



ARTIFICIAL INTELLIGENCE

Copyright © 2018, Kaopan Boonyuen. All rights reserved.
Creative Commons Attribution Non-Commercial-ShareAlike license
Attribution required. A formal source code license is not available at this
time due to restrictions on the use of the TensorFlow library. Please contact
me if you have any questions or comments regarding this license.

Machine Learning Talks

<https://github.com/kaopanboonyuen/mltalks>

Reference:

1. <https://pytorch.org/tutorials>
2. <https://stanford.edu/~shervine/teaching/cs-229/>
3. <http://introtodeeplearning.com/>
4. <https://www.simplilearn.com/tutorials/deep-learning-tutorial/introduction-to-deep-learning>
5. <https://www.geeksforgeeks.org/introduction-deep-learning/>

About Me



Kao
Panboonyuen
kao-panboonyuen 
AI Research Scientist

Name: Kao

Contact: teerapong.pa@chula.ac.th

panboonyuen.kao@gmail.com

Education: Ph.D. in Computer Eng., Chula

Position: AI Team Lead, MARS

PostDoc, Chula

Interests: Computer Vision, Deep Learning
Machine Learning



Teerapong Panboonyuen

[FOLLOW](#)

AI/ML Senior Team Lead at MARS, Post-doc at Chula
Verified email at chula.ac.th - [Homepage](#)

[Artificial Intelligence](#) [Machine Learning](#) [Deep Learning](#) [Computer Vision](#) [Remote Sensing](#)

[GET MY OWN PROFILE](#)

TITLE

CITED BY

YEAR

Road segmentation of remotely-sensed images using deep convolutional neural networks with landscape metrics and conditional random fields

111

2017

T Panboonyuen, K Jitkajornwanich, S Lawawirojwong, P Srestasathien, ...
Remote Sensing 9 (7), 680

Semantic segmentation on remotely sensed images using an enhanced global convolutional network with channel attention and domain specific transfer learning

76

2019

T Panboonyuen, K Jitkajornwanich, S Lawawirojwong, P Srestasathien, ...
Remote Sensing 11 (1), 83

An enhanced deep convolutional encoder-decoder network for road segmentation on aerial imagery

35

2018

T Panboonyuen, P Vateekul, K Jitkajornwanich, S Lawawirojwong
Recent Advances in Information and Communication Technology 2017 ...

Transformer-based decoder designs for semantic segmentation on remotely sensed images

16

2021

T Panboonyuen, K Jitkajornwanich, S Lawawirojwong, P Srestasathien, ...
Remote Sensing 13 (24), 5100

Object detection of road assets using transformer-based YOLOX with feature pyramid decoder on thai highway panorama

12

2022

T Panboonyuen, S Thongbai, W Wongweeranimit, P Santitamont, ...
Information 13 (1), 5

Semantic segmentation on medium-resolution satellite images using deep convolutional networks with remote sensing derived indices

12

2018

S Chantharaj, K Pornrathanapong, P Chitsimpachayakun, T Panboonyuen, ...
2018 15th International joint conference on computer science and software ...

Real-time polyps segmentation for colonoscopy video frames using compressed fully convolutional network

12

2018

I Wichakarn, T Panboonyuen, C Udomcharoenchaikit, P Vateekul
MultiMedia Modeling: 24th International Conference, MMM 2018, Bangkok ...

Semantic Labeling in Remote Sensing Corpora Using Feature Fusion-Based Enhanced Global Convolutional Network with High-Resolution Representations and Depthwise Atrous Convolution

10

2020

T Panboonyuen, K Jitkajornwanich, S Lawawirojwong, P Srestasathien, ...
Remote Sensing 12 (8), 1233

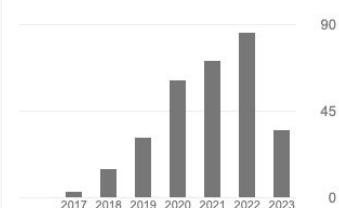
Cited by

All Since 2018

Citations 303 299

h-index 8 8

i10-index 8 8



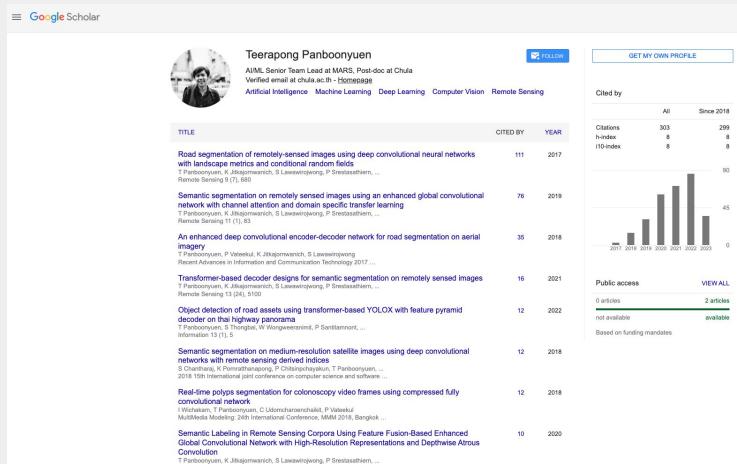
Public access [VIEW ALL](#)

0 articles 2 articles

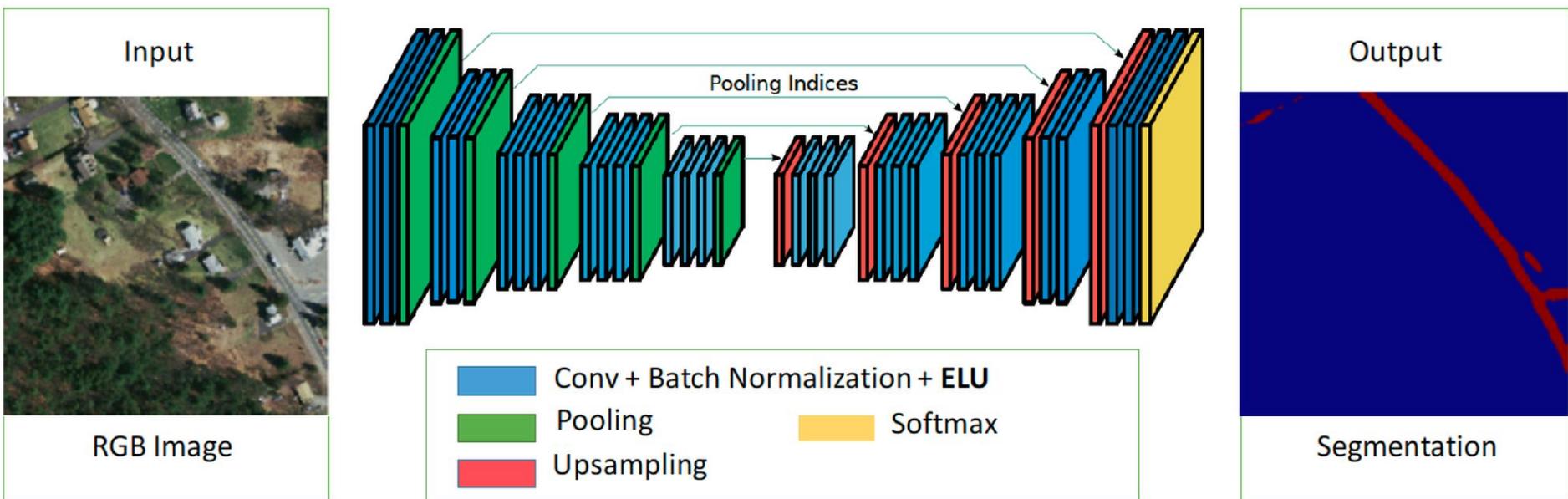
not available available

Based on funding mandates

Featured Publications

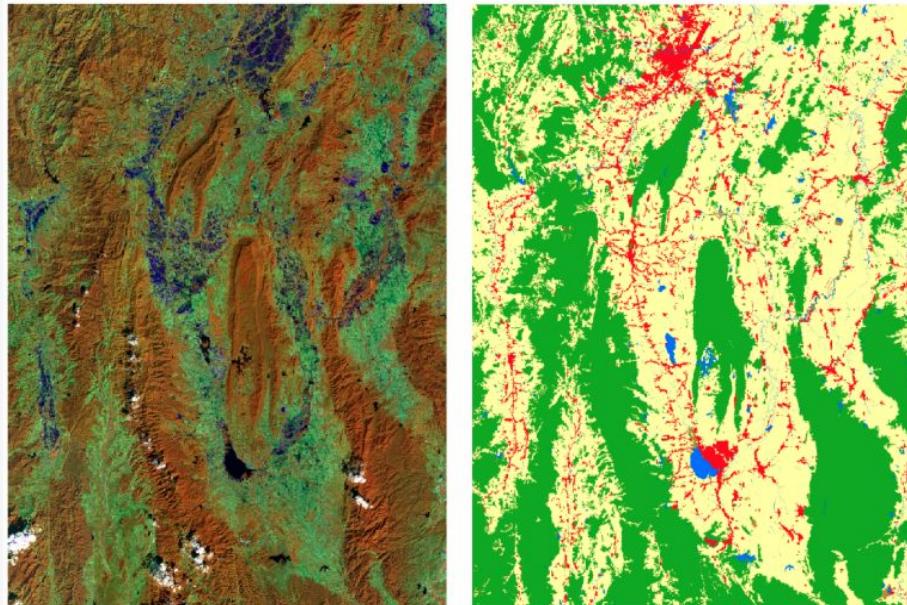


[1] Panboonyuen, Teerapong, et al. "Road segmentation of remotely-sensed images using deep convolutional neural networks with landscape metrics and conditional random fields." *Remote Sensing* 9.7 (2017): 680.



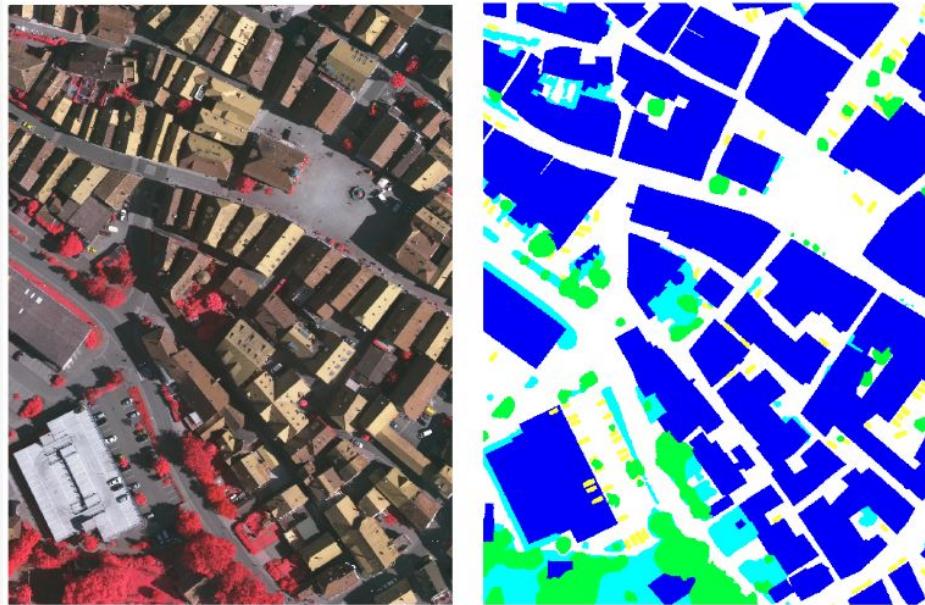
[2] Panboonyuen, Teerapong, et al. "Semantic segmentation on remotely sensed images using an enhanced global convolutional network with channel attention and domain specific transfer learning." Remote Sensing 11.1 (2019): 83.

Figure 6. Sample satellite images from Nan, a province in Thailand (**left**), and corresponding ground truth (**right**). The label of medium resolution dataset includes five categories: agriculture (yellow), forest (green), miscellaneous (brown), urban (red), and water (blue).



X

Figure 8. The sample input tile from **Figure 7 (left)** and corresponding ground truth (**right**). The label of the Vaihingen Challenge includes six categories: impervious surface (imp surf, white), building (blue), low vegetation (low veg, cyan), tree (green), car (yellow), and clutter/background (red).



X



30 FPS



[4] Wichakam, I., Panboonyuen, T., Udomcharoenchaikit, C., & Vateekul, P. (2018). **Real-time polyps segmentation for colonoscopy video frames using compressed fully convolutional network**. In MultiMedia Modeling: 24th International Conference, MMM 2018, Bangkok, Thailand, February 5-7, 2018, Proceedings, Part I 24 (pp. 393-404). Springer International Publishing.

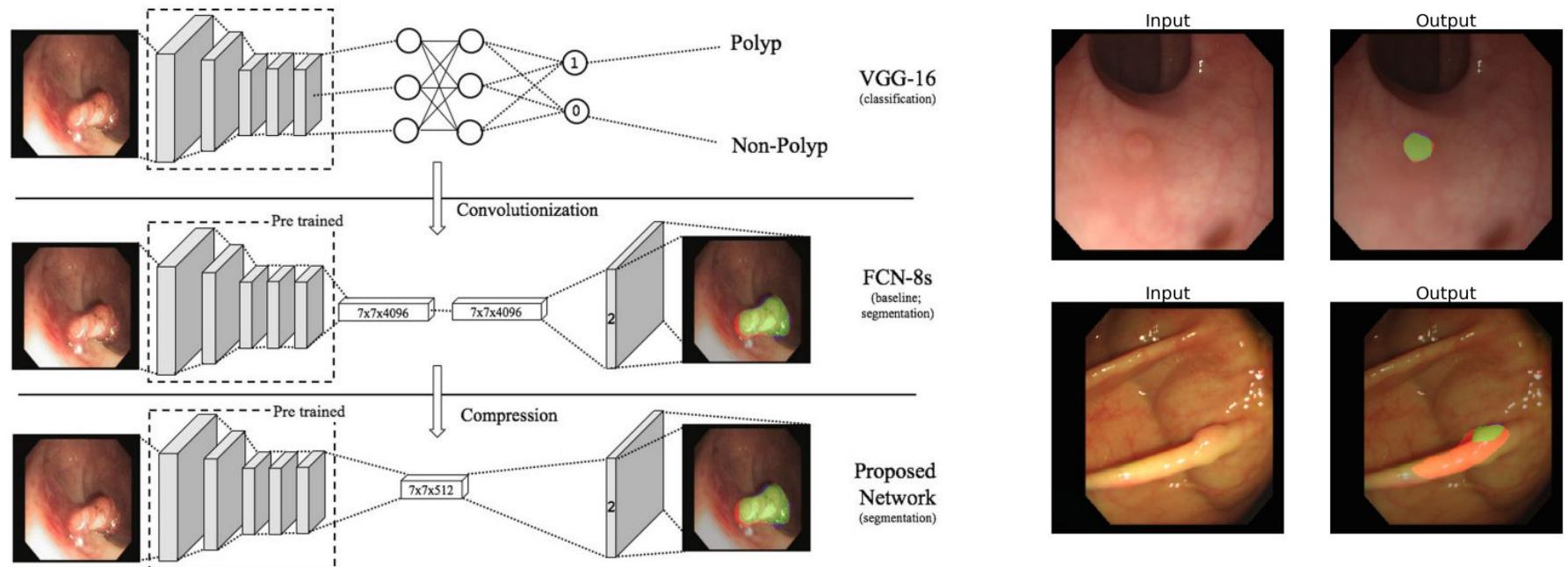
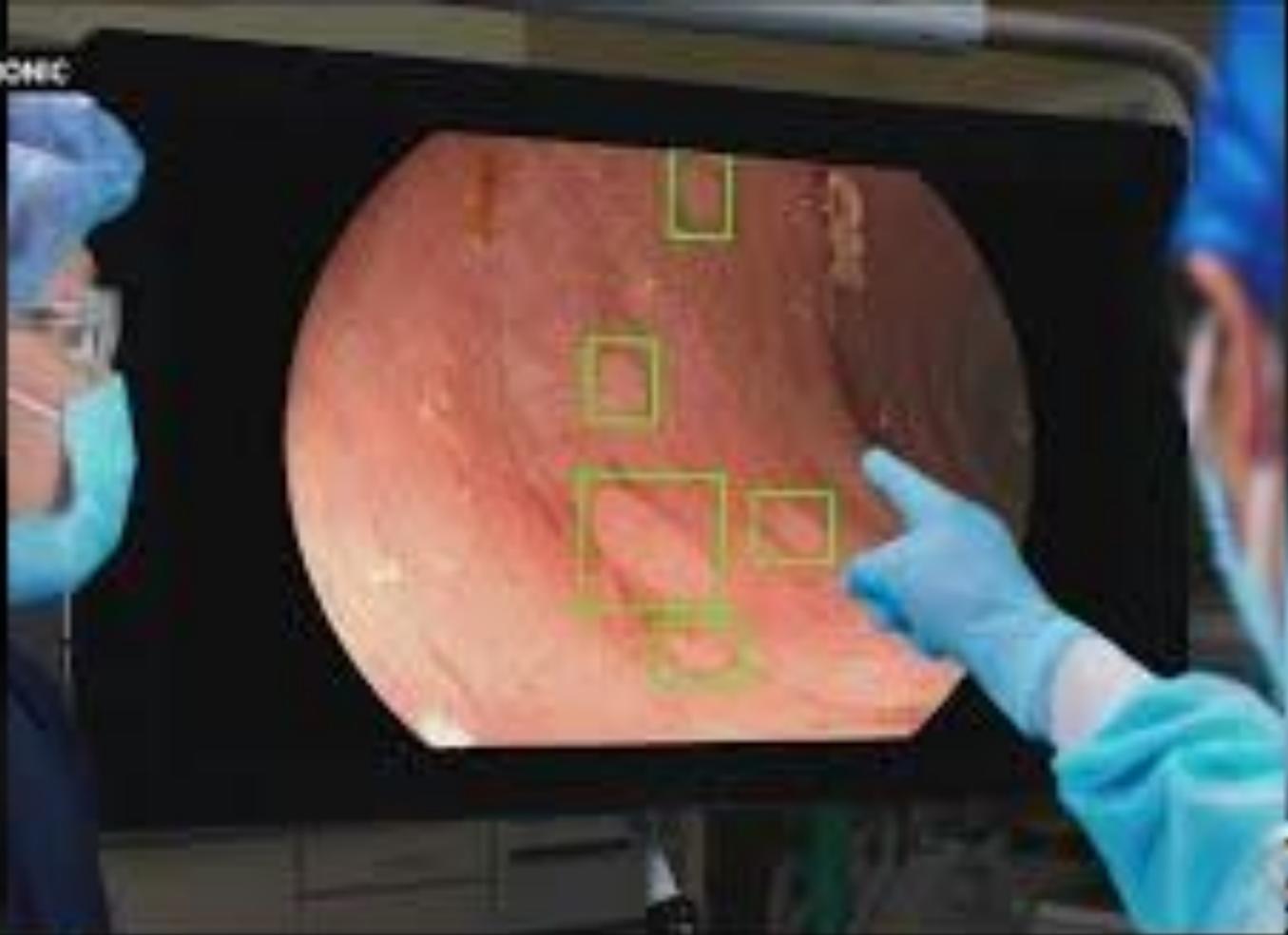


Fig. 1. Overview of our compressed network which is compressed from the original FCN-8s [8] based on VGG-16 [4] architecture.

CREDIT: MEDTRONIC



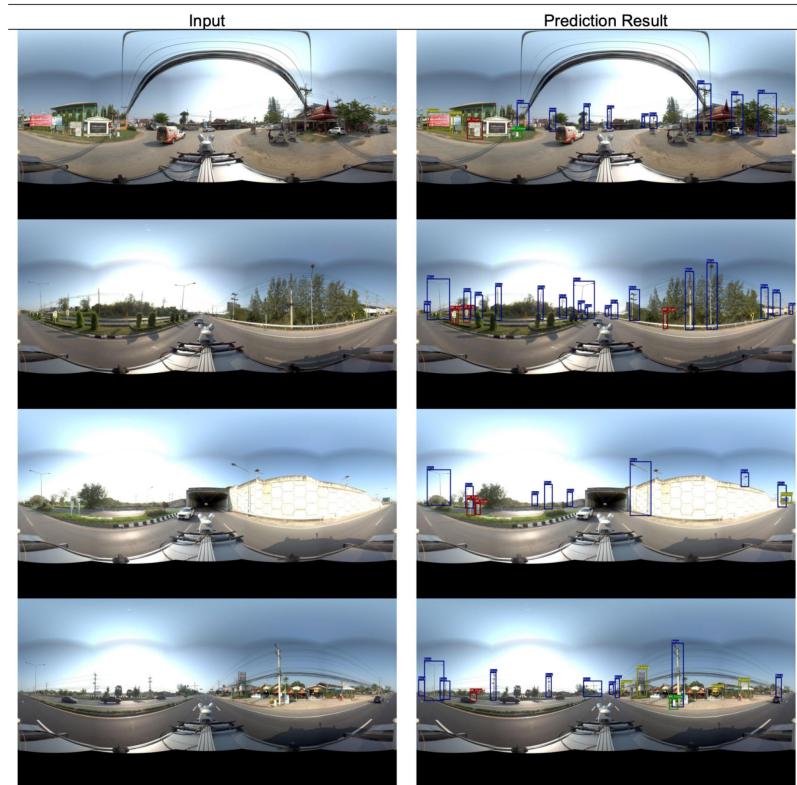
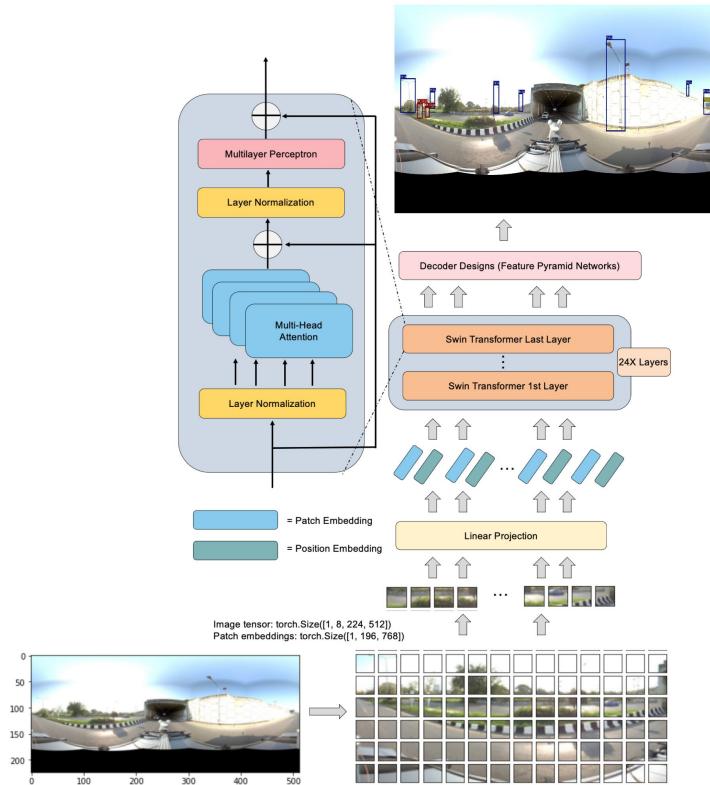
abc 2

HEADLINES

TASK FORCE

83° 5:24 PM

[5] Panboonyuen, Teerapong, et al. "Object detection of road assets using transformer-based YOLOX with feature pyramid decoder on thai highway panorama." Information 13.1 (2022): 5.



[6] Thitisiriwech, K., Panboonyuen, T., Kantavat, P., Iwahori, Y., & Kijsirikul, B. (2022). The Bangkok Urbanscapes Dataset for Semantic Urban Scene Understanding Using Enhanced Encoder-Decoder With Atrous Depthwise Separable A1 Convolutional Neural Networks. *IEEE Access*, 10, 59327-59349.

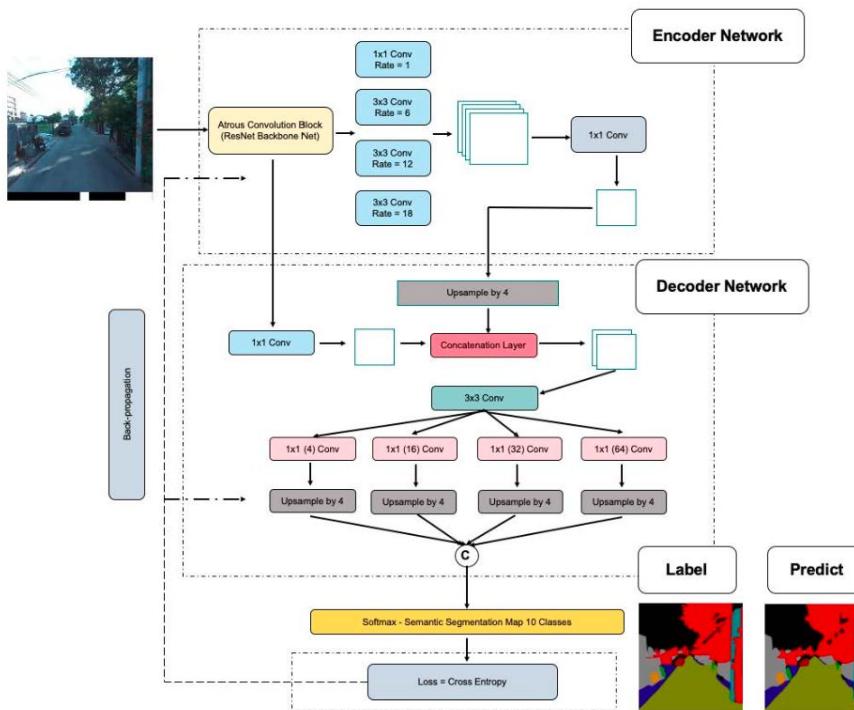


FIGURE 6. An overview of enhanced DeepLab-V3+ (Encoder-Decoder with atrous separable convolutional for semantic segmentation [15]) with ResNet-101 backbone [41] (DeepLab-V3-A1-ResNet-101).

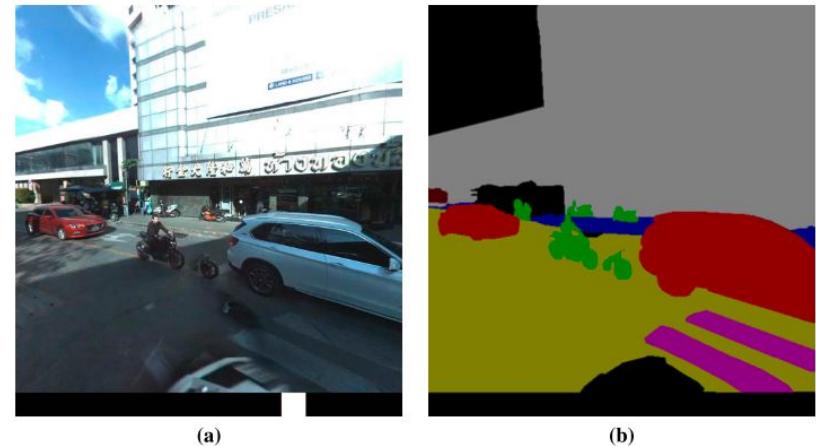


FIGURE 8. Sample 1: The example of Sukhumvit's large road from the training set of the Bangkok Urbanscapes dataset. The input image is shown in (a), and the ground truth is shown in (b).

Void	Building	Wall	Tree	VegetationMisc	Fence
Sidewalk	ParkingBlock	Column_Pole	TrafficConc	Bridge	SignSymbol
Misc_Text	TrafficLight	Sky	Tunnel	Archway	Road
RoadShoulder	LaneMkgsDriv	LaneMkgsNonDriv	Animal	Pedestrian	Child
CartLuggagePram	Bicyclist	MotorcycleScooter	Car	SUVPickupTruck	Truck_Bus
Train	OtherMoving				

FIGURE 3. The semantic color codes of the CamVid dataset. Each color is encoded with respect to the semantic class in the ground truth images.

[6] Vajeethaveesin, T., Panboonyuen, T., Lawawironjwong, S., Srestasathiern, P., Jaiyen, S., & Jitkajornwanich, K. (2022). **A performance comparison between GIS-based and neuron network methods for flood susceptibility assessment in ayutthaya province.** Trends in Sciences, 19(2), 2038-2038.

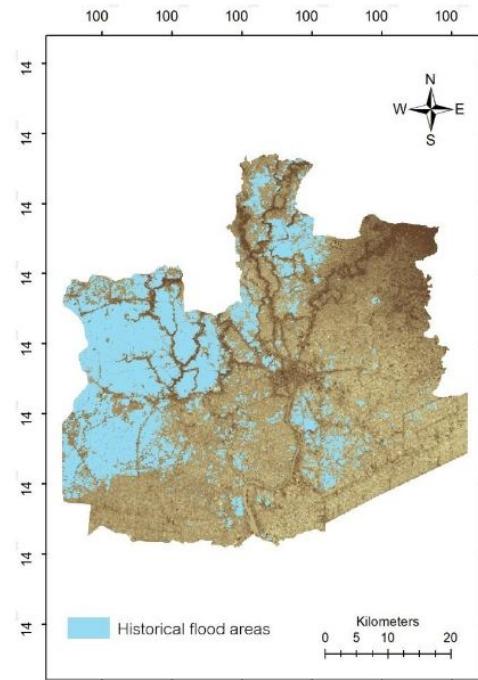


Figure 2 The historical flood zone shows flood events, which were captured in 2016 from GISTDA. Most of the flooding areas are located in the west of the study area.

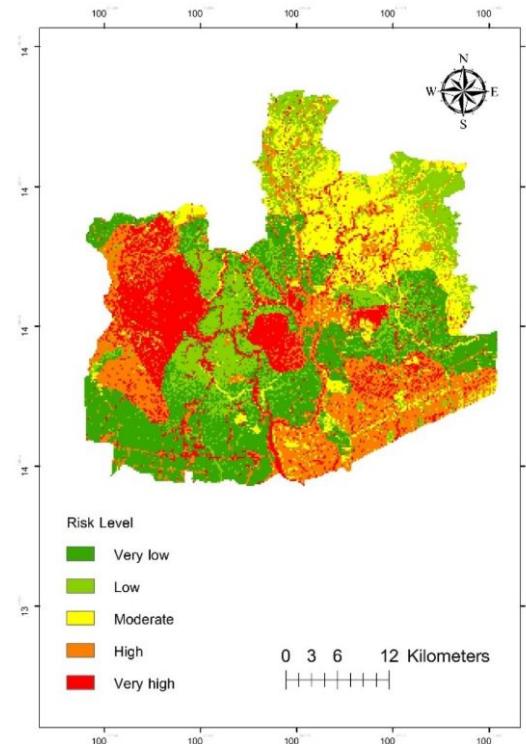
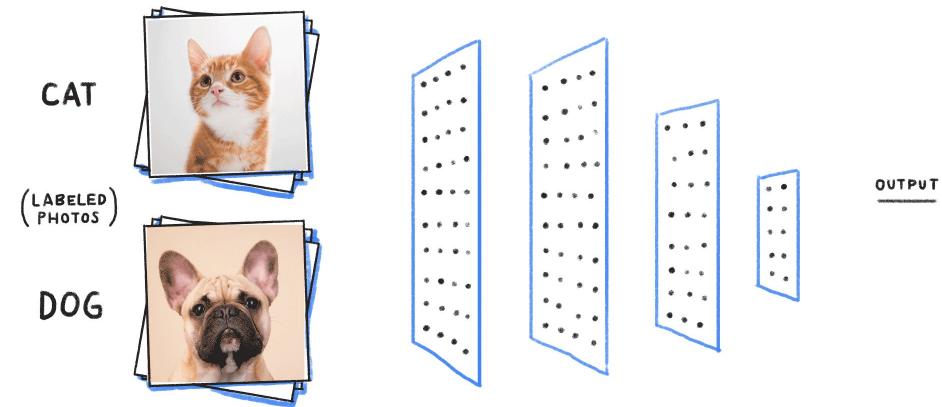


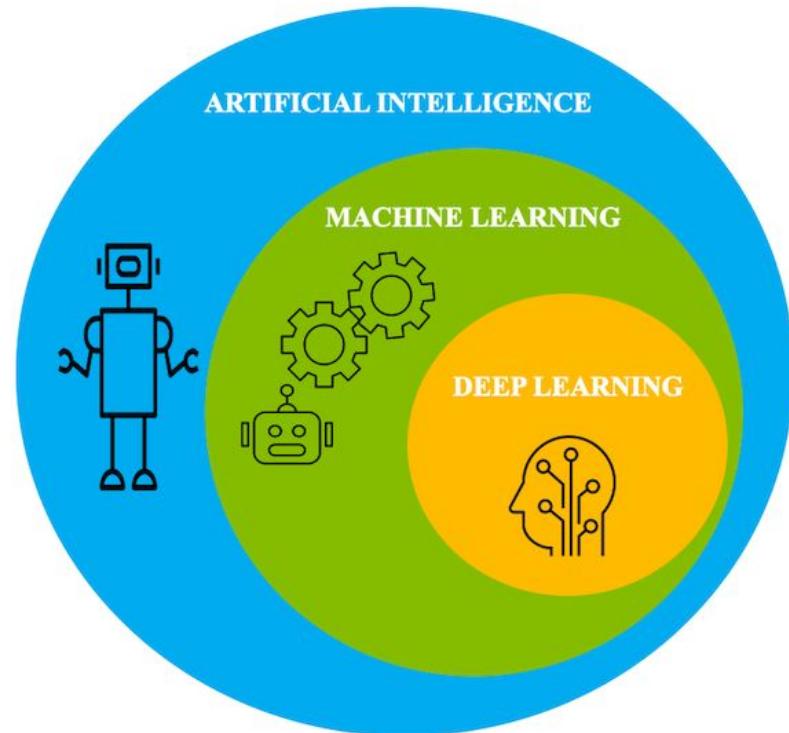
Figure 4 Flood susceptibility map from the flood risk assessment model (FRAM).

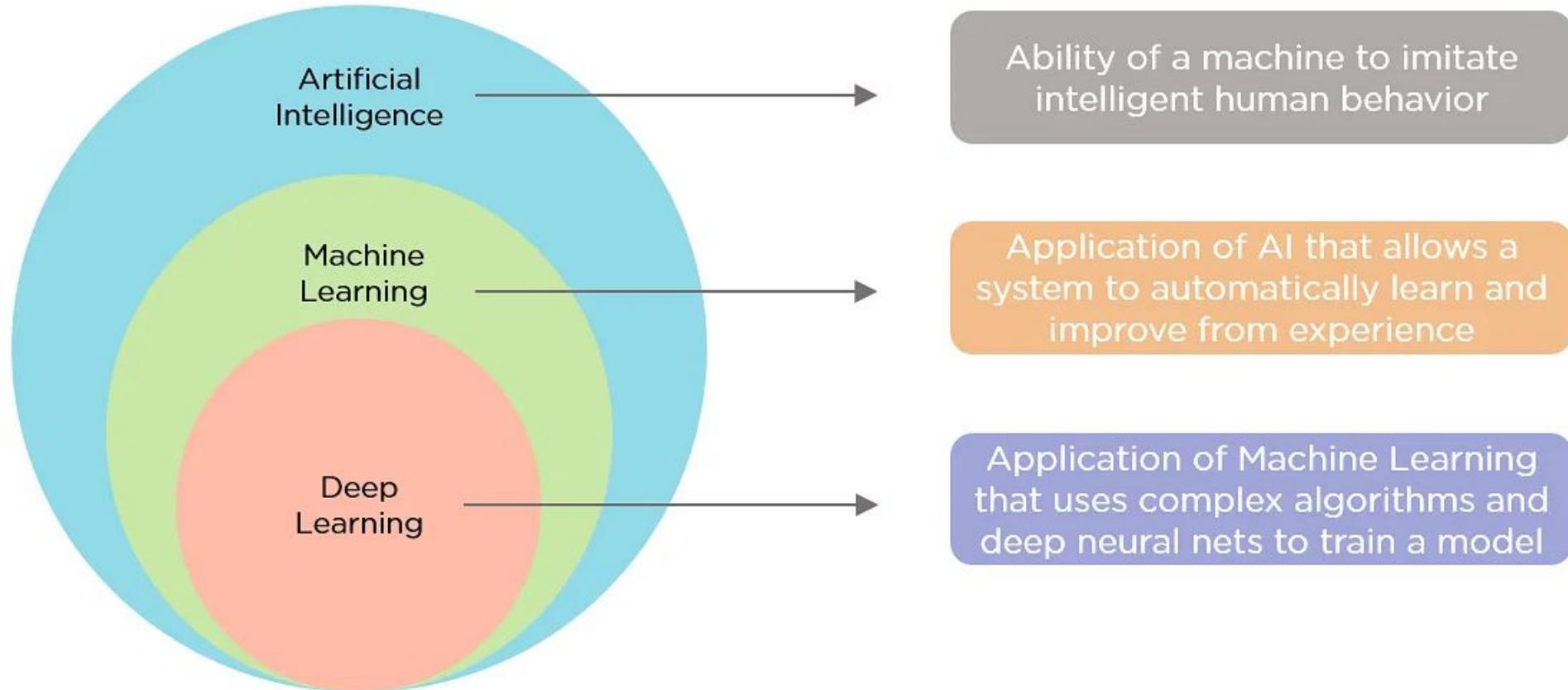
Basic AI



AI vs Machine Learning vs Deep Learning

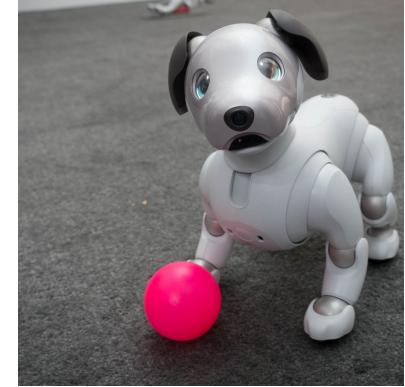
- **Artificial Intelligence (AI)** is the concept of creating smart intelligent machines.
- **Machine Learning (ML)** is a subset of artificial intelligence that helps you build AI-driven applications.
- **Deep Learning** is a subset of machine learning that uses vast volumes of data and complex algorithms to train a model.





Applications of Artificial Intelligence

- Machine Translation such as Google Translate
- Self Driving Vehicles such as Google's Waymo
- AI Robots such as Sophia and Aibo
- Speech Recognition applications like Apple's Siri or OK Google



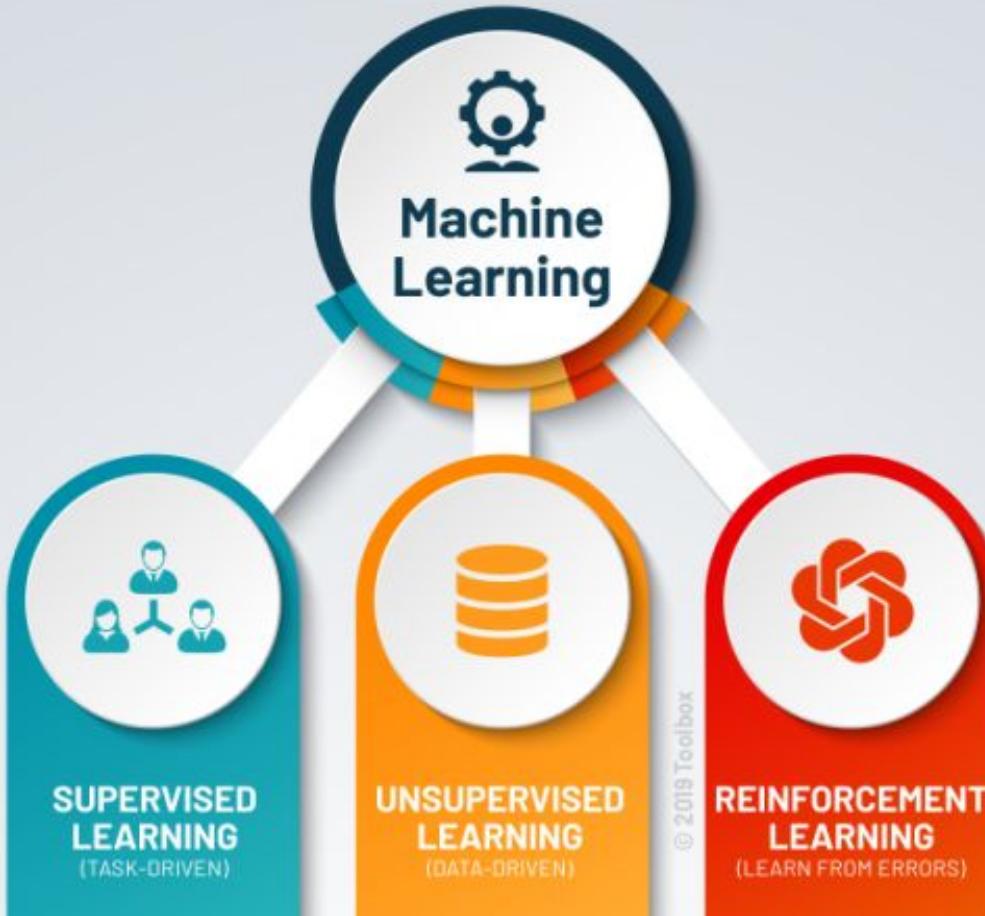
<https://www.youtube.com/watch?v=kJoAcEl2PXQ&t=14s>

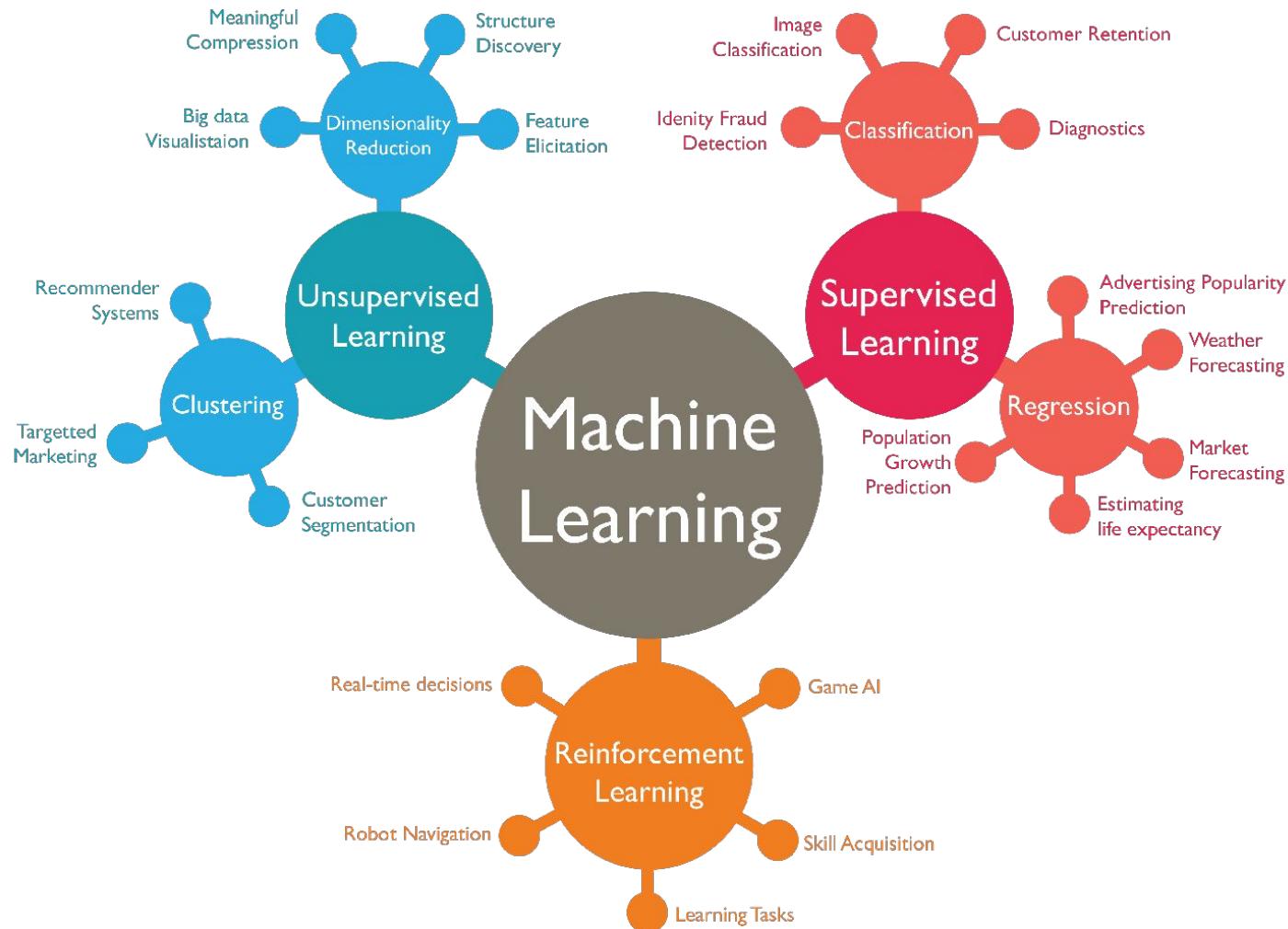


Sophia is a humanoid robot with unique features developed by Hanson Robotics Limited



TYPES OF MACHINE LEARNING





Supervised Learning

Input



Annotations

These are
grapes

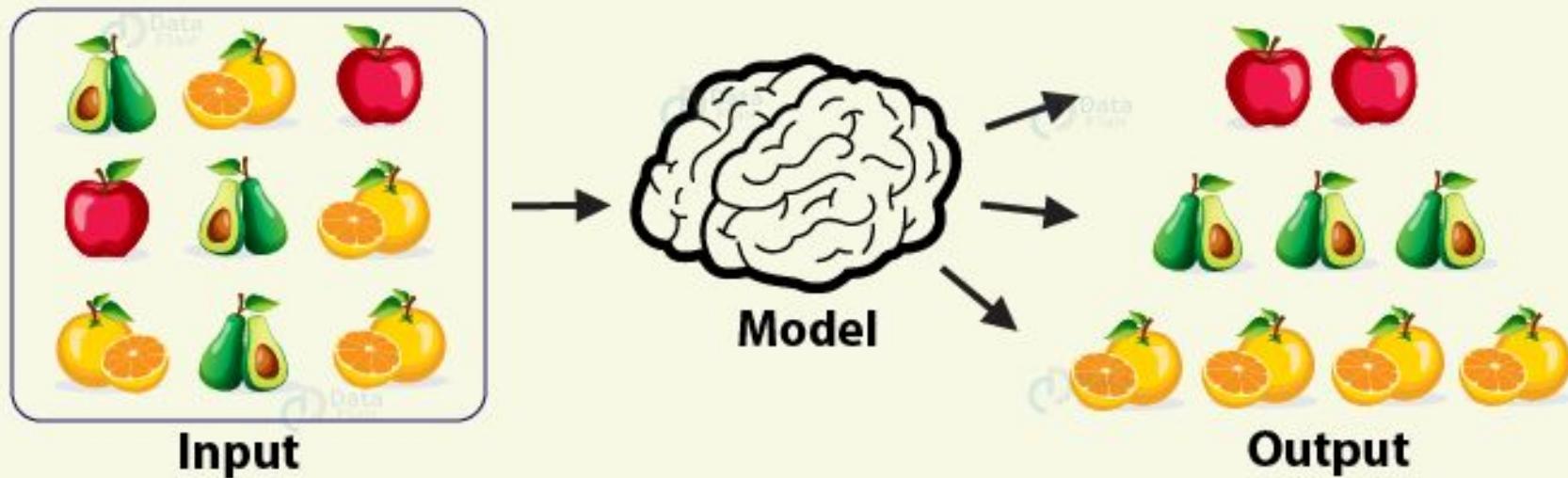
Model



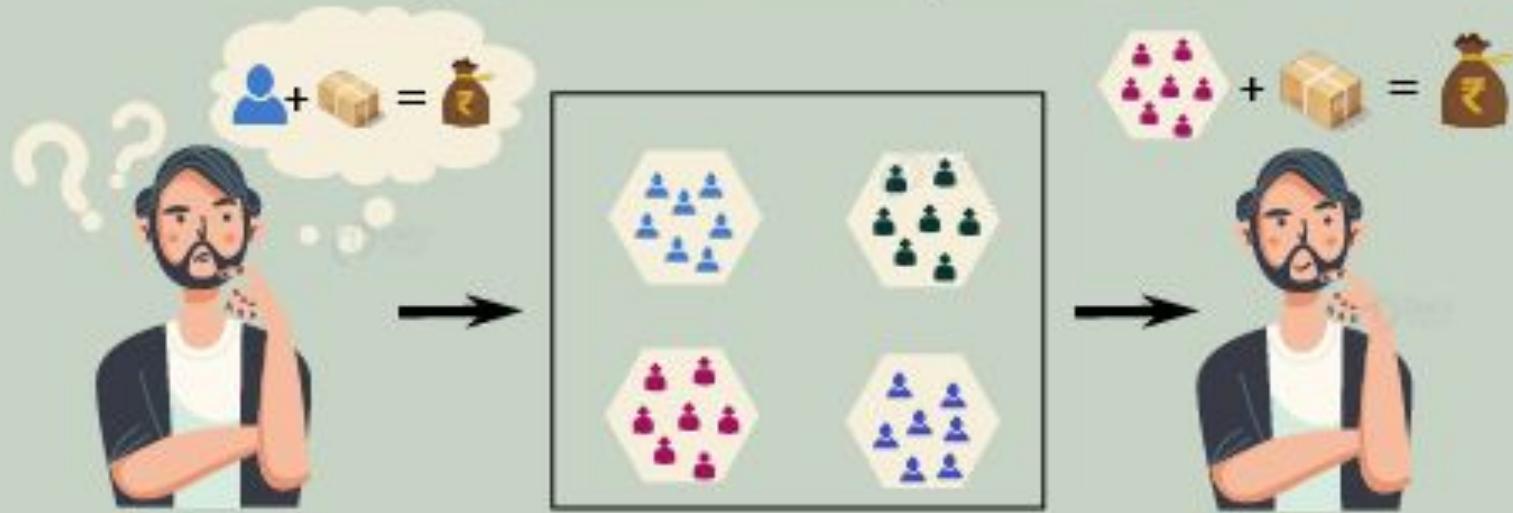
It's Grapes

Prediction

Unsupervised Learning



Unsupervised Learning in real life

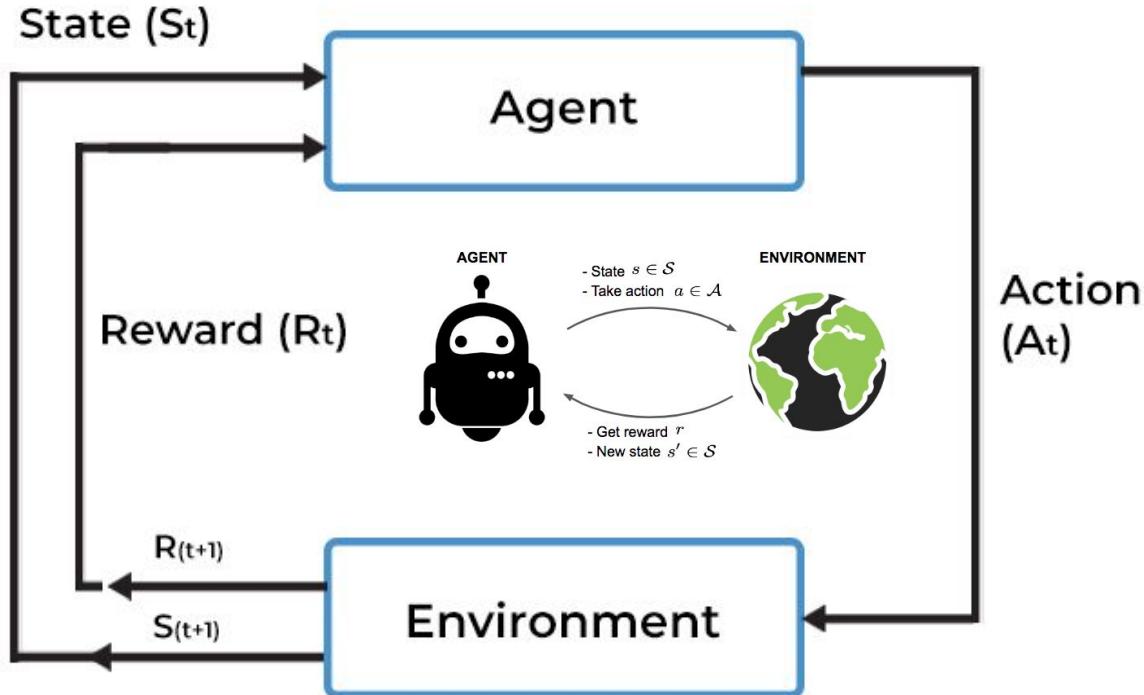


Identifying the potential customer base for selling the product

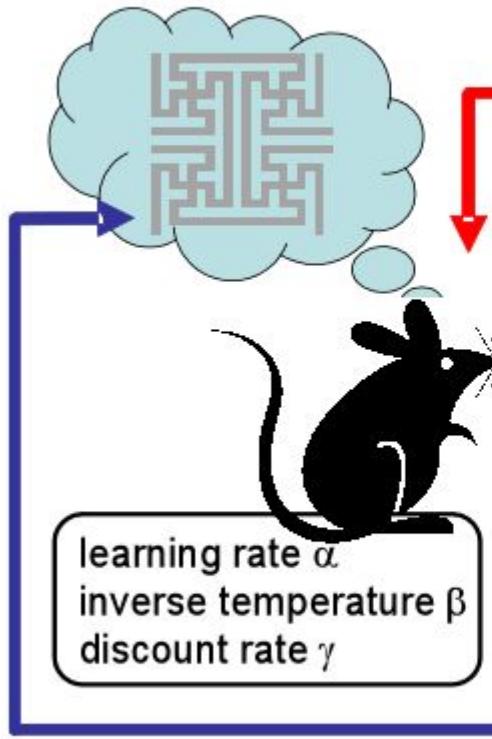
Implementing Clustering Algorithms to group the customer base

Selling product to the identified customer group

REINFORCEMENT LEARNING MODEL



internal state



reward

environment



observation

SUPER MARIO WORLD™



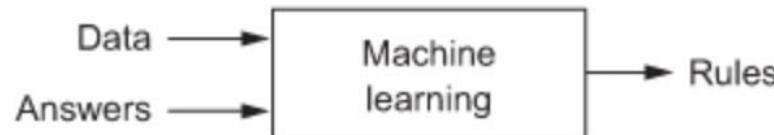
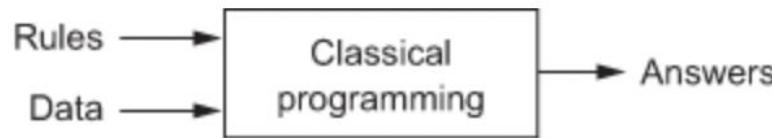
© 1990, 1991 Nintendo



What is Deep Learning?

Deep Learning is a subset of Machine Learning that uses mathematical functions to map the input to the output.

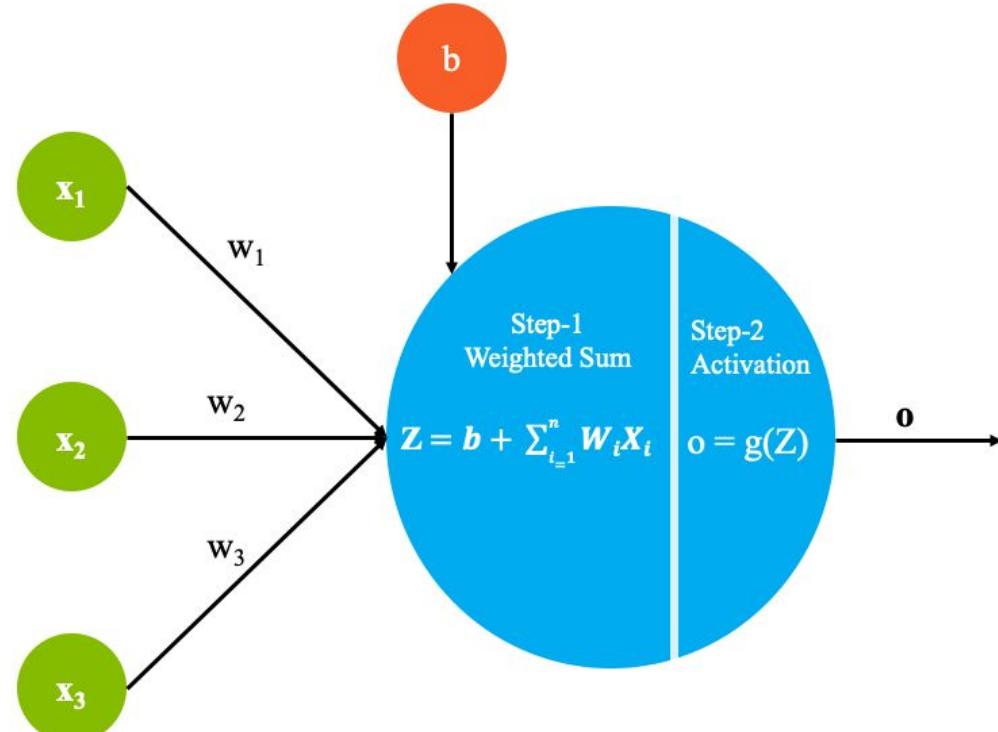
These functions can extract non-redundant information or patterns from the data, which enables them to form a relationship between the input and the output.



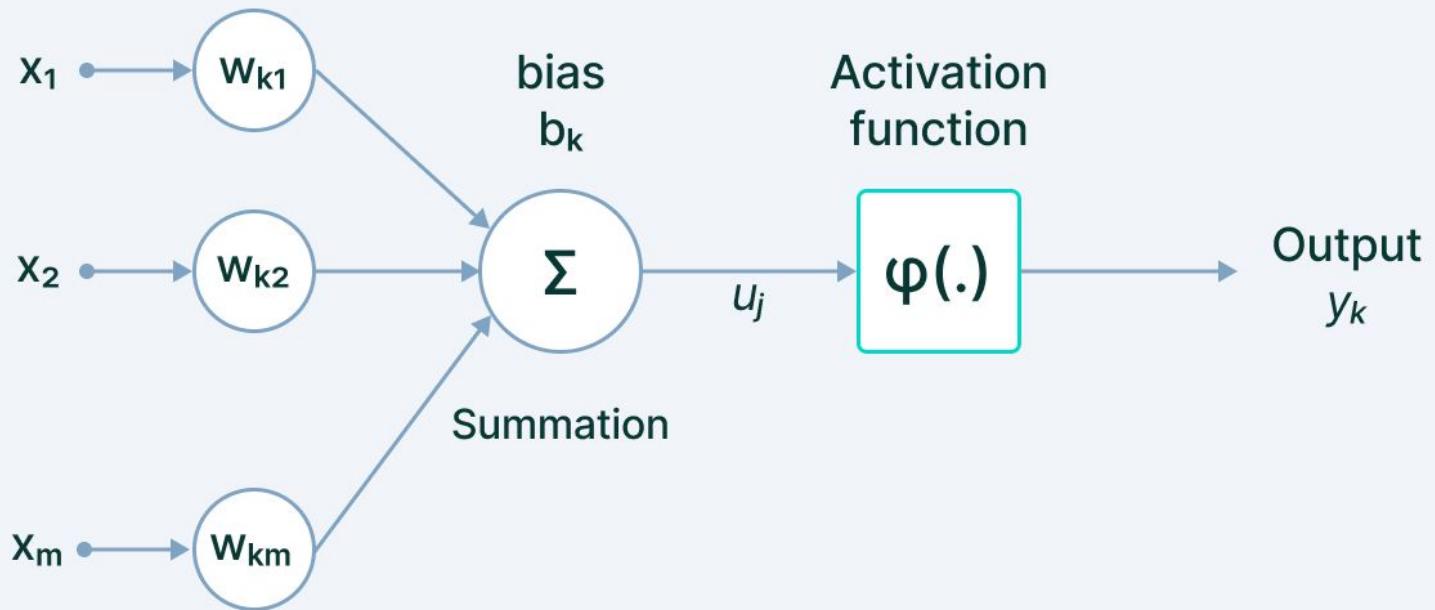
Artificial Neural Network (ANN)

This section provides an overview of the architecture behind deep learning, artificial neural networks (ANN), and discusses some of the key terminology.

As shown in the following figure, each perceptron is made up of the following parts:



Neuron



Sigmoid

$$y = \frac{1}{1 + e^{-x}}$$

Tanh

$$y = \tanh(x)$$

Step Function

$$y = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

Softplus

$$y = \ln(1 + e^x)$$

$$y = \frac{1}{1 + e^{-x}}$$

ReLU

$$y = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases}$$

Softsign

$$y = \frac{x}{(1+|x|)}$$

ELU

$$y = \begin{cases} \alpha(e^{x-1}) - 1, & x < 0 \\ x, & x \geq 0 \end{cases}$$

Log of Sigmoid

$$y = \ln\left(\frac{1}{1 + e^{-x}}\right)$$

$$y = \max(0.1x, x)$$

Swish

$$y = \frac{x}{1 + e^{-x}}$$

Sinc

$$y = \frac{\sin(x)}{x}$$

Leaky ReLU

$$y = \max(0.01x, x)$$

Mish

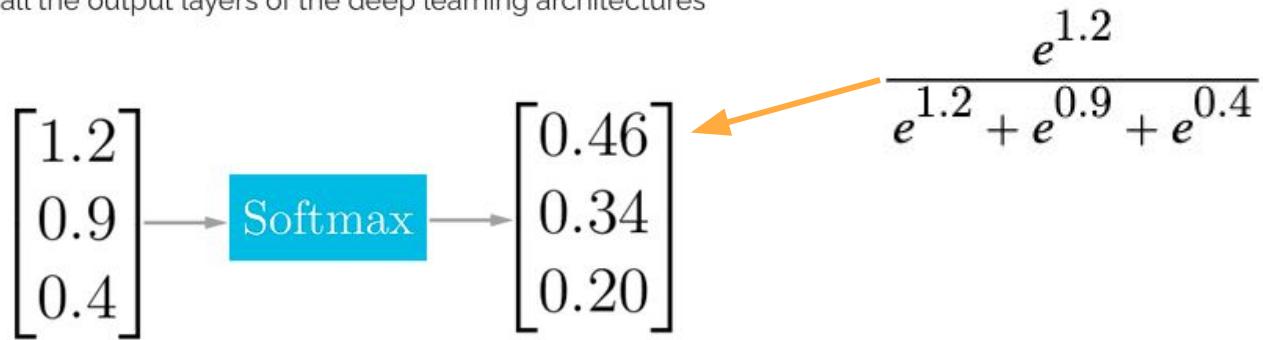
$$y = x(\tanh(\text{softplus}(x)))$$

Softmax Function

The softmax function is used to compute probability distribution from a vector of real numbers. The Softmax function produces an output which is a range of values between 0 and 1, with the sum of the probabilities being equal to 1. The Softmax function is computed using the relationship:

$$f(x_i) = \frac{\exp(x_i)}{\sum_j \exp(x_j)}$$

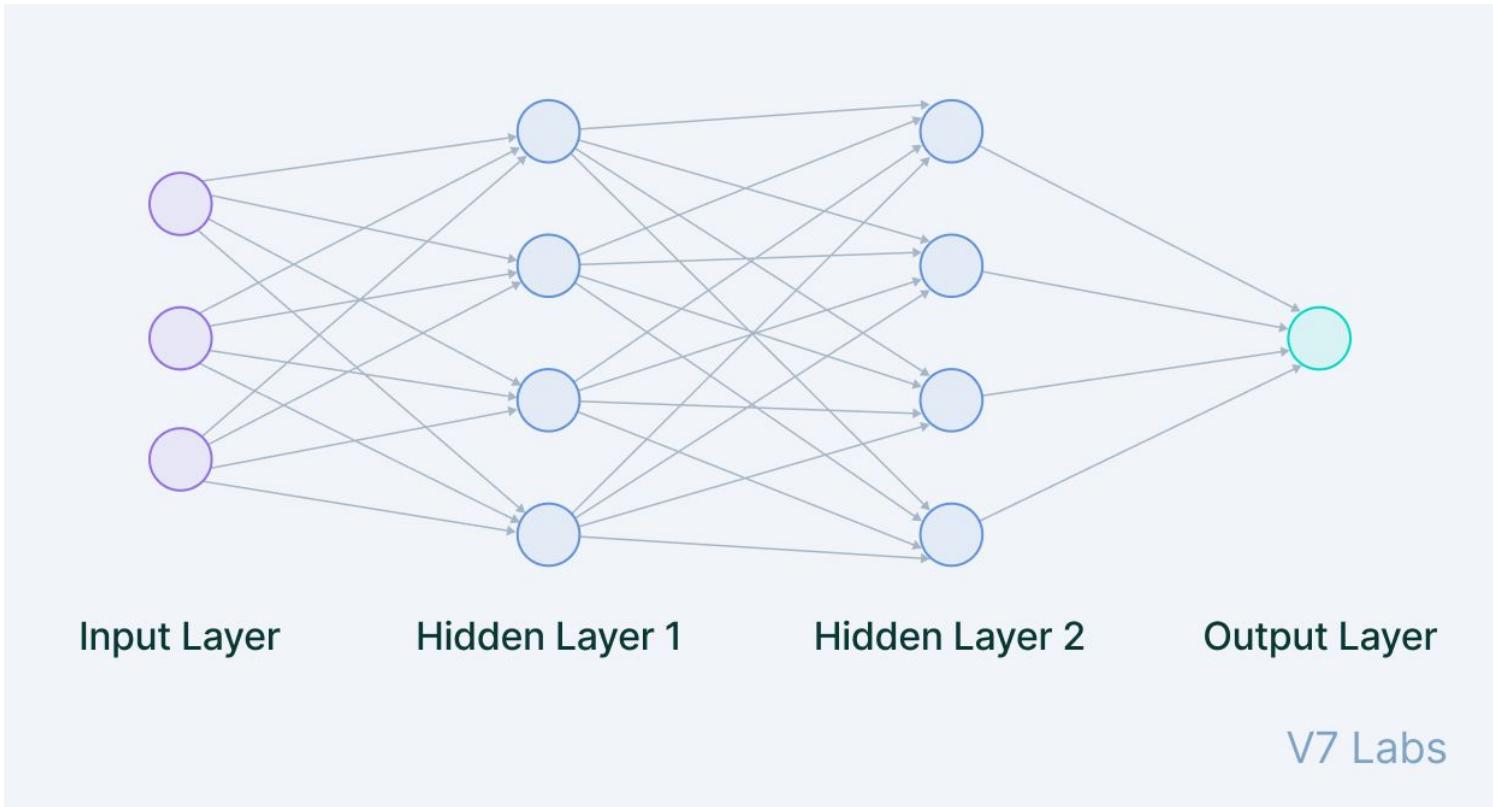
The Softmax function is used in multi-class models where it returns probabilities of each class, with the target class having the highest probability. The Softmax function mostly appears in almost all the output layers of the deep learning architectures

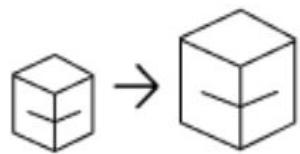


The softmax function squashes the outputs of each unit to be between 0 and 1, just like a sigmoid. It also divides each output such that the total sum of the outputs is equal to 1.

Credits

How does Deep Learning work?

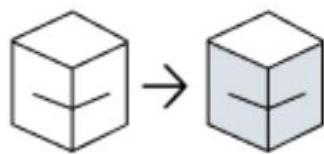


 \rightarrow 

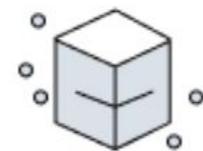
Take input and
multiply by the
neuron's weight



Add bias



Feed the result, x ,
to the activation
function: $f(x)$



Take the output
and transmit to the
next layer of
neurons

V7 Labs

Forward propagation in neural networks

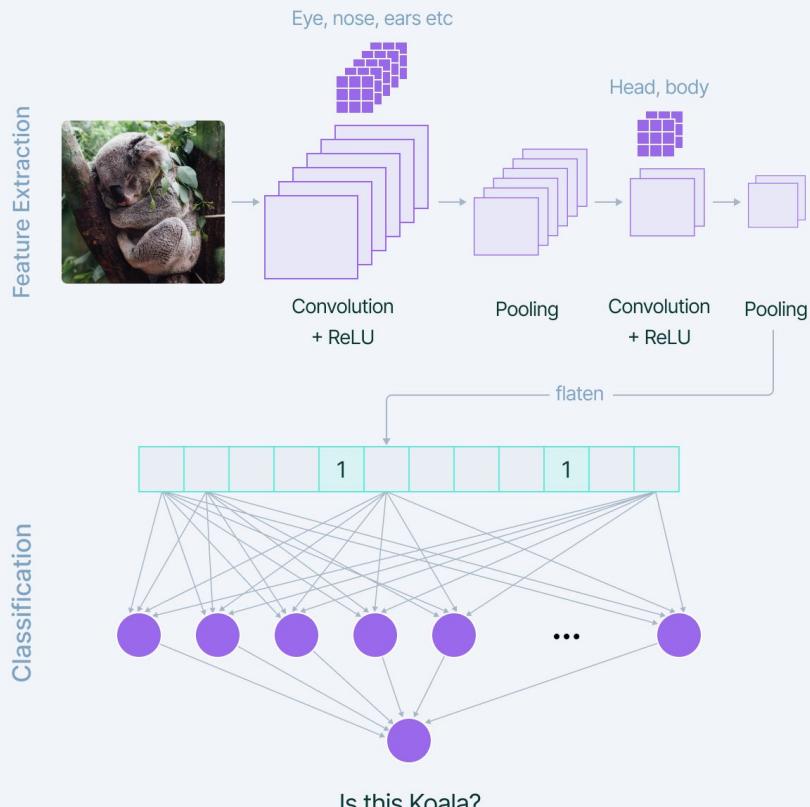
The Convolutional Neural Networks or CNNs

The Convolutional Neural Networks or CNNs are primarily used for tasks related to computer vision or image processing.

CNNs are extremely good in modeling spatial data such as 2D or 3D images and videos.

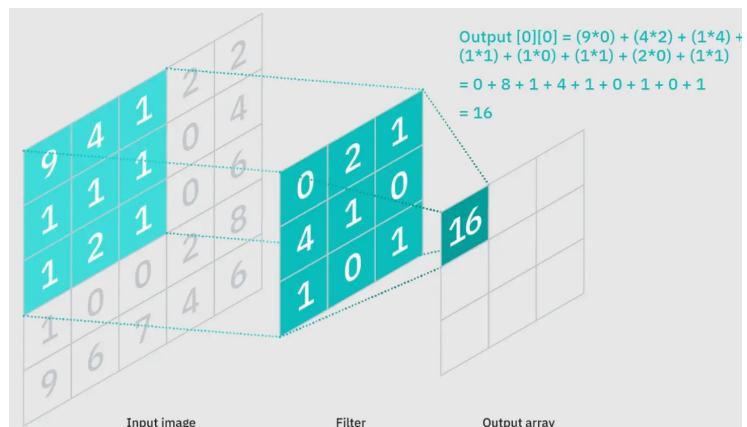
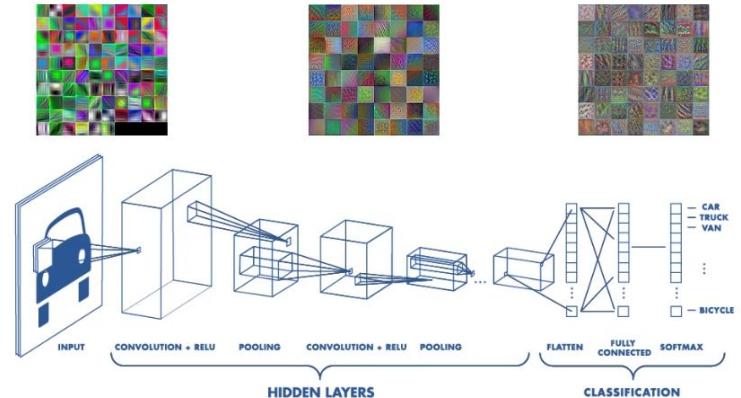
They can extract features and patterns within an image, enabling tasks such as image classification or object detection.

Convolutional Neural Networks



V7 Labs

Convolutional Neural Network



The Convolution Operation

We slide the 3x3 filter over the input image, element-wise multiply, and add the outputs:

1 <small>$\times 1$</small>	1 <small>$\times 0$</small>	1 <small>$\times 1$</small>	0	0
0 <small>$\times 0$</small>	1 <small>$\times 1$</small>	1 <small>$\times 0$</small>	1	0
0 <small>$\times 1$</small>	0 <small>$\times 0$</small>	1 <small>$\times 1$</small>	1	1
0	0	1	1	0
0	1	1	0	0



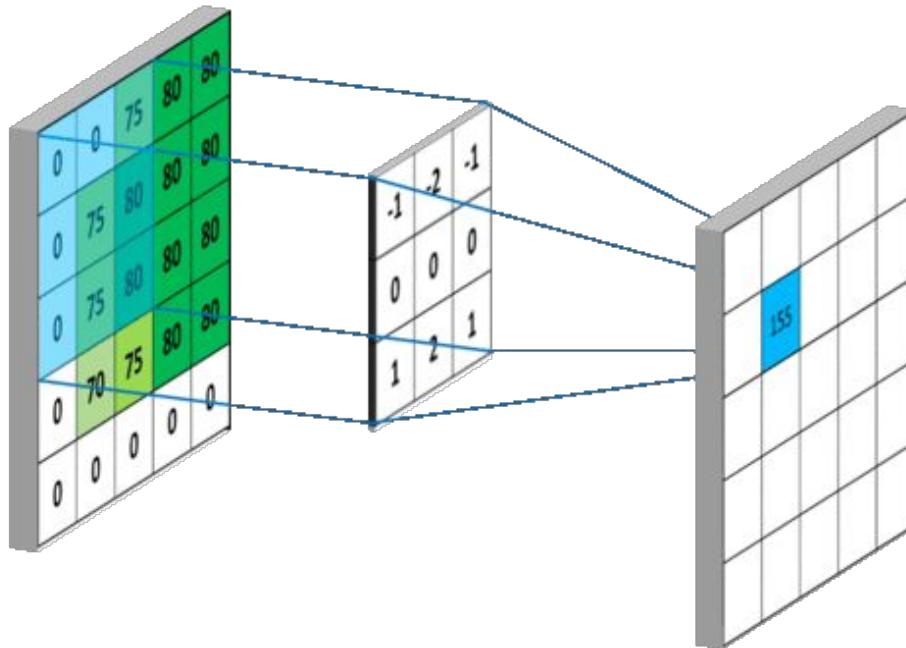
1	0	1
0	1	0
1	0	1

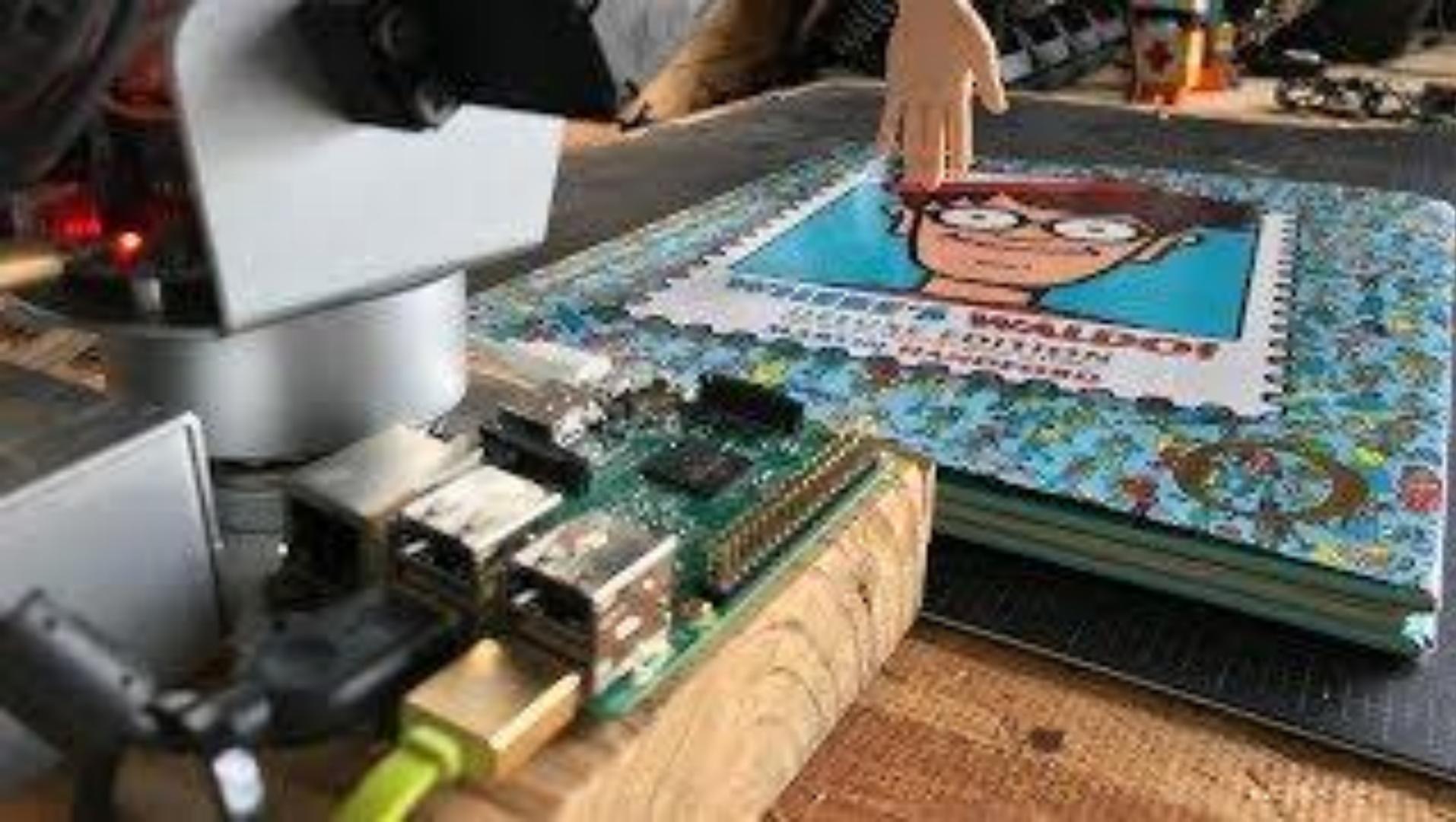
filter



4		

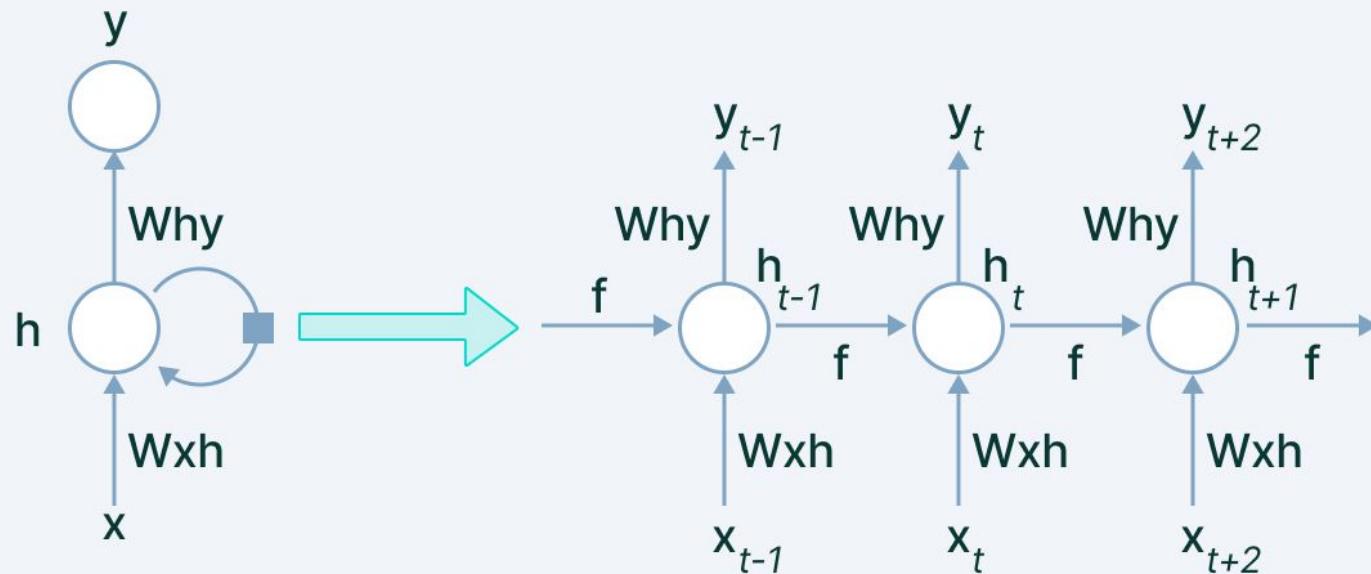
feature map

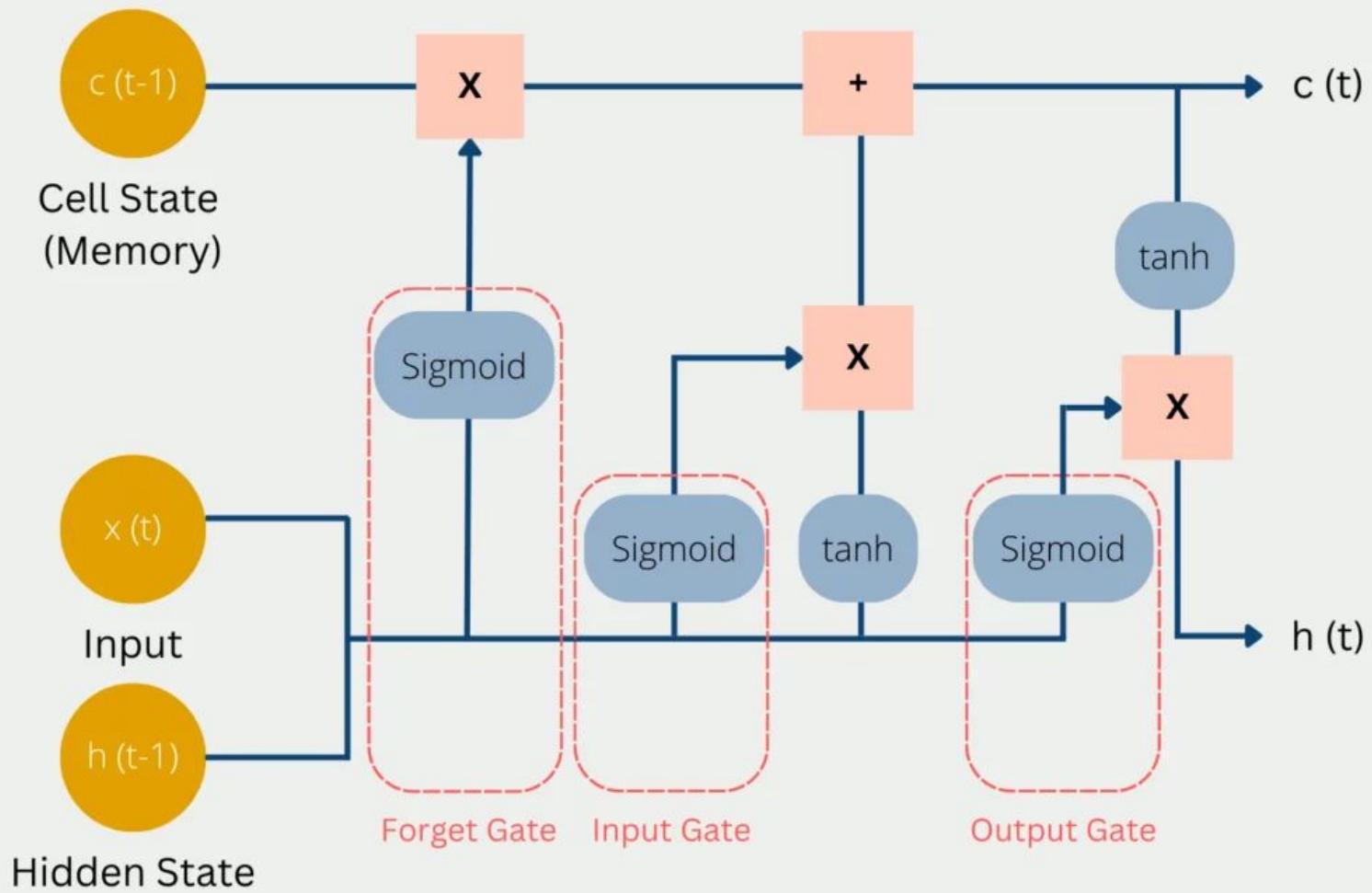




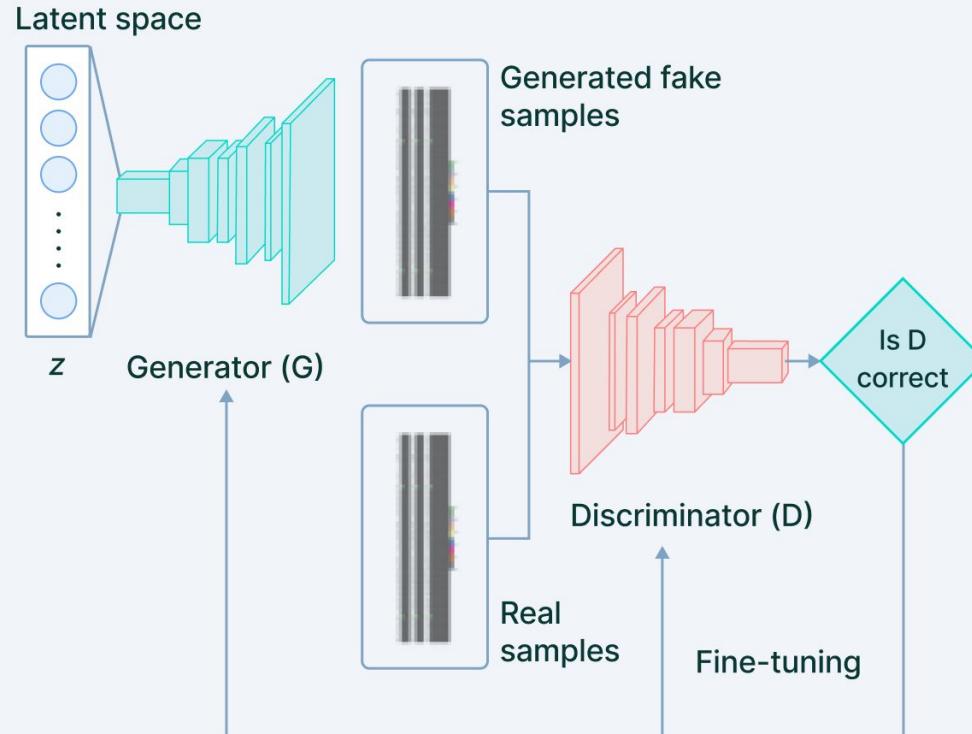


The Recurrent Neural Networks (RNN)





Generative Adversarial Networks



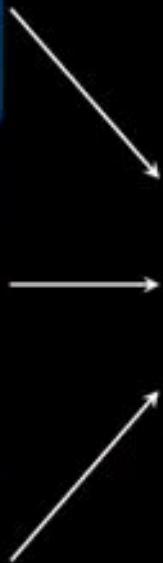
Coarse styles
 $(4^2 - 8^2)$



Middle styles
 $(16^2 - 32^2)$



Fine styles
 $(64^2 - 1024^2)$

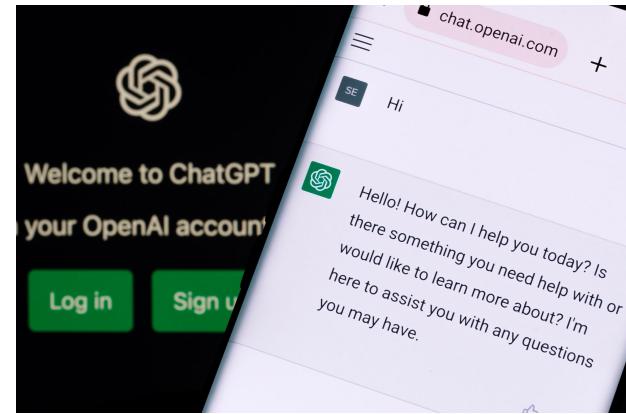




DIEP
NEP.

THIS IS NOT MORGAN FREEMAN.

<https://chat.openai.com/>



GPT-3 (Generative Pre-trained Transformer 3) uses a process called tokenization to break down text. Many words map to single tokens, though longer or more complex words often break down into multiple tokens. On average, tokens are roughly 4 characters long.

[Clear](#)[Show example](#)**Tokens****56****Characters****258**

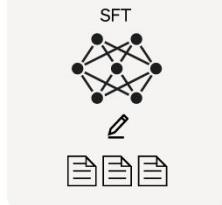
GPT-3 (Generative Pre-trained Transformer 3) uses a process called tokenization to break down text. Many words map to single tokens, though longer or more complex words often break down into multiple tokens. On average, tokens are roughly 4 characters long.

[TEXT](#)[TOKEN IDS](#)

Step 1

Collect demonstration data and train a supervised policy.

A prompt is sampled from our prompt dataset.



A labeler demonstrates the desired output behavior.

This data is used to fine-tune GPT-3.5 with supervised learning.

Step 2

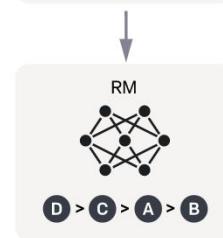
Collect comparison data and train a reward model.

A prompt and several model outputs are sampled.



A labeler ranks the outputs from best to worst.

This data is used to train our reward model.



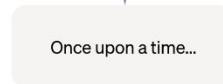
Step 3

Optimize a policy against the reward model using the PPO reinforcement learning algorithm.

A new prompt is sampled from the dataset.



The PPO model is initialized from the supervised policy.



The policy generates an output.

The reward model calculates a reward for the output.



The reward is used to update the policy using PPO.

A woman with long brown hair, wearing a black hoodie, is looking up at her smartphone with a surprised expression. A white speech bubble originates from the phone, containing the text "I'll change ur life!".

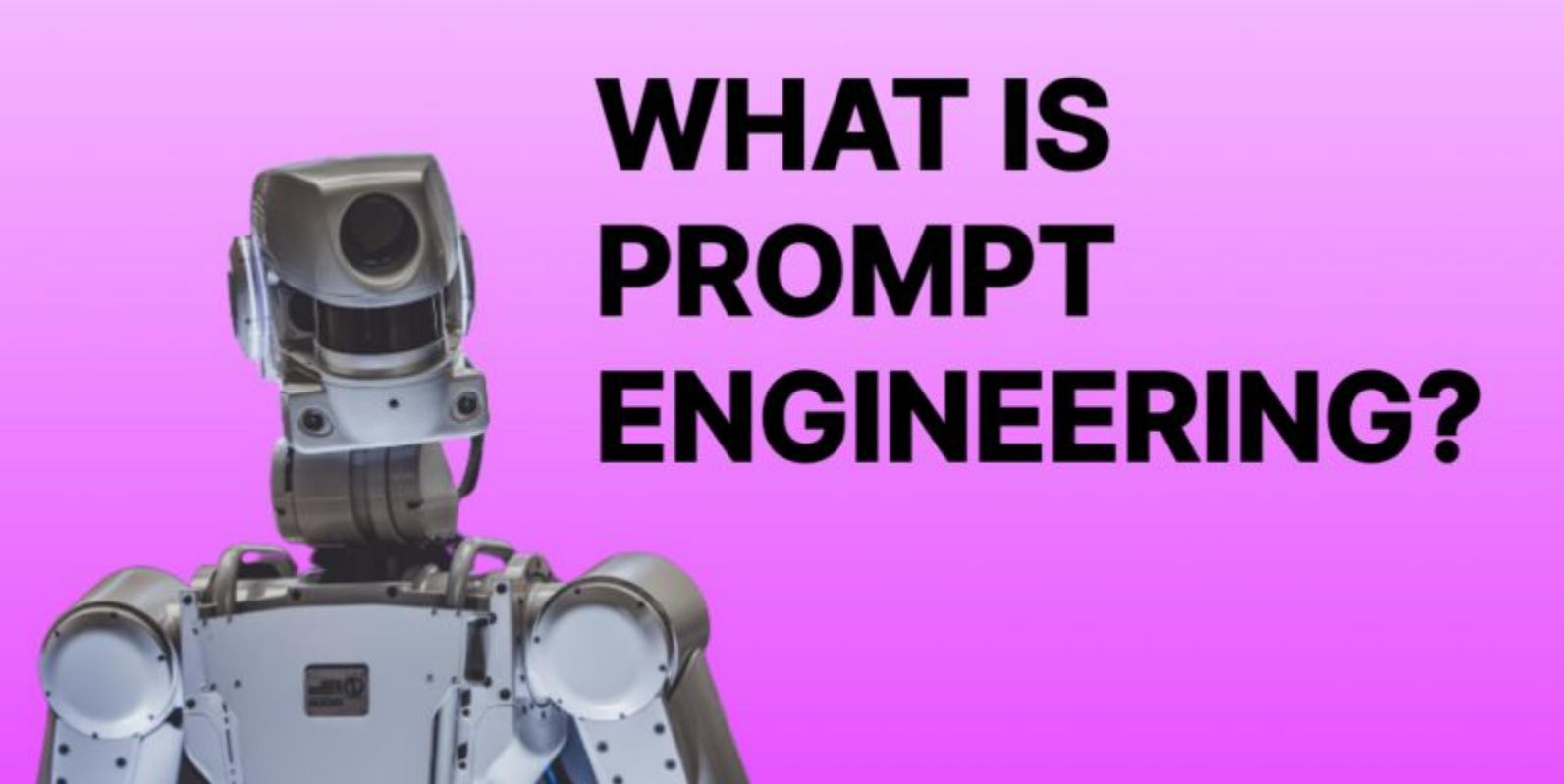
I'll change ur life!



What is GPT?

Generative Pre-trained Transformers, commonly known as GPT, are a family of neural network models that uses the transformer architecture and is a key advancement in artificial intelligence (AI) powering generative AI applications such as ChatGPT.

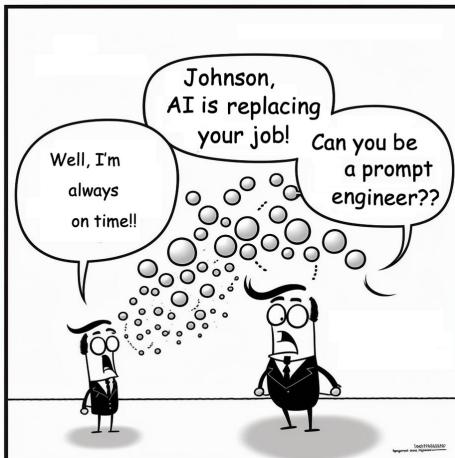
GPT models give applications the ability to create human-like text and content (images, music, and more), and answer questions in a conversational manner. Organizations across industries are using GPT models and generative AI for Q&A bots, text summarization, content generation, and search.



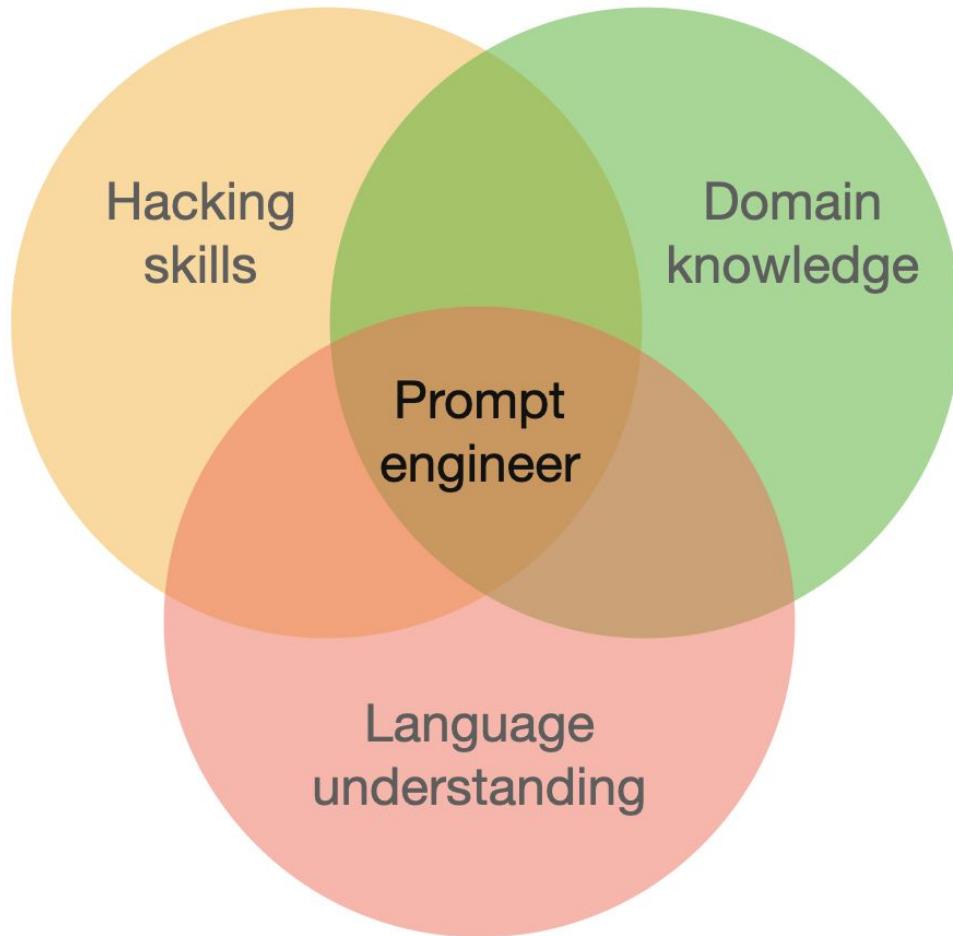
WHAT IS PROMPT ENGINEERING?

Introduction to prompt engineering

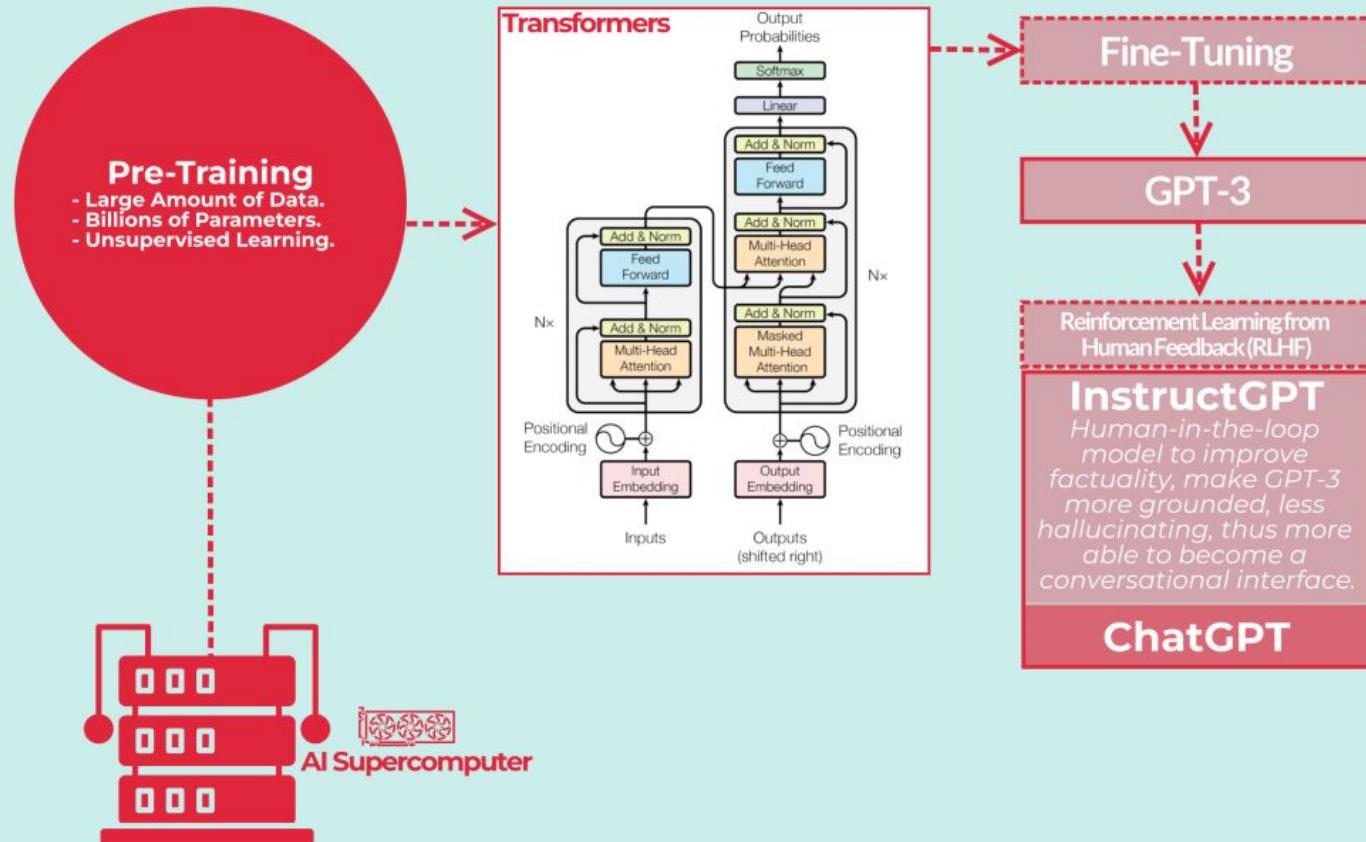
<https://learn.microsoft.com/en-us/azure/cognitive-services/openai/concepts/prompt-engineering>



Instruction Type	Prompt	Completion
Simple instructions	Write an introduction for a weekly newsletter.	Welcome to our weekly newsletter. This week we have a special focus on productivity. We'll be sharing tips and tricks on how to be productive in all areas of your life. We hope you enjoy and please share with your friends!
Complex instructions	Write an introduction for a weekly newsletter to my company named Contoso. It should mention the great all hands meeting we had last week and thank the team for all their hard work over the tough past few months.	Welcome to Contoso's weekly newsletter! We had a great all hands meeting last week and I want to thank the team for all their hard work over the tough past few months. We're looking forward to a bright future and I'm confident that with our team's dedication and commitment, we will achieve great things. Thank you for being a part of our team and I look forward to working with you all in the coming weeks and months. Sincerely, [Your name].
Complex Instructions	Write an introduction for a weekly newsletter, including the following: <ul style="list-style-type: none">- My company is named contoso- An exuberant greeting- Thanking the team for their hard work over a couple tough months- Positive outlook for the coming quarter- Signed by the SLT	Hello Contoso team! I wanted to take a moment to thank you for all your hard work over the past few months. It's been a tough couple of months, but you've all pulled together and done an amazing job. I'm really proud of you all. Looking ahead to the next quarter, I'm feeling really positive. I think we're going to achieve some great things. Thanks again, The Senior Leadership Team.



The Current AI Paradigm

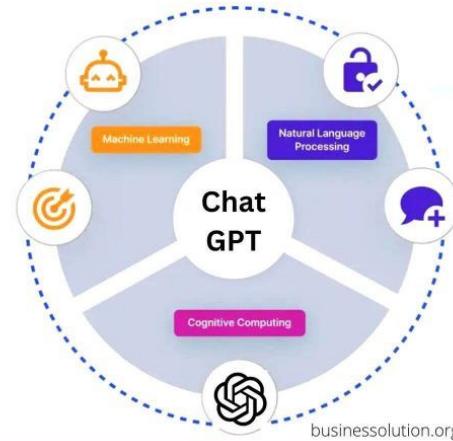


50 Chat GPT Use Cases [+Prompts]

<https://businesssolution.org/chat-gpt-use-cases/>

**+50 ChatGPT
Use Cases**

+Prompts



DALL-E 2 AI IMAGE GENERATORS



TEXT PROMPT

a teapot in the shape of an avocado. a teapot imitating an avocado.

AI-GENERATED IMAGES



DALL-E mini

DALL-E mini is an AI model that generates images from any prompt you give!

voldemort as a cat person

Run



► Bias and Limitations

DALL·E mini

AI model generating images from any prompt!

lord voldemort as the baby on the cover of nevermind by nirvana

Run



DALL·E mini

AI model generating images from any prompt!

Thanos eating pizza

Run



DALL-E Sample

Prompt: /imagine Cute thanos from AVENGERS eating pizza on a beach, anime, HD



DALL-E Sample

Prompt: an astronaut lounging in a tropical resort in space, vaporwave



Let's Play: <https://openai.com/dall-e-2>



Research ▾ Product ▾ Developers ▾ Safety Company ▾

DALL·E 2

DALL·E 2 is an AI system that can create realistic images and art from a description in natural language.

[Try DALL·E ↗](#)

[Follow on Instagram ↗](#)

GPT-4 can solve difficult problems with greater accuracy, thanks to its broader general knowledge and problem solving abilities.

Creativity Visual input Longer context

GPT-4 can accept images as inputs and generate captions, classifications, and analyses.

Input

What can I make with these ingredients?



Input

What can I make with these ingredients?



Output

There are many options for what you can make with these ingredients! Some possibilities include:

- Pancakes or waffles
- Crepes
- French toast
- Omelette or frittata
- Quiche
- Custard or pudding
- Cake or cupcakes
- Muffins or bread- Cookies or biscuits

These are just a few examples, but the possibilities are endless!

GPT-4 surpasses ChatGPT in its advanced reasoning capabilities.

ChatGPT

Input

Andrew is free from 11 am to 3 pm, Joanne is free from noon to 2 pm and then 3:30 pm to 5 pm. Hannah is available at noon for half an hour, and then 4 pm to 6 pm. What are some options for start times for a 30 minute meeting for Andrew, Hannah, and Joanne?

Output

Here is a list of their availability:

- Andrew: 11 am to 3 pm
- Joanne: noon to 2 pm, and 3:30 pm to 5 pm
- Hannah: noon to 12:30 pm, and 4 pm to 6 pm

Based on their availability, there is a 30-minute window where all three of them are available, which is from 4 pm to 4:30 pm. So, the meeting can be scheduled at 4 pm. ■

GPT-4

Input

Andrew is free from 11 am to 3 pm, Joanne is free from noon to 2 pm and then 3:30 pm to 5 pm. Hannah is available at noon for half an hour, and then 4 pm to 6 pm. What are some options for start times for a 30 minute meeting for Andrew, Hannah, and Joanne?

Output

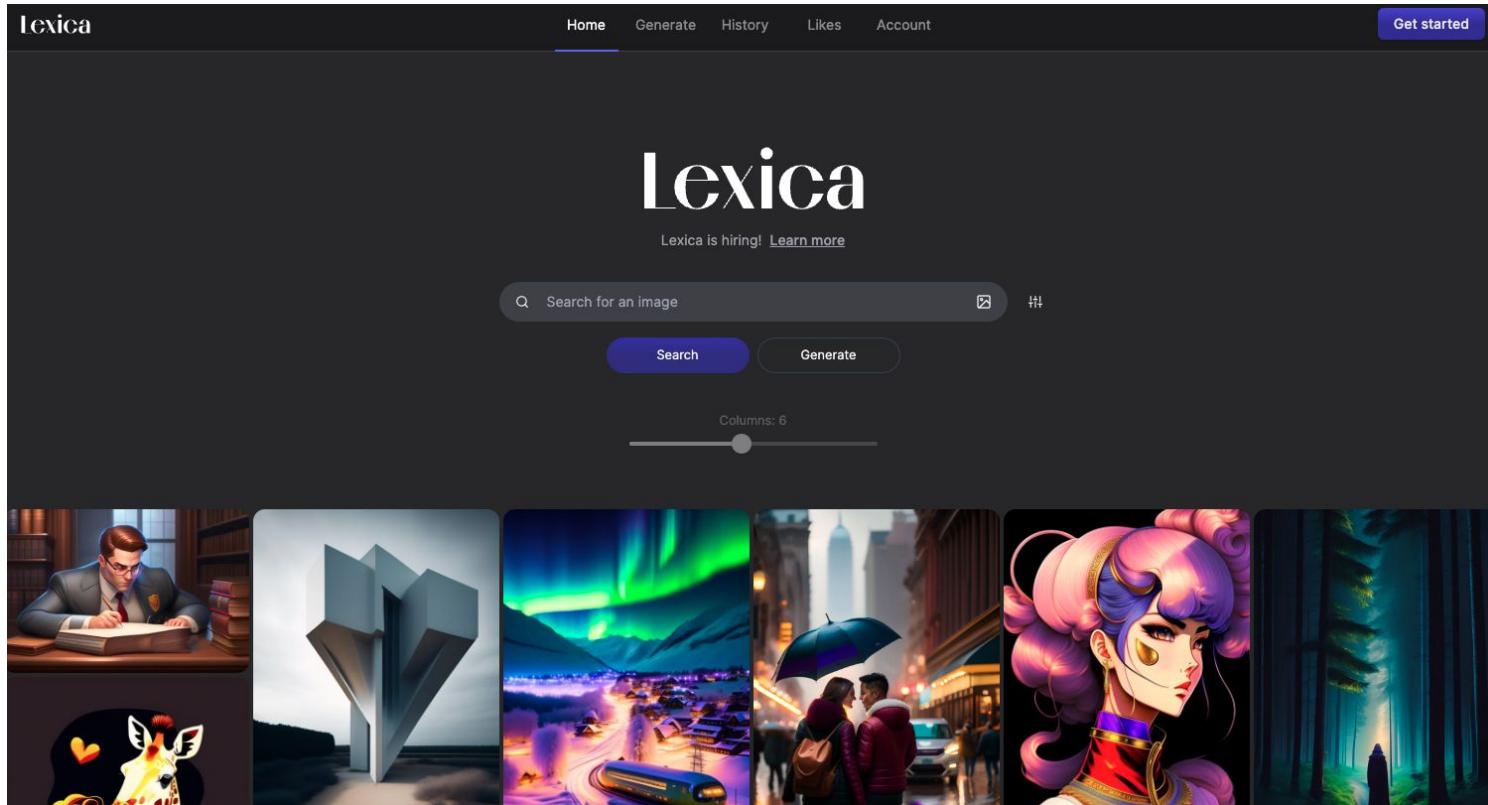
Andrew: 11 am - 3 pm

Joanne: 12 pm - 2 pm, 3:30 pm - 5 pm

Hannah: 12 pm - 12:30 pm, 4 pm - 6 pm

Common availability for a 30-minute meeting: 12 pm - 12:30 pm ■

<https://lexica.art/>



www.midjourney.com

Midjourney

Community Showcase

Help & FAQ

About & Careers

Sign In

Community Showcase



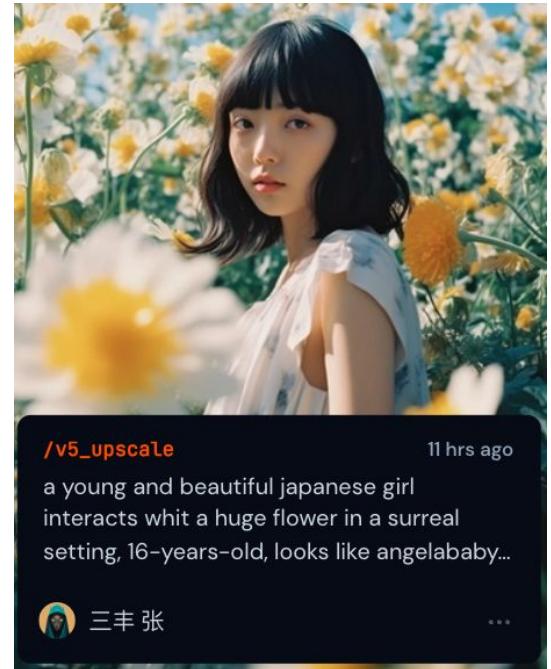
Recent Top

/v5_upscale
a young and beautiful japanese girl interacts whit a huge flower in a surreal setting, 16-years-old, looks like angelababy...

三丰张 11 hrs ago

/v5_upscale
colorized paddy field at night, mosque, in the style of graphic novel inspired illustrations, intricate landscapes...

mauripan 17 hrs ago





Describe your image

Cute thanos from AVENGERS eating pizza on a beach, anime, HD

Negative prompt

text, blurry

Generate

Dimensions + Upload image

768 x 1152

Advanced settings



Outpaint



AnyFace: Free-style Text-to-Face Synthesis and Manipulation

Jianxin Sun^{1,2*}, Qiya Deng^{1,2*}, Qi Li^{1,2†}, Muyi Sun¹, Min Ren^{1,2}, Zhenan Sun^{1,2}

¹ Center for Research on Intelligent Perception and Computing, NLPR, CASIA

² School of Artificial Intelligence, University of Chinese Academy of Sciences (UCAS)

{jianxin.sun, dengqiyao, muyi.sun, min.ren}@cripac.ia.ac.cn, {qli, znsun}@nlpr.ia.ac.cn



Source



(1)

(1) - (2)

(1) - (3)

(1) - (4)

(1) - (5)

(a) One caption vs Multi-caption

She graduated with a PhD.



He looks very knowledgeable.



(b) Open-world text descriptions

He has black hair and beard.



Source

(c) Face manipulation

Figure 1. Our AnyFace framework can be used for real-life applications. (a) Face image synthesis with optical captions. The top left is the source face. (b) Open-world face synthesis with out-of-dataset descriptions. (c) Text-guided face manipulation with continuous control. Given source images, AnyFace can manipulate faces with continuous changes. The arrow indicates the increasing relevance to the text.

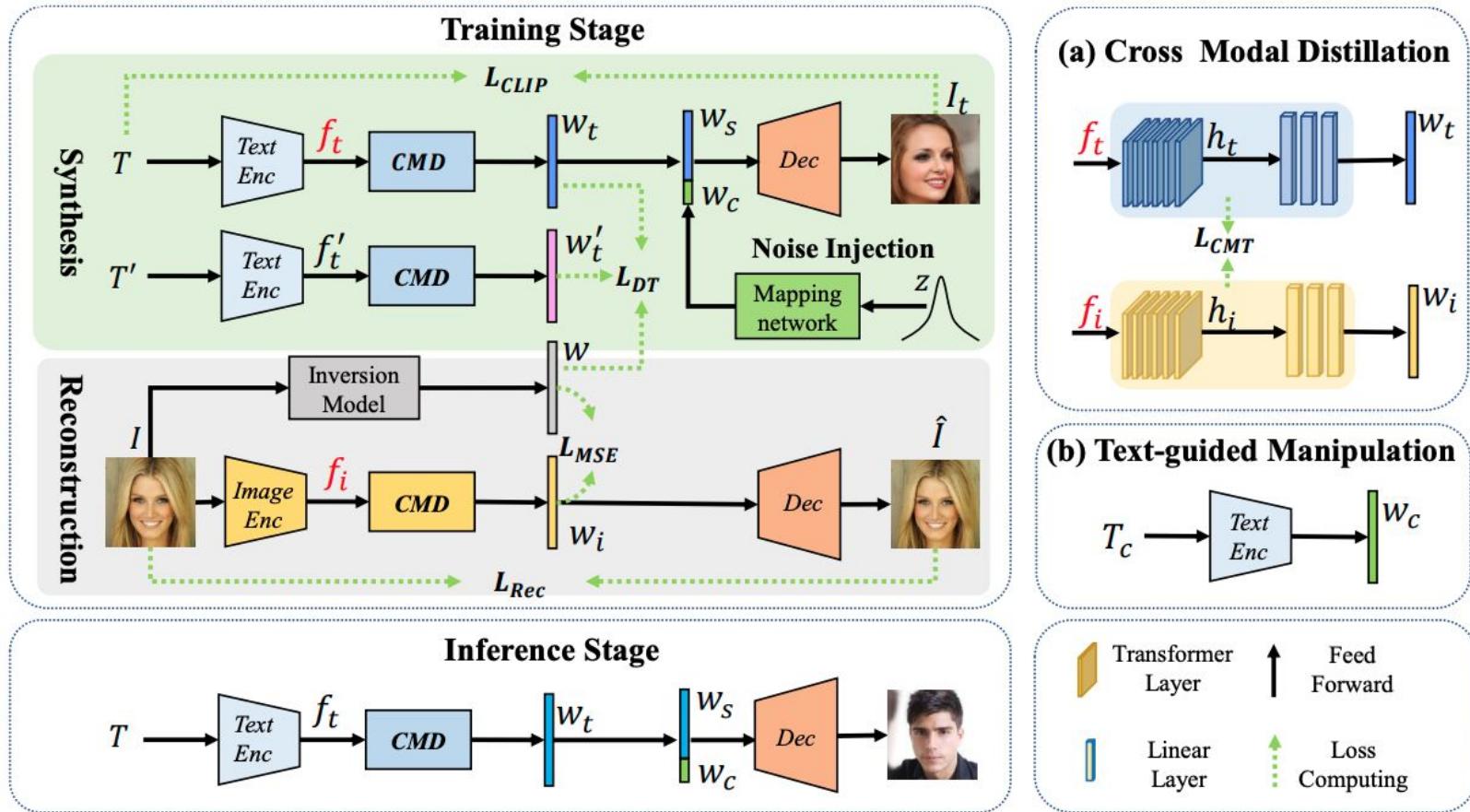


Figure 2. Overview of AnyFace. It consists of a face synthesis network and a face reconstruction network. Text Enc and Image Enc represent CLIP text and image encoder respectively. Dec is the pre-trained StyleGAN decoder. (a) Detailed architecture of Cross Modal Distillation (CMD) module. (b) For text-guided face manipulation, the content code w_c is replaced by the latent code of source image.

The person wears lipstick.
She has blond hair, and
pale skin. She is attractive.



The woman has wavy hair,
black hair, and arched eyebrows. She is young. She is wearing heavy makeup.



She is wearing lipstick. She has high cheekbones, wavy hair, bushy eyebrows, and oval face. She is attractive.



He has mouth slightly open,
wavy hair, bushy eyebrows, and oval face. He is attractive, and young. He has no beard.



AttnGAN

SEA-T2F

TediGAN-B

Ours w/o L_{DT}

Ours w/o L_{CMT}

Ours

Figure 4. Qualitative comparisons with state-of-the-art methods. The leftmost is the input sentences, columns from left to right represent the results of AttnGAN, SEA-T2F, TediGNA-B, AnyFace without \mathcal{L}_{DT} , AnyFace without \mathcal{L}_{CMT} and our AnyFace respectively.

Image Captioning: Example Results

Captions generated using [neuraltalk2](#)
All images are [CC0 Public domain](#):
[cat suitcase](#), [cat tree](#), [dog](#), [bear](#),
[surfers](#), [tennis](#), [giraffe](#), [motorcycle](#)



A cat sitting on a suitcase on the floor



A cat is sitting on a tree branch



A dog is running in the grass with a frisbee



A white teddy bear sitting in the grass



Two people walking on the beach with surfboards



A tennis player in action on the court



Two giraffes standing in a grassy field



A man riding a dirt bike on a dirt track

<https://murf.ai/>

The MURF.AI homepage features a dark blue background with a subtle pattern of white media-related icons. At the top left is the MURF.AI logo, consisting of a stylized yellow 'M' icon followed by the text 'MURF.AI'. The top right contains navigation links: 'Products ▾', 'Create ▾', 'Enterprise', 'Pricing', and 'Resources ▾'. A central headline reads 'Go from text to speech with a versatile AI voice generator' and 'AI enabled, real people's voices'. Below this, a subtext states 'Make studio-quality voice overs in minutes. Use Murf's lifelike AI voices for podcasts, videos, and all your professional presentations'. Five circular profile pictures of diverse individuals are displayed, with the fourth one featuring a yellow sound icon. A large white waveform graphic is centered below the profiles. At the bottom center is a yellow button labeled 'OPEN STUDIO'.

Products ▾ Create ▾ Enterprise Pricing Resources ▾

Go from text to speech with a versatile AI voice generator

AI enabled, real people's voices

Make studio-quality voice overs in minutes. Use Murf's lifelike AI voices for podcasts, videos, and all your professional presentations

OPEN STUDIO

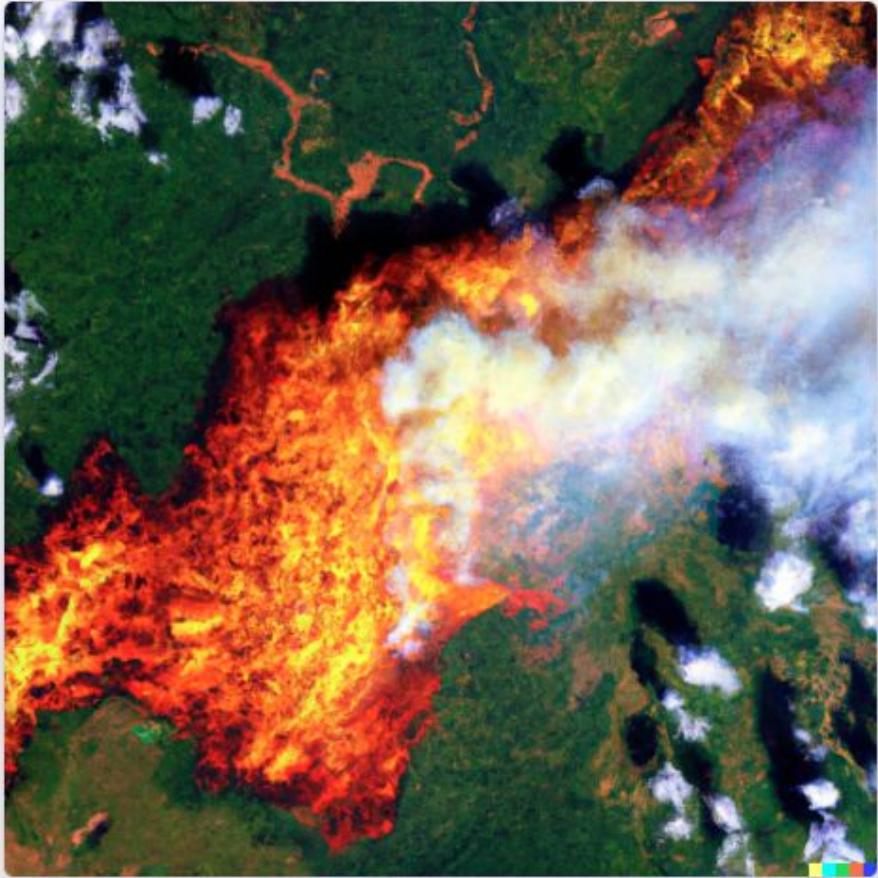
<https://labs.openai.com/s/q74WPnaXW3mldsf498zyGma>



“A satellite image showing
Amazon rainforest after
deforestation ”



Vinícius × DALL·E
Human & AI



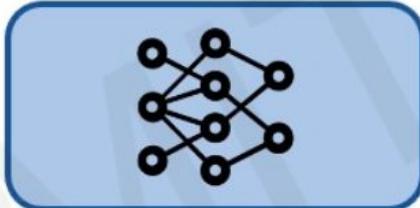
“A satellite image showing
Amazon rainforest under fire”



Vinícius × DALL·E
Human & AI

Generating Images from Natural Language

“A photo of an astronaut riding a horse.”



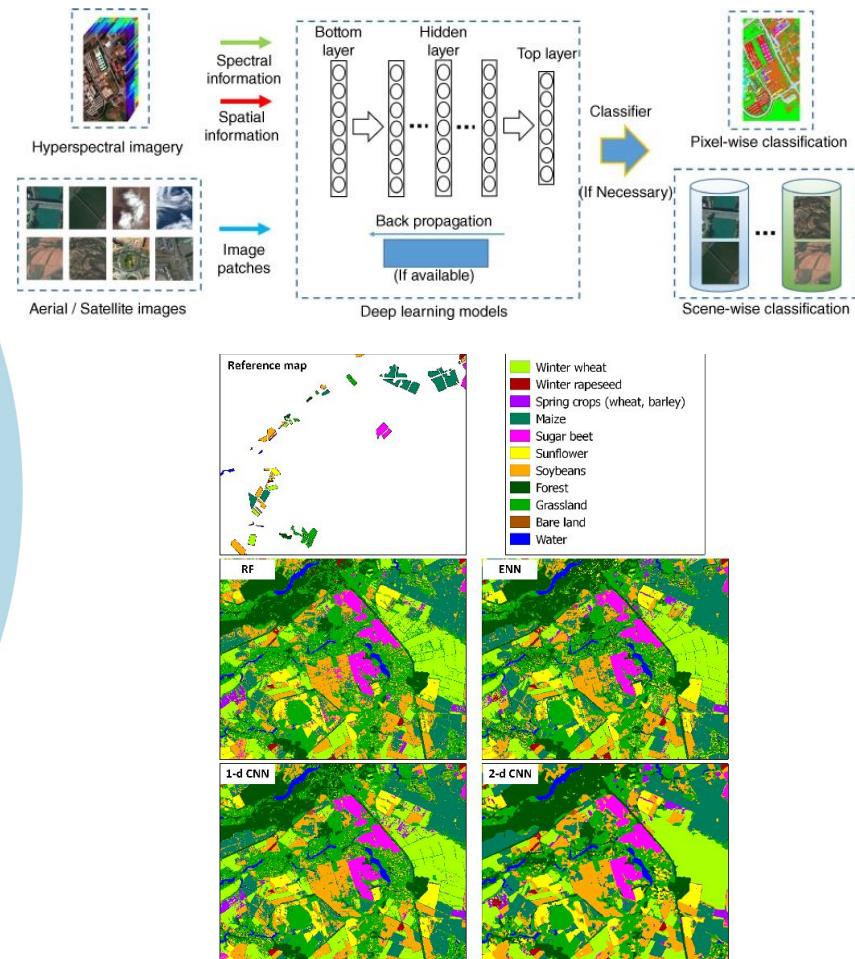
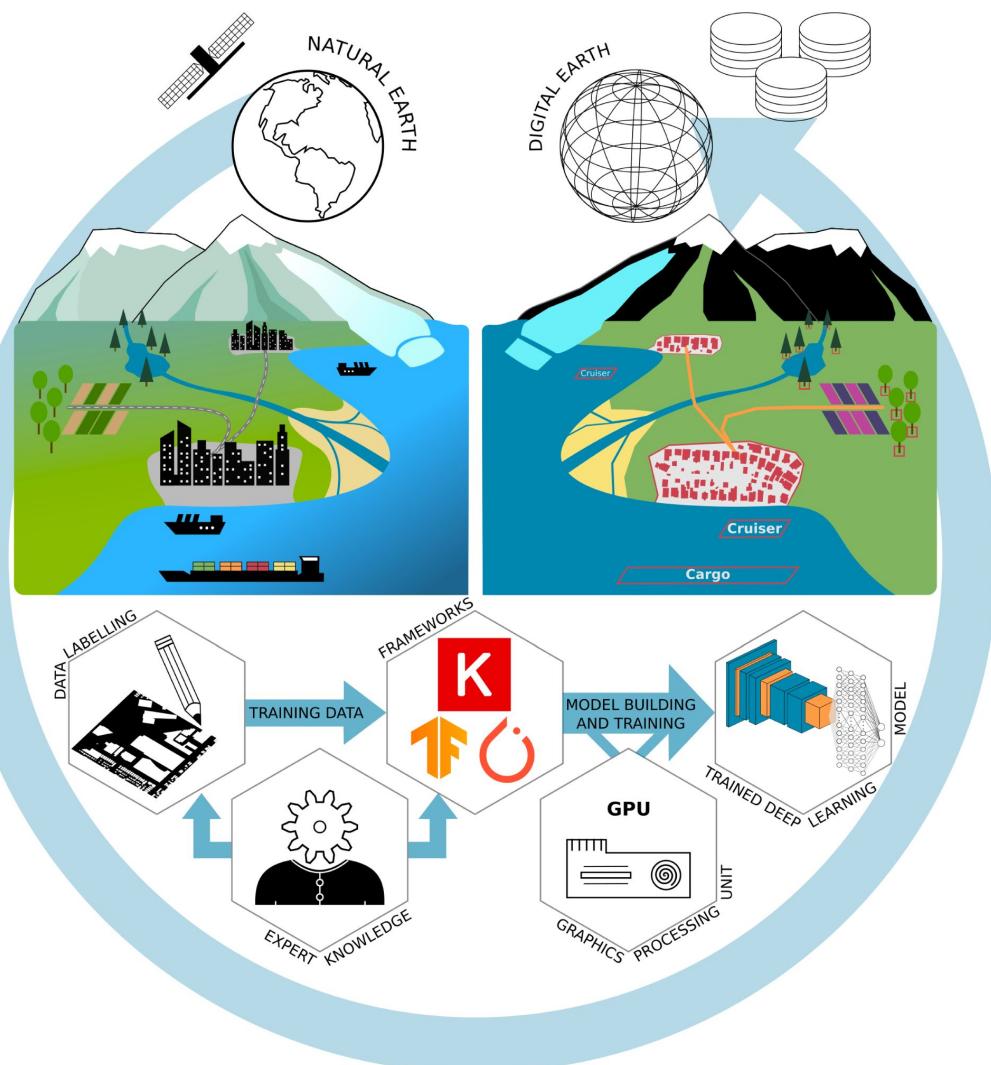
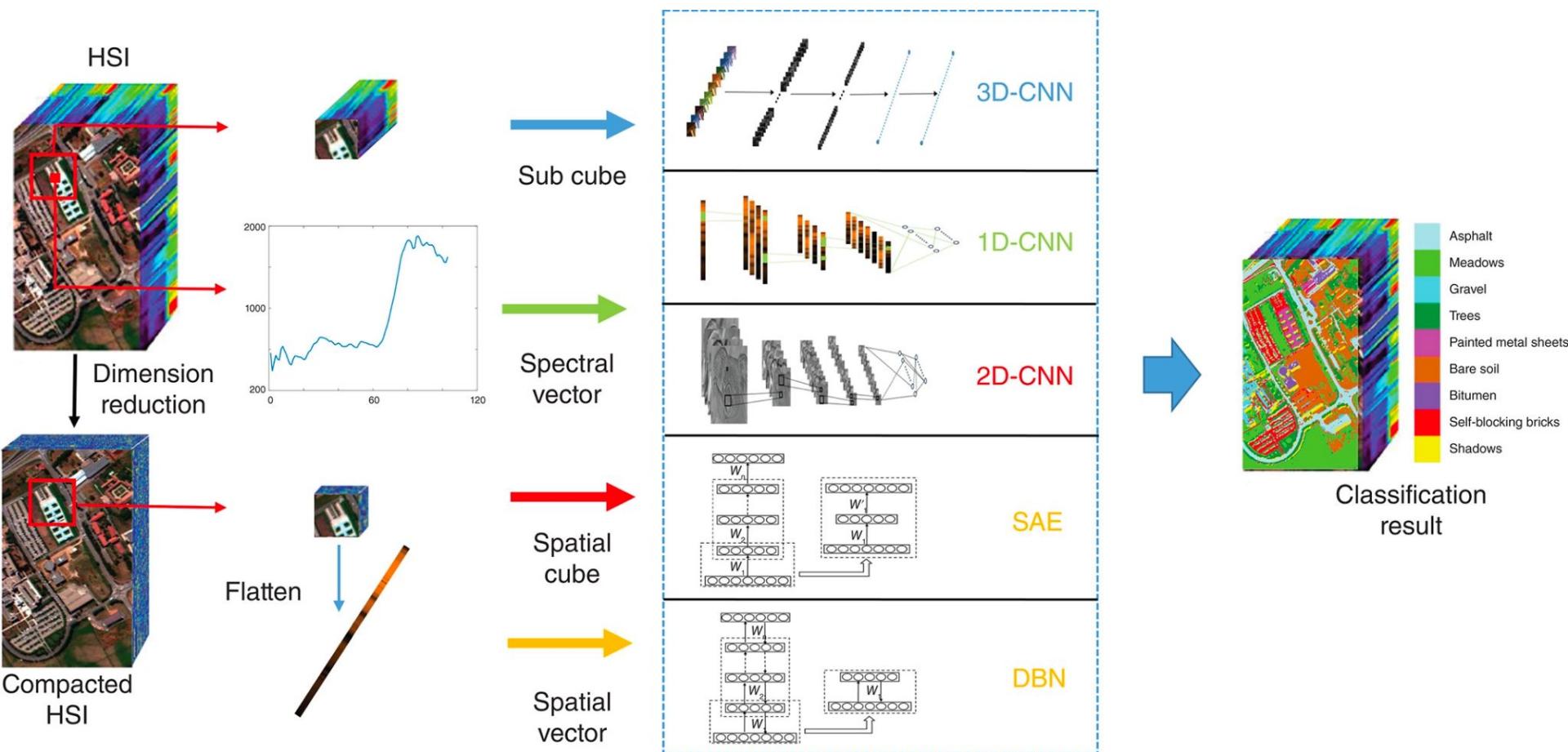


Fig. 3. Example of classification result for the Kviv region for 2015 based



Why Now?

Neural Networks date back decades, so why the resurgence?

1. Big Data

- Larger Datasets
- Easier Collection & Storage

IM⁺GENET



WIKIPEDIA
The Free Encyclopedia



2. Hardware

- Graphics Processing Units (GPUs)
- Massively Parallelizable

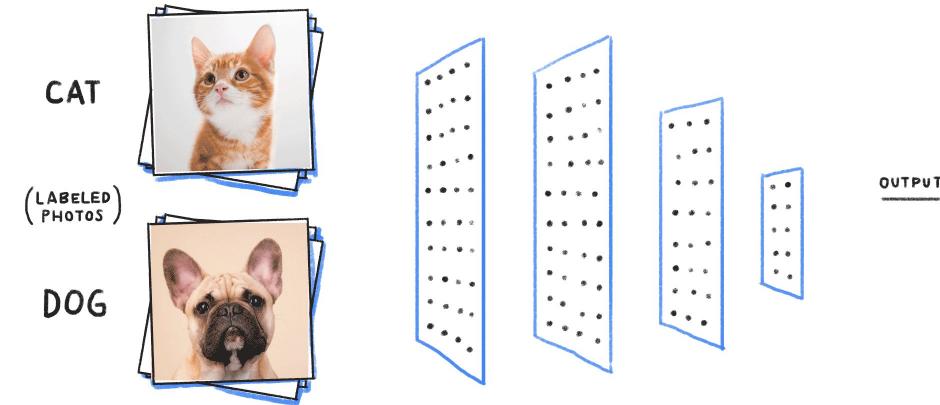


3. Software

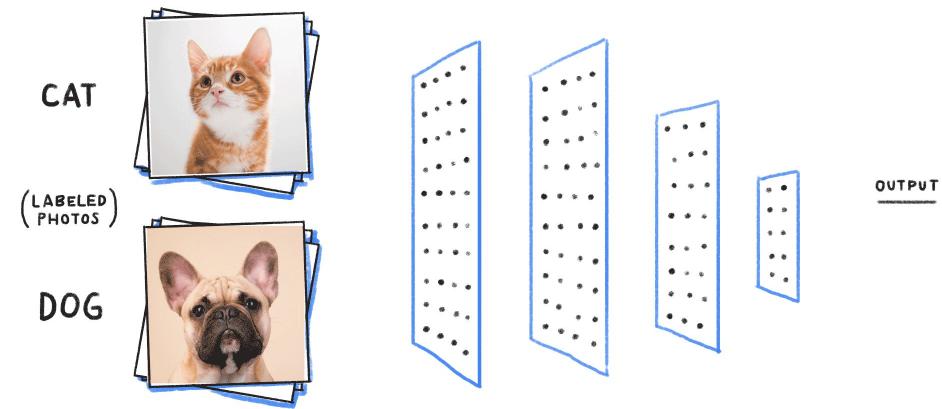
- Improved Techniques
- New Models
- Toolboxes



Machine Learning Glossary



Exercises



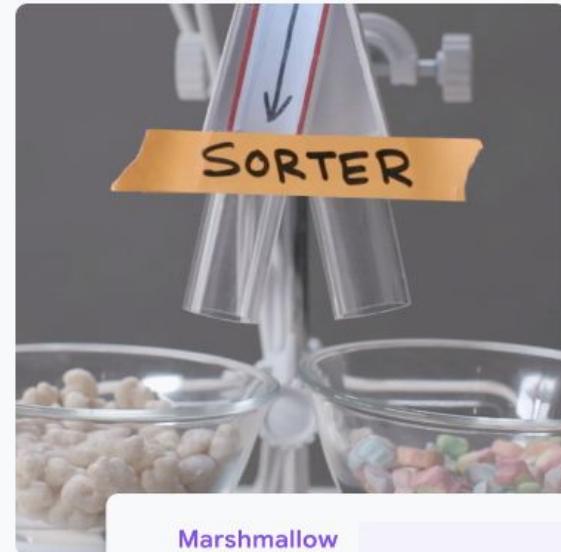
<https://teachablemachine.withgoogle.com/>

Teachable Machine

Train a computer to recognize your own images, sounds, & poses.

A fast, easy way to create machine learning models for your sites, apps, and more – no expertise or coding required.

Get Started



Marshmallow

Not Marshmallow

100%

Annual Crop 🖌

20 Image Samples

Webcam Upload



Forest 🖌

20 Image Samples

Webcam Upload



River 🖌

20 Image Samples

Webcam Upload



Add a class

Training

Model Trained

Advanced

Epochs: 50



Batch Size: 16



Learning Rate:

0.001



Reset Defaults



Under the hood



Preview

Export Model

Input ON

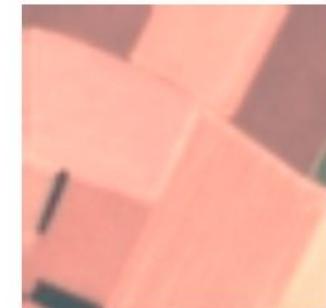
File



Choose images from your files,
or drag & drop here



Import images from
Google Drive



Output

Annual
Crop

100%

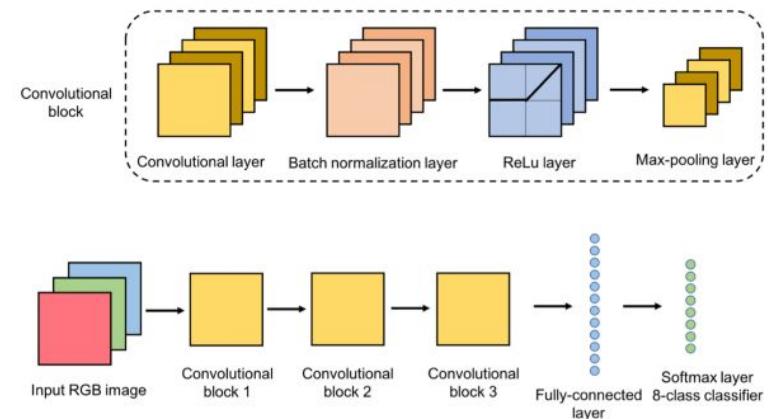
Basic Python and Pandas

Code

- Basic Python:  [Open in Colab](#)
- Basic Pandas:  [Open in Colab](#)

Deep Learning (Sample Code):

- Basic Convolution Neural Networks:  [Open in Colab](#)



Deep Learning (Sample Code):

- Sample BiLSTM with Conv1D Code (Air Pollution Data Set): [Github](#) [Notebook](#)

