

Model

noise

GSM

Encoder

input

normalize

mirror-pad

 $64 \times 5 \times 5 / 2$

conv

 $128 \times 5 \times 5 / 2$

conv

 $128 \times 3 \times 3$

conv

 $128 \times 3 \times 3$

conv

sum

 $128 \times 3 \times 3$

conv

 $128 \times 3 \times 3$

conv

sum

 $96 \times 5 \times 5 / 2$

conv

round

code

output

clip

denormalize

 $12 \times 3 \times 3 / 2$

subpix

 $256 \times 3 \times 3 / 2$

subpix

sum

 $128 \times 3 \times 3$

conv

 $128 \times 3 \times 3$

conv

sum

 $128 \times 3 \times 3$

conv

 $128 \times 3 \times 3$

conv

 $512 \times 3 \times 3 / 2$

subpix

Decoder