## Transfusion: Understanding Transfer Learning for Medical Imaging

随着深度神经网络被应用于越来越多的领域，迁移学习（transfer learning）已经成为开发深度学习模型的一种非常流行的技术。在迁移学习中，神经网络的训练分为两个阶段：1）*预*训练，通常在代表大量多种标签/类别的大规模基准数据集（例如[ImageNet](https://link.zhihu.com/?target=http%3A//www.image-net.org/)）上训练神经网络；和2）*微调*，对预训练的网络在感兴趣的特定目标任务上进一步训练，可能需要比预训练数据集更少的标签样本。预训练步骤可帮助网络学习可在目标任务上重用的通用特征（general features）。

将随机初始化训练并直接应用于任务的模型与在ImageNet上预训练的，将迁移学习用于相同任务的模型进行了比较.

结论：

* 出人意料的是，迁移学习不会显著影响医学影像任务的性能，从头开始训练的模型几乎与标准ImageNet迁移模型一样好。
* 在医学影像任务上，小得多的CBR模型的性能可与标准ImageNet体系结构相当。
* 由于CBR模型比标准ImageNet模型要小得多和浅得多，因此它们在ImageNet分类上的表现要差得多，这表明ImageNet的性能*并不*能代表医学任务的性能。
* 这两个医学任务的大小比ImageNet小得多（〜200k对比〜1.2m训练图像），在非常少的数据范围内，可能只有数千个训练样本。我们在非常少的数据环境下评估了迁移学习，发现对于大型模型（ResNet），迁移学习和从头开始训练在性能上存在较大差距，而对于较小的模型（CBR）则并非如此，这表明为ImageNet设计的大型模型，对于非常少的数据而言，它可能参数过量了。