

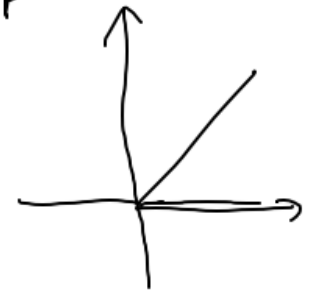
Residual Neural Network

2015-16



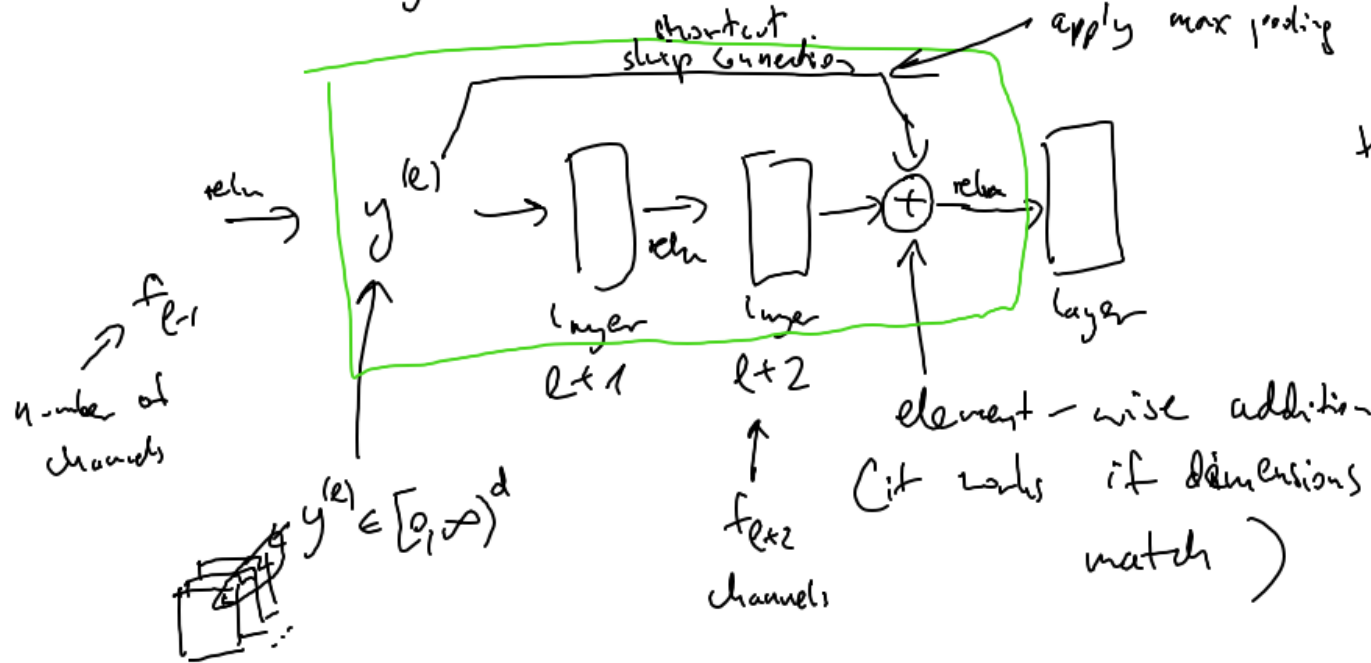
try to add some extra layers

it's quite difficult to model the identity map



The idea: $y^{(l)}$ - output from the l -th layer

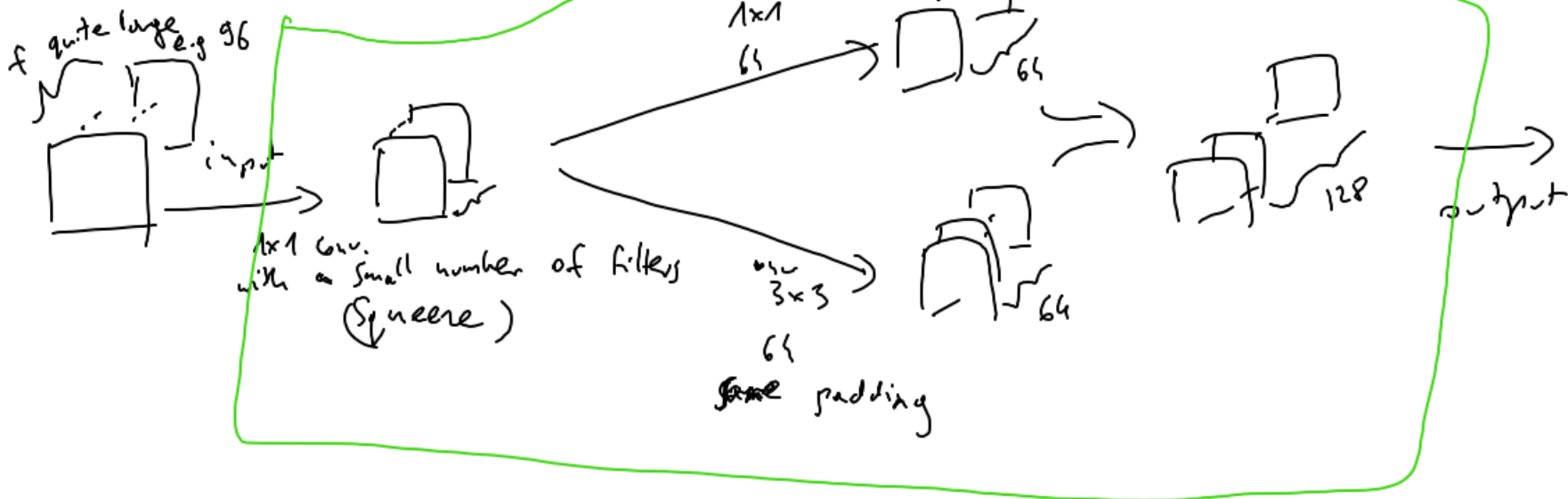
use 1×1 convolutions with the right # of filters



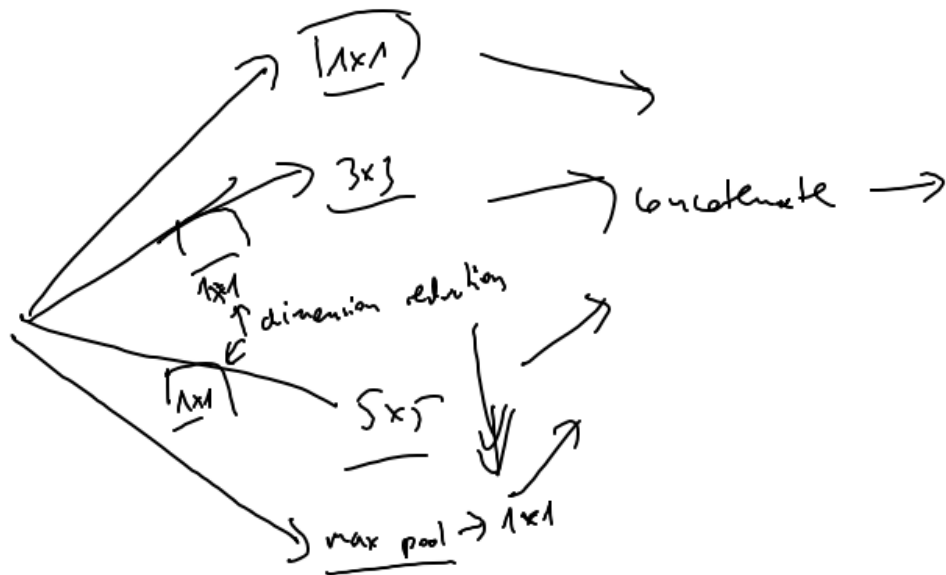
the identity map corresponds to zero weights in the layers $l+1, l+2$. If we take small weights in the layers $l+1, l+2$, then \square is close to identity.

SqueezeNet

fire module



Inception v3



Optimisers

- stochastic gradient descent (SGD)



mini-batch



learning rate is typically decreased
in the process of training
(or increase the batch size)

- in with momentum

take a convex combination (i.e. weighted mean) of the current gradient
and the previous one and use it for adjusting
the weights
(or the previous
convex comb.)

- adaptive... (e.g. Adam)



Face recognition (or fingerprints...)

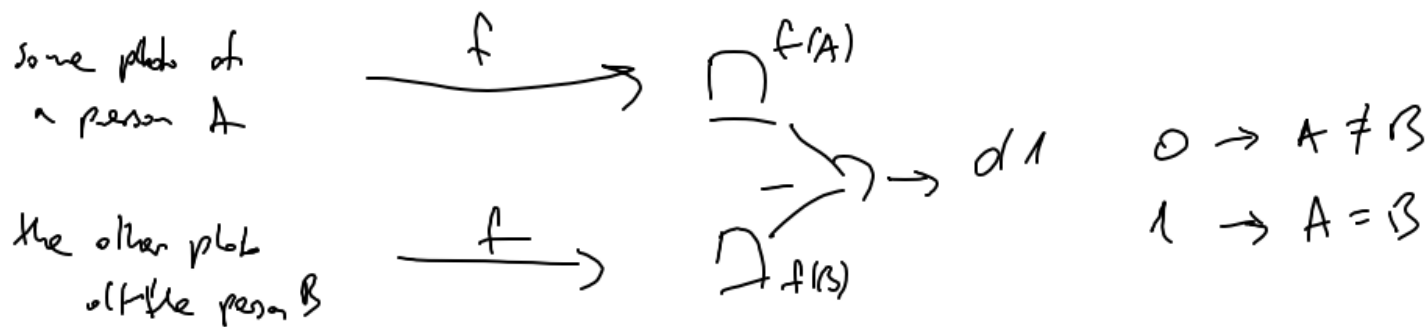


a single photo
of a person

the idea: train the network that will produce some vector of numbers

$$\{\text{face}\} \xrightarrow{f} \mathbb{R}^d$$

with the property that ~~the~~ different photos of the same person
are mapped to similar vectors



the training is done on triplets: A, N, P anchor, negative, positive

\downarrow \downarrow \downarrow

Person 1 Person 2 Person 1

