

Core VLM Components

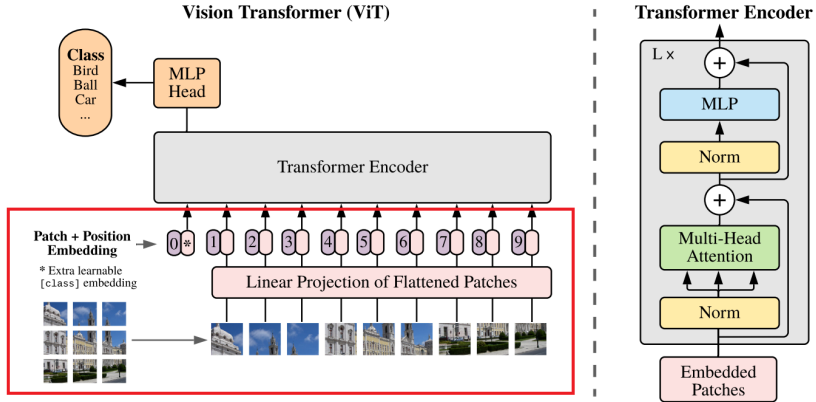


Figure 14: Vision Transformer (ViT) Architecture