

Vertical Stream

Horizontal Stream

Input:  
Frame

$C_{(1,1)}$

$C_{(1,2)}$

...

$C_{(1,n)}$

$F$

$L_1$

$C_{(2,1)}$

$C_{(2,2)}$

...

$C_{(2,n)}$

$F$

$L_2$

⋮

⋮

...

⋮

⋮

⋮

$C_{(m,1)}$

$C_{(m,2)}$

...

$C_{(m,n)}$

$F$

$L_m$

Concatenation

$V$

$\pi$

$C_{(1,1)}$

$F$

$L_1$

Concatenation

Output

Conv Layer

Flatten Operation

LSTM Layer

Concatenation

Output