# 小论文之实验

本文所提出的方法在两个公开的可见光-红外行人重识别数据集SYSU-MM01和RegDB上进行实验。

SYSU-MM01数据集的行人图像由四个可见光相机和两个红外相机拍摄收集，其中包含491个行人身份，共有287628张可见光图像和15792张近红外图像，涵盖室内和室外环境。

训练集包含395个行人身份，共有19659张可见光图像和12792张红外图像；测试集包含 96 个行人身份，其中有3803张红外图像作为查询图像。图库集由测试模式决定，其中测试模式有两种，分别是全局搜索模式和室内搜索模式。在全局搜索模式下，所有由可见光摄像头拍摄的图像将构成图库集。在室内搜索模式下，仅使用两个室内可见光摄像头拍摄的图像来构成图库集。

RegDB数据集的行人图像由一个可见光摄像头和一个红外摄像头拍摄收集，其中包含412个行人身份，每个行人身份有10张可见光图像和10张红外图像。数据集被随机分为两半：206 个行人身份的图像用于训练，另外的 206 个行人身份的图像用于测试。测试有两种评估模式，分别是可见光图像作为查询集，红外图像作为图库集和红外图像作为查询集合，可见光图像作为图库集。

在实验过程中，本文使用累积匹配特征(cumulative matching characteristics, CMC) 和平均精度(mean average precision, mAP)作为模型性能的评估指标。

实验使用pytorch搭建环境，总共进行了80个Epoch的模型训练。首先，输入图像的尺寸被调整为3×384×192，并采用随机水平翻转和随机擦除进行数据增强。在训练过程中，采用warm-up策略来平滑训练梯度，初始学习率设置为1×10−2，经过10个epoch线性增加到1×10−1，再在第20个epoch时将学习率衰减为1×10−2，最后在第60个epoch时进一步衰减至1×10−3。每个批次随机选择4个行人，每个行人分别选取4张可见光图像和4张红外图像进行训练。模型采用SGD优化器进行优化，动量参数设置为0.9。三元组损失函数的边距超参数设定为0.3，权重超参数设定为1，而分布一致性损失函数的权重超参数设定为0.5。

本文所提方法与主流的可见光-红外跨模态行人重识别方法进行对比实验，包括：Zero-padding、HCML、BDTR、D2RL、AlignGAN、X-Modality、DDAG、LADN、AGW、NFS、CM-NAS、FMCNet、TSME、MID。其中，基于模态共享特征提取的方法有：Zero-padding、HCML、BDTR、DDAG、LADN、AGW、NFS、CM-NAS；基于模态特定信息补充的方法有：BDTR、D2RL、AlignGAN、X-Modality、FMCNet、TSME、MID。

表1和表2是在SYSU-MM01数据集和RegDB数据集上的实验结果。通过表1和表2可知，本文所提方法具有出不错的性能效果。具体而言，相比于基于模态共享特征提取的CM-NAS方法，本文所提方法在SYSU-MM01数据集的all-search模式Rank-1和mAP分别提升了5.09%和3.91，在in-door模式下Rank-1和mAP分别提升了5.41%和4.52%，而在RegDB数据集的可见光图像检索红外图像的测试模式下，Rank-1和mAP分别提升了9.04%和5.32%，在红外图像检索可见光图像的测试模式下，Rank-1和mAP分别提升了5.78%和2.87%。和基于模态特定信息补偿的FMCnet相比，本文所提方法在SYSU-MM01数据集的in-door模式下Rank-1和mAP分别提升了4.27%和3.38%，而在在RegDB数据集的可见光图像检索红外图像的测试模式下，Rank-1提升了2.72%。

消融实验

本文所提方法在SYSU-MM01数据集上all-search和indoor-search两种模式下进行消融实验。表3是对AMG模块和MAML模块的消融实验结果，由表3可知，辅助模态生成器和混合注意力引导学习模块对模型性能有积极影响。具体而言，在两种测试模式下，基线模型的Rank-1和mAP值分别为58.66、53.76和61.63、67.66；加入混合注意力引导学习模块后，Rank-1和mAP值分别提高到了62.40、58.44和64.09、70.79；在此基础上加入辅助模态生成器，Rank-1和mAP值分别提高到了67.08、63.93和72.42、77.47。

都显示出出色的性能

RegDB 数据集 [27] 由 412 个身份组成，每个身份都包含一对重叠的 VIS 和 IR 相机的 10 个 VIS 图像和 10 个 IR 图像

RegDB数据集是由一个可见光摄像头和红外摄像头捕获的小规模数据集，包含总共412个行人身份，每个行人分别有10张可见光图像和10张红外图像

SYSU-MM01 是一个由四个可见光相机和两个近红外相机收集的大规模数据集，涵盖室内和室外环境。

并且采用累 积匹配特征(cumulative matching characteristics, CMC) 和平均精度(mean average precision, mAP)来对模型进 行评估

从表 1 中，我们得到以下观察结果。 (1) 所提出的 MMN 方法在 RegDB 和 SYSUMM01 数据集上都明显优于其他最先进的方法。具体来说，所提出的 MMN 方法在所有搜索设置的 SYSUMM01 数据集上获得 Rank-1 准确率 = (70.6%) 和 mAP = (66.9%)，并且在可见光到红外设置的 RegDB 数据集上获得 Rank-1 准确率 = (91.6%) 和 mAP = (84.1%)。结果表明，MMN 可以通过利用辅助 M 模态图像来学习 VIS 和 IR 之间更好的模态共享特征。 （2）所有基于 GAN 的方法（例如 D2RL 和 AlignGAN）都采用编码器-解码器架构来生成新特征或新图像，以减少 VIS 和 IR 图像之间的模态差异。一方面，这些方法忽略了这样一个事实：由于非线性关系，可见光和红外之间的模态差异非常复杂。另一方面，基于 GAN 的方法在通道和空间级别上重建模态信息，但模态差异主要保留在通道级别。这些方法在一定程度上破坏了空间结构信息，导致生成的图像与真实图像存在较大差距。与这些方法相比，我们的方法将可见光和红外图像转换为统一的M模态图像，并且可以有效地保留原始空间信息。因此，所提出的MMN方法比基于GAN的方法具有很大的性能优势。 (3)虽然X模态也利用辅助模态来协调可见光和红外模态，但我们的方法将这两种模态投影到统一的M模态图像空间中，这是减少可见光和红外图像之间模态差异的更有效方法。

Zero-padding[1]和D2RL[7]均为单流网络结构，D2RL[7]进一步利用生成对抗网络将跨模态问题转为单模态识别问题，提升了网络性能。BDTR[4]、DPMBN[8]、MSPAC-MeCen[15]、TSLFN[6]采用双流网络，将可见光图像和红外图像分别输入到网络中进行特征提取。MACE[13]、DDAG[12]、DF2AM[19]采用双流共享网络，提升网络提取不同模态共享特征的能力。CDP+DHSM[11]随机选择一个光谱作为中间模态，XCMG[9]根据可见光图像和红外图像生成中间模态来辅助模型学习不同模态共享特征，NFS[20]受NAS方法启发，让模型实现特征的自动选择。本文所提出的方法选择的是参数不共享的双流网络，融入MHSA让模型进行特征学习时考虑到空间信息，并在训练的时候加上辅助模态，抑制模型在可见光图像上对颜色信息的提取，进一步减小模态差异。对比XCMG[9]方法，在SYSU-MM01数据集上Rank-1和mAP均提高了12.28%和9.63%，在RegDB数据集上Rank-1和mAP均提高了25.02%和15.92%。对比最新方法NFS[20]，在SYSU-MM01数据集上Rank-1和mAP均提高了5.29%和4.91%，在RegDB数据集上Rank-1和mAP均提高了6.69%和4%。

所提方法在 SYSU-MM01 数据集上的两种查询模 式下的结果如表 4 所示，具体来看，在 All-search 模 式下，Rank-1 和 mAP 指标均能够超过次优化方法 SFANet，Rank-1和 mAP分别增加了3.73%和4.22%； 在 Indoor-search 模式下，Rank-1 增加了 2.11%。

如表5所示，所提方法在RegDB数据集上获得了不错的效果。具体来看，从可见光图像到红外图像查询设置，所提方法比次优化方MAUMG相比，Rank1增加了2.34%。

我们将所提出的 PMCM 与几种最先进的 VI-ReID 方法进行了比较，包括 Zero-Padding [Wu et al, 2017a]、cmGAN [Dai et al, 2018]、D2RL [Wang et al, 2019b]、AlignGAN [Wang et al, 2019a]、Hi-CMD [Choi et al, 2020]、cm-SSFT [Lu et al, 202] 0]、CMM [Ling 等人，2020]、DDAG [Ye 等人，2020b]、X-Modality [Li 等人，2020]、MACE [Ye 等人，2020a]、MMN [Zhang 等人，2021]、MPANet [Wu 等人，2021]、LbA [Park 等人，2021]、SMCL [Wei 等人， 2021]、MSA [Miao 等人，2021]、JCCL [Zhao 等人，2021]、DCLNet [Sun 等人，2022]、FMCNet [Zhang 等人，2022]、MAUM [Liu 等人，2022b] 和 MID [Huang 等人，2022]。

SYSU-MM01 上的比较结果如表 1 所示。我们观察到，我们的 PMCM 在所有评估指标上都远远优于现有的 SOTA。具体来说，PMCM在全搜索和单次模式下实现了75.54%的Rank-1准确率和71.16%的mAP，分别超过了第二好的方法MAMU 3.68%和2.37%。此外，我们在最后四行中将我们的方法与其他基于中间模态的方法进行了比较。我们观察到我们的方法优于混合方法。RegDB 上的比较结果如表 2 所示，其中我们在比较方法中获得了最佳性能，在 Visible2Infrared 模式下，rank-1 = 93.09% 和 mAP = 89.57%，在 Infrared2Visible 模式下，rank-1 = 91.44% 和 mAP = 87.15%。具体来说，PMCM 在 Visible2Infrared 模式下的 Rank-1 和 mAP 优于第二佳方法 FMCNet 3.93% 和 5.14%，在 Infrared2Visible 模式下优于第二好的方法 FMCNet 3.06% 和 3.29%。两个数据集的一致改进证明了我们方法的优越性，可以学习更多模态不变和判别性特征。

在本小节中，我们将我们提出的策略应用于不同的基线模型进行实验，并将我们的方法与各种现有的跨模态人员 ReID 方法进行比较。我们将现有方法分为两类：（1）基于生成的方法：cmGAN [15]、D2RL [18]、HiCMD [19]、JSIA-ReID [20]、AlignGAN [17]、TS-GAN [21]、DG-VAE [59]、GECNet [22]、FMCNet [25]、TSME [23]。

一些方法引入中间模态来消除跨模态差异，包括 X-Modality [87]、MID [88]。 （2）非生成式方法：BDTR [51]、AGW [3]、CMAlign [76]、SIM [77]、NFS [78]、CIMA [80]、MSO [13]、cm-SSFT [37]、CM-NAS [81]、MCLNet [83]、SMCL [82]、DDAG [41]、MPANet [42]、DMiR [74]、DML [ 79]、HAT [75]、SPOT [14]、MAUM [86]、CMT [85]、DCLNet [84]。表一和表二展示了这些最先进的方法在两个数据集 SYSU-MM01 和 RegDB 上的实验结果。

表中，R1、R10 和 R20 分别代表 Rank-1、Rank-10 和 Rank-20 准确率，mAP 代表平均精度得分。

我们选择两种最具代表性的方法作为基线来证明我们提出的策略的有效性。

首先，我们选择最广泛使用的AGW作为比较的基线，提取全局特征作为表示。我们训练 AGW† 模型作为基线模型。我们使用重新实现模型的实验结果，该模型比原始论文的实验结果提高了性能。其次，我们选择性能优异的MPANet模型作为实验的基线，该模型使用局部特征。通过大量的实验，我们可以发现我们对每种方法的策略都有很大的改进。

表中的\*表示重新排序后的结果。

与 SYSU-MM01 的比较。我们在表 I 中显示了 SYSU-MM01 数据集上的实验结果。从表中我们可以注意到，我们提出的 GC-IFS 策略在所有四种测试设置中都显示出出色的性能，并且优于现有的 SOTA 方法。首先，我们对多个基线进行实验，观察到 GC-IFS 均显着改善了基线。具体来说，(a) 与基线 AGW† 相比，我们的模型 GC-IFSAGW 在 allsearch 模式的单次设置中，Rank1 上超过 15.22%，mAP 超过 12.33%。在室内搜索模式的单次设置中，我们的模型在 Rank-1 和 mAP 上分别优于基线 AGW† 13.68% 和 11.24%。对于重新排序的结果，我们的模型 GC-IFSAGW⋆ 在最困难的测试模式下达到了 85.63% 的 Rank1 准确率和 83.00% 的 mAP。 (b) 与基线 MPANet 相比，GCIFSMPA 在最具挑战性的 allsearch 模式单次设置下实现了 74.83% 的 Rank-1 准确率和 71.54% 的 mAP，分别比 MPANet 高 4.25% 和 3.30%。同时，我们的模型 GC-IFSMPA⋆ 达到了 88.95% 的 Rank-1 准确率和 86.67% 的 mAP。所提出的策略的有效性得到了证明

RegDB 上的比较。我们在表二中展示了该模型在 RegDB 数据集上的实验结果。我们观察到 GC-IFS 策略也比 SOTA 方法获得了更好的性能。我们还在几个基线模型上进行了实验，(a) 与基线 AGW† 相比，GC-IFSAGW 优于基线 AGW†，在可见光到热模式下具有 +19.36% mAP 和 +18.27% Rank-1，在热到可见模式下具有 +19.53% mAP 和 +18.23% Rank-1。此外，我们的模型 GC-IFSAGW⋆ 在可见光到热模式下达到 97.21% Rank-1 和 97.50% mAP，在热到可见模式下达到 96.22% Rank-1 和 96.82% mAP。 (b) 与基线 MPANet 相比，GC-IFSMPA 在可见光到热敏模式下实现了 94.40% Rank-1 和 92.19% mAP，在 Rank-1 上超出基线 10.70%，在 mAP 上超出基线 11.29%，在热敏模式下实现 92.87% Rank-1 和 91.00% mAP。本文已被 IEEE Transactions on Circuits and Systems for Video Technology 接受发表。这是作者的版本，尚未经过完全编辑，内容在最终发布之前可能会发生变化。引文信息：DOI 10.1109/TCSVT.2023.3287300 到 Visible 模式，在 Rank-1 上超过基线 MPANet 10.07%，在 mAP 上超过 10.30%。 GC-IFSMPA⋆在可见光到热模式下达到97.50% Rank-1和97.80% mAP，在热到可见模式下达到96.62% Rank-1和97.08% mAP。对于 Rank-20，我们的模型甚至可以达到 100% 的准确率。

表：SYSU-MM01数据集

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Method** | **All search** | | | | **Indoor search** | | | |
| Rank-1 | Rank-10 | Rank-20 | mAP | Rank-1 | Rank-10 | Rank-20 | mAP |
| Zero-padding | 14.80 | 54.12 | 71.33 | 15.95 | 20.58 | 68.38 | 85.79 | 26.92 |
| HCML | 14.30 | 53.20 | 69.20 | 16.20 | 24.50 | 73.30 | 86.70 | 30.10 |
| BDTR | 17.0 | 55.40 | 72.00 | 19.70 | -- | -- | -- | -- |
| D2RL | 28.90 | 70.60 | 82.40 | 29.20 | -- | -- | -- | -- |
| AlignGAN | 42.40 | 85.00 | 93.70 | 40.70 | 45.90 | 87.60 | 94.40 | 54.30 |
| X-Modality | 49.92 | 89.79 | 95.96 | 50.73 | -- | -- | -- | -- |
| DDAG | 54.75 | 90.39 | 95.81 | 53.02 | 61.02 | 94.06 | 98.41 | 67.98 |
| LADN | 53.60 | 90.92 | 96.30 | 52.00 | 57.00 | 92.80 | 98.30 | 63.20 |
| AGW | 47.50 | 84.39 | 92.14 | 47.65 | 54.17 | 91.14 | 95.98 | 62.97 |
| NFS | 56.91 | 91.34 | 96.52 | 55.45 | 62.79 | 96.53 | 99.07 | 69.79 |
| CM-NAS | 61.99 | 92.87 | 97.25 | 60.02 | 67.01 | 97.02 | 99.32 | 72.95 |
| FMCnet | 66.34 | -- | -- | 62.51 | 68.15 | -- | -- | 74.09 |
| TSME | 64.23 | 95.19 | 98.73 | 61.21 | 64.80 | 96.92 | 99.31 | 71.53 |
| MID | 60.27 | 92.90 | -- | 59.40 | 64.86 | 96.12 | -- | 70.12 |
| **ours** | 67.08 | 92.74 | 97.08 | 63.93 | 72.42 | 97.92 | 99.73 | 77.47 |

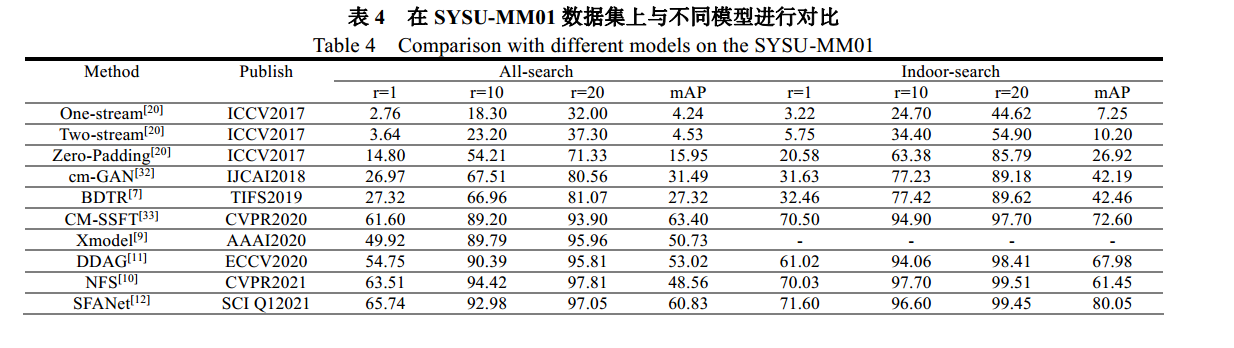


表2 Regdb数据集

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Method** | **Visible to Infrared** | | | | **Infrared to Visible** | | | |
| Rank-1 | Rank-10 | Rank-20 | mAP | Rank-1 | Rank-10 | Rank-20 | mAP |
| Zero-padding | 17.80 | 34.20 | 44.40 | 18.90 | 16.60 | 34.70 | 44.30 | 17.80 |
| HCML | 24.40 | 47.50 | 56.80 | 20.80 | 21.70 | 45.00 | 55.60 | 22.20 |
| BDTR | 33.47 | 58.42 | 67.52 | 31.83 | 68.06 | 85.15 | 90.31 | 61.80 |
| D2RL | 43.40 | 66.10 | 76.30 | 44.10 | -- | -- | -- | -- |
| AlignGAN | 57.90 | -- | -- | 53.60 | 56.30 | -- | -- | 53.40 |
| X-Modality | -- | -- | -- | -- | 62.21 | 83.13 | 91.72 | 60.18 |
| DDAG | 69.34 | 86.19 | 91.49 | 63.46 | 68.06 | 85.15 | 90.31 | 61.80 |
| LADN | 75.70 | -- | -- | 72.90 | 75.30 | -- | -- | 73.00 |
| AGW | 70.05 | -- | -- | 66.37 | -- | -- | -- | -- |
| NFS | 80.54 | 91.96 | 95.07 | 72.10 | 77.95 | 90.45 | 93.62 | 69.79 |
| CM-NAS | 82.80 | 95.10 | 97.70 | 79.30 | 81.70 | 94.10 | 96.90 | 77.60 |
| FMCnet | 89.12 | -- | -- | 84.43 | 88.38 | -- | -- | 83.86 |
| TSME | 87.35 | 97.10 | 98.90 | 76.94 | 86.41 | 96.39 | 98.20 | 75.70 |
| MID | 87.45 | 95.73 | -- | 84.85 | 84.29 | 93.44 | -- | 81.41 |
| **ours** | 89.66 | 96.61 | 98.45 | 83.43 | 85.19 | 94.61 | 97.31 | 79.44 |

# 消融实验

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Method** | **All search** | | | | **Indoor search** | | | |
| Rank-1 | Rank-10 | Rank-20 | mAP | Rank-1 | Rank-10 | Rank-20 | mAP |
| Baseline | 58.66 | 90.55 | 96.89 | 53.76 | 61.63 | 94.64 | 98.34 | 67.66 |
| Baseline+MAML | 62.40 | 90.82 | 96.06 | 58.44 | 64.09 | 94.57 | 97.19 | 70.79 |
| **Baseline+MAML+AMG** | 67.08 | 92.74 | 97.08 | 63.93 | 72.42 | 97.92 | 99.73 | 77.47 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Method** | **All search** | | | |
| Rank-1 | Rank-10 | Rank-20 | mAP |
| Baseline | 58.66 | 90.55 | 96.89 | 53.76 |
| Baseline+MAML | 62.40 | 90.82 | 96.06 | 58.44 |
| **Baseline+MAML+AMG** | 67.08 | 92.74 | 97.08 | 63.93 |

本文在SYSU-MM01[1]和RegDB[2]两个公开数据集进行实验，使用Rank-n、mAP作为评价指标。

SYSU-MM01数据集包括由2台红外相机和4台可见光相机拍摄到的491个行人、303420张图像。其中三个相机采集室内图像，三个相机采集室外图像，使其能更好的模拟真实场景。使用296个身份进行训练，99个身份进行验证，96个身份进行测试。在测试阶段，使用可见光图像作为图像数据库，红外图像作为待检索的目标图像库。

RegDB数据集使用双可见光和热成像相机同时捕捉人体的可见光和红外图像，捕获了一个412人的图像数据库，对于每个人，捕捉了10张可见光图像和相应的10张红外图像。每个身份的红外图像和可见光图像的相机视角、天气状况、姿态都是一一对应的，相对SYSU-MM01数据集而言，降低了跨模态行人重识别任务难度。使用206个身份进行训练，206个身份进行测试。在测试阶段，使用可见光图像作为图像数据库，红外图像作为待检索的目标图像库。

为了评估模型性能，并与之前的工作进行比较，在两个公开的可见光红外行人重识别数据集上进行实验，这两个数据集分别为RegDB[31]和SYSU-MM01[20]。

SYSU-MM01数据集是由两个红外摄像机和四个可见光摄像机拍摄组成的大型数据集，其中包括287628张可见光图像和15792张近红外图像，包含491个行人身份。训练集有395个行人身份，包含22258张可见光图像和11909张近红外图像；测试集有96个行人身份，具有两种不同的评估模式，分别是全局搜索和室内搜索。这两种搜索模式的查询集均为近红外图像，有3803张。在全局搜索模式下，图库集包含四个可见光摄像头采集的所有可见光图像。在室内搜索模式下，图库集只包含两个可见光摄像头采集的可见光图像。

RegDB数据集是由一个可见光摄像头和红外摄像头捕获的小规模数据集，其中总共包含412个行人身份，每个行人有10张可见光图像和10张红外图像。该数据集具有两种评估模式，分别是可见光图像查询红外图像和红外图像查询可见光图像。

Thesis revision 研究生毕业论文修改

I am a PhD candidate majoring in cancer biology and trying to revise my thesis.In this process,I need you to act as an expert professor of the same research filed and I will provide you with some paragraphs in Chinese and your task is to accurately, logically and academically revise the sentences. I want you to give your output in Chinese and in a markdown table where the first column is the original sentence and the second is the first version of revision and third column is the detailed reason to make this revision, and give each row only one sentence. tyou understand the above task, please reply with ves. and then I will provide you with the paragraphs