cs-13. 人工知能体験

（コンピューターサイエンス）

金子邦彦

資料中の図などは Google Nano Banana 2 を用いて作成

「コンピューターサイエンス」第１３回の内容

AIは、コンピュータに人間のような知能・知識・学習の能力を持たせる技術

AI（人工知能）

コンピュータビジョン（画像から意味を理解）

(1) 画像分類

ネコ

(2) 物体検出

(3) セグメンテーション

画像全体 → 位置（矩形） → 画素単位

顔情報処理
- 顔検出
- 顔ランドマーク検出
- 表情推定
- 3次元再構成

モデルのパラメータ調整

信頼度しきい値・IoUしきい値

検出精度（見逃し）

検出精度（見逃し）

誤検出

└ しきい値の調整で、検出精度と誤検出のバランスが変わる

13-1. 人工知能、機械学習、ディープラーニング

AIの基礎

AIの定義

コンピュータが人間のような知的能力を持つことを目指す技術

AIの3要素

①知能
思考や判断
などの能力

→

②知識
情報を収集し、
処理する能力

→

③学習
知的な能力が
向上する能力

3要素が相互に作用し、知的な処理を実現する

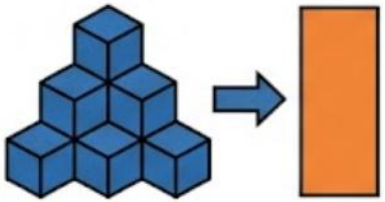
AIの利点

AIの3要素（知能・知識・学習）が実現する成果

1
24時間365日稼働

休まず連続して動作

2
大量データを高速処理

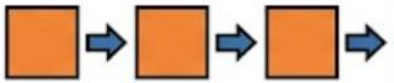


膨大な情報を短時間で

3
細かいパターンを検出

人が見落としがちなパターンを検出

4
反復作業を効率化



繰り返し作業を自動化

機械学習

機械学習とは、ルールを教えるのではなく、データからコンピューター自身にパターンやルールを見つけさせる仕組み

これまでのやり方

人がルールを考える

↓

コンピューターは
ルール通りに動くだけ

人が答え方を全部教える

機械学習

データを与える

↓

コンピューターが自分で
パターンやルールを
見てつける

↓

パターン・ルールが出てくる

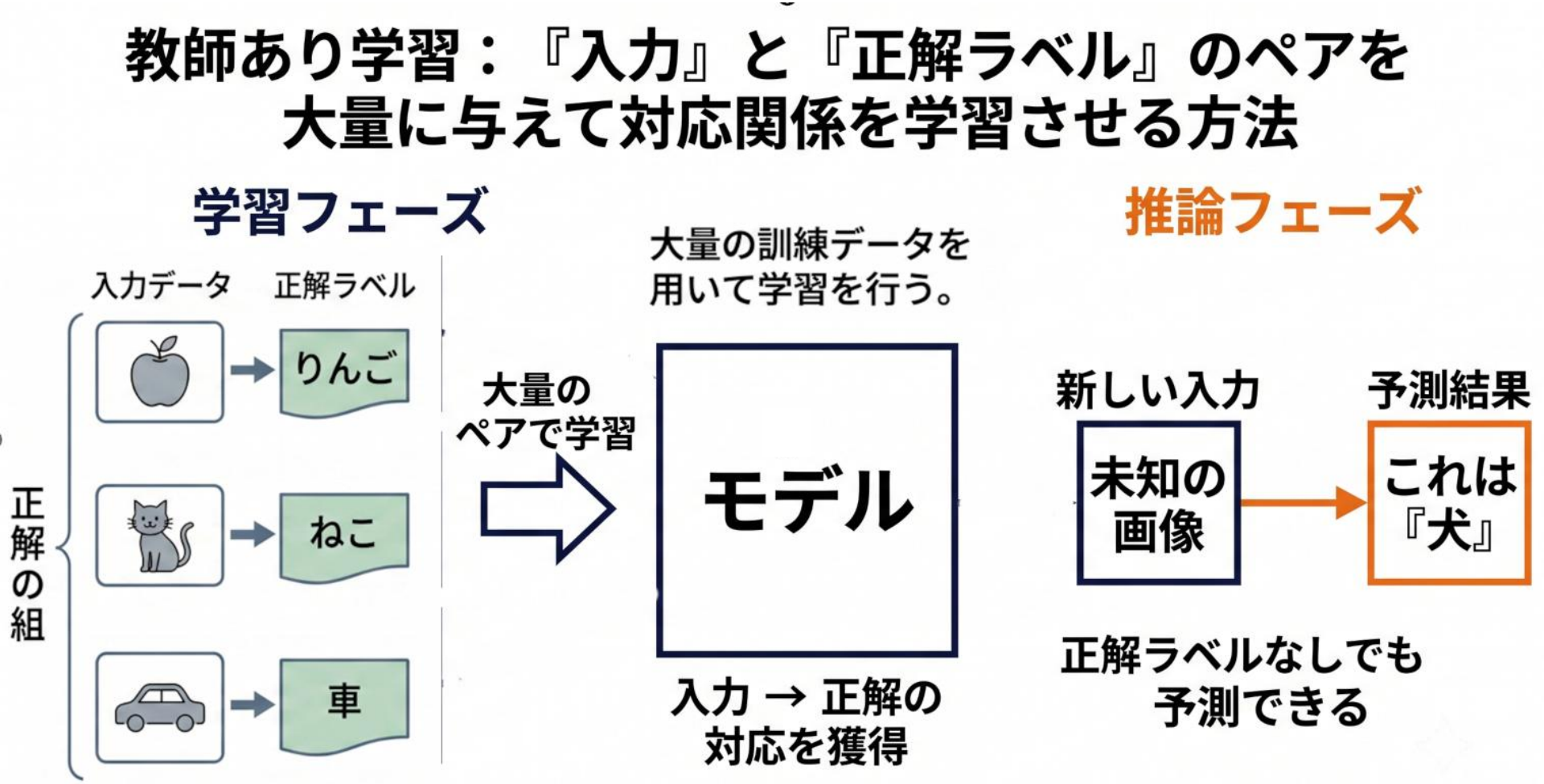
人はデータを渡す

自分でルールを
書いていないのに、
コンピューターが
分類してくれた

これは猫っぽい

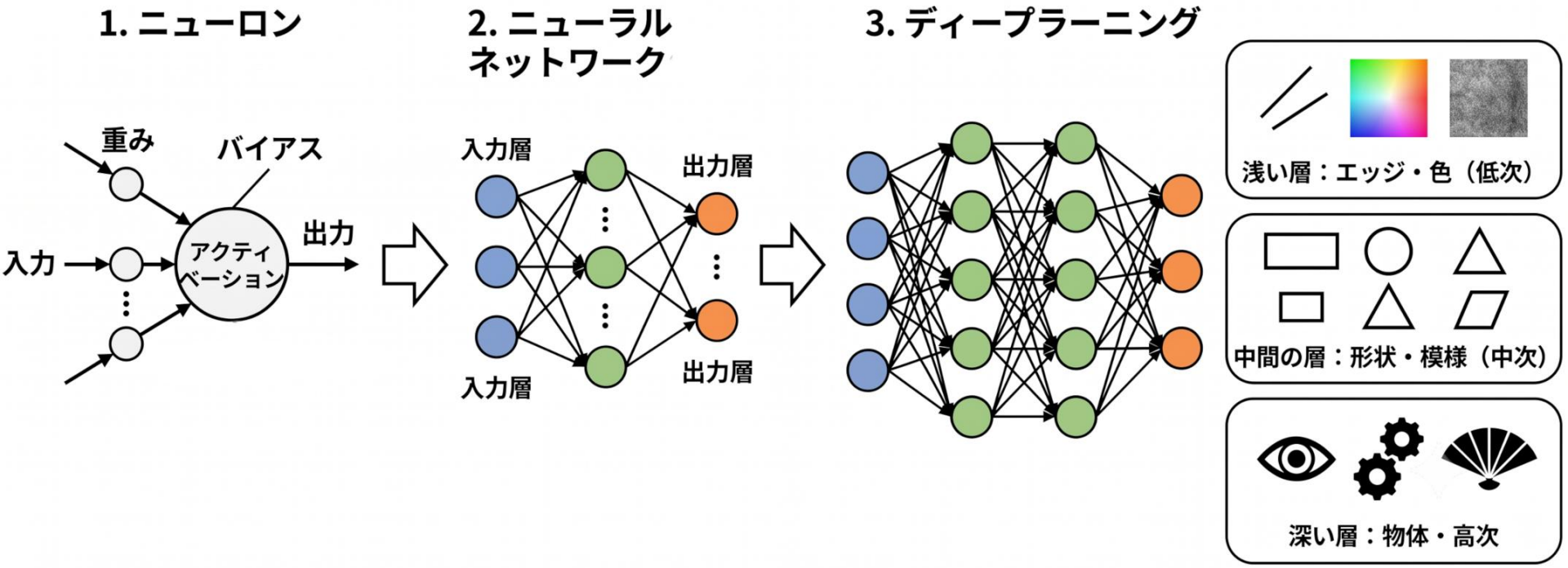
与えるのはデータ

機械学習の「教師あり学習」の仕組み



ニューロン、ニューラルネットワーク、ディープラーニング

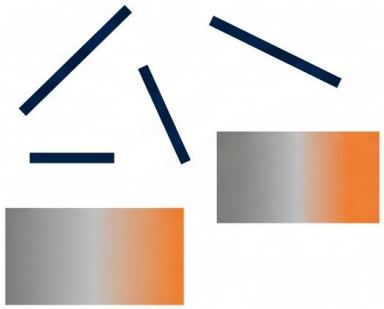
ニューロンを基本単位とし、その層を重ねた構造がニューラルネットワークであり、層を深く重ねたものがディープラーニングである



ディープラーニング

ニューラルネットワークは、層が深くなるほど単純なパターンを組み合わせて複雑なパターンをとらえる

浅い層



エッジ・色の変化など単純で局所的なパターン（低次の特徴）

組み合わせる

中間の層

エッジを組み合わせた形状・模様のパターン

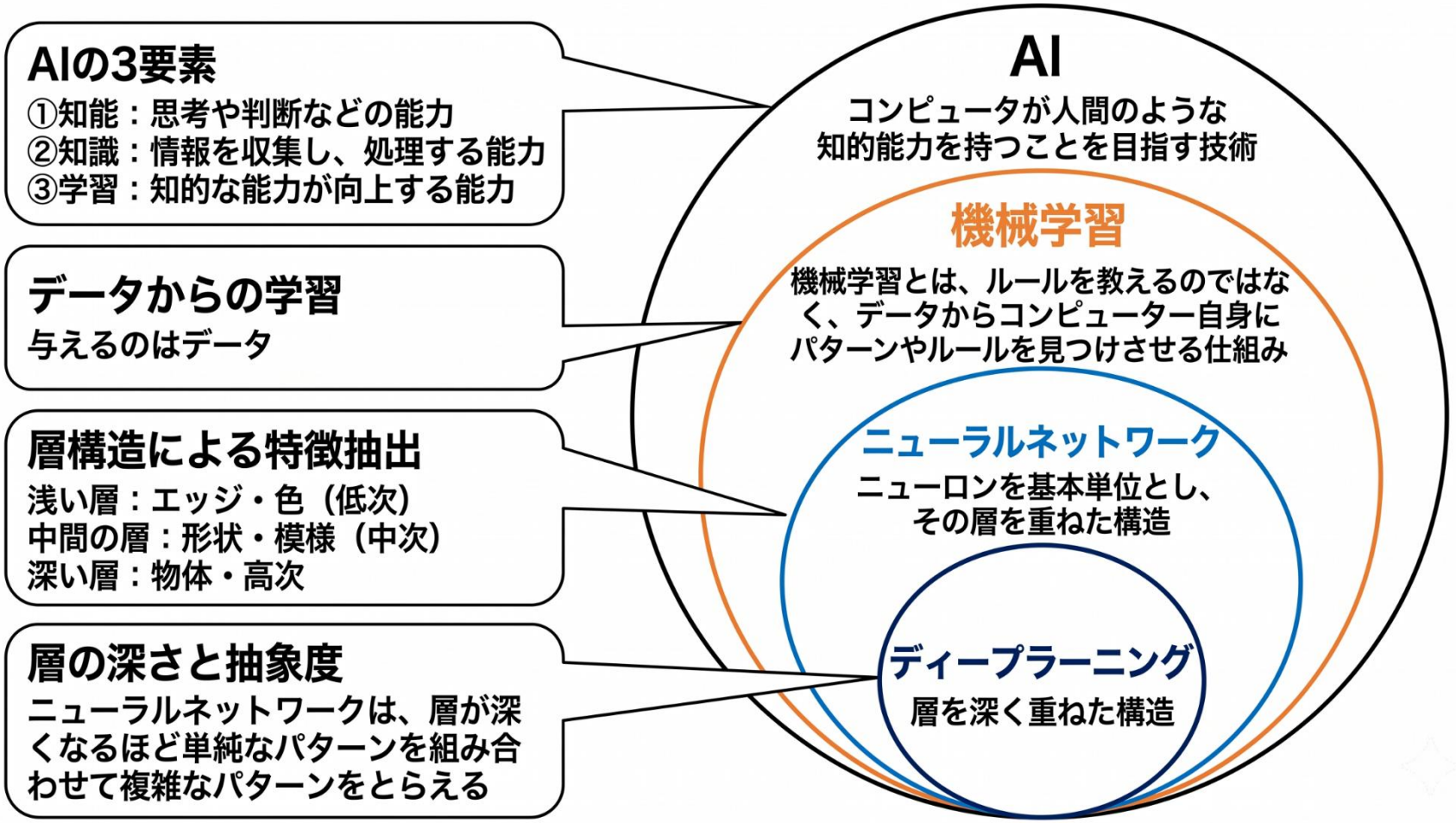
さらに組み合わせる

深い層

形状を組み合わせた物体の部分など複雑なパターン（高次の特徴）

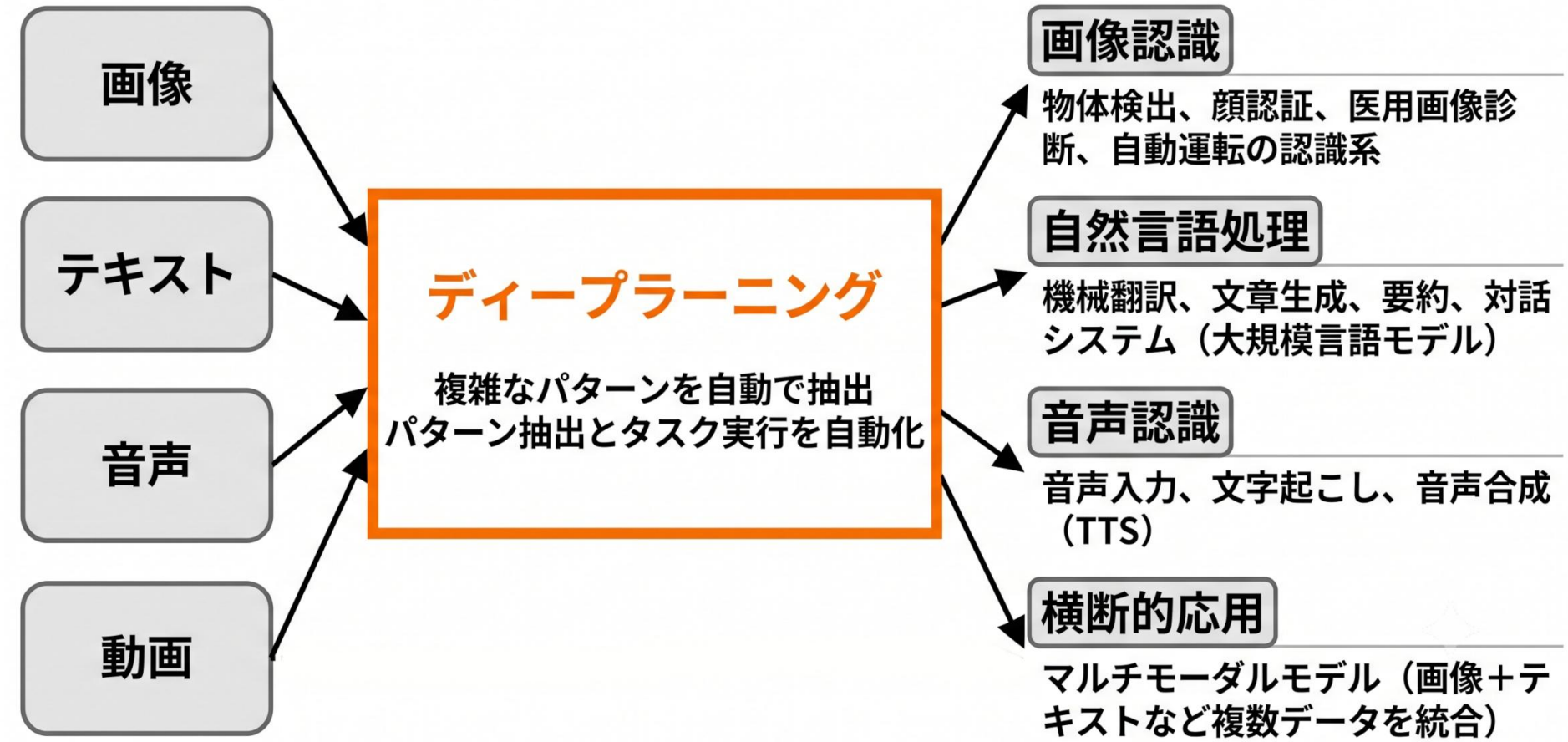
単純・局所的 → 層が深くなる → 複雑・全体的

AI・機械学習・ディープラーニングの関係



AIは最上位の概念であり、機械学習はその実現手法の一つである。ニューラルネットワークは機械学習の手法の一つであり、層を深く重ねたニューラルネットワークがディープラーニングである

ディープラーニングの応用分野



13-2. コンピュータビジョン

コンピュータビジョン

コンピュータビジョン ＝ コンピュータが『視覚』を持つこと

入力

実世界の様子
（カメラ映像・画像）

コンピュータが理解

理解

- 何があるか（認識）
- 何が起きているか（状況把握）
- 何が起きそうか（予測）

理解を活用

活用

現実世界での判断・行動に役立てる

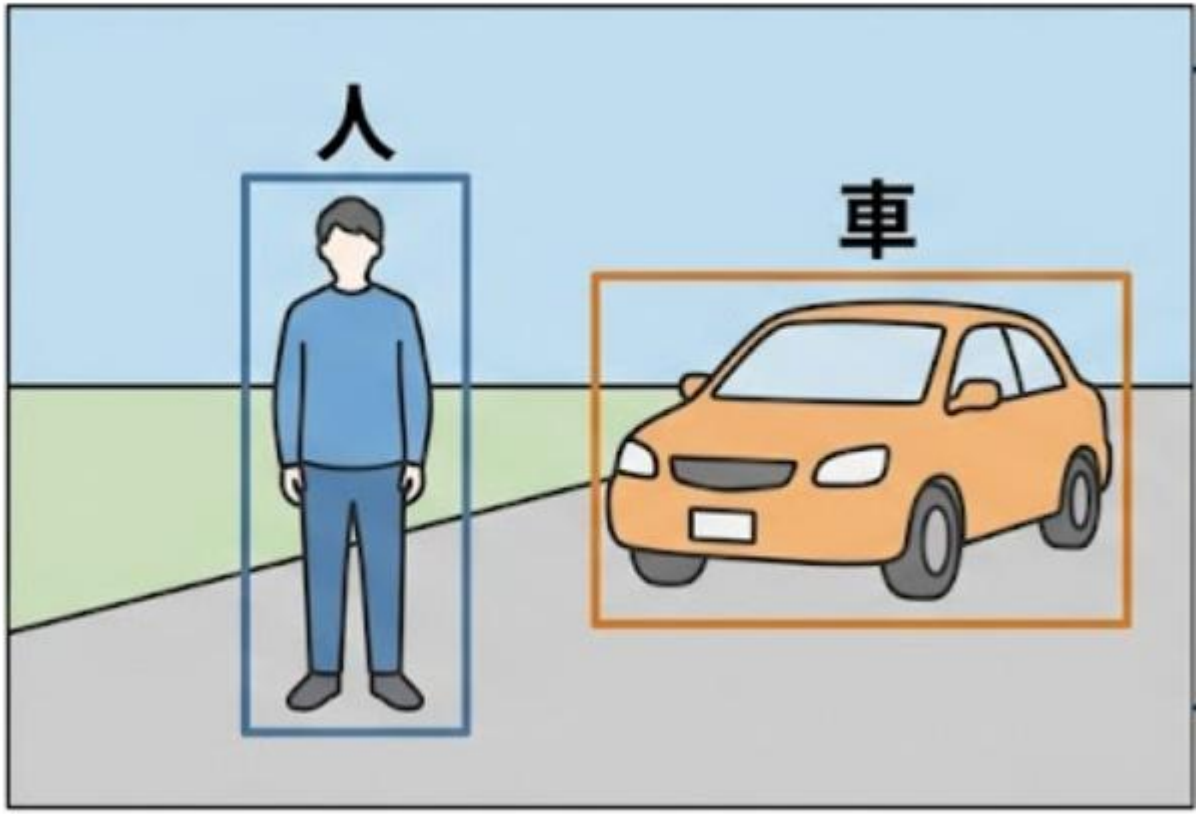
例：自動運転、検査の自動化

画像分類

画像分類：画像全体を見て『何が写っているか』を1つのクラスラベルで答える

ポイント：出力は『どこに写っているか』ではなく、『何が写っているか』というラベル
画像全体に対して1つの答えを返す（物体検出のように位置は示さない）

物体検出

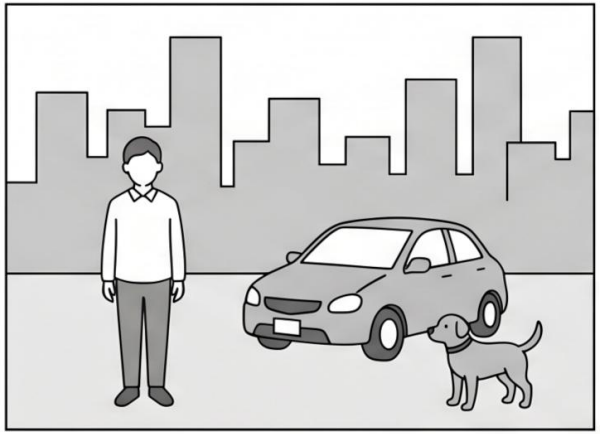


物体の位置と大きさを
四角形で特定する

物体検出

物体検出：画像の中の「何が」「どこに」「どの大きさで」あるかを特定する

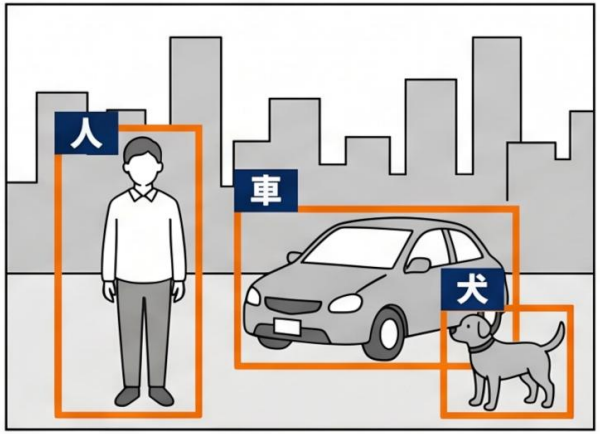
入力



物体検出

位置・大きさ・種類を同時に特定

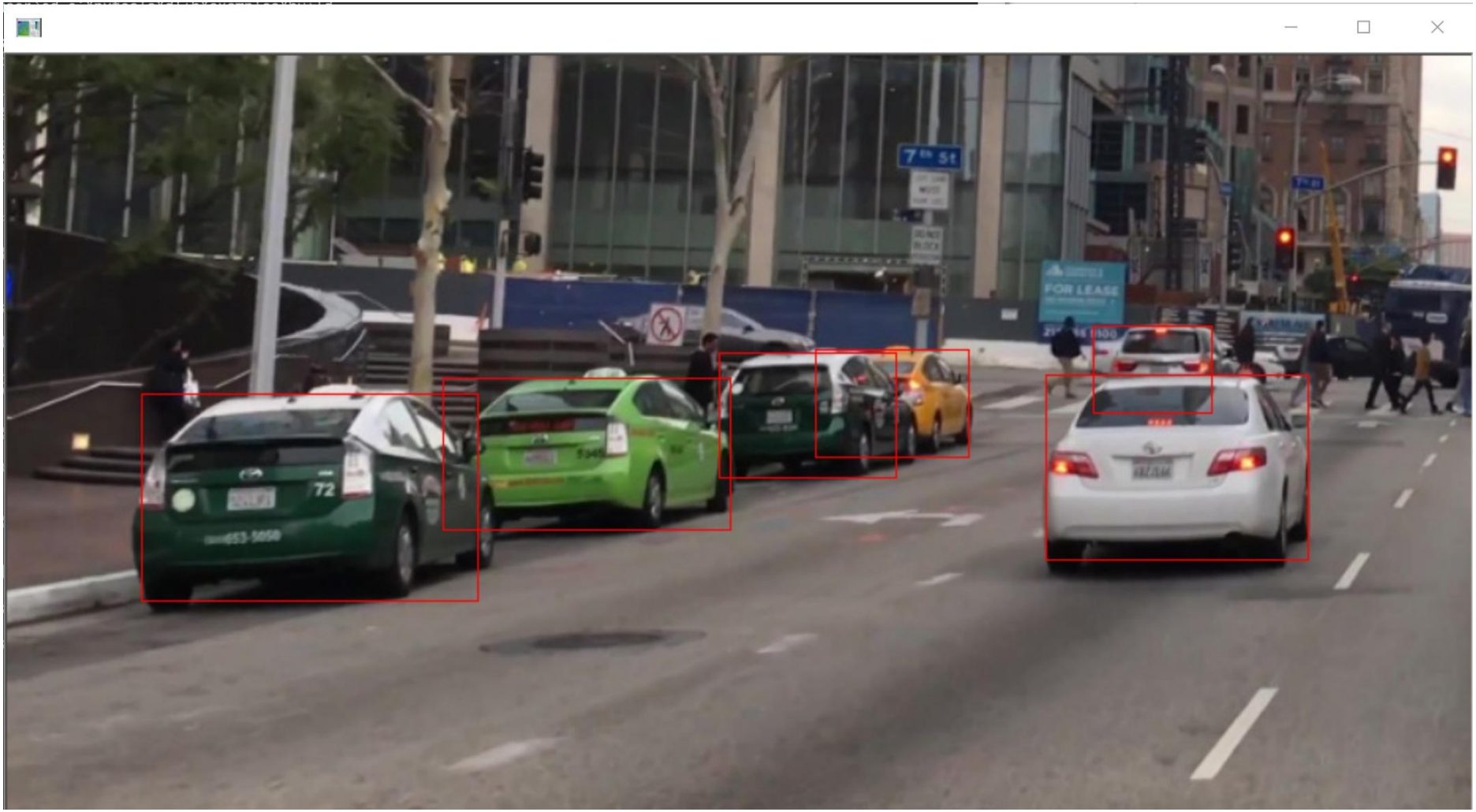
出力



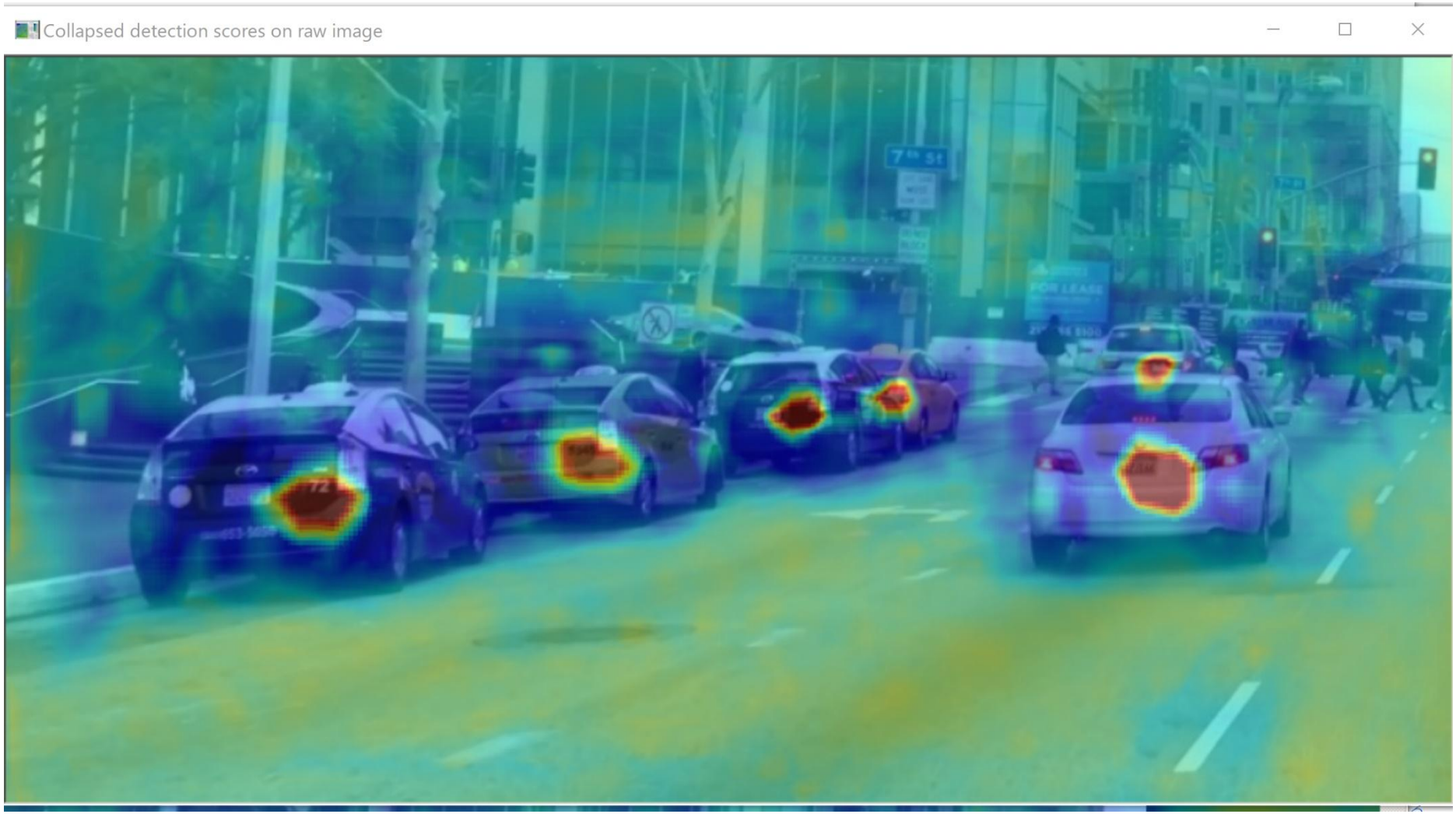
バウンディングボックス＝物体を囲む最小の四角形

出力されるのは「物体を囲む四角形（位置と大きさ）」と「ラベル（種類）」
四角形は物体を囲む最小の長方形で、位置と大きさを表す

物体検出

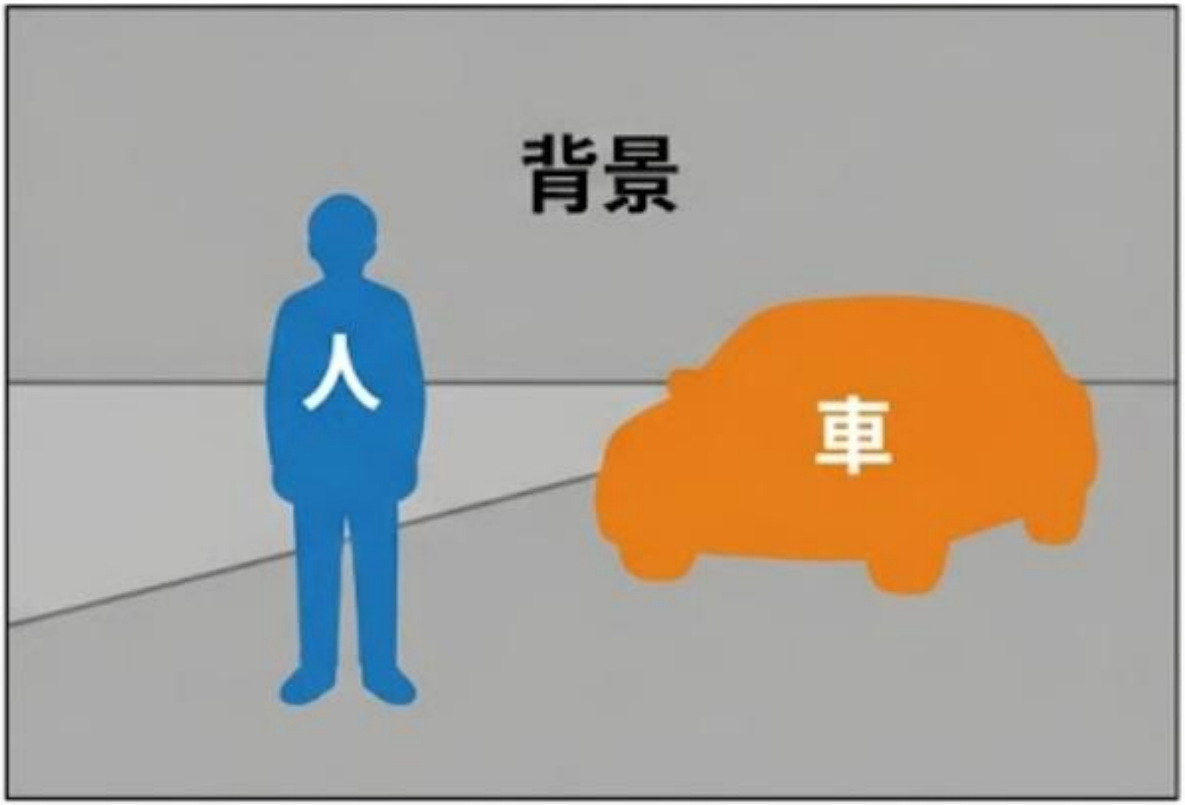


物体検出における特徴マップ



セグメンテーション

画像を『背景・人・物体』などの領域ごとにピクセル単位で分類する技術



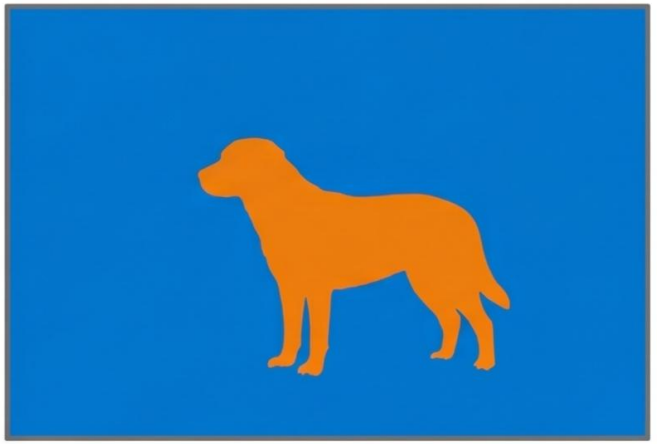
輪郭に沿ってピクセル単位で塗り分ける

セグメンテーション

セグメンテーション：画像を意味のある領域に分割し、物体の形を詳細に解析する

入力画像

領域に分割



物体の形を領域として抽出

位置が分かる

物体が画像のどこにあるかを正確に把握できる

形・大きさを数値化

物体の形や大きさを数値として測る基礎になる

セグメンテーションの例

元画像

セグメンテーションの結果

セグメンテーションの種類

※ パノプティックセグメンテーションは、セマンティックセグメンテーションとインスタンスセグメンテーションの統合

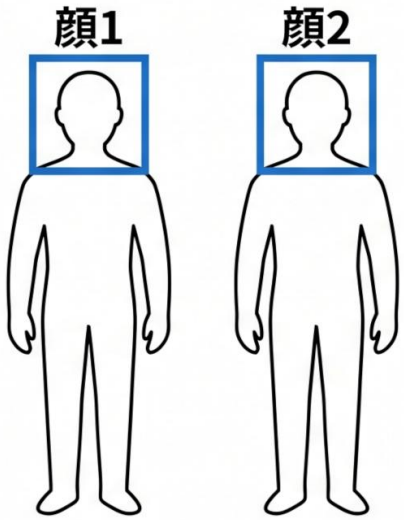
セマンティックセグメンテーションの例

インスタンスセグメンテーションの例

顔情報処理：顔検出と顔ランドマーク検出

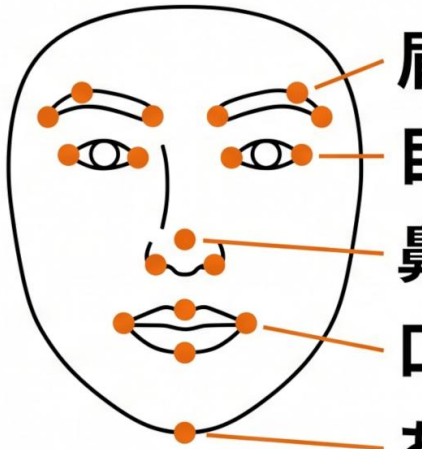
顔情報処理は『顔検出』→『顔ランドマーク検出』の順に進む

1. 顔検出



- 画像の中から顔を見つける
- 顔の数を数える
- 顔情報処理の最初の処理

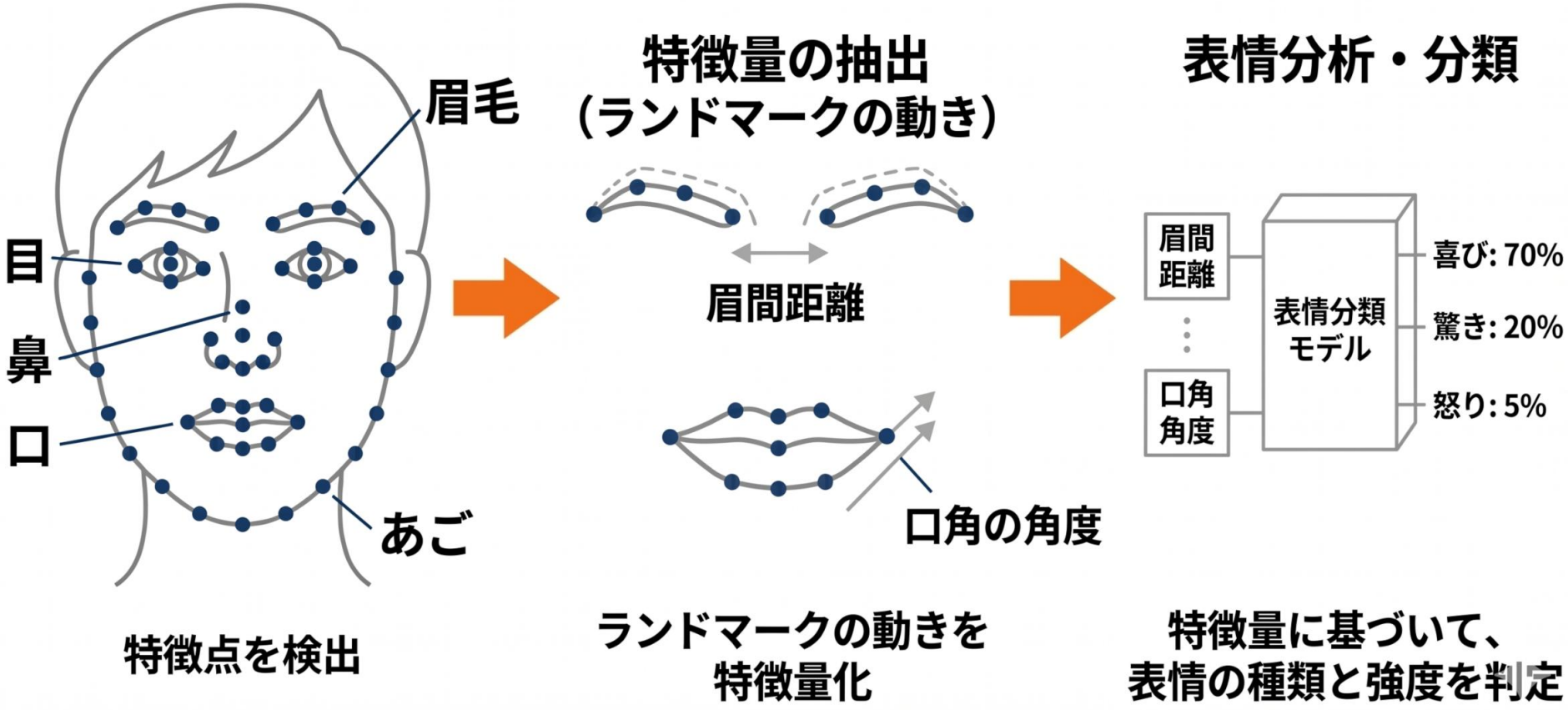
2. 顔ランドマーク検出



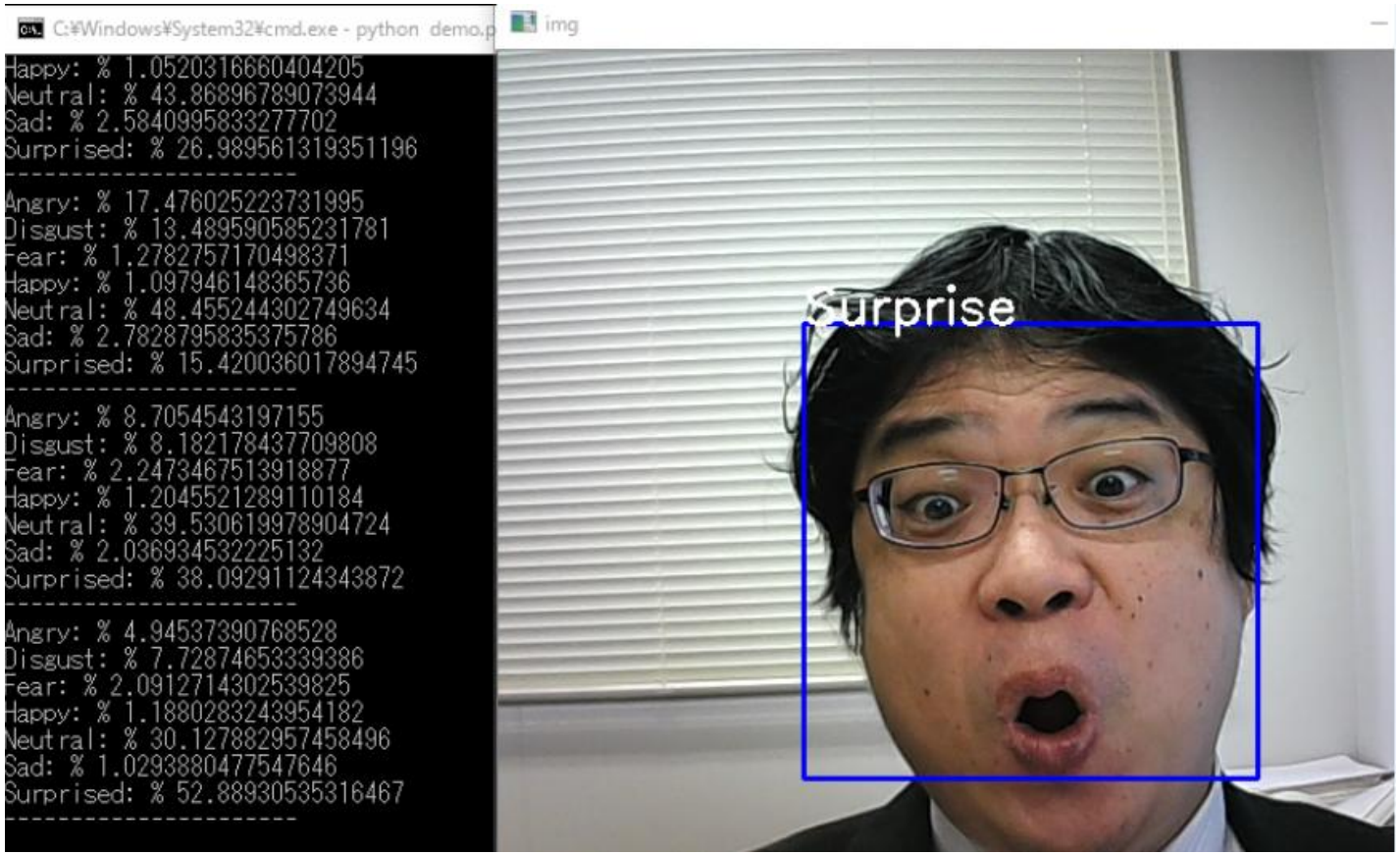
- 目，鼻，口，眉，あごなどの位置を特定
- 用途：位置合わせ/表情分析/人物特定

表情の推定

検出された顔ランドマークの相対的な位置関係から、表情や感情を分析する



表情の推定

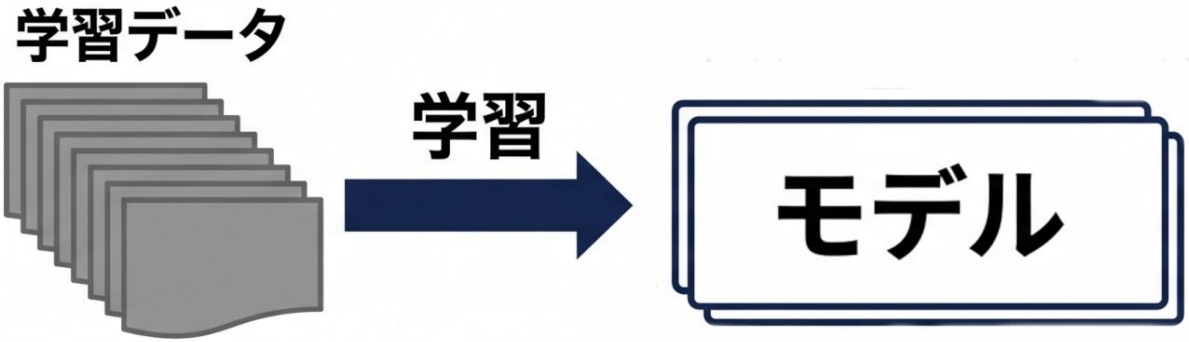


- ここでは，**７種の表情** Angry, Disgust, Fear, Happy, Neutral, Sad, Surprised の**それぞれの確率**を推定
- https://github.com/ezgiakcora/Facial-Expression-Keras で公開されている成果物

13-3. モデルとパラメータ

モデル

モデルとは、学習によって作られた『判断のルール』そのものである



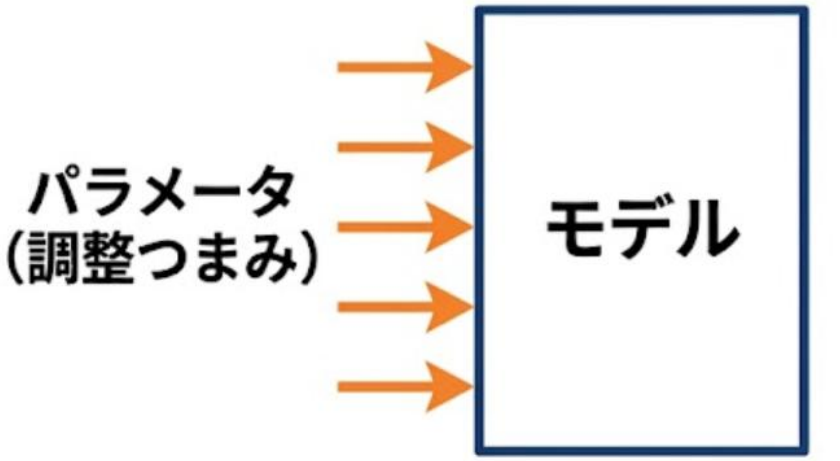
同じ課題でも、種類・サイズの異なる複数のモデルがある

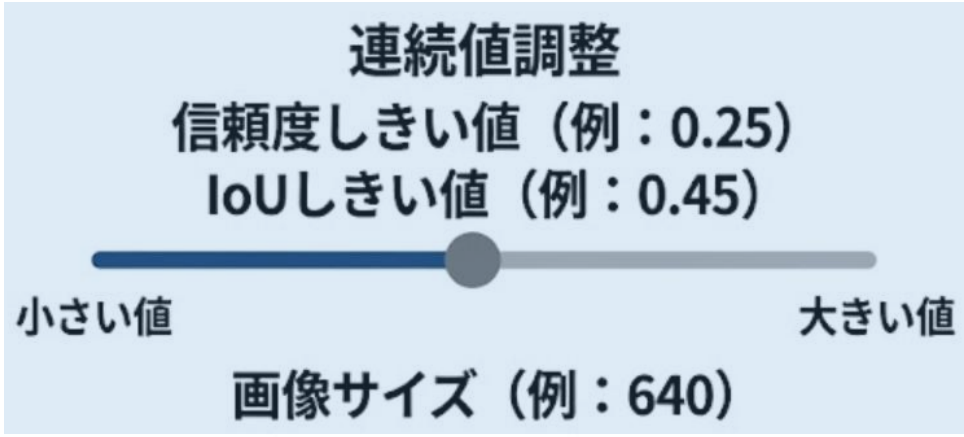
小さいモデル	大きいモデル	別の種類のモデル
軽い・速い	重い・遅い	構造が異なる
得意：シンプルな判断	得意：複雑な判断	得意：特定の課題に強い
不得意：複雑な判断	不得意：動作が重い・コスト大	不得意：それ以外は弱い

万能なモデルは存在せず、目的に合わせて選ぶ

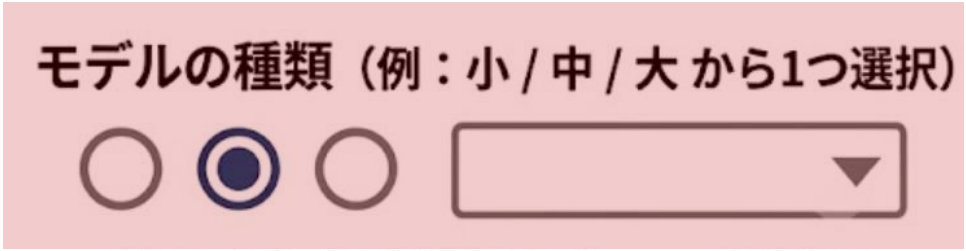
パラメータの調整、モデルの種類選択

パラメータ ＝ モデルの動作を調整する『数値』や『選択肢』の設定値





数値を連続的に細かく調整



決められた選択肢から一つを選ぶ

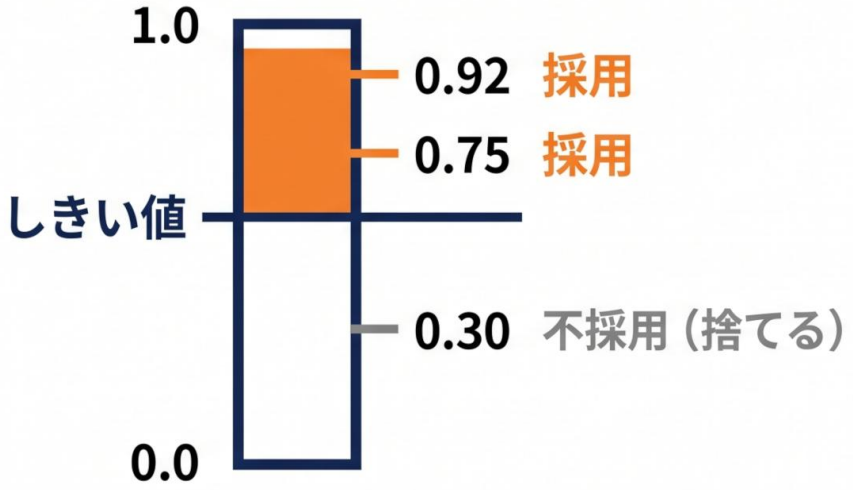
パラメータの調整、モデルの種類選択
のイメージ

信頼度しきい値、IoUしきい値

物体検出の2つのしきい値：信頼度は"採用するか"、IoUは"重なる枠をまとめるか"を決める

信頼度しきい値

AIが付けた"物体である確率"の採用ライン



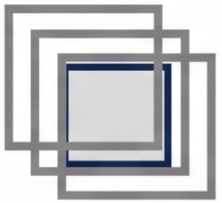
上げる → 確実な枠だけ残る（見逃し増加）
下げる → 枠が増える（誤検出増加）

IoUしきい値

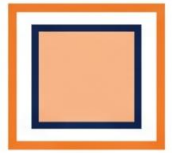
重なった複数の枠を1つにまとめる基準

$$IoU = \frac{\text{重なりの面積}}{\text{全体の面積}}$$

重なり=intersection
全体=union



IoUで統合(NMS)



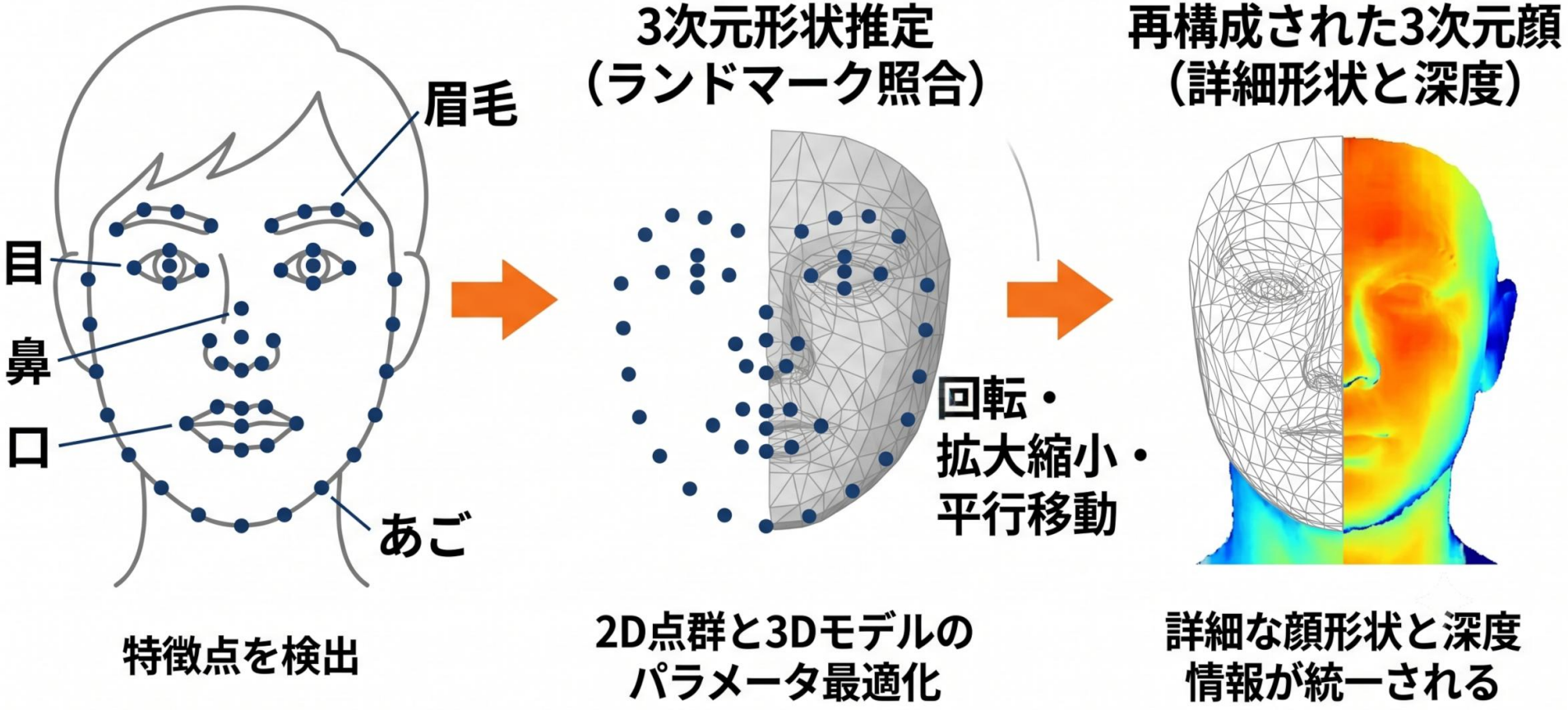
重複した枠
1つに統合

上げる → 統合されにくい(枠が重複しやすい)
下げる → 統合されやすい（枠がまとまる）

流れ：候補枠 →【信頼度で採用/不採用】→【IoUで重なりを統合】→ 最終結果

顔ランドマークと顔の3次元再構成

顔ランドマーク（目・鼻・口・眉毛・あご）を基準点として、2次元画像から顔の3次元形状を再構成する



演習

演習１．Rembg（背景除去セグメンテーション）

https://huggingface.co/spaces/KenjieDec/RemBG

- **概要**：**画像から被写体を切り抜き、背景を透明化**するセグメンテーションのデモ。u2net など**16種のモデルを選択可能**、モデルによって切り抜き精度が変わる。
- **手順**：内蔵のサンプル画像を選ぶ（または自分の画像をアップロード）→ モデルを選択 → Process をクリック → 出力を確認 → モデルを変えて再実行し結果を見比べる。
- **ヒント**：髪の毛や動物の毛など「輪郭が複雑な画像」でモデルを切り替えると差が出やすい。
- **考察ポイント**：モデルを変えると「細部の切り抜き精度」と「処理の速さ」に変化があるか。

演習１．Rembg（背景除去セグメンテーション）

RemBG

Gradio demo for RemBG. To use it, simply upload your image, select a model, click Process, and wait.

Input Image

Output Image

Output Mask

演習２．DETR 物体検出

https://huggingface.co/spaces/ClassCat/DETR-Object-Detection

- **概要**：**物体検知モデル DETR のデモ**。ResNet-50 と ResNet-101 の**2モデルを選べる**（サンプル画像：cats.jpg、detectron2.png、cat.jpg、hotdog.jpg を内蔵）。
- **手順**：モデル（ResNet-50／101）を選ぶ → サンプルのサムネイルをクリック（または画像をアップロード）→ Infer をクリック → CPU で約10秒待つ → 検出枠を確認。
- **ヒント**：**重なった個体を別々に検出できているか**見てみる。
- **考察ポイント**：ResNet-50 と 101 で検出漏れや枠の精度がどう変わるか。

演習２．DETR 物体検出

DETR Object Detection

1. Select a model.

Model name

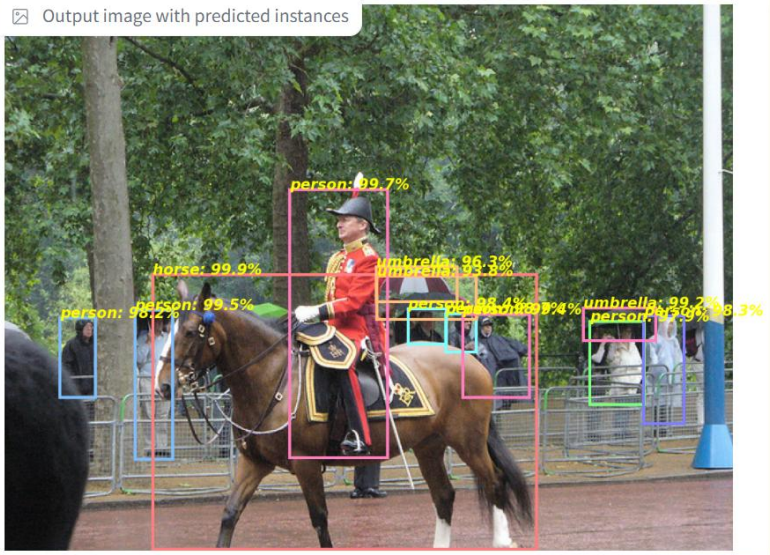
detr-resnet-50　detr-resnet-101

2-a. Select an example by clicking a thumbnail below.

2-b. Or upload an image by clicking on the canvas.

Input image

Output image with predicted instances



Examples

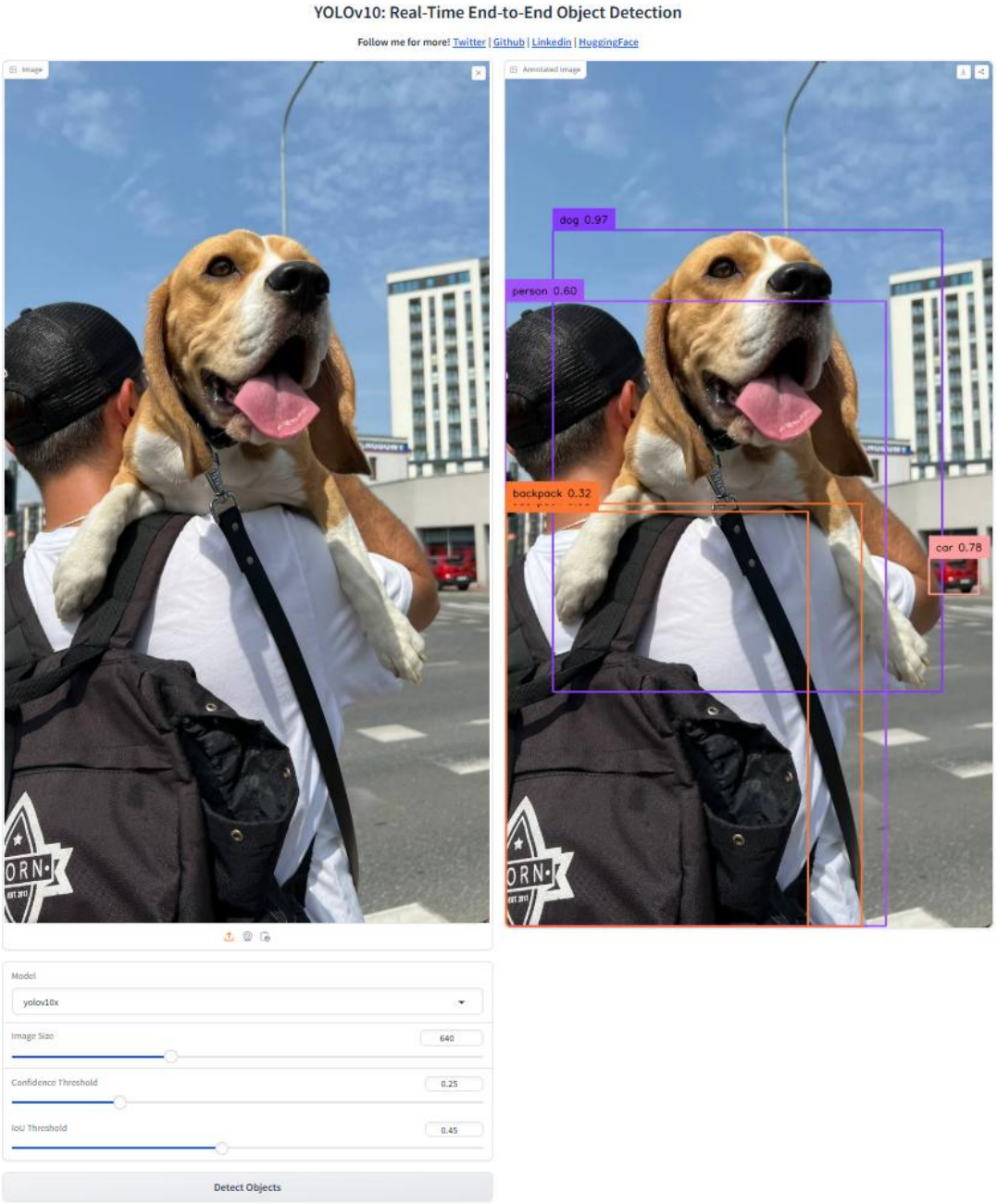



演習３．YOLOv10 物体検出

https://huggingface.co/spaces/kadirnar/Yolov10

- **概要：物体検知モデル YOLOv10 のデモ。モデルの種類（n／s／m／b／l／x）**、画像サイズ、信頼度しきい値、IoUしきい値の4つを調整できる（サンプル画像3種を内蔵、初期値は image size 640・confidence 0.25・IoU 0.45）。
- **手順**：サンプル画像を選ぶ → モデルを選択 → 画像サイズ・信頼度・IoUしきい値を設定 → 実行 → 各パラメータを変えて検出結果の変化を観察。
- **ヒント**：同じ画像を、異なるモデルで処理し、精度の差があるか確認。
- **考察ポイント**：画像サイズを大きくすると小さな物体の検出がどう改善／悪化するか。

演習３．YOLOv10 物体検出

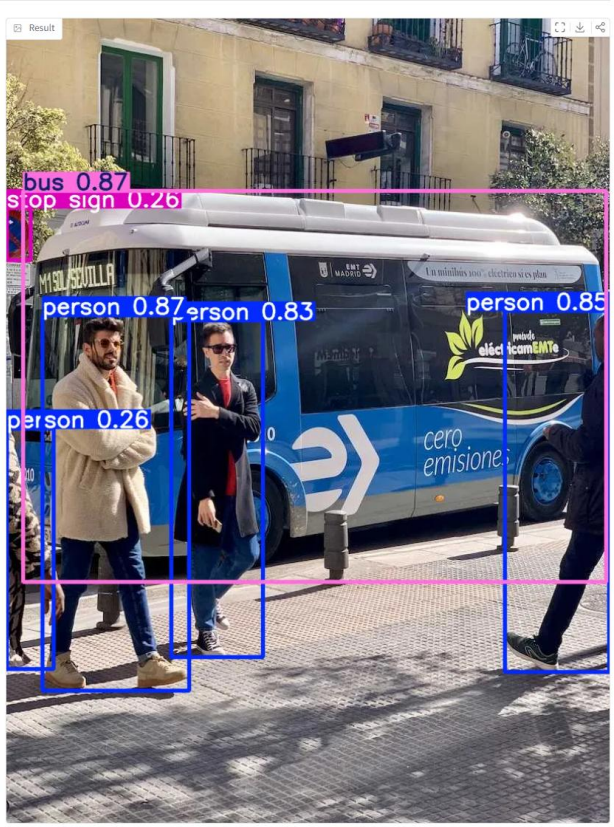


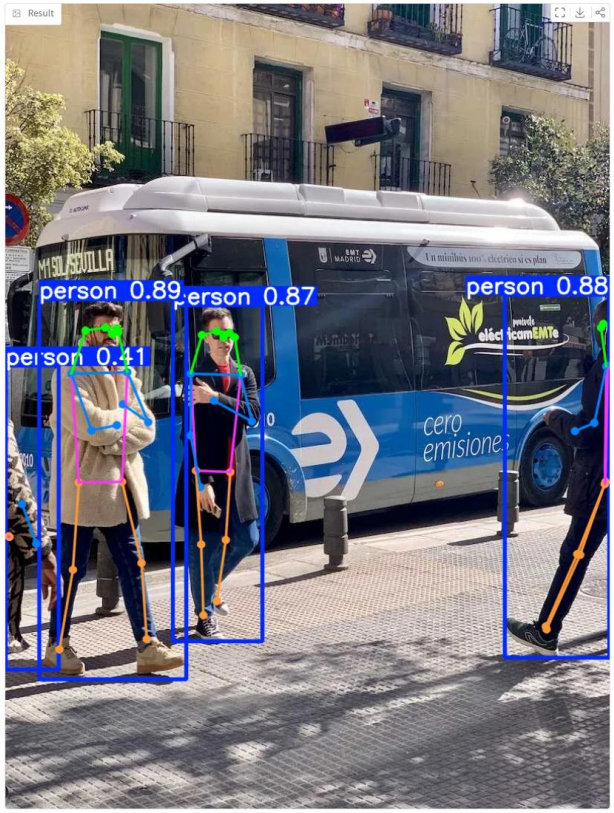
演習４．YOLOv8 の種々の機能

https://huggingface.co/spaces/Ultralytics/YOLOv8

- **概要：物体検知・セグメンテーション・姿勢推定・回転矩形・分類まで扱える多機能デモ。モデル種別（yolov8n／yolov8n-seg など15種）**に加え、信頼度しきい値・IoUしきい値・画像サイズを調整でき、画像・動画・Webカメラに対応（bus.jpg 等のサンプル内蔵）。
- **手順**：サンプル画像を選ぶ → モデルを選択 → 信頼度・IoUしきい値を設定 → 実行 → しきい値を上下させて検出数の変化を観察 → 検知モデルとsegモデルを切り替えて出力の違いを見る。
- **ヒント**：yolov8n を yolov8n-seg に替えると、**物体検出から、セグメンテーションに変わる**。
- **考察ポイント**（1行）：信頼度しきい値を下げると「見逃し」が減る代わりに「誤検出」が増える。

演習４．YOLOv8 の種々の機能



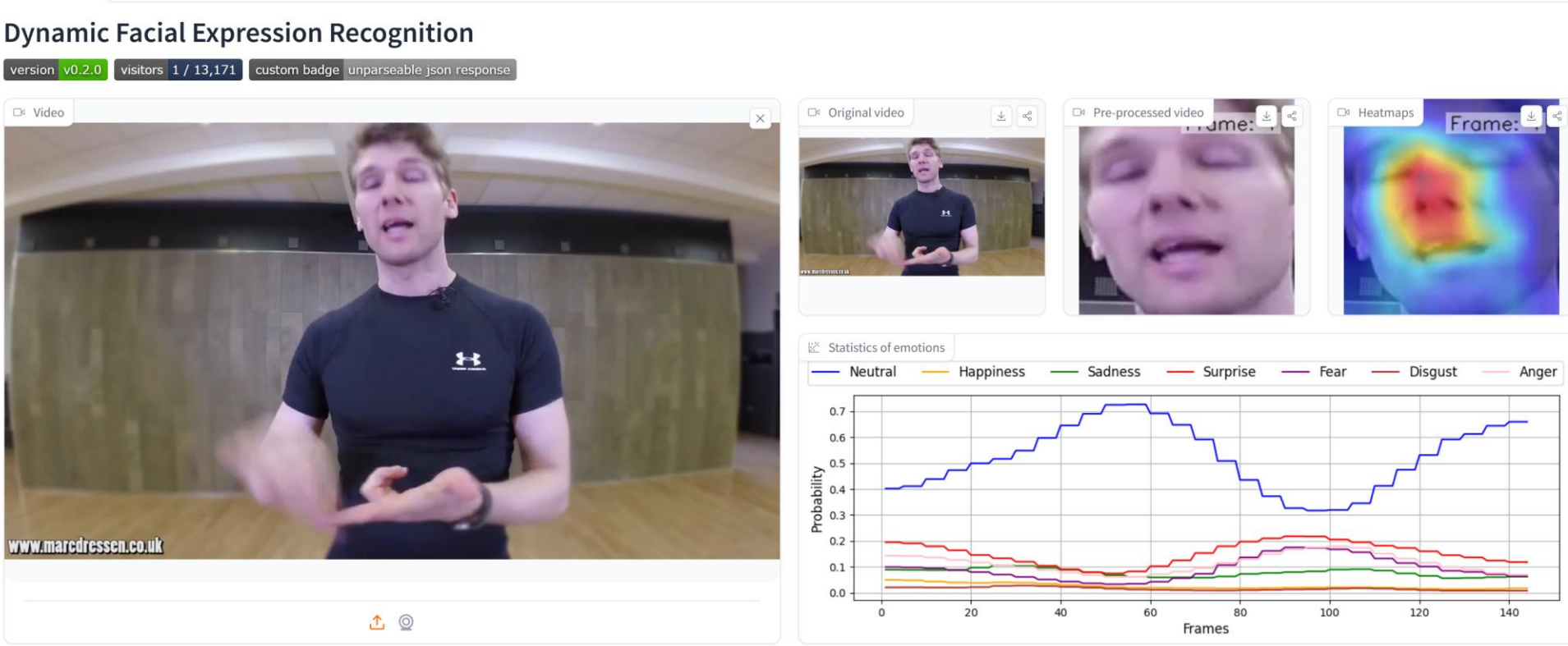


演習５．表情認識

https://huggingface.co/spaces/ElenaRyumina/Facial_Expression_Recognition

- **概要**：顔情報処理（表情認識）のデモ。**顔検出→表情分類に加え、判断根拠を示すヒートマップ**と、動画では**感情の時系列統計グラフ**まで出力する（静止画7種・動画2種のサンプル内蔵）。
- **手順**：静止画タブでサンプル画像を選び表情の判定結果とヒートマップを確認 → 動画タブに切り替えサンプル動画を実行 → 時系列で感情がどう変化するかグラフで観察。
- **ヒント**：ヒートマップで「顔のどの部分（目・口など）」が判断根拠になっているか確認できる。
- **考察ポイント**：横顔・部分的に隠れた顔など「難しい」顔写真を準備、試してみる

演習５．表情認識


Dynamic Facial Expression Recognition
version v0.2.0
visitors 1 / 13,171
custom badge unparseable json response
Video
www.marcdressen.co.uk
Original video
Pre-processed video
Heatmaps
Frame:
Statistics of emotions
Neutral
Happiness
Sadness
Surprise
Fear
Disgust
Anger
Probability
Frames