

2 different

Feature	Swin Transformer	ViT
Attention	within each local window	Globally
complexity	linear computation ~ to input image size	quadratic computation ~
Input	hierarchical feature	fixed-size patches