

# 機械学習を用いた変形AR マーカの位置姿勢推定

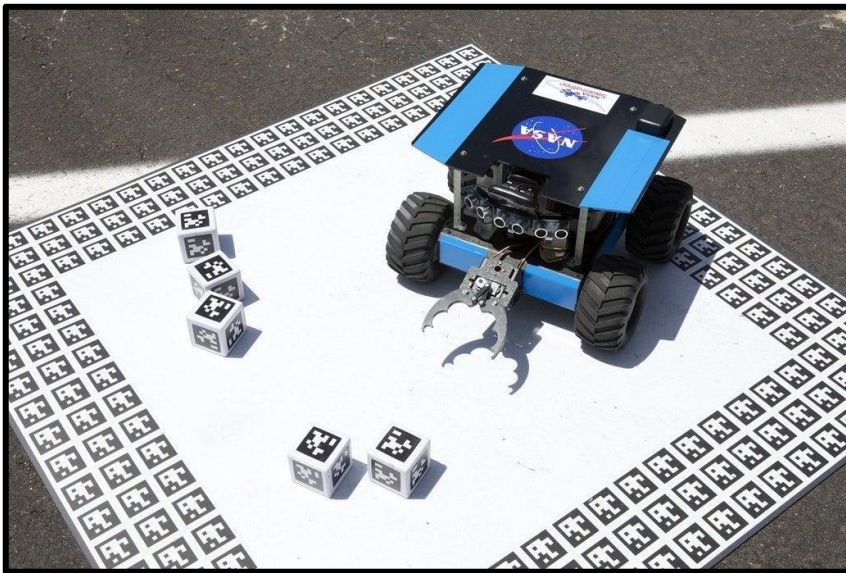
2021/2/12

機械知能研究室

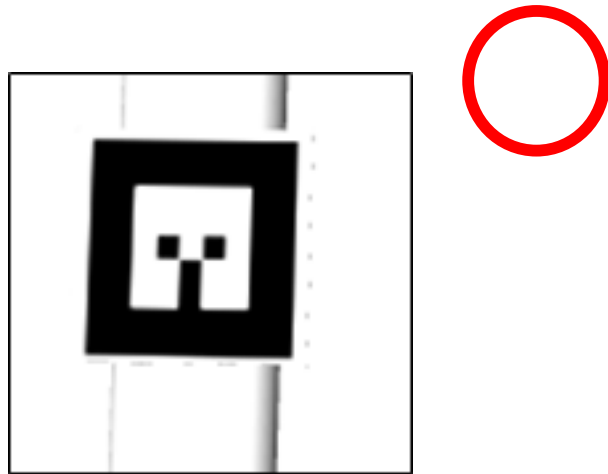
ER17013榎元洋平

ER17076安井理

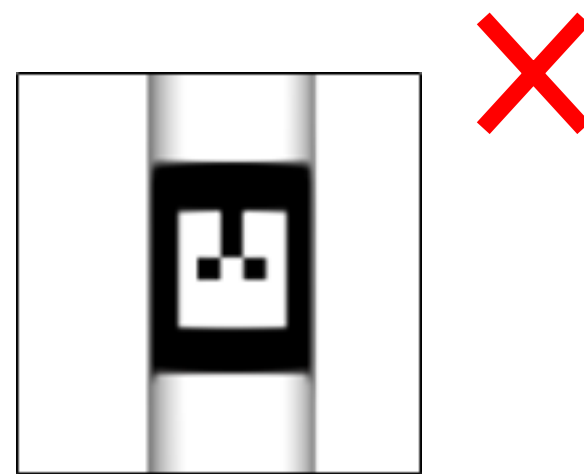
- 2次元コードが普及し様々な場所で活用されている
  - 大きさを定義することで3次元位置・姿勢の推定を行える
  - 特殊なパターンによりどの角度からでも検出可能



- 2次元コードに歪みが生じた場合検出精度が著しく低下
- 機械学習により変形したマーカを検出する方法は提案されているが姿勢推定までは至っていない

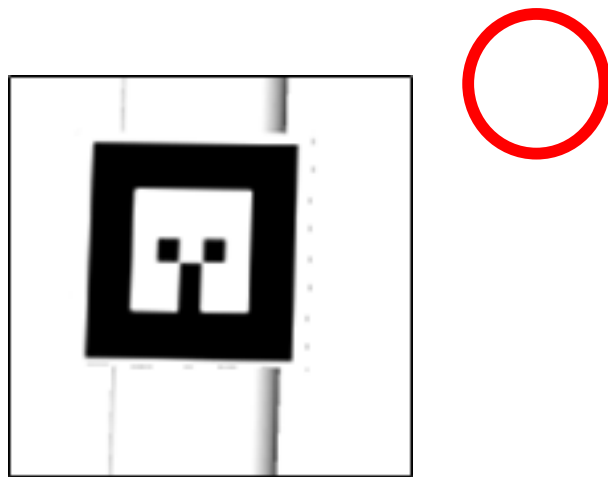


認識可能なマーカ

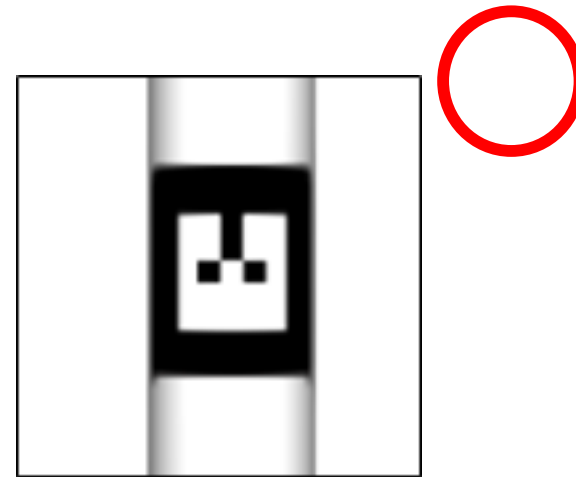


歪みにより認識不可

- 2次元コードに歪みが生じた場合検出精度が著しく落ちる
- 機械学習により変形したマーカを検出する方法は提案されているが姿勢推定までは至っていない

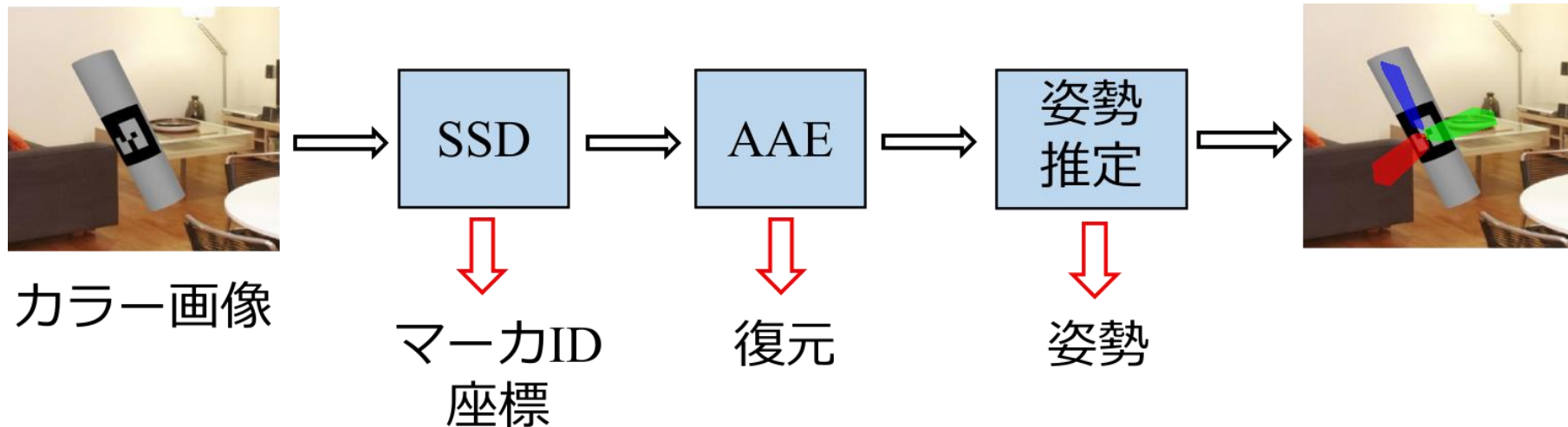


認識可能なマーカ

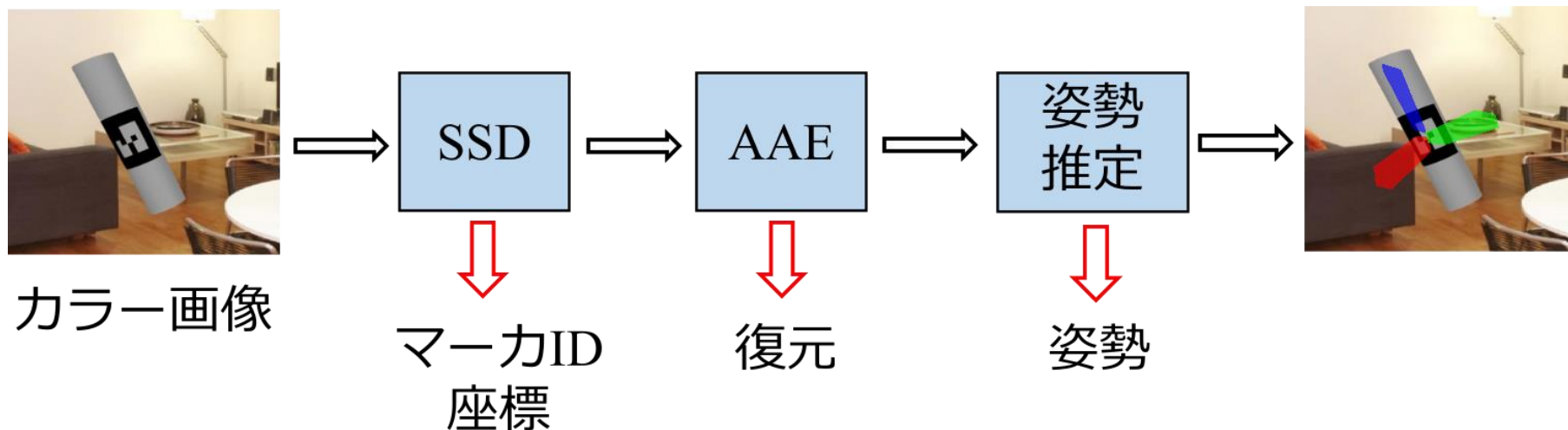


歪みにより認識不可

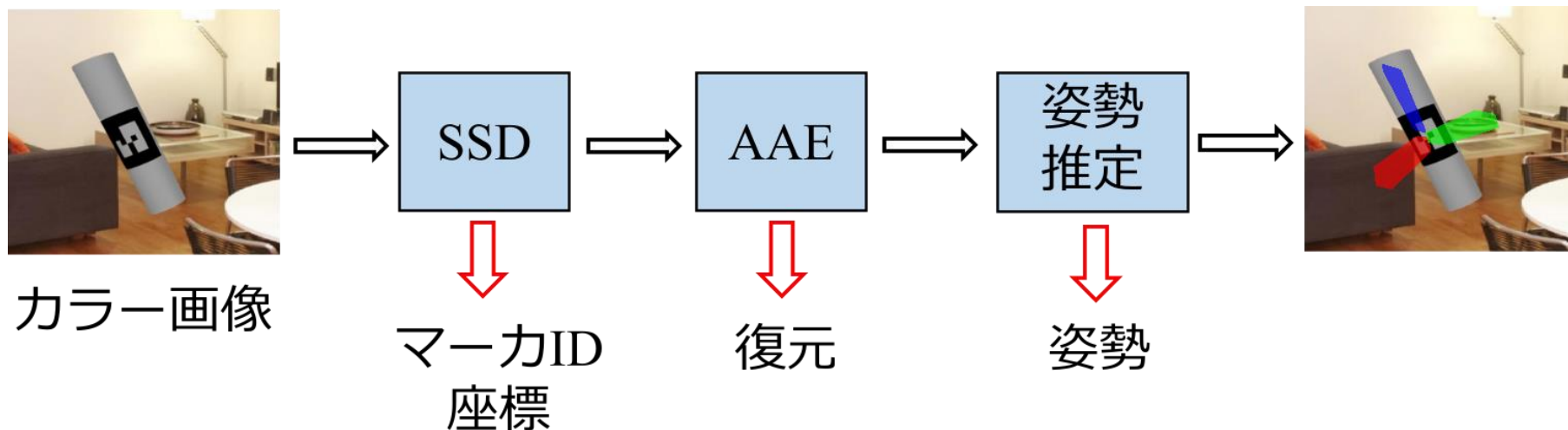
- 変形の加わったARマーカをカラー画像から  
「検出・平面状ARマーカへの復元・姿勢推定」を行う



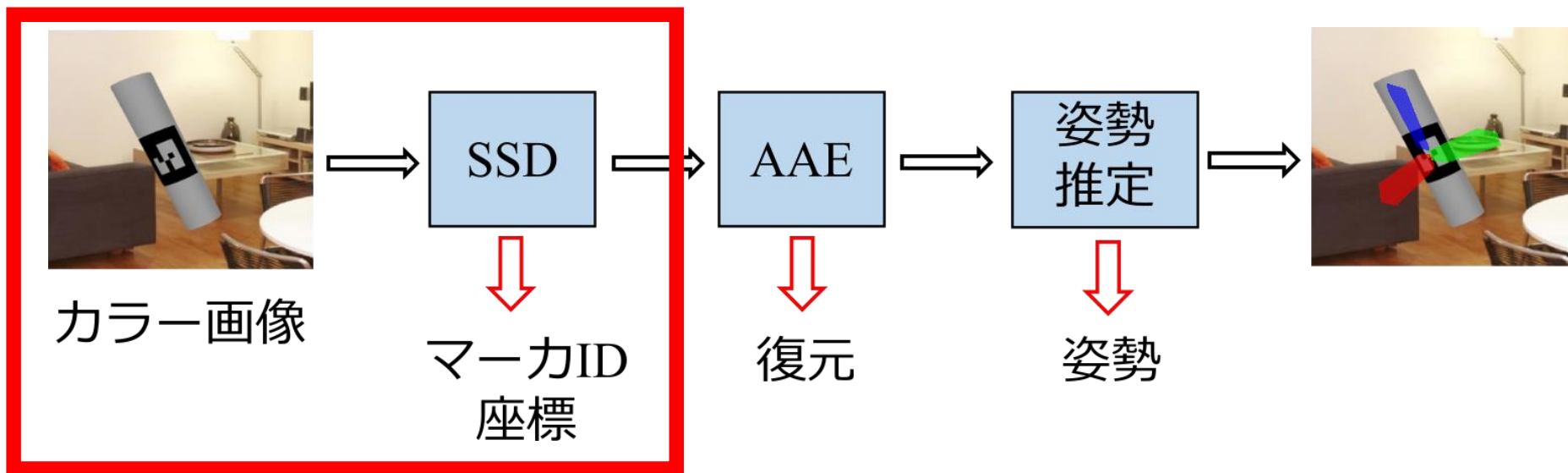
- SSD(Single Shot MultiBox Detector)によってARマーカを検出しID・座標を検出
- 変化の加わったARマーカをAAE(Augumented Autoencoder)での平面化を行い，取得した潜在変数より姿勢推定を行う
- SSDによるARマーカの検出法を提案



- SSD(Single Shot MultiBox Detector)によってARマーカを検出しID・座標を検出
- 変化の加わったARマーカをAAE(Augmented Autoencoder)により平面化を行い，取得した潜在変数より姿勢推定を行う
- SSDによるARマーカの検出法を提案



- SSD(Single Shot MultiBox Detector)によってARマーカを検出しID・座標を検出
- 変化の加わったARマーカをAAE(Augumented Autoencoder)での平面化を行い，取得した潜在変数より姿勢推定を行う
- SSDによるARマーカの検出法を提案



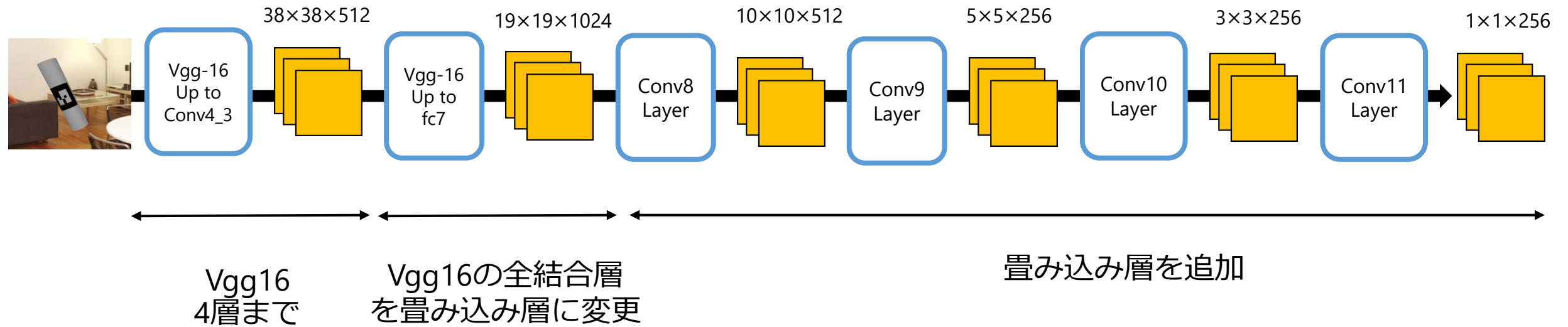


# SSD(Single Shot MultiBox Detector)

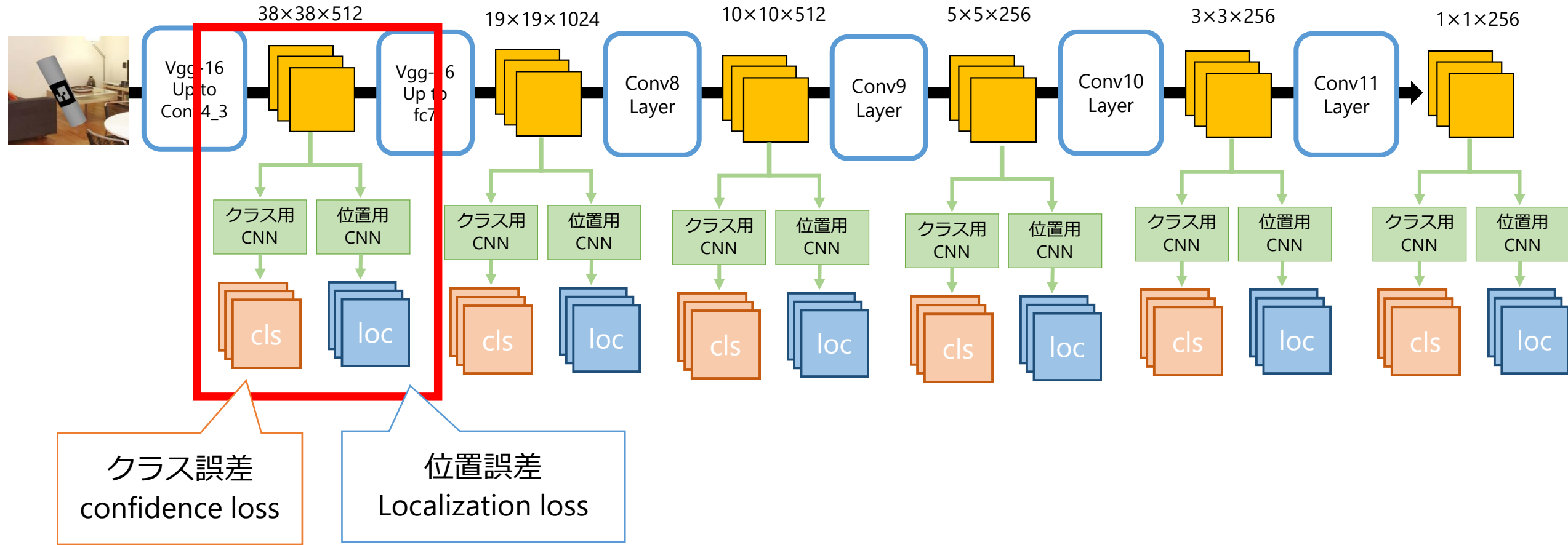
- 物体検出アルゴリズム
- バウンディングボックスで囲み位置とクラスを推定



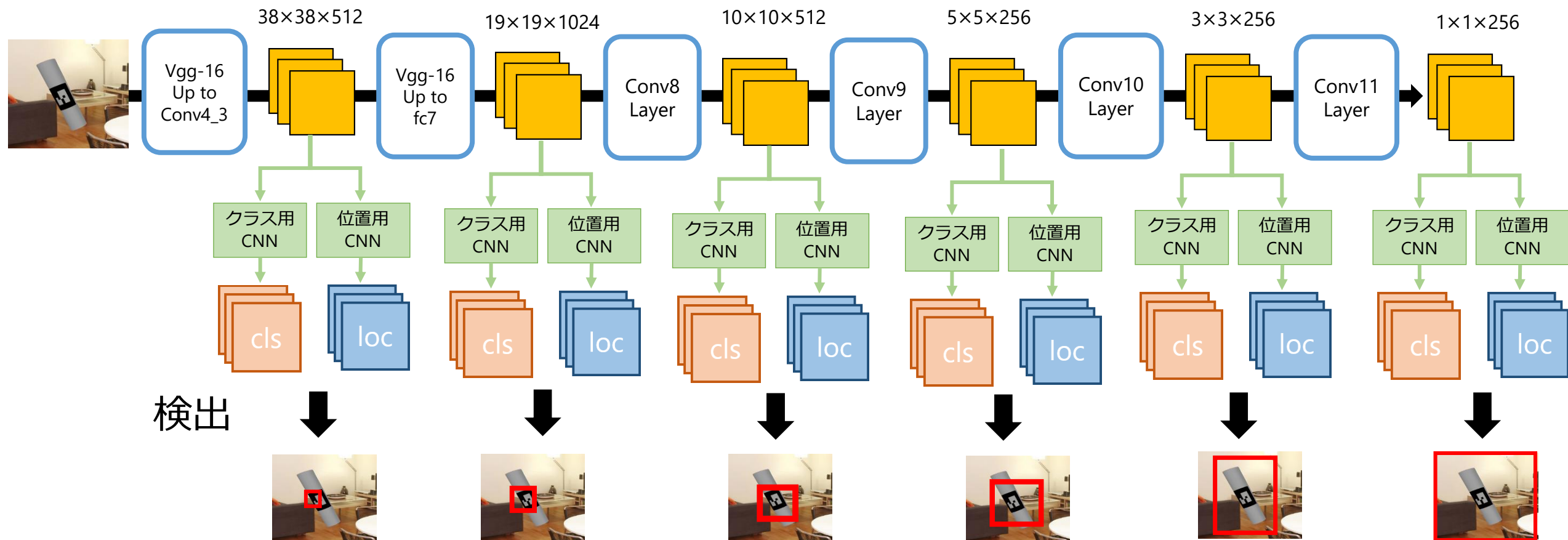
- サイズの異なる畳み込み層により様々な大きさの物体の特徴抽出



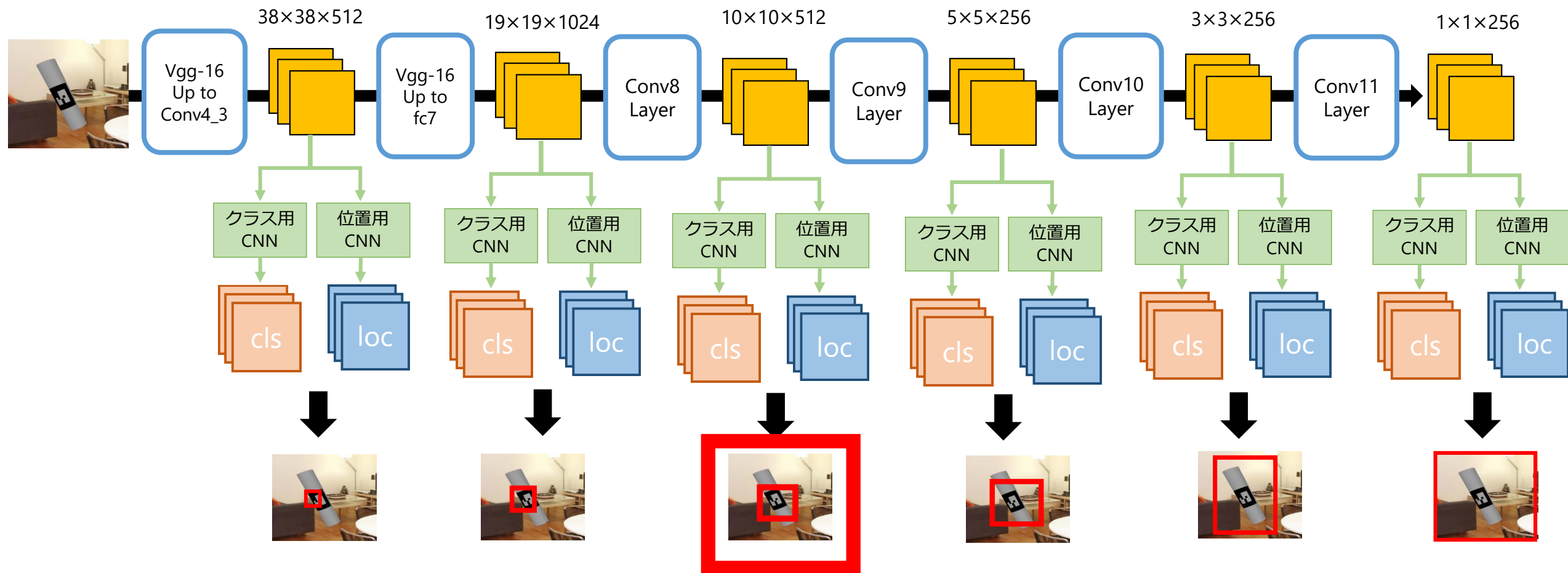
- 各特徴マップ毎にクラス特徴量と位置特徴量を算出



- 各畳み込み層で物体検出を行う

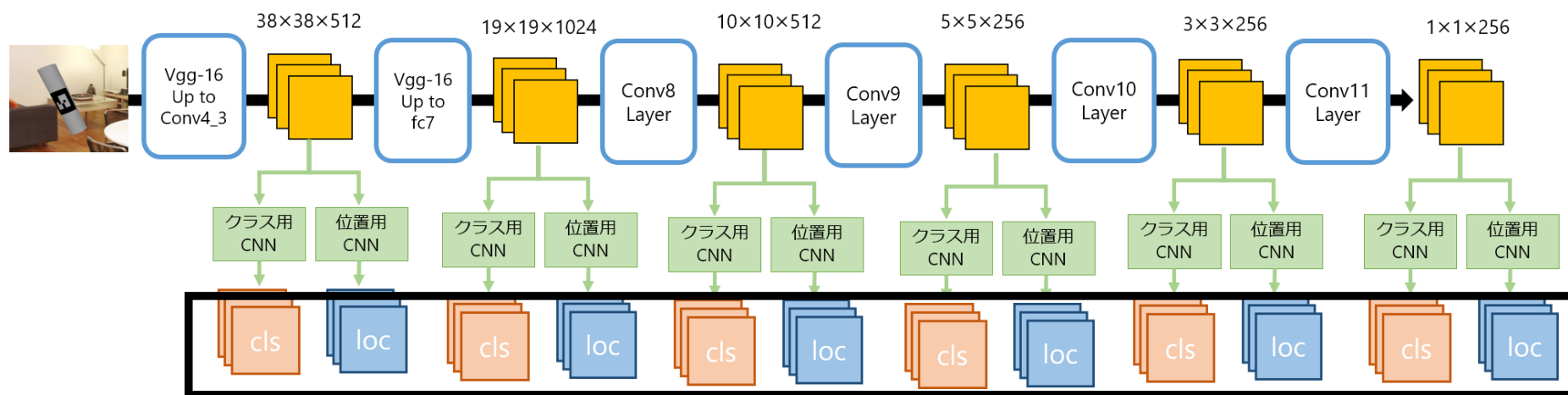


- 複数の検出結果から正しい位置との一致度が高いものを選択



IoU(重なり度合い)が高いものを選択

- ・クラス誤差，位置誤差の損失関数から正解データとの誤差を計算



①正解データと誤差を計算

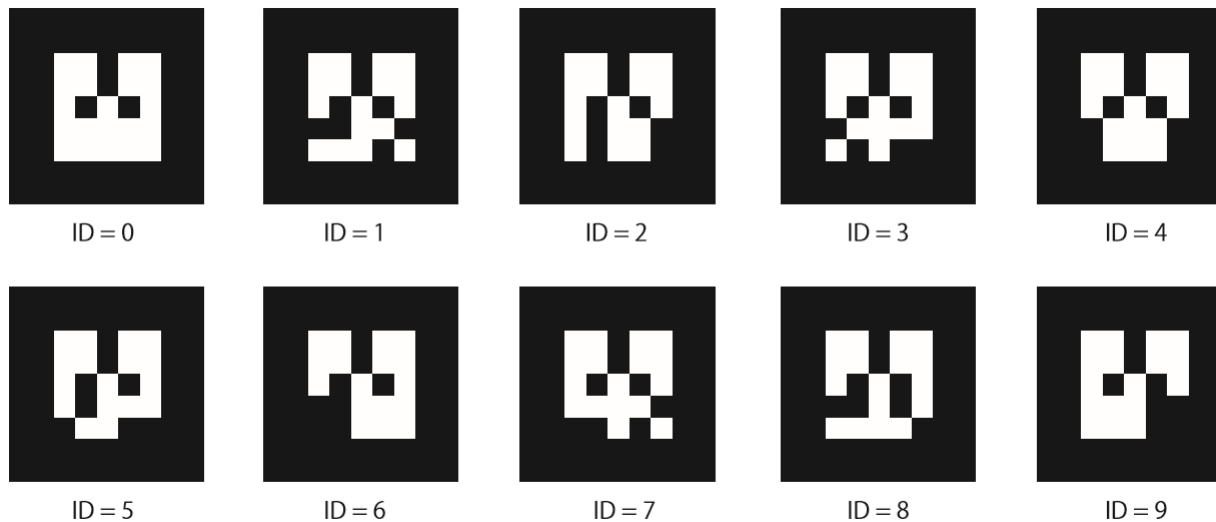
正解データ

②損失関数の計算

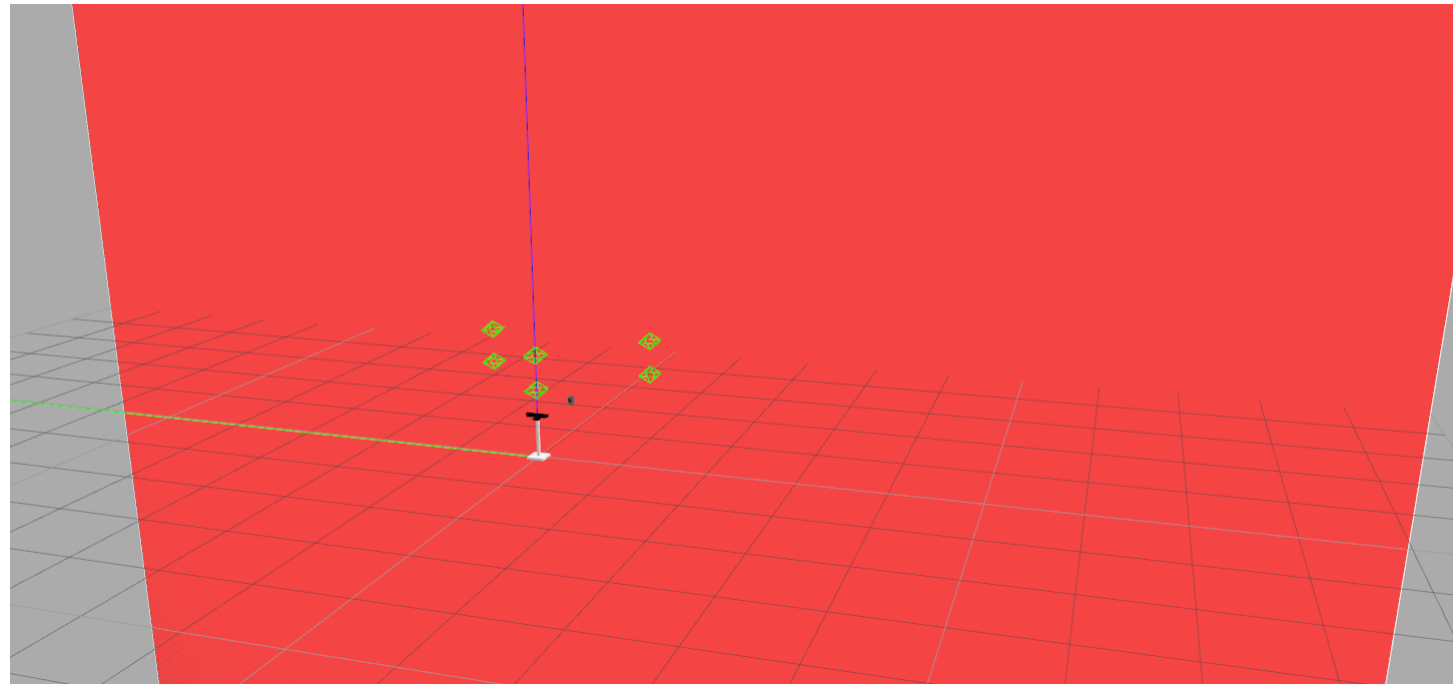
$$L(x, c, l, g) = \frac{1}{N} (L_{CLS}(x, c) + \alpha L_{LOC}(x, l, g))$$

③誤差逆伝搬法

- ROSパッケージのar\_track\_alvarのARマーカを用いた
- ARマーカの大きさ：縦横50mm
- ARマーカの種類：10種類（ID:0~9）
- 半径20,25,30,35,40,45,50mmの円柱に張り付ける



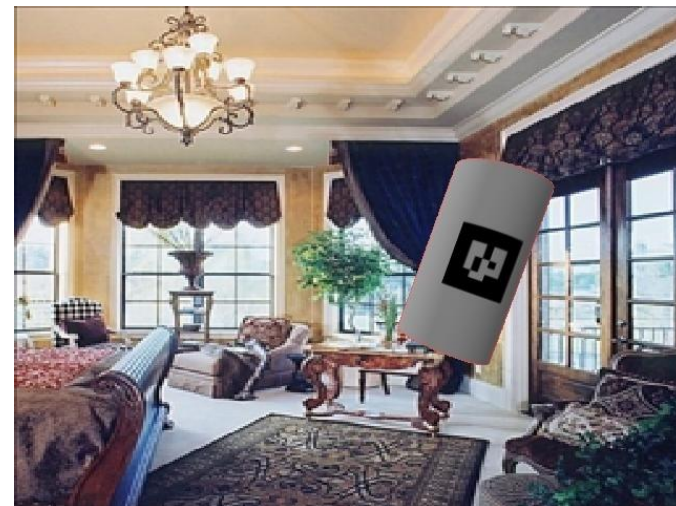
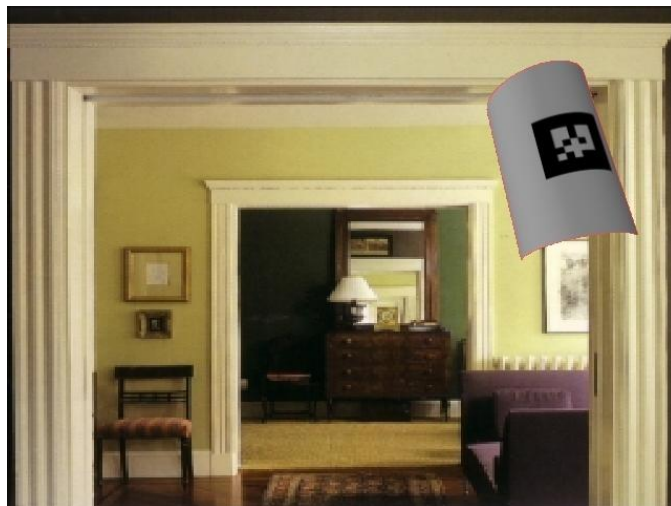
- SSDには大量の学習データが必要
- GazeboによってARマーカを自動生成
- Gazeboで図のように赤い板を設置した環境で撮影
- 背景が均一した赤色になるように撮影





- gazeboで撮影した画像を透過
  - 背景画像と合成しデータセットを作成
- ➡実環境でも検出できるように学習データを作成





- 距離の違いによるmAP・IoUの比較
  - カメラとARマーカの距離の違いによって検出精度の違い
- テスト画像
  - カメラとARマーカの距離の違いによって検出精度の違い
  - テスト画像1200枚, 学習データ6250枚
  - カメラとの距離0.4~0.5m, 0.5~0.6m, 0.6~0.7m, 0.7~0.8mの画像を各300枚

- 変形度合いによるmAP・IoUの比較
  - ARマーカの変形度合いの違いによって検出精度の違い
- テスト画像
  - テスト画像300枚, 学習データ6250枚
  - ARマーカが貼付された半径20,30,40の円柱の画像を各100枚用意

- カメラとの距離が0.4~0.7までは検出精度は変わらない
- 0.7以降は精度が下がった

カメラとの距離[m]	mAP	IoU
0.4~0.5	0.81	0.71
0.5~0.6	0.83	0.73
0.6~0.7	0.81	0.70
0.7~0.8	0.79	0.67

- 変形度合いが大きいほど検出精度は下がった

変形度(円柱の半径[mm])	mAP	IoU
40	0.85	0.73
30	0.83	0.71
20	0.79	0.68



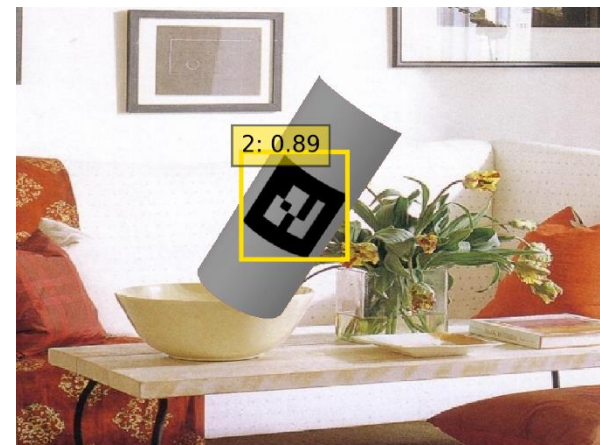
## • 検出例



ID:0



ID:1

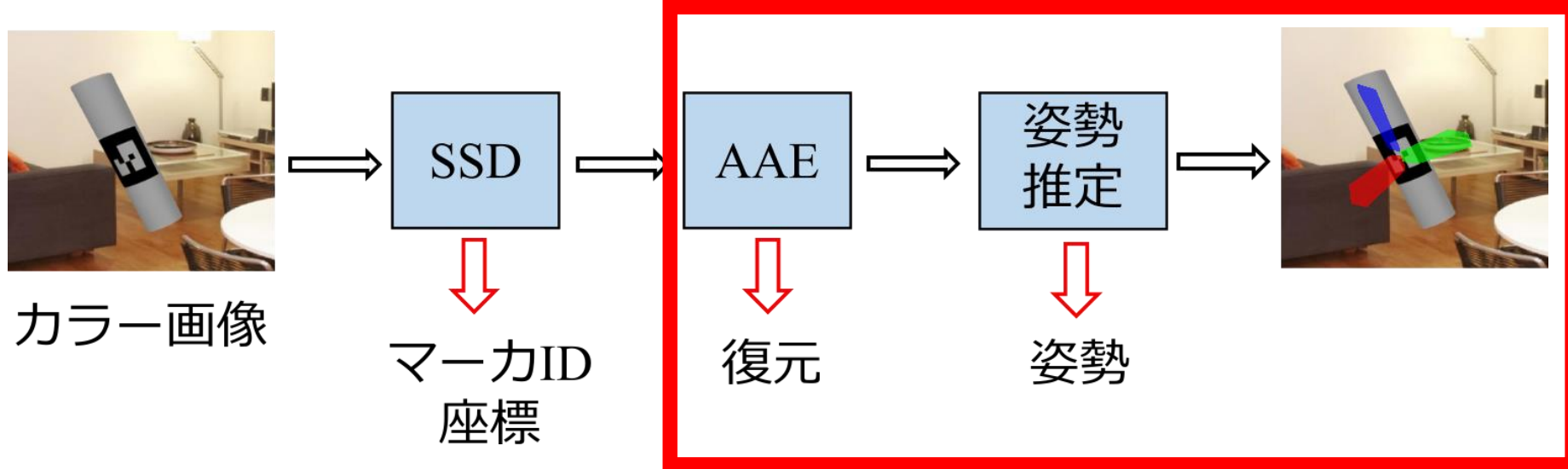


ID:2

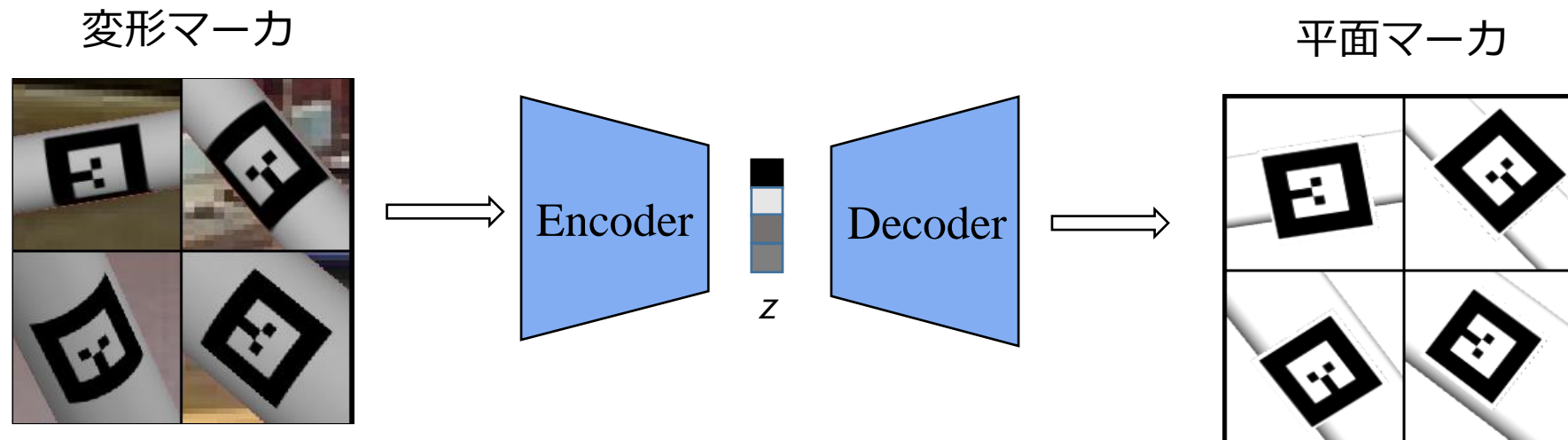
- 変形ARマーカの検出法を提案
  - 距離, 変形度による精度を比較
- 今後の課題
  - リアルタイム, 実環境での検出



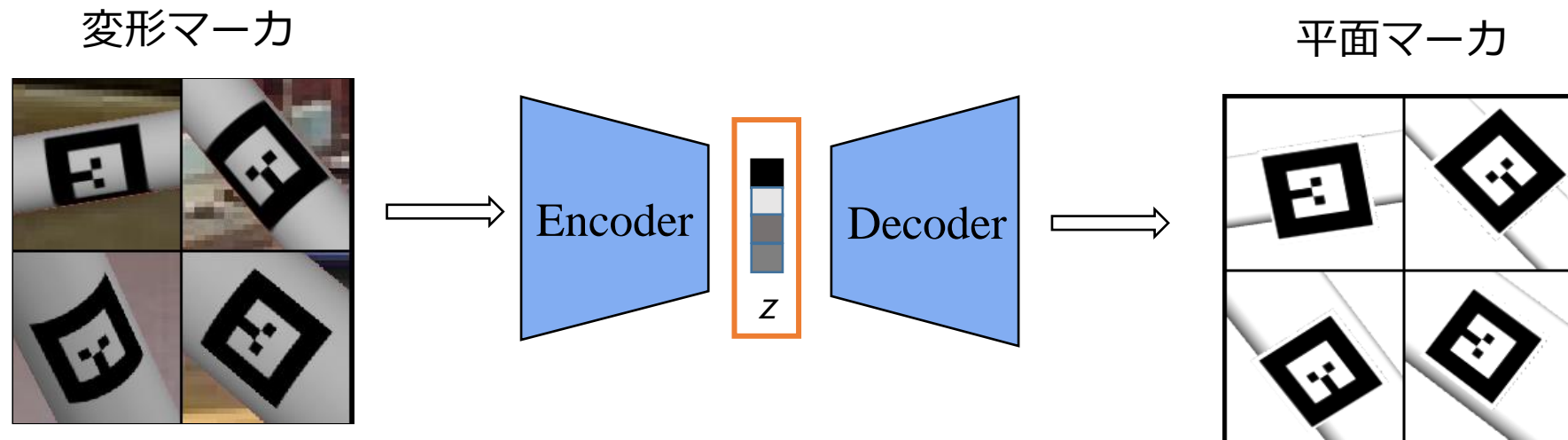
- SSD(Single Shot MultiBox Detector)によってARマーカを検出しID・座標を検出
- 変化の加わったARマーカをAAE(Augumented Autoencoder)での平面化を行い，取得した潜在変数より姿勢推定を行う
- AAEを用いた復元・姿勢推定の提案



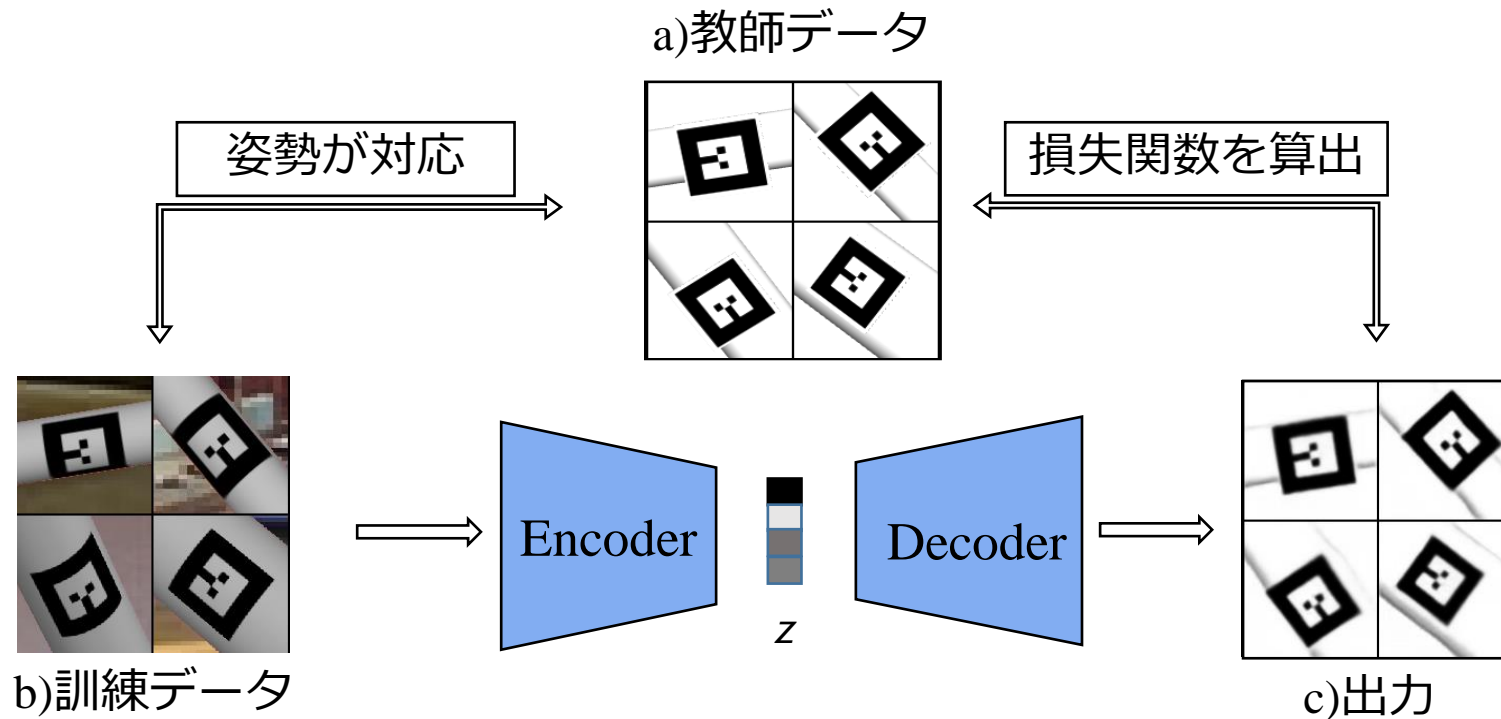
- 変化の加わったARマーカを平面状ARマーカへ復元
  - 背景画像や光，遮蔽物など環境ノイズを加え学習を行いノイズによらない本質的な潜在表現を取得できるよう学習
  - エンコーダーにより128次元まで圧縮されたマーカの姿勢情報を潜在変数 $z$ として取得



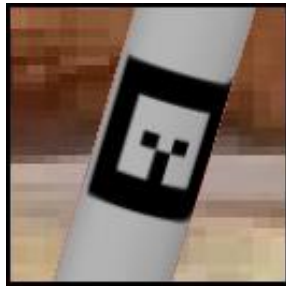
- 変化の加わったARマーカを平面状ARマーカへ復元
  - 背景画像や光，遮蔽物など環境ノイズを加え学習を行いノイズによらない本質的な潜在表現を取得できるよう学習
  - エンコーダーにより128次元まで圧縮されたマーカの姿勢情報を潜在変数 $z$ として取得



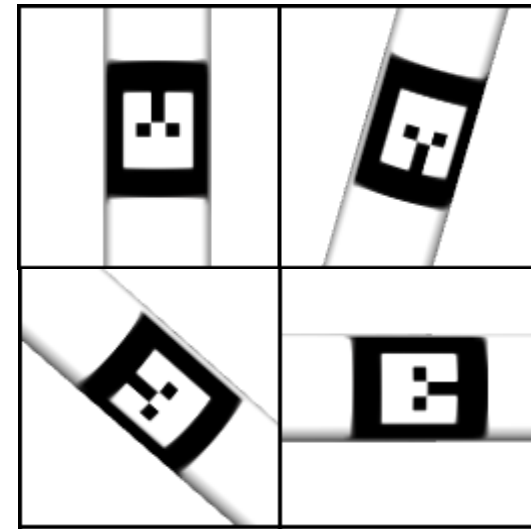
- 入力(b):教師データ(a)と姿勢が対応する背景付きの変形ARマーカ画像
- 出力(c):教師データ(a)との損失関数が小さくなるように学習
- 学習により歪みのないARマーカが復元可能



- 学習済みのエンコーダーによって得られる潜在変数 $z$ を使用し類似度計算を行い姿勢を推定
- 推定対象画像とデータベース(DB)画像の2つを使用

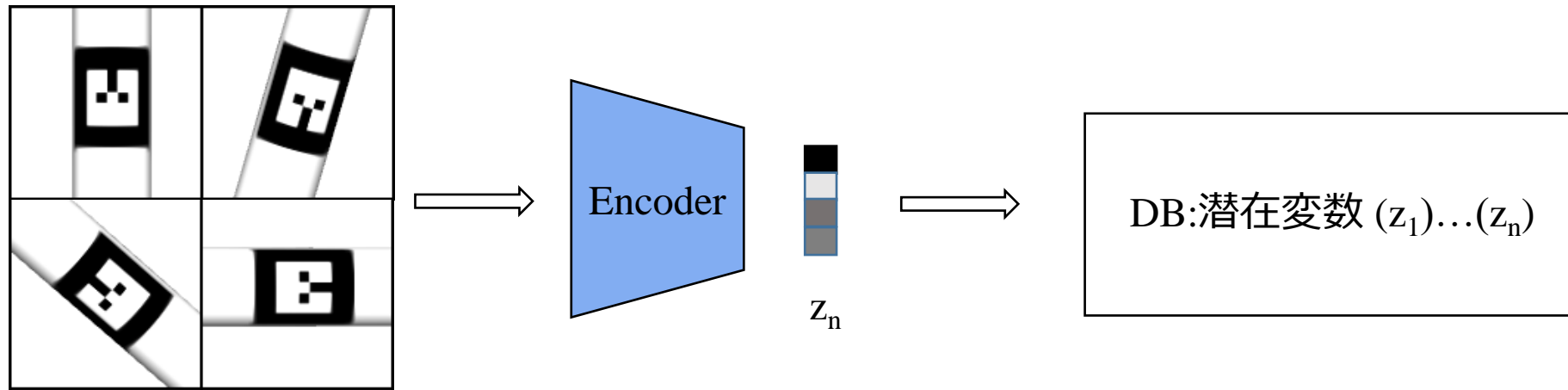


推定対象画像



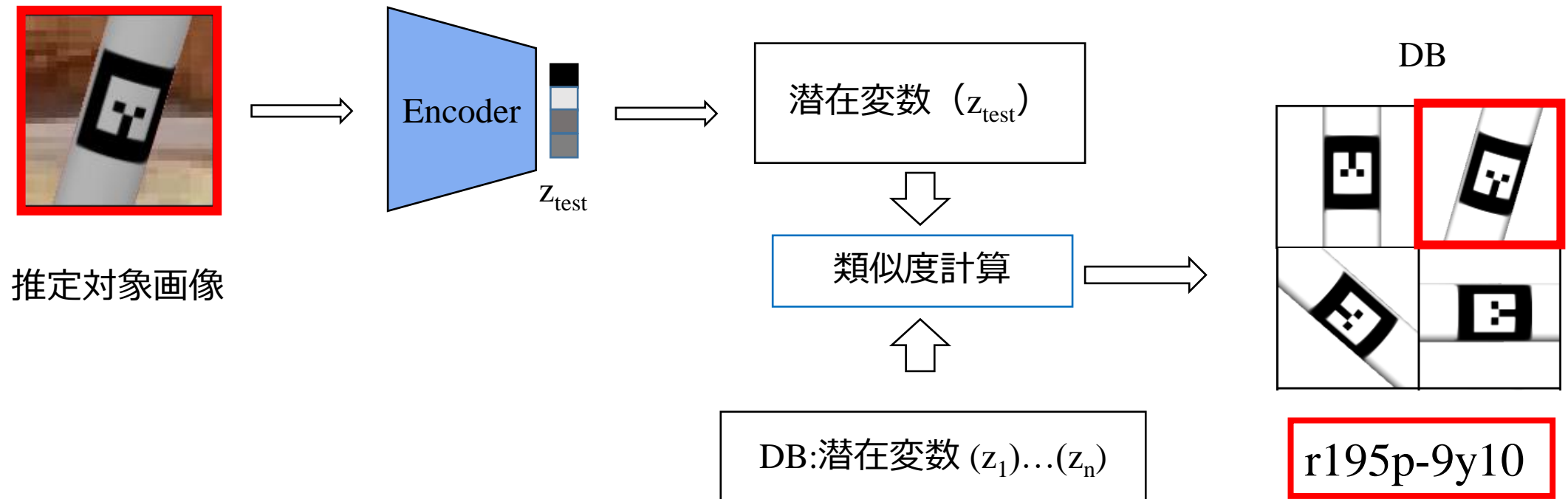
DB画像

- 各姿勢の変形ARマーカの画像をあらかじめAAEに入力
  - 各姿勢画像(n枚)それぞれの潜在変数( $z_n$ )をDBとして保存

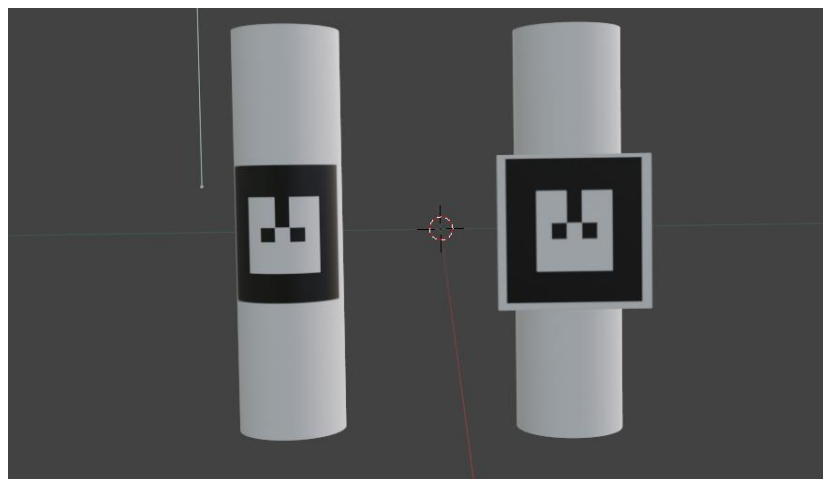


各姿勢の画像 (n枚)

- 推定対象画像をエンコーダーに入力し潜在変数を取得
- 取得した潜在変数とDBの類似度を計算
  - 最も近いDBの姿勢を推定姿勢として決定
  - 推定姿勢[roll,pitch,yaw]は度数法で標記



- ARマーカのサイズは縦横50mm
- 半径20, 30, 40mmの円柱に貼り付けたモデルを使用
- 学習モデルの種類
  - 「訓練データモデル・教師データモデル」2種類それぞれID0~9×半径3種類の合計60種類のモデルを使用



使用モデルの例 (ID=0, 半径=20)



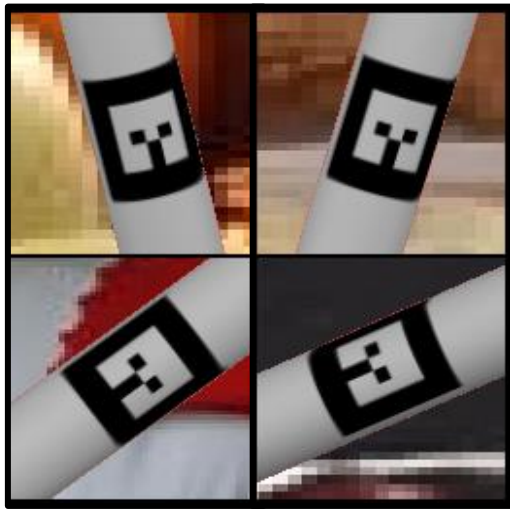
- センサシミュレーションにより学習用画像を作成
  - 学習画像は1種類あたりランダム姿勢1500枚撮影
  - 「訓練データ・教師データ」の60種類, 合計90,000枚用意
  - 画像生成後, 変形ARマーカの背景にはテクスチャを付け現実環境を仮定



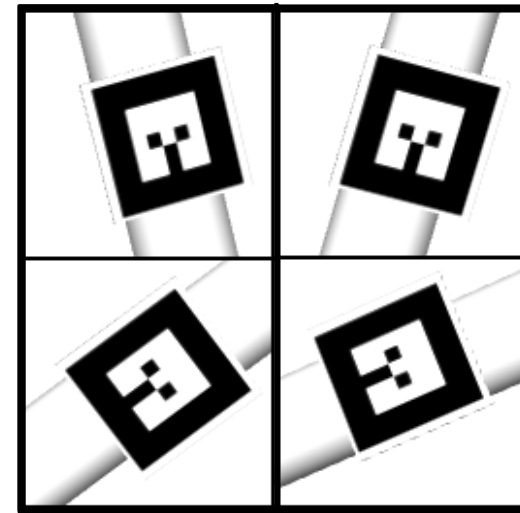
- 平均絶対誤差（MAE）を用いて推定精度を評価
  - 推定姿勢[roll, pitch, yaw]の誤差を算出
- 姿勢推定で使用するDB
  - モデル姿勢範囲： roll:0~360°pitch:-35~35°yaw:-15~15°
  - 分解能 3 度となるよう36,000枚の姿勢画像をDBに使用

- 評価データ

- 画像：20, 30, 40の各半径100枚
- モデル姿勢：ARマーカが半分以上見える範囲内からランダム

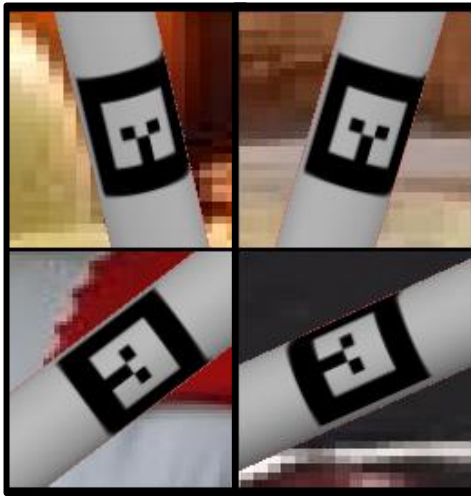


評価データ

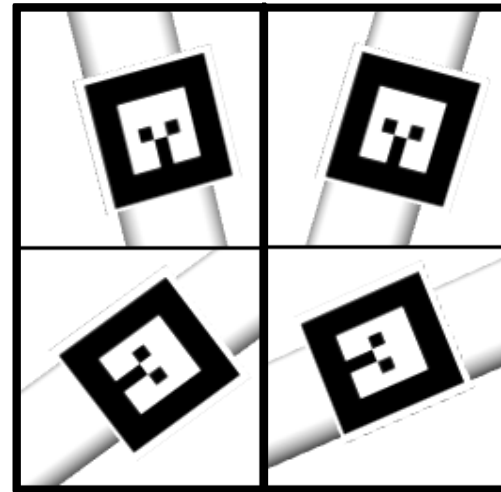


評価データの正解画像

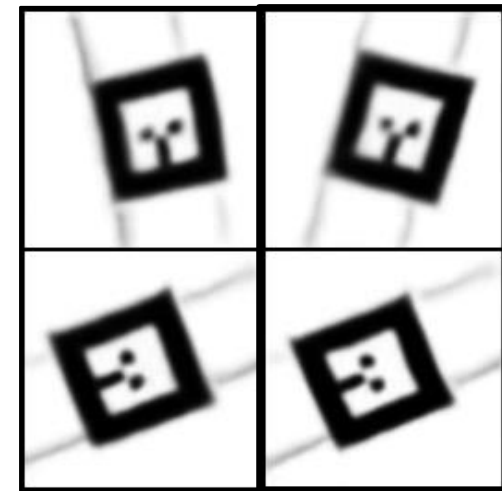
- AAEによる評価データの復元画像
  - 変化を取り除き平面化を行えていることを確認



評価データ



正解画像



復元画像

提案手法における姿勢推定精度 MAE

円柱半径[mm]	roll	pitch	yaw	姿勢平均
20	5.30	3.64	3.42	4.12
30	5.78	4.49	3.71	4.66
40	6.52	4.51	3.73	4.91

- 姿勢推定のMAEは、4～5前後となりズレはあるが姿勢推定は可能
- 半径が小さいモデルほど推定精度が高い
- ➡半径が小さいモデルほど姿勢ごとの画像特徴が明確になる為

提案手法における姿勢推定精度 MAE

円柱半径[mm]	roll	pitch	yaw	姿勢平均
20	5.30	3.64	3.42	4.12
30	5.78	4.49	3.71	4.66
40	6.52	4.51	3.73	4.91

- 姿勢推定のMAEは、4～5前後となりズレはあるが姿勢推定は可能
  - 半径が小さいモデルほど推定精度が高い
- ➡半径が小さいモデルほど姿勢ごとの画像特徴が明確になる為

提案手法における姿勢推定精度 MAE

円柱半径[mm]	roll	pitch	yaw	姿勢平均
20	5.30	3.64	3.42	4.12
30	5.78	4.49	3.71	4.66
40	6.52	4.51	3.73	4.91

- 機械学習による姿勢推定を提案
  - AAEを用いて変形ARマーカの平面化
  - 潜在変数を用いた姿勢推定が可能であることを確認
- 今後の課題
  - 推定精度を上げるため
    - DBの分解能を1度で用意
    - 学習データのバリエーションを増やし潜在変数を明確に表現
  - 実環境下での姿勢推定