**开题检查记录表**

|  |  |  |  |
| --- | --- | --- | --- |
| 学生姓名 | 黄思雨 | 学 号 | 20201105047 |
| 专 业 | 计算机科学与技术 | 开题日期 | 2023年10月25日 |
| 指导教师 | 殷雁君 | 职 称 | 教授 |
| 设计（论文）题目 | 基于深度学习的蒙古文手写字元识别系统 | | |
| 指导教师评语：  该生开题报告选题具有一定的实用价值，符合学生专业发展方向，有助于提高学生的基本知识和技能以及有助于提高学生研究能力。选题采用的研究方法难度适中、制定研究计划基本合理，学生能够在预定时间内完成课题设计。同意开题。  签字： | | | |
| 检查记录（答辩记录）： | | | |
| 是否通过开题： □通过   □不通过 | | | |
| 答辩组长签字： | | 答辩组成员： | |
| 答辩秘书签字： | |  | |



**内蒙古师范大学计算机科学技术学院**

**毕业设计（论文）开题报告**

**题 目： 基于深度学习的蒙古文手写字元识别系统**

**专 业 计算机科学与技术 (师范类)**

**姓 名 黄思雨**

**学 号 20201105047**

**指导教师 殷雁君**

**日 期 2023.10.25**

**计算机科学技术学院制**

**题目：基于深度学习的蒙古文手写字元识别系统**

1．课题来源及研究的目的和意义

1.1研究背景

手写体识别是指利用计算机技术和人工智能方法对人手写的文字进行自动识别和理解。手写体文字作为人类交流的重要方式之一，具有独特的艺术性和个性化特征。手写体识别研究对文字处理领域有着重要的推动作用，手写体识别技术能够将人的手写输入转化为机器可识别的文本信息。手写体识别技术在教育应用中也具有广泛的应用前景。

蒙古文手写体字元识别难度较高，因为蒙古文手写体字元中存在不工整、蒙古文手写体字元外观类似、字迹不清晰、有涂抹痕迹等现象。同时，目前蒙古文手写字元数据库较少，特别笔迹难以采集，使得相应的蒙古文手写字元的数据库构建不足，因此蒙古文手写体字元识别是一项具有挑战性的问题。

1.2研究目的

蒙古文手写字元识别系统的研究目的，是要解决蒙古文手写体字元的识别问题，为蒙古文手写体数字化的研究打下坚实的基础。传统上，人们一直在用机器识别蒙古文手写体字符。随着电子计算机技术的迅速发展，在人们学习使用汉字系统和计算机的过程中，如何有效地利用手写字元来解决蒙古文手写体识别问题，已成为一个亟待解决的问题。从理论上讲，蒙古文手写体字元识别问题可以看作是一个最优问题。蒙古文手写体字元是一种特殊的文字符号，它以字为基础，把不同形式、不同结构的文字符号进行组合而形成。在计算机系统中，由于汉字字库没有对应于各种蒙古文手写体字元的字库，因此没有现成的蒙古文手写体字元识别系统供开发使用。

1.3研究意义

目前，蒙古文手写字元识别系统的研究已经取得了很大的进展，但是由于蒙古文手写字元识别系统具有特殊性，因而，该系统的研究在理论和技术上还存在许多不足之处，有待进一步的深入研究。随着蒙古文手写字元识别系统功能的不断完善以及识别率的不断提高，相信其将会在各个领域得到更广泛的应用。

蒙古文手写字元识别系统的研究将会为文字信息处理技术提供一个新的应用领域。在文字识别领域中，该研究具有重要意义。随着科学技术的不断进步以及计算机技术水平的不断提高，相信在不远的将来，蒙古文手写字元识别系统将会取得更大进步。

1. 国内外在该方向的研究现状及分析

2.1国内研究现状

李敏在《基于深度学习的联机蒙古文手写识别系统研究》(2019)中分析到当今主流蒙古文输入法的输入方式是基于键盘输入,在产品研发方面已趋于成熟,但是面向手写识别输入方式的蒙古文输入法研究目前仍处于初级阶段。

范道尔吉在《蒙古文脱机手写识别研究》（2020）中分析到手写识别一直是模式识别的一个重要研究领域,得到了学术界的广泛研究和关注。蒙古文脱机手写识别起步较晚、相关研究较少,且蒙古文具有词汇量巨大、书写自由、字符变形严重等特点,这些都给蒙古文脱机手写识别带来了巨大挑战。因此,本文将传统蒙古文作为对象,开展脱机手写识别研究。

石佳钰在《融合边缘注意力的手写蒙古文字元数据增强方法》（2023）中分析研究针对手写蒙古文字元数据集样本少且多样性差的问题，提出融合边缘注意力的条件手写蒙古文字元生成模型。模型在条件生成对抗网络的基础上引入了边缘注意力机制，使得数据生成模块对手写蒙古文字元边缘变化更加敏感，增加特征多样正则项在一定程度损失避免模式崩溃并使得生成样本更具多样性。

2.2国外研究现状

Cui, ShaoDon等人在《An end-to-end network for irregular printed Mongolian recognition》（2021）中分析研究蒙古语是中国内蒙古地区通用的语言。在识别过程中,由于拍摄角度等原因,图像和文字会发生形变,会对识别造成一定的困难。本文提出了一种用于印刷体蒙古文文本识别的三元组注意力迁移网络( TAMN )，该网络使用空间变换网络对变形的蒙古文图像进行校正。实验结果表明,空间变换网络能够有效地识别变形蒙古文图像,识别准确率可达90.30 %。

Sharma Abhinav等人在《Gujarati Script Recognition》（2023）中分析研究了字符识别是将印刷或手写文本从图像中提取到机器可读的格式。提取的文本可以很容易地编辑、修改和有效地存储。本文针对文本识别任务分析了两种不同的模型：基于CNN的EfficientNet B3和YOLO v4。该系统是使用EfficientNet B3模型开发的,具有更好的准确性和效率。系统的输入是具有光学文本的图像,系统产生一个可编辑的文本文档,其中包含图像中识别文本的内容。该系统已经成功实现了从报纸文章的图像创建数字图书馆的任务。

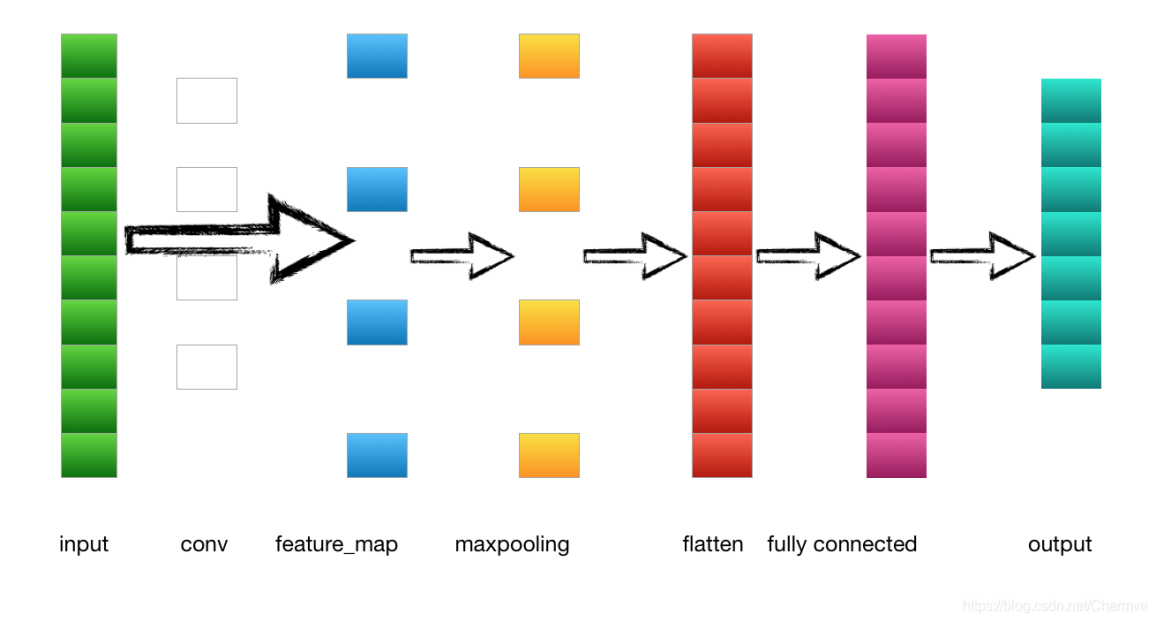
Jija Das Gupta等人在《Ensemble classifier-based off-line handwritten word recognition system in holistic approach》中分析研究了一种新的基于集成分类器的脱机手写体字识别系统。在这里，每个手写词都使用两个手工特征进行识别，即( i )基于Arnold变换的局部方向特征(取决于草书词的笔画方向分布)和( ii )基于曲率的方向特征(基本上是曲波索引直方图)，以及使用深度卷积神经网络( DCNN )的一个机器生成特征。这些特征由三个分类器分别识别。最后，结合三个分类器的决策来预测最终的词类水平。

1. 主要研究内容

3.1相关知识学习

3.1.1卷积神经网络

卷积神经网络是一类典型的深度神经网络，该网络最早于1984年由日本学者K.Fukushima发现并验证，它具有局部感知、权值共享等特性，该网络在 计算机视觉研究领域发挥了重要作用。

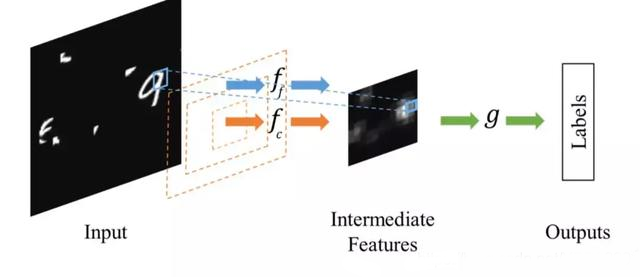


3.1.2注意力机制

研究发现，人类视觉系统的接收能力是有限的，但它同时具有强大的视觉信息处理能力。在视觉处理的过程中，眼睛接收数据后，视觉系统将注意力迅速聚焦于局部重要的区域，这种筛除机制可以很大程度上降低视觉系统需要处理的数据量，使我们在处理复杂视觉信息时可以过滤不重要的信息，把资源更合理 的分配到局部的重要区域，为后续的分析和推理提供了更精确的信息。

3.1.3空间注意力

不是图像中所有的区域对任务的贡献都是同样重要的，只有任务相关的区域才是需要关心的。空间注意力模型就是寻找网络中最重要的部位进行处理。



3.2数据准备

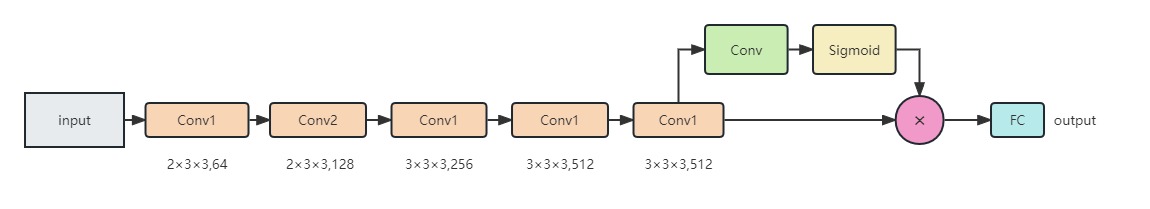
首先，需要准备用于训练和测试的蒙古文字元图像数据集。可以自己收集并标注的蒙古文手写字元数据集。确保数据集中的文字图像具有标签信息，即每个图像对应一个类别标签。

3.3数据预处理

对蒙古字元图像数据进行预处理是很重要的一步。预处理操作包括将图像转换为灰度图像、调整图像尺寸、归一化像素值。确保所有图像在输入模型之前具有相同的尺寸和像素值范围。

3.4构建VGG16模型

VGG16由多个卷积层和全连接层组成。可以使用深度学习框架，根据VGG16的网络结构与空间注意力相结合构建模型。首先添加卷积层，然后添加池化层，最后添加全连接层。确保根据数据集的类别数量设置最后一层的输出节点数。



3.5模型训练

在准备好的数据集上对构建的VGG16模型进行训练和调优。通过定义适当的损失函数和优化器，迭代地调整模型的权重，以使其逐渐适应训练数据。通过尝试不同的超参数设置，如学习率、批量大小等，以优化模型的性能。

3.6模型评估

在训练完成后，使用测试集对模型进行评估。将测试集中的蒙古文字元图像输入训练好的VGG16模型，得到预测结果。通过计算预测结果与真实标签之间的准确率，评估模型在文字识别任务上的性能。

3.7模型应用

在模型训练和评估完成后，可以将VGG16模型应用于实际的蒙古文手写字元识别任务。将待识别的蒙古文手写字元图像输入模型，通过计算模型的输出，得到对应的类别标签。

以上列出的内容代表了开发基于深度学习的蒙古文手写字元识别系统所需的主要步骤。这个系统的研究将结合卷积神经网络、空间注意力等多个领域的技术，以实现文字识别的高准确度和实用性。

1. 研究方案

对于事物来说特征的重要性是不同的，反映在卷积网络中即每张特征图的重要性是具有差异性的。注意力机制的核心思想是通过一定手段获取到每张特征图重要性的差异，将神经网络的计算资源更多地投入更重要的任务当中，并利用任务结果反向指导特征图的权重更新，从而高效快速地完成识别任务。

本项目因注意力机制其优秀的效果与即插即用的便利性，与卷积神经网络相结合，共同搭建完善VGG16模型，来完成蒙古文手写字元的识别任务。

首先自己收集并标注的蒙古文手写字元数据集，包括各种手写字元、书写痕迹和书写风格。标记信息，确保每个图像对应一个类别标签。并使用PyTorch的torchvision库加载和预处理数据集。将图像转换为灰度图像、调整图像尺寸、归一化像素值，确保所有图像在输入模型之前具有相同的尺寸和像素值范围。使用VGG16作为基础模型，根据VGG16的网络结构构建模型，首先添加卷积层，然后添加池化层，最后添加全连接层，确保根据数据集的类别数量设置最后一层的输出节点数。

使用训练集对模型进行训练，同时在验证集上进行性能监控。通过尝试不同的超参数设置，如学习率、批量大小等，以优化模型的性能，以防止过拟合。然后使用测试集来评估模型的性能，通过计算预测结果与真实标签之间的准确率及精确度，评估模型在文字识别任务上的性能。根据结果，对模型的调优和改进。在模型训练和评估完成后，可以将VGG16模型应用于实际的蒙古文手写字元识别任务。将待识别的蒙古文手写字元图像输入模型，通过计算模型的输出，得到对应的类别标签。

1. 进度安排，预期达到的目标

|  |  |
| --- | --- |
| 时间安排 | 进度安排 |
| 2023.9.21---2023.10.9 | 确定论文选题 |
| 2023.10.9---2023.10.24 | 查阅参考文献，撰写开题报告 |
| 2023.10.24---2023.10.28 | 整理完善开题报告，确定论文大纲 |
| 2023.10.28---2023.10.29 | 开题答辩 |
| 2023.10.29---2023.12.10 | 查阅论文相关资料及写作资料，学习相关理论知识 |
| 2023.12.10---2024.1.10 | 撰写论文初稿，搭建VGG16模型 |
| 2024.1.10---2024.2.1 | 与指导老师沟通并修改论文相关事宜；细化模型中的功能 |
| 2024.2.1---2024.2.15 | 对论文中存在问题进行修改；实现系统中的功能测试 |
| 2024.2.15---2024.3.20 | 撰写并完成毕设论文 |

1. 课题已具备和所需的条件、经费

外部条件：网络在线资源、图书馆资料

硬件条件：笔记本电脑

软件条件：PyCharm、Window 10

1. 研究过程中可能遇到的困难和问题，解决的措施

（1）语言特点和文化差异：

困难： 蒙语文具有独特的字符集和语法结构，可能需要特殊处理。

解决方案：字符编码：将蒙语文字符正确地编码成神经网络可处理的格式，例如使用Unicode或UTF-8编码。

（2）数据问题：

困难： 蒙语文手写字元的数据可能有限，并且质量可能参差不齐。

解决方案：尝试联系相关同学及老师，寻求合作或数据分享。

数据增强技术：使用图像旋转、翻转、缩放等技术增加数据多样性。

1. 主要参考文献

[1] 李敏.基于深度学习的联机蒙古文手写识别系统研究[J].内蒙古大学内蒙古自治区211工程院校,2019.9.

[2] 范道尔吉.蒙古文脱机手写识别研究[J].内蒙古大学内蒙古自治区211工程院校，2020.12.16.

[3]石佳钰，殷雁君，张文轩，智敏.融合边缘注意力的手写蒙古文字元数据增强方法.内蒙古师范大学学报(自然科学汉文版). 2023,52(02).

[4]Cui, ShaoDong;Su, YiLa;Qing dao er ji, Ren;Ji, YaTu.An end-to-end network for irregular printed Mongolian recognition.Journal | [J] International Journal on Document Analysis and Recognition (IJDAR). Volume 25 , Issue 1 . 2021.

[5]Sharma Abhinav;Soneji Dhiren;Ranade Aabha;Serai Dhwani;RL Priya;Lifna CS;Dugad Shashikant R.Gujarati Script Recognition.Journal | [J] Procedia Computer Science. Volume 218 , Issue . 2023. PP 2287-2298.

[6]Jija Das Gupta;Soumitra Samanta;Bhabatosh Chanda.Ensemble classifier-based off-line handwritten word recognition system in holistic approach.Journal | [J] IET Image Processing. Volume 12 , Issue 8 . 2018. PP 1467-1474.

[7] 张驰 , 郭媛 , 黎明 . 人工神经网络模型发展及应用综述 [J]. 计算机工程与应用 ,2021,57(11).

[8] 张宸嘉 , 朱磊 . 俞璐 . 卷积神经网络中的注意力机制综 述 [J]. 计算机工程与应用 ,2021,57(20).

[9]武子毅 , 刘亮亮 , 张再跃 . 基于集成注意力层卷积神经网络的汉字识别 [J]. 计算机技术与发展 ,2018,8.

[10]章琳,袁非牛,张文睿,等. 全卷积神经网络研究综述[J]. 计算机工程与应用,2020,56(01):25 - 37.

[11]徐奇.基于改进卷积神经网络的手写体汉字识别[J]. 电子技术与软件工程. 2022(09)

[12] 丁蒙 , 戴曙光 , 于恒 . 卷积神经网络在手写字符识别中的应用 [J]. 软件导刊 ,2020,1.

[13] 任欢，王旭光. 注意力机制综述[J]. 计算机应用，2021，41（21）：1-6.

[14] 朱张莉，饶元，吴渊，等.注意力机制在深度学习中的研究进展[J].中文信息学报，2019，33（6）：1-11.