

揚げ物調理支援のための画像による油中食材の追跡

工藤 央一郎[†] 平井 重行[‡]京都産業大学大学院 先端情報学研究科[†]京都産業大学 情報理工学部[‡]

1. はじめに

飲食業界の人手不足で調理場担当者は多忙を極め、食材の下準備や補充を並行して行い、調理に専念できないことが多い。唐揚げ等の揚げ物には食中毒等の防止のため最低限の調理時間が設けられるが、注文毎にフライヤーへ逐次投入する現場も多い。キッチンタイマーでの個別管理には限界があり、揚げ物調理は担当者の目視で管理する店も多く、少ない人数での調理作業の効率化が課題となっている。

本研究は、このような調理現場での揚げ物管理を自動化し、調理場全体の効率向上を目指す。食材の揚がり具合に応じ、取り上げる食材情報の人への提示や、ロボットで自動的に取り出すことを想定している。我々は現在、この自動化のために、フライヤーの油中での個々の食材を画像処理で判別および追跡し、逐次投入される食材の揚がる時間を個別管理する手法について取り組んでいる。本稿は、深層学習による油中食材の領域判別とその追跡手法について述べる。

2. 揚げ物調理のデータ収集環境

調理動画の収録は、家庭用キッチンのコンロにて、業務用天ぷら鍋を用いて行った。業務用フライヤーではないものの、油面の面積や一定



図 1. キッチンでのデータ収集の様子

の深さを確保し、飲食店での揚げ物調理現場に近い環境とした。画像はコンロ上の換気扇にカメラを取り付け、調理中の様子を直上付近から撮影した(図1参照)。

3. 深層学習による食材領域の抽出

油中にある食材領域の抽出について、食材の個数判別と領域分割の二段階の処理で行う。

食材の個数判別は、ResNet50[1]を用いた画像分類モデルで処理する。この学習には、食材個数 0~5 個の 6 種類、解像度 1024×768 の画像を計 60000 枚用い、エポック数 100 で学習した。

領域分割には、Mask-RCNN (Mask Region-with Convolution Neural Network) [2]を用いた。Mask-RCNN は物体検出とセグメンテーションを同時に行うもので、これにより個別食材の領域を抽出する。学習には個数判別と同じ画像を使用した。こちらは食材個数 1~5 個それぞれのデータセット (5 種類、各 10000 枚) とし、意図的に偏りをもたせた個数毎の学習モデルを作成した。ネットワークには ResNet101[1]を用い、エポック数 100 で 5 種類の各モデルを学習した。なお、この領域分割の 5 種類の学習モデルは、前段での食材個数判別の結果 (個数) を踏まえ、後段の処理として用いる。

これら学習済みモデルに対し、未学習の調理画像を入力した際の出力画像を図2に示す。ここでは、5 個の各食材がそれぞれ異なる色で塗られた領域として抽出されている。

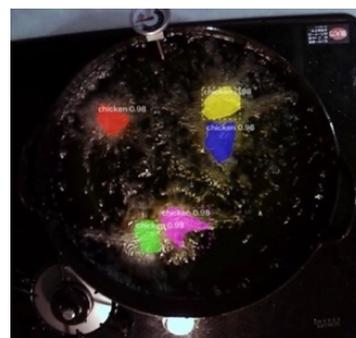


図 2. Mask-RCNN による処理結果

Attempts to Track Deep-Frying Foods in A Fryer by Image Processing for Cooking Support

[†]Eiichiro Kudo, Division of Frontier Informatics, Graduate School of Kyoto Sangyo University

[‡]Shigeyuki Hirai, Faculty of Information Science and Engineering, Kyoto Sangyo University

4. 面積値を用いた判別個数の尤度処理

食材4個の調理動画を時系列画像として入力した場合の判別個数の結果を図3に示す。5個と3個の誤りが頻繁に出力されており、調理し始めの左側で特に多い。そこで、正解(4個)と不正解(3ないし5個)の画像において、抽出された食材領域の面積の違いを確認した。

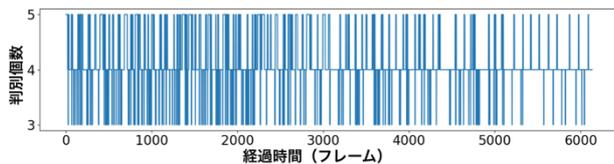


図3. 調理動画(正解個数4個)の個数判別結果

図4は、食材個数4の正解画像で抽出された食材領域の画素面積の分布(左側)と、それを正規化したもの(右側)である。食材個数毎にこれらを作成し、それぞれを確率密度関数として尤度計算に用いる。この関数に正解個数のモデル(Class4)と不正解の個数のモデル(Class3, Class5)で領域抽出を行った場合の面積をそれぞれ入力し、尤度として比較したものを図5に示す。この図より、正解個数のモデルで抽出された領域は尤度が高く出やすいことがわかる。

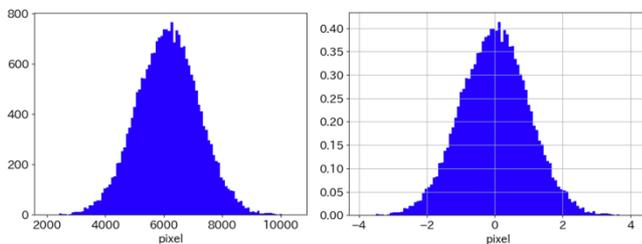


図4. 食材4個の調理画像の面積分布

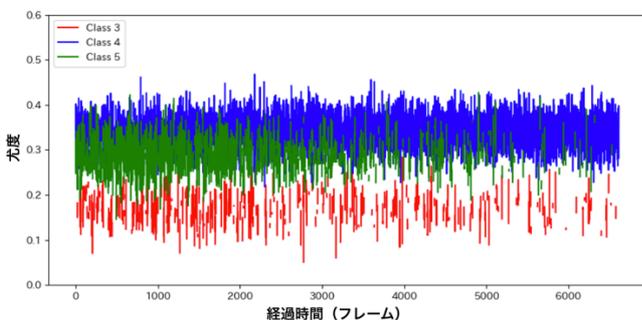


図5. 各学習モデルでの出力面積の平均

これらより、画像内の食材領域の平均面積について、前段の処理結果の判別個数の±1個のモデル(計3個のモデル)で常に求め、尤度が最も高い個数の個別学習モデルの結果を採用するこ

ととした。図3で示した調理動画にこの処理を施した結果を図6に示す。これにより動画の大部分で正解の判別個数を出力できるようになった。

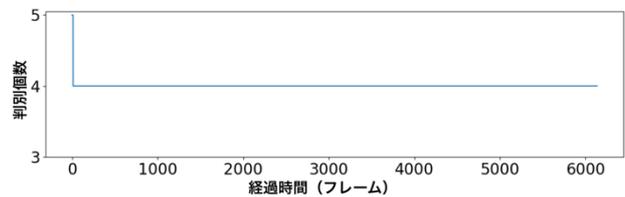


図6. 重み付け処理後の判別個数

5. 食材の追跡処理

後段処理で抽出された食材領域を個別追跡するために各領域にIDを割り当てる。ここでは、フレーム毎に各抽出領域の重心を計算し、前後フレーム間の各重心位置の相対変化量の総量が最も小さい組み合わせを求める。これにより、前後フレーム間での食材領域のID割り当てを引き継ぎ、油中で移動する食材を追跡する。その他、散発的に起こる未検出画像等に対する処理を施すことで、食材個数が変化しない場面では、食材追跡が安定して行えるようになった。

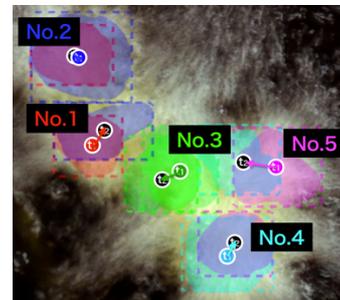


図7. フレーム間のID割り振り

6. まとめ

本研究では、揚げ物調理画像に対し深層学習を用いて個別食材追跡の処理を実装した。今後は個数判定の精度向上を行いつつ、食材個数が変化する場面においても追跡をロバストに行う処理を実現し、揚げ物調理支援システムとして開発していく予定である。

参考文献

- [1] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, "Deep Residual Learning for Image Recognition" CoRRabs/1512.03385 (2015).
- [2] Kaiming He, Georgia Gkioxari, Piotr Dollár, Ross Girshick, "Mask R-CNN" CoRRabs/1703.06870 (2017).