

物体認識システムを用いたゲームの開発

小助川克也^{†1} 小泉康一^{†2} 大槻正伸^{†2}

概要：ごみ分別の効率化のために、ごみ分別をプログラムで支援することを考えた。そこで、畳み込みニューラルネットワークを用いて、ごみの写真を入力すると、どのごみでありそうかを数値で出力するごみ識別システムを作成した。しかし、本ごみ識別システムでは識別できるごみの種類が少なく、実用的なものにはならなかった。また複雑な背景の写真では識別率が低く、最終的な完成度が低くなってしまった。そこで、幼少期からごみ分別や環境問題に対する意識を向上させるために、本ごみ識別システムを利用した、子供向けのごみ分別に関する教育ゲームや対戦ゲームを作成した。まず、表示されるごみの種類に合うごみを選んで Web カメラで撮影するごみ当てゲームを作成した。次に、Web カメラでごみを撮影すると、キャラクターのステータスが生成され、生成されたキャラクターで対戦する対戦ゲームを作成した。識別精度があまり高くない不完全な物体認識システムをゲームに活かすことができた。不完全なシステムを出力装置として見た場合、偏りのある乱数発生器として見ることができるのではないかと考えられ、コイン投げ、ダイスなどのほぼ均等な、フェアな乱数発生器とは異なる利用法があるのではないかと提案する。

キーワード：物体認識、畳み込みニューラルネットワーク、ごみ分別

1. はじめに

ごみ分別の効率化のために、ごみ分別をプログラムで支援することを考えた。そこで、ごみの写真を入力すると、そのごみの種類を識別するごみ識別システムを作成した。しかし識別できるごみの種類が少なく、また複雑な背景の写真では正答率が低かった。そのため、本研究で作成したごみ識別システムは実用性が低くなってしまった。そこで、本システムをごみ分別支援のためとするのではなく、本システムを利用した教育ゲームを作成することで、幼少期から環境への意識向上を目指す。また、正しく識別できない本システムを利用して、対戦ゲームを作成する。

2. アルゴリズムについて

まず、作成したごみ識別システムに用いたアルゴリズムである畳み込みニューラルネットワーク（以下、CNN と呼ぶ）について説明する。CNN は物体認識などに用いられる機械学習の一手法である。画像認識に用いられる CNN は、生物の視覚神経系を真似ることによってその基本構造を構成する [1]。本研究では、画像を入力し、画像の特徴を抽出する畳み込み層や、画像の解像度を下げるプーリング層を繰り返す行い、最後に全結合層というものを経て、画像にどのごみが写っているかを数値で出力するという CNN を構成する。プーリング層は画像の解像度を下げることで、画像に写っているものの位置ずれに強くする役割を果たす。図 1 に構成する CNN のイメージ図を示す。

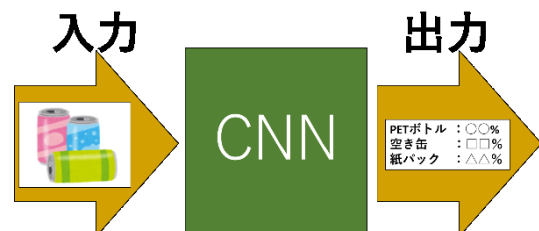


図 1 構成する CNN

3. ごみ識別システムについて

一枚の画像を入力すると、画像に写っているごみがどのごみでありそうかを数値で出力するごみ識別システムを作成した。本研究ではまずペットボトル、空き缶、紙パックの 3 種類のごみを識別するように作成した。

3.1 作成したモデルの構造について

まず、ごみを識別するための CNN の構造を考えた。この構造をモデルと呼ぶ。本研究では入力画像の画質が低いモデルと高いモデルの 2 つを作成した。2 つのモデルの出力を足し合わせることで、識別精度が向上すると考えたためである。各モデルは一つの画像を入力とし、どのごみでありそうかを 3 つの数値で出力する。3 つの出力値はそれぞれペットボトル、紙パック、空き缶に対応している。3 つの出力値の中で最も大きい値に対応しているごみが、識別結果となる。例えばある画像を入力したとき、出力が (1.0, 3.0, 2.0) である場合、一番大きい値である 2 番目に対応する紙パックが識別結果となる。各モデルの構造は付録に示す。

一つは深層学習が大きな注目を浴びるきっかけとなった AlexNet を参考にした [2]。入力画像の大きさは 227×227

^{†1} 福島工業高等専門学校 電気工学科
Electrical Engineering,
National Institute of Technology, Fukushima College
^{†2} 福島工業高等専門学校 電気電子システム工学科
Electrical and Electronic System Engineering,
National Institute of Technology, Fukushima College

である。このモデルをモデル1と呼ぶ。

もう一つのモデルはある程度低画質の画像を入力とした。入力画像の大きさは32×32である。このモデルをモデル2と呼ぶ。

ごみ識別システムは入力画像をまず227×227と32×32にリサイズしてからそれぞれモデルに代入する。それぞれの出力値を平均0分散1に正規化してから、足し合わせる。

3.2 モデルの学習について

作成したモデルがごみを識別できるようにするために、モデルを学習させる。モデルの学習にあたり、Webカメラを使用して、背景が単純になるようにごみの写真を多数枚撮影し、ごみの写真とその写真に写っているごみの種類をまとめたデータセットを作成した。作成したデータセットを学習データセットとテストデータセットに半分ずつ分けた。学習データセットを用いて二つのモデルを学習させた。

モデル1は1回の学習につき、123枚をミニバッチとして同時にモデルに代入して58000回学習させた。ミニバッチとは1回の学習に使用する画像の集まりである。学習100回毎にテストデータセットを代入して求めた正解率のグラフを図2に示す。このグラフを見ることで、学習中のモデルの学習の様子を調べることができる。モデルの学習が進むと正解率が上がっていく。

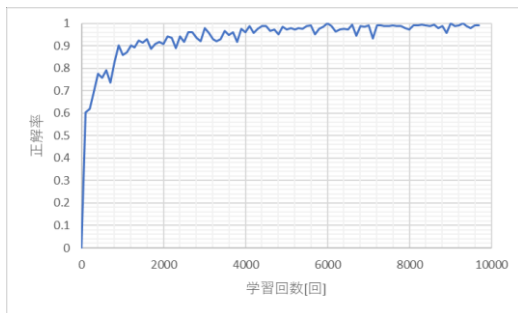


図2 モデル1の学習曲線

学習したモデル1は、テストデータセットの画像ではおおよそ95%を超える正解率で、ごみを正しく識別できた。

モデル2は1回の学習につき、32枚をミニバッチとして同時にモデルに代入して9700回学習させた。学習100回毎にテストデータセットを代入して求めた正解率のグラフを図3に示す。モデル2も、テストデータセットでもおおよそ95%を超える正解率で、ごみを正しく識別できている。

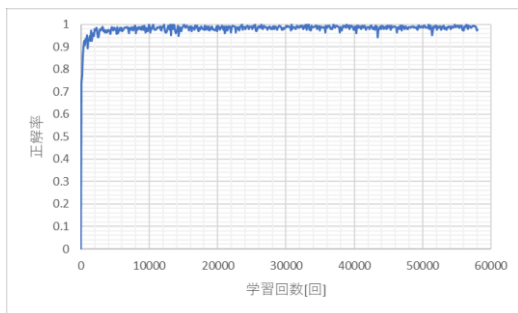


図3 モデル2の学習曲線

3.3 識別精度について

作成したごみ識別システムの識別精度を測定した。測定のために、データセットとは別に、ごみの種類毎に50枚ずつ計150枚の単純な背景のごみの写真を新たに撮影した。また、同様に複雑な背景のごみの写真を撮影した。それらの写真をごみ識別システムに入力し、正しいごみの種類を識別できるか測定した。その測定結果を表1に示す。

表1 ごみ識別システムの識別精度

正解率[%]	ペットボトル	紙パック	空き缶	全体
単純な背景	82	40	94	72
複雑な背景	64	8	90	54

単純な背景の写真では正解率が72%と、ある程度正しく識別できた。しかし、複雑な背景の写真では正解率が54%と、正しく識別できないときがあった。また、紙パックがほかのごみと比べ、正解率が低かった。ほかのごみと比べ、撮影に使用したごみの数が少なかったため、データセットの多様性が低かったためである。

4. ごみ分別アプリについて

本ごみ識別システムを用いてごみ分別アプリを作成した。スマートフォンでごみの写真を撮影すると、ごみの種類を表示するアプリである。図4に本アプリのイメージ図を示す。



図4 ごみ分別アプリ

図5に本アプリのシステムの流れを示す。

ユーザがごみの写真を撮影すると、サーバに送信される。サーバ内にあるごみ識別システムがごみの種類を識別し、識別結果を送り返す。アプリは送られたごみの種類を画面に表示する。

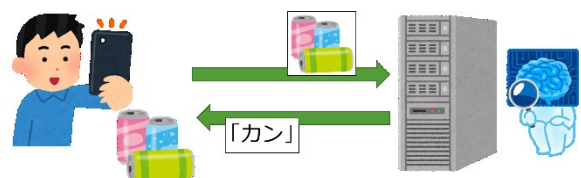


図5 ごみ分別アプリのシステムの流れ

しかし、ごみ識別システムの識別精度があまり高くなくことや、識別できるごみの種類が少ないなど実用性が低くなってしまった。そこで、このごみ識別システムでごみ分別の支援するのを断念し、本システムを子供向けのゲームに利用することで、ごみ分別や環境問題に対する意識を向上させることに活かすことを考えた。

5. 作成した教育ゲームについて

ごみの分別について学べる教育ゲームを作成した。問題文として表示されたごみの種類に合うごみをカメラに映して当てるゲームで、対象者は4歳から6歳の幼児である。図6に本ゲームのイメージ図を示す。

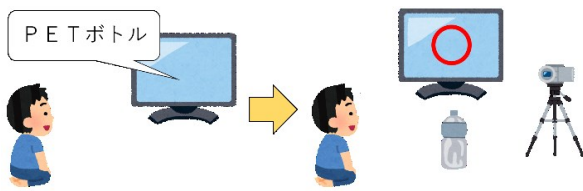


図6 教育ゲーム

図7に実際のゲーム画面を示す。

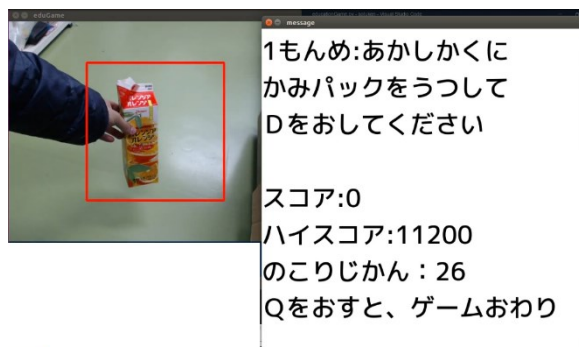


図7 教育ゲームの画面

本ゲームが入ったPCとWebカメラ1台,ペットボトル,紙パック,空き缶を用いる。ゲームの具体的な処理の流れを示す。まず,ごみを文字で画面に表示する。プレイヤーはごみを選びWebカメラで撮影する。撮影された画像はごみ識別システムに入力され,ごみの種類を識別する。識別結果が問題と合っている場合,正解となり,次の問題が表示される。制限時間以内に5問正解すると,ゲームクリアとなる。

本ゲームを他の知育ゲームと比べる。図8に他の知育ゲームの例として,穴の形に合うパズルを当てはめる知育ゲームのイメージ図を示す。他の知育ゲームと比べ,本アプリはプログラムであり,ゲームに使用のごみはゲーム終了後に捨てることのできるため,場所を取らないことが特徴である。昨今の住宅事情を考えると,場所を取らないことは重要である。

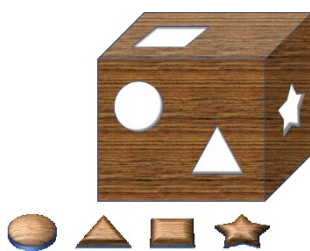


図8 他の知育ゲームの例

6. 作成した対戦ゲームについて

識別精度があまり高くない本ごみ識別システムを利用した対戦ゲームを作成した。二人のプレイヤーがごみの写真を撮影し,その写真からキャラクターのステータスを生成し,そのキャラクターで対戦するゲームである。このゲームの対象者は6から10歳程度の小学生である。本ゲームはバーコードバトラー[3]というゲームを参考にした。図9に本ゲームのイメージ図を示す。



図9 対戦ゲーム

図10に実際のゲーム画面を示す。

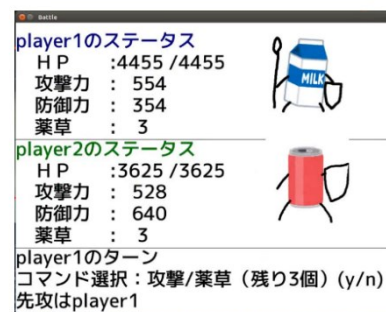
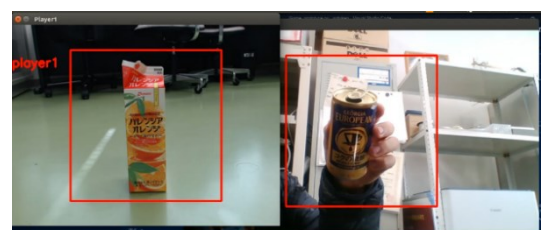


図10 対戦ゲームの画面

二人のプレイヤーがそれぞれWebカメラでごみの写真を一枚ずつ撮影する。撮影したごみの写真を物体認識システムに入力し,その出力値からキャラクターのステータスを生成する。プレイヤーは攻撃と回復のコマンドを順々に選んで対戦する。

6.1 バーコードバトラーについて

バーコードバトラーは,1991年にエポック社から発売されたゲーム機である[3]。様々な商品についているバーコードを読み取り,その数値情報からキャラクターを生成し,そのキャラクターで対戦して遊ぶゲームである。基本ステータスである生命力HP,攻撃力ST,防御力DFのほか,特殊能力などの隠しステータスが生成される。本研究で作成した対戦ゲームの計算はバーコードバトラーを参考にしている[4]。

6.2 本ゲームの具体的な流れ

本対戦ゲームの具体的な流れについて説明する。まず、2人のプレイヤーが、それぞれごみの写真を撮影する。撮影された写真はごみ識別システムに入力される。識別結果がペットボトル、紙パック、空き缶のどれであったかにより、生成されるキャラクターのタイプが決定する。次に、ごみ識別システムの3つの出力値を用いて、キャラクターのタイプに依る計算式によって、キャラクターの体力であるHP、攻撃力、防御力の基本ステータスのほか、隠しステータスである速さと特殊能力が決定する。図10のように、ゲーム画面には基本ステータスは表示されるが、隠しステータスは表示されない。その後、隠しステータスである速さを基に先攻が決まる。そして、先攻となった方のプレイヤーから順々に“攻撃”か“回復”のコマンドを選んで、対戦を行う。“攻撃”は相手を攻撃して、相手のHPを減らすことができる。相手に与えるダメージは、自分の攻撃力や相手の防御力の他、自分と相手のキャラクターのタイプによって決定する。また、ある確率で攻撃が回避されることがある。この確率は隠しステータスである速さや特殊能力によって決定する。“回復”は、薬草を1つ消費して、自分のHPを回復することができる。対戦開始時にプレイヤーはそれぞれ薬草を3つずつ持っている。どちらかのプレイヤーのキャラクターのHPが0になると、対戦終了となる。

6.3 本ゲーム内の計算

対戦ゲームのステータス生成に用いられるごみ識別システムは入力される画像は必ずごみが映っているという前提条件があるため、ごみが映っていない写真を撮影しても、キャラクターのステータスが生成されてしまう。そこで、ごみが映っていない写真を撮影されたときは、生成されるキャラクターのステータスが低くなりやすいような計算式でステータスを生成するようにした。ごみ識別システムはごみが映っていない画像を入力したときに出力される3つの確率はそれぞれ近い値になりやすい。そこで基本ステータスは以下の式で求めることにした。ただし、 y_{max} をごみ識別システムの3つの出力値の最大値、 y_{min} を最小値、基礎値は定数、乱数は0.95から1.05までの範囲の疑似乱数とする。

$$(\text{基本ステータス}) = (y_{max} - y_{min}) * (\text{基礎値}) * (\text{乱数})$$

ごみが映っていない画像を入力した場合、 y_{max} と y_{min} は近い値になりやすいため、基本ステータスが低い値になりやすい。例えばごみが映っていない画像を入力した場合、ごみ識別システムの出力が(1.0, 0.8, 0.9)となった場合、最大値と最小値の差は0.2となり、これは3つの識別を行うための差としては小さい。したがって生成されるステータスは低くなる。

7. まとめ

識別精度があまり高くない物体認識システムでもゲームに利用することで、教育や娯楽に活かすことができた。しかし、作成したごみ識別システムの識別精度を高めることができなかった。識別精度を上げて、複雑な背景であっても教育ゲームの判定が正確に行われるようにするべきである。また、作成した教育ゲーム及び対戦ゲームを多数の人にテストプレイしてもらい、ゲームバランスの調整やプログラムの修正を行うべきである。

不完全なシステムでも、その不完全さを不確定要素とする偏りのある乱数発生器として見ることで、ゲームに活用できるのではないかとこのことを提案する。

参考文献

- [1] 小高知宏, : 機械学習と深層学習—C言語によるシミュレーション—, オーム社, 2016.
- [2] A., Krizhevsky, I., Sutskever and G. E., Hinton, : ImageNet Classification with Deep Convolutional Neural Networks, Neural Information Processing Systems, pp.1097—1105, 2012.
- [3] 沿革 | エポック社, https://epoch.jp/epoch_info/outline.html, 2019年1月31日閲覧.
- [4] バーコードバトルII 解析ページ, <http://barcodebattler.net/>, 2019年2月1日閲覧.

付録

付録 A.1 作成したモデルの構造

図11に作成したモデル1の構造を示す。

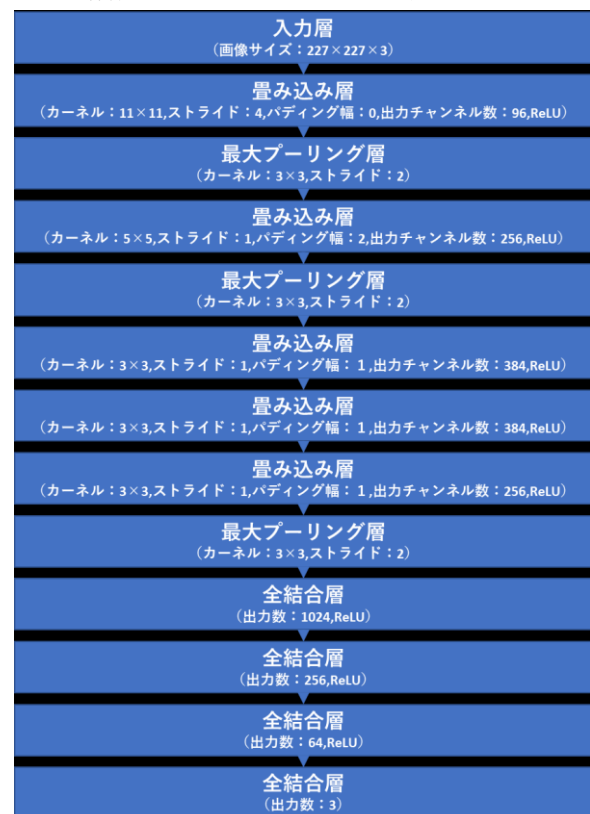


図11 モデル1の構造

図12に作成したモデル2の構造を示す。



図 12 モデル 2 の構造