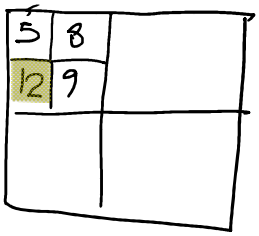
CNN

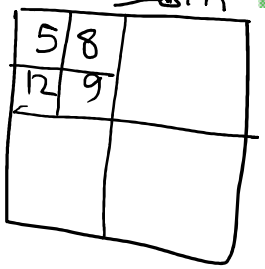
Sunday, February 18, 2024 5:19 PM

Pooling: (A.K.A Downsampling)

Max: Max(R) = 12

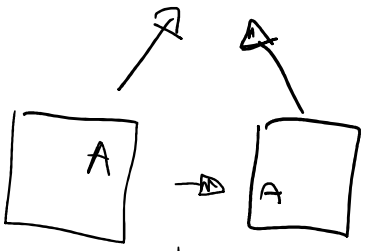


Avg: $\frac{\sum_{i}^{n} R_i}{n} = 8.5$



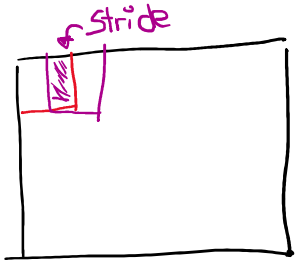
Advantage:

- Translational Invariance

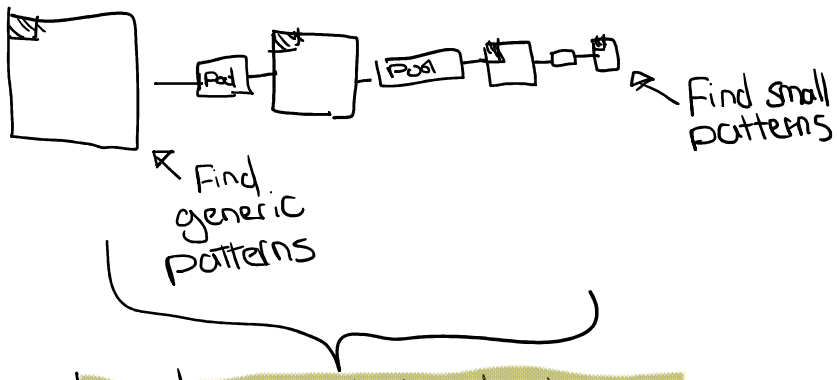


Location doesn't matter

- Pools can have a non-squared window
- Boxes can overlap



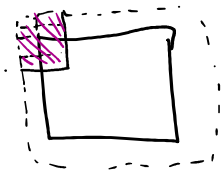
Conv. followed by pooling:



We lose data but we gain features

Note:

- We can use conv. stride instead of pooling



- An image often has large patches of stuffs: Eg.

- If we downsample image, we lose neighbors that could be high correlated but that info doesn't gives us useful information: At all. we do this to make calculations fast.

High correlation between neighbors

- CNN's can handle images of different sizes:

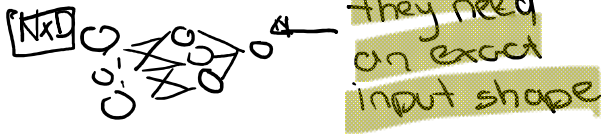
Eg. CNN 32x32, 2 stride

32x32 → 16x16 → 8x8 → 4x4 → 2x2

Input: 64x64

64x64 → ... → 4x9

- ANN's can't do that:



- We use **Global Max Pooling:**