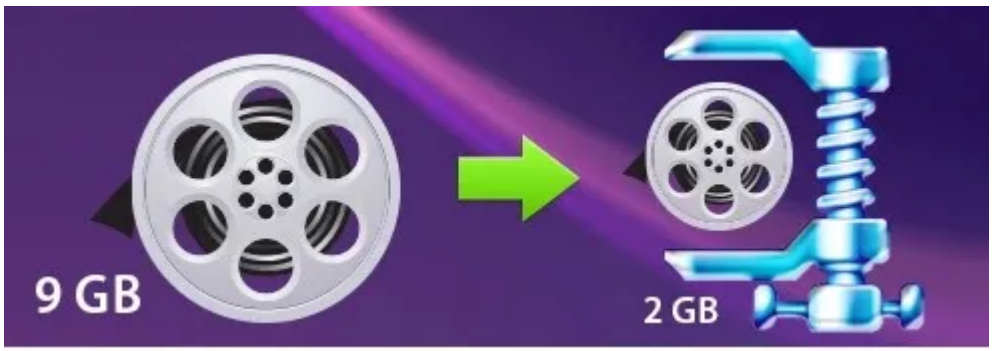
The concept of video compression

ali abdari · Follow
3 min read · Jun 22, 2019

compressed videos are everywhere on our computer or cellphone, ...

But why we use compression for videos?

The main Goal for compression is to transmit videos faster and save the space of our computers memories. If we do not use compression, we should initialize about 200GB for a one hour video by 720 quality. But with compression, we can save it just by 1GB.



reduce the size of the video file for easier storing

Video is redundant data because adjacent frames are very similar.

So every compression method tries to reduce the redundancy of video.

And now how to overcome it?

The Solution is so simple, we keep similarities just once and just coding the differences between frames.

For compression we have three types of frames: I frames, P frames and B frames.

I frames will be encoded like a single image(like jpeg compression). P frames will be encoded by using previous frames. and B(bidirectional) frames will be encoded by

frames (until next I frame), we just encode the differences.

In the compression process, we divide frames to blocks. In older compression methods like MPEG2, equal size of blocks were used. But in last methods like H264, for a more complex area smaller block size used. And for a lighter areas bigger block size will be used. it is obvious that when we have many structures in some area, we should capture it by smaller block size. And in a more uncomplicated area, we do not need to use high precision.

after initializing the blocks, we should find the best match for them. It means we should find the most similar block to current block from previous or next reference frames(for P and B frames)

We have just two component in a compressed video; DCT coefficients and Motion Vectors(MVs).

Motion Vectors

After finding the best match for a special block, we save the distance between them. for example, a block transmits 2 to the right and 3 in the downside in comparison with the reference frame. so we save a vector of (3,-2) for it. we call these vectors as motion vectors. MVs are very similar to optical flows but the precision of optical flow is more than MVs. becuase optical flows are pixel-based and MVs are block-based. it means for every pixel between two frames we have a special optical flow, but in a compressed video just for every block that contains many pixels we have one motion vector.

frame n and n+1 and Motion Vectors of 16*16 blocks

DCT coefficients

Also when we find the best matches, we should subtract two blocks to achieve the difference between them. The difference will be encoded and saved. Then we utilize Discrete Cosine Transform(DCT) (For better result) to obtain DCT coefficients and after that, we utilize quantized matrix to save just important values. We dived the DCT coefficients to the quantized matrix.

Decompressing Videos

For watching a video, we must decode it. I frames will be decoded like jpg images. But for decoding P and B frames we should use the reference frames. The image of DCT coefficients is in the frequency domain and it is not clear. So after multiplying them to the quantized matrix, we apply IDCT (inverse of DCT) to receive the spatial domain images.

Compression | Realtime Video Processing | Convolutional Network

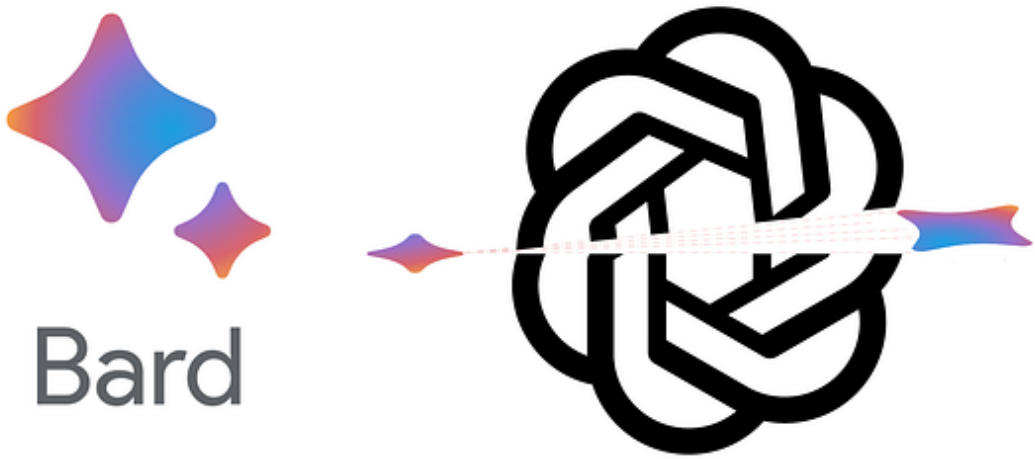
Follow

Written by ali abdari

4 Followers

Machine learning and computer vision researcher

Recommended from Medium



AL Anany

The ChatGPT Hype Is Over—Now Watch How Google Will Kill ChatGPT.

It never happens instantly. The business game is longer than you know.

· 6 min read · Sep 1

20K 626

Lessig

ChatGPT, or: How I Learned to Stop Worrying and Love AI

In my first book, Code and Other Laws of Cyberspace (1999), I told the story of why I had become a lawyer. My uncle, Richard Cates, had...

7 min read · Nov 29

4.97K 124

Lists

Staff Picks

531 stories · 518 saves

Stories to Help You Level-Up at Work

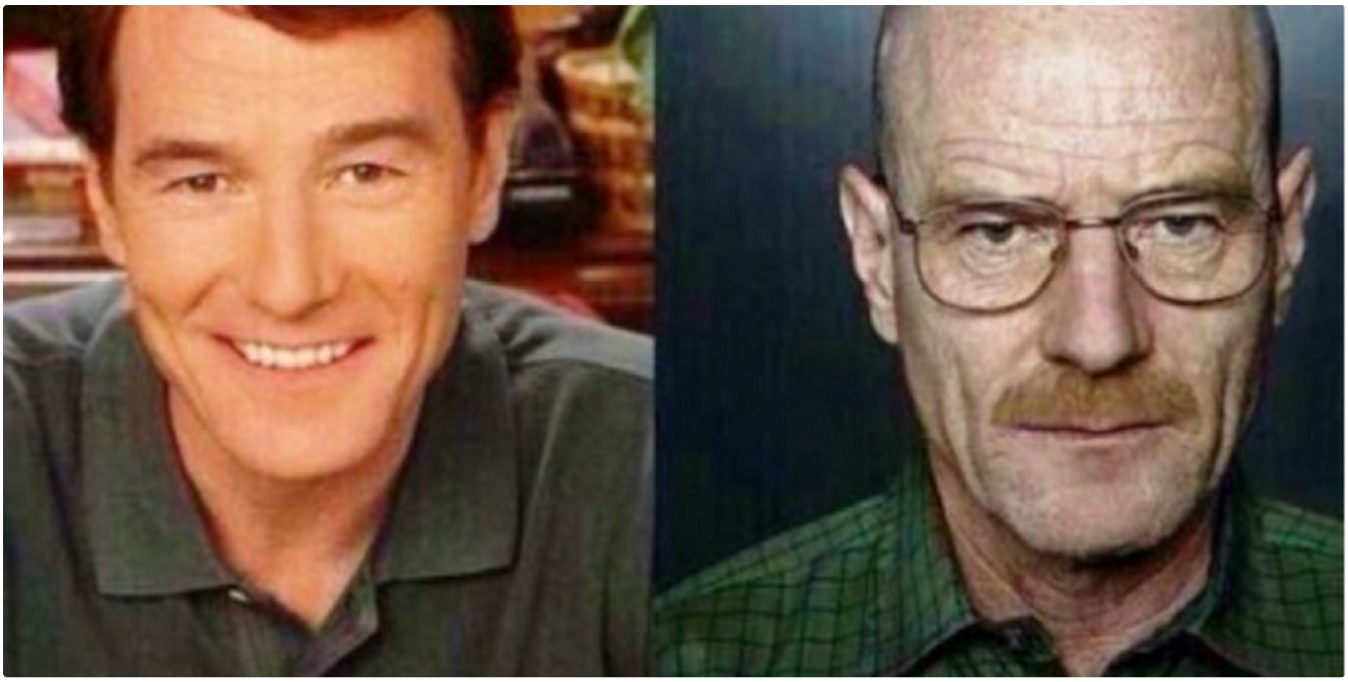
19 stories · 361 saves

Self-Improvement 101

20 stories · 1022 saves

Productivity 101

20 stories · 928 saves



David Goudet

This is Why I Didn't Accept You as a Senior Software Engineer

An Alarming Trend in The Software Industry

· 5 min read · Jul 26

5.7K 61

Vaishnav Manoj in DataX Journal

JSON is incredibly slow: Here's What's Faster!

Unlocking the Need for Speed: Optimizing JSON Performance for Lightning-Fast Apps and Finding Alternatives to it!

16 min read · Sep 28

10.7K 125

Unbecoming

10 Seconds That Ended My 20 Year Marriage

It's August in Northern Virginia, hot and humid. I still haven't showered from my morning trail run. I'm wearing my stay-at-home mom...

· 4 min read · Feb 16, 2022

71K 1017

Alberto Romero

The Best AI Model in the World: Google DeepMind's Gemini Has Surpassed GPT-4

Google has long been an object of ridicule. No more

10 min read · 5 days ago

1.6K 36