



华南理工大学
South China University of Technology

本科毕业设计（论文）

基于在线签名的动静态特征提取方法实现

学 院	软件学院
专 业	软件工程
学生姓名	宋雨杭
学生学号	201736721010
指导教师	吴秋霞
提交日期	2021 年 5 月 24 日

摘 要

签名作为个人信誉的象征，从古至今在世界范围内都作为身份认证的凭证而备受重视。古代签订协议需要“签字画押”，现代的信用卡消费需要签字生效。随着科技发展，纸质文件逐渐被电子文件所取代，签名也由纸面转移到了屏幕。在这样的时代背景下，我们亟需一种优良的方法来自动鉴别在线签名的真伪，验证用户身份。

本文提出了一种新型的基于在线签名的动静态特征提取方法，分别利用在线签名的轨迹信息和签名图像，提取出动、静态特征的特征向量，然后对其进行有效的融合，最终实现在线签名的验证。对于静态特征，本文通过基于卷积神经网络（CNN）的自动编码器来识别、提取。对于动态特征，本文通过基于门控循环单元（GRU）的自动编码器来提取。最后通过孪生网络（Siamese Network），采用早融合和晚融合的策略对上述特征进行融合，判断签名相似度以鉴别真伪签名，取得了理想的效果。

本文在 SCUT-MMSIG、SigWiComp2013、SVC2004 与 MCYT-Signature-100 数据集上进行了实验，将本文提出的特征提取方法与已知文献方法的运行效果进行对比。其中在 SCUT-MMSIG 的空中签名数据集上，本文的动静态特征融合方法取得了最低 9.24% 的等误率（EER），为该数据集上目前已知的最佳成绩。同时，本文还对比了不同特征融合方法对等误率的影响，并分析了其内在原因。最后通过改变训练集签名数量，本文对比了不同参考签名数量对等误率的影响。

关键词：在线签名验证；自动编码器；孪生网络；卷积神经网络；门控循环网络

Abstract

As an important identity credential, signatures are widely used in our daily lives. Ever since ancient times, people have been signing contracts, and we are still using signatures for credit card transactions today. With the rapid development of information technology, paper-based documents are gradually being replaced by electronic ones, and the signing process is being transferred onto the digital screen. With the development of our times, it is necessary to incorporate a well-designed online signature verification system so as to boost automation and verify user identity.

In this paper, we proposed a novel online signature verification system to extract both dynamic and static features from online signatures. Notably, we utilized tracking information and signature image from online signatures to generate feature vector of dynamic and static features. Then, we fused those features that were extracted and applied them to verify user identity. Firstly, a convolutional neural network (CNN) based autoencoder was trained, and was later utilized to extract static features from signature images. After, a gated recurrent unit (GRU) based autoencoder was trained as a means to process the tracking data and extract the dynamic features. Finally, a siamese network was trained. Particularly, the features were fused using multiple strategies and considered as an input of the siamese network. Notably, the output of the siamese network is able to distinguish genuine ones from those that are forged.

Multiple experiments have been carried out on our online signature verification system. Specifically, our system was benchmarked on SCUT-MMSIG, SigWiComp2013, SVC2004, MCYT-Signature-100 and our results were compared with the one in the literature. Remarkably, our proposed system reached as low as a 9.24% equal error rate (ERR) on the SCUT-MMSIG dataset in-air signature subset. To the best of our knowledge, this is the best score on this dataset currently. Additionally, the effect of the different feature fusion method was evaluated, and our explanation was discussed. Then, through changing the number of

reference signatures, the EER change was analyzed.

Keywords: Online Signature Verification; Autoencoder; Siamese Network; Convolutional Neural Network; Gated Recurrent Unit

目 录

摘 要.....	II
Abstract.....	III
目 录.....	V
第一章 绪论.....	1
1.1 引言.....	1
1.2 研究背景.....	1
1.3 研究现状.....	2
1.4 论文结构.....	3
第二章 动静态特征的提取方法.....	5
2.1 数据的预处理.....	5
2.1.1 静态特征提取的数据预处理.....	5
2.1.2 动态特征提取的数据预处理.....	6
2.2 基于 CNN 的静态特征提取方法.....	6
2.3 基于 GRU 的动态特征提取方法.....	10
2.4 基于孪生网络的动静态特征融合.....	12
2.4.1 孪生网络.....	12
2.4.2 动静态特征的融合.....	15
2.5 本章小结.....	17
第三章 数据集与实验设计.....	19
3.1 数据集.....	19
3.1.1 SCUT-MMSIG 数据集-空中签名子集.....	19
3.1.2 SigWiComp2013 数据集.....	21
3.1.3 SVC2004 数据集.....	22
3.1.4 MCYT-Signature-100 数据集.....	23
3.2 特征提取模型的参数选择.....	24

3.2.1 基于卷积神经网络（CNN）的自动编码器训练参数	24
3.2.2 基于门控神经网络（GRU）的自动编码器训练参数	25
3.2.3 孪生网络的训练参数.....	25
3.3 评价指标.....	27
3.4 本章小结	28
第四章 实验结果.....	29
4.1 SUCT-MMSIG 空中签名数据集上的实验结果.....	29
4.1.1 不同特征融合方式的对比.....	30
4.1.2 不同参考签名数量对等误率的影响.....	31
4.2 SigWiComp2013 数据集上的实验结果	33
4.3 SVC2004 数据集上的实验结果	34
4.4 MCYT-Signature-100 数据集上的实验结果	35
4.5 本章小结	36
结论.....	37
1. 论文工作总结.....	37
2. 工作展望.....	37
参考文献.....	38
致谢.....	41

第一章 绪论

1.1 引言

在信息社会，小到注册社交账户，大到获取学历，都离不开身份认证的环节。在信息爆炸的时代下，高效、准确的确认用户身份，对于维持信息化社会的正常运转至关重要。如今签名作为使用最广泛的身份认证手段，在政府、金融、贸易等各个方面都得以应用。随着电子化时代的到来，我们越来越多的通过触摸屏、数位板等方式进行电子化的签名。相较于传统的纸质签名，电子化的签名富含丰富的轨迹信息数据，且可以非常容易的保存、分析、对比验证，这也使得每次签名后都校验签名的真实性成为了可能。

1.2 研究背景

古往今来，身份认证一直是人类社会活动中的重要话题。人们一直为了探索高效、精准的身份验证技术而不断努力。传统验证方式如密码、PIN 虽已广泛应用，但其易遗忘、易泄露等缺点导致其难以满足现代社会的需求。

近年来，生物技术的发展使得通过生物手段的身份认证成为可能。生物认证方法主要分为两种：基于生理特征的认证和基于行为特征的认证^[1]。前者直接通过直接扫描人体的一部分来获取每个人独一无二的标志，以此来鉴别用户。代表性的有指纹、人脸和虹膜识别技术，这些技术随着智能手机的普及已经深入大众生活。但其随之暴露出的种种问题也饱为诟病。由于它们都使用的是生物的固有特征，并且指纹、人脸特征是人体无法更改的生物信息，一旦泄露或落入不法分子手中，后果将不堪设想。并且，这类基于生理特征的认证技术难以判断认证对象是否有意识。这类认证技术的环境下，趁对方睡着时偷用指纹查看对方手机之类的行为并不是一件令人惊奇的事情。

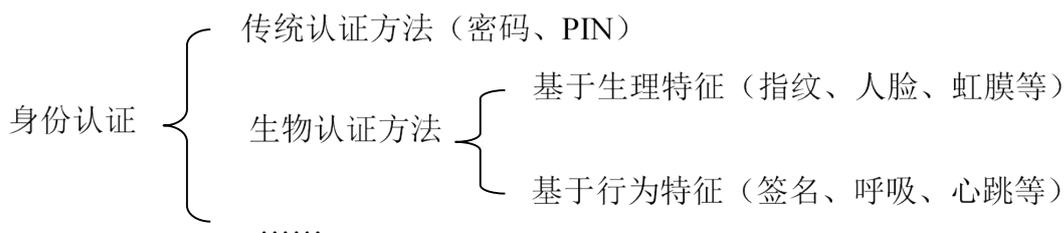


图 1-1 身份认证的分类

基于行为特征的生物认证技术通过提取生物动作、行为特征来验证身份^[11]。代表性的有签名、呼吸、心跳检测技术等。呼吸和心跳检测需要专业的设备，而签名认证操作简单，设备要求低，成为比较理想的行为特征认证方式。签名的采集方法多种多样，数位板、触控屏幕、甚至通过摄像头进行指尖签名都可以采集签名，这大大降低了实际应用中数据采集的成本。在验证的时候，只需要将已知的签名特征与现场的真实签名作对比即可。签名难以复制和模仿，具有很强的唯一性，同时又需要被认证的个体意识清醒才能完成，这使得其安全性大大增加。现在银行等金融机构开卡时都会收集客户的电子签名，签名认证的安全性与可靠性可见一斑。

综上，从认证信息的可靠性、信息泄露的风险性、信息收集的便利性等角度来看，电子签名的综合性能远远超过了其他现有的认证手段，是现代不可替代的身份认证手段。

1.3 研究现状

目前签名验证的研究领域已有大量的研究成果，可以实现签名的自动化验证，其实现的思路和方法各不相同。

根据已有的文献资料，签名验证可以分为在线签名验证和离线签名验证。在线签名验证需要收集签名过程中的多种信息，包括但不限于笔触坐标、压力、方位角、倾角、时刻等，以一系列点的信息来记录特征。在验证的时候，需要提取出这些信息的特征，并将其与真实签名进行对比，以判断真伪。因此，在线签名验证系统在收集真实签名的时候，不能仅仅收集签名完成后的图像，还要收集一系列的轨迹信息。与在线签名验证相对应的是离线签名验证。离线签名验证只使用签名完成后的位图作为输入，从图像中提取特征信息，与真实签名的图像进行对比。离线签名验证可用的信息非常有限，但优点在于设备兼容性强，只要能得到签名图像，就可以验证身份。

从特征提取的角度出发，可将签名验证的方法分为基于全局特征的方法和基于局部特征的方法。前者将整个时间序列作为一组全局特征来表示签名，例如一系列的坐标、压力、倾角、加速度、曲率等，也被称作函数式特征。而后者使用签名轨迹中的一些局部属性来代表签名，只从整个时间序列中筛选出一些统计信息，而不考虑全局的序列信息，也被称作参数式特征。

从聚类的角度出发，可将签名验证的方法分为基于距离的聚类和基于模型的聚类方法。例如 DTW 动态时间规划就是一种典型的距离匹配算法，它可以将两种不同时长的签名的样本点对齐，并计算两者的距离。基于距离的聚类算法还包括最长公共子序列、编辑距离等，它们需要用固定长度的全局特征向量来表示每个签名，然后用欧拉距离等方式计算差距。而基于模型的方法则使用诸如隐马尔可夫模型 HMM、高斯混合模型 GMM 等方式来构建签名的统计特征，用以鉴别签名。

最传统的签名认证方法是动态时间规划 (DTW)，它可以将两个不同长度的样本对齐后计算其距离^[7]。由于直接使用 DTW 匹配的效果并不理想，学者们提出了很多改进措施。Harma 等人^[5]尝试先使用高斯混合模型 GMM 来提取统计特征，再使用 DTW 进行匹配验证。Okawa 等人^[4]提出了一种基于时间序列平均的方法，用以降低时间计算复杂度和错误率。除了上述基于 DTW 的匹配方法外，Barkoula 等人^[8]还通过计算最长公共子序列来进行签名验证。而 Gupta 等人^[9]则使用编辑距离的方式来计算两个签名之间的差异，设定一个阈值以判断真伪。上述方法都是传统的方法，没有涉及到深度学习。但由于在线签名验证的特殊性，样本数量较少，所以这些传统的方法也取得了非常好的效果，值得我们关注。

采用深度学习的签名验证方法也有很多。Lai 等人^[10]将现有的多种数据库整合到一起，训练了一个循环神经网络 RNN，可以将长度不一的签名转换为一个统一长度的向量表达，通过直接对比向量间的距离就能验证签名。Vorugunti 等人^{[11][2][6]}在 2019 年提出使用特征融合这种深度学习的方法来提取签名特征。他通过聚类的方法，从多个角度来表达用户签名的特征信息。他后续又持续改进了这个神经网络，提出使用 DeepFuseNet，将多个特征进行深度融合，以此获得更好的效果。Ahrabian 等人^[3]采用长短记忆网络 (LSTM) 构造自动编码器 (Autoencode) 这种无监督聚类的方法来识别并提取特征向量，并使用孪生网络进行分类训练，该研究对本文产生了很大的启发。

1.4 论文结构

本论文共分为五章。

第一章绪论，简述了在线签名验证的动静态特征提取方法选题背景和研究意义，以及在线签名验证领域已有的研究成果。

第二章详细描述了本论文所提出的在线签名动静态特征的提取方法与签名验证的框架，包括数据的预处理、所使用的网络结构与原理、特征融合的方法等。

第三章详细描述了实验设计，包括所使用的数据集、神经网络的训练参数等。

第四章汇总了本论文提出的在线签名验证框架在不同数据集、不同参数下的实验效果，对实验数据进行总结分析，并与已有的研究成果进行对比。

最后，第五章对本论文的在线签名的动静态特征提取工作进行了总结，并提出未来的展望。

第二章 动静态特征的提取方法

本章将介绍基于在线签名的动静态特征提取方法。2.1 节将会介绍数据的预处理过程，该小节通过轨迹数据生成合适的静态图片，用于静态特征的提取，同时还对动态特征序列进行填充，用于动态特征的提取。2.2 节将会介绍基于卷积神经网络（CNN）的静态特征提取方法。该小节将训练一个基于 CNN 的自动编码器模型（Autoencoder），从而实现静态特征的提取与降维。2.3 节将会介绍基于门控循环单元（GRU）的动态特征提取方法。该小节会训练一个基于 GRU 的自动编码器模型，从而将长度不一的签名序列转换为统一长度的特征向量，作为签名的动态特征。2.4 节将会介绍孪生网络模型与动静态特征融合。该小节通过 Early Fusion 和 Late Fusion 两种方法来构造孪生网络，用以融合上述的动、静态特征。最后，使用该模型来实现一次学习（One-Shot Learning），以实现在少量参考签名的情况下也能正常判断真伪签名。

2.1 数据的预处理

2.1.1 静态特征提取的数据预处理

对于在线签名的静态特征提取部分，需要将静态图片输入到基于 CNN 的自动编码器中进行提取。本论文实验所采用的数据集，除 SigWiComp2013 之外，都没有提供在线签名的图片，所以需要通过轨迹数据绘制出签名图像。

本论文使用 Pillow 绘制出分辨率为 128x128 的空白图像，然后依据动态序列绘制签名图像，并设置线条粗细为 4 像素。实际测试中发现，如果线条设置过粗（例如线条宽度大于 10 像素），会使得线条大量重叠，重叠的部分无法再分辨出线条，这会导致原本签名中的细节信息丢失，不利于后续的身份验证。但线条又不能过细（例如宽度为 1 像素），否则在自动编码器训练的过程中，由于黑色线条部分过少，这些黑色线条就如同噪点一样，它们会在编码器进行编码压缩的时候被移除。最终会导致在任何输入下，解码器重建的图像都是一个灰色图。这也就是 Autoencoder 的降噪特性。通过实验发现，在绘制图像分辨率为 128x128 的情况下，4 像素粗细的线条在本论文提出的框架中表现良好。

为了便于后续的研究，我将所有签名读入后并绘制成图像后，按原来的文件名重新

保存。值得一提的是，平时常见的 JPG 格式是有压缩的图像格式，会造成图像信息损失，不适合于本论文的研究场景，所以本论文所有转换的图像均使用 PNG 格式保存。

由于各个数据库的数据格式不尽相同，我会在 3.1 节介绍本论文使用的数据库时，详细阐述每个数据库需要进行的预处理。

2.1.2 动态特征提取的数据预处理

对于在线签名的动态特征提取部分，需要将签名的轨迹序列信息输入基于 GRU 的自动编码器。对于有多种指标的数据集，通常每个指标的范围都不一样，但对它们进行缩放不会对最终结果造成影响，因此我们先将所有值都进行归一化处理到[0,1]区间，便于后续的研究。对于基于 GRU 的自动编码器模型，由于需要在一个训练批次内输入多条签名，所以还需要对签名进行填充，使得所有签名长度一致。这样就可以在一个批次中同时训练多个签名，加快了训练速度。

2.2 基于 CNN 的静态特征提取方法

卷积神经网络 (CNN) 是最广为人知的深度神经网络算法之一。它通常由多个卷积层、池化层、全连接层等所构成，常常被用于图像处理、分类、自然语言处理等多个方面。

自动编码器 (Autoencoder) 则是一种能自动学习如何将数据有损压缩并重建的无监督神经网络。它主要由编码器 (Encoder)、解码器 (Decoder) 两部分构成。它在神经网络一侧进行编码 (Encode)，在另一侧进行解码 (Decode)。在训练时，自动编码器通过编码器的神经网络，对输入的数据进行有损压缩，得到一个维度较低的瓶颈层

(Bottleneck Layer)，有时也被称作隐藏层。然后再通过解码器的神经网络，使用瓶颈层向量来尝试重建原始的输入数据。重建出来的数据与原始数据的差异，就是重建损失 (Reconstruction Loss)。在训练的过程中，我们期望重建损失尽可能的小，即自动编码器的最终的输出与原始输入相比，差异越小越好。由于整个过程的输入与目标输出是一样的，不需要额外的先验知识，所以自动编码器属于无监督学习。

理想情况下，自动编码器右侧的解码器可以根据瓶颈层中的信息，重建出原始的输入数据，因此，瓶颈层中的向量被视作原始输入数据的特征，以此实现数据降维。常见

的降维方法还有 PCA、LDA 等，但这些简单的数学方法只适用于线性数据，而自动编码器可以处理非线性数据^[13]，因此更符合我们的应用场景。

对于在线签名的静态特征，我们使用基于 CNN 的自动编码器模型来提取，构造的自动编码器网络构造如图 2-1 所示。输入的 128x128 像素的灰度图像，经过卷积神经网络的编码器编码后，瓶颈层输出一个 128 维的特征向量。然后这个向量再经过卷积神经网络的解码器解码之后，重新还原回 128x128 维度的向量。

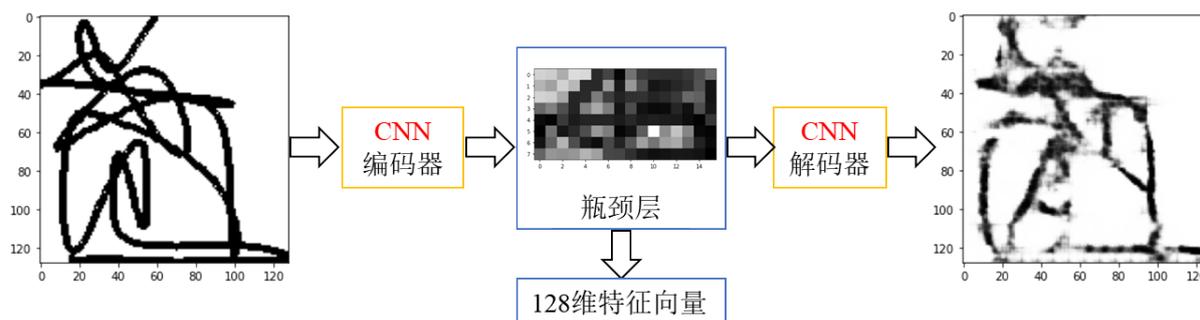


图 2-1 基于 CNN 的自动编码器结构示意图

对于编码器部分，首层是 32 过滤器、3x3 卷积核的卷积层；第 2 层是 2x2 窗口的池化层；第 3 层是 64 过滤器、3x3 卷积核的卷积层；第 4 层是 2x2 窗口的池化层；第 5 层是 2x2 窗口的池化层；然后是 32 过滤器、3x3 卷积核的卷积层；接下来是 Flatten 层，将输入的向量“压平”；最后是全连接层，输出的向量即为瓶颈层向量，可作为输入图片的代表特征。整个编码器的结构如图 2-2 所示。

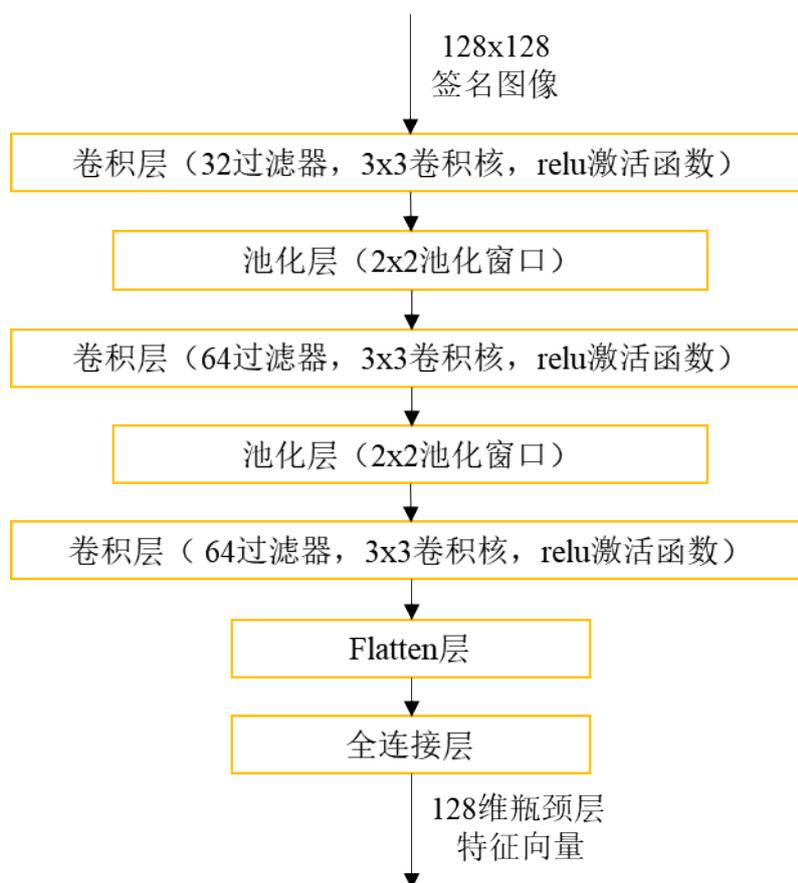


图 2-2 编码器结构示意图

对于解码器部分，首层是重构层 (Reshape Layer)，将输入的 128 维向量转换为 8×16 的向量，但内容不会发生改变；其次是 64 个过滤器、 3×3 卷积核的反卷积层，卷积的步长为 2×1 ，即第一维拉伸至原来的 2 倍，第二维不发生变化；然后是批量标准化层，它可以使得输出向量的均值接近于 0，方差接近于 1，它的输出向量维度不会发生变化；第 4 层是 64 个过滤器、 3×3 卷积核的反卷积层，积的步长为 2×2 ，即两个维度都拉伸至原来的 2 倍；第 5 层同样是批量标准化层；第 6 层是 32 个过滤器、 3×3 卷积核的反卷积层，卷积的步长为 4×4 ，即两个维度都拉伸至原来的 4 倍；最后一层是 32 个过滤器、 1×1 卷积核的卷积层，并指定 sigmoid 激活函数，最终输出 128×128 维向量，与自动编码器的原始输入形状相同。整个解码器的结构如图 2-3 所示。

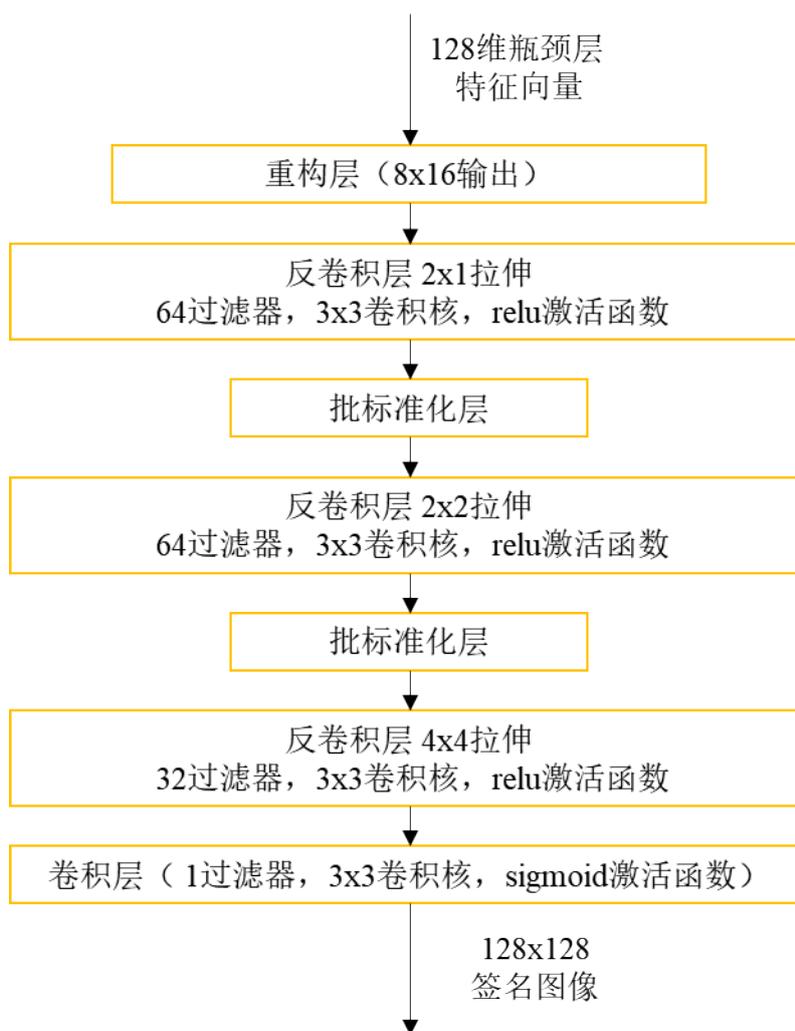


图 2-3 解码器结构示意图

理想状态下，训练完成之后的自动编码器的输出向量会非常的接近输入向量。此时需要固定自动编码器各个节点的权重，然后去掉自动编码器右侧的解码器，就得到了一个输入 128×128 像素的灰度图，输出 128 维特征向量以描述静态特征的神经网络。其结构如图 2-4 所示。

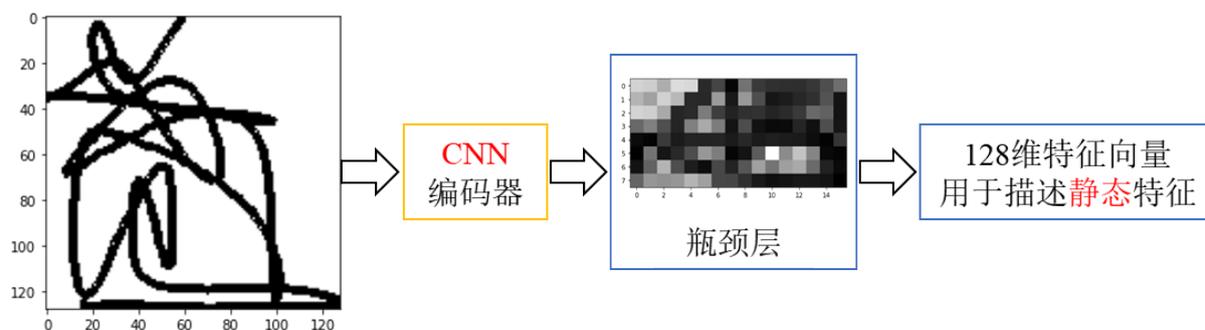


图 2-4 去除解码器后得到的静态特征提取器

2.3 基于 GRU 的动态特征提取方法

在卷积神经网络中，神经元的信号只能向一个方向单向传播，这导致其对序列的处理能力较差，无法联系上下文。而循环神经网络（RNN）在每个时刻的输出则可在下一个时刻直接作用于自己。每个 RNN 单元会维持一个隐藏状态 S ，其输出取决于输入 x 和上一时刻的隐藏层，如公式 2-1 所示。

$$S_t = f(U \cdot X_t + W \cdot S_{t-1}) \quad (2-1)$$

$$O_t = g(V \cdot S_t)$$

这种结构使得循环神经网络能够很好的联系上下文，因此具有很好的序列处理能力。但这也带来了一个新问题：此时的循环神经网络深度就等同于输入序列的长度。如果输入过长的序列长度，就等同于建了一个深度非常大的网络，因此带来了梯度消失或梯度爆炸的问题。此时容易陷入局部最优解，而难以找到全局最优。同时，实际使用中 RNN 网络还有难以记忆较远之前访问的数据点的问题。

1997 年 Sepp Hochreiter 提出了长短记忆网络 Long Short-term Memory (LSTM)，循环神经网络的梯度消失与梯度爆炸问题才得以解决^[15]。它通过控制输入门、遗忘门、输出门来决定记忆或者丢弃某些信息，以此来避免梯度消失的情况。2014 年，Kyunghyun Cho^[14]提出了门控循环单元 (GRU)。它通过合并遗忘门和输入门，简化了 LSTM 的状态，使得 GRU 网络比 LSTM 网络的计算效率更高。与此同时，GRU 网络的计算效果却和 LSTM 相差无几。

基于 GRU 的自动编码器模型如图 2-5 所示。构建基于 GRU 的自动编码器模型与上述基于 CNN 的自动编码器模型结构较为类似，其主要区别在于接受的输入向量形状不同，以及 GRU 编码器内部结构不同。（注：该图中的演示图像由轨迹序列重建出来，实际处理中并不涉及图像，而是使用其轨迹序列。）

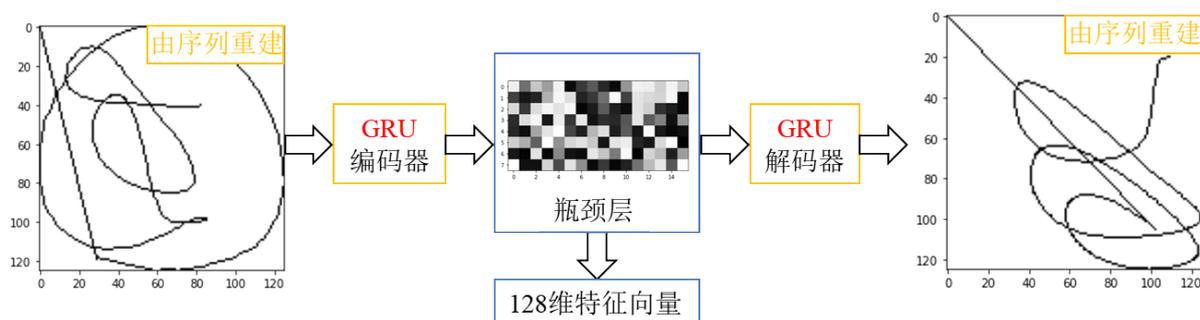


图 2-5 基于 GRU 的 Autoencoder 模型

GRU 编码器由 2 个 128 单元的 GRU 层构成。为了避免填充数据影响模型精度，还需额外添加一个遮罩层。遮罩层的信息会自动向下传递给后续的 GRU 层。GRU 编码器结构如图 2-6 所示。

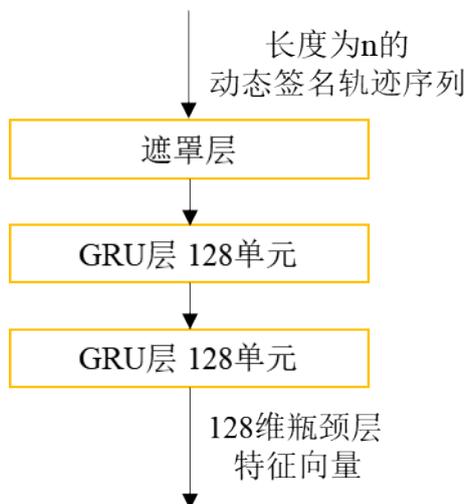


图 2-6 GRU 编码器的构造

GRU 解码器同样也由 2 个 128 单元的 GRU 层组成。为了使得瓶颈层能够作为 GRU 解码器输入，还需要在 GRU 解码器前添加一个重复层，使瓶颈层重复 MaxN 次（MaxN 为最长序列的长度）。GRU 解码器结构如图 2-7 所示。

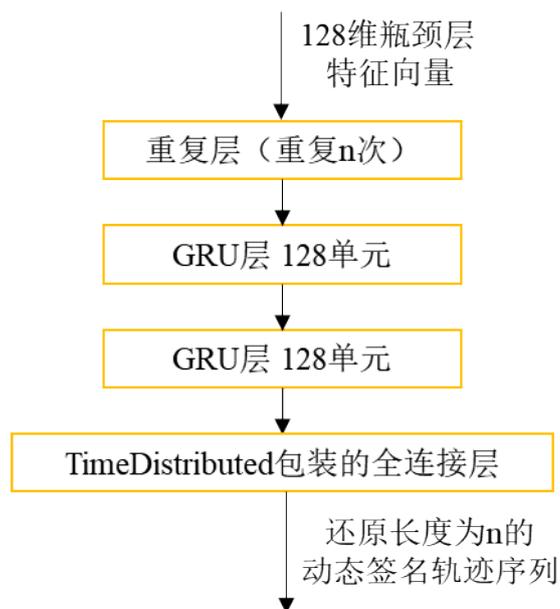


图 2-7 GRU 解码器的构造

由于每个签名的长短并不一致，而每个训练批次（Batch）的向量长度又必须相同，因此可以选择逐个训练，或者添加填充（Padding）。出于效率的考虑，本论文实验时选择了后者。先找到训练集中最长的序列长度，然后通过填充 0 向量使得所有训练序列长度一致。

同样，在完成训练后将节点权重固定，移除自动编码器右侧的解码器，即可得到一个输入变长签名填充后的序列、输出固定 128 维特征向量的神经网络，如图 2-8 所示。

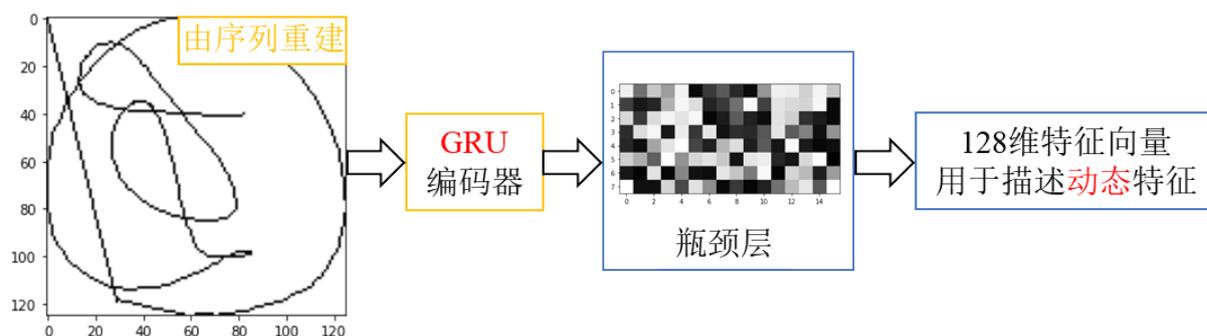


图 2-8 移除解码器后得到的神经网络模型

2.4 基于孪生网络的动静态特征融合

2.4.1 孪生网络

孪生网络（Siamese Network）是一种能够衡量两个输入的相似度的神经网络，由两

个相同且共享权重的左右子网络构成。它将两个输入向量分别通过左右子网络运算，得到新的输出向量，然后将这两个向量相减，并通过一个全连接层来衡量输入的相似度。在训练过程中，左右子网络会使得两个同一目标值的输出距离尽量接近、不同目标值的输出距离尽量远。

使用孪生网络的一个重要原因是孪生网络训练不需要依赖大量的样本。训练在线签名识别系统的一个特殊之处就在于，样本量通常非常小。这是因为在实际的应用中，在线签名验证系统很难从用户侧获取大量的签名作为参考签名。根据已有文献，一般只使用 5 个签名作为验证系统的参考签名。这类参考样本极少的极端情况被称作一次学习 (One-Shot Learning)。而孪生网络只需要少量样本就能完成训练。同时，如果需要添加、移除分类，传统的神经网络就需要将所有数据重新进行训练。而孪生网络并不是直接被训练来用于分类，而是训练出一个相似函数，能够衡量两个输入的相似性。因此，训练好的孪生网络可以直接在新添加的分类下使用，不需要再次训练^[19]。之所以其被称作一次学习，是因为在训练好孪生网络后，如果要鉴别新的签名，我们可以最少只提供一个新的签名作为参考，就能在后续鉴别测试签名的真伪。

孪生网络的示意图如图 2-9 所示。两个输入向量 X_1 、 X_2 经过 CNN 子网络之后，得到两个输出向量 $G_w(X_1)$ 、 $G_w(X_2)$ ，然后通过一个距离函数 D 即可得到这两个输出向量的距离。这个距离就可作为它们相似度的评判标准。

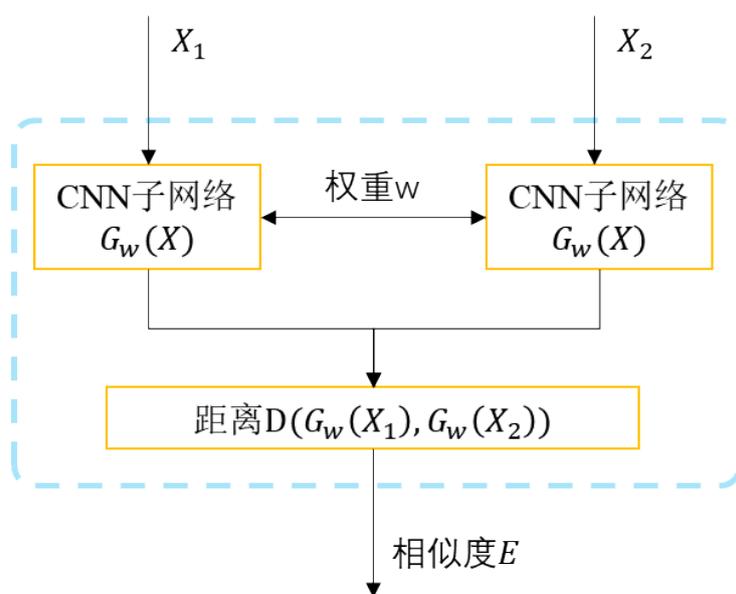


图 2-9 孪生网络结构示意图

图 2-9 中所述孪生网络的 CNN 子网络结构如下。它接受 128 维特征向量的输入，然后经过一个 Dropout 层，通过每次随机丢弃 20% 的神经元节点来避免过拟合；第 2 层是 64 过滤器、10x10 卷积核的一维卷积层，激活函数为 tanh；第 3 层是一维池化层；第 4 层是 128 过滤器、7x7 卷积核的一维卷积层，激活函数为 tanh；第 5 层是一维池化层；第 6 层是 128 过滤器、4x4 卷积核的一维卷积层，激活函数为 tanh；第 7 层是一维池化层；第 8 层是 128 个过滤器、4x4 卷积核的一维卷积层，激活函数为 tanh；接下来第 9 层是 Flatten 层，将输入的向量“压平”；然后是 Dropout 层，通过每次随机丢弃 20% 的节点来避免过拟合；最后是全连接层，其输出作为孪生网络子网络的输出。该网络的整体结构如图 2-10 所示。

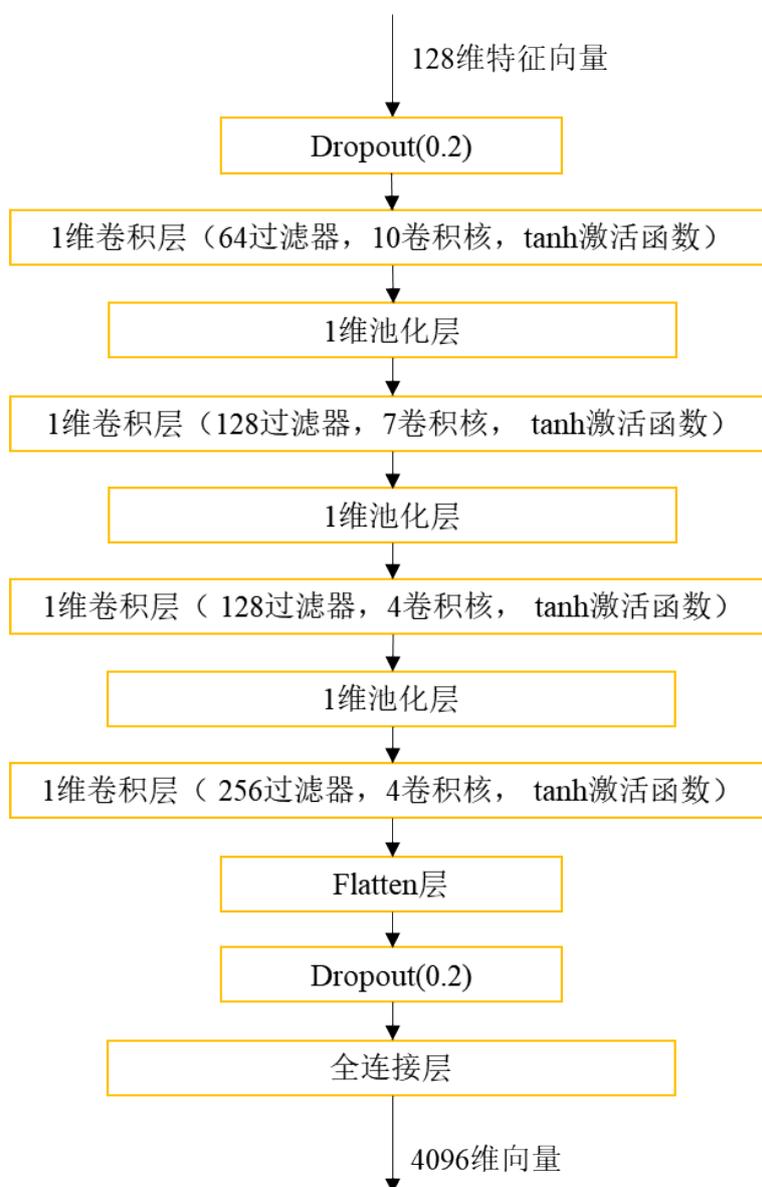


图 2-10 孪生网络的 CNN 子网络结构

2.4.2 动静态特征的融合

如前文所述，本论文实现了两个自动编码器，分别是基于 CNN 的自动编码器与基于 GRU 的自动编码器。它们分别对应于静态、动态特征的提取。这就意味着有动态、静态两种特征向量。特征融合是指，在形成最终的真伪判定之前，需要将两种特征向量结合起来的过程。

依据特征融合处于模型中的位置，可以将特征融合分为早融合（Early Fusion）、晚融合（Late Fusion）。早融合是特征级别的融合，先融合特征向量，再进行分类器的训练。而晚融合则是在已经完成特征分类后，再将其得到的结果分数进行融合，以期提升分类效果。

对于不同的融合方式，我们构建的孪生网络模型各有不同。本论文将在第四章中比较不同动静态特征融合方式对特征提取效果的影响。

2.4.2.1 早融合（Early Fusion）

使用早融合方法时，需要对每个签名采用 2.2、2.3 小节中的方法，提取出动态特征向量与静态特征向量，并将它们拼接到一起，构成一个 256 维向量以作为孪生网络的输入。

此时孪生网络的构造如图 2-11 所示。它接受两个 256 维特征向量作为输入，然后经过一个 Dropout 层，每次随机丢弃 20% 的神经元节点来避免过拟合。接着通过 CNN 神经网络得到两个输出向量，将其相减后再经过一个 Dropout 层，最后经过一个全连接层，输出签名的特征相似度 E 。

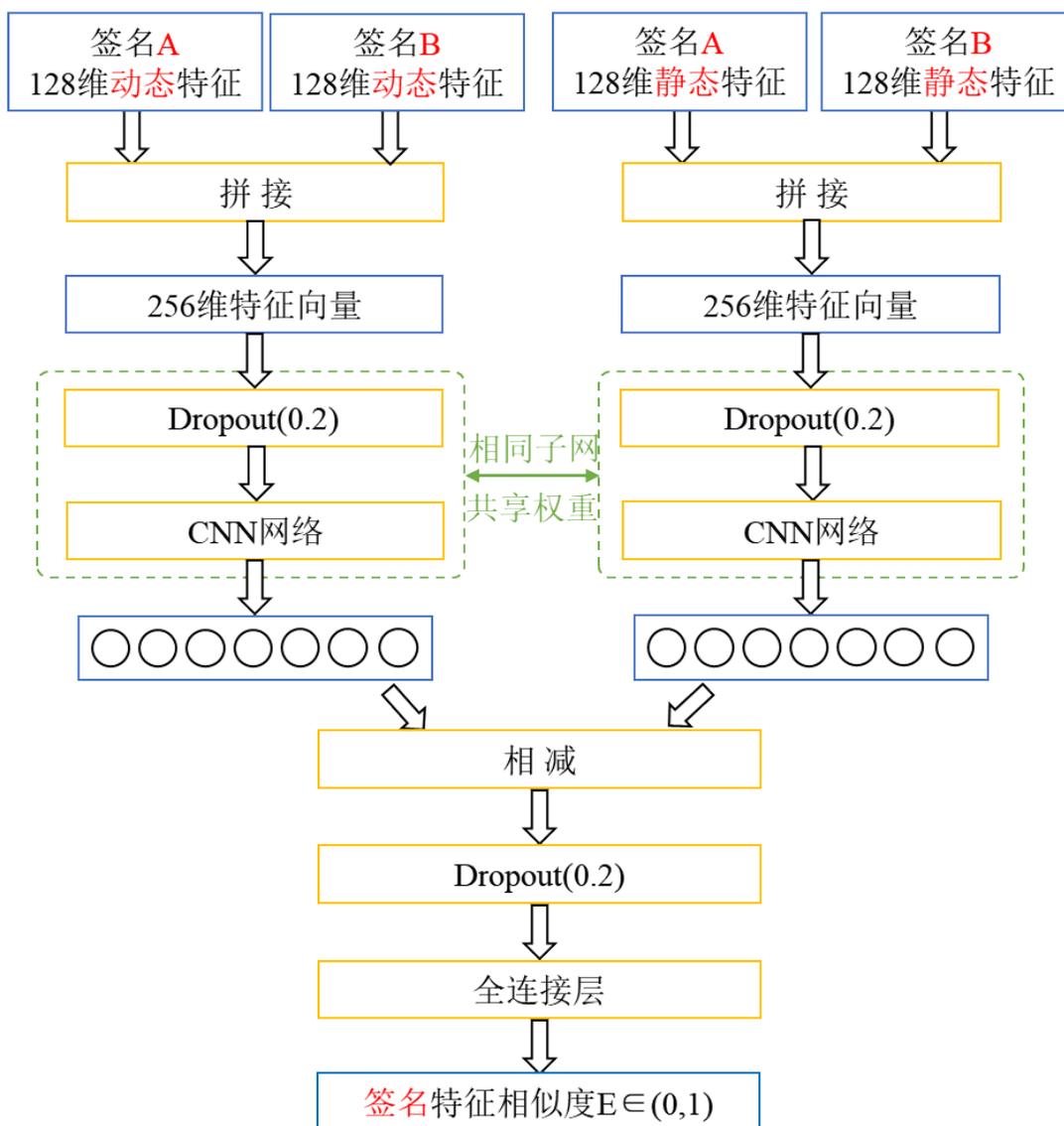


图 2-11 早融合方式的孪生网络示意图

2.4.2.2 晚融合 (Late Fusion)

在进行晚融合的时候，我们需要分别训练两个不同的孪生网络，各自处理动态特征与静态特征。计算得到待测签名与参考签名的动、静态相似度 E_1 、 E_2 ，最后将两个网络的输出的相似度 E_1 、 E_2 求平均，即可得到签名相似度 E 。

此时，对于处理动态特征的孪生网络，它接受两个 128 维特征向量的输入，然后经过一个 Dropout 层，每次随机丢弃部分的神经元节点来避免过拟合。接着通过 CNN 神经网络得到两个输出向量，相减后再经过一个 Dropout 层，最后经过一个全连接层，输出签名的动态特征相似度 E_1 。同理，处理静态特征的孪生网络会输出静态特征相似度 E_2 。我们将两个相似度 E_1 、 E_2 求平均，即可得到签名相似度 E 。

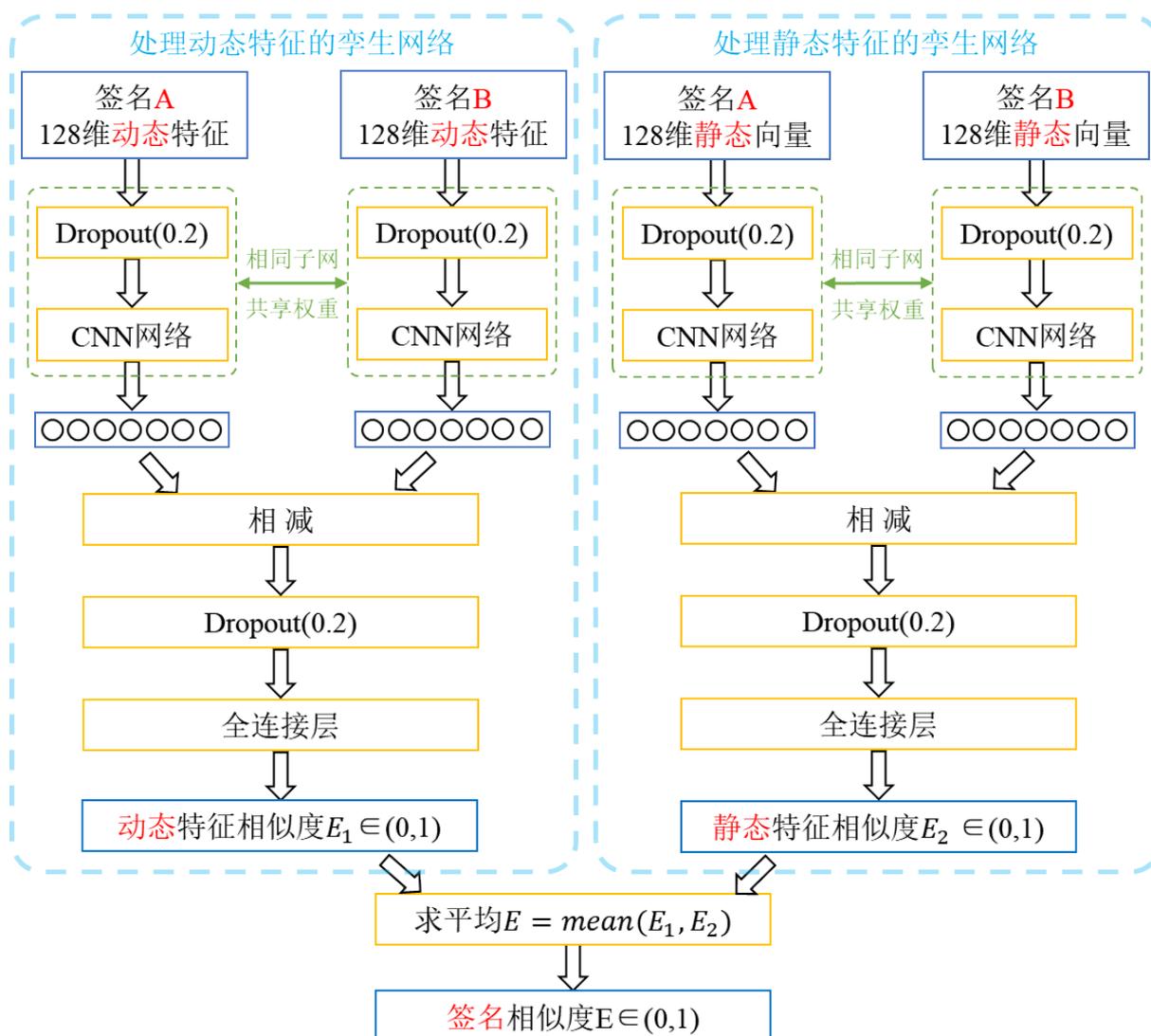


图 2-12 晚融合方式的孪生网络示意图

由图 2-12 可以清晰地看出，孪生网络需要成对的特征向量和对应的数据标签作为输入，我们需要手动生成这样的数据对。我将在第三部分实验设计与参数选择中详细讲解如何构建孪生网络的训练数据。

2.5 本章小结

本章对动静态特征提取的整体方法进行了讲解。首先介绍了数据预处理操作，包括静态签名图像的生成以及动态签名轨迹的填充。然后介绍了论文所有需要用到的神经网络模型的基本概念与实验模型的构造。接着详细描述了基于 CNN 自动编码器的静态特征提取、基于 GRU 自动编码器的动态特征提取，以及基于孪生网络的特征融合。文中采用了两种不同的特征融合策略，也就构建了两种不同的孪生网络模型。下一章节将会

主要介绍在上述结构的神经网络下进行的实验参数设定。

第三章 数据集与实验设计

本章将重点介绍与本论文相关的实验设计。3.1 节将介绍实验所用的数据集，包括其原始数据的构成、需要进行的预处理等；3.2 节将介绍自动编码器的训练参数，包括数据填充等；3.3 节将介绍训练孪生网络的训练参数以及训练所需数据的生成方式；3.4 节将介绍两种特征融合方式具体的实现；3.5 节将介绍衡量签名验证系统的等误率及其计算方法。

3.1 数据集

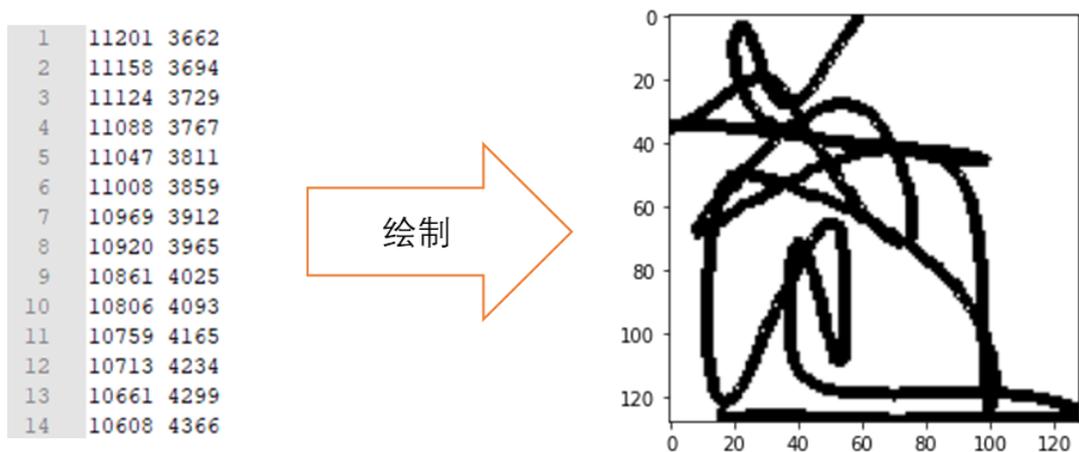
3.1.1 SCUT-MMSIG 数据集-空中签名字集

SCUT-MMSIG (Multimodal Signature Database of BIP Lab in SCUT)是一个多模态的签名数据库，由手机、手写笔、空中签名三个子库构成^[16]。本论文使用的是空中签名字库，共 55 个用户，每人对应 40 条签名。其中前半是真实签名，后半为熟练的模仿者伪造的签名。每个签名均以轨迹序列信息的形式保存，每个时间点由一个二维向量表达，分别表示 x 、 y 坐标。数据集文件为 txt 格式，每一行有两个参数即为上述的 x 、 y 坐标。

该空中签名数据库使用摄像头采集数据，通过跟踪用户在空中的手指运动轨迹来记录签名信息。其最大的特点是签名轨迹是完全连续不中断的，没有抬笔信息，因为用户在空中签名的时候，摄像头无法分辨用户是否抬笔。另外，该数据集制作过程中，出于空中签名的复杂度考虑，数据集内的所有签名均为单汉字的签名。SCUT-MMSIG 数据集不包含静态的签名图片，因此需要根据签名轨迹信息自行生产静态图片。

观察 SUCT-MMSIG 的空中签名数据集可发现，每张签名图像的 x 、 y 所在区间变化非常大。这是因为采集过程中，每个受试者签名的位置差别很大，书写签名的尺寸也不尽相同。而空中签名使用的是摄像头采集数据，签名者在签名过程中自己实际难以把控签名的位置大小等，更加剧了这一现象。所以在绘制图像前，必须将所有签名的坐标归一化，这样才能消除签名在图像中所在位置、签名大小所带来的影响。为了方便计算，本论文先将每个签名的 x 、 y 坐标都归一化到 $[0,1]$ 区间，然后将其乘以目标图像的长、宽。这样就能保证签名处于图像中心并占满图像，避免大量留白。绘制静态图象时，设

置绘图粗细为 4 像素，绘制完成后的静态图像效果如图 3-1 所示。



用户2-签名1-轨迹信息

绘制得到的用户2-签名1-签名图像

图 3-1 SCUT-MMSIG 空中签名数据集的轨迹信息及绘制的签名图象示例

在划分训练集与测试集时，本论文参考了房育勋^[12]的“混合阶段测试方法”。我们先用洗牌算法分别将真实签名、伪造签名顺序打乱。这样每次按顺序取出的签名就相当于随机抽取的签名。我们取每个用户前 5 个真实签名、5 个伪造签名作为训练集。然后再取 5 个真实签名作为测试集的参考签名。其余的则作为测试集的待测签名。在实验时，我们需要将测试集中 5 个参考签名与 25 个测试签名做对比，判断每个测试签名的真伪。数据集划分的示意图如图 3-2 所示。

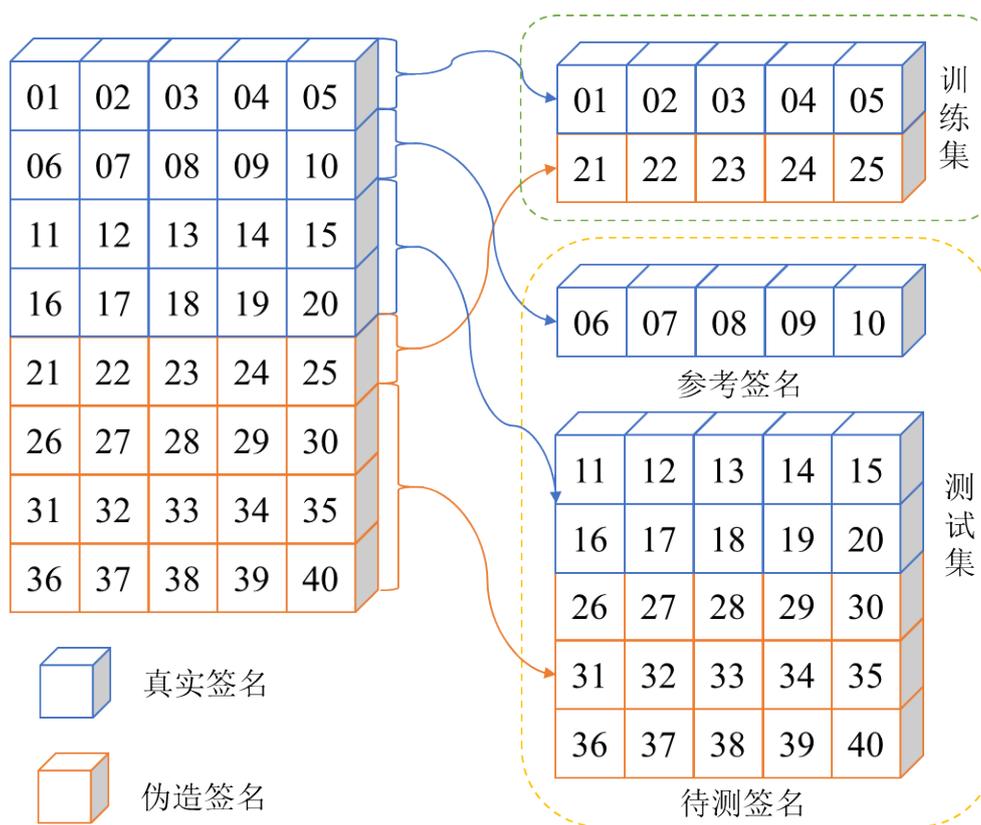
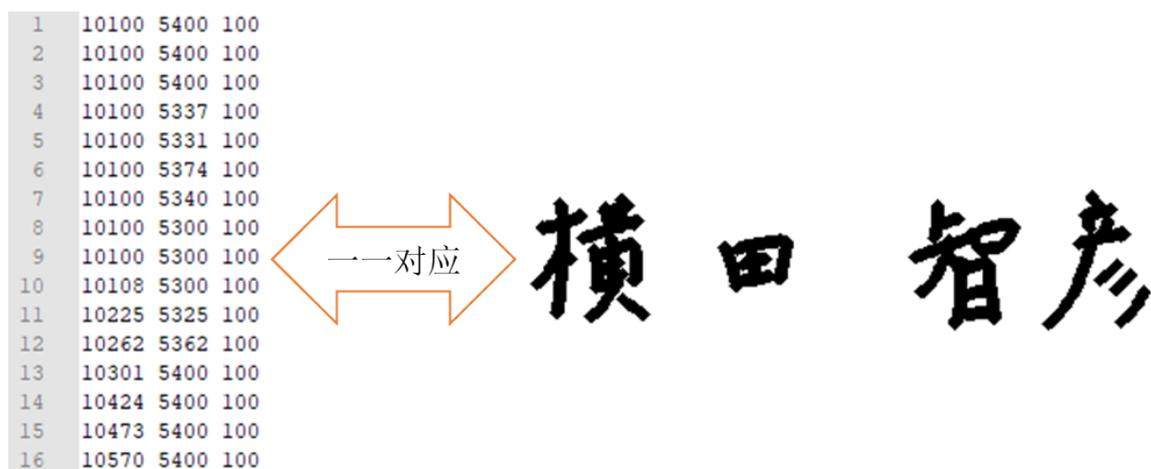


图 3-2 SCUT-MMSIG 空中签名数据集的训练集、测试集的划分示意图

3.1.2 SigWiComp2013 数据集

SigWiComp2013 是 ICDAR2013 签名识别竞赛所提供的的一个日语签名的数据集，由在线签名和离线签名部分构成^[17]。两部分签名一一对应，即所有签名既有静态的签名图像，也有对应的动态轨迹信息。该数据库包含 11 个人的签名数据，每人对应 42 个亲笔签名和 36 个伪造签名。伪造签名由熟练训练者仿写完成，每个伪造者对每个伪造对象各模仿了 9 次。数据集文件格式为 HWR，可以作为文本文件的形式读取。每一行有三个整数，用空格隔开，分别表示该采样点的 x、y 坐标以及笔压。当第三个参数笔压为 0 时即为抬笔状态。

SigWiComp2013 数据集在采集手写签名时，设备的采样率过高，这导致数据集过于庞大。最长的签名长度达到 17889 个采样点。由于循环神经网络的深度等同于输入序列的长度，过长的签名序列不仅会大大减缓训练速度，还可能带来梯度消失或者梯度爆炸的问题。为了有效提升训练效果，本文实验时对其进行了 10:1 降采样，即每 10 个点选取它们中第一个点作为代表，其他的数据点将直接舍去。



用户10-签名31-轨迹信息

用户10-签名31-签名图像

图 3-3 SigWiComp2013 签名轨迹信息与图像示例

离线签名信息部分，每个签名均以单独的 PNG 格式图像表示。该数据集中的签名图像大小均不统一，宽度在 110 像素左右，长度在 360 像素左右到 480 像素左右不等，但都已进行归一化处理，签名内容均在图像的中心，四周有少量等宽的白边。本论文使用双线性插值的方式，将所有图像都缩放到 128x128 像素，以便于研究。

与 SCUT-MMSIG 数据集的处理方法类似的，划分训练集与测试集时，我们先用洗牌算法分别将真实签名、伪造签名顺序打乱，以保证每次顺序取出的签名就相当于随机抽取。然后我们取每个用户前 5 个真实签名、5 个伪造签名作为训练集。在剩余的签名中，再取 5 个真实签名作为测试集的参考签名。剩余 32 个真实签名和 31 个伪造签名作为测试集的待测签名。在实验时，我们需要将测试集中的 5 个真实参考签名与 63 个测试签名做对比，判断每个测试签名的真伪。

3.1.3 SVC2004 数据集

SVC2004 是首届国际签名验证竞赛所提供的数据集^[20]，其数据集分为两部分：任务一的数据集仅包含签名序列的坐标信息；任务二的数据集在任务一的数据基础上，包含笔的倾角、压力等额外的信息。本论文采用任务二的数据集进行实验。该数据集共有 40 个用户数据，每个用户包含 40 条签名信息。其中前半是真实签名，后半是熟练伪造的签名。数据集中的用户签名由多种语言构成，包括简体中文、繁体中文、英文等。数据集文件为 txt 格式，首行为一个整数 n ，表示该数据集中轨迹点的个数。接下来 n

行，每行有 7 个参数，依次是 x 坐标、y 坐标、时间戳、抬笔状态、方位角、倾角、笔压。

因为这个数据库中包含了笔压数据，所以在绘制的过程中，可以将归一化至[0,1]后的笔压数据乘以预定的绘图线条粗细，以此动态地调整绘图中笔触的大小。绘制好后的图像如图 3-4 所示。

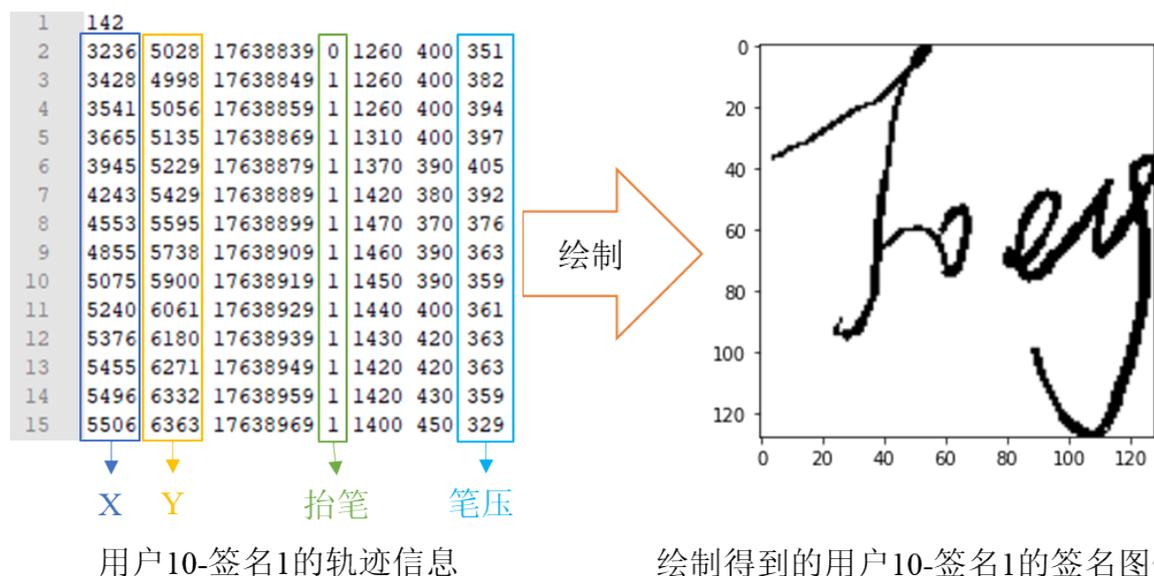


图 3-4 SVC2004 数据集的轨迹信息及绘制的签名图像示例

同样的，在划分训练集与测试集时，先用洗牌算法分别将真实签名、伪造签名顺序打乱，以保证顺序取出的签名就等效于随机抽取。我们取每个用户前 5 个真实签名、5 个伪造签名作为训练集。再取 5 个真实签名作为测试集的参考签名，其余的作为测试集的待测签名。在实验时，我们需要将测试集中的 5 个参考签名与 25 个待测签名做对比，判断每个待测签名的真伪。

3.1.4 MCYT-Signature-100 数据集

MCYT-Signature-100 数据集是 2003 年西班牙马德里理工大学生物统计实验室设计并制作的签名数据集^[21]。数据集包含 100 个用户签名信息，以文件名区分签名真伪，签名编号前是 v 的表示真实签名，签名编号前为 f 的表示虚假签名。该数据集需要使用 matlab 读取，为了方便研究，本论文将数据集进行预处理并转换为通用的 txt 格式。处理后，首行两个整数分别表示采样点数量和采样率；剩余每行有 5 个整数，依次是 x 坐标、y 坐标、笔压、方位角、高度，笔压为 0 时为抬笔状态。与前述数据集类似，该数

数据集包含了笔压数据，所以在绘制的过程中，可以将归一化至[0,1]后的笔压数据乘以预定的绘图线条粗细，以此动态地调整绘图中笔触的大小。绘制好后的图像如图 3-5 所示。

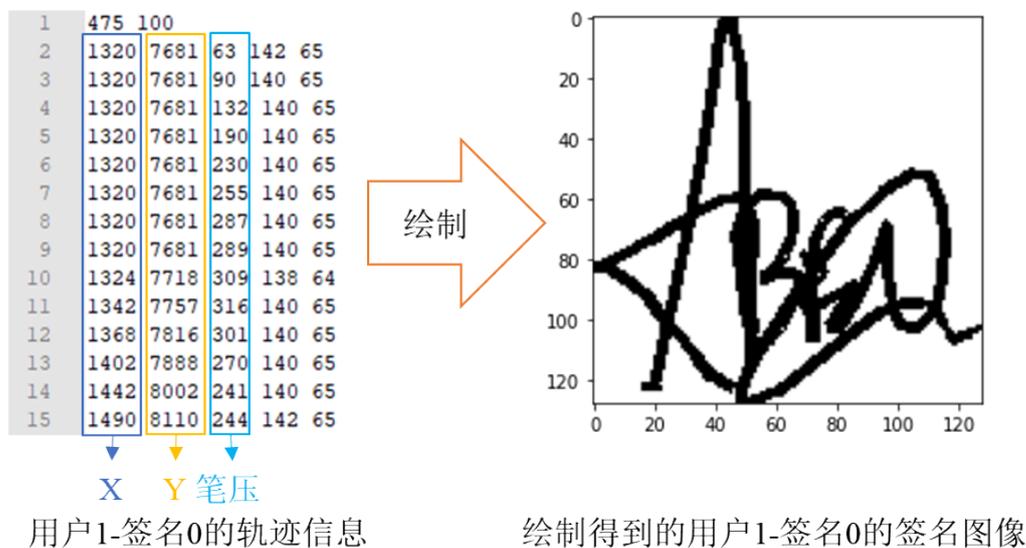


图 3-5 MCYT-Signature-100 数据集的轨迹信息及绘制的签名图象示例

在划分训练集与测试集时，先用洗牌算法分别将真实签名、伪造签名顺序打乱，以保证顺序取出的签名就等效于随机抽取。我们取每个用户前 5 个真实签名、5 个伪造签名作为训练集。再取 5 个真实签名作为测试集的参考签名，其余的作为测试集的待测签名。在实验时，我们需要将测试集中的 5 个参考签名与 35 个待测签名做对比，判断每个待测签名的真伪。

3.2 特征提取模型的参数选择

3.2.1 基于卷积神经网络（CNN）的自动编码器训练参数

对于基于 CNN 的自动编码器，我们主要需要将图像归一化居中之后，再放缩到 128x128 像素。同时要注意 PNG 图像有 3 通道。我们绘制的黑白签名图像这 3 个通道的数据是完全一致的，只需要抽出其中的一个通道，得到 128x128x1 的输入向量。

训练自动编码器时，使用 MSE 损失函数，使用 Adam 优化器来自动调整训练速率，进行 500 轮训练。每个批次大小（Batch Size）为 128 个，并开启洗牌算法，每轮训练都将训练数据打乱后输入。

3.2.2 基于门控神经网络（GRU）的自动编码器训练参数

对于基于 GRU 的自动编码器，因为我们需要一次将多个签名数据输入训练，所以我们需要对长度不一的签名进行填充（padding），全部填充 0 向量。同时还需要在 GRU 网络中的每个层对填充的 0 向量进行过滤。对于不同的数据集，其最长的长度不同，需要填充的长度也就不同，具体数据如表格 3-1 所示。

表 3-1 不同数据集填充长度与填充向量内容

	填充内容	填充至长度
SCUT-MMSIG	[0 0]	1369
SIGWICOMP2013	[0 0 0]	1789
SVC2004	[0 0 0 0 0 0 0]	713

在训练 GRU 模型时，还需要对其添加遮罩（Masking），使得 GRU 模型能够忽略我们填充进去的 0 向量，这样填充就不会对模型精度造成影响。

值得注意的是，如果使用 Keras 框架，我们需要使用它的重复层（Repeat Layer）将瓶颈层的内容重复叠放 n 次（ n 为上述填充完成后的长度）。但该重复层是不支持传递遮罩的，神经网络输入层指定的遮罩将会在这一层失效，并且无法向下一层传递。这种情况下通常训练出来的损失值会小一些，但其实是因为填充的内容也被当作了数据的一部分，大量的 0 向量填充干扰了模型训练，这与我们的预期不符。所以这里需要手动创建重复层，将其嵌入一个 GRU 层中，这样遮罩信息能够以 GRU 层的处理逻辑，正常向下传递。

在训练自动编码器时，我们使用训练集中的全部签名进行训练，输入与输出相同。自动编码器的主要工作是将图像或序列转化为定长的特征向量，因此训练自动编码器时，不需要区分真实签名与伪造签名，只需要将所有签名都放入自动编码器训练即可。

3.2.3 孪生网络的训练参数

由于孪生网络特殊的构造，它每次需要两组输入，然后输出一个 $[0, 1]$ 的数来判断真伪。我们按如下方式制作训练数据：

- (1) 所有训练集中真实签名之间的组合，其目标输出为 1
- (2) 所有训练集中真实签名与虚假签名的组合，其目标输出是 0

因为孪生网络左右子网络实际是同一个 CNN 网络，所以上述组合是无序组合，不需要再将他们交换顺序并添加到训练集。训练数据制作的示意图如图 3-6 所示，图中为了演示方便，使用的是该用户签名的图片，但实际孪生网络接受的输入是签名的特征向量。

X1	X2	Y
 用户 1-真实签名 1	 用户 1-真实签名 2	1
 用户 1-真实签名 1	 用户 1-真实签名 3	1
.....		
 用户 1-真实签名 1	 用户 1-伪造签名 1	0
 用户 1-真实签名 1	 用户 1-伪造签名 2	0
.....		

图 3-6 孪生网络的两个输入与输出示意图。

另一点值得注意的是，有文献将训练集中所有虚假签名的组合也添加进入了训练数据，并设定其目标输出是 1。本文进行多次实验后认为这种做法欠妥，因为不同伪造签名之间并不一定相似，特别是有多个伪造者伪造签名时，将它们添加进训练数据将会降低系统的最终表现。

本论文构造的孪生网络的训练目标只有 0 和 1，所以使用 Binary Crossentropy 损失函数，使用 Adam 优化器来自动调整训练速率，进行 500 轮训练。每个批次大小（Batch Size）为 128 个，并开启洗牌算法，每轮训练都将训练数据打乱后输入。

3.3 评价指标

在身份认证领域中，一个常见的衡量标准是等误率 Equal Error Rate (EER)。它能在错误接受率与错误拒绝率相同的情况下，衡量一个模型识别的好坏。

计算 EER 就需要先计算误识率 (FAR) 和拒识率 (FRR)。其中，FAR 是指在某一设定阈值的条件下，所有伪造签名被误判为真实签名的数量，占总签名数的比例。同理，FRR 是指在某一设定阈值的条件下，所有真实签名被误判为伪造签名的数量，占总签名数的比例。显然，但我们将判断的阈值从小到大移动时，这时被系统接受的签名数量将会越来越多，FAR 会不断上升，FRR 会不断下降。我们期望良好的系统 FRR、FAR 应该尽可能接近，所以我们可以寻找这两个比率变化的交点。这个交点的横坐标值就是此时的选定的阈值，而纵坐标的值就是该系统的等误率。等误率能够良好的反应该系统在 FRR、FAR 相同的情况下的表现状况。系统的等误率越低，就意味着系统的整体识别效果越好。

图 3-7 是在 SCUT-MMSIG 空中签名数据集中进行某一次实验时，FRR 和 FAR 变化图，图中绿色的交点即为此次训练得到的在线签名验证系统的等误率。系统实现过程中，只需要将每个待测签名按得分排序，然后依次计算出每种阈值情况下的 FAR、FRR，最后找到 FAR 大于 FRR 的第一个点，就是图中的等误率了。

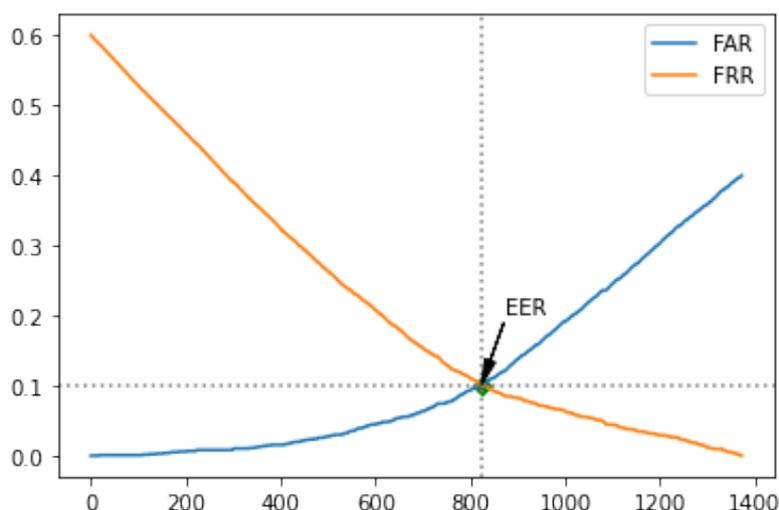


图 3-7 SCUT-MMSIG 数据集上一次实验的 FAR、FRR、EER 示意图

3.4 本章小结

本章主要介绍了实验设计的细节。首先介绍了实验所采用的数据集并列举了部分数据作为参考，详细阐述了每个数据集需要进行的数据预处理。其次介绍了 GRU 自动编码器的填充方式以及每个数据集需要填充到的长度。然后介绍了孪生网络训练数据的生成方式，包括训练数据的组合，数据标签的定义。接下来介绍了特征融合的具体实现方法。最后介绍了实验效果的衡量方式和等误率的计算方法。下一章将会介绍上述实验条件下的实验结果及分析。

第四章 实验结果

在第三章所提出的动、静态特征提取与特征融合方法的基础上，本章在不同的数据集下进行了实验，并将实验结果与前人研究的结果进行对比。同时，还对比了不同的特征融合方法对等误率的影响，分析了单个用户的等误率分布情况，衡量了本文所述方法的实际表现情况。

4.1 SUCT-MMSIG 空中签名数据集上的实验结果

本论文的实验方法在 SUCT-MMSIG 空中签名数据集上的实验结果与其它已知论文的结果对比如表 4-1 所示。其他学者的部分实验数据的选取与本论文区别较大，无法直接对比，且本实验无法在那样的实验数据条件下进行，故不在此列出。

表 4-1 SUCT-MMSIG 空中签名数据集不同方法等误率的对比

实验模型	等误率 (EER) %
本论文模型 (晚融合模型)	9.24
CHUANG LI - STROKE-BASED RNN ^[18]	13.90
房育勋-跨阶段测试 ^[12]	16.10

在本论文的实验条件中，特征融合后，SCUT-MMSIG 空中签名数据集的等误率低于该数据集目前已知的所有论文成果。其 FRR、FAR 的变化曲线如图 4-1 所示。

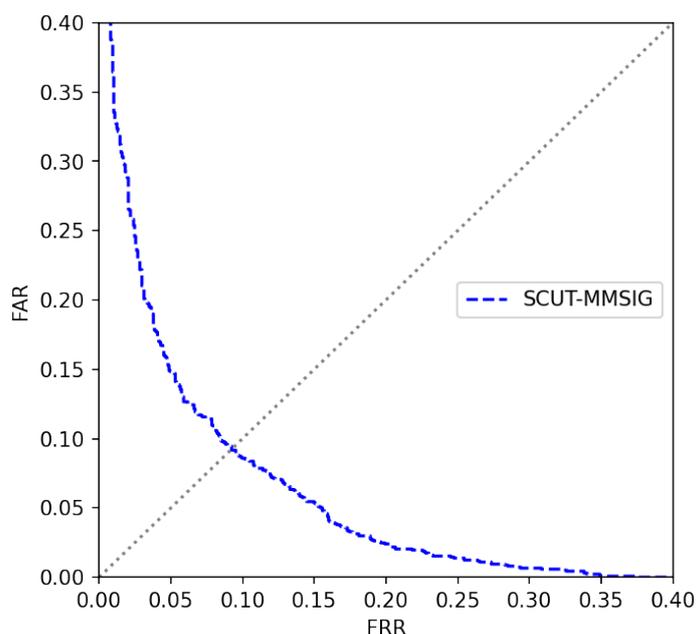


图 4-1 SUCT-MMSIG 空中签名数据集 FRR、FAR 的变化曲线

我们还可以计算出每个用户的等误率，以观察本系统在不同用户之间的均衡情况。绘制出的每个用户等误率的频次直方图，如图 4-2 所示。可以看出单个用户识别的等误率分布是比较均匀的，大多数集中在系统等误率附近，且最高用户等误率的没有偏离整个系统等误率太多，符合我们的预期。

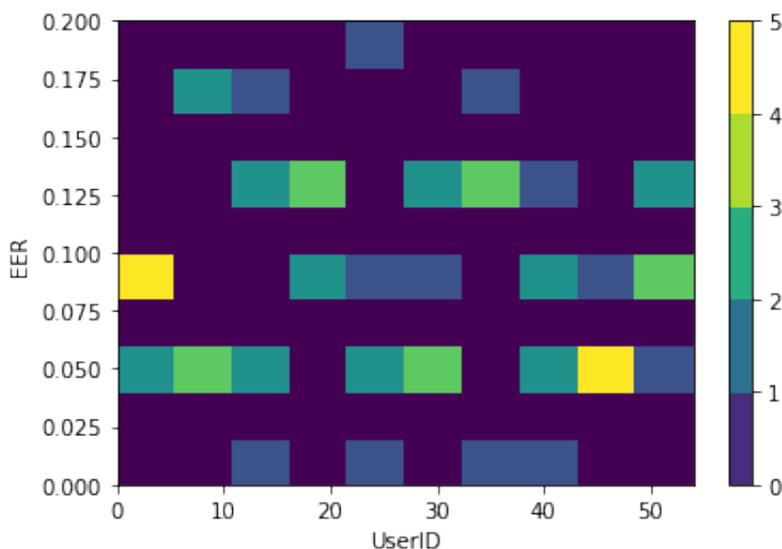


图 4-2 SUCT-MMSIG 空中签名数据集上，每个用户的等误率频次直方图

4.1.1 不同特征融合方式的对比

表 4-2 为 SCUT-MMSIG 空中签名数据集中，使用单一特征向量和使用两个特征向量、不同融合方法获得的系统等差率对比。

表 4-2 早融合与晚融合对等误率 EER 影响的对比

特征融合方式	等误率 (EER) %
仅动态特征	10.11
仅静态特征	13.52
使用早融合	12.15
使用晚融合	9.24

图 4-3 是 FRR、FAR 的变化曲线。曲线与 $y=x$ 相交的点，即为 FRR 与 FAR 相等的点，也就是 EER。由图 4-3 可知，晚融合 (Late Fuse) 的特征融合方式下，等误率最低。同时还可发现，晚融合在不同阈值之下的表现均优于其它融合方式或单一特征向量。因

此可得出结论，晚融合的方式对于降低等误率是有效的。

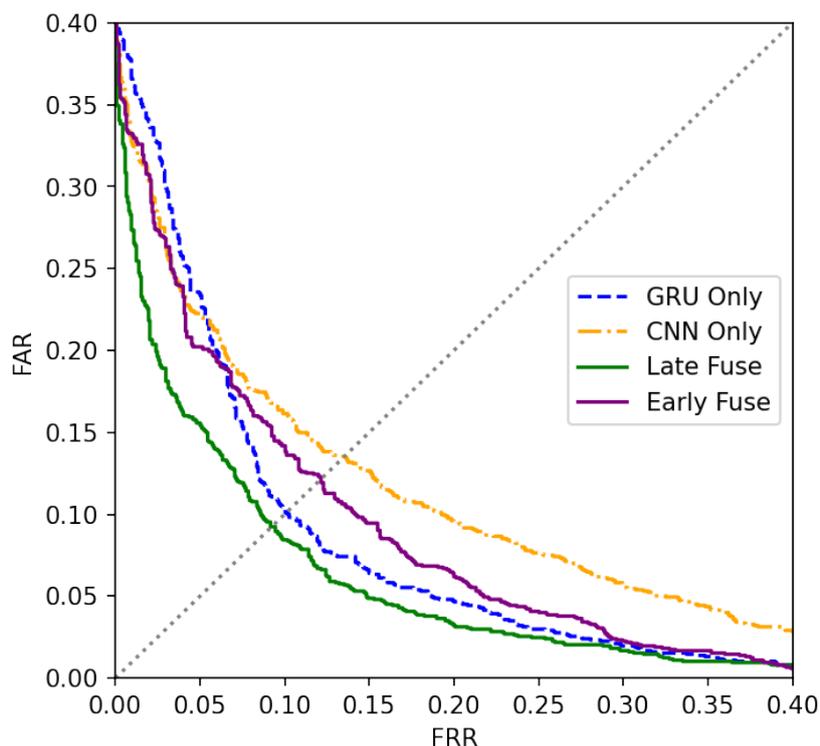


图 4-3 不同特征融合方式对 FAR、FRR 变化曲线的影响

同时我们还可以发现，使用早融合的情况下，等误率介于早融合和晚融合之间。推测原因是在早融合下，特征向量由 128 维增加到了 256 维，这会增加孪生网络的复杂度。但与此同时，训练数据十分有限无法增加，孪生网络处于欠拟合状态，无论怎样增加训练次数都无法提升准确性。

4.1.2 不同参考签名数量对等误率的影响

在签名识别领域，通常会选用 5 个真实签名来作为参考签名。过多的参考签名意味着实际使用过程中采集困难，过少的参考签名又会让签名认证系统的准确率大打折扣。本文在不同的参考签名数量下进行实验，探究参考签名数量对等误率的影响。值得注意的是，这里参考签名数量的改变仅针对测试集中的参考签名数改变。在训练集中的签名数量不变的条件下，已经训练好的神经网络可以在不同数量的参考签名下使用。

表 4-3 不同参考签名数量特征融合后的等误率对比

参考签名数量	等误率 (EER) %
1	16.27
2	15.64
3	13.60
4	12.90
5	9.24
6	13.95

由表 4-3 的等误率变化数据绘制出散点图如图 4-4 所示。

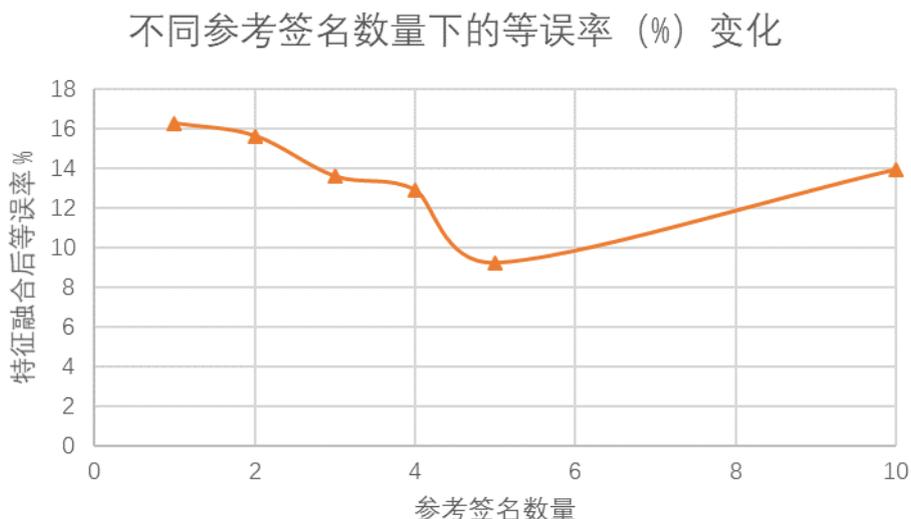


图 4-4 不同参考签名数量对应等误率变化

从图 4-4 中可以发现，在参考签名过多或过少的情况下，等误率都不是最优。如果参考签名过少，在匹配时如果恰好真实的测试签名与该参考签名的距离过大，就会被拒识，导致等误率偏高。而当参考签名过多时，误差的累积会导致假签名的得分偏高，会比较容易的和某一个真实签名匹配上。而真实签名不一定能和其它所有真实签名匹配上，即真实签名之间不一定都非常相似，因此真实签名的得分增长较少，最终就导致了假签名相似度得分高于真签名，等误率上升。

4.2 SigWiComp2013 数据集上的实验结果

表 4-4 为 SigWiComp2013 数据集上的实验结果与当年该竞赛最优的前 5 名方法的效果对比^[17]。从等误率来看，本模型仅次于竞赛中的前三名。不过后来又经过很多学者的研究，该数据集上有大量的研究成果出炉，难以将它们全部收集对比，因此本论文只与该签名验证比赛的汇总结果做了对比。

表 4-4 SigWiComp2013 日语签名数据集不同方法等误率的对比

实验模型	等误率 (EER) %
本论文模型	22.44
SABANCI UNIVERSITY - 1	9.28
SABANCI UNIVERSITY - 2	10.18
SABANCI UNIVERSITY - 3	13.05
TEBESSA UNIVERSITY - 1	23.3
TEBESSA UNIVERSITY - 2	25.41

本模型在 SigWiComp2013 数据集上的 FAR、FRR 变化曲线如图 4-5 所示。

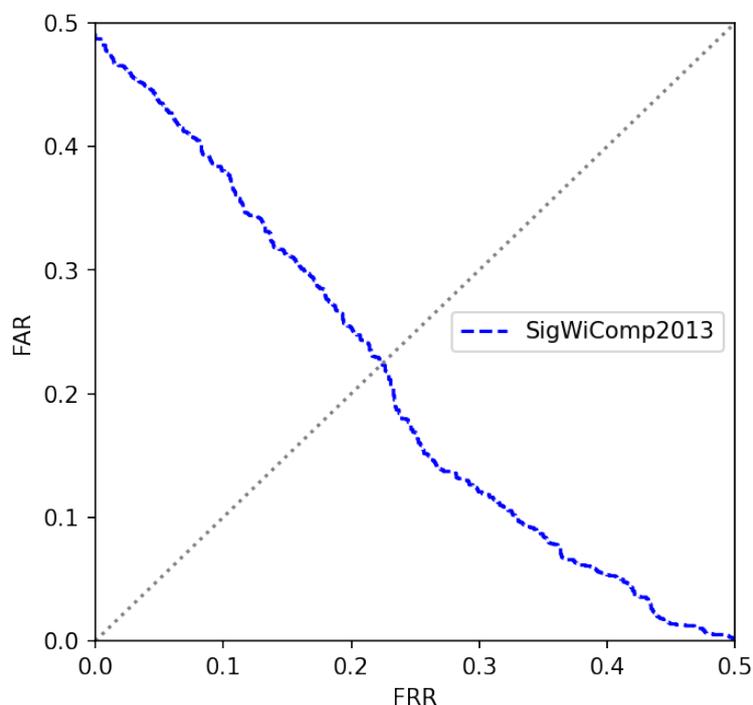


图 4-5 SigWiComp2013 数据集 FRR、FAR 曲线的变化曲线

我们同样也计算出每个用户的等误率，并绘制出的每个用户等误率的频次直方图，

如图 4-6 所示。可以观察到，少量用户的等误率极大的偏离了平均等误率，这也是在这个数据集上表现不如 MMSIG 的原因。另外 SigWiComp2013 数据集只有 11 个用户的签名数据，也会造成实验结果的不稳定。

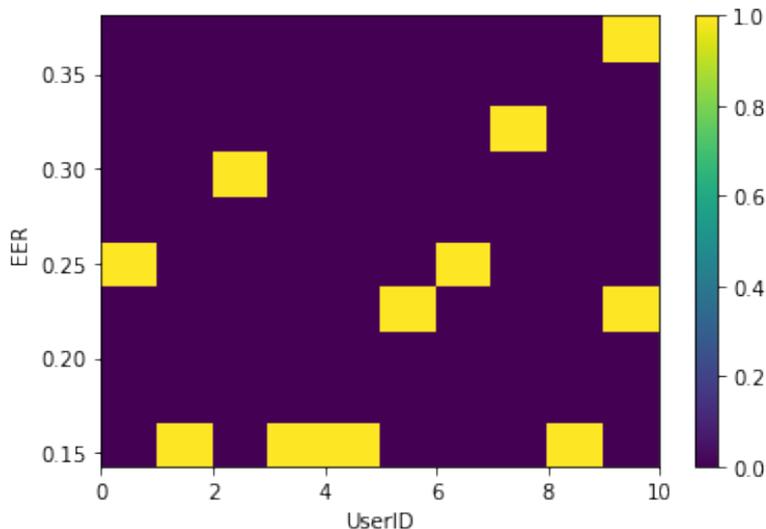


图 4-6 SigWiComp2013 数据集上，每个用户的等误率频次直方图

4.3 SVC2004 数据集上的实验结果

本论文在 SVC2004 数据集上进行了实验，等误率为 12.85%。绘制出 FRR、FAR 的变化曲线如图 4-7 所示。

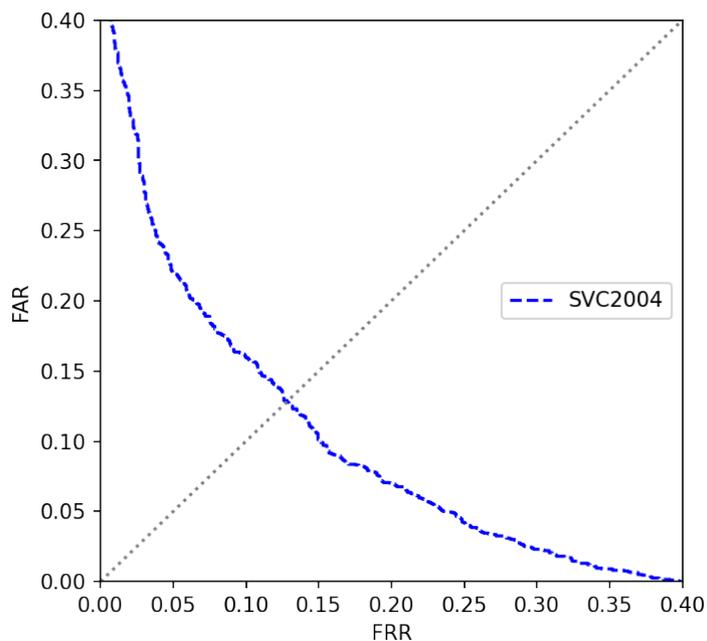


图 4-7 SVC2004 数据集 FRR、FAR 曲线的变化曲线

绘制出的每个用户等误率的频次直方图如图 4-8 所示。可以看出我们的算法在部分用户的数据上效果非常理想，甚至低于 2% 的等误率。但由于在少量用户数据上表现不佳，拉高了平均等误率。

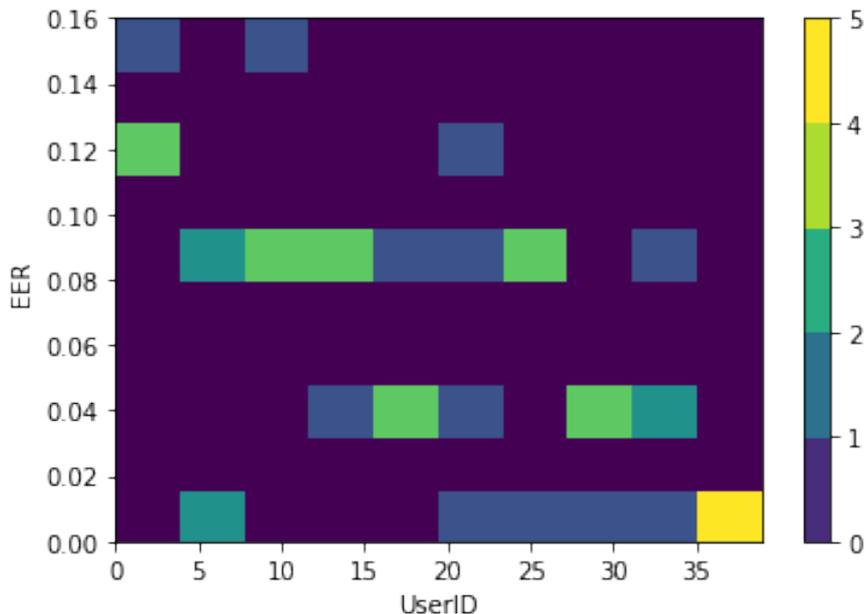


图 4-8 SVC2004 数据集上，每个用户的等误率频次直方图

4.4 MCYT-Signature-100 数据集上的实验结果

本论文在 MCYT-Signature-100 数据集上进行了实验，等误率为 15.17%。绘制出 FRR、FAR 的变化曲线如图 4-9 所示。

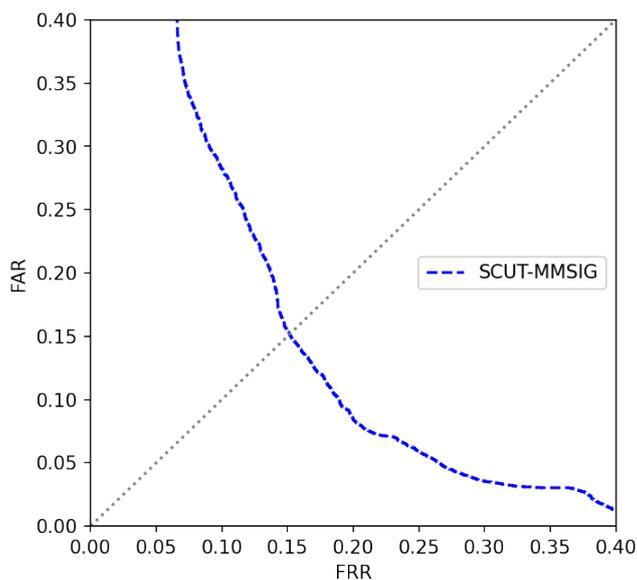


图 4-9 MCYT-Signature-100 数据集拒识率、FAR 曲线的变化曲线

绘制出的每个用户等误率的频次直方图如图 4-10 所示。可以看出该数据集下，大部分用户的签名识别效果都非常良好，等误率处于 5% 以内，只有少部分用户等误率超过了 5%。

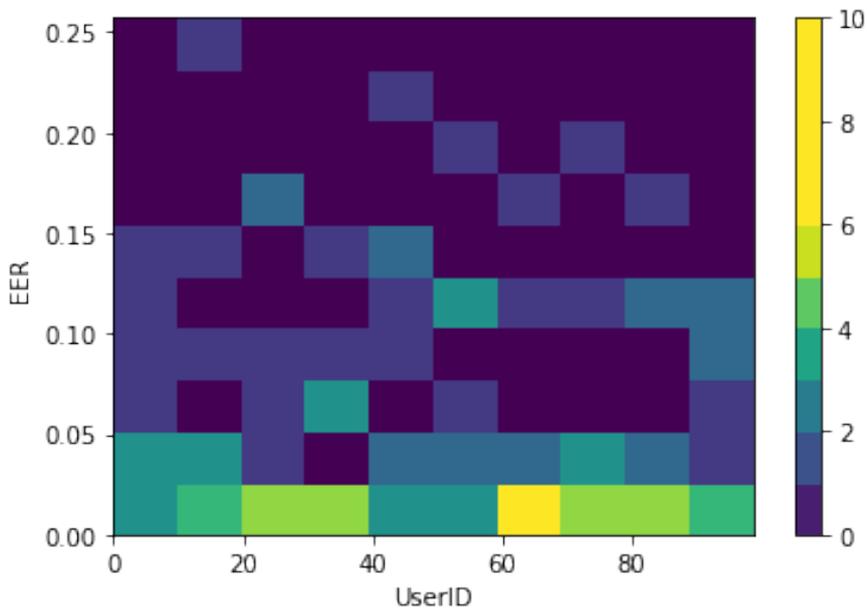


图 4-10 MCYT-Signature-100 数据集上，每个用户的等误率频次直方图

4.5 本章小结

本章中，我们根据前面章节所描述的神经网络构造和实验设置，进行了相关实验，并绘制出相应的图表、进行分析。同时还将本论文提出方法的实验结果与前人的成果作对比，在 SCUT-MMSIG 空中签名数据库上，本论文达到了低至 9.24% 的等误率，为目前该数据集上的最佳算法。最后，我们还对比了不同参考签名数量对等误率的影响，并绘制图表直观反映等误率的变化。

结论

1. 论文工作总结

本论文主要提出了一个新型的在线签名动静态特征的提取方法。

对于静态特征的提取，本论文训练了一个基于 CNN 的自动编码器，它能将输入的签名图像压缩为一个特征向量，再重新还原为原始图像。训练完成之后移除解码器部分，瓶颈层的输出即为签名的静态特征。

对于动态特征的提取，本论文训练了一个基于 GRU 的自动编码器，它能将输入的变长轨迹序列转换为一个定长的特征向量，再重新还原为原始序列。训练完成之后移除解码器部分，瓶颈层的输出即为签名的动态特征。

最后我们训练了一个基于 CNN 的孪生网络，并进行对上述的动、静态特征进行融合，并输出签名的相似度来完成签名真伪的鉴别。

本文在 SCUT-MMSIG、SigWiComp2013、SVC2004 和 MCYT-Signature-100 数据集上分别进行了实验，并与其它学者的成果作对比。其中在 SCUT-MMSIG 空中签名数据集中达到了该数据集历史最佳的效果。同时本文还对比研究了不同特征融合方式、不同参考签名数量对等误率的影响。

2. 工作展望

(1) 对于在线签名验证系统的效果评估方面，学界还亟需一个统一的标准。各个期刊文章都采用自己的衡量标准，导致各学者的成果难以横向对比。且目前还缺乏对在线签名验证方法运行效率的评判标准。

(2) 通过对比单一动、静态特征的等误率可发现，本论文所采用的静态特征提取方案，也就是基于 CNN 的自动编码器还有不少提升的空间。可能可以通过调整 CNN 的深度、过滤器数量、卷积核大小等参数来提升 CNN 对静态特征提取的能力。也可以考虑将其换为诸如 Inception Net 等 CNN 的衍生模型，从而更好的提取静态特征。

(3) 在特征融合方面，还可以探究更多的特征融合方案，并对比它们对等误率的影响。

参考文献

- [1] Vorugunti S C, Pulabaigari V, Gorthi R, et al. “OSVFuseNet: Online Signature Verification by Feature Fusion and Depth-Wise Separable Convolution Based Deep Learning.” *Neurocomputing*, vol. 409, 2020, pp. 157–172., doi:10.1016/j.neucom.2020.05.072.
- [2] Vorugunti S C, Pulabaigari V, Mukherjee P, et al. “DeepFuseOSV: Online Signature Verification Using Hybrid Feature Fusion and Depthwise Separable Convolution Neural Network Architecture.” *IET Biometrics*, vol. 9, no. 6, 2020, pp. 259–268., doi:10.1049/iet-bmt.2020.0032.
- [3] Ahrabian K, Babaali B. “Usage of Autoencoders and Siamese Networks for Online Handwritten Signature Verification.” *Neural Computing and Applications*, vol. 31, no. 12, 2018, pp. 9321–9334., doi:10.1007/s00521-018-3844-z.
- [4] Okawa M. “Template Matching Using Time-Series Averaging and DTW With Dependent Warping for Online Signature Verification.” *IEEE Access*, vol. 7, 2019, pp. 81010–81019., doi:10.1109/access.2019.2923093.
- [5] Sharma A, Sundaram S. “A Novel Online Signature Verification System Based on GMM Features in a DTW Framework”, *IEEE Transactions on Information Forensics and Security*[C], 2017: 705–718.
- [6] Vorugunti S C, Guru D S, Pulabaigari V, et al. “An Efficient Online Signature Verification Based on Feature Fusion and Interval Valued Representation of Writer Specific Features.” *2019 IEEE 5th International Conference on Identity, Security, and Behavior Analysis (ISBA)*, 2019, doi:10.1109/isba.2019.8778566.
- [7] Martens, R, Claesen L. “On-Line Signature Verification by Dynamic Time-Warping.” *Proceedings of 13th International Conference on Pattern Recognition*, 1996, doi:10.1109/icpr.1996.546791.
- [8] Barkoula K, Economou G, Fotopoulos S. “Online Signature Verification Based on Signatures Turning Angle Representation Using Longest Common Subsequence

-
- Matching.” *International Journal on Document Analysis and Recognition (IJ DAR)*, vol. 16, no. 3, 2012, pp. 261–272., doi:10.1007/s10032-012-0193-9.
- [9] Gupta G K, Joyce R C, “Using Position Extrema Points to Capture Shape in on-Line Handwritten Signature Verification.” *Pattern Recognition*, vol. 40, no. 10, 2007, pp. 2811–2817., doi:10.1016/j.patcog.2007.01.014.
- [10]Lai S, Jin L, Yang W. “Online Signature Verification Using Recurrent Neural Network and Length-Normalized Path Signature Descriptor.” 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), 2017, doi:10.1109/icdar.2017.73.
- [11]汤磊. 在线签名认证算法研究及其在智能手机上的应用[D]. 华南理工大学, 2017.
- [12]房育勋. 在线签名认证若干关键问题研究[D]. 华南理工大学, 2018.
- [13]Will B. “Auto-Encoder: What Is It? And What Is It Used For? (Part 1).” *Medium*, Towards Data Science, 1 July 2019, towardsdatascience.com/auto-encoder-what-is-it-and-what-is-it-used-for-part-1-3e5c6f017726.
- [14]Cho K, Merriënboer B, Gulcehre C. “Learning Phrase Representations Using RNN Encoder–Decoder for Statistical Machine Translation.” *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2014, doi:10.3115/v1/d14-1179.
- [15]Hochreiter S, Schmidhuber J. “Long Short-Term Memory.” *Neural Computation*, vol. 9, no. 8, 1997, pp. 1735–1780., doi:10.1162/neco.1997.9.8.1735.
- [16]Lu, X, Fang Y, Kang W. “SCUT-MMSIG: A Multimodal Online Signature Database.” *Biometric Recognition*, 2017, pp. 729–738., doi:10.1007/978-3-319-69923-3_78.
- [17]Malik M I, Liwicki M, Alewijnse L, et al. “ICDAR 2013 Competitions on Signature Verification and Writer Identification for On- and Offline Skilled Forgeries (SigWiComp 2013).” 2013 12th International Conference on Document Analysis and Recognition, 2013, doi:10.1109/icdar.2013.220.
- [18]Li C, Zhang X, Lin F, et al. “A Stroke-Based RNN for Writer-Independent Online

Signature Verification.” 2019 International Conference on Document Analysis and Recognition (ICDAR), 2019, doi:10.1109/icdar.2019.00090.

- [19]Lamba H. “One Shot Learning with Siamese Networks Using Keras.” Medium, Towards Data Science, 17 Feb. 2019, towardsdatascience.com/one-shot-learning-with-siamese-networks-using-keras-17f34e75bb3d.
- [20]Yeung D Y, Chang H, Xiong Y, et al. “SVC2004: First International Signature Verification Competition.” Biometric Authentication, 2004, pp. 16–22., doi:10.1007/978-3-540-25948-0_3.
- [21]Ortega-Garcia J, Fierrez J, Simon D, et al. “MCYT Baseline Corpus: a Bimodal Biometric Database.” IEE Proceedings - Vision, Image, and Signal Processing, vol. 150, no. 6, 2003, p. 395., doi:10.1049/ip-vis:20031078.

致谢

回首大学四年时光，有过迷茫的夜晚，也有过努力后的雨过天晴。很幸运遇到了我的导师和各位朋友，在我困惑时不吝指点，在我低谷时伸出援手。

首先要感谢在签名验证领域默默深耕的前辈们，是你们的研究基础让我站在巨人的肩膀上看世界。

然后要感谢我的导师吴秋霞老师，她为我指明了研究的方向，关注我的研究进展，并为我提出了很多建设性的建议。我还要感谢苏欣佩学姐，为我解答了很多疑惑，并帮我梳理逻辑、调整文章结构。

最后，我还要感谢所有在我大学四年里为我默默付出过的老师、同学和我的父母，是你们给了我精神和生活上的支持，没有你们就没有今天的我。