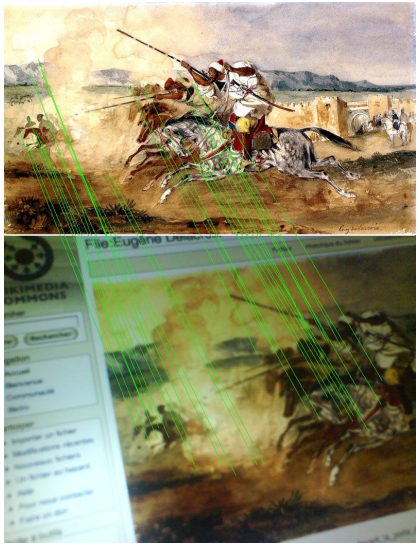
Computer Vision and Deep Neural Network

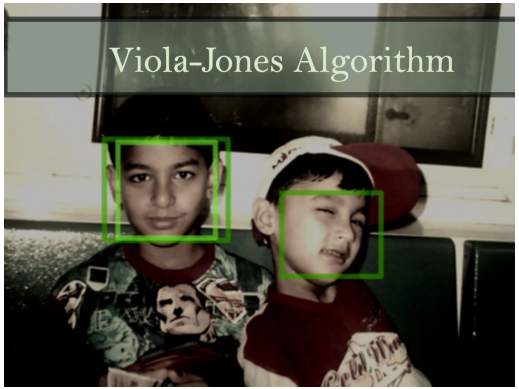
Why deep makes a difference

Computer Vision - State Of The Art before 2012

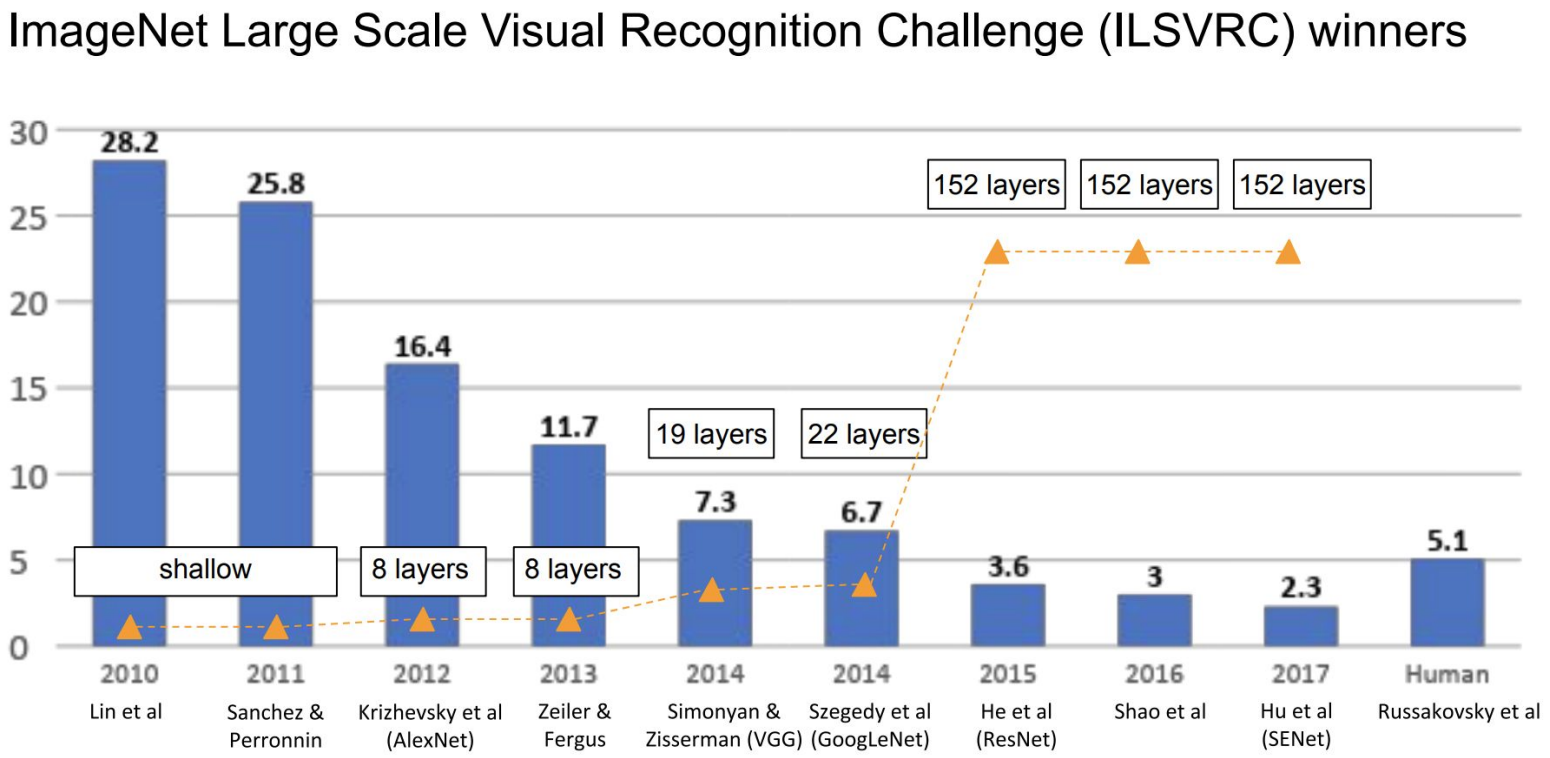
Combination of multiple operations to build a CV system

Filtering :
Average, Gaussian, Laplacian, Median, Bilateral
Thresholding :
Otsu, Ridler-Calvard
Edge Detector :
Prewit, Sobel, Marr-Hildreth, Canny
Corner Detector :
Harris, Förstner
Texture :
Co-occurrence, Haralick
Feature Description :
SIFT, SURF, HOG, Haar wavelets
Shallow Machine Learning:
AdaBoost, Decision Tree,, Linear/Logistic Regression, SVM





Computer Vision breakthrough

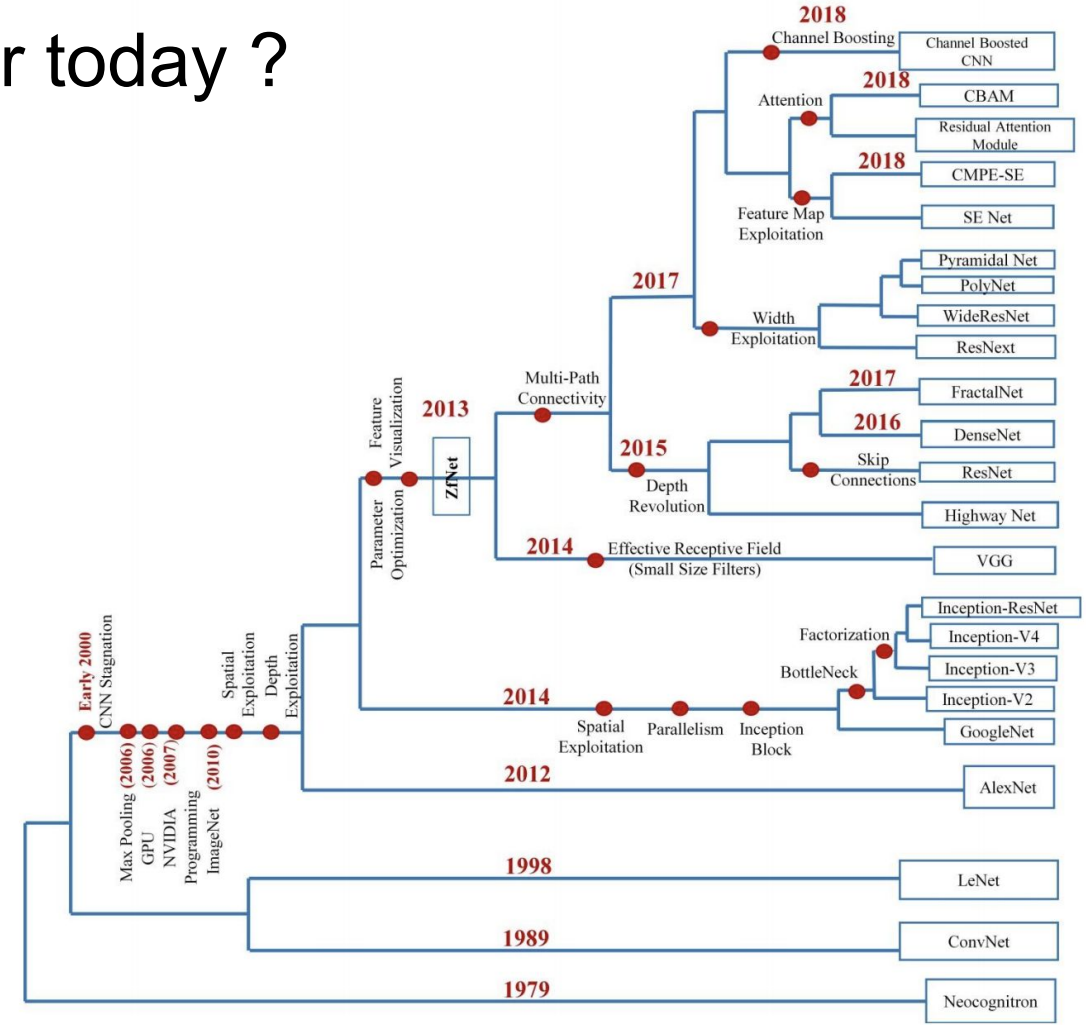


Why CNNs are better today ?

End to end learning (automatic features extraction, comparison, classification)

Transfer Learning

Big improvements on CNN architecture (Pooling, Activation Function, Dropout, NetworkInNetwork, Batch Normalisation, Skip Connection, Depth-wise separable convolution, Attention, Channel Boosting)

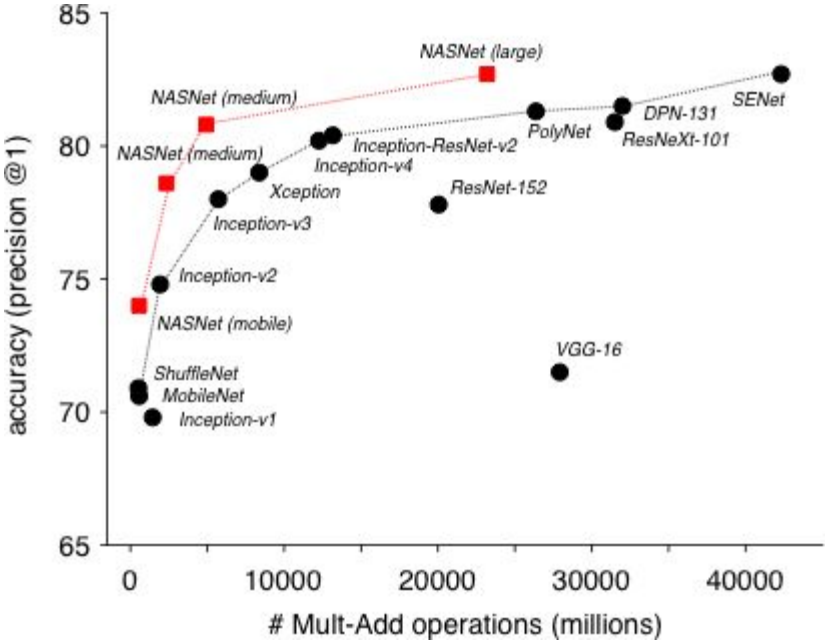


From handcrafted to automatic CNN architecture

Designing neural nets is extremely time intensive

NASNet [https://arxiv.org/pdf/1707.07012.pdf] : Search for an architectural building block on a small dataset and then transfer the block to a larger dataset.

Papers on Neural Architecture Search : https://www.ml4aad.org/automl/literature-on-neural-architecture-search/

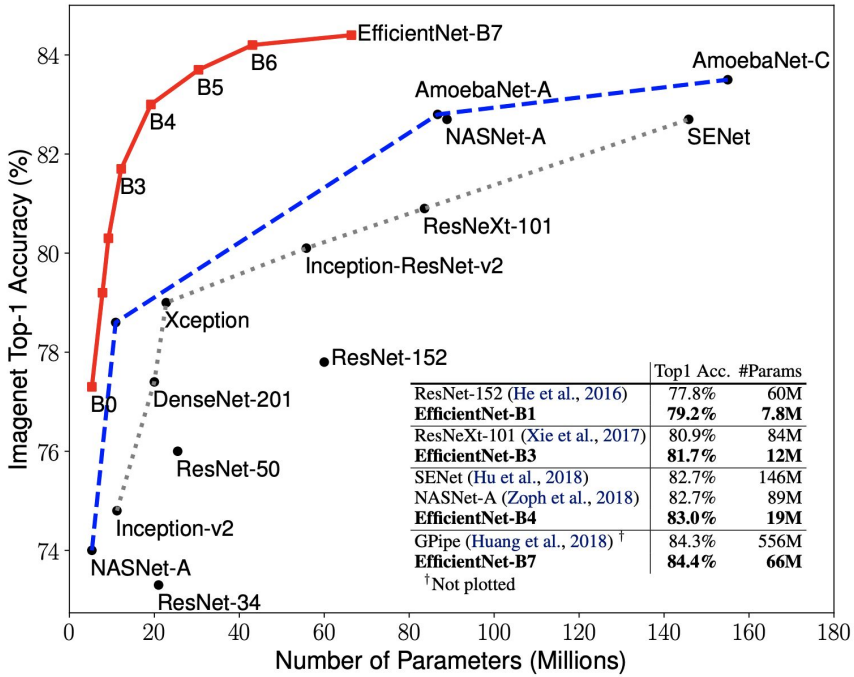


SOTA NAS - EfficientNet

Scaling existing models in terms of depth, width, input image resolution to improve the performance : compound scaling

Instead of scaling only one model attribute, strategically scaling all three of them together delivers better results.

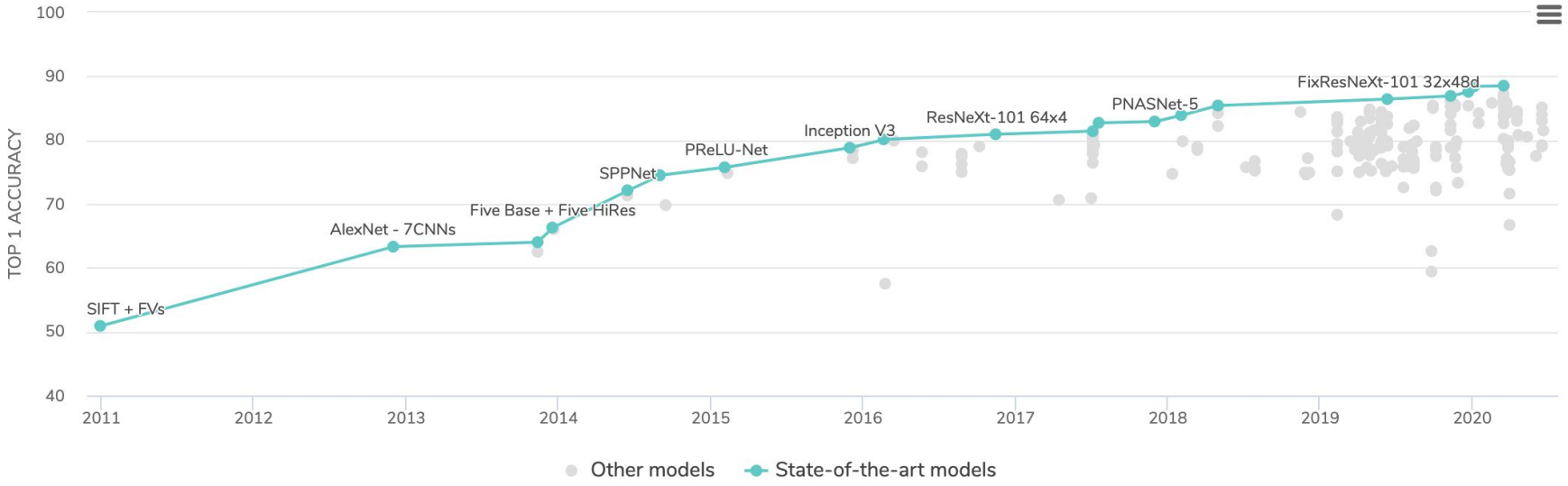
Using a multi-objective Neural Architecture Search that optimizes both accuracy and floating-point operations.



	Top1 Acc.	#Params
ResNet-152 (He et al., 2016)	77.8%	60M
EfficientNet-B1	**79.2%**	**7.8M**
ResNeXt-101 (Xie et al., 2017)	80.9%	84M
EfficientNet-B3	**81.7%**	**12M**
SENet (Hu et al., 2018)	82.7%	146M
NASNet-A (Zoph et al., 2018)	82.7%	89M
EfficientNet-B4	**83.0%**	**19M**
GPipe (Huang et al., 2018) †	84.3%	556M
EfficientNet-B7	**84.4%**	**66M**

†Not plotted

Image Classification on ImageNet



https://paperswithcode.com/sota/image-classification-on-imagenet

Computer Vision and Deep Neural Network

Why deep makes a difference