**江西理工大学 信息工程学院（人工智能学院） 学院 2025 届**

**本科毕业设计（论文）选题申请表**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 指导教师 | 梁苗苗 | | 学历/学位 | 博士研究生毕业 | | | 职称 | 副教授 |
| 学生姓名 | 朱金康 | | 专业 | 人工智能 | | | 班级 | 21人智2班 |
| 申请设计  （论文）题目 | | 面向大规模人脸识别模型的深度学习方法研究 | | | | | | |
| 课题性质 | | 理论（应用）研究 | | | 来源 | 其他 | | |
| 课题简介 | | 人脸识别技术是人工智能领域的一个重要分支，随着大数据技术的发展，人脸识别技术在处理大量面部特征数据方面得到了显著提升。在无约束条件的环境中，由于人脸图像在现实世界中的呈现具有高度的可变性，包括头部姿势、年龄、遮挡、光照条件和人脸表情等，所以人脸识别也是最有挑战性的生物识别方法之一。卷积神经网络（CNN）是人脸识别方面最常用的一类深度学习方法，其主要优势是通过大量数据训练学习稳健的人脸表征，而其中模型框架和损失函数对识别精度起着关键作用。因此，基于卷积神经网络的方法，如SphereFace、CosFace、ArcFace、AdaptiveFace等均从模型与损失监督出发，对人脸识别的准确率和效率不断提高。 | | | | | | |
| 课题要求（包括应具备的条件） | | 论文要求综述人脸识别任务的研究背景、意义以及研究现状，对现有经典的基于深度模型的人脸识别方法（如SphereFace、CosFace、ArcFace、AdaptiveFace等）分章节做出原理详述以及所使用损失函数对比分析，最后通过实验展示和总结各方法在人脸识别方面的性能。要求学生系统学习Python、数字图像处理、模式识别、计算机视觉等课程，为此次毕业论文中系统开发打下理论基础。此外在实践课程和第二课堂培养了Python语言编程与实践能力，为毕业论文的深入研究提供良好的基础保障。 | | | | | | |
| 学生确认签名：  2024年11月30日 | | | | | | | | |
| 指导教师确认签名：  2024年12月3日 | | | | | | | | |

**注:** 1．.本表为学生选择本科毕业设计（论文）题目时专用，由指导教师填写前面内容，学生签名后生效。

2．有关内容的填写见背面的**填表说明**，课题性质和来源在表中填写具体内容。

3．各学院可根据具体情况自拟表格。

4．本表正反面打印，学生放入毕业设计资料袋中一份。

**江西理工大学 信息工程学院（人工智能学院） 学院(学部) 2025 届**

**本 科 毕 业 设 计（论文）任 务 书**



**题 目：**面向大规模人脸识别模型的深度学习方法研究

**专题题目**（若无专题则不填）：

**原始依据**(包括设计（论文）的工作基础、研究条件、应用环境、工作目的等)：

工作基础：系统学习Python、数字图像处理、模式识别、计算机视觉等课程，为此次毕业论文中系统开发打下理论基础。此外在实践课程和第二课堂培养了Python语言编程与实践能力，为毕业论文的深入研究提供良好的基础保障。

研究条件：硬件条件——计算机；软件条件——Python

应用环境：生产开发、科学研究。

工作目的：

（1）使学生提高综合运用所学专业知识的能力；

（2）使学生掌握从事学术研究的基本方法；

（3）使学生具备基本的学术论文撰写的能力；

（4）培养和考核学生的实践动手能力，为参加工作奠定基础。

**主要内容和要求：**（包括设计（研究）内容、主要指标与技术参数，并根据课题性质对学生提出具体要求）：

论文要求综述人脸识别任务的研究背景、意义以及研究现状，对现有经典的基于深度模型的人脸识别方法（如SphereFace、CosFace、ArcFace、AdaptiveFace等）分章节做出原理详述以及所使用损失函数对比分析，最后通过实验展示和总结各方法在人脸识别方面的性能。

**日程安排：**

第1～2周：论文课题的选择和内容的理解；

第3～5周：收集课题相关文献资料，整理大规模人脸识别方面的研究现状；

第6～11周：复现部分典型方法代码，直观展示方法实验效果；

第12～15周：书写毕业论文，并对其进行修改和完善；

第16周：做好毕业论文答辩前的各项准备工作；

第17周：打印、装订整理毕业论文，毕业答辩。

**主要参考文献和书目：**

[1] Wang M, Deng W. Deep face recognition: A survey[J]. Neurocomputing, 2021, 429: 215-244.

[2] Liu F, Chen D, Wang F, et al. Deep learning based single sample face recognition: a survey[J]. Artificial Intelligence Review, 2023, 56(3): 2723-2748.

[3] Guo G, Zhang N. A survey on deep learning based face recognition[J]. Computer vision and image understanding, 2019, 189: 102805.

[4] Schroff F, Kalenichenko D, Philbin J. FaceNet: A unified embedding for face recognition and clustering[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 815-823.

[5] Deng J, Guo J, Xue N, et al. ArcFace: Additive angular margin loss for deep face recognition[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019: 4690-4699.

[6] Wu X, He R, Sun Z, et al. A light CNN for deep face representation with noisy labels[J]. IEEE Transactions on Information Forensics and Security, 2018, 13(11): 2884-2896.

[7] Cao Q, Shen L, Xie W, et al. VGGface2: A dataset for recognising faces across pose and age[C]//2018 13th IEEE international conference on automatic face & gesture recognition (FG 2018). IEEE, 2018: 67-74.

[8] Liu W, Wen Y, Yu Z, et al. SphereFace: Deep hypersphere embedding for face recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 212-220.

[9] Wang H, Wang Y, Zhou Z, et al. Cosface: Large margin cosine loss for deep face recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 5265-5274.

[10] Liu H, Zhu X, Lei Z, et al. AdaptiveFace: Adaptive margin and sampling for face recognition[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019: 11947-11956.

**指导教师**（签字）**：**

**2024年12月23日**



注：本表可自主延伸，各专业根据需要调整。

**承 诺 书**

**本人郑重承诺：**根据所学专业人才培养方案和布置下达的毕业设计（论文）任务书要求，在学校规定的任务期限内认真自主完成。提交的毕业设计（论文）查重文稿，确系本人经指导教师指导完成并同意提交的终稿。本人已知晓“毕业设计（论文）初检查重率大于等于70%、二检大于等于50%及三检大于等于30%”或认定存在剽窃、伪造、篡改、买卖、代写等学术不端行为将取消毕业答辩资格的相关规定；若本人的毕业设计（论文）经上级部门抽检结论为不合格，接受并同意学校注销并追回已发放的毕业证和学位证。如发生上述情况，本人愿意承担因本科毕业设计（论文）质量达不到要求而产生的一切后果。

所在学院：信息工程学院（人工智能学院）

所学专业：人工智能

所在班级：21人智2班

本人学号：1520203586

姓 名： （签字确认）

指导教师： （签字确认）

2024年12月23日

说明：承诺书在指导教师下达毕业设计（论文）任务书时提交。本承诺书纸质稿一式二份，学生本人留存一份，学院留存一份。各学院（学部）按班级整理汇总全院的承诺书电子稿发教务处一份存档备查。

**江西理工大学 信息工程学院（人工智能学院） 学院(学部) 2025 届**

**本 科 毕 业 设 计（论文）开 题 报 告**

**学号：1520203586 姓名：朱金康 班级：21人智2班**



**题 目：面向大规模人脸识别模型的深度学习方法研究**

**专题题目**（若无专题则不填）**：**

**本课题来源及研究现状：**

面向大规模人脸识别模型的深度学习方法研究的课题来源：

随着社会对安全问题的关注增加，如机场、车站等公共场所的安全监控需求日益增长，迫切需要高效准确的人脸识别技术来提升公共安全。智能手机和平板电脑等移动设备越来越依赖于生物特征识别技术（包括人脸识别）作为便捷且安全的解锁方式。社交媒体平台和在线服务也利用人脸识别技术为用户提供照片标记、好友推荐等功能，增强用户体验。

面向大规模人脸识别模型的深度学习方法研究的研究现状：

近年来，深度卷积神经网络已经成为人脸识别的核心技术。通过使用大量标注数据进行训练，这些网络能够自动提取出高维度的人脸特征表示，并实现非常高的识别精度。为了提高人脸识别模型的判别能力，研究人员设计了各种专门的损失函数，例如中心损失（Center Loss）、三元组损失（Triplet Loss）等，这些损失函数有助于模型学习到更加分离和紧凑的特征空间。大规模人脸数据库（如 LFW、CASIA-WebFace、MS-Celeb-1M 等）的出现为人脸识别提供了丰富的训练资源，促进了算法的发展，但是随着对精度的追求导致现在的数据集的规模和样本类别数越来越大。导致训练的过程对硬件性能的要求非常的高。特别是全连接层，因为模型占用的显存的大小是固定的，而全连接层的大小会随着数据集内的人脸身份的增加而增加，所以对于较大的数据集，这点是不可接受的显存占用。之前的研究主要的思路是尝试将整个全连接层分配到不同的 GPU 上面，但是这样会带来沉重的通信成本。

**课题研究目标、内容、方法和手段：**

课题研究目标：

L-Softmax（Large-Margin Softmax），是一种改进的 Softmax 损失函数，旨在通过增加分类边界的角度间隔（angular margin），来增强深度学习模型在人脸识别任务中的判别能力。实现类内更加紧凑，类间更加分散的人脸识别任务，并应用在面向大规模人脸识别中，和其相似思想的角边距损失函数还有 ArcFace、AMFace、CosFace、SphereFace 本课题旨在通过动态队列池来降低训练过程中的硬件资源的消耗，并且在训练的过程中结合角边距损失函数，以求尽量不影响训练结果的正确率。

课题研究内容：

由于 L-Softmax 增强了类别间的区分度，它有助于减少过拟合现象，特别是在训练集和测试集之间存在差异的情况下。这种特性使得基于 L-Softmax 训练的模型具有更好的泛化能力，可以更好地处理未见过的数据。通过增加类别间的角度间隔，L-Softmax 能够有效降低误识率，提高人脸识别的准确性和可靠性。尤其是在大规模数据集上，使用 L-Softmax 往往能取得比传统 Softmax 损失函数更好的性能。本课题打算通过大规模人脸数据集来训练出识别准确率较高，鲁棒性较好的模型，主要的思路先构造一个类中心池，也就是前文提到的动态队列池，里面每个位置存储每一类的类中心向量，所以现在问题变成了在尽量不影响正确率的前提下，怎么样能让每个 batch 训练过程中内部的类别在动态队列池中尽量都存在，这样才能够将训练过程进行下去，实现的方式就是使用一个最近最少使用算法（LRU）来管理我们的动态队列池，该方法的原理是这样的，当新的类别进来后 该类别放在队列的最前端，同时其他没被使用的类别自然向后移动，当超过队列的最大长度之后，尾部的类别被抛弃，当查询的类别不是新的类别，将该类别提到最前端即可。这样我们的全连接层的大小就可以通过自己的设置来配置。

课题研究方法和手段：

首先获取大规模人脸数据集，对该数据集进行预处理，生成读取更快的 lmdb 数据库形式，设计模型的架构，设计损失函数，设计较为合理的超参数，特别是动态队列池大小这个参数，测试不同的参数对训练正确率的影响，测试不同的角边距损失函数对正确率的影响，并在公开的人脸识别数据集上进行基准测试。为了检验模型的泛化能力，可以尝试在不同的数据分布或场景下进行验证，例如从实验室环境到现实世界的迁移。

**设计（论文）提纲及进度安排：**

2024.12-2025.1 阅读相关文献，明确研究方向

2025.1-2025.3 构建好一个网络架构，构思创新点，并且训练出简单模型进行测试

2025.3-2025.5 构建动态队列池，选择不同的超参数跑公开数据集，得到较好的模型，检测模型

2025.5-2025.6 检测模型泛化能力，完成毕业答辩

**主要参考文献和书目：**

Large-Margin Softmax Loss for Convolutional Neural Networks

S. Ren, K. He, R. Girshick, and J. Sun, “Faster r-cnn: Towards real-time object detection with region

proposal networks,” Advances in neural information processing systems,vol. 28, pp. 91–99, 2015.

W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, “Ssd: Single shot multibox

detector,” in European conference on computer vision,pp. 21–37, Springer, 2016.

**指导教师审核意见：**

对大规模人脸识别任务背景及意义较了解，研究现状的概述不够深入全面，研究内容、方法手段以及进度安排清晰合理，参考文献太少且格式不规范，同意开题。

**指导教师（**签字）**： 2025年1月18日**

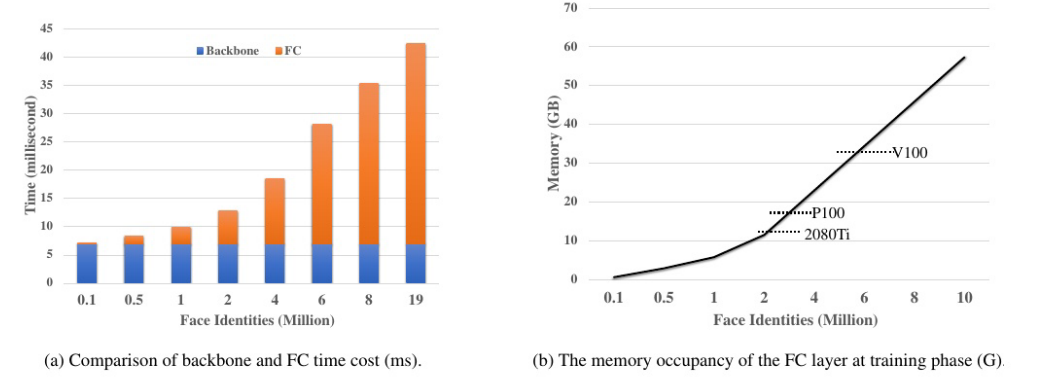


注：本表可自主延伸

一种大规模人脸识别的有效训练方法

**摘要：**由于超大规模和良好标记的数据集，人脸识别在深度学习时代取得了重大进展。然而，在超大规模的数据集上进行训练非常耗时，并且占用了大量的硬件资源。因此，设计一种高效的训练方法是必不可少的。全连接层(FC)的百万级维数导致了大量的计算和内存开销。为此，我们提出了一种新的训练方法，称为快速人脸分类(F2C)，以减轻时间和成本，而不牺牲性能。该方法采用动态类池(DynamicClass Pool, DCP)动态存储和更新身份特征，可以看作是FC层的替代品。DCP具有体积小、不依赖于整个人脸身份的特点，有效地节省了时间和成本。我们进一步在多个人脸基准和私有数据集上验证了所提出的F2C方法，并显示了可比较的结果，同时在识别精度和硬件成本方面，速度比最先进的基于fc的方法更快。此外，通过设计基于身份和基于实例的双重数据加载器，进一步改进了该方法，提高了DCP参数更新的效率。

**简介：**深度神经网络(Deep Neural Networks, DNNs)在计算机视觉任务中取得了许多令人瞩目的成果[22,23,33,34,35,36]。人脸识别可以看作是计算机视觉中最热门的研究课题之一。在过去的十年中，已经发布了许多大规模且标记良好的数据集[8,11,43,45,49]。训练过程人脸识别的目的是学习与身份相关的嵌入空间，在该嵌入空间中，类内距离减小，类间距离增大。先前的工作[8,38,39]已经证明，在大数据集上的训练可以获得比小数据集更大的改进。为此，学术界和工业界收集了包括10个甚至1亿个人脸身份在内的超大规模数据集。谷歌收集了2亿张人脸图像，由800万个身份组成[24]。清华引入WebFace260M[49]，包括2.6亿张人脸，这是目前最大的公开人脸数据集，实现了最先进的性能。总的来说，这些超大规模的数据集大大提升了人脸识别性能。但是，随着人脸身份的增长和硬件的限制，在训练阶段主要存在两个问题。第一个问题源于训练时间和硬件资源占用。如图1所示，当人脸身份达到1000万时，FC层的时间成本和GPU内存占用都远远大于骨干网。为了解决这些问题，以前的许多方法[1,16]都侧重于减少FC层的时间和资源成本。以前的方法可以归纳为两类。一个[1]试图将整个FC分配给不同的gpu，引入了沉重的通信成本。另一个[46]试图通过从FC层随机选择一定比例的神经元来降低计算成本，但它仍然需要存储整个FC参数。当标识达到1000万或1亿个时，存储整个FC参数的成本非常高。如何有效降低高维FC层带来的计算和内存成本?一个直观的想法是减少FC的大小或设计一个替代范例，这在以前几乎没有被探索过。第二个问题与FC参数的更新效率和速度有关。正如[6]所指出的，类中心的最优解实际上是这个类的所有样本的均值。具有罕见样本且采样频率非常低的身份将很少有机会通过其样本更新类中心，这可能会妨碍特征表示。



为了解决上述问题，我们提出了一种针对超大规模人脸数据集的高效训练方法，称为Faster faceClassification(F2C)。在F2C中，我们首先引入了名为Gallery Net (G-Net)和Probe Net (P-Net)的双主干，分别生成身份中心和提取人脸特征。G-Net具有与P-Net相同的结构，并以移动平均的方式继承了P-Net的参数。考虑到超大规模训练中最耗时的部分位于FC层，我们提出了动态类池(Dynamic Class Pool, DCP)来存储G-Net的特征，并计算每个小批中具有正样本(其身份出现在DCP中)的logits。DCP可以看作是FC层的替代品，它的尺寸比FC小得多，这就是为什么F2C相比FC层可以大大减少时间和资源成本的原因。对于负样本(其身份不出现在DCP)，我们最小化负样本和DCP之间的余弦相似度。为了提高DCP参数更新的效率和速度，我们设计了基于身份和基于实例的双重数据加载器。双数据加载器根据实例和身份从给定数据集中加载图像，生成用于训练的批次。最后，我们在几个人脸基准上进行了充分的实验，证明F2C可以达到与普通基于fc的方法相当的结果和更高的训练速度。F2C在识别精度和硬件成本方面也比以前的方法获得了更好的性能。我们的贡献可以总结如下。

1)我们提出了一种用于超大规模人脸识别训练的高效训练方法F2C，该方法旨在减少训练时间和硬件成本，同时保持与最先进的基于fc的方法相当的性能。

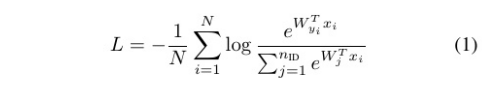
2)我们设计了DCP动态存储和更新身份特征，这是FC层的替代方案。DCP的大小比FC小得多，并且独立于整个人脸身份，因此可以大大减少训练时间和硬件成本。

3)设计了基于身份和基于实例的双重数据加载器，提高了DCP参数的更新效率。

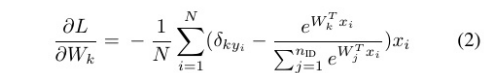
**相关工作：**人脸识别。由于大规模的数据集、先进的架构和损失函数，人脸识别已经取得了巨大的进步。大规模数据集在促进人脸识别性能[5]方面发挥着最关键的作用。这些数据集可以根据人脸身份的数量分为三个区间:1-10K, 11-100K， >100K。VGGFace[21]、VGGFace2[3]、UMD-Faces[2]、CelebFaces[28]、CASIA-WebFace[45]属于第一个区间。IMDB-Face[30]和MS1MV2[5]的人脸识别在11K ~ 100K之间。Glint360k[1]和Webface260M[49]的人脸身份约为0.36M和4M。许多先前的工作[1,46,49]表明，在更大的人脸身份数据集上进行训练可以比在较小的数据集上获得更好的性能。因此，使用WebFace260M作为训练数据集，在IJBC[19]和NIST-FRVT挑战中获得了最先进的性能。基于这些数据集，已经提出了各种用于提高性能的CNN架构，如VG-GNet[26]、GoogleNet[29]、ResNet[10]、AttentionNet[32]和MobileFaceNet[4]。对于损失函数，对比损失[28,44]和三重损失[26]可能是很好的候选者。但它们的缺点是计算成本高，收敛速度慢。为此，研究人员试图探索新的度量学习损失函数来提升人脸识别性能。已经利用了几种基于边缘的softmax损失[5,17,31,340,41]并获得了最先进的结果。综上所述，目前的方法和大规模数据集在人脸识别方面已经取得了优异的性能，但训练时间和硬件成本仍然是训练阶段的瓶颈，特别是对于百万规模甚至更多的人脸身份数据集的训练。大规模FC层的加速。如图1a所示，当人脸识别达到10M时，时间成本主要集中在FC层，而不是卷积层。自2001年以来，研究人员在加速大规模FC训练方面进行了一些尝试。一个直观的想法是设计一个近似函数来降低计算成本，层次Softmax (HSM)[7]试图将多类分类器重新表述为二元分类器的层次结构。因此，可以通过给定样本只需要沿着从根到相应类的路径遍历的方式来降低训练成本。但是，所有的类中心都存储在RAM中，随着人脸身份的增加，检索时间也不容忽视。Zhang等人。[46]提出了一种可以在每个迷你批中识别少量“活动类”的方法，该方法动态地构建动态类层次结构。然而，当人脸身份太大时，识别“活动类”也很耗时。一些公司，如谷歌和微软，试图将所有类别平均划分为多个gpu。服务器间的通信成本也不容忽视。为了解决这个问题，Partial FC[1]尝试在单个GPU服务器上训练大规模数据集，在每次迭代中随机使用10%的身份。然而，它仍然受到单个机器中gpu内存的限制。如图1b所示，Partial FC只能在人脸识别数量不是超大(<10M)的情况下工作，否则gpu仍然会耗尽内存。有几种基于成对的方法[12]，利用人脸对训练大规模数据集，而时间复杂度为O(Nk)，其中k表示成对的大小。最新的相关工作VFC[16]通过构建一些虚拟FC参数来降低计算成本，但其性能与普通FC相比要低得多。与以往的工作不同，我们的F2C可以大大降低FC训练成本，并且与普通的基于FC的方法相比，可以达到相当的性能。

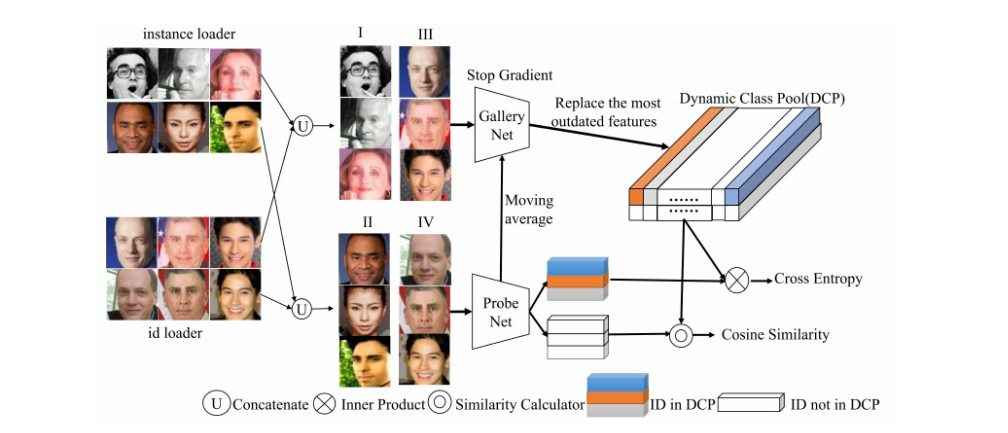
**更快的人脸分类：**在本节中，我们首先给出F2C的概述，以便简要了解我们的方法。然后介绍我们的动机和用于超大规模数据集训练的关键模块。之后，我们展示了对这些模块的理论/实证分析。最后，我们展示了训练细节，以便更好地复制。我们要解决的问题是在不造成明显性能下降的情况下，加快超大规模人脸数据集(人脸身份> 10M)的训练速度，降低硬件成本。为此，我们提出了用于超大规模人脸数据集训练的F2C框架。如图2所示，给定超大规模的人脸数据集，我们利用基于实例的加载器来生成一个实例批处理，就像数据加载器通常做的那样。同时，基于身份的加载器从同一身份中随机选择两张图像，形成配对身份批处理。随后，我们将来自实例和配对身份批次的图像混合在一起，如图2所示，并将它们输入G-Net和P-Net。受MoCo[9]的启发，G-Net具有与P-Net相同的结构，并以移动平均方式继承P-Net的参数。利用G-Net和P-Net分别生成身份中心和提取人脸特征进行人脸识别。然后，我们引入DCP作为FC层的替代品。DCP在每次迭代中随机初始化并由G-Net的特征更新。DCP的更新策略遵循以下原则:用当前的特性替换DCP中最过时的部分特性。对于正样本，我们使用常见的交叉熵损失。对于负样本，我们最小化负样本和DCP之间的余弦相似度。整个F2C同时通过交叉熵损失和余弦相似度进行优化。在深入研究F2C之前，我们通过重新思考与FC层合作的损失函数来提供一些动机。为方便起见，我们

将Softmax考虑如下:



其中N为批量大小，nID为整张脸身份的数量。对于训练过程的每次迭代，分类器{Wj}nj=1 的更新按照以下公式执行:



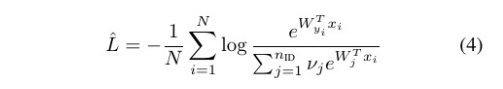


F2C的流水线。我们使用实例和id数据加载器生成混合批(I∪III, II∪IV)，然后分别输入G-Net和P-Net。G-Net的特征将以LRU的方式更新DCP, P-Net的特征将与DCP一起计算损失。

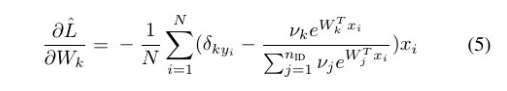
显然，所有分类器{Wj}nj=1 将在每次迭代中更新，这意味着每个分类器都有相同的优化机会。人脸识别的目标是通过同一身份的特征被拉到一起，属于不同身份的特征被推到一边的机制来区分不同身份的人。由于超大规模数据集训练的主要问题是FC层的爆炸性大小，我们可以将整个FC视为一组分类器。为了降低计算成本，在训练过程中，我们可以直观地在每次迭代中对分类器的固定比例进行优化。具体来说，我们使用如下的向量来表示给定的分类器是否在优化队列中。



其中C是一个常数，表示优化队列的长度，νi= 0/1表示分类器Wi是(不在)/(在)优化队列中。我们为这个设置画出相应的目标。



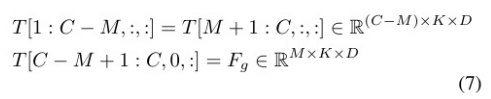
分类器根据以下等式进行更新:



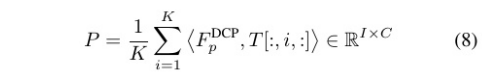
形式上，方程5与方程2类似，向量V的选择机制将直接影响分类器的更新过程。我们应该设计可行的选择机制，以便在每次迭代只更新部分分类器的约束下，更好地优化分类器。然而，这种直截了当的方法仍然承受着存储整组分类器的沉重压力。事实上，在我们的新框架中，我们只提供有限的空间来动态存储固定比例的分类器/特征。

在本小节中，我们将详细介绍我们的双数据加载器。为方便起见，我们将批大小表示为M。实际上，我们利用基于实例的加载器从给定的人脸数据集中随机采样M张图像以获得实例批。同时，基于身份的加载器通过随机选择M个身份而不从整个身份中替换来提供身份批处理，并为每个身份采样两张图像。我们将实例批分成两部分，每部分取M/2张图像。对于配对身份批，我们按人脸身份进行分割，形成具有相同人脸身份集的两部分。我们将这四部分混合得到I∪III;II∪IV(如图2所示)，其中∪表示集合的并集运算。

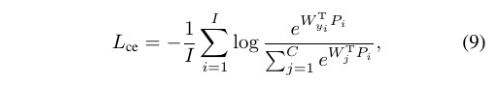
在本小节中，我们将详细介绍动态类池(DCP)。受对象检测任务中的滑动窗口[14]的启发，我们可以利用一个动态身份组，通过迭代滑动整个人脸的身份。我们把这个滑动身份群称为DCP，它可以看作是FC层的替代品。首先，我们定义一个张量T，其大小为C × K ×D，用高斯分布初始化，其中C是DCP可以容纳的人脸身份的容量或数量，K表示属于同一身份的特征的数量(我们设置默认值为K = 2)。我们在DCP中存储Fg，并在每次更新中使用Fg更新DCP中最过时的特征。更新规则类似于最近最少使用(LRU)1策略，其可以表述为:



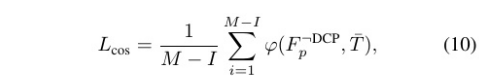
对于当前批处理，通过更新DCP，我们为DCP中的每个身份获得伪特征中心，包括包含在II∪IV中的身份。如公式6所述，与DCP相比，P-Net的特征可以分为两种类型。一个是FpDCP，另一个是-FpDCP。对于FpDCP，我们可以通过以下公式计算其对数:



其中h·，·i表示内积运算，P表示FpDCP的logits。因此，我们可以将Cross-Entropy loss表述为:

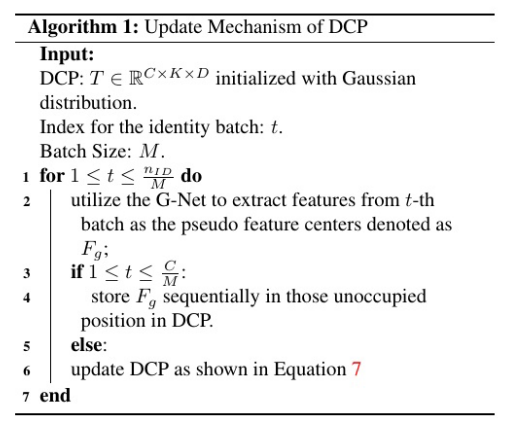


对于id不在DCP中的-FpDCP，我们添加了一个约束来最小化-FpDCP和T之间的余弦相似度，可以表示为:



总损失为Ltotal = Lce + Lcos。

DCP如公式4和9所示，我们用于DCP的交叉熵损失通常与FC的损失相似。通过在方程3中对向量V的特殊可以用方程4的形式表示Lce。为了进一步验证这一机制对DCP训练的影响，我们进行了一些实证分析。

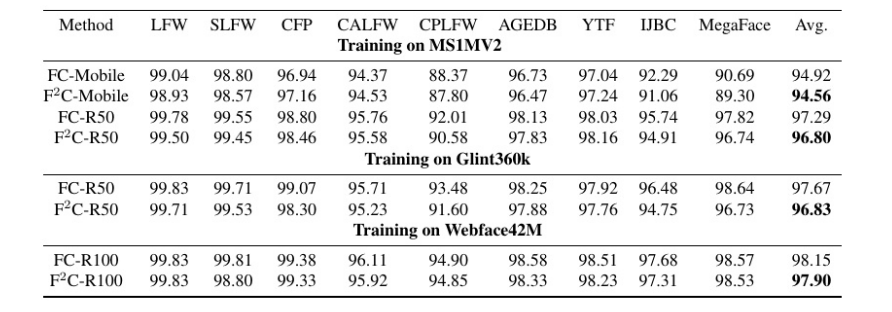


**实验：**在本节中，我们首先简要回顾人脸识别领域的几个基准数据集。然后，我们进行消融研究，以评估每个模块的有效性以及F2C中超参数的设置。最后，我们将F2C与相关的最新方法进行了比较。

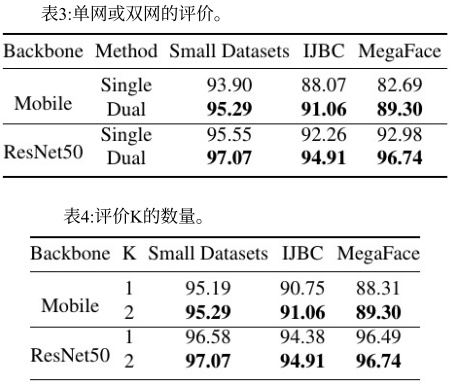
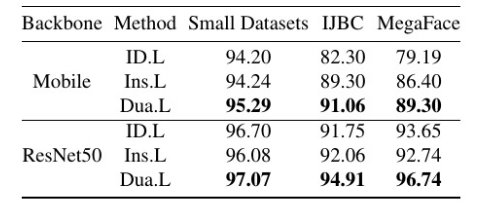
我 们 分 别 使 用 MobileFaceNet、ResNet50 和ResNet100在MS1MV2、Glint360k和Webface42M上训练F2C(Webface42M是原始Web-face260M的清洗版本，具有2M ID和约42M图像)。我们主要在以下9个学术数据集中展示了F2C的性能:LFW[11]、SLFW[11]、CFP[25]、CALFW[48]、CPLFW[47]、AGEDB[20]、YTF[42]、IJBC[19]和MegaFace[13]。LFW收集自Internet，其中包含13233张图片，5749个id。SLFW与LFW相似，但SLFW的规模小于LFW。CFP收集名人的正面和侧面照片。CALFW是LFW的跨年龄版本。CPLFW与CALFW类似，但CPLFW包含更多的姿态变化图像。AGEDB包含精确到年份的图像注释，无噪声标签。YTF包括来自YouTube的3425个视频和1595个id。IJBC由IJBB更新而来，包括21294张3531个对象的图像。megface旨在评估百万干扰物尺度下的人脸识别性能，包括一个大型画廊集和一个探针集。在这项工作中，我们使用Facescrub作为megface的探针集作为图库。

我们选择了3种不同的骨干网，并使用MS1MV2、Glint360k和Webface42M作为训练数据集，在FC和F2C之间的9个学术基准上评估了性能。如表1所示，F2C可以达到与FC相当的性能。我们还提供了这些数据集之间的平均性能，并在最后一列中进行了演示，其中F2C仅低于FC 1%。请注意，DCP的大小仅占全部面部身份的10%。我们对F2C的超参数和设置进行消融研究。在这里，

我们演示了使用MobileFaceNet和ResNet50在MS1MV2上的实验。单加载器还是双加载器?如方法学部分所述，双加载程序可以提高DCP的更新效率。为了评估加载器在F2C中的影响，我们使用了基于身份和基于实例的加载器的不同组合，并在表2中展示了结果。小数据集代表LFW、SLFW、CFP、CALFW、本小节中的CPLFW、AGEDB和YTF。我们展示了在Small Datasets上的平均精度。除非指定，否则IJBC使用TPR@FAR=1e-4度量，而Megafce默认为FPR@FAR=1e-6。使用基于实例的加载器或基于身份的加载器进行训练可以在小数据集上获得可比的结果。基于实例的加载器在IJBC和megface上的性能大大优于基于身份的加载器。可以这样解释，仅仅使用身份加载器并不能保证所有的图像都被采样。使用双数据加载器与每个单加载器相比，可以明显提高性能，这与我们的分析是一致的。请注意，为了进行公平的比较，结果是使用相同数量的样本馈送到模型，而不是使用相同数量的epoch获得的



探讨K对DCP的影响。K表示属于同一恒等式的特征的个数。我们对表4中的K = 1和K = 2求值。由于DCP中的特征代表类别中心，因此直观的感觉是K越大，中心估计越可靠。实验结果也支持了我们的直觉。然而，我们必须在性能和存储之间做出权衡。更大的K意味着更好的性能，但代价是GPU内存和服务器之间的通信。因此，我们在DCP中默认设置K = 2。



双数据加载器内的比率。我们将基于实例和基于身份的加载器之间的大小比例默认设置为1:1。为了进一步探索双数据加载器中比率的影响，我们在表5中展示了实验。我们利用ResNet50作为主干来训练MS1MV2数据集。我们发现，双数据加载器内的默认比率在大多数数据集上达到了最高的结果，特别是在挑战IJBC和megface上面。

**结论：**本文提出了一种用于超大规模人脸识别训练的高效训练方法F2C，其主要创新点是动态类池(Dynamic ClassPool, DCP)用于存储和更新人脸身份特征，以替代FC和双加载器来帮助DCP高效更新。综合实验和分析的结果表明，我们的方法可以减少硬件成本和训练时间，并获得与最先进的基于fc的方法相当的性能。更广泛的影响。所提出的方法在人脸训练数据集上进行了验证，由于种类繁多，该方案可以扩展到其他数据集和情况。然而，它不包含任何作者进行的涉及影响伦理或人权的任何研究。

**江西理工大学 信息工程学院（人工智能学院） 学院(学部) 2025 届**

**本科毕业设计(论文)中期进展情况表**



|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 姓名 | 朱金康 | 学号 | 1520203586 | 班级 | | 21人智2班 |
| 填写时间 | 2025年4月12日 | 设计开始时间（周） | 第1周 | | 已进行时间（周） | 第8周 |
| 进 展  情 况 | 1、完成了训练网络的搭建；  2、完成了训练数据集的加载和生成对应的 lmdb 数据库方便加载数据；  3、简单了训练了几轮，对网络和动态队列池存在的问题进行了调整；  4、收集了很多模型方便对不同参数的训练情况进行对比。 | | | | | |
| 存 在  问 题 | 1、正确率太低；  2、动态队列池的样本更新存在问题；  3、负样本的处理不太合理。 | | | | | |
| 改 进  措 施 | 1、调整网络结构和层数；  2、更改了动态队列池的类中心向量的存储方式；  3、看了前沿论文对负样本的处理进行了一定的更改。 | | | | | |
| 指导教师  意 见 | 同意  签 名： 2025年4月16日 | | | | | |

本科毕业设计(论文)指导记录表

|  |  |  |  |
| --- | --- | --- | --- |
| 学院 | 信息工程学院（人工智能学院） | 专业 | 人工智能 |
| 学生姓名 | 朱金康 | 学号 | 1520203586 |
| 指导教师 | 梁苗苗 | | |
| 课题题目 | 面向大规模人脸识别模型的深度学习方法研究 | | |
| 指导记录1  主题：  论文大纲与开题指导  指导时间：  2024年12月1日  主要工作：  按照任务书的要求和文献参考，确定论文的大纲以及论文的具体内容，指导填写开题报告。  指导教师签名：  2025年4月30日 | | | |
| 指导记录2  主题：  开题报告的修正  指导时间：  2024年12月4日  主要工作：  根据导师的要求，进行开题报告的规范化的修改。  指导教师签名：  2025年4月30日 | | | |
| 指导记录3  主题：  开题报告的修正  指导时间：  2024年12月4日  主要工作：  根据导师的要求，进行开题报告的规范化的修改。  指导教师签名：  2025年4月30日 | | | |
| 指导记录4  主题：  外文译文的整理  指导时间：  2025年1月6日  主要工作：  根据导师的指导选取了合适的论文并进行了翻译整理上面的知识点，整理了很多的相关的外文文献。  指导教师签名：  2025年4月30日 | | | |
| 指导记录5  主题：  进度报告与论文的书写规范  指导时间：  2025年3月10日  主要工作：  汇报当前的进度，配置好了实验要用到的环境，搭建好了实验用到的网络，正在调整代码里面存在的 bug，根据老师的指导明确了后续的进展方向。  指导教师签名：  2025年4月30日 | | | |
| 指导记录6  主题：  论文初稿撰写  指导时间：  2025年3月20日  主要工作：  完成了论文的初稿部分的研究背景，现状，研究意义部分的撰写。  指导教师签名：  2025年4月30日 | | | |
| 指导记录7  主题：  动态队列池的完善  指导时间：  2025年3月24日  主要工作：  解决了代码里面出现的 bug，实现了人脸识别需要的队列池，动态的更新里面的人脸数据。  指导教师签名：  2025年4月30日 | | | |
| 指导记录8  主题：  参数调整和结果分析  指导时间：  2025年4月7日  主要工作：  对实验里面遇到的超参数进行了参数的调整，观察合适的参数并分析了原因，对实验的结果的好坏采用了常见的量来量化，以求得到当前模型的好坏程度.  指导教师签名：  2025年4月30日 | | | |
| 指导记录9  主题：  FFC 新方法的构建与训练  指导时间：  2025年4月14日  主要工作：  构建了 FFC 的新方法进行人脸识别，构建相关的网络和用到的模块。  指导教师签名：  2025年6月5日 | | | |
| 指导记录10  主题：  FFC 方法对大量数据集进行训练  指导时间：  2025年4月21日  主要工作：  使用新的方法进行训练，采用 LRU 来构建动态的人脸数据库实现大范围的人脸训练，并且减少对资源的消耗。  指导教师签名：  2025年6月5日 | | | |

**江西理工大学 信息工程学院（人工智能学院） 学院(学部) 2025 届**

**本科毕业设计（论文）成绩指导教师评审表**

**学号：1520203586 姓名：朱金康 班级：21人智2班**

**指导教师评阅意见及评分**

|  |  |  |  |
| --- | --- | --- | --- |
| **评价内容** | **评分标准** | | **得 分** |
| **调查论证与开题报告**  **（20分）** | 1.选题得当，能概括整篇内容(18-20分)；  2.选题得当，能较好概括整篇内容(14-17分)；  3.选题较好，基本能反映整篇内容(10-13分)；  4.选题不够鲜明，不能较好反映整篇内容(0-9分)。 | | 19 |
| **论文（设计）格式**  **（20分）** | 1.格式规范，注释清晰、引用得当(18-20分)；  2.格式较规范，注释清晰、引用得当(14-17分)；  3.格式较规范，注释清晰、引用一般(10-13分)；  4.格式不够规范(0-9分)。 | | 16 |
| **论文（设计）撰写质量**  **（20分）** | 1.结论正确，能全面总结、归纳研究成果(18-20分)；  2.结论正确，较全面总结、归纳研究成果(14-17分)；  3.结论正确，能基本反映研究成果(10-13分)；  4.结论基本正确，研究结果论述不够全面(0-9分)。 | | 16 |
| **创新能力**  **（20分）** | 1.能按期完成且有创新性(18-20分)；  2.能按期完成且有一定的创新性(14-17分)；  3.基本按期完成，但缺乏创新性(10-13分)；  4.基本按期完成，但完成情况不好(0-9分)。 | | 13 |
| **图纸（小论文）的质量**  **（20分）** | 1.在省部级及以上刊物上公开发表(18-20分)；  2.论文质量较高，能较好反映专业学术水平(14-17分)；  3.论文质量一般，能基本反映专业学术水平(10-13分)；  4.论文质量较差，不能较好反应专业学术水平(0-9分)。 | | 15 |
| **指导教师评阅意见：**  选题“面向大规模人脸识别模型的深度学习方法研究”符合专业人才培养目标与基本教学内容，学生受到人脸识别理论与建模的全面综合实验实践训练，培养了学生独立解决实际问题的工作能力。论文格式有待完善。同意申报答辩。 | | | 总分 |
| 79 |
| 是否同意该学生参加答辩（累计缺席及耽误时间超过教学计划规定时间的2∕3以上，不得参加答辩）。 | | 是☑ | 否□ |

**指导教师**（签字）：

**2025年5月18日**

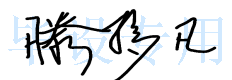
**江西理工大学 信息工程学院（人工智能学院） 学院(学部) 2025 届**

**本科毕业设计（论文）成绩评阅人评审表**

**学号：1520203586 姓名：朱金康 班级：21人智2班**

**评阅人评阅意见及评分**

|  |  |  |  |
| --- | --- | --- | --- |
| **评价内容** | **评分标准** | | **得 分** |
| **调查论证与开题报告**  **（20分）** | 1.选题得当，能概括整篇内容(18-20分)；  2.选题得当，能较好概括整篇内容(14-17分)；  3.选题较好，基本能反映整篇内容(10-13分)；  4.选题不够鲜明，不能较好反映整篇内容(0-9分)。 | | 16 |
| **论文（设计）格式**  **（20分）** | 1.格式规范，注释清晰、引用得当(18-20分)；  2.格式较规范，注释清晰、引用得当(14-17分)；  3.格式较规范，注释清晰、引用一般(10-13分)；  4.格式不够规范(0-9分)。 | | 14 |
| **论文（设计）撰写质量**  **（20分）** | 1.结论正确，能全面总结、归纳研究成果(18-20分)；  2.结论正确，较全面总结、归纳研究成果(14-17分)；  3.结论正确，能基本反映研究成果(10-13分)；  4.结论基本正确，研究结果论述不够全面(0-9分)。 | | 16 |
| **创新能力**  **（20分）** | 1.能按期完成且有创新性(18-20分)；  2.能按期完成且有一定的创新性(14-17分)；  3.基本按期完成，但缺乏创新性(10-13分)；  4.基本按期完成，但完成情况不好(0-9分)。 | | 16 |
| **图纸（小论文）的质量**  **（20分）** | 1.在省部级及以上刊物上公开发表(18-20分)；  2.论文质量较高，能较好反映专业学术水平(14-17分)；  3.论文质量一般，能基本反映专业学术水平(10-13分)；  4.论文质量较差，不能较好反应专业学术水平(0-9分)。 | | 16 |
| **评阅人意见：**  论文呢中体上以及达到了本科论文的水平，格式上需要继续修改，文中有些图的字体过小，文中设及的数字，英文都应该统一为新罗马字体。 | | | 总分 |
| 78 |
| 是否同意该学生参加答辩。 | | 是☑ | 否□ |

**评阅人（**签字）：

**2025年5月20日**

**江西理工大学信息工程学院(学部)2025届**

**本 科 毕 业 设 计（ 论 文 ）答 辩 记 录 及 成 绩 表**

**学号： 1520203586 姓名： 朱金康 班级：21人智2班**



|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **题目** | 面向大规模人脸识别模型的深度学习方法研究 | | | | | |
| 答  辩  情  况  记  录 | **问题1：队列池大小如何确定，其对性能有什么影响？**  回答：队列大小通常设置为总类别数的 10%，过大的话效果不好，违反了节省时间和资源的初衷，太小的话，数据的多样性不足，样本不均衡，模型会偏向多数类，会导致过拟合，泛化能力严重下降。  **问题2：测试是如何进行的？**  回答：选用的 LFW 数据集，每个测试样本分为 图片A、图片B、还是是否为同一个人，通过模型生成特征向量，对比特征向量的余弦相似度判断是否为同一人，得到正确率。  **问题3：FC 和 FFC 的区别是什么？**  回答：FC是经典的全连接层，FFC是在其基础上改进的动态全连接层，前者输出类别是整个样本的种类数，后者输出类别的动态池的大小，故可以动态的调节动态池的大小来调节全连接层占用显存的大小对训练过程进行控制。 | | | | | |
| **答**  **辩**  **成**  **绩** | 答辩评分：83.3 答辩结果：通过  答辩小组成员（签字）：  答辩小组组长（签字）： | | | | | |
| **综合**  **成绩** | 指导老师成绩  (30％) | 23.7 | 评阅老师成绩（30%） | 23.4 | 答辩成绩  (40％) | 33.3 |
| **总成绩** | 80.4 | **教研室(系)主任（签字）**： | | | | |

**注：**1**.** 评阅成绩为指导教师和评阅人平均分；

2**.** 答辩结果为“通过”或“不通过”，其中60分及以上为通过，否则为不通过；

3**.** 本表可自动延伸。