自然環境映像を対象とした画像検索における

画像特徴ベクトルのインデクシング手法

2021065　　瀬尾　幸斗

（指導教員：鷹野　孝典教授）

# はじめに

地球温暖化の影響により，自然環境観測が重要視されている．しかし，観測されたアーカイブはデータ量が非常に膨大で処理コストが高いため，先行研究では，画像を入力としてニューラルネットワークにより抽出した低次元の画像特徴ベクトルを用いた画像シーン検索手法を提案してきた2)．しかし，VGG畳み込みニューラルネットワークなどの中間層を多く持つニューラルネットワークを用いて特徴ベクトルを抽出する際に，中間層によって得られる特徴ベクトルの性質にばらつきが生じるなどの課題があった．

# 提案手法

本研究では，画像特徴を抽出するネットワークとしてオートエンコーダ，および距離学習ニューラルネットワークを適用した画像特徴の分類手法を提案する．以下に手順を示す．

1. 画像集合を用いて事前学習済みオートエンコーダモデル()を作成する．
2. STEP-1で作成したのエンコーダ出力から画像特徴ベクトルまたは画像特徴テンソルを抽出する．



**図1　中間層による画像特徴ベクトル抽出**



**図2　畳み込み層による画像特徴テンソル抽出**

1. STEP-2で抽出した画像特徴に対し，距離学習ニューラルネットワークを適用することで，正例，負例のペアからベクトル間距離を学習させ，ランキングや分類に適した特徴ベクトルに変換する．



**図3　距離学習による画像特徴の最適化**

1. STEP-3で得られた特徴ベクトルにクラスタリングを適用し，平均や重心といったクラスタの代表となる値を算出し, 代表特徴ベクトルとしてクラスタ内の画像特徴ベクトルに対しインデックスとして設定する．



**図4　代表特徴ベクトルによるインデクシングの例**

# 実験1

提案手法を用いた実験システムでの特徴抽出を行うことで，画像分類やランキングに適した，距離の最適化された特徴ベクトルを得られるかを確認する．データセットには国土地理院の航空画像を用いる2)．M1,M2と比較して距離学習を適用したM3,M4,M5,M6の再現率，適合率が高いことから距離学習には一定の効果があることが確認できる．また，入力がベクトルであるM3,M4よりも，入力がテンソル形式のM5,M6が高精度であることがわかる．



**図5　各モデルの再現率（左）と適合率（右）**

また，M5,M6の学習時におけるエポック数と損失値の結果から，オートエンコーダと距離学習ネットワークを分離して構成することは，分類精度および学習の安定性の観点から一定の効果があると考えられる．

**表1　平均再現率と平均適合率の比較 (M5とM6)**

|  |  |  |
| --- | --- | --- |
|  | M5 | M6 |
| エポック数 | 600 | 1000 |
| 平均再現率 | 0.736 | 0.733 |
| 平均適合率 | 0.736 | 0.750 |

|  |  |
| --- | --- |
| グラフ, ヒストグラム  自動的に生成された説明 | グラフ, ヒストグラム  自動的に生成された説明 |
| M5の損失値変化 | M6の損失値変化 |

**図5　学習における損失値の比較**

# 実験2

提案手法により抽出した特徴ベクトルの距離関係をt-SNEを用いて可視化することで，提案手法の実現可能性を確認する．M1，M2と比較して，M3，M4はクラスを考慮した描画がされていることがわかる．また，M5，M6においても，クラスを考慮した描画がされている．このことから，距離学習ニューラルネットワークを適用することで，画像特徴をクラス毎の距離情報を抽出しやすい特徴ベクトルへと変換できることがわかる．

**表2　次元毎の画像特徴ベクトルのプロット**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **1000次元** | **500次元** | **100次元** |
| **M1** |  |  |  |
| **M2** |  |  |  |
| **M3** |  |  |  |
| **M4** |  |  |  |
| **M5** |  |  |  |
| **M6** |  |  |  |

# まとめと今後の展望

実験より，提案手法を用いて，距離学習ニューラルネットワークを適用することにより，画像特徴の分類精度を向上させることができることを示し，提案手法の実現可能性を確認した．今後の予定としては，提案手法を適用して，自然環境観測において異なる複数地点に設置された観測カメラから記録・蓄積されるアーカイブを対象としたシーン検索システムの構築に応用していくことを検討している．

文献

1. Hiroki Mimura, Masaya Tahara, Kosuke Takano, Nobuya Watanabe, Kin Fun Li: Video Indexing for Live Nature Camera on Digital Earth, International Conference on Advanced Information Networking and Applications, pp.660-667 (2023).
2. 国土地理院，CNNによる水田抽出のための教師画像データ，国土地理院技術資料 H1-No.26 (2023).