

LEARNING ENRICHED FEATURES FOR FAST IMAGE RESTORATION AND ENHANCEMENT*

Fredy Abel Huanca Torres[†], Jose Edison Perez Mamani[‡] y Henryr Ivan Arias Mamani[§]

Abstract

Image restoration is an important task in surveillance, computational photography, medical imaging, and remote sensing. Recently, convolutional neural networks (CNNs) have achieved dramatic improvements over conventional approaches. This paper presents a novel architecture with the collective goals of maintaining spatially precise high-resolution representations and receiving contextual solid information from the low-resolution representations. The core of the approach is a multi-scale residual block containing several key elements: parallel multi-resolution convolution streams, information exchange across the multi-resolution streams, spatial and channel attention mechanisms, and attention-based multi-scale feature aggregation. Extensive experiments on real image benchmark datasets demonstrate that MIRNet achieves state-of-the-art results for various image processing tasks.

Keywords Image denoising, super-resolution, and image enhancement

INTRODUCCIÓN

La restauración de imágenes es una tarea importante en vigilancia, fotografía computacional, imágenes médicas y teledetección. Recientemente, las redes neuronales convolucionales (CNN) han logrado mejoras dramáticas sobre los enfoques convencionales. Este artículo presenta una arquitectura novedosa con los objetivos colectivos de mantener representaciones espacialmente precisas de alta resolución y recibir información contextual sólida de las representaciones de baja resolución. El núcleo del enfoque es un bloque residual multiescala que contiene varios elementos clave: flujos paralelos de convolución multiresolución, intercambio de información a través de los flujos multiresolución, mecanismos de atención espacial y de canales, y agregación de características multiescala basada en la atención. Extensos experimentos en conjuntos de datos de referencia de imágenes reales demuestran que MIRNet logra resultados de vanguardia para una variedad de tareas de procesamiento de imágenes.

TRABAJO RELACIONADOS

analiza los métodos existentes basados en CNN para la restauración y mejora de imágenes. Compara dos diseños de arquitectura: un codificador-decodificador y procesamiento de características de alta resolución (escala única). La sección también destaca algunas limitaciones de los métodos

basados en CNN existentes, como la pérdida de detalles espaciales finos en los modelos de codificador-decodificador.

Image denoising: Mientras que las técnicas clásicas modifican los coeficientes de transformación o los píxeles de vecindad promedio, los algoritmos basados en la autosimilitud como NLM y BM3D se han mostrado muy prometedores para lograr una eliminación de ruido de alta calidad. Además, se han desarrollado numerosos algoritmos basados en parches que explotan la redundancia y la autosimilitud en las imágenes. Más recientemente, los enfoques basados en el aprendizaje profundo han logrado avances significativos en la eliminación de ruido de imágenes, superando los métodos tradicionales hechos a mano y brindando resultados favorables.

Super-resolution (SR): En la era anterior al aprendizaje profundo, se desarrollaron una variedad de algoritmos de súper resolución (SR), aprovechando la teoría del muestreo, la interpolación guiada por bordes, las imágenes previas naturales, los ejemplos de parches y las representaciones dispersas. Sin embargo, con la llegada del aprendizaje profundo, ha habido un cambio significativo hacia enfoques de SR basados en datos que arrojan resultados superiores. Estos enfoques difieren en sus diseños de arquitectura, con métodos tempranos que generan directamente imágenes de alta resolución (HR) a partir de entradas de baja resolución (LR), mientras que las redes SR recientes emplean el marco de aprendizaje residual para capturar detalles de imágenes de alta frecuencia y agregarlos al Entrada LR para el resultado final súper resuelto. Otras redes SR incorporan el aprendizaje recursivo, la reconstrucción progresiva, las conexiones densas, los mecanismos de atención, el aprendizaje de múltiples ramas y las redes antagónicas generativas (GAN).

Image enhancement. En el contexto de la mejora de imágenes, es común que las imágenes capturadas por las cámaras carezcan de viveza y contraste debido a varios factores, como condiciones de iluminación desfavorables y limitaciones de los dispositivos de la cámara. La ecualización de histogramas es un método ampliamente utilizado para la mejora de imágenes, pero a menudo conduce a resultados con una mejora insuficiente o excesiva. Inspirándose en la teoría Retinex, se han propuesto varios algoritmos de mejora que pretenden imitar la visión humana. Recientemente, las redes neuronales convolucionales (CNN) han demostrado un gran éxito al abordar los desafíos de mejora de imágenes tanto generales como con poca luz. Los enfoques notables incluyen redes inspiradas en Retinex, redes de codificador-decodificador y redes adversarias generativas (GAN). Estas técnicas han demostrado resultados prometedores en la mejora de la calidad de las imágenes.

* Work supported by ...

[†] Universidad Nacional San Agustín de Arequipa, fhuancat@unsa.edu.pe

[‡] Universidad Nacional San Agustín de Arequipa, jperezma@unsa.edu.pe

[§] Universidad Nacional San Agustín de Arequipa, hariasm@unsa.edu.pe

MÉTODO PROPUESTO

BLOQUE RESIDUAL MULTIESCALA

General Layout

These instructions are a typical implementation of the requirements. Manuscripts should have:

- Either A4 (21.0 cm × 29.7 cm; 8.27 in × 11.69 in) or US letter size (21.6 cm × 27.9 cm; 8.5 in × 11.0 in) paper.
- Single-spaced text in two columns of 82.5 mm (3.25 in) with 5.3 mm (0.2 in) separation. More recent versions of Microsoft Word have a default spacing of 1.5 lines; authors must change this to 1 line.
- The text located within the margins specified in Table 1.

Table 1: Margin Specifications

Margin	A4 Paper	US Letter Paper
Top	37 mm (1.46 in)	0.75 in (19 mm)
Bottom	19 mm (0.75 in)	0.75 in (19 mm)
Left	20 mm (0.79 in)	0.79 in (20 mm)
Right	20 mm (0.79 in)	1.02 in (26 mm)

Fonts

In order to produce good Adobe Acrobat PDF files, authors using the ‘jacow’ L^AT_EX template are asked to use only the fonts defined in the ‘jacow’ class file (v2.3 of 2019/01/15) in standard, bold (i. e., `\textbf`) or italic (i. e., `\textit`) form and symbols from the standard set of fonts. In Word use only Symbol and, depending on your platform, Times or Times New Roman fonts in standard, bold or italic form.

The layout of the text on the page is illustrated in Fig. ?? . Note that the paper’s title and the author list should be the width of the full page. Tables and figures may span the whole 170 mm page width, if desired (see Fig. ??), but if they span both columns, they should be placed at either the top or bottom of a page to ensure proper flow of the text (which should flow from top to bottom in each column).

Title and Author List

The title should use 14 pt bold uppercase letters and be centred on the page. Individual letters may be lowercased to avoid misinterpretation (e. g., mW, GeV, SPring-8, Swiss-FEL). To include a funding support statement, put an asterisk after the title and the support text at the bottom of the first column on page 1—in Word, use a text box; in L^AT_EX, use `\thanks`. See also the subsection on footnotes.

The names of authors, their organizations/affiliations and postal addresses should be grouped by affiliation and listed in 12 pt upper- and lowercase letters. The name of the submitting or primary author should be first, followed by the coauthors, alphabetically by affiliation. Where authors have multiple affiliations, the secondary affiliation may be indicated with a superscript, as shown in the author listing of this paper. See ANNEX A for further examples.

Section Headings

Section headings should not be numbered. They should use 12 pt bold uppercase letters and be centred in the column. All section headings should appear directly above the text—there should never be a column break between a heading and the following paragraph.

Subsection Headings

Subsection headings should not be numbered. They should use 12 pt italic letters and be left aligned in the column. Subsection headings use Title Case (or Initial Caps) and should appear directly above the text—there should never be a column break between a subheading and the following paragraph.

Third-level Headings These should use 10 pt bold letters and be run into the paragraph text. In L^AT_EX these headings are created with L^AT_EX’s `\subsubsection` command. In the Word templates authors must bold the heading text themselves. This heading should be used sparingly. See Table 2 for its style details.

Paragraph Text

Paragraphs should use 10 pt font and be justified (touch each side) in the column. The beginning of each paragraph should be indented approximately 0.33 cm (0.13 in). The last line of a paragraph should not be printed by itself at the beginning of a column nor should the first line of a paragraph be printed by itself at the end of a column.

Figures, Tables and Equations

Place figures and tables as close to their place of mention as possible. Lettering in figures and tables should be large enough to reproduce clearly. Use of non-approved fonts in figures can lead to problems when the files are processed. L^AT_EX users should be sure to use non-bitmapped versions of Computer Modern fonts in equations (Type 1 PostScript or OpenType fonts are required. Their use is described in the help pages of the JACoW website [1]).

Each figure and table must be numbered in ascending order (1, 2, 3, etc.) throughout the paper. Figure captions are placed below figures, and table captions are placed above tables.

Figure captions are formatted as shown in Figs. ?? and ??, while table captions take the form of a heading, with initial letters of principle words, capitalized, and without a period at the end (see Tables 1 and 2). Any reference to the contents of the table should be made from the body of the paper rather than from within the table caption itself.

Single-line captions are centred in the column, while captions that span more than one line should be justified. The L^AT_EX template uses the ‘booktabs’ package to format tables.

When referring to a figure from within the text, the convention is to use the abbreviated form [e. g., Fig. 1] *unless* the reference is at the start of the sentence, in which case

“Figure” is written in full. Reference to a table, however, is never abbreviated [e. g., Table 1].

If a displayed equation needs a number (i. e., it will be referenced), place it in parentheses, and flush with the right margin of the column. The equation itself should be indented and centred, as far as is possible:

$$C_B = \frac{q^3}{3\epsilon_0 mc} = 3.54 \mu\text{eV/T} \quad (1)$$

When referencing a numbered equation, use the word “Equation” at the start of a sentence, and the abbreviated form, “Eq.”, if in the text. The equation number is placed in parentheses [e. g., Eq. (1)], which can be achieved in L^AT_EX using `\eqref{eq:label}`.

Units

Units should be written using the standard, roman font, not the italic font, as shown in Eq. (1). An unbreakable space should precede a unit (in L^AT_EX use a “\,”, the template uses the ‘siunitx’ package to format units). Some examples are: 3 keV, 100 kW, 7 μm. When a unit appears in a hyphenated, compound adjective that precedes a noun, it takes on the singular form, e. g., the 3.8-metre long undulator.

References

All bibliographical and web references should be numbered and listed at the end of the paper in a section called **REFERENCES**. When citing a reference in the text, place the corresponding reference number in square brackets [1]. The reference citations in the text should be numbered in ascending order. Multiple citations should appear in the same bracket [3, 4] and with ranges where appropriate [1–4, 10].

A URL or DOI may be included as part of a reference, but its hyperlink should NOT be added. The usual practice is to use a monospaced font for the URL so as to help distinguish it from normal text. In L^AT_EX the ‘url’ package is used with its default font now being switched to “newtxtt” which offers a better distinction between “0” and “O”.

For authors to properly cite the resources used when researching their papers is an obligation. In the interest of promoting uniformity and complete citations, the IEEE Editorial Style for Transactions and Journals, which itself adheres to the Chicago Manual of Style, has been adopted [2]. When citing a periodical, the official abbreviation of the journal should be used [3]. Please consult the appended material, **ANNEX B**, for details. The onus is on authors to pay attention to the details of the said style to ensure complete, accurate and properly formatted references.

Footnotes

Footnotes on the title and author lines may be used for acknowledgements and e-mail addresses. A non-numeric sequence of characters (*, #, †, ‡, ¶) should be used to indicate the footnote. These “pseudo footnotes” should only appear at the bottom of the first column on the first page.

Any other footnote in the body of the paper should use the normal numeric sequencing (i. e., 1, 2, 3) and appear at the bottom of the same column in which it is used. All footnotes are of 8pt font size.

Acronyms

Acronyms should be defined the first time they appear, both in the abstract and in the rest of the paper.

STYLES

Table 2 summarizes the fonts and spacing used in the styles of a JACoW template. In L^AT_EX, these are implemented in the ‘jacow’ class file.

PAGE NUMBERS

DO NOT include any page numbers. They will be added when the final proceedings are produced.

Table 2: Summary of Styles

Style	Font	Space Before	Space After
PAPER TITLE	14 pt UPPERCASE EXCEPT FOR REQUIRED lowercase letters Bold	0 pt	3 pt
Author list	12 pt UPPER- and lowercase	9 pt	12 pt
<i>Abstract Title</i>	12 pt <i>Initial Caps Italic</i>	0 pt	3 pt
Section Heading	12 pt UPPERCASE bold	9 pt	3 pt
<i>Subsection Heading</i>	12 pt <i>Initial Caps Italic</i>	6 pt	3 pt
Third-level Heading	10 pt Initial Caps Bold	6 pt	0 pt
Figure Captions	10 pt	3 pt	≥3 pt
Table Captions	10 pt	≥3 pt	3 pt
Equations	10 pt base font	≥6 pt	≥6 pt
References when ≤ 9	9 pt <code>\bibliography{9}</code>	0 pt	3 pt
Refs. ≥ 10	9 pt <code>\bibliography{99}</code>	0 pt	3 pt

TEMPLATES

Template documents for the recommended word processing software are available from the JACoW website [1] and exist for L^AT_EX, Microsoft Word (Mac and PC) and LibreOffice/Apache OpenOffice for US letter and A4 paper sizes. Use the correct template for your paper size and platform.

Fonts are embedded by default with pdfL^AT_EX. Using L^AT_EX with ‘dvips’, make sure that ‘ps2pdf’ has the option `-dEmbedAllFonts=true`. Fonts of included figure graphics in PDF or EPS are often not embedded. So make sure that this is done when generating them or reprocess them in ‘Ghostscript’ with the switch `-dEmbedAllFonts=true` set.

CHECKLIST FOR ELECTRONIC PUBLICATION

Authors are requested to go over the following checklist for electronic publication:

- Use only Times or Times New Roman (standard, bold or italic) and Symbol fonts for text, 10 pt except references, which should be 9 pt.
- Figures should use Times or Times New Roman (standard, bold or italic) and Symbol fonts when possible—6 pt minimum, with fonts embedded.
- Check that citations to references appear in sequential order and that all references are cited.
- Check that the PDF file prints correctly.
- Check that there are no page numbers.
- Check that the margins on the printed version are within ± 1 mm of the specifications.
- L^AT_EX users can check their margins by invoking the `boxit` option.

Please also check the list of common oversights which can be found in ANNEX C.

CONCLUSION

La restauración de imágenes es una tarea importante en vigilancia, fotografía computacional, imágenes médicas y teledetección. Recientemente, las redes neuronales convolucionales (CNN) han logrado mejoras dramáticas sobre los enfoques convencionales. Este artículo presenta una arquitectura novedosa con los objetivos colectivos de mantener representaciones espacialmente precisas de alta resolución y recibir información contextual sólida de las representaciones de baja resolución. El núcleo del enfoque es un bloque residual multiescala que contiene varios elementos clave: flujos paralelos de convolución multiresolución, intercambio de información a través de los flujos multiresolución, mecanismos de atención espacial y de canales, y agregación de características multiescala basada en la atención. Extensos experimentos en conjuntos de datos de referencia de imágenes reales demuestran que MIRNet logra resultados de vanguardia para varias tareas de procesamiento de imágenes.

APPENDIX

Any appendix should be in a separate section directly preceding the **REFERENCES** section. If there is no **REFERENCES** section, this should be the last section of the paper.

REFERENCES

- [1] JACoW, <http://www.jacow.org>
- [2] *IEEE Editorial Style Manual*, IEEE Periodicals, Piscataway, NJ, USA, Oct. 2014, pp. 34–52.
- [3] <https://woodward.library.ubc.ca/researchhelp/journal-abbreviations/>