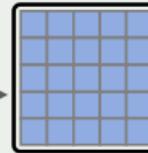


**Structured
Attention Masks**

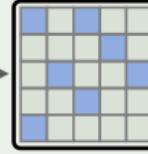
**Speaker
Embedding**

Concatenate

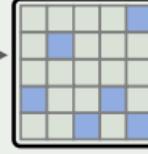
Global



Speaker

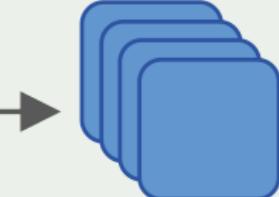


Reply



**Emotion
Classification
(FFNN)**

\mathbf{p}_i^e



Enhanced Feature

**Cause
Classification
(FFNN)**

\mathbf{p}_i^c

ECSR-MoE

FG

Gated Fusion

→

Residual Connection

FG

→

FG

→

FG

→

FG

→

FG

→

FG