

VGGNet に基づく中国書道作品の漢字認識システム

崔, 文一
九州大学大学院芸術工学府

井上, 光平
九州大学大学院芸術工学研究院 : 准教授

<https://hdl.handle.net/2324/4123607>

出版情報 : 画像の認識・理解シンポジウム. (IS2-3-9), pp.1-3, 2020-08-04. 画像の認識・理解シンポジウム (MIRU)
バージョン :
権利関係 :

VGGNet に基づく中国書道作品の漢字認識システム

崔 文一^{1,a)} 井上 光平¹

概要

中国の書道は、手書き漢字の芸術の一種である。書道作品の漢字の分類方法は主に二つの種類、すなわち文字内容と書体種類がある。本稿では、VGGNet に基づく中国漢字の内容と書体を同時に認識できるシステムを提案する。本システムは、中国書道作品のより良い理解を促進することを目的としている。

1. はじめに

中国の書道作品はレイアウトと字形が不規則のため、特徴を抽出することが難しい。畳み込みニューラルネットワーク (Convolutional Neural Network: CNN) とは、「畳み込み層」や「プーリング層」などの幾つかの特徴的な機能を持った層を積み上げることで構成され、特に画像認識の分野で優れた性能を発揮しているネットワークである。CNN によって、漢字をクラスによる分類をすることができる。しかし、二つのクラスを分類するとき(多クラス分類)、CNN の訓練を二回繰り返す必要があるため、計算時間と計算量がかかる。本稿では、多クラス分類できるネットワークを利用して、漢字の内容と書体を同時に認識できるシステムを提案する。

1.1 VGGNet

VGGNet は、Oxford 大学の研究グループが 2014 年に考案したニューラルネットワークのモデルである。図 1 に VGGNet の 16 層バージョンの構造を示す。

VGGNet はシンプルな構造で高い性能を持っている。その特徴は主に以下の 2 点である。

- 1) 「3×3」の小さいサイズのフィルタを使って、計算時間と計算量を省ける。
- 2) 畳み込み層を 2、3 回繰り返して、その後にプーリング層を行う流れを繰り返して、精度を上げる。

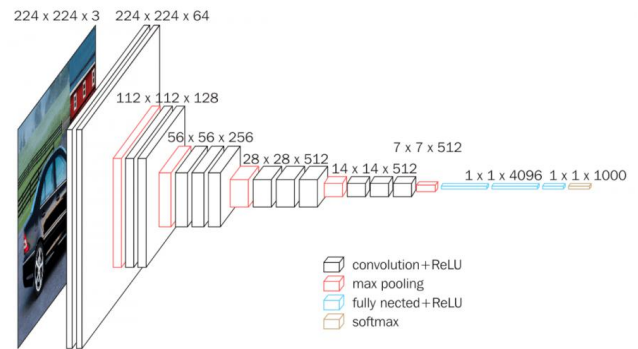


図 1 : VGGNet の構造

1.2 ラベルの作成規則

本節では、本システムのラベルの作成規則を紹介する。本システムのラベルは二つの種類がある。各文字に内容と書体に関する二つのラベルを与える。

種類に応じてラベルの命名方法が違う。以下にラベルの命名方法を示す

1) 書体種類

中国書道の書体は主に、篆書体、隸書体、楷書体、行書体、草書体の五種類がある。略称によってラベルを命名する。

以下にラベルを示す。

- 篆書体 → TST
- 隸書体 → RST
- 楷書体 → KST
- 行書体 → GST
- 草書体 → SST

2) 文字内容

中国漢字は個数が多いため、UNICODE によってラベルを命名する。Unicode とは、文字コードの国際的な業界標準の一つで、以下にラベルの例を示す。

- 天 → u5929
- 地 → u5730
- 玄 → u7384
- 黄 → u9ec4
- 宇 → u5b87

2. 提案システム

本システムは Python プログラミング言語で作成し、主に keras ライブラリの関数を使用する。システムは以下の手順に従って動作する。

- 1) 多クラス分類の VGGNet を作成する。図 2 に本システムのネットワークの構造を示す。

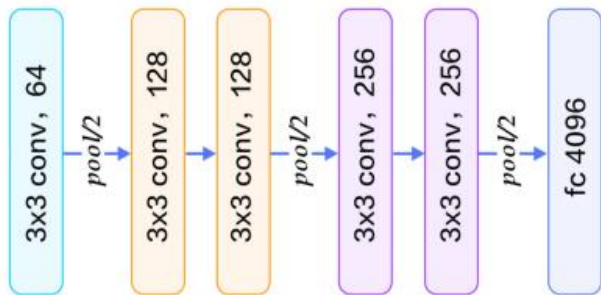


図 2：本システムのネットワークの構造

- 2) ファイルから文字の画像とラベルを読み取り、データベースを作る。各文字のラベルは集めてラベルセットを作る。本システムは 48 個漢字の 5 書体、合計 240 種漢字の画像を使用してデータベースを作る。各種の漢字が 40 枚以上の画像を使う。
- 3) データベースを使って、VGGNet モデルを訓練する。図 3 にモデルの loss と accuracy を示す。図中の train_loss と train_acc はそれぞれトレーニングデータに対する損失と正答率であり、val_loss と val_acc はそれぞれテストデータに対する損失と正答率である。本システムでは 80% のデータをトレーニングデータにして、20% のデータをテストデータにする。



図 3：Loss & Accuracy

- 4) 文字の画像を入力し、訓練されたモデルを使って、文字を認識する。各ラベルの可能性を得る。
- 5) ラベルセットを書体と内容によって二つのセットに

分けて、各セットの中で可能性が一番高いのラベルを選択し、結果を出力する。

3. 実験例

篆書体の『天』、隸書体の『地』、楷書体の『玄』、行書体の『黄』、草書体の『宇』を対象として、システムを実行する。実行結果を図 4(a)、図 4(b)、図 4(c)、図 4(d)、図 4(e)に示す。この結果から、本システムの認識は正確である。



図 4(a)：『天』の実行結果



図 4(b)：『地』の実行結果



図 4(c)：『玄』の実行結果



図 4(d)：『黄』の実行結果

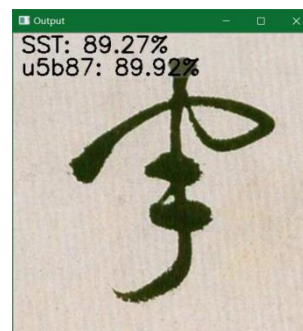


図 4(e)：『宇』の実行結果

4. むすび

VGGNet に基づく中国書道作品の漢字認識システムを提案し、各書体の漢字を使用して実験例を示した。今後、漢字の数量を増やし、より詳細な実験を行って、システムの認識率を向上させたい。

References

¹九州大学大学院芸術工学研究院

^{a)}cui.wenyi.333@s.kyushu-u.ac.jp

- [1] Karen Simonyan & Andrew Zisserman , Very deep convolutional networks for large-scale image recognition, arXiv:1409.1556v6 [cs.CV] 10 Apr 2015
- [2] Neurohive, (Nov 2018) , VGG16 - Convolutional Network for Classification and Detection, Retrieved from <https://neurohive.io/en/popular-networks/vgg16/>