**深度学习可以从组织学角度预测肠胃癌的微卫星标记的不稳定（MSI）**

**朱梦瑶（大二 郭春明组）**

*作者：Kather, J.N., Pearson, A.T., Halama, N. et al 期刊：nature medicine*

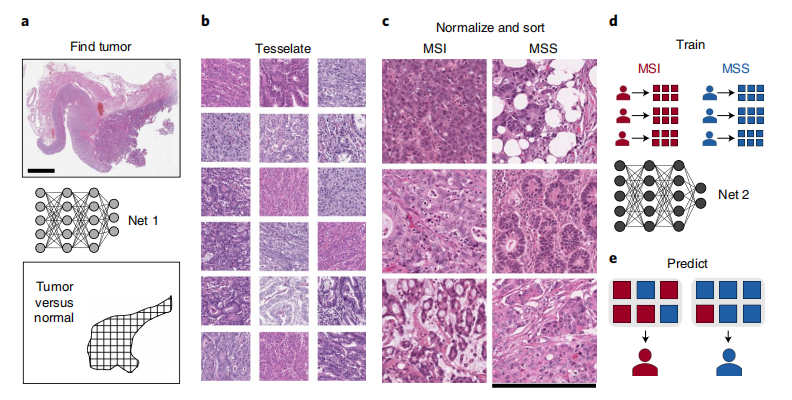
*发表日期：2019年7月 原文链接：https://www.nature.com/articles/s41591-019-0462-y#Sec9*

**背景介绍**

MSI是肿瘤的一个分子标志物，在一些肿瘤中，微卫星序列由于不稳定常出现重复、缺失等现象，这部分肿瘤便称为MSI肿瘤。

免疫疗法对于MSI肿瘤在临床上已经取得了很好的成效，MSI肿瘤约占胃腺癌（STAD）和结直肠癌（CRC）的15%。但并非所有病人都进行了MSI检测，因为这需要额外的基因或者免疫组化检测。而通过深度残留学习，可以实现在H&E染色组织上直接预测肠胃癌的微卫星标记的不稳定，可以节约大量时间金钱人力物力，为更广泛的肠道肿瘤患者提供免疫治疗。

**主要步骤**



将卷积神经网络训练成肿瘤检测器resource

A convolutional neural network was trained as a tumor detector for STAD and CRC

将肿瘤H&E染色区域切割成正方形的块

Scale bar, 4 mm. **b**,**c**, Tumor regions were cut into square tiles

对其颜色进行标准化，分为MSS和MSI

color-normalized and sorted into MSI and MSS

训练另一个卷积神经网络来分类MSI和MSS

Scale bar, 256 µm. **d**, Another network was trained to classify MSI versus MSS

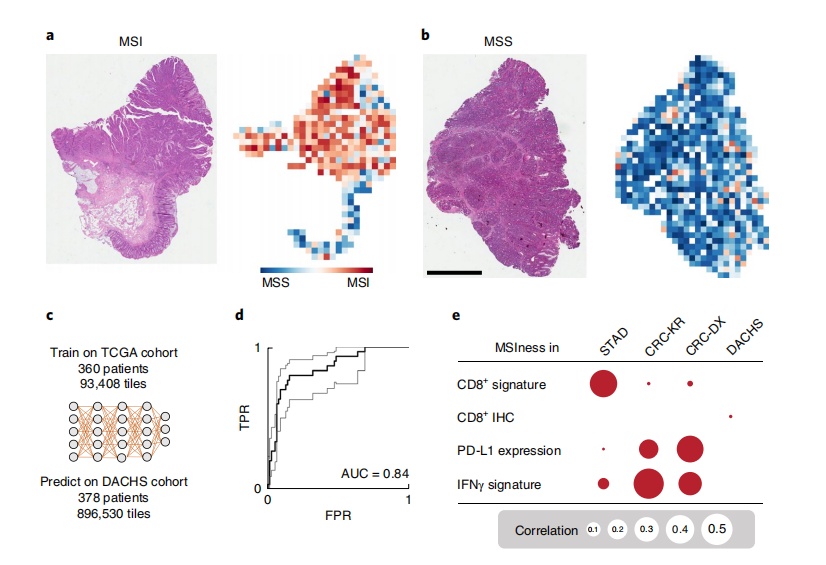
应用

This automatic pipeline was applied to held-out patient sets.

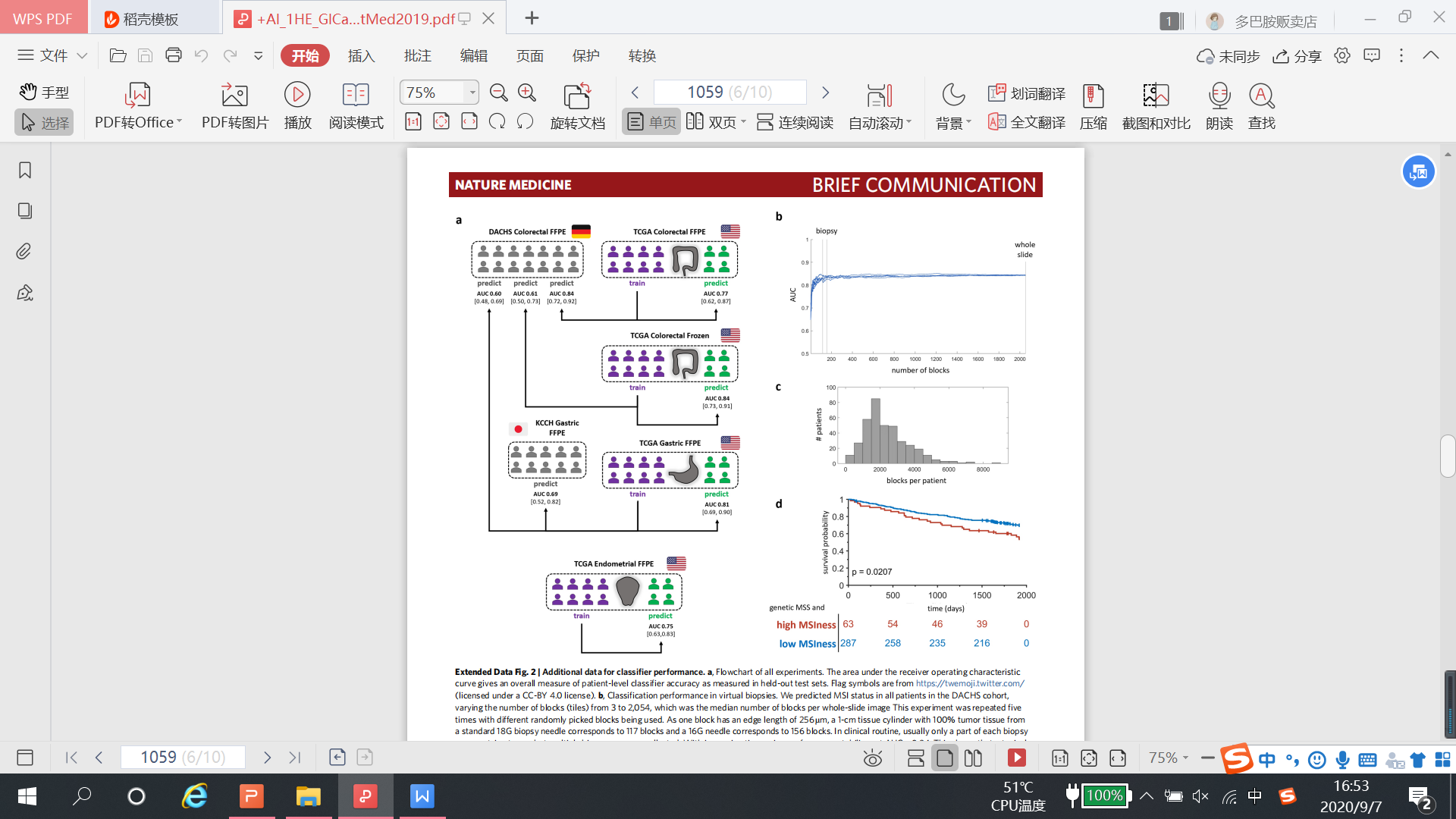
注：本文中提到，卷积神经网络resent18相较于Alexnet、VGG19更不容易过度拟合，故此实验采用的是resent18

resource收集标本，将图像区域手动标注并分类为肿瘤和非肿瘤两类（致密组织和松散组织）70%训练，5%验证，15%测试图像，用于将卷积神经网络训练成肿瘤检测器。

**实验结果：**



**目前存在的问题**

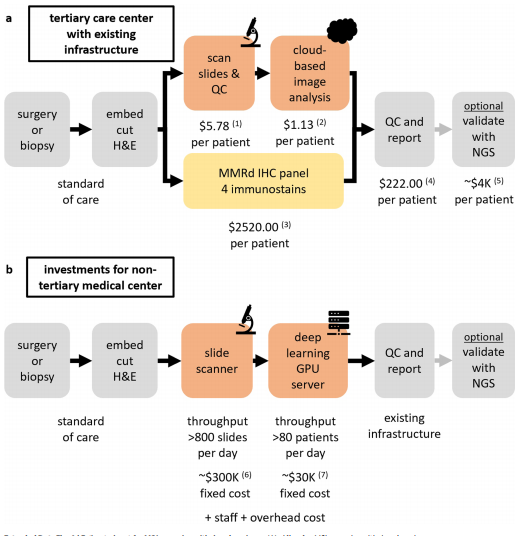


\*分类能力不会超出训练集中的癌症类型和种族，要很大的训练队列来提高分类性能，因为网络可以学习到罕见的变异体，所以需要庞大的病理组织标本。

\*所需组织大小。深度学习预测肿瘤的MSI所需组织下限大约256um，活检足以预测MSI

\*一套经过训练的卷积神经网络只能用于采集样本的相同人群（例如用美国人的样本训练出的卷积神经网络来诊断日本人，AUC值会大幅降低，诊断结果也变得不可靠）

#### 可节约的成本



#### Data availability

All whole-slide images for datasets are available at <https://portal.gdc.cancer.gov/>. Training images for tumor detection are available at <https://doi.org/10.5281/zenodo.2530789>. Training images for MSI detection are available at <https://doi.org/10.5281/zenodo.2530835> and <https://doi.org/10.5281/zenodo.2532612>. Source data for Fig. [1](https://www.nature.com/articles/s41591-019-0462-y" \l "Fig1) are available in public repositories at <https://doi.org/10.5281/zenodo.2530789>, <https://doi.org/10.5281/zenodo.2530835> and <https://doi.org/10.5281/zenodo.2532612>. Source Data for Figs. [1](https://www.nature.com/articles/s41591-019-0462-y" \l "Fig1), [2](https://www.nature.com/articles/s41591-019-0462-y" \l "Fig2) and Extended Data Figs. [1](https://www.nature.com/articles/s41591-019-0462-y" \l "Fig3), [2](https://www.nature.com/articles/s41591-019-0462-y" \l "Fig4) containing the raw data for these figures are available in the online version of the paper.