

Layer No.	Layer Type	Kernel Number	Kernel Size	Stride	Activation	Padding	Output Size
1	Conv2D	64	3×3	1×1	ReLU	Yes	$M \times M$
2	Conv2D	64	3×3	1×1	ReLU	Yes	$M \times M$
3	MaxPooling2D	-	2×2	2×2	-	No	$M/2 \times M/2$
4	Conv2D	128	3×3	1×1	ReLU	Yes	$M/2 \times M/2$
5	Conv2D	128	3×3	1×1	ReLU	Yes	$M/2 \times M/2$
6	MaxPooling2D	-	2×2	2×2	-	No	$M/4 \times M/4$
7	Flatten	-	-	-	-	-	$(M/4 * M/4 * 128)$
8	Dense	1	$S \times (M/4 * M/4 * 128)$		-	-	S

S , grid number; M , microphone number.