6ZE-01

物体検知と音声認識を用いた非接触調理支援システムの研究

竹内　龍太郎†　　神戸　英利‡
東京電機大学大学院理工学研究科†‡

1. 研究背景・課題

調理をする際には,携帯情報端末機器(以下スマホと示す)でレシピサイトを参照することが一般的になっている [1].近年,レシピサイトは文字形式のレシピだけでなく,一連の調理方法の概要や各調理工程調を簡潔に示す料理動画も掲載している.このような料理動画は文字形式のレシピに対し,調理方法を理解しやすく,ユーザに対し視覚的かつ効率的な支援を可能にしている.一方,料理レシピを閲覧しているスマホには,食中毒の原因にもなる黄色ブドウ球菌を含む一般細菌が多く常在している [2].そのため,調理中のスマホ操作は衛生的ではなく,ノロウイルスや食中毒などの感染リスクを高める.また,料理レシピや料理動画を参考にする場合,料理に慣れていないと何度も見直すことになるため,直接スマホを操作する機会が多くなる.調理中は食材を扱っているため,手が汚れている状況や濡れていることがある.このような手指の状態で料理レシピを表示するスマホを操作することは,不衛生であり,調理をする環境として不適切である.

2. 研究目的

本研究では, 料理レシピを閲覧している携帯情報端末機器等の機能(カメラ・マイク等)のみを用いて, 非接触での調理工程認識や該当料理動画の提示及び非接触制御を行い, 端末への接触回数減少や調理時間の短縮可能にし, 衛生面に配慮したユーザ中心の効率的な調理支援を目的とする.

3. 提案手法

3.1 システム概要

システム概要図を図1に示す.

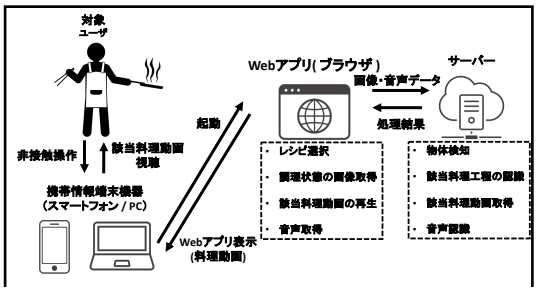

図 1 システム概要図

Research on non-contact cooking support system using object detection and voice recognition.
Ryutaro Takeuchi†, Hidetoshi Kambe†
†Graduate School of Science and Engineering,
Tokyo Denki University

はじめにユーザが Web アプリを起動し,料理一覧から料理選択を行う.次に,Web アプリを閲覧しているスマホの外カメラを用いて調理状態のリアルタイム物体検知を行い,食材と調理器具をリストに格納する.検知結果をレシピデータの調理手順と照合し,該当調理工程を認識した場合,認識した調理工程をポップアップにて表示する.その後,認識された調理工程に該当する料理動画をスマホのディスプレイに表示し,調理が完了するまでリピート再生を行う.なお,この料理動画の再生速度の変更,指定時間のスキップ送り,再生,一時停止等の料理動画の制御はマイクを用いた音声認識で非接触制御を行う.生成された該当料理動画を模倣し,調理を進める.調理工程認識及び料理動画の模倣を全調理工程が終了するまで繰り返し行う.

3.2 食材・調理器具検知

スマホの外カメラで調理状態を撮影した画像をサーバに送信し,リアルタイムにサーバ上で食材・調理器具の検知を行う.その後,検知結果とレシピデータを照合し,調理手順を認識する.そのために,リアルタイムに様々なクラスを検知可能な一般物体検知アルゴリズムである YOLO を用いる.YOLO は候補領域検出とクラス分類の両方を一度の CNN 演算で行うため高速であり,尚且つ高い認識精度である.よって,本研究では物体検知アルゴリズムとして YOLO を用いる.

3.3 料理動画再生

ユーザへの料理レシピ提示手法として効率的かつ視覚的に理解しやすい料理動画を用いた料理レシピの提示を行う.該当料理レシピを全再生し続けるのではなく,認識された該当調理工程のみの料理動画再生を行う.なお,この料理動画については該当している調理工程が終了するまでリピート再生を行う.

3.4 音声認識

ユーザが効率的に調理を行えるように Web アプリ制御や料理動画制御をスマホのマイクから入力された音声を用いて,音声認識を行う.この音声認識は,リアルタイムで音声をテキストへ変換し,文字列として取得する.Web アプリの画面遷移を行う指定コマンドや再生,一時停止,動画の再生速度変更,指定時間のスキップ送り・戻しなどの動画制御を行う指定コマンドと取得した音声文字列が一致した場合,そのコマンドの制御を行う.

4. 実装

本研究では,ユーザ評価を基に Web アプリを開発・改善した.スマホのカメラを用いてリアルタイム調理状態を取得し,そのデータを基に食材・調理器具の検知を行い,調理工程を認識する.そして,該当調理工程の料理動画をリピート再生し,スマホのマイクを用いた音声認識を行うことで,料理動画の非接触制御を行うシステムを実装した.本システムでは開発言語として Python, JavaScript を用いて Web アプリ開発を行い,ユーザーに対し直感的かつ動的な調理支援を提供するために React, Flask, SQLite, Web Speech API, YOLO, WebRTC 等の技術を組み合わせた.

ユーザは Web アプリから料理レシピを選択し,料理開始ボタンをクリックすると React ベースの非接触調理支援アプリに遷移する.最初の調理工程の必要食材と器具が表示され,YOLO と WebRTC を使ったリアルタイム物体検知を行う.必要アイテムが検出されると,ポップアップが表示され,「次へ」と音声入力すると,該当する調理工程の動画がリピート再生される.その他にも動画の再生,停止,速度変更などの動画制御を音声認識で行うことが可能である.各工程が完了した場合,「完了」と音声入力し,次の調理工程の検知画面に戻る.全工程終了後はレシピ一覧画面に自動的に遷移する.これらのプロセスを図 2 に示す.

図 2 アプリケーションの画面遷移フロー

5. 評価

本研究のシステムに対して目的と照合し,評価を得た.有用性評価では,スマホを直接制御して調理する場合,本システム(ver.1)を用いて調理をする場合,ユーザ評価を元に改良した本システム(ver.2)を用いて調理をする場合を比較し,スマホに接触する回数と調理時間の変化を観察し,評価を行った.アンケート対象者は男女含めた 20 代の全 10 名である.

表 1 端末への接触回数と調理時間の変化

	端末を直接制御して調理		本システムver.1を用いて調理		本システムver.2を用いて調理	
名前	端末接触回数	調理時間	端末接触回数	調理時間	端末接触回数	調理時間
A	30	22分12秒	0	15分4秒	0	12分32秒
B	34	20分38秒	4	22分37秒	0	14分11秒
C	41	25分42秒	0	19分48秒	0	16分22秒
D	40	21分48秒	0	18分52秒	0	15分48秒
E	50	20分38秒	0	18分52秒	0	14分50秒
F	27	25分12秒	0	17分29秒	0	14分09秒
G	38	33分40秒	0	18分26秒	0	16分09秒
H	32	24分24秒	5	18分52秒	0	15分45秒
I	37	18分17秒	0	14分30秒	0	11分28秒
J	96	12分49秒	0	9分37秒	0	9分17秒
平均	42.5	22分32秒	0.9	17分24秒	0	14分03秒

表 1より端末への接触回数は,スマホを直接制御して調理する場合の平均が 42.5 回であり,本システム(ver.1) を用いて調理をする場合の平均が 0.9 回であり,,本システム(ver.2)を用いて調理をする場合の平均が 0 回である.スマホ接触回数は全員大幅に減少している.調理時間は,スマホを直接制御して調理する場合の平均が 22 分 32 秒であり,本システム(ver.1)を用いて調理をする場合の平均が 17 分 24 秒であり,本システム(ver.2)を用いて調理をする場合の平均が 14 分 03 秒である.スマホを直接操作する調理時間と本システム(ver.2)を用いる調理時間の差は平均して 8 分 29 秒短縮されている.

全ユーザのスマホへの接触回数が大幅に減少していることから本システムは衛生面への配慮があると言える.また,本システムを用いることで調理時間が短縮されることから,本システムの有用性および効率的な調理が可能になることがわかる.

5. 考察

本研究により,衛生面に配慮したユーザ中心の調理支援システムを構築し,調理状態を撮影した画像より YOLO を用いたリアルタイム物体検知を行い,検知結果から調理工程の認識,該当調理工程の料理動画の音声認識制御をすることが可能になり,スマホへの接触回数の減少及び調理時間の短縮が見られた.しかし,食材・調理器具検出においては,未検知・誤検知があるため改善が必要である.これらの対策として,未検知画像を学習データとして用いたモデルの作成の必要があると考える.物体検知を用いた非接触の調理工程認識における問題点の改善を行うことで,より調理時間の短縮が期待できる.

6. まとめと今後の展望

本研究では,衛生面に配慮したユーザ中心の効率的な調理支援を目的とし,調理状態を撮影した画像より YOLO を用いたリアルタイム物体検知を行い,検知結果から調理工程の認識,該当調理工程の料理動画の音声認識制御を行うことができた.しかし,物体検知および音声認識の処理が滞ることが多くあり,精度や処理速度の向上が必要である.

参考文献

[1] スリーエム株式会社, “料理を参考にするのは文字派と動画派どっちが多い？｜スリーエム株式会社のプレスリリース,” PRTIMES, [オンライン]. Available: https://prtimes.jp/main/html/rd/p/000000011.000009321.html. [アクセス日: 11 4 2022].

[2] 宇. 賀. 山. 美. 森岡 郁晴, “タッチパネルを有する機器の細菌汚染状況と清掃状況および汚染意識,” 日本衛生学会, 2015. [オンライン]. Available: https://www.jstage.jst.go.jp/article/jjh/70/3/70_242/_pdf. [アクセス日: 11 12 2022].