Similarity Scoring

CLS

RCS Module
(Response Correlation
Spreading)

SRR Module
(Spatial Response
Refinement)

Anomaly Heatmap

Vision
Transformer
(VIT) Encoder

Normal Text Prompt
"a standard,working cirtuit board"

Anomaly Text Prompt
"a faulty,admaged circuit board"

Text Encoder

Mid-layer Self-Attention Correlation

Frozen $\oplus$ Element-wise sum

Learnable

RCS

Top anomaly responses

Patch-level anomaly
response map P

patch

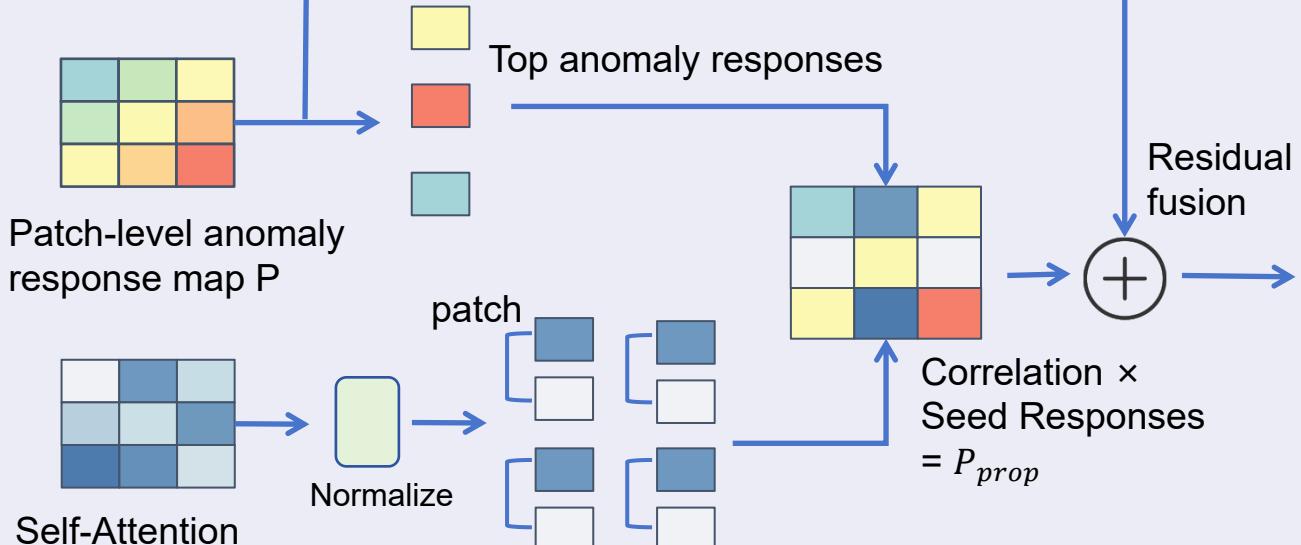Residual
fusion

Correlation ×
Seed Responses
= P_{prop}

Self-Attention

Normalize

SRR

Intensity-Gated Adaptiver Spatial Smoothing

Low Response
(Unchanged)

Medium
(Smoothing)

High (Stronger
Smoothing)

● Local neighborhood lightweight aggreation

● Noise Suppression
Improved local spatial consistency

● Reduce scattered responses near unstable
boundaries