Applications

CNN

general

most famous commonly employed algorithm

main benefit --> automatically identifies the relevant features

key benefits:
- equivalent representations
- sparse interactions
- parameter sharing

extensively applied in:
- computer vision
- speech recognition
- face recognition

common usage type

convolution layers + FC layers

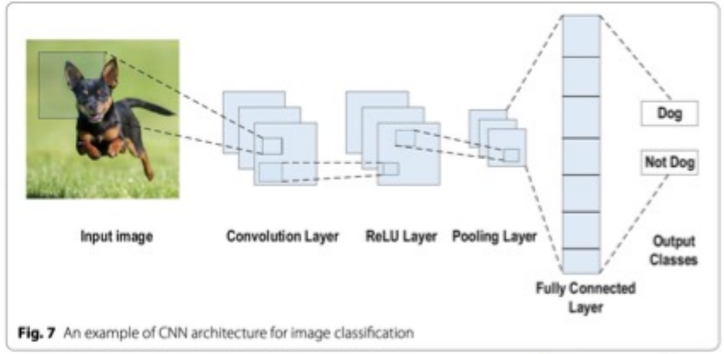


Fig. 7 An example of CNN architecture for image classification

benchmark

Dataset	Num. of classes	Applications	Link to dataset
ImageNet	1000	Image classification, object localization, object detection, etc.	http://www.image-net.org/
CIFAR10/100	10/100	Image classification	https://www.cs.toronto.edu/~kriz/cifar.html
MNIST	10	Classification of handwritten digits	http://yann.lecun.com/exdb/mnist/
Pascal VOC	20	Image classification, segmentation, object detection	http://host.robots.ox.ac.uk/pascal/VOC/voc2012/
Microsoft COCO	80	Object detection, semantic segmentation	https://cocodataset.org/#home
YFCC100M	8M	Video and image understanding	http://projects.dfki.uni-kl.de/yfcc100m/
YouTube-8M	4716	Video classification	https://research.google.com/youtube8m/
UCF-101	101	Human action detection	https://www.crcv.ucf.edu/data/UCF101.php
Kinetics	400	Human action detection	https://deepmind.com/research/open-source/kinetics
Google Open Images	350	Image classification, segmentation, object detection	https://storage.googleapis.com/openimages/web/index.html
CalTech101	101	Classification	http://www.vision.caltech.edu/Image_Datasets/Caltech101/
Labeled Faces in the Wild	–	Face recognition	http://vis-www.cs.umass.edu/lfw/
MIT-67 scene dataset	67	Indoor scene recognition	http://web.mit.edu/torralba/www/indoor.htm

common usage type

convolution layers + FC layers

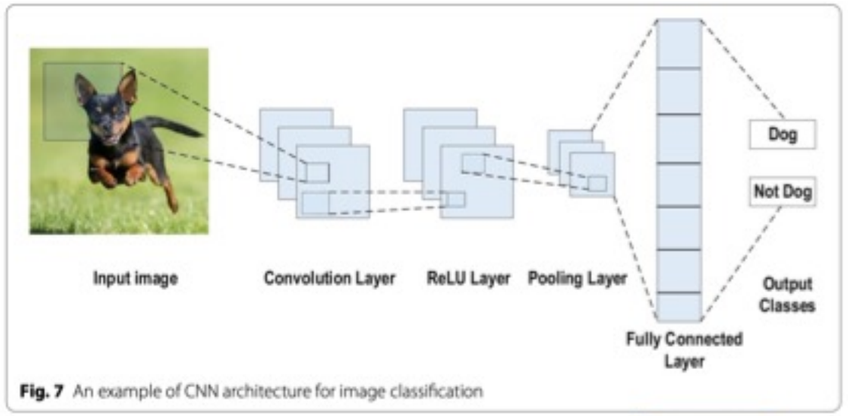


Fig. 7 An example of CNN architecture for image classification

CNN Layers & important components:

Convolutional

Kernel (conv. filters)

learns to extract significant features

weights are adjusted during the training

Convolution

mathematical operation

convolution with input & filter --> output

output shape will change according to filter size.

if padding is valid, then the output size will be same.

padding --> add edge

valid

same

Pooling

reduce size - downsampling

decrease the computational cost

sometimes, it can decrease the performance due to missing relevant info!

Benefits

weight sharing

sparse connectivity

Convolution

mathematical operation

convolution with input & filter --> output

output shape will change according to filter size.

if padding is valid, then the output size will be same.

padding --> add edge

valid

same

Pooling

reduce size - downsampling

decrease the computational cost

sometimes, it can decrease the performance due to missing relevant info!

Benefits

weight sharing

sparse connectivity

Activation function

Non-linearity

ability to learn extra complicated features

allows back-propagation

Common functions

sigmoid

tanh

relu

Most common

lower computational load

dying relu

leaky relu

inputs are never ignored

parametric linear units

similar with leaky relu - leak factor is updated during the training

nosiy relu

Optimiser

gradient descent

learning rate --> step size

momentum

batch gradient descent

accuracy

stochastic gradient descent

training speed

momentum factor: low --> loses ability to avoid local minimum. high-->rapid

mini-batch gradient descent

$$\Delta w_{ij^t} = \left(\eta * \frac{\partial E}{\partial w_{ij}}\right) + (\lambda * \Delta w_{ij^{t-1}})$$

Adam

more memory efficient

less computational power

integrates the pros of momentum and RMSprob

CNN Layers & important components:

After flattening

End of the CNN

Fully connected

CNN classifier

Loss functions

hinge-loss function (binary class.)

cross-entropy or softmax loss func.

euclidean loss (mean squared error)

Backpropagation is about understanding how to **change** the **weights** and **biases** in a network based on the changes of the **cost** function (Error).

performance

back propagation

Dropout

Drop-weights

Data augmentation

Batch normalisation

prevent vanishing gradient

can control poor weight initialisation

reduce the time

reduce internal co-variate shift

Regularisation

Overfitting

train & validation scores

CNN

Optimiser

gradient descent

learning rate --> step size

increase training time

hyperparameters tuning

use transfer learning

regularisation

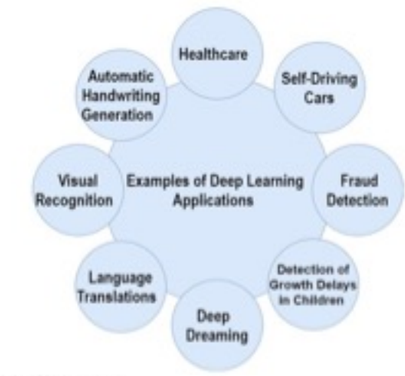
expand dataset with data augmentation

increase the depth-width

performance

	to symmetry mapping-based skip links					
Inception-ResNet-v2	Introduced the concept of residual links	164	ImageNet	3.52	229 × 229 × 3	2016
FractalNet	Introduced the concept of Drop-Path as regularization	40,80	CIFAR-10 CIFAR-100	4.60 18.85	32 × 32 × 3	2016
WideResNet	Decreased the depth and increased the width	28	CIFAR-10 CIFAR-100	3.89 18.85	32 × 32 × 3	2016
Xception	A depthwise convolutionfollowed by a pointwise convolution	71	ImageNet	0.055	229 × 229 × 3	2017
Residual attention neural network	Presented the attention technique	452	CIFAR-10, CIFAR-100	3.90, 20.4	40 × 40 × 3	2017
Squeeze-and-excitation networks	Modeled interdependencies between channels	152	ImageNet	2.25	229 × 229 × 3 224 × 224 × 3 320 × 320 × 3	2017
DenseNet	Blocks of layers; layers connected to each other	201	CIFAR-10, CIFAR-100,ImageNet	3.46, 17.18, 5.54	224 × 224 × 3	2017
Competitive squeeze and excitation network	Both residual and identity mappings utilized to rescale the channel	152	CIFAR-10 CIFAR-100	3.58 18.47	32 × 32 × 3	2018
MobileNet v2	Inverted residual structure	53	ImageNet	–	224 × 224 × 3	2018
CapsuleNet	Put attention to spatial relationships between features	3	MNIST	0.25%	28 × 28 × 1	2018
EfficientNet	Compound coefficient to scale up models	–	ImageNet	3.1	224 × 224 × 3	2019

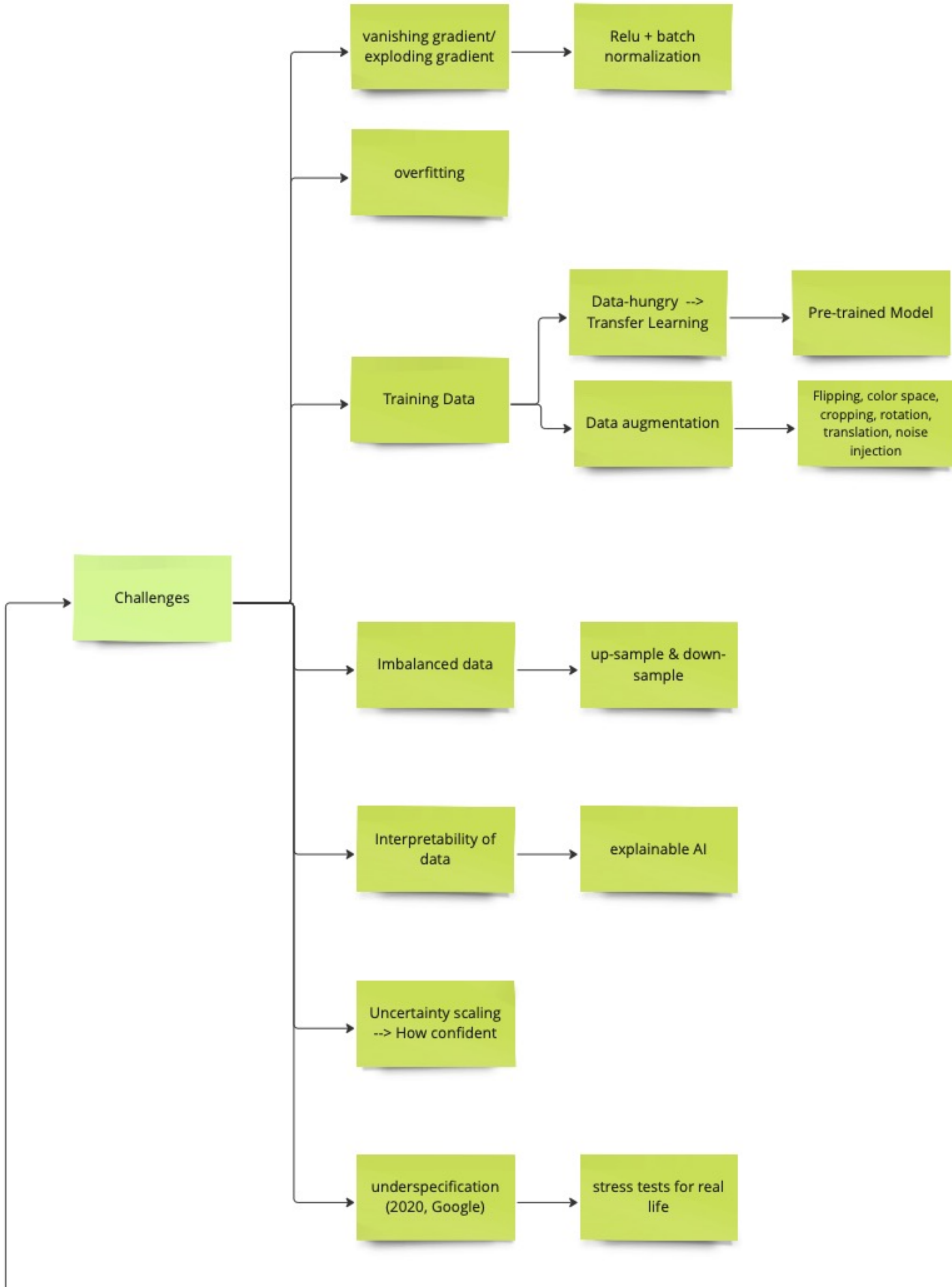
architectures (2012-2020)



Applications

Challenges

most famous commonly employed algorithm


vanishing gradient/ exploding gradient
Relu + batch normalization
overfitting
Training Data
Data-hungry --> Transfer Learning
Pre-trained Model
Data augmentation
Flipping, color space, cropping, rotation, translation, noise injection
Challenges
Imbalanced data
up-sample & down-sample
Interpretability of data
explainable AI
Uncertainty scaling --> How confident
underspecification (2020, Google)
stress tests for real life