

色覚障害者のための食べ時判定アプリの提案

The proposal of eating app for color blind people

下地康平[†] 室賀翼[†] 吉野太雄[†] 増淵凌也[†] 木崎悟[‡]
Kohei Shimoji[†] Tsubasa Muroga[†] Taiyu Yoshino[†] Ryoya Masubuchi[†] Satoru Kizaki[‡]

[†] 日本工学院八王子専門学校 ITスペシャリスト科

[‡] 日本工学院八王子専門学校 ITカレッジ

[†] IT Specialist Department, Nippon Engineering College of Hachioji.

[‡] IT College, Nippon Engineering College of Hachioji.

要旨

本論文では、AIの画像認識技術を使いその食べ時を知らせてくれるソリューションを提案する。料理が皆で楽しめるようになり障害者の抱える問題意識を解決することが目的である。五感による情報入手方法で視覚が占める割合は約9割にも及び、日本には色覚障害者が人口の20分の1もいる。この2つの観点より日本には視覚による弊害を抱えた方達が多いと考える。また料理など色の変化によって行う作業は難しいのではないだろうか。

1. はじめに

人間の五感による情報入手方法は、視覚・聴覚・嗅覚・味覚・触覚であるが、そのうち視覚が占める割合はおおよそ8割以上に及ぶ[1]。世界には視覚に障害を抱えた人も多く、その中で色覚障害者の割合は、日本においては男性で約5%、女性で約0.2%の割合であるが、フランスや北欧では男性で約10%、女性で約0.5%であり、アフリカ系の人では2~4%程にものぼる[2]。この2つの観点より、世界中には視覚による弊害を抱えた人たちが多いと考える。

また、視覚の色の情報量が少ないため、料理をする際に例えば、「肉の焼き加減がわからない」「食材の鮮度が見分けられない」といった意見も見受けられる[3]。このような事例から、料理をする際の色の変化によつての作業は困難である。

これらを解決する手段として、我々はAIによる画像認識技術を使い、その食べ時を知らせてくれるソリューションを提案する。色覚障害者の方が食事を楽しめるように、様々な問題意識を解決することが目的である。AIは身近な存在になり、家電製品や日常品などにも進出している[4]。今やAIなしで語れない時代になりつつある。与えられた画像から何が映っているかを認識することができる技術である画像認識技術も進歩している。

なお、本アプリケーションのターゲットは色覚障害者を中心に対象としているが、他にも料理に苦手意識を抱えているユーザーにも使ってもらえることを想定している。料理に苦手意識を抱えたユーザーの場合、色の判別はできても、どのくらいで食べ時になるのか判断しづらいため、本アプリケーションを適していると言える。

2. 視覚における情報の割合と重要性

人間の五感における情報入手方法は第1章で述べたように、視覚、聴覚、嗅覚、味覚、触覚であるが、視覚が最も高く次いで聴覚なのだが、そのうち視覚が占める割合は約9割に上る[5] (表1参照)。

表1 五感の情報取得の割合

順位	経路	割合
1	視覚	87.0%
2	聴覚	7.0%
3	嗅覚	3.5%
4	触覚	1.5%
5	味覚	1.0%

実際に相手に対する第一印象（以下、印象）は一般的に 3 秒で決まると言われている。人間の印象を決めるときに視覚、聴覚、言語の 3 要素に分かれている。しかし、与える影響力は違う。印象において影響を大きく与えているのは、視覚情報と聴覚情報の 2 つである[6]。これだけで全体の 93% を占めており、どんな話をしているかはさほど重要ではないことがわかる（表 2 参照）。

表 2 印象を決める 3 要素と割合

順位	要素	割合
1	視覚（見た目）	55%
2	聴覚（声のトーン、スピード）	38%
3	触覚	7%

上記の 2 つの表からもわかるように人間は視覚に重きを置いた生活をしていると言える。このことから、生活の障害を改善する優先順位においても視覚が高いと言える。

3. AI ライブラリによる画像認識機能の精度・性能

3.1. 画像認識の精度について

画像認識機能を搭載する際、使用した Computer Vision API[7]は画像から豊富な情報を抽出して視覚データを分析及び処理を行う。また、機械による画像のモデレートを実施し、サービスのキューレーションを支援する。Computer Vision API は Microsoft Azure が提供する画像認識 AI で以下の機能を持つ。画像認識の精度の検証のため Microsoft, Google, IBM, Amazon の画像認識 API と比較した(図 1, 表 3 参照)。この結果、Amazon 以外で焼肉の判定ができることがわかった。



図 1 比較に使用した画像

対象画像	利用 API	認識結果
焼肉	Microsoft Azure Computer Vision	○
	Google Cloud Vision API	○
	IBM Watson Visual Recognition	○
	AWS Amazon Rekognition	×

表 3 各種 API の検証結果と比較

Microsoft Azure の Computer Vision API のデモを使用し解析した結果、手書きのメモ（英語）に対しても正確な結果が表示され精度の面において優秀だと言えた。また、様々な情報を返してくれるサービスであると分かった。画像一つに対して、複数の結果が回答され正確性の割合が表示される。コスト面においても 5000 トランザクションまで無料で利用することができ、コスト面と正確性を比較させ考えた結果、Computer Vision API を利用することにした。

3.2. AI ライブラリによる画像認識

本研究では、4 種類の API を比較して、Microsoft 社の Computer Vision API による画像認識サービスを用いた。Computer Vision API とは、Microsoft Azure というクラウドサービスの 1 つである。このサービスを使用すると開発者は、画像を処理して情報を返す高度なアルゴリズムを利用することができる。今回は、この API を使い Android 端末向けのアプリケーションを開発した。開発環境は Android Studio[8]を使い、プログラミング言語は Java を用いている。端末のカメラ機能を利用して、AI による色判別によって対象の食べ時を知らせるシステムとなっている。

具体的なシステム構成を以下に示す(図 2 参照)．送信された画像を学習済みモデルで分析し、結果を JSON 形式で返す．Microsoft が膨大なデータを使って学習したモデルをもとに分析するため、高い精度で画像の情報を抽出できることが大きな特徴である．今回は、写真内の配色を特定し、利用している．JPEG ファイルや PNG ファイルなどの一般的なイメージファイルをアップロード又は、ファイル URL を指定する．与えられた画像や動画、音声 Microsoft Azure が提供するクラウドに転送し画像を処理し情報を返すためのアルゴリズムが実装されている．

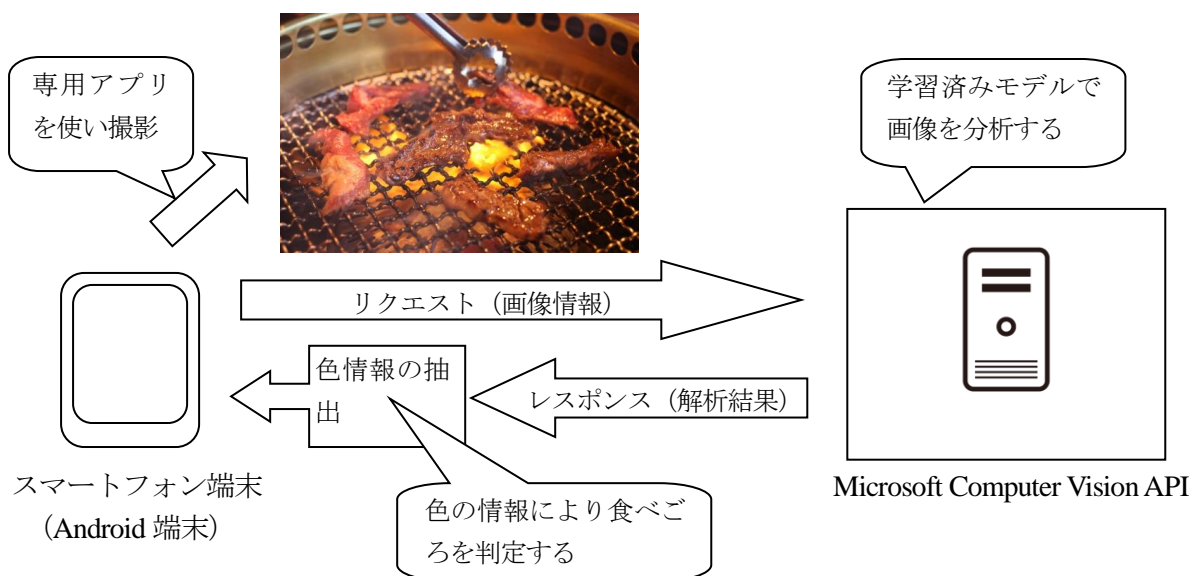


図 2 本アプリケーションのシステム構成

3.3. 食べ時を判定するアルゴリズム

Android で撮影した画像を API にリクエストをすると、JSON 形式の結果が返る．返される情報は 11 種類あるが、視覚的特徴のタグ付け(食べ物であるか)、配色の検出(赤や黒、茶)の 2 種類により情報を抽出し、食べ時であるかを判別している．「食べ物であるか」をまず判定し、「red」がキーワードに含まれる場合は食べ時ではない．「black」とキーワードに含まれたら「焼きすぎ」という判定をしている．「yakniku」と正常に検出した場合は「食べ時である」と表示する(図 3 参照)．検出した物体からそのものの可能性を提示するが、図 4 場合、81%の確率で焼肉であると判定することができる．抽出された JSON 形式の情報からキーワード化したワードが出たら、食べ時であるという判定をする．

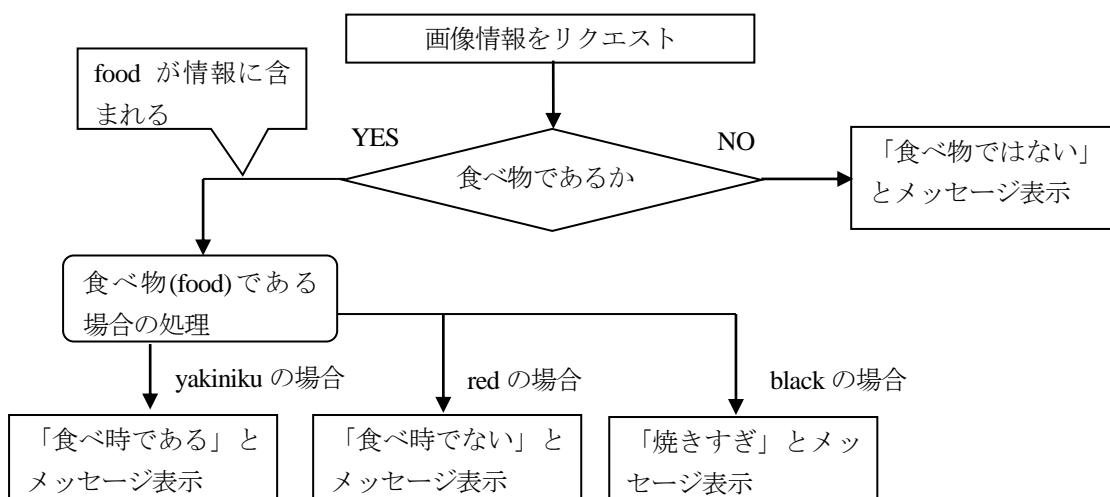
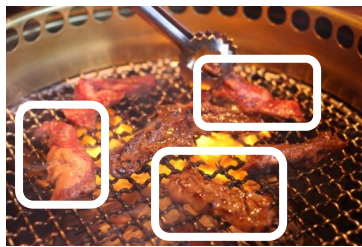


図 3 食べ時を判定するアルゴリズム



JSON 形式（一部抜粋）

```
{ "name": "roasting", "confidence": 0.926457763 },
{ "name": "yakniku", "confidence": 0.8101791 },
{ "name": "meat", "confidence": 0.8432825 }
```

図 4 JSON 形式での画像情報抽出

4. まとめ

本論文では、AI の画像認識技術を使いその食べ時を知らせてくれるソリューションを提案した。料理が皆で楽しめるようになり障害者の抱える問題意識を解決することが目的である。五感による情報入手方法で視覚が占める割合は約 9 割にも及び、日本には色覚障害者が人口の 20 分の 1 もいる。この 2 つの観点より日本には視覚による弊害を抱えた方達は多いと考える。また料理など色の変化によって行う作業は難しいのではないだろうか。

具体的には Microsoft Computer Vision API という技術を使い、Java によるプログラミングから JSON 形式のデータを処理して、色情報を判断するというアルゴリズムを実装した。多数の焼肉や生肉の画像で検証、解析した結果、焼肉の場合、meat, yakiniku, steak, food, 生肉の場合は red meat, red, pink と検出されることがわかった。この結果より、それらのタグが返されたときに食べ時か否かを判別することにした。また、焼きすぎの場合は black, 全く関係ない対象を撮影するなど、設定したキーワード以外が検出された場合は食べ物ではないと判断することにした。

5. 今後の展望

今回、画像を解析するにあたり 4 種類（Microsoft, Google, IBM, Amazon の画像認識 API）を比較したが Google が提供している Google Cloud Vision API が最も優秀な結果を出した。そのため、今後使用する画像認識 API も高精度な API を使用することも視野に入れている。

また、開発時間が不足しており、焼き加減のみしか対応させることができなかった。茹で加減や蒸し加減なども判定できるように修正していきたい。

参考文献

- [1] 加藤宏, “「視覚は人間の情報入力 80%」説の来し方と行方”, 筑波技術大学テクノレポート, Vol25(1), 2017.12.
- [2] カラーバリアフリー, “色使いのガイドライン”, 国立遺伝学研究所, 2005.4.
- [3] 厚木保健福祉事務所大和センター, “料理教室を開催するときの色覚にかかる対応のポイント”, 神奈川県ホームページより, <http://www.pref.kanagawa.jp/docs/e7a/cnt/f38/p389836.html>, 2019.6.14.
- [4] 総務省, “AI・IoT 導入状況と予定”, 平成 30 年版 情報通信白書 第 1 部特集人口減少時代の ICT による持続的成長, <http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h30/html/nd132210.html>.
- [5] 照明学会編, “屋内照明ガイド”, 電気書院, 1978.
- [6] 平野喜久, “天使と悪魔のビジネス用語辞典～世の中にあふれるビジネス用語をまじめに笑い飛ばそう～(メラビアンの法則)”, <http://www2u.biglobe.ne.jp/~hiraki/d74.htm>, 2005.
- [7] Computer Vision API, URL: <https://azure.microsoft.com/ja-jp/services/cognitive-services/computer-vision/>.
- [8] Android Studio, URL: <https://developer.android.com/>.