

CONTENTS

第1部 行動観察について

第2部 AIを活用した調理動画解析



)第1部 行動観察について

Technology Research本部 人間科学研究所

森梓

1. これまでの行動観察

う調理動画から、ヒトやアイテムの動きを目視・手作業で分析。



時間

座標(ヒト・アイテム)

アイテム

動作

アイテムの状態

設備へのアクセス

調理台の使われ方

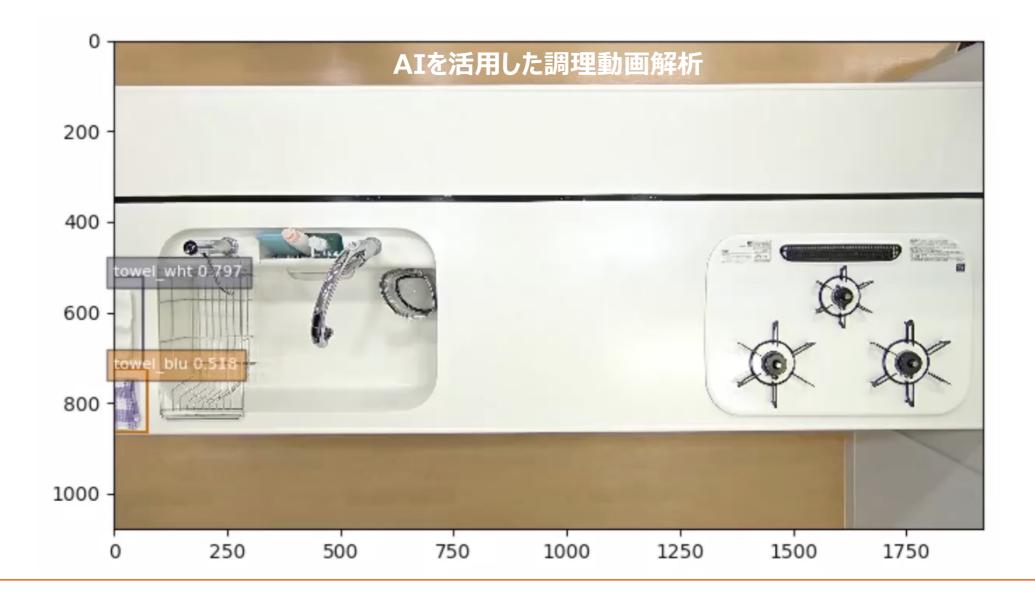
1. これまでの行動観察

- 沙 膨大な工数やバラつきによって、分析の深掘りに限界があった。
-) 自動化によって、私たちが取り組みたいことに集中できる。



自動化(工数
$$\frac{5時間}{3ヶ月} = \frac{1}{100}$$
)

2. AIを活用した行動観察





)第2部 AIを活用した調理動画解析

マーケティング本部 デジタルテクノロジーセンター

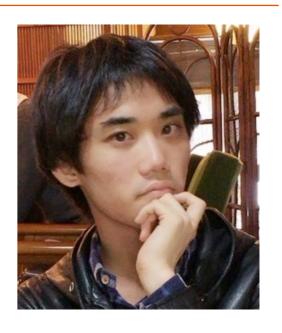
原田篤

自己紹介

原田篤

株式会社LIXIL マーケティング本部 デジタルテクノロジーセンター

Research Software Engineer



材料研究所に入社配属





気づいたら・・・



エンジニアに・・・!!!





ABEJAさんとの出会い

Deep Learning Training Course に参加







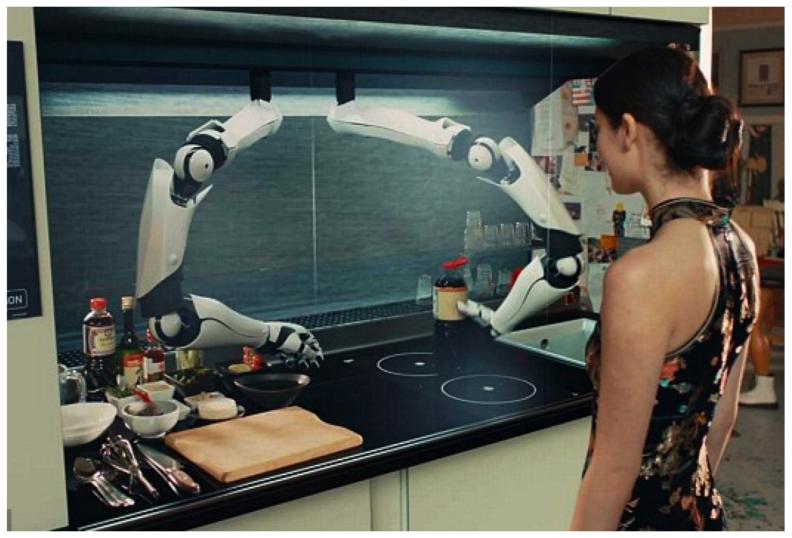
Deep Learningの基本的なところをハンズオン形式で学べる

今日語ること

LIXILがどのようにして 調理動画解析を実現しているか



キッチンとAI・・・ロボット



http://www.moley.com/



キッチンとAI・・・スマートスピーカー





キッチンとAI・・・開発をアシスト









調理行動観察に画像認識を使うメリット?

従来の取り組み



- ・人手で物の動きを解析
- 人によりばらつく
- ・深堀まで辿り着かない

ありたい姿



- ・画像認識で物の動きを追う
- ・一定の基準で解析できる
- ・考える時間がある



「人工知能の民主化」!!!

「人工知能の民主化」を実現する革新的AIプラットフォーム事業のPMを募集



株式会社ABEJA











https://www.wantedly.com/projects/188020



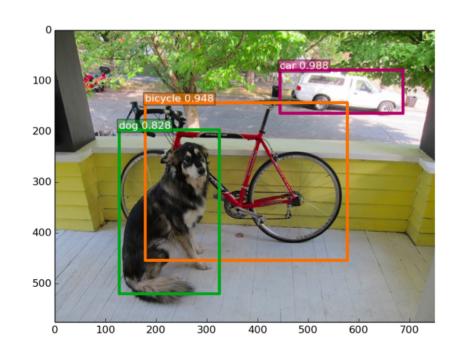
「人工知能の民主化」の流れに乗りたい 自分たちでやってみよう!!!

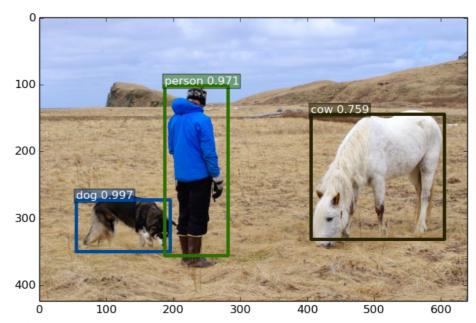
手法検討 編



画像認識だとR-CNN系が使えそう

2017年年始頃の出来事



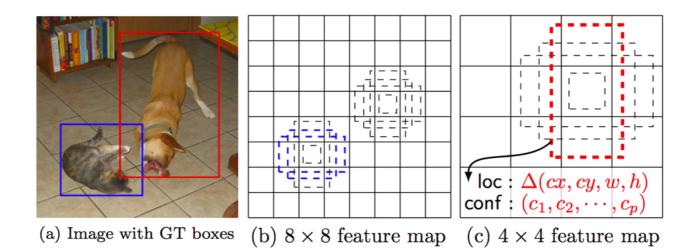


- ・物体の位置がわかる
- ・物体の種類がわかる
- ・光とか色に結構ロバストっぽい





その中でもSSDが早くて精度が良さそう



System	VOC2007 test mAP	FPS (Titan X)	Number of Boxes	Input resolution
Faster R-CNN (VGG16)	73.2	7	~6000	~1000 x 600
YOLO (customized)	63.4	45	98	448 x 448
SSD300* (VGG16)	77.2	46	8732	300 x 300
SSD512* (VGG16)	79.8	19	24564	512 x 512

https://github.com/weiliu89/caffe/tree/ssd





しかもGithubにすぐに使えそうなコードがある! Let's do it!

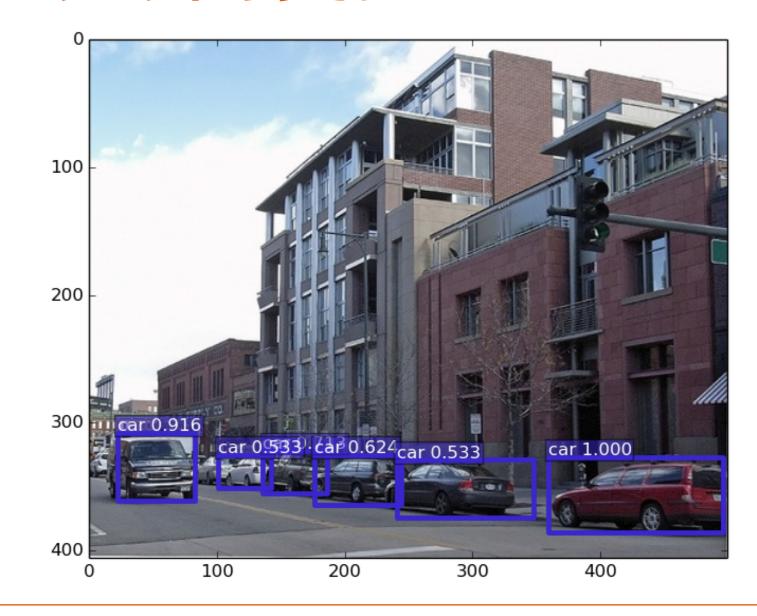




データセット準備 編



一般物体は割とデータセットありそう





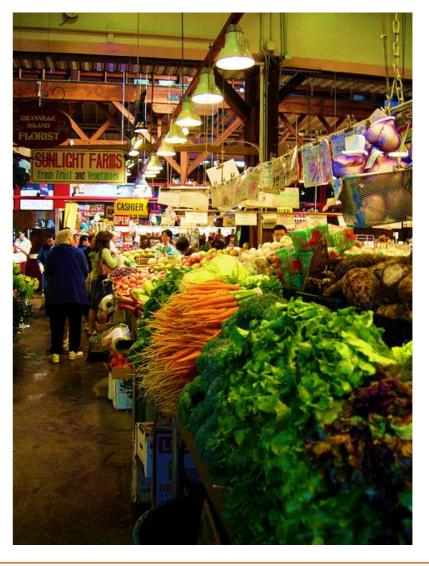


食材とかは?

食卓に並ぶ野菜



市場



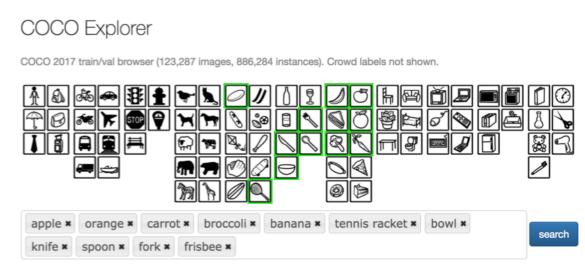


調理中の食材とか調理道具のデータセットがない

一応入ってる公開のデータセットもあるが・・・



http://mscoco.org/





- ・日本特有の物体が少ない
- ・調理途中の画像が少ない
- ・そもそも種類が全然足りない



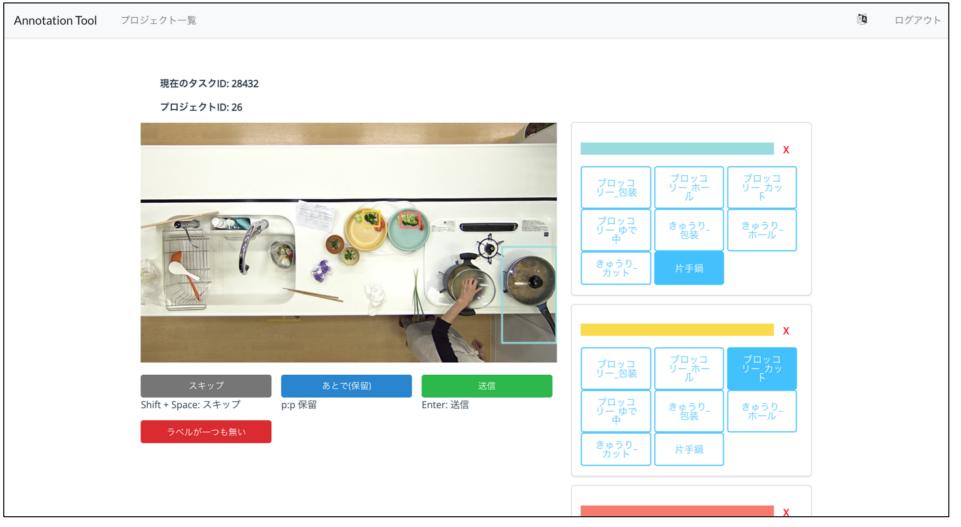
画像認識を自社で活用しようとしてわかったこと

一般物体に適用できるデータセットはたくさんある

・自社特有の課題に適用したい場合にはデータセットが必要

1から作る覚悟

ABEJAプラットフォームでアノテーションできる







そのままBPOとして業務委託も可能



運用環境構築 編



1から構築した例

非エンジニア 動画解析システム エンジニア ユーザー docker S3に認識モデルや Lambdaから パラメータ ミドル環境 ユーザーが 撮影データを格納 EC2を立ち上げ ロード ロード アップロード 学習させる場合は 機械学習エンジニアがサポート ケースごとにEC2インスタンスを生成、使い捨て Slackに通知 LambdaからWebhook通知 +ab|eau Tableauで可視化 解析結果や認識モデルはS3に格納



ABEJAプラットフォームを使った場合

非エンジニア 動画解析システム エンジニア ユーザー 取得 蓄積 1. 多数のIoTデバイスや外部シス 2. 取得した大量データをセキュア テムからAPI経由で簡単にデ に蓄積 ータを取得 3. データのバリデーション ユーザーが 4.アノテーション アップロード 学習させる場合は **ABEJA Platformで** 機械学習エンジニアがサポート AIの継続的 インテグレーション を実現 推論・再学習 学習 Slackに通知 9. デプロイされたモデルによる推論 5. Deep Learning • Machine 10.新たな教師データやチューニング Learningのモデル構築 されたモデルを用いた再学習 6. 教師データを用いた学習済み モデルの生成 デプロイ 7. 生成された学習済みモデルの 評価 8. 推論を行うクラウドまたはエッジ Tableauで可視化 環境へのモデルのデプロイ



ABEJAプラットフォームの優位性

・運用を見据えた学習・推論環境が手軽に構築できる

・特にこれから環境を構築する企業には適している

・細かい部分の保守管理を気にしなくて良い



まとめ



私たちの課題

・行動観察を効率的に進めるために画像認識を適用したい

・特有の事例にはデータセットを1から作る必要がある

・学習・推論共に現場で運用できる環境が必要

この取り組みの結果

データセットを自分たちで作ることができた

・運用を見据えた動画解析システムが構築できた

・元々3ヶ月かかっていた作業がAIを使って1/100になった



