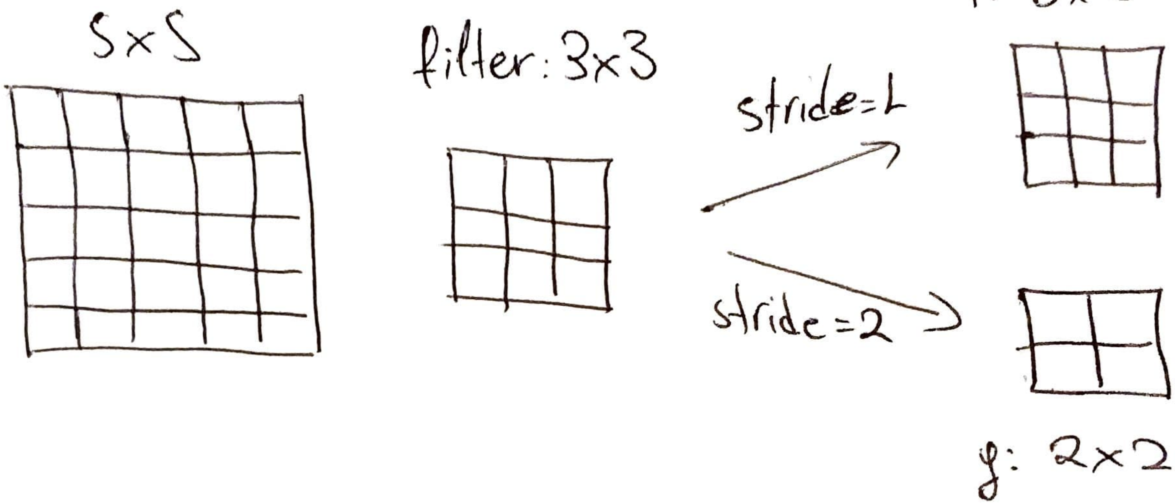
Problem-12 | (1)

Πώς μπορούμε να υπολογίσουμε το gradient ενός conv. layer όταν έχουμε stride>1

Λύση: Αρχικά, στην πράξη το αποτέλεσμα που θα προκύψει θα είναι περίπου το ίδιο με τη διαφορά ότι το μέγεθος του τελικού πίνακα θα είναι μικρότερο.

π.χ.



Problem-13

Υπολογισε το gradient ενος max-pooling με overlapping

Λύση: Γενικά: $\delta^{l}_{i,j} = \frac{\partial E}{\partial x^{l}_{i,j}} = m'(x^{l}_{i,j}) \sum_{i,j} \delta^{l+1}_{i,j} w^{l,l+1}_{i,j}$

↑ max pooling function

Έστω ότι έχουμε input $a=[i_0, i_1, i_2, i_3, i_4, i_5]$ κ' 1 max pooling function $m=[m_0, m_1]$ όπου $m_0=\max(a_0, i_1, i_2)$ κ' $m_1=\max(i_2, i_3, i_4)$

Απο την επίδραση της m στο i ~~...~~ προκύπτει

$$\boxed{m=[i_1^*, i_2^*]}$$

όπου i_1^* είναι το max των i_0, i_1, i_2 κ' το i_2^* το max των i_2, i_3, i_4

π.χ. για $a=[1,2,3,4,5] \Rightarrow m=[3,5]$

Η gradient της max pooling operation θα είχε μόνο 2 στοιχεία

$$\nabla_m F=[x,y]$$

Αν όμως επιτρέψουμε overlapping (σύμφωνα με το παράδειγμα) τότε θα προέκυπτε:

$$m=[0,0,i_1^*,0,i_2^*]$$

Άρα: η τελική gradient θα ήταν της μορφής: $\boxed{\nabla_m F=[0,0,x,0,y]}$