

基于深度学习的 EMG 信号跨域动作识别技术方案

宫文娟

青岛软件学院、计算机科学与技术学院
中国石油大学（华东）

2022-8-9

技术方案

1	问题描述	1
2	解决方案	1
2-1	基于迁移学习的动作识别	1
2-2	基于元学习的动作识别	1
2-3	基于多模态融合的动作识别	2
2-3-a	特征级融合	2
2-3-b	模型级融合	2
2-3-c	特征的模型级融合	2
3	研究方法	2
3-1	数据收集	2
3-2	EMG 信号预处理	3
3-3	数据校准	3
3-4	数据增强	3
3-5	模型结构设计	4
3-6	优化目标函数设计	4
3-6-a	分类损失	4
3-6-b	回归损失	4
3-6-c	无监督学习网络损失	4
3-7	方法验证	4
	附录 A: EMG 信号处理相关工作	4
	附录 B: 迁移学习、少样本学习和多模态融合前期工作	4
	参考文献	4

图

1	基于深度学习的迁移学习示意图	1
2	元学习整体示意图	1
3	基于关系网络的元学习网络架构示意图（来自网络）	1
4	特征融合示意图	2
5	模型融合示意图	2
6	相关工作中基于深度学习的 EMG 信号动作识别中采用的模型类别及效果	3
7	语音处理中最常用的特征-梅尔频谱倒谱系数的计算流程示意图（来自网络）	4

表格

1	EMG 动作识别公开数据库	2
---	---------------	---

基于深度学习的 EMG 信号跨域动作识别技术方案

1. 问题描述

- 目标:** 针对 EMG 信号, 我们将设计深度学习算法, 解决跨年龄、跨性别、不同体型、不同时段、不同使用阶段、传感器偏移等情况下的跨域动作识别问题。
- 定义:** 目前大部分的 EMG 信号动作识别问题都定义成分类问题; 识别问题也可以建模成查询问题, 对于大量数据的查询, 聚类处理和决策森林模型可能会提高识别效率。
- 注 1:** 经过阅读 EMG 信号处理领域相关工作, 我们发现目前大部分工作都集中在动作识别方向。与动作识别非常密切的一个研究方向是动作检测, 即在 EMG 信号序列中将行为信号所在的时间窗口定位出来。
- 注 2:** 需要对相关医学文献进行调研, 了解不同动作中肌肉运动导致的 EMG 信号特点, 以便进一步对信号进行校准。

2. 解决方案

1. 基于迁移学习的动作识别

基于深度学习和迁移学习的方法在 EMG 信号手部运动解码、跨受试者动作识别等方向都有很好的效果。



图 1. 基于深度学习的迁移学习示意图

基于深度学习的方法是一种非常自然的迁移学习的框架。利用现有的数据进行训练和优化的深度学习模型, 在处理另一个受试者的数据时, 可以将分类头网络 (如图 1 中的网络分类头模块) 进行重新训练。

我们对现有的基于深度学习的方法进行了调研和总结, 如图 6。在选择基于深度学习的迁移学习的网络模型时, 我们可以参考相关工作中各个模型的效果。另外, 考虑传感器端部署问题, 我们可以对其他适合在嵌入式设备上运行的网络进行试验, 例如 MobileNet。

注: 深度学习模型特征提取 + 传统机器学习模型决策的方法很有效; 多流网络似乎效果不错; 太复杂的网络似乎不适合 EMG 信号, 可能更适合 HD-EMG 信号, 例如 3D CNN 对 HD-EMG 效果还行。

2. 基于元学习的动作识别

元学习主要解决少样本学习问题, 本课题我们将利用元学习解决如何将训练好的模型迁移到未知受试者 EMG 数据数据上的问题。

元学习主要由两个流程构成: 元学习训练和元学习测试。在元学习的训练优化过程由很多 episode 构成, 每一

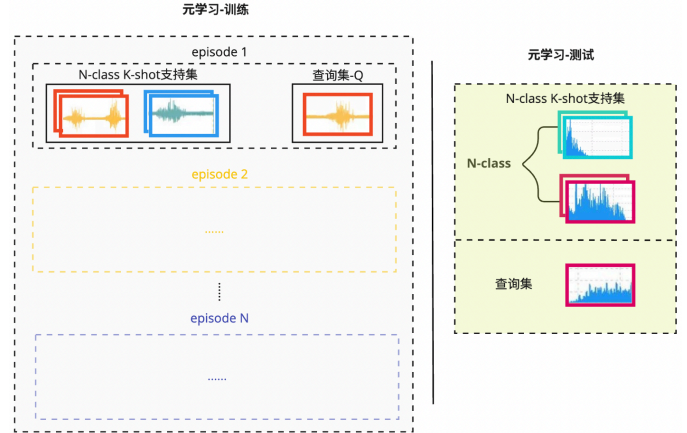


图 2. 元学习整体示意图

个 episode 计算一次整体损失并对网络模型进行一次优化; 每个 episode 都由一个支持集和一个查询集构成, 支持集包含 N 个类别, 每个类别包含 K 条数据, 支持集由 Q 条数据构成。元学习测试也被构建成为一个 N -class K -shot 的分类任务, 包含一个支持集和一个查询集。图 2 给出了 N -class K -shot 的少样本学习的元学习训练和元学习测试过程示意图。

元学习训练和元学习测试过程中的训练数据都由 N 个类别构成, 但是他们的类别一般是不同的, 在训练过程中, 模型学习了如何区分 N 个不同类别, 而不是学习某个具体类别的特征。基于元学习的思想, 我们将在元学习训练过程学习收集到的受试者的 EMG 数据, 在元学习测试过程中对不同受试者的 EMG 数据进行分类。

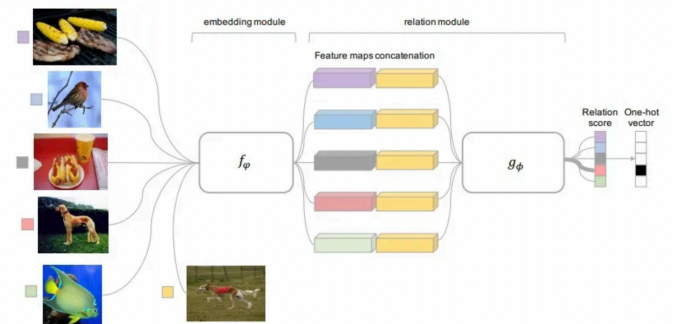


图 3. 基于关系网络的元学习网络架构示意图 (来自网络)

基于元学习的少样本学习主要有两种类别的方法构成: 基于度量和基于梯度的方法。图 3 给出了一种基于度量的

元学习方法的架构示意图。

3. 基于多模态融合的动作识别

不同传感器的信号或者从同一类信号提取的不同的特征和对其采用的不同的深度学习模型属于多模态信息，对这些信息进行有效地融合可以提高任务效果。例如，EMG 数据和 IMU 数据融合在手语识别、手势识别等方向取得了很好的效果。多模态融合技术可以在特征级和模型级实现。

a) 特征级融合：特征级的融合将不同类型的特征进行串联，然后送到深度学习网络进行下一步处理。对于 EMG 信号动作识别，我们可以采用的特征有：

- EMG 瞬时采样数据
- 滑动窗口数据
- 瞬时采样 + 滑动窗口。最简单的特征级融合方法是把不同的特征进行串联，然后将融合的特征作为一个特征向量整体进行处理，图 4 给出了特征融合示意图。

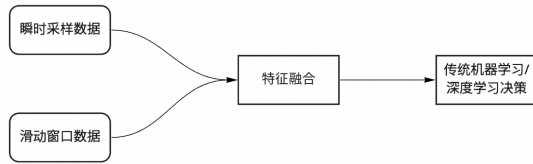


图 4. 特征融合示意图

- 瞬时采样与滑动窗口相结合。现有工作中，针对瞬时采样数据和滑动窗口采样数据是进行分别处理的，我们觉得可以将这两类方法进行结合。例如，可以先采用滑动窗口取出原始数据，然后在整个滑动窗口中取极值，例如取极大值，然后继续对这些数值进行处理。

b) 模型级融合：在模型级融合方法中，我们可以将针对不同任务的模型进行融合，例如可以对不同网络类型进行融合，例如可以将基于 MobileNet 和 ResNet 的动作分类模型进行融合；也可以将手部动作分类和手部动作角度回归任务进行融合。例如，针对不同任务的模型级融合技术方案主要包含如下模块。

- 动作分类。基于现有工作，我们可以为动作分类选择合适的网络模型，可选的网络模型如图 6 所示。
- 动作角度回归。现有工作中，有几项是针对手部进行连续角度和手指弯曲角度估计的。他们采用的都是 RNN、LSTM 等时序处理模型。另外，研究表明肌肉力量/转矩估计和肢体角度估计两个问题很相似，对手部进行动作角度回归可以参考针对肌肉力量/转矩估计的大量相关工作。

注¹：手腕运动回归是虚拟键盘和虚拟手写输入两项应用必须解决的问题。目前的相关工作有对两个和三个自由度的手腕运动进行估计的工作；根据直觉，手腕的运动应该是 4DOF 的：三维空间中的位移加旋转，如果在键盘和虚拟手写板平面上估计，可能运动自由度会降低。

注²：角度估计任务需要了解人体肌肉，Hill 模型是计算肌肉力量/转矩的基础模型。

- 多模型决策级融合。传统机器学习方法普遍地使用特征和模型融合技术；针对于基于深度学习的方法，本

研究小组以往的经验表明模型融合对多模态信息处理效果提升有很大的帮助。图 5 给出了模型融合方法的示意图。最简单常用的模型融合方式是通过超参数将不同网络的决策结果进行加权平均，得到的结果作为最终决策的依据。

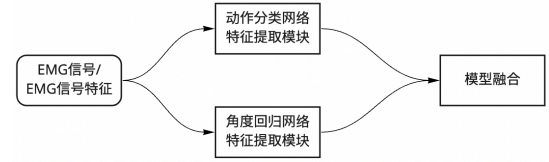


图 5. 模型融合示意图

模型级融合处理的模型甚至不需要都是深度学习模型，可以将传统机器学习模型的决策结果与深度学习模型的决策结果进行融合。

c) 特征的模型级融合：对于不同类型的特征融合，除去采用特征级融合方式，也可以将不同的特征利用不同的模型进行处理，例如将单点采样数据通过 CNN 进行处理后得到的决策结果，与将滑动窗口采样数据通过机器学习模型处理后的决策结果进行融合。

3. 研究方法

依据深度学习方法实现的一般流程，我们设计了 EMG 信号动作识别的研究方案，主要包含以下几个步骤。

1. 数据收集

现有的可以进行试验的数据库包括以下几种。

EMG 数据库	数据库构成			
	类型	类别数	健康者	截肢者
Ninapro DB1,DB2,DB3	手势识别	50	67	22
Ninapro DB8	手指角度回归	*	*	*
CSL	高密度 EMG	*	*	*
CapMyo	高密度 EMG	*	*	*

表 1. EMG 动作识别公开数据库

针对 EMG 信号动作识别任务，我们需要收集跨年龄、跨性别、不同体型、不同时段、不同使用阶段、传感器偏移等情况下的 EMG 跨域数据。

注¹：可以收集传感器错位、手环带错等异常情况下的数据。让模型学习这些数据，增强手环使用的鲁棒性。

注²：手环在使用过程中也可以随时采集数据，对本地模型进行微调，并将数据传输到服务器，增加训练数据库数据，以便后期升级时更新模型。

注³：相关工作表明手臂位置会影响信号，收集数据时需要注意这一点。

注⁴：相关工作中将疲劳程度划分成三类：疲劳、不疲劳、疲劳到疲劳的转换阶段，然后对疲劳程度进行估计。因此，收集数据时建议按照三种疲劳程度收集数据。

收集的数据集将划分成训练和测试数据。收集的数据可以统一划分也可以按照单个跨域任务划分。例如，针对跨年龄任务划分训练测试数据集时，可以取出某个年龄段的

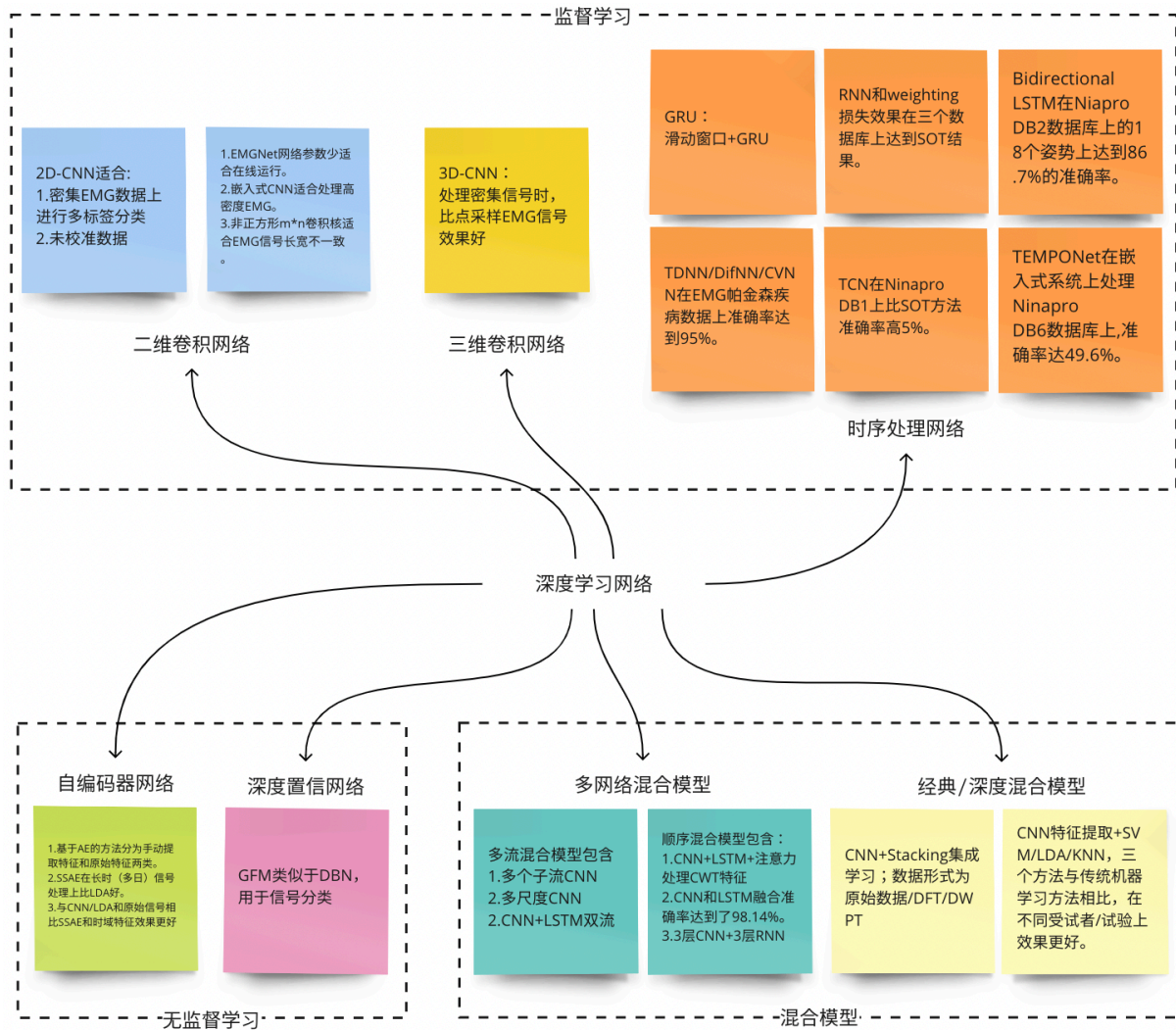


图 6. 相关工作中基于深度学习的 EMG 信号动作识别中采用的模型类别及效果

数据作为测试数据，将其他年龄段的数据作为训练数据；针对跨性别任务划分数据时，取出一个性别的数据做训练，另一个性别的数据做测试。

注：完成单个任务学习后，可以将适合不同任务的特征提取深度学习模型进行特征级/模型级融合。

2. EMG 信号预处理

主要是对原始数据进行滤波实现去噪，以及选择合适的滑动窗口大小对数据进行采样。

如果采用深度学习模型，还需要对信号进行规整，转换成类似于二维图像的形式以方便将数据以图像的规格输入到深度学习模型。

我们觉得可以引入语音处理中的特征提取方法，例如学习传统语音处理特征（梅尔频谱倒谱系数）的特征提取和表示方法，对原始特征进行提取之后输入到深度学习网络。我们可以采用梅尔频谱倒谱系数的整个流程或者梅尔谱图的计算流程，但是去掉梅尔滤波的过程，因为梅尔滤波模

拟了人耳的听觉特点对低频的数据进行增强，而对高频数据进行过滤，EMG 信息处理中不需要这个步骤。

注 1：这类特征提取方法优点是准确率高，能够过滤掉噪音等信息并保持时序关系，但是缺点是耗时。而且在短时傅里叶变换时需要窗口大小进行大量实验。

注 2：如果能增加一个模型，判断手环佩戴是否正确、传感器是否贴合或者移位，可以过滤掉一部分干扰信号。如果对手环的位置进行估计，还可以对用户提出调整建议。

3. 数据校准

在利用模型处理 EMG 信号之前，将数据信号归一到一个范围可以在一定程度上解决数据的跨域迁移问题。

4. 数据增强

可以通过添加随机噪音对数据进行增强。

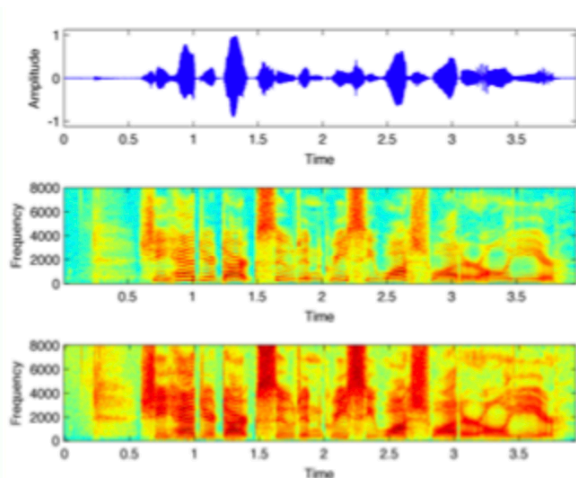


图 7. 语音处理中最常用的特征-梅尔频谱倒谱系数的计算流程示意图 (来自网络)

参考文献

- [1] Dezhen Xiong, Daohui Zhang, Xingang Zhao, and Yiwen Zhao, *Deep Learning for EMG-based Human-Machine Interaction: A Review*. IEEE/CAA Journal of automatica sinca, Vol.8, No. 3, 2021.
- [2] Wenjuan Gong, Yue Zhang, Wei Wang, Peng Cheng, Jordi González, *Meta-MMFNet: Meta-Learning Based Multi-Model Fusion Network for Micro-Expression Recognition*. ACM Transactions on Multimedia Computing, Communications, and Applications (2022).
- [3] Wenjuan Gong, Qingshuang Yu, *A Deep Music Recommendation Method Based on Human Motion Analysis*. IEEE Access 9: 26290-26300 (2021).

5. 模型结构设计

注: 现有的工作表明, 对 EMG 数据不适合使用复杂的模型。所以, 在模型结构设计上, 应该尽量采用相对简单的模型

6. 优化目标函数设计

- a) 分类损失: 多类别分类问题最常用的损失就是 softmax 损失。
- b) 回归损失: 平均均方误差损失对回归问题非常有效。
- c) 无监督学习网络损失: 对比损失通常用来优化无监督学习网络模型参数。

7. 方法验证

将训练好的模型对未知的受试者进行识别, 通过动作识别的准确率、Recall 等标准对方法进行验证。

附录 A EMG 信号处理相关工作

1. EMG 信号处理综述 [1]
2. Research page from Ulysse-Cote-Allard: <https://www.researchgate.net/profile/Ulysse-Cote-Allard>

附录 B 迁移学习、少样本学习和多模态融合前期工作

1. 在微表情识别中采用了特征融合和模型融合方法, 效果都很好 [2]。
2. 采用基于元学习的少样本学习方法进行微表情识别, 利用某些受试者的数据进行训练, 用剩下一个受试者的数据进行测试 [2]。
3. 对于舞蹈动作、音乐进行多模态融合, 效果很好 [3]; 对人脸、整体场景、人体姿态进行多模态融合, 效果良好; 也有对视频和音乐进行融合的工作。