

Top 9 Pose Estimation Models of 2022

Ritesh Kanjee · [Follow](#)

Published in Augmented Startups · 10 min read · Apr 25, 2022

25

2



What is Pose Estimation?

“Pose estimation?”...The term pose may mean different things to different individuals, but we are not discussing Arnold Classics, Olympia, or pageant shows. So, what exactly is pose estimation? So, let’s deep dive and explore the topic.

Pose estimation has garnered immense attention in the field of computer vision. The increasing interest is the ability to use computer vision techniques to identify and track the movement of a person or an object in real-time, which offers a lot of usefulness across industries. In the ever-evolving era of advanced technologies, pose estimation can become an effective tool in sports bio mechanics, animation, gaming, robotics, medical rehabilitation, and surveillance.

Essentially, pose estimation predicts different poses based on a person's body parts and joint positioning in an image or video. For instance, we can automatically detect the joints, arms, hips, and spine position while performing a squat. Right now, some of us may wonder how it is useful? However, consider the example of an athlete rehabilitating after an injury or undergoing strength training; the pose estimation may help sports analysts analyze vital points from the starting position to the end position of a squat. As a result, these analysts can correct the postures and help prevent training injuries (figure 1).

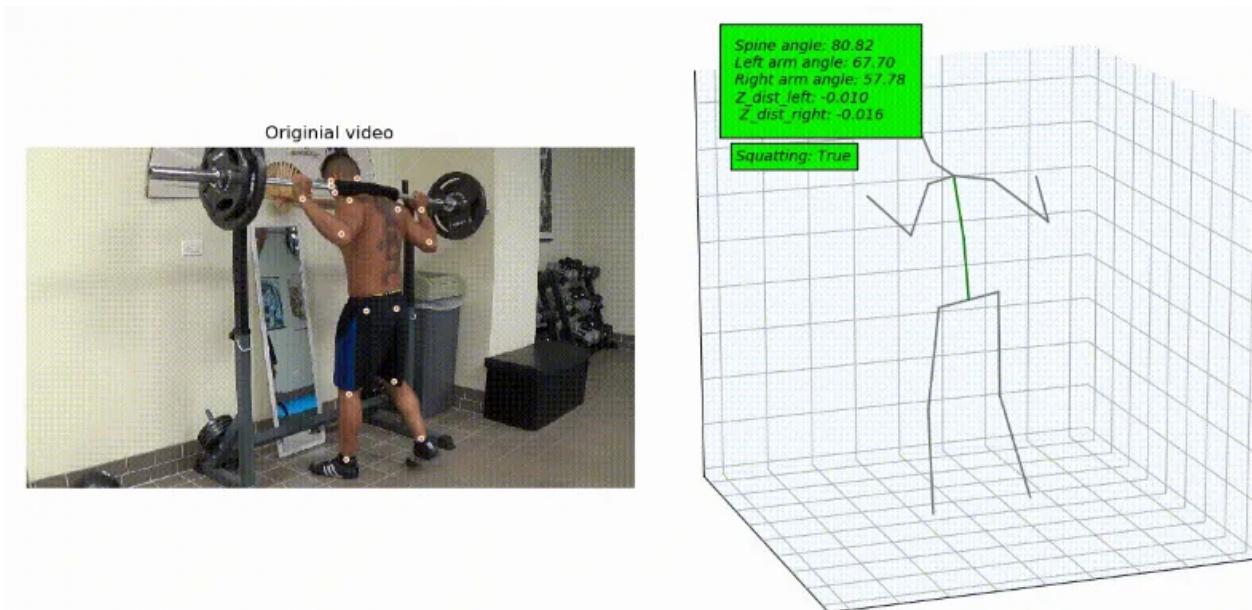


Figure.1: Human pose estimation during squatting (adapted from mobidev). [Source](#)

Essentials of Pose Estimation

Pose estimation for detecting human figures or objects from images and videos. However, we should be aware that there are different categories of pose estimation.

When working with humans, pose analysis is performed by determining various body joints. For example, it can be determined by how someone's elbow is positioned or the knee joint position. This form of pose detection falls under the category of human pose estimation. The pose estimation models are provided input in well-processed images or videos. The model offers an output about different key points based on the information in the input image. In general, the key points are provided with an ID and a confidence score determining the probability that a key point exists in a particular position of the given input. Now, if we recall the previous image of the athlete performing a squat, we can assign various IDs such as:

ID	Body Part
0	Left elbow
1	Right elbow
2	Left Wrist
3	Right Wrist
4	Left Shoulder
5	Right Shoulder
6	Spine
7	Left Hip
8	Right Hip
9	Left Knee
10	Right Knee
11	Left Ankle
12	Right Ankle

On the contrary, unlike human subjects, pose estimation can be performed for objects that are mainly rigid; therefore, they fall into the category of rigid pose estimation.

2D and 3D Aspects of Pose Estimation

Pose estimation can be performed in two ways, namely 2D and 3D. Perhaps, some of us relate these concepts of 2D and 3D to the field of animation. However, 2D aspects of pose estimation are associated with predicting key points from the images based on the pixel values. Therefore, most 2D human pose estimation techniques implement feature extraction methods to provide the appropriate key points of the human body.

Similarly, 3D pose estimation is associated with predicting the spatial positions of a specific person or object from images and videos. With the advent of deep learning, these models have significantly improved their performance, but they are more complex to work with as the datasets need to be curated with appropriate 3D structural information of the human body, including the background and the lighting conditions. Besides, there are new approaches for single pose and multi-pose estimation related to detecting one person or object or tracking multiple people and objects, respectively.

Pose Estimation Models

Various researchers have proposed different pose estimation models. Before we dive in, it is essential to understand that human pose estimation models are basically of three types:

- a) kinematic,
- b) planar, and
- c) volumetric.

The Kinematic models can be used for both 2D and 3D pose estimation. Essentially, this model focuses on the different joint and limb positions to provide the structural information of the human body. Therefore, such

models effectively identify various relations between human body parts. However, kinematic models have few limitations when representing texture or shape-based information. Next, we discuss the planar model that emphasizes 2D pose estimation. Ideally, the human body parts are represented using rectangles to provide approximate body contour. Finally, the volumetric pose estimation model focuses on 3D pose estimation. These are end-to-end deep learning models trained with complex datasets comprising high-resolution data of full-body scans to derive human body mesh of various shapes and poses.

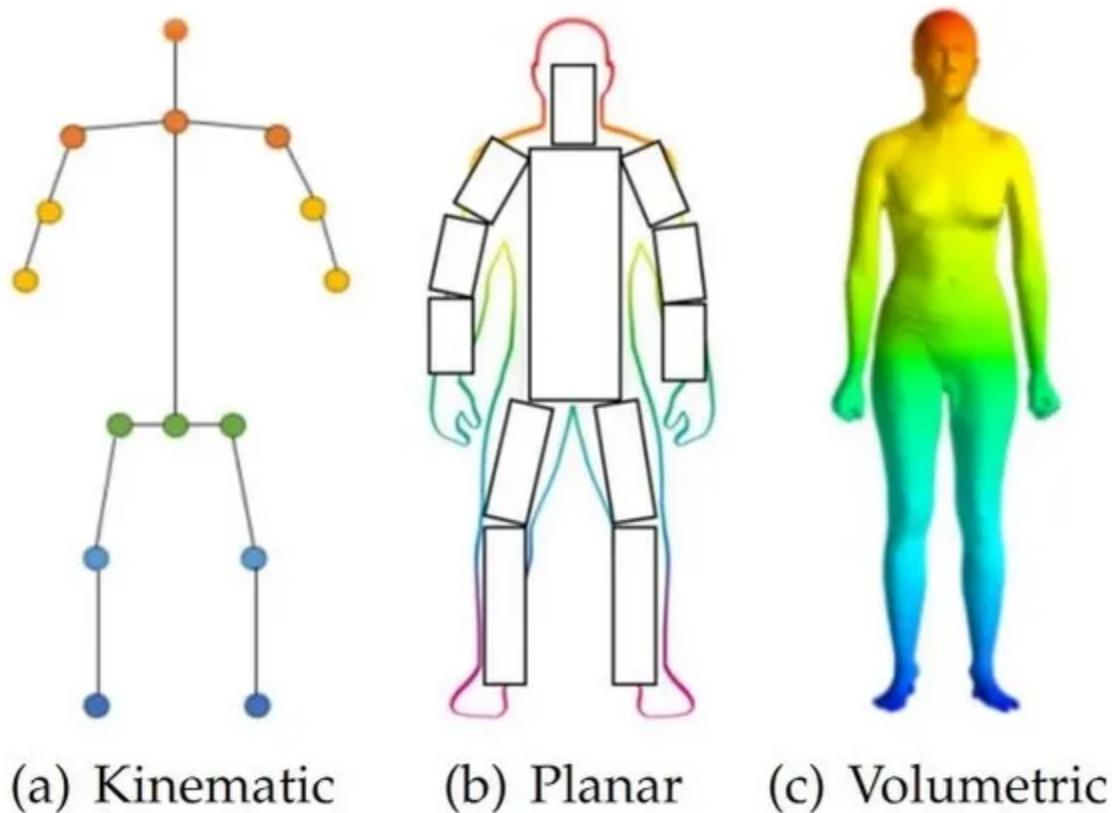


Figure.2: Different Types of Pose Estimation Models. [Source](#)

Although it is impossible to cover a wide range of models, we will discuss some of the most reliable and robust models proposed with different methods in recent years.

Let's check out some of the popular pose estimation that are being used in 2022.

1. OpenPose

The OpenPose is the first real-time pose estimation model developed at Carnegie Mellon University. The model mainly focuses on detecting key points of the human body such as the hand, facial, and foot of multiple people in a real-time scenario. In general, the image is processed with the help of a Convolutional Neural Network (CNN) to generate feature maps of the specific input. Further, the feature map is processed through different stages of the CNN pipeline to achieve the confidence map and affinity field.

For more insights, you can find additional information on their [GitHub repository](#).

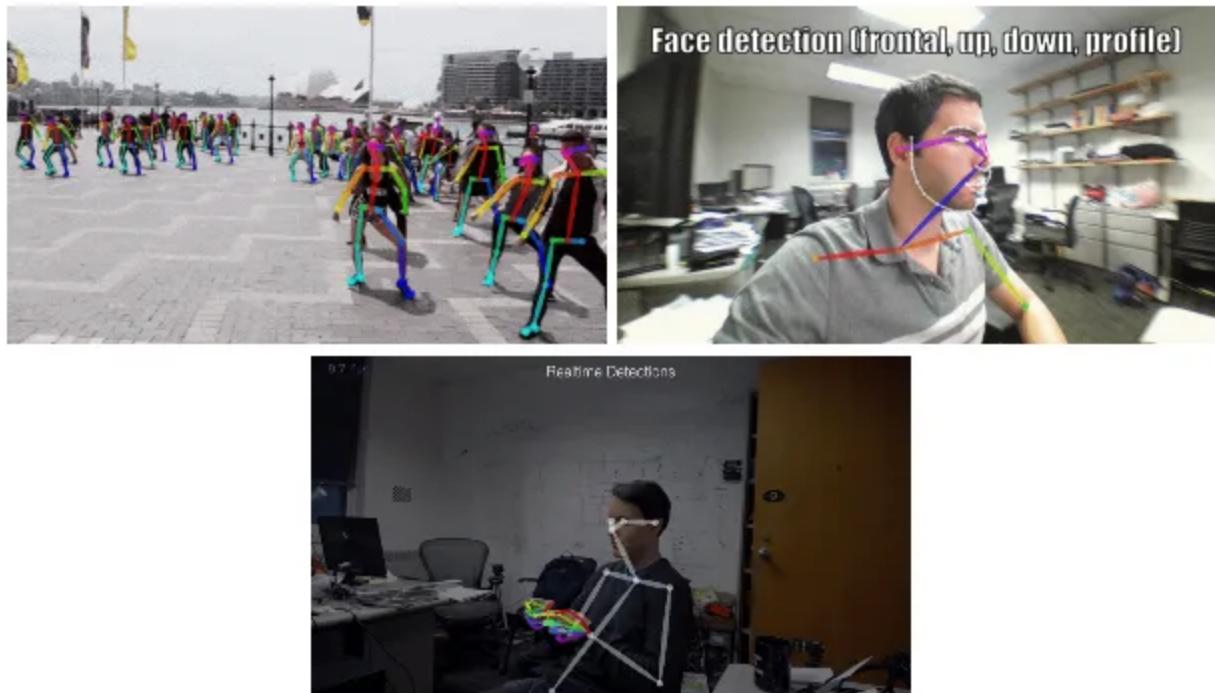


Figure.3: OpenPose Testing Results. [Source](#)

2. MoveNet

MoveNet is developed by Google research using TensorFlow.js. The researchers claim this model to be ultra-fast and highly accurate, capable of detecting 17 critical key points of the human body. However, the model has two versions, namely the Lightning, targeted for applications with low latency requirements. On the other hand, the Thunder version is designed for applications that focus on achieving higher accuracy. Additionally, both the models are capable of real-time detection and have proven effective for detecting live fitness, sports, or healthcare-based applications.



Figure.4: MoveNet Pose Detection. [Source](#)

3. PoseNet

PoseNet is yet another popular pose detection model. This model can detect poses in real-time and works efficiently for single and multi-pose detection of human beings. The PoseNet is a deep learning model that uses TensorFlow to detect different body parts and provides comprehensive

skeletal information by joining other key points. Moreover, there are 17 key points provided by PoseNet for various parts from the eyes to the ankles of a human body. A confidence score is generated to determine how accurately the model has recognized a specific key point from the image to identify the model's accuracy. All the information about the tests and configuration are accessible via [GitHub](#).



Figure.4: PoseNet-17 Key Points. [Source](#)

4. DCPose

DCPose stands for Deep Dual Consecutive Network, developed to detect human pose from multiple frames. The framework leverages deep learning techniques to overcome critical challenges in multi-frame human pose estimation, such as motion blur, defocused video, and occlusions occurring due to the dependency on each video frame. In addition, various temporal references are provided between these video frames to facilitate accurate keypoint detection. Further, the temporal merger acts as an encoder to enable broader searching scope, while a residual fusion module is responsible for computing the residuals in different directions. ([GitHub](#))



Figure.5: DCPose Pose Estimation Results. [Source](#)

5. DensePose

DensePose is a human pose estimator that aims to map various human-based pixels from an RGB image regarding the 3D surface of a human body. This model can be implemented for single pose and multi-pose estimation necessities. DensePose uses ground truth in the form of a large-scale dataset comprising image-to-surface annotated information. Besides, a Recurrent Neural Network (RCNN) is proposed that is capable of regressing different body part-related UV coordinates among each human subject at multiple frames per second. ([GitHub](#))



Figure 6: DensePose: Pose Estimation from RGB images. [Source](#)

6. HigherHRNet

[HigherHRNet](#) is a popular bottom-up pose estimation model proposed to address some challenges in predicting the correct poses of shorter people because of scaling differences. Feature pyramids are integral components that allow the proposed method to learn from scale-aware representations that help estimate precise key points to determine variations in pose estimation for a shorter person. The feature pyramids mainly comprise feature map outputs generated by the HRNet model, including high-resolution outputs produced by a transposed convolution. Additionally, the authors have revealed that the model has outperformed some of the existing bottom-up methods by up to 2.5% AP for medium-sized people. Besides, the

model also performs effectively when estimating poses from a crowded scene. ([GitHub](#))

7. Lightweight OpenPose

Lightweight OpenPose is an optimized version of the OpenPose approach that focuses on real-time inference without dropping much on the accuracy aspects of the model. The model can detect human poses from each person in the image through different key points. The authors claim that the model achieved 40% AP for the single-scale inference without any post-processing involved. ([GitHub](#))

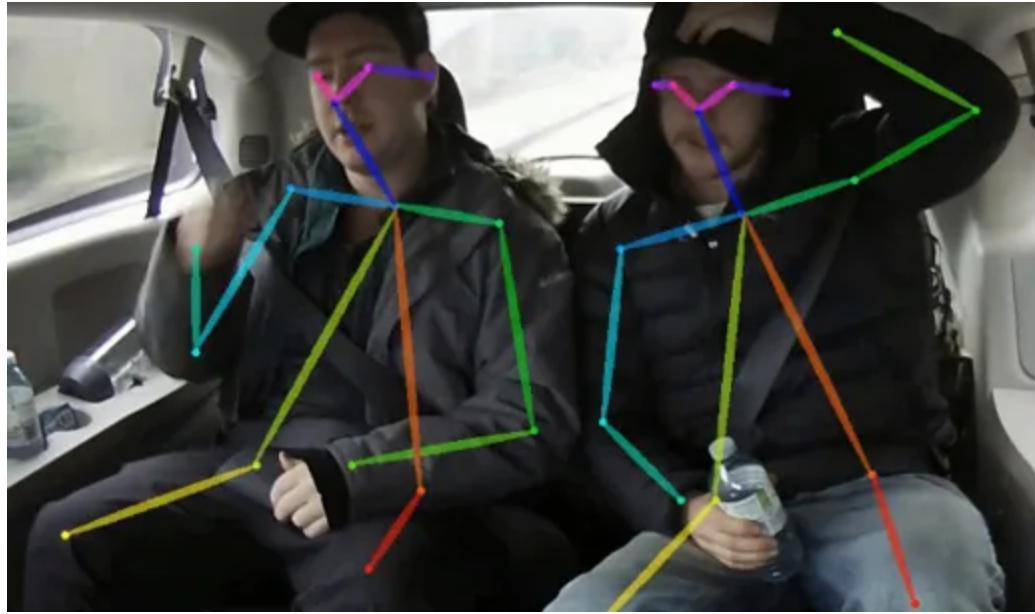


Figure 7: Lightweight OpenPose Results. [Source](#)

8. AlphaPose

AlphaPose is an exciting offering for pose estimation. Whether you are trying to detect multiple individuals in a shopping street, a flash mob, or street performers, it is now possible with the help of this model. In addition, the AlphaPose estimator is the first open-source offering to achieve over 70 mAP and 80mAP on the COCO dataset and MPII dataset, respectively. Ideally,

the model can match the poses of a person across different frames and is significantly capable of performing well as an online pose tracker. ([GitHub](#))



Figure.8: AlphaPose Tracking Results. [Source](#)

9. TransPose

TransPose is a pose estimation model that implements a CNN-based feature extraction approach, a transformer encoder, and prediction capabilities. The model has built-in features like a transformer that can capture information from long-range spatial relationships between different key points. The final output provides critical information about the predicted key points' location and the various dependencies they rely on. Besides, the model has produced excellent results of 75.8AP on the COCO dataset and has achieved superior performance with appropriate transfers per the MPII benchmark. [GitHub](#)



Figure.9: TransPose Estimation Results. [Source](#)

Conclusion

Pose detection is an ever-evolving area of research in the computer vision field. From providing real-life applications to applications running on servers in the cloud, pose estimation has immensely gained traction in the industry. In fact, advanced pose estimation models are faster and smaller to be effective on mobile devices, which provides ample opportunities.

These models can be effective for sports analysts in real-time and even reliable for medical rehabilitation, personal trainers, and realistic gaming. Although various applications have been developed, each new model aims at improving some of the limitations of the previous models. However, with

Open in app ↗

[Sign up](#) [Sign In](#)

 Medium

 Search

 Write



enable the possibility of implementing state-of-the-art pose detection applications effectively across industries.

Until a few years ago, very few of us were aware of computer vision capabilities because creating AI models was complicated. Considering the AI industry now, we are witnessing an AI revolution with tremendous capabilities, and many more job opportunities are emerging in the marketplace. So, perhaps, it is right to say that the best time for investing in a career in AI is now when the industry is rapidly growing.

So how do aspirants upskill themselves? We know there are multiple online courses, but selecting from so many may seem confusing. So are you excited about computer vision and its pose estimation capabilities? Do you want to be out there among the top computer vision experts? If you are reading this, chances are you are looking for what we offer? So why not use the most comprehensive courses to pass on the knowledge from one enthusiast to another? Check out our [computer vision courses](#) and multiple hands-on projects to level up your skills.



With several years of experience under the belt, you can be assured to master the skills of computer vision, irrespective of being a beginner or an advanced learner. We have a wide range of courses covering state-of-the-art models to solve various in-demand problems in the industry, such as lane detection, car velocity calculation, object size measurement, automatic number plate recognition, and more. Now, you can enroll and fulfill the journey towards becoming a professional computer vision engineer. We have got you covered.

Computer Vision

Opencv

Pose Estimation

Deep Learning

Machine Learning



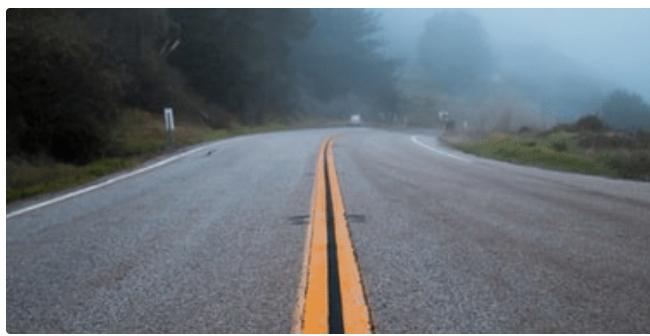
Written by Ritesh Kanjee

1.4K Followers · Editor for Augmented Startups

[Follow](#)


CEO Augmented Startups — M(Eng) Electronic Engineer, YouTuber 100'000+ Subscribers.

More from Ritesh Kanjee and Augmented Startups

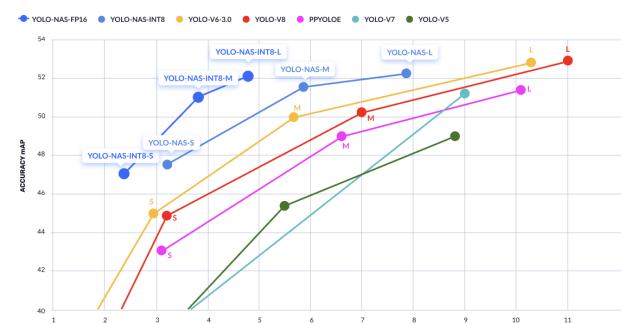


 Ritesh Kanjee in Augmented Startups

Top 10 AI Projects—Supercharge your Résumé with in July 2023 !

What if I told you that the secret to standing out in your job search or freelance bids lies i...

6 min read · Jul 17



 Ritesh Kanjee in Augmented Startups

How YOLO-NAS is Leaving YOLOv8 in the Dust—And Why You Need t...

Ritz here. You know, I've been around the block a time or two when it comes to workin...

6 min read · May 3

 992 18 + 347 4 + Aditya Singh in Augmented Startups

Top 5 Object Tracking Methods

Object tracking aims at estimating bounding boxes and the identities of objects in videos. ...

8 min read · Nov 8, 2021

 30 + Ritesh Kanjee in Augmented Startups

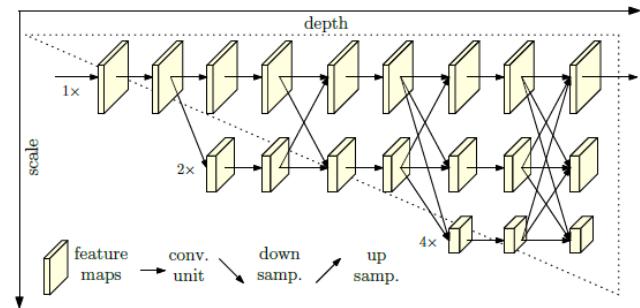
Rise of YOLO-NAS: Discover the 8 Features That Propel It to Object...

Ladies and gentlemen, boys and girls, welcome to the world of YOLO-NAS, where...

7 min read · May 5

 253 +[See all from Ritesh Kanjee](#)[See all from Augmented Startups](#)

Recommended from Medium



Asad iqbal

Computer Vision Top Projects

Computer Vision has been undergoing rapid advancements in the market, increasingly...

4 min read · Sep 15

223 2



4 min read · May 7

4

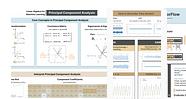


Lists



Predictive Modeling w/ Python

20 stories · 557 saves



Practical Guides to Machine Learning

10 stories · 634 saves



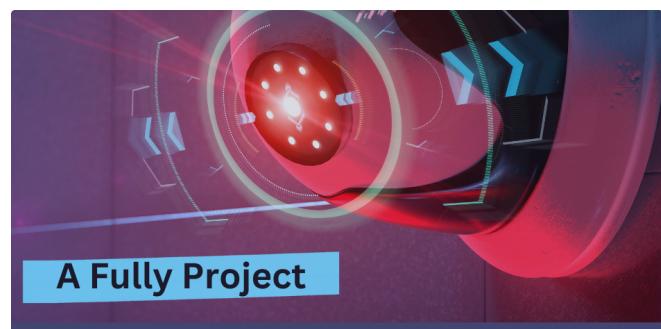
Natural Language Processing

787 stories · 363 saves



The New Chatbots: ChatGPT, Bard, and Beyond

12 stories · 178 saves





Ali Mustofa

YoloV8 Pose Estimation and Pose Keypoint Classification using...

Yolov8 Pose estimation is a task that involves identifying the location of specific points in a...

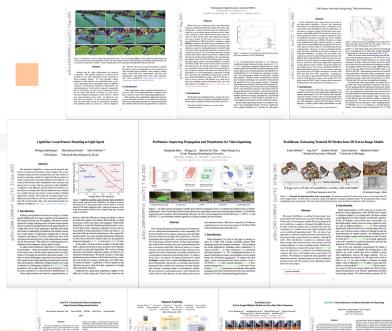
◆ · 5 min read · Jun 30

 79
 



ICCV 2023 Survival Guide

10 papers you won't want to miss




Jacob Marks, Ph.D. in Voxel51

ICCV 2023 Survival Guide: 10 Papers You Won't Want to Miss

10 Computer Vision Papers You Won't Want to Miss

14 min read · Sep 27

 549
 


[See more recommendations](#)



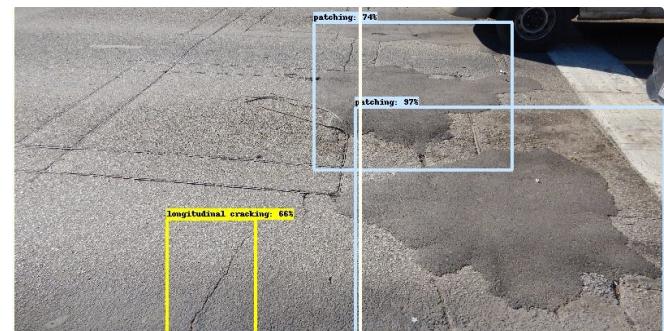
Gal Ben-Evgi

Building a Camera Recorder with Human Detection using Python a...

Build a Camera Recorder with Human Detection using Python and OpenCV:...

6 min read · Aug 1

 66
 




Prashant Nair

Road Pavement Pothole & Crack Object Detection Using Faster R...

Making Deep Learning is exciting !!

11 min read · May 21

 19
 
