

推荐论文：

以下论文是帮助你们快速了解唇读和深度学习的工作，在此之前，希望你们可以先阅读实验室的唇读专著以构建相关概念。

下列论文我都列出了推荐理由，经典的论文太多了，很多相关领域我也是非常熟悉，比如语音识别，比如自然语言处理，很多。但唇读是必须要了解的，相关领域建议了解，因为你们不再是复刻别人的工作，你们需要去发现问题并解决它们，而这需要大量的阅读其它领域的解决方案以充实自己的知识库。对于相关领域的可以泛读，但是当你思考发现它对你解决问题有用时，建议细读。

建议每天每周都要抽出点时间阅读了解学术界、工业界的新动态（从公众号上，从学术会议或者其它渠道），你可能会得到新的启发。

一、唇读论文推荐

1、Ma P, Petridis S, Pantic M. Visual speech recognition for multiple languages in the wild[J]. Nature Machine Intelligence, 2022, 4(11): 930–939.

帝国理工大学mpc团队（截止2024. 7. 21），阅读亮点在于介绍了当前最先进唇读模型及其优化策略。

2、Prajwal K R, Afouras T, Zisserman A. Sub-word level lip reading with visual attention[C]//Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition. 2022: 5162–5172.

唇读经典模型之一，阅读亮点是其子词策略。

3、Ma P, Wang Y, Petridis S, et al. Training strategies for improved lip-reading[C]//ICASSP 2022–2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2022: 8472–8476.

帝国理工大学mpc团队（截止2024. 7. 21），阅读亮点在于唇读模型的各种训练优化策略。

4、Zhao Y, Xu R, Wang X, et al. Hearing lips: Improving lip reading by distilling speech recognizers[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2020, 34(04): 6917–6924.

唇读中的各种level的跨模态知识蒸馏，值得关注。

5、Zhang Q, Wang S, Chen G. Speaker-independent lipreading by disentangled representation learning[C]//2021 IEEE international conference on image processing (ICIP). IEEE, 2021: 2493–2497.

唇读中的说话人个性化解耦经典方案，值得关注。

6、Zhao Y, Xu R, Song M. A cascade sequence-to-sequence model for chinese mandarin lip reading[C]//Proceedings of the 1st ACM International Conference on Multimedia in Asia. 2019: 1–6.