IMAGENET CLASSIFICATION WITH DEEP CONVOLUTIONAL NEURAL NETWORKS

ALEX KRIZHEVSKY
ILYA SUTSKEVER
GEOFFREY E. HINTON

ICT3115 - SISTEMAS URBANOS INTELIGENTES

PAULA CORONADO
GASPAR MASIHY

Presentación de hoy

ESTRUCTURA

Introducción
Arquitectura de la red
Reducción sobreajuste
Aprendizaje
Resultados
Discusión

Introducción

DATOS

Base de Datos: subset de ImageNet

1.200.000 de imágenes de entrenamiento

50.000 imágenes de validación

150.000 imágenes de testeo

ImageNet Classification with Deep Convolutional Neural Networks

Alex Krizhevsky
University of Toronto
kriz@cs.utoronto.ca

Ilya Sutskever
University of Toronto
ilya@cs.utoronto.ca

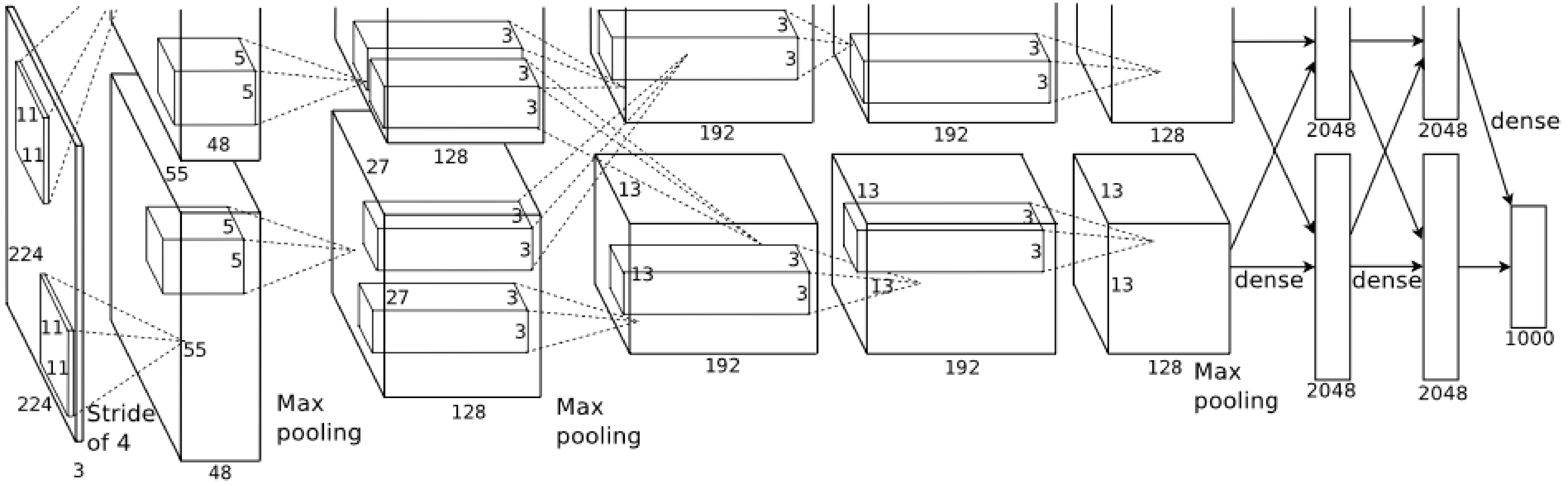
Geoffrey E. Hinton
University of Toronto
hinton@cs.utoronto.ca

Abstract

We trained a large, deep convolutional neural network to classify the 1.2 million high-resolution images in the ImageNet LSVRC-2010 contest into the 1000 different classes. On the test data, we achieved top-1 and top-5 error rates of 37.5% and 17.0% which is considerably better than the previous state-of-the-art. The neural network, which has 60 million parameters and 650,000 neurons, consists of five convolutional layers, some of which are followed by max-pooling layers, and three fully-connected layers with a final 1000-way softmax. To make training faster, we used non-saturating neurons and a very efficient GPU implementation of the convolution operation. To reduce overfitting in the fully-connected layers we employed a recently-developed regularization method called “dropout” that proved to be very effective. We also entered a variant of this model in the ILSVRC-2012 competition and achieved a winning top-5 test error rate of 15.3%, compared to 26.2% achieved by the second-best entry.

1 Introduction

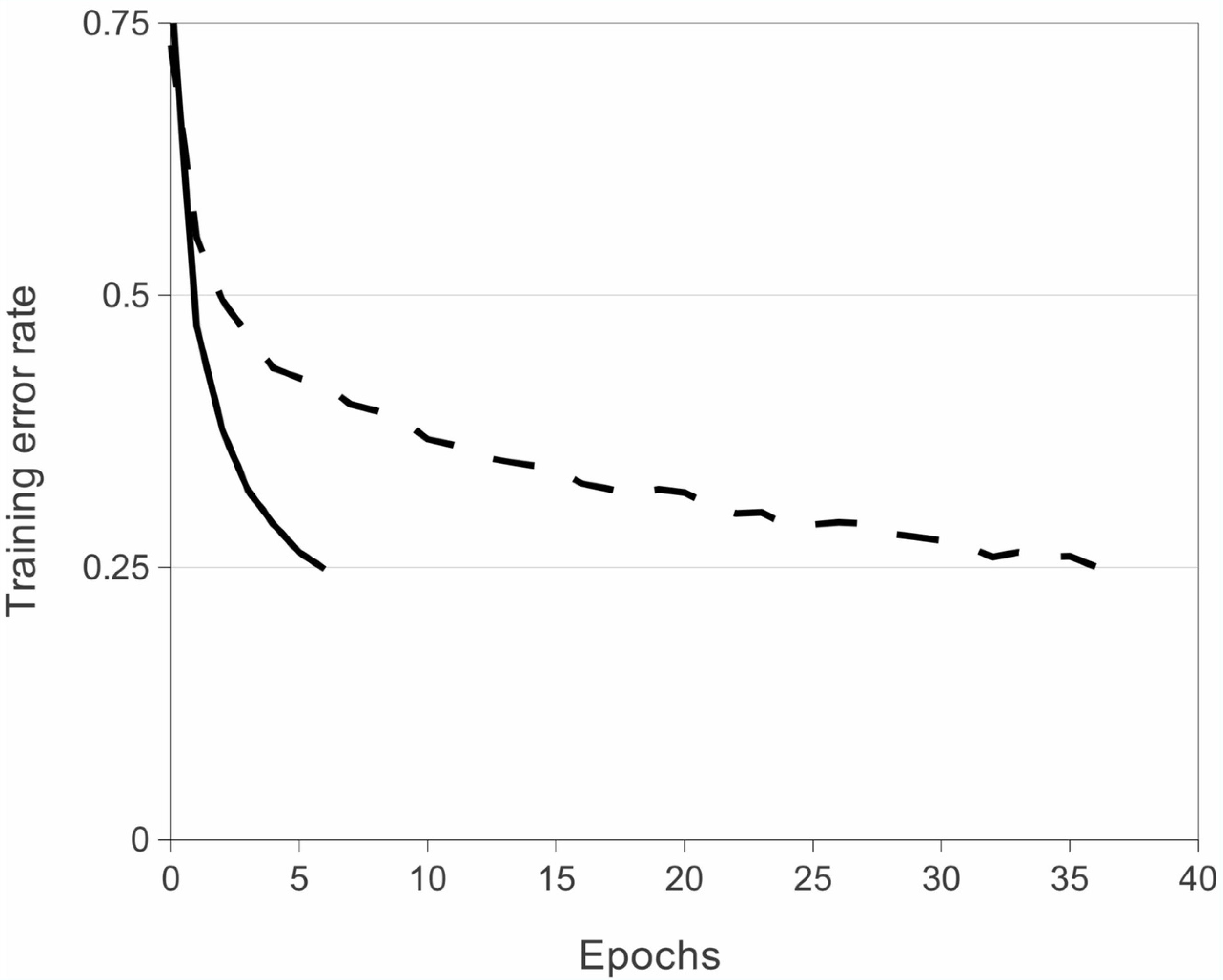
Arquitectura de la red



Datos CIFAR -10

alcanza una tasa de error del 25% 6 veces más rápido.

Relus aprenden más rápido

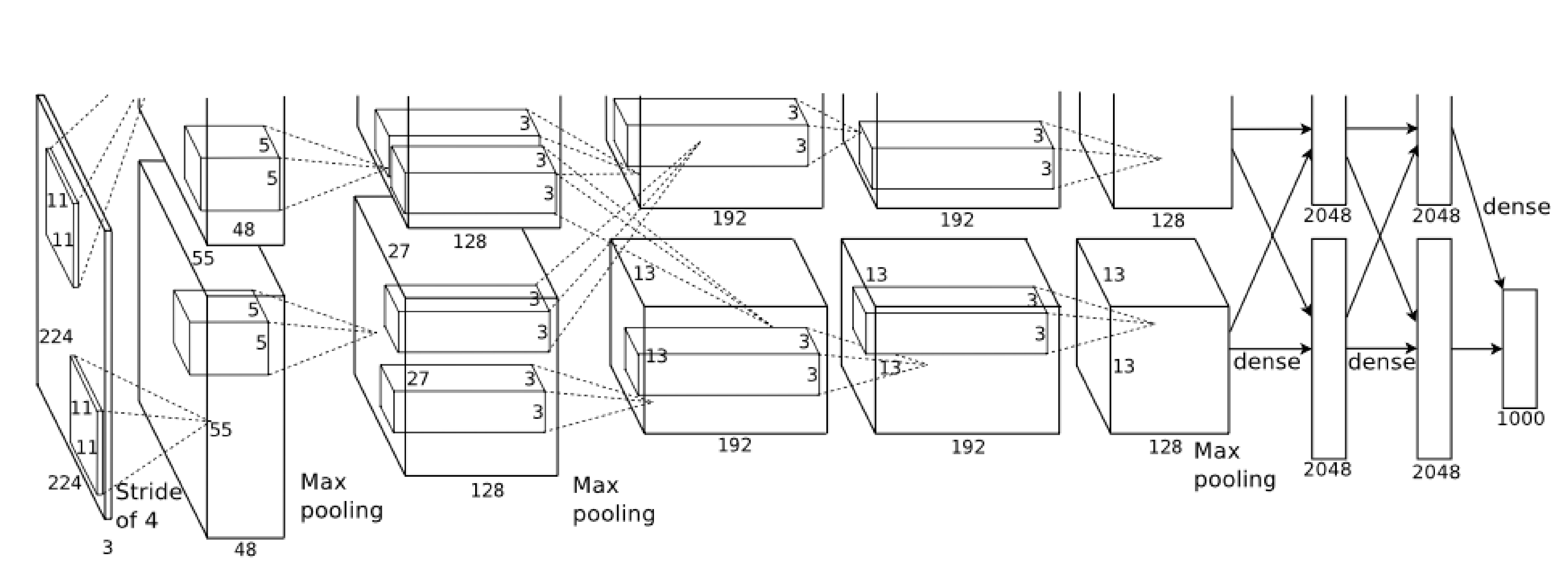


8 CAPAS APRENDIDAS

5 convolucionales
3 totalmente conectadas

CAPAS

1º 96 núcleos
2º 256 núcleos
3º 384 núcleos
4º 384 núcleos
5º 256 núcleos



REDUCCIÓN SOBREAJUSTE

DATA AUGMENTATION

DROPOUT

Aprendizaje

$$
\begin{aligned}
v_{i+1} &:= 0.9 \cdot v_i - 0.0005 \cdot \epsilon \cdot w_i - \epsilon \cdot \left\langle \frac{\partial L}{\partial w}\Big|_{w_i} \right\rangle_{D_i} \\
w_{i+1} &:= w_i + v_{i+1}
\end{aligned}
$$

ºDescenso de gradiente estocástico

º Descenso de peso de 0,0005

ºPesos inicializados a partir de distribución Gaussiana : media 0, Desviación estándar 0.01

º Sesgos inicializados : 2-4-5- MLP -> 1 ; 1-3 ->0

º Tasa de aprendizaje inicial : 0.01

º Entrenamiento de red 90 ciclos aproximadamente

º 5 - 6 días

ε : Tasa de aprendizaje

i : Indice de iteración

v : variable de impulso

<∂L/ ∂w> Di : Media del i esimo lote de la derivada del objetivo con respecto a w, evaluada en wi.

Resultados

Model	**Top-1**	**Top-5**
Sparse coding [2]	*47.1%*	*28.2%*
SIFT + FVs [24]	*45.7%*	*25.7%*
CNN	**37.5%**	**17.0%**

Discusión

APARTE DE ESPERAR A TENER CPUS MÁS EFICIENTES:
(1) ¿SE PUEDE MEJORAR LA PRECISIÓN DEL MODELO DE CLASIFICACIÓN?
(2) ¿QUÉ DICE LA LITERATURA?

IMAGENET CLASSIFICATION WITH DEEP CONVOLUTIONAL NEURAL NETWORKS

ALEX KRIZHEVSKY
ILYA SUTSKEVER
GEOFFREY E. HINTON

ICT3115 - SISTEMAS URBANOS INTELIGENTES

PAULA CORONADO
GASPAR MASIHY