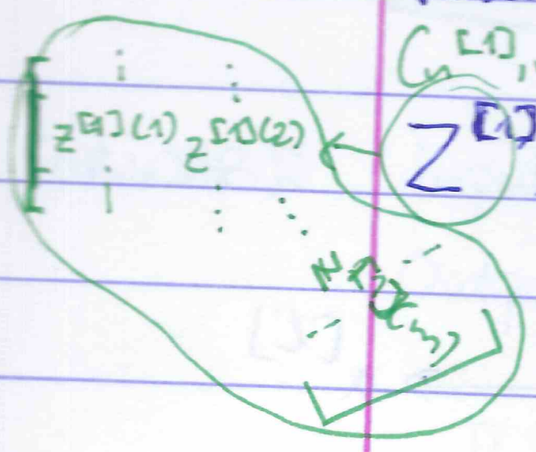
Vectorised:



$$Z^{[1]} = W^{[1]} X + b^{[1]}$$

$(n^{[1]}, m) \quad (n^{[1]}, n^{[0]}) \quad (n^{[0]}, m) \quad (n^{[1]}, 1)$

↓ when you add this to b, Python broadc.

$(n^{[1]}, m)$

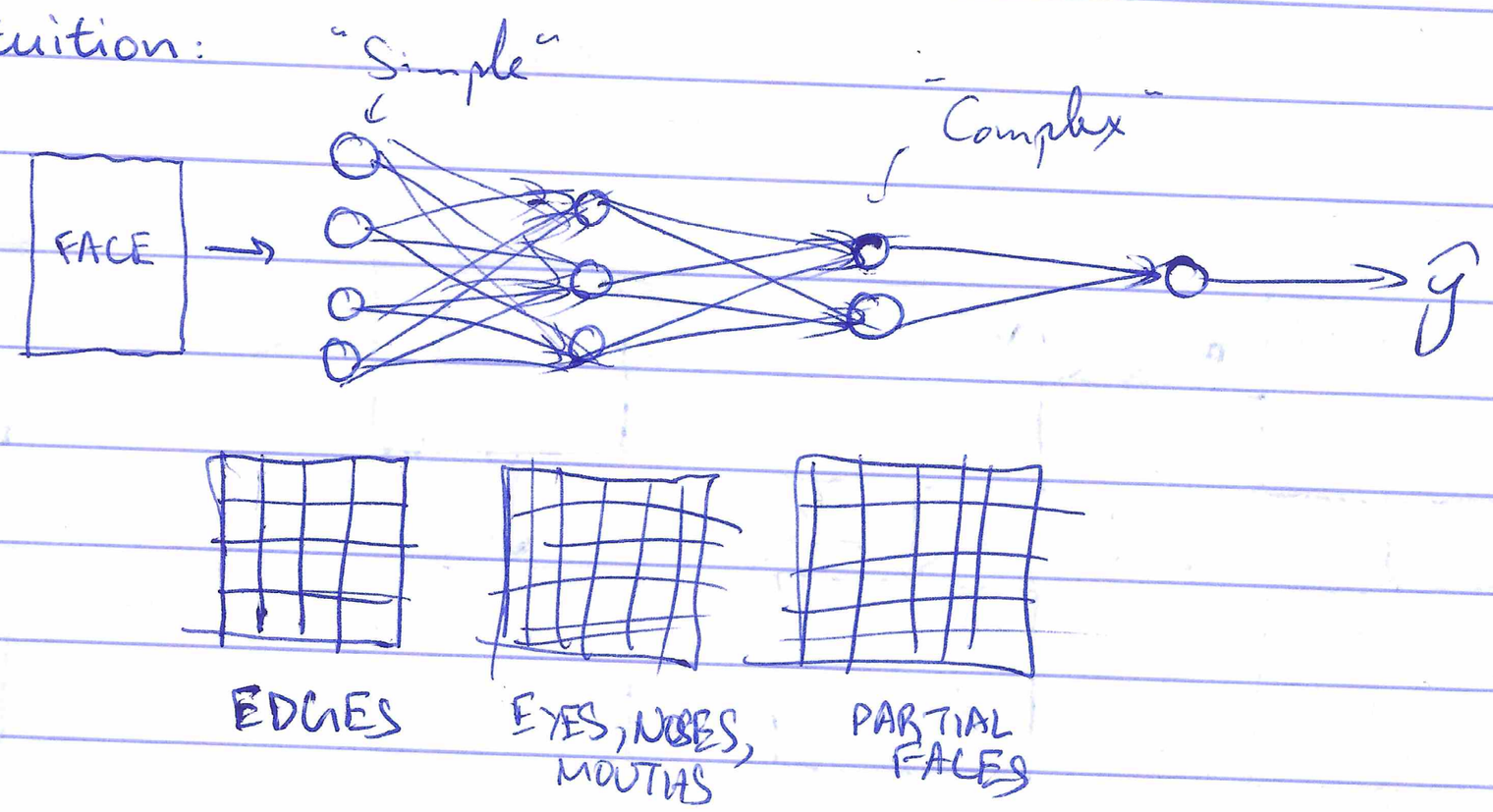
$Z^{[l]}, A^{[l]} : (n^{[l]}, m)$

$l = 0, \; A^{[0]} = X \bullet : (n^{[0]}, m)$

$dZ^{[l]}, dA^{[l]} : (n^{[l]}, m)$

Why Deep Representations?

- Intuition:



- Circuit theory & DL:
  - There are functions that you can compute w/ a "small" L-layer deep NN that shallower networks require exponent more hidden units to compute