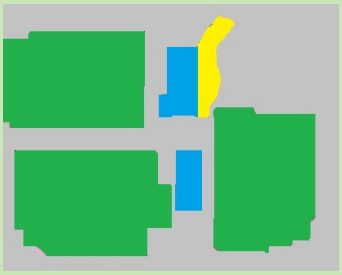
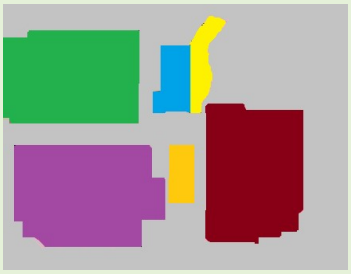
Fundamental Tasks in Computer Vision

Task	Models	Example
Image Classification: Assigning a single label to an input image based on its content using machine learning techniques. **Single Label**	AlexNet VGG (VGG16, VGG19) ResNet (ResNet50, ResNet101, ResNet152) Inception (InceptionV3, InceptionResNetV2) MobileNet EfficientNet	 Swimming Pool
Image Classification with Localization: Identifying the main object then assign a single label to the bounding box of the image. **Single Label + Bounding box**	YOLO (You Only Look Once) Faster R-CNN SSD (Single Shot MultiBox Detector) RetinaNet	
Object Detection: Locating and classifying multiple objects within an image by drawing bounding boxes around them. **Multiple (Labels + Bounding box)**	Faster R-CNN YOLO SSD RetinaNet EfficientDe	
Semantic Segmentation: Assigning a class label to each pixel in an image, allowing for fine-grained segmentation based on object semantics. **Class label to each pixel**	U-Net FCN (Fully Convolutional Network) DeepLab (DeepLabv3, DeepLabv3+)	
Instance Segmentation: Detecting and segmenting individual objects within an image at the pixel level, distinguishing between multiple instances of the same class. **Different labels for different instances of the same object** Notice how each house is assigned a different class, and so is each swimming pool.	Mask R-CNN SOLO (Segmenting Objects by Locations) BlendMask HTC (Hybrid Task Cascade)	

Pratik Dhungana
https://www.linkedin.com/in/pratikdh