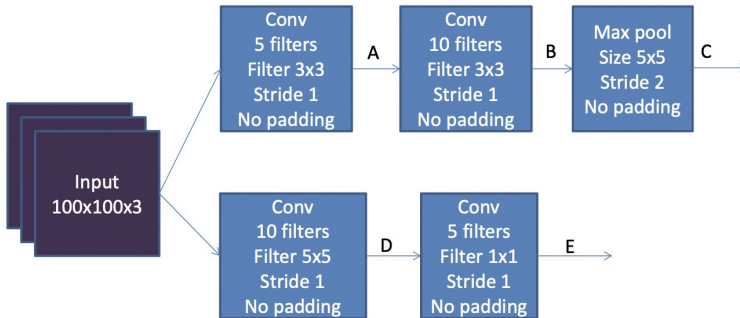


**T1.** Given the following CNN architecture specifications, determine the size of all the intermediate feature maps.



A :  $5 \times 98 \times 98$

B :  $10 \times 96 \times 96$

C :  $46 \times 46$

D :  $10 \times 96 \times 96$

E :  $5 \times 96 \times 96$

T2. Consider a CNN with the following layers:

1. conv2d layer with  $5 \times 5$ , stride 1
2. maxpooling of size  $3 \times 3$  (stride is usually the same as the size for pooling layers) 3
3. conv2d layer with  $7 \times 7$ , stride 2
4. max pooling of size  $2 \times 2$  **stride 2**
5. conv2d layer with  $3 \times 3$ , stride 1

What is the receptive field of this network?

**stride**

1.  $5 \times 5$

1

↓

2.  $7 \times 7$

3

↓

3.  $25 \times 25$

6

↓

4.  $31 \times 31$

12

↓

5.  $55 \times 55$

**T3.** Can we use CNN on a time-related input (stride over time steps)? Discuss the different inductive biases imposed by the model when you use a single CNN layer on a time-related input vs. when you use a single GRU layer on a time-related input. Come up with an example when one might be better than the other.

ได้, คิดว่า GRU จะมีประสิทธิภาพกว่า เช่น

อยากทำนายเวลาที่คนจะกลับบ้าน ทราบว่ามี input เป็น  
สถานที่ ที่คนๆ นั้น อยู่ ใน 24 ชม.

CNN จะทำนาย เวลา ใกล้เคียง เกิดเป็น routine ของ  
คนๆ นั้นก็จริง แต่ RNN ที่ รู้ลำดับที่ก่อนหน้าได้  
ดีกว่านะ เก่งกว่า