

CNN を用いた毒きのこの画像分類

北海道科学大学 ○岩田慶介, 大江亮介, 川上敬

要 旨

日本では毎年 40~160 人ほどが毒きのこにより食中毒を引き起こしている。このような問題を解決するために、本研究では、畳み込みニューラルネットワーク(CNN)を用いて毒きのこの画像分類を行うスマートフォンアプリを作成する。この分類器の作成には CNTK を使い、モデルには ImageNet で学習済みの VGG16 を利用する。モデルの有効性を検証するため、特徴が類似するペアを含んだ 10 種類のきのこ画像を学習し、分類精度を確かめる実験を行う。

1. 緒 言

きのこは、古くから日本で親しまれている食材の一つである。しかし、日本では毒きのこの誤食による食中毒事故が後を絶たない。日本では平成 24 年から平成 29 年の間に 214 件の食中毒事故が発生しており、その患者数は 600 人以上にのぼる。このような食中毒事故の原因として、きのこは種類が多く、色や形が似ているものが多く存在することが挙げられる。また、きのこは雨などの天候や成長過程によって見た目に変化することがあり、判別を難しくする要因の一つとなっている。

最も手軽なきのこの判別方法は、目視による判別である。しかし目視による判別は、正しい知識や経験が必要になるため、初心者が誤った判別を行うことがある。そこで本研究では、毒きのこの画像分類を行うスマートフォンアプリを作成する。分類器の作成には CNN を採用する。トレーニングデータは、特徴が類似するグループを含んだ 10 種類のきのこ画像を各クラス 174 枚用意する。テストデータはトレーニングデータとは別に画像を各クラス 10 枚用意した。

2. Deep Learning 技法

2.1 畳み込みニューラルネットワーク (CNN)

CNN とは、画像の特徴量をニューラルネットワークによって学習する Deep Learning の手法である。CNN では入力画像から特徴マップを生成するために、畳み込み層と Pooling 層で構成された畳み込みネットワークを用いる。画像を 3 次元ベクトルとして扱うため、各チャンネルの空間的関連性をくみ取り学習することが可能となる。

2.2 VGGNet

VGGNet は、2014 年の ILSVRC で 2 位となったネットワークである。3x3 畳み込み層と全結合 3 層で構成されており、Pooling 層で特徴マップを 1/2 にした後、チャンネル数を倍に増やす。VGGNet は畳み込み層の深さが違うモデルが複数あり、本研究で使用するモデルは文献 2) の Model-D とする。

本研究では、学習済み VGG16 を採用し fine-tuning を行う。VGG16 は 3x3 の畳み込み 13 層と全結合 3 層で構成されており、Input 層は 224x224x3 となっている。ただし、元の VGG16 は 1000 クラス分類のため、全結合層は自作した 10 クラス分類のものに入れ替える。以下に全結合層の構成を示す。

GlobalAveragePooling2D
FC-4096
FC-4096
FC-10
soft-max

図 1 全結合層の構成

3. 検証実験

3.1 学習データ

本研究では、色や形が似ているきのこの判別も行う。特徴が非常に似ているきのこのグループとして A グループ(ツキヨタケ、ムキヨタケ、ヒラタケ)と B グループ(キタマゴタケ、タマゴタケモドキ)を用意した。また、特徴が少し似ているグループとして C グループ(タマゴタケ、ベニテングタケ、ヒメベニテングタケ)を用意した。これに、D グループ(ドクツルタケ、クサウラベニタケ)を加えた 10 種類を学習データとして使用する。

3.2 学習データセット

集めた学習データをもとに、トレーニングデータを作成する。このとき、不均衡データが存在すると学習に悪影響を与えるため、最もデータの多いクラスに合わせてアップサンプリングを行う。アップサンプリングは ImageDataGenerator を利用して行い、ランダムに回転、水平垂直シフト、左右反転を適用しデータ拡張を行う。拡張後の各クラスの画像枚数は 174 枚である。テストデータは、トレーニングデータとは別のものを各クラス 10 枚ずつ用意する。

3.3 学習

ニューラルネットワークライブラリの Keras および、4.1 で作成したデータセットを使用して学習を行う。モデルは 10 層目までの重みを固定して再学習せず、10 層目以降の重みのみを学習対象とする。最適化手法を Adam(学習率=0.0001)とし、epoch 数は 50、バッチサイズが 4 で学習を行う。トレーニングデータの入力には ImageDataGenerator を使用し、入力画像を 224x224x3 にリサイズする。また、テストデータを用い、学習後のモデルの識別率を検証する。

3.4 結果と考察

上記の設定で学習を 3 回行い、それらの平均を識別率とする。以下に結果を示す。

表1 10クラス分類の識別率

グループ	種類	識別率(%)	平均識別率(%)
A	ツキヨタケ	36.67	58.89
	ムキタケ	63.33	
	ヒラタケ	76.67	
B	キタマゴタケ	70	58.34
	タマゴタケモドキ	46.67	
C	タマゴタケ	70	85.56
	ベニテングタケ	90	
	ヒメベニテングタケ	96.67	
D	ドクツルタケ	90	88.34
	クサウラベニタケ	86.67	

表2 識別率70%未満クラスの誤識別の内訳

	ツキヨタケ	ムキタケ	タマゴタケモドキ
誤識別の内訳	ヒラタケ:9	ツキヨタケ:8	ドクツルタケ:5
	ムキタケ:7	ヒラタケ:2	キタマゴタケ:1
	クサウラ:2	クサウラベニタケ:1	クサウラベニタケ:1
	ベニテングタケ:1		

分類器全体の識別率は72.6%であった。各クラスの識別率ではヒメベニテングタケが96.6%と最も高く、ツキヨタケが36.6%と最も低かった。グループごとの識別率は、Dグループが88.3%と最も高く、Bグループが最も低かった。この結果から、CNNを用いた類似クラスの分類では、精度が落ちることがわかった。

また、表2に示す識別率が70%を下回ったクラスの誤識別の内訳を見ると、ツキヨタケはムキタケとヒラタケへの誤認識が多く、ムキタケはツキヨタケへの誤認識が多い。ツキヨタケ、ムキタケ、ヒラタケは類似するAグループであることから、類似クラス同士が互いの識別率に悪影響を与える可能性がある。しかし、タマゴタケモドキでは、別グループのドクツルタケへの誤認識が多くなっているため、今後さらなる検証が必要である。

4. 実装

4.1 システム構成

Android 端末側は画像の送信と結果の表示を行う。画像は、PNG形式のものをバイト配列に変換し送信する。分類結果はUTF-8の文字列を受信し表示する。

サーバ側では、画像データを受信し分類を行った後、分類結果をAndroid 端末へと送信する。サーバOSはUbuntuを採用し、言語はPythonを使用した。Pythonをサーバ言語とすることにより、通信部分と分類器を同じ言語で実装することが可能となり、処理のオーバーヘッドを抑制できる。

データの送受信はソケット通信で行う。ソケットタイプはIPv4のTCP通信である。通信データはbyte型でやり取りする。本稿では、ローカルネットワーク内にシステムを構築し、端末とサーバ間での通信を行う形で動作の検証を行った。図2に本システムの構成を示す。

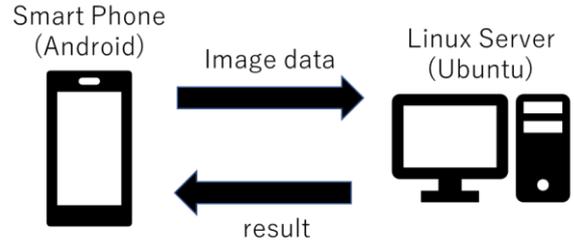


図2 システム構成図

4.2 実機テスト

サーバに作成した分類器を実装する。分類器はKerasのpredictメソッドを使用し、最も確率の高いクラスを診断結果として表示する。図3は、分類器を実装したAndroidアプリケーションの動作画面例である。実機テストの結果、正常に動作することが確認できた。



図3 実機テスト

5. 結 言

本研究では、CNNを用いた毒きのこの画像分類器を作成し、スマートフォンアプリの作成を行った。全体の識別率は72.6%であったが、一部のきのこの識別率は50%を下回っていた。今後は幼菌と成菌のクラス分けや、別のモデルも比較検討し、分類器の精度をより高めることが課題となる。

参 考 文 献

- 1) 厚生労働省：過去のキノコを原因とする食中毒発生状況
https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryoku/shokuhin/kinoko/
- 2) Karen Simonyan & Andrew Zisserman, VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION, ICLR 2015