Object Detection & Tracking In Space

Team Names

1-Ibrahim Anwer EL-SADAT Ibrahim

2- Ahmed Ibrahim Ali El-Barbry

3- Zakaria Ahmed Zakaria Abd El-Gawad

4- Ibrahim Nashat Mohamed Ali

5- Abanoub Fekry Bedier Gendy

6- Ahmed muhammed Abd_Elrahman Muhammed

Supervised by

Eng : Mohamed Embaby

Smart-AI

Abstract:

This project was an attempt at developing an object detection and tracking system using modern computer vision technology. The project delivers an implemented tracking system. It consists of a servo motor, stand and yolo algorithm and is applicable to areas such as unsupervised surveillance or semi-autonomous control. It is stable and is applicable as a stand alone system or one that could easily be embedded into an even larger system.

The project was implemented in 5 months, and involved research into the area of computer vision and robotic automation. It also involved the inclusion of cutting-edge technology of both the hardware and software kind. The results of the project are expressed in this report, and amount to the application of computer vision techniques in tracking animate objects in both a 2 dimensional and 3 dimensional scene.

Introduction:

A few years ago, the creation of the software and hardware image processing systems was mainly limited to the development of the user interface, which most of the programmers of each firm were engaged in. The situation has been significantly changed with the advent of the Windows operating system when the majority of the developers switched to solving the problems of image processing itself. However, this has not yet led to

the cardinal progress in solving typical tasks of recognizing faces, car numbers, road signs, analyzing remote and medical images, etc. Each of these "eternal" problems is solved by trial and error by the efforts of numerous groups of the engineers and scientists.

As modern technical solutions are turn out to be excessively expensive, the task of automating the creation of the software tools for solving intellectual problems is formulated and intensively solved abroad. In the field of image processing, the required tool kit should be supporting the analysis and recognition of images of previously unknown content and ensure the effective development of applications by ordinary programmers. Just as the Windows toolkit supports the creation of interfaces for solving various applied problems.

Object recognition :

describe a collection of related computer vision tasks that involve activities like identifying objects in digital photographs. Image classification involves activities such as predicting the class of one object in an image. Object localization is refers to identifying the location of one or more objects in an image and drawing an abounding box around their extent.

Object detection :

the work of combines these two tasks and localizes and classifies one or more objects in an image. When a user or practitioner refers to the term "object recognition",

they often mean "object detection". It may be challenging for beginners to distinguish between different related computer vision tasks. So, we can distinguish between these three computer vision tasks with this example:

Image Classification:

This is done by Predict the type or class of an object in an image.

Input:

An image which consists of a single object, such as a photograph.

Output:

A class label (e.g. one or more integers that are mapped to class labels).

Object Localization:

This is done through, Locate the presence of objects in an image and indicate their location with a bounding box.

Input:

An image which consists of one or more objects, such as a photograph.

Output:

One or more bounding boxes (e.g. defined by a point, width, and height).

Object Detection:

This is done through, Locate the presence of objects with a bounding box and types or classes of the located objects in an image.

Input:

An image which consists of one or more objects, such as a photograph.

Output:

One or more bounding boxes (e.g. defined by a point, width, and height), and a class label for each bounding box. One of the further extension to this breakdown of computer vision tasks is object segmentation, also called “object instance segmentation” or “semantic segmentation,” where instances of recognized objects are indicated by highlighting the specific pixels of the object instead of a coarse bounding box. From this breakdown, we can understand that object recognition refers to a suite of challenging computer vision tasks.

For example, image classification is simply straight forward, but the differences between object localization and object detection can be confusing, especially when all three tasks may be just as equally referred to as object recognition.

Humans can detect and identify objects present in an image. The human visual system is fast and accurate and can also perform complex tasks like identifying multiple objects and detect obstacles with little conscious thought. The availability of large sets of data, faster GPUs, and better algorithms, we can now easily train computers to detect and classify multiple objects within an image with high accuracy. We need to understand terms such as object detection, object localization, loss function for object detection and localization, and finally explore an object detection algorithm known as "You only look once" (YOLO).

Image classification :

involves assigning a class label to an image, whereas object localization involves drawing a bounding box around one or more objects in an image. Object detection is always more challenging and combines these two tasks and draws a bounding box around each object of interest in the image and assigns them a class label. Together, all these problems are referred to as object recognition.

Object recognition :

refers to a collection of related tasks for identifying objects in digital photographs. Region-based Convolutional Neural Networks, or R-CNNs, is a family of techniques for addressing object localization and recognition tasks, designed for model performance. You Only Look Once, or YOLO is known as the second family of techniques for object recognition designed for speed and real-time use.

Computer Vision :

branch of the science of computers and software systems which can recognize as well as understand images and scenes. Computer Vision is consists of various aspects such as image recognition, object detection, image generation, image super-resolution and many more. Object detection is widely used for face detection, vehicle detection, pedestrian counting, web images, security systems and self-driving cars.

In this project, we are using highly accurate object detection-algorithms and methods such as R-CNN, Fast-RCNN, Faster-RCNN, RetinaNet and fast yet highly accurate ones like SSD and YOLO.

Using these methods and algorithms, based on deep learning which is also based on machine learning require lots of mathematical and deep learning frameworks understanding by using dependencies such as TensorFlow, OpenCV, imageai etc, we can detect each and every object in image by the area object in an highlighted rectangular boxes and identify each and every object and assign its tag to the object. This also includes the accuracy of each method for identifying objects.

Motivation :

Image recognition/image processing is in the forefront of Artificial Intelligence today. It is however far from perfection. Seemingly simple scenarios, such as object detection, face recognition, removing motion blur, etc.

and more complex scenarios such as compression artefacts, scratch detection, sensor noise, and spilling detection are applications of image recognition/image processing.

Digitized images are often represented as a two-dimensional (2D) array of pixels values.

Each pixel value which makes up the color scheme of the image is often influenced by an array of factors such as light intensity. Visual scene is projected unto a surface, where receptors (natural or artificial) produce values that depend on the intensity of incident light. These exciting concepts are however hard to implement. Forming an image leads to loss of details of information while collapsing a three-dimensional (3D) image into a two-dimensional image. Many other factors are responsible for why image recognition/ image processing is hard. Some of such factors are noise in the image (pixels values that are off from its surrounding pixels), mapping from scene to image etc. In recent years, during the ImageNet Large Scale Visual Recognition Competition (ILSVRC, 2015), computers were going better than humans in the image classification task. In 2016, a faster object detector, YOLO, was proposed to implement object detection in real-time situation. Our motivation is to apply YOLO to object detection task of URL links within an image scene. We will also be comparing the speed and accuracy of this with an OCR software.

What is Object Detection?

Object detection is a computer vision implementation that makes a system (an algorithm) about to estimate the location of objects in a digitized scene such as an image or video. Usually, a bounded box is wrapped around the detected object which helps humans locate the object quicker than unprocessed images. For this discourse, an object is the representation of a physical object (URL) in an image. In image processing, it is an identifiable portion of an image that can be interpreted as a single unit. This creates a sharp contrast to the layman's idea that an image or an object are interchangeable. Usually, an image may contain one or more objects, the discernibility of which is of upmost importance. For instance, in a single image the objects contained can range from a single unit to as many objects of as numbers, bordering on infinity. Although "detection" could mean locating a hidden concealed object, detection may also mean the ability of an intelligence to signify the existence and identification of an object. The object in question does not have to be hidden. This later form is the form in which we based this thesis. Interpreting the object localization for object detection can be done in various ways, including creating a bounding box around the object or marking every pixel in the image which contains the object (called segmentation). Thus, given an image or video stream, an object detection model should be able to

identify which of a known set of objects might be present and provide information about their positions within the image. Object Detection is widely employed in various computer vision tasks such as image annotation, activity recognition, face detection, face recognition, video object co-segmentation. It is also used in tracking objects. For example, tracking a ball during a football match, tracking movement of a cricket bat, or tracking a person in a video are some of its various uses. Fundamentally, two approaches to image detection exist, they are machine learning-based approaches and deep learning-based approaches. In more traditional ML-based approaches, computer vision techniques are employed to analyze various features of an image, such as the color histogram or edges, to identify groups of pixels that may belong to an object. These features are then inputted into a regression model that predicts the location of the object along with its label. Some Machine learning approaches are Viola–Jones object detection framework based on Haar features, Scale-invariant feature transform (SIFT) and Histogram of oriented gradients (HOG) features. On the other hand, deep learning-based approaches employ convolutional neural networks (CNNs) to perform end-to-end, unsupervised object detection, in which features do not need to be defined and extracted separately. Some Deep learning approaches are: Region Proposals (R-CNN,Fast R-CNN,Faster R-CNN, cascade R-CNN.), Single Shot MultiBox Detector (SSD), You Only

Look Once (YOLO), Single-Shot Refinement Neural Network for Object Detection (RefineDet),Retina-Net and Deformable convolutional networks.

Image Recognition :

Object detection is often confused with image recognition (Figure 1). A picture of a dog receives the label "dog". A picture of two dogs, still receives the label "dog". Object detection, on the other hand, draws a box around each dog and labels the box "dog". The model predicts where each object is and what label should be applied. In that way, object detection provides more information about an image than recognition. Recognition in this context is the ability of an intelligent system to identify an object based on certain similarities that it shares with another object that the intelligence has previously encountered.

Recognition may be based on inference or relation, that is, a situation whereby an intelligence is able to recognize an object because it recognizes similarities in form and properties. Recognition may also occur because the Artificial intelligence has encountered the exact specimen at a previous instance. In human beings' recognition is a cognitive process that happens seamlessly and almost instantly without any hitch. The human brain is capable of learning and adapting information with minimal effort such that even humans that are still in the developmental stages of their existence can recognize objects and patterns easily.

Humans can recognize a multitude of objects in images with little effort, even though the image of the objects may vary somewhat in different viewpoints, in many different sizes and scales or even when they are translated or rotated. Objects can even be recognized when they are partially obstructed from view. Artificial intelligence, however, does not innately possess this cognitive ability. For Artificial Intelligence to acquire this level of skill they must acquire training. This training is usually acquired by 'teaching' the A.I. using coding, datasets, and databases. This task is still a challenge for computer vision systems given these A.I. systems need to be trained for each class of object it is meant to recognize.

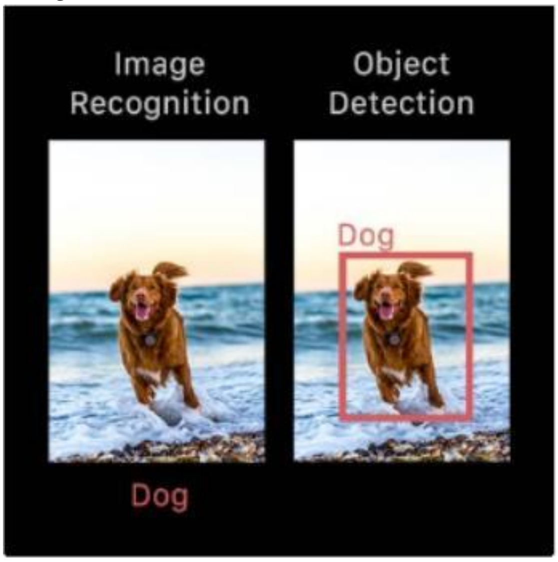


Figure 1

The object recognition problem can be defined as a labelling problem based on models of known objects. Formally, given an image containing one or more objects of interest (and background) and a set of labels corresponding to a set of models known to the system, the system should be able to accurately assign correct labels to regions, or a set of regions, in the image. The object recognition problem is closely tied to the segmentation problem: without at least a partial recognition of objects, segmentation cannot be done, and without segmentation, object recognition is not possible.

Object recognition :

difficult computational problem. The core problem is that each object in the world can cast an infinite number of different 2-D images onto the retina as the object's position, pose, lighting, and background vary relative to the viewer. Yet the brain solves this problem effortlessly. Progress in understanding the brain's solution to object recognition requires the construction of artificial recognition systems that ultimately aim to emulate our own visual abilities, often with biological inspiration.

Object Tracking :

The image segmentation step allows us to separate foreground objects from the scene background. However, we are still working with full images, not the individual points of motion desired for human motion detection. The problem of computing the motion in an image is known as finding the optical flow of the image. There are a variety of well-understood techniques for doing so, but the KanadeLucas-Tomasi method stands out for its simplicity and lack of assumptions about the underlying image. The Kanade-Lucas-Tomasi algorithm uses the image’s gradients to predict the new location of the feature already detected — iterating until the new location is converged upon. Since this approach is based on a Taylor series expansion, it makes no assumptions about the underlying image.

to the viewer. Yet the brain solves this problem effortlessly. Progress in understanding the brain's solution to object recognition requires the construction of artificial recognition systems that ultimately aim to emulate our own visual abilities, often with biological inspiration.

Object Tracking :

The image segmentation step allows us to separate foreground objects from the scene background. However, we are still working with full images, not the individual points of motion desired for human motion detection. The problem of computing the motion in an image is known as finding the optical flow of the image. There are a variety of well-understood techniques for doing so, but the KanadeLucas-Tomasi method stands out for its simplicity and lack of assumptions about the underlying image. The Kanade-Lucas-Tomasi algorithm uses the image's gradients to predict the new location of the feature already detected — iterating until the new location is converged upon. Since this approach is based on a Taylor series expansion, it makes no assumptions about the underlying image.

Artificial Intelligence

Artificial intelligence (AI) intelligence demonstrated by machines, as opposed to the natural intelligence displayed by animals including humans. AI research has been defined as the field of study of intelligent agents , which refers to any system that perceives its environment and takes actions that maximize its chance of achieving its goals.

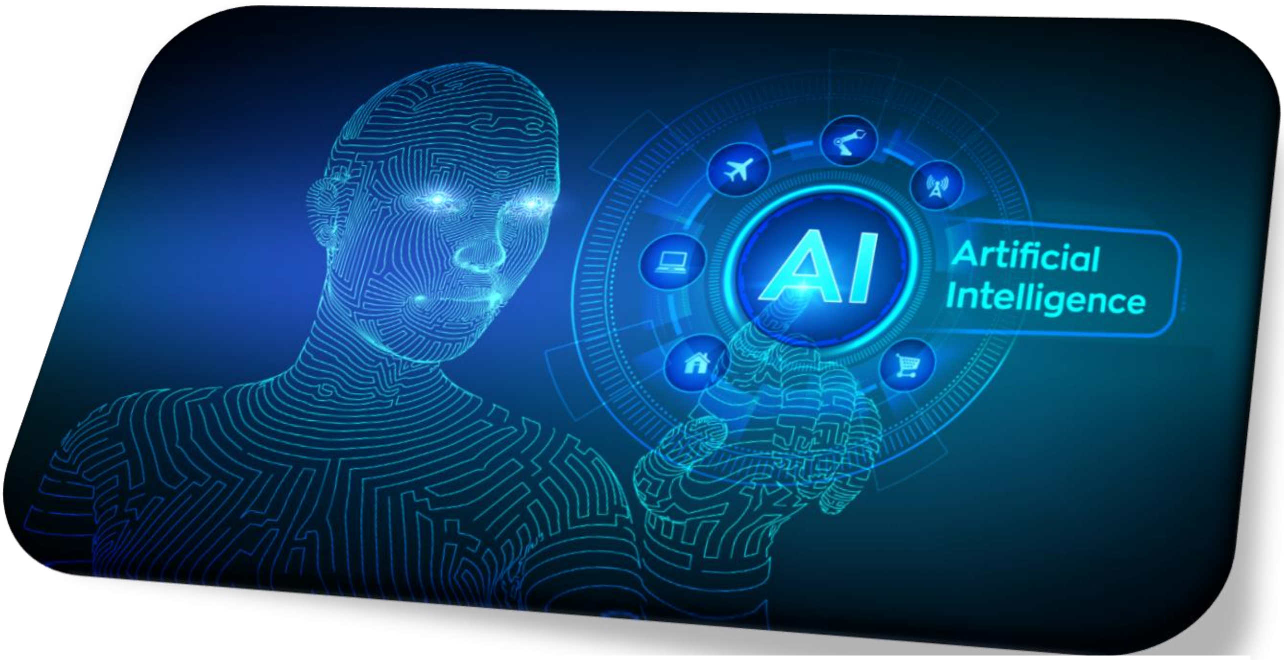


Figure 2

The term "artificial intelligence" had previously been used to describe machines that mimic and display "human" cognitive skills that are associated with the human mind such as "learning" and "problem-solving". This definition has since been rejected by major AI researchers who now describe AI in terms of rationality and acting rationally, which does not limit how intelligence can be articulated.

Types of Artificial Intelligence

Reactive AI :

The most basic type of artificial intelligence is reactive AI, which is programmed to provide a predictable output based on the input it receives. Reactive machines always respond to identical situations in the exact same way every time, and they are not able to learn actions or conceive of past or future.

Examples of reactive AI include:

- **Deep Blue, the chess-playing IBM supercomputer that bested world champion Garry Kasparov**

- **Spam filters for our email that keep promotions and phishing attempts out of our inboxes**

Reactive AI was an enormous step forward in the history of artificial intelligence development, but these types of AIs can't function beyond the tasks they were initially designed for. That makes them inherently limited and ripe for improvement. Scientists developed the next type of AI from this foundation.

Figure 3

Limited Memory AI :

Limited memory AI learns from the past and builds experiential knowledge by observing actions or data. This type of AI uses historical, observational data in combination with pre-programmed information to make predictions and perform complex classification tasks. It is the most widely-used kind of AI today.

For example, autonomous vehicles use limited memory AI to observe other cars' speed and direction, helping them "read the road" and adjust as needed. This process for understanding and interpreting incoming data makes them safer on the roads.

However, limited memory AI – as its name suggests – is still limited. The information that autonomous vehicles work with is fleeting, and it is not saved in the car's long-term memory.

Theory of Mind AI :

Want to hold a meaningful conversation with an emotionally intelligent robot that looks and sounds like a real human being? That's on the horizon with theory of mind AI.

With this type of AI, machines will acquire true decision-making capabilities that are similar to humans. Machines with theory of mind AI will be able to understand and remember emotions, then adjust behavior based on those emotions as they interact with people.

There are still a number of hurdles to achieving theory of mind AI, because the process of shifting behavior based on rapidly shifting emotions is so fluid in human communication. It is difficult to mimic as we try to create more and more emotionally intelligent machines.

That said, we are making progress. The Kismat robot head, developed by Professor Cynthia Breazeal, could recognize emotional signals on human faces and replicate those emotions on

its own face. Humanoid robot Sophia, developed by Hanson Robotics in Hong Kong, can recognize faces and respond to interactions with her own facial expressions.

Going Further with Artificial Intelligence :

Will we continue pushing the limits of AI, and develop a fifth type? How much progress will we make in the next decade toward theory of mind and self-aware AI? Maybe there will be a super-intelligent AI, that even surpasses the current intelligence of humans?

Only time will tell – but understanding the distinctions between the different types of AI will help you make sense of AI advancements as science continues to push the limits.

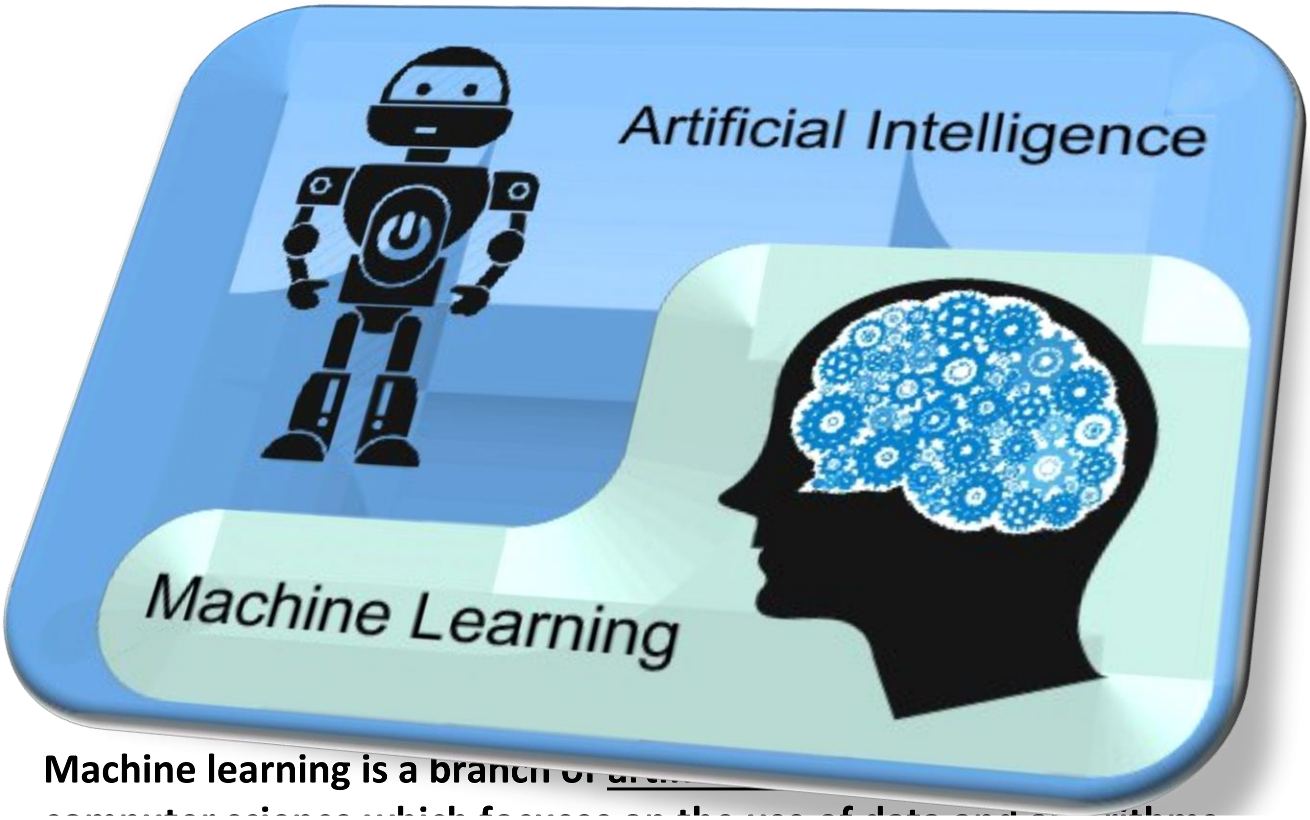


Figure 4

Machine learning is a branch of [illegible] computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy.

Over the last couple of decades, the technological advances in storage and processing power have enabled some innovative products based on machine learning, such as Netflix's recommendation engine and self-driving cars.

Machine learning is an important component of the growing field of data science. Through the use of statistical methods, algorithms are trained to make classifications or predictions, and to uncover key insights in data mining projects. These insights subsequently drive decision making within applications and businesses, ideally impacting key growth metrics. As big data continues to expand and grow, the market demand for data scientists will increase. They will be required to help identify the most relevant business questions and the data to answer them.

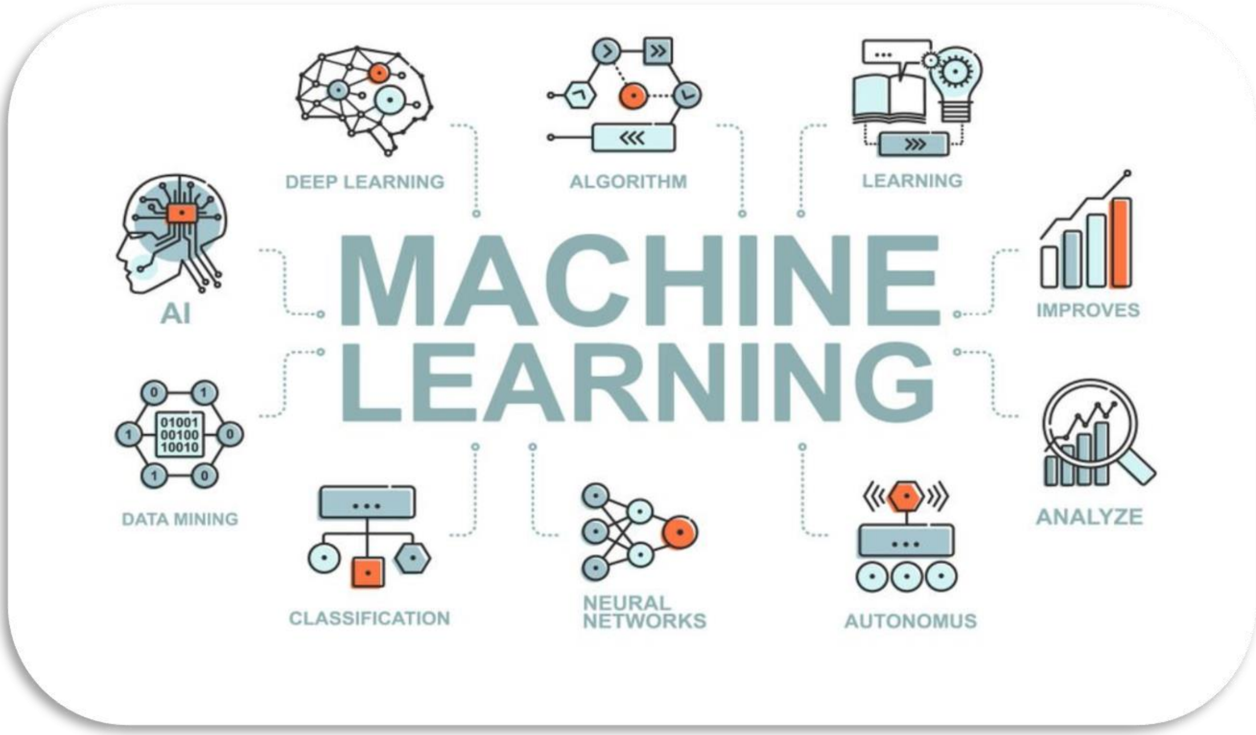


Figure 5

COMMON PRACTICAL ISSUES IN MACHINE LEARNING :

1) Lack Of Quality Data:

One of the main issues in Machine Learning is the absence of good data. While upgrading, algorithms tend to make developers exhaust most of their time on artificial intelligence. Data quality is fundamental for the algorithms to work as proposed. Incomplete data, unclean data, and noisy data are the quintessential foes of ideal ML. Different reasons for low data quality

2) FAULT IN CREDIT CARD FRAUD DETECTION :

Although this AI-driven software helps to successfully detect credit card fraud, there are issues in Machine Learning that make the process redundant. It is tough for the system to spot anything without adequate amounts of data, hence making them blind to any illegal connections. Hence detecting fraud without possessing a significant amount of data is close to impossible.

3) GETTING BAD RECOMMENDATIONS :

Proposal engines are quite regular today. While some might be dependable, others may not appear to provide the necessary results.

Machine Learning algorithms tend to only impose what these proposal engines have suggested. So if there is any modification in the necessity of the result, then the recommendation will be of no use.

Creating a complex algorithm, collecting large amounts of data, and implementing the algorithm, leading to nothing but incorrect results in case of changed priorities is one of the biggest issues with Machine Learning

4) TALENT DEFICIT :

Albeit numerous individuals are pulled into the ML business, however, there are still not many experts who can take complete control of this innovation. It is quite rare to find a trained professional who is capable of comprehending the problems in Machine Learning and being able to reach out to a reliable software solution for the same.

5) IMPLEMENTATION :

Organizations regularly have examination engines working with them when they decide to move up to ML. The usage of fresher ML strategies with existing procedures is a complicated errand. Keeping up legitimate documentation and interpretation needs to go a long way to facilitating maximum usage. There are issues in Machine Learning when it comes to implementation

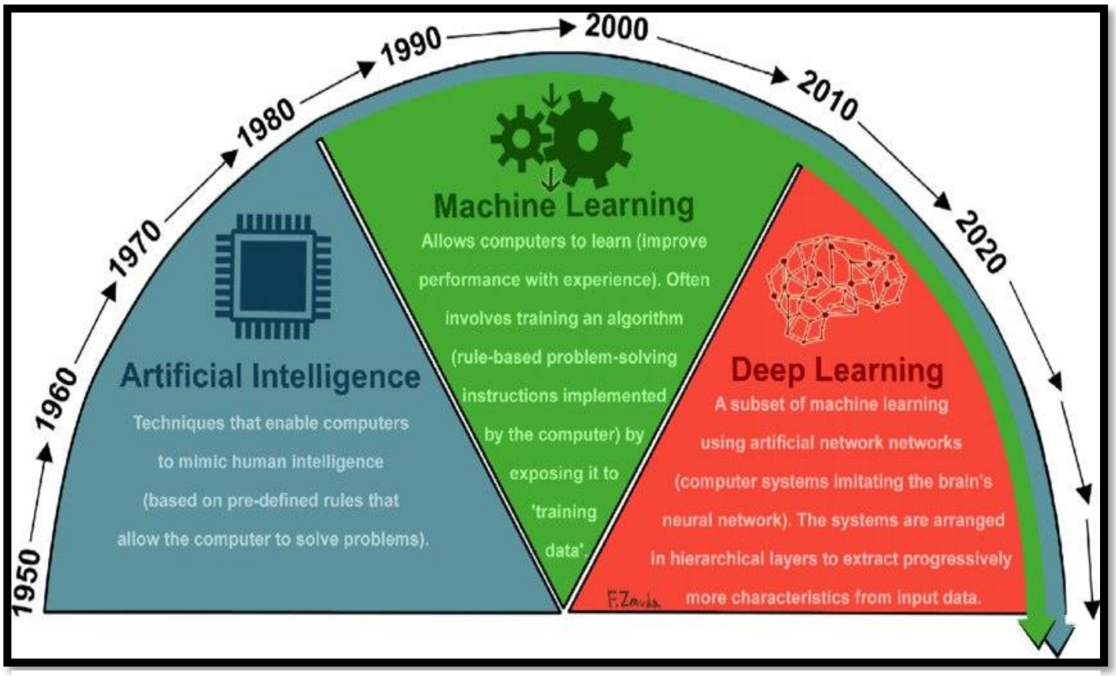


Figure 6

Figure 7

Deep Learning

Deep learning (also known as deep structured learning) is part of a broader family of machine learning methods based on artificial neural network with representation learning . Learning can be supervised, semi- supervised or unsupervised .

Deep-learning architectures such as deep neural networks , deep belief networks , deep reinforcement learning , recurrent neural networks, convolutional neural networks and Transformers have been applied to fields including computer vision , speech recognition , natural language processing , machine translation , bioinformatics , drug design , medical image analysis, climate science, material inspection and board game programs, where they have produced results comparable to and in some cases surpassing human expert performance.

Artificial neural networks (ANNs) were inspired by information processing and distributed communication nodes in biological systems. ANNs have various differences from biological brains. Specifically, artificial neural networks tend to be static and symbolic, while the biological brain of most living organisms is dynamic (plastic) and analogue.

The adjective "deep" in deep learning refers to the use of multiple layers in the network. Early work showed that a linear perceptron cannot be a universal classifier, but that a network with a nonpolynomial activation function with one hidden layer of unbounded width can. Deep learning is a modern variation which is concerned with an unbounded number of layers of bounded size, which permits practical application and optimized implementation, while retaining theoretical universality under mild conditions. In deep learning the layers are also permitted to be heterogeneous and to deviate widely from biologically

informed connectionist models, for the sake of efficiency, trainability and understandability, hence the "structured" part .

Neural networks

Artificial neural networks :

Artificial neural networks (ANNs) or connectionist systems are computing systems inspired by the biological neural networks that constitute animal brains. Such systems learn (progressively improve their ability) to do tasks by considering examples, generally without task-specific programming. For example, in image recognition, they might learn to identify images that contain cats by analyzing example images that have been manually labeled as "cat" or "no cat" and using the analytic results to identify cats in other images. They have found most use in applications difficult to express with a traditional computer algorithm using rule-based programming.

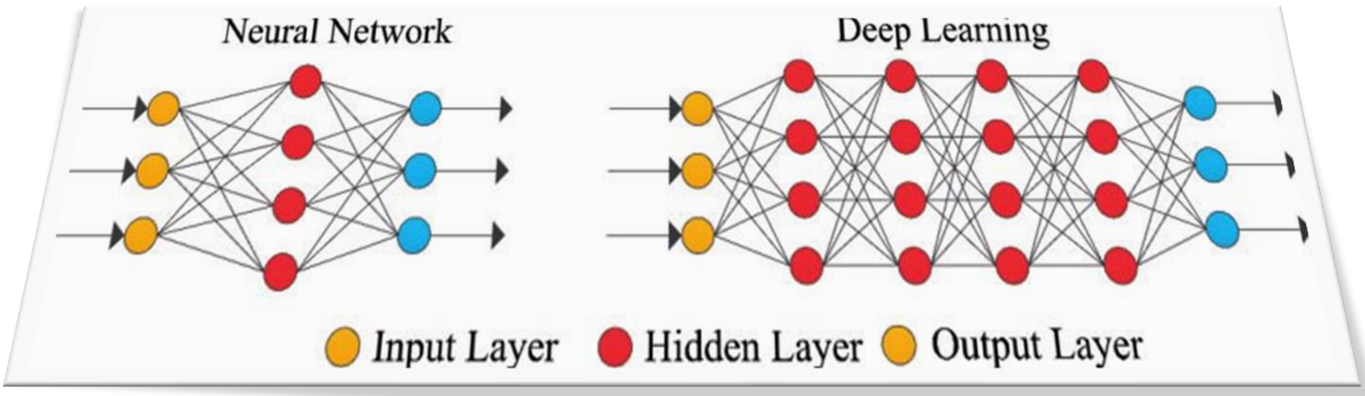
An ANN is based on a collection of connected units called artificial neurons, (analogous to biological neurons in a biological brain). Each connection (synapse) between neurons can transmit a signal to another neuron. The receiving (postsynaptic) neuron can process the signal(s) and then signal downstream neurons connected to it. Neurons may have state, generally represented by real numbers, typically between 0 and 1. Neurons and synapses may also have a weight that varies as learning proceeds, which can increase or decrease the strength of the signal that it sends downstream. Typically, neurons are organized in layers. Different layers may perform different kinds of transformations on their inputs.

Signals travel from the first (input), to the last (output) layer, possibly after traversing the layers multiple times.

The original goal of the neural network approach was to solve problems in the same way that a human brain would. Over time, attention focused on matching specific mental abilities, leading to deviations from biology such as backpropagation, or passing information in the reverse direction and adjusting the network to reflect that information.

Neural networks have been used on a variety of tasks, including computer vision, speech recognition, machine translation, social network filtering, playing board and video games and medical diagnosis.

As of 2017, neural networks typically have a few thousand to a few million units and millions of connections. Despite this number being several order of magnitude less than the number of neurons on a human brain, these networks can perform many tasks at a level beyond that of humans (e.g., recognizing faces, or playing "Go").



A deep neural network (DNN) is an artificial neural network

Figure 7

layers. There are different types of neural networks but they always consist of the same components: neurons, synapses, weights, biases, and functions. These components as a whole function similarly to a human brain, and can be trained like any other ML algorithm.

For example, a DNN that is trained to recognize dog breeds will go over the given image and calculate the probability that the dog in the image is a certain breed. The user can review the results and select which probabilities the network should display (above a certain threshold, etc.) and return the proposed label. Each mathematical manipulation as such is considered a layer,[citation needed] and complex DNN have many layers, hence the name "deep" networks.

DNNs can model complex non-linear relationships. DNN architectures generate compositional models where the object is expressed as a layered composition of primitives. The extra layers enable composition of features from lower layers, potentially modeling complex data with fewer units than a similarly performing shallow network. For instance, it was proved that sparse multivariate polynomials are exponentially easier to approximate with DNNs than with shallow networks.

Deep architectures include many variants of a few basic approaches. Each architecture has found success in specific domains. It is not always possible to compare the performance of multiple architectures, unless they have been evaluated on the same data sets.

DNNs are typically feedforward networks in which data flows from the input layer to the output layer without looping back. At first, the DNN creates a map of virtual neurons and assigns random numerical values, or "weights", to connections between them. The weights and inputs are multiplied and return an output between 0 and 1. If the network did not accurately recognize a particular pattern, an algorithm would adjust the weights. That way the algorithm can make certain parameters more influential, until it determines the correct mathematical manipulation to fully process the data.

Recurrent neural networks (RNNs), in which data can flow in any direction, are used for applications such as language modeling. Long short-term memory is particularly effective for this use.

Convolutional deep neural networks (CNNs) are used in computer vision. CNNs also have been applied to acoustic modeling for automatic speech recognition (ASR).

Hardware:

Since the 2010s, advances in both machine learning algorithms and computer hardware have led to more efficient methods for training deep neural networks that contain many layers of non-linear hidden units and a very large output layer. By 2019, graphic processing units (GPUs), often with AI-specific enhancements, had displaced CPUs as the dominant method of training large-scale commercial cloud AI. OpenAI estimated the hardware computation used in the largest deep learning projects from AlexNet (2012) to AlphaZero (2017), and found a 300,000-fold increase in the amount of computation required, with a doubling-time trendline of 3.4 months.

Special electronic circuits called deep learning processors were designed to speed up deep learning algorithms. Deep learning processors include neural processing units (NPUs) in Huawei cellphones and cloud computing servers such as tensor processing units (TPU) in the Google Cloud Platform. Cerebras Systems has also built a dedicated system to handle large deep learning models, the CS-2, based on the largest processor in the industry, the second-generation Wafer Scale Engine (WSE-2).

Atomically thin semiconductors are considered promising for energy-efficient deep learning hardware where the same basic device structure is used for both logic operations and

data storage. In 2020, Marega et al. published experiments with a large-area active channel material for developing logic-in-memory devices and circuits based on floating-gate field-effect transistors (FGFETs).

In 2021, J. Feldmann et al. proposed an integrated photonic hardware accelerator for parallel convolutional processing. The authors identify two key advantages of integrated photonics over its electronic counterparts: (1) massively parallel data transfer through wavelength division multiplexing in conjunction with frequency combs, and (2) extremely high data modulation speeds. Their system can execute trillions of multiply-accumulate operations per second, indicating the potential of integrated photonics in data-heavy AI applications.

Applications :

Automatic speech recognition :

Large-scale automatic speech recognition is the first and most convincing successful case of deep learning. LSTM RNNs can learn "Very Deep Learning" tasks that involve multi-second intervals containing speech events separated by thousands of discrete time steps, where one time step corresponds to about 10 ms. LSTM with forget gates is competitive with traditional speech recognizers on certain tasks.

The initial success in speech recognition was based on small-scale recognition tasks based on TIMIT. The data set contains 630 speakers from eight major dialects of American English, where each speaker reads 10 sentences. Its small size lets many configurations be tried. More importantly, the TIMIT task concerns phone-sequence recognition, which, unlike word-sequence recognition, allows weak phone bigram language models. This lets the strength of the acoustic modeling aspects of speech recognition be more easily analyzed. The error rates listed below, including these early results and measured as percent phone error rates (PER), have been summarized since 1991

The debut of DNNs for speaker recognition in the late 1990s and speech recognition around 2009-2011 and of LSTM around 2003–2007, accelerated progress in eight major areas:

Scale-up/out and accelerated DNN training and decoding

Sequence discriminative training

Feature processing by deep models with solid understanding of the underlying mechanisms

Adaptation of DNNs and related deep models

Multi-task and transfer learning by DNNs and related deep models

CNNs and how to design them to best exploit domain knowledge of speech

RNN and its rich LSTM variants

Other types of deep models including tensor-based models and integrated deep generative/discriminative models.

All major commercial speech recognition systems (e.g., Microsoft Cortana, Xbox, Skype Translator, Amazon Alexa, Google Now, Apple Siri, Baidu and iFlyTek voice search, and a range of Nuance speech products, etc.) are based on deep learning.

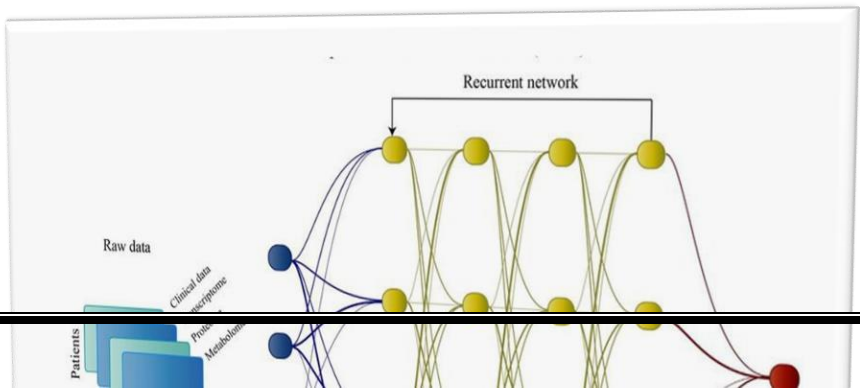


Figure 8

Image recognition :

Computer vision :

A common evaluation set for image classification is the MNIST database data set. MNIST is composed of handwritten digits and includes 60,000 training examples and 10,000 test examples. As with TIMIT, its small size lets users test multiple configurations. A comprehensive list of results on this set is available.

Deep learning-based image recognition has become "superhuman", producing more accurate results than human contestants. This first occurred in 2011 in recognition of traffic signs, and in 2014, with recognition of human faces.

Deep learning-trained vehicles now interpret 360° camera views. Another example is Facial Dysmorphology Novel

Analysis (FDNA) used to analyze cases of human malformation connected to a large database of genetic syndromes.

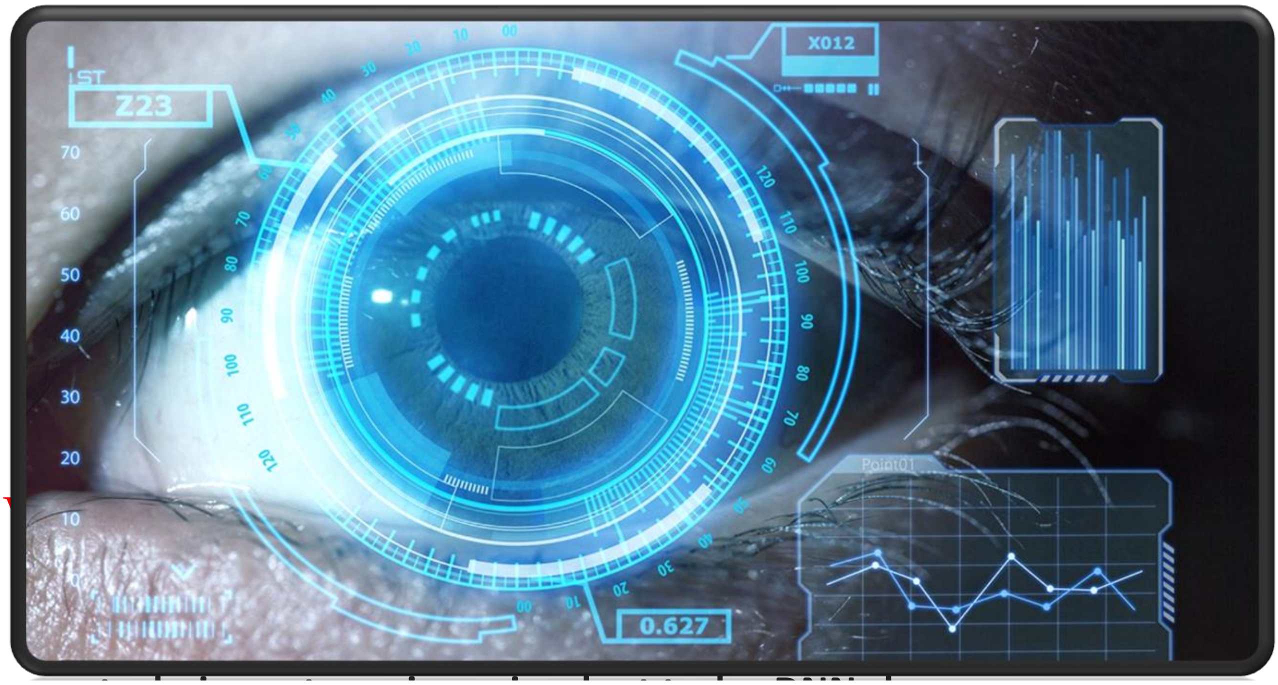


Figure 9

themselves capable, for example of identifying the style period of a given painting Neural Style Transfer – capturing the style of a given artwork and applying it in a visually pleasing manner to an arbitrary photograph or video generating striking imagery based on random visual input fields.

Natural language processing :

Neural networks have been used for implementing language models since the early 2000s. LSTM helped to improve machine translation and language modeling.

Other key techniques in this field are negative sampling and word embedding. Word embedding, such as word2vec, can be thought of as a representational layer in a deep learning architecture that transforms an atomic word into a positional representation of the word relative to other words in the

dataset; the position is represented as a point in a vector space. Using word embedding as an RNN input layer allows the network to parse sentences and phrases using an effective compositional vector grammar. A compositional vector grammar can be thought of as probabilistic context free grammar (PCFG) implemented by an RNN. Recursive auto-encoders built atop word embeddings can assess sentence similarity and detect paraphrasing. Deep neural architectures provide the best results for constituency parsing, sentiment analysis, information retrieval, spoken language understanding, machine translation, contextual entity linking, writing style recognition, Text classification and others. Recent developments generalize word embedding to sentence embedding.

Google Translate (GT) uses a large end-to-end long short-term memory (LSTM) network. Google Neural Machine Translation (GNMT) uses an example-based machine translation method in which the system "learns from millions of examples." It translates "whole sentences at a time, rather than pieces. Google Translate supports over one hundred languages. The network encodes the "semantics of the sentence rather than simply memorizing phrase-to-phrase translations". GT uses English as an intermediate between most language pairs.

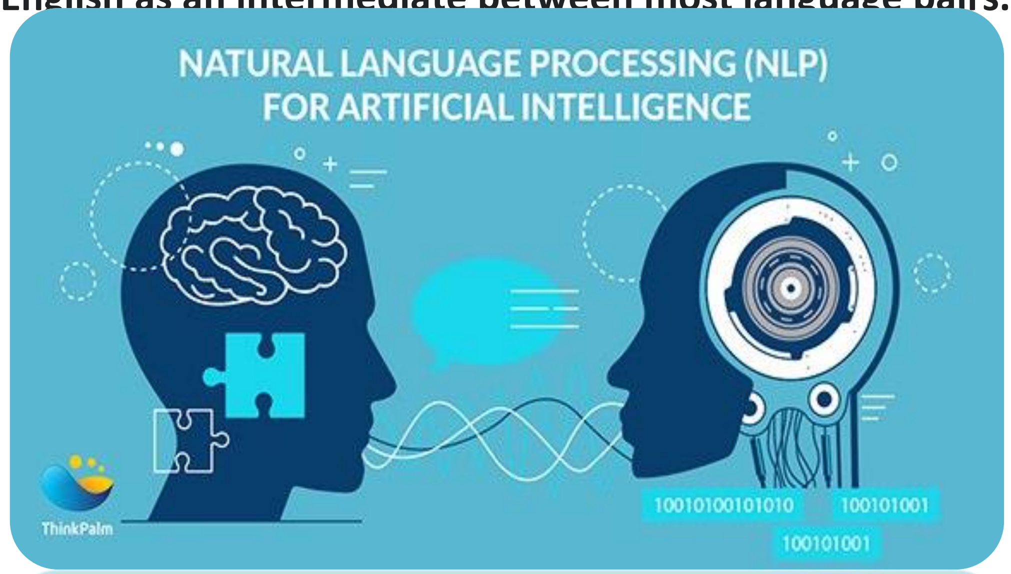


Figure 10

Introduction to Convolutional Neural Networks (CNN)

In the past few decades, Deep Learning has proved to be a very powerful tool because of its ability to handle large amounts of data. The interest to use hidden layers has surpassed traditional techniques, especially in pattern recognition. One of the most popular deep neural networks is Convolutional Neural Networks.

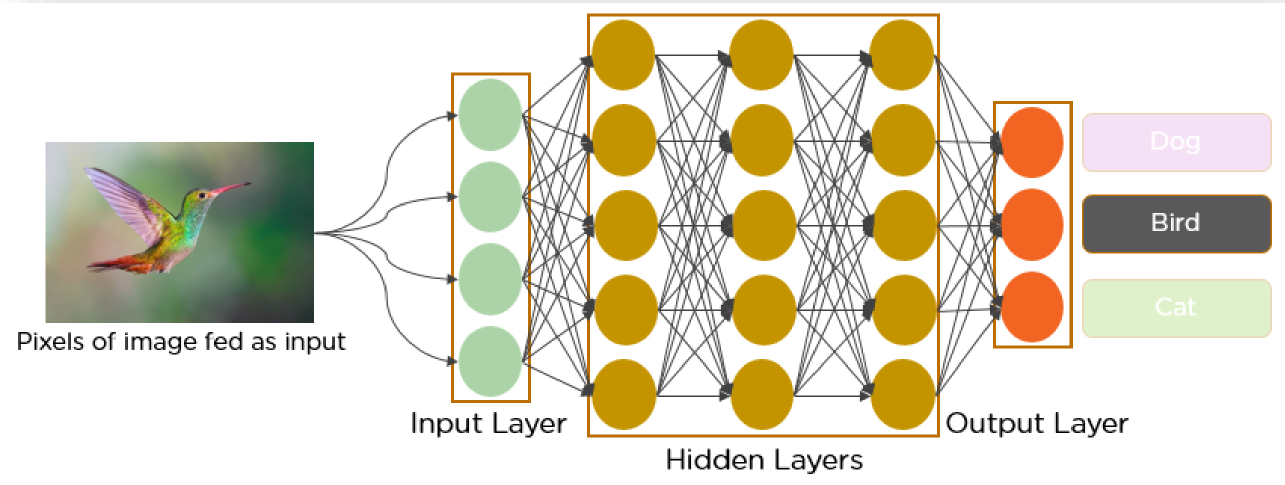


Figure 11

computer vision took a quantum leap when a group of researchers from the University of Toronto developed an AI model that surpassed the best image recognition algorithms and that too by a large margin.

The AI system, which became known as AlexNet (named after its main creator, Alex Krizhevsky), won the 2012 ImageNet computer vision contest with an amazing 85 percent accuracy. The runner-up scored a modest 74 percent on the test.

At the heart of AlexNet was Convolutional Neural Networks a special type of neural network that roughly imitates human vision. Over the years CNNs have become a very important part of many Computer Vision applications and hence a part of any computer vision course online. So let's take a look at the workings of CNNs.

Background of CNNs

CNN's were first developed and used around the 1980s. The most that a CNN could do at that time was recognize handwritten digits. It was mostly used in the postal sectors to read zip codes, pin codes, etc. The important thing to remember about any deep learning model is that it requires a large amount of data to train and also requires a lot of computing resources. This was a major drawback for CNNs at that period and hence CNNs were only limited to the postal sectors and it failed to enter the world of machine learning.

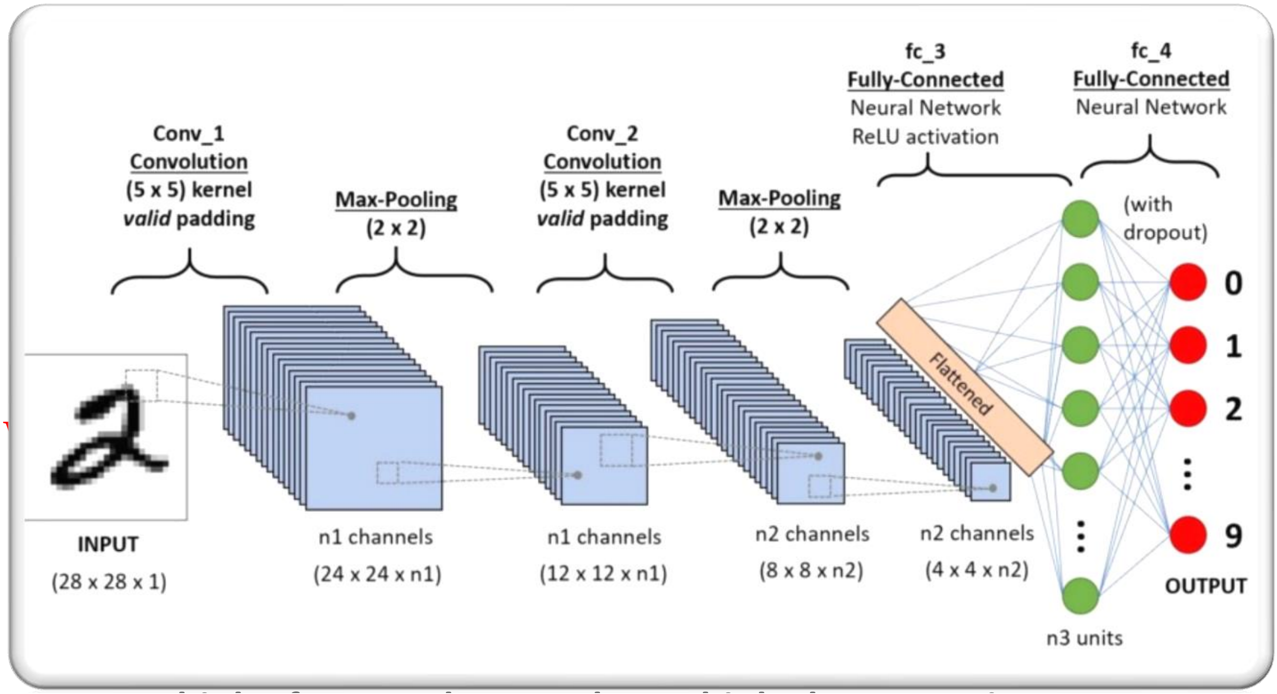


Figure 12

multiplications but that is not the case with ConvNet. It uses a special technique called Convolution. Now in mathematics convolution is a mathematical operation on two functions that produces a third function that expresses how the shape of one is modified by the other.

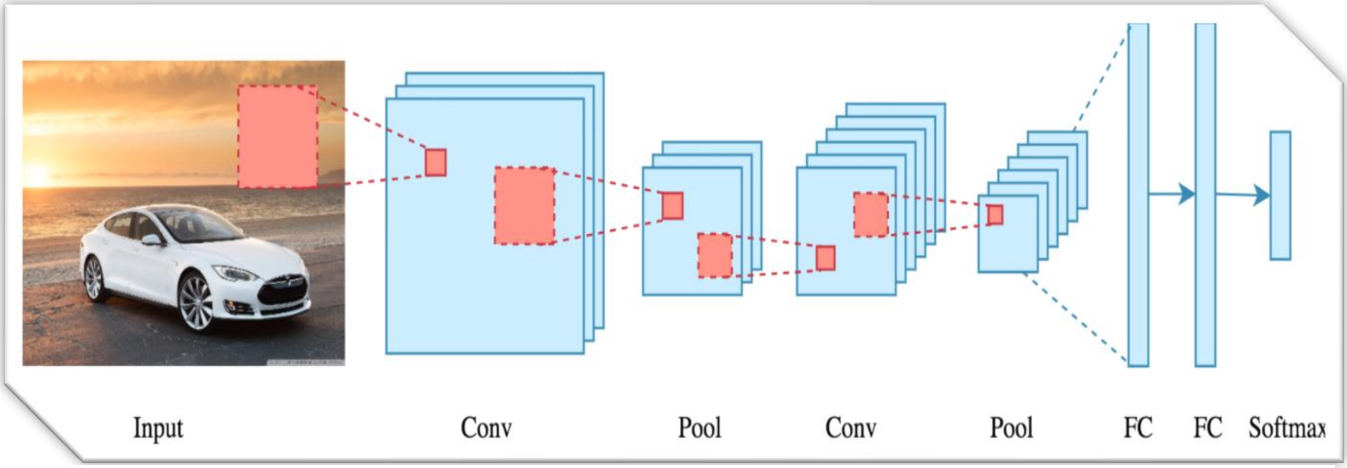


Figure 13

But we don’t really need to go behind the mathematics part to understand what a CNN is or how it works.

Bottom line is that the role of the ConvNet is to reduce the images into a form that is easier to process, without losing features that are critical for getting a good prediction.

How does it work?

Before we go to the working of CNN’s let’s cover the basics such as what is an image and how is it represented. An RGB image is nothing but a matrix of pixel values having three planes whereas a grayscale image is the same but it has a single plane. Take a look at this image to understand more.

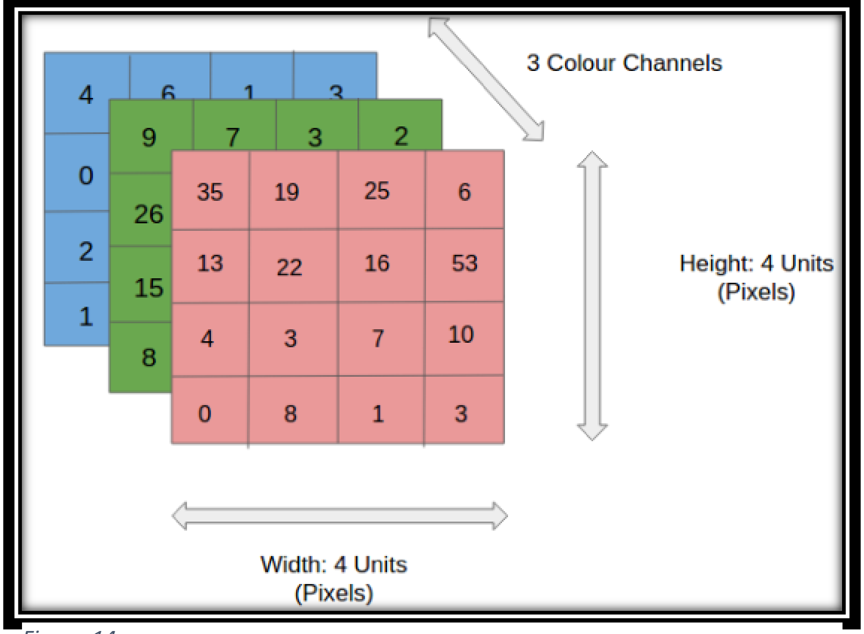


Figure 14

For simplicity, let's stick with grayscale images as we try to understand how CNNs work.

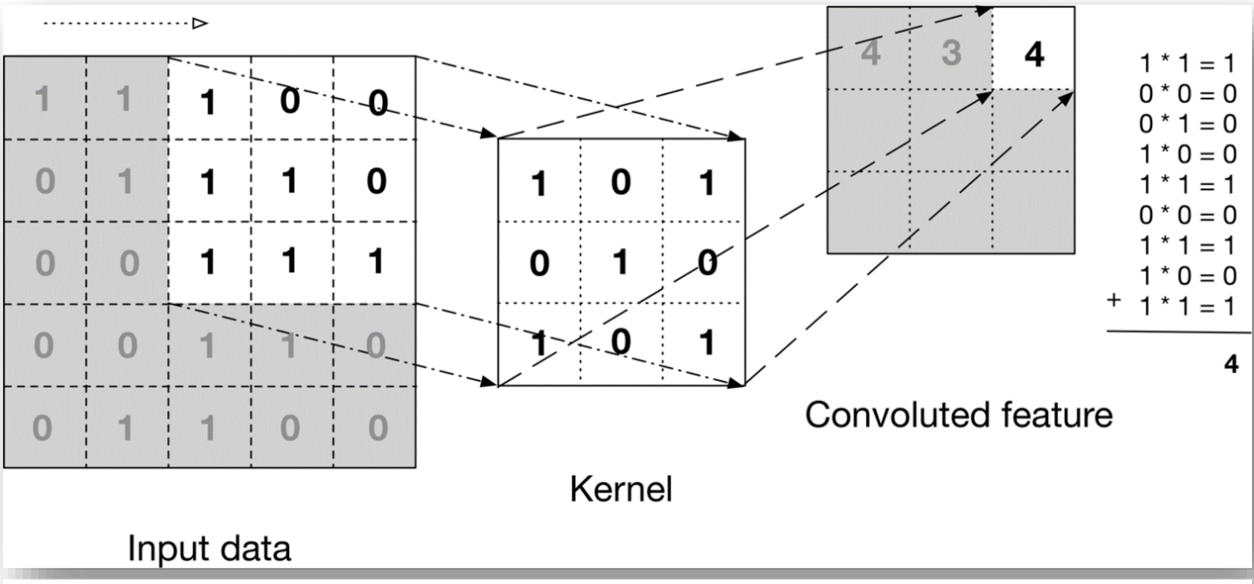


Figure 15

The above image shows what a convolution is. We take a filter/kernel(3×3 matrix) and apply it to the input image to get the convolved feature. This convolved feature is passed on to the next layer.

Convolutional neural networks are composed of multiple layers of artificial neurons. Artificial neurons, a rough imitation of their biological counterparts, are mathematical functions that calculate the weighted sum of multiple inputs and outputs an activation value. When you input an image in a ConvNet, each layer generates several activation functions that are passed on to the next layer.

The first layer usually extracts basic features such as horizontal or diagonal edges. This output is passed on to the next layer which detects more complex features such as corners or combinational edges. As we move deeper into the network it can identify even more complex features such as objects, faces, etc

Introduction to YOLO Algorithm for Object Detection

YOLO is an algorithm that uses neural networks to provide real-time object detection. This algorithm is popular because of its speed and accuracy. It has been used in various applications to detect traffic signals, people, parking meters, and animals.

This article introduces readers to the YOLO algorithm for object detection and explains how it works. It also highlights some of its real-life applications

Introduction to object detection :

Object detection is a phenomenon in computer vision that involves the detection of various objects in digital images or videos. Some of the objects detected include people, cars, chairs, stones, buildings, and animals.

This phenomenon seeks to answer two basic questions:

1. ***What is the object?* This question seeks to identify the object in a specific image.**
2. ***Where is it?* This question seeks to establish the exact location of the object within the image.**

Object detection consists of various approaches such as fast R-CNN, Retina-Net, and Single-Shot MultiBox Detector (SSD). Although these approaches have solved the challenges of data limitation and modeling in object detection, they are not able to detect objects in a single algorithm run.

YOLO algorithm has gained popularity because of its superior performance over the aforementioned object detection techniques.

How does object detection work ?

To explore the concept of object detection it's useful to begin with image classification. Image classification goes through levels of incremental complexity.

1. **Image classification**
   aims at assigning an image to one of a number of different categories (e.g. car, dog, cat, human, etc.), essentially answering the question *"What is in this picture?"*. One image has only one category assigned to it.
2. **Object localization**
   allows us to locate our object in the image, so our question changes to *"What is it and where it is?"*.
3. **Object detection**
   provides the tools for doing just that – finding all the objects in an image and drawing the so-called bounding boxes around them.

In a real real-life scenario, we need to go beyond locating just one object but rather multiple objects in one image. For example, a self-driving car has to find the location of other cars, traffic lights, signs, and humans and take appropriate action based on this information.

In the case of bounding boxes, there are also some situations where we want to find the exact boundaries of our objects. This process is called instance segmentation, but this is a topic for another post.

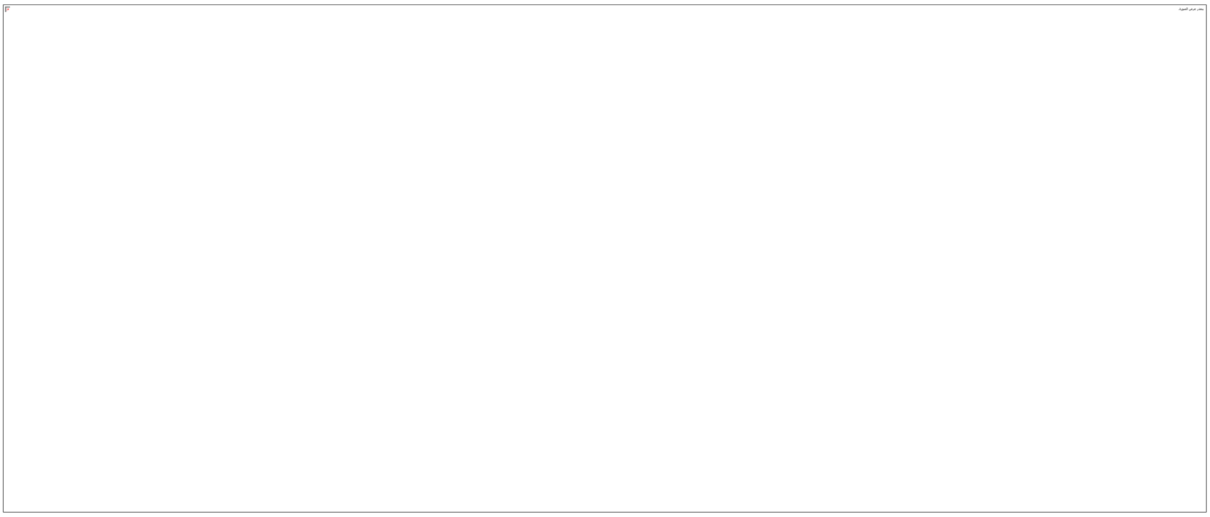


Figure 16

What is YOLO ?

YOLO is an abbreviation for the term 'You Only Look Once'. This is an algorithm that detects and recognizes various objects in a picture (in real-time). Object detection in YOLO is done as a regression problem and provides the class probabilities of the detected images.

YOLO algorithm employs convolutional neural networks (CNN) to detect objects in real-time. As the name suggests, the algorithm requires only a single forward propagation through a neural network to detect objects.

This means that prediction in the entire image is done in a single algorithm run. The CNN is used to predict various class probabilities and bounding boxes simultaneously.

The YOLO algorithm consists of various variants. Some of the common ones include tiny YOLO and YOLOv3

Why the YOLO algorithm is important :

YOLO algorithm is important because of the following reasons:

- **Speed:** **This algorithm improves the speed of detection because it can predict objects in real-time.**
- **High accuracy:** **YOLO is a predictive technique that provides accurate results with minimal background errors.**
- **Learning capabilities:** **The algorithm has excellent learning capabilities that enable it to learn the representations of objects and apply them in object detection.**

How the YOLO algorithm works :

YOLO algorithm works using the following three techniques:

- **Residual blocks**
- **Bounding box regression**
- **Intersection Over Union (IOU)**

Residual blocks

First, the image is divided into various grids. Each grid has a dimension of S x S. The following image shows how an input image is divided into grids

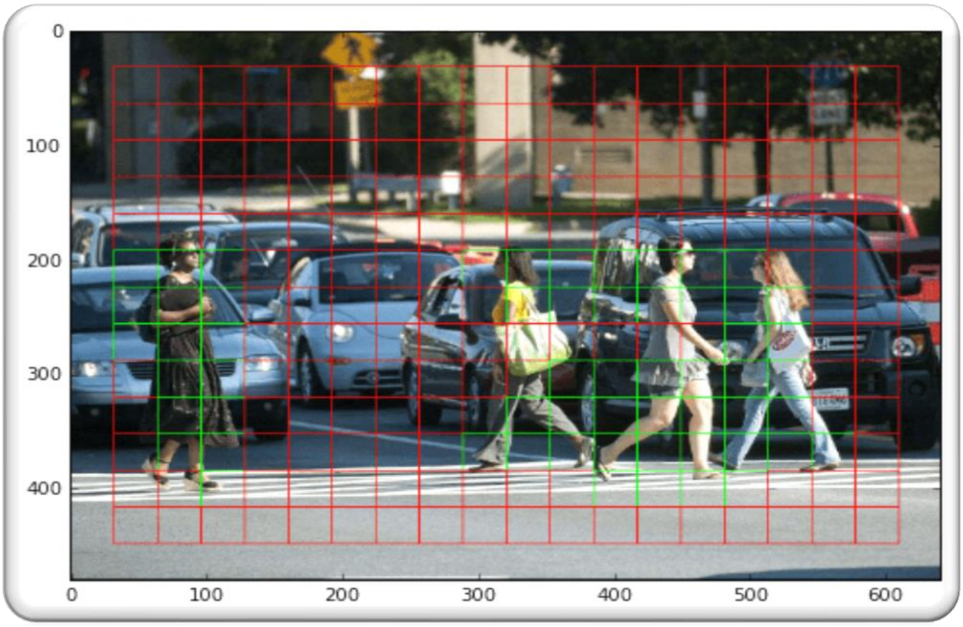


Figure 17

In the image above, there are many grid cells of equal dimension. Every grid cell will detect objects that appear within them. For example, if an object center appears within a certain grid cell, then this cell will be responsible for detecting it.

Bounding box regression

A bounding box is an outline that highlights an object in an image.

Every bounding box in the image consists of the following attributes:

- Width (bw)
- Height (bh)
- Class (for example, person, car, traffic light, etc.)- This is represented by the letter c.
- Bounding box center (bx,by)

The following image shows an example of a bounding box. The bounding box has been represented by a yellow outline.

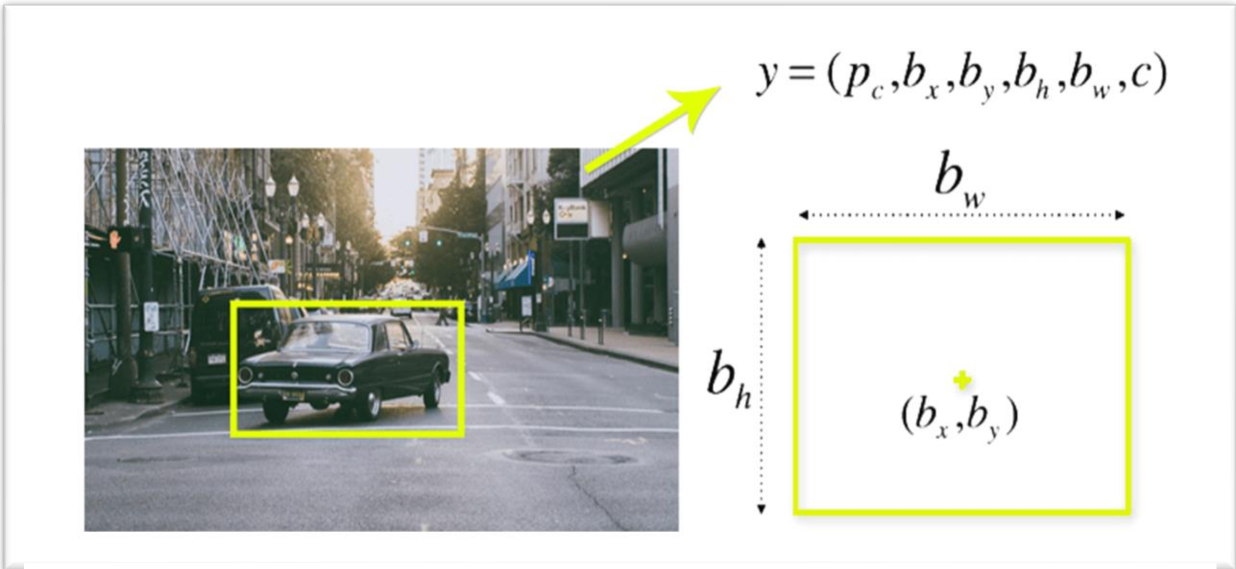


Figure 18

YOLO uses a single bounding box regression to predict the height, width, center, and class of objects. In the image above, represents the probability of an object appearing in the bounding box.

Intersection over union (IOU)

Intersection over union (IOU) is a phenomenon in object detection that describes how boxes overlap. YOLO uses IOU to provide an output box that surrounds the objects perfectly.

Each grid cell is responsible for predicting the bounding boxes and their confidence scores. The IOU is equal to 1 if the predicted bounding box is the same as the real box. This mechanism eliminates bounding boxes that are not equal to the real box.

The following image provides a simple example of how IOU works.

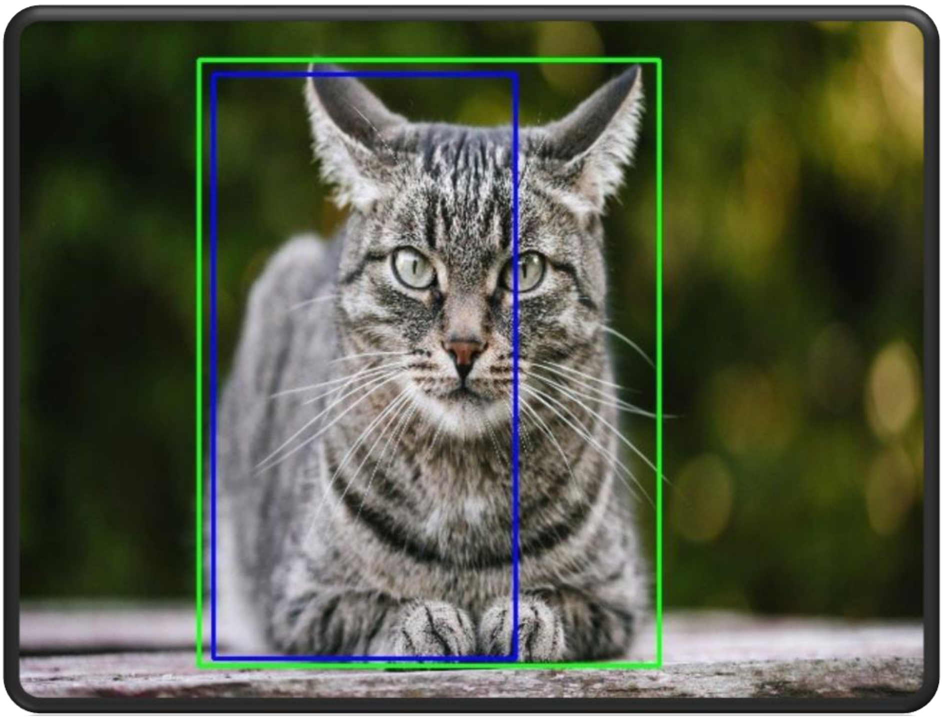


Figure 19

Object Detection is a deep learning algorithm use convolutional neural network (CNN) that can detect objects using computer vision with package cv2 that makes computer see as same as human and make action depend on that.

We choose Yolo version 3 as one of the most algorithm pioneer in this field .

Yolo is short for 'You Only Look Once' You can use Yolo version 3 in real time to detect object in video .

Object Detection systems are used by artificial intelligence programs to perceive specific objects in a class as subjects of interest .

The system sort objects in images into group where objects with similar characteristics are placed together , while others neglected unless programmed to do other wise.

Q) but why named 'you only look once' ?

A) because prediction is based on a convolutional layer that uses 1 x 1 convolutions , that means the size of prediction map is exactly the size of feature map before it

Yolo version 3 can detect 80 objects in file is called coco.names and other two files to algorithm it self named weight and config .

Figure 20

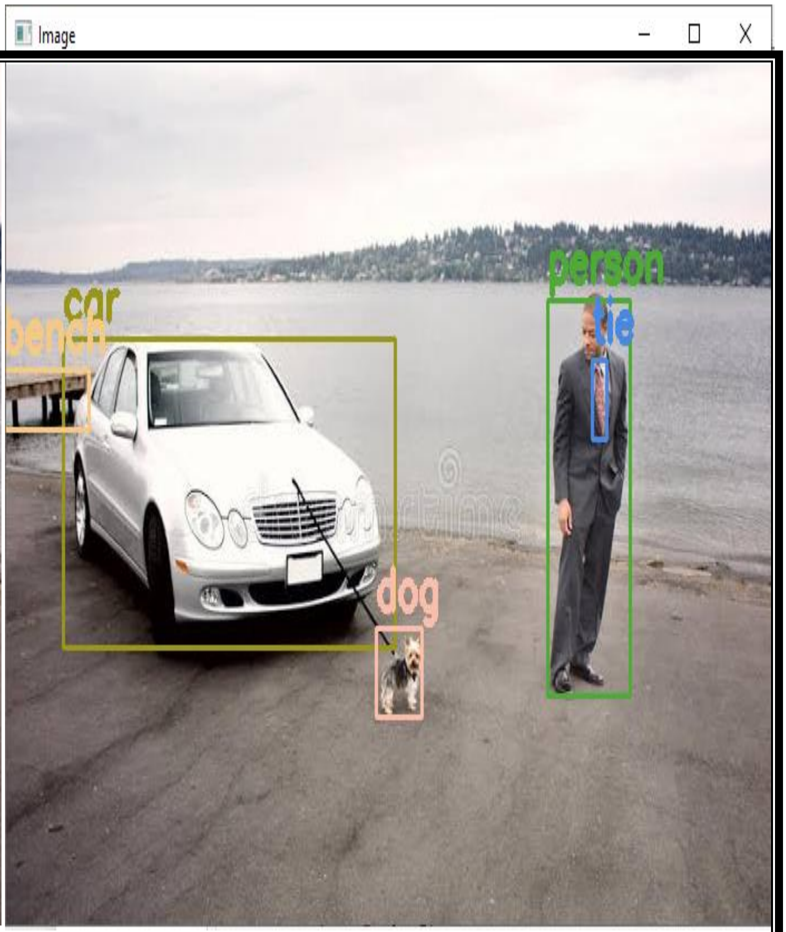


Figure 2111

Before Prediction **After Prediction**

We suffer some problems in our model and we think with solution :

- **First problem that we face is model can't predict actually object position**
- **We solve it and make sure that rectangle in the right position**
- **Second problem that we face**
- **is the model suffer delay when object moves**
- **We solve it by assign to confidence threshold equal 0.5 and nms Threshold 0.2**

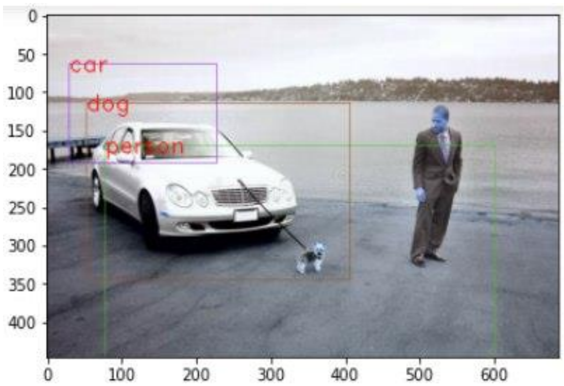


Figure 22

Space environment

Space environment is a branch of astronautics, aerospace engineering and space physics that seeks to understand and address conditions existing in space that affect the design and operation of spacecraft. A related subject, space weather, deals with dynamic processes in the solar-terrestrial system that can give rise to effects on spacecraft, but that can also affect the atmosphere, ionosphere and geomagnetic field, giving rise to several other kinds of effects on human technologies.

Effects on spacecraft can arise from radiation, space debris and meteoroid impact, upper atmospheric drag and spacecraft electrostatic charging. Radiation in space usually comes from three main sources:

The Van Allen radiation belts

Solar proton events and solar energetic particles; and

Galactic cosmic rays.

For long-duration missions, the high doses of radiation can damage electronic components and solar cells. A major concern is also radiation-induced "single-event effects" such as single event upset. Crewed missions usually avoid the radiation belts and the International Space Station is at an altitude well below the most severe regions of the radiation belts.

During solar energetic events (solar flares and coronal mass ejections)

particles can be accelerated to very high energies and can reach the Earth in times as short as 30 minutes (but usually take some hours).

These particles are mainly protons and heavier ions that can cause radiation damage, disruption to logic circuits, and even hazards to astronauts. Crewed missions to return to the Moon or to travel to Mars will have to deal with the major problems presented by solar particle events to radiation safety, in addition to the important contribution to doses from the low-level background cosmic rays. In near-Earth orbits, the Earth's geomagnetic field screens spacecraft from a large part of these hazards - a process called geomagnetic shielding.

Space debris and meteoroids can impact spacecraft at high speeds, causing mechanical or electrical damage. The average speed of space debris is 10 km/s (22,000 mph; 36,000 km/h)

[1] while the average speed of meteoroids is much greater.

For example, the meteoroids associated with the Perseid meteor shower travel at an average speed of 58 km/s (130,000 mph; 210,000 km/h).

[2] Mechanical damage from debris impacts have been studied through space missions including LDEF, which had over 20,000 documented impacts through its 5.7-year mission.

[3] Electrical anomalies associated with impact events include ESA's Olympus spacecraft, which lost attitude control during the 1993 Perseid meteor shower.

[4] A similar event occurred with the Landsat 5 spacecraft

[5] during the 2009 Perseid meteor shower.

[6] Spacecraft electrostatic charging is caused by the hot plasma environment around the Earth.

The plasma

encountered in the region of the geostationary orbit becomes heated during geomagnetic substorms caused by disturbances in the solar wind. "Hot" electrons (with energies in the kilo-electron volt range) collect on surfaces of spacecraft and can establish electrostatic potentials of the order of kilovolts. As a result, discharges can occur and are known to be the source of many spacecraft anomalies.

Solutions devised by scientists and engineers include, but are not limited to, spacecraft shielding, special "hardening" of electronic systems, various collision detection systems. Evaluation of effects during spacecraft design includes application of various models of the environment, including radiation belt models, spacecraft-plasma interaction models and atmospheric models to predict drag effects encountered in lower orbits and during reentry.

The field often overlaps with the disciplines of astrophysics, atmospheric science, space physics, and geophysics, albeit usually with an emphasis on application.

The United States government maintains a Space Weather Prediction Center at Boulder, Colorado. The Space Weather Prediction Center (SWPC) is part of the National Oceanic and Atmospheric Administration (NOAA). SWPC is one of the National Weather Service's (NWS) National Centers for Environmental Prediction (NCEP).

Space weather effects on Earth can include ionospheric storms, temporary decreases in ozone densities, disruption to radio communication, to GPS signals and submarine positioning. Some scientists also theorize links between sunspot activity and ice ages.

Space Weather Impacts on satellite :

Space near Earth contains a hostile environment for spacecraft. Satellites in space are exposed to such hazards as single-event effects from cosmic rays, internal charging from Van Allen radiation belt electrons, and surface charging from energetic electrons in hot plasma injected into the inner magnetosphere during geomagnetic storms and sub storms. These geophysical phenomena are highly variable, and are collectively known as space weather. Problems associated with these hazards include loss of mission, subsystem failure, mission degradation, loss of data, phantom commands, spurious signals, and single-event effects (upsets, latch up, and burnout).

Here, we describe the physical phenomena and give numerous examples of their effects on communications satellites.

Sunspots are magnetic storms on the surface of the Sun.

The sunspots number is an indicator for of solar activity of the sun.

The regular increase or decrease in sunspots numbers over time, approximately 11-year cycle, is called the solar cycle or sunspot cycle.

More sunspots mean more solar activity.

The highest designated number of any given cycle is 7

years which is a solar maximum, While the lowest designated number is 4 years which is a solar minimum

Sunspot Number (SSN) :

On the surface of the Sun, we can often see sunspots: dark areas where the strong magnetic field inhibits the flow from the Sun's interior. Sunspots are cooler and therefore appear darker than their surrounding "active regions". The active regions around the cool sunspots are very bright and may remain active for a number of days or even weeks. The sunspots number is a useful way to describe solar activity. In order to account for the statistical appearance of sunspots; the relative

sunspot number {R=k(10g+f)}, where (k is a correction factor usually less than unity which depends on the sensitivity of the observing equipment), (g is the number of sunspots groups A group may includes one or more sunspots), and (f is the number of individual spots) Some days in solar minimum, it could be spotless

<u>Solar Radiation Index (F10.7):</u>

The solar flux on the radio wavelength of 10.7 cm (2800 MHz) is well correlated with X-ray, EUV, and UV fluxes. The flux on wavelength of 10.7 cm is measured daily with a reflector of 1.8 m diameter at the Algonquin Radio Observatory, near Ottawa at 17:00 UT. It corresponds to a radio emission line for iron and is normally reported in solar flux units where one solar flux unit is of $10^{-22} Wm^{-2} Hz^{-1}$. The 10.7 cm radio flux is known as F10.7

index which varies from a minimum near 65 (corresponding to sunspot number of zero at solar minimum) to a maximum of about 200, corresponding to a sunspot number of about 150 to 160. Because of the correlation with X-ray, EUV and UV fluxes, the F10.7 is one of the most commonly used indicators for solar activity. The F10.7 index displays similar variations as the sunspot number

Solar Wind Speed Index (Vx) :

The solar wind streams off of the Sun in all directions at speeds of about 400 km/s (about 1 million miles per hour). The source of the solar wind is the Sun's hot corona. The temperature of the corona is so high that the Sun's gravity cannot hold on to it. The solar wind is not uniform. Although it is always directed away from the Sun, it changes speed and carries magnetic clouds with it, interacting regions. The solar wind speed is high (800km/s) over coronal holes and low (300km/s) over streamers.. These wind speed variations buffet the Earth's magnetic field and can produce storms in the Earth's magnetosphere.

Solar Wind IMF Index (Bz) :

The magnetic field of our Sun doesn’t stay around the Sun itself. The solar wind carries it through the Solar System until it reaches the heliopause. The heliopause is a place where the solar wind comes to a stop and where it collides with the interstellar medium. Because the Sun turns around its axis (once in about 25 days) the interplanetary

magnetic field has a spiral shape which is called the Parker Spiral. During solar minimum, the magnetic field of the Sun looks similar to Earth’s magnetic field. Around solar maximum, when the Sun reaches its maximum activity, many sunspots are visible on the visible solar disk. These sunspots are filled with magnetism and large magnetic field lines which run

material along them. There are 4 components in the sun's magnetic field (Bt, Bx, By, and Bz).

How do we protect the satellite?

All satellites are full of electronic devices. With their help, operators control spacecrafts' functionality by changing the angle of solar cells, controlling the orbit, taking pictures, sending and receiving messages. It requires a lot of effort to design special coatings, which would protect these devices from solar radiation.

How do we protect the satellite from space radiation ?

The scientists suggested a new material for power electronics for use on orbital spacecrafts. It's a semiconductor based on gallium oxides and aluminum that is better at resisting space radiation.

Define mobile applications.

In the modern age of Information and communication system, people are habituated to use mobile and computer application.

But Mobile Application uses, and development is a new and rapidly growing sector.

There is a global positive impact of mobile application.

Using mobile application developed country are becoming facilitate and people, society of developing country are upgrading themselves and making a new type of IT infrastructure. Mobile applications are running on a small hand hold mobile device which is moveable, easy to use and accessible from anywhere and any place. Now a day, so many people are using mobile application to contact friends, browse internet, file content management, document creating and handling, entertainment etc.

From everywhere user can get facility of mobile application.

People can do many things of his daily life and business life.

Not only the mobile application has an impact for user but also it plays an important role in business.

Mobile applications **are consisting of software/set of programs that runs on a mobile device and perform certain tasks for the user.**

Mobile application is a new and fast developing Segment of the global Information and Communication Technology.

The mobile application has wide used for its vast functioning area like calling, messaging, browsing, chatting, social network communication, audio, video, game etc. In large number of mobile applications some are preinstalled in phone and others user can download from internet and install it in mobile phone.

the different mobile applications are run able in different managed platforms like iPhone, BlackBerry, Android, Symbian, windows.

also, some virtual machine such as Java/J2ME, BREW, Flash Light, Silverlight.

From the last few years every mobile company are making the smartphone and feature phone.

And increasing the computing power of those mobile phone, rapidly increase the smart mobile application.

Most of the people from developed countries including America and Europe can't imagine leaving home without mobile.

Not only the developed countries but also the developing countries the mobile application uses rate growth rapidly.

All of the smartphone and feature phone have smart environment and vast number of mobile applications.

Most of the applications are connect people to the world via Internet/ mobile broadband.

Mobile App Usage & Download Statistics

We kick things off with stats about usage trends and download habits. These numbers highlight what's going on in the mobile app market today and even speak to the psychology of mobile users.

A study by App Annie's found that the average person spent 4.8 hours a day on their mobile phone last year—up 30% from 2019. Social, photo and video apps like Facebook, TikTok and YouTube were by far the most popular mobile apps.

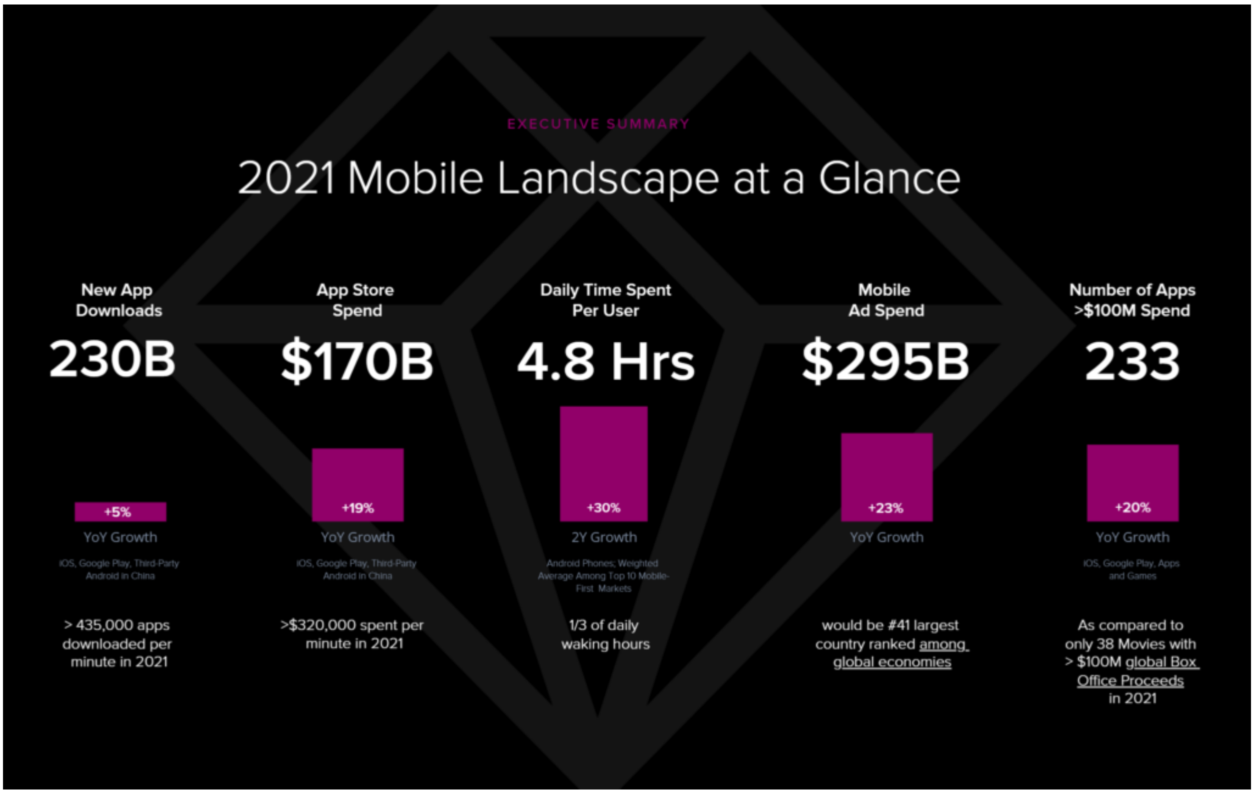


Figure 23

According to the 2019 Mary Meeker report, the number of hours spent on the internet continues to increase every

year, but the split between desktop and mobile is becoming more and more pronounced. In 2018, Americans were spending 3.6 hours per day on mobile

(12 times more than a decade prior) and just 2 hours per day on a desktop or laptop (which has stayed consistent since 2008).

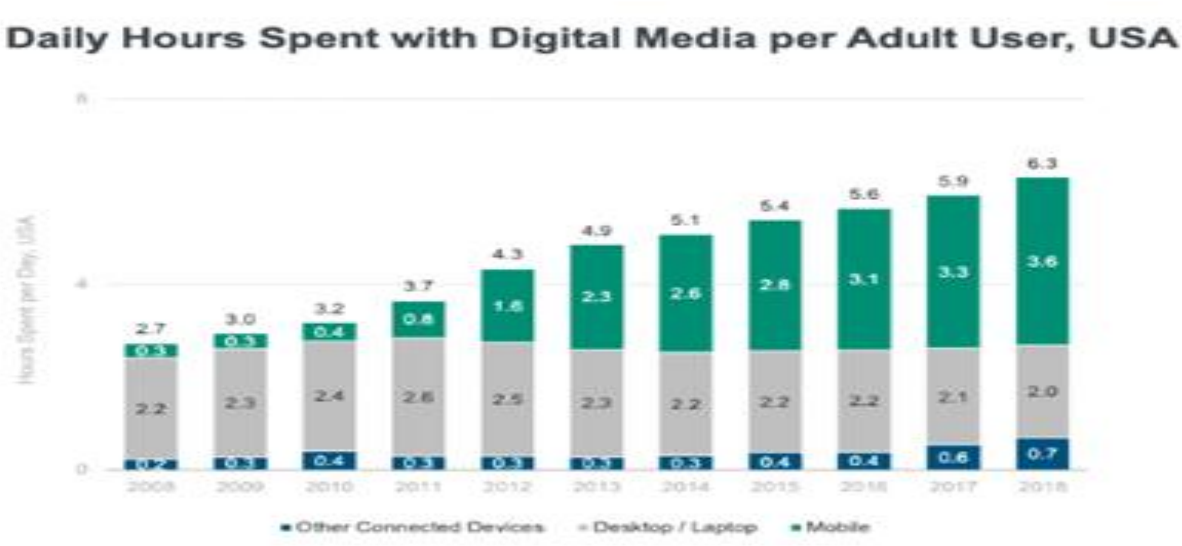


Figure 24

How often do you purge your phone of previously installed mobile apps? Many users do, primarily because those apps are not being used. The reason they aren't being used, though, is what businesses need to focus on. Among the other top reasons for uninstalling: not enough space on their phone, excessive advertising, and excessive notifications. As Clever Tap points out, all these responses speak to one thing: a poor user experience. Functionality is everything, and poor UX can guarantee the failure of your app.

Figure 25

Not only are social apps the most frequently downloaded, but they are also where smartphone users are spending the biggest chunk of their time (50% of total usage time, to be exact). In second place are video and entertainment apps, like Netflix and TikTok, coming in at 21% of total usage time. As App Annie reports, the lines are beginning to blur between social apps and entertainment apps as new

generations turn to them for similar purposes.

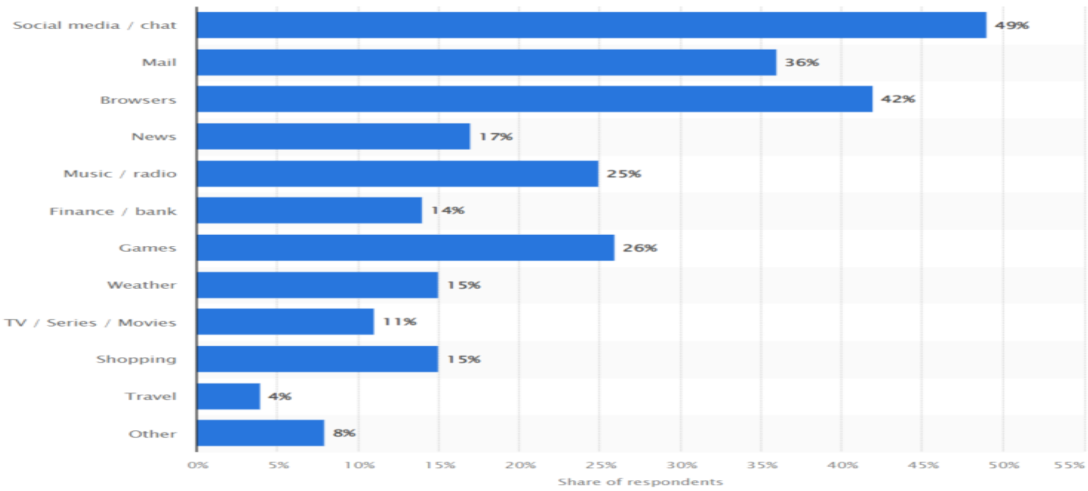
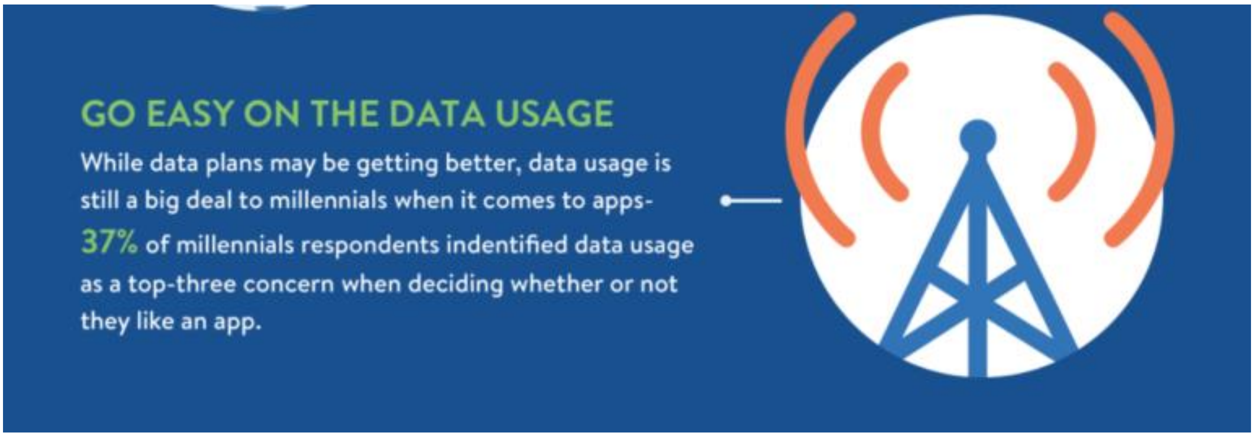


Figure 26

Millennials & Gen Z App Usage Statistics

Millennials and gen Z are the two most tech-savvy generations with rapidly growing purchasing power. If your app targets these groups, understanding what makes them tick will help you persuade them to adopt your app.

When asked about the top factors when deciding if they'll keep a mobile app, 37% of millennials said excessive data usage was a dealbreaker for them. A perfect example of this is Pokémon Go—when it first launched, the app quickly became a viral hit, but users were constantly frustrated with how much data it ate up, and many deleted it. As you're building your own mobile app, keep in mind that high data usage may result in users hitting the delete button fast.



As ComScore points out, smartphone users between the ages of 18 and 24 are the heaviest mobile app users

About EGSA Application:

From what we have seen previously, we know the importance of smart phone applications, their impact on users, statistics about the number of smart phone users, and the number of applications published on platforms such as Google Play,

It is also useful and important to take care of creating a smartphone application that is responsible for displaying the elements that the AI model does not recognize,

We will take you on a journey to learn about the role that artificial intelligence goes through before displaying the image on the mobile application,

First, the artificial intelligence identifies the elements that appear in front of it through a camera, then it identifies all the elements in front of the camera, and then determines the name of each of these elements, and when the artificial intelligence model finds an element that it cannot identify, and does not know what type of this element is nor His name, never seen before,

The artificial intelligence model sends commands to the camera to photograph this element and then sends it to a place called Firebase, which is responsible for storing data on the cloud, and then the phone application takes this picture and displays it on the application on a page through which the user of this application can identify this element And enter its name in the entry field for the elements, and then it sends it to the firebase and then gets the artificial intelligence model again and stores it in a special database of the model, and thus the model can identify the element if it sees it again, and if it finds this element again, this element will have previously defined,

This cycle is repeated when the AI model finds another component that it could not identify.

About EGSA Application Components:

On boarding screens:

Through the onboarding screens, the user will be able to look at what the Egyptian Space Agency is interested in,

Such as satellites and how they revolve around the earth, and that the satellites have missions that they accomplish when they are launched and revolve around the earth.

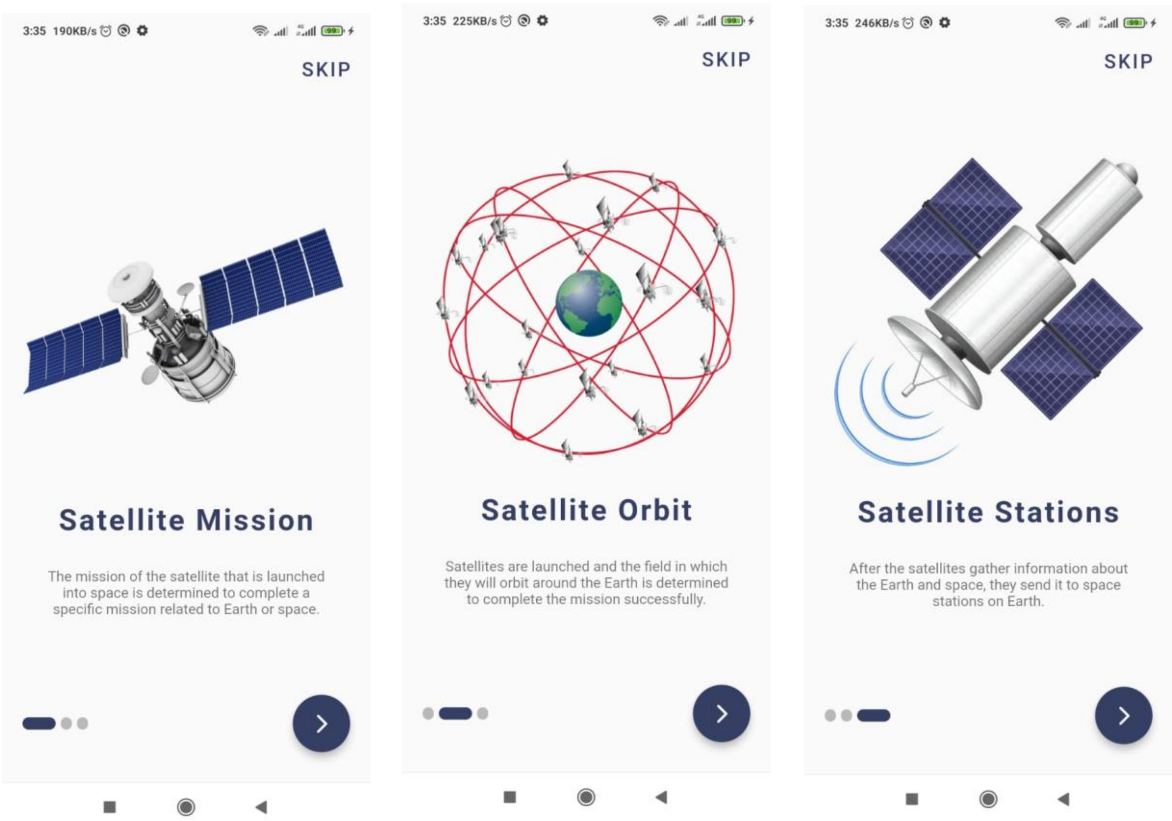


Figure 27

Login and Register:

The user can log into his account through the application to ensure the security of his account and data.

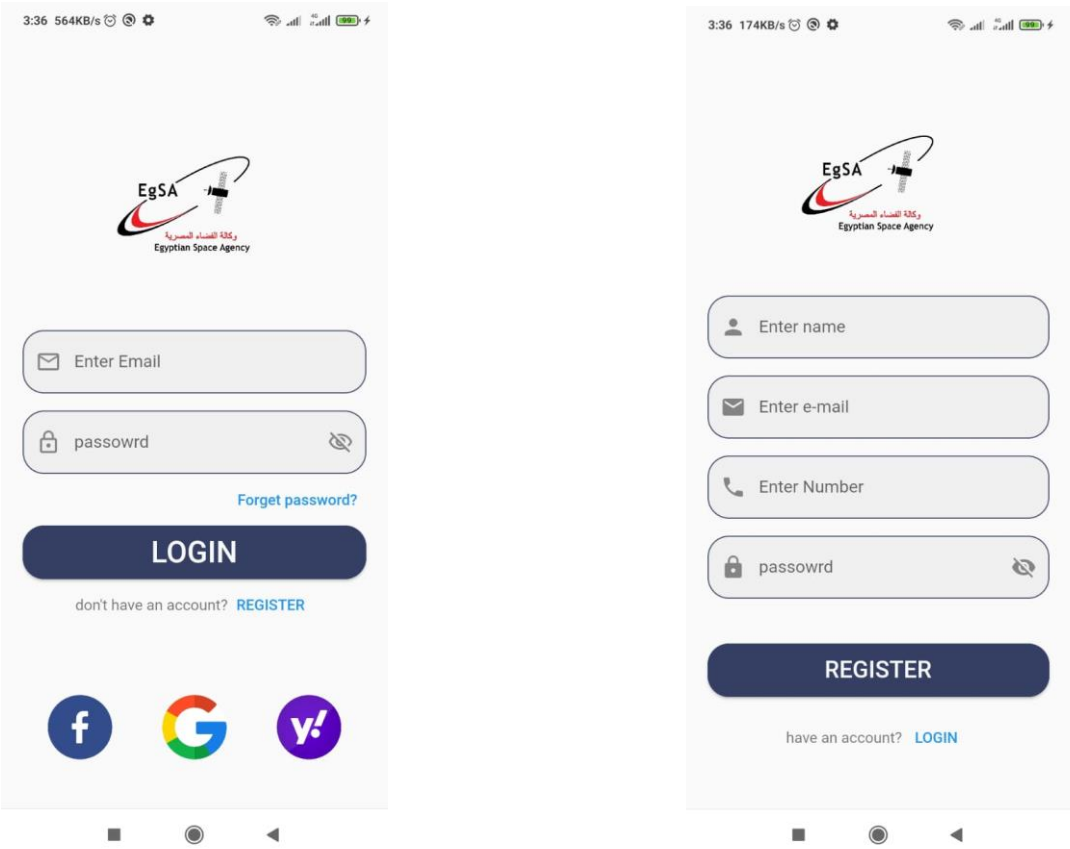
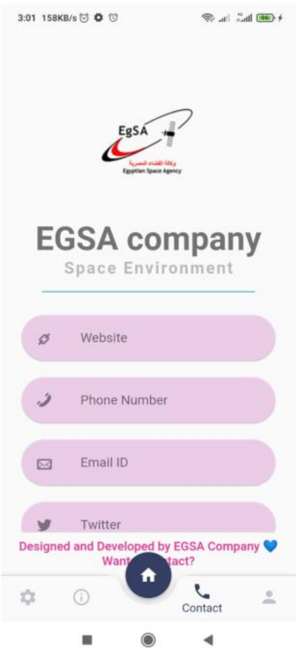
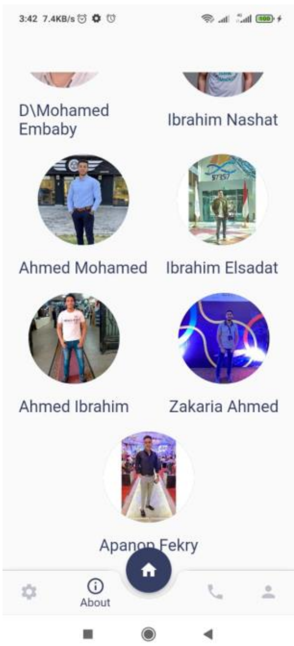


Figure 28

Main Pages:

These pages display the unknown photos taken by the camera, and show ways to communicate with the Egyptian Space Agency, and about team members.





Embedded system

In the beginning, the project was working to identify the object through the Raspberry Pi, Motor Servo and Camera.

But using the deep model wanted to work with GPU because it uses higher quality.

Raspberry Pi does not support the system that we want to achieve.

Because his capabilities are reduced

Because yolo model works very hard and wants to use gpu.

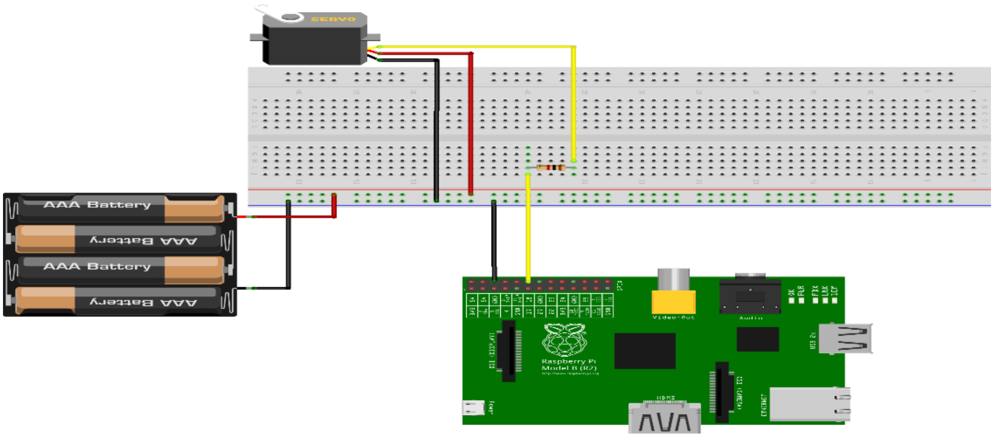


Figure 29

As shown in a way, is to connect the motor with the Raspberry Pi and take the decision to rotate when the model is given the order for it

This was canceled, as we explained, due to the lack of capabilities and the lack of a gpu.

Alternative idea:-

It is working on using socket programming.

Socket programming is a means of communicating data between two computers across a network. Connections can be made using either a connection-oriented protocol or a connectionless protocol.

In our case, we will use TCP/IP which is a connection-oriented protocol.

Before exchanging data, computers must establish a link that is for connection-oriented protocols. UDP (User Datagram Protocol) is the only option for connectionless protocol.

To demonstrate sockets further, we shall use the Client/Server architecture. Client and server communicate by writing to and reading from the socket connection.

Definition of a socket

A socket is a communication endpoint that serves as a link between two machines on a network.

It has a port number, which the TCP/IP layer can use to identify the application that receives the data. An endpoint usually includes a port number and an IP address.

What is TCP ?

Transmission Control Protocol (TCP) is a widely used protocol for data transmission on a network that supports client/server end points.

Two categories of Sockets:

1. **A server socket - It awaits a request from a client.**
2. **A client socket - It establishes communication between client and server**

The client has to know two things about the server:

1. **The server's IP address.**
2. **The port number.**

Creating a socket connection

In Java, we create a socket connection by doing the following steps:

The server constructs a ServerSocket object to specify the port number on which our conversation will occur. Exception handling methods are in use whenever an I/O error occurs.

The accept() method is called by the server to validate an incoming request to the socket.

A client then creates a Socket object by specifying the server name and the port number.

The Socket class constructor attempts to connect the client to the server using the provided port number.

If the connection is successful, the client and server can then communicate using I/O streams. The client and server socket classes are responsible for the I/O streams.

The client's OutputStream communicates with the server's InputStream, and the server's OutputStream communicates with the client's InputStream.

A stream is basically a collection of sequenced data.

How to use socket programming in the project?

It is through the link of the machine-running model. It uses the Python language with its link through the IP with the C code, so that the action is taken by the motor, so that it can take the movement or the correct direction that the model is following.

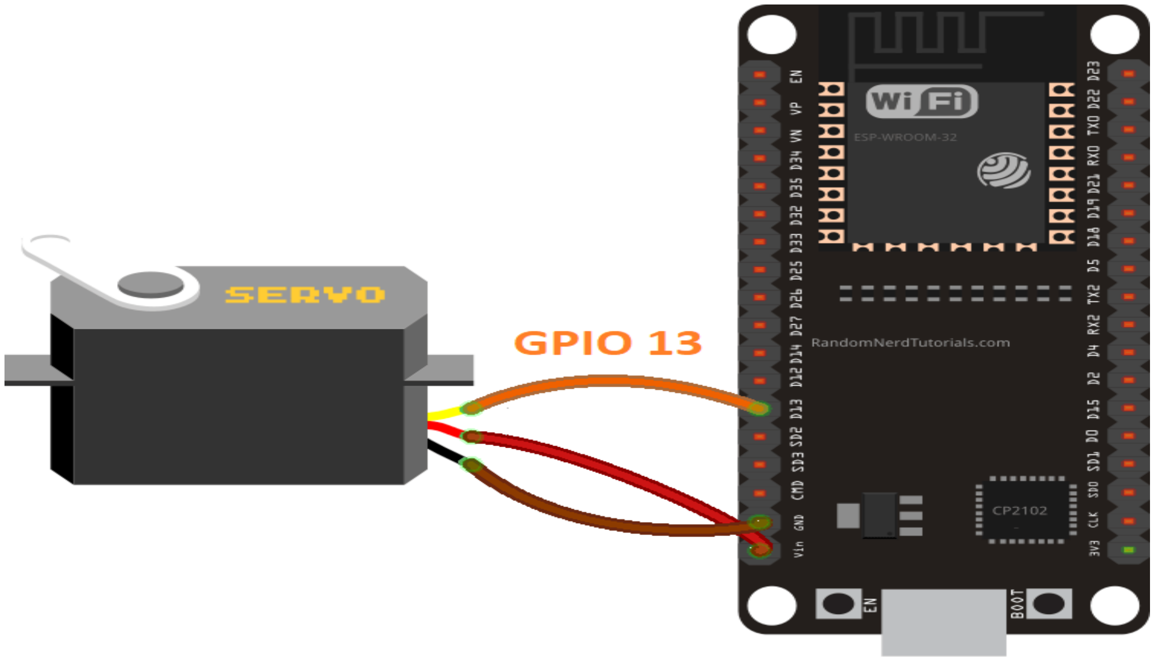


Figure 30

And both the device that runs the model and the ESP32 chip must be on the same network, so they call me on the same IP so that I can send an instantaneous reading.

And that the element of time plays a very big role in the movement's currency, and this is what we are researching and working on so that no mistake occurs between reading.

What is ESP32?

Created by Espressif Systems, ESP32 is a low-cost, low-power system on a chip (SoC) series with Wi-Fi & dual-mode Bluetooth capabilities! The ESP32 family includes the chips ESP32-D0WDQ6 (and ESP32-D0WD), ESP32-D2WD, ESP32-S0WD, and the system in package (SiP) ESP32-PICO-D4.

At its heart, there's a dual-core or single-core Tensilica Xtensa LX6 microprocessor with a clock rate of up to 240 MHz.

ESP32 is highly integrated with built-in antenna switches, RF balun, power amplifier, low-noise receive amplifier, filters, and power management modules.

Engineered for mobile devices, wearable electronics, and IoT applications, ESP32 achieves ultra-low power consumption through power saving features including fine resolution clock gating, multiple power modes, and dynamic power scaling.

ESP32 is created by Espressif Systems with a series of SoC (System on a Chip) and modules which are low cost with low power consumption.

This new ESP32 is the successor to the well-known ESP8266(became very popular with its inbuilt WiFi). ESP32 not only has Built in WiFi but also has Bluetooth and Bluetooth Low Energy. In other words we can define ESP32 as “ESP8266 on Steroids”.

ESP32 chip ESP32-D0WDQ6 is based on a Tensilica Xtensa LX6 dual core microprocessor with an operating frequency of up to 240 MHz.

The small ESP32 package has a high level of integrations such as:

- **Antenna switches**
- **Balun to control RF**
- **Power amplifier**
- **Low noise reception amplifier**
- **Filters and power management modules**

On top of all that, it achieves very low power consumption through power saving features including clock synchronization and multiple modes of operation. The ESP32 chip's quiescent current is less than 5 µA which makes it the ideal tool for your battery powered projects or IoT applications .

Difference between ESP32 and ESP8266

As already mentioned above that ESP32 **is the successor of ESP8266, Lets learn what are the differences between ESP32 and ESP8266 with their features and specifications below.**

- **ESP32 has an additional core compared to ESP8266**
- **Faster Wi-Fi**
- **Increased number of GPIO (input/output) pins**
- **Compatibility with Bluetooth 4.2 and Bluetooth low energy (low energy).**

Additionally, the ESP32 comes with touch-sensitive pins that can be used to “wake up” the ESP32 from deep sleep mode and a built-in Hall effect sensor.

Although both boards are extremely cheap, the ESP32 is slightly more expensive than ESP8266. ESP32 deserves it as it has more features than ESP8266.
We have differentiated the main technical specifications between the ESP8266 and ESP32 in the below table.

Wireless Connectivity:

The ESP32 SoC chip has WiFi connectivity, being compatible with 802.11 b / g / n in the 2.4GHz band, reaching speeds of up to 150 Mbits/s. It also includes Bluetooth communication compatible with Bluetooth v4.2 and Bluetooth Low Energy (BLE).

Figure 31

Project Results

In the beginning, the work was to turn on the camera that tracks anything that is not recognized through a servo motor connected to Raspberry Pi

But I have a big problem, which is that the user model is

Figure 3216

Figure 33

running very slowly

Because it needs a GPU and Raspberry Pi does not have this system.

And work on another, better and more accurate solution.

I use esp32, the python code connects to me via socket, it connects to the same network and sends the angle on which the object moves.

The angle of movement was taken by means of x, y that appear in the model and showed through it the angle that is sent to the motor to move.

Project installation steps:-

Figure 34

Figure 35

The stand was used to adjust the angle at which the servo moves the camera in order to avoid any error in photography.

A MG 996R servo motor was used in order to be quick in response and bear the weight of the camera during rotation, and it works to rotate up to a 360-degree angle.

Model Result

The model works to identify a lot of the object, which is through the verses or the readings that enter through it and when it is recognized.

It tracks the thing and sends a large number of pictures to the space station.

This puts it above the satellite in terms of trying to transform space and identifying many of the problems facing space that may affect the satellite.

Figure 36

The spy satellite that can see INSIDE your apartment and take 'crystal clear' pictures even through clouds

- **Capella Space's Capella-2 satellite is capable of 50cm x 50cm resolution imaging with its Spotlight mode**
- **This allows for long exposures up to 60 seconds over an area which results in 'crystal clear' imagery**
- **Unlike other commercial satellites, Capella-2 can see through dense clouds and operate at night**
- **This technology is called Synthetic Aperture Radar, which has been used by NASA since the 1970s**
- **It shoots powerful radio signals at a point of interest to 'illuminate' a target on the ground**
- **Echos from each pulse are then collected and interpreted to create the detailed images**
- **Because the satellite sends its own signal down instead of collecting light it can penetrate walls**
- **The firm has one satellite in orbit, but has signed contracts with a number of US government groups**

A new satellite that is orbiting the Earth can create high resolution images of nearly any place on our planet using radar – and is powerful enough to penetrate the walls of buildings.

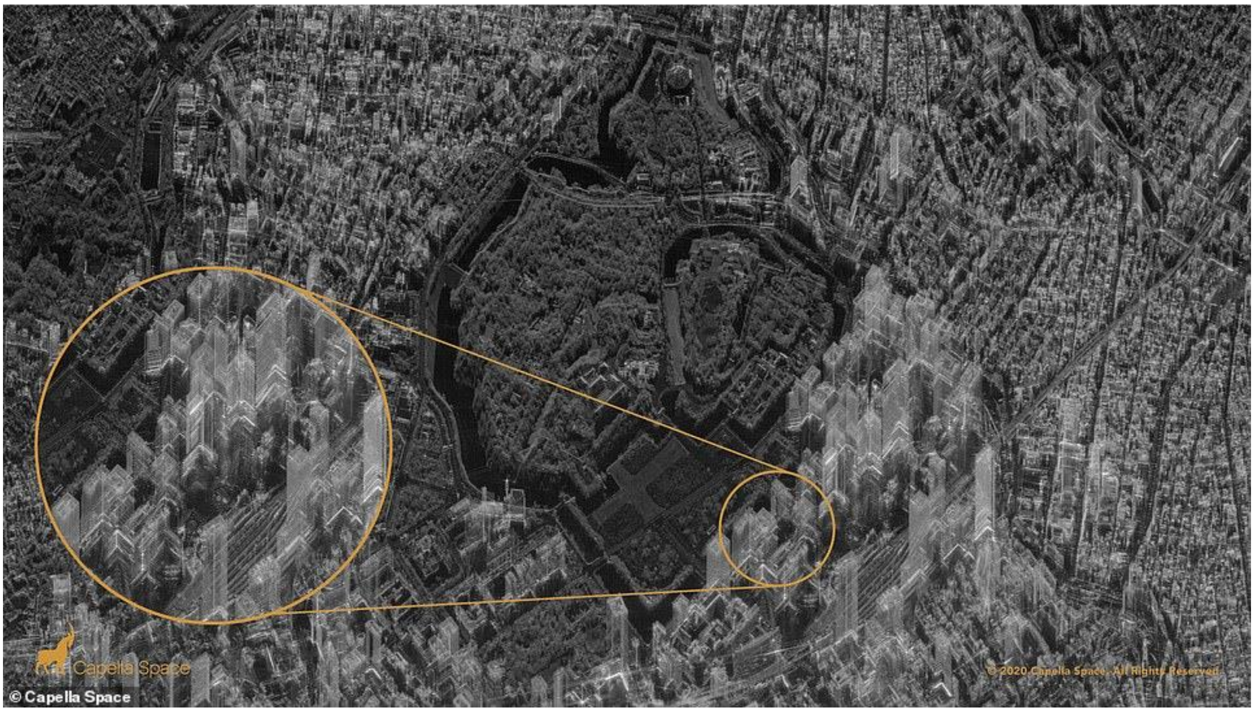


Figure 37

Capella-2, designed by Capella Space, uses Synthetic Aperture Radar (SAR), which can view the Earth regardless of air visibility, cloud covering or time of day - and is the same technology NASA has used since the 1970s.

SAR shoots powerful radio signals to 'illuminate' a point of interest, and collects data on the echo of each pulse that bounces back, interpreting them to create a detailed image.

A new satellite is orbiting the Earth that is capable of snapping high resolution images of nearly any place on our

planet – and is powerful enough to penetrate the walls of buildings. Pictured is a detailed image of Chiyoda City skyscrapers in Tokyo.

But Capella insists that the technology can't be used to spy on people in their homes, and although the radar waves can penetrate walls, they say they cannot image anything inside.

The company explains: 'The technology uses radio waves, which are capable of traveling through walls (like cell phones and Wi-Fi).

'But, as we know, even cell phone and Wi-Fi signals weaken as they move further away from a cellular tower or a Wi-Fi access point. Radar signals are the same: They can travel through walls, but are far too weak to image or see anything indoors.'

An image of Chiyoda City skyscrapers in Tokyo appears to make the towers see-through, with streets visible on the other side.

However, the firm says the layover effect is caused by 'imagery distortion giving the buildings a ghost-like appearance.'

It continues: 'What appears as a black and white optical satellite image is in reality a visual representation of the radar data — the reflectance of radio waves against the Earth's surface and manmade objects.'

Capella says their innovation is a way to help people around the world utilize space to improve their businesses and lives by monitoring everything from climate to crop fields and infrastructure.

The average commercial satellite is unable to peer through clouds or shoot detailed images of interest points at night, but Capella uses SAR that can snap images regardless of weather or light conditions.

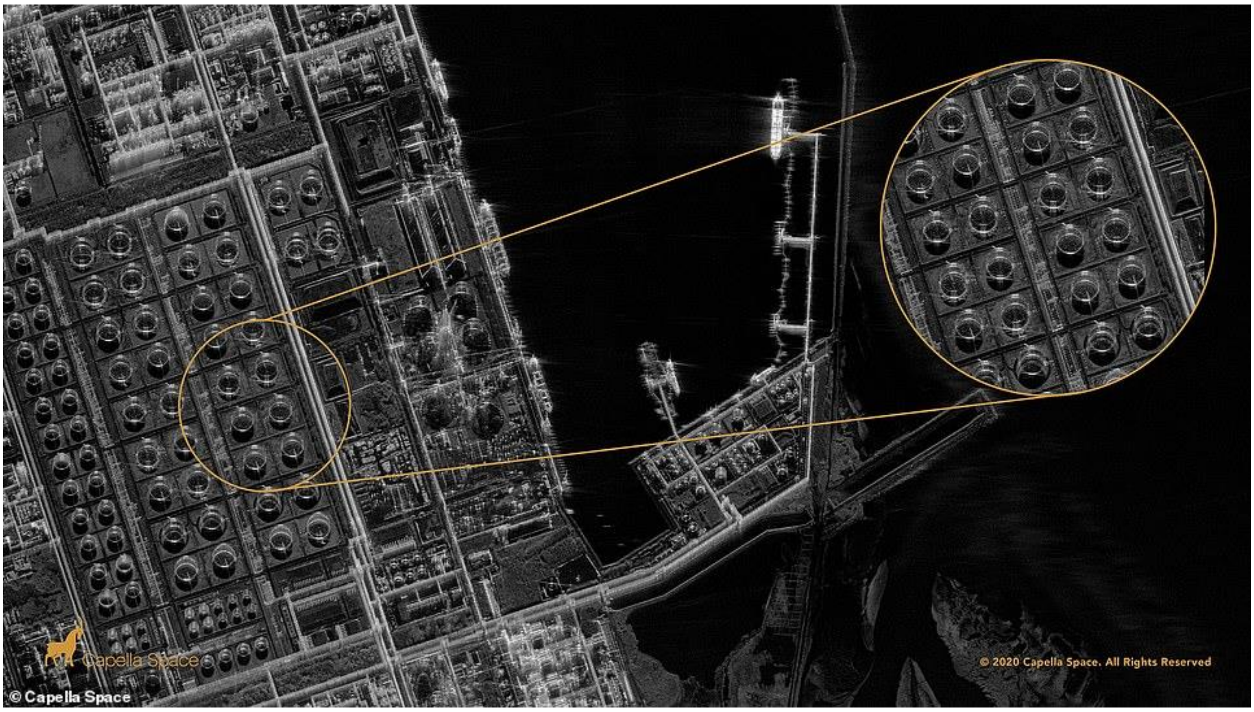


Figure 38

The firm says it is working on creating a satellite constellation of 36 devices that, combined, can monitor 'anywhere in the world every ever hour.'

As for now, Capella Space is offering governments or private companies to request images of anything from around the world.

Capella Space CEO Payam Banazadeh, a former system engineer at the NASA Jet Propulsion laboratory, said: 'Everything we do is through the lens of how we can help our customers make better decisions with data and grapple with the accelerating pace of change in the world today.'

'After launching our prototype satellite in late 2018 we evolved our satellites with customer requirements in mind, focusing on resolution, quality, latency, and user experience.

'Last week we unveiled our real-time tasking capabilities that minimizes latency and today we are unveiling our high resolution and high-quality imaging capabilities.

'Our 50 cm x 50 cm Spot imagery is the highest resolution SAR imagery in the commercial market that is allowed by US regulations.

'All of these breakthroughs, innovations and milestones will ultimately help our US Government, commercial and international ally customers make critical, high-impact and potentially life-saving decisions.'

future work

1-Working on Raspberry Pi and providing GPU

2-And work on a more accurate camera in quality and make it used in a lot of important things in space

3-Put the camera on the satellite to visualize and predict what is happening in space

References:

- Internet Trends 2019, Mary Meeker.
- report on the State of Mobile in 2022, data.io
- International Journal of Engineering and Technology · January 2010, Mobile application and its global impact, by (Dr. MD Rashedul Islam, University of Asia Pacific)&(Tridib Mazumder, M.S. Ramaiah Institute of Technology).
- YOLO & Semantic Segmentation, by Matthijs Hollemans , 2022.
- Pattern Recognition and Machine Learning, Christopher Bishop,2019.
- Artificial Intelligence: A Modern Approach, Stuart Russell, 1994.
- Designing Neural Networks using Genetic Algorithms, Satoshi Mizuta, 2001.
- OBJECT DETECTION AND IDENTIFICATION A Project Report, Mukkamala Rohith Sri Sai K L University, 2019.
- Understanding Socket Programming in Java, 2021.
- The Internet of Things with ESP32
- SPACE WEATHER PREDICTION CENTER NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION, Homepage | NOAA / NWS Space Weather Prediction Center.
- NASA