

# Image-based Freeform Handwriting Authentication with Energy-oriented Self-Supervised Learning

王婧瑶, 郑昌文, 牟伦田, 高文, 牟伦田

IEEE Transactions on Multimedia  
王婧瑶 Jingyao\_wang0728@163.com

## 概述

自由笔迹认证通过识别笔迹风格验证身份, 近年来在防欺诈与文化保护等领域受到关注。现实中该任务面临数据损坏、特征复杂与标注困难等挑战。为此, 本文提出一种面向能量的双分支对比自监督框架, 包含: 能量转换预处理、双分支预训练、少样本微调及快速识别。框架兼顾鲁棒性与效率, 并构建模拟真实损坏的数据集。实验证明该方法具备优异的性能与广泛适应性。

## 动机与分析

自由笔迹认证利用个体独特的书写风格进行身份验证, 广泛应用于电子安全与健康领域。相比固定签名, 自由笔迹更具挑战, 因其内容多样、结构复杂, 且常伴随划痕、污渍等数据损坏。标注成本高也限制了监督学习的可行性。为此, 理想模型应在无需高质量数据、固定内容或大量标注的前提下, 实现高效、鲁棒的认证能力。

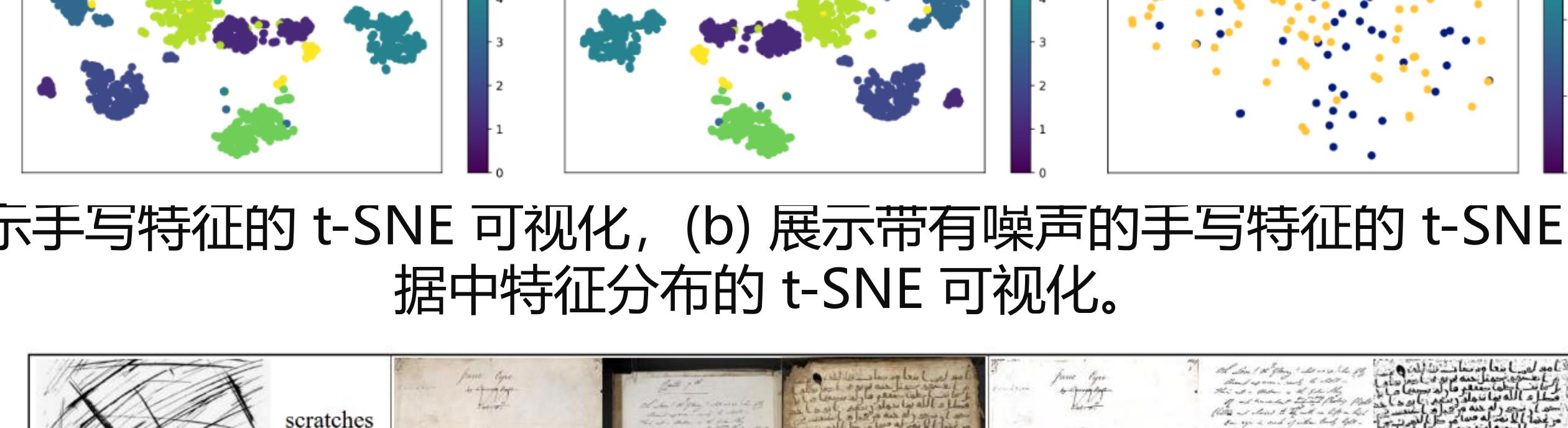


图1 特征可视化。 (a) 展示手写特征的 t-SNE 可视化, (b) 展示带有噪声的手写特征的 t-SNE 可视化, (c) 展示真假数据中特征分布的 t-SNE 可视化。



图2 笔迹缺陷及预处理结果。 (a) 所收集手稿中常见的缺陷和损坏; (b) 所收集现实世界中的笔迹数据样本; (c) 基于能量算子在两轮迭代后的样本。

## 方法

为应对自由笔迹认证中的挑战, 本文提出以能量引导的双分支自监督框架, 涵盖预处理、预训练、微调与应用四阶段。预处理阶段通过能量算子将笔迹转化为能量分布, 抑制噪声并保留关键信息; 预训练阶段结合动量对比学习与自适应匹配机制, 提取高维特征并识别关键区域; 微调与部署阶段提供简便接口, 支持少样本快速适应多场景应用。

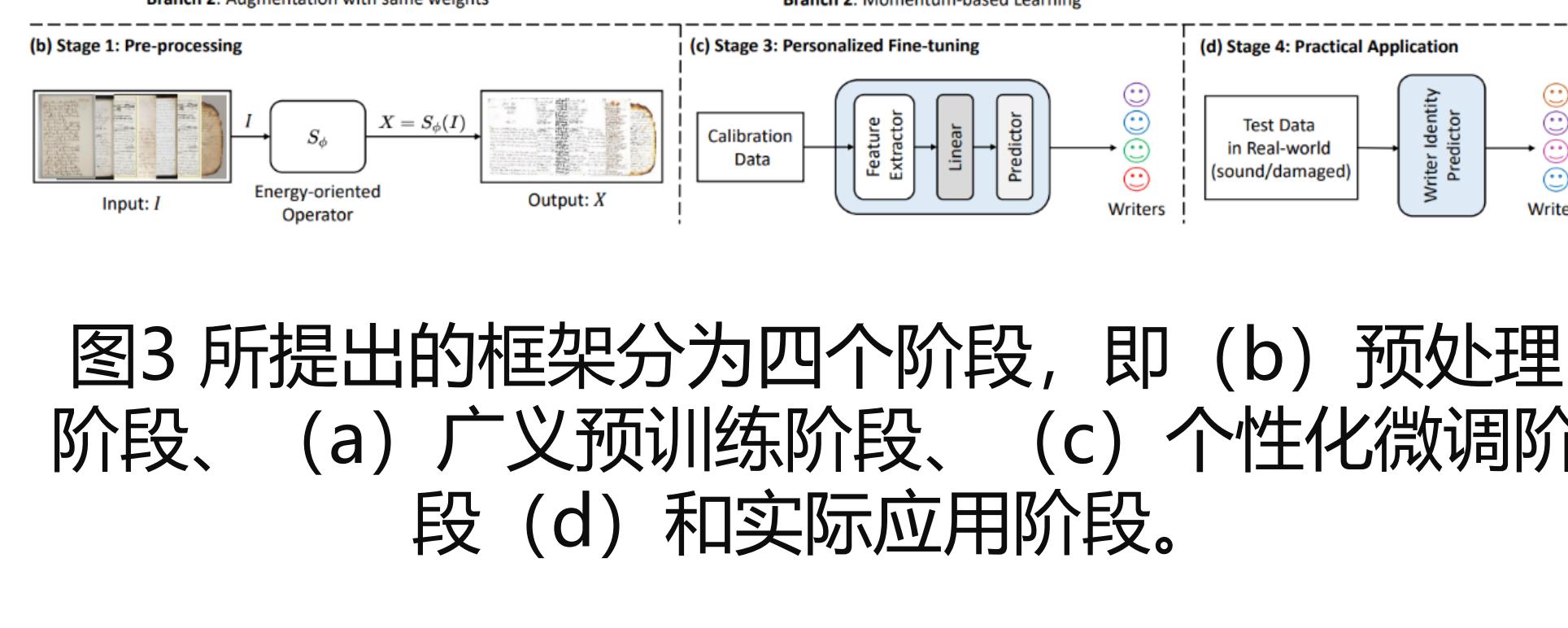


图3 所提出的框架分为四个阶段, 即 (b) 预处理阶段、(a) 广义预训练阶段、(c) 个性化微调阶段 (d) 和实际应用阶段。

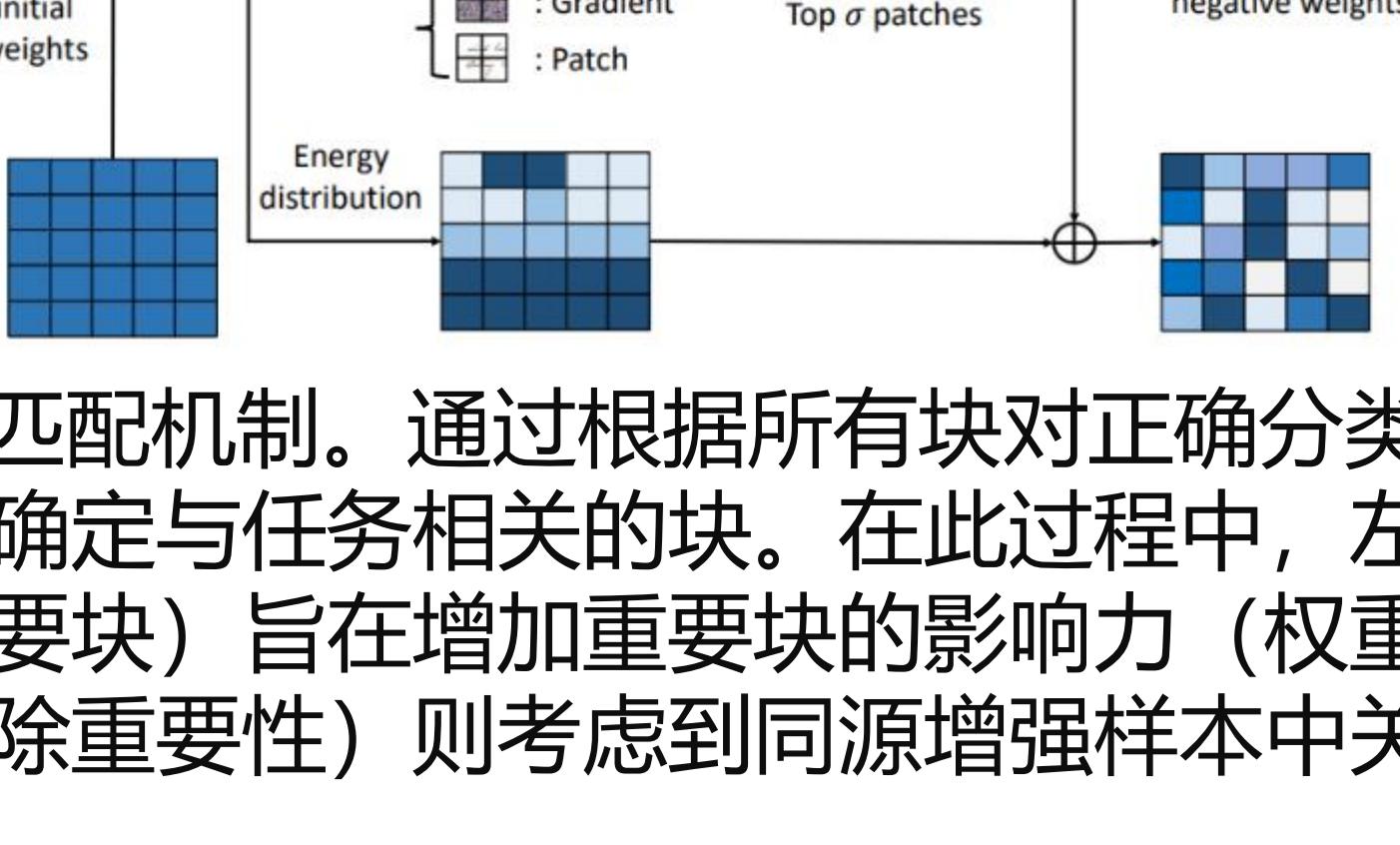


图4 自适应匹配机制。通过根据所有块对正确分类结果的贡献重新加权来确定与任务相关的块。在此过程中, 左侧步骤(即重新加权重要块)旨在增加重要块的影响力(权重), 而右侧步骤(即移除重要性)则考虑到同源增强样本中关键块的差异。

## 实验

为验证方法有效性, 本文在多个基准数据集上开展定性与定量实验, 并构建含800页无标签手写内容的新数据集, 模拟真实与伪造场景。结果显示, 所提方法在有无噪声条件下均优于现有基线, 具备更清晰的聚类边界与更强的鲁棒性, 体现出其在自由笔迹认证中的显著优势。

Methods	IAM-FHA		CEDAR-FHA		CVL-FHA		QUWI-FHA		ICDAR2013-FHA		EN-HA	
	Top 1	Top 5										
SimCLR [57]	61.412	80.515	75.832	82.012	62.301	85.298	49.785	56.353	51.936	69.501	45.893	66.289
BYOL [69]	53.238	79.789	71.520	79.233	49.872	82.011	50.542	68.487	55.293	71.220	44.732	71.458
Barlow Twins [70]	49.947	86.289	71.389	72.332	80.299	42.891	57.777	50.829	68.839	48.203	68.128	
MOCO [61]	64.852	82.545	69.298	70.122	71.513	81.022	61.825	49.063	74.202	56.256	77.197	
NN-LBP [72]	18.512	31.293	24.355	39.053	13.523	28.238	9.328	17.938	19.544	35.205	10.083	21.279
NN-LPQ [72]	18.148	32.932	25.534	42.200	30.856	10.254	17.652	17.025	34.545	12.877	22.830	
NN-LTP [72]	17.843	29.842	24.378	37.234	14.784	30.239	9.010	16.382	21.019	38.231	11.793	24.873
Cohinge [73]	19.622	35.215	40.420	51.527	18.164	34.055	15.058	22.024	22.017	44.789	14.109	26.724
QuadHinge [73]	20.984	36.492	40.281	52.098	16.373	37.017	15.441	25.093	25.234	44.897	15.389	27.018
COLD [73]	11.893	27.809	39.839	48.879	17.132	35.500	13.201	20.865	20.052	39.581	10.035	25.284
CC-Pairs [74]	13.480	27.652	30.939	48.039	19.892	30.180	12.209	24.278	27.492	41.033	20.840	45.923
CC-Triplets [74]	15.415	34.732	34.893	51.289	19.010	31.122	12.757	25.207	28.565	44.028	19.982	49.284
SherlockNet (Ours)	<b>83.290</b>	<b>96.101</b>	<b>87.663</b>	<b>95.019</b>	<b>81.004</b>	<b>95.129</b>	<b>60.437</b>	<b>76.748</b>	<b>71.178</b>	<b>91.908</b>	<b>82.782</b>	<b>94.355</b>

表1 不同方法在六个基准数据集上的准确率(%) , 采用随机拼接的手写页面作为训练数据

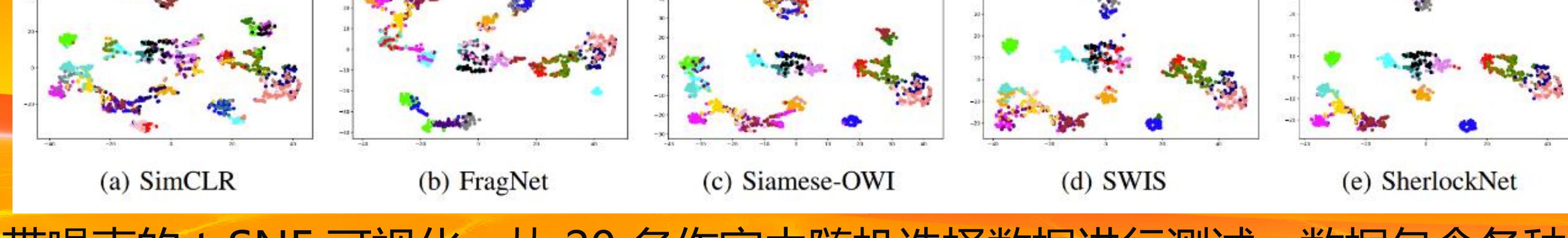


图5 EN-HA 上带噪声的 t-SNE 可视化, 从 20 名作家中随机选择数据进行测试, 数据包含各种噪声(黑色)。