# What is the Semantic Segmentation?

Semantic segmentation is a computer vision task that involves classifying and labeling each pixel in an image with a specific semantic category. In simpler terms, it is the process of dividing an image into meaningful segments and assigning a label to each segment based on its content. Unlike basic image classification, where the goal is to assign a single label to the entire image, semantic segmentation provides a more detailed understanding by segmenting the image at the pixel level.

The output of semantic segmentation is a pixel-wise classification map, where each pixel in the input image is assigned a label representing the category it belongs to. This enables computers to understand and differentiate between different objects, structures, or regions within an image. Semantic segmentation has various applications, including autonomous vehicles, medical image analysis, object detection, and augmented reality, where precise understanding of the visual scene is crucial.

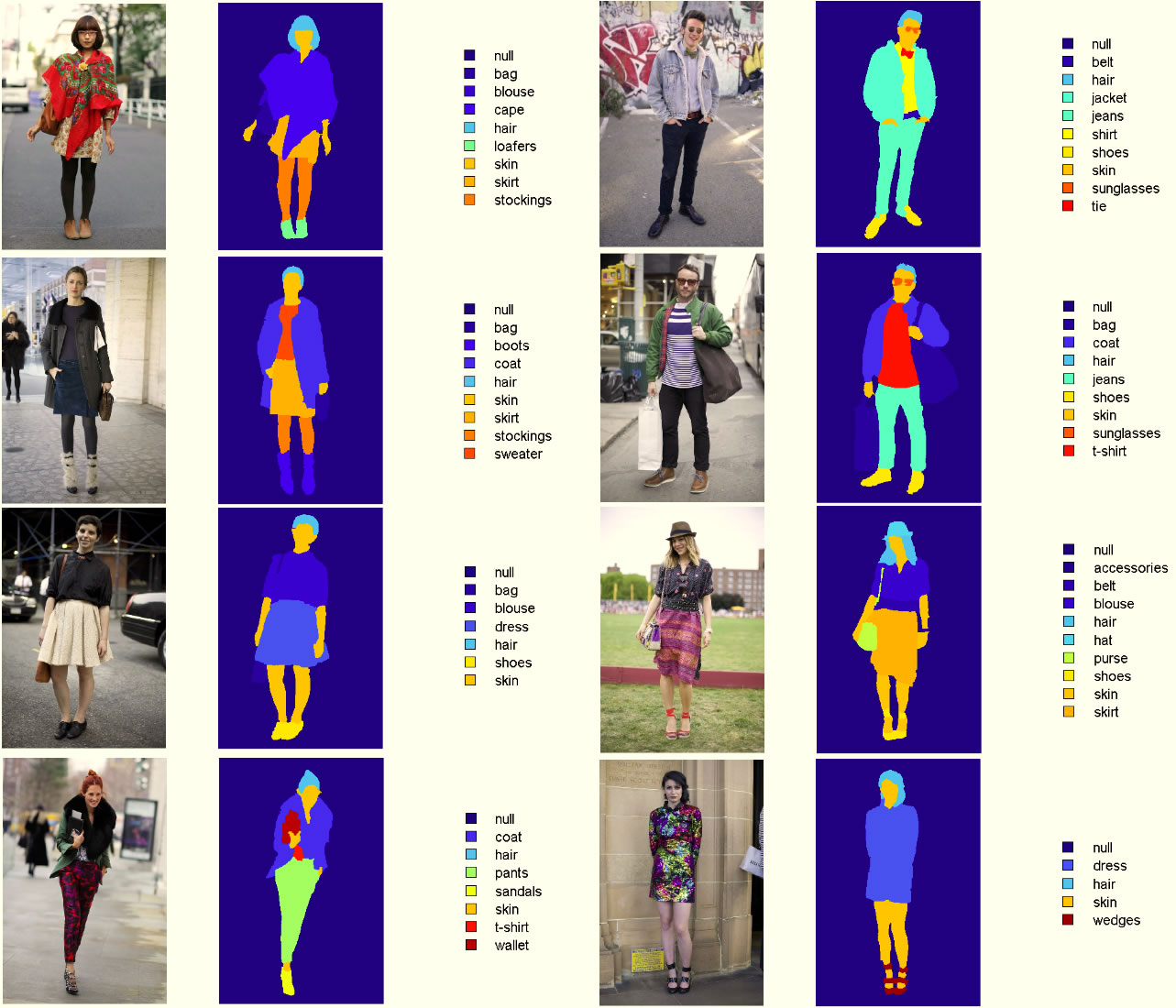
A close-up of a diagram

Description automatically generated

# In our Project

Semantic segmentation of fashion images, also known as clothing parsing, is to segment different clothing items out by classifying each pixel to a clothing category. It has been a hot topic in computer vision, due to the huge potential of related applications, such as clothing retrieval and recommendation. A lot of research work has been carried out in this field.

In this study, we deploy Deeplab V2, a state-of-the-art semantic segmentation model with deep learning, which has remarkable performance on some challenging datasets. For currently proposed semantic segmentation models of fashion images, they are trained on relatively small datasets, such as the Fashionista and CFPD datasets. And the quality of these datasets are not quite high with some noisy annotations and inconsistent contours. We train Deeplab V2 with ModaNet dataset, which is a brand new, high-quality and currently largest fashion dataset for semantic segmentation. Having adequate fashion images helps to improve the performance of the model.



# About Deeplab V2

DeepLabv2 is an architecture for semantic segmentation that build on DeepLab with an atrous spatial pyramid pooling scheme. Here we have parallel dilated convolutions with different rates applied in the input feature map, which are then fused together. As objects of the same class can have different sizes in the image, ASPP helps to account for different object sizes.

