

基于动态结构相似度的多源指纹配准方法

王海霞^{1,2)}

颜诗琪^{1,2)}

于 洋³⁾

陈 朋^{1,2)}

梁荣华^{1,2),3)}

张怡龙^{1,2)}

刘义鹏^{1,2)}

¹⁾(浙江工业大学计算机科学与技术学院 杭州 310023)

²⁾(浙江工业大学软件学院 杭州 310023)

³⁾(浙江科技大学信息与电子工程学院 杭州 310023)

摘要 指纹作为最普遍的身份标识之一,一般采集自手指表皮。近年来,成像技术的发展成功实现了皮下内部指纹的获取。通常,外部指纹的图像质量较高,而内部指纹完整性较好,内外指纹之间具有互补的特性。将外部指纹和内部指纹图像中的有效信息提取并融合,可以生成信息量更完整、对后续应用更有益的新的指纹表达。由于采集方式的差异,内外指纹之间存在相对畸变,因此,在融合前必须先进行指纹的配准,恢复其在空间上的一致性。然而,内外指纹之间的质量差异较大,指纹内部重复模式较多,加上融合对高配准精度的追求,使内外指纹的配准成为一项具有挑战性的任务。本文提出了一个无监督密集配准网络 U-NeXt,它具有较好的全局信息捕获能力和对细节纹理等高频信息的感知能力,能够减少局部误配准的发生,提高配准精度,实现内外指纹之间的亚像素级变形场估计。此外,针对传统相似性度量面对指纹图像时缺乏语义结构理解、难以准确评估指纹图像间复杂关系的问题,本文还设计了一个动态结构相似度损失函数,能够辅助网络挖掘更深层次的潜在语义关联,提高配准的准确性。本文利用同步采集系统获取的内外指纹公开数据集进行了一系列实验与评估,与配准前相比,所提出的方法使内外指纹的平均结构相似度提升了 30.31%,使融合指纹的 EER(Equal Error Rate, EER)降低了 74.96%,并且在相对畸变较严重和内外指纹质量差异较大的情况下依然有较好的表现,为中下游任务奠定了良好的基础。

关键词 外部指纹;内部指纹;光学相干断层扫描;密集配准;深度学习;动态损失

中图法分类号 TP391 **DOI号** 10.11897/SP.J.1016.2025.01812

Multi-Source Fingerprint Registration Method Based on Dynamic Structural Similarity

WANG Hai-Xia^{1,2)} YAN Shi-Qi^{1,2)} YU Yang³⁾ CHEN Peng^{1,2)} LIANG Rong-Hua^{1,2),3)}

ZHANG Yi-Long^{1,2)} LIU Yi-Peng^{1,2)}

¹⁾(College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023)

²⁾(College of Software, Zhejiang University of Technology, Hangzhou 310023)

³⁾(School of Information and Electronic Engineering, Zhejiang University of Science & Technology, Hangzhou 310023)

Abstract Fingerprints have become one of the most widespread biometric identifiers due to their universality, uniqueness, and permanence. Traditionally, fingerprint images are captured from the surface of the finger. However, with advancements in imaging technologies in recent years, it has become possible to obtain subcutaneous internal fingerprints, which are located beneath the

收稿日期:2024-08-29;在线发布日期:2025-05-20。本课题得到浙江省自然科学基金(LR24F030003)、国家自然科学基金面上项目(62376250, 62276236, 62076220)、浙江省领军创新团队(2021R01002)资助。王海霞(通信作者),博士,教授,中国计算机学会(CCF)会员,主要研究领域为生物特征感知与识别、计算光学。E-mail:hxwang@zjut.edu.cn。颜诗琪,硕士研究生,主要研究领域为生物特征识别与图像处理。于 洋,博士,讲师,主要研究领域为生物特征识别与图像处理。陈 朋,博士,教授,主要研究领域为计算机视觉与信号处理。梁荣华,博士,教授,主要研究领域为多模态数据的感知与理解。张怡龙,博士,副研究员,中国计算机学会(CCF)会员,主要研究领域为生物医学光学成像与光学生物传感。刘义鹏,博士,副教授,中国计算机学会(CCF)会员,主要研究领域为生物特征模式识别与医学图像处理。

outer skin layer. Typically, external fingerprint images exhibit high clarity and contrast but are less stable, being susceptible to external environmental factors and the condition of the finger. In contrast, internal fingerprints show better integrity and stability, though their image quality is lower than that of external fingerprints. These two types of fingerprints possess complementary characteristics. By extracting and fusing the useful information from both external and internal fingerprint images, it is possible to create a new fingerprint representation that is more comprehensive and beneficial for subsequent applications. Owing to the distinct characteristics of acquisition devices employed, there are relative deformations between internal and external fingerprints. The actual physical locations represented by the same coordinates in the images differ. Therefore, before fusion, it is essential to perform registration to restore spatial consistency between the two types of fingerprints. However, the registration process becomes challenging because of the significant differences in image quality between external and internal fingerprints, as well as the presence of repetitive ridge patterns within the fingerprints. Moreover, the pursuit of high registration accuracy required for successful fusion further complicates the registration process. This paper proposes an unsupervised dense registration network, U-NeXt, which is designed to address the challenges of fingerprint registration. Building on the traditional U-Net architecture, U-NeXt replaces the conventional convolutional layers with ConvNeXt as the encoder to capture the spatial correspondences between the target and reference images, which enables the network to better capture global information and mitigate the disadvantages caused by repetitive texture patterns. Additionally, the network incorporates a high-frequency feature enhancement convolution module (HFFEC) within the skip connections to enhance the network's ability to perceive fine-grained fingerprint texture details and improve the registration accuracy of local features. Furthermore, traditional similarity metrics often lack the capability to understand the semantic structures of fingerprint images and struggle to accurately assess the complex relationships between them. To address this limitation, this paper introduces a dynamic structural similarity loss function based on seed points. This loss function assists the network in uncovering deeper semantic relationships and improves the accuracy of registration. A series of experiments and evaluations were conducted using a publicly available dataset of paired internal and external fingerprints obtained through a synchronized acquisition system. The proposed method was compared with several existing fingerprint registration methods. The results show that, the proposed method reduced the structural similarity between internal and external fingerprints by 30.31% on average after registration. Moreover, the Equal Error Rate (EER) of the fused fingerprints was reduced by 74.96%, representing the largest improvement among all the comparison methods. Notably, even under conditions with significant distortions or substantial quality differences between the internal and external fingerprints, the proposed method still performed well, providing a solid data foundation for downstream tasks.

Keywords external fingerprint; internal fingerprint; optical coherence tomography; dense registration; deep learning; dynamic loss

1 引言

指纹被誉为“人体身份证”，早在中国古代的秦朝时期，就已有与其相关的记载。指纹作为身份标

识具备普遍性、唯一性和永久性等特点^[1]，被广泛应用于身份认证、执法和金融交易等领域。传统的指纹采集设备如光学器件^[2]、电容器件^[3]和射频器件^[4]等，采集外部表皮指纹用于识别，速度快且成本低，使外部指纹成为多数应用场景中的首选。外部指纹

是指人类手指末端指腹上由凹凸皮肤形成的纹路,其图像通常具有较大的有效面积和高对比度,但容易受到指尖状况和外部环境的影响^[5]。例如,指尖的干湿情况、表皮的磨损与划痕,以及手指或采集设备上的污渍等,这些都会导致外部指纹图像有效信息的缺失和污染。

外部指纹结构来源于表皮和真皮交界处的活性表皮层,该层是外部指纹的母版,从中可以重建内部指纹^[6]。光学相干断层扫描技术(Optical Coherence Tomography, OCT)是一种基于低相干干涉测量原理的非破坏性、高分辨率、无创成像技术。利用OCT,能够获取指尖的三维体数据^[7],从这些数据中重建出内部指纹。内部指纹不易受到外部磨损、划痕和污染的影响,适应环境能力强,且不易被指纹膜等伪造^[5]。然而,已有研究表明,内部指纹受采集方式和提取算法的影响,其图像通常具有散斑噪声和较低的对比度^[7-8],因此成像质量普遍低于外部指纹。综合来看,内外指纹之间具有互补的优势,将外部指纹和内部指纹图像中的有效信息提取并融合,可以生成信息量更完整、对后续应用更有益的新的指纹表达。

与其它融合任务^{[9][10]}相同,当指纹图像中的相同物理位置在内外指纹中都具有相同的像素坐标时,融合指纹才是具有应用意义的。由于目前人们主要采用接触式手段采集指纹,在按压过程中,指纹图像会产生非线性形变^[11],且会受不同按压姿势和力度的影响,因此不同时刻采集的指纹之间存在较大的相对畸变,这给融合工作带来了不小的阻碍。最近,Sun等人^[8]提出了一种内外指纹同步采集系统,它集成了全内反射(Total Internal Reflection, TIR)系统和OCT系统,可以同时获取指尖同一区域的外部和内部指纹,相比传统多次采集的方式,具有相对一致的采集面积与接触形变。然而,两个系统在采集上存在明显差异,其一是采集系统的畸变不同,其二是OCT相较于TIR具有更长的采集时间。前者引入畸变矫正误差,而后者在采集时间上的较大差异间接导致了指纹之间接触形变的微小不同。因此,需要通过指纹配准消除内外指纹之间的相对畸变,使指纹重新处于同一坐标系下。由于多源指纹之间质量与风格差异较大,而指纹本身是一种纹理细密且模式重复率高的图像,指纹配准面临诸多挑战,包括匹配区域之间的类内差异大、非匹配区域之间的类间差异小以及搜索空间大等^[12]。考虑到成本效益比,大部分场景对指纹配准精度的

要求并不严格,例如在指纹的识别与匹配中,所采用的指纹特征具有较强的鲁棒性,并且算法本身存在容错率,可以容忍指纹之间存在一定程度的相对畸变。而融合对相对畸变较为敏感,因此对配准的精度有更高的要求,也增加了配准的难度。

基于稀疏特征的配准方法主要以细节点作为配准特征^[13-15],依赖于细节点提取算法。然而,对于一些质量较差的指纹,能够提取到的细节点数量较少,并且存在分布不均匀的问题。这些方法往往只能获取较为稀疏的形变参数量,导致配准精度较低。相比之下,密集指纹配准^[16]旨在寻找两张指纹的同位点,获取具有非线性皮肤失真的两个指纹之间的像素级位移度量,而不仅仅是稀疏的位移度量。一种基于相位的配准算法^[17]通过计算两枚指纹之间的相位差来重建畸变场,可以有效地消除指纹间的相对畸变,但其只能获得垂直于脊方向的小于一个脊周期的位移。另一种基于图像域的非刚性指纹配准算法^[18]将图像场与传统的基于模型的算法相结合,直接将指纹脊线方向引入模型中进行变形失真估计,从而更好地利用指纹的方向信息,简化变形模型,但难以解决较大的局部配准误差。一种基于快速数字图像相关性和异常值拒绝的两步配准^[19]方法考虑了内外指纹的融合,配准后的指纹在融合中有着较好的表现,但由于缺乏对指纹语义信息的利用,局部配准表现欠佳。

近年来,深度学习在图像配准领域引起了广泛关注。相比传统的方法,深度学习模型不依赖于个体对特征的定义,可以直接学习图像之间的端到端变换。Cui等人提出的基于监督学习的方法^[12,20]表现出更快的速度,但难以获得包含足够多样皮肤变形类型的大型且多样化的训练集,导致泛化能力较差,配准结果和匹配分数不够理想。Guan等人^[21]提出了聚合相位特征的双分支稠密配准网络,该网络引入了一种双分支结构,分别提取高分辨率的相位特征和低分辨率的纹理特征,并在多个阶段进行信息交互,以综合利用两者的优势。对于密集配准而言,监督学习的方法需要获取到图像之间的密集变形场,然而具有变形场真值的指纹数据集较为稀少,而手工标注准确的变形场是极其困难的。无监督学习是解决缺少标注数据的重要方法,已广泛应用于光流估计和医学图像配准等类似任务。Yu等人提出了将空间变换网络与卷积网络相结合的方法^[22],以指纹图像之间的相似性代替大量真实变形场数据。然而,该方法通过将网络输出的稀疏变形场转换为密

集变形场对指纹进行配准,因此精度有限,甚至存在加剧形变的风险。Jia 等人提出了一种无监督指纹密集配准框架^[23],首次将薄板样条(Thin Plate Spline, TPS)应用于指纹配准网络,利用定位网络获得关键点实现配准,并通过递归训练获得最好的结果,虽然多次迭代提高了精度,但是该方法仍是基于稀疏的特征点实现的。

纵观现有基于深度学习的研究,不难发现,实现高精度的密集指纹配准面临一个矛盾。监督学习方法由于具备真实的变形场作为参考,能够提供绝对准确的范本指导,其预测精度在很大程度上取决于训练样本的数量。然而,高质量的精细配准数据集极难获取,标注成本极高,从而限制了监督方法的实际应用。无监督方法无需依赖真实变形场,而是利用相关性或相似度量作为指导信息,但在实际应用中容易受到匹配区域内较大类内差异及非匹配区域间较小类间差异的干扰,导致错误的配准结果。为了缓解这一问题,以往一些无监督指纹配准方法未直接利用神经网络预测密集变形场,而是采用稀疏参数作为中间表示,并通过插值得到密集形变参数,以避免非匹配区域的强制拟合对指纹匹配造成的不利影响。然而,这种策略在一定程度上限制了配准精度的上限。如何在不依赖真实变形场的情况下,实现稳定且高精度的指纹配准,仍然是一个亟待解决的挑战。

针对上述的问题,本文提出了一种面向融合的基于无监督学习的内外指纹配准方法,旨在消除以不同方式采集的多源指纹之间的相对畸变,确保在图像空间内,代表实际物理空间中相同位置的像素具有一致的坐标值,实现多源指纹之间的高精度配准。该方法主要利用一个密集配准网络估计从待配准指纹变换到参考指纹的变形场,其结果能够达到亚像素精度。基于指纹的特征分布,本文设计了一个可以在网络训练时阶段性变化的高斯动态结构相似度损失函数,以代替变形场真值在网络中的作用,弥补指导信息的缺失,降低局部误配准的发生率。本文的主要贡献如下:

(1)提出了一种具有全局特征捕获能力的无监督指纹密集配准网络模型 U-NeXt。该模型在传统 U-Net 的基础上,以 ConvNeXt 作为编码器来捕获待配准图像和参考图像之间的空间对应关系,能够较好地捕获全局信息,降低重复模式带来的弊端;设计了一种高频特征增强卷积模块 HFFEC 并加入到跳跃连接中,增强网络对指纹纹理细节等高频信息

的感知,以提高局部细节的配准精度。

(2)设计了一种基于种子点的动态结构相似度损失。在传统结构相似度损失中引入从细节点筛选而来的种子点,并使损失函数能够在训练过程中基于种子点动态变化,为网络的学习提供辅助性的语义信息。动态结构相似度损失能够有效减少由于粗配准精度较低、局部纹理模式重复及局部质量差异造成的局部误配准。

本文的其余部分安排如下。第 2 节回顾了相关工作;第 3 节阐述了我们所提出的内外指纹配准方法;第 4 节展示了实验评估和与现有方法的对比;最后,第 5 节总结了我们的工作。

2 相关工作

2.1 内外指纹的同步采集

不同时刻采集得到的指纹通常存在较大的几何失真,给配准工作增加了难度。此外,对于接触式采集而言,不同时刻手指在传感器上的接触区域不同,这可能造成采集到的指纹重叠面积较小,信息可交互量下降,不利于融合的开展。因此,一个能够同步采集到内外指纹的系统可以在硬件层面上为后续工作提供较大的帮助。

外部指纹的采集方式已经发展得较为成熟,光学器件^[2]、电容器件^[3]和射频器件^[4]等都可以将指尖表皮的表面信息捕获到二维图像上,并且采集速度快,图像对比度高。其中,TIR 技术是一种强大的光学技术,可以基于两种不同折射率介质界面中发生的光学现象,对物体表面进行成像,因此常用于外部指纹的采集。

OCT 是近年来备受关注的新型指纹扫描方法^[24],目前常见于医学成像和身份识别领域,例如视网膜成像^[25]和内部指纹成像^[26]等。OCT 技术可以分为三类,分别是时域 OCT (TD-OCT)、频域 OCT (FD-OCT) 和全场 OCT (FF-OCT)。其中,综合考虑成像速度、信噪比、成像质量和穿透深度,FD-OCT 是 3D 指纹成像的最佳选择。根据传感器的不同,FD-OCT 又可以进一步分为扫频 OCT (SS-OCT) 和光谱域 OCT (SD-OCT),SD-OCT 由于具有更高的轴向分辨率、光源简单以及成本低等优点,更适合用于指纹成像^[27]。从 OCT 捕获到的指尖 3D 体数据中可以提取到活性表皮层的轮廓,即内部指纹,现已存在一些从体数据中获取内部指纹的研究。刘等人^[28]提出了一种基于深度学习的亚表

面指纹重构方法,以HDCRes-34为基础网络将每个OCT横截面图像分割成三层,比较三个层次亚表面指纹信息,针对不同指纹特征,使用不同权重进行融合,得到高质量的亚表面指纹图像。Ding等人^[29]提出了一个BCLU-Net网络模型,它通过残差学习、双向卷积长期记忆和混合扩展卷积实现OCT体数据分割和内部指纹重建。

Sun等人^[8]实现了一种集成了TIR和OCT的内外部指纹同步采集系统,其结构如图1所示。其中,OCT部分采用了基于光纤的高速SD-OCT的设计。该系统利用一个梯形棱镜来结合TIR和OCT光路,以实现同步采集和相同的扫描区域。在OCT光路中,系统采用了小光斑光纤准直器和长焦透镜来确保足够的可获取面积,并利用离轴但不失焦的测量设置^[30]来减少棱镜表面的强反射。OCT子系统通过线扫描获取皮下体数据,因此采集速度慢于TIR子系统,TIR采集一张指纹的时间通常小于1秒钟,而OCT采集一个体数据需要60秒以上。由于内部指纹采集时间慢于外部指纹,为了减轻内部指纹采集过程中由于手指的运动产生的形变,系统利用快速线扫描相机减少内部指纹采集时间,并采取接触式采集的方式减少手指运动。此外,两条光路的采集特性导致了不同类型的畸变,系统利用稀疏网格校准板进行了初步的校正。指纹的同步获取为获得高质量融合指纹提供了新的研究机会,减小了配准算法的压力。

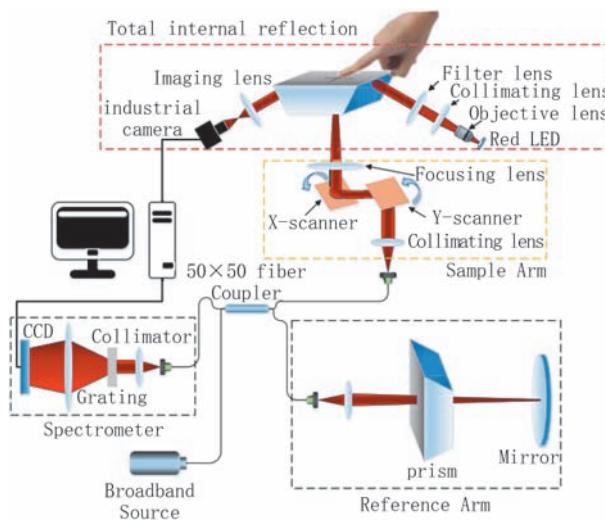


图1 指纹同步采集系统构成^[8]

2.2 指纹配准方法

指纹配准通常是指纹匹配、识别等任务的前置步骤,旨在提高下游结果的准确率。面对不同的需

求,指纹配准也具有多种形式,以参与计算的形变参数量进行划分,可将其分为稀疏配准^[13-15]和密集配准^[16]两类。其中,稀疏配准的精度有限但计算效率较高,目前大多指纹的下游研究对变形失真具有一定的容忍度,因此稀疏配准得到了较多的应用。刚性配准是一种最简单的稀疏配准,只需要估计平移和旋转两个参数量。Tico等人^[31]利用细节点的方向场推导细节点之间相似性函数实现了两个指纹之间的刚性空间变换。Yin等人^[32]在分析传统的指纹姿态和指纹奇异点提取算法的基础上,设计了一个统一的深度网络进行指纹姿态与奇异点的联合提取以获取指纹的中心位置和角度,从而实现指纹的姿态校正。若要消除更复杂的形变,例如非线性形变,则需要估计更多的形变参数。一些研究^[13-15]利用TPS函数^[33]对指纹的弹性变形进行建模,以一组匹配细节点作为控制点拟合平滑的变换函数,当细节点匹配准确时,非线性变形模型比刚性模型的对齐精度更高。但是,有多种原因会导致基于匹配细节点的TPS或其他非线性配准方法难以准确估计实际的变形场,例如细节点位置不够精确、细节点对错误匹配、匹配点数量不足以及匹配点的分布不均匀等。

密集配准通常利用像素间的相似信息进行匹配,将图像中的所有点都纳入搜索范围,而不是依靠较为稀疏的细节点建立对应关系,从而实现较高精度的配准。传统的密集配准算法针对密集的特征表示(灰度图像、脊线二值图像、相位图),估计密集的变形场。Si等人^[16]提出了一种基于双分辨率块的配准方法,利用图像相关性确定输入块与参考块之间的候选变换集,从低分辨率到高分辨率逐步细化,再通过一个区域增长风格的算法来最小化能量函数,实现全局优化,获得最佳的配准结果。Cui等人^[17]提出了一种基于相位调制解调机制的密集配准方法,将指纹图像在局域上视为二维的正弦波图像,计算两张图像的相位场,通过求解相位差推算出变形场。Yu等人^[19]提出了一种基于快速数字图像相关性(F-DIC)和异常值拒绝(OR)的两步配准,首先通过快速傅里叶变换快速估算出整数像素位移,然后利用逆复合高斯-牛顿算法精细化到亚像素级位移,在F-DIC基础上,采用基于几何约束的策略来识别并纠正位移场中的异常值,实现从全局到局部的精确对齐。这些密集配准方法摆脱了对细节点的依赖,提高了精度的上限,但是它们的算法通常较为复杂,需要耗费大量的时间及计算资源,并且具有算法本身的局限性。

2.3 基于深度学习的图像配准方法

近年来,深度学习技术的引入为图像配准带来了革命性的进步^[34],其卓越的自动特征提取能力、强大的数据处理能力以及在多模态图像融合中的高效性能使得深度学习逐渐成为图像配准的主流方法之一。其中,医学图像配准在所有配准任务中最具有代表性,其挑战性与复杂性推动了深度学习在图像配准领域的研究。Yang等人^[35]使用全卷积网络(FCN)预测2D/3D大脑体数据的配准变形场,该方法采用类似U-Net的架构,并基于大型微分度量映射,使用图像体像素的初始动量值作为网络输入,通过演化这些值以获得变形场估计。Uzunova等人^[36]利用统计外观模型(SAMs)生成真值数据,并使用卷积神经网络(CNN)来估计2D大脑MR和2D心脏MR的配准变形场,同时调整FlowNet^[37]实现端到端配准。尽管监督学习方法表现良好,但真值的获取始终是一个难题,且配准质量依赖于真值的可靠性,这促使了部分监督和无监督方法的发展。Balakrishnan等人^[38]提出了一种基于U-Net的无监督医学图像配准框架Voxelmorph,该方法用相似度度量代替真值,通过U-Net估计变形场,并利用空间变换网络(Spatial Transformer Networks, STN)^[39]完成CT图像的配准。Wang等人^[40]提出了一种递归可变形金字塔网络,用于无监督的脑部MRI图像配准,充分利用传统金字塔结构和递归策略的优势,实现从粗到细预测变形场,并在训练时采用归一化互相关评估相似性。目前同样存在一些基于深度学习的指纹配准研究,本文重点关注其中的可变形配准方法。Cui等人首次将深度学习应用于指纹的密集配准^[12,20],提出了一种多尺度的指纹融合方法。首先,他们利用基于细节点的配准方法进行初步配准。在[20]一文中,他们采用典型的孪生网络作为

回归器,预测输入指纹切片之间的位移场;而在[12]一文中,他们通过U-Net结构的编码-解码器估计整张指纹的位移场。Guan等人^[21]提出了一种聚合相位特征的双分支的稠密配准网络,引入了双分支结构,通过相位特征与纹理特征的跨分辨率融合,结合传统方法与深度学习优势实现指纹密集配准。这几种网络都采用监督学习方式进行训练,因此需要大量的位移场真值。为了解决真值缺乏的问题, Yu等人^[22]首次提出了一种基于无监督学习的指纹配准网络。该方法在卷积网络中加入STN,以相似度及平滑正则项代替真值,估计多尺度的稀疏形变参数,再通过插值得到变形后的指纹图像。Jia等人^[23]提出了一种无监督指纹密集配准框架,利用定位网络从特征图或原图中提取稀疏关键点,再利用TPS对指纹进行插值变换,通过多次递归变换后的指纹,可能得到不同的关键点,从而使稀疏关键点累积为密集关键点,更好地模拟真实变形场。无监督学习虽然摆脱了对真值的依赖,但是其大多以图像相关性或相似度作为损失函数,容易受到指纹间较大的质量差异与指纹内部的高自相似性等因素的干扰,导致局部区域的误配准。

3 面向融合的内外指纹配准方法

本文提出方法的整体流程如图2所示。为了减轻内外指纹图像之间存在的显著风格差异,在配准前,首先利用风格迁移网络^[19]对内外指纹图像的风格进行统一。然后,利用基于细节点的传统方法实现对两组指纹图像的粗略配准,初步完成内外指纹图像的对齐。最后,通过一个端到端的配准网络U-NeXt估计内部指纹到外部指纹的亚像素级变形场,实现从内部指纹到外部指纹的像素坐标映射。



图2 内外指纹配准流程示意图

3.1 数据预处理

由于TIR和OCT之间的采集方式不同,内外指纹呈现出两种显著不同的图像风格。外部指纹通常具有较高的对比度,脊线较细且纹理清晰,而内部

指纹则受噪声影响较大,对比度较低,脊线较粗且连续性较差。在无真实变形场可供参考的情况下,图像相似度成为评估配准质量的核心标准。然而,由于内外指纹风格的差异,对应区域之间的相似度降

低,从而增加了误配准的风险。此外,风格差异也对指纹融合产生不利影响,因此,消除内外指纹之间的风格差异是提高配准精度的一个关键。

鉴于外部指纹的平均质量高于内部指纹,本文采用风格迁移的方式,使内部指纹的图像风格更接近外部指纹。具体而言,本文使用Yu等人提出的基于结构约束的CycleGAN网络^[19]进行指纹风格迁移,该方法能够在保持原始脊谷结构的同时,实现多传感器指纹图像风格的统一。如图2所示,经过风格迁移后,内部指纹的脊谷纹理与细节构成保持不变,而整体风格更接近外部指纹。

原始内外指纹之间的相对畸变通常是由采集系统固有的形变以及指尖运动引起的非线性形变共同作用的结果,若直接进行密集配准,则搜索空间较大,计算成本高,且易陷入局部最优解。因此,通常采用先粗后细的两阶段策略,其中粗配准负责实现大致对齐,以减少后续密集配准的计算负担。

粗配准通常基于稀疏特征点匹配,对噪声和局部细节的敏感度较低,能够稳定且高效地完成大范围的初步对齐。本文采用与[20]相似的步骤。首先,利用VeriFinger11.1SDK^[41]提取内外指纹的细节点,并采用细节点圆柱编码(Minutia Cylinder-Code,MCC)^[42]计算所有细节点对之间的相似度得分,通过得分高低筛选出一批可信度较高的细节点对。接下来,利用AdaLAM算法^[43]过滤外点,留下最有可能匹配的细节点对。最后,以匹配细节点对作为控制点计算TPS模型,实现粗配准。在理想情况下,匹配的细节点对数量充足且分布均匀,则粗配准能够较好地初步对齐内外指纹,从而减少后续精配准的计算压力。然而,该方法对细节点匹配的依赖性较强,如果匹配细节点的数量较少或分布不均匀,则难以实现有效的初步对齐,甚至可能加重相对畸变。基于同步采集系统获取的内外指纹,尽管仍然存在一定的相对畸变,但相比传统接触式指纹采集设备,其形变程度相对较小。对于接触形变较为轻微的指纹,粗配准时仅需校正不同采集系统之间的刚性位移。因此,根据匹配细节点对的数量和分布情况以及相对畸变的程度,本文提出了两种粗配准方案:(1)匹配细节点对数量充足且分布均匀时,直接采用匹配的细节点对作为控制点,计算TPS变形模型,完成粗配准;(2)匹配细节点对数量较少且分布集中时,计算匹配细节点对的平均位移和平均旋转量,对待配准指纹进行刚性变换以避免过拟合或错误匹配导致的形变加剧。以内部指纹为例,粗

配准阶段的公式化表达如下:

$$\begin{aligned}\bar{M} &= \text{AdaLAM}(\text{MCC}(M)) \\ s &= \text{BB}(\bar{M}) \\ r &= \text{NND}(\bar{M}) \\ f_{OCT} &= \begin{cases} \text{TPS}(f_{OCT}^0, \bar{M}), & \text{if } n \geq N \text{ and } s \geq S \text{ and } r \geq R \\ \text{RT}(f_{OCT}^0, \bar{M}), & \text{if } n < N \text{ or } s < S \text{ or } r < R \end{cases}\end{aligned}\quad (1)$$

其中, M 代表由 VeriFinger 提取的细节点对, \bar{M} 代表经过 MCC 排序筛选和 AdaLAM 过滤外点后的细节点对, n 代表 \bar{M} 中细节点对的个数; $BB(\cdot)$ 代表最小外接矩形(Bounding Box)计算, 可以粗略估计 \bar{M} 在图像上的分布面积 s ; $NND(\cdot)$ 代表最邻近距离(Nearest Neighbor Distance)计算, 输出 \bar{M} 中所有点与之最接近的点的距离的平均值 r , 可以大致判断点的分布情况。 f_{OCT}^0 代表风格转换后的内部指纹图像, $TPS(\cdot)$ 代表薄板样条插值函数, $RT(\cdot)$ 代表刚性变换(Rigid Transformation), 而 N, S, R 分别代表细节点对数量、最小外接矩形面积和平均最邻近距离的阈值。只有 n, s, r 都超过阈值, 指纹图像才会基于 TPS 进行粗配准。由于本文以内部指纹作为待配准指纹, 因此在计算 s 与 r 的过程中, 只考虑内部指纹一侧的细节点。

3.2 基于深度学习的密集配准

粗配准在全局范围内对指纹形变进行了初步校正, 实现了内外指纹的大致对齐。然而, 由于其精度有限, 仍需进一步进行精配准, 以消除局部区域的细微形变, 确保指纹的高精度对齐。在精配准阶段, 核心任务是利用神经网络估计指纹之间的密集变形场, 并基于该位移场将待配准指纹的像素逐一映射到参考指纹的标准坐标系中。本文选择外部指纹作为参考图像, 则

$$f_{TIR} \approx f_{OCT} \circ \phi \quad (2)$$

其中, f_{TIR}, f_{OCT} 分别表示外部指纹和粗配准后的内部指纹, ϕ 是亚像素变形场, \circ 表示将变形场应用于图像, 实现 f_{OCT} 的每个像素到 f_{TIR} 中的相应位置的映射。

图3展示了精配准阶段的整体框架, 框架的主体是一个变形场估计网络U-NeXt, 该网络主要进行以下设计。首先采用ConvNeXt^[44]块替代了传统U-Net编码器侧的卷积块, 使网络能够捕获更广泛的特征。其次, 设计了HFFEC模块加入跳跃连接, 从而增强网络对高频信息的感知能力。此外, U-NeXt引入了循环一致性^[45]作为训练机制, 从而

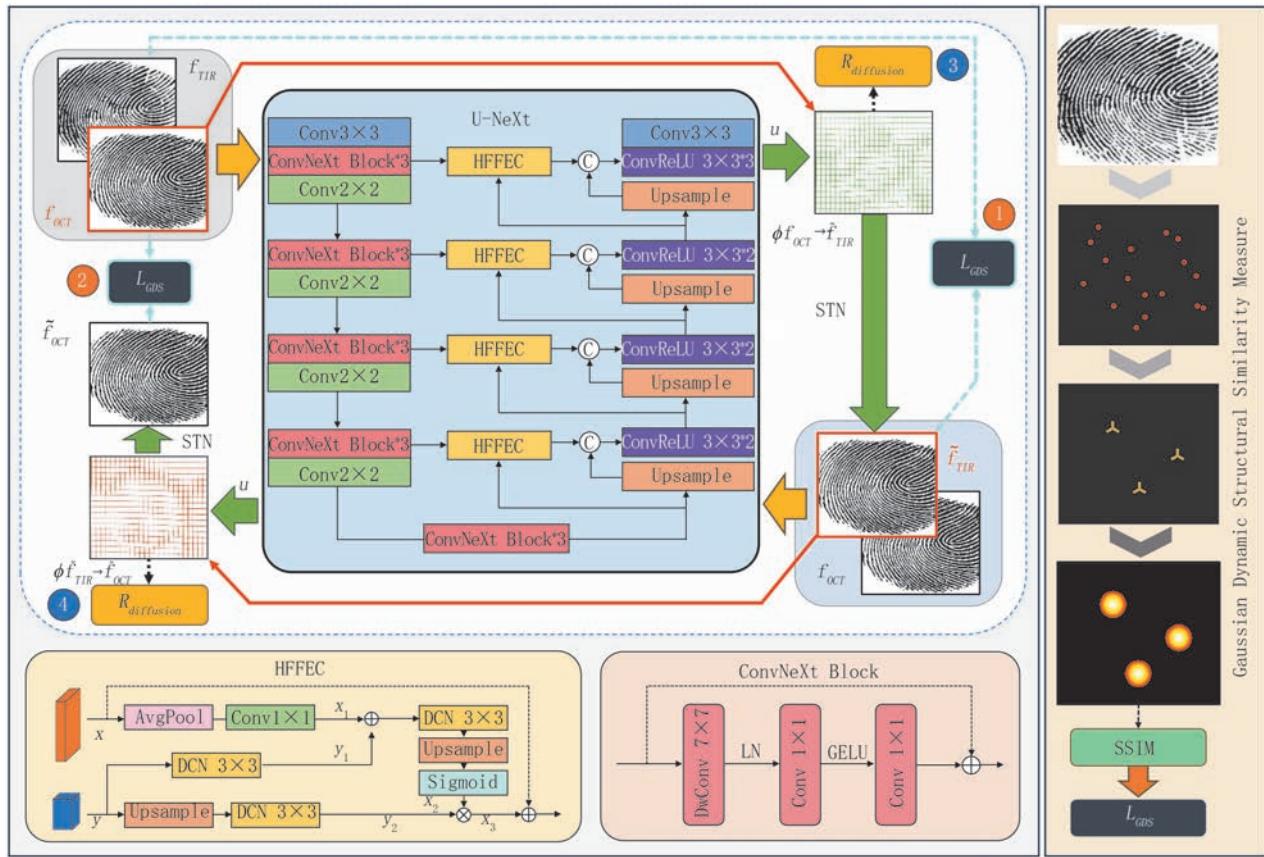


图3 整体框架

在一定程度上保证变形场的可逆性,得到更可靠的配准结果。最后,损失函数部分在两个位移场的平滑正则项损失 $R_{diffusion}$ 基础上,加入了基于种子点的动态结构相似度损失 L_{GDS} ,它能够基于种子点随着训练轮次动态调整,从而降低局部误配准的概率。

3.2.1 具有全局特征捕获能力的纯卷积网络U-NeXt
 U-NeXt采用了形似U-Net的编码-解码器架构。由于指纹图像中经常存在重复的脊谷模式,仅关注局部信息容易导致网络在训练时陷入局部最优,从而错误地对齐非匹配区域。传统的U-Net通过连续的卷积层进行特征提取,但由于卷积核大小固定,难以捕获图像的全局信息。为了使网络能够捕获全局特征,本文用ConvNeXt块替代了传统U-Net编码器侧的卷积块。ConvNeXt是一种纯卷积神经网络模块,由ResNet50改进而来,并借鉴了Swin Transformer的设计,具备精度高、效率高、可扩展性强和设计简单的特点。与传统卷积相比,ConvNeXt使用较大的卷积核,使得单个卷积层能够覆盖更大的感受野,从而捕获更广泛的全局上下文信息。同时,ConvNeXt利用深度可分离卷积减少计算量,以保持对空间特征的有效捕捉能力。因此,本文以

ConvNeXt块作为编码器,捕获待配准指纹图像和参考指纹图像之间的空间对应关系。在解码器部分,采用连续的卷积层将编码器提供的信息处理成密集位移场。为了增强网络对指纹纹理细节等高频信息的感知,提高纹理细节的配准精度,本文对传统的跳跃连接进行了改进,设计了一个高频信息增强卷积模块HFFEC,将低级特征与高级特征进行融合,在保留原始输入数据中的细节和语义信息的同时,增强了原始特征图中的高频特征。

网络首先在通道维度上对内外指纹进行了拼接,使输入维度变为 $2 \times H \times W$ 。得到拼接图像后, U-NeXt最开始通过一个卷积核大小为 3×3 的卷积层对输入进行特征提取。需要说明的是,在原本的ConvNeXt中,起始部分的“Patchify”Stem将输入图像分割为较小的非重叠块($P \times P$),这导致ConvNeXt作为编码器只能提供最高达 $H/P \times W/P$ 分辨率的特征图,无法提供高分辨率特征图以及在较低层次聚合局部信息。因此,在U-NeXt中,通过修改分割步长为1实现了 $H \times W$ 分辨率的特征图提取。之后,特征图会通过五个连续的卷积阶段,每个阶段的通道数分别为48、96、192、384和768,每个阶

段都由三个相同的 ConvNeXt 块组成。在每个 ConvNeXt 块中,首先,通过卷积核大小为 7×7 深度可分离卷积提取全局和局部特征,同时使用 LayerNorm 对特征进行归一化,紧接着通过点卷积(1×1 卷积)在通道维度上进行重塑和信息交互,调整特征图的通道数为原来的 4 倍;然后,通过 GELU 激活函数,引入平滑的非线性,增强对复杂特征的表达能力;最后,通过线性层(1×1 卷积)将通道数缩减至原始数量,并利用残差连接将输入直接加到线性层的输出中,减轻梯度消失的同时增强特征的多层次表达。在前 4 个阶段中每个阶段的最后,通过步幅为 2 的降采样卷积来降低特征图的分辨率,将特征图大小缩减为原来的 $1/2$,并将通道数调整为原来的 2 倍以提取更加抽象的高层次语义信息。利用多个编码阶段,网络捕捉到了多尺度的特征信息。

网络通过跳跃连接维护编码器与解码器之间的定位信息流,而位于跳跃连接上的 HFFEC 模块则是一种旨在增强高频特征感知并实现多尺度信息高效融合的结构。在 HFFEC 中,首先通过平均池化对低级特征 $x \in \mathbb{R}^{C \times H \times W}$ 进行下采样,并使用 1×1 卷积对其进行通道扩张和特征变换,得到 $x_1 \in \mathbb{R}^{2C \times H/2 \times W/2}$ 。同时,高级特征 $y \in \mathbb{R}^{2C \times H/2 \times W/2}$ 被送入两个并行分支,其中第一个分支通过一个 3×3 的可变形卷积(DCN)捕捉灵活的局部空间信息,得到 $y_1 \in \mathbb{R}^{2C \times H/2 \times W/2}$,第二个分支则先对输入特征进行上采样以提高分辨率,然后通过一个 3×3 的可变形卷积增强特征和降低通道数,得到 $y_2 \in \mathbb{R}^{C \times H \times W}$ 。接下来, x_1 通过逐像素加法与 y_1 融合,并经过另一个 3×3 的可变形卷积进一步整合信息,在进行上采样处理后,通过 Sigmoid 函数来区分不同频率的特征,得到 $x_2 \in \mathbb{R}^{C \times H \times W}$ 。通过将 x_2 乘以 y_2 实现对高频特征的自适应关注和增强,得到 $x_3 \in \mathbb{R}^{C \times H \times W}$ 。最后,将原始低级特征加入到增强后的特征图 x_3 中,以突出原始特征图中显著差异信息,增强对复杂细节的关注。

解码器由级联的卷积核大小为 3×3 的上采样和卷积层组成。解码阶段中的每个上采样后的特征图与 HFFEC 模块的输出进行连接,再进入到两个连续的卷积层。在解码器的最后阶段,特征图尺寸恢复至 $48 \times H \times W$,再经过一个 3×3 的卷积层将维度进一步缩减至 16,最后利用 3×3 卷积层来生成输出位移场 u ,除了最后一个卷积层外,每个卷积层后面都有一个 Leaky 整流线性单元激活层。 u 与输入大小相同,两个通道分别代表 x 方向和 y 方向的位

移,而变形场是位移场与恒等映射相加的结果: $\phi(x, y) = (x, y) + u(x, y)$ 。由于 STN 采用变形场实现源图像到目标图像的转变,STN 因此将位移场 u 转换为应用于内部指纹的非线性变形场 ϕ ,从而实现内部指纹与外部指纹的配准。

此外,本文在 U-NeXt 的训练模式中加入了循环一致性。循环一致性旨在确保两个相互转换的操作在数学上保持一致性,具体来说,该性质要求图像经过某种变换从原始域转换至目标域后,再通过逆变换返回时,其结果应与原始图像保持高度一致或极为相似。这一约束在不增加额外模型参数的前提下,有效排除非循环一致的变换,从而促使网络学到更加合理、更加平滑的变形场映射。循环一致性训练模式如图 4 所示,网络首先将外部指纹 f_{TIR} 作为参考图像,并利用网络生成的变形场 $\phi_{f_{OCT} \rightarrow \tilde{f}_{TIR}}$ 将内部指纹 f_{OCT} 配准到 f_{TIR} ,得到配准后的图像 \tilde{f}_{TIR} ,随后,以 f_{OCT} 作为新的参考图像,将 \tilde{f}_{TIR} 配准到 f_{OCT} ,对应的变形场为 $\phi_{\tilde{f}_{TIR} \rightarrow \tilde{f}_{OCT}}$,还原后的图像为 \tilde{f}_{OCT} ,最后,网络通过循环一致性损失鼓励 \tilde{f}_{OCT} 与 f_{OCT} 的差异得进一步缩小,损失的具体表达将在下一小节中详细说明。

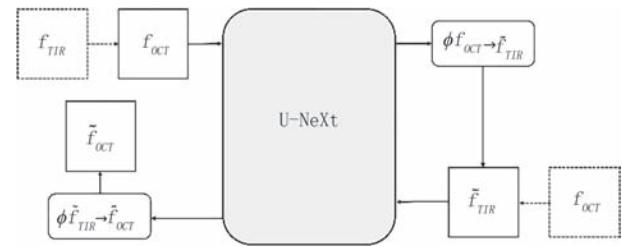


图 4 循环一致性训练模式

3.2.2 动态结构相似度损失函数

在无监督学习中,由于缺乏变形场真值,无法像监督学习那样通过大量的真实变形场数据来拟合得到高精度的变换模型。在这种情况下,图像相似度量成为了评估配准质量的核心标准。然而,指纹脊线的重复模式会导致局部区域的高自相似性,噪声和污染则进一步降低了指纹匹配区域之间的相似度。与同源指纹相比,内外指纹之间的图像差异通常较大,这些因素会干扰相似度量的判断。因此,尽管相似度的提升通常与配准质量的提高相关,但它并不总能保证配准结果的正确性,反而可能会导致局部最优解,进而引发局部误配准现象。局部误配准指的是相似度较高的非匹配区域通过配准形成错误的空间对应关系,常见的情形是邻近脊线错位对

齐。本文总结了两种主要导致局部误配准的原因:(1)粗配准阶段效果不佳,未能有效消除严重的相对畸变,导致精配准阶段存在较大的搜索空间,从而加剧了模式重复的干扰;(2)内部指纹或外部指纹局部质量差异较大,造成匹配区域之间的相似度较低。为了减少上述问题对精配准的影响,本文提出了一种基于种子点的动态结构相似度损失函数。

动态结构相似度损失的核心思想是在传统结构相似度损失中引入以指纹中粗配准后置信度较高的种子点为中心生成的随时间离散变化的蒙版,其演变模式类似于热传导中的热量扩散,即服从二次高斯函数分布,因而具备平滑、局部集中的快速衰减特性,可以作为一种时空引导信息。在这种设计下,各像素处变形场的自由度与其置信度成反比,而置信度则以种子点为中心逐渐向周边扩散,类似于热量的扩散过程,从而为指纹图像中每个像素构建了时空依赖关系,不再仅依赖局部邻域相似性进行定位,进而有效缓解纹理重复模式和指纹局部质量差异带来的配准误差。该方法主要包含种子点的确定以及动态扩散过程的设计两个关键部分。

通过大量观察,我们发现指纹中的奇点及部分细节点的邻域与周围区域具有显著差异。这些点通常具有独特的结构(如分叉点或端点),且周围较少出现重复的纹理模式。本文从粗配准中获得的匹配细节点对中筛选出一些特殊的细节点,这些点的邻域在粗配准阶段已经获得较好的对齐,这些点被定义为种子点。种子点筛选的标准如下:(1)奇点或其邻域存在细节点时,这些细节点具有最高优先级,选择与奇点最近的细节点作为种子点;(2)指纹差异较大的区域,如果在该区域外脊线法线方向两端存在细节点,则将其纳入候选种子点,但位于脊线平行方向两端及区域内部的细节点不适合作为种子点;(3)种子点应尽量分散且数量适中,以确保覆盖指纹中的关键区域。简单来说,种子点即是少数的分散于风险区域(容易发生误配准的区域)的正确搜索方向两端的可靠细节点。完成筛选后,保存种子点在指纹图像上对应的坐标。

在形变失真较轻的情况下,相似度量能够快速且稳定地建立像素之间的对应关系。因此,优先对种子点邻域进行配准,利用变形场的连续性和光滑性,使其周围像素在匹配过程中受到牵引,从而在已获得高置信度配准的区域附近进一步搜索匹配对象,并逐步向外扩展,形成置信度的逐层传递。基于这一思想,仅依赖静态相似度量作为损失函数是不

足的,还需要引入动态发展的机制。

借鉴热传导理论中的高斯函数应用,本文将二次高斯函数的变化机制引入到损失函数设计中,以模拟热量从种子点向周围区域扩散的过程。该函数的平滑曲线及其与时间相关的方差变量能够有效控制数据的分布。

首先,以种子点为中心在平面空间建立二维高斯函数 G ,与热扩散不同的是,由于不存在能量守恒的规则,本文设置峰值系数为 1,即无论在什么方差下,函数的最大值都为 1:

$$G(x, y) = \begin{cases} e^{-\frac{(x-d_x)^2+(y-d_y)^2}{2\sigma^2}}, & e^{-\frac{(x-d_x)^2+(y-d_y)^2}{2\sigma^2}} \geq c \\ 0, & e^{-\frac{(x-d_x)^2+(y-d_y)^2}{2\sigma^2}} < c \end{cases} \quad (3)$$

其中 $x \in [0, H]$, $y \in [0, W]$ 分别代表像素的纵横坐标, (d_x, d_y) 表示种子点坐标, σ 表示高斯函数的标准差, 其大小决定了数据的分布, c 代表截止阈值, 小于 c 的区域将直接取零, 实验中将其设置为 0.3。在本文中, 标准差 σ 是一个与训练轮次 ep 相关的变量, 拥有初始值 γ , 可随着轮次的增加而逐渐增大增加, σ 的表达如下:

$$\sigma = \delta \times ep \mid \alpha + \gamma \quad (4)$$

其中 \mid 表示整除, δ 表示变化系数, α 表示间隔系数, 这些系数共同决定了以 (d_x, d_y) 为圆心的“高斯圆”面积的增长速率, 在本文中, γ, δ 和 α 分别被设置为 35、3 和 8。接下来, 这个基于二次高斯函数生成的图像将作为蒙版与结构相似度损失结合, 将网络的反馈视野限制于以 (d_x, d_y) 为圆心的圆内, 随着 σ 的增大, 置信区间也会逐渐增大, 越靠近圆心, 蒙版的权重越大, 即越靠近种子点的区域置信度越高, 并且不易受到低置信度区域的影响而发生变化。结构相似度图 M_{SSIM} 的计算方式如下:

$$M_{SSIM} = ssim(f_{TIR}, f_{OCT} \circ \phi) \quad (5)$$

其中 $ssim(\cdot)$ 表示结构相似性计算, 它的结果是一张与输入图像大小相同的相似度图 M_{SSIM} 。将高斯蒙版作用于相似度图, 计算每个有效像素的均值, 得到高斯动态结构相似度损失函数 L_{GDS} :

$$L_{GDS}(f_{TIR}, f_{OCT}, \phi) = 1 - \overline{avg}(M_{SSIM}, M_{G_s}(S)) \quad (6)$$

其中 S 代表种子点集合, $S \in \overline{M}$, $M_{G_s}(\cdot)$ 表示二次高斯蒙版, 它由各种子点通过式(3)得到的热力图叠加而来, 图 5 展示了包含若干种子点的高斯蒙版随方差变化的过程。 $\overline{avg}(A, B)$ 表示对 A 和 B 的哈达玛积在 B 中不为 0 的元素进行均值计算。

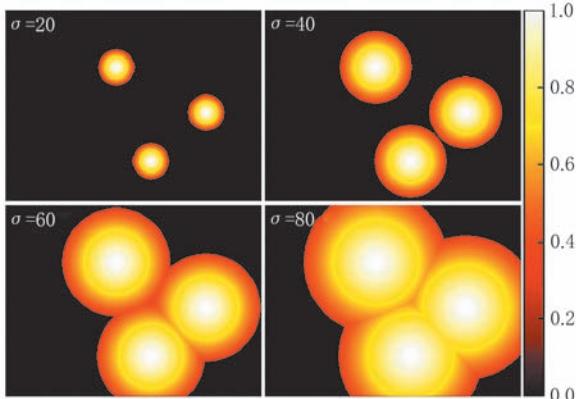


图5 不同标准差下的热力图

高斯动态结构相似度损失二次利用了粗配准阶段得到的指纹细节点,加强了两个配准阶段的联系,通过动态的二次高斯蒙版为网络的学习提供了辅助性的语义信息,更好地利用了指纹特征。除了上述的动态结构相似度损失,还需要加入一个正则项对位移场进行约束,保证其平滑性:

$$R_{\text{diffusion}}(\phi) = \sum_{p \in \Omega} \|\nabla u(p)\|^2 \quad (7)$$

其中, $R_{\text{diffusion}}(\cdot)$ 表示扩散正则化, $\nabla u(p)$ 表示位移场 u 的平面梯度, p 表示像素位置, Ω 表示图像域, 平面梯度使用前向差分来近似, 即 $\frac{\partial u(p)}{\partial |x, y|} \approx u(p_{|x, y|}) - u(p_{|x, y|})$ 。

引入循环一致性后, 完整的损失函数如下:

$$\begin{aligned} L(f_{\text{TIR}}, f_{\text{OCT}}, \phi) = & L_{\text{GDS}}(f_{\text{TIR}}, f_{\text{OCT}}, \phi_{f_{\text{OCT}} \rightarrow \tilde{f}_{\text{TIR}}}) + \\ & L_{\text{GDS}}(f_{\text{OCT}}, \tilde{f}_{\text{TIR}}, \phi_{\tilde{f}_{\text{TIR}} \rightarrow \tilde{f}_{\text{OCT}}}) + \lambda R_{\text{diffusion}}(\phi_{f_{\text{OCT}} \rightarrow \tilde{f}_{\text{TIR}}}) + \\ & \lambda R_{\text{diffusion}}(\phi_{\tilde{f}_{\text{TIR}} \rightarrow f_{\text{OCT}}}) \end{aligned} \quad (8)$$

其中, λ 表示正则化超参数, 本文中将其设置为 20。

4 实验结果与分析

4.1 数据集及配置细节

实验中使用的指纹数据来自外部和内部指纹同步采集系统所获取的公开数据集 ZJUT_EIFD^[46]。该数据集中, 内外指纹之间的相对畸变主要来源于系统畸变和采集时差。在实验过程中, 其中一枚指纹作为待配准指纹, 另一枚作为参考指纹, 本文选择内部指纹作为待配准指纹。ZJUT_EIFD 中共包含三个子集, 分别为 datasetA、datasetB 和 datasetC。其中, datasetA 采集自年轻的白领群体, 包含 151 个

手指, 每个手指分别采集了 8 次, 共计 1208 对内外指纹; datasetB 采集自蓝领工作者, 包含 59 个手指, 每个手指分别采集了 6 次, 共计 354 对内外指纹; datasetC 作为 datasetA 的补充, 同样采集自年轻的白领群体, 包含 180 个手指, 每个手指分别采集了 6 次, 共计 1080 对内外指纹。ZJUT_EIFD 数据集的研究者基于 BCL-U Net^[29] 对 OCT 体数据进行皮肤层次分割以提取内部指纹, 利用网格校正板校准 TIR 和 OCT 成像的几何畸变后, 截取两种指纹图像的共同区域并将所有指纹图像统一降采样至 500 dpi。为了确保粗配准过程中有效区域不被丢失, 并为后续网络训练提供便利, 本文对上述数据集进行了填充, 通过添加与边缘像素值相同的边框, 将图像尺寸扩展至 384×288 像素。在实验中, 本文以 datasetA 中四分之三的数据作为训练集用于对网络进行训练。

本文的粗配准方法是基于 Intel 酷睿 i5-10400 2.9 GHz 处理器和 16 GB 内存环境, 在 PyCharm Community Edition 软件中运行的。密集配准网络采用了 PyTorch1.9 和 Cuda11.0 框架, 并使用 NVIDIA RTX2080 (11 GB 显存) 显卡和 i9-10900X CPU 进行训练。为了优化网络性能, 训练时采用 Adam 优化器, 设置学习率为 0.001, 并使用线性下降策略来更新学习率。

4.2 配准精度评估

本文从图像全局相似度、指纹方向场、标记点偏差以及指纹匹配四个方面评估配准精度。通过多种相似度分数粗略分析全局配准情况, 通过对方向场分析脊线对齐情况, 通过计算手工标记点坐标误差分析微观细节的配准情况, 最后通过内外指纹的匹配情况评估配准对指纹应用性能的影响。本文将提出的方法与配准前(No-reg)以及几种非刚性方法进行了对比, 这些方法分别是经典的非刚性指纹配准方法(OFD)^[18]、基于 STN 的密集指纹配准自监督网络(Fcn-stn)^[22]、端到端指纹密集配准网络(DRN)^[12]以及用于高效指纹密集配准的相位聚合双分支网络(PDRNet)^[21]。实验中所有方法统一在应用了 3.1 小节中介绍的风格迁移及粗配准方案后的数据集上展开。在 3.2.2 小节中, 本文提到了两种容易造成误配准的情况, 其中, datasetB 中的数据采集自蓝领工作者并且为提高采集速度降低了 OCT 分辨率, 因此内外指纹之间的质量差异较大, 可用于展现不同方法在质量差异较大的情况下的表现。此外, 本文从数据集中挑选了粗配准后相对畸

变依然比较严重的内外指纹 150 对并命名为 datasetSD, 用于展现不同方法在畸变严重情况下的表现。

4.2.1 全局相似度评估

本文对配准后的 OCT 指纹图像与 TIR 指纹图像进行相似度量化分析, 选取了结构相似性指数 (Structural Similarity Index Measure, SSIM)^[47]、相关系数 (Correlation Coefficient, CC)^[48] 和均方误差 (Mean Squared Error, MSE)^[49] 作为评估指标。其

中 SSIM 是一种基于人眼视觉特性的感知模型, 通过亮度、对比度及结构三个维度的联合度量来表征图像相似性, 其数值域为 [0, 1], 数值越趋近于 1 表明图像相似度越高; CC 基于积差法构建, 通过计算两图像像素值与其均值的离差乘积来表征线性相关性, 其数值域同样为 [0, 1], 数值越大表示图像间相关性越强; MSE 则通过计算两图像对应像素差值平方的均值来量化差异程度, 其数值与图像差异呈正相关。表 1 统计了各指标的定量评估结果。

表 1 全局相似度得分均值统计

方法名	全局相似度指标											
	datasetA			datasetB			datasetC			datasetSD		
	SSIM	CC	MSE									
No-reg	0.5468	0.6815	4956.62	0.5614	0.6820	4697.84	0.5872	0.7356	4593.86	0.5241	0.6768	5297.08
OFD	0.6732	0.7927	3238.65	0.6140	0.7383	3821.38	0.6493	0.7998	3624.39	0.6564	0.7902	3474.77
Fcn-stn	0.5855	0.7182	4290.89	0.5754	0.6926	4408.53	0.5882	0.7565	4329.55	0.5734	0.7225	4483.90
DRN	0.6910	0.8099	3065.49	0.6294	0.7452	3697.38	0.6330	0.8207	3543.89	0.6946	0.8203	3074.21
PDRNet	0.7253	0.8468	2546.31	0.6423	0.7684	3517.41	0.6873	0.8325	3149.33	0.7363	0.8342	2872.13
Proposed	0.7796	0.8723	2104.82	0.6884	0.8044	2892.80	0.7383	0.8652	2670.17	0.7811	0.8837	2053.26

从表 1 中可以看出, 未配准的指纹在三项指标中的得分均为最低, 在这些方法中, OFD 依赖于方向场和梯度场的计算精度, 对噪声比较敏感, 并且容易忽略细小特征; Fcn-stn 的得分最低, 因为网络预测的是指纹之间的稀疏形变参数, 对局部细节的配准效果较差; DRN 的网络模型较为简单, 可能无法有效建模复杂的非线性形变, 导致配准精度不足; PDRNet 将位移场离散为固定间隔的分类任务可能导致连续位移的量化误差, 降低配准精细度。

U-NeXt 同时具备全局上下文信息捕捉能力与高频局部细节感知能力, 具有较高的配准精细度, 因此相比其他方法得到了更高的相似度得分。图 6 展示了这些配准方法分别在四个数据集上的相关系数累积分布曲线 (Cumulative Distribution Function, CDF)。在四幅图表中, 各配准方法对应的曲线均较未配准的曲线向右移动, 其中本文提出的方法所得到的曲线位于最右侧, 其起始和截止位置均较晚, 这表明得分分布更偏向于高分区, 并且曲线上升的平均速率更快, 表明得分的方差较小。图 7 还展示了不同方法下的三组配准实例, 其中第一行为外部指纹图像, 图片下方标注了配准后内部指纹的 SSIM 分数。实际上, 这些相似度指标在评估指纹配准精度时存在一定局限性: 一方面, 许多配准方法在训练过程中直接以相似度作为损失函数, 这导致网络在

训练时过度关注提高整体指纹图像的相似性, 而相似度指标通常只能反映整体的一致性, 而难以捕捉到局部细节上的偏差或错误, 因此即使整体相似度较高, 也可能存在局部误配准的情况; 另一方面, 不同指纹图像区域的纹理、对比度和结构差异较大, 相似度指标在面对这些变化时可能存在敏感性不足的问题。即使如此, 越精确的配准结果在大部分时候依然拥有越高的相似度得分, 相似度指标依然是评价全局配准精度的有效手段, 尽管这并不全面。

4.2.2 指纹方向场差异评估

指纹图像由连续的脊线和谷线组成, 脊线的走向和分布形成了独特的模式。方向场描述了指纹脊线的方向信息, 内外指纹方向场 O_{OCT} 和 O_{TIR} 之间的差异体现了两张指纹之间脊线纹理的对齐情况。方向场差异定义如下: $D(O_{TIR}, O_{OCT}) = \frac{1}{hw} \sum |O_{TIR} - O_{OCT}|$, 即为每个两个方向场之差的绝对值, 其中 h 和 w 分别表示方向场的高和宽。通常, 指纹对齐得越准确, 它们的方向场越相似, $D(O_{TIR}, O_{OCT})$ 的值则越小。由于指纹的方向场在计算时容易受到噪声等因素的影响, 因此在计算方向场前, 预先对指纹进行了二值化处理。表 2 统计了各个方法得到的 $D(O_{TIR}, O_{OCT})$ 的均值, $D(O_{TIR}, O_{OCT})$ 的数值不仅反映了全局配准效果, 同时也与局部细节的处理能力

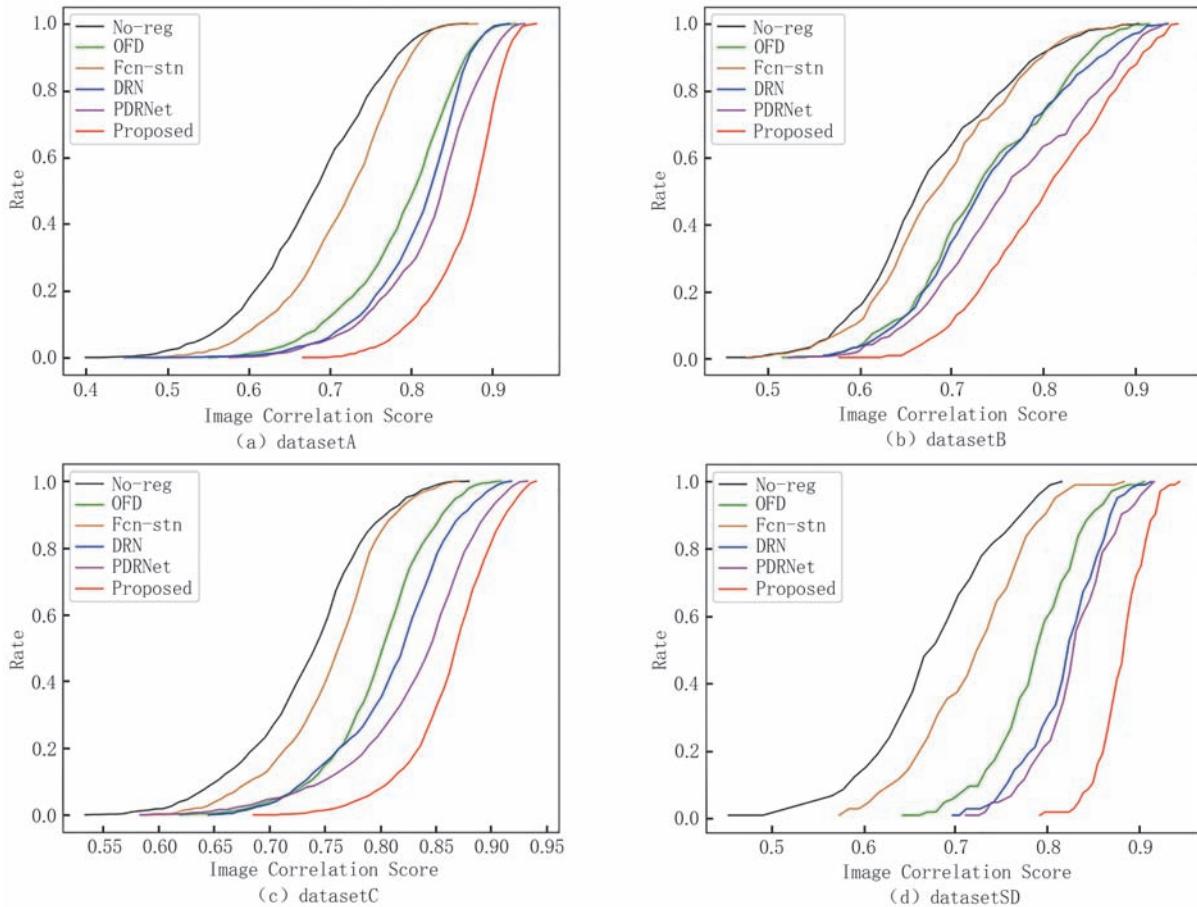


图6 四个数据集下的相关系数分布曲线

密切相关。在所有方法中, Fcn-stn更关心全局一致性, 局部的配准精度较低, 因此有时会加剧内外指纹之间的方向场差异, 其余方法则在不同程度上降低了这种差异。本文提出的U-NeXt在所有对比方法中均实现了最小的方向场差异, 并且在指纹质量差异较大的datasetB和畸变较为严重的datasetSD上, 其差异得分依然最低, 表明U-NeXt对指纹脊线具有较高的对齐精度。图8展示了部分内外指纹脊线的对齐情况及对应的 $D(O_{TIR}, O_{OCT})$ 得分, 其中红色表示外部指纹, 绿色表示内部指纹, 而重叠部分显示为灰色。当误配准发生时, 会产生伪重合现象, 此时局部脊线也呈现灰色; 因此, 若图中出现带状的非重叠区域, 通常表明其周围存在误配准问题。在所有方法中, 本文所提出的方法生成的叠加图显示出最大的灰色连续区域, 表明配准后内外指纹脊线的重合度较高, 局部误配准现象较少。

4.2.3 基于标记点偏差的配准误差评估

本小节以一种更具体的方式, 从配准误差的角度评估不同方法的配准结果。本文随机挑选了100对指纹, 并从这些指纹中各挑选了8处特征进行了仔

细的手工标记, 由于多数细节点在粗配准中已经得到了较好的对齐, 因此本文尽可能挑选了一部分未经使用的细节特征作为标记点。理想情况下, 内外指纹中相应的标记点应该完全对齐, 配准精度越低, 则标记点之间的偏差越大, 偏差的值不仅体现了匹配精度, 更能在一定程度上反映了局部误配准发生的情况。接下来本文通过计算这些匹配标记点之间的偏差来评估各种方法的配准误差, 配准误差表示为匹配标记点之间的欧氏距离。图9展示了标记点偏差的箱线图(a)与累积分布曲线(b), 在(a)中, 本文的方法在几种方法中方差最小且均值最低, 在(b)中, 本文的方法的累积分布曲线比其他方法更快地达到截止值。所有方法的平均配准误差分别为4.70(No-reg)、4.33(OFD)、4.19(Fcn-stn)、3.73(DRN)、3.33(PDRNet)和2.31(Proposed), 本文的方法在其中拥有最小的配准误差。相比之下, 其他方法对于局部细节的配准精密程度不一, 其中OOFD和Fcns-tn的局部精度较差, 并且易造成二次畸变, 因此误差区间较未配准前更大, 部分误差值可能大于配准前, 但平均误差低于配准前。

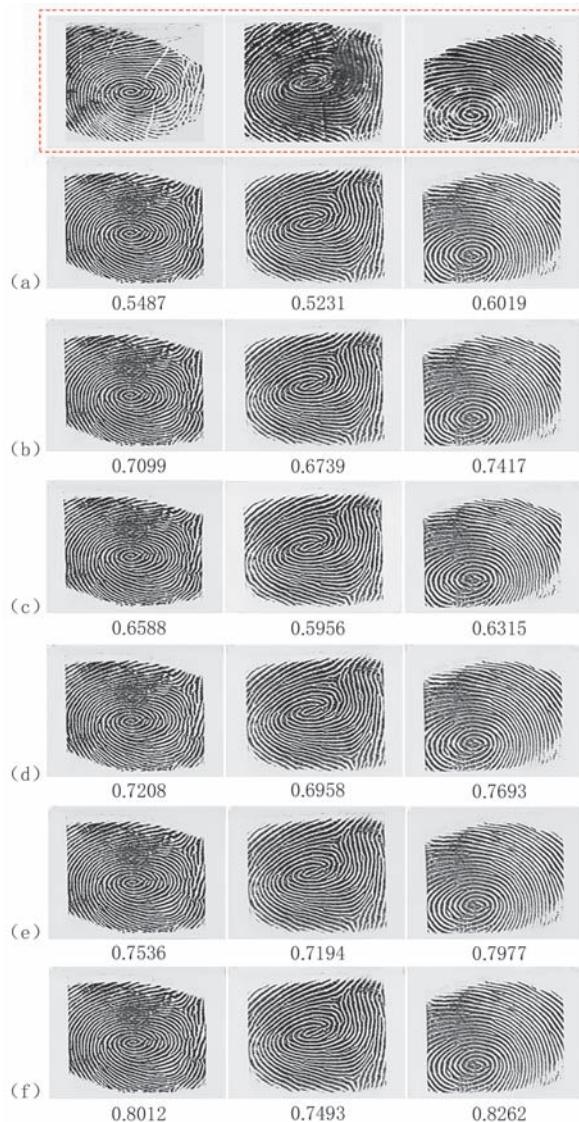


图7 配准实例((a)~(f)依次代表No-reg、OOF、Fcns-tn、DRN、PDRNet和U-NeXt)

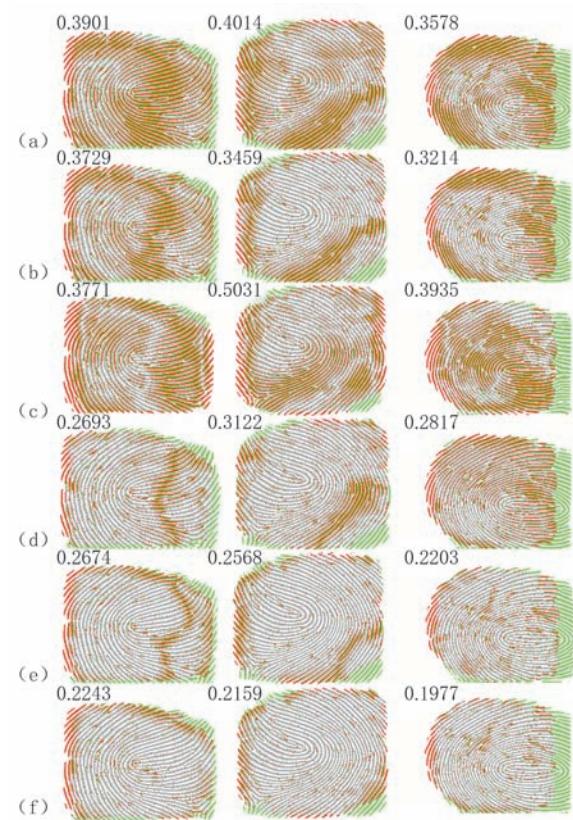


图8 脊线叠加图((a)~(f)依次代表No-reg、OOF、Fcns-tn、DRN、PDRNet和U-NeXt)

从而提高匹配成功的概率。本文利用VeriFinger11.1 SDK对配准后的内外指纹匹配性能进行了评估,表3统计了不同方法配准后的内外指纹匹配分数均值,图10进一步展示了不同方法配准后的匹配分数箱线图。几乎所有配准方法均提升了匹配分数,其中本文提出的方法带来的提升幅度最大。在指纹存在较大扭曲或质量差异显著的情况下,本文的方法相比其他方法依然能够取得更高的匹配分数,如图10中的(b)、(d)所示。这说明本文的方法能够有效对齐匹配特征,提高匹配的准确性,即使在复杂或恶劣条件下仍具有较好性能。此外,几乎所有方法在配准后的匹配分数方差均较配准前有所增加,这是由于不同指纹图像的形变程度不同,配准对匹配分数的提升幅度有所差异。然而,由于整体匹配分数呈正向提升,方差的增加并不意味着负面影响,而是反映了配准对匹配性能的差异化改善。

4.3 融合效果评估

指纹融合对指纹之间的配准精度有着极高的要求,因此本文将配准好的指纹应用于指纹的融合,从融合指纹的角度进一步评估配准的效果。通常,配

表2 方向场差异均值统计

方法名	datasetA	datasetB	datasetC	datasetSD
No-reg	0.1577	0.2601	0.1698	0.2050
OOF	0.1460	0.2127	0.1597	0.1883
Fcn-stn	0.1553	0.2171	0.1699	0.2004
DRN	0.1299	0.1923	0.1480	0.1649
PDRNet	0.1211	0.1785	0.1432	0.1526
Proposed	0.0990	0.1553	0.1347	0.1188

4.2.4 内外指纹匹配评估

配准后,指纹的特征表示在形态、位置和分布上均发生了变化。当前大多数指纹匹配算法主要基于细节点进行匹配,部分方法则依赖方向场或深度网络学习的特征表示。无论采用何种方法,配准精度的提高都会增强两张指纹之间特征表示的关联性,

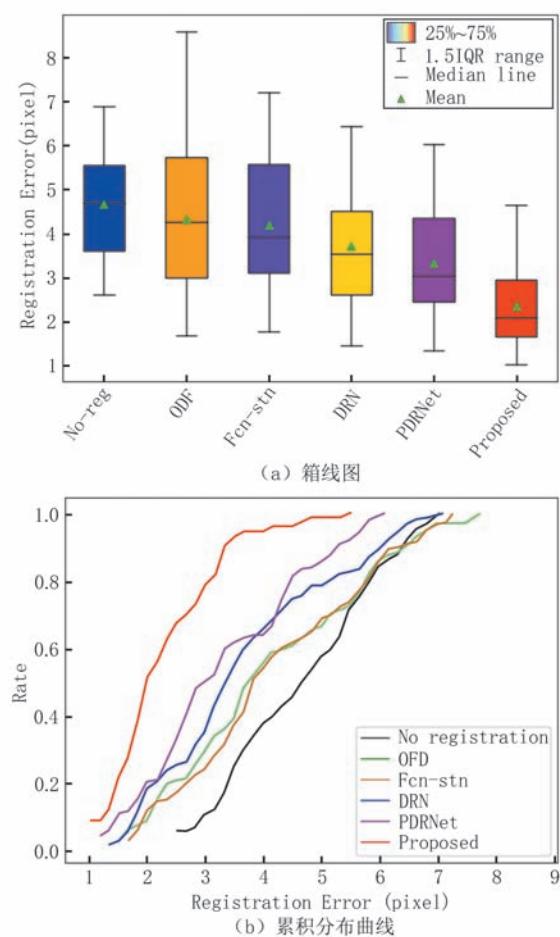


图9 配准误差评估

表3 内外指纹匹配均值统计

方法名	datasetA	datasetB	datasetC	datasetSD
No-reg	199.55	126.02	157.37	192.68
OFD	204.17	129.73	163.29	194.81
Fcn-stn	193.17	121.42	156.77	185.33
DRN	212.88	132.77	169.98	203.27
PDRNet	221.42	136.73	174.71	213.08
Proposed	231.58	143.34	181.27	222.36

准的精度越高,融合指纹中的伪影越少,融合指纹的质量越高,并且指纹的保真度越高,这意味着匹配的准确率越高。实验中采用了Shi等人提出的图像融合方法^[50],本文针对融合指纹进行了两组实验,其一是利用开源指纹图像质量评估器NFIQ 2.0(NIST Fingerprint Image Quality 2.0)^[51]对融合指纹进行评分,其二是验证融合指纹的匹配能力。

NFIQ2.0综合了局部清晰度评分、方向确定度、脊谷均匀性、频域分析和方向流等评估指标,得分越高代表指纹图像质量越好。通过使用NFIQ2.0评估器对融合前后的指纹图像进行评分,

可以定量地评估配准对指纹融合效果的影响。图11展示了各个方法下融合指纹的NFIQ 2.0分数箱线图,在所有方法中,本文方法得到的NFIQ 2.0分数较高且分布集中,在提高融合指纹质量和缩小整体质量差异两方面都有着最好的表现。表4给出了各种方法下配准指纹的融合NFIQ 2.0分数均值,本文的方法取得了其中最高的NFIQ 2.0均分。图12展示了各方法配准后得到融合指纹实例,红框指示了融合指纹中存在的明显缺陷。OOF和Fcn-stn由于精度较低,融合指纹中出现较多的伪影,而其余对比方法配准后得到的融合指纹中伪影的面积有所减少,但由于局部误配准的存在,其边缘依然保留了少量伪影。本文的方法配准后得到的融合指纹局部伪影较少,细节纹理较为清晰,对融合指纹的质量提升最明显。

为了进一步评估配准对融合指纹的影响,本文使用VeriFinger11.1SDK对融合指纹进行了匹配性能评估,本文将datasetA、datasetB和datasetC三个数据集中的指纹进行了混合,共计390个手指。在生成真匹配对时,每个手指统一选取了6次不同采集得到的指纹,因此共计5850($6 \times 5/2 \times 390$)对真匹配;在生成假匹配对时,为了避免全量配对造成真假匹配对数量失衡,本文采用随机抽样平衡真假匹配对数量,得到5850对假匹配。基于匹配得分,绘制了各种配准方法下融合指纹匹配结果的检测误差权衡(Detection Error Tradeoff, DET)曲线,如图13所示,DET曲线显示了在不同的判定阈值下错误匹配率(False Match Rate, FMR)和错误不匹配率(False Non-Match Rate, FNMR)之间的关系。当FMR与FNMR相等时,匹配精度可以表示为等错误率(EER)。EER值越低,匹配精度越高。从图13中可以看出,本文的方法的DET曲线位于最下方且EER值最小,这说明基于本文的方法配准得到的融合指纹有着最佳的匹配性能,能够更好地区分真匹配指纹与假匹配指纹。

4.4 消融实验

为了验证配准网络U-NeXt及动态相似度损失的有效性,本文设置了七组消融实验,本文从先前的实验中挑选了几个具有代表性的指标进行了评估,分别是结构相似指数(SSIM)、方向场误差(Orient-diff)、标注点误差(Reg-error)和融合指纹EER分数,结果如表5所示,第七行的各维数据均最优。以第一行作为基线,通过比较第一、二行数据,可以发现SSIM损失在本研究中综合表现略优于NCC损

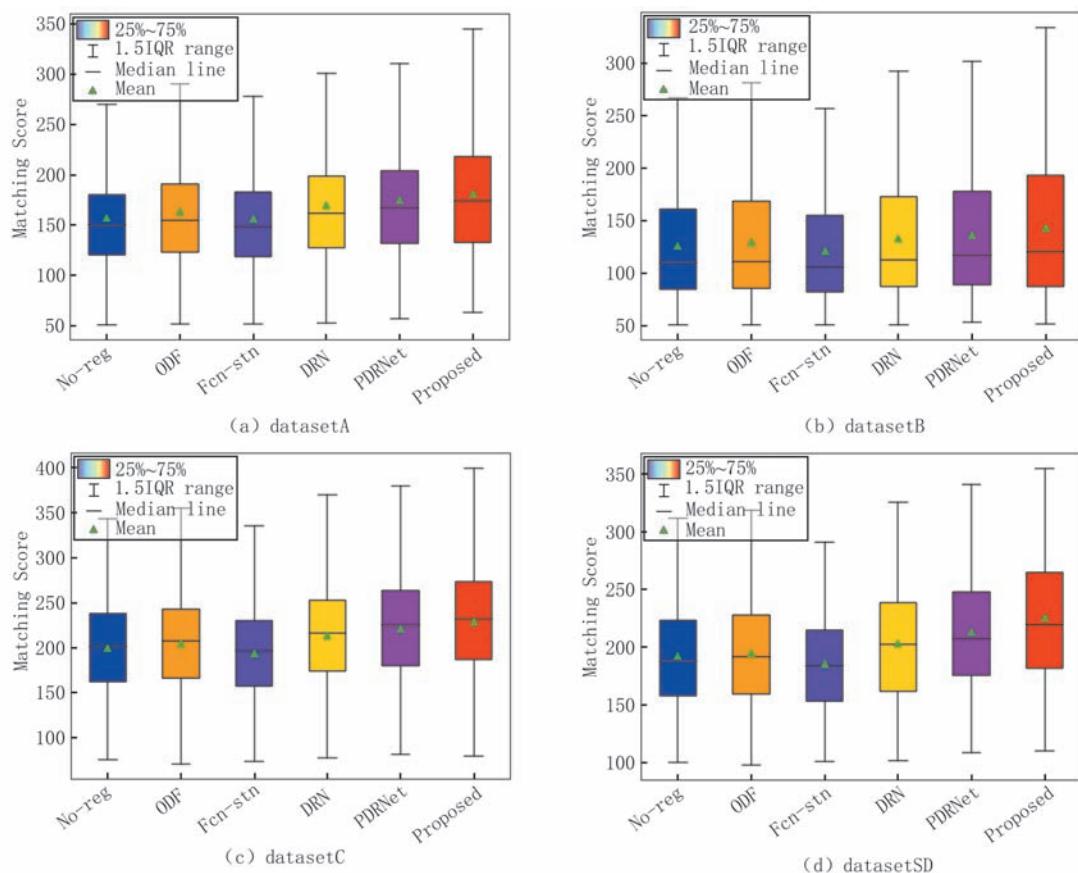


图10 四个数据集下的内外指纹匹配分数箱线图

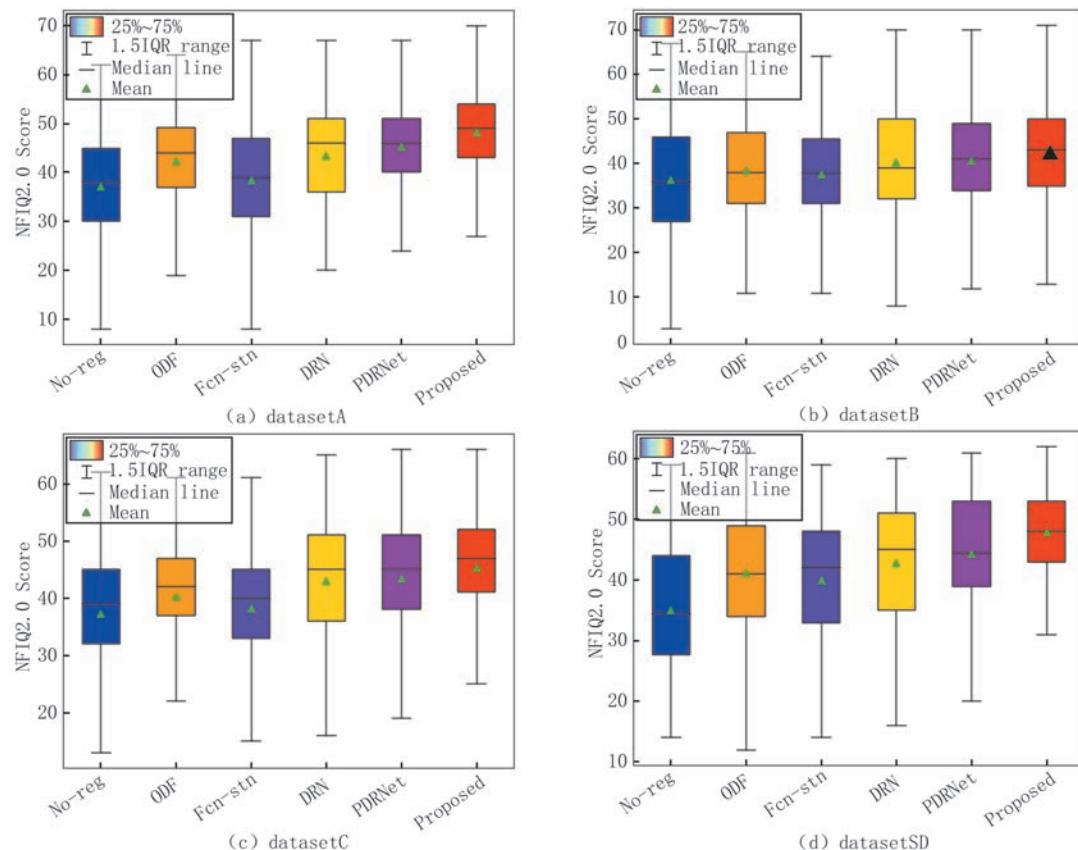


图11 四个数据集下的融合指纹NFIQ箱线图

表4 融合指纹的NFIQ平均分数统计

方法名	datasetA	datasetB	datasetC	datasetSD
No-reg	37.7147	36.3808	37.3102	35.8104
OFD	42.8699	38.3591	40.3491	41.9023
Fcn-stn	38.8994	37.5634	38.1675	37.0500
DRN	43.2516	40.1826	42.9620	42.8557
PDRNet	45.2217	40.6873	43.5222	44.3173
Proposed	48.5890	42.1609	45.3373	47.5642

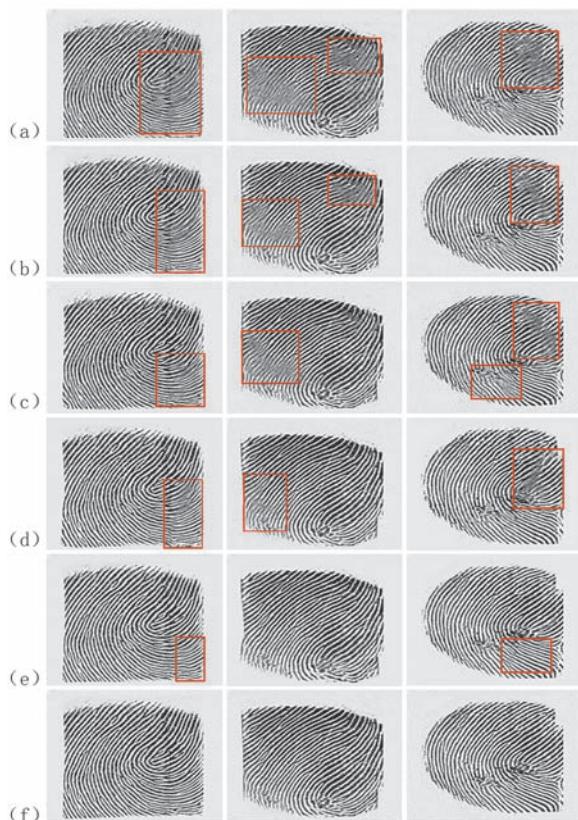


图12 融合指纹实例((a)~(f)依次代表No-reg、OFD、Fcns-tn、DRN、PDRNet和U-NeXt)

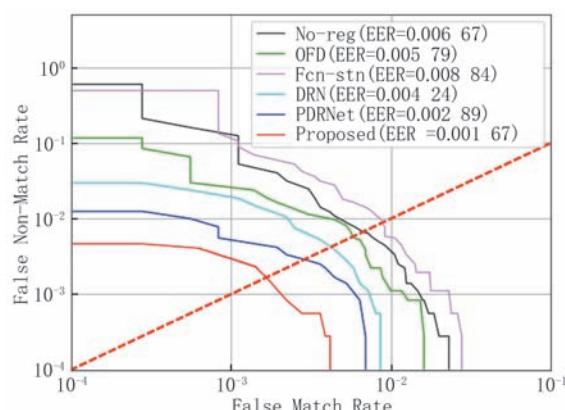


图13 融合指纹DET曲线

失;第三行在基线的基础上将损失替换为动态相似度损失,实验数据证明了动态相似度损失的有效性,

相较另外两个指标,标注点误差的下降较为突出,证明其在降低局部误配准方面具有较好的效果;第四行数据相较基线的提升证明了循环一致性损失的有效性;第五行的实验数据证明了ConvNeXt在网络中的有效性;第六行的实验数据证明了HFFEC模块的有效性;第七行的实验数据则证明了整体相较于基线及基线与单独组件的组合具有更好的配准效果。本文的方法能够有效提高内外指纹的相似度,提升融合指纹的质量,提高内外指纹及融合指纹在匹配中的表现。

表5 消融实验

	SSIM	Orien-diff	Reg-error	EER
U-Net+ L_{SSIM}	0.7082	0.1248	3.5374	0.004 15
U-Net+ L_{NCC}	0.7063	0.1236	3.6084	0.004 23
U-Net+ L_{GDS}	0.7170	0.1193	2.8163	0.003 16
U-Net+ L_{SSIM} + Cycle Consistency	0.7261	0.1207	3.0035	0.003 87
U-Net+ L_{SSIM} + ConvNeXt	0.7696	0.1064	2.6155	0.001 83
U-Net+ L_{SSIM} + HFFEC	0.7534	0.1090	2.7942	0.002 54
U-NeXt (proposed)	0.7796	0.0990	2.3126	0.001 67

5 总 结

本文提出了一个面向融合的无监督多源指纹配准方法,旨在实现内部指纹与外部指纹的像素级精密配准。为了降低网络的配准压力,本文预先进行了内部指纹到外部指纹的风格迁移,在保证脊谷结构不变的前提下统一内外指纹的图像风格,并采用基于细节点的粗配准方法初步对齐内外指纹。本文提出了一种具有动态结构相似度损失的无监督密集配准网络模型U-NeXt,通过预测内部指纹的亚像素级变形场,实现内部指纹到外部指纹的像素映射。U-NeXt结合了U-Net和ConvNeXt,具有全局特征捕获能力,并通过在跳跃连接中加入高频特征增强模块HFFEC提高其对纹理细节的感知能力,此外还引入了循环一致性思想作为训练机制进一步保证变形场的平滑性。为了弥补无监督学习中指导信息的缺失,本文根据指纹特性,设计了一种动态结构相似度损失,能够在训练过程中辅助网络挖掘潜在的语义信息,降低局部误配准的发生。本文基于公开数据集ZJUT-EFID的TIR外部指纹和OCT内部指纹进行了一系列实验与评估,并与其他配准

方法进行了对比,证明了所提出方法的有效性。本文提出的方法在多源指纹的配准任务上表现良好,并且在局部失真较大、指纹质量显著的情况下仍具有鲁棒性,为指纹融合或指纹拼接等中下游任务奠定了良好的数据基础。然而,该方法依然存在提升空间,例如完全端到端配准、跨设备的泛化性等。未来我们将针对以上几点继续改进,特别是在提高方法的泛化性这一方面,由于数据集的指纹数量有限而现实中采集到的指纹图像风格、形变模式多样,在未来的工作中,我们考虑通过引入变形场增强模块实现半监督来提高方法的泛化能力。

参 考 文 献

- [1] Maltoni D, Maio D, Jain A K, et al. *Handbook of Fingerprint Recognition*. London, UK: Springer, 2009
- [2] Soifer V, Kotlyar V, Khonina S, et al. Optical-digital methods of fingerprint identification. *Optics and Lasers in Engineering*, 1998, 29(4-5): 351-359
- [3] Yang J, Oh Y W, Siracosit S, et al. Highly sensitive mutual-capacitive fingerprint sensor with reference electrode. *IEEE Electron Device Letters*, 2022, 43(11): 1973-1976
- [4] Jagannath A, Jagannath J, Kumar P S P V. A comprehensive survey on radio frequency (RF) fingerprinting: traditional approaches, deep learning, and open challenges. *Computer Networks*, 2022, 219: 109455
- [5] Bossen A, Lehmann R, Meier C. Internal fingerprint identification with optical coherence tomography. *IEEE Photonics Technology Letters*, 2010, 22(7): 507-509
- [6] Liu F, Zeng W, Li Y, et al. A lightweight and noise-robust method for Internal OCT fingerprint reconstruction. *IEEE Transactions on Information Forensics and Security*, 2024, 19: 5492-5505
- [7] Wang H, Yang X, Chen P, et al. Acquisition and extraction of surface and internal fingerprints from optical coherence tomography through 3D fully convolutional network. *Optik*, 2020, 205: 164-176
- [8] Sun H, Zhang Y, Chen P, et al. Synchronous fingerprint acquisition system based on total internal reflection and optical coherence tomography. *IEEE Transactions on Instrumentation and Measurement*, 2020, 69(10): 8452-8465
- [9] Lin S, Han Z. Images fusion based on deep stack convolutional neural network. *Chinese Journal of Computers*, 2017, 40(11): 2506-2518 (in Chinese)
(蔺素珍, 韩泽. 基于深度堆叠卷积神经网络的图像融合. *计算机学报*, 2017, 40(11): 2506-2518)
- [10] Li H, Liu F, Yang S. Remote sensing image fusion based on deep support value learning networks. *Chinese Journal of Computers*, 2016, 39(8): 1583-1596 (in Chinese)
(李红, 刘芳, 杨淑媛, 等. 基于深度支撑值学习网络的遥感图像融合. *计算机学报*, 2016, 39(8): 1583-1596)
- [11] Grosz S A, Engelsma J J, Liu E, et al. C2CL: Contact to contactless fingerprint matching. *IEEE Transactions on Information Forensics and Security*, 2021, 17: 196-210
- [12] Cui Z, Feng J, Zhou J. Dense registration and mosaicking of fingerprints by training an end-to-end network. *IEEE Transactions on Information Forensics and Security*, 2020, 16: 627-642
- [13] Zhao D Q, Su F, Cai A. Fingerprint registration using minutia clusters and centroid structure // *Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06)*. HongKong, China, 2006, 4: 413-416
- [14] Uz T, Bebis G, Erol A, et al. Minutiae-based template synthesis and matching using hierarchical delaunay triangulations // *Proceedings of the 2007 First IEEE International Conference on Biometrics: Theory, Applications, and Systems*. Washington, USA, 2007: 1-8
- [15] Ross A, Dass S C, Jain A K. Fingerprint warping using ridge curve correspondences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 28(1): 19-30
- [16] Si X, Feng J, Yuan B, et al. Dense registration of fingerprints. *Pattern Recognition*, 2017, 63: 87-101
- [17] Cui Z, Feng J, Li S, et al. 2-D phase demodulation for deformable fingerprint registration. *IEEE Transactions on Information Forensics and Security*, 2018, 13(12): 3153-3165
- [18] Lan S, Guo Z, You J. A non-rigid registration method with application to distorted fingerprint matching. *Pattern Recognition*, 2019, 95: 48-57
- [19] Yu Y, Wang H, Chen P, et al. A new approach to external and internal fingerprint registration with multisensor difference minimization. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 2020, 2(4): 363-376
- [20] Cui Z, Feng J, Zhou J. Dense fingerprint registration via displacement regression network // *Proceedings of the 2019 International Conference on Biometrics*. Crete, Greece, 2019: 1-8
- [21] Guan X, Feng J, Zhou J. Phase-aggregated dual-branch network for efficient fingerprint dense registration. *IEEE Transactions on Information Forensics and Security*, 2024, 19: 5712-5724
- [22] Yu Y, Wang H, Zhang Y, et al. A STN-based self-supervised network for dense fingerprint registration // *Proceedings of the 15th Chinese Conference on Biometric Recognition (CCBR 2021)*. Shanghai, China, 2021: 277-286
- [23] Jia Y, Cui Z, Su F. Unsupervised fingerprint dense registration // *Proceedings of the 17th Chinese Conference on Biometric Recognition*. Xuzhou, China, 2023: 3-12
- [24] Yu Y, Wang H, Sun H, et al. Optical coherence tomography in fingertip biometrics. *Optics and Lasers in Engineering*, 2022, 151: 106868
- [25] Hu K, Jiang S, Liu D, Gao X. Segmentation of retinal layer boundary in OCT images based on end-to-end deep neural network and graph search. *Journal of Software*, 2021, 35(06): 3036-3051 (in Chinese)
(胡凯, 蒋帅, 刘冬, 等. 基于端到端深度神经网络和图搜索的 OCT 图像视网膜层边界分割方法. *软件学报*, 2021, 35(06): 3036-3051)
- [26] Wang H, Pan D, Liang R, Zhang Y, et al. Anti-spoofing

research of OCT fingerprint based on attention residual network. *Journal of Computer-Aided Design & Computer Graphics*, 2022, 34(03):392-402 (in Chinese)
(王海霞, 潘栋, 梁荣华, 等. 基于注意力残差网络的OCT指纹防伪方法. *计算机辅助设计与图形学学报*, 2022, 34(03): 392-402)

[27] Liu F, Shen C, Liu H, et al. A flexible touch-based fingerprint acquisition device and a benchmark database using optical coherence tomography. *IEEE Transactions on Instrumentation and Measurement*, 2020, 69(9): 6518-6529

[28] Liu F, Zhang W, Liu H, et al. Deep learning based fingerprint subsurface reconstruction. *Chinese Journal of Computers*, 2021, 44(10): 2033-2046 (in Chinese)
(刘凤, 张文天, 刘浩哲, 等. 基于深度学习的亚表面指纹重构. *计算机学报*, 2021, 44(10): 2033-2046)

[29] Ding B, Wang H, Chen P, et al. Surface and internal fingerprint reconstruction from optical coherence tomography through convolutional neural network. *IEEE Transactions on Information Forensics and Security*, 2020, 16: 685-700

[30] An L, Wang R. Use of a scanner to modulate spatial interferograms for in vivo full-range Fourier-domain optical coherence tomography. *Optics letters*, 2007, 32(23): 3423-3425

[31] Tico M, Kuosmanen P. Fingerprint matching using an orientation-based minutia descriptor. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, 25(8): 1009-1014

[32] Yin Q, Feng J, Lu J, et al. Joint estimation of pose and singular points of fingerprints. *IEEE Transactions on Information Forensics and Security*, 2020, 16: 1467-1479

[33] Bookstein F L. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1989, 11(6): 567-585

[34] Kong Q, Wu F, Fan B. Image matching in deep learning era: methods, applications and challenges. *Chinese Journal of Computers*, 2024, 47(7): 1485-1520 (in Chinese)
(孔庆群, 吴福朝, 樊彬. 基于深度学习的图像匹配:方法、应用与挑战. *计算机学报*, 2024, 47(7): 1485-1520)

[35] Yang X, Kwitt R, Niethammer M. Fast predictive image registration//Deep Learning and Data Labeling for Medical Applications: First International Workshop, LABELS 2016, and Second International Workshop, DLMIA 2016, Held in Conjunction with MICCAI 2016, Athens, Greece, 2016: 48-57

[36] Uzunova H, Wilms M, Handels H, et al. Training CNNs for image registration//Proceedings of the 20th International Conference on Medical Image Computing and Computer Assisted Intervention(MICCAI 2017), CityQuebec, Canada, 2017: 223-231

[37] Dosovitskiy A, Fischer P, Ilg E, et al. Flownet: learning optical flow with convolutional networks//Proceedings of the IEEE International Conference on Computer Vision. Santiago, Chile, 2015: 2758-2766

[38] Balakrishnan G, Zhao A, Sabuncu M R, et al. Voxelmorph: A learning framework for deformable medical image registration. *IEEE Transactions on Medical Imaging*, 2019, 38(8): 1788-1800

[39] Jaderberg M, Simonyan K, Zisserman A, et al. Spatial transformer networks//Proceedings of the 29th International Conference on Neural Information Processing Systems (NIPS'15), Montréal, Canada, 2015, 2:2017-2025

[40] Wang H, Ni D, Wang Y. Recursive deformable pyramid network for unsupervised medical image registration. *IEEE Transactions on Medical Imaging*, 2024, 43(6): 2229-2240

[41] IncNeurotechnology. VeriFinger SDK 6.2. Accessed: Aug. 24, 2020, Available: <http://www.neurotechnology.com>

[42] Cappelli R, Ferrara M, Maltoni D. Minutia cylinder-code: a new representation and matching technique for fingerprint recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(12): 2128-2141

[43] Cavalli L, Larsson V, Oswald M R, et al. Adalam: revisiting handcrafted outlier detection. *arXiv preprint arXiv:2006.04250*, 2020

[44] Liu Z, Mao H, Wu C Y, et al. A convnet for the 2020s//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. LA, USA, 2022: 11976-11986

[45] Zhu J, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks//Proceedings of the IEEE International Conference on Computer Vision. Venice, Italy, 2017: 2223-2232

[46] Sun H, Wang H, Zhang Y, et al. ZJUT-EIFD: A synchronously collected external and internal fingerprint database. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 46(4): 2267-2284

[47] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 2004, 13(4): 600-612

[48] Madhuri G S, Gandhi M P I. Image registration with similarity measures using correlation techniques-a research study//Proceedings of the 2015 IEEE International Conference on Computational Intelligence and Computing Research. Tamilnadu, India, 2015: 1-4

[49] Tan H L, Li Z, Tan Y H, et al. A perceptually relevant MSE-based image quality metric. *IEEE Transactions on Image Processing*, 2013, 22(11): 4447-4459

[50] Shi M, Wang H. Internal and external fingerprint fusion combining cross attention and quality optimization. *Journal of Chinese Computer Systems*, 2024, 45(8): 1914-1920 (in Chinese)
(石梦菠, 王海霞. 交叉注意力与质量优化结合的内外指纹融合. *小型微型计算机系统*, 2024, 45(8): 1914-1920)

[51] Tabassi E, Olsen M, Bausinger O, et al. NFIQ 2 NIST fingerprint image quality. 2021, Available: <https://www.nist.gov/services-resources/software/nfiq-2>



WANG Hai-Xia, Ph. D. , professor.

Her main research interests include biometric sensing and recognition and computational optics.

YAN Shi-Qi, M. D. candidate. Her main research interests include biometric pattern recognition and image processing.

YU Yang, Ph. D. , Her main research interests include

biometric pattern recognition and image processing.

CHEN Peng, Ph. D. , professor. His main research interests include computer vision and signal processing.

LIANG Rong-Hua, Ph. D. , professor. His main research interests include perception and understanding of multimodal data.

ZHANG Yi-Long, Ph. D. , associate research fellow.

His main research interests include biomedical optical imaging and optical biosensing.

LIU Yi-Peng, Ph. D. , associate professor. His main research interests include biometric pattern recognition and medical image processing.

Background

Fingerprint is one of the most commonly used biometric feature for personal identification, which is typically collected from the surface of the finger. In recent years, the development of optical coherence tomography realizes the measurement of internal fingerprints. Generally, external fingerprints exhibit higher image quality, whereas internal fingerprints exhibit greater integrity. The complementary nature of external and internal fingerprints allows for the extraction and fusion of valuable information from both, thereby creating a new, more comprehensive fingerprint representation that benefits subsequent applications.

However, owing to the distinct characteristics of acquisition devices employed, there are relative deformations between external and internal fingerprints. Therefore, registration must be conducted prior to fusion in order to restore spatial consistency. Fingerprint registration is a small subset of image registration and presents several challenges. Firstly, in contrast to other natural or medical images, fingerprint images possess distinctive textures and more repetitive patterns, which introduces low intra-class differences

but high inter-class differences. Secondly, current mainstream contact-based acquisition results in relative elastic deformation between fingerprints. Thirdly, fingerprints are susceptible to the influence of the environment, which leads to the quality difference. All these challenges make sub-pixel accuracy registration between external and internal fingerprints a challenging endeavor. In recent years, deep learning-based methods have garnered significant interest as a means of achieving rapid registration. However, the lack of a large number of ground truth displacement fields means that supervised learning cannot be used to achieve internal and external fingerprint registration. Existing unsupervised fingerprint registration researches are relatively rare and prone to local registration errors.

In this paper, we propose an unsupervised dense registration network, U-NeXt, which effectively captures global information, reduces local misregistration, and achieves subpixel-level deformation field estimation between external and internal fingerprints, thus lays a good foundation for fingerprint fusion.