeM池化也被证明对于提高模型的鲁棒性和泛化能力是非常有帮助的。

GeM Pooling的原理是通过对每个通道上的特征进行幂指数的非线性变换，得到一个代表整个特征图的池化结果。其具体计算公式如下：

$$\text{GeM}(X\_c) = \left(\frac{1}{HW}\sum\_{i=1}^H\sum\_{j=1}^W X\_{cij}^p\right)^{1/p}$$

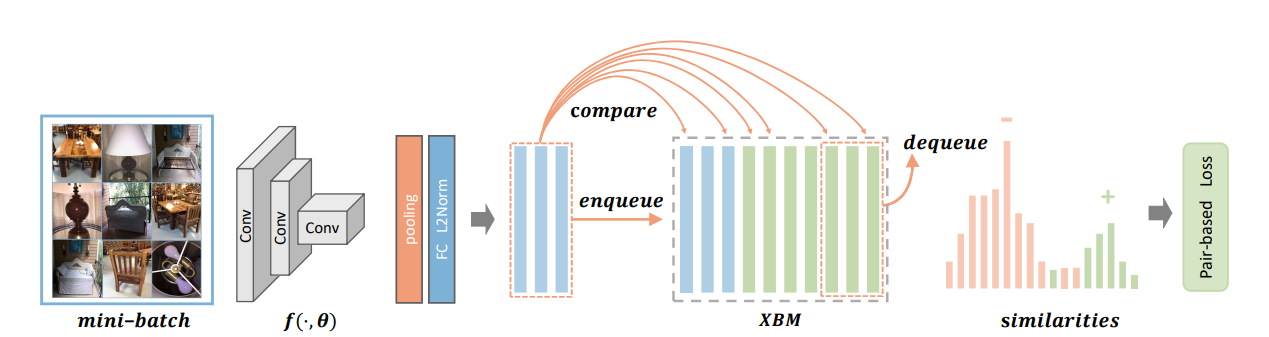
其中$X\_c$表示特征图中的第$c$个通道，$H$和$W$分别表示特征图的高和宽，$X\_{cij}$表示第$c$个通道中第$(i,j)$个位置的特征值，$p$是GeM的超参数，通常称为幂指数，表示对每个像素进行非线性变换的程度。在 GeM中，$p$ 用于控制平均池化的范围。$p$ 越大，池化区域的重点将放在 feature map 的最大值上，而 $p$ 越小，池化的区域会更加关注全局平均值。因此，通过调整 $p$ 值，可以实现对网络特征提取的不同调整和优化。当$p=1$时，GeM Pooling等价于平均池化，即对每个通道上的特征取平均值；当$p=\infty$时，GeM Pooling等价于最大池化，即对每个通道上的特征取最大值。因此，GeM Pooling可以看作是平均池化和最大池化的一种平滑形式。

### Cross-batch memory

本课题的模型训练是基于对比学习的方法，而挖掘信息丰富的负样本对于对比学习至关重要，但是这项任务本质上受到（mini-batch）小批量训练的限制，小批量训练允许在存储器中只存储部分数据，对于模型训练需要的存储变得更高效，但这样也有缺席，即训练模型在每次迭代中只能访问小批量内的样本，这就导致了在构造负样本时，每次迭代都只能使用有限个图像与当前图像构成的负样本对。

为了解决这种局限性，本课题计划引入cross-batch memory（XBM）机制，XBM是一种用于解决深度神经网络中内存效率和泛化性能之间的平衡问题的方法，它的核心思想是将当前批次的激活作为记忆池（memory bank），并将记忆池中的激活用于下一个批次的训练中。这样一来，模型可以从前几个批次中学习到的信息和当前批次中的信息进行联合训练，从而提高模型的泛化性能。

XBM的优势主要表现在两个方面。首先，Cross-batch memory可以提高模型的泛化性能。通过使用记忆池来融合前几个批次中的信息和当前批次中的信息，模型可以更好地适应不同的数据分布，从而提高泛化性能。其次，Cross-batch memory可以提高模型的内存效率。由于Cross-batch memory使用当前批次的激活作为记忆池，而不是使用全部的训练数据，因此可以大大降低内存使用量。

可以发现，在模型训练的中后期，每次迭代产生的embedding的变化很小，这表明在训练过程中，之前的迭代中计算的embedding特征可以被近似XBM机制设置一个动态队列，用于存储之前迭代产生的embedding， 允许模型在当前小批次训练过程中构造出足够的负样本对，将小批次训练本身导致的局限性问题降到最低。

XBM的主要机制见上图，它可以概括为：

1. 具体来说，Cross-batch memory机制包括以下几个步骤：
2. 存储历史样本特征。在模型的训练过程中，对每个batch的特征进行存储。这里存储的特征可以是任意网络层的输出，如卷积层、池化层、全连接层等。
3. 计算当前batch的特征。在前馈过程中，计算当前batch的特征表示，同样可以是任意网络层的输出。
4. 计算历史样本和当前batch的相似性。使用存储的历史样本特征和当前batch的特征表示，计算它们之间的相似性。这里可以使用余弦相似度等度量方式。
5. 更新当前batch的特征。根据当前batch和历史样本之间的相似性，对当前batch的特征进行加权更新，以得到更好的特征表示。具体来说，可以使用如下公式进行特征更新：

$x'i = x\_i + \alpha \sum{j=1}^{m} w\_{ij}h\_j$

其中，$x\_i$是当前batch的第$i$个样本特征表示，$h\_j$是历史样本的第$j$个特征表示，$w\_{ij}$是当前batch的第$i$个样本和历史样本的第$j$个之间的权重，$\alpha$是学习率。

1. 进行前向传播和反向传播。使用更新后的当前batch的特征表示进行正常的前向传播和反向传播，更新网络参数。

Cross-batch memory机制的核心思想是在不同batch之间共享信息，以增强模型的泛化性能。通过存储历史样本的特征表示，并根据相似性信息对当前batch的特征进行更新，可以减少过拟合，提高模型的泛化能力。此外，Cross-batch memory机制还可以在数据量较少的情况下提高模型的性能，使得模型可以从有限的数据中学习更多的信息。

本课题计划在使用对比学习的方法训练模型时加入XBM机制，XBM 可以直接集成到基于对比学习的模型训练框架中，构造多样化的负样本对，显著提高模型性能。特别是，在没有花里胡哨的情况下，XBM 在概念上很简单，易于实现，使用XBM仅需要额外的 0.2 GB 左右的GPU 内存，对于内存十分友好。

### Contrastive loss

构造损失（Contrastive Loss）是一种基于对比学习（Contrastive Learning）的损失函数。在对比学习中，模型通过学习将来自同一样本的不同变体（如不同的裁剪、旋转、变形等）的特征嵌入映射到相近的位置，而将来自不同样本的特征嵌入映射到较远的位置。这样一来，模型就能够根据特征之间的距离，来区分不同样本或者同一样本的不同变体。构造损失的作用就是通过最小化同一样本的不同变体之间的距离，来促进模型学习更加鲁棒和具有判别性的特征表示。

本文的loss函数选择了经典的contrastive loss，对于给定的一对样本$(x\_i,x\_j)$，将它们分别映射到特征空间中，得到它们的特征向量$f\_i$和$f\_j$，然后计算它们在特征空间中的距离$d\_{ij}$，距离的计算可以使用欧氏距离或余弦距离等度量方式。如果它们属于同一类别，那么我们希望它们的距离$d\_{ij}$尽可能地小，如果它们属于不同类别，那么我们希望它们的距离$d\_{ij}$尽可能地大。Contrastive loss的目标是最小化同类别样本的距离，最大化不同类别样本的距离，它的具体的损失函数可以表示为：



其中 $D(x\_1, x\_2)$ 是样本 $x\_1$ 和 $x\_2$ 之间的距离，通常使用欧几里得距离或余弦相似度计算。$m$ 是一个超参数，用来控制同类样本之间的距离，它也被称为negative margin，表示当两个样本属于不同样本时，它们之间的距离应该超过m。当 $y=1$ 时，表示两个样本属于同一类，Contrastive loss 会最小化样本之间的距离，即拉近样本。当 $y=0$ 时，表示两个样本属于不同类，Contrastive loss 会惩罚距离小于 $m$ 的样本对，即推远样本。整个模型的损失函数通常由分类损失和 Contrastive loss 组成。

综上所述，本课题采取对比学习的框架，将XBM机制应与对比学习相结合，提高模型训练效率，避免了小批量训练带来的局限性，本课题的训练框架见下图

## 基于faiss的向量搜索

向量搜索是一种基于向量空间模型的搜索方法，它通过计算向量之间的相似度来实现搜索。向量搜索是指在大规模高维向量数据集中快速搜索最相似的向量的过程。与传统的关系型数据库不同，向量搜索并不是以文本或数字作为查询的关键字，而是以高维向量作为查询条件。向量搜索技术广泛应用于图像、语音、自然语言处理等领域，其中在图像领域中应用尤为广泛。在进行图像复制检测过程中，需要进行以下几个步骤：

1. 通过模型对每张图像进行特征提取，得到一个高维向量表示。
2. 将这些向量存储在HDF5文件中，方便后续读取建立向量索引。
3. 将待检测图像的向量作为查询条件，使用向量搜索算法在向量索引中查找与之最相似的向量。
4. 根据相似向量的相似度值，判断查询的图像是否是某张图像变形后的副本。

在这个过程中，向量搜索算法扮演了一个非常重要的角色，向量搜索是图像复制检测技术的核心部分之一，是一种基于向量相似度的检索方法，适用于大规模的高维向量数据集。在图像复制检测中，向量搜索使用深度学习模型生成的图像特征向量作为检索对象，这些特征向量可以表示图像的高层语义信息。基于向量相似度的搜索可以快速找到与待检测图像相似的特征向量，进而找到图像库中与待检测图像相似的图像，从而更加精准地检测出是否存在图像复制行为。在本文中，我们选择FAISS库实现对图像特征向量的搜索。

FAISS（Facebook AI Similarity Search）是Facebook AI Research团队开源的高效相似度搜索库，它提供了一系列的向量索引算法和搜索工具，可以快速地在大规模的高维向量空间中进行相似度搜索。FAISS提供了丰富的接口和功能，可以被广泛地应用于文本搜索、图像搜索、音频搜索、推荐系统等领域。

在本文中，我们通过FAISS库中实现构建图像复制检测系统的向量索引。向量索引是将一组向量构建为索引结构，以便能够高效地搜索与输入向量相似的向量。在将ISC2021的参考集中所有图像的特征向量放进构建成向量索引后，对查询图像的特征向量进行向量搜索。向量搜索是指在向量空间中搜索与查询向量最相似的向量。在图像复制检测中，通过章中的模型将图像转换为向量表示，以便能够在向量空间中进行相似度搜索。使用向量搜索算法可以快速地找到与查询图像最相似的图像，从而实现图像复制检测的功能。FAISS提供了多种向量搜索算法，包括暴力搜索、倒排索引搜索、基于图的搜索、聚类树搜索等，每种算法都有自己的特点和适用范围。

由于本文中涉及到的图像数并不是特别大，加上本文中重点也不落在向量搜索的性能上，因此本文选用了基于线性扫描的索引类型，并使用余弦相似度来计算特征向量之间的相似度，在构建索引时，我们将向量集合存储在一个简单的平坦数组中。余弦相似度是一种常用的相似度计算方法，它可以用于计算文档或者物品向量之间的相似度。余弦相似度的计算公式为：

cos(\theta) = \frac{\boldsymbol{a} \cdot \boldsymbol{b}}{|\boldsymbol{a}| |\boldsymbol{b}|} = \frac{\sum\_{i=1}^{n} a\_i b\_i}{\sqrt{\sum\_{i=1}^{n} a\_i^2} \sqrt{\sum\_{i=1}^{n} b\_i^2}}

其中，$\boldsymbol{a}$和$\boldsymbol{b}$是两个向量，$n$是向量的维度。$\boldsymbol{a} \cdot \boldsymbol{b}$表示向量$\boldsymbol{a}$和向量$\boldsymbol{b}$的点积。$|\boldsymbol{a}|$和$|\boldsymbol{b}|$分别表示向量$\boldsymbol{a}$和向量$\boldsymbol{b}$的模长。在搜索时，向量索引将查询向量与存储在索引数组中的所有向量进行点积计算，并返回与查询向量最相似的向量。这种点积计算是通过计算查询向量和存储向量之间的余弦值来完成的。

总之，向量搜索是在图像复制检测系统中占据着重要地位，可以用于很多领域的相似性搜索和聚类。通过选择合适的相似度计算方法和索引结构，可以实现高效的向量搜索。

# 训练

## 数据集

## ISC2021数据集

本文中使用的数据集来自于Facebook在2021年举办的比赛 (ISC2021)。该数据集由三部分组成：

参考数据集(reference set)：1000000张图像，其中图像未做任何处理，由原始图像组成；

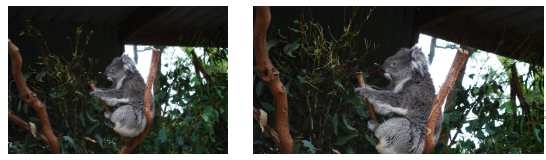
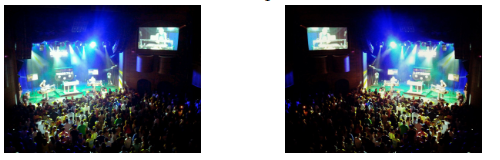
测试数据集(test set)：50000张图像，由转换后的图像组成，其中10000张图像由参考数据集中的原始图像转换而来，剩下40000张图像是干扰项。

训练数据集(training set)：以与参考集相同的方式收集的 1000000张图像。预期用途是用于各种训练任务，尤其是那些依赖于参考图像数据分布的任务，例如通过数据增强进行模型训练、分数归一化和 PCA 训练.

其中查询数据集中的图像分为两种，一种图像的源图像是参考数据集中的图像，剩下的图像则是干扰项，它们的源图像不存在于参考数据集中，没有可供检测的原始图像。该数据集来源于真实工业环境，是社交媒体在日常中所处理的图像的小规模示例，它模拟了在真实检测过程中可能会遇到的各种问题。

ISC2021的数据来源主要是两个来源：YFCC100M [63] 公开数据集和 DeepFake 检测挑战 (DFDC) 公开数据集 [13]。ISC2021从 YFCC100M 中选择图像，为了防止涉及到公众个人隐私等道德法律问题，被挑选出的图像不包含任何可识别的人物，即使有些图像中包含了行人，其对应的区域也不大于整个图像区域的 0.5%，无法被识别出来任何有效人物信息。然而，在真实的社交媒体中上传的图像包含着许多人类图像，因此，为了使数据集更真实，该数据集从Facebook举办的竞赛 DeepFake 检测挑战赛 (DFDC) 中公开的数据集中获得真实来自于付费演员的人脸图像，以此来模拟现实中真实的数据。

为了保证数据集对真实情况的良好模拟，ISC2021对图像进行了不同类型的转换，并且这些转换并不是统一难度的转换，而是分成多个难度等级进行转换，从简单的转换到转换后图像与原始图像极度不相似，甚至连人类难以评估的图像复制转换。

旋转 水平翻转+IG滤镜（调色）

### 模糊+文字+特效覆盖 饱和度+文字+高斯模糊

### 

截屏上传 局部+覆盖+文字

## 实验环境与工具

## 评价指标uAP

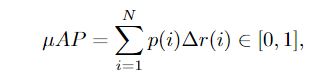
本文中主要使用三个指标对模型性能进行评测，分别是recall，precision和uAP。

Precision（精确率）：它是指所有被分类器判定为正类的样本中，实际上属于正类的样本的比例。计算公式为：$Precision = \frac{TP}{TP+FP}$，其中TP是真正例，FP是假正例。

Recall（召回率）：它是指所有实际为正类的样本中，被分类器判定为正类的样本的比例。计算公式为：$Recall = \frac{TP}{TP+FN}$，其中TP是真正例，FN是假反例。

Recall@Precision90（召回率@精确率90%）：它是指在精确率达到90%的情况下，分类器能够正确识别的正样本的比例，即召回率。该指标经常被用于要求在精确率较高的情况下仍能保持较高的召回率的模型。

Micro-Average Precision（uAP）是在深度学习中用于评估模型性能的指标之一。在本文中，每张图像都被视为单独的一类，在对test set中的图像进行检测时，对于每张检测图像，都会给reference set中与检测图像最相似，也就是置信值最高的图像作为结果对，结果对的格式为(confidence, reference\_id)，其中reference\_id表示对检测出的目标图像的id，confidence代表结果的可信度，越高越准确。

在对test set中所有图像都检测后，设定confidence在某个threshold以上的检测结果才有效，将所有结果对排序后可以获得不同threshhold下所对应的precision和recall，从而得到一个总体的Precision-Recall曲线，此时uAP与Precision-Recall曲线下的面积AP是等同的，计算公式如下：

其中 p(i) 是将所有结果对排序后，在列表中排在第 i 位的结果对在PR曲线中所对应的precision，Δr(i) 是第 i 位和第 i-1 位结果对所对应的的recall的差值，N 是所有结果对的总数。

与传统的mean Average Precision (mAP)不同，uAP的优势在于它将所有类别的结果平等地对待，因此更适合处理类别不平衡的情况。此外，uAP还能够更精确地反映模型的性能，因为它考虑了每个类别的性能，而不是简单地对所有类别进行平均。

## 实验结果分析与对比

### 训练时数据增广

本文中使用的对比学习模型在第三章已做过详细介绍，本小节将详细介绍该模型训练过程。

在训练时，对于trainin set中每一张图像，都会先做进行随机裁剪（random crop）和数据增强 （data augmentation）变换来形成正向样本，每张图像变换后会生成4张新图像，其中1张是对原图像做随机裁剪，将图像缩放到大小为指定尺寸后，再对图片进行随机水平翻转（random horizontally flip），剩下3张图像则是对原图像做随机数据增强，如增亮，像素化，模糊，增加文字等，数据增强的类别在3.1中给出。

测试时，将每张图片按原比例缩放到短边的大小为256，然后只对图片做中心裁剪 （center crop）最后得到的图像大小为227\*227。

### 实验细节和参数设置

本文模型所使用的卷积神经网络主体为ResNet50，这部分的神经网络称为backbone，并将ResNet50使用Barlow-Twins进行预训练，预训练数据集是使用广泛的公开图像数据集ImageNet[63]。在训练深度神经网络时，训练的batch大小设为128，使用1000000张图像作为样本，训练所使用的硬件配置为4张NVIDIA GeForce GTX 1080 Ti。

在ResNet50网络后增加GeM池化层，根据经验，GeM池化层的p参数被初始化为3。在池化层后面使用一个head结构，该head结构由一个全连接层（fully connected layer）和一个BN层（Batch Normalization layer）组成，用于将经过池化后提取出的图像特征映射为256维的特征向量（embedding），最后对获得的特征向量再进行L2归一化操作，使得每个特征向量的范数为1，这是提高特征向量之间相似度计算准确度的重要操作。

模型训练时采用SGD优化器[65]，与其他优化算法相比，SGD使用的内存较少，可以在较小的GPU内存上运行大型模型。同时它的可扩展性好，收敛速度也比较快，适合在本章节训练中使用。包括深度神经网络训练时的迭代次数（iter，深度神经网络训练一个batch就是一 次iter）、学习率以及衰减率，其中backbone部分的网络和head部分的网络的学习 率是分开设定的。

在训练时，给定初始学习率的值为0.1，随着训练的进行，学习率会根据给定的调度策略逐步递减，能够更加精细地调整模型，以加速模型的收敛，学习率调整方式见下：

cur\_lr = init\_lr \* (1 - (cur\_epoch / total\_epochs))

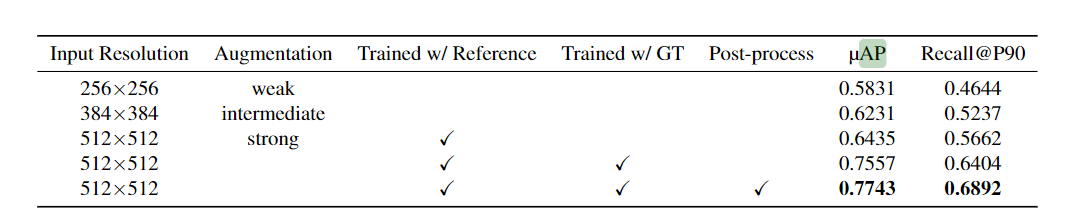
其中cur\_lr表示当前epoch要使用的学习率，init\_lr是初始学习率，cur\_epoch表示当前是第几个epoch，total\_epoch表示最终要训练的epoch次数。该计算公式表示学习率会随着训练的进行而逐渐减小，以更好的调整参数变化。

受渐进式学习（progressive learning） [18] 的启发，本文中模型训练分多个步骤进行。渐进式学习是指将训练数据按照难易程度划分成多个阶段，逐步增加模型的复杂度和训练难度，使模型能够逐步学习到更加复杂和抽象的特征表示。

在本章节中，随着训练步骤的进行，输入图像的分辨率和对应用到图像上的数据增强幅度都会逐步增加。具体来说，在前5个epoch的训练中，使用的图像输入分辨率为256\*256，并且使用较弱幅度的数据增强进行训练，在接下来10个epoch中，使用的图像输入分辨率仍为256\*256，但数据增强幅度变大，比如图像颜色变换程度更剧烈，模糊程度更高，文字或者表情包覆盖面积更大等等。随后的epoch中则会逐步将图像输入分辨率逐渐增大，同样对图像的数据增强程度也在变化。

### 实验对比与分析

上面章节中已经介绍过本文中使用的图像数据增强的具体类别，在图像复制检测中，数据的重要性毋庸置疑，可以说图像复制检测技术就是由“数据驱动”。为了检测数据增强的程度对于训练结果的影响，并且为了验证渐进式学习的作用，本小节对不同程度的数据增强造成的影响做了对比实验，见表



根据上表可以推断，数据增强是提高模型效果的一种重要手段，并且相较于较基础级别的图像变换增强，应用种类更多，幅度更大的复杂变换能够明显提高模型检测的性能。这是因为在现实应用中，社交媒体图像具有十分复杂多变的特征，提高模型的鲁棒性显得极为重要。因此在训练时为模型提供复杂的数据增强能够有效有助于降低过拟合的风险，提高模型的泛化能力，通过对训练集进行多次随机变换，模型可以学习到更多不同样式的图像，使其能够更好地处理未见过的图像。

对于应用于现实问题的图像复制检测模型来说，选择合适的损失函数对于模型训练来说非常重要，损失函数定义了我们所期望的模型输出和真实值之间的距离。如果损失函数选择不合适，可能会导致模型无法生成到正确的特征，从而无法根据样本的特征搜索到正确的目标图像。本文中的模型是基于深度度量学习构造的模型，而深度度量学习中常用的损失函数有Contrastive loss、triplet loss和ms loss等，其中contrastive loss、triplet loss都是pair\_based（基于对）的损失函数，适用于复杂图像复制检测场景中。与Contrastive loss不同，triplet loss是用三元组训练，每个三元组包含一个锚点（anchor）、一个正样本（positive）和一个负样本（negative），其中正样本和锚点相似，负样本和锚点不相似，目标是让正样本和负样本之间的距离大于正样本和锚点之间的距离。而contrastive loss是用成对的样本训练，每个样本对包含两个向量，目标是让同类样本之间的距离尽可能小，不同类样本之间的距离尽可能大。

对比实验使用了ISC2021的子集，从训练集中选取10000张图像，从测试集中选取1000张图像，作为对比实验的数据集。为了确保实验的公平性，除了损失函数不同以外，训练模型与数据集都是一致的，在训练了同样的epoch后，获得了以下数据：

从表中可知相较于triplet loss，contrastive loss更适用于本文中的训练场景，这可能是因为在大规模图像复制相似度检测这一现实应用中，contrastive loss相较于triplet loss有以下优点：

1. 计算效率更高：在大规模图像相似度检测中，数据量通常非常大，因此计算效率是非常重要的考虑因素。相对于Triplet Loss，Contrastive Loss的计算效率更高。因为Triplet Loss需要比较三个样本之间的距离，而Contrastive Loss只需要比较两个样本之间的距离。这意味着，Contrastive Loss相对于Triplet Loss需要更少的计算量和存储空间，从而更适用于大规模图像相似度检测。
2. 容易实现：Contrastive Loss的实现相对简单，只需要计算同类样本之间的距离和不同类样本之间的距离，并将其作为损失函数的输入即可。相比之下，Triplet Loss需要选取三个样本，并计算它们之间的距离。这个过程需要更多的代码实现和计算，因此Contrastive Loss更容易实现。
3. 适用于不同的任务需求：Contrastive Loss适用于学习样本之间的相似度，而Triplet Loss适用于学习同一类别图像之间的相对位置关系。在大规模图像相似度检测中，我们通常需要检索与查询图像最相似的图像，这就需要学习样本之间的相似度。因此，Contrastive Loss更适用于大规模图像相似度检测这一任务需求。
4. 易于调参：在使用Triplet Loss时，我们需要选择合适的三元组样本，这个过程需要一定的经验和技巧。而在使用Contrastive Loss时，我们只需要选择合适的阈值即可。这个阈值决定了同类样本之间的距离和不同类样本之间的距离的界限。因此，Contrastive Loss更易于调参。

## 本章小结

本章主要介绍了评测个体认知结构发展水平的方法以及评测模型在网络教学中的应用。首先，从网络教的用户学发展水平的网络教学系统整体实现框架。

# 端到端系统图像复制检测系统

本章基于学系统确实能提高学习效率，也验证了本文提出的基于个体认知结构的评测模型的可行性。

## 系统需求分析

对于设计一个图像复制检测系统，我们需要考虑以下几个方面的系统需求：

数据预处理：由于处理大量图像需要消耗大量计算资源，因此需要在处理前对数据进行预处理。这包括图像质量检查、大小归一化、格式转换等。这些预处理步骤可以加快后续对图像的处理。

特征提取：在实现图像复制检测系统时，需要对图像进行特征提取，可以选择使用深度学习模型如ResNet、VGG等提取图像特征，也可以使用传统方法如SIFT、HOG等算法提取图像特征。同时还需要对提取出来的特征进行归一化和降维处理，方便后续的检索操作。

特征存储：提取出来的特征需要进行存储，以便后续的检索。存储的方式可以选择保存为本地文件，也可以使用数据库进行存储。

向量检索：对于大规模图像数据，需要使用高效的索引结构进行存储和搜索。需要根据具体情况选择合适的索引结构进行索引建立和搜索。首先，需要将所有图像的特征向量添加到索引中。然后，对于每个查询图像，计算其特征向量与索引中所有特征向量的相似度，并返回相似度最高的图像作为检测结果。

用户界面设计：用户界面设计是图像复制检测系统的重要组成部分，需要提供用户友好的界面，方便用户进行图像上传、检索操作，以及展示检索结果。可以选择使用Web应用、桌面应用或移动应用等不同的方式进行用户界面的设计。

## 系统总体设计

在本文中的图像复制检测系统采用B/S架构，后端使用Python Flask框架搭建服务器，前端使用HTML、CSS、JavaScript等技术实现web界面。服务器上运行图像特征提取模块、相似度匹配模块、数据库存储模块等功能模块，同时支持多用户上传、查询、管理功能。

本文中设计了用于检索目标图像的原始图像的检索系统，可以用于社交媒体中对负面图像的检测，以及用于原创图片的版权保护。系统包括三个模块：目标图像上传模块，内部检索模块，结果数据显示模块。

1）用户界面展示模块：本文中的图像复制检测系统需要提供友好的用户界面，包括图像数据的输入界面、处理流程的设置界面、结果展示界面等。用户界面包括上传图像功能、查询图像功能和管理图像库特征功能。用户可以通过上传图像对图像进行相似度匹配，查询是否存在相似的图像。管理图像库特征功能包括添加、删除和查询已上传的图像库特征。

1. 内部检索模块：该模块用于处理通过web界面上传的图像，对上传的查询图像检索是否有匹配的目标图像，它主要包括下面两个子模块：

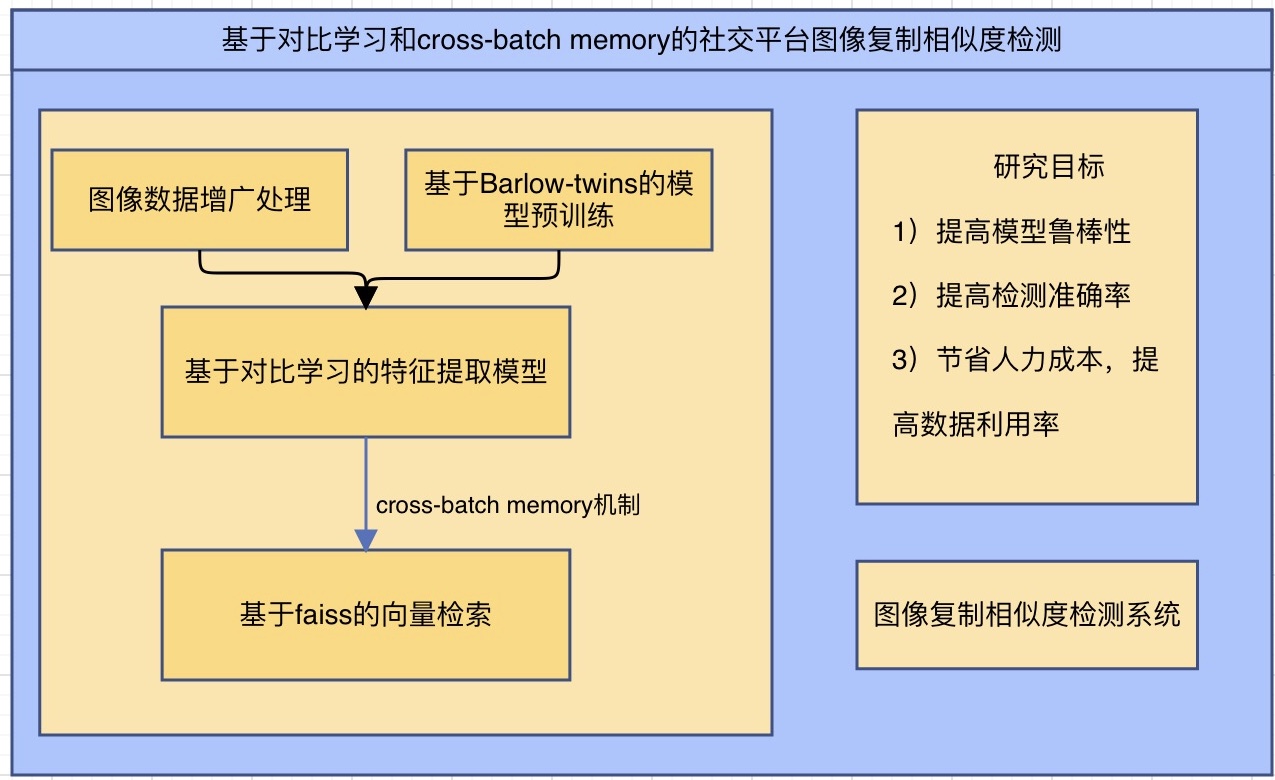
特征提取模块：本文采用基于对比学习和cross-batch memory的深度学习模型对图像进行特征提取，该模型在章节已完成训练，可以直接提取图像的特征向量。

相似度匹配模块：本文使用faiss库实现图像相似度匹配功能。我们选择使用余弦相似度作为相似度度量，将通过特征提取模块生成的参考集的特征向量导入构建相应的向量索引，对于给定的图像特征向量进行快速搜索，找到高于给定阈值且相似度最高的图像作为目标图像，认定其为查询图像的原始图像（或者副本图像）。

数据存储模块：该模块对内部检索模块提取出的图像特征向量数据进行存储，在本文设计的系统中，由于系统并发量较小，图像数据规模不是很大，因此生成的特征向量将直接存储在格式为HDF5 (Hierarchical Data Format 的文件中，这是一种用于存储和管理大量科学数据的文件格式，可以存储各种类型的数据，它采用一种基于树形结构的层次化数据模型来组织数据，每个数据集可以包含多个数据项。在本文中通过设计HDF5的数据结构来存储图像的特征向量，并选用faiss库读取HDF5文件将特征向量加载到索引内，

1. 服务接口模块：Web服务接口是连接前端和后端的重要纽带，它需要提供对外的API接口以便前端调用。这些API接口需要提供上传图片、查询图片、删除图片等基本功能，同时也需要提供一些高级查询功能，如根据特征向量查询相似图片。

综上所述，本文中的图像复制检测系统运行流程为：用户通过web界面上传图像，图像被特征提取模块处理成特征向量，存储到本地HDF5文件中。用户查询图像时，图像的特征向量通过相似度匹配模块查询相似的图像，最终将查询到的图像以及对应的相似度展示在web界面上，如果未查询到目标图像则返回提示，最终完成本次查询。

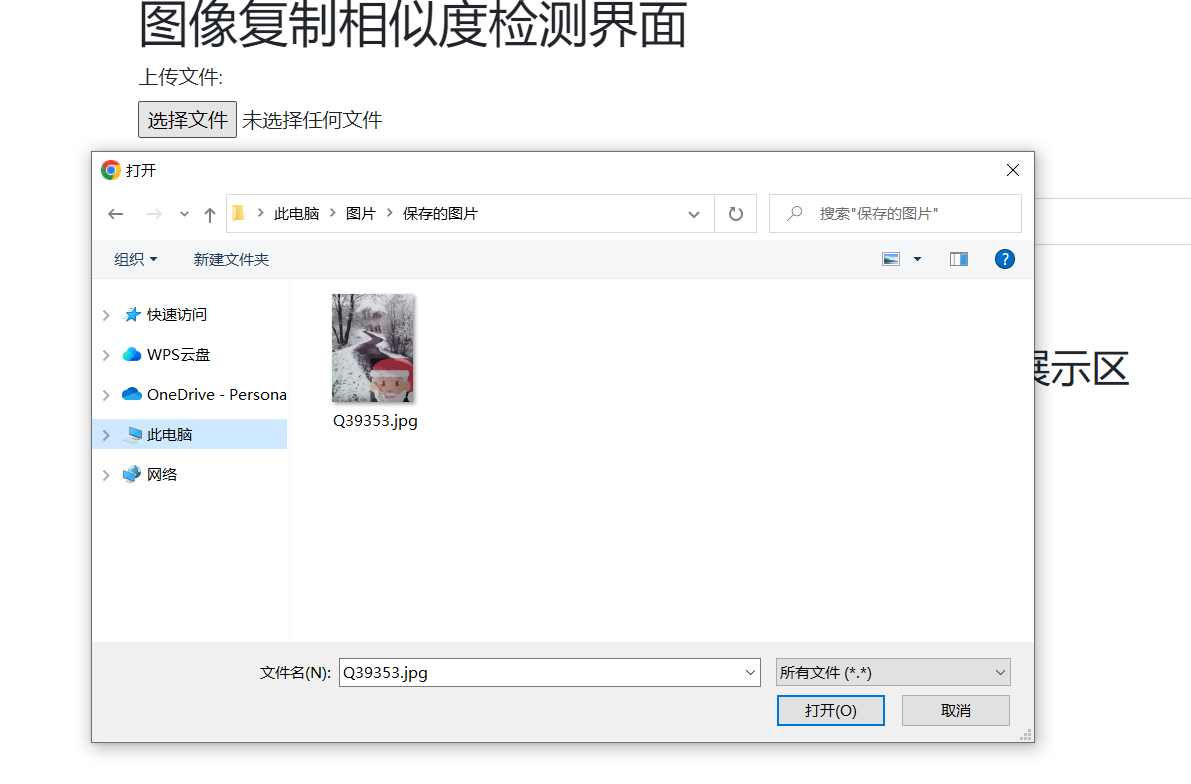
 基于以上理论研究和构思，本课题设计并实现一种端到端的社交平台图像复制相似度检测系统，该系统提供一个web页面上传待检测图像，对于该图像，系统将使用上述经过训练的特征提取模型计算出该图像的特征向量，获得的向量会经由相似性搜索在给定的数据集中进行搜索，获得的图像即为图像复制检测结果列表，并展示到图形界面中。该系统模型如下：

## 系统具体实现

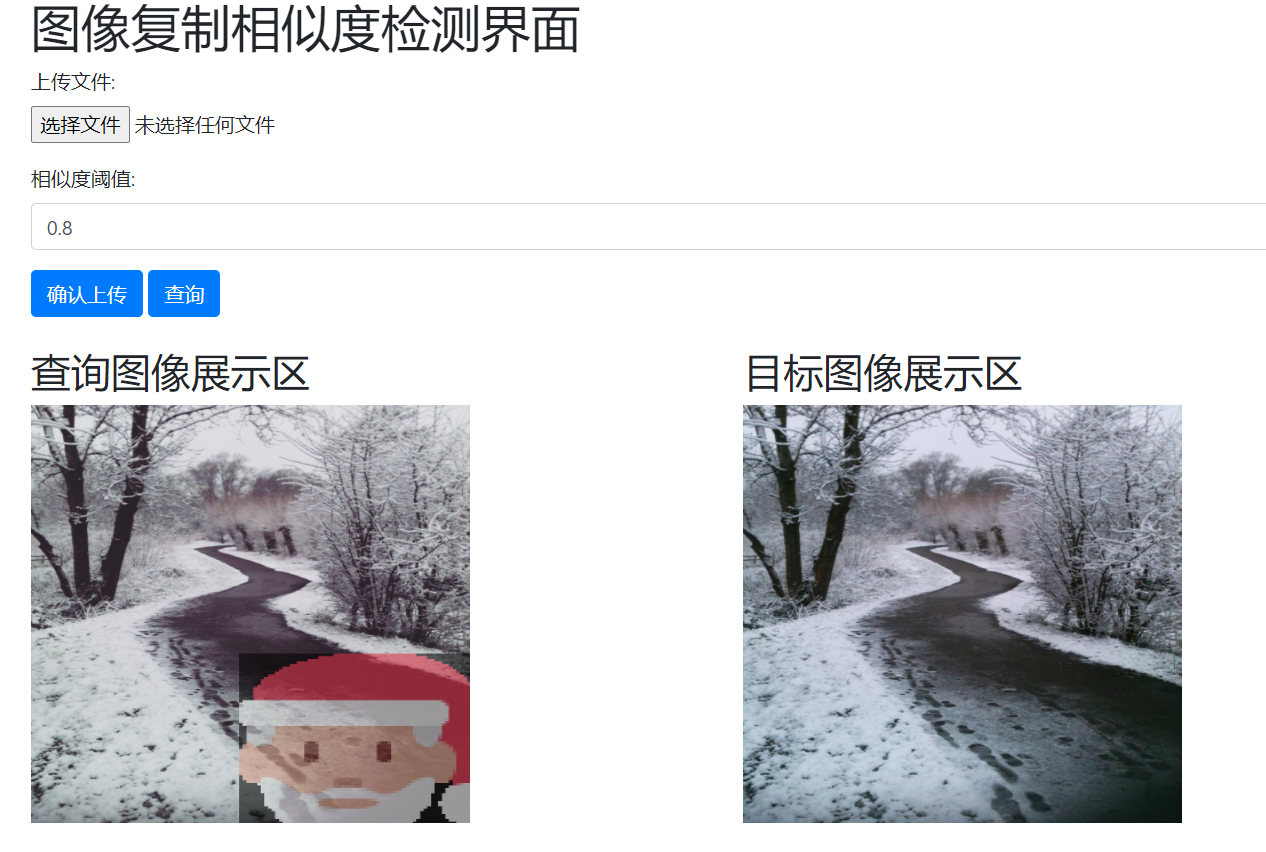
根据上面对图像复制检测系统的描述，我们设计了如下用户界面，接下来通过介绍该用户界面的组成模块来介绍系统的具体实现：

图像上传模块：该模块用来上传图像文件，也就是要查询的图像。

相似度阈值设置模块，该模块可以输入一个范围在0-1之间的数字，用来设置相似度阈值，当查询到的目标图像与查询图像之间的相似度高于这个阈值时，认为本次搜索到了有效的结构，否则本次搜索无结果。

查询选项：点击“查询”按钮后，会将上传的查询图像送入图像处理模块来获得目标图像，在web界面上传的图像会在服务器上生成特征向量，并搜索该特征向量与之前通过h5文件中导入的特征向量中哪些特征向量最为相似，然后返回相似度与最相似的图像。

图像展示模块：图像展示模块由两部分组成：查询图像展示区和目标图像展示区，这两个图像展示区横向并列，分别用于展示查询图片和目标图像。如果本次搜索没有有效结果，则会弹出“查询图像无副本”的提示。如果本次搜索有结果，目标图像就会在右侧展示出。





# 总结与展望

致 谢

时光荏苒，我在东南大学的三年硕士研究生生活即将告一段落。

最后，我要谢谢自己。不忘初心，方得始终。



**参考文献**

[1] 林克. 21世纪:教育竞争什么?[J]. 内蒙古教育, 2004(5):1.

[2] 陈自成, 宋昱霖, 邢宝峻, 等. 论“传统教育”与“互联网教育”的火花碰撞[J]. 通讯世界, 2016(4):279.

[3] YUAN L, POWELL S. MOOCs and Open Education: Implcations for Higher Education[J]. 2013.

[4] 杨秦飞, 杨孝斌, 罗永超. 皮亚杰的发生认识论原理及其教学启示[J]. 凯里学院学报, 2015,33(3):25-28.

[5] 李红. 论认知结构——系统的分析[J]. 心理学探新, 1988(4):1-6.

[6] R·赖丁, S·雷纳. 认知风格与学习策略[M]. 华东师范大学出版社, 2003.

[7] ARIAS. 2016 DLIAward Recipient – Tougaloo College[EB/OL]. https://onlinelearningconsortium.org/about/2016-dliaward-recipient-tougaloo-college/.

[8] 温善毅. 网络教学平台中数据挖掘云系统设计与开发[D]. 江西师范大学, 2015.

[9] 于杭路佳. 数据挖掘在网络教学评价中的应用[D]. 北京邮电大学, 2010.

[10] 段伟丽. 不同认知风格的汉语教师在课堂教学策略运用上的差异研究[D]. 华东师范大学, 2014.

[11] 石进芳. 网络环境下理工科学生自主学习能力培养——一项基于元认知策略训练的实证研究[J]. 中国外语, 2010,7(2):63-67.

[12] 沈晓屹. 支持隐性知识的知识建模与评测服务的研究与实现[D]. 东南大学, 2007.

[13] TRUONG H M. Integrating learning styles and adaptive e-learning system: Current developments, problems and opportunities[J]. Computers in Human Behavior, 2015,55(PB):1185-1193.

[14] LATHAM A, CROCKETT K, MCLEAN D. An adaptation algorithm for an intelligent natural language tutoring system[J]. Computers & Education, 2014,71(1):97-110.

[15] SIEMENS G, BAKER R S J D. Learning analytics and educational data mining:towards communication and collaboration: International Conference on Learning Analytics and Knowledge, 2012[C].

附 录