

DNN による星座画像識別の困難さとその解消法 Difficulty on constellation image recognition and its solution

黒田 修二郎[†]
Kuroda Shujiro

和田 俊和[†]
Wada Toshikazu

1. はじめに

近年、Deep Neural Network(DNN)を用いたパターン認識の研究が盛んに行われている。その適用範囲は音声認識、画像認識、動画認識など、多岐にわたっている。入力画像から道路や障害物、人物や車などの領域を検出する Semantic Segmentation や、対象検出・認識の技術は、自動運転などに応用可能であり、実用化を目指した研究が盛んに行われている。

近年の DNN を用いた画像認識の研究で取り扱われている対象は目、鼻、口などから構成される人間や動物の顔のように、局所的な特徴が十分に含まれる対象である。このような豊かな局所特徴が利用可能な状況下では、ボトムアップ的な認識を実行することが出来る。しかし、局所特徴が乏しい対象の場合はどうであろうか。本研究では局所特徴が乏しく、むしろ構造的な特徴の方が識別にとって重要であるような対象として、星座を取り上げ、画像識別用の DNN である VGG-16 [1]を用いて学習可能であるかを検討する。十分な識別精度を示す識別対象として犬を例に、星座画像との相違点について図 1 を用いて説明する。犬の画像の場合、局所領域のみを確認しても目としての特徴を確認できる。しかし、星座画像では局所領域のみを確認しても、点としての特徴しか確認できず、特徴の正体を把握することは困難である。88 クラスの星座画像識別実験の結果、テストデータに対しては 5%程度の識別精度という結果が得られた。この問題を解決するために、星座画像の構造的な特徴が局所的に観測しやすくなる様に星座画像を変換する手法を提案する。この手法を用いた実験でのテストデータに対する識別精度は 31%程度にまで向上することが確認され、提案手法の有効性が確認された。

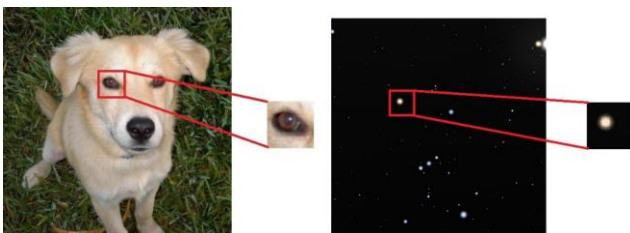


図 1 犬の画像と星座画像の局所領域

2. DNN を用いた画像認識

この章では、近年の DNN を用いた画像認識でよく用いられる Convolutional Neural Network(CNN)モデルの一例として VGG-16 を取り上げる。図 2 に示す人の顔の画像の特徴抽出の例のように、VGG-16 による特徴抽出は入力層から出力層にかけて、入力画像に対する Convolution と Max

Pooling の処理によってボトムアップ的に特徴を抽出する。入力層付近の層では、エッジなどの単純な特徴を抽出する。続いて層が進むにつれてエッジなどの特徴をまとめて構成される目、鼻、口などの人間の顔のパーツの特徴が抽出される。そして出力層付近の層では、これまでに抽出してきた特徴をまとめた、人間の顔全体としての特徴を抽出する。

VGG-16 では、入力層から出力層にかけて、抽出した特徴をボトムアップ的にまとめていくことで入力画像全体としての特徴を抽出するため、局所的な情報に重要な意味の少ない星座画像の識別は困難であると考えられる。

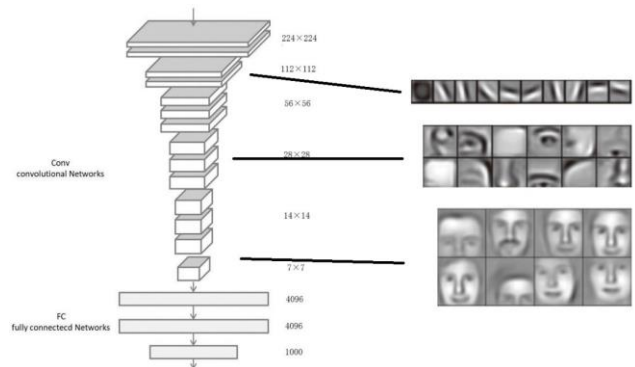


図 2 VGG-16 による人間の顔の特徴抽出 [3]

3. VGG-16 を用いた星座画像識別の予備実験

本研究では、識別対象として星座画像を用いる。理想としては実際の夜空を撮影した星座画像を用いることが望ましいが、星座毎に学習に必要な枚数を確保することは困難である。そこで、星座画像としてオープンソースのプラネタリウムアプリケーションである Stellarium [2]のスクリーンショット画像を使用する。各星座を捉えたスクリーンショット画像を撮影し、画像の回転や範囲内のトリミングを行うことで、星座画像の準備を行う。また、星座画像は実験で用いる VGG-16 の標準入力サイズである 224×224 の画像サイズにリサイズを行う。

準備した星座画像を用いて VGG-16 の Fine Tuning による学習を行い、星座画像の識別が困難であるかを確認する。実験に用いる学習データは 1 クラス 700 枚の合計 61600 枚とし、テストデータは 1 クラス 16 枚の合計 1408 枚とする。Fine Tuning の初期値には ImageNet 学習済みモデルを利用し、学習エポックは 100、識別クラス数は 88 クラスとする。実験に用いる識別精度 Acc は式(1)として定義する。式(1)内の True, False は共に 0 以上の整数値であり、True は入力画像に対し出力の予測結果が正しい場合に 1 加算する。また、False は入力画像に対し出力の予測結果を誤った場合に 1 加算する。Fine Tuning 後のテストデータに対する識別精度は Acc = 0.0482 となった。このことから星座画像の学習は困難であることが確認できる。

$$\text{Acc} = \frac{\text{True}}{\text{True} + \text{False}} \quad (1)$$

[†]和歌山大学 Wakayama University

4. 提案手法による星座画像の変換と変換画像を用いた実験

この章では、提案手法による星座画像の変換方法と変換画像を用いた実験を行う。

4.1 提案手法

既存の DNN では、局所的な特徴が十分存在する対象に対し高い識別精度を示している。しかし、星座画像は局所的な特徴が少なく、VGG-16 では識別が困難であったため、星座の構造的特徴を局所的な情報に反映するような画像変換を提案する。星座画像において、星座を構成する星の領域は、星以外の夜空の領域に比べ比較的高い輝度値を持つ。そこで、画素間の L1 距離と L2 距離を用いた画像変換を行う。尚、画像変換は RGB の各チャンネルでそれぞれ行う。

L1 距離を用いた画像変換では、入力画像の全画素に対し上下左右の四方向の隣接画素との画素値の比較を行い、画素値の変換を行う。右方向の変換は式(2)を満たす場合、式(3)を用いて行う。式(2)内の $P(x,y)$ は (x,y) の画素値を表す。尚、左方向の変換は $(x-1,y)$ 、上方向の変換は $(x,y-1)$ 、下方向の変換は $(x,y+1)$ との画素値の比較を行い、変換を行う。

$$p(x,y) - 1 > p(x+1,y) \quad (2)$$

$$P(x+1,y) = P(x,y) - 1 \quad (3)$$

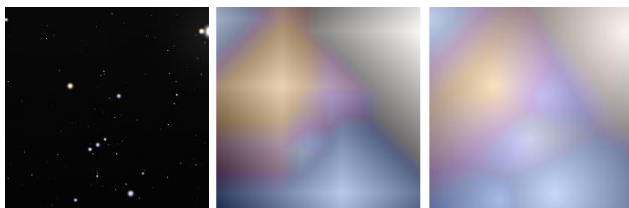
L2 距離を用いた画像変換では、Daniel P.Huttenlocher らの手法[4]の考えを基に変換を行う。各星座画像に対し、10 個程度の L2 距離を計算するための基点を定める。基点は星座画像をグレースケール変換した後に、適切な閾値を用いて星座画像内の星の領域を白とする 2 値画像に変換する。作成した 2 値画像を用いて、白の領域が大きいものから領域の重心を計算し、L2 距離を計算するための基点とする。基点から注目画素の L2 距離を計算し、画素値の変換を行う。 (x,y) と $(x+i,y+j)$ の 2 点間の L2 距離を式(4)の D として定義する。式(4)~式(6)内の i と j の値は、 (x,y) 以外の画素を参照するために用いる整数値である。L2 距離を用いた画像変換は式(5)を持たず場合、式(6)を用いて行う。尚、式(5)と式(6)内の α の値は $(0 < \alpha \leq 2)$ の範囲で変化させて比較を行ったが、識別精度に大きな差は見られなかったため、実験結果は $\alpha = 1$ を用いた場合の結果を示す。

$$D = \sqrt{i^2 + j^2} \quad (4)$$

$$P(x,y) - \alpha D > P(x+i,y+j) \quad (5)$$

$$P(x+i,y+j) = P(x,y) - \alpha D \quad (6)$$

図 3 に提案手法の画像変換を星座画像に対して行った結果の例を示す。図 3 を確認すると、提案手法による画像変換で、異なる星と星が影響しあう境界領域が作成されることを確認できる。この境界領域が、提案手法によって星座の構造的特徴を局所的な情報に反映した領域である。



(a)入力画像 (b)変換画像(L1) (c)変換画像(L2)

図 3 提案手法による星座画像の変換の例

4.2 変換画像を用いた実験

第 4 章 1 節で説明した提案手法による画像変換を行った変換画像を用いて学習を行う。学習のパラメータ等は第 3 章の予備実験と同じものとする。

L1 距離を用いた変換画像での Fine Tuning 後のテストデータに対する識別精度は $\text{Acc} = 0.2947$ となった。また、L2 距離を用いた変換画像での Fine Tuning 後のテストデータに対する識別精度は $\text{Acc} = 0.3110$ となった。このことから、予備実験に比べ大幅に識別精度が向上したことが確認できる。L2 距離と L1 距離では、L2 距離を用いた変換画像の識別精度がわずかに高い。本研究においては L2 距離を用いた変換画像における約 31%が識別精度の最大値となった。

L2 距離を用いた変換での識別精度は約 31%という VGG-16 の識別精度としては低い結果になった。しかし、クラス毎の識別成功数を確認すると、識別成功数に偏りが存在する事を確認した。識別成功数が多いクラスでは学習データ、テストデータ共に、提案手法による画像変換がうまく出来ており、識別成功数が少ないクラスでは、テストデータに対し、提案手法による画像変換がうまく出来ていない事を確認した。提案手法による画像変換の成功・失敗の例を、図 4 に示す。図 4 の入力画像 1 と入力画像 2 はオリオン座の画像を異なる条件下で撮影した画像であり、変換画像 1 と変換画像 2 はそれぞれの入力画像に対応する L2 距離を用いた変換画像である。変換画像 2 は変換画像 1 よりも星の情報が少なくなっていることが確認できる。

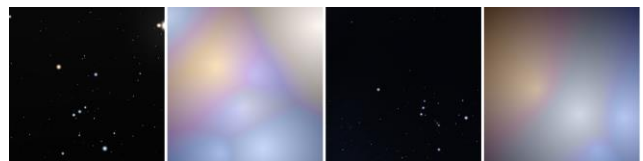


図 4 提案手法による画像変換の成功・失敗例

5. おわりに

本研究では、まず、局所的な領域の属性に多様性が少ない星座画像のボトムアップ的な認識は VGG-16 を用いても困難である事を示した。この問題に対し、星座画像の変換を行う事で識別精度が向上することを確認した。しかし、本研究での提案手法はあくまで変換の一例であり、他の変換方法を考える余地は十分に存在する。また、本研究では画像の変換法について述べたが、トップダウン的な認識を行うアーキテクチャについても検討する余地がある。

参考文献

- [1] Karen Simonyan, Andrew Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", arXiv.org, (2014)
- [2] Fabien Chéreau, Alexander Wolf, Guillaume Chéreau, Georg Zotti, Marcos Cardinot, Hans Lambermont, Khalid AlAjaji, "stellarium v0.19.0", stellarium.org, (2019)
- [3] Honglak Lee, Roger Grosse, Rajesh Ranganath, Andrew Y. Ng, "Unsupervised learning of hierarchical representations with convolutional deep belief networks", dl.acm.org, (2011)
- [4] P.Felzenszwalb, D.Huttenlocher, "Distance Transforms of Sampled Functions", Theory of Computing Vol.8 No.19, (2012)