

パターン認識を用いた家電製品からの情報収集に関する研究

2016SE013 服部華之

指導教員：沢田篤史

1 はじめに

近年スマート家電や製品の自動化が進んでいる、スマート家電など家電の連携自体は古くからアイデアがあったが、ネットワーク接続の方法やコストがネックとなっていた。しかしスマートフォンの普及で状況は変わった。スマート冷蔵庫はインターネットに接続することで例えばレシピ検索などをすることが出来る。一方で、中にある食材の種類、個数を把握して調理や購買行動に活かすことは十分に出来ていない。スマート冷蔵庫では製品の在庫管理を冷蔵庫内のカメラの画像から確認して把握することしかできないという問題がある。

本研究の目標はスマート冷蔵庫から家庭にある食材状況を正確に把握しリアルタイムで購買行動に反映させることを可能にするシステムを構築することである。本研究ではその一例として冷蔵庫の中のペットボトルの個数、種類を正確に把握する技術の確立を目指す。CNN など領域抽出である既存の技術を組み合わせ、その実現を試みる。

本研究で提案する情報収集システムが実現すると、冷蔵庫の中にある食品（ペットボトル）の種類や個数を正確に把握出来るようになる。そのデータを使用しリアルタイムでの把握が可能にすることができる。画像処理を用いて情報収集を行う方式を採用したことで、特徴量の抽出方法を変えることが出来るので他の食品、食材にも発展させることが出来ると期待できる。本研究では、ペットボトルの画像認識をし、そのペットボトルが何本あるのか、何の種類であるのかを把握するために CNN の構造を検討し、プログラミング言語 python を利用してシステムを構築する。そして得られた在庫状況をユーザに知らせるシステムを構築するために、Apache と PHP スクリプトを用いた Web サービスを作成する。

2 家電からの情報抽出に関する課題

CNN を使用した画像認識は今日の生活の中で多種多様な用途に使われている [1]。近年、家電がインターネットと繋がることで様々な使用方法が飛躍的に増加した。スマートフォンと連携することが出来る IoT 家電の出現など身の周りの生活を手助けする家電も増えてきた。しかしながら家電製品からの情報をもとにしてその得られた情報をリアルタイムで使用し生かす技術は十分に実現されていない。そこで家電製品から様々な情報をリアルタイムで収集した上で、それを生かすことが出来るようにするという要求が生じている。

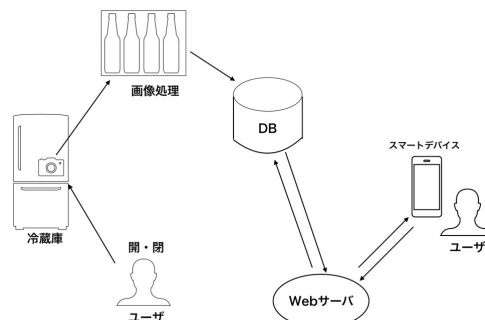


図 1 全体像

3 画像認識を用いた家電からの情報抽出

全体像の図を図 1 に示す。冷蔵庫の中のカメラで冷蔵庫内のペットボトルの写真を撮影し、取得した画像の処理をさせる。画像処理をさせ、ペットボトルの個数、種類を在庫データベース（以降、DB）に格納する。ユーザがスマートデバイスで web サーバを通じて DB にアクセスし、在庫状況を取得する。外部使用者が冷蔵庫を開けて閉めた際は、閉めたら撮影を開始し、同じ処理をさせ、変化があればユーザに通知する。

画像認識機械は事前に収集された「実例」から対象に関する知識を「学習」する必要があるので物体を認識するには膨大なデータが必要である。そして得られた情報うまく処理させるには CNN のモデル形成が重要である。中山英樹らの研究で CNN の技術が使われている [1]。そこで CNN の説明を以下に述べる。

3.1 畳み込みニューラルネットワーク (CNN) を用いたペットボトルの在庫状況抽出

以下に CNN モデルを示す。以下が CNN モデル図である。モデル構造は 1. 畳み込み層 [5x5] 2. 畳み込み層 [3x3] 3. プーリング層 [2x2] 4. 畳み込み層 [3x3] 5. プーリング層 [2x2] 6. 全結合層 7. Softmax Regression 層からなる。図 3.2 では畳み込み層を順に conv1, conv2, conv3 と示している。プーリング層を順に pool1, pool2 と示している。全結合層を順に fn1, fn2 と示している。まず conv1 で物体の線分要素を抽出する。conv2 でペットボトルの形状などの特徴を抽出する。conv3 で要素同士の繋がりを抽出する。また、conv1, conv2, conv3 の後にそれぞれ maxpooling を行う。maxpooling を行うことで、データの扱いをしやすくし、畳み込み層で抽出した特徴を際立たせる。畳み込み fn1 で 1 次元配列に変換し、fn 2 で 11 のカテゴリーに分類する。fn2 の出力に softmax 関数を用い、そのカテゴリーに分類される確立を出力する。maxpooling

層のフィルタを 2×2 にした pooling 層では、特徴を際立たせてデータを扱いやすくする。pooling のフィルタサイズを大きくしすぎると、粗い特徴になってしまうことと、画像認識の cnn の pooling 層は 2×2 が推奨されているのと、画像サイズが小さくなりすぎてしまうので 2×2 を設定する。最適化手法はニューラルネットワークに用いられるものとして代表的な Adam を用いる。

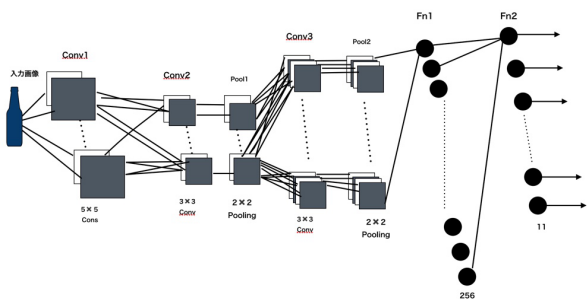


図 2 CNN モデル図

3.2 web サーバを用いた通信システム構築

ユーザが冷蔵庫の中の在庫状況確認をする場合 controller に冷蔵庫内のライトに ON 通知を送る。カメラに起動依頼の指令をだす。カメラで撮影した後、controller へ処理終了通知をし controller がライトへ OFF 通知をする。画像認識部から在庫 DB に在庫状況のデータ転送を行なう。在庫 DB から情報提供部へ処理結果を渡し、その処理結果を情報提供部がユーザへ表示をする。

外部操作者が操作した場合の説明をする。まず外部操作者が冷蔵庫を開けて閉めた際の通知を controller へ送る。controller はライトへ ON 通知をし、カメラへ起動依頼を送る。カメラで撮影した画像を画像認識部へ処理依頼をする。処理が終了したら controller へ処理終了通知をし controller はライトにライト OFF 通知を送る。その後、画像認識部から在庫 DB へ、ペットボトルの個数、種類、情報を転送をする。在庫 DB は画像認識部から送られてきたデータと送られる前のデータと違えば情報提供部へ通知を行う。同じであれば通知は行わない。情報提供部は以前のデータと照らしあわせて違えば、情報取得者(ユーザ)へ表示通知を行う。

4 実験による評価

本研究では学習データを下準備する為にペットボトルの画像をペットボトル飲料を販売している会社のホームページの商品一覧からダウンロードを行うプログラムを作成しデータを集めた。しかし本研究の目的はペットボトルの画像を読み込んだ際にその画像に写っているものがペットボトルであるのかを判別するだけではなく、その判別したペットボトルがなんの種類のペットボトル飲料なのかを判別する。そのためにはたくさんのペットボトル飲料の画像を用意する必要がある。そこでただ単ににペットボトル

の画像をダウンロードしては時間が足りない為、Web ページからダウンロードする方法をとった。しかしその方法でも学習データとしては足りなく、過学習を起してしまうので、画像の増幅をするプログラムを作成した [2][3]。画像の増幅にはガンマ変換、平滑化、回転、射影変換、ノイズ付加の既存の技術を組み合わせさせる。そうすることで約 60 枚程度の画像が約 6000 枚の量にすることができた。

5 考察

本研究での CNN による画像認識を行う判別器を作成し、在庫状況を把握することで Apache と PHP スクリプトを用いた Web サービスを作成することでリアルタイムで購買行動に反映させることは可能だと推測される。

6 おわりに

近年スマート家電や製品の自動化が進んでいる。スマート家電の一例としてスマート冷蔵庫を取り上げる。スマート冷蔵庫はインターネットに接続することで例えばレシピ検索などをすることが出来る。

一方で、中にある食材の種類、個数を把握して調理や購買行動に活かすことは十分に出来ていない。スマート冷蔵庫では製品の在庫管理を冷蔵庫内のカメラの画像から確認して把握することしかできないという問題がある。

本研究の目的は、ペットボトルの画像認識をさせ個数処理をさせる。そして得られた在庫情報をリアルタイムでユーザに知らせ、購買行動に反映させることである。

本研究では、CNN の構造を考え、プログラミング言語 python を利用し、そして得られた在庫状況をリアルタイムでユーザに知らせるためのシステムを構築するために Apache でサーバを活用し PHP でプログラムを作成する。

本研究では CNN による画像認識を行う判別器を作成するためのデータを集めることと、CNN モデルの作成はできたが実装までは出来なかった。今後判別器を完成させ、提案したコミュニケーション図の考察を行う必要がある。

参考文献

- [1] 中山英樹, "深層畳み込みニューラルネットワークによる画像特徴抽出と転移学習", 信学技報, vol.115, no.146, pp. 55-59, 2015
- [2] P.Y.Simard, D.Steinkraus, and J.C.Platt, "Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis," Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR), vol.2, pp.908-914,2003.
- [3] A.Krizhevsky, I.Sutskever, and G.E.Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Proceedings of the 25th International Conference Neural Information Processing Systems, pp.1097-1105, 2012.