Bottom-Up Processing

Input Frame

Goal: **Key**

Gray ceiling

Chandelier

4 big windows

TV

Green plant in a white vase

Black lamp

Key

3 brown sofas

Coffee table with a laptop on the surface

Brown box

Wooden floor

Walkable path

Textured carpet

Top-Down Selective Attention

Input Frame

Goal: **Key**

Encoder with Codebook bottleneck

Flat surface might contain the key

Key

Flat surface might contain the key

Walkable path to the flat surfaces