

CNN を用いた影の物体推定

2018SE070 太田嵯教

指導教員：野呂昌満

1 はじめに

現在、走行中の車両や固定カメラの死角を推定する研究が盛んに行われている [1]. 事前に危険を察知し、円滑な交通環境を実現する技術や軍事的など多分野に渡って活用されている. 柴田ら [2] は、死角は完全になくならないことを問題点として上げている. 死角を推測する手法として移動センサを取り上げている. 移動センサでリスクマップに基づき死角をカバーすることで、歩行者などの存在を車両に予告し、事故を未然に防ぐことができる. リスクマップとは、災害による被害を予測して図にしたものである. ボストン大学の Vivek Goyal ら [3] は、デジタルカメラを用いて影を推測する研究を行っている. 画像をモニターで写し、モニターから発した光をマット (白い壁) に映した. モニターからの光がカメラに影響しないようにカメラとモニターの間に壁を作り、それを死角とした. 映った影を撮影し、色と形をアルゴリズムミックスに解いている.

本研究の目的は、死角から出ている影を識別し死角に隠れた物体を推定して死角を減少させることである. 性能による変化がないと考えたので、死角の推定に機械学習を用いる. 細部の認識が可能であれば大枠の認識が可能と考え、推定する影を帽子の有無とする. 移動物体の影を推定する方法を提案し、妥当性の検証を行う. 本研究の技術課題を以下に示す.

1. ニューラルネットワークの設計と実現
2. 妥当性・精度の検証

2 既存研究ならびに関連研究

2.1 背景差分法

背景差分法は、移動物体検出法の 1 つである. 入力画像と背景画像の差の絶対値を計算することで差分画像が得られる.

2.2 二値化処理

二値化処理とは、画像を白と黒の 2 色 (2 階調) に変化させる処理である. 階調とは、画像の濃度の多様性のことである. 差分画像に対して背景との区別を明確にするため二値化処理を行う. 以下の図 1 に差分画像と二値化処理を示す.

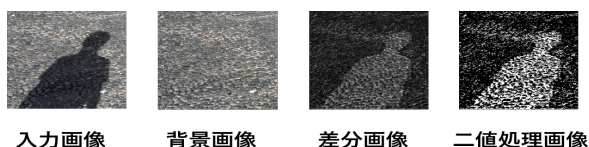


図 1 差分画像と二値化画像

3 提案手法の概要

本研究では、影の物体推定に CNN を用いた手法を提案する. 影画像を「帽子あり」と「帽子なし」とラベル付けし、2 種類に分類する. ニューラルネットワークの設計を行い、データセットの作成を行う. 検証には移動物体の影を推定するため、差分画像に二値化処理して非移動物体の影を除く.

4 ニューラルネットワークの設計

解像度が良く画像として実用的であるため、 220×300 の RGB 画像を入力層とした. 影の認識、影の頭部、帽子を被っているかの特徴抽出に最低 3 層必要と考えて実験を行った. 最終的に畳み込み層 5 層、プーリング層 5 層として設計したニューラルネットワークが一番精度が良かったため使用する. 1 層目、2 層目を $32 \times 3 \times 3$ と指定し、3 層から 5 層目を $64 \times 3 \times 3$ と隠れ層を指定して畳み込みを行い、出力層を 2×128 と設計している. 過学習を抑えるために、各層の間にドロップアウト層を追加している. [帽子あり] と [帽子なし] の 2 クラスに分類する. 設計したニューラルネットワークを用いて実験を行う. 以下の図 2 に設計した CNN を示す.

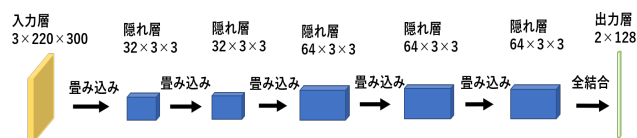


図 2 本研究で設計した CNN

5 データセットの作成

本研究のデータセットには、Web で収集した影画像を用いる. 事前処理として腰部から頭部までの領域をトリミング処理を行った. データセットを 220×300 に画像サイズを統一させた. コントラスト値を変化させ、データ数を増加させた. 「帽子あり」の影画像 6480 枚と「帽子なし」の影画像 7840 枚をトレーニングセットとし、テストセットには 6672 枚のデータセットを作成した.

以下の図 3 に、本研究で用いる影画像を示す.



図 3 影の画像

6 実験

6.1 プロトタイプの実装環境

プロトタイプの実装環境を以下の表 1 に示す。

表 1 実装環境

コンポーネント	コンポーネント名	バージョン
OS	Windows 10 Pro	20H2
使用言語	Python	3.9.7
深層学習フレームワーク	Keras	2.7.0
	TensorFlow	2.7.0
数値計算ライブラリ	Numpy	1.20.3
描画ライブラリ	Matplotlib	3.4.3

6.2 ニューラルネットワークの実現

学習回数を 50 回とし、バッチサイズ 16 と指定した。約 2 万枚のデータセットを用いて学習を行った。

以下に損失関数と正答率を表した図 4 を示す。青線がトレーニングセットに対して、オレンジ線がテストセットに対する結果である。損失関数は約 3% と低く、正答率は約 97% と高い数値で収束している。

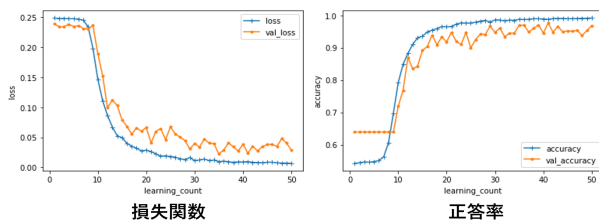


図 4 損失関数と正答率

6.3 背景差分と二値化処理画像の検証

検証の前処理として、背景画像と入力画像を差分する。差分した画像に二値化処理を適用する。移動物体が白く浮き出するため、ネガポジ反転処理を行い移動物体を黒くして影のようにした。以上のように前処理した画像を、設計したニューラルネットワークを用いて検証する。以下の図 5 に前処理の画像を示す。

学習していない太陽に映った影画像に前処理を行い精度の検証を行った。検証画像を図 6、結果を以下の表 2 に示す。

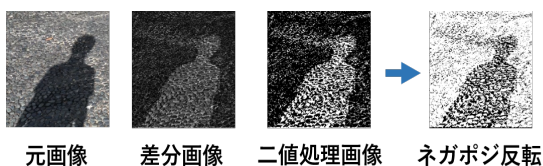


図 5 前処理画像

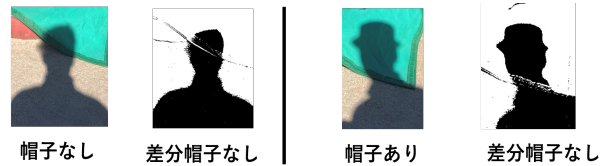


図 6 太陽の影と背景差分の画像

表 2 検証結果

画像	帽子あり精度値	帽子なし精度値	判定
帽子なし	26%	74%	なし
差分なし	97%	3%	あり
帽子あり	51%	49%	あり
差分あり	99%	1%	あり

7 考察

ニューラルネットワークの設計では、各層のドロップアウトの割合を全層 50% 程度が適切と予想していたが、1 層目から 4 層目の各層に 25%、5 層目と全結合層の層に 50% 層との結合を無効化している設計が良い精度がでた。

実験では損失関数は低く正答率は高い値で収束し、過学習が抑えることができた。妥当なモデルであると考え。

前処理画像を用いた精度の検証では、前処理をした画像の方が精度の絶対値の差が元画像よりも大きい。検証した部分の特徴が明確に出るため、帽子を被ってなくいなくても髪型による影響が大きい。背景差分法を適用する際、風の影響で撮影するカメラが揺れて地面の一部が抽出される。地面の一部が分類したい頭部へ影響していると考え。アフロなどの帽子のような髪型の判別が難しいと考える。

8 おわりに

本研究では、CNN を用いて動的な影を推定することを目的とした。今後は、データ数増加の方法をコントラスト以外を適用させ、精度変化を比較する。人数推定やデータセットの収集が困難だった動物や植物など色々な影の形で推定できるか試していきたい。

参考文献

- [1] 嶋田淳, 城戸恵美子, 朴信映, 吉武良治: "前方死角モニタ使用による運転者の運転不安感低減効果について - 見通しの悪い T 字路右折時の主観評価 -", ヒューマンファクターズ Vol.25, No.2 ('21), 2020.
- [2] 柴田一騎, 宮野竜也, 神保智彦: "高応答な死角捕捉を実現する被覆制御開発と実機検証", 計測自動制御学会論文集 Vol.54, No.2, 201/208, 2018.
- [3] Jeff Hecht: "Simple Camera and an Algorithm Let You See around Corners", <https://www.scientificamerican.com/article/a-simple-camera-and-an-algorithm-let-you-see-around-corners/>, 2019, 2021/09 閲覧.