Page No.
Date

Deep Learning

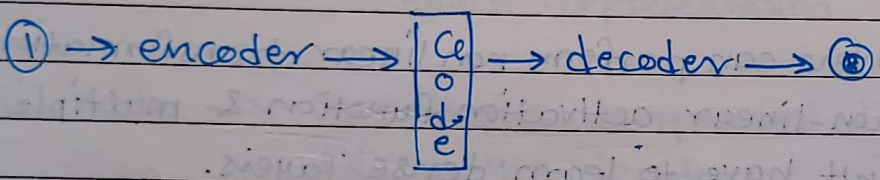
Unit 5 :- Encoder decoder models

- Autoencoder
- undercomplete vs overcomplete.
- Application based questions
- machine translation, autocomplete etc.

* Autoencoders

i) Unsupervised neural networks or self supervised

ii) It has two parts ① encoders
② decoders



Derives important features from data & remove noise.

middle / core layers are of more value than output layer.

Performs the task of data encoding (representation learning)

The code layer

i) Autoencoders work by adding a bottleneck in the network.
ii) This bottleneck forces the network to create compressed (encoded) version of the original input.

The target output of the nueral network in input.

If number of nuerons in the middle layer is less than the number of neurons in the input layer, the network extracts the more effective information.

* Autoencoders are preferred over PCA because:-

i) Autoencoders can perform non linear transformations with a non-linear activation function & multiple layers.
ii) It doesn't have to learn dense layers.
iii) It can use convolutional layer to learn which is better for video, image & series data.
iv) It can make use of pre-trained layers from another model to apply transfer learning to enhance the encoder/decoder.

Page No.
Date

* Applications of autoencoders

i) Image coloring
ii) Feature variation. → Extracts required features & removes noise
iii) Dimensionality reduction
iv) Denoising Image
v) watermark removal

* Architecture of Autoencoders

It consists of 3 layers :- i) Encoder
ii) Code
iii) Decoder

i) Encoder → compresses the image into latent space representation.

ii) Code → Represents compressed input which is fed to the decoder.

iii) Decoder → This layer decodes the encoded image back to original dimension. It is reconstructed from latent space representation.

The encoders are set of convolutional blocks followed by pooling modules that compress the input to the model into a compact section called the bottleneck. The bottlenecks are followed by decoder that consists of a series of upsampling modules to bring the compressed feature back into the form of an image.

Bottleneck → It is the most important part of nueral network, & ironically the smallest one. It exists to restrict the flow of information to the decoder from the encoder. It is designed such a way that maximum information possessed by an image is captured in it.

* Properties of Autoencoders

i) Data-specific :- Autoencoders are only able to compress data similar to what they have been trained on

ii) Lossy - The decompressed outputs will be degraded compared to original i/p.

iii) Learned automatically from examples :- It is easy to train specialized instances of algorithm that will perform well on specific type of ~~data~~ input.

Page No.
Date

* Representation learning

It is defined as set of techniques that allow a system to discover the representations needed for feature ~~extraction~~ detection or classification from raw data.

* Types of Autoencoders

i) Convolutional Autoencoder
ii) Denoising autoencoder

iii) Sparse Autoencoders

- It is simply an autoencoder whose training criterion involves sparsity penalty.
- Fewer nodes activating while still keeping its performance would gurantee that the autoencoder is actually learning latent representation instead of redundant information in our input data.

iv) Variational Autoencoders

Page No.
Date

* Hyper parameters of Autoencoders

i) Code size → smaller size results in more compression
ii) Number of layers → it can consist as many layer as you want.
iii) Number of nodes per ~~year~~ :- Decreases in encoder layer. & increases in decoder
iv) Loss function :- We either use mean-squared error or binary cross-entropy.

Page No.
Date

RNN (Reaurrent Neural Network)

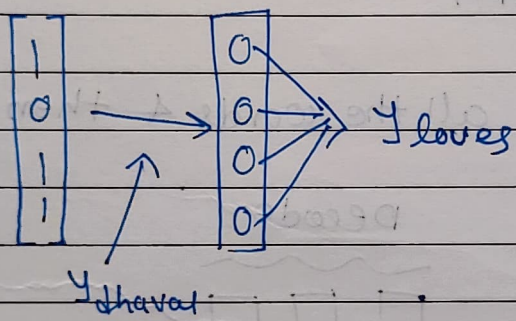
* Issues for using simple Neural Network

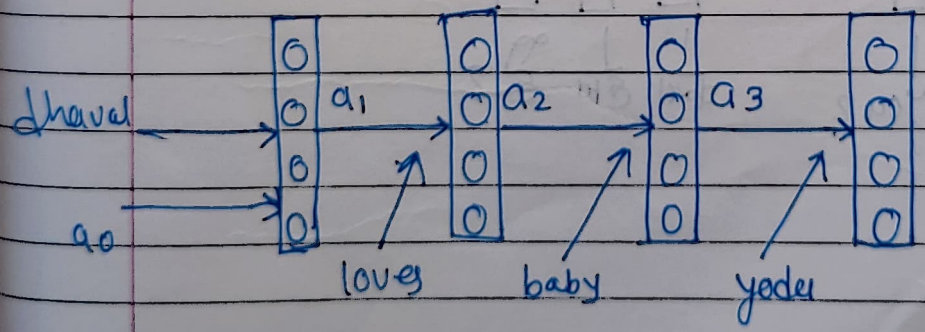
i) No fixed size of neurons in layer
ii) Too much computation
iii) Parameters are not shared.

* Named Entity Recognition

{Vectorising words}

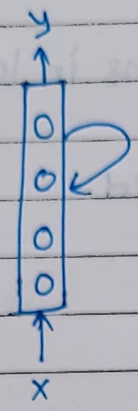
Dhaval loves baby yoda





{This are not different levels}

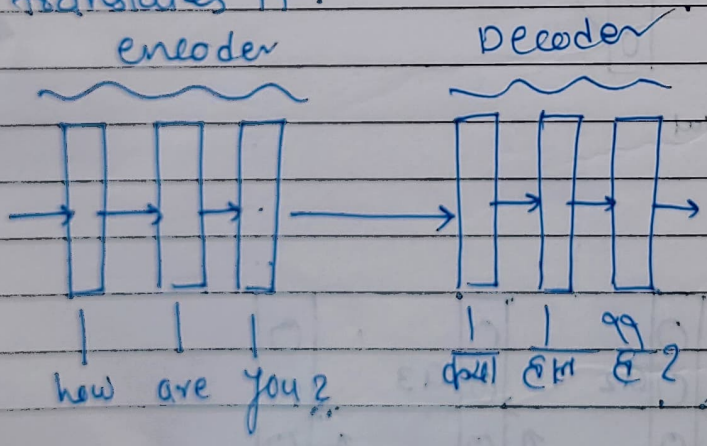
a_1 → output of 1st layer (Activation func)
a_2 → output of 2nd layer



* Training

Language translation.

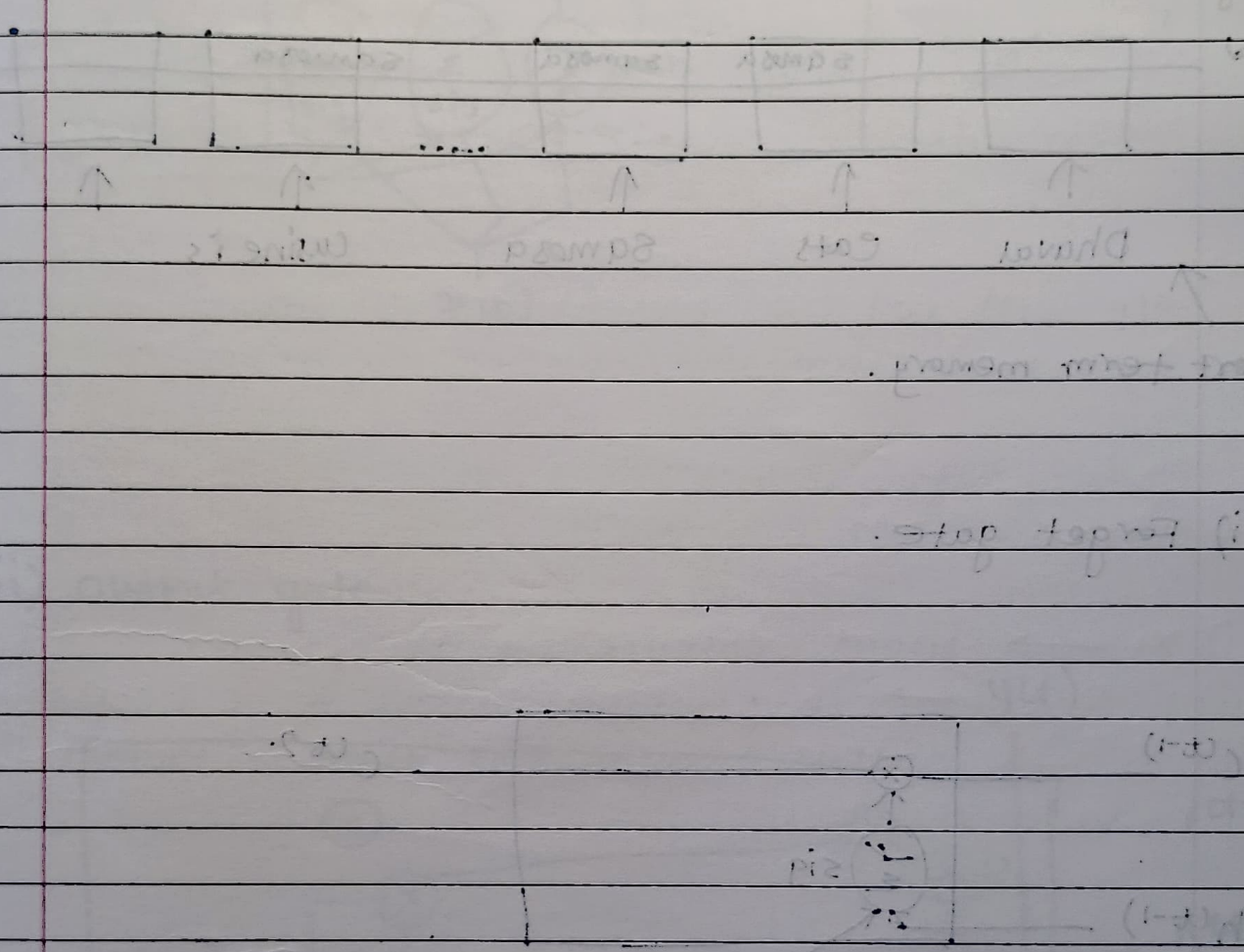
First you supply all the words & then the model translates it.



Page No.
Date

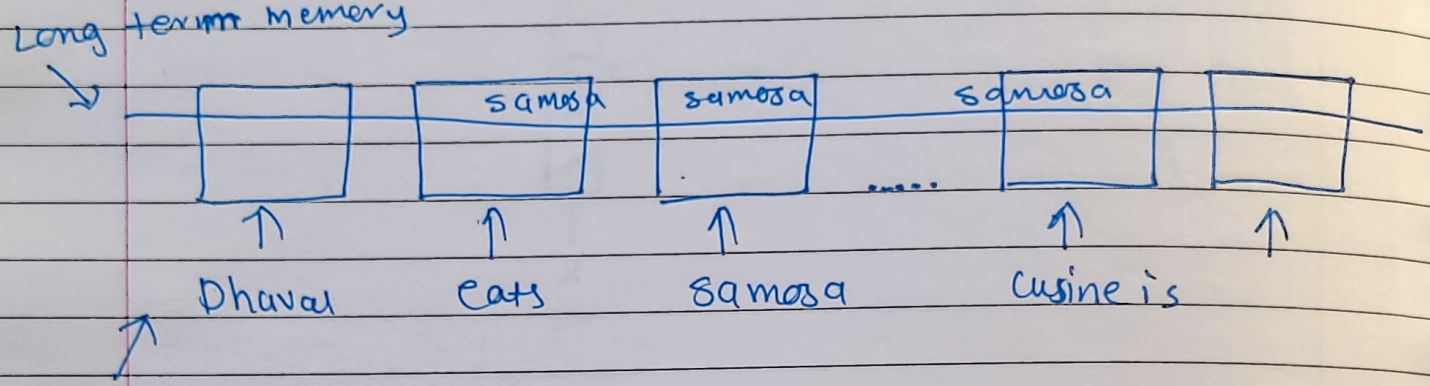
i) Autocompletion
ii) Language translation
iii) NER
iv) Sentiment Analysis

} Applications of RNN

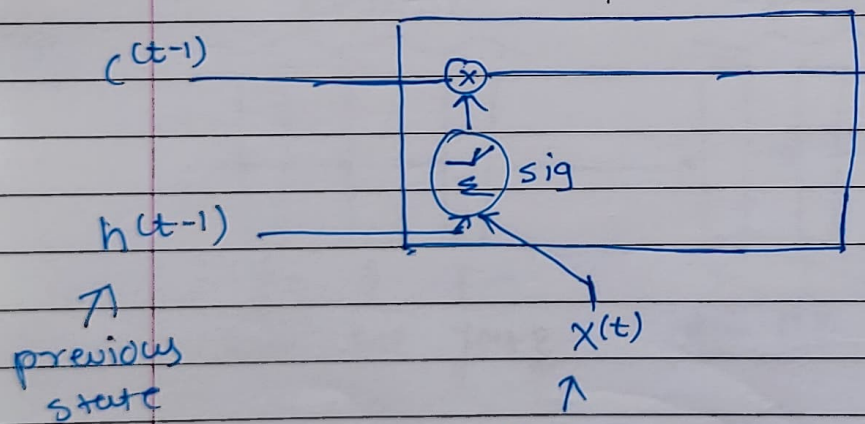


LSTM (Long short term memory)

- Special case of RNN

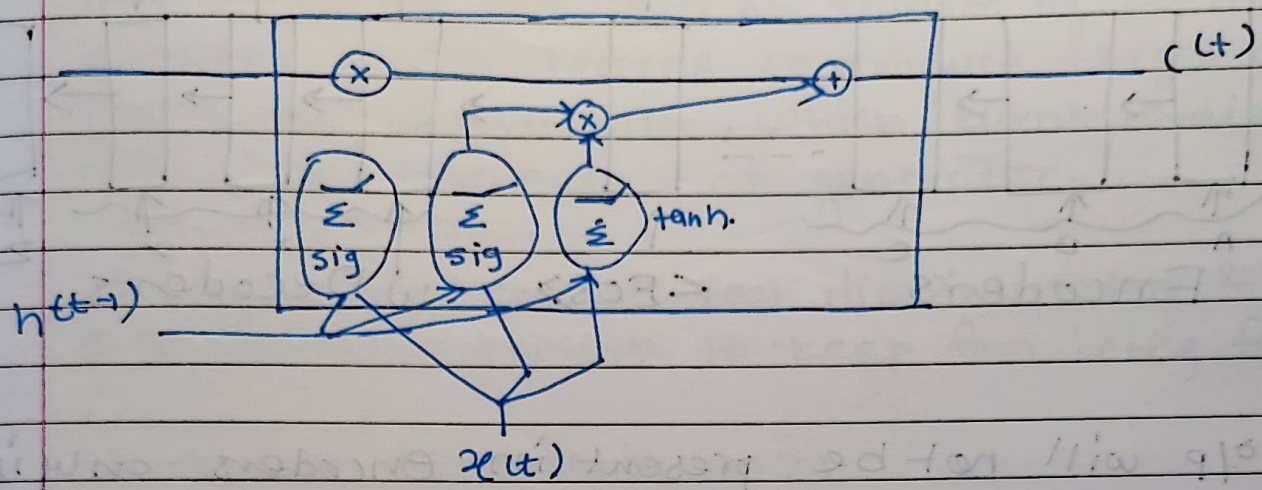


i) Forget gate.

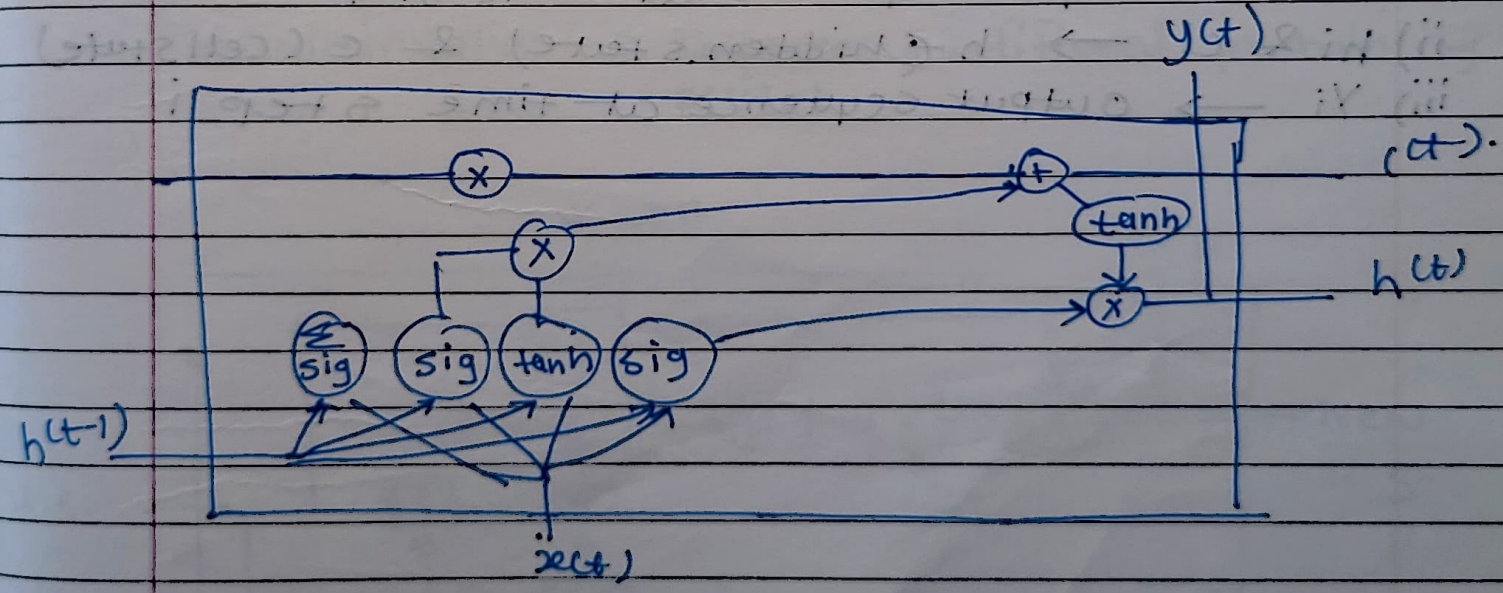


Apply sigmoid funcn.

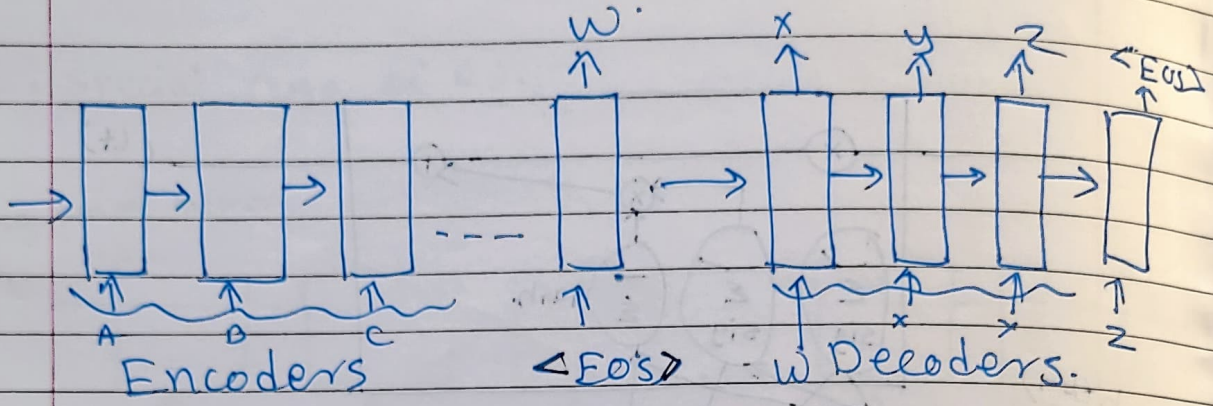
ii) Input gate.



iii) output gate



* Encoder Decoder (Unit 6)



o/p will not be present in encoders only in last time stamp

* ~~Intro~~ LSTM

i) X_i ⟶ input sequence at time step i
ii) h_i & c_i ⟶ h (hidden state) & c (cell state)
iii) Y_i ⟶ output sequence at time step i

Page No.
Date

* Generative Adversarial Networks.

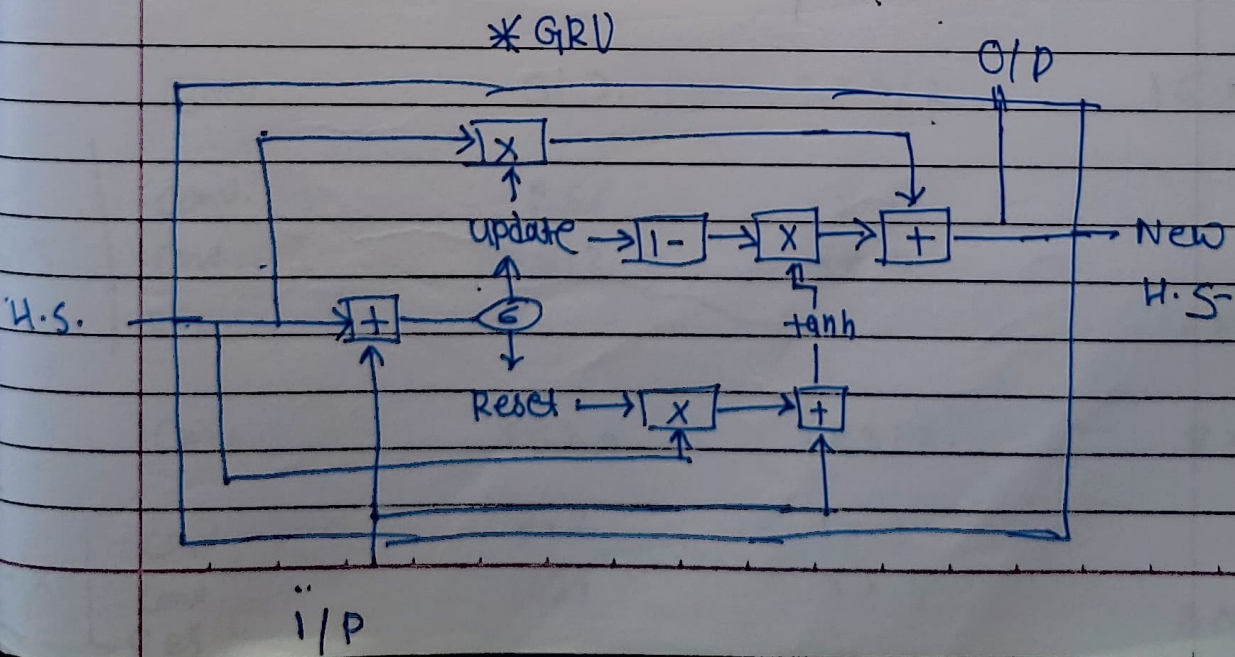
GAN's achieve this level of realism by pairing generator, which learns to produce target output, with ~~des~~ discriminator, which learns to distinguish true data from output of generator.

The generator tries to fool discriminator & discriminator tries ~~to~~ to keep from being fooled

X → set of data instances.
Y → set of labels.

Generative → Capture the joint probability P(X,Y) or just P(x) if there are no labels.

Discriminative → Capture conditional probability P(Y|X).



Page No.
Date

* YOLO Architecture

1st	2nd	3rd	4th
Conv. layer 7x7x64 ÷ 2.	Conv. layer 3x3x192.	Conv. layer 1x1x128 3x3x256. 1x1x26 3x3x512	Conv. layer 1x1x256 } x4 3x3x512 } 1x1x512. 3x3x1028
maxpool 2x2 ÷ 2.	max pool 2x2 ÷ 2	maxpool 2x2 ÷ 2.	maxpool 2x2 ÷ 2.

5th	6th
1x1x512 } x2. 3x3x1024 } 3x3x1024 3x3x1024 ÷ 2.	3x3x1024 3x3x1024.

1X
2X
8X
8X
4X

Page No.
Date

Darknet -53

Convolutional	32	3x3	256x256
Convolution	64	3x3/2.	128x128
convolution	32	1x1	
conv.	64	3x3.	
1X Residual			128x128
Conv.	128	3x3/2	64x64.
conv.	64	1x1	
2X Conv.	128	3x3.	
Residual.			64x64
Conv.	256	3x3/2	32x32
conv.	128	1x1.	
8X Conv.	256.	3x3	
Res.			32x32
Conv.	512.	3x3/2	16x16.
conv.	256	1x1.	
8X Conv.	512	3x3	
Res.			16x16
Conv.	1024	3x3/2.	8x8.
conv.	512.	1x1	
4X Conv.	1024	3x3	
Res			8x8

Page No.
Date

Avgpool
connected
Softmax.

Global.
1000