

ロボット・ネットワーク領域

# 魚眼カメラを用いた エレベータ内の乗員の頭部検出

---

工学部 機械機能工学科  
中村 真吾

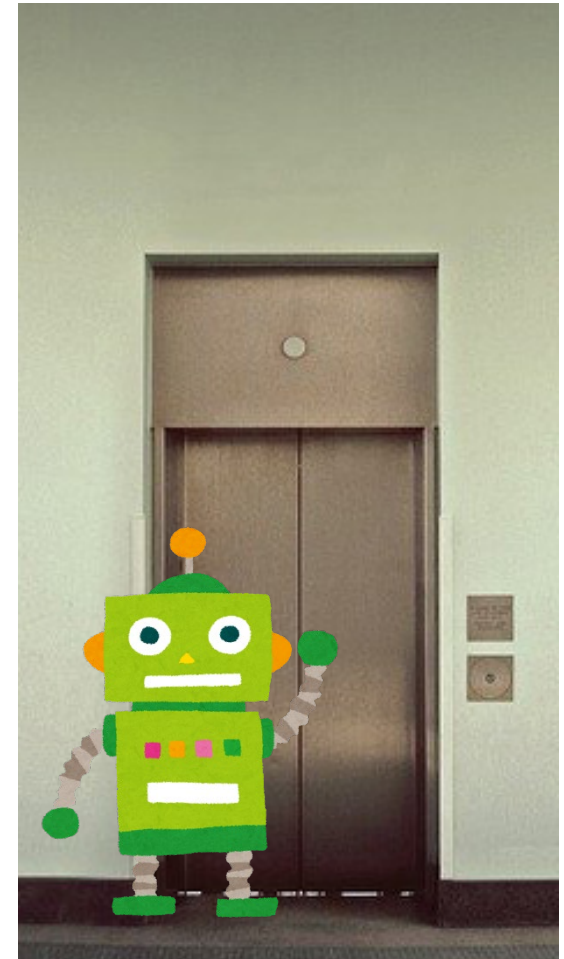
アーバン・エコ・モビリティ研究拠点の形成  
～私立大学研究ブランディング事業シンポジウム～  
2021年3月2日

# はじめに

- 街中・施設内での人とロボット共存
  - シームレスなロボットの移動技術
  - エレベータを利用した移動技術
- エレベータ搭乗時の状況把握
  - ロボットの搭乗できる空間
  - 他乗員の状況(人数、位置)



エレベータ内のセンシング技術



# エレベータ内のセンシング

## ■ 取得すべき内容

### 他乗員の情報

≫ 人数

≫ 位置

≫ 向き

本発表の内容

≫ 体積

### その他(荷物など)

## ■ センサデバイス

### 魚眼カメラ

≫ EV内を一度で撮影



# 魚眼カメラ(全方位カメラ)

## ■ エレベータ内全領域の撮影



RICOH THETA V



# 既存の人検出器

## ■ Deep Neural Network (DNN)による人検出

### □ YOLO<sup>[1]</sup>



[1] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779-788, 2016.

### □ SSD<sup>[2]</sup>



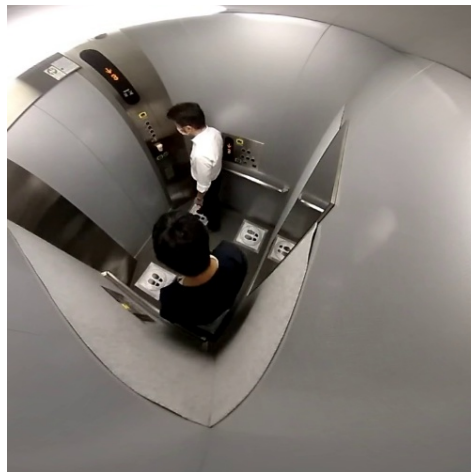
[2] Liu W. et al. (2016) SSD: Single Shot MultiBox Detector. In: Leibe B., Matas J., Sebe N., Welling M. (eds) *Computer Vision – ECCV 2016*. ECCV 2016. Lecture Notes in Computer Science, vol 9905. Springer, Cham.

複数人の検出、真上からの検出が困難

# 新たな人検出器

## ■ 頭部を検出

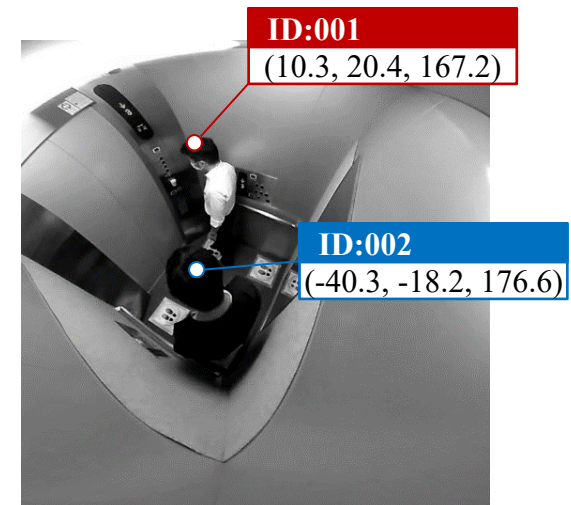
- 重なりが生じてても、カメラで捉えることが可能
- 身長計測技術と組み合わせることで、3次元位置が測定可能



画像



新たな人検出器  
(頭部検出器)



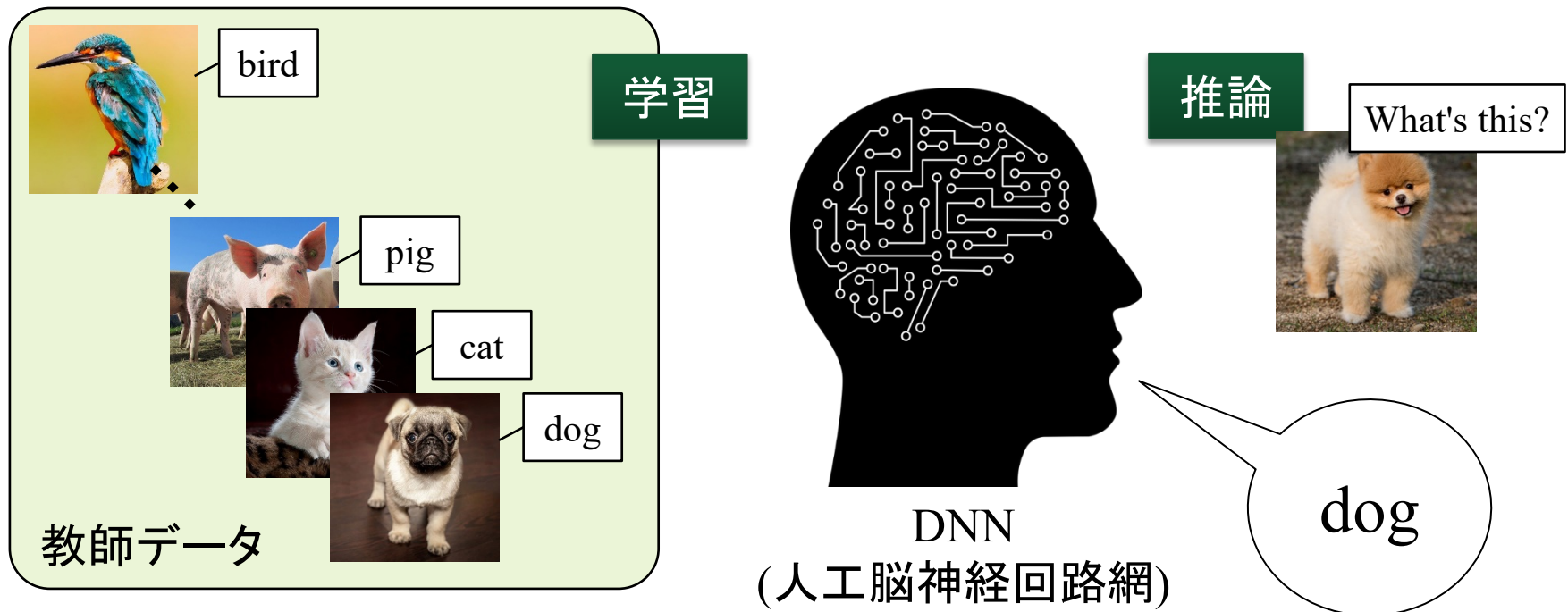
頭部検出

頭部検出器に Deep Neural Network (DNN) を使う

# Deep Neural Network (DNN) とは

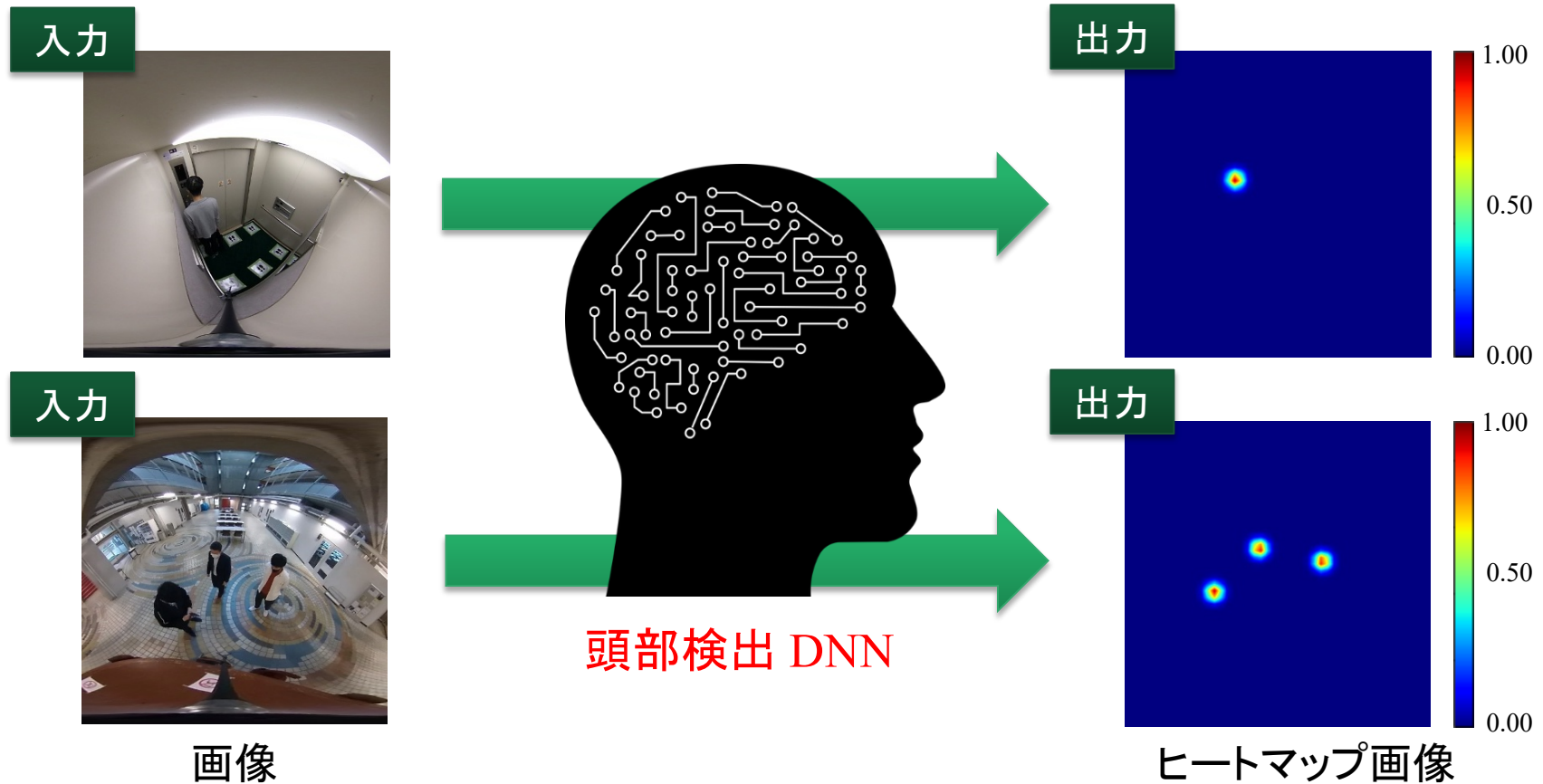
## ■ 機械学習の一手法

- 教師データを**学習**し、未知の入力に対する**推論**を行う
- 脳神経回路を模倣した数理モデルから構成される
- ネットワーク構造が深いものを、特に "**Deep**" NN と呼ぶ



# 頭部検出をする DNN

- 入力: 画像、出力: 頭部位置のヒートマップ画像
  - 複数人の検出が可能

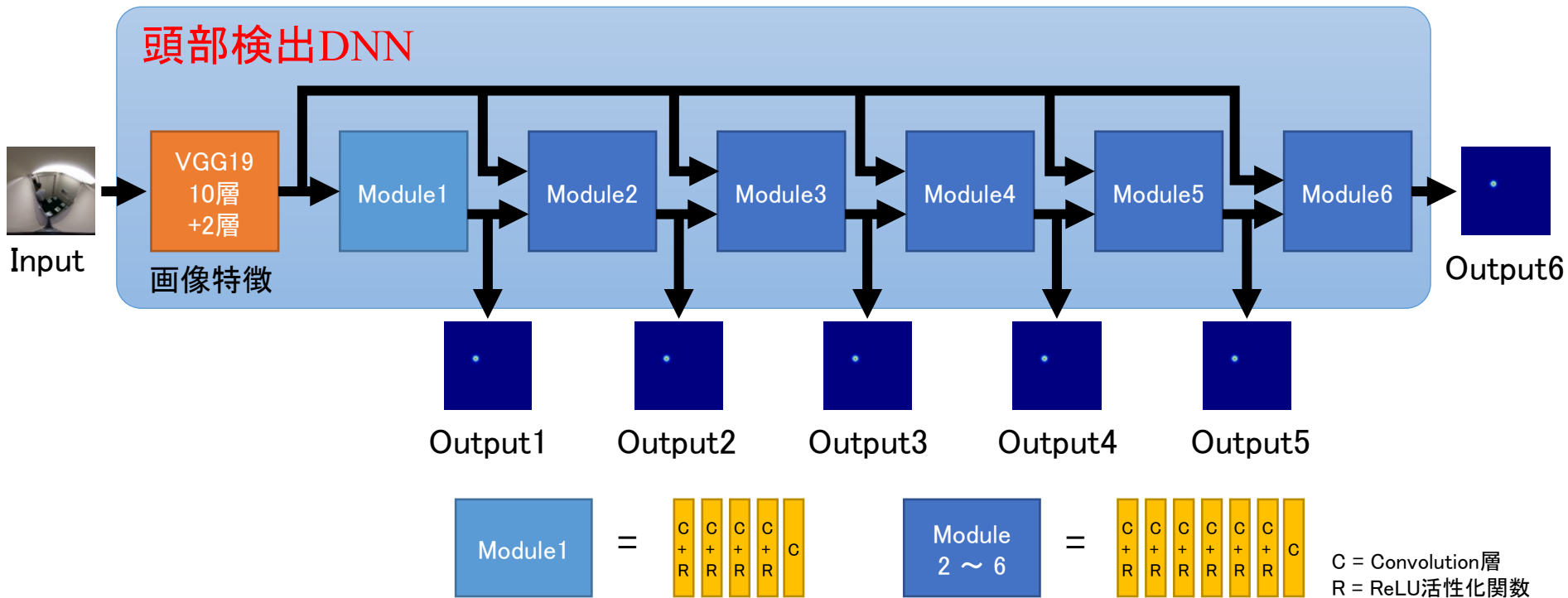




# 頭部検出をする DNN の構造

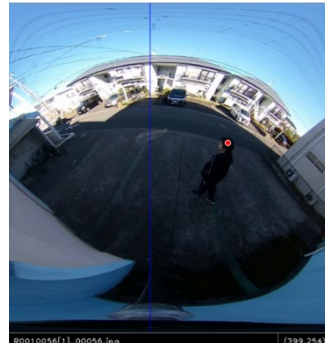
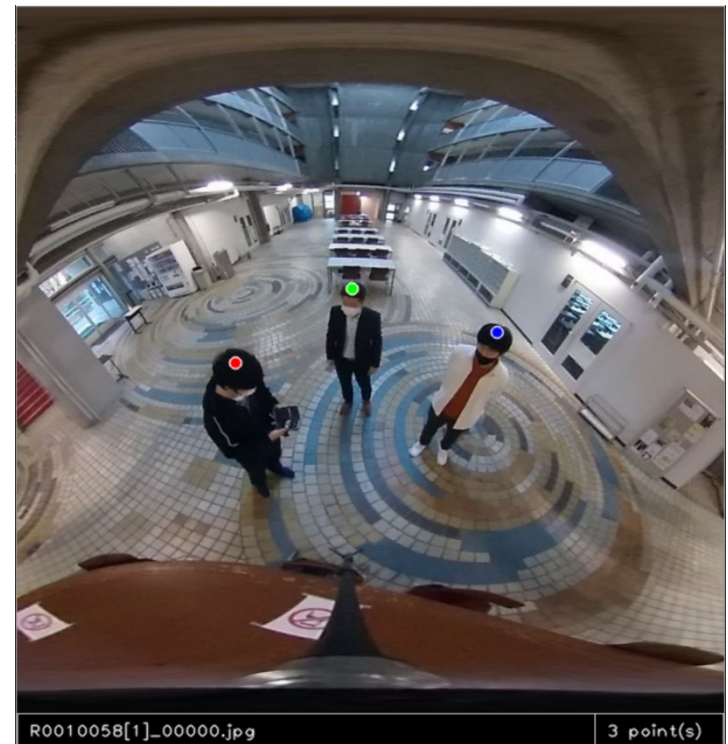
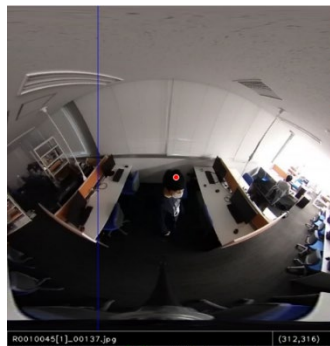
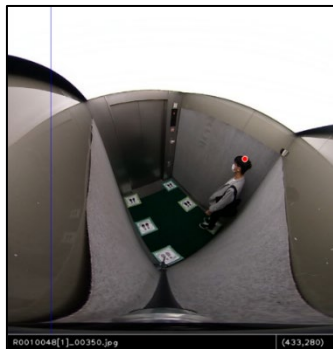
## ■ VGG19 + ヒートマップ出力モジュール × 6

- Convolutional NN をベースとする
- 特徴量抽出にVGG19を使用
- 6モジュールに、画像特徴を入力することで学習を促進



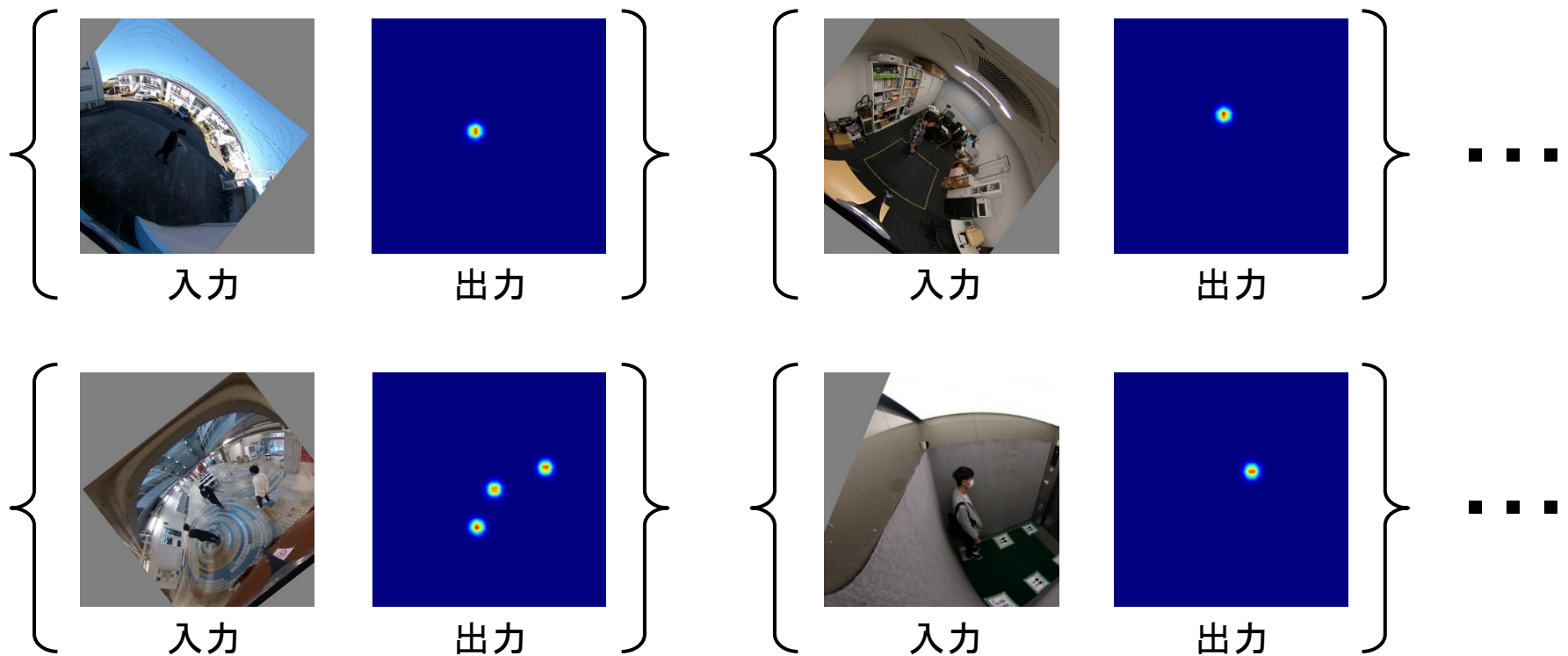
# 教師データの作成（動画、アノテーション）

- 20本の上方アングルの動画を撮影
- 合計14,830枚の画像にアノテーション
- 様々な状況下（屋外、屋内、一人、複数人）



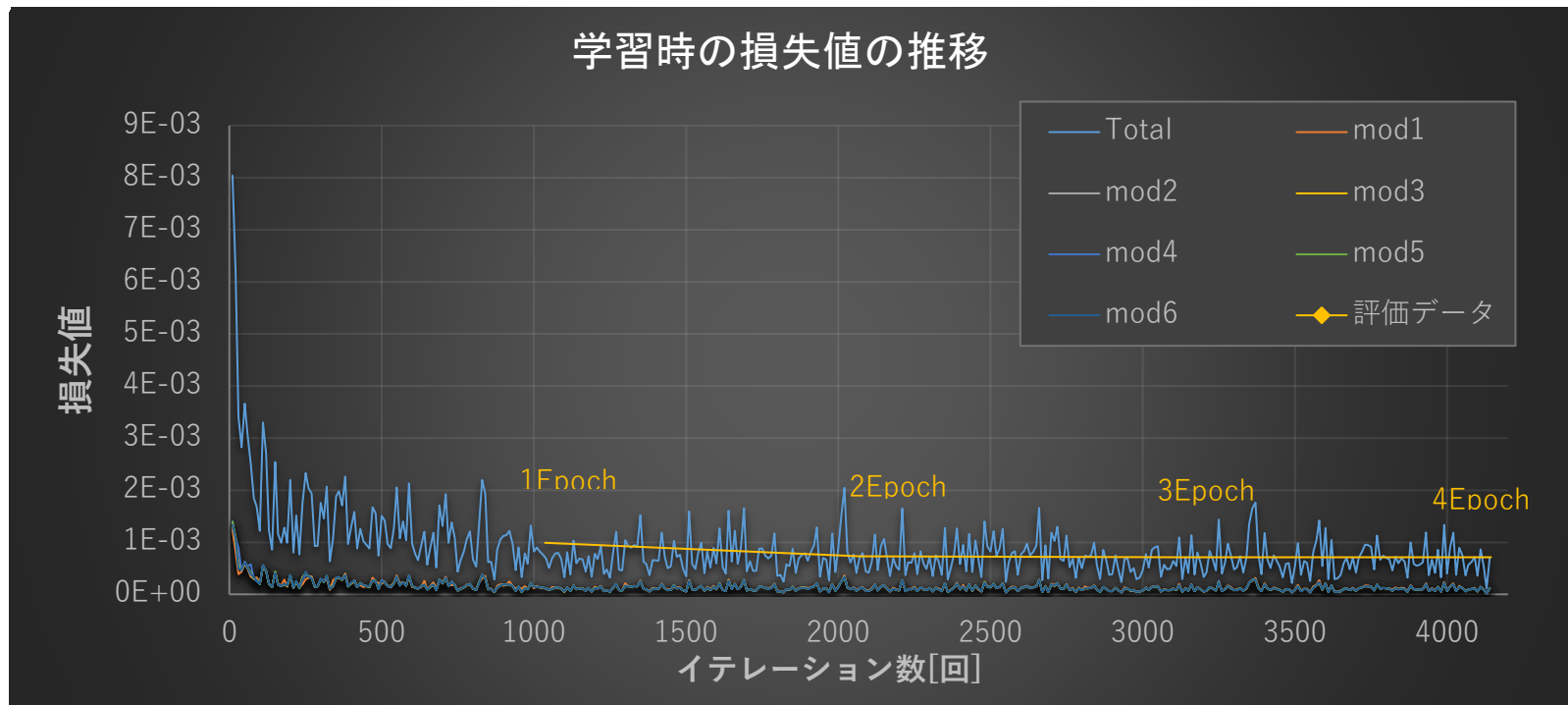
# 教師データの作成 (Augmentation、出力画像)

- オグメンテーションを施し、データを水増し
  - ランダムに、(a) 拡大・縮小 (b) 回転 (c) 左右反転
- 対応する出力画像(ヒートマップ)を作成



# 学習実験と結果

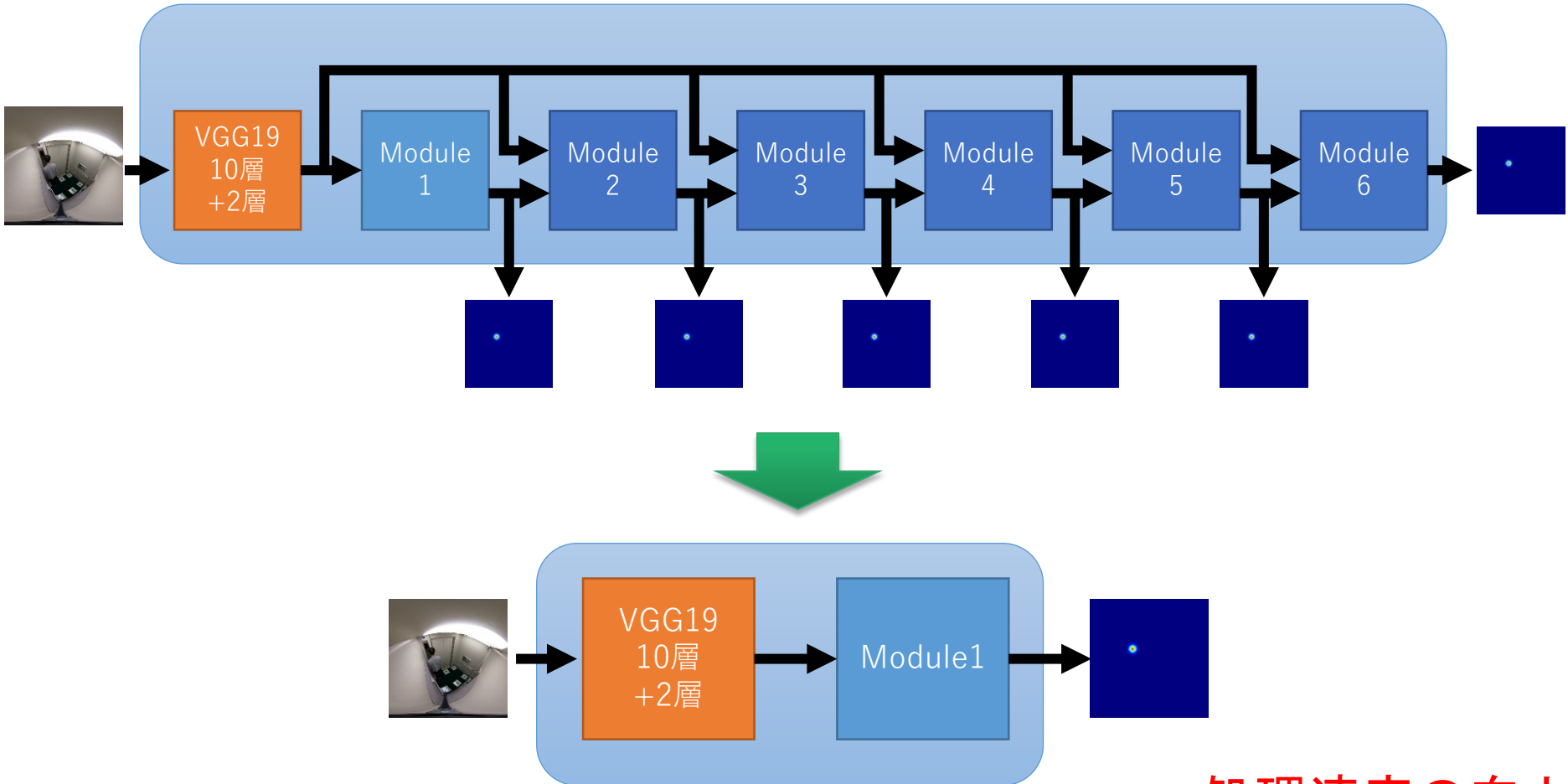
- 学習データ:(8282枚)
- 検証データ:(3602枚)
- バッチサイズ:8
- イテレーション数(4144回)
- 損失関数:各出力と教師データのMSE
- 最適化アルゴリズム:Adam
- 学習係数: $1.0 \times 10^{-5}$  (1エポック毎に1/10)



6つのモジュールの出力結果が同じ ➡ 1モジュールで十分

# DNN構造の簡素化

## ■ 6モジュールから1モジュールへ

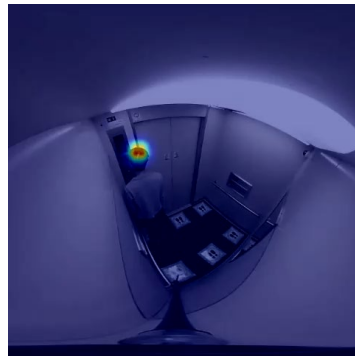


処理速度の向上

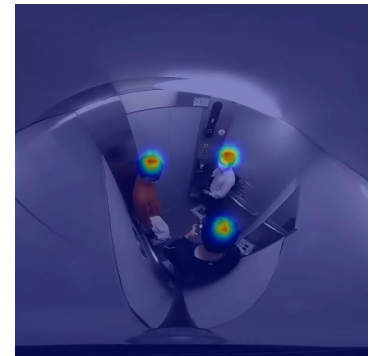
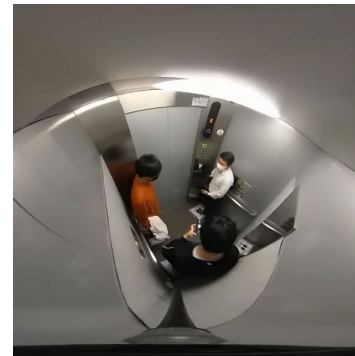
# 動画への適用

## ■ 学習した1モジュールDNNに対して動画を適用

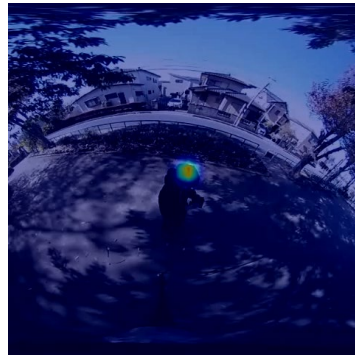
□ 入力動画(左) 入力と出力を重ねた動画(右)



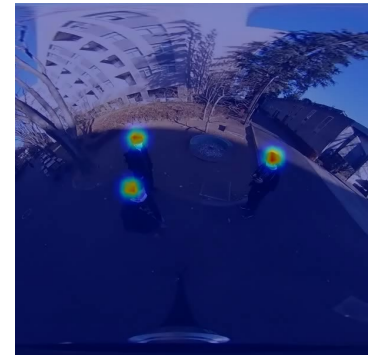
エレベータ(一人)



エレベータ(複数人)



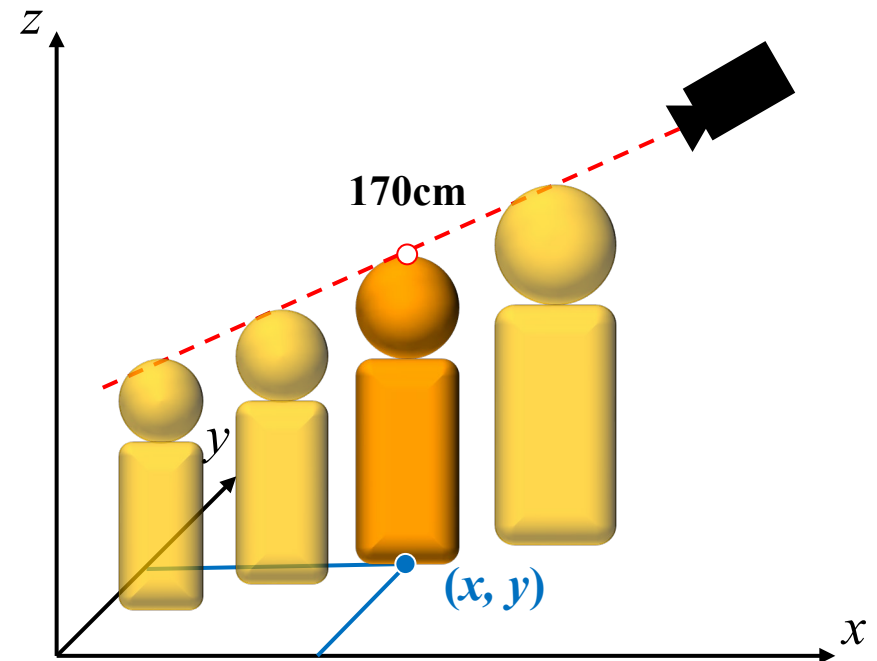
屋外(一人)



屋外(複数人)

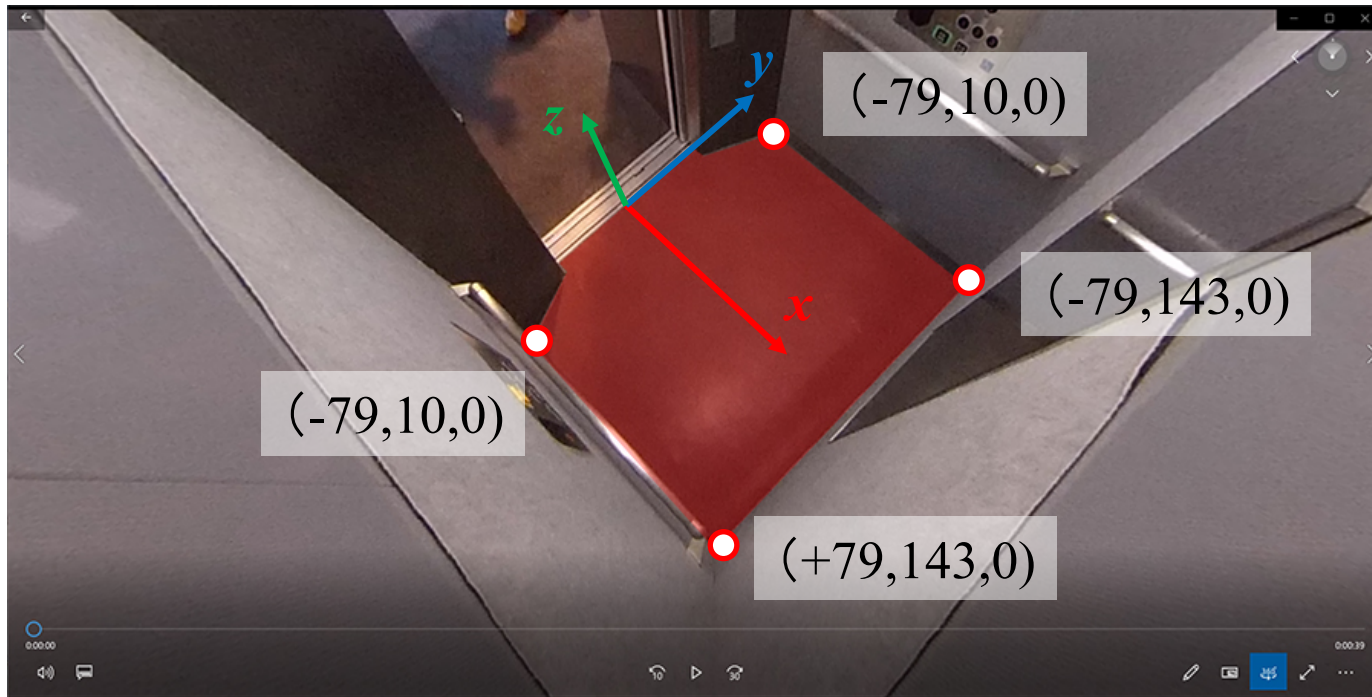
# 乗員位置の推定

- 乗員の身長計測ができていると仮定
- 身長と頭部位置から、水平面上の位置を推定



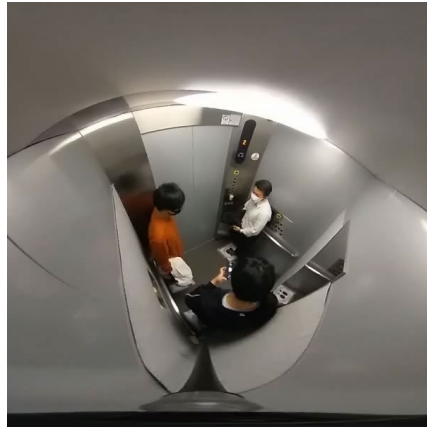
# カメラキャリブレーション

- カメラの位置の推定
- エレベータ内床上的の4点の世界座標値を与える
  - 簡単な線形方程式

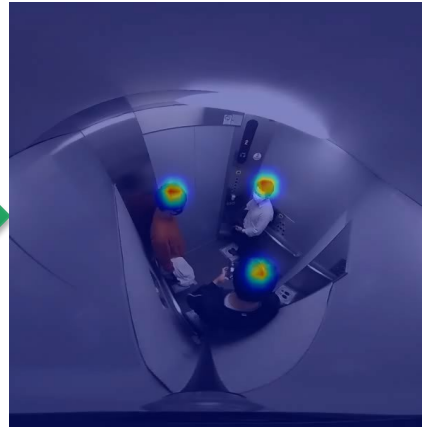
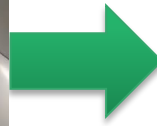




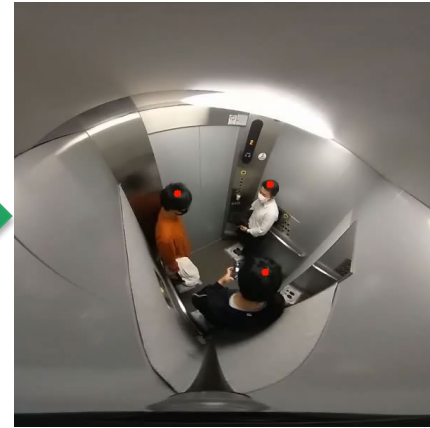
# 総合処理



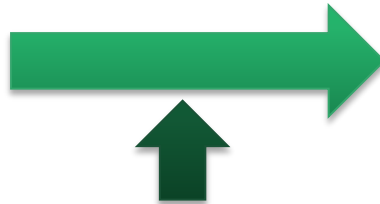
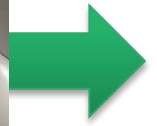
入力



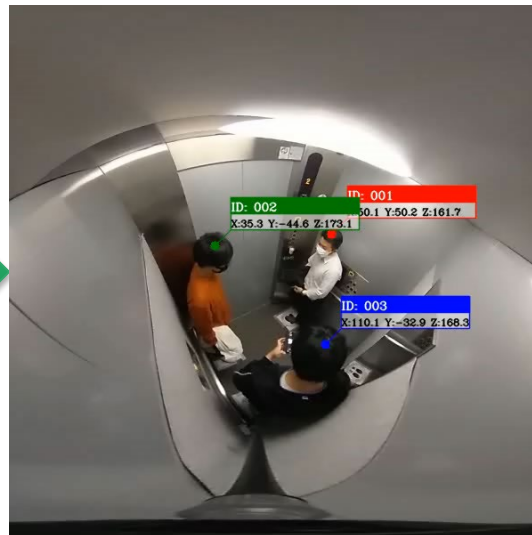
ヒートマップ



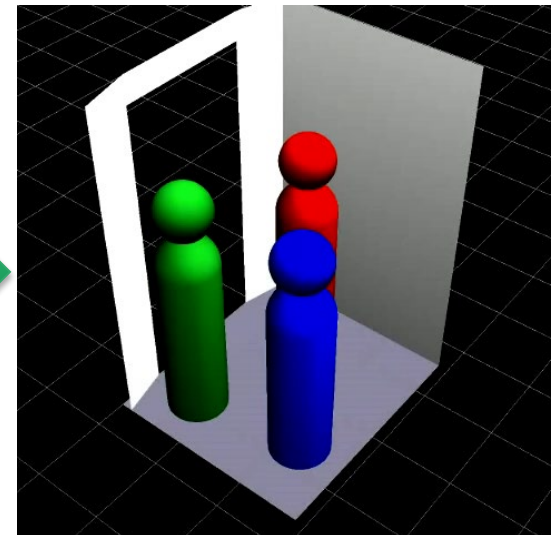
頭部位置推定



身長情報  
(別システム)



3次元位置推定



3次元CGによる再現

# まとめ

- 魚眼カメラを使い、エレベータ内の乗員頭部の検出
- 身長計測技術と合わせ、乗員位置の推定
  
- 今後
  - 位置以外の情報(向き、ボリューム)
  - ロボットの搭乗可能性の評価
  - 黒髪以外のデータ適用
  
- 応用
  - セキュリティ技術
  - マーケティング調査