

# Arithmer Dynamics

## AI Systems



## Dynamics技術紹介セミナー

2021/01/21  
Dynamics Team

## 【動画】

連続的に見ると残像効果で動いているように見える、少しずつ変化させた一つながりの画像。映画・アニメーションやビデオカメラで撮影した映像など。(デジタル大辞泉)

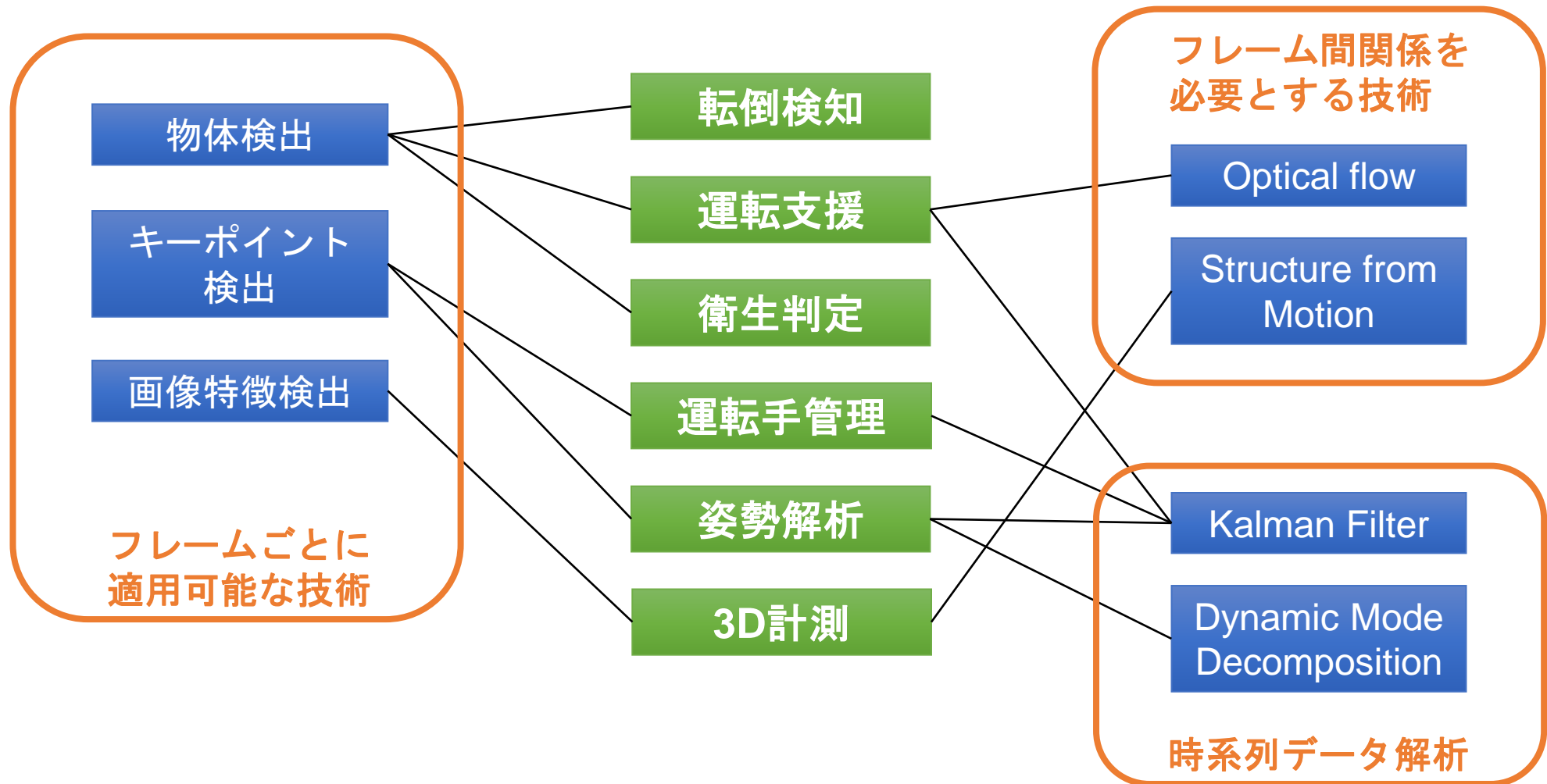
## 動画＝画像(フレーム)の集合

画像解析で使える技術は、動画解析でも(各フレームに対して)使える。  
e.g., 物体検出, 画像分類, 画像異常検知...

## 動画＝画像の時系列データ

各フレームを1つのデータと見れば、動画は時系列データであるとも見なせる。  
e.g., 時系列解析

動画解析は性質上  
既に複合的な側面を持っている



## 乗用車・大型車両における周囲の危険を検知

- 車のフロント・サイド画像から人・自転車などを検知し、ある範囲に入ったときに危険と判定.
- アラートによる通知の他、危険があった箇所を集約してDynamical Hazard Mapを作成可能.
- リアルタイムでの解析・撮りためた動画に対する事後解析の両方が可能.
- AndroidやJetson Nanoなど、エッジデバイスでも使用可能.

## 強み

- (Smartphone as an edge) リアルタイム解析ではスマートフォン (Android) をエッジデバイスとして利用可能であり、導入コストが低い.
- 大型車両のサイドも外部カメラ接続により対応可能である.



フロント



サイド

## 乗用車・大型車両における周囲の危険を検知

## 使用技術

物体検出: 物体の位置と種類を検出  
→人・自転車・標識などを検出可能

Optical Flow: 画像内の特徴点の移動量を検出.  
→車が走っているかどうかを動画のみから検出可能

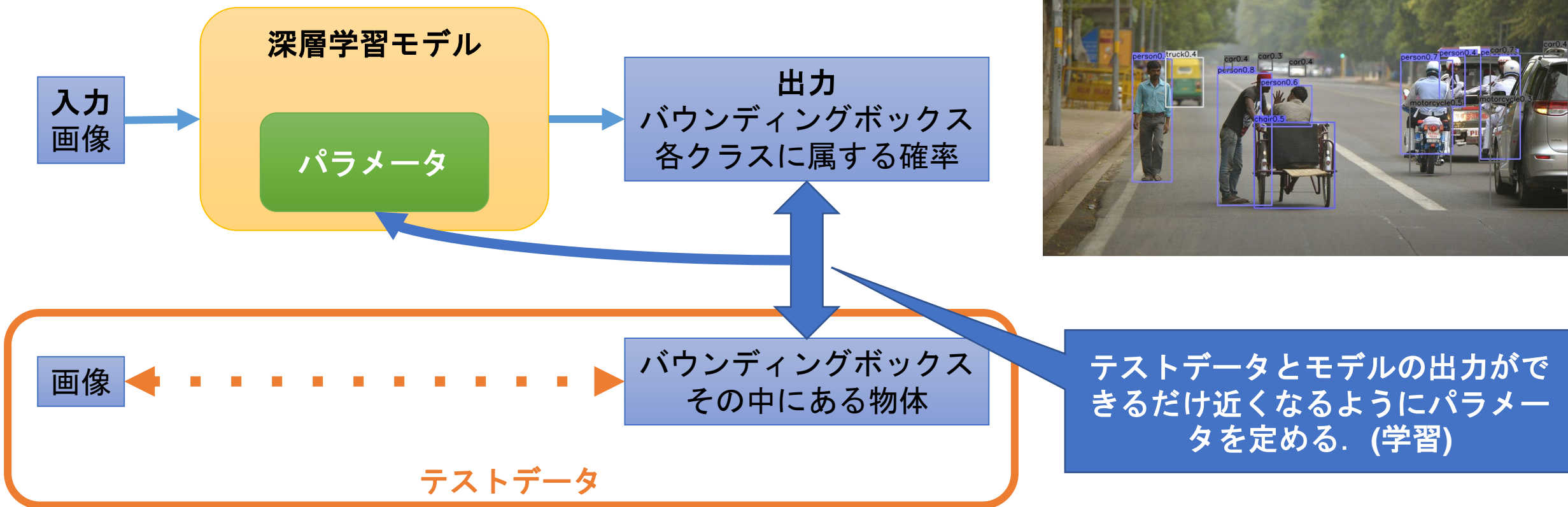
※精度は高くないため、GPSや加速度センサなどが使える場合はそちらを使用している。

## 導入事例

- ・ 現在テストユーザとなってくださる運輸会社がいくつかある



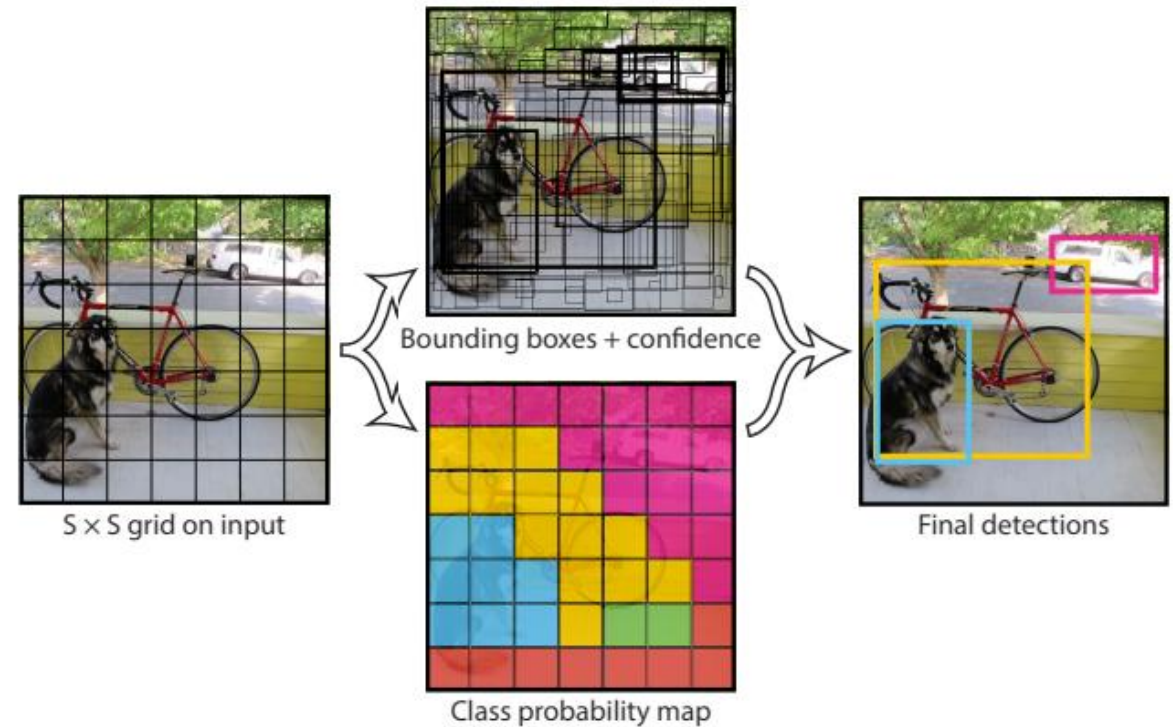
赤信号無視・一時停止無視



汎化: テストデータで学習をさせると, (テストデータと同じ確率分布から得られる)テストデータにないデータでも精度の良い出力が出せる.

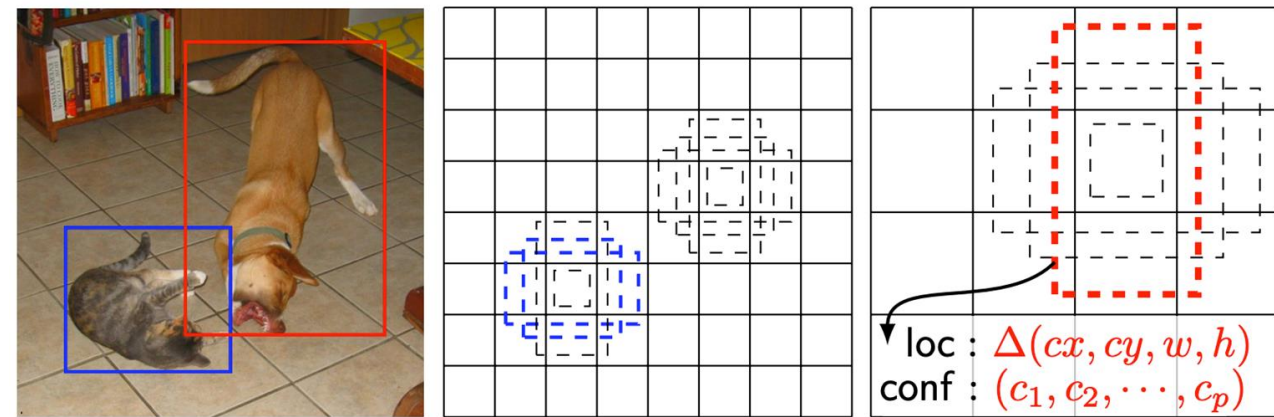
物体検出モデルとしてよく使っているもの: **SSD, yolo, CenterNet**

- 画像を粗いグリッドに切る.
- 各グリッドに対して, そのグリッドを中心とするようなバウンディングボックスの候補を作っておく.
- 各バウンディングボックスについて, 次を出力するような深層学習モデルを作る.
  - そのバウンディングボックスに物体がある確率
  - 物体がある場合, それがどのクラスに属するかの(条件付き)確率
- 使うときは,  
バウンディングボックスに物体がある確率×各クラスに属する(条件付き)確率が指定した値より大きいバウンディングボックスと分類クラスを選ぶ.



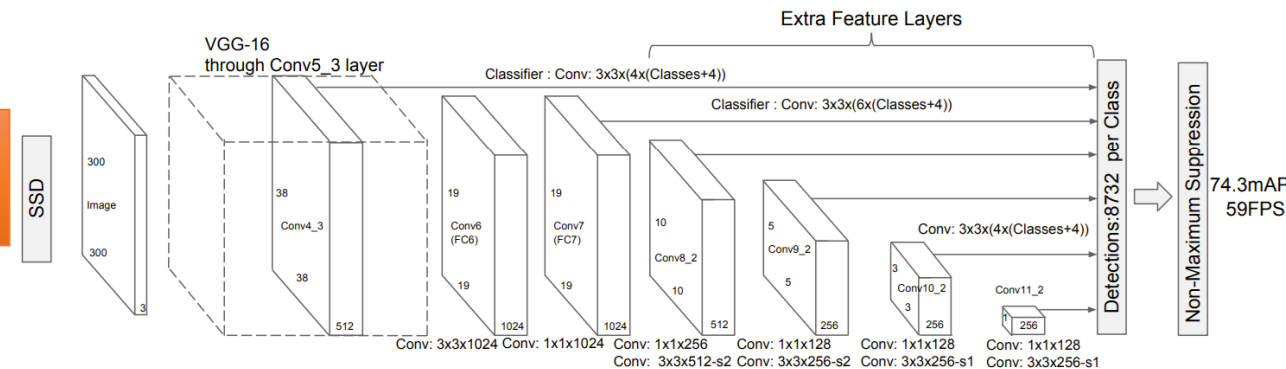
Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).

- 基本的な考え方はyoloとほぼ同じ
- ただし, yoloでは予めグリッドの切り方を固定しているのに対して, **SSDでは粗いスケールから細かいスケールまでグリッドの切り方を複数用意する.**
- 各スケールごとに, 用意するバウンディングボックスの大きさも変わる.
  - こうすることで, 大きな物体も小さな物体も検知できるようにする.



(a) Image with GT boxes (b)  $8 \times 8$  feature map (c)  $4 \times 4$  feature map

軽量化がやりやすい反面, 出力となるバウンディングボックスは(動画として見ると)ぶれやすい.



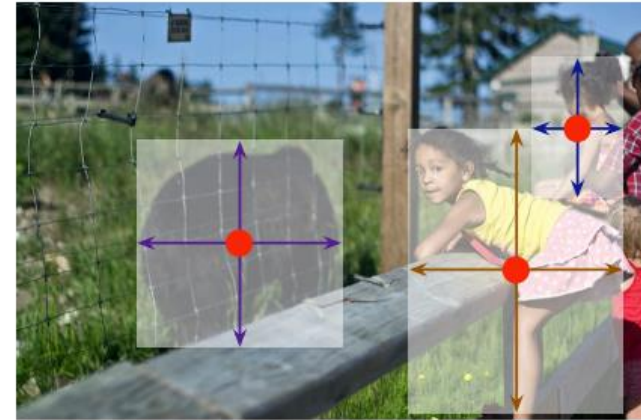
各スケールの特徴量に合わせて  
各バウンディングボックスの確度を出力

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016, October). SSD: Single shot multibox detector. In *European conference on computer vision* (pp. 21-37). Springer, Cham.



- バウンディングボックスの候補を前もって指定しない。
- バウンディングボックスを, 「中心+幅+高さ」で表現する。
- 各クラスごとに, その画像の中でバウンディングボックスの中心がある確率を(粗いグリッドの)ヒートマップで出力する。
- それとは別に, バウンディングボックスの幅と高さも数値として出力。
- 他にもコスト関数の作り方などで色々工夫している。

学習・検知速度も早く, 精度も良い。ブレも少なく, 実用性が高い。



Zhou, X., Wang, D., & Krähenbühl, P. (2019). Objects as points. *arXiv preprint arXiv:1904.07850*.

- 2つの連続する画像の間で、点がどのように動いたかを推定する.
- 2つの画像内の特徴点を比較して結びつける.
- 以下の仮定のもとで移動量を計算する.
  - 連続するフレーム間で明るさは変わらない
  - 隣接するピクセルは似たように動く.

ただし、絶対的な速度は得られず、周囲の環境にも依存するので、GPSなどが使える場合はそちらを優先したほうがいい。



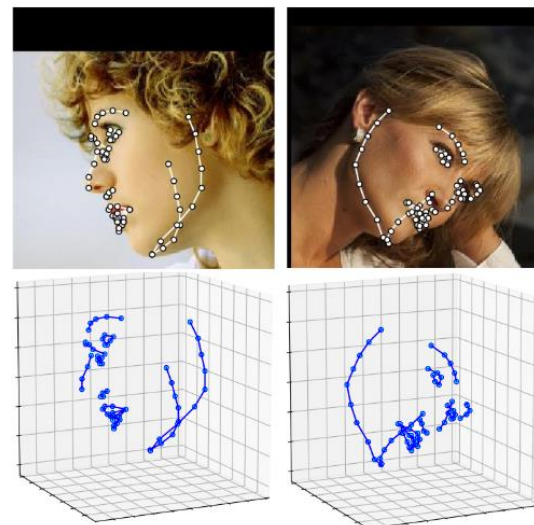
運転手を撮影し，居眠り・ながら運転・よそ見などを判断

## 使用技術

物体検出: スマートフォンなどの物体を検出する。  
顔特徴点検出: 顔の特徴点を検出し，目が閉じているか，よそ見をしているかを判断

## 強み

- いくつかの運転手に関わる危険要素を組み合わせることができる
- 動画のみから判定することができる。



人間に、愛を。  
未来に、AIを。

Arithmer 株式会社

〒106-6040

東京都港区六本木一丁目6番1号 泉ガーデンタワー 38/40F(受付)

03-5579-6683

<https://arithmer.co.jp/>

Arithmer

