LeNet - 5

goal : recognize handwritten digits
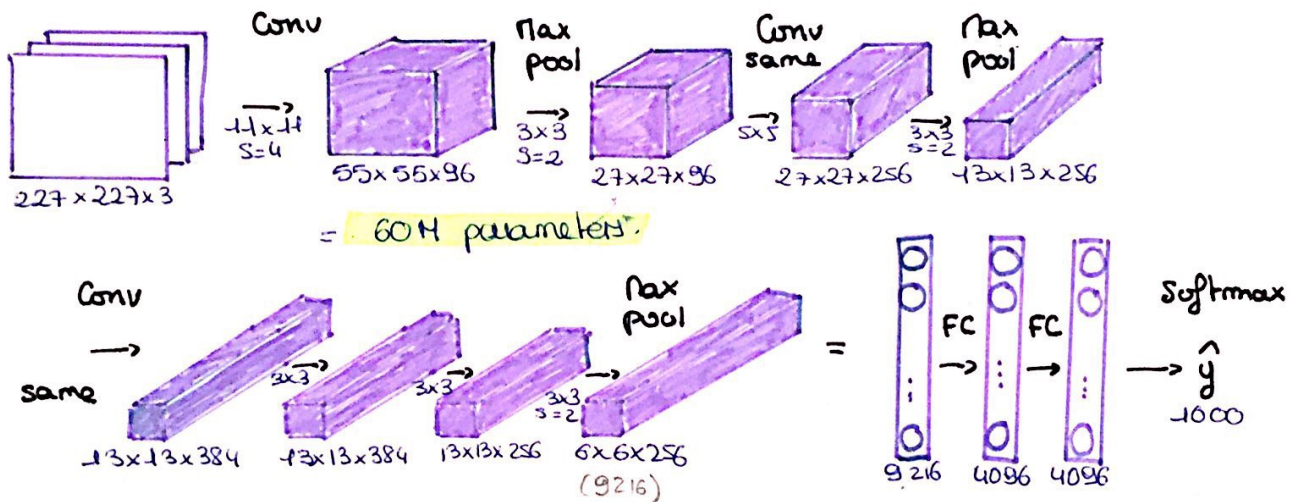
Activation = sigmoid / tanh.

$32 \times 32 \times 1$ — Conv (valid) 5×5, $S=1$ → $28 \times 28 \times 6$ — Avg pool $f=2$, $S=2$ → $14 \times 14 \times 6$ — Conv 5×5, $S=1$ → $10 \times 10 \times 16$ — Avg pool $f=2$, $S=2$ → $5 \times 5 \times 16$ (400) → FC 120 → FC 84 → Softmax $\hat{y}$ →0

= 60K parameters

n_H et n_W ↘ n_C ↗

---
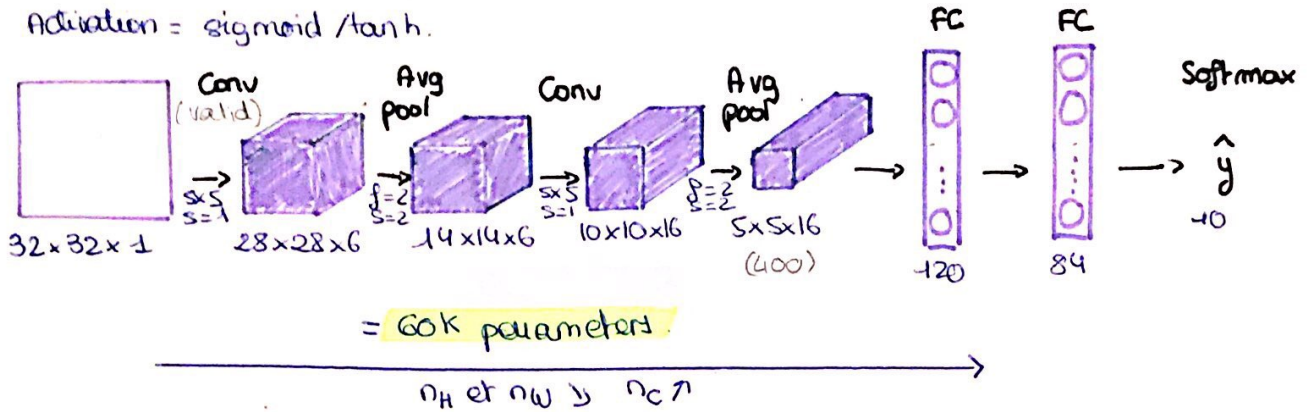
activation : Relu

AlexNet

Multiple GPUs
Local Response Normalisation (LRN).

$227 \times 227 \times 3$ — Conv 11×11, $S=4$ → $55 \times 55 \times 96$ — Max pool 3×3, $S=2$ → $27 \times 27 \times 96$ — Conv same 5×5 → $27 \times 27 \times 256$ — Max pool 3×3, $S=2$ → $13 \times 13 \times 256$

= 60 M parameters

Conv same → $13 \times 13 \times 384$ — 3×3 → $13 \times 13 \times 384$ — 3×3 → $13 \times 13 \times 256$ — Max pool 3×3, $S=2$ → $6 \times 6 \times 256$ (9216) = 9216 → FC 4096 → FC 4096 → Softmax $\hat{y}$ →1000

---

uniform architecture.

VGG - 16

Conv = 3×3 filter, $S=1$ same   MaxPool = 2×2, $S=2$

$224 \times 224 \times 3$ — Conv 64 ×2 → $224 \times 224 \times 64$ — Pool → $112 \times 112 \times 64$ — Conv 128 ×2 → $112 \times 112 \times 128$ — Pool → $56 \times 56 \times 128$

— Conv 256 ×3 → $56 \times 56 \times 256$ — Pool → $28 \times 28 \times 256$ — Conv 512 ×3 → $28 \times 28 \times 512$ — Pool → $14 \times 14 \times 512$

— Conv 512 ×3 → $14 \times 14 \times 512$ — Pool → $7 \times 7 \times 512$ — FC → 4096 — FC → 4096 → $\hat{y}$ Softmax

= 138 n parameters

Scanned with CamScanner