Formatting instructions for NIPS 2018

Anonymous Author(s)
Affiliation
Address
email

Abstract

The growing tourism industry demands that tourists be provided with immediate information about the places they visit. To address this issue we provide a framework which aims to detect and correctly recognize a landmark from images taken by tourists. We test our model on the google landmark recognition challenge hosted on kaggle. In the google landmark recognition challenge hosted on kaggle. Unlike previous challenges, this task requires classification of 15k classes. We demonstrate the efficacy of our models by training it on the given set of images with ground truth and testing it on unlabelled dataset. We obtain xyz score on the kaggle leader board on the testing set.

1 Introduction

Neural networks have made a lot of progress in recent years owing to the presence of multiple image classification dataset like **?**,**?**,cifar1000. Many architectures have been designed to specifically address these problems. Networks such as VGG16 **?** and AlexNet **?** have set the premise for more research. In this project we compare the performance of traditional networks such as the VGG-16 with that of more sophisticated designs such as customized DenseNets adopted from the works of **?** and demonstrate that a customized DenseNet works better. We further utilize the concept of ensemble learning, by using a multi-pathway Neural network where one of the inputs use clustering information of all the images, and a third branch which is pretrained on the image net dataset. These along with the features extracted form the raw images demonstrate the best performance obtaining a score of xyz on the Kaggle leaderboard.

1.1 Method

A neural network based classifier such as the VGG-16 is normally used when it comes to image classification. The standard VGG-16 architecture Figurefig1 consists of multiple convolution blocks of same filter size followed by a pooling layer to downsample the image by half along both the axis.The first two blocks contain two convolution layers each with filters of size 64 and 128 respectively followed by pooling layers. The third, fourth and fifth branches contain three layers of filters of size 256, 512, and 512. Following this the original VGG paper consists of fully connected layers.

In order to reduce the number of parameter in the above architecture, we replace the fully connected layers with convolution layers of filter size (1x1). This has two improvement; firstly it reduces the number of parameter drastically since convolution layers share parameters. Secondly the utilization of convolution layers ensure that the spacial relation of the pixels are taken into account during the classification.

We also investigate deeper networks, namely a 36 layer densely connected convolution network. Motivated by the densely connected architecture of **?** where every layer is connected to every other layer in a block, our version of dense network consists of five densely connected block. The first

Submitted to 32nd Conference on Neural Information Processing Systems (NIPS 2018). Do not distribute.

block consists of four convolution layers with 12 feature maps in each block except for the last, which consists of 36 layers. The second till fifth block also consists of 12 convolution filters in each of the seven layers except on the eight layer which increases as 64, 128, 256, 512 for each of the ultimate layer of a dense block. We also introduce batch normalization **?** in order to ensure that the relu activation layers do not blow out of proportion. Figurefig2.

Our intuition suggests that the performance of our classifier can be improved by using hierarchical classification first before fine tuning those features. We anticipate that landmarks which are water based will be different from those which are land based and so on. Thus we cluster our 15000 classes into 100 clusters. This number was chosen arbitrarily and we would like to experiment more on this later. Thus we send to our 36 layer dense-net, the cluster information that image belongs to along with the raw image.

One of the challenges with the task at hand is that many of the images are dominated by faces of the person taking the picture Figurefig3-picsofhumanfaces and other objects. This makes the classification task harder since the image contains a lot of noise which we would like to remove.

In order to overcome this issue, we decided to introduce pretrained network to generate a probability map of the contents of the image. We use a pretrained VGG-16, trained on the Image-Net dataset. The probability map output by this network is used to improve our classification accuracy. The intuition is that some landmarks, some local features may dominate which can help identify them. With these two modification we observe that our performance improves by xyz.

1.2 Retrieval of style files

The style files for NIPS and other conference information are available on the World Wide Web at

`http://www.nips.cc/`

The file `nips_2018.pdf` contains these instructions and illustrates the various formatting requirements your NIPS paper must satisfy.

The only supported style file for NIPS 2018 is `nips_2018.sty`, rewritten for LaTeX 2_ε. **Previous style files for LaTeX 2.09, Microsoft Word, and RTF are no longer supported!**

The LaTeX style file contains three optional arguments: `final`, which creates a camera-ready copy, `preprint`, which creates a preprint for submission to, e.g., arXiv, and `nonatbib`, which will not load the `natbib` package for you in case of package clash.

New preprint option for 2018 If you wish to post a preprint of your work online, e.g., on arXiv, using the NIPS style, please use the `preprint` option. This will create a nonanonymized version of your work with the text "Preprint. Work in progress." in the footer. This version may be distributed as you see fit. Please **do not** use the `final` option, which should **only** be used for papers accepted to NIPS.

At submission time, please omit the `final` and `preprint` options. This will anonymize your submission and add line numbers to aid review. Please do *not* refer to these line numbers in your paper as they will be removed during generation of camera-ready copies.

The file `nips_2018.tex` may be used as a "shell" for writing your paper. All you have to do is replace the author, title, abstract, and text of the paper with your own.

The formatting instructions contained in these style files are summarized in Sections 2, 3, and 4 below.

2 General formatting instructions

The text must be confined within a rectangle 5.5 inches (33 picas) wide and 9 inches (54 picas) long. The left margin is 1.5 inch (9 picas). Use 10 point type with a vertical spacing (leading) of 11 points. Times New Roman is the preferred typeface throughout, and will be selected for you by default. Paragraphs are separated by ½ line space (5.5 points), with no indentation.

The paper title should be 17 point, initial caps/lower case, bold, centered between two horizontal rules. The top rule should be 4 points thick and the bottom rule should be 1 point thick. Allow ¼ inch

space above and below the title to rules. All pages should start at 1 inch (6 picas) from the top of the page.

For the final version, authors' names are set in boldface, and each name is centered above the corresponding address. The lead author's name is to be listed first (left-most), and the co-authors' names (if different address) are set to follow. If there is only one co-author, list both author and co-author side by side.

Please pay special attention to the instructions in Section 4 regarding figures, tables, acknowledgments, and references.

3 Headings: first level

All headings should be lower case (except for first word and proper nouns), flush left, and bold.

First-level headings should be in 12-point type.

3.1 Headings: second level

Second-level headings should be in 10-point type.

3.1.1 Headings: third level

Third-level headings should be in 10-point type.

Paragraphs There is also a `\paragraph` command available, which sets the heading in bold, flush left, and inline with the text, with the heading followed by 1 em of space.

4 Citations, figures, tables, references

These instructions apply to everyone.

4.1 Citations within the text

The `natbib` package will be loaded for you by default. Citations may be author/year or numeric, as long as you maintain internal consistency. As to the format of the references themselves, any style is acceptable as long as it is used consistently.

The documentation for `natbib` may be found at

`http://mirrors.ctan.org/macros/latex/contrib/natbib/natnotes.pdf`

Of note is the command `\citet`, which produces citations appropriate for use in inline text. For example,

```
\citet{hasselmo} investigated\dots
```

produces

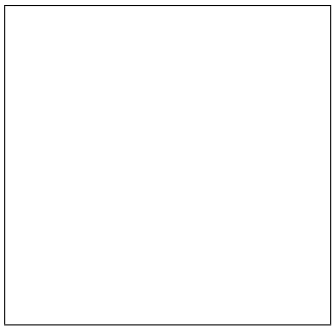
Hasselmo, et al. (1995) investigated…

If you wish to load the `natbib` package with options, you may add the following before loading the `nips_2018` package:

```
\PassOptionsToPackage{options}{natbib}
```

If `natbib` clashes with another package you load, you can add the optional argument `nonatbib` when loading the style file:

```
\usepackage[nonatbib]{nips_2018}
```



[width=0.8]/fig/vgg$_1$6$_b$$aseline.png$

Figure 1: VGG-16 Basic

As submission is double blind, refer to your own published work in the third person. That is, use "In the previous work of Jones et al. [4]," not "In our previous work [4]." If you cite your other papers that are not widely available (e.g., a journal paper under review), use anonymous author names in the citation, e.g., an author of the form "A. Anonymous."

4.2 Footnotes

Footnotes should be used sparingly. If you do require a footnote, indicate footnotes with a number[1] in the text. Place the footnotes at the bottom of the page on which they appear. Precede the footnote with a horizontal rule of 2 inches (12 picas).

Note that footnotes are properly typeset *after* punctuation marks.[2]

4.3 Figures

All artwork must be neat, clean, and legible. Lines should be dark enough for purposes of reproduction. The figure number and caption always appear after the figure. Place one line space before the figure caption and one line space after the figure. The figure caption should be lower case (except for first word and proper nouns); figures are numbered consecutively.

You may use color figures. However, it is best for the figure captions and the paper body to be legible if the paper is printed in either black/white or in color.

4.4 Tables

All tables must be centered, neat, clean and legible. The table number and title always appear before the table. See Table **??**.

Place one line space before the table title, one line space after the table title, and one line space after the table. The table title must be lower case (except for first word and proper nouns); tables are numbered consecutively.

Note that publication-quality tables *do not contain vertical rules.* We strongly suggest the use of the `booktabs` package, which allows for typesetting high-quality, professional tables:

`https://www.ctan.org/pkg/booktabs`

This package was used to typeset Table **??**.

[1] Sample of the first footnote.

[2] As in this example.

5 Final instructions

Do not change any aspects of the formatting parameters in the style files. In particular, do not modify the width or length of the rectangle the text should fit into, and do not change font sizes (except perhaps in the **References** section; see below). Please note that pages should be numbered.

6 Preparing PDF files

Please prepare submission files with paper size "US Letter," and not, for example, "A4."

Fonts were the main cause of problems in the past years. Your PDF file must only contain Type 1 or Embedded TrueType fonts. Here are a few instructions to achieve this.

- You should directly generate PDF files using `pdflatex`.
- You can check which fonts a PDF files uses. In Acrobat Reader, select the menu Files>Document Properties>Fonts and select Show All Fonts. You can also use the program `pdffonts` which comes with `xpdf` and is available out-of-the-box on most Linux machines.
- The IEEE has recommendations for generating PDF files whose fonts are also acceptable for NIPS. Please see `http://www.emfield.org/icuwb2010/downloads/IEEE-PDF-SpecV32.pdf`
- `xfig` "patterned" shapes are implemented with bitmap fonts. Use "solid" shapes instead.
- The `\bbold` package almost always uses bitmap fonts. You should use the equivalent AMS Fonts:

  ```
  \usepackage{amsfonts}
  ```

  followed by, e.g., `\mathbb{R}`, `\mathbb{N}`, or `\mathbb{C}` for $\mathbb{R}$, $\mathbb{N}$ or $\mathbb{C}$. You can also use the following workaround for reals, natural and complex:

  ```
  \newcommand{\RR}{I\!\!R} %real numbers
  \newcommand{\Nat}{I\!\!N} %natural numbers
  \newcommand{\CC}{I\!\!\!\!C} %complex numbers
  ```

  Note that `amsfonts` is automatically loaded by the `amssymb` package.

If your file contains type 3 fonts or non embedded TrueType fonts, we will ask you to fix it.

6.1 Margins in LaTeX

Most of the margin problems come from figures positioned by hand using `\special` or other commands. We suggest using the command `\includegraphics` from the `graphicx` package. Always specify the figure width as a multiple of the line width as in the example below:

```
\usepackage[pdftex]{graphicx} ...
\includegraphics[width=0.8\linewidth]{myfile.pdf}
```

See Section 4.4 in the graphics bundle documentation (`http://mirrors.ctan.org/macros/latex/required/graphics/grfguide.pdf`)

A number of width problems arise when LaTeX cannot properly hyphenate a line. Please give LaTeX hyphenation hints using the `\-` command when necessary.

Acknowledgments

Use unnumbered third level headings for the acknowledgments. All acknowledgments go at the end of the paper. Do not include acknowledgments in the anonymized submission, only in the final paper.

References

References follow the acknowledgments. Use unnumbered first-level heading for the references. Any choice of citation style is acceptable as long as you are consistent. It is permissible to reduce the font

size to `small` (9 point) when listing the references. **Remember that you can use more than eight pages as long as the additional pages contain *only* cited references.**

[1] Alexander, J.A. & Mozer, M.C. (1995) Template-based algorithms for connectionist rule extraction. In G. Tesauro, D.S. Touretzky and T.K. Leen (eds.), *Advances in Neural Information Processing Systems 7*, pp. 609–616. Cambridge, MA: MIT Press.

[2] Bower, J.M. & Beeman, D. (1995) *The Book of GENESIS: Exploring Realistic Neural Models with the GEneral NEural SImulation System.* New York: TELOS/Springer–Verlag.

[3] Hasselmo, M.E., Schnell, E. & Barkai, E. (1995) Dynamics of learning and recall at excitatory recurrent synapses and cholinergic modulation in rat hippocampal region CA3. *Journal of Neuroscience* **15**(7):5249-5262.