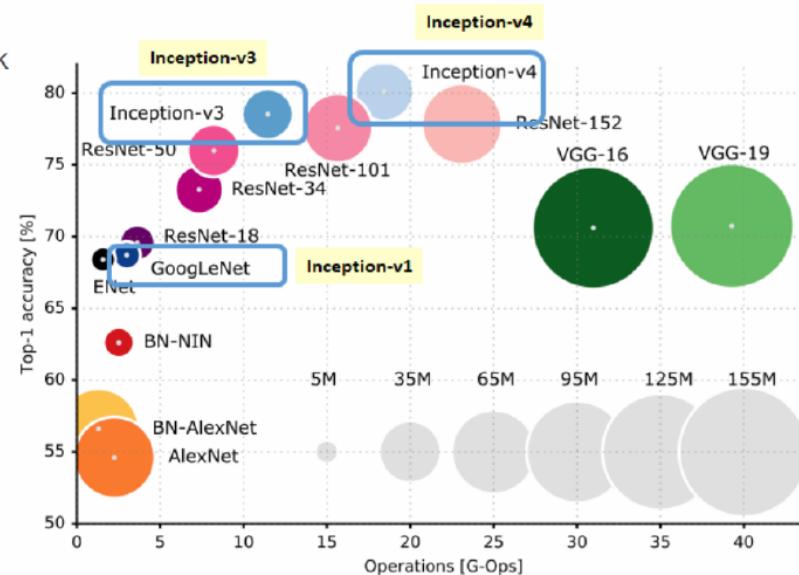
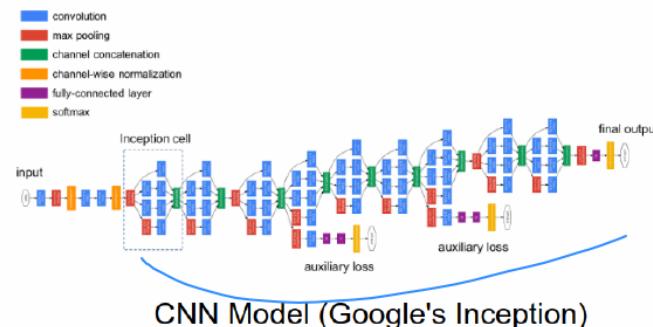


Who can see what you share here? Recording On



# Advances in Deep Learning

- Fuel for DL Revolution
  - Algorithm: Very deep neural network
  - Big Data
  - Computing power (GPU-based)

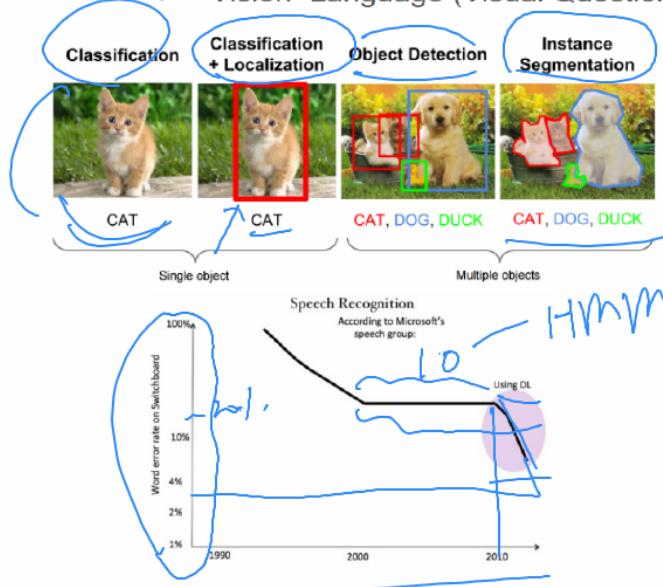


So if it can map is matrix or any algebra operations onto the rendering pipeline of the Gpu, you can utilize this massive parallel computing power

# Advances in Deep Learning



- Major advances in Vision, Speech, NLP - Superhuman in some cases
  - Beyond Classification: Detection, Segmentation (Semantic / Instance); Enhanced speech
  - Vision+Language (Visual Question Answering): Question answering based on images.



When people kept on using things like hidden market model, there was really no improvement in performance

You are screen sharing

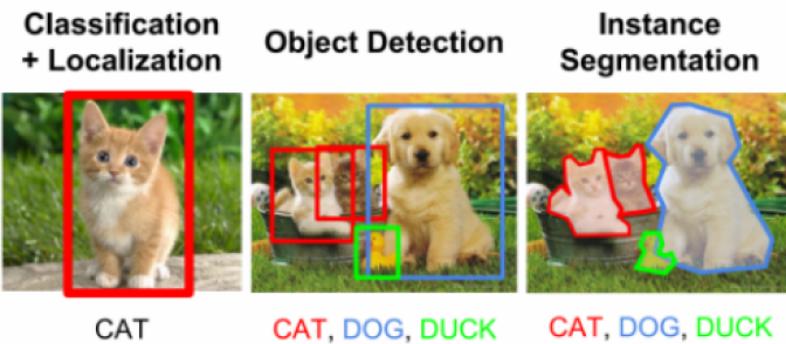
Stop Share

Mouse Select T Text Draw Stamp Spotlight Eraser Format Undo Redo Clear Save

Who can see what you share here? Recording On

# Beyond Classification: Detection, Segmentation (Segmentation)

## Vision+Language (Visual Question Answering): Question answering based on images.



object

Multiple objects

### Speech Recognition

According to Microsoft's speech group:

Using DL

1990 2000 2010 →

 <p>What vegetable is on the plate? Neural Net: broccoli Ground Truth: broccoli</p>	 <p>What color are the shoes on the person's feet? Neural Net: brown Ground Truth: brown</p>	 <p>How many school busses are there? Neural Net: 2 Ground Truth: 2</p>	 <p>What sport is this? Neural Net: baseball Ground Truth: baseball</p>
 <p>What is on top of the refrigerator? Neural Net: magnets Ground Truth: cereal</p>	 <p>What uniform is she wearing? Neural Net: shorts Ground Truth: girl scout</p>	 <p>What is the table number? Neural Net: 4 Ground Truth: 40</p>	 <p>What are people sitting under in the back? Neural Net: bench Ground Truth: tent</p>

For example,

You are screen sharing Stop Share

Mouse Select T Draw Stamp Spotlight Eraser Format Undo Redo Clear Save

Who can see what you share here? Recording On



Choe, Yoonsuck

The first level is composed of 24 rooms.

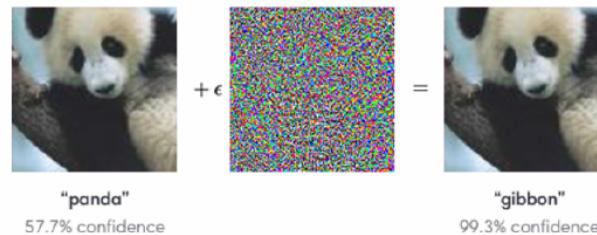
		0		2	
8	9	10	11	12	13
16	17	18	19	20	21
					23

So as you can see, that key was used to open that and then now you're coming from the name

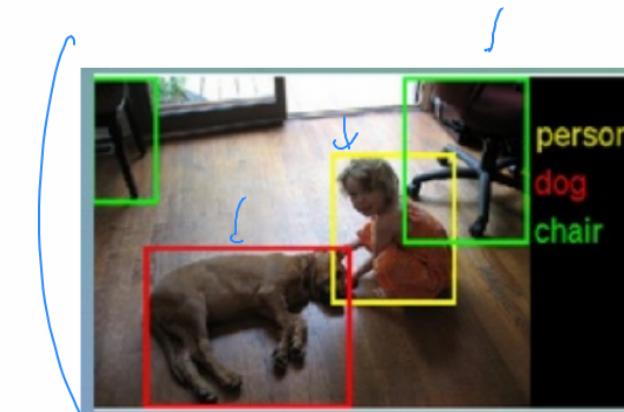


# Limitations of Deep Learning

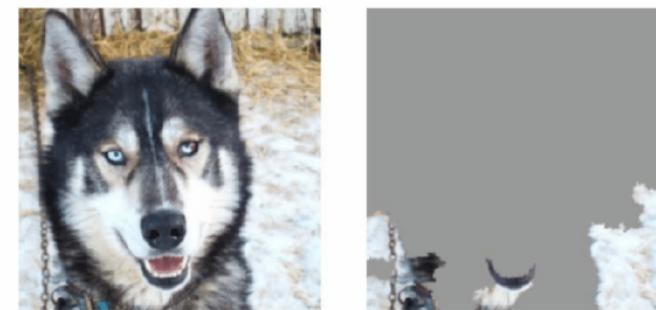
- Limitations of deep learning
  - Need large data, Cannot do complex reasoning
  - Lack of common sense
  - Lack of interpretability, Susceptible to noise
  - Learn spurious features



Sensitive to small perturbations



What to rescue first in case of fire?



I learned snow field feature, not husky feature

So yeah, you have your clear preferences, and humans can easily answer discussion. But

You are screen sharing

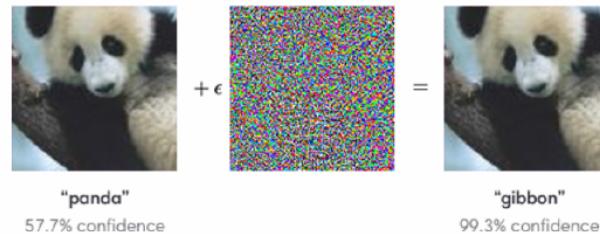
Stop Share

Mouse Select T Text Draw Stamp Spotlight Eraser Format Undo Redo Clear Save Who can see what you share here? Recording On Saved as PNG Show in Folder

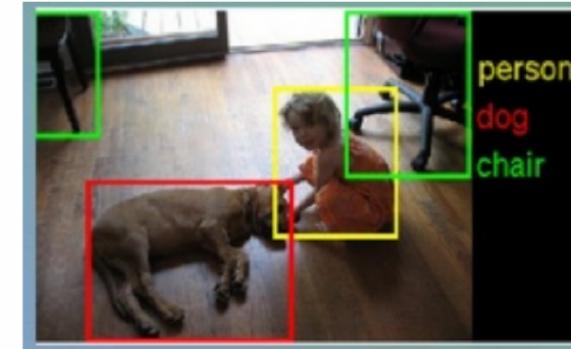
# Limitations of Deep Learning

Choe, Yoonsuck

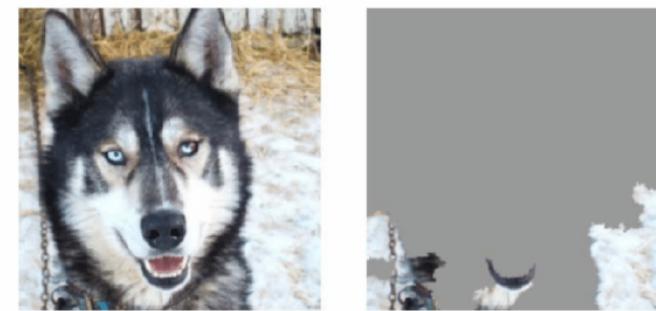
- Limitations of deep learning
  - Need large data, Cannot do complex reasoning
  - Lack of common sense
  - Lack of interpretability, Susceptible to noise
  - Learn spurious features



Sensitive to small perturbations



What to rescue first in case of fire?



I learned snow field feature, not husky feature

You are screen sharing Stop Share

Mouse Select T Text Draw Stamp Spotlight Eraser Format Undo Redo Clear Save

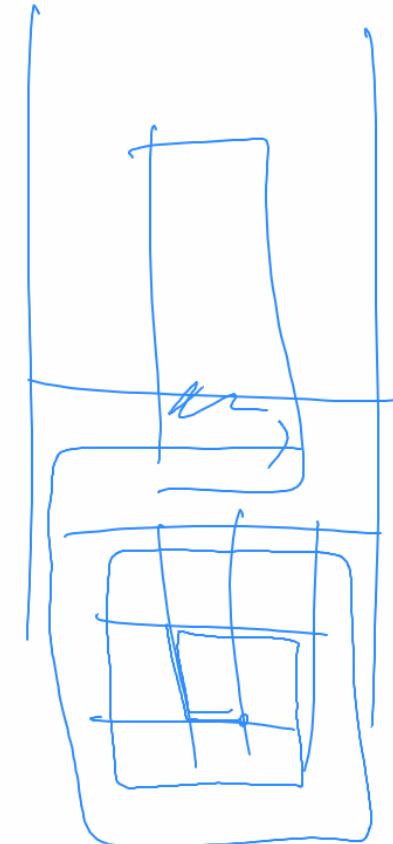
Who can see what you share here? Recording On

# Search and Game Playing



Choe, Yoonsuck

- Uninformed search
- Informed search
- Iterative improvement, Constraint satisfaction
- Game playing



Alright. So search can be very effective, and it could be part  
of a intelligent approach

You are screen sharing

Stop Share

Mouse Select T Text Draw Stamp Spotlight Eraser Format Undo Redo Clear Save

Who can see what you share here? Recording On

Choe, Yoonsuck

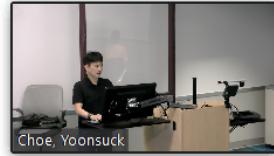
# Search and Game Playing

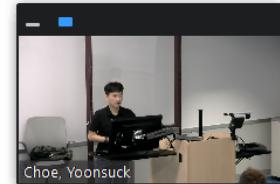
- ↳ • Uninformed search
- ⇒ • Informed search : domain knowledge - heuristic
  - Iterative improvement, Constraint satisfaction
  - Game playing



# Search and Game Playing



- Uninformed search
- Informed search : domain knowledge - heuristic
- Iterative improvement, Constraint satisfaction
- Game playing



## History of AI (III)

XOR

The 60's and 70's: Hard-coded

- ELIZA: pattern-matching-based NLP
- Genetic algorithms
- Knowledge-based systems: scientific domain, engineering domain, natural language
- Minsky and Papert (1969): limitations of perceptron

$x$	$y$	$x \oplus y$
0	0	0
0	1	1
1	0	1
1	1	0

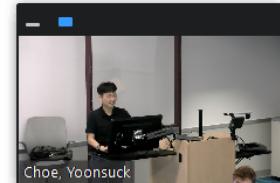
The 80's: Symbolic Logic-based

- 5th generation AI – Prolog.
- Neural networks: Neocognitron, Convolutional Neural Networks, Back Propagation, etc.

Is the people to figure out how to derive a learning rule, to train the connection weights which can allow you to solve this problem, and much more



Who can see what you share here? Recording On



## History of AI (III)

XOR

The 60's and 70's: Hard-coded

- ELIZA: pattern-matching-based NLP
- Genetic algorithms
- Knowledge-based systems: scientific domain, engineering domain, natural language
- Minsky and Papert (1969): limitations of perceptron

$x$	$y$	$x \oplus y$
0	0	0
0	1	1
1	0	1
1	1	0

The 80's: Symbolic Logic-based

- 5th generation AI – Prolog.
- Neural networks: Neocognitron, Convolutional Neural Networks, Back Propagation, etc.

5 days from now. it's a little bit hard to see what's the correspondence between this Xo Sing and the text right so hopefully, you know, you remember

You are screen sharing Stop Share

Mouse Select T Text Draw Stamp Spotlight Eraser Format Undo Redo Clear Save Who can see what you share here? Recording On

Choe, Yoonsuck

## History of AI (VI)

Mid 2010's – present: Age of Deep Learning

- Deep learning
  - Vision : Convolutional Neural Networks
  - Speech : Recurrent Neural Networks
  - Natural language : Recurrent Neural Networks,  
Attention-based (e.g. Transformer)
  - Deep reinforcement learning: Complex games, Robot control, autonomous driving
- Artificial General Intelligence (AGI)
  - Reasoning, Common sense
  - Neurosymbolic AI

Transformers,  
more recently

And then also things like attention. So I actually I said it here and

You are screen sharing Stop Share

Mouse Select T Text Draw Stamp Spotlight Eraser Format Undo Redo Clear Save Who can see what you share here? Recording On

Choe, Yoonseok

## History of AI (VI)

Mid 2010's – present: Age of Deep Learning

- Deep learning
  - Vision : Convolutional Neural Networks
  - Speech : Recurrent Neural Networks
  - Natural language : Recurrent Neural Networks,  
Attention-based (e.g. Transformer)
  - Deep reinforcement learning: Complex games, Robot control, autonomous driving
- Artificial General Intelligence (AGI)
  - Reasoning, Common sense
  - Neurosymbolic AI

Transformers,  
more recently

And you can have layers of piece and so on so it's a very interesting topic, but it's a little bit too complicated

Supervised learning

## Machine Learning Primer

41	34	38	4	38	232	228	211	249	253
37	14	5	42	21	248	237	290	248	246
8	45	6	25	7	237	252	241	258	256
8	9	34	41	7	221	254	252	242	227
28	40	27	28	11	259	231	217	244	231
29	15	39	37	20	217	213	222	255	245
1	31	36	18	0	249	217	245	252	237
16	11	40	8	18	231	244	235	249	232
23	9	15	35	31	243	231	256	242	238
37	12	17	16	17	241	213	215	244	231

Learn from data pair in the form  
of ( $x$ =PATTERN ,  $y$ =LABEL).

Can recognize novel input  $x'$  not  
seen during training

Data set 1 : Logic operation

x1	x2	y
0	0	0
0	1	1
1	0	1
1	1	1

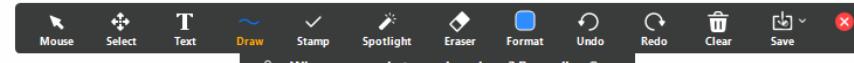
\* Deep learning is a form of  
machine learning

Data set 2 : Animal Recognition

x	y
	Dog
	Cat
	Dog
	Cat

So actually, this is a specific instance called supervised  
machine learning Supervised learning





Who can see what you share here? Recording On

*Supervised learning*

## Machine Learning Primer

41	34	38	4	38	232	228	217	249	253
17	14	5	42	21	248	232	290	248	240
8	45	6	25	7	237	252	241	258	256
8	9	34	41	7	221	254	252	242	227
28	40	27	28	11	259	231	217	244	231
29	15	39	37	20	217	213	222	255	245
1	31	36	18	0	249	217	245	252	237
16	11	40	8	18	231	244	235	249	232
23	9	15	25	19	243	231	256	242	238
17	12	17	16	17	241	213	215	244	231

Learn from data pair in the form  
of ( $x$ =PATTERN ,  $y$ =LABEL).

Can recognize novel input  $x'$  not  
seen during training

Data set 1: Logic operation

$x_1$	$x_2$	$y$
0	0	0
0	1	1
1	0	1
1	1	1

\* Deep learning is a form of  
machine learning

When I show you my face alright. So this tiny patch of  
correspond to this matrix

 → Cat

$x$	$y$
	Dog
	Cat
	Dog
	Cat



Choe, Yoonsuck



Who can see what you share here? Recording On

Supervised learning

## Machine Learning Primer

41	34	38	4	38	232	228	217	249	253
17	14	5	42	21	248	237	290	240	240
8	43	6	25	7	237	252	241	252	256
8	9	34	41	7	221	254	252	242	227
28	40	27	28	11	259	231	217	244	231
29	15	39	37	20	217	213	222	255	245
1	31	36	18	0	249	217	245	252	237
16	11	40	8	18	231	244	237	249	252
23	9	15	25	24	243	231	256	242	238
17	12	17	16	17	241	213	215	214	231

Learn from data pair in the form of ( $x$ =PATTERN,  $y$ =LABEL).

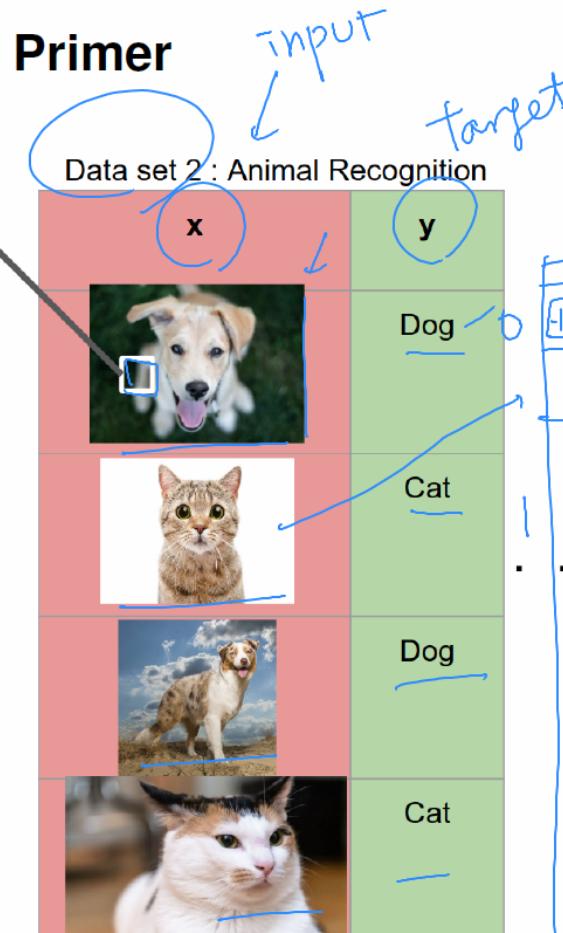
Can recognize novel input  $x'$  not seen during training

Data set 1: Logic operation

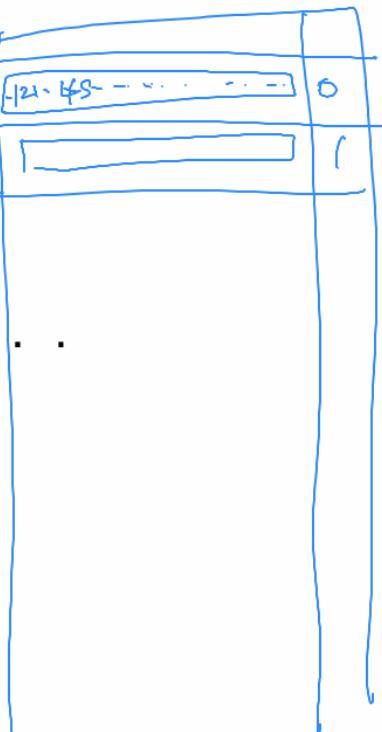
$x_1$	$x_2$	$y$
0	0	0
0	1	1
1	0	1
1	1	1

\* Deep learning is a form of machine learning

Well, 1 2145 that's the theta that then this would be another very, very long, vector and so on,



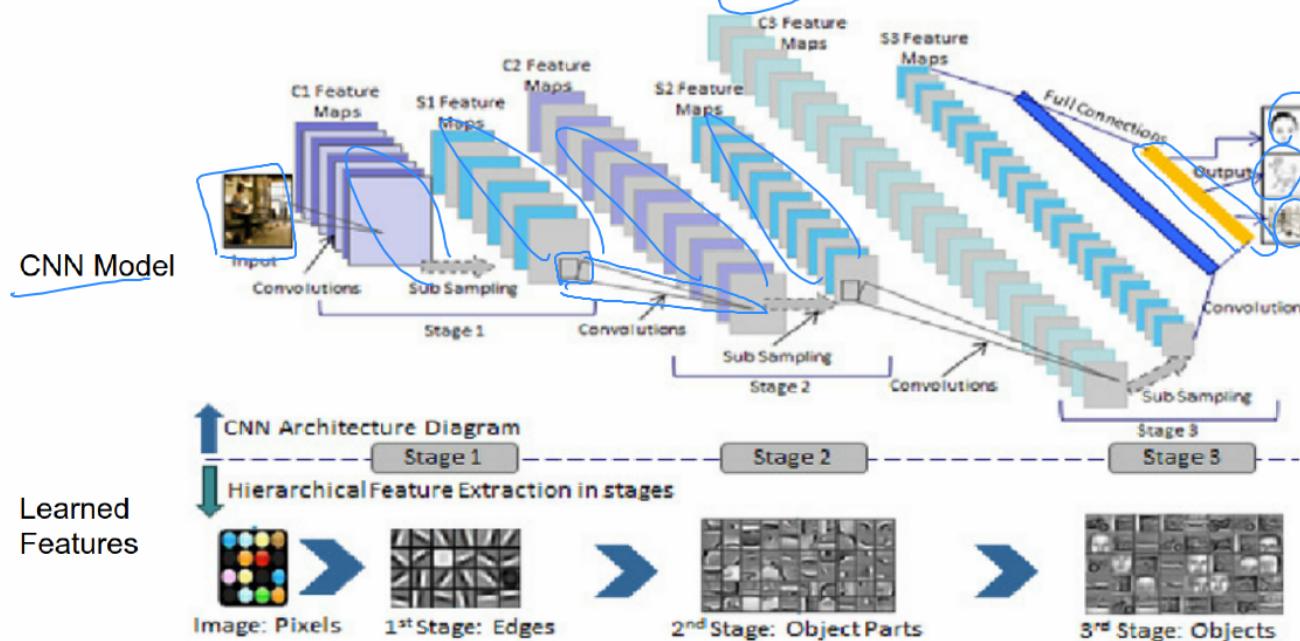
→ Cat





## Deep Learning: mid 2010's – present

- Rapid advances in Deep Learning: CNN, RNN, Attention, Deep RL

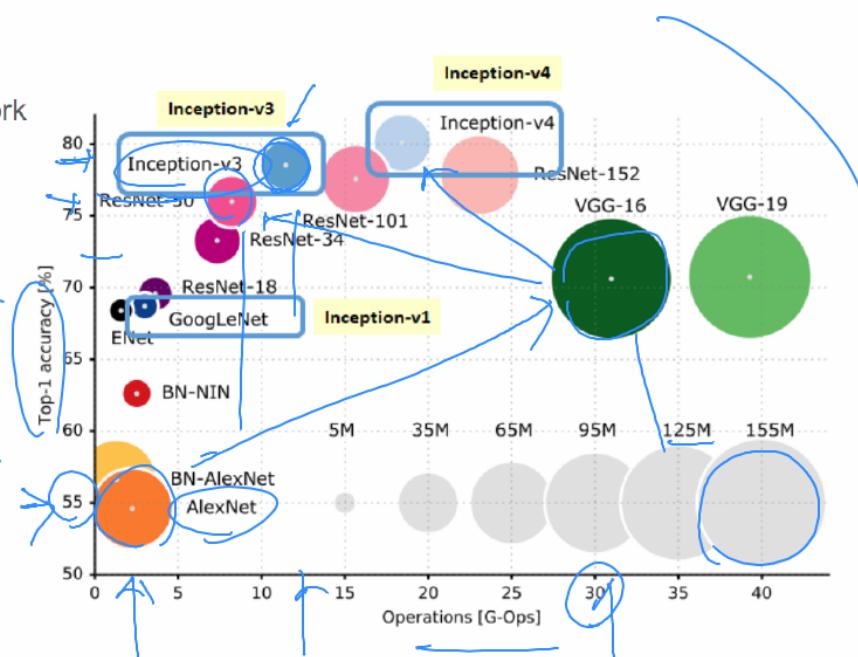
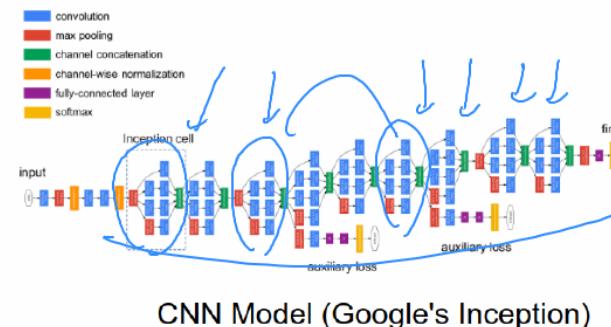


Then in the end you get the output, which will tell you, Is it a human, or is it the dog, or is it a book or something

# Advances in Deep Learning



- Fuel for DL Revolution
  - Algorithm: Very deep neural network
  - Big Data
  - Computing power (GPU-based)



Pervador. Yeah, but don't worry I don't think that there will be that moment just yet.