
Exploración, Transformación y Clustering de Imágenes de Flores: Consideraciones para Datos de Alta Dimensionalidad

Alejandro Uribe

Maestría en Minería de Datos y Descubrimiento del Conocimiento

Buenos Aires, Argentina

jauriberamirez@gmail.com

Abstract

El estudio se adentró en el mundo de las imágenes florales, llevando a cabo transformaciones de datos y aplicando técnicas de agrupamiento sin recurrir a modelos previamente entrenados. A través de este enfoque analítico, se buscó desentrañar las relaciones inherentes entre las imágenes, identificar patrones visuales y agruparlas en categorías cohesivas. Los descubrimientos proporcionaron valiosas perspectivas sobre la diversidad y similitudes presentes en las imágenes florales, estableciendo así una base sólida para investigaciones futuras en visión por computadora y clasificación de imágenes florales. Este enfoque subrayó que, incluso sin el entrenamiento de modelos, era posible extraer información significativa y relevante de conjuntos de datos visuales.

1 Introducción

El procesamiento de imágenes se presentaba como un desafío debido a la alta dimensionalidad que caracterizaba a estas representaciones visuales. La esencia de una imagen digital residía en una matriz de dimensiones $N \times M$, donde cada elemento individual, o píxel, encerraba información específica sobre un color determinado. Cada píxel traducía la intensidad de la luz en ese punto, usualmente variando en un rango de $[0, 255]$, lo que equivalía a 8 bits de información. Para representar imágenes a color, se utilizaba un modelo basado en la percepción humana, que operaba mediante un sistema aditivo fundamentado en los componentes primarios del color: Rojo, Verde y Azul (RGB, por sus siglas en inglés). De esta manera, la representación digital de una imagen en color requería tres matrices distintas de dimensiones $N \times M$, correspondientes a los canales de Rojo, Verde y Azul.

En este contexto, el presente trabajo se propuso adentrarse en el universo del procesamiento de imágenes. Para ello, se plantearon una serie de manipulaciones que tenían por finalidad preparar el conjunto de datos para la detección y exploración de agrupamientos naturales. Esta exploración no solo se concentró en aspectos técnicos y de comparabilidad de las imágenes, sino que se extendió hacia la búsqueda de características distintivas a través de análisis y transformaciones pertinentes. El conjunto de datos base estaba conformado por 210 imágenes a color de flores pertenecientes a 10 especies diferentes, acompañadas de etiquetas que indicaban la especie a la que pertenecía cada imagen. A lo largo del trabajo, se abordaron distintas etapas de preprocesamiento, manipulación y análisis de datos, con el objetivo de proporcionar un enfoque completo y efectivo para la detección de patrones naturales en este conjunto de imágenes.

2 Trabajo relacionado

Lorem Ipsum



Figure 1: Sample figure caption.

3 Preprocesamiento de imágenes

Lorem Ipsum

4 Análisis de componentes principales

Lorem Ipsum

5 Submission of papers to NeurIPS 2023

Please read the instructions below carefully and follow them faithfully. **Important:** This year the checklist will be submitted separately from the main paper in OpenReview, please review it well ahead of the submission deadline: <https://neurips.cc/public/guides/PaperChecklist>.

5.1 Style

Papers to be submitted to NeurIPS 2023 must be prepared according to the instructions presented here. Papers may only be up to **nine** pages long, including figures. Additional pages *containing only acknowledgments and references* are allowed. Papers that exceed the page limit will not be reviewed, or in any other way considered for presentation at the conference.

The margins in 2023 are the same as those in previous years.

Authors are required to use the NeurIPS L^AT_EX style files obtainable at the NeurIPS website as indicated below. Please make sure you use the current files and not previous versions. Tweaking the style files may be grounds for rejection.

5.2 Retrieval of style files

The style files for NeurIPS and other conference information are available on the website at

<http://www.neurips.cc/>

The file `neurips_2023.pdf` contains these instructions and illustrates the various formatting requirements your NeurIPS paper must satisfy.

The only supported style file for NeurIPS 2023 is `neurips_2023.sty`, rewritten for L^AT_EX 2_ε. **Previous style files for L^AT_EX 2.09, Microsoft Word, and RTF are no longer supported!**

The L^AT_EX 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.

Preprint option If you wish to post a preprint of your work online, e.g., on arXiv, using the NeurIPS 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, as long as you do not say which conference it was submitted to. Please **do not** use the `final` option, which should **only** be used for papers accepted to NeurIPS.

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 `neurips_2023.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 ??, ??, and ?? below.

72 **6 General formatting instructions**

73 The text must be confined within a rectangle 5.5 inches (33 picas) wide and 9 inches (54 picas) long.
74 The left margin is 1.5 inch (9 picas). Use 10 point type with a vertical spacing (leading) of 11 points.
75 Times New Roman is the preferred typeface throughout, and will be selected for you by default.
76 Paragraphs are separated by 1/2 line space (5.5 points), with no indentation.

77 The paper title should be 17 point, initial caps/lower case, bold, centered between two horizontal
78 rules. The top rule should be 4 points thick and the bottom rule should be 1 point thick. Allow 1/4 inch
79 space above and below the title to rules. All pages should start at 1 inch (6 picas) from the top of the
80 page.

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

85 Please pay special attention to the instructions in Section ?? regarding figures, tables, acknowledg-
86 ments, and references.

87 **7 Headings: first level**

88 All headings should be lower case (except for first word and proper nouns), flush left, and bold.
89 First-level headings should be in 12-point type.

90 **7.1 Headings: second level**

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

92 **7.1.1 Headings: third level**

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

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

96 **8 Citations, figures, tables, references**

97 These instructions apply to everyone.

98 **8.1 Citations within the text**

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

102 The documentation for `natbib` may be found at

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

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

106 `\citet{hasselmo}` investigated\dots

107 produces

108 Hasselmo, et al. (1995) investigated...

109 If you wish to load the `natbib` package with options, you may add the following before loading the
110 `neurips_2023` package:



Figure 2: Sample figure caption.

111 `\PassOptionsToPackage{options}{natbib}`

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

114 `\usepackage[nonatbib]{neurips_2023}`

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

120 8.2 Footnotes

121 Footnotes should be used sparingly. If you do require a footnote, indicate footnotes with a number¹
122 in the text. Place the footnotes at the bottom of the page on which they appear. Precede the footnote
123 with a horizontal rule of 2 inches (12 picas).

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

125 8.3 Figures

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

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

132 8.4 Tables

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

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

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

140 <https://www.ctan.org/pkg/booktabs>

141 This package was used to typeset Table ??.

142 8.5 Math

143 Note that display math in bare TeX commands will not create correct line numbers for sub-
144 mission. Please use LaTeX (or AMSTeX) commands for unnumbered display math. (You
145 really shouldn’t be using `$$` anyway; see <https://tex.stackexchange.com/questions/503/why-is-preferable-to> and [https://tex.stackexchange.com/questions/40492/](https://tex.stackexchange.com/questions/40492/what-are-the-differences-between-align-equation-and-displaymath)
146 [what-are-the-differences-between-align-equation-and-displaymath](https://tex.stackexchange.com/questions/40492/what-are-the-differences-between-align-equation-and-displaymath) for more infor-
147 mation.)
148

¹Sample of the first footnote.

²As in this example.

Table 1: Sample table title

Part		
Name	Description	Size (μm)
Dendrite	Input terminal	~ 100
Axon	Output terminal	~ 10
Soma	Cell body	up to 10^6

8.6 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.

9 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 file uses. In Acrobat Reader, select the menu Files>Document Properties>Fonts and select Show All Fonts. You can also use the program `pdf fonts` which comes with `xpdf` and is available out-of-the-box on most Linux machines.
- `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}{\mathbb{R}} %real numbers
\newcommand{\Nat}{\mathbb{N}} %natural numbers
\newcommand{\CC}{\mathbb{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.

9.1 Margins in L^AT_EX

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 L^AT_EX cannot properly hyphenate a line. Please give LaTeX hyphenation hints using the `\-` command when necessary.

10 Supplementary Material

Authors may wish to optionally include extra information (complete proofs, additional experiments and plots) in the appendix. All such materials should be part of the supplemental material (submitted separately) and should NOT be included in the main submission.

References

References follow the acknowledgments in the camera-ready paper. 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. Note that the Reference section does not count towards the page limit.

- [1] Alexander, J.A. & Mozer, M.C. (1995) Template-based algorithms for connectionist rule extraction. In G. Tesauero, 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.