

---

# Autoencoders for Image Reconstruction and Contrastive Learning

---

**Kevin Lewis**

Department of Computer Science  
The Ohio State University  
281 W Lane Ave, Columbus, OH 43210  
lewis.3164@osu.edu

**Shashank Raghuraj**

Department of Computer Science  
The Ohio State University  
281 W Lane Ave, Columbus, OH 43210  
raghuraj.2@osu.edu

**Michael Lin**

Department of Computer Science  
The Ohio State University  
281 W Lane Ave, Columbus, OH 43210  
lin.3976@osu.edu

## Abstract

The abstract paragraph should be indented  $\frac{1}{2}$  inch (3 picas) on both the left- and right-hand margins. Use 10 point type, with a vertical spacing (leading) of 11 points. The word **Abstract** must be centered, bold, and in point size 12. Two line spaces precede the abstract. The abstract must be limited to one paragraph.

## 1 Introduction

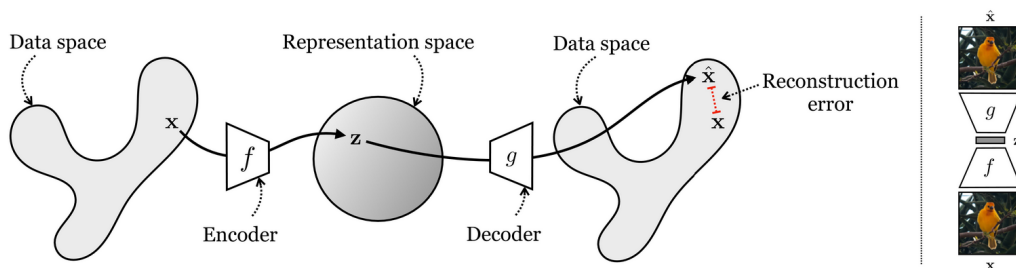


Figure 1: A digram of image reconstruction using Autoencoders, taken from [1]. The encoder learns to compress the image into a smaller feature vector, while the decoder learns to to reconstruct the original image using this representation. The goal is to reconstruct the original image as closely as possible.

Image understanding is useful for many machine learning applications such as classification or generation. Raw image data however is high-dimensional and contains sparse signal, which can lead to low performance when using traditional ML models. In this paper we study a class of neural networks called Autoencoders [1], which are designed to compress images through an encoder and reconstruct the original image through a decoder. We further investigate the effectiveness of

contrastive learning [1] with the goal of training the encoder to produce representations that are invariant to fixed transformations, such as the shape of objects or their colors.

## 2 Problem Formulation

Consider a dataset of points  $\mathbf{x} \in \mathbb{R}^N$ .

**For Image Reconstruction**, given datapoints  $\mathbf{x} \in \mathbb{R}^M$  draw from some distribution, we are given encoder  $f : \mathbb{R}^N \rightarrow \mathbb{R}^M$ , decoder  $g : \mathbb{R}^M \rightarrow \mathbb{R}^N$ , where  $N \ll M$ . Our goal is to then learn  $f^*, g^*$  such that

$$f^*, g^* = \operatorname{argmax}_{f, g} \mathbb{E}_{\mathbf{x}} \|g(f(\mathbf{x})) - \mathbf{x}\|_2^2$$

from [1]. Intuitively, we should expect our encoder  $f(\mathbf{x}) = \mathbf{z}$  to learn a useful representation  $\mathbf{z} \in \mathbb{R}^M$  that retains most of the information in the original image. We then want  $g(\mathbf{z}) = \hat{\mathbf{x}}$  to learn how to use this representation to reconstruct the original image.

**For Contrastive Learning**, given some datapoint  $\mathbf{x} \in \mathbb{R}^d$ , similar (positive) datapoints  $\mathbf{x}^+$ , and dissimilar (negative) datapoints  $\mathbf{x}^-$ , we aim to minimize  $D(\mathbf{x}, \mathbf{x}^+)$  and maximize  $D(\mathbf{x}, \mathbf{x}^-)$ , where  $D$  is some distance function. In practice, we may take  $\mathbf{x}^+ = T(\mathbf{x})$ , where  $T$  is a chosen transformation applied to our dataset (e.g. cropping) to preserve certain features, while  $\mathbf{x}^-$  may be randomly sampled from our dataset [1].

## 3 Methodology

We first

## 4 Experiments

To evaluate the performance of our autoencoder, we first generate a dataset of 50,000 images of simple shapes (e.g. triangles, circles, squares) of different colors, sizes, and orientation, and save them using the Pillow library.

Papers to be submitted to NeurIPS 2025 must be prepared according to the instructions presented here. Papers may only be up to **nine** pages long, including figures. Additional pages *containing references, checklist, and the optional technical appendices* do not count as content pages. Papers that exceed the page limit will not be reviewed, or in any other way considered for presentation at the conference.

The margins in 2025 are the same as those in previous years.

Authors are required to use the NeurIPS L<sup>A</sup>T<sub>E</sub>X 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.

### 4.1 Retrieval of style files

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

<https://neurips.cc>

The file `neurips_2025.pdf` contains these instructions and illustrates the various formatting requirements your NeurIPS paper must satisfy.

The only supported style file for NeurIPS 2025 is `neurips_2025.sty`, rewritten for L<sup>A</sup>T<sub>E</sub>X 2<sub>ε</sub>. **Previous style files for L<sup>A</sup>T<sub>E</sub>X 2.09, Microsoft Word, and RTF are no longer supported!**

The L<sup>A</sup>T<sub>E</sub>X 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_2025.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 5, 6, and 7 below.

## 5 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 1/2 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 1/4 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 7 regarding figures, tables, acknowledgments, and references.

## 6 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.

### 6.1 Headings: second level

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

#### 6.1.1 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.

## 7 Citations, figures, tables, references

These instructions apply to everyone.

## 7.1 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\dotso
```

produces

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

If you wish to load the `natbib` package with options, you may add the following before loading the `neurips_2025` 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]{neurips_2025}
```

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” and include a copy of the anonymized paper in the supplementary material.

## 7.2 Footnotes

Footnotes should be used sparingly. If you do require a footnote, indicate footnotes with a number<sup>1</sup> 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.<sup>2</sup>

## 7.3 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.

## 7.4 Tables

All tables must be centered, neat, clean and legible. The table number and title always appear before the table. See Table 1.

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.

---

<sup>1</sup>Sample of the first footnote.

<sup>2</sup>As in this example.



Figure 2: Sample figure caption.

Table 1: Sample table title

Part		
Name	Description	Size ( $\mu\text{m}$ )
Dendrite	Input terminal	$\sim 100$
Axon	Output terminal	$\sim 10$
Soma	Cell body	up to $10^6$

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.

## 7.5 Math

Note that display math in bare TeX commands will not create correct line numbers for submission. Please use LaTeX (or AMSTeX) commands for unnumbered display math. (You really shouldn't be using \$\$ anyway; see <https://tex.stackexchange.com/questions/503/why-is-preferable-to> and <https://tex.stackexchange.com/questions/40492/what-are-the-differences-between-align-equation-and-displaymath> for more information.)

## 7.6 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.

## 8 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}{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.

## 8.1 Margins in L<sup>A</sup>T<sub>E</sub>X

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<sup>A</sup>T<sub>E</sub>X cannot properly hyphenate a line. Please give LaTeX hyphenation hints using the `\-` command when necessary.

## Acknowledgments and Disclosure of Funding

Use unnumbered first level headings for the acknowledgments. All acknowledgments go at the end of the paper before the list of references. Moreover, you are required to declare funding (financial activities supporting the submitted work) and competing interests (related financial activities outside the submitted work). More information about this disclosure can be found at: <https://neurips.cc/Conferences/2025/PaperInformation/FundingDisclosure>.

Do **not** include this section in the anonymized submission, only in the final paper. You can use the `ack` environment provided in the style file to automatically hide this section in the anonymized submission.

## References

### References

- [1] A. Torralba, P. Isola, and W. Freeman. *Foundations of Computer Vision*. Adaptive Computation and Machine Learning series. MIT Press, 2024. ISBN 9780262378666. URL <https://mitpress.mit.edu/9780262048972/foundations-of-computer-vision/>.

## A Technical Appendices and Supplementary Material

Technical appendices with additional results, figures, graphs and proofs may be submitted with the paper submission before the full submission deadline (see above), or as a separate PDF in the ZIP file below before the supplementary material deadline. There is no page limit for the technical appendices.