

作业 6

吴同

2022 年 5 月 31 日

调研报告部分

1 调研报告

调研：基于咳嗽数据的新冠检测方法

随着新冠疫情的常态化，核酸检测作为一种预筛选的抗疫手段已经成为人们生活的一部分。然而传统的核酸检测既需要消耗大量的物资和试剂，给社会带来了极大的负担；也需要人们到检测点进行检测或者购买检测物资，给人们的生活带来了不便。因此，一种新的、人们方便进行且不会消耗大量物资的预筛选方法已然成为社会迫切的需求。而新冠作为呼吸道疾病，对人们的呼吸系统会有病理性的影响，因此，咳嗽数据或许能成为检测新冠的一种方法。

在这种迫切期待下，Imran et al. (2020) 首先开发出了 AI4COVID-19 系统，基于收集的新冠患者的咳嗽数据和百日咳、支气管炎患者的咳嗽数据以及正常人或者其它疾病患者的咳嗽数据训练系统，取得了良好的结果。他们提出的模型采用了三个独立的分类器，两个采用深度迁移学习并以原始数据的梅尔频谱作为输入，分别输出 4 种分类和 2 种分类；另外一个采用传统机器学习，利用梅尔频率倒谱系数和 PCA 分析提取特征，最后利用 SVM 进行分类。该模型的最终输出采用“第二意见”机制对三个分类器的输出进行综合。此外，Laguarta et al. (2020) 基于已有的阿尔兹海默症模型，构建出了一个新的基于咳嗽数据的新冠检测模型。在该模型下，咳嗽数据首先被分成 6s 的块，随后得到其梅尔频率倒谱系数，输入到肌肉退化模型，然后经过三个并行的模块进行进一步的特征提取，分别为：声带模型、情绪模型和肺及呼吸道模型。将三个模型的输出拼接并经过后续的稠密神经网络得出最终的诊断。该模型也取得了较好的效果。

对于 AI4COVID-19 模型，其训练的数据较少，且数据种类少，数据质量难以保证，因此其对于新冠的检测准确率难以达到较高水平。但是，其最终的结果采用了“第二意见”机制，可以有效避免假阳的出现。对于基于阿尔兹海默症模型的新冠检测模型，它采用了更加高层次的特征，例如声带模型和肺及呼吸道模型，因此准确率应该更高。但其情绪模型的特征提

取效果应该会较差，因为 1) 新冠和情绪的关系比较弱，而且干扰较大；2) 咳嗽数据难以反映情绪。

咳嗽作为新冠检测的数据，已然取得了较好的效果。然而，咳嗽仅是呼吸道特征表征的一种方式，其它各种设计呼吸道的表征方式如呼吸、说话，都有可能成为新的表征数据。此外，除了音频数据，例如图像、视频等视觉数据也有可能成为新的数据来源（有研究表明部分新冠患者皮肤会有异常,Recalcati et al. (2020)）。进一步，如果能够普及一些廉价的生物特征检测设备，如检测手环等，其它的生物数据都有可能成为新的新冠检测数据。

编程部分

2 编程作业报告

2.1 模型的训练和验证

模型训练一共进行了 100 轮，训练集的 loss 及验证集单词识别正确率的可视化结果如图 1。验证集的最终准确率大致在 50%，最后一轮训练的验证集的准确率为 49%。由图 1 可以看出，在前 10 轮训练中，验证集的准确率基本没有提升，且基本为 0，可见此时模型并没有效果。10 轮训练后，准确率开始不断上升。在整个训练过程中，loss 下降速率在不断减小。在 80 轮训练后，loss 降低速率非常小，以至于 loss 于之前下降速率相比基本不变；并且准确率的提升速率也很小，准确率基本不变。可见，此时训练的提升能力已经很小了，可以停止训练。

2.2 使用模型预测

模型预测的对象分别为“PARKING”和“recommend”字样，分别如图 2、3。预测结果分别如图 4、5。其中，“PARKING”字样的预测结果正确，“recommend”字样的预测结果为“recomment”，最后一个字母发生了错误。由“recommend”字样预测可视化结果可以看出，最后一个字母预测为 t 的概率仅稍大于 d，可见该模型的预测能力还是较强的。

2.3 transformer 并行训练分析、transformer 解码和 CTC 解码的区别分析

Transformer 有一个很好的优点，是可以让数据并行训练，这主要取决于解码器的原理和结构。解码器解码过程中采用的也是 attention 机制，利用已解码的结果和编码器的输出得到解码器的输出。为了保持因果特性，

attention 机制往往使用掩膜，使之有时序的特性。但实际上，解码器使用的还是 attention 机制，将特征进行加权，在训练时由于本身具有所有目标数据，因此完全不需要等待上次的输出结果再进行下一次的训练，所以 transformer 可以并行训练。但是再测试过程中，由于没有目标数据，还是得等到上一次结果的输出再进行下一步的预测。

transformer 解码和 CTC 解码的区别有：1)transformer 解码可以提前终止，而 CTC 会对所有序列进行译码，这是因为 transformer 中有 <sos> 和 <eos> 两个符号作为解码过程的开始和结束，而 CTC 则是利用 blank 字符进行多余的去除，但是并不会提前结束。2)transformer 解码的每一步都是最优，因此可能得到的是局部最优，而 CTC 是全局最优，这是因为 transformer 每次都需要前一次的预测结果进行下一次预测，因此只能在每一步择优，而 CTC 采用动态规划得到全局最优，从这里看，transformer 解码带有时序的概念，而 CTC 没有。

2.4 问题及建议

本次作业没有遇到太大的问题。感谢老师和助教辛苦付出！

参考文献

- Imran A, Posokhova I, Qureshi H N, et al. Ai4covid-19: Ai enabled preliminary diagnosis for covid-19 from cough samples via an app[J]. Informatics in Medicine Unlocked, 2020, 20: 100378.
- Laguarta J, Hueto F, Subirana B. Covid-19 artificial intelligence diagnosis using only cough recordings[J]. IEEE Open Journal of Engineering in Medicine and Biology, 2020, 1: 275-281.
- Recalcati S, et al. Cutaneous manifestations in covid-19: a first perspective[J]. J Eur Acad Dermatol Venereol, 2020, 34(5).

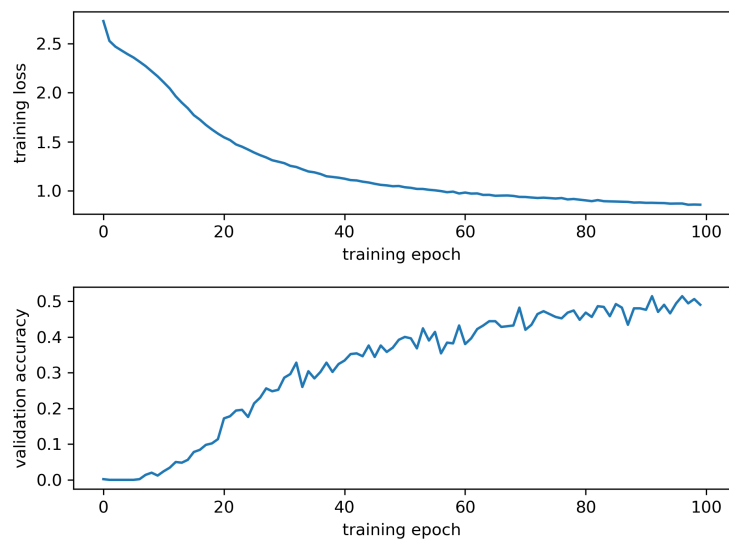


图 1: 训练和验证过程的 loss 及 accuracy 的变化曲线



图 2: “PARKING” 字样



图 3: “recommend” 字样

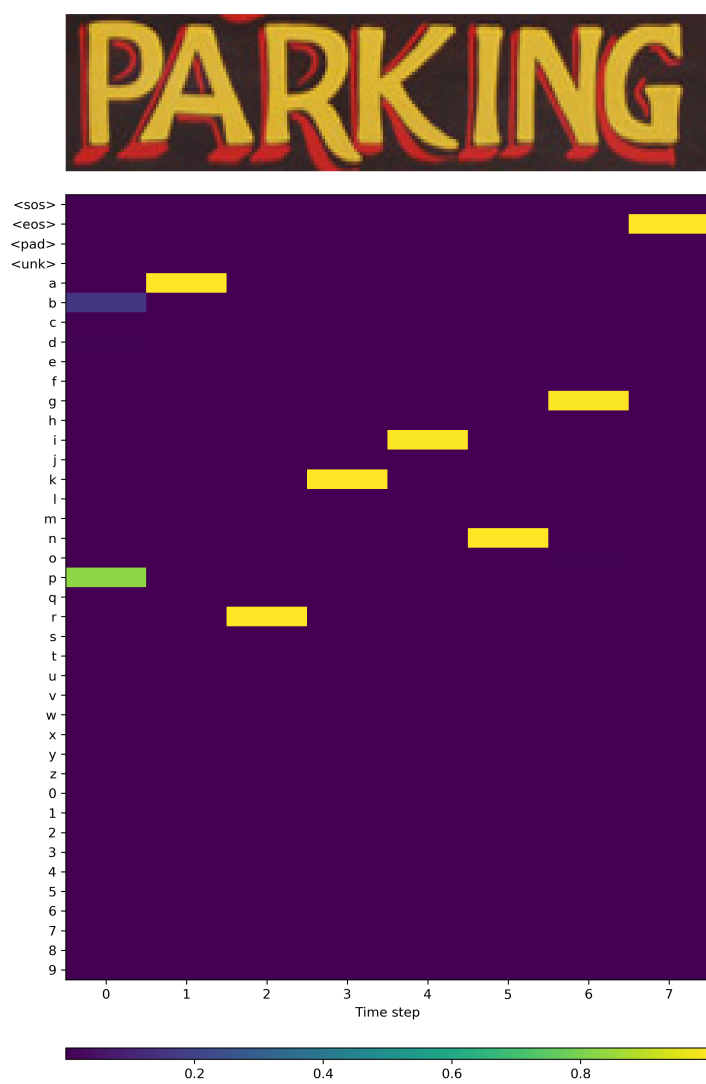


图 4: “PARKING” 字样预测可视化结果

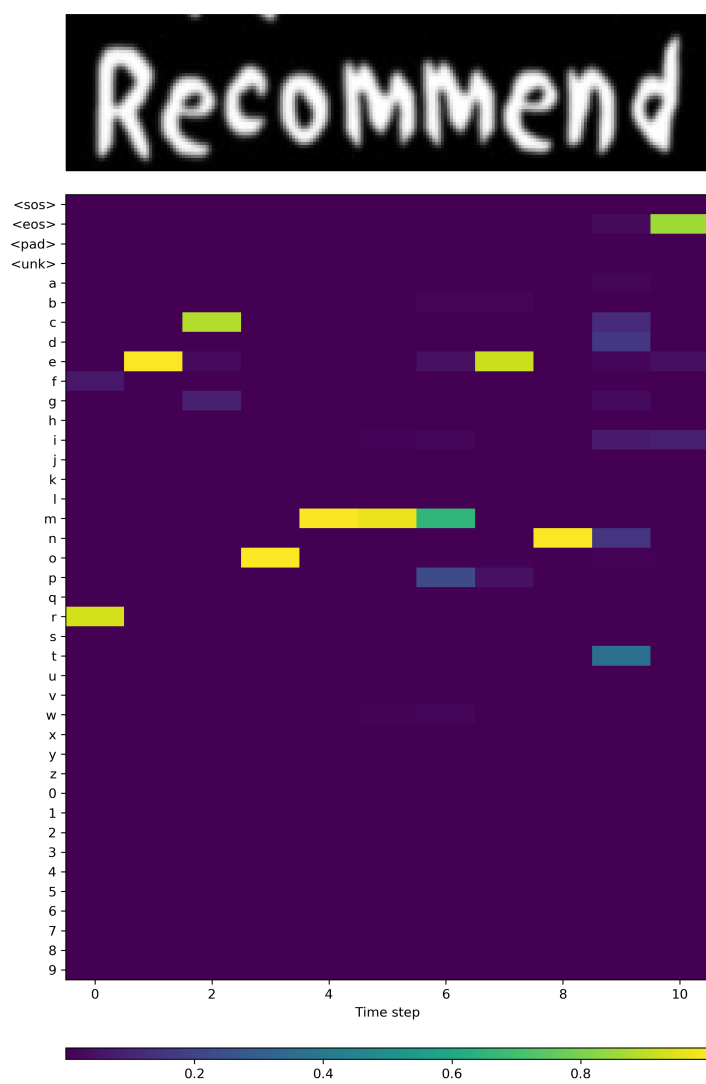


图 5: “recommend” 字样预测可视化结果