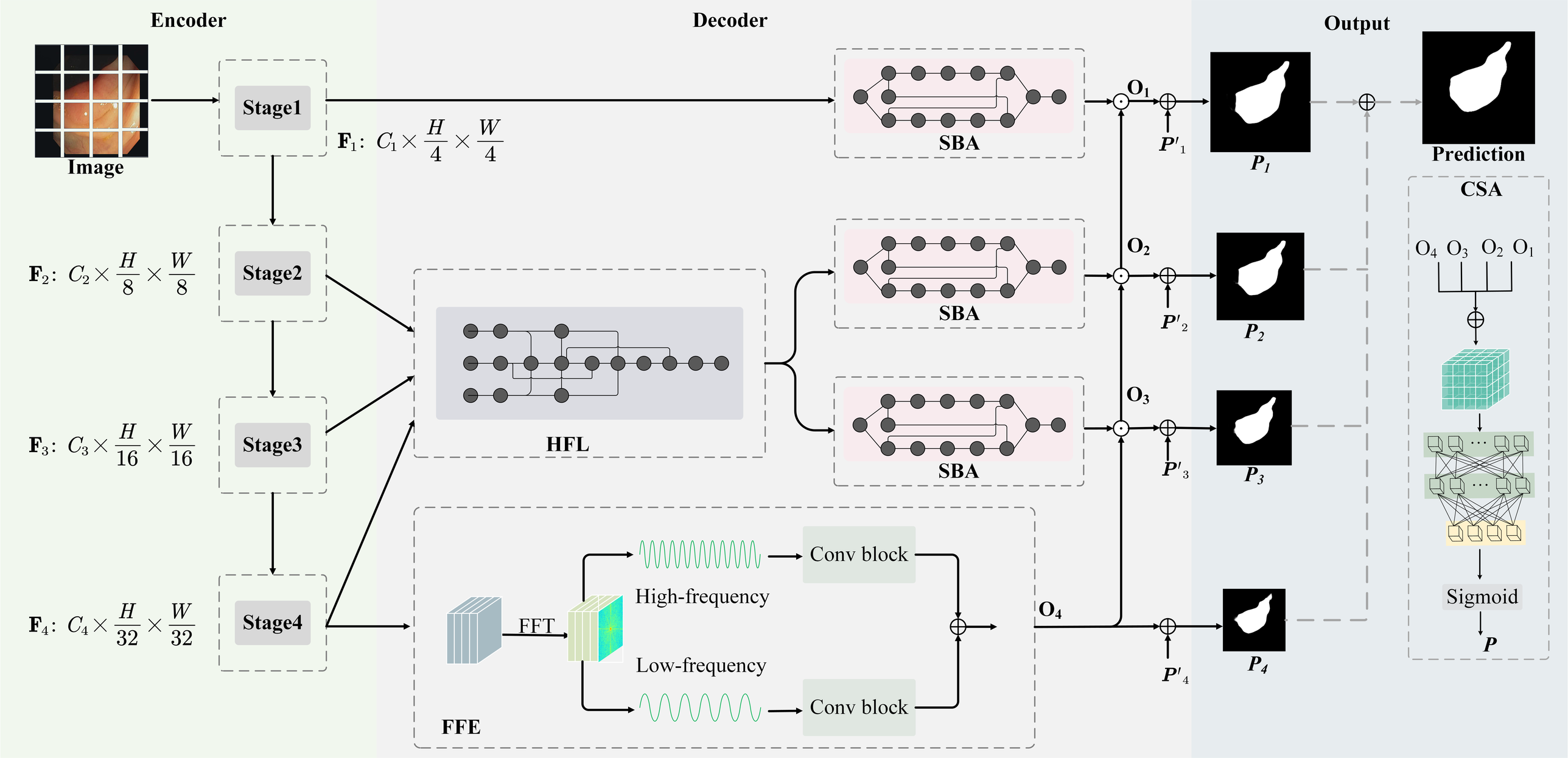
Encoder
Decoder
Output
Image
Stage1
Stage2
Stage3
Stage4
$\mathbf{F}_1$: $C_1 \times \frac{H}{4} \times \frac{W}{4}$
$\mathbf{F}_2$: $C_2 \times \frac{H}{8} \times \frac{W}{8}$
$\mathbf{F}_3$: $C_3 \times \frac{H}{16} \times \frac{W}{16}$
$\mathbf{F}_4$: $C_4 \times \frac{H}{32} \times \frac{W}{32}$
HFL
SBA
SBA
SBA
FFE
FFT
High-frequency
Low-frequency
Conv block
Conv block
O_1
O_2
O_3
O_4
P'_1
P'_2
P'_3
P'_4
P_1
P_2
P_3
P_4
Prediction
CSA
O_4 O_3 O_2 O_1
Sigmoid
P