

Learning Longterm Representations for Person Re-identification Using Radio Signals

*2020 IEEE/CVF Conference on Computer Vision and Pattern
Recognition (CVPR)*

Lijie Fan and Tianhong Li, Member, MIT CSAIL

汇报人：潘承安

本文用RF信号实现了人物重识别的新方法方法，与以往利用图像或视频数据进行人物重识别的方法相比，利用RF信号具有以下优势：

- 1、由于RF信号的穿透性，可以在有遮挡和黑暗环境下进行人物识别，对环境不敏感。
- 2、无需获取人物的具体图像，保护对象隐私。
- 3、可以提取人的内在特征（如身体形状、大小）而不是根据所穿衣物。

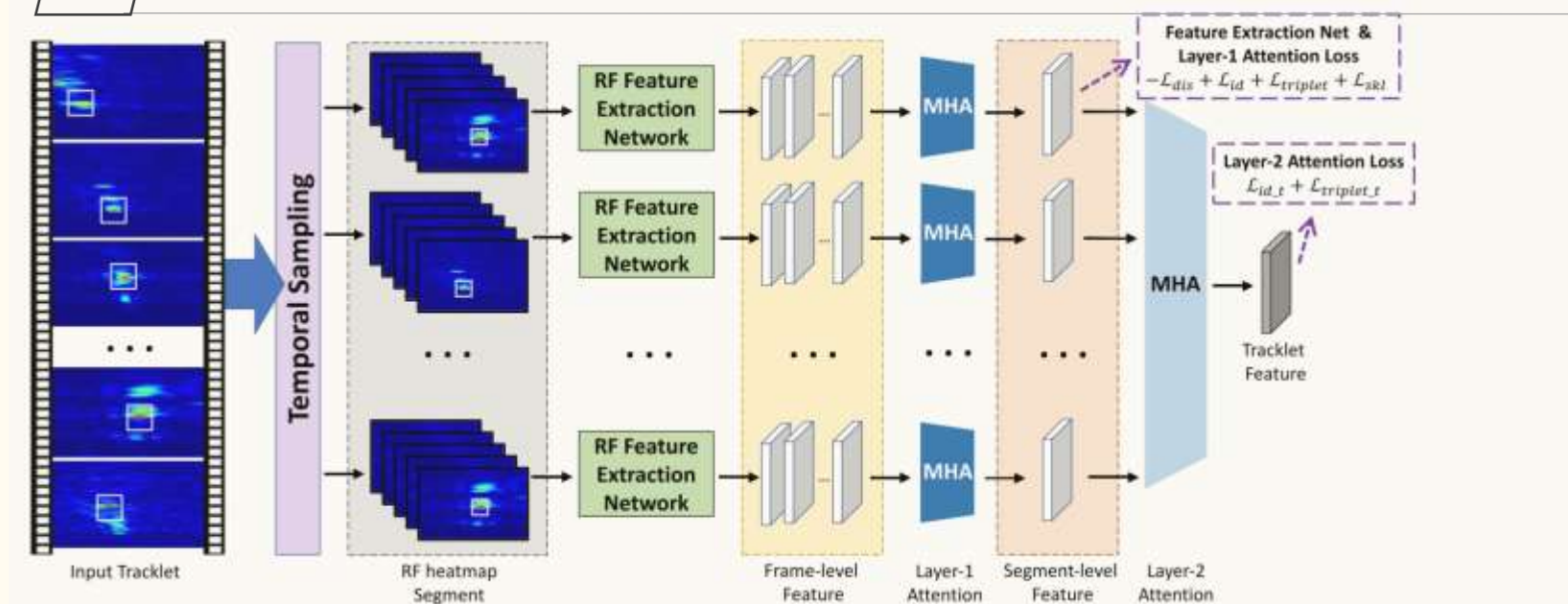
此外，本文还添加了

- ①跟踪人物骨架和行走模式的框架，作为副任务辅助进行人物识别，用于长期识别人物特征。

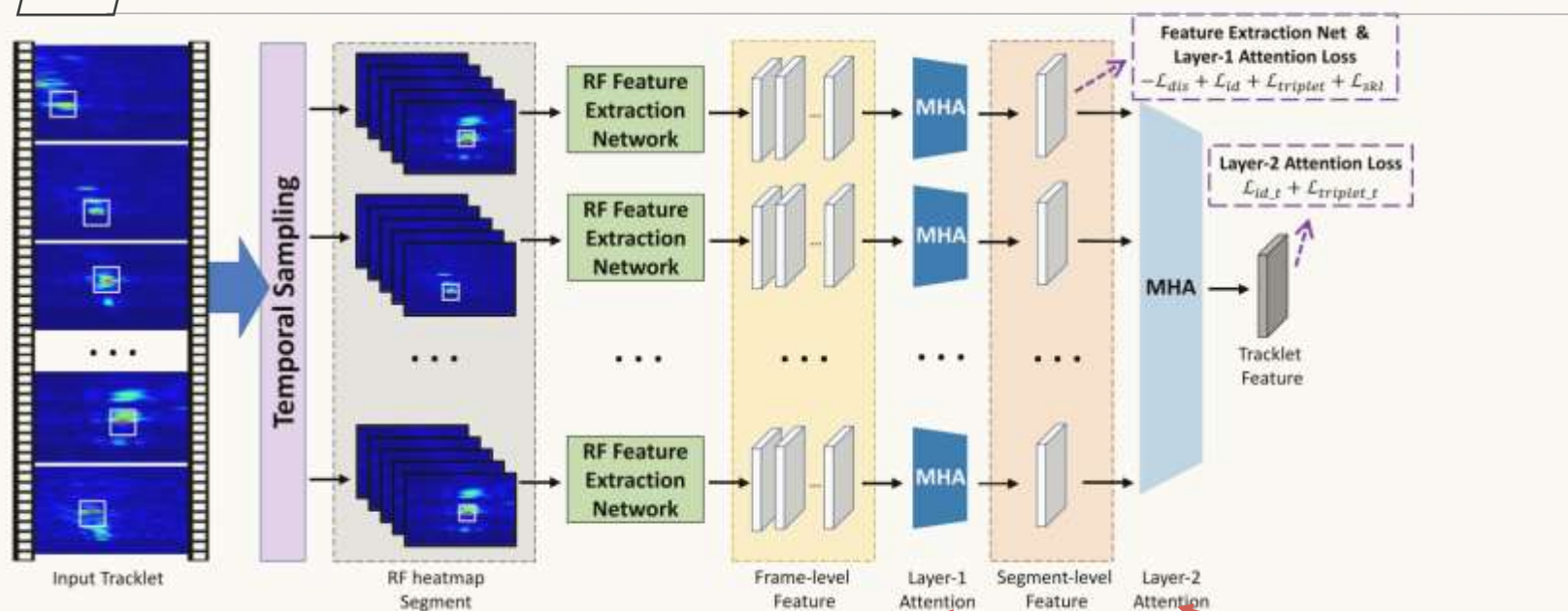


- ②环境区分器，帮助模型无法根据环境区分特征，降低了环境因素对模型准确率的影响。

- ③分层注意力模块，解决了人体的镜面反射导致信息不足的问题。



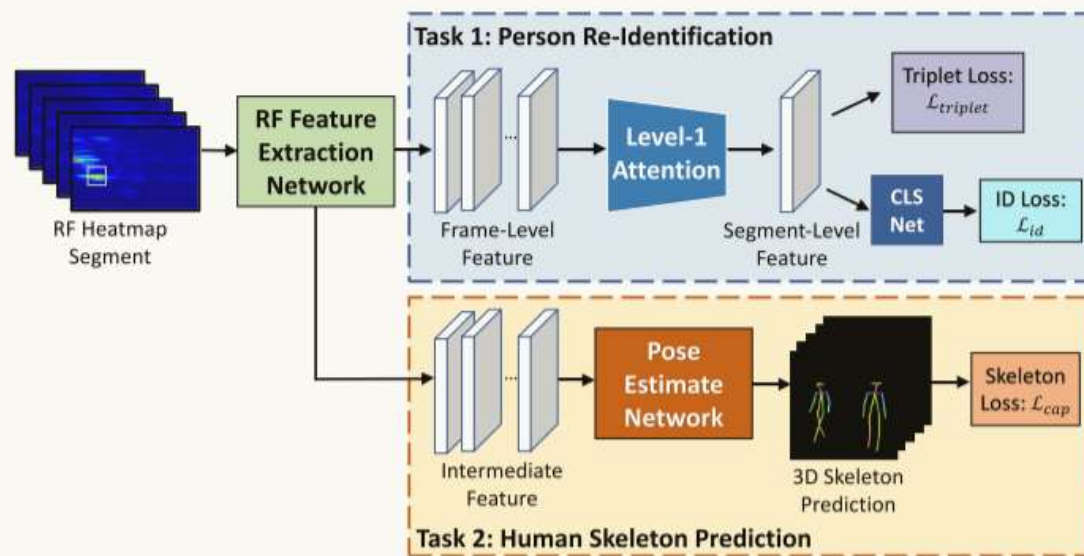
输入为RF轨迹，对每个轨迹均匀采样25个持续3秒（90帧）的片段。
该网络首先使用时空卷积从RF帧中提取全局特征。然后，在特征图中裁剪出轨迹周围的关注区域。最后，将裁剪的特征馈送到子网络中以生成帧级可识别特征。



RF轨迹片段中有两种与人员识别相关的信息：形状和行走方式。文章提出了一个两步分层注意力模块来聚合每个轨迹的信息。

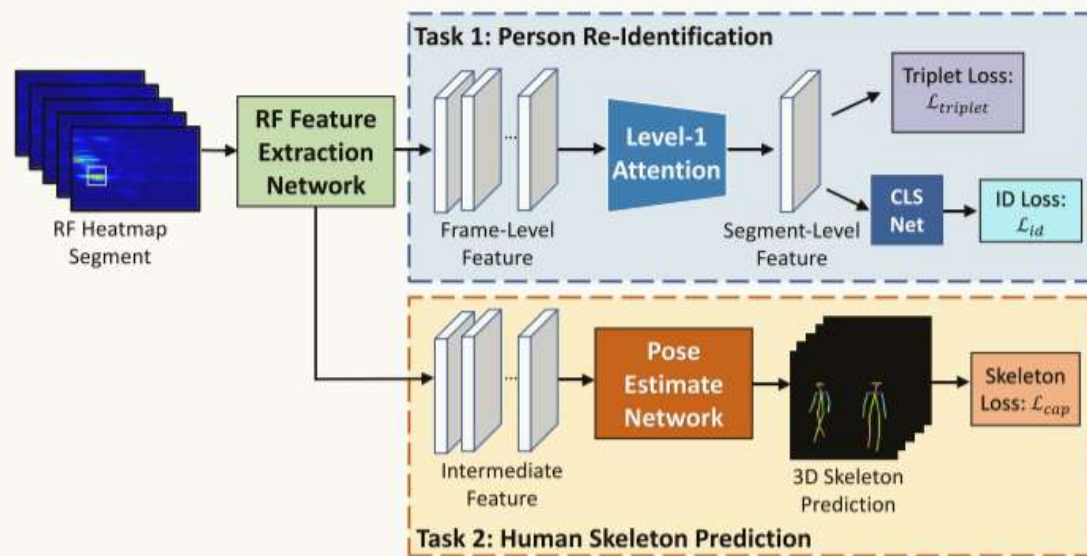
人的粗略形状可以通过聚合来自几秒的RF信号的信息来获得。因此，第一层注意力模块被添加到帧级特征上，以聚合每个90帧段（3秒）内的形状信息。

另一方面，行走风格是只能从较长的时间跨度推断的特征。因此使用第二个注意力模块来关注来自轨迹片段上不同片段的特征，并将它们聚合以生成每个轨迹片段的一个最终特征向量。

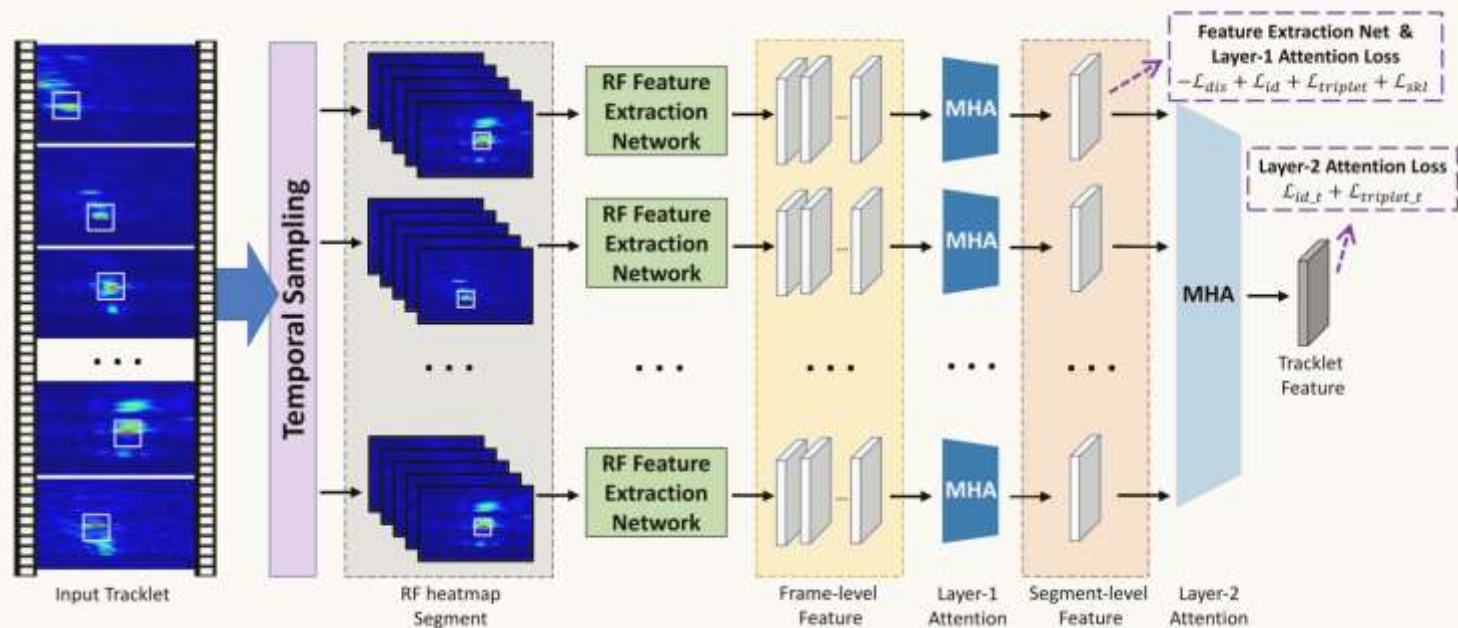


首先添加了在人物重识别中广泛使用的两个损失：鉴定损失 L_{id} 和三元组损失 $L_{triplet}$ 。对于 L_{id} ，分段级特征进一步通过具有两个全连接层的另一分类网络来执行分类。该任务有助于模型从RF信号中学习人类可识别信息。

三元组损失通过公式 $L_{triplet} = \max(d_p - d_n + \alpha, 0)$ 计算得到。其中 d_p 和 d_n 分别是来自同一人的分段级特征和来自不同人的特征的欧式距离。 α 是三重态损失的裕度。这种损失迫使来自不同人的特征彼此远离，而来自同一个人的特征彼此接近。



此外，文章迫使模型学习从RF信号推断人的骨架，并添加额外的骨架损失 L_{skl} 用于监督。具体来说，我们将特征提取网络中帧级特征之上两层的中间特征送到生成3D人体骨架的姿态估计子网络中。骨架损失使用二进制交叉熵损失。这部分使得特征中包含用于骨架生成的足够信息，可以帮助重识别任务捕获人的身高和行走风格。骨架损失还充当对提取的特征的正则化器以防止过拟合。



RF轨迹可以具有与环境强烈相关的可识别模式。模型可能使用环境特征或人们的行走路径作为快捷方式来进行识别。这将严重损害模型在不同环境中泛化的能力，或者当人们不遵循通常的路径时识别他们的能力。因此将来自每个无线电位置的信号视为一个环境，并训练一个鉴别器来识别信号的环境。该模型以对抗性的方式进行训练，因此最终模型将消除依赖于环境的特征。该函数对段级特性进行操作。

鉴别器的损失将从多任务模块的损失中被减去，将RF特征提取网络表示为 F ，将环境SNR表示为 D ，则SNR损失为： $L_{dis} = -\sum_{c=1}^M y_c \log(D(F(x))_c)$ ，其中 x 是输入RF轨迹， M 是环境的总数， y_c 是当前轨迹所属环境的指示符。优化目标为： $\min_F \max_D V(F, D) = -L_{dis} + L_{id} + L_{triplet} + L_{skl}$

文章采用两个数据集：RRD-Campus和RRD-Home。

RRD-Campus是指在我们校园收集的RF-ReID数据集。它包含RF信号和RGB视频数据。视频数据既用于真值标记，也用于与基于RGB的视频和基于图像的ReID方法进行比较。数据包含同一批人经常经过的地方，总共包含100个身份。平均而言，每个身份有8.63个RF 轨迹，每个轨迹的超过11.0秒。

RRD-Home基于先前使用RF信号进行帕金森病分析的工作中的数据集。该数据集是通过在19个不同的家庭中部署RF设备来收集的，以获得用于移动人员的RF轨迹，其中每个家庭均包含帕金森患者和正常人。要求帕金森患者佩戴加速度计以收集加速度数据。该数据集包含19个不同家庭的38个不同身份。数据跨度为127天，平均每家一周。每个身份平均具有165.91个RF轨迹。每个RF 轨迹超过9.24秒。

与其他模型的比较。

传统的基于RGB的ReID模型侧重于从衣服中提取特征，当同一个人穿着不同的衣服时会失败。相比之下，RF-ReID专注于一个人的形状和行走方式，这些在很长一段时间内仍然有效。

Methods	Modality	RRD-Campus			RRD-Home		
		mAP	CMC-1	CMC-5	mAP	CMC-1	CMC-5
Luo <i>et al.</i> [30]	RGB Image	41.3	61.4	84.3	-	-	-
Gao <i>et al.</i> [11]	RGB Video	48.1	69.2	89.1	-	-	-
RF-ReID (separate)	RF Signal	59.5	82.1	95.5	46.4	74.6	89.5
RF-ReID (combined)	RF Signal	60.7	83.6	96.5	49.4	75.8	92.5

消融实验

多任务学习模块

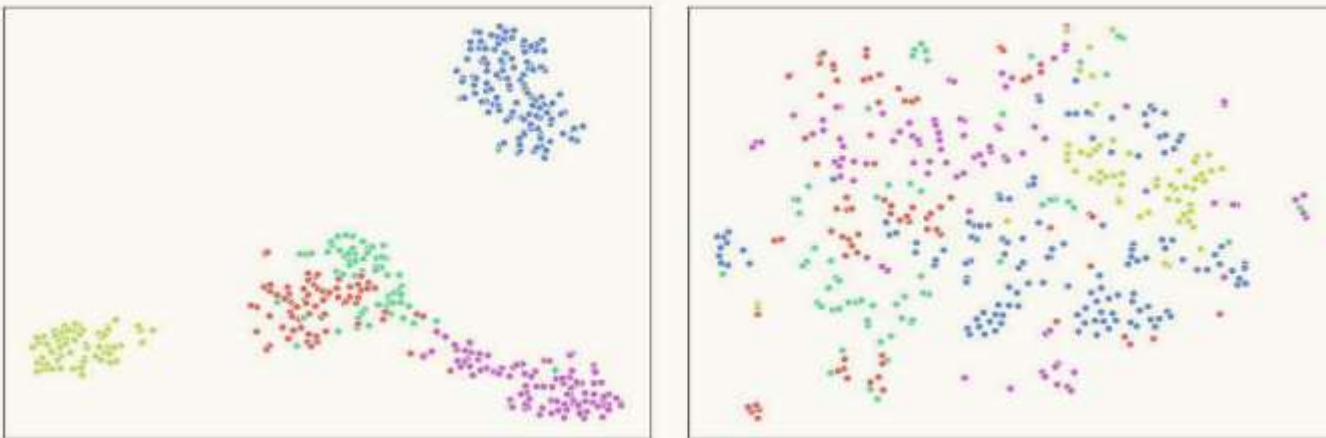
Method	mAP	CMC-1	CMC-5	MPJPE
w/o skl loss	57.3	78.2	94.4	-
w/ skl loss	59.5	82.1	95.5	7.44

分层注意力模块

Method	mAP	CMC-1	CMC-5
Avg Pool+Avg Pool	55.4	74.5	93.8
Avg Pool+2nd Att.	58.3	80.3	94.4
1st Att.+Avg Pool	58.6	81.0	94.3
1st Att.+2nd Att	59.5	82.1	95.5

环境鉴别器

Method	mAP	CMC-1	CMC-5
w/o discriminator	56.7	74.2	93.3
w/ discriminator	59.5	82.1	95.5



RF-ReID从不同环境中提取的特征分布。颜色表示所属的环境。左边的子图显示了没有约束的话来自相同环境的特征由于环境相关信息而被强烈聚类。右边的子图显示了具有环境鉴别器的特征分布。在这里，特征的间隔更均匀，表明它们更具有环境不变性。



对于大部分情况，RF-ReID都能准确地进行身份重识别，即使在黑暗或遮挡（第三行）环境中也能展现出优异的表现。但是在某些情况下可能会失败，如最后一行所示。左边的示例中的人正在骑自行车。自行车改变了人的行走方式；由于它的金属框架，还干扰了RF的反射，导致不准确的预测。另一个失败的情况是当人在滑板上时。由于RF-ReID专注于人的行走风格，因此无法正确识别此人，这也是RF-ReID的局限性所在。

通过使用RF信号进行人物身份重识别的新方法可以从RF信号中提取**长期识别特征**，从而使人能够在几天，几周等时间内重新识别。

此外，与相机不同，RF信号**不会泄露私人或个人信息**。因此这种重识别方法可用于医疗保健应用和临床研究。

最后，RF信号可以在存在**遮挡和光线不足**的情况下工作，能够在这种情况下进行识别。

本文的工作为需要重识别的许多应用提供了可行的新方法。

**THANK
YOU**

