

1. Student 1

✓ P1

大家好，我们是Group 2，我们的组员是xxx，xxx还有我xxx。我们组的project主题是古埃及文字识别。

✓ P2

假如我们能穿越到7000多年以前，来到非洲东北部尼罗河中下游的土地上，我们便来到了古埃及文明。

✓ P3

古埃及文明有着在古文明中非常发达的农业技术、天文学和数学知识，还有先进的医学。说到古埃及文明，大家想到的肯定会是金字塔、木乃伊等等。但今天我们研究的对象是古埃及的象形文字。

✓ P4

距今5000多年前，古埃及出现了象形文字，即埃及文字，**埃及文字是世界上最古老的文字**，由法老王那默尔的铠甲关节板上的最早期象形刻记起（公元前3,100年），到现在用在教堂内的古埃及文字。后来被欧洲人称作Hiéroglyphe——这是希腊语“神圣”与“铭刻”组成的复合词，意思是“神的文字”，即“神碑体”。古埃及人认为他们的文字是月神、计算与学问之神图特（Thoth）造的。

通常古埃及文字被书写在一种称作“纸草”（papyrus）的纸张上，但这些草纸无法保存千年，所以目前我们能见到的古埃及文字大多出现在石板的雕刻上。

✓ P5

这是5个常见的古埃及文字，意思分别是坐着的男人、秃鹫、眼镜蛇、沙草、水波。古埃及文字有复杂的书写系统，书写顺序不同，句子的含义也不同。古埃及文字还有表音文字系统，有的文字其实并没有实际意义，只是表达语气。

2. Student 2

✓ P6

我们的任务目标是对古埃及象形文字做识别工作。通常来说，古埃及象形文字常见于陵墓或庙宇中的装饰雕刻或者是在古埃及文物上的绘制图案。我们希望实现给算法输入一个含有古埃及文字的场景图片，比如一块古埃及石板的照片，我们的算法能够识别出图片中的每一个象形文字分别是什么意思。对于这样一个任务，我们的pipeline分为两个部分，第一部分是对输入的图片做图像分割，把输入的图片中每一个象形文字都分割出来。算法的第二部分是把分割出来的每一个象形文字去交给我们的识别网络去做识别。

✓ P7

我们最初希望用神经网络去完成图像分割的部分，我们用Labelme工具，标注了一些古埃及石板的图片，作为我们的迷你数据集。我们为了简化任务难度，我们选了8种在古埃及石板上比较常见的图案，作为我们要分类的类别。我们也只标注了这些类别的图案。

✓ P8

我们在经过pre-trained的solo网络用我们的数据集做了迁移训练，由于我们的数据集太小，分割的效果依然不是很好。（可以结合图片即兴发挥，随便bb几句）由于我们算法的分类表现首先是依赖于算法对图像的分割工作做的如何，所以我们决定采用备用方案，我们写了一个交互的节目让用户去手动框选需要识别的象形文字。

3. Student 3

✓ P9

从石板图片中获取待处理的区域后，我们通过两个阶段的处理对文字进行匹配。

第一个阶段主要完成的是字符类别的分类。

第二个阶段则根据第一阶段的输出结果，遍历对应的字符集，选择最接近的字符作为最终的输出结果。

举个例子，我们从石板中分割出图上的区域，stage1对图片进行分析，预测这张图片属于‘Bird’，因此stage2中对鸟类的字符进行匹配，选择出了如上图的结果。

✓ P10

stage1中类别划分主要是通过Resnet50 实现的。我们在经过pre-trained的model上使用了自制的数据集进行了迁移学习。

✓ P11

我们在测试集上对模型进行了准确率检测。

可能由于数据集本身的原因，模型在有些类，例如snake，上面的准确率并不高。

但也可以观察到，大部分类上模型预测还是达到了不错的效果。（展示一下图片）

✓ P12

在stage1的基础上，我们进一步进行stage2中的更细节的字符匹配。

这一步我们最基本的想法是，在stage1预测出的类对应字符库中，选择与输入差距最小的字符作为最终的结果。

在这里我们采取了hog(方向梯度直方图)，来描述图片的特征。

✓ P13

那么如何定义两张图片的差距呢？

最直观的想法是直接计算两张图片hog的差。

但我们输入的图片噪音可能会比较大。

除了线条之外，图片中存在其他的噪音，可能会干扰图片的差距，从而导致匹配不稳定。

因此我们考虑使用canny边缘检测，只保留特征比较显著的区域，然后计算canny算子处理之后的图片的hog，补充描述图片的特征。

✓ P14

因此我们将两张图片的差距分为两个组成部分，diff1和diff2。

diff1是原图的hog的差距，diff2是经过canny算子处理之后的结果图的hog差距。

我们将这两部分赋以不同的权重，得到一个复合的特征描述。

✓ P15

由于输入图片的质量不均，同样的canny可能在不同的输入图片上效果参差不齐。

因此我们计算了3个不同尺度下的边缘检测图片，然后选取其中差距最小的一个尺度，作为diff2的值。

以上就是我们分类部分的主要算法。

4. Demo

5. Student 1

最后一页（在demo之后，你接着讲future work）

我们目前的算法，由于没有现成的数据集，我们制作的数据集太小，所以我们训练出来的模型在预测准确率上仍有很大的提升空间。如果我们可以提高我们算法在图像分割这一部分的准确率，我们便不再需要用户手动分割图像。同时我们目前由于不同类别的古埃及文字的数据数量差别很大，导致我们在有些类别的文字的分类准确度上表现得很糟糕，对于有较多数据的鸟类图案，我们的算法能很准确的识别和分类并匹配。所以下一步工作应该是扩大我们的分类数据集，同时提高我们数据集中的类别数量，这样我们就不再是只能对石板图片中部分类别的文字进行识别，而是能对整块石板上的文字都做识别。