8. 顔情報処理
（顔検出、顔ランドマーク、感情認識、顔認識）

（ディープラーニング，Python を使用）

（全１５回）

https://www.kkaneko.jp/ai/ae/index.html

金子邦彦

BY NC SA

①顔情報処理の基本的な概念の理解

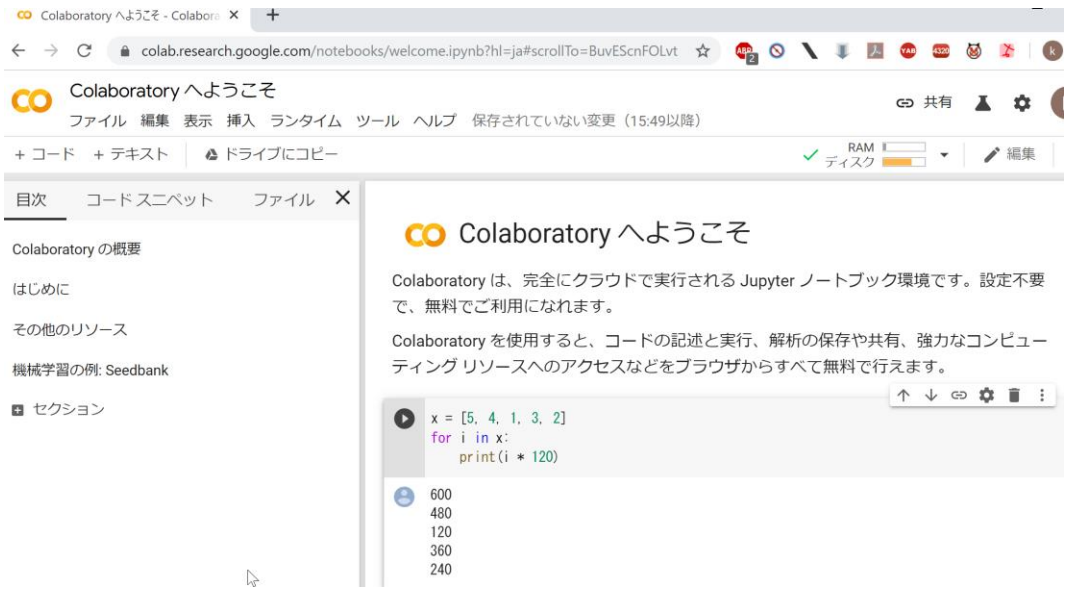
②テクノロジーの理解

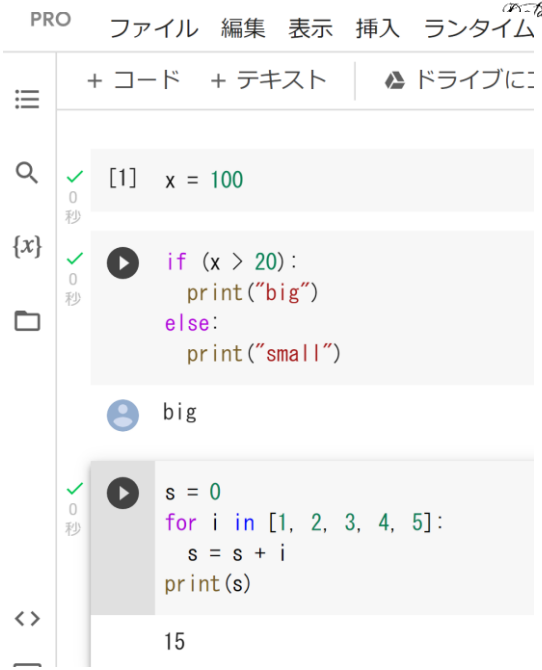
③実用的なスキルの獲得

アウトライン

1. イントロダクション
2. 顔情報処理の概要
3. ディープラーニングによる顔検出
4. 顔ランドマーク
5. ディープラーニングによる顔認識
6. ディープラーニングによる感情認識その他

Google Colaboratory

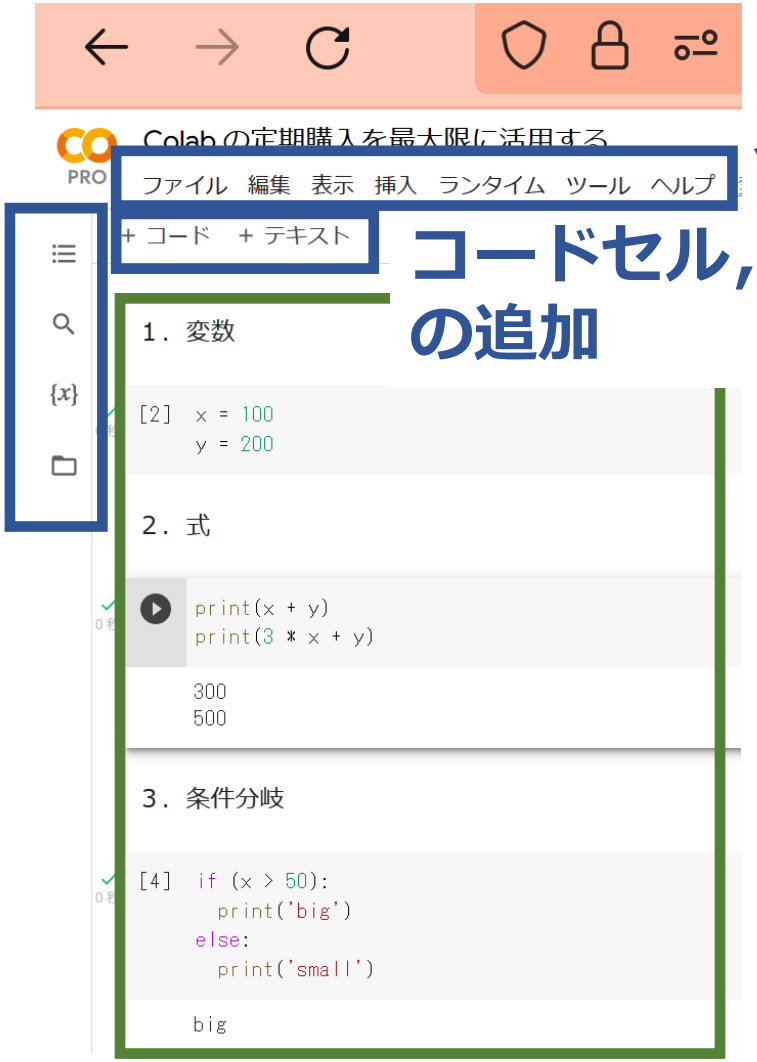




URL: https://colab.research.google.com/

- **オンラインで動く**
- **Python のノートブック**の機能を持つ
- Python や種々の機能が**インストール済み**
- **本格的な利用**には，**Google アカウントが必要**

Google Colaboratory の全体画面



メニュー

コードセル，テキストセルの追加

メニュー（目次，検索と置換，変数，ファイル）

コードセル，テキストセルの並び

Web ブラウザの画面

Google Colaboratory のノートブック

コードセル，**テキストセル**の２種類

- **コードセル**： Python プログラム，コマンド，実行結果
- **テキストセル**：説明文，図

２．式 ← テキストセル

```
[5] print(x + y)
    print(3 * x + y)
```

← コードセル

```
300
500
```

３．条件分岐 ← テキストセル

```
if (x > 50):
  print('big')
else:
  print('small')
```

← コードセル

```
big
```

8-1. イントロダクション

画像理解の主な種類

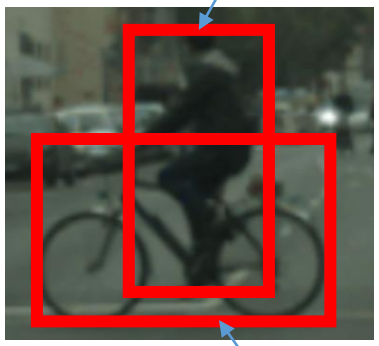
① **画像分類**

「何があるか」を理解

person
bicycle

② **物体検出**

場所と大きさも理解



③ **セグメンテーション**

画素単位で理解

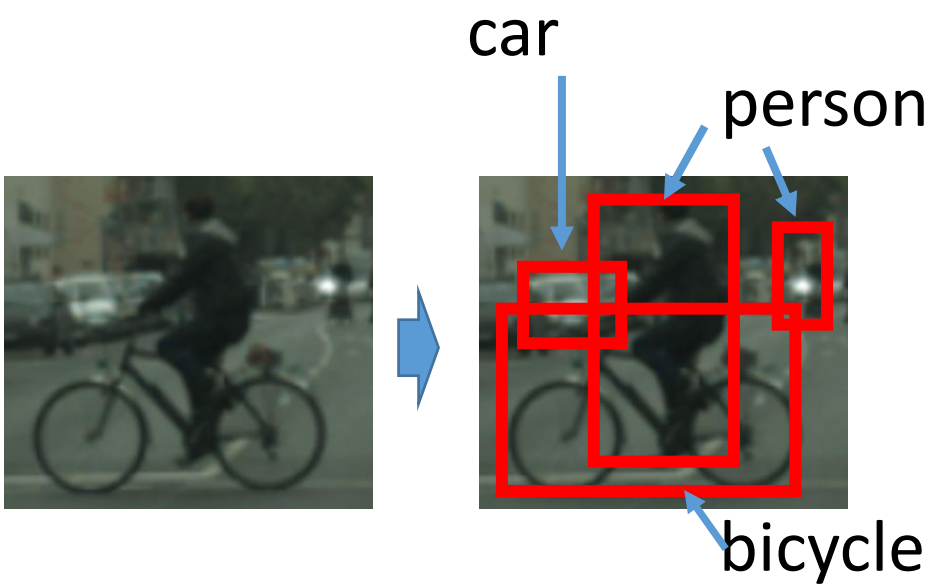
① 画像分類

```
Score 0.9827020168304443, Label lab_coat
Score 0.0030872616916894913, Label syringe
Score 0.0024311079178005457, Label beaker
Score 0.0016609227750450373, Label stethoscope
Score 0.00037950885598547757, Label plate
```

画像分類の結果は，ラベルと**確率**

※ 5 つの候補 (top 5) が表示されている

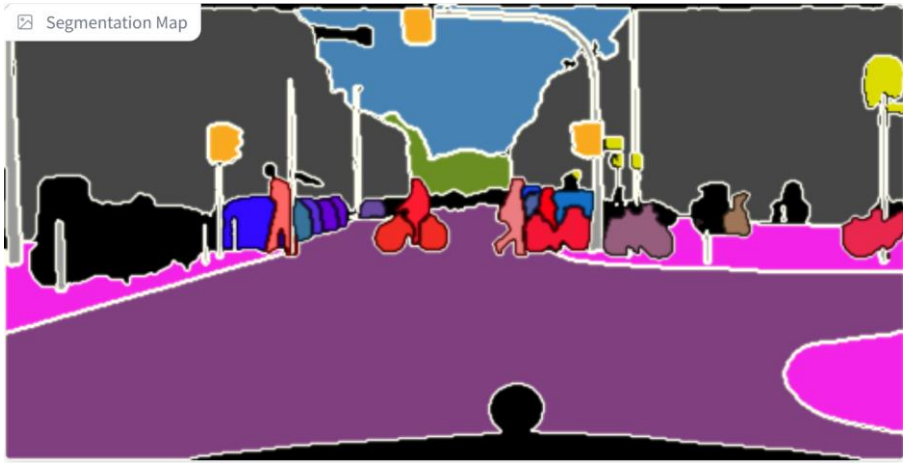
② 物体検出



バウンディングボックス，
ラベルを得る

バウンディングボックスは，
物体を囲む最小のボックス（四角形）

③ セグメンテーション



物体の形を画素単位で抜き出し

ラベルを得ることもできる

人工知能

知的なITシステム

機械学習

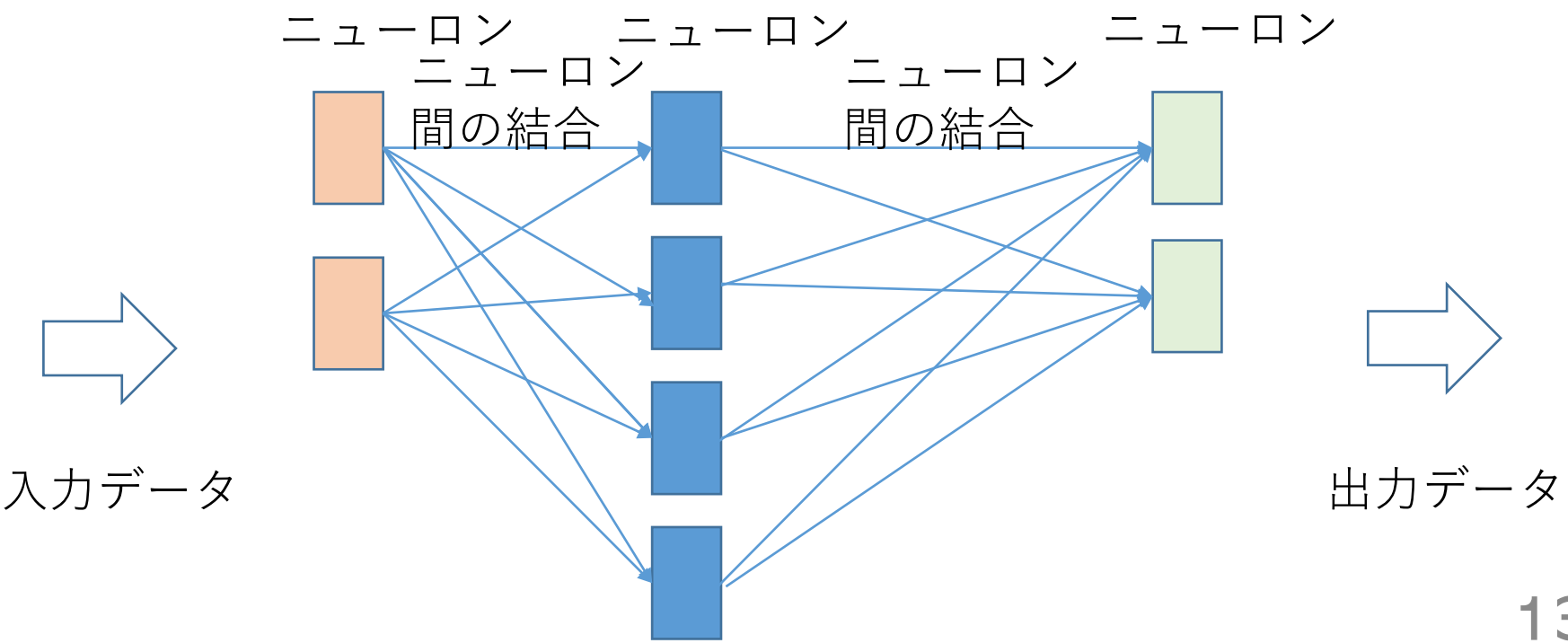
データから**学習**し、知的能力を向上

ディープラーニング

データから**学習**し、複雑なタスクを実行。**多層のニューラルネットワーク**を使用

ニューロンとニューラルネットワーク

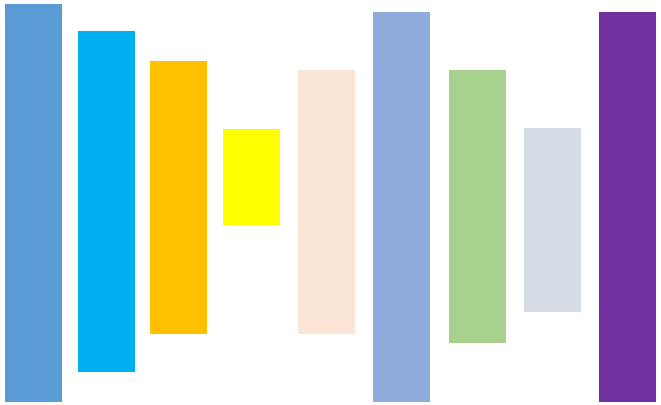
- **ニューロン**は、**ニューラルネットワーク**の基本的な構成要素
- 一つ一つの**ニューロン**は、データの受け取り、処理、伝達を行う
- **ニューラルネットワーク**は、これらの**ニューロン**が多数組み合わさったもの



ディープラーニング

ディープラーニングに「**ディープ**」とついているのは、**多層のニューラルネットワーク**を使用するため

層の数が少ない



層の数が多い（ディープ）

教師あり学習と教師なし学習

<u>教師あり学習</u>

- **訓練データ**には、**入力データ（x）**、それに**対応する正解（y）**が含まれる

  例：手書き文字 (x) とその認識結果 (y)

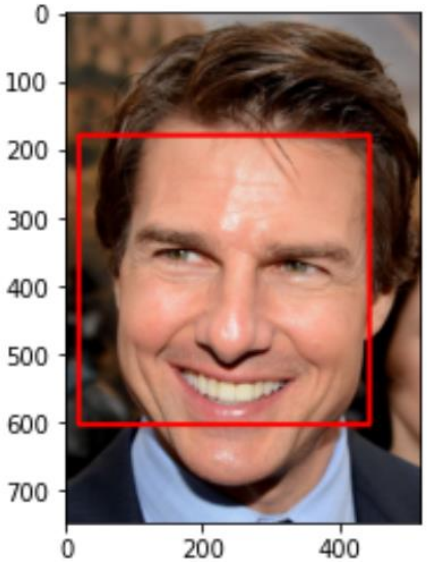
<u>教師なし学習</u>

- 訓練データに**正解（y）が含まれない．入力データ（x）のみ**が使用される

※ 教師あり学習と教師なし学習の両方の混在させる「半教師あり学習」もある

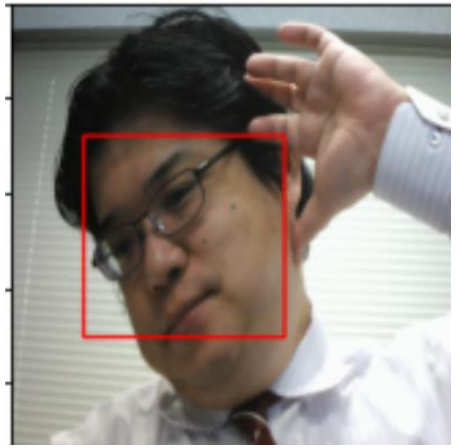
8-2. 顔情報処理の概要

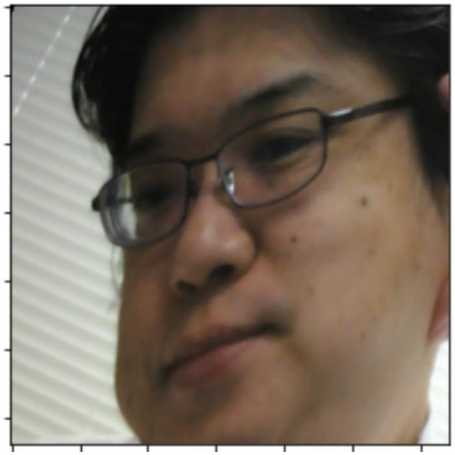
顔情報処理のバリエーション①



- **顔検出**

物体検出と同様に顔を検出

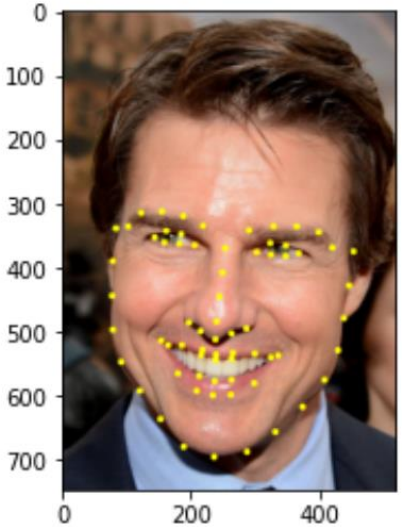




- **顔のアラインメント**

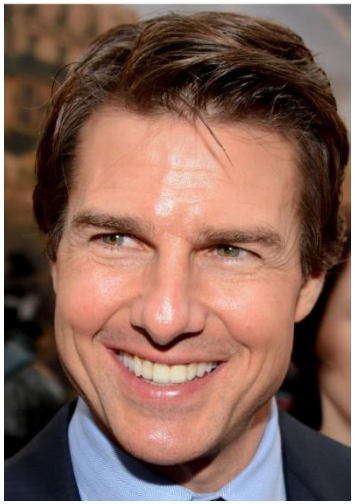
顔の向き，大きさをそろえる

顔情報処理のバリエーション②



- **顔ランドマーク**

顔の重要な特徴点（目、鼻、口など）の位置を示すポイント

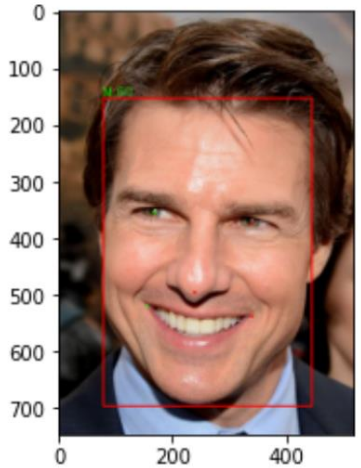


```
[ 0.00455161  0.18412845  0.03993299 -0.00525442 -0.07139583  0.13386713
 -0.00805785 -0.05532961  0.23680274 -0.08285979  0.23483047  0.00531268
 -0.20885959 -0.08143201 -0.07200904  0.08402673 -0.11594398 -0.08790354
 -0.14109443 -0.06426463 -0.01909781  0.04940728 -0.09086579  0.04928832
 -0.21027854 -0.29970455 -0.0172264  -0.11317077  0.01727176 -0.16984728
 -0.09449591  0.13467805 -0.11701771 -0.05477776 -0.00075595  0.02513538
 -0.14347562 -0.03627753  0.16524641 -0.01949969 -0.12205475  0.08029588
  0.0263175   0.27922362  0.26064846  0.00236352 -0.00611193 -0.09485988
  0.21062998 -0.27925995  0.02801728  0.21664803  0.10859946  0.12886725
  0.13740893 -0.2259727  -0.04096841  0.08036234 -0.15912706  0.04860799
 -0.05707375 -0.0537034  -0.0382829  -0.04455713  0.08876641  0.08084331
 -0.08434555 -0.12838244  0.14743124 -0.16956419  0.05809632  0.13801341
 -0.08743402 -0.27705559 -0.17667291  0.12853563  0.40629697  0.25780377
 -0.18405011  0.0308622  -0.04818188 -0.08573647  0.05904602 -0.00438196
 -0.12593931 -0.09314068 -0.07387269  0.06559635  0.18202811  0.02111054
 -0.0110196   0.23951159 -0.00382505 -0.03864232  0.03268222  0.00397856
 -0.16739431  0.05639653 -0.11059792 -0.01466753  0.09382112 -0.13222337
  0.00125152  0.08060078 -0.16358295  0.22464372 -0.00883166  0.03960528
  0.02616193 -0.04999878 -0.14178708  0.06188779  0.25730038 -0.27515596
  0.19043125  0.04089454  0.02481355  0.17113039  0.10915946  0.05673687
 -0.00430411 -0.0710548  -0.14189517 -0.08640159  0.0943445  -0.08253401
  0.04771694 -0.04451045]
```

- **顔の数値化**

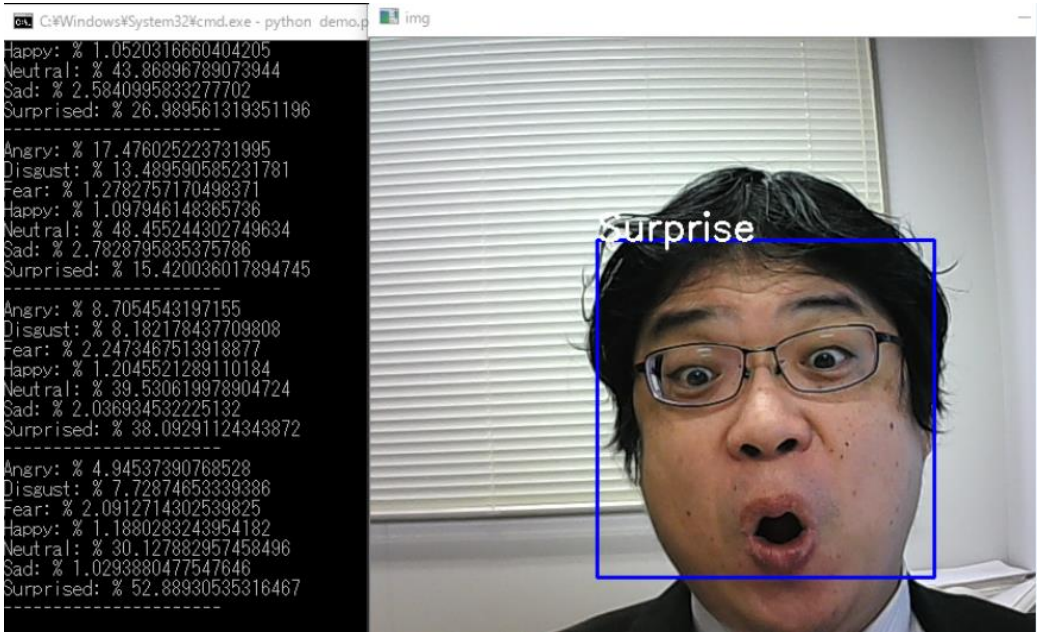
顔の特徴を数値化したもの．数値は複数になる．顔による本人確認や、顔認識に使用

顔情報処理のバリエーション③



gender： 1
age： 60

- **性別、年齢の推定**



- **表情の推定**

https://github.com/ezgiakcora/Facial-Expression-Keras で公開されている成果物を利用

顔情報処理のバリエーション

- 顔検出
- 顔のアラインメント
- 顔ランドマーク
- 顔の数値化：顔による本人確認や顔認識の基礎
- 性別、年齢の推定
- 表情の推定
- その他

演習 1

オンラインデモによる

顔検出の体験

【トピックス】

- **顔検出**

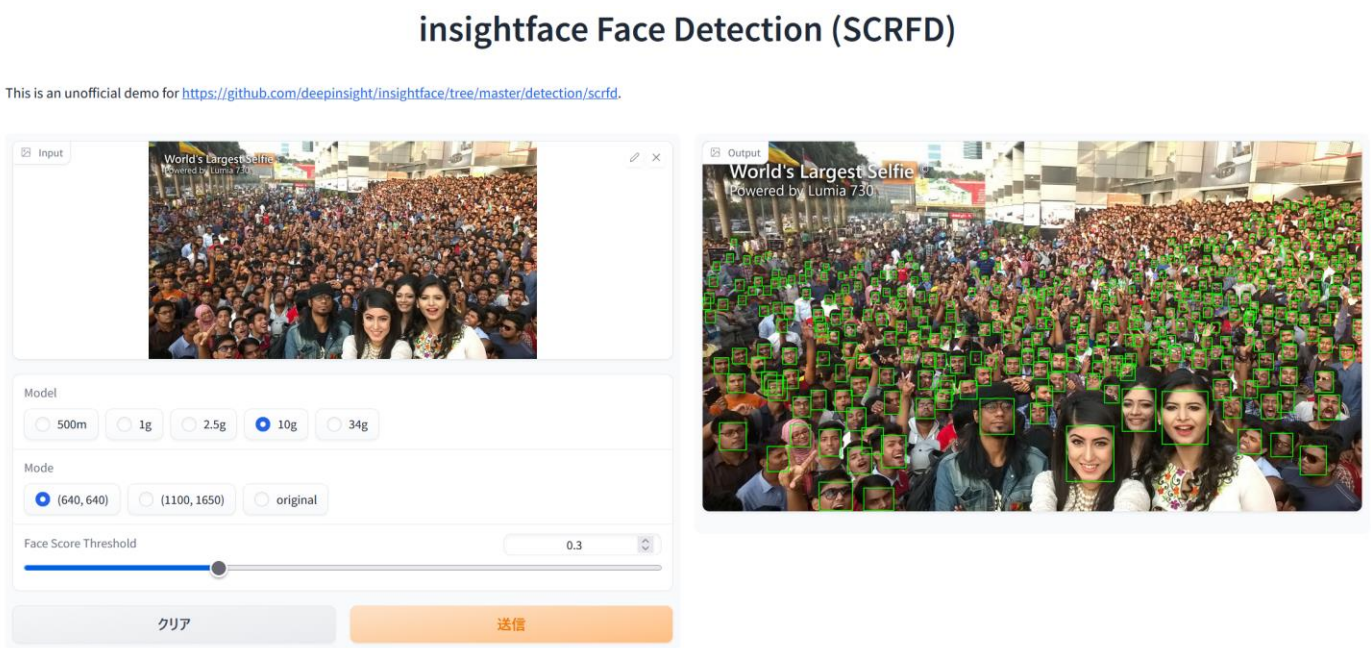
① 次のデモページを開く

https://huggingface.co/spaces/hysts/insightface-SCRFD

　オンラインサービスであり、混雑時などは動かない場合がある。授業中で動かなかった場合には、後日試してほしい

② 下の画像をクリック。画像をアップロードすることもできる。

③ 「**送信**」 をクリック。結果を確認。



自習１

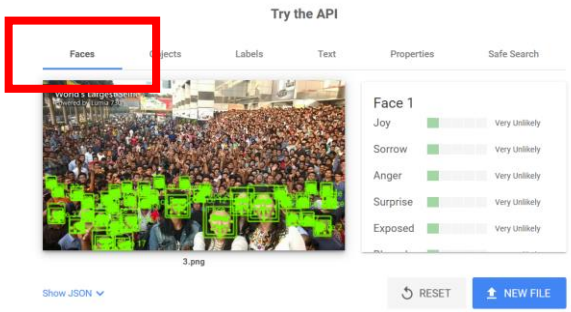
- まず、自分で顔画像のファイルを準備する
- 前ページで説明したデモページを使い、アップロードして**顔検出**を試してみる

https://huggingface.co/spaces/hysts/insightface-SCRFD

自習２

- まず、自分で顔画像のファイルを準備する
- 次のデモページを使う
- 「Drag image file here of Browse from your computer」をクリック。画像をアップロードして、**表情の推定**を試してみる
- そのとき、「私はロボットではありません」と表示されたらチェック

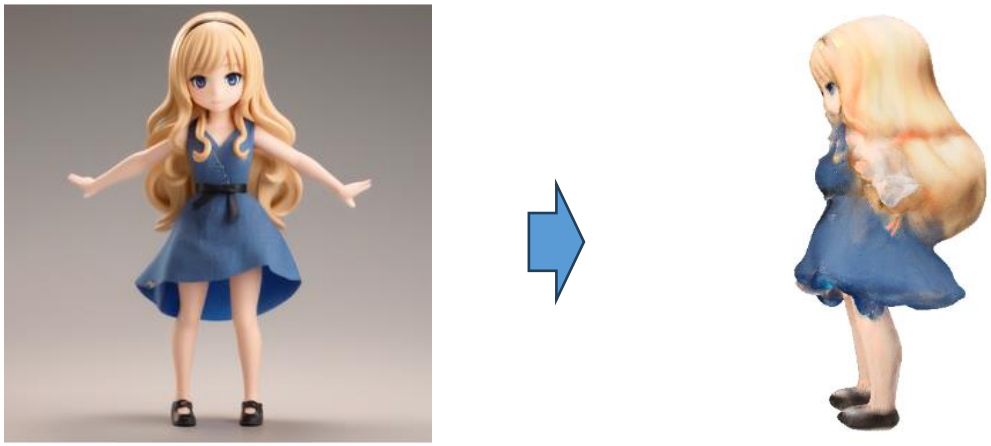
https://cloud.google.com/vision/docs/drag-and-drop?hl=ja



自習３

次の Google Colaboratory で、画像の３次元化を試してみる。
１つの画像を、ディープラーニングにより３次元化している。

https://colab.research.google.com/drive/1sLpYmmLS209-e5eHgcuqdryFRRO6ZhFS?usp=sharing



自分で画像を準備し、アップロードして実行してみることもできる

顔情報処理の重要性

多岐にわたる分野で重要な役割を果たしている

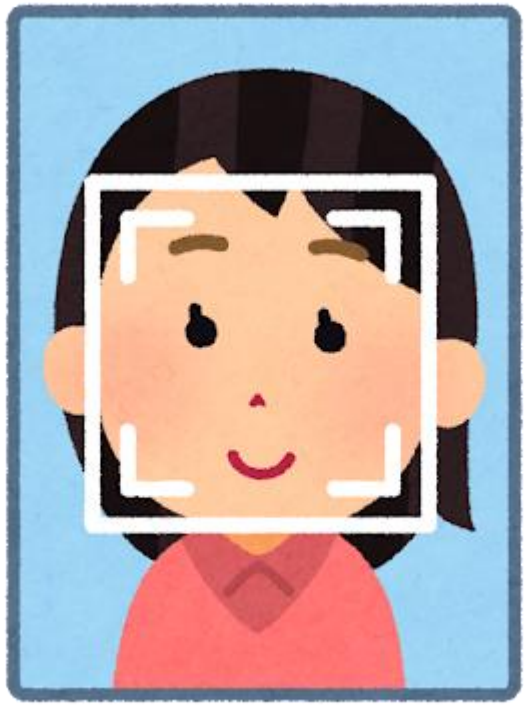
- セキュリティと監視
- 顔認証システム
- パーソナライズされたサービス
- 感情認識

その他

セキュリティと監視

公共の場での監視、身元確認や不審者の追跡

顔認証システム



スマートフォンやパソコンでの顔によるロック解除機能

顔情報処理の歴史

- 誕生（1960年代）

顔認識システムの誕生

- 進化（1980年代）

顔認識アルゴリズム（Eigenfaces など）

- ディープラーニングの登場（2010年代）

精度の向上、応用の広がり（群衆のカウント、マスク有り顔の処理、顔画像からの３次元顔生成などさまざまな応用）

群衆のカウント

- 群衆のカウント（画像内の人数を数える）
- 監視等に役立つ.

元画像

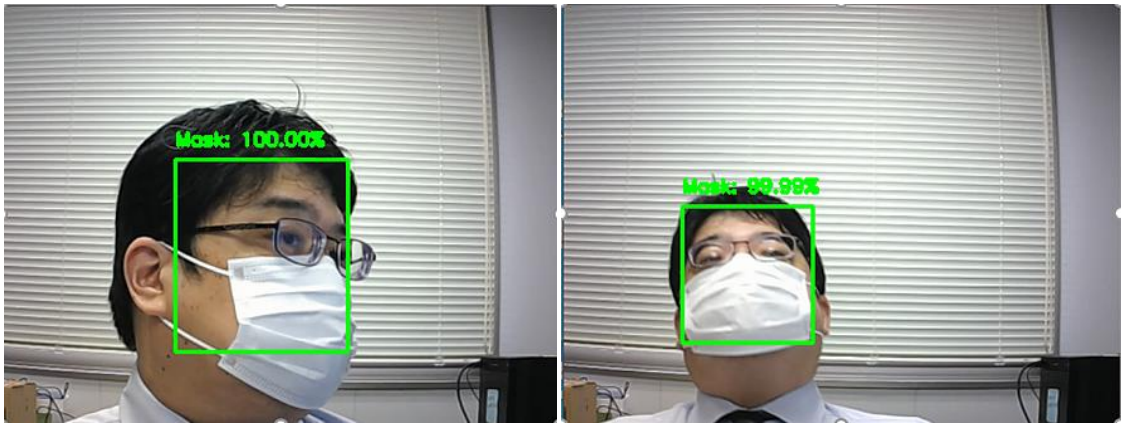
FIDTM 法による群衆のカウント

FIDTM 法（2021年発表）は，
それ以前の手法よりも，さまざまな大きさの顔を精度よく検出できるとされている

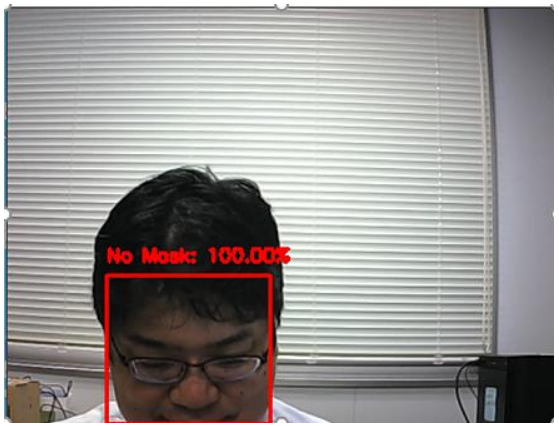
マスク有りの顔，マスク無しの顔検出

Chandrika Deb の顔マスク検出法では，次を同時に実行

- 顔検出
- マスク有りの顔とマスク無しの顔の画像分類



マスク有りの顔検出



マスク無しの顔検出

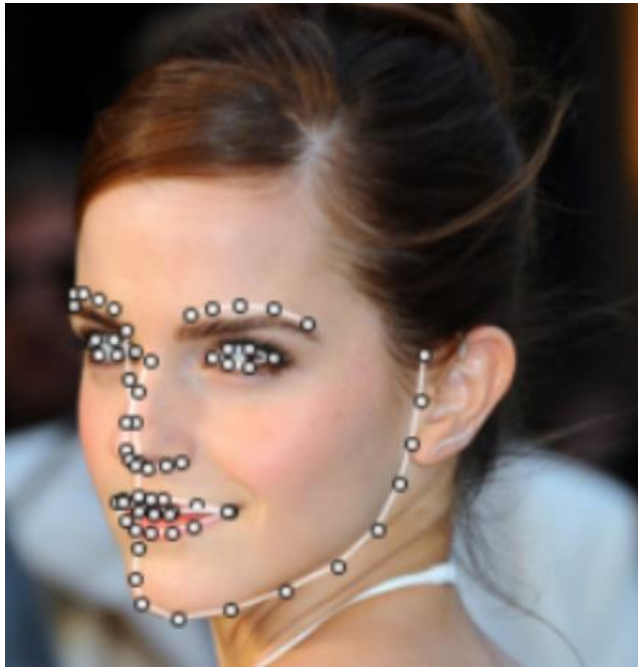
訓練データ：顔のデータセット（マスク有り: 2165 枚，マスク無し 1930 枚）

顔写真からの３次元再構成

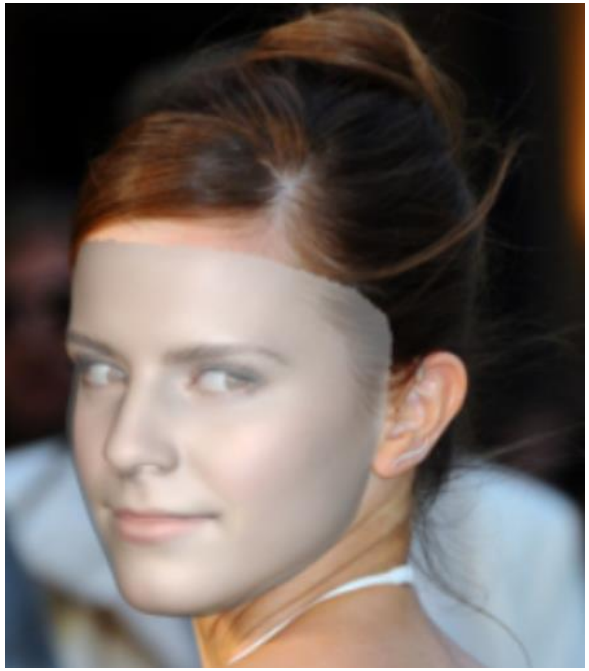
3DFFA 法（2022年発表）

- 元画像から，３次元の顔を生成（３次元再構成）
- 顔検出，顔ランドマーク（顔の目印となるポイント）の検出ののち，顔ランドマークに顔の３次元モデルをあてはめる

元画像

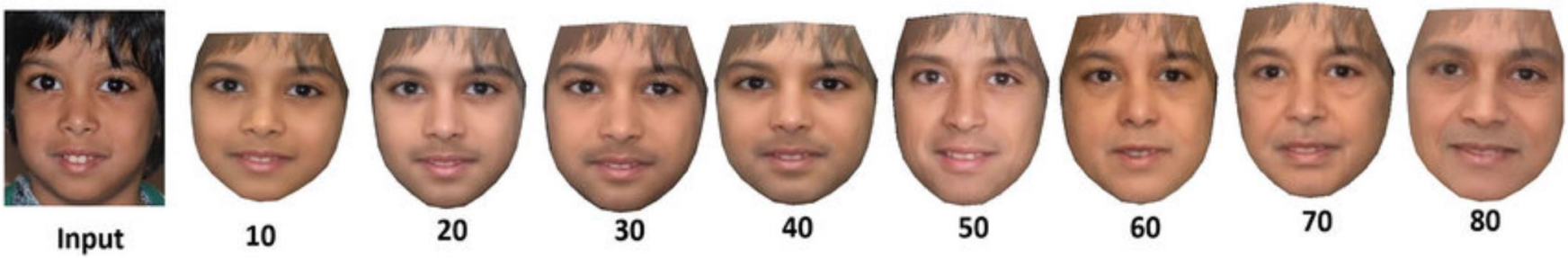


顔ランドマーク



３次元再構成

Face aging



Given the frontal view of a single image of a child, an ageing algorithm can generate his faces with varied ages

元の顔画像から，さまざまな年齢の顔画像を生成

Ali Elmahmudi, H. Ugail,
A framework for facial age progression and regression using exemplar face templates, 2020.

ここまでのまとめ

<u>顔情報処理のバリエーション</u>

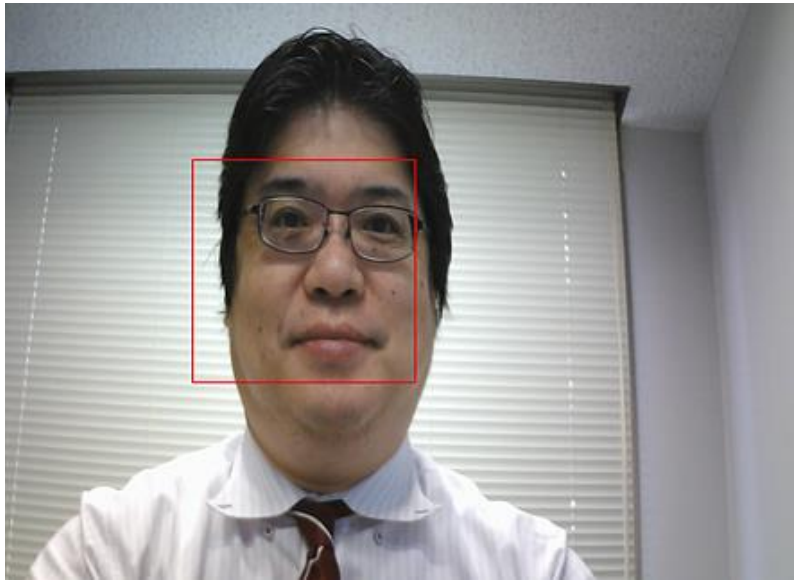
- 顔検出：物体検出の一環として顔を特定
- 顔のアラインメント：顔の向きや大きさの調整
- 顔ランドマーク：目、鼻、口などの重要な特徴点の位置
- 顔の数値化：顔特徴の数値化による本人確認、顔認識
- 性別、年齢の推定
- 表情の推定

<u>顔情報処理の重要性</u>

- セキュリティと監視
- 個人認証システム
- パーソナライズされたサービス
- 感情認識

8-3. ディープラーニングによる顔検出

顔検出



- 顔検出は、画像内で**顔と思われる領域を特定**
- **顔の特徴**（目、鼻、口など）を**識別**し、**それらが集まる領域を顔**と判断。

ディープラーニングによる顔検出の仕組み

- ニューラルネットワークを使用
- **畳み込みニューラルネットワーク（CNN）**が広く用いられている。

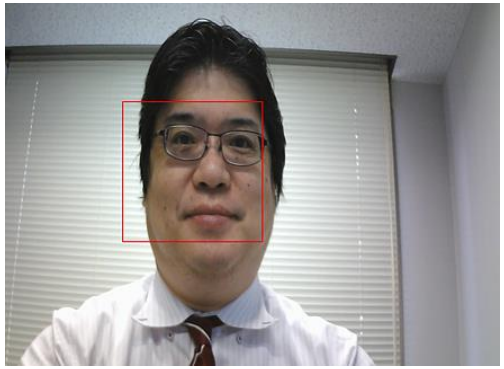
- **大量の画像データからの学習**
- 学習のあと、新しい画像に適用して顔検出を行う

顔検出での訓練データ

データ
（入力）

AI

顔検出結果



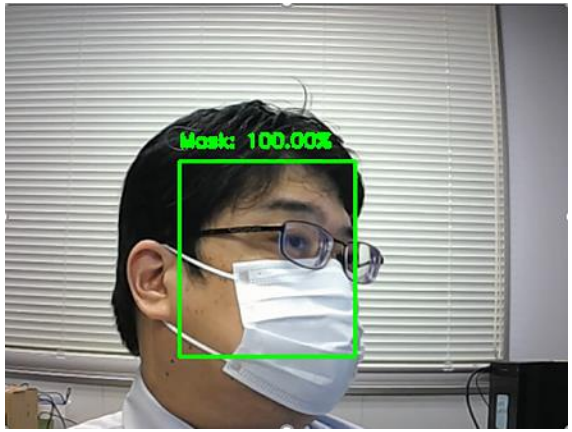
顔画像と
バウンディングボックス

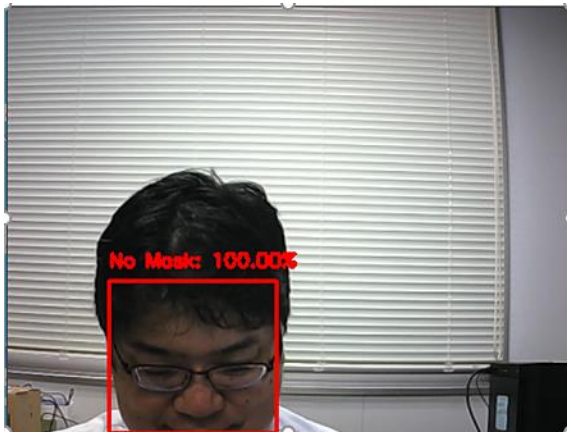
訓練データ

顔画像の例

顔検出の難しさ

- 照明のバリエーション
- さまざまな顔の向き
- 顔の個性
- 表情
- 付属物（例：メガネ、帽子）





ディープラーニングにより、マスクあり顔でも
高い精度で検出できるようになってきた

演習 2

マスク有り顔と、マスクなし顔の検出

【トピックス】

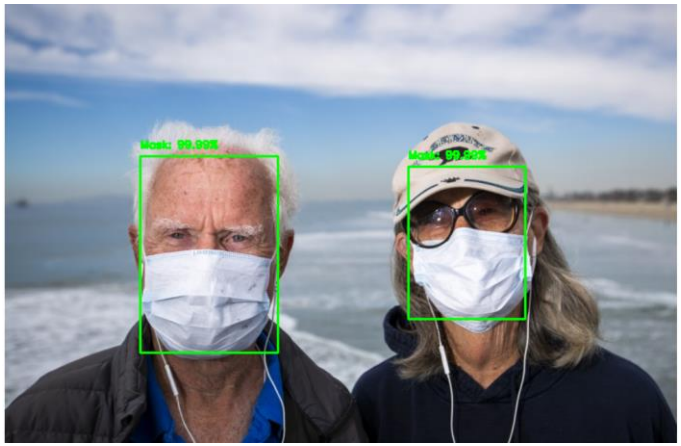
- **顔検出**
- **マスク有り顔と、マスクなし顔**
- **Google Colaboratory**

① Google Colaboratory のページを開く

https://colab.research.google.com/drive/1iYEI9O_cxWw4VyyafaYB4ne6CNdNJYeS?usp=sharing

② **マスク有り顔、マスクなし顔の顔検出**のプログラムや説明や実行結果が掲載されている

このページでは、顔画像を用いた学習も行う。自分で準備した画像で試すことも可能。



学習で用いた顔画像の顔のサイズにより、
大きい顔のみが検出されている

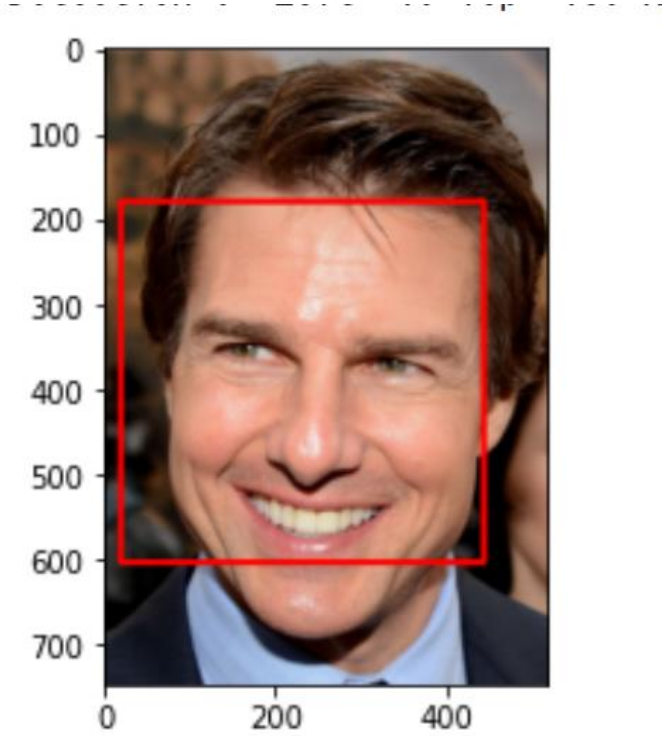
自習４

- まず、自分で顔画像のファイルを準備する
- 前ページで説明したデモページを使う
- このとき、このページに書いている注意事項をよく読んでください。**ランタイムのタイプ**を「**T4 GPU**」に設定することを説明しています。
- メニューの「ランタイム」で「すべてのセルを実行」
- コードセルが上から順に実行されるので確認

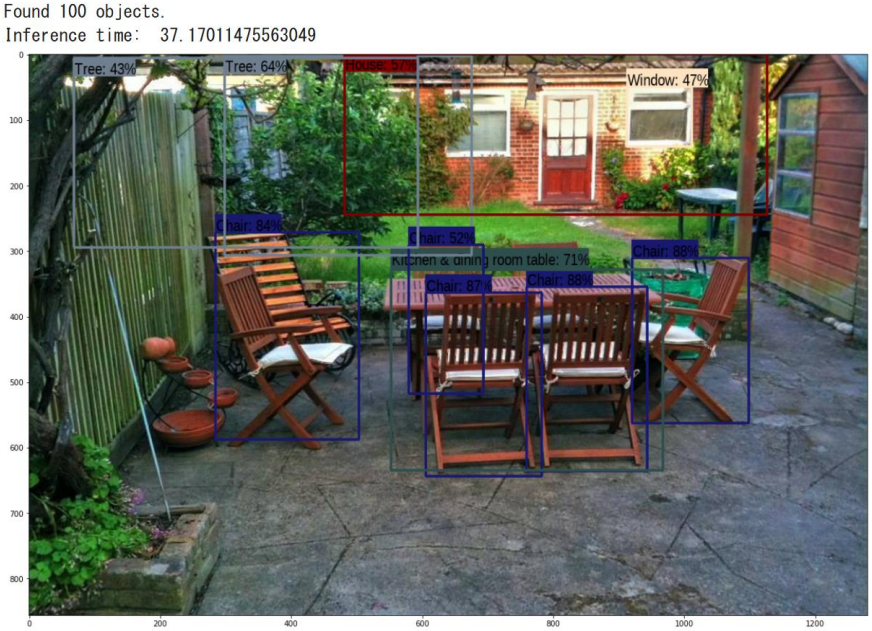
- 「５．画像のアップロード」のセルの実行が始まったら、画像をアップロードして**顔検出**を試してみる

https://colab.research.google.com/drive/1iYEI9O_cxWw4VyyafaYB4ne6CNdNJYeS?usp=sharing

顔検出と物体検出



顔検出

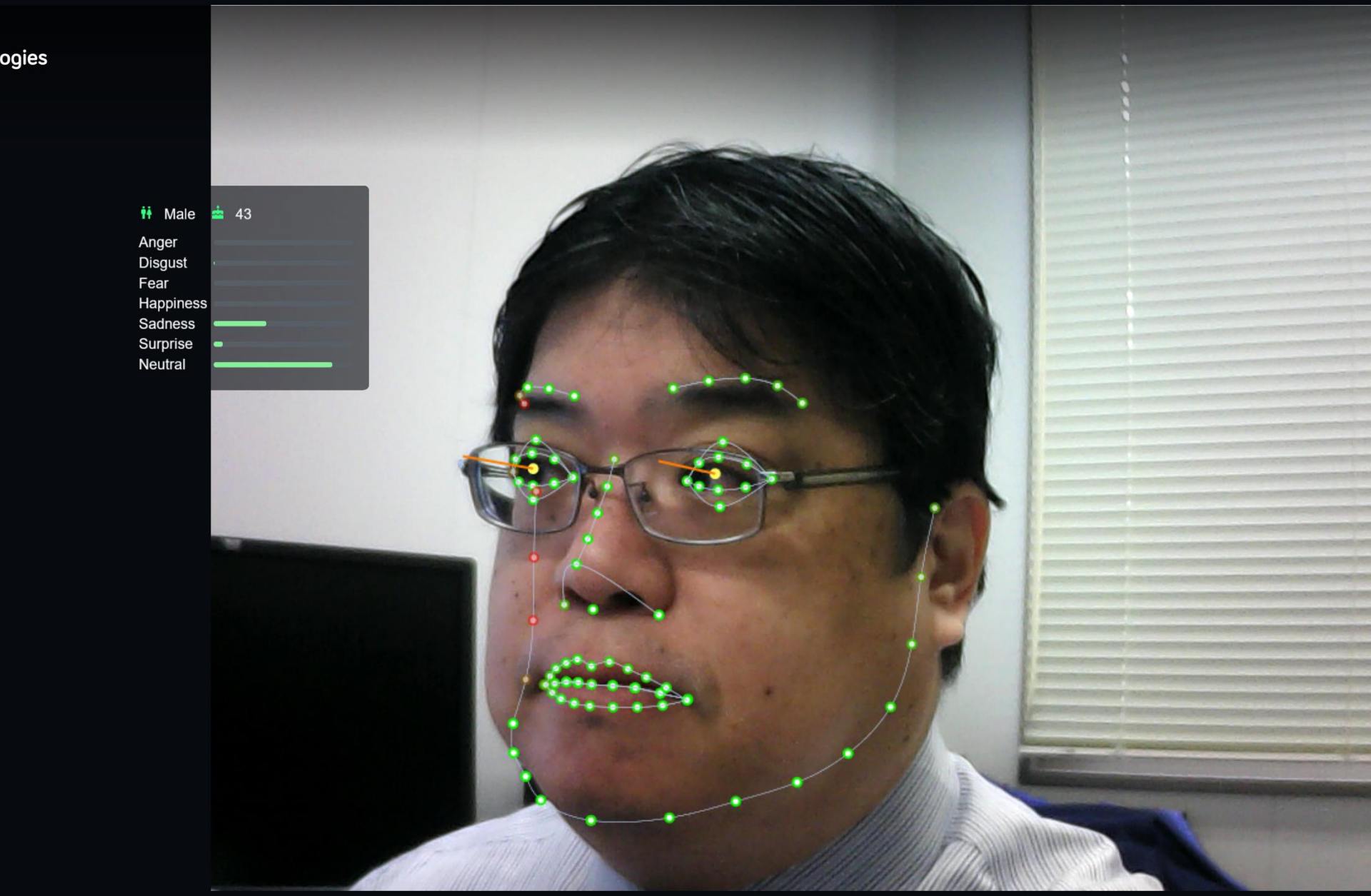


物体検出

顔検出と物体検出

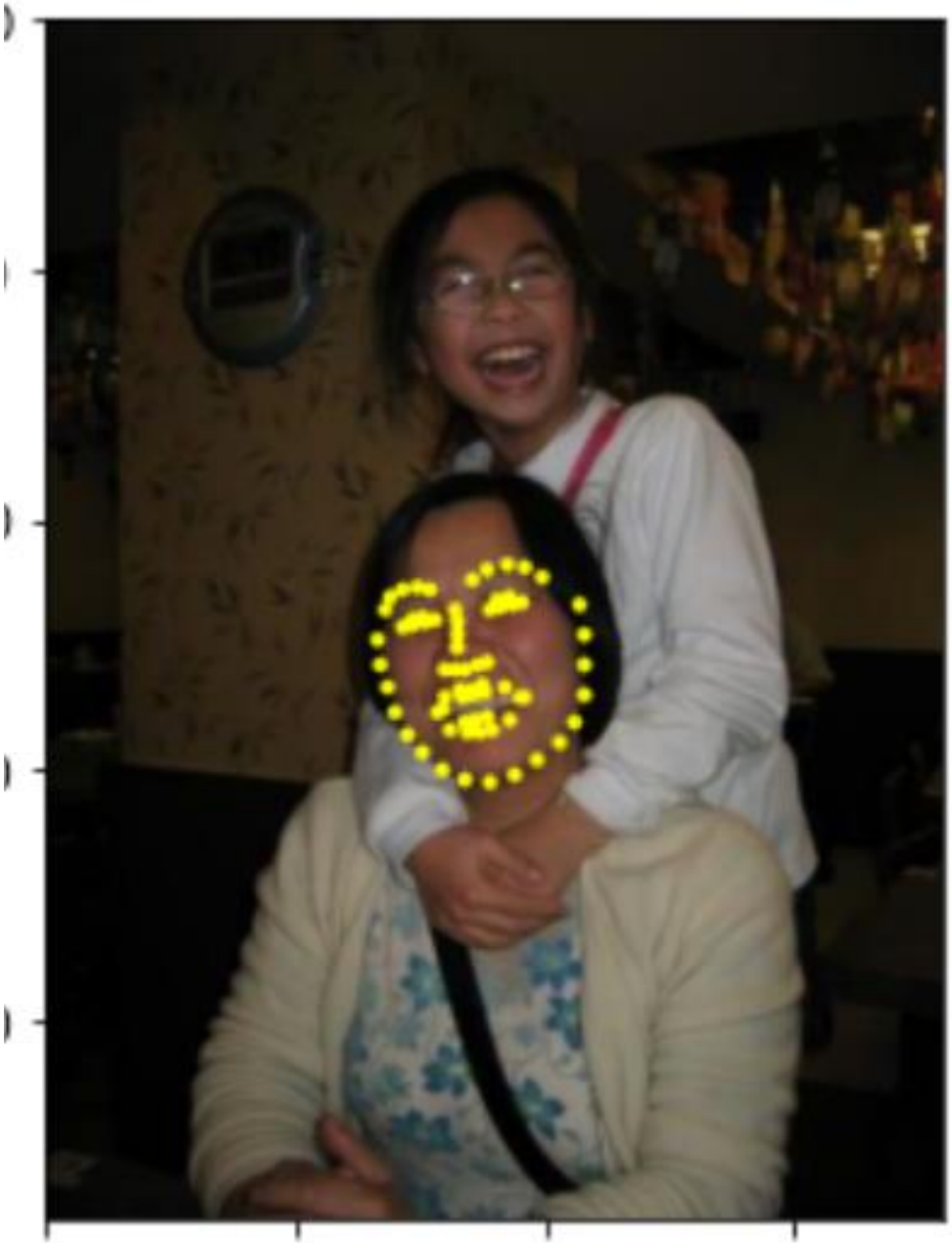
顔検出	物体検出
顔のみの検出	さまざまな種類の物体を検出
顔の位置とサイズを識別	物体の種類、位置、サイズを識別
特徴抽出の対象は、顔の重要な特徴点である顔ランドマーク	特徴抽出の対象は、形状、色、テクスチャなどさまざま（多様な物体を検出するため）
代表手法: RetinaFace, MTCNN	代表手法：YOLO

8-4. 顔ランドマーク



https://visagetechnologies.com/demo/

顔ランドマーク



顔ランドマーク

顔の重要な特徴点（目、鼻、口など）の位置を示すポイント

【用途】

顔の表情

顔の向き

顔識別　など

顔ランドマークの検出

畳み込みニューラルネットワークによる**顔ランドマークの検出**は 2013年に登場

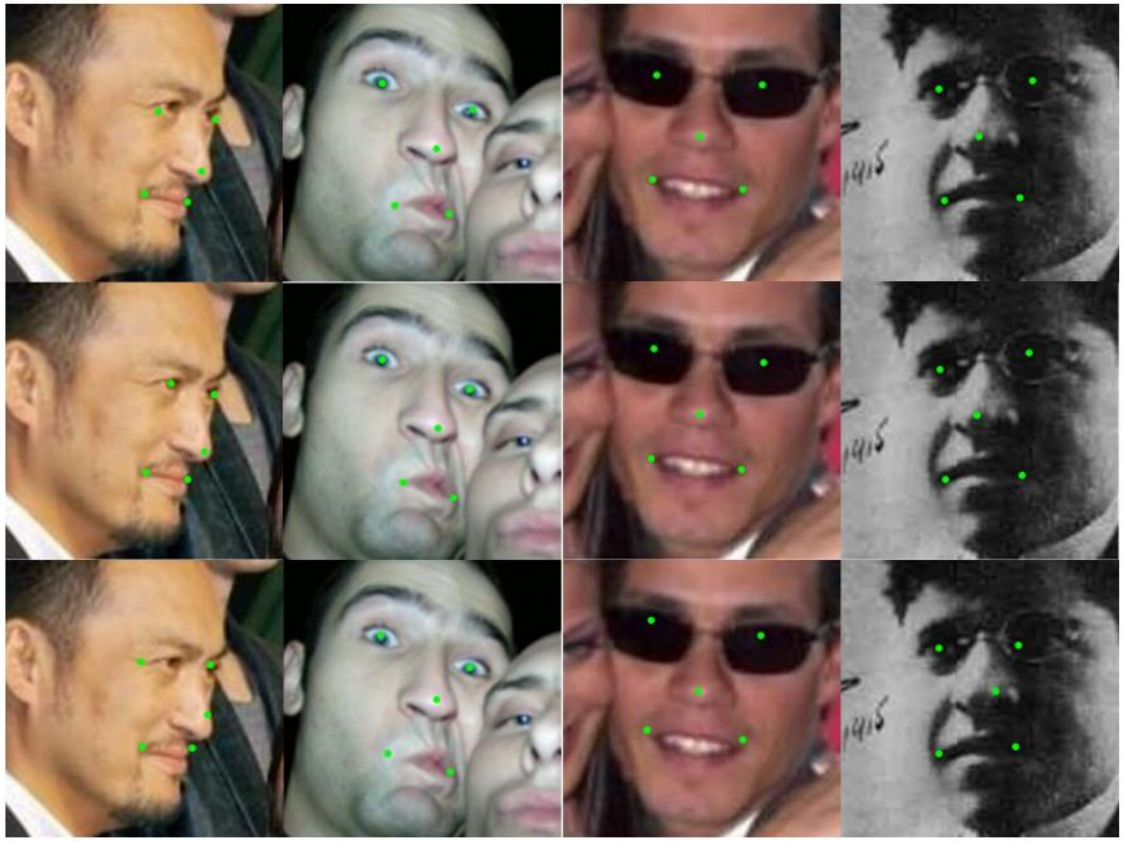


Figure 1: Examples of facial point detection. First row: ini-

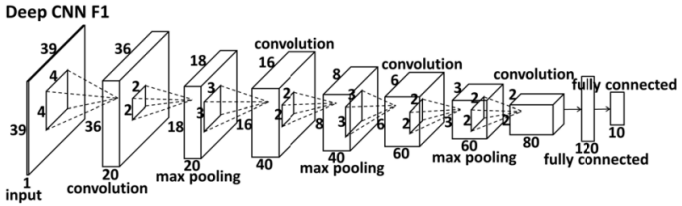


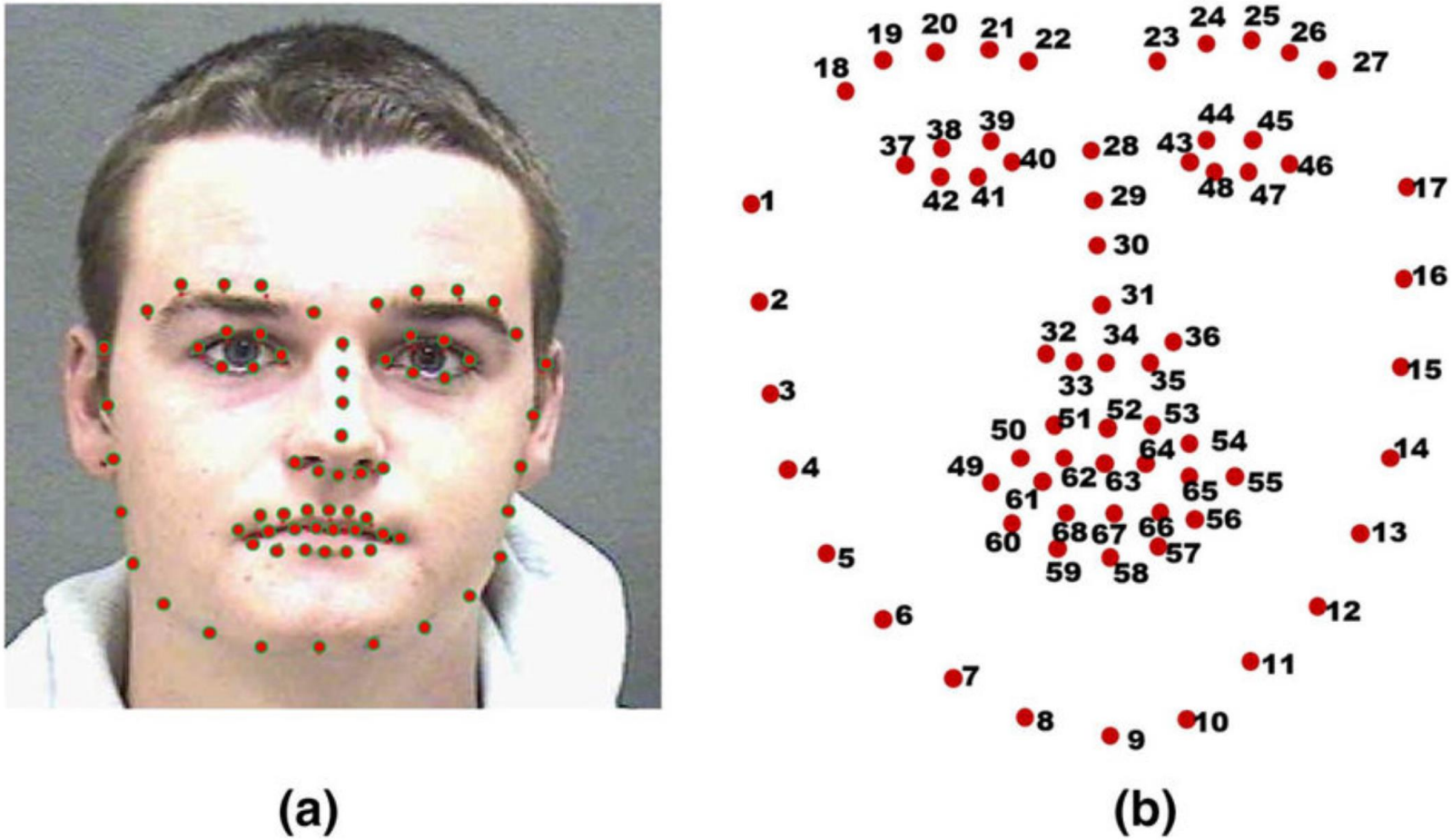
Figure 3: The structure of deep convolutional network F1.

畳み込みニューラルネットワークを使用

Yi Sun, Xiaogang Wang, Xiaoou Tang
Deep Convolutional Network Cascade for Facial Point Detection, 2013

６８ランドマーク

顔の 68個のランドマーク。目、眉毛、鼻、口、顔の輪郭など



Identification of facial landmarks using Dlib. a Facial landmarks. b The position and order of 68 points on the face

https://www.researchgate.net/figure/dentification-of-facial-landmarks-using-Dlib-a-Facial-landmarks-b-The-position-and_fig2_343699139

8-5. ディープラーニングによる顔認識

顔認識

顔認識は、画像または動画から人の顔を認識

比較

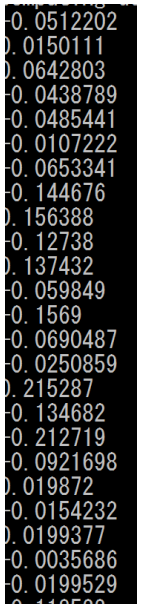
データベース

データベース内の顔との比較により、人物を特定

顔の数値化

- **顔の数値化**は，複数の数値（ふつう１００以上）の組み合わせ
- **顔画像**から，**顔ランドマーク**を求め，顔を数値化する

顔検出，
顔ランドマーク



数値化

同一人物かの判定（顔の数値を利用）

２つの別の写真あるいは動画を**照合**し，**同一人物である**
かを判定する

顔の数値

比較
(距離計算)

顔の数値

同一人物である
または
同一人物でない

顔認識（顔の数値を利用）

精度の向上のため,
同一人物の
多数の写真と比較

顔の数値

比較（距離計算）

本人である
または
本人でない

顔の数値

顔の数値

顔の数値

距離学習

距離学習は，ディープラーニングでの学習で、距離に関する学習を行うこと。顔認識で使用。

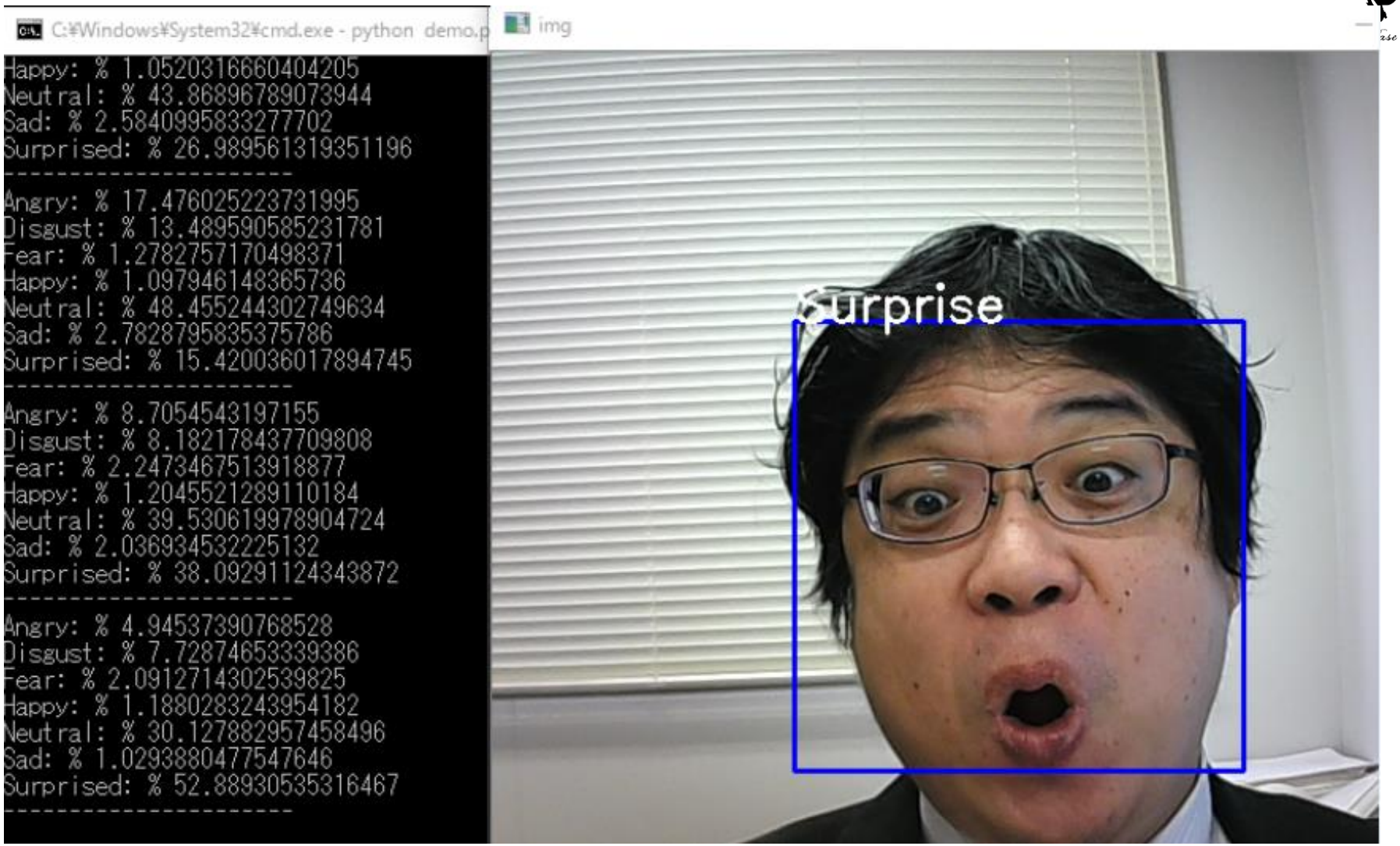
目的：同一人物の顔を近く、異なる人物の顔を遠くなるようにすること

仕組み：畳み込みニューラルネットワークなどを使用

効果：顔認識、その他、画像分類やパターン認識など

8-5. ディープラーニングによる感情認識その他

感情認識

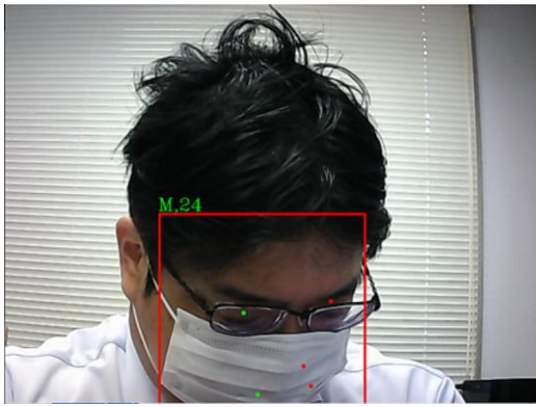


- ここでは， **７種の表情** Angry, Disgust, Fear, Happy, Neutral, Sad, Surprised の**それぞれの確率**を推定
- https://github.com/ezgiakcora/Facial-Expression-Keras で公開されている成果物

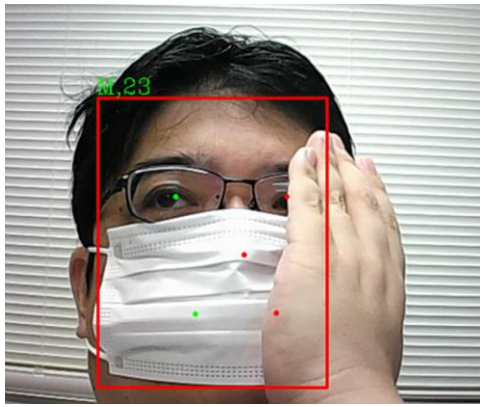
感情認識

- 感情：喜び、悲しみ、驚き、怒り、恐怖、嫌悪など
- ディープラーニング、顔ランドマークを使用

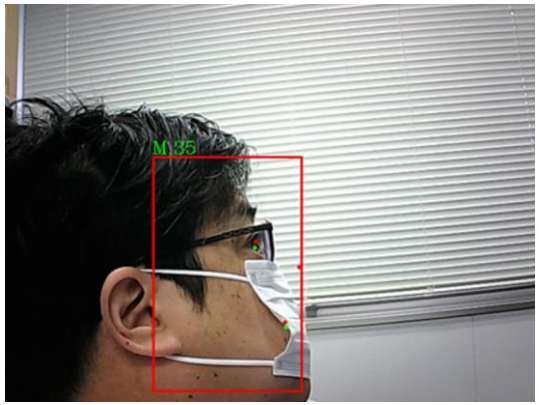
年齢推定，性別推定の例



推定結果：24歳男性



推定結果：23歳男性



推定結果：35歳男性

InsightFace による処理結果

【関連情報】

https://www.kkaneko.jp/ai/win/insightface.html
https://colab.research.google.com/drive/1S55yEFiQpdIRdjWbdH0zzEYD5VAfklHd?usp=sharing

自習５

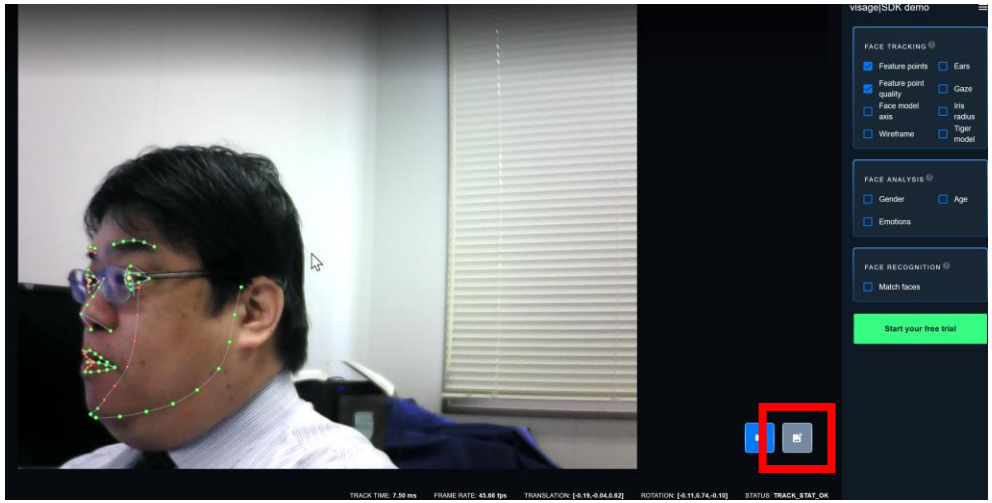
- まず、**自分で顔画像のファイルを準備**する
- 次のページを開く。Visage Technologies 社のデモページ

https://visagetechnologies.com/demo/

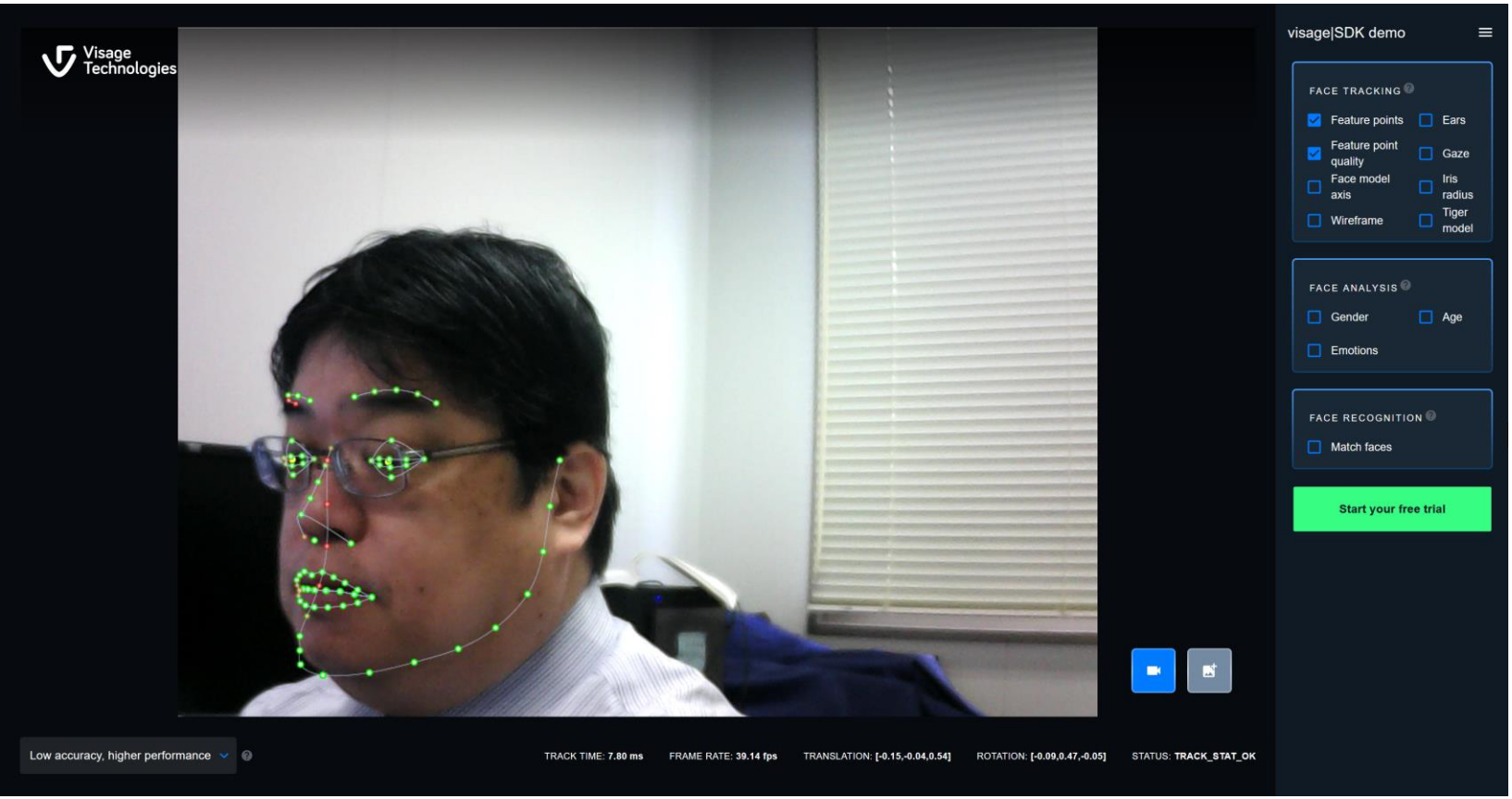
- 「**Try out demo**」をクリック

Try out demo

- **画像をアップロード**するので、**右下のボタン**をクリックして、画像をアップロード。（パソコンのカメラにも対応）



さまざまな設定



Gaze
視線

Gender
性別

Emotions
感情

Age
年齢

全体まとめ①

<u>顔情報処理のバリエーション</u>

- 顔検出：物体検出の一環として顔を特定
- 顔のアラインメント：顔の向きや大きさの調整
- 顔ランドマーク：目、鼻、口などの重要な特徴点の位置
- 顔の数値化：顔特徴の数値化による本人確認、顔認識
- 性別、年齢の推定
- 表情の推定

<u>顔情報処理の重要性</u>

- セキュリティと監視
- 個人認証システム
- パーソナライズされたサービス
- 感情認識

全体まとめ②

ディープラーニングによる顔情報処理

- **顔検出**：畳み込みニューラルネットワーク（CNN）を使用して顔を特定。
- **顔ランドマークの検出**：CNNによる顔の重要な特徴点の検出。
- **顔認識**：データベース内の顔との比較により人物を特定。
- **感情認識**：ディープラーニングと顔ランドマークを使用して感情を推定。

など