

# Computer Vision OverVIEW

## (Apr 2024)

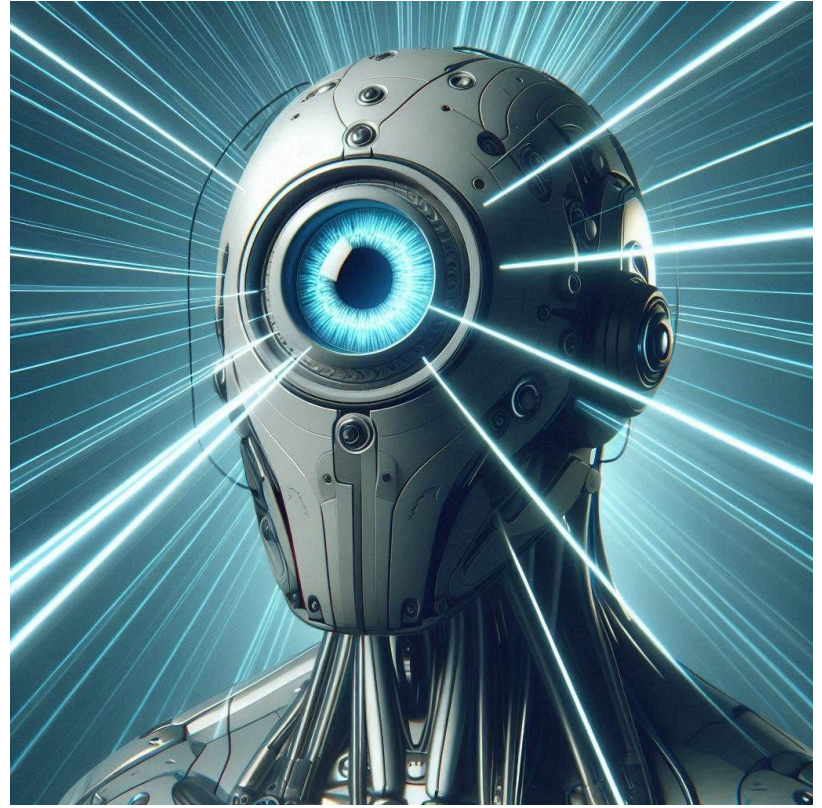
Roger Lam  
[lamroger.com](http://lamroger.com)

# Lots going on!

While not as hyped as LLMs and agents, Computer Vision has had equally impressive innovations.

Even better, applications are clear and being applied today.

We'll also review popular architectures, libraries, tutorials, and places to follow for more.



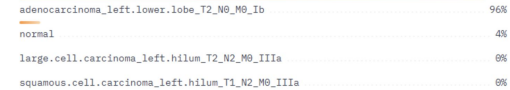
# A lotttt of applications today.

Snapchat filters, radiology anomaly detection, autonomous vehicles, robotics, content moderation, agriculture monitoring, traffic flow analysis, defect detection in manufacturing and lots more!

If you use your eyes for your job, you can probably apply CV.



adenocarcinoma\_left.lower.lobe\_T2\_N0\_M0\_Ib



Lung cancer classification model to analyze CT medical imaging



3 erosomyia\_sp



<https://viso.ai/applications/computer-vision-applications/>  
<https://intellias.com/top-computer-vision-applications-for-industries/>

# Fast.ai for quick transfer learning learning


Fast.ai has a Practical Deep Learning for Coders course - a great first step into Deep Learning and computer vision.

With a few labels, you can fine-tune a model to recognize and categorize your new classes.

Tell me which dinosaur is this?

Prehistoric dinosaurs finder using fastai library deployed on [JarvisLabs.ai](#) dashboard by Gradio

IMG ✕



OUTPUT 0.3s

**stegosaurus**


stegosaurus 100%

triceratops 0%

trex 0%

Clear Submit

Examples



# Roboflow for latest tutorials and image processing SaaS

Roboflow is a company that provides tools to manage your image data.

I used to label and augment images of coffee beans and exported really easily. Highly recommend.

They also have possibly the largest collection of CV tutorials and blog posts.

## Latest Posts



### What is Handwriting Recognition?

In this guide, we go over an overview of handwriting recognition, including the use cases, challenges, and ways of using of handwriting recognition, as well as a tutorial.

Leo Ueno | APR 12, 2024



### Using Computer Vision to Create AR Experiences

This blog post is contributed by Hart Woolery, Founder and CEO, 2020CV. Introduction I've always been fascinated by special effects in movies and video games. My career took an

Trevor Lynn | APR 10, 2024



### Estimate Bottle Orientation with Computer Vision

Learn how to estimate the orientation of a water bottle using computer vision.

Contributing Writer | APR 10, 2024

# Huggingface also has great articles - and releases!

Don't forget Huggingface!

[The State of CV at Hugging Face](#) goes over their transformers package and how it's expanded to support CV tasks with examples of one-liners to run pretrained models very easily.

They also just released Idefics2 - an open license 8B Vision-Language Model.

## Blog, Articles, and discussions

**Everything** Community Guide Open Source Collab Partnerships Research NLP Audio  
CV RL Ethics Diffusion Game Development RLHF Leaderboard Case Studies

Welcome **Idefics** to the hub!



**Introducing Idefics2: A Powerful 8B Vision-Language Model for the community**

By Leyo · April 14, 2024 · △ 20

# NeRF - Neural Radiance Fields (2020)

From images combine with a viewing direction, you can create 3D representations.

Now, faster and better.

<https://github.com/awesome-NeRF/awesome-NeRF> for more papers.



# Object Detection - YOLO, for example (2015)

YOLO (You only look once) - a single neural network that predicts bounding boxes and class probabilities from full images in one evaluation.

My coffee recognizer is on YOLOv3.

Very fun to try!



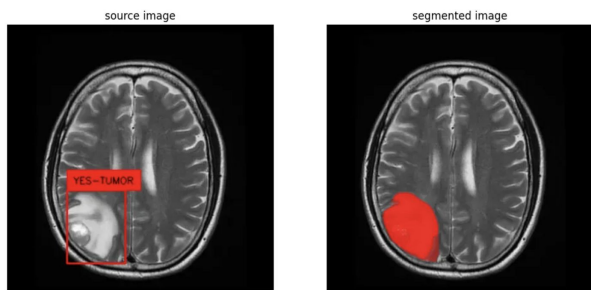
<https://blog.roboflow.com/how-to-train-yolov8-on-a-custom-dataset/>



# Image Segmentation - Segment Anything Model (2023)

SAM from Meta AI was another mind-blowing moment.

A promptable zero-shot image segmentation model



<https://blog.roboflow.com/how-to-use-segment-anything-model-sam/>

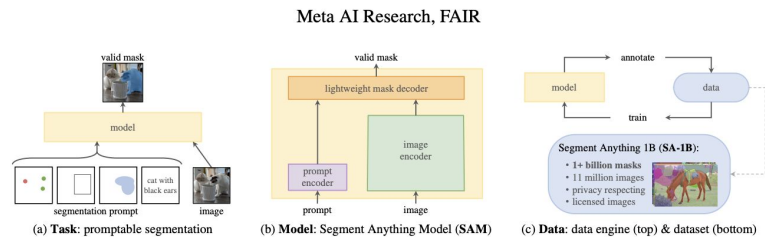


Figure 1: We aim to build a foundation model for segmentation by introducing three interconnected components: a promptable segmentation *task*, a segmentation *model* (SAM) that powers data annotation and enables zero-shot transfer to a range of tasks via prompt engineering, and a *data engine* for collecting SA-1B, our dataset of over 1 billion masks.

<https://ai.meta.com/research/publications/segment-anything/>

# Adobe integrating OpenAI, Runway, Pika

In the commercial world, we see Adobe embracing other vendor models.

Adobe has developed its own Firefly model but also playing as a platform is a smart move.

The gap between research and industry is very narrow and hard to compete against without deep investment in team and infrastructure.

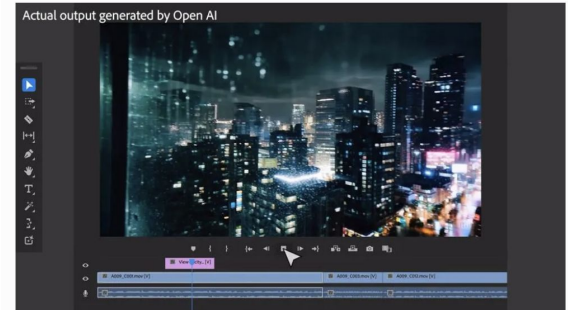
## Adobe Announces New AI Tools For Premiere Pro

**Mark Sparrow** Senior Contributor @  
*Technology journalist specializing in audio, computing and photography*

Follow



Apr 15, 2024, 09:00am EDT



Adobe's new generative AI video tools and new audio workflows in Premiere Pro will be generally ... [+] ADOBER

Adobe announced generative AI tools within its video editing software on Monday. The creative software giant said new technology in [Adobe Premiere Pro](#) will streamline editing and enable users to add or remove objects in a scene or extend an existing clip.

# Use your imagination!

Get creative! What would be fun to analyze?

Categorize your pets. Count cells in your research projects. Facial recognition for opening your smart door.

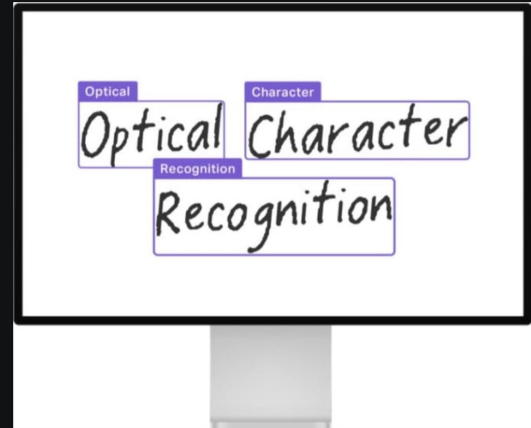
Or check out Neptune AI's [15 CV projects you can do now](#).

## 3. Text Recognition using OpenCV and Tesseract (OCR)

Here, you use OpenCV and OCR (Optical Character Recognition) on your image to identify each letter and convert them into text. It's perfect for anyone looking to take information from an image or video and turn it into text-based data. Many apps use OCR, like Google Lens, PDF Scanner, and more.

Ways to detect text from images:

- Use OpenCV – popular,
- Use Deep Learning models – the newest method,
- Use your custom model.



Source

## More info

[CS231n: Convolutional Neural Networks for Visual recognition](#)

[State of CV at Hugging Face](#)

[Sebastian Raschka on CV 2023](#)

[Nvidia Developer Blog](#)

[Latent Space Discord](#) - Helpful people like Bryan Blackbee