**大作业视觉任务说明**

#### IMG_256**本次大作业，我们参考了CVPR 2018 best paper *[Taskonomy: Disentangling Task Transfer Learning](https://taskonomy.vision/)*的多任务学习模型，用一个整体大网络，实现了24种视觉任务。该模型考量了不同视觉任务之间的联系，在训练时采用迁移学习的方法，让视觉任务成为一个整体工程。本次作业由于我们只需要使用模型，故不多介绍其对迁移学习领域起到的推动作用。下面我们详细介绍一下这些视觉任务。**

**任务按照输入图片的数量分为两类：有两种任务，Pairwise-Fixated-Camera-Pose（成对固定相机姿势估计），Pairwise-Nonfixated-Camera-Pose（成对非固定相机姿势估计））需要两张输入图，且两张图为相关的同一场景的不同角度；类似的，Triplet-Fixated-Camera-Pose（三固定相机姿势估计）需要三张场景相关图片作为输入。为方便助教测试，我们提供了一组来自同一场景的test.png来测试。测试时不建议用完全不相关的图片来考察鲁棒性，因为这没什么意义。这些多图输入任务都是CV中场景理解和重建方向典型的多目视觉任务，对于场景级应用有至关重要的作用。其它21种任务均只需要单图作为输入，下面我们逐一简单介绍。**

**Autoencoder（自编码器）是图重构任务，它会先把图encode成一张低阶表示图，然后学习一个模型来重建，可以用于低维重构（这个任务的输出不仔细看和原图会有点像）； Curvatures（曲率计算）来计算曲率，看结果即可，可应用于一些物理视觉应用；Colorization（着色）可以像p图一样让图片更富有色彩感；Denoising-Autoencoder（降噪自编码器）可以为图片降噪，比如很多曝点的图便可以用这种方法修复；Depth（深度估计）可以估计图片大概的深度，生成的是深度热图；Edge Detection-2D/3D（2D/3D边缘检测）两者均为边缘计算所用，从两种维度来处理；Euclidean-Distance（欧氏距离深度估计）可以估计其深度的欧式距离；Inpainting（图像修复）用来修复图片；Jigsaw-Puzzle（分块）可以把图按物体分布大致分块；Keypoint-2D/3D（2D/3D关键点检测）用来从两个维度检测关键点；Object-Classification（分类）在ImageNet上完成对1000种类的图进行分类；Reshading（重阴影）可以通过阴影来分析场景，属于补充工作；Room-Layout（室内布局）可以用来衡量室内的物体分布结构；Scene-Classification（场景识别）用来识别场景；Segmentation-2D/3D/semantic（2D/3D/语义分割）从三种不同角度对图片进行分割；Surface-Normal（表面法向量）和深度估计相关性很大；Vanishing-Point（消失点）来考量图中的消失点以对图片进行更好的把握。**

**这些任务基本涵盖了当今有输入图条件下所能完成的所有工作，是对于视觉感官的整体工程，也是逼近人眼级的史诗级项目。选取这个模型，我们认为可以在对深度学习cv任务的探索中迈出重要一步，也让我们的大作业在众多检测、分割任务网站中更富有亮点。**