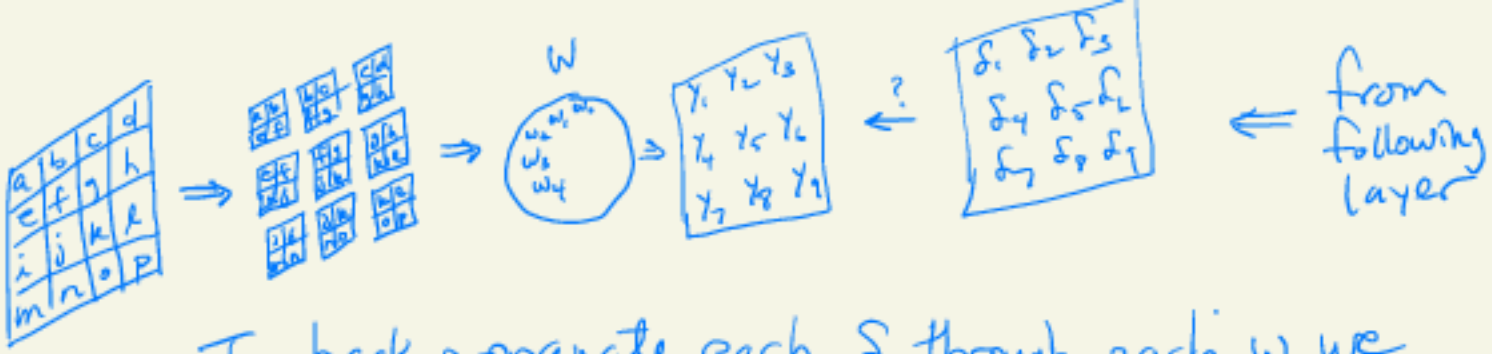
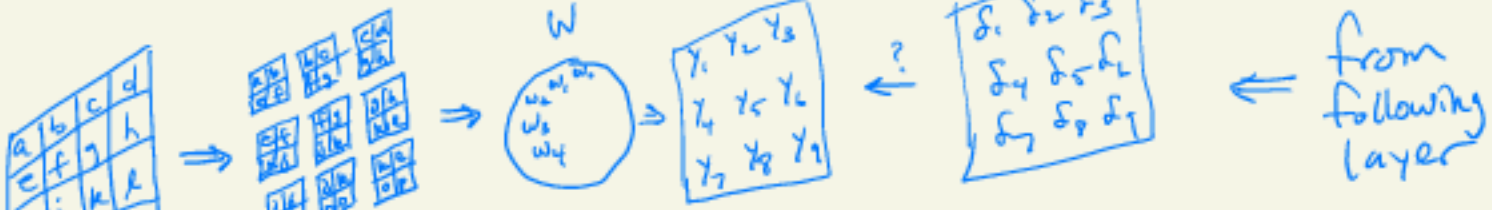
# Back propagation through a Convolutional Layer




To back-propagate each $\delta$ through each $w$, we need to know which $\delta$s are dependent on each input $a, b, \ldots, p$.



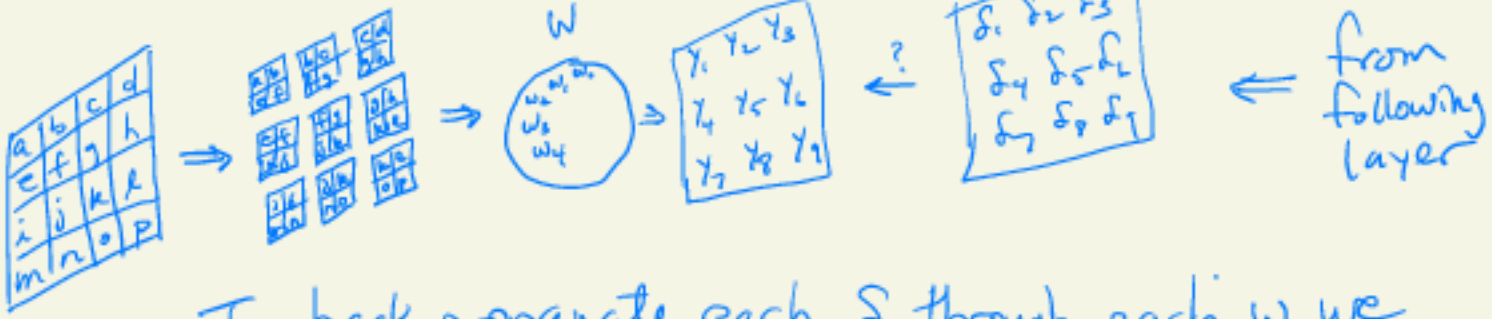

To back-propagate each $\delta$ through each $w$, we need to know which $\delta$s are dependent on each input $a, b, \ldots p$.

$$\begin{bmatrix} a & b \\ e & f \end{bmatrix} * \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} = y_1 \qquad \begin{bmatrix} b & c \\ f & g \end{bmatrix} * \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} = y_2 \qquad \begin{bmatrix} c & d \\ g & h \end{bmatrix} * \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} = y_3$$

$$\begin{bmatrix} e & f \\ i & j \end{bmatrix} * \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} = y_4 \qquad \begin{bmatrix} f & g \\ j & k \end{bmatrix} * \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} = y_5 \qquad \begin{bmatrix} g & h \\ k & l \end{bmatrix} * \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} = y_6$$

$$\begin{bmatrix} i & j \\ m & n \end{bmatrix} * \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} = y_7 \qquad \begin{bmatrix} j & k \\ n & o \end{bmatrix} * \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} = y_8 \qquad \begin{bmatrix} k & l \\ o & p \end{bmatrix} * \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} = y_9$$

Start with input a. Where does it appear?




To back-propagate each $\delta$ through each $w_i$, we need to know which $\delta$s are dependent on each input $a, b, \ldots p$.

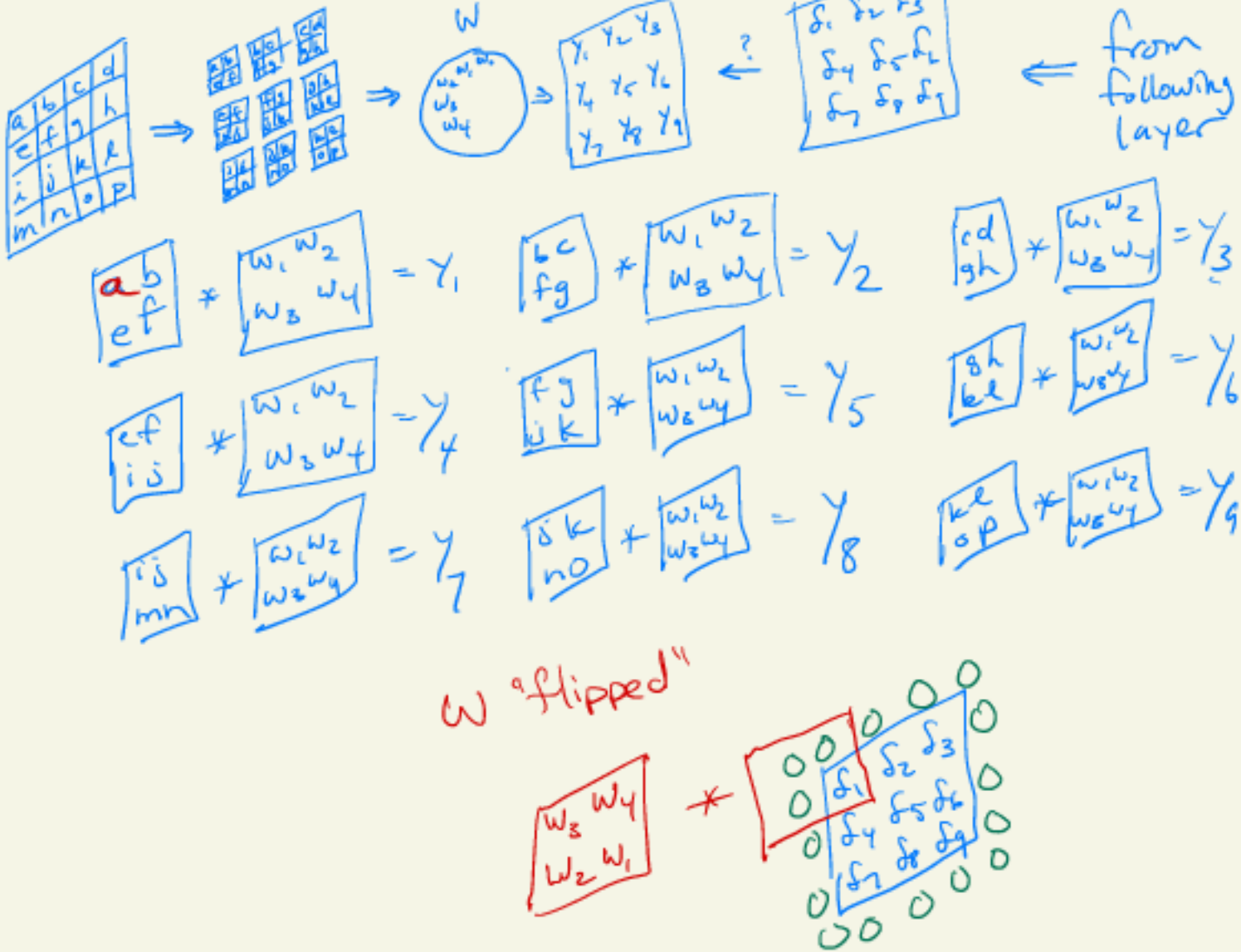
$$\begin{bmatrix} a & b \\ e & f \end{bmatrix} * \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} = y_1 \qquad \begin{bmatrix} b & c \\ f & g \end{bmatrix} * \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} = y_2 \qquad \begin{bmatrix} c & d \\ g & h \end{bmatrix} * \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} = y_3$$

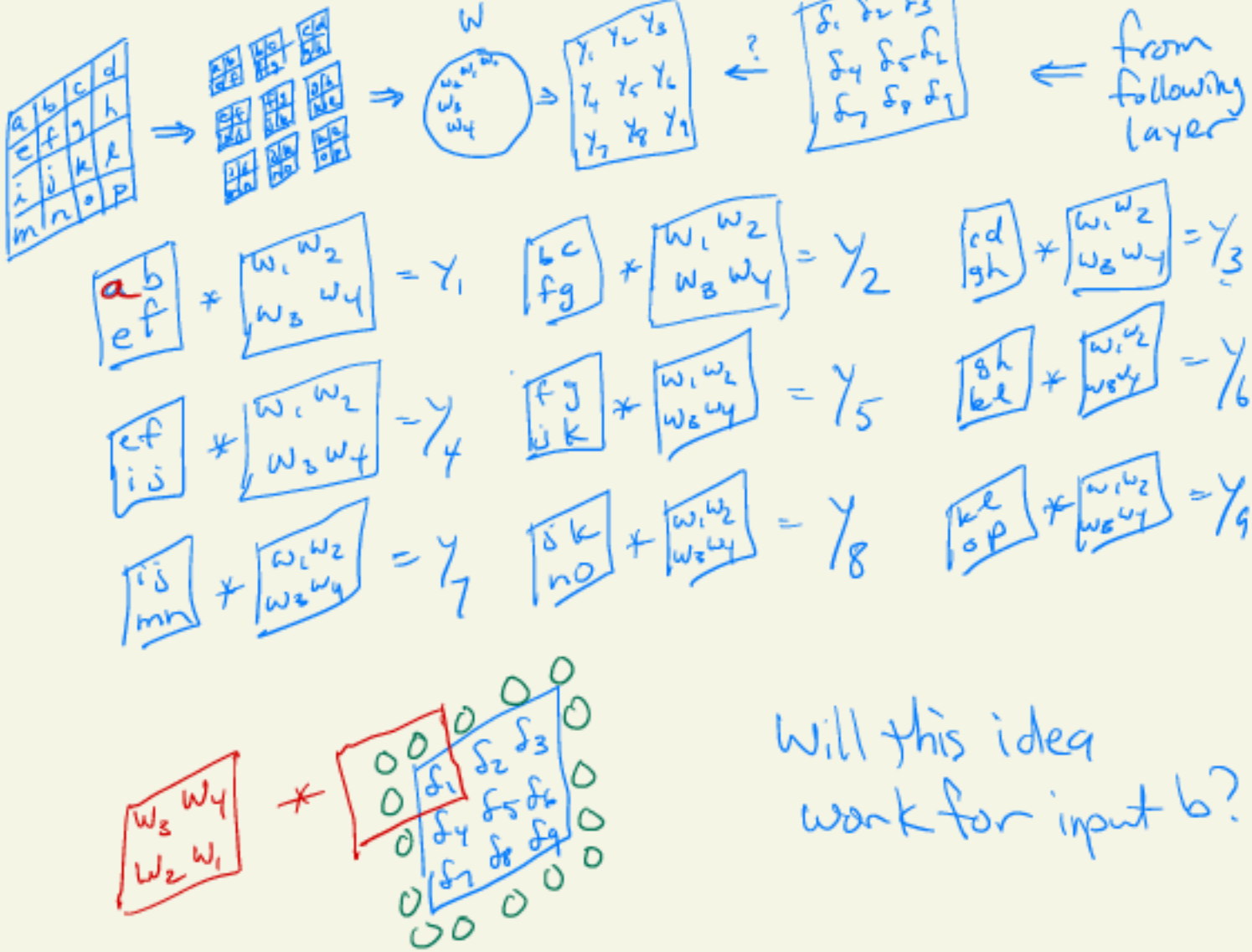
$$\begin{bmatrix} e & f \\ i & j \end{bmatrix} * \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} = y_4 \qquad \begin{bmatrix} f & g \\ j & k \end{bmatrix} * \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} = y_5 \qquad \begin{bmatrix} g & h \\ k & l \end{bmatrix} * \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} = y_6$$

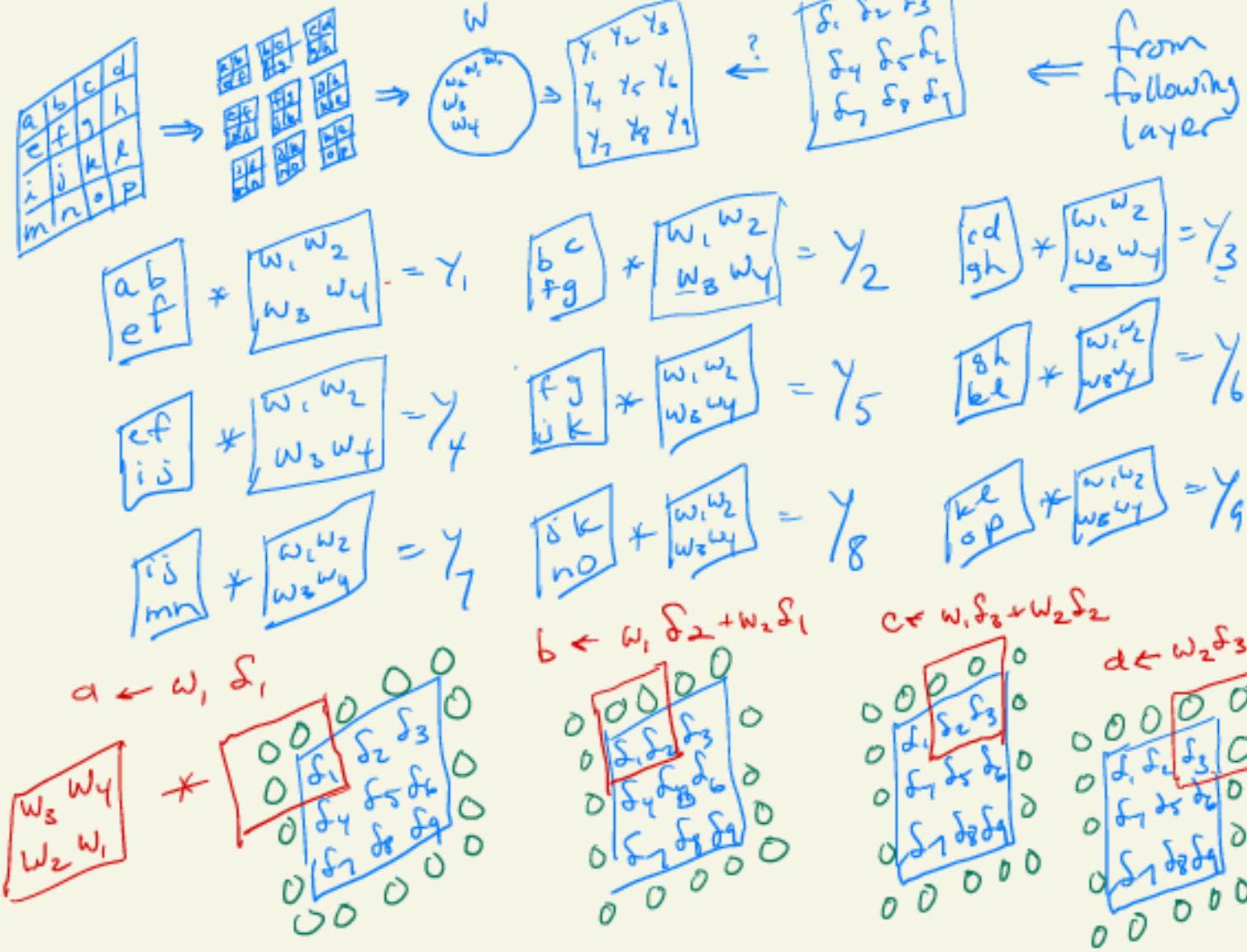
$$\begin{bmatrix} i & j \\ m & n \end{bmatrix} * \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} = y_7 \qquad \begin{bmatrix} j & k \\ n & o \end{bmatrix} * \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} = y_8 \qquad \begin{bmatrix} k & l \\ o & p \end{bmatrix} * \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} = y_9$$

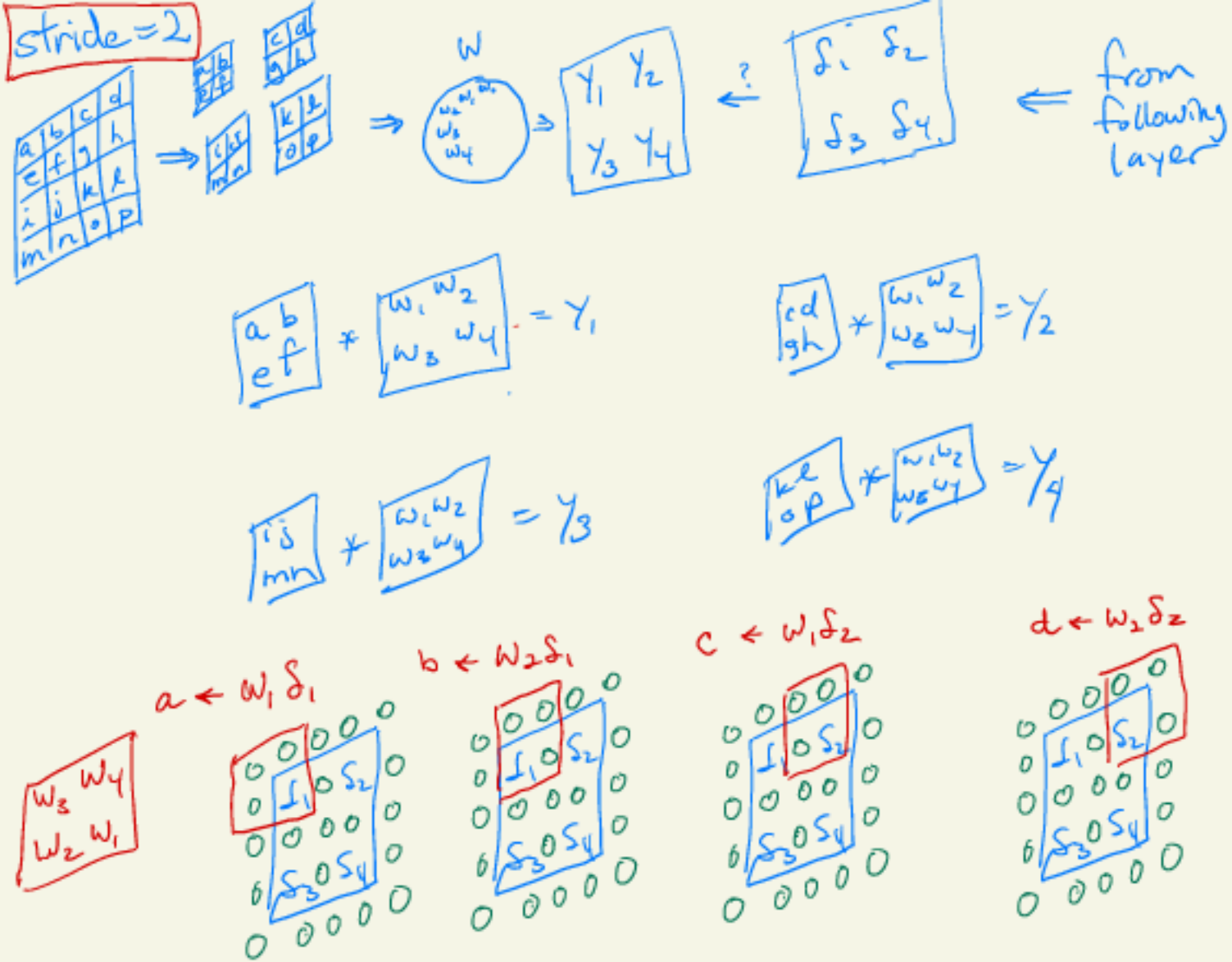
Start with input a. Where does it appear?

Can we construct a matrix operation that will multiply only a and $w_1$?


W
?
from following layer
W "flipped"


W
?
from following layer
Will this idea work for input b?


W
?
from following layer
a ← w1 δ1
b ← w1 δ2 + w2 δ1
c ← w1 δ3 + w2 δ2
d ← w2 δ3


stride=2
a b c d
e f g h
i j k l
m n o p
W
$w_1 w_2$
$w_3 w_4$
$Y_1$ $Y_2$
$Y_3$ $Y_4$
?
$\delta_1$ $\delta_2$
$\delta_3$ $\delta_4$
from following layer
$\begin{matrix} a & b \\ e & f \end{matrix} * \begin{matrix} w_1 & w_2 \\ w_3 & w_4 \end{matrix} = Y_1$
$\begin{matrix} c & d \\ g & h \end{matrix} * \begin{matrix} w_1 & w_2 \\ w_3 & w_4 \end{matrix} = Y_2$
$\begin{matrix} i & j \\ m & n \end{matrix} * \begin{matrix} w_1 & w_2 \\ w_3 & w_4 \end{matrix} = Y_3$
$\begin{matrix} k & l \\ o & p \end{matrix} * \begin{matrix} w_1 & w_2 \\ w_3 & w_4 \end{matrix} = Y_4$
$a \leftarrow w_1 \delta_1$
$b \leftarrow w_2 \delta_1$
$c \leftarrow w_1 \delta_2$
$d \leftarrow w_2 \delta_2$
$\begin{matrix} w_3 & w_4 \\ w_2 & w_1 \end{matrix}$