

令和5年度 静岡県支援機関等連携促進会議

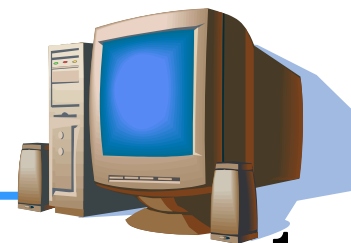
深層学習の仕組みと応用

— 画像解析における事例 —

2023年 7月 13日 (木)

静岡理工科大学 総合技術研究所

工藤 司



略歴

1978年 北海道大学大学院工学研究科修士課程情報工学専攻入学

1980年 三菱電機株式会社

- ✓ 研究所:スーパーコンピュータ開発 (日本初)
- ✓ SE部門:津波予報システム (日本初), などの開発

2005年 三菱電機インフォメーションシステムズ株式会社

- ✓ 品質保証部:システム開発業務のプロセス改善

2010年 静岡理工科大学 情報学部 教授

2020年 静岡理工科大学 情報学部 特任教授

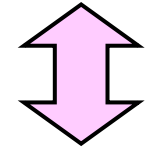
- ✓ 経営情報システム研究室 (データベース, 深層学習の応用)

2023年 静岡理工科大学 総合技術研究所 客員教授

- ✓ 新キャンパス推進準備: 企業のDXに関する技術支援

2024年4月に静岡駅前到新キャンパス誕生！

地域・社会



■地域協働センター

- ✓ 産官学連携推進
- ✓ 地域連携推進
- ✓ 社会人教育推進

■静岡理工科大学 情報学部 サテライト研究室

■静岡デザイン 専門学校



1. 深層学習のインパクト

— あいまいな対象へ —

2. 深層学習の仕組み

— 画像解析における事例 —

3. 物体検出

— YOLOによるジェスチャ認識 —

4. おわりに

1. 深層学習のインパクト

— あいまいな対象へ —

深層学習（人工知能）普及のきっかけ

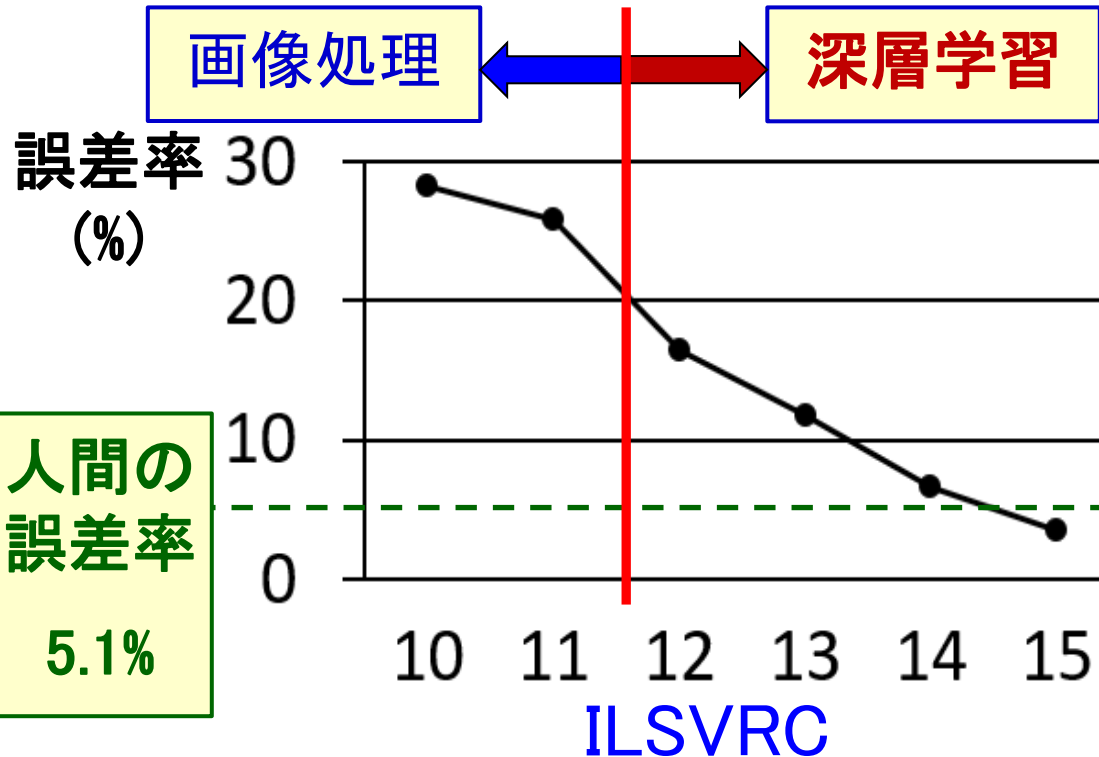
- ILSVRC（大規模画像認識の競技会）での成果
- 大量の学習データが収集可能（ビッグデータの活用）



■ 様々な分野へ応用

- ✓ 碁, 将棋
- ✓ 機械翻訳
- ✓ 顔認証

.....



【出典】 M. VERHELST, B. MOONS: Embedded Deep Neural Network Processing: Algorithmic and Processor Techniques Bring Deep Learning to IoT and Edge Devices. IEEE Solid-State Circuits Magazine, Vol.9, No.4, pp.55-65, 2017. により作成

基本的な深層学習モデル

1. 多クラス分類モデル

対象がどのクラスに分類されるか推定（顔認証）

2. 回帰モデル

連続量の推定（在庫数の推定）

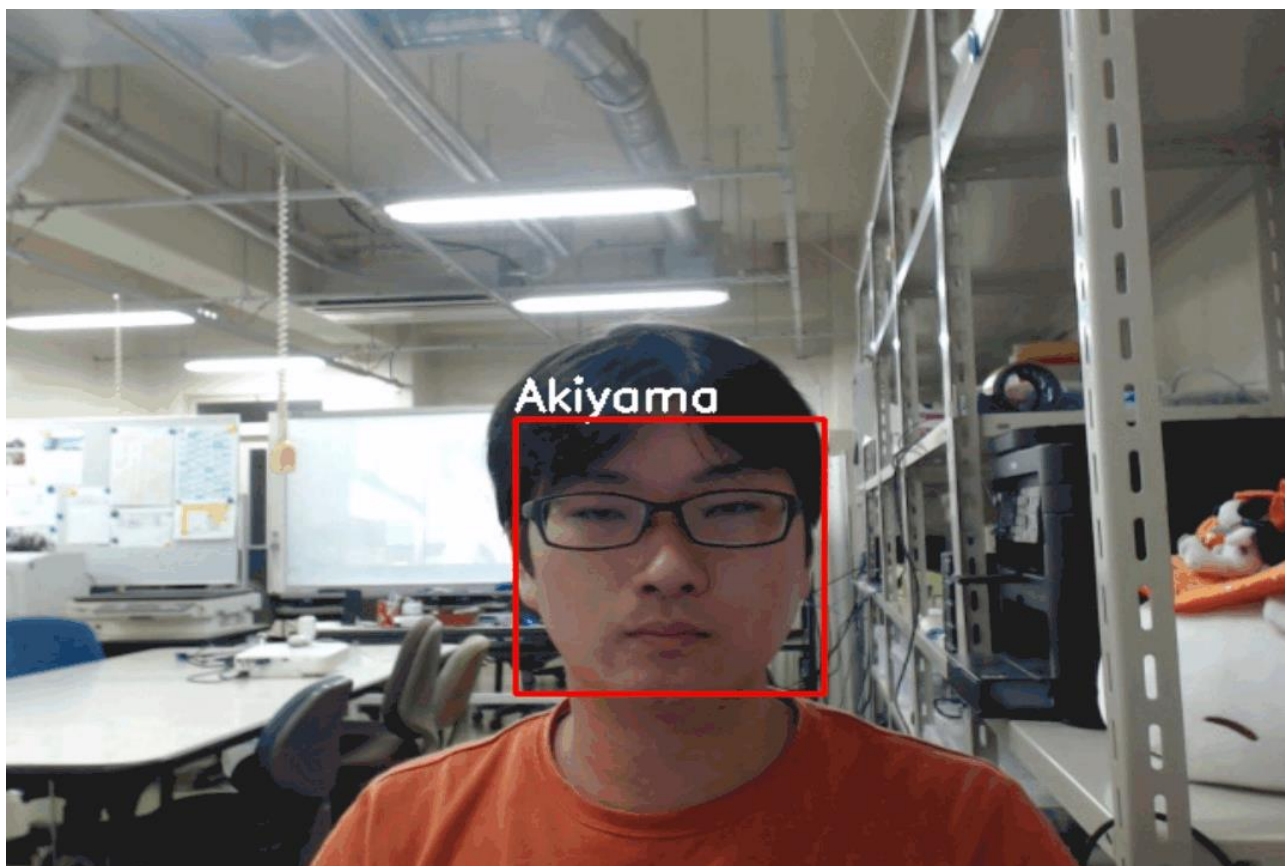
3. GAN (Generative Adversarial Networks; 敵対的生成ネットワーク)

画像, 文章, 音楽などを自動的に生成・変換
(作画, 作詩, 作曲)

多クラス分類モデル

■ 動画からの顔認証

1. 顔検出機能で、顔の領域を抽出
2. **多クラス分類モデル**で対象者を判別



回帰モデル

- 連続量の推定で使用：在庫数の推定
 - **多クラス分類モデル**：全ての数量の学習が必要
 - **回帰モデル**：学習した数量の間も、補完して推定

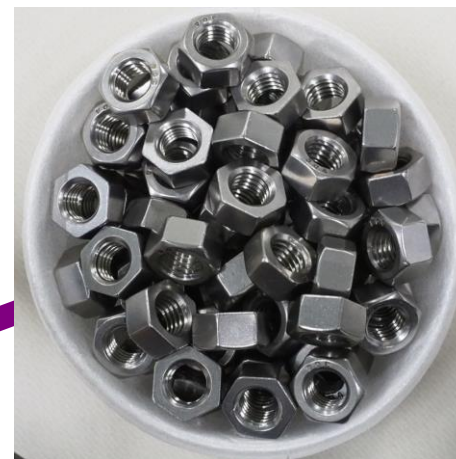
10



45



80



GAN（敵対的生成ネットワーク）

■ GAN（プログレッシブGAN）による自動生成画像の例



【出典】 T. Karras et. al.: Progressive Growing of GANs for Improved Quality, Stability, and Variation – Official TensorFlow implementation of the ICLR 2018 paper

https://github.com/tkarras/progressive_growing_of_gans

■ 自動生成・変換を行う様々な分野の研究

- アニメーションの製作，画像合成（草原の車），
- 画像変換（モノクロからカラー画像作成），などなど

データ処理における**深層学習**のインパクト

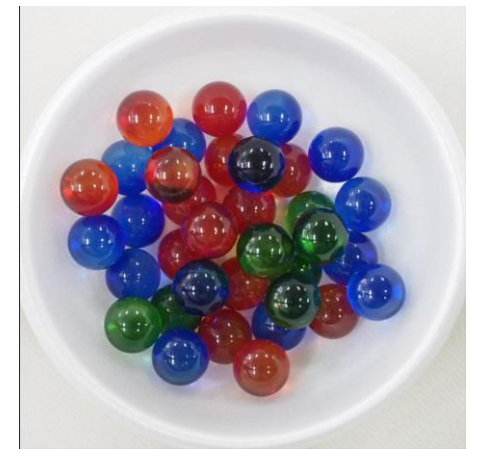
■ 従来データ処理：計算機

- ✓ 数値として計算できるものが対象
- ✓ 数値, 文字.
画像・音声(数値として処理)



■ 深層学習：特徴を自動的に識別し推定

- ✓ あいまいな対象まで適用を拡大
- ✓ 人の顔 (顔認証)
画像の分類, 個数の推定
文章解析, 音声認識, など



2. 深層学習の仕組み

— 画像解析における事例 —

背景：企業の生産管理システム構築支援

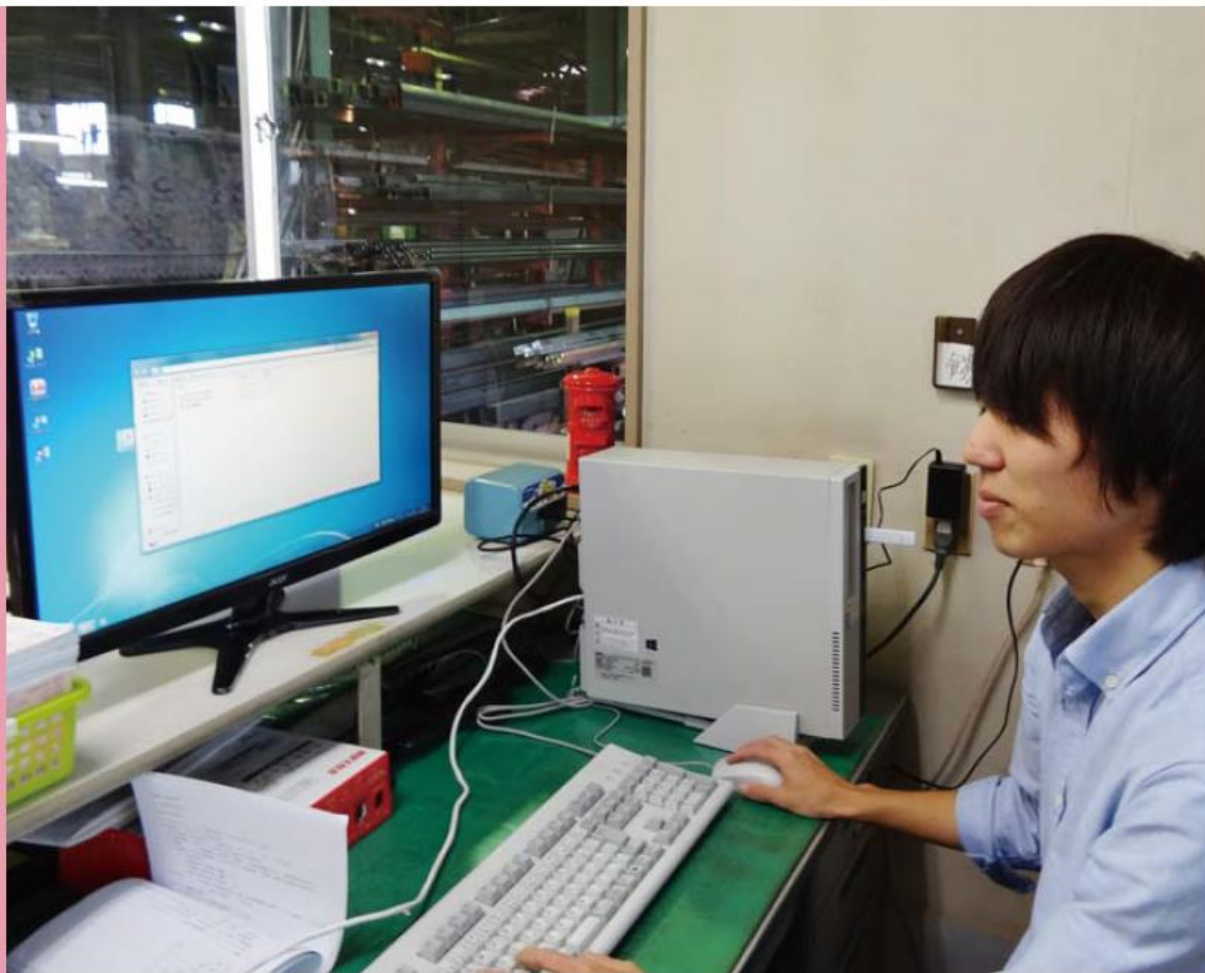
経営・社会システム 分野

データベースや統計をはじめとする
アプリケーションソフトウェアの高度利用技術や
ユーザー動向、事業経営などを学びます。

キーワード データマイニング／ビッグデータ／トレンド分析／価値／商品開発／事業企画

【本学 大学案内より】

情報に溢れたこの時代。
企業経営や人々の生活に供する
情報の入力と運用方法を探る。

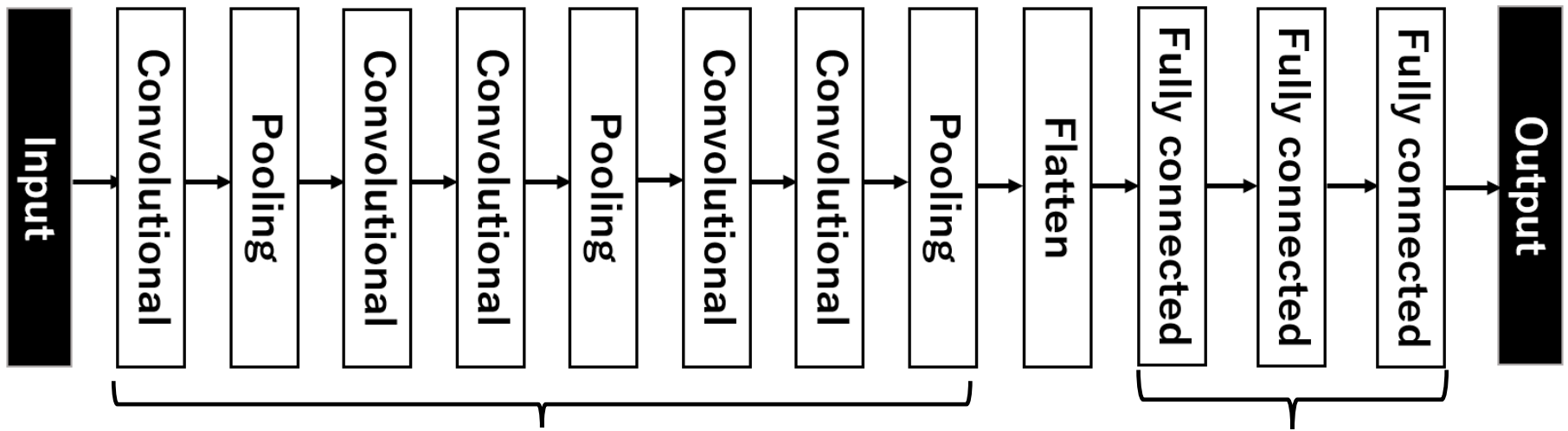


工場における在庫管理の課題

- バラ積みの部品の在庫は、**目視で数えられない**
 - 取り出して数えるため、大きな手間、**棚数は数千**。



画像認識における深層学習の応用

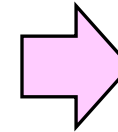
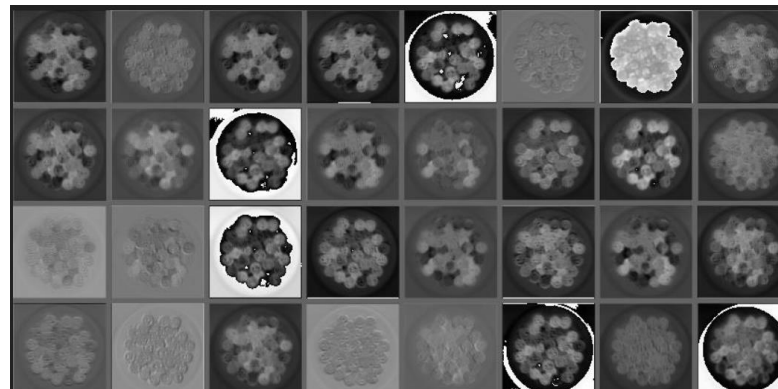
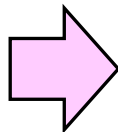
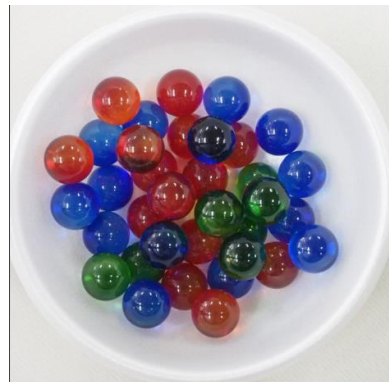


畳込み層

全結合層

画像から様々な特徴を抽出

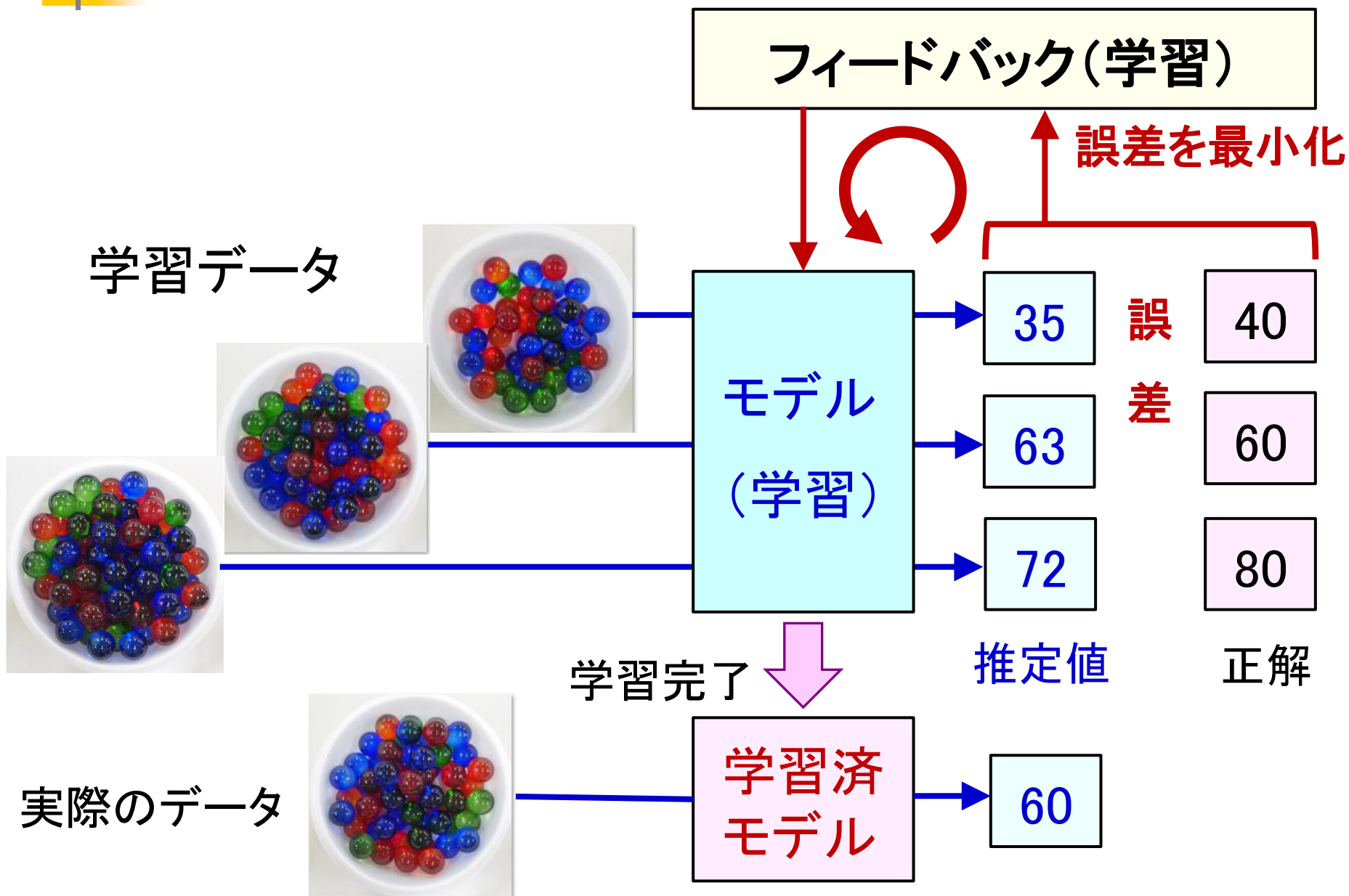
特徴から入力を推定



数量推定

60個

深層学習 (教師あり学習)



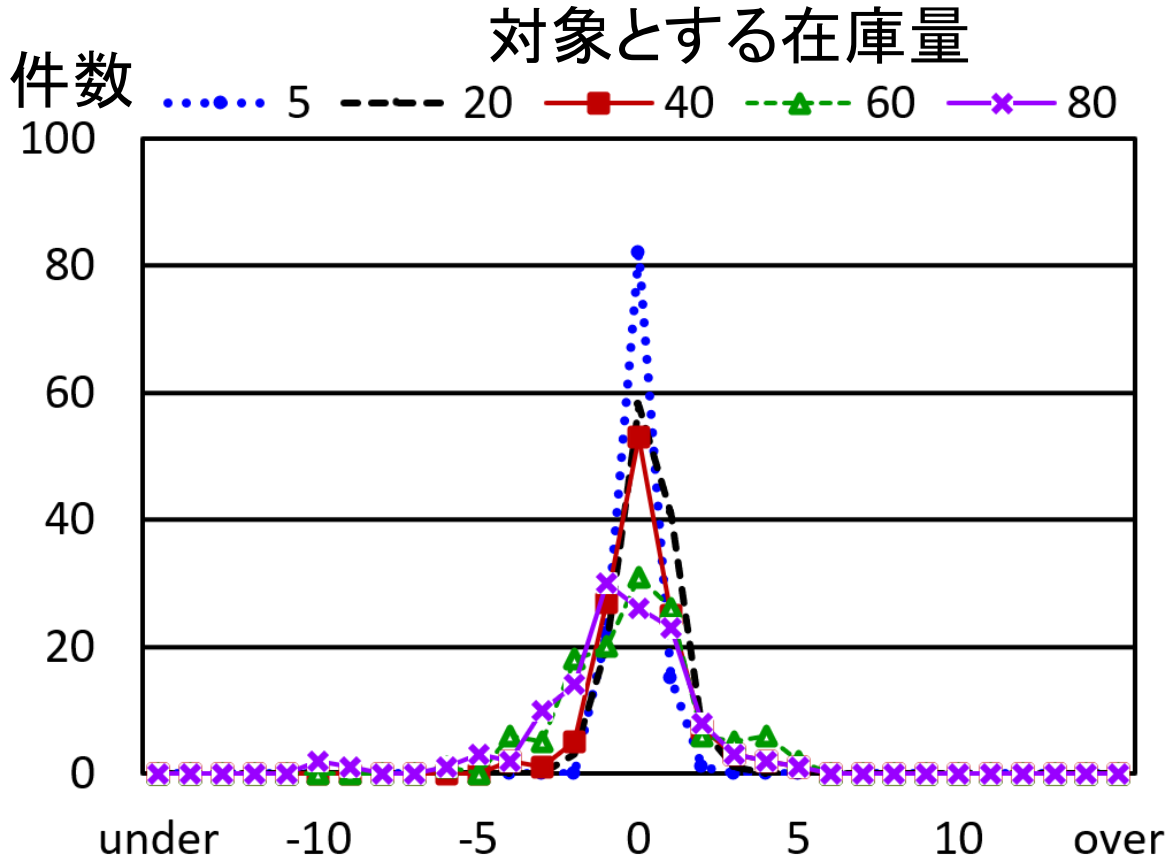
深層学習の在庫量推定モデルの改善

- モデルとパラメータの調整による改善
 - 概ね、**誤差±5以内**にできることを確認



対象（ナット）

(5, 10, 15, ..., 80個)

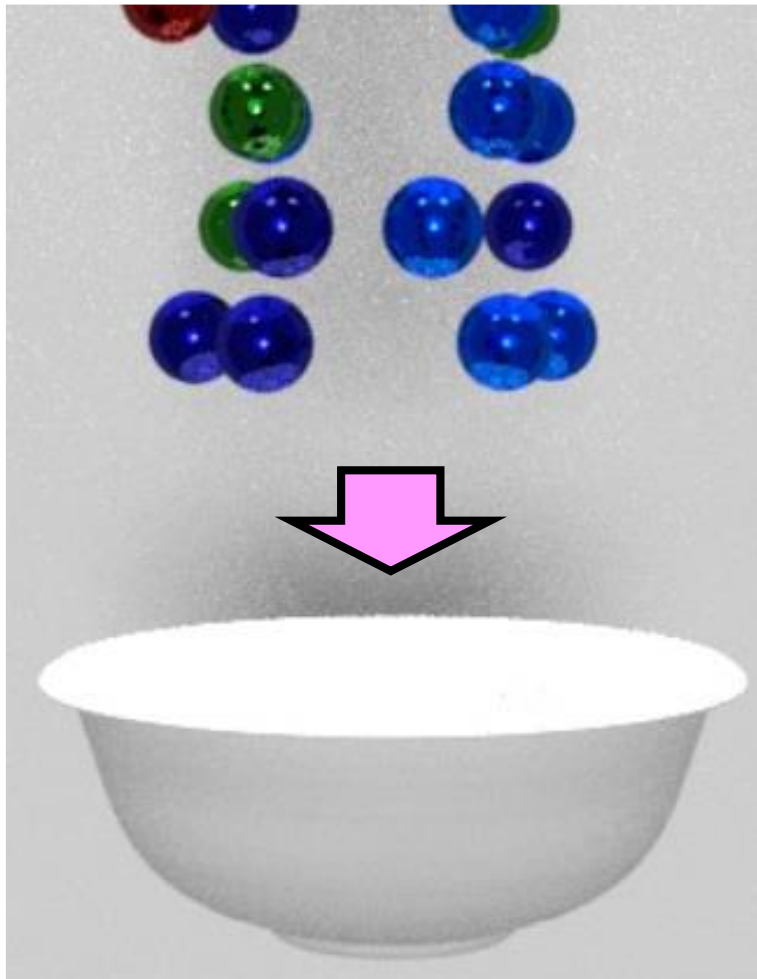


誤差分散のヒストグラム

深層学習の課題：大量の学習データを準備



CGを活用した学習データの自動作成



落下後の最終状態

1. 対象をランダムに配置, 物理シミュレーションで位置決定
2. これをプログラムで自動実行

CG ⇔ 写真 の 自動画像変換

- Cycle-GAN: ある画像から別の種類の偽画像を生成敵対的生成ネットワーク (Generative Adversarial Network, GAN) の1種で, 2017年に提案された.
- Cycle-GANによる, シマウマから馬への変換事例

Zebras ⇄ Horses



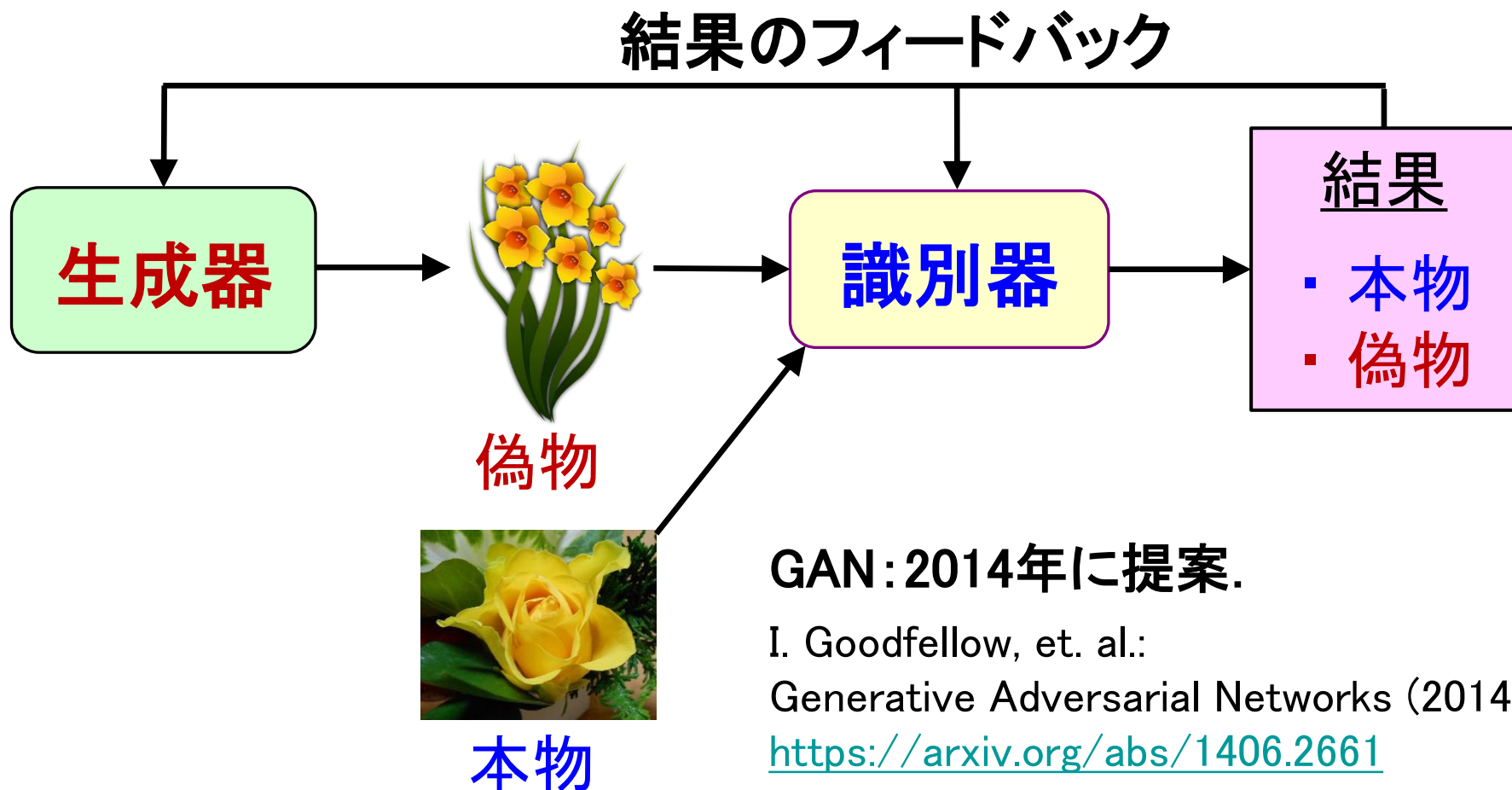
zebra → horse

【出典】 J. Y. Zhu, et. al.: Unpaired image-to-image translation using cycle-consistent adversarial networks,” Proc. IEEE Int. Conf. on computer vision, pp. 2223–2232 (2017).

<https://github.com/junyanz/pytorch-CycleGAN-and-pix2pix>

GANの基本的な仕組み

- **生成器**は**識別器**を騙すように、
識別器は騙されないように、訓練する。

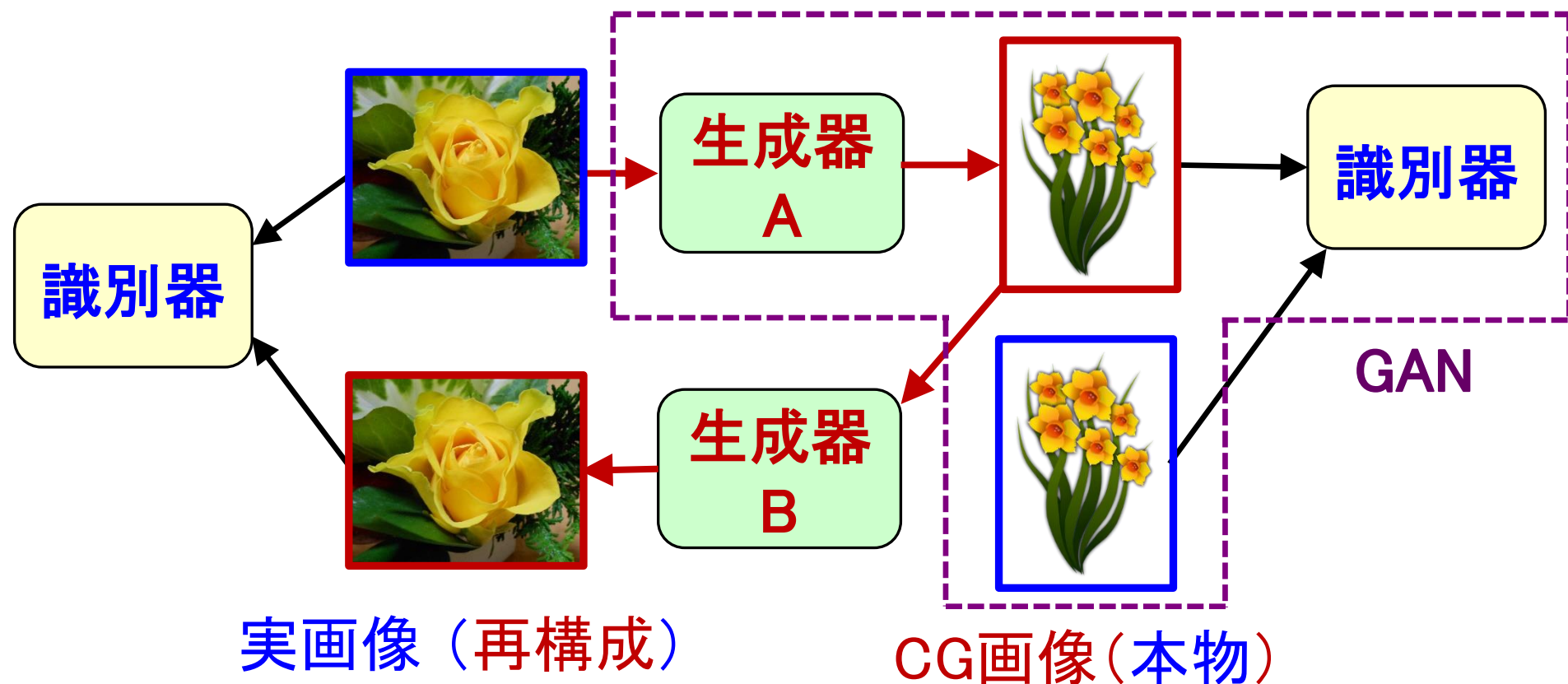


Cycle-GANの基本的な構成

- 実画像(本物)からCG画像(偽物)を作成,
さらに, 実画像(再構成)を作成

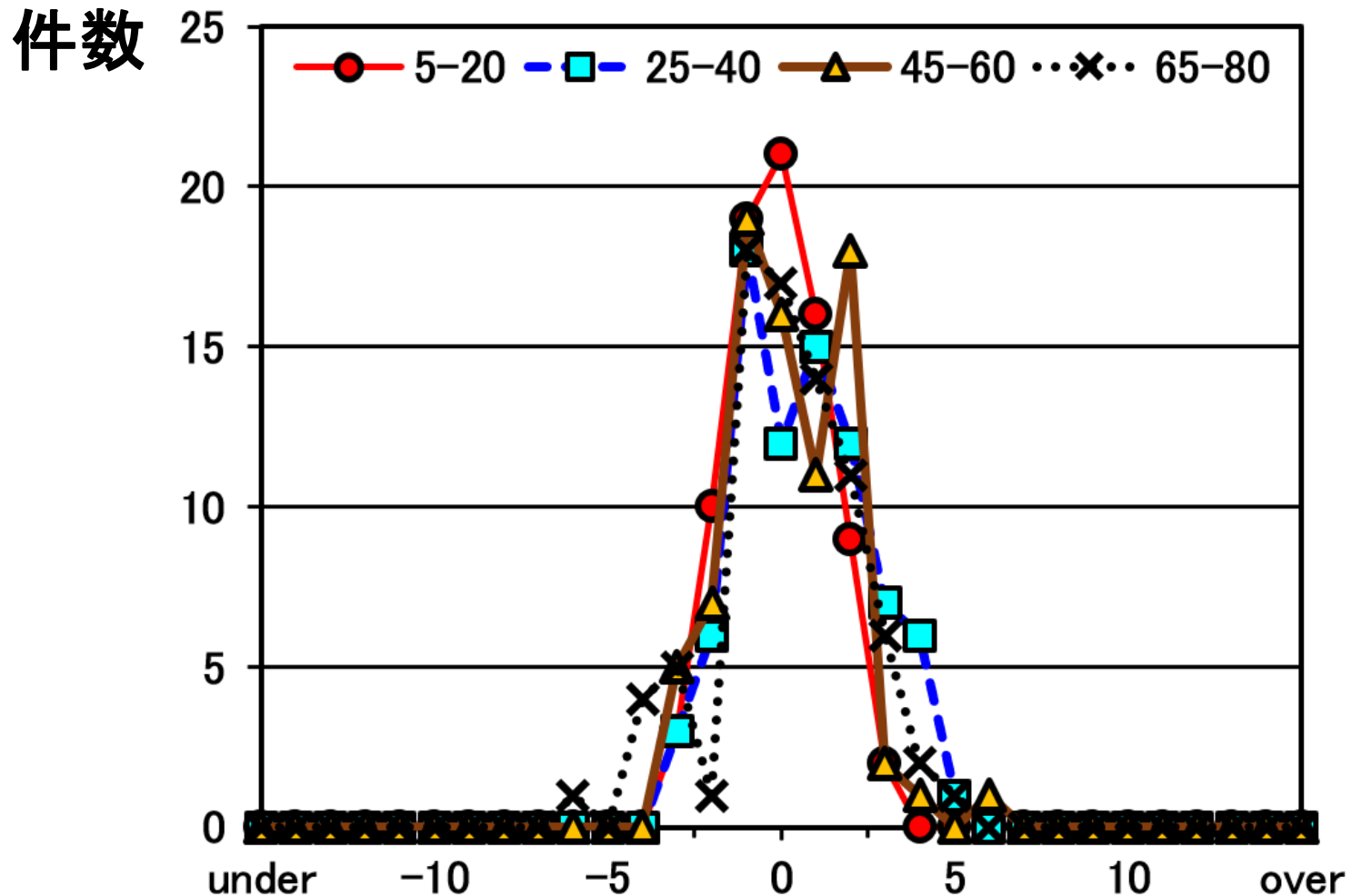
実画像 (本物)

CG画像 (偽物)



在庫モデルを改良：推定精度の評価

■ 目標誤差（±5以内）を概ね達成



誤差分散のヒストグラム

3. 物体検出

— YOLOによるジェスチャ認識 —

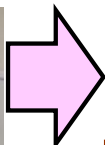
数千のコンテナの在庫管理自動化研究

■ 在庫管理の自動化方式

- (1) 作業員でウェアラブルカメラで、動画を自動収集.
- (2) **コンテナの画像**から、**在庫推定**.

■ 課題: 対象コンテナを、どうやって検出するか?

Camera



(1) 動画収集

【課題】対象コンテナの検出

(2) 在庫推定

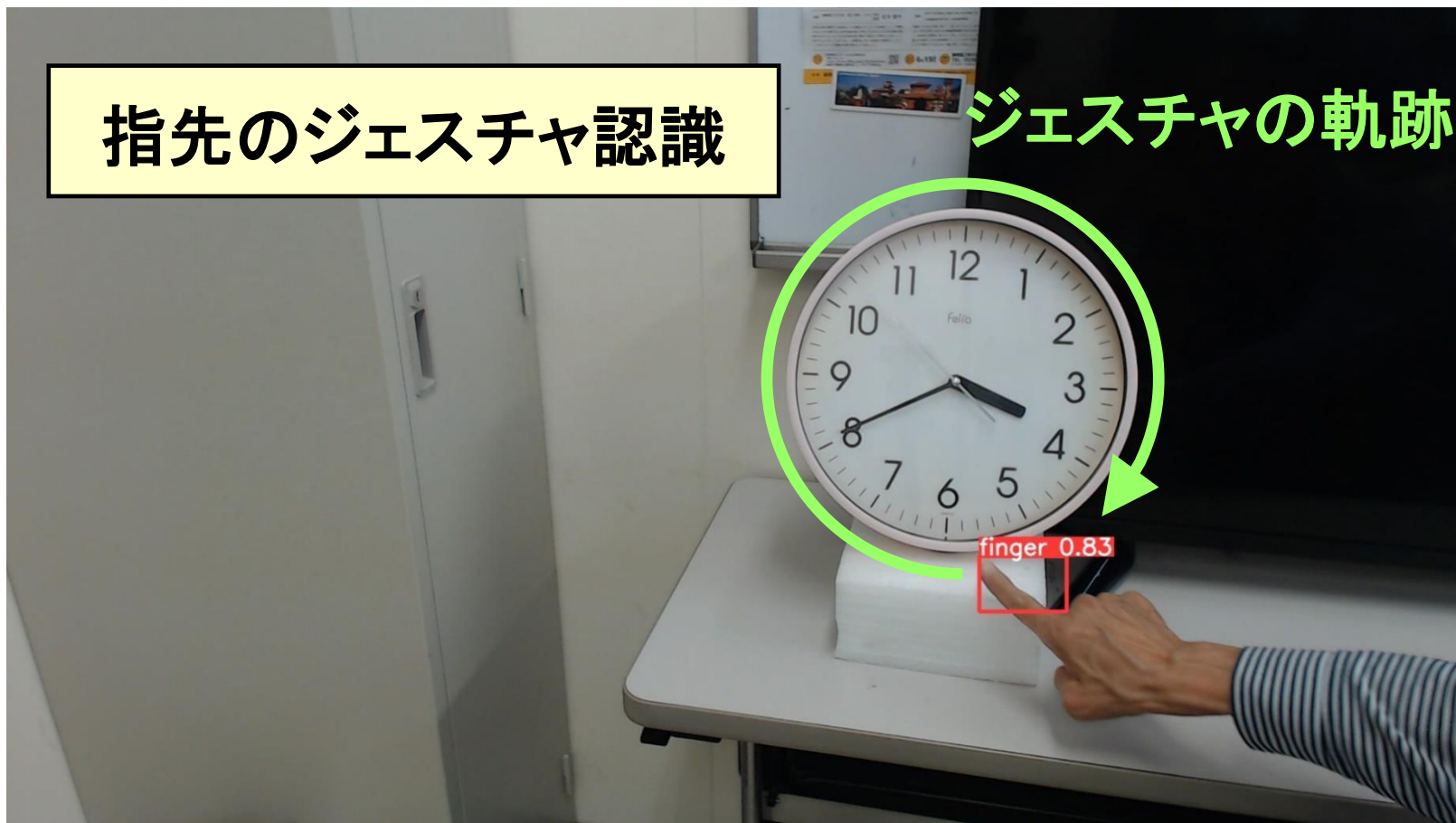
ジェスチャによる対象領域の抽出

■ **物体検出**: 画像から対象の物体を検出する技術

➤ **指先を物体検出, ジェスチャの軌跡**で対象を指定

指先のジェスチャ認識

ジェスチャの軌跡



物体検出モデル:YOLO

■ You Look Only Once (YOLO)

- ✓ 物体検出と物体認識を同時にリアルタイムで実行
- ✓ プログラム, 多数の学習データが公開
 - 特定の対象の物体検出は容易に実現

【出典】Bambach, S., et.al.: Lending A Hand: Detecting Hands and Recognizing Activities in Complex Egocentric Interactions, IEEE Int. Conf. Computer Vision (ICCV), pp. 1949–1957 (2015), <https://public.roboflow.com/object-detection/hands>.



□ EgoHands Datasetによる評価

- ✓ 訓練データ(画像数): 4,800枚
- ✓ 自身と相手の, 左右の手を検出

■ 課題

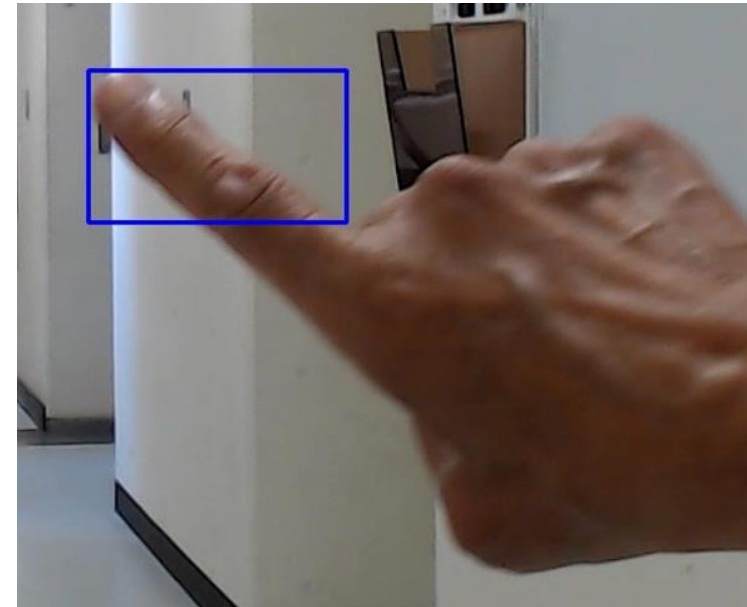
1. 手では指示精度が劣化
2. 誤検出(ノイズ)の存在

「手の検出」から「指の検出」へ: 転移学習

■ 転移学習

ある対象(手)で学習したモデルを活用して、別の対象(指)を学習させる方法.

➤ 少ない学習データで, モデルの学習が可能.



■ 訓練データ: 正解情報を付加

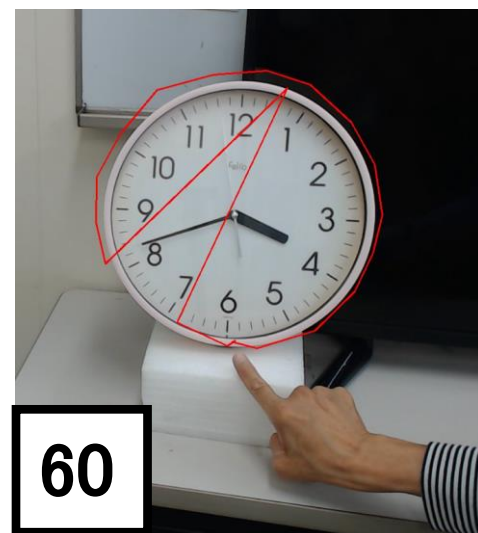
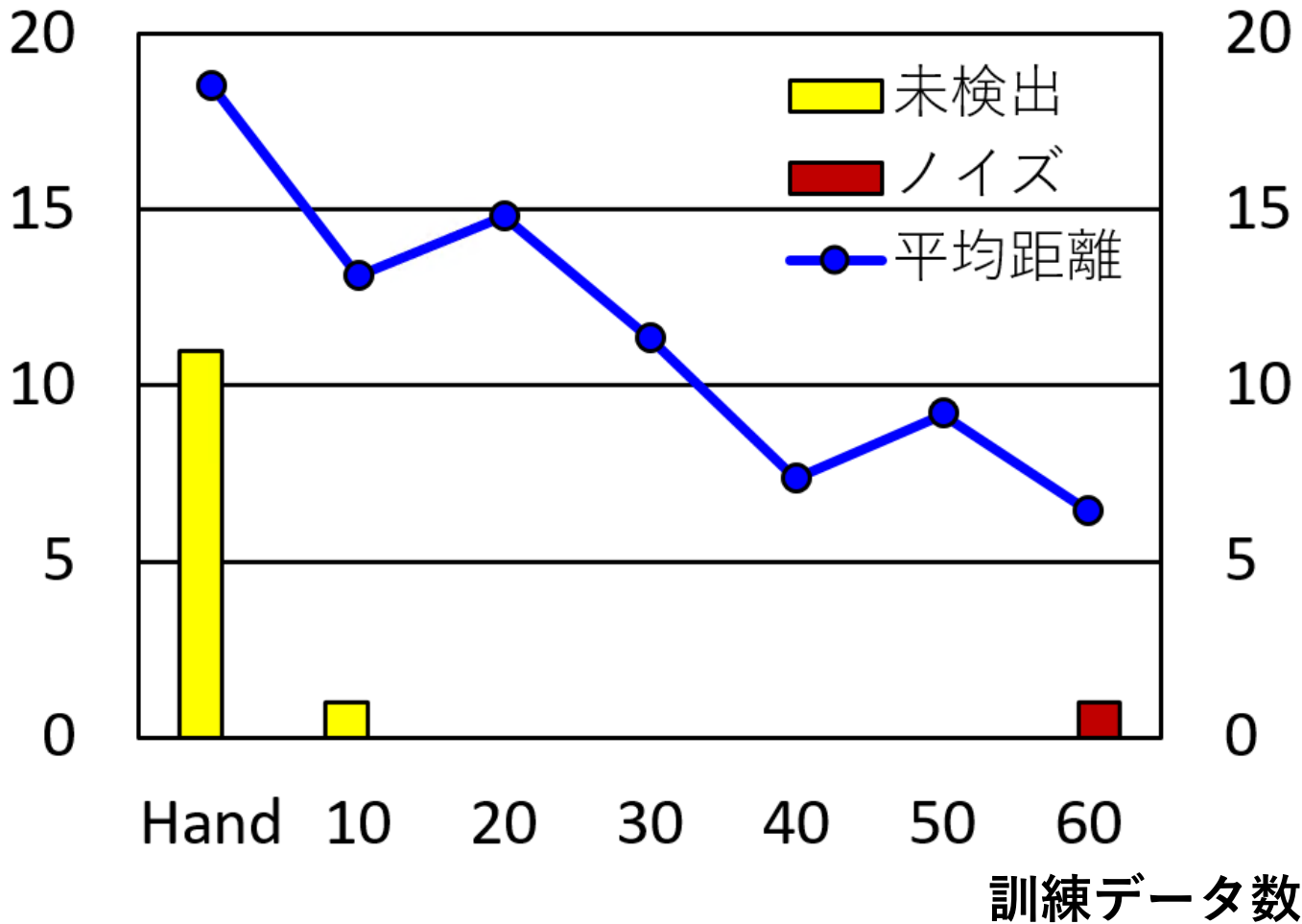
■ 拡大図(正解位置)

転移学習の精度評価 (テストデータ数: 20件)

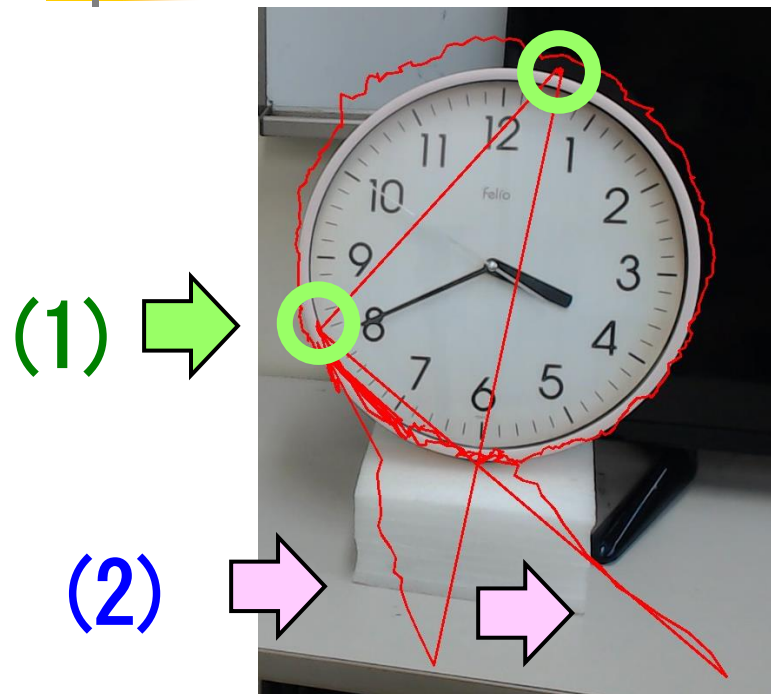
1. 40件で精度改善が収束; 2. 10件でも精度改善

平均誤差 (ピクセル)

未検出／ノイズ件数



ノイズと不要区間の除去



■課題:

(1) ノイズ (誤検出点)

(2) 不要な区間の軌跡

= 対象までの指の移動

■ 対策(提案方式):

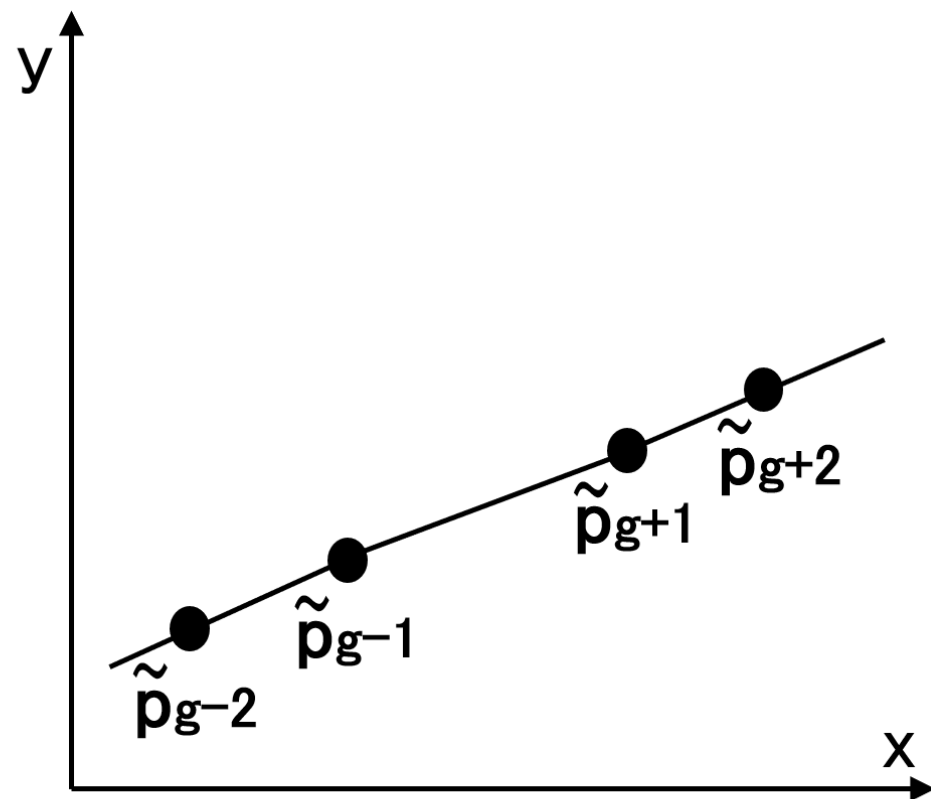
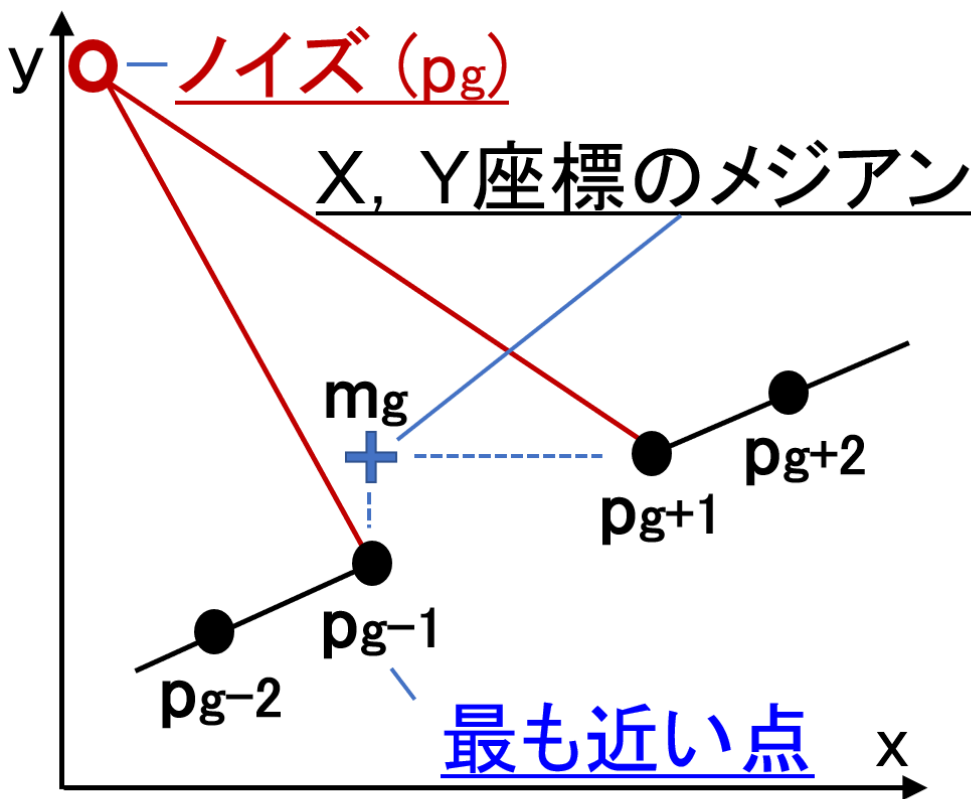
(1) **メジアン**によるノイズ除去

(2) **停止動作**によるジェスチャ区間の指示

ノイズの除去方式

(1) メジアンによるノイズ除去

- X, Y 座標のメジアンに最も近い点を選択

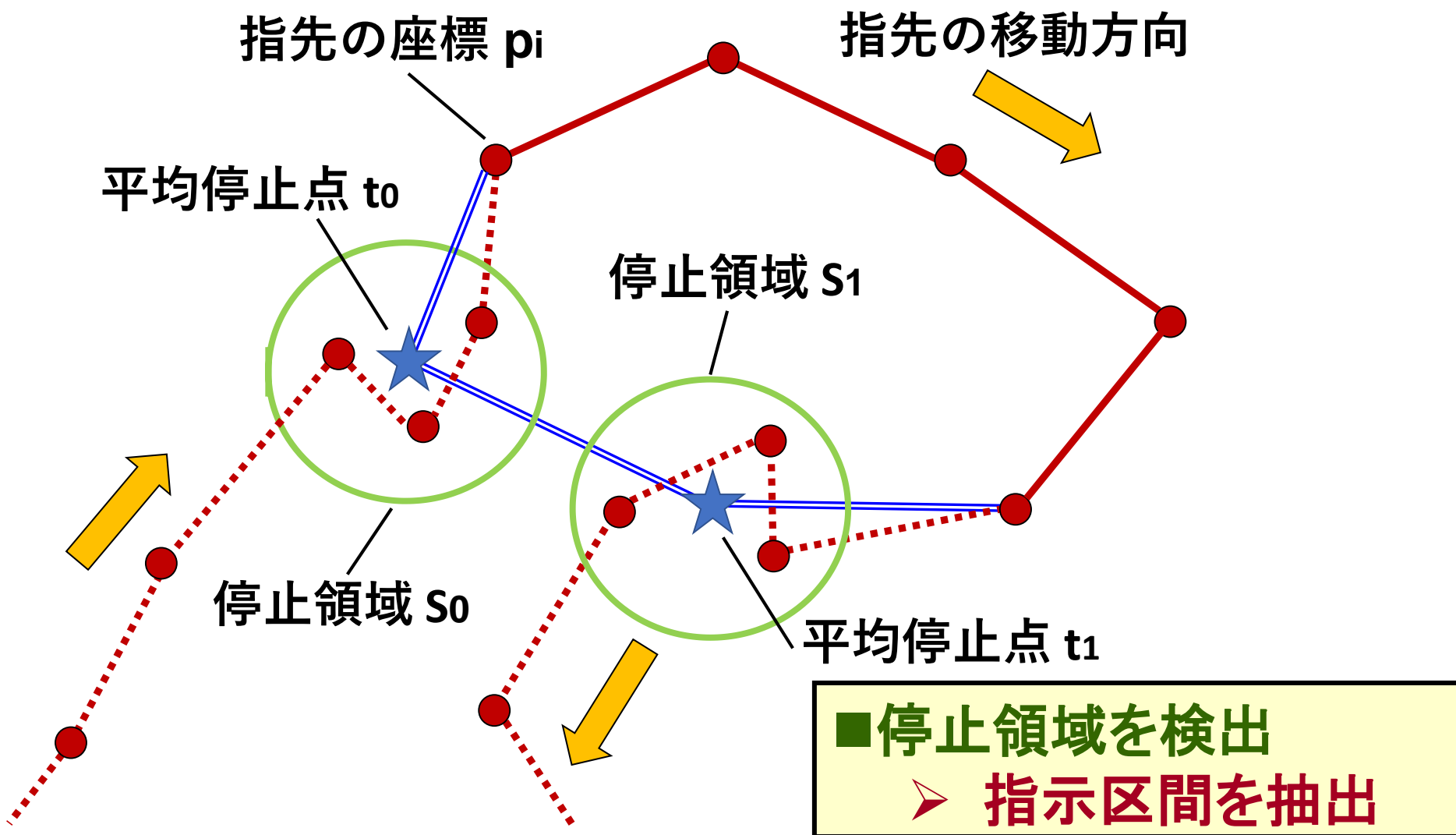


(1) ノイズを含む軌跡

(2) 補正された軌跡

不要区間の除去方式

(2) ジェスチャの停止動作で、対象区間を指定

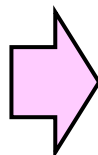


対象領域抽出の評価

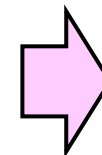
- 提案方式により、任意の形状の対象領域を抽出



(1) YOLO



(2) ノイズ除去

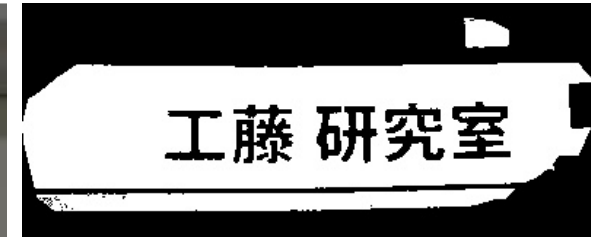
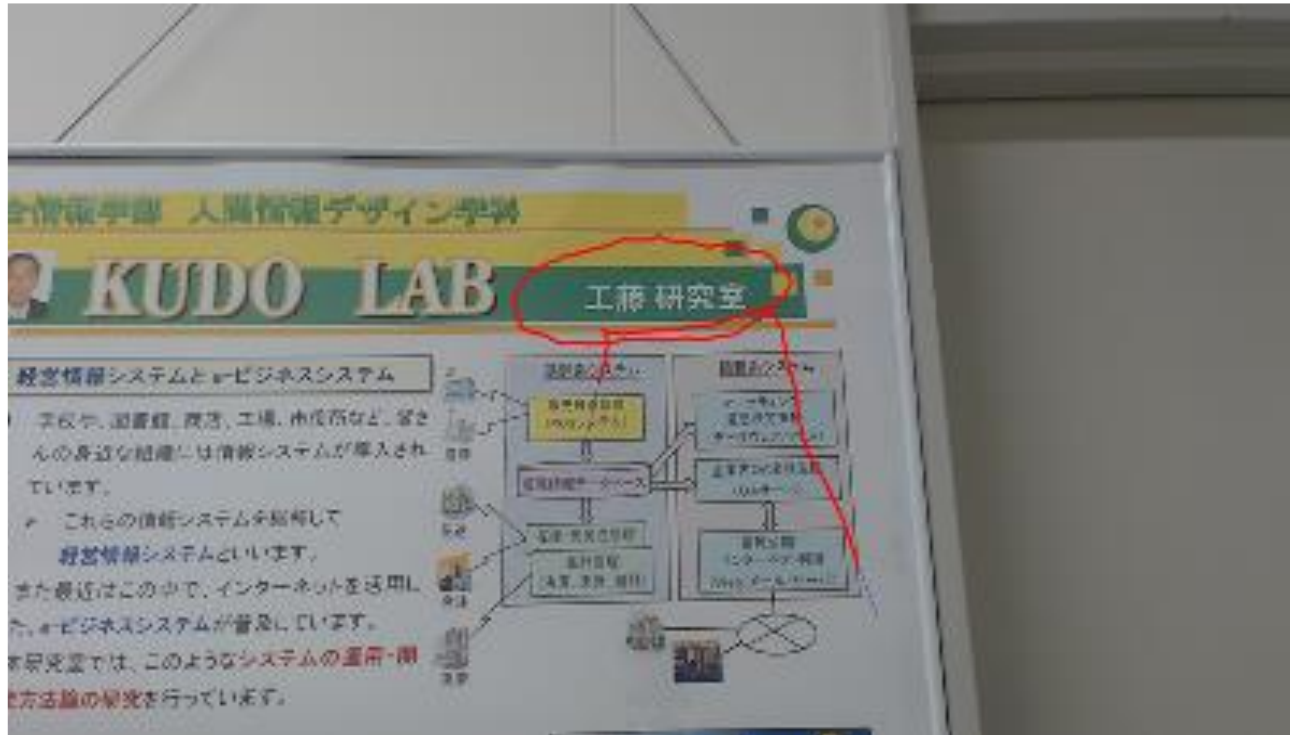


(3) 不要軌跡除去
対象領域抽出

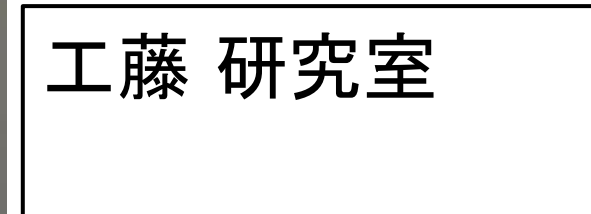
ジェスチャによる対象領域抽出の応用実験

■ 実験内容

1. 研究室紹介パネルから、**研究室名**の部分を指定
2. OCRにより研究室名を自動認識



(2) **抽出領域**を
モノクロ変換



(1) **ノイズ除去後のジェスチャの軌跡**

(3) **OCRの結果**

4. おわりに

深層学習：研究対象から当たり前前の技術へ

「犬が人を噛んでもニュースにならないが、
人が犬を噛んだらニュースになる。」

- 技術が普及し一般化すると、マスコミは取り上げない。
- 周りを見回してみると、当たり前になった多くの技術。
 - ✓ 顔認証, 音声認識, 文字認識 (AI-OCR)
 - ✓ 自然言語処理 (機械翻訳, ChatGPT), などなど。
- 業務システムの「**技術の1つ**」として**活用の時代へ**
 - 例えば, **YOLOのジェスチャ認識を応用すると...**

応用1：ピッキング経路指示書

- 多品種少量の組み立て工場では、
効率的な部品のピッキング経路の情報共有が課題。

■ 作成方法1：順次，ジェスチャで対象を指示

- ✓ 背景に，ジェスチャの軌跡（円）を描画，
- ✓ 中にピッキングの順番（撮影順序）を表示



(1) 在庫棚の順番指示



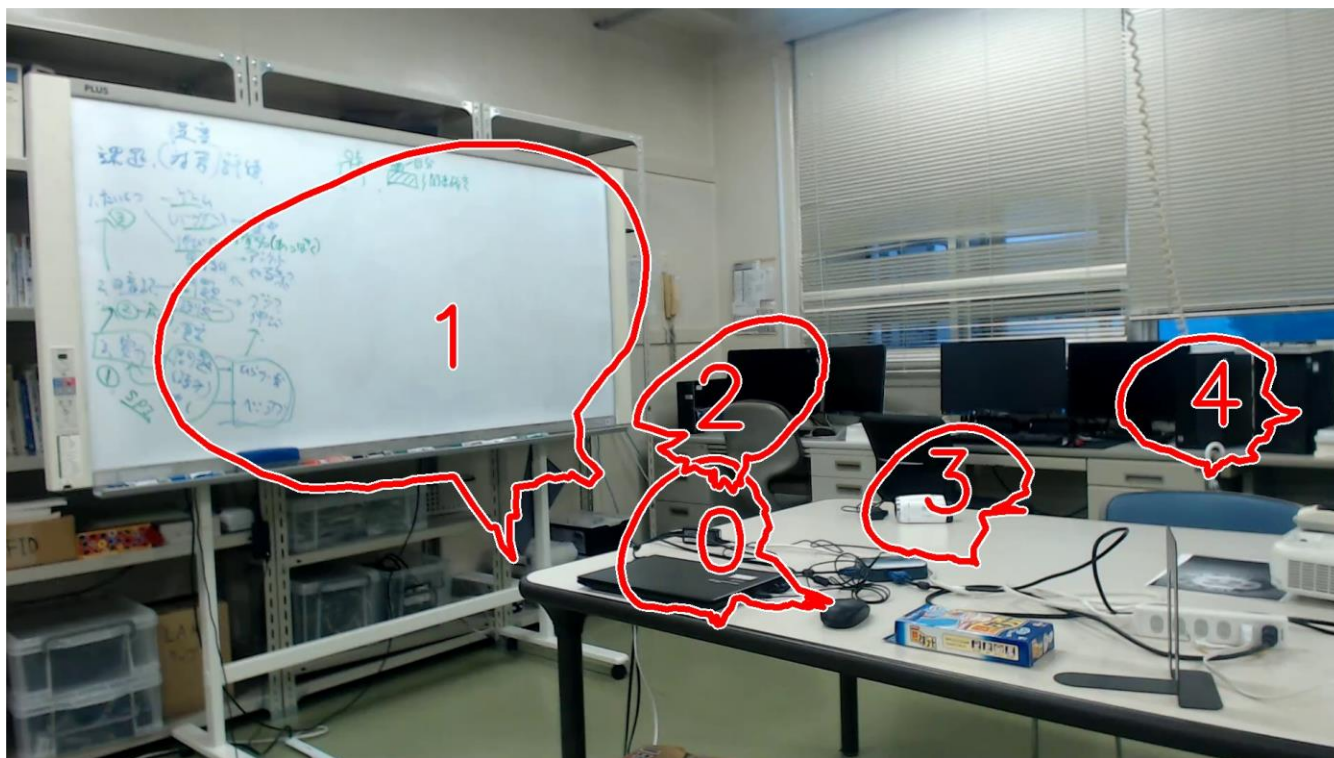
(2) 各棚のコンテナの順番指示

応用1：ピッキング経路指示書

■ 大学の研究室における実験結果

✓ 実験方法：指定順に機器をジェスチャで指示，
自動的に対象領域と通番を表示。

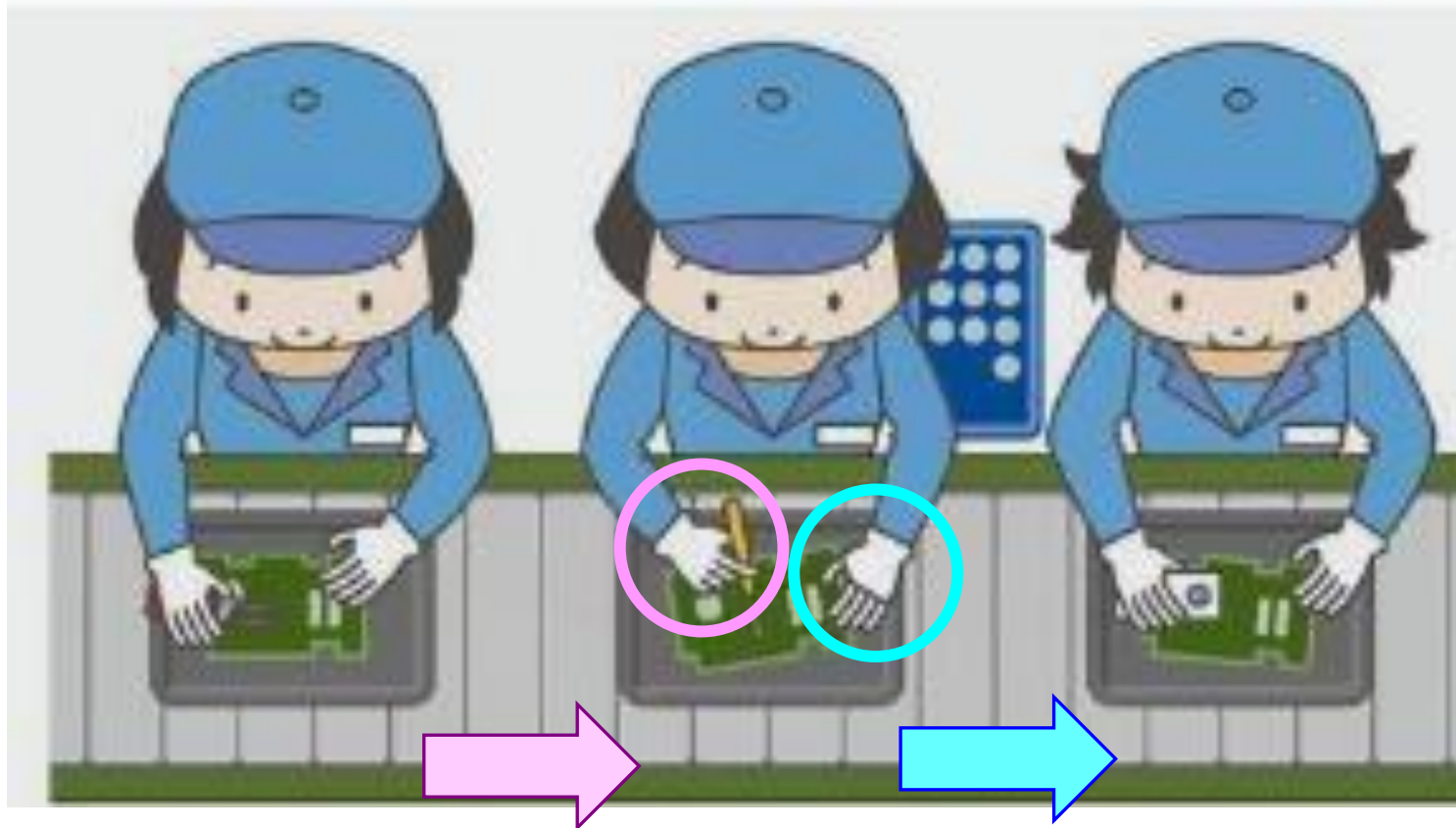
➤ 結果：厳密な精度が不要のため，効率的に指示可能。



応用 2 : 生産ラインの作業分析

■ 右手, 左手の動きを監視

- 作業者のスキル(作業時間), 出来高を把握
- **作業分析**: どの手順で時間を要しているか



■ 深層学習の技術を応用するために

➤ **メリット**, **デメリット**を理解し, 応用可否を判断.

➤ **メリット**

✓ **あいまいな対象**の領域へ応用可能

✓ **入力**(在庫画像)と**正解**(在庫数)があれば応用可能

➤ **デメリット**

✓ **大量の学習データ**が必要

✓ **近似的に推定** (厳密な判定には補完機能が必要)

深層学習の仕組みと応用

— 画像解析における事例 —

ご静聴ありがとうございました

2023年 7月 13日

静岡理工科大学 総合技術研究所

工藤 司

