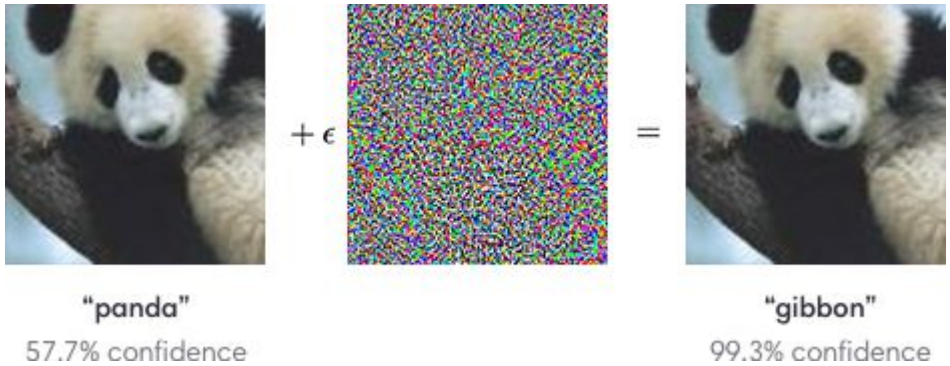
From Neural Networks to Reasoning Machines

Inherent Limitations of Deep Neural Networks

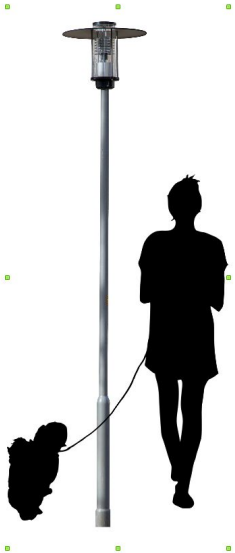
State of Deep Neural Networks (DNNs)

- The Good
  - Best solution available for image, speech and language recognition.
  - Training using DNN models is automated and it is much easier than developing algorithms
- The Bad
  - Fragile as demonstrated by Panda/Gibbon example.
  - Requires a lot of training data. A child picks up the concept of an elephant with one picture.
- The Ugly
  - DNNs Perform very badly in novel situations that were not explicitly trained for.



Perception and Reasoning, Comparison of Humans and Dogs

- Dogs
  - Dogs are nature's statisticians.
  - Unlike some robots dogs do not run into solid objects, even if they had never encountered the objects and the environment before.
- However, dogs cannot project that the leash will get caught in a tree/lamp post if the post is between them and the human.
  - No amount of training will fix this because their reasoning machinery is not equipped with the necessary tools. It is also unlikely to achieve this machinery with DNNs.



- Humans
  - Not as good of statisticians as dogs
  - They have the ability to project mental models and object into future time and space.
- By projecting the trajectory of the leash, they can foresee that the leash will get caught in tree/lamp post.
- Projections are made by the automatic brain system, but, since dogs cannot do that, it shows that this machinery is very specialized and cannot be achieved by a general statistical DNN.

Modeling of the environment

- Our brain recreates the environment like a holodeck.
  - When we perceive our surrounding we create a duplicate of the external objects in our consciousness. I.e., we recognize solid objects and mentally recreate them.
  - Look at a chair and pay attention that your consciousness is creating a three dimensional version of a chair.
  - We can recreate the same model when we recall the past or create one when we are dreaming or imagining.
- .The model has many dimensions
  - Geometry
  - Solid/liquid/gas
  - Rigid/soft, temperature
  - Movement/speed

Reasoning

- Consciousness reasons by creating models which are virtual objects and applies rules such as the laws of physics and simulates the behavior of perceived objects. Examples:
  - Two solid objects cannot occupy the same space. A dog knows that it should not collide with solid objects. This is why as long as it correctly perceives solid objects in its environment it will not run into them. The detection of objects is similar to ML classification task. The class of object is also enriched with attribute like the dimensions, or, coordinates of the solid objects. The fundamental rule of preventing its solid body getting in the space of other solid objects like walls allows dogs not requiring to be trained for different environments and they avoid collision in novel situations.
  - Human drivers project the position of their cars and moving solid objects in the environment and they can detect if these solid objects are going to collide. Once a potential collision is detected they try to solve avoiding collisions by using different alternatives in their skillset such as changing the speed of their car or swerving and changing the direction of their movement.

Using Modeled objects and laws of physics to Process Language

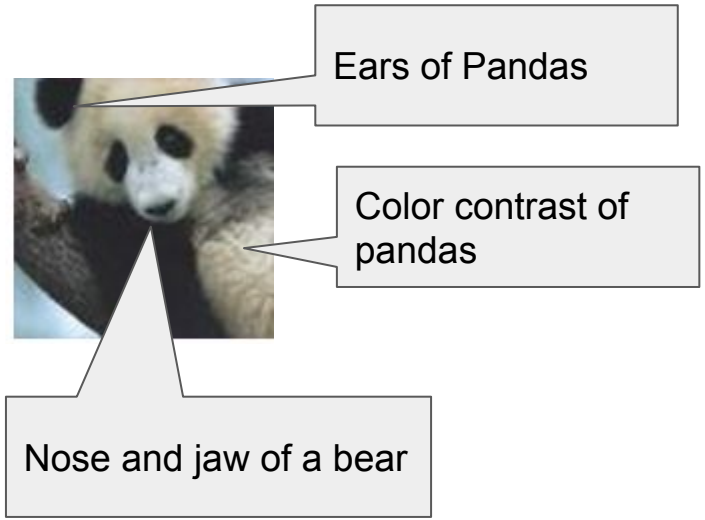
Consider the two sentences “The pen is in the box” and “the box is in the pen”

Humans can resolve the different context using the models for different meanings of pen. As an exercise, try searching for “pen in the box” and you will see no search engine can solve for its correct meaning.

- Writing pen is a solid object
- Box has empty space
- Pen can get into the box

- Box is inside a space
- Animal pen has space
- Box can get into an animal pen

Using Models to Classify Images



Humans construct models of objects by extracting their salient parts.

In the first step they break the objects they perceive into parts and match against the generalizations they have acquired such as mouth being the food intake orifice be it a mammal or a fish.

Their mental model of animal species resembles a tree classifier.

In the second step the put the parts together using the machinery that allows them to create mental objects.

This hybrid approach makes human classification to need less sample data, because when the creation of the concept of elephants is done by differentiating it from every other class. This approach is much more robust than a simple convolutional network that is purely statistical and less prone to noise (panda/gibbon) confusion.

Human Consciousness Compared to other Animals

Humans do not simply have a higher IQ, our machinery has made a quantum leap. Some of these differences are:

1. Humans generalize concepts to a higher degree. This is a quantum leap apart and not simply a better version of generalization.
2. Power of analogy and transferring abstraction from one object to another. We also share this with animals.
3. The capability of creating a replica of objects in our mind and manipulate the mental objects. This is where humans diverge from other animals. Other animals can only use tools that are in their field of vision and do not create whole objects at will and manipulate their mental objects.
4. Our procedural or algorithmic brain that allows us to do tasks like multiplying multiple digit numbers. This procedural capability is also unique to us humans and it allows us to go beyond statistics. We can also use this to manipulate very abstract concepts.

Knowledge Acquisition

Our primary mechanism of acquiring concepts is through analogies. For example we use the phrase “long time ago”. The word “long” is an attribute of physical dimension and means a large quantity. We make an analogy of large length to the passage of time.

Summary

- Once objects are detected and classified we can implementing reasoning by applying rules and laws like the laws of motion.
- This will help handling novel situations.
- This reasoning will help us with avoiding collisions, understanding language and classification as it was demonstrated in earlier examples.
- The detection must be richer than a discrete classification with limited cardinality. For example detection of objects should have the volume and position attributes.
- Next, we can apply rules such as laws of physics to the recognized objects and simulate their dynamics.