

Summer Internship 2021

University of Strathclyde, Glasgow, UK

Professor: Shishir Nagaraja



Problem Statement

The topic of my research is “Profile Cloning detection using openCV”. My research is focused on the images used in profile cloning. It aims to detect morphing and also give original images that were used in creating that morphed images.

I thus explores some of the fake image creation techniques and the major ones were **morphing/photoshopping** and **deep fakes**.

Deep fake images and videos



Deep fake of Elon Musk
singing song



Easy to detect deep fake of
Donald Trump and Elon Musk



Advanced deep fake.
Source: kaggle.com

Existing search engines on detection


The topic of my research is “Profile Cloning detection using openCV”. My research is focused on the images used in profile cloning. It aims to detect morphing and also give original images that were used in creating that morphed images.


I thus explores some of the fake image creation techniques and the major ones were **morphing/photoshopping** and **deep fakes**.

Detection - 1



Tested at: <https://deepware.ai/>

 **NO DEEFAKE DETECTED**



Name: result.mp4

Size: 288.9 KB

User

Source

2021-05-28 17:16:35 UTC


6 seconds ago

[New Scan](#)

DETAILS

Deepware aims to give an opinion about the scanned video and is not responsible for the result. As Deepware Scanner is still in beta, the results should not be treated as an absolute truth or evidence.

1.00



Model Results

Deepware: NO DEEFAKE DETECTED(45%)

Seferbekov: NO DEEFAKE DETECTED(36%)

Ensemble: NO DEEFAKE DETECTED(38%)

Video

Duration: 11 sec

Resolution: 256 x 256

Frame Rate: 23.98 fps

Codec: h264

Audio

Duration: 11 sec

Channel: stereo

Sample Rate: 44 khz

Codec: aac

Detection - 2



Submission date: 2021 May 28

File: 0004STET6P.jpeg

image / jpeg

Visual threats:

▲ Detected

Detection of

Result

Confidence ⓘ

Model Generator ⓘ

AI face swap ⓘ

✓ Undetected

-

-

GAN-generated face ⓘ

▲ Detected

99.9%

Stylegan

Tested at: <https://sensity.ai/>

Detection - 3



Tested at:
<https://images.google.com/>



Image size:
167 × 144

No other sizes of this image found.

Possible related search: **suit separate**

<https://www.menswearhouse.com> › Suits › Suit Separates


Suit Separates - Men's Suit Separate Combinations | Men's ...

Shop for men's suit separates including dress slacks, vests & tuxedos. See the latest **suit separate** combinations in the latest styles from Men's Wearhouse.

<https://www.amazon.in> › Dockers-Mens-Suit-Separate-P...

Buy Dockers Men's Suit Separate Pant, Grey Sharkskin, 40x32 ...

Buy Dockers Men's **Suit Separate** Pant, Grey Sharkskin, 40x32 from Casual Trousers at Amazon.in. 30 days free exchange or return.

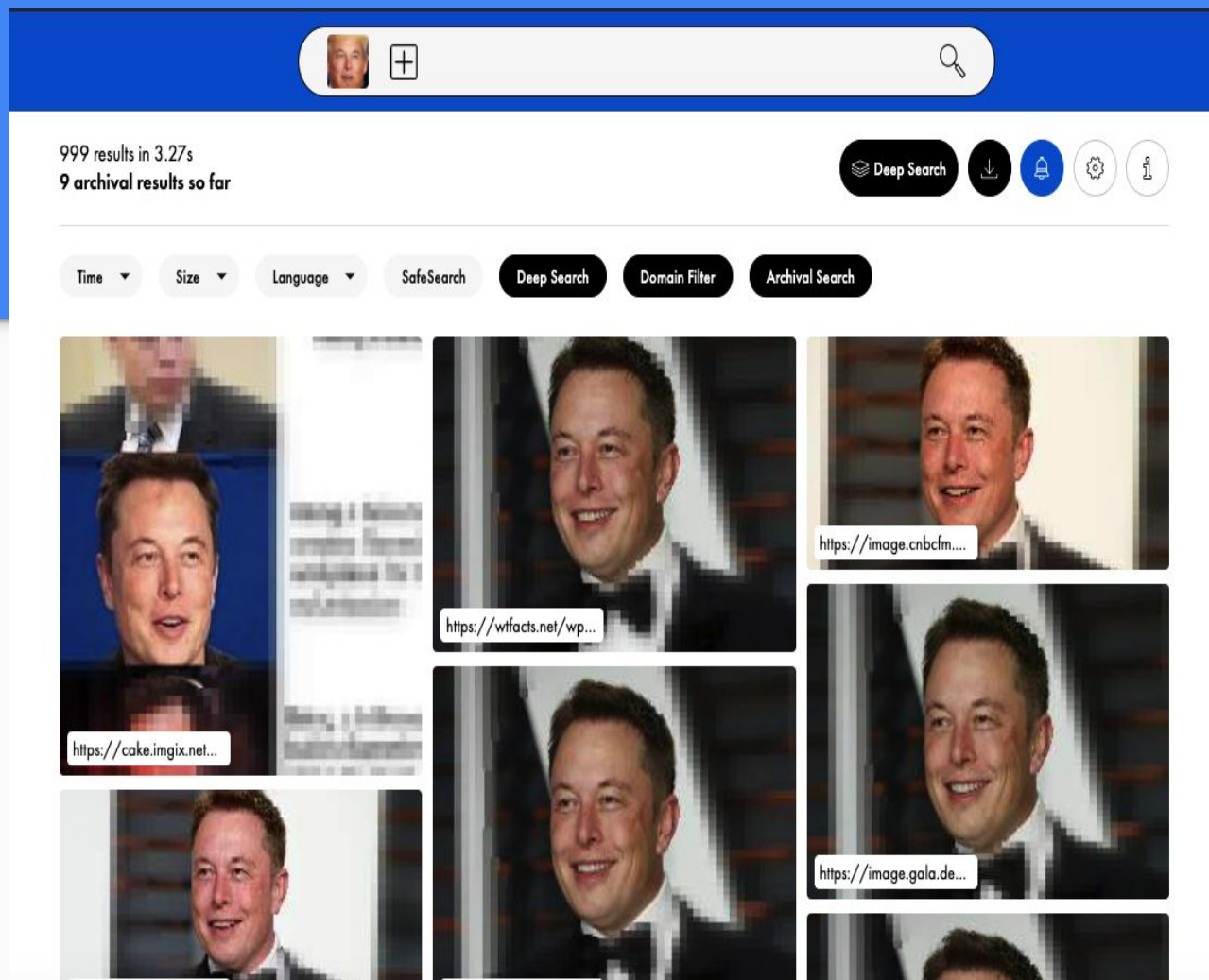
 **Visually similar images**



Detection - 4



Tested at:
<https://pimeyes.com/en>



The screenshot shows the Pimeyes search interface. At the top, there is a search bar with a small profile picture of Elon Musk and a plus icon. To the right of the search bar is a magnifying glass icon. Below the search bar, it displays "999 results in 3.27s" and "9 archival results so far". On the right side, there are several icons: a "Deep Search" button, a download icon, a bell icon, a settings icon, and a user profile icon. Below these icons, there are filters: "Time", "Size", "Language", "SafeSearch", "Deep Search", "Domain Filter", and "Archival Search". The main area displays a grid of search results, each showing a small image of Elon Musk and a URL. The URLs visible are: <https://cake.imgix.net...>, <https://wtfacts.net/wp...>, <https://image.cnbcfm...>, <https://image.gala.de...>, and <https://image.cnbcfm...>.

Research developments in this field

Next I looked up on existing research work done in this area and read research papers published regarding the same.

Following are some of the insights from the readings.

Paper-1

Link: <http://www.ws.binghamton.edu/fridrich/Research/copymove.pdf>

Using some image processing tools, specific areas are detected and other noise is blacked out. Then comparisons are made to detect the forgery/find the original.

Talks about block matching algorithm that compares images in chunks and not as a whole.

Paper-2

Link:

https://www.researchgate.net/publication/264276516_Detection_of_Clones_in_Digital_Images

This paper discusses detection of cloning. Compare small blocks from the image and then compare those blocks in separate threads to get the output.

There were other methods like DCT and PCA that calculated the gray factor of each block to compare the blocks.

Paper-3

Link: <https://ieeexplore.ieee.org/document/8782292>

This paper discusses the efficient methods to detect cloning of an image. It talked about doing this in 3 ways which are as follows:

- Take the image and convert it to grayscale.
- Divide into disjoint blocks and extract the main features.
- Do the matching process (comparison) and detect the duplicates.

Facts on profile cloning (non technical)

- Facebook alone removed over **580 million profiles** in the first quarter of 2018 alone while Twitter removed **70 million accounts** from May to June of 2018.
- **Profile photo and location** are the most important information that is checked by any user for checking the originality of any profile.
- The main problem with deep fakes is that it can preserve the pose, facial expressions and lighting of the photographs.
- They also **reduce the quality** a bit to make some of the current deep fake detection algorithms give false negatives.

Retrieving original from deep fake-1



Tested at:
<https://images.google.com/>



Image size:
167 × 144

No other sizes of this image found.

Possible related search: **suit separate**

<https://www.menswearhouse.com> › Suits › Suit Separates


Suit Separates - Men's Suit Separate Combinations | Men's ...

Shop for men's suit separates including dress slacks, vests & tuxedos. See the latest **suit separate** combinations in the latest styles from Men's Wearhouse.

<https://www.amazon.in> › Dockers-Mens-Suit-Separate-P...

Buy Dockers Men's Suit Separate Pant, Grey Sharkskin, 40x32 ...

Buy Dockers Men's **Suit Separate** Pant, Grey Sharkskin, 40x32 from Casual Trousers at Amazon.in. 30 days free exchange or return.

 **Visually similar images**



Revisit detection - 4

In the previous image, the most part of the image was taken by the face only that resulted in different results in different search engines (google and pimeyes).

So, I took an image where the face did not take much area and other features of the image (which are crucial for cloning/deep fake detection) get more area.

The results were much better in that case verifying the algorithm described in the research papers.

Retrieving original from deep fake-2



Tested at:
<https://images.google.com/>

Google

Photo...929087.jpg x virat kohli

Pages that include matching images

<https://twitter.com/imvkohli/status/12009087>

Virat Kohli on Twitter: "Quarantine diaries. Un-ironed T-shirt ...

1200 × 900 · 16-Nov-2020 — **Virat Kohli Fans™** @vk_devotee 16 Nov 2020. More. Copy link to Tweet; Embed Tweet. Replying to @imVkohli. ❤️ 2 replies 0 retweets 7 ...

<https://www.indiatvnews.com/Sports/Cricket>

'Quarantine diaries': Virat Kohli shares picture from Australia ...

715 × 402 · 17-Nov-2020 — **Virat Kohli** took to Twitter to post a picture of him in quarantine as the Indian cricketers prepare for the upcoming series against Australia.

[https://www.filmfare.com/news/bollywood/throwback...](https://www.filmfare.com/news/bollywood/throwback)

Throwback: Virat Kohli is stunned that Katrina Kaif spoke to ...

600 × 450 · 14-Apr-2021 — Indian skipper **Virat Kohli** and actor and producer Anushka Sharma are one of India's most-loved and followed couples. The celebrity couple ...

<https://www.dnaindia.com/Sports/Cricket>

THIS is how Virat Kohli is spending his quarantine in Australia

640 × 360 · 17-Nov-2020 — The decision of **Virat Kohli** to leave the Australian series has drawn mixed reactions from Indian and foreign audiences. **Virat Kohli's** presence for ...

<https://zeenews.india.com/Cricket>

Retrieving original from deep fake-3



Tested at:
<https://images.google.com/>

Google

Photo...929376.jpg X effort to level the playing field barron trump will X

<https://www.ndtv.com> > World

Donald Trump Changes Campaign Logo After Being Mocked ...

360 × 222 · 17-Jul-2016 — The campaign of US Republican Presidential nominee Donald Trump unveiled a new logo on Saturday, a day after an original logo was ...

<https://imgflip.com> > ...

It seems Biden is afraid of a real debate. - Imgflip

500 × 955 — It seems Biden is afraid of a real **debate**. | BIDEN WON'T DEBATE. share. 49,236 views · 42 upvotes · Made by Perspicacity 12 months ago in politics · Joe Biden ...

<https://www.ndtv.com> > World

Saddam Hussein Killed Terrorists 'So Good': Donald Trump ...

360 × 222 · 07-Jul-2016 — Donald Trump, who frequently criticizes US foreign policy under President Barack Obama and former Secretary of State Hillary Clinton, ...

<https://www.ndtv.com> > World

Donald Trump's Plans Could Cripple Economy, Says New ...

360 × 222 · 13-Jul-2016 — An analysis by the New York City Council finds that Donald Trump's lightning-rod proposals to deport illegal immigrants and temporarily ban ...

<https://www.ndtv.com> > World

Donald Trump Promises To Herald US Economic Resurgence

360 × 222 · 29-Jun-2016 — Criticising the current US trade policies that has wiped

Getting original from deep fake image

After output of certain search engines on deep fakes, we could see original images as output.

Features apart from faces play vital role in such cases. This is because google and other search engines just outputs visually similar images and then looks for the tags related to those images to give the text corresponding to "Possible related search".

Kaggle deep fake result -1



Tested at:
<https://images.google.com/>

The screenshot shows a Google Images search interface. At the top, the search bar contains the word "lady". Below the search bar, the results are displayed. The first result is a video thumbnail for "Lady - Kenny Rogers - YouTube" with a duration of 11-Jan-2008. Below this, there is a section for "Visually similar images" showing a grid of 16 small images of the same woman. To the right of the main results, there is a sidebar with a definition of the word "lady" and a section for "People also search for" with links to "Knight", "Gentleman", "Viscount", and "Goddess".

Google image2.jpeg x lady

Q All Images Maps Shopping More Settings Tools

About 2 results (2.31 seconds)

Image size: 1024 x 1024
No other sizes of this image found.

Possible related search: **lady**

<https://www.youtube.com/watch>
Lady - Kenny Rogers - YouTube
11-Jan-2008 — lady kenny rogers If you enjoyed my channel and decide to tip in crypto. I thank you for your support. Bitcoin wallet: ...

<https://www.youtube.com/watch>
Lady - Kenny Rogers - YouTube
28-May-2018 — Youtube RadioUtopia Video Creations Subscribe <https://bit.ly/2Mi6DG7> ...

Visually similar images

Lady

The word lady is a term of respect for a girl or woman, the equivalent of gentleman. Once used to describe only women of a high social class or status, the female equivalent of lord, now it may refer to any adult woman. Informal use of this word is sometimes euphemistic or, in American slang, condescending. Wikipedia

People also search for View 5+ more

[Knight](#) [Gentleman](#) [Viscount](#) [Goddess](#)

Feedback

Kaggle deep fake result -2



Tested at:
<https://images.google.com/>

The screenshot shows a Google Images search interface. The search bar contains the text 'boy'. Below the search bar, it says 'About 2 results (0.66 seconds)'. The first result is a small thumbnail of a young boy, with the text 'image size: 1024 x 1024' and 'No other sizes of this image found.' below it. To the right of the thumbnail, there is a section titled 'Boy' with a definition: 'A boy is a young male human. The term is usually used for a child or an adolescent. When a male human reaches adulthood, he is described as a man. Wikipedia'. Below the definition is a 'Feedback' link. Underneath the first result, there are two links to YouTube videos: 'Charlie Puth - BOY [Official Audio] - YouTube' and 'TREASURE - 'BOY' MV - YouTube'. At the bottom, there is a section titled 'Visually similar images' showing a grid of 12 small thumbnails of various young boys.

Face extraction

This step is required as a pre-processing step before executing the NMF algorithm.

This step takes images as input and gives the faces as the output in .npz format using deep learning.

Output

Extracted faces from
images of the dataset.



Extracting images using NMF Technique

The output we are expecting is the NMF components and the eigenfaces.

This output represents the features corresponding to the images in the dataset.

The idea to use this features to search for the images in the search engines.

Comparison-1

For this, we will use the notion of “norm” in a vector space. Using linear algebra, we know that “norm” behaves similarly as the distance and the difference in norm would thus become the parameter for comparison.

We used the standard norm function from opencv called **cv.NORM_L2**.

Matching the features

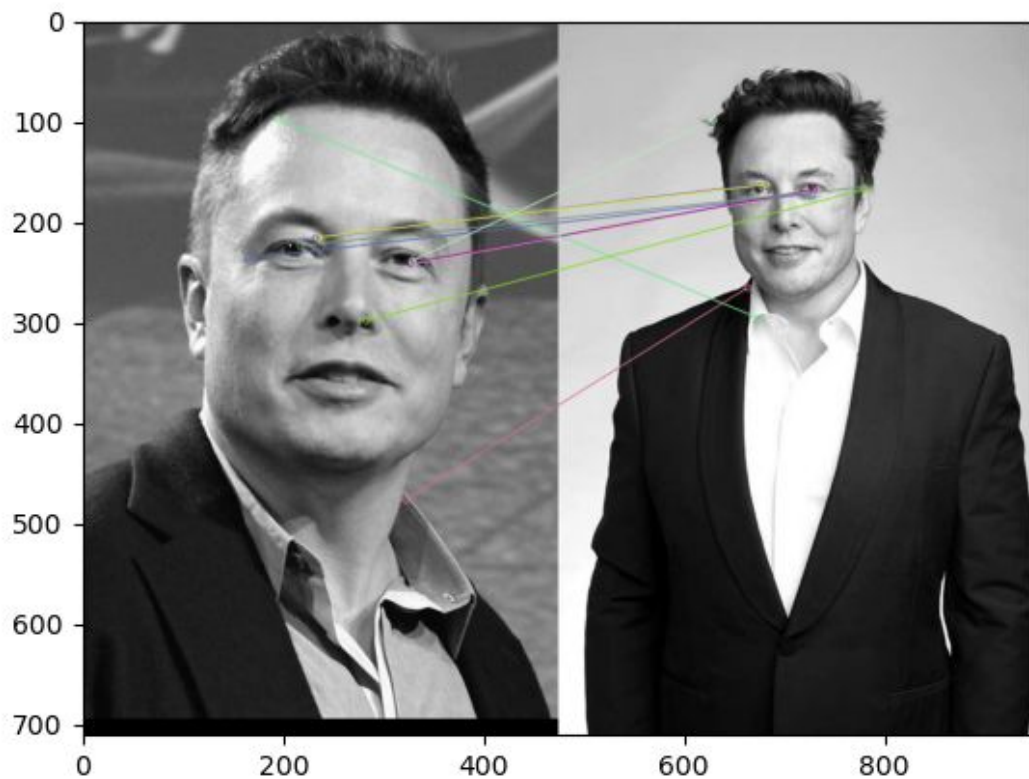
For this, we will use the ORB Descriptors.

This will return the similar parts from two images and arrange them in increasing order of norm.

It is thus stored in decreasing order of similarity. You can choose to find first “M” similar features from the image.

Output-1

Feature comparison of
two images using norm



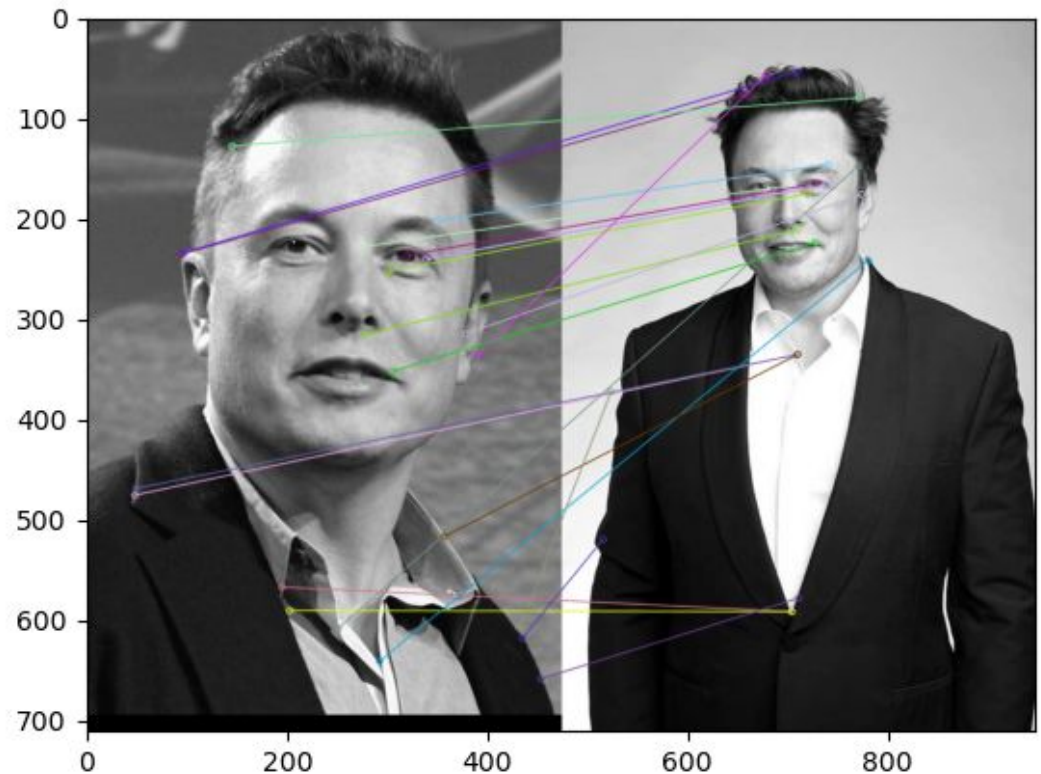
Comparison-2

Another way is to use the Brute Force Matcher to get k best matches.

It uses SIFT Descriptors and machine learning to do feature comparison.

Output-2

Feature comparison of two images using BF Matcher.



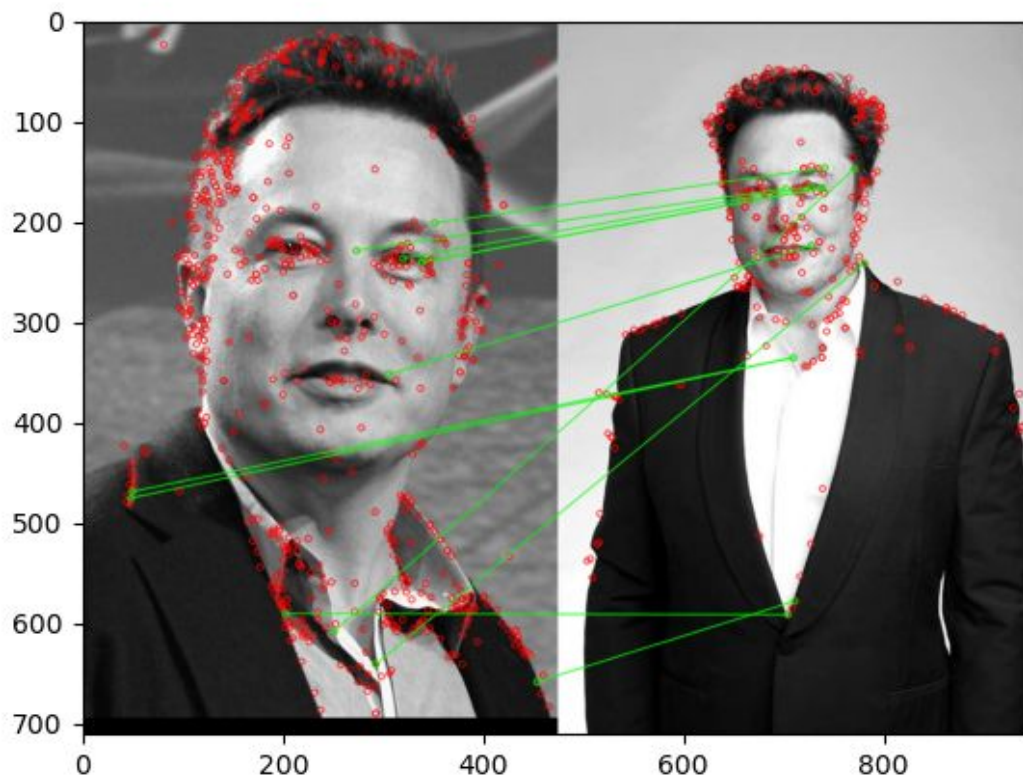
Comparison-3

Another way is to use the FLANN based matcher.

FLANN stands for Fast Library for Approximate Nearest Neighbours. For large datasets and for higher dimensions, it works better than BF Matcher.

Output-3

Feature comparison of
two images using
FLANN matcher.



Key features of the comparison algorithm

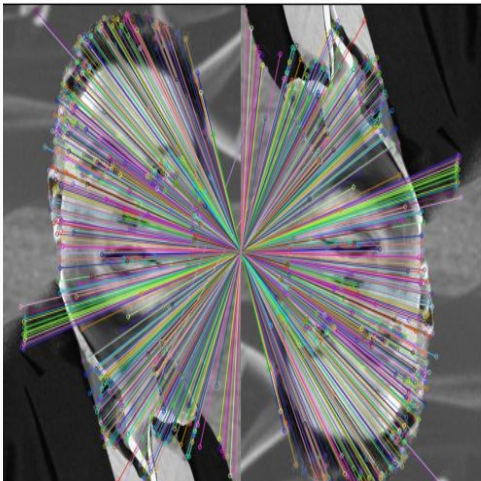
Some of the main features are:

- The algorithm is fast and takes <1 sec/image.
- The algorithm is able to take into consideration the rotations and alignment of view.
- The algorithm compares the features and 2 out of 3 algorithms can be given a mathematical quantity for the similarity extent.

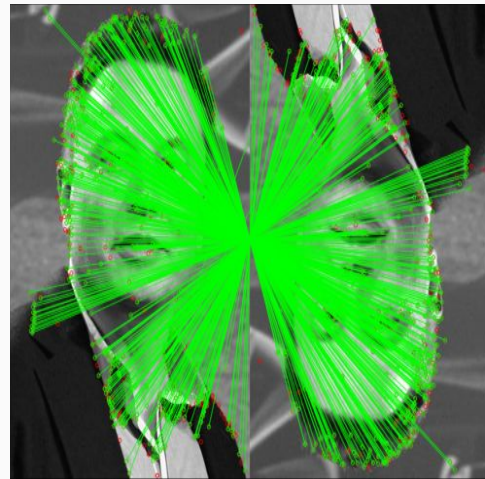
Output with rotated but same image



Algorithm-1



Algorithm-2



Algorithm-3

Output with rotated and different image



Algorithm-1



Algorithm-2



Algorithm-3

References

- Research paper on FLANN Algorithm:
https://www.researchgate.net/publication/339170738_FLANN_Based_Matching_with_SIFT_Descriptors_for_Drowsy_Features_Extraction
- Comparison of various feature detection algorithms:
https://www.researchgate.net/publication/312634655_Comparison_of_OpenCV's_feature_detectors_and_feature_matchers
- https://www.researchgate.net/publication/292995470_Image_Features_Detection_Description_and_Matching

Integrating the methods (engineering)

Taking the directory names and method number as input from command line and calling corresponding functions.

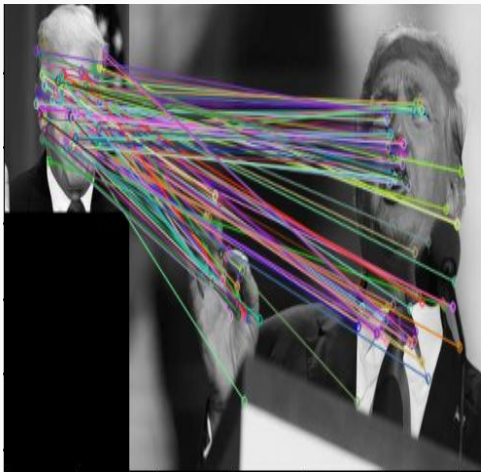
Also ensuring all the exception cases and adding comments for the user.

Also did some optimisations that were possible as we were calling the method multiple times in a loop.

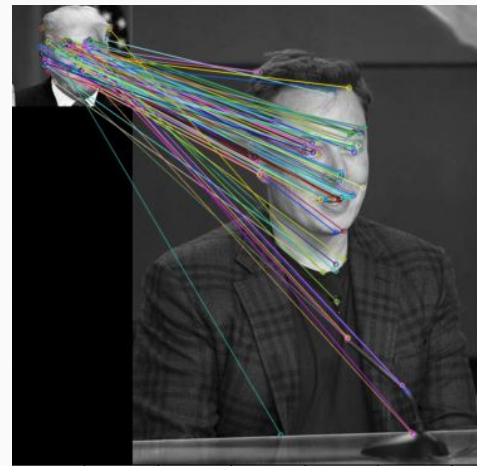
Test-1 (Method1)



Test image



Original-1 (wrong)
Similarity score: 116/260
Score of correct image: 95/260

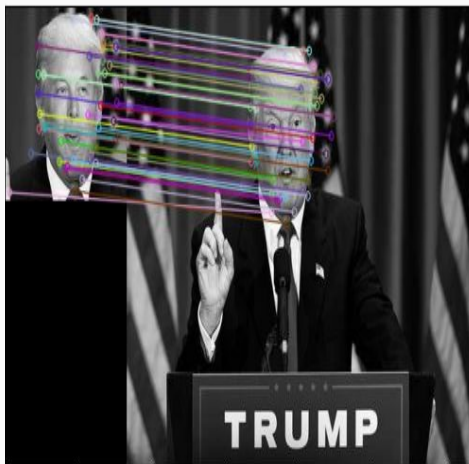


Original-2 (wrong)
Similarity score: 116/260
Score of correct image: 105/260

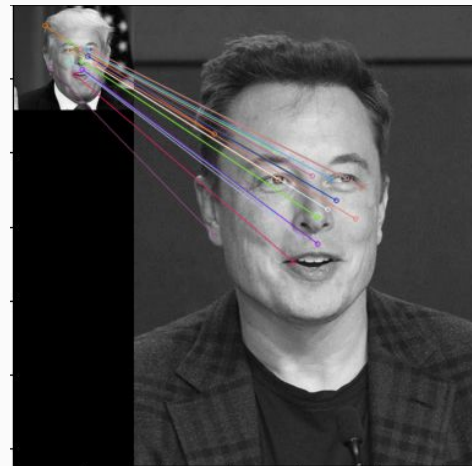
Test-1 (Method2)



Test image



Original-1 (correct)
Similarity score: 49/99

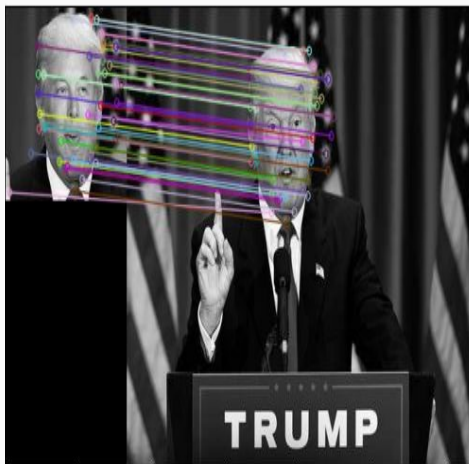


Original-2 (correct)
Similarity score: 18/99

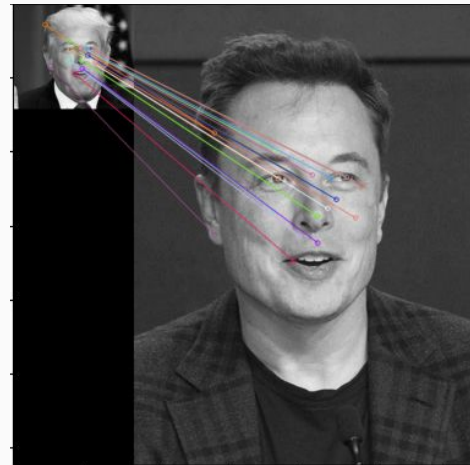
Test-1 (Method3)



Test image



Original-1 (correct)
Similarity score: 149/198

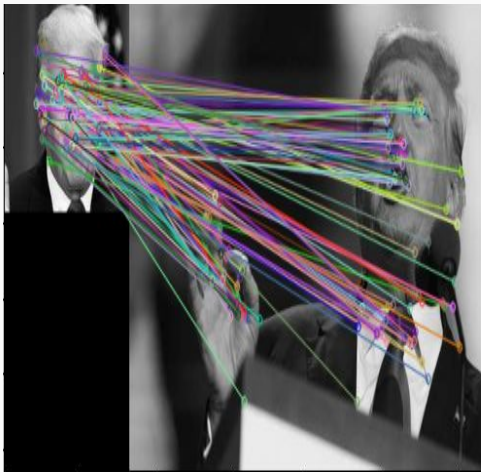


Original-2 (correct)
Similarity score: 116/198

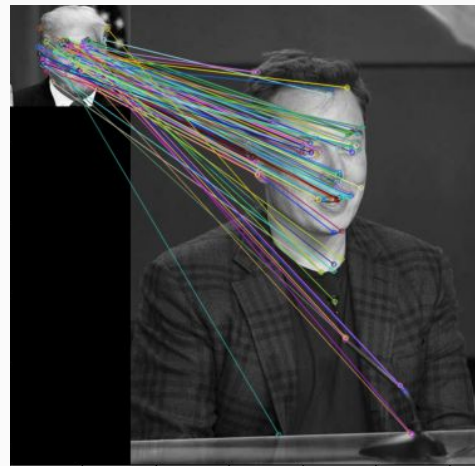
Test-2 (Method1)



Test image



Original-1 (wrong)
Similarity score: 147/500
Score of correct image: 118/500

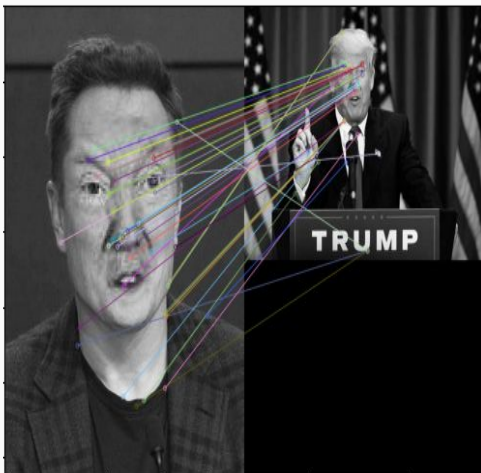


Original-2 (correct)
Similarity score: 155/500

Test-2 (Method2)



Test image



Original-1 (correct)
Similarity score: 42/443



Original-2 (correct)
Similarity score: 184/443

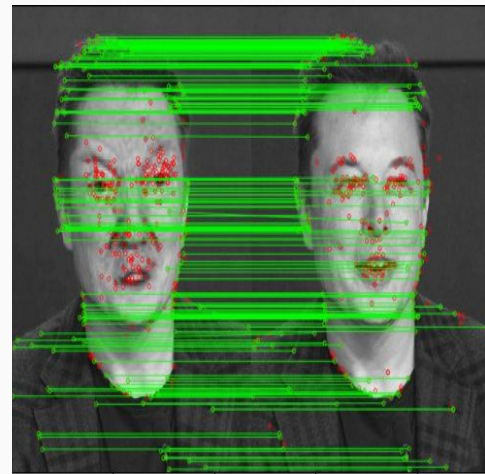
Test-2 (Method3)



Test image



Original-1 (correct)
Similarity score: 479/886



Original-2 (correct)
Similarity score: 622/886

Notes and Observations

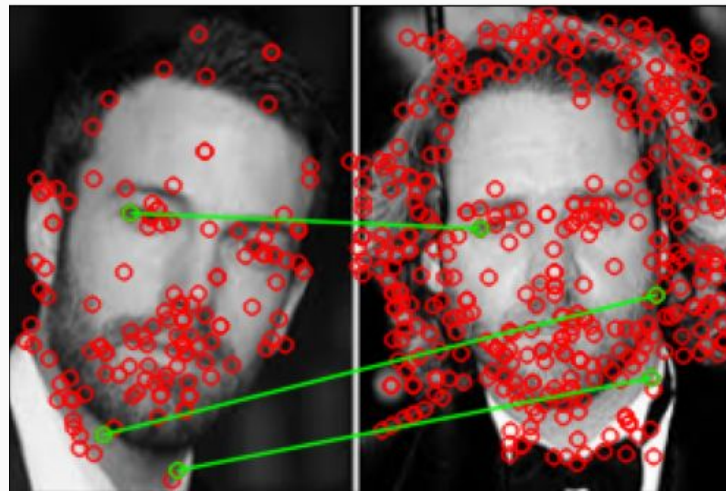
When face of person B is put over image of person A, then major features are corresponding to image of person hence the similarity score corresponding to image of person A is higher (refer to slide 5 and 8 where person A refers to Donald Trump and Elon Musk respectively).

When features other than face (stage, dress and miscellaneous objects) are significant in the image, method 2 and method 3 gives greater similarity corresponding to those features and hence correct output.

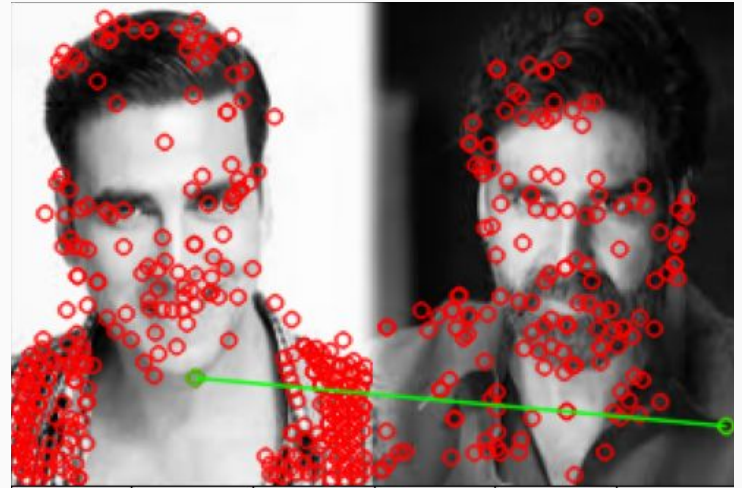
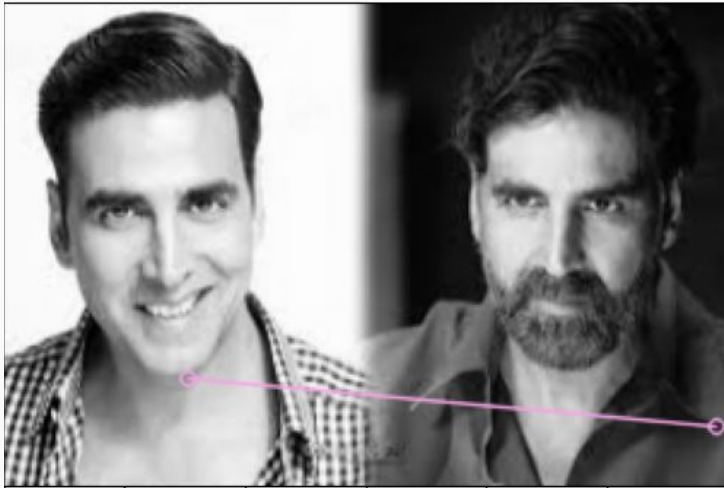
Beard output-1



Beard output-2



Beard output-3



Note -2

Between method 2 and method 3, method 2 focuses more on explicit and distinct features in the image like facial features.

Method 3 on the other hand focuses more on the comparison of outlining structure and skeleton of the features (outlined shape and size).

Summary

- Method 2 and method 3 takes around 16-18 seconds on average to process the dataset of size 300 while method 1 takes nearly 3 seconds.
- Method 1 has lower accuracy than method 2 and method 3.
- Images that compare beard and no beard **showed mixed output**. In some cases, other features like eyes and nose came out properly while in some cases the features are incorrectly marked.

Generating deep fakes

1. Used pre-trained model by stylegan for faces and provided seed images to get the corresponding output images.
2. Created a google colab notebook for the same.

Link: https://colab.research.google.com/drive/1z8Q7qWbcWA-SvtfzX_sEqq7agyXpNDPE?usp=sharing

Seed images



Output images



Testing the algorithm

The seed and output were not matching since one output took features (learning by the model) from all the seed images.

Some features match out but overall the image does not seem to be related at all.

Images with people of different age and gender are paired with maximum similarity score.

Algorithm output - 1



Algorithm output -2



Deep learning models for custom deep fake

The model is hungry for images and needs to take a lot of images as input. We could include out 2-3 images in the image dataset but the output would not resemble very much with the original image (as seen earlier).

Google colab notebook link for one such

training:https://colab.research.google.com/drive/1sAg9vBwYbgtxJltg_XHScSSjFD_P6UG2O?usp=sharing

Source:https://github.com/jeffheaton/present/blob/master/youtube/gan/colab_gan_train.ipynb

Minutes of the meeting - 1

1. Deep fakes are not mainly used in profile cloning. Hence, our code should be targeting morphed/photoshopped images.
2. We thus need to explore some morphing techniques for generating images for profile cloning.
3. We also have to establish that morphing and photoshopping are relevant methods and widely used in profile cloning techniques.
4. We will keep deep fakes into consideration as additional method of generating fake images.

Morphing in profile cloning

This paper talks about using morphing techniques to create fake documents like passport.

It also compares the detection ability of humans and computers.

Clean photoshopped images are very difficult to detect and is widely used in creating fake profiles and documents.

Link: <https://cognitiveresearchjournal.springeropen.com/articles/10.1186/s41235-019-0181-4>

Image morphing survey

These papers talk about various morphing techniques and use/misuse of the technique.

Link: https://www.researchgate.net/publication/2801109_Image_Morphing_A_Survey

Link: https://www.researchgate.net/publication/2855203_Recent_Advances_in_Image_Morphing

Fraudulent ID using face morphs and their detection. Link: <https://bit.ly/3hOvc3b>

Created morphed images

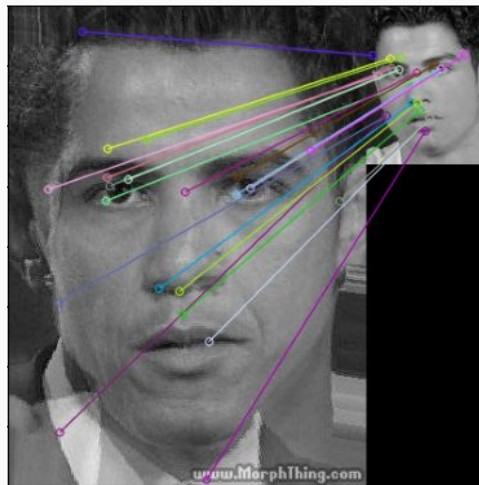
I used <https://www.morphthing.com/morph> currently for creating morphs for the testing.

Some morphs are as good as natural while others have some distortions that could be removed using photoshop techniques.

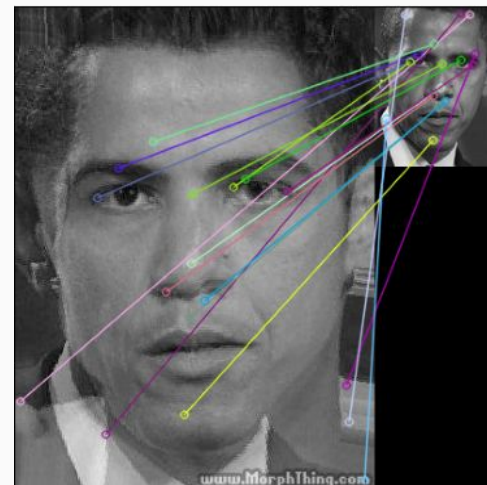
Morphing testing - 1



Test image



Original-1 (correct)



Original-2 (correct)

Morphing testing on google search



Tested at:
<https://images.google.com/>

About 178 results (1.07 seconds)



Image size:
300 × 400

No other sizes of this image found.

Possible related search: *harry styles and obama combined*

<https://www.eonline.com> > News

Harry Styles Supposedly Had an Affair With This Man - E! Online

12-Dec-2014 — The hoax National Examiner cover features **Obama** with a line of text that reads, "[**Obama**] had sex with THIS MAN," with an arrow that points to ...

<https://www.wattpad.com> > stories > hobama

Hobama Stories - Wattpad

Harry styles and his "friend" Barack **Obama** have been photographed together at the local Starbucks, read this magical love story about harry+**Obama**=❤️.

Pages that include matching images

<https://www.morphthing.com> > celebrity > 46089605-B...

- MorphThing.com



300 × 400 — Barack **Obama** 50%; Harry Styles 50%. Name Combinations. 1st Name. Bararry; Barry; Bary; Hack; Harack; Harrack. 2st Name. Morph parents. Barack **Obama** ...

<https://in.linkedin.com> > shanvaan-ls

Morphing testing on search engines

Google images showed output of “Harry Styles and Obama combined”. Though it was able to guess that **it is a combination of two images**. But the originals were guessed incorrectly.

Pimeyes showed Barack Obama images as output.

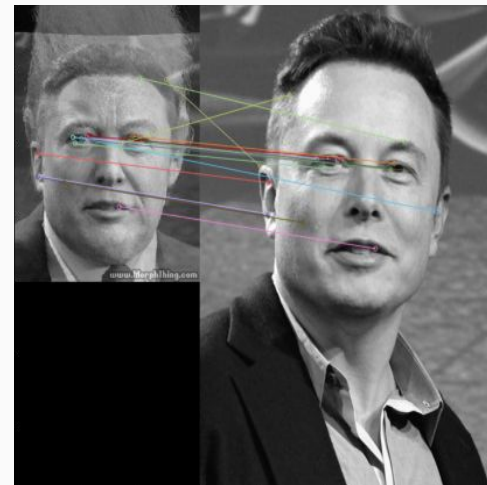
Morphing testing - 2



Test image



Original-1 (correct)

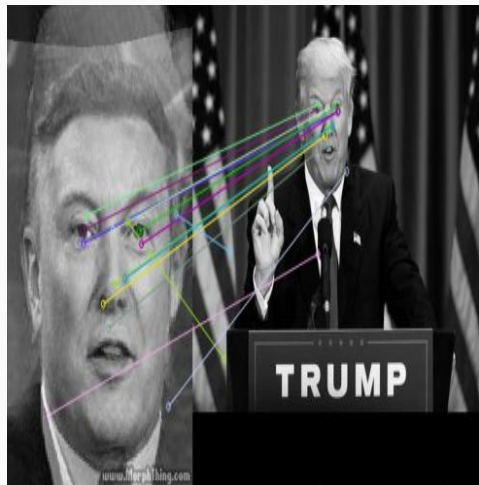


Original-2 (correct)

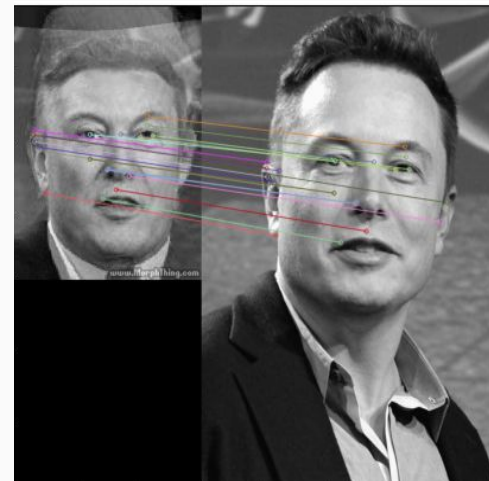
Morphing testing - 3



Test image



Original-1 (correct)



Original-2 (correct)

Getting feature-wise score


By simply sorting the features array in method -1 and taking distance array in method-2 and method-3, we are able to quantify similarity for every feature separately.





We can also implement mathematical tools like power law to further emphasise if a particular feature has high score (so that it is not averaged out if other features are not taken from a particular image).

Morphing testing on google search




Tested at:
<https://images.google.com/>

Google 

Q All  Images  Maps  Shopping  More Tools


About 152 results (0.69 seconds)

 Image size:
100 × 133
No other sizes of this image found.

Possible related search: **ashton hilton dermatology**

<https://hiltondermatology.com> › board-certified-dermat...
Our Dermatologists | Hilton Dermatology + Aesthetics
A **dermatologist** who has been "board certified" has undergone the strict training required by the American Board of **Dermatology**. ... Dr. Ashton Taylor Hilton, MD.

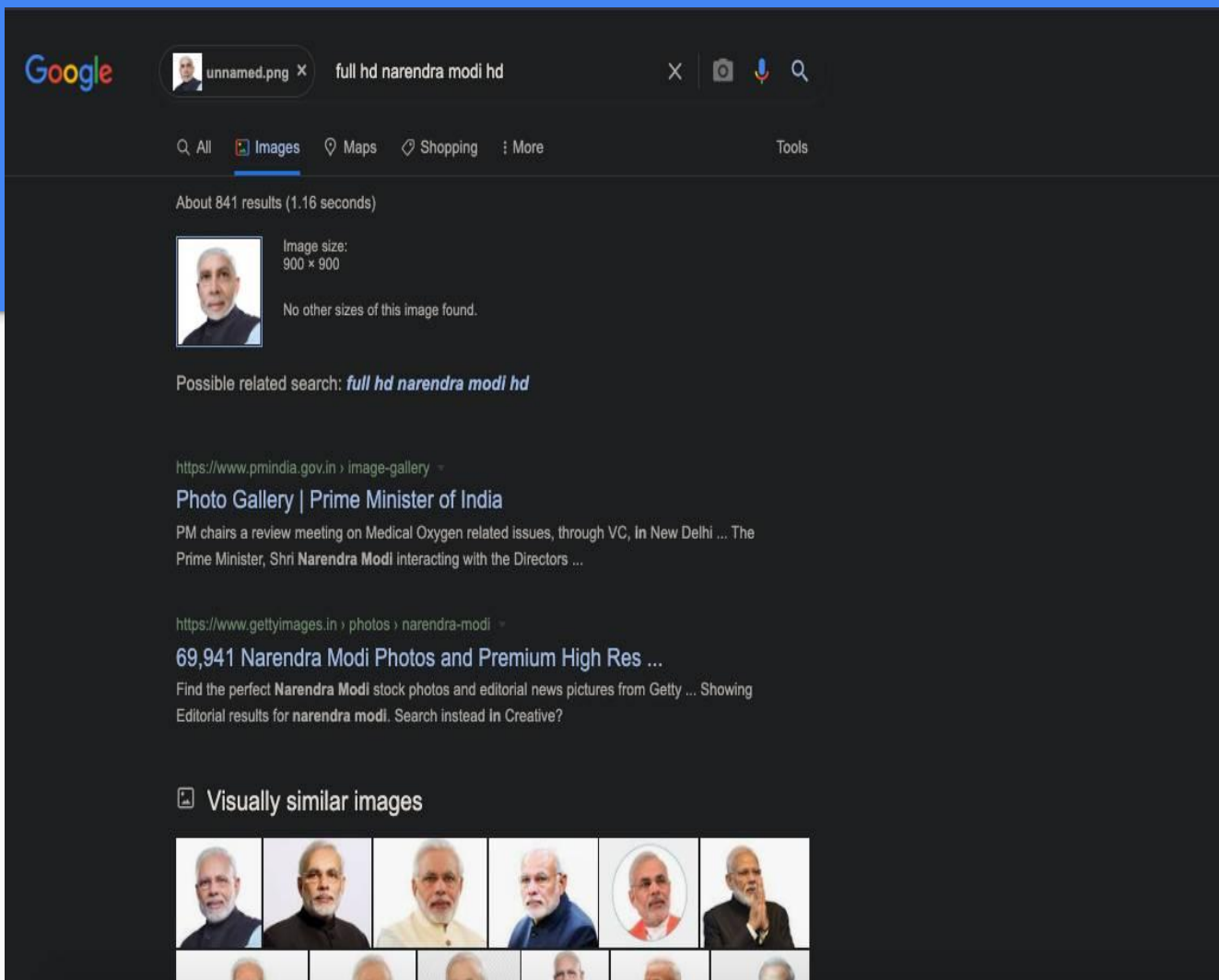
<https://health.usnews.com> › Doctors
Dr. Ashton T. Hilton, MD | Slidell, LA | Dermatologist | US ...
Dr. Ashton T. Hilton is a **dermatologist** in Slidell, Louisiana and is affiliated with Slidell Memorial Hospital. He received his medical degree from Louisiana State ...

Pages that include matching images
<https://www.doximity.com> › States › Louisiana › Slidell
Dr. Ashton Hilton, MD – Slidell, LA | Dermatology - Doximity
 320 × 320 — Dr. Ashton Hilton, MD is a board certified dermatologist in Slidell, Louisiana. He is affiliated with Slidell Memorial Hospital.

Morphing testing on google search



Tested at:
<https://images.google.com/>




Google

unnamed.png x full hd narendra modi hd

Q All Images Maps Shopping More Tools

About 841 results (1.16 seconds)

 Image size: 900 × 900
No other sizes of this image found.

Possible related search: **full hd narendra modi hd**

<https://www.pmindia.gov.in> > image-gallery


Photo Gallery | Prime Minister of India


PM chairs a review meeting on Medical Oxygen related issues, through VC, in New Delhi ... The Prime Minister, Shri **Narendra Modi** interacting with the Directors ...

<https://www.gettyimages.in> > photos > narendra-modi

69,941 Narendra Modi Photos and Premium High Res ...

Find the perfect **Narendra Modi** stock photos and editorial news pictures from Getty ... Showing Editorial results for **narendra modi**. Search instead in Creative?

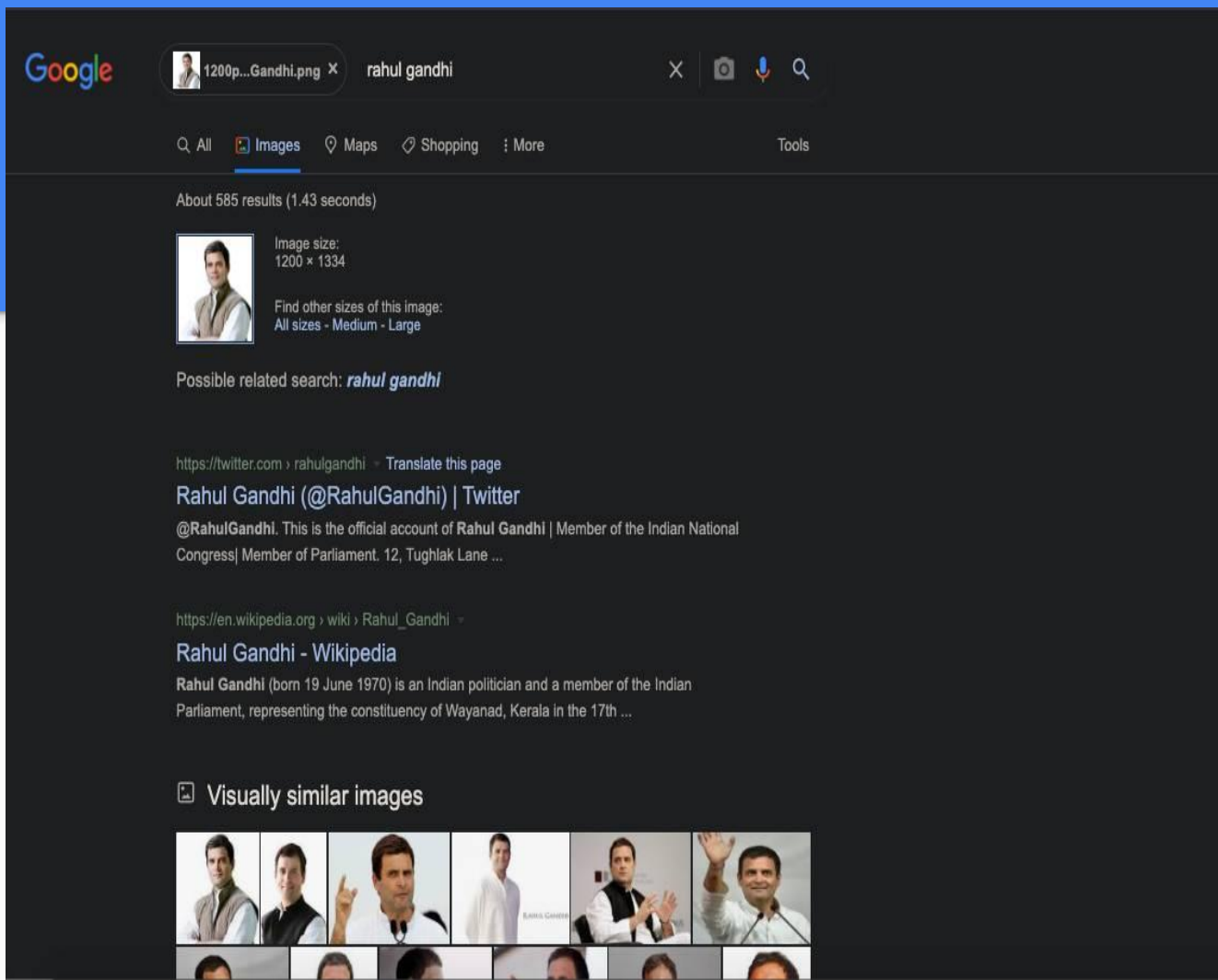
 **Visually similar images**



Morphing testing on google search



Tested at:
<https://images.google.com/>

A screenshot of a Google Images search for 'rahul gandhi'. The search bar shows the query and a thumbnail of the image. Below the search bar, the 'Images' tab is selected. The results show a single image of Rahul Gandhi with a size of 1200 x 1334. Below the image, there are links to his Twitter profile and Wikipedia page. At the bottom, there is a section for 'Visually similar images' showing a grid of related photos.

Google 1200p...Gandhi.png x rahul gandhi

Q All Images Maps Shopping More Tools

About 585 results (1.43 seconds)

Image size:
1200 x 1334

Find other sizes of this image:
All sizes - Medium - Large

Possible related search: *rahul gandhi*

<https://twitter.com/rahulgandhi> Translate this page
Rahul Gandhi (@RahulGandhi) | Twitter
@RahulGandhi. This is the official account of Rahul Gandhi | Member of the Indian National Congress| Member of Parliament. 12, Tughlak Lane ...

https://en.wikipedia.org/wiki/Rahul_Gandhi
Rahul Gandhi - Wikipedia
Rahul Gandhi (born 19 June 1970) is an Indian politician and a member of the Indian Parliament, representing the constituency of Wayanad, Kerala in the 17th ...

Visually similar images