

**© 2019 - Đại Học Khoa Học Tự Nhiên, Đại Học Quốc Gia Hà Nội**

**Khoa Toán - Cơ - Tin Học**

technIcal report

**PES-TR??**

**Đại học Khoa Học Tự Nhiên**

**12 2019**

**Báo cáo Xử Lý Ảnh**

THIS PAGE LEFT BLANK INTENTIONALLY

**(GROUP NAME THAT WROTE THE REPORT)**

Đề tài

**DeepWriter:**

**Hệ thống DeepCNN nhận diện tác giả,**

**độc lập với nội dung văn bản**

Thành viên

Đỗ Tất Thành

Lê Văn Anh Tú

**ACKNOWLEDGMENTS (Optional)**

The following are examples of acknowledgments.

The Task Force is truly grateful for the support of our sponsoring subcommittee and committee.

The Task Force gratefully acknowledges the contributions of T. Edison, G. Westinghouse, N. Tesla, A. Volta and A. Ampere to the electric power industry.

The Task Force gratefully acknowledges the principal authors/contributors of the following sections:

* Section 1: T. Edison
* Section 2: G. Westinghouse and N. Tesla
* Section 3: A. Ampere, N. Tesla and A. Volta
* Appendix A: T. Edison

**KEYWORDS**

Provide up to 10 keywords (in alphabetical order) to help identify the major topics of the paper.

**CONTENTS**

[1. INTRODUCTION 1](#_Toc440176258)

[2. EASE OF USE 1](#_Toc440176259)

[2.1 Template 1](#_Toc440176260)

[2.2 Maintaining the Integrity of the Specifications 1](#_Toc440176261)

[3. TECHNICAL REPORT PREPARATION 1](#_Toc440176262)

[3.1 Abbreviations and Acronyms 2](#_Toc440176263)

[3.2 Units 2](#_Toc440176264)

[3.3 Equations 2](#_Toc440176265)

[3.4 Footnotes 3](#_Toc440176266)

[3.5 Some Common Mistakes 3](#_Toc440176267)

[4. USING THE TEMPLATE 3](#_Toc440176268)

[4.1 Identify the Headings 4](#_Toc440176269)

[4.1.1 Members and Contributors 4](#_Toc440176270)

[4.2 Figures and Tables 4](#_Toc440176271)

[4.2.1 Figures 4](#_Toc440176272)

[4.2.2 Tables 4](#_Toc440176273)

[4.2.3 Positioning Figures and Tables 5](#_Toc440176274)

[5. REFERENCES 5](#_Toc440176275)

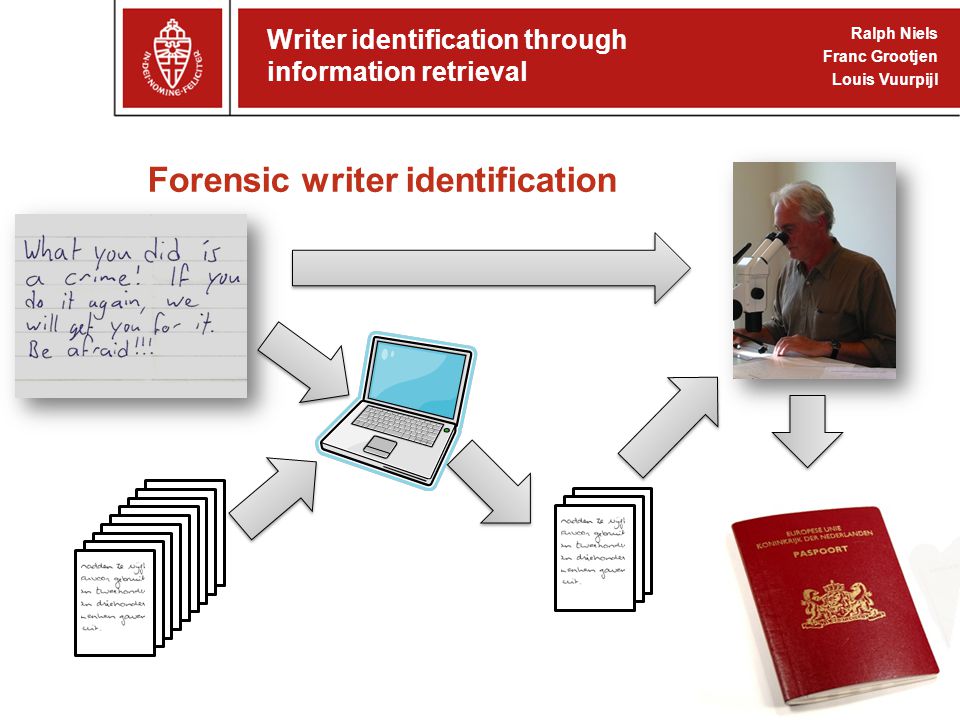
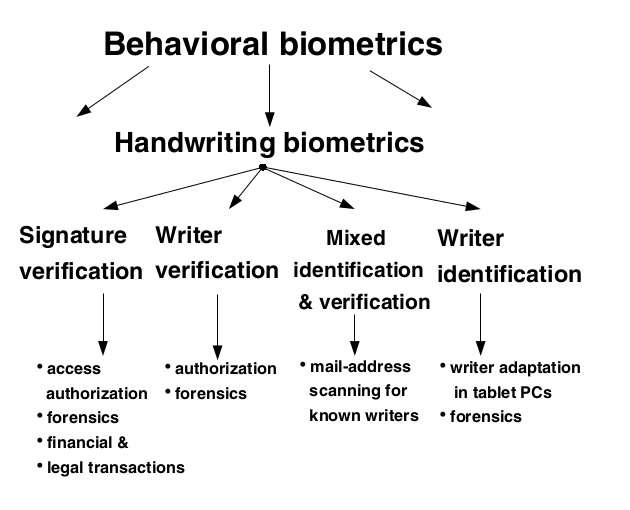
[APPENDIX A SAMPLE HEADING 8](#_Toc440176276)

[A.1 Sample Heading 8](#_Toc440176277)

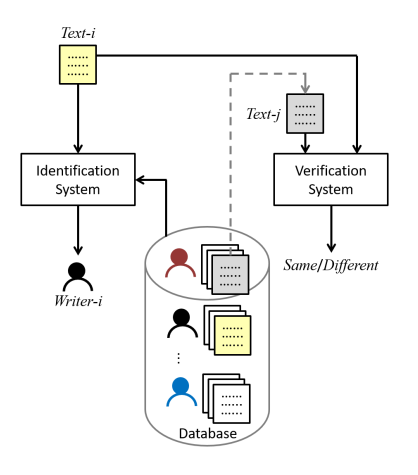
[A.1.1 Sample Heading 8](#_Toc440176278)

THIS PAGE LEFT BLANK INTENTIONALLY

# GIỚI THIỆU ĐỀ TÀI

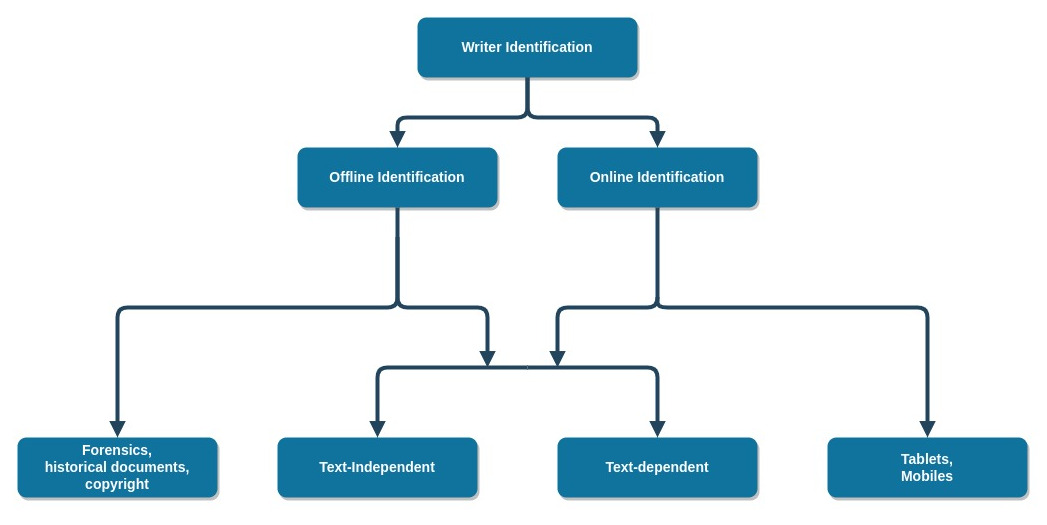
Bài toán được đưa ra để giải quyết là một bài toán về nhận diện tác giả thông qua hình ảnh về văn bản viết tay. Chữ viết tay là một loại dữ liệu định dạng cá nhân dựa trên hành vi. Tác giả sẽ được nhận diện thông qua việc phát hiện những đặc trưng đặc biệt của tác giả ẩn trong chữ viết tay, mà những người viết khác không có. Nhận diện tác giả thông qua chữ viết tay có ứng dụng lâu đời trong khoa học

hình sự, được tính là bằng chứng trong điều tra và xét xử. Ngoài ra, nhận diện tác giả thông qua chữ viết tay cũng có ứng dụng trong nghiên cứu lịch sử, nhằm tra soát nguồn gốc tác giả của những tài liệu lịch sử. Bài toán cũng tương tự như một số bài toán khác như xác thực thông qua chữ viết tay hay chữ ký.



Bài toán nhận diện tác giả thông qua chữ viết tay thường được chia làm hai hướng chính: Online Identification và Offline Identification. Online Identification là bài toán nhận diện khi thông tin có được từ người viết đang trong quá trình viết, trên các thiết bị đặc biệt có thể lưu lại thông tin về quá trình viết (máy tính bảng...). Khi đó, các thông tin có thể thu thập được ngoài bản chữ viết tay còn có

thông tin theo thời gian của vị trí đầu bút, tốc độ di chuyển, cách viết nét cong, lực ghì, hướng... Offline Identification là bài toán nhận diện chữ viết tay khi chỉ có hình ảnh về văn bản chữ viết tay, thiếu đi thông tin về quá trình viết, do đó thường khó hơn Online Identification. Bài toán nhận diện tác giả được tìm hiểu là bài toán Offline Identification. Trong Offline Identification lại được chia làm nhiều hướng chính dựa theo đặc điểm của phương pháp phân tích: Text-dependent và Text-indipendent. Text-dependent phụ thuộc vào nội dung văn bản, thường là cố định về nội dung, ví dụ như bản khai theo mẫu cho các văn bản hành chính. Ở chiều ngược lại, Text-Indipendent không phụ thuộc vào nội dung văn bản cố định, do đó có khả năng áp dụng rộng rãi hơn. Bài toán đang được tìm hiểu ở đây thuộc về loại Text-Indipendent.



# DeepWriter: Hệ thống DeepCNN nhận diện tác giả, độc lập với nội dung văn bản

## Template (Heading 2)

This template has been tailored for output on US letter-sized paper.

## Maintaining the Integrity of the Specifications (Heading 2)

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the heading margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

# TECHNICAL REPORT PREPARATION (Heading 1)

Please use automatic hyphenation and check your spelling. Additionally, be sure your sentences are complete and that there is continuity within your paragraphs. Check the numbering of your graphics (figures and tables) and make sure that all appropriate references are included. Please take note of the following items when proofreading spelling and grammar.

## Abbreviations and Acronyms (Heading 2)

Define abbreviations and acronyms the first time they are used in the text. Abbreviations such as IEEE, SI, ac, dc, and rms do not have to be defined. Do not use abbreviations in the title or section headings unless they are unavoidable.

## Units (Heading 2)

Metric units are preferred for use in IEEE publications in light of their global readership and the inherent convenience of these units in many fields. In particular, the use of the International System of Units (Systeme Internationale d’Unites or SI Units) is advocated. This system includes a subsystem of units based on the meter, kilogram, second, and ampere (MKSA). U.S. Customary units, or British units, may be used as secondary units (in parentheses). An exception is when U.S. Customary units are used as identifiers in trade, such as 3.5-inch disk drive.

Avoid combining SI and U.S. Customary units, such as current in amperes and magnetic field in oersteds. This often leads to confusion because equations do not balance dimensionally. If you must use mixed units, clearly state the units for each quantity that you use in an equation.

Do not mix complete spellings and abbreviations of units: “Wb/m2” or “webers per square meter,” not “webers/m2.” Spell out units when they appear in text: “. . . a few henries,” not “. . . a few H.”

Use a zero before decimal points: “0.25,” not “.25.” Use “cm3,” not “cc.”

## Equations (Heading 2)

The equations are an exception to the prescribed specifications of this template. You will need to determine whether or not your equation should be typed using either the Times New Roman or the Symbol font (please no other font). To create multileveled equations, it may be necessary to treat the equation as a graphic and insert it into the text after your paper is styled. Use of the Microsoft Equation Editor or the *MathType* commercial add-on for MS Word for math objects in your paper is permissible (Insert | Equation *or* MathType Equation). “Float over text” should *not* be selected. Number equations consecutively. Equation numbers, within parentheses, are to position flush right, as in (1), using a right tab stop. To make your equations more compact, you may use the solidus   
( / ), the exp function, or appropriate exponents. Italicize Roman symbols for quantities and variables, but not Greek symbols. Use a long dash, as shown in (1), rather than a hyphen for a minus sign. Punctuate equations with commas or periods when they are part of a sentence, as in

−δ 

Note that the equation above is centered using a center tab stop. Be sure that the symbols in your equation have been defined before or immediately following the equation. Use “(1),” not “Eq. (1)” or “equation (1),” except at the beginning of a sentence: “Equation (1) is . . . .”

## **Footnotes** (Heading 2)

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the page in which it was cited. Do not put footnotes in the reference list. Use letters for table footnotes.

## Some Common Mistakes (Heading 2)

* The word “data” is plural, not singular.
* The subscript for the permeability of vacuum μ0, and other common scientific constants, is zero with subscript formatting, not a lowercase letter “o.”
* In American English, commas and periods are located inside quotation marks; semicolons and colons are located outside quotation marks. Question and exclamation marks are located within quotation marks only when they are part of the quote. A parenthetical phrase or statement at the end of a sentence is punctuated outside of the closing parenthesis (like this). (A parenthetical sentence is punctuated within the parentheses.)
* A graph within a graph is an “inset,” not an “insert.” The word “alternatively” is preferred to the word “alternately” (unless you really mean something that alternates).
* Do not use the word “essentially” to mean “approximately” or “effectively.”
* In the title, if the words “that uses” can accurately replace the word “using,” capitalize the “u”; if not, keep “using” lower-cased.
* Be aware of the different meanings of the homophones “affect” and “effect,” “complement” and “compliment,” “discreet” and “discrete,” “principal” and “principle.”
* Do not confuse “imply” and “infer.”
* The prefix “non” is not a word; it should be joined to the word it modifies, usually without a hyphen.
* There is no period after the “et” in the Latin abbreviation “et al.”
* The abbreviation “i.e.” means “that is,” and the abbreviation “e.g.” means “for example.”

# USING THE TEMPLATE (Heading 1)

This document should be used as a template for preparing a PES Technical Report. You may type over sections of the document, cut and paste into it, and/or use markup styles.

Duplicate the template file by using the Save As command.

## Identify the Headings (Heading 2)

Headings are organizational devices that guide the reader through your paper. There are two types: component headings and text headings.

Component headings identify the different components of your paper and are not topically subordinate to each other. Examples include Acknowledgments and References and, for these, the correct style to use is “Heading 1.” Use “Figure caption” for your figure captions, and “table heading” for your table title. Run-in headings will require you to apply a style (such as italic) to differentiate the heading from the text.

Text headings organize the topics on a relational, hierarchical basis. For example, the paper title is the primary text heading because all subsequent material relates and elaborates on this one topic. If there are two or more sub-topics, the next level heading should be used and, conversely, if there are not at least two sub-topics, then no subheadings should be introduced. Styles named “Heading 1,” “Heading 2,” and “Heading 3” are prescribed.

### Members and Contributors (Heading 3)

The title of the group that prepared the document, including the Chair, Co-Chair(s), Editor(s), Members and/or Contributors shall be listed on page iii.

## Figures and Tables (Heading 2)

### Figures (Heading 3)

* Figures should be numbered consecutively using Arabic numerals.
* Use bold 12 point Times New Roman for figure captions.
* Use words rather than symbols or abbreviations when writing figure axis labels to avoid confusing the reader. As an example, write the quantity “Magnetization,” or “Magnetization, M,” not just “M.”
* If including units in the label, present them within parentheses.
* Do not label axes only with units. In the example, write “Magnetization (A/m)” or “Magnetization {A[m(1)]},” not just “A/m.”
* Do not label axes with a ratio of quantities and units. For example, write “Temperature (K),” not “Temperature/K.”

### Tables (Heading 3)

* Tables should be numbered consecutively using Arabic numerals.
* Use bold 12 point Times New Roman for table titles (labels).
* Use bold 10 point Times New Roman for table headings and subheadings.
* Use 10 point Times New Roman for text (table copy) within the table.
* Use words rather than symbols or abbreviations when writing table headings/subheadings to avoid confusing the reader. As an example, write “Transient Recovery Voltage,” not just “TRV.” Do not label headings/ subheadings only with units. If abbreviations must be used, define the abbreviation as a footnote to the table or in the text immediately preceding the table.
* If including units in the heading/subheading, present them within parentheses. For example, “Temperature (K).”

### Positioning Figures and Tables (Heading 3)

* Figures and tables should be centered in the page (see Table 1 and Fig. 1).
* Figure captions should be centered below the figures.
* Table labels should be centered above the tables.
* Insert figures and tables after they are cited in the text as close to the citation as possible.
* Use the abbreviation “Fig. 1,” even at the beginning of a sentence.

TABLE 1. Table Type Styles

| Table Heading | Table Column Heading | | |
| --- | --- | --- | --- |
| Table column subheading | Subheading | Subheading |
| copy | More table copya |  |  |

a. Example of a Table footnote.

We suggest that you use a text box to insert a graphic (which is ideally a 300 dpi TIFF or EPS file, with all fonts embedded) because, in an MSW document, this method is somewhat more stable than directly inserting a picture.

To have non-visible rules on your frame, use the MSWord “Format” pull-down menu, select Shape Outline > No Outline.

Fig. 1. Example of a figure caption.

# REFERENCES (Heading 1)

Linjie Xing, Yu Qiao, “Deepwriter: a multi-stream deep cnn for text-independent writer identification”, Conference: 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR), October 2016

<https://en.wikipedia.org/wiki/Vanishing_gradient_problem>

Vinod Nair, Geoffrey E. Hinton, “Rectified linear units improve restricted boltzmann machines”, ICML'10 Proceedings of the 27th International Conference on International Conference on Machine Learning, 807-814, 2010-06-21

Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, Ruslan Salakhutdinov, “Dropout: a simple way to prevent neural networks from overfitting”, Journal of Machine Learning Research 15 (2014) 1929-1958, 2014

Jason Yosinski, Jeff Clune, Yoshua Bengio, and Hod Lipson, “How transferable are features in deep neural networks?”, Advances in Neural Information Processing Systems 27 (NIPS 2014), 2014

Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton, "Imagenet classification with deep convolutional neural networks", 2012 ILSVRC (ImageNet Large-Scale Visual Recognition Challenge), 2012

Matthew D. Zeiler, Rob Fergus, "Visualizing and understanding convolutional networks", European Conference on Computer Vision 2014, pp 818-833, 2014

Karen Simonyan, Andrew Zisserman, "Very deep convolutional networks for large-scale image recognition", ICLR 2015, 2015

Alex Krizhevsky, "One weird trick for parallelizing convolutional neural networks", Google Inc., April 29, 2014

Arshia Rehman, Saeeda Naz, Muhammad Imran Razzak, Ibrahim A. Hameed, "Automatic visual features for writer identification: a deep learning approach", IEEE Access ( Volume: 7 ), pg 17149 - 17157, 21 January 2019, <https://doi.org/10.1109/ACCESS.2018.2890810>

Connor Shorten, Taghi M. Khoshgoftaar, "A survey on image data augmentation for deep learning", T.M. J Big Data (2019) 6: 60. <https://doi.org/10.1186/s40537-019-0197-0>

Zhun Zhong, Liang Zheng, Guoliang Kang, Shaozi Li, Yi Yang, “Random erasing data augmentation”, 2017, arXiv:1708.04896

See the footnotes below for samples of the correct formats for the following types of references:

* periodicals[[1]](#footnote-0),[[2]](#footnote-1)
* books[[3]](#footnote-2),[[4]](#footnote-3),[[5]](#footnote-4)
* technical reports[[6]](#footnote-5),[[7]](#footnote-6)
* unpublished papers (even if they have been submitted for publication)[[8]](#footnote-7),[[9]](#footnote-8),[[10]](#footnote-9)
* papers published in translation journals[[11]](#footnote-10) (Please give the English citation first, followed by the original foreign-language citation.)
* papers that are in press (papers accepted for publication, but not yet published)[[12]](#footnote-11)
* papers from conference proceedings (published)[[13]](#footnote-12)
* dissertations[[14]](#footnote-13)
* standards[[15]](#footnote-14)
* patents[[16]](#footnote-15)

To create a footnote, place your cursor where the footnote citation should appear (in general, the footnote citation should follow the sentence punctuation) and click “Insert Footnote” in the Footnotes section of the References tab. Then insert the text of the footnote. Unless there are six authors or more, give all authors’ names; do not use et al. Capitalize only the first word in a paper title, except for proper nouns and element symbols. The footnotes will be numbered automatically.

1. SAMPLE HEADING (Heading A1)

Text

* 1. Sample Heading (Heading A2)

Text

* + 1. Sample Heading (Heading A3)

Text

* + - 1. Sample Heading (Heading A4)

Text

* + - 1. Sample Heading (Heading A4)

Text

1. J. F. Fuller, E. F. Fuchs, and K. J. Roesler, “Influence of harmonics on power distribution system protection,” *IEEE Trans. Power Delivery*, vol. 3, pp. 549–557, Apr. 1988. [↑](#footnote-ref-0)
2. R. J. Vidmar. (1992, Aug.). On the use of atmospheric plasmas as electromagnetic reflectors. *IEEE Trans. Plasma Sci.* [Online]. *21(3)*, pp. 876–880. Available: http://www.halcyon.com/pub/journals/  
   21ps03-vidmar [↑](#footnote-ref-1)
3. E. Clarke, *Circuit Analysis of AC Power Systems*, vol. I. New York: Wiley, 1950, p. 81. [↑](#footnote-ref-2)
4. G. O. Young, “Synthetic structure of industrial plastics,” in *Plastics*, 2nd ed., vol. 3, J. Peters, Ed. New York: McGraw-Hill, 1964, pp. 15–64. [↑](#footnote-ref-3)
5. J. Jones. (1991, May 10). *Networks*. (2nd ed.) [Online]. Available: http://www.atm.com [↑](#footnote-ref-4)
6. E. E. Reber, R. L. Mitchell, and C. J. Carter, “Oxygen absorption in the Earth's atmosphere,” Aerospace Corp., Los Angeles, CA, Tech. Rep. TR-0200 (4230-46)-3, Nov. 1968. [↑](#footnote-ref-5)
7. S. L. Talleen. (1996, Apr.). The Intranet Architecture: Managing information in the new paradigm. Amdahl Corp., Sunnyvale, CA. [Online]. Available: http://www.amdahl.com/doc/products/bsg/intra/  
   infra/html [↑](#footnote-ref-6)
8. D. Ebehard and E. Voges, “Digital single sideband detection for interferometric sensors,” unpublished, presented at the 2nd Int. Conf. Optical Fiber Sensors, Stuttgart, Germany, 1984. [↑](#footnote-ref-7)
9. Process Corp., Framingham, MA. “Intranets: Internet technologies deployed behind the firewall for corporate productivity,” unpublished. Presented at INET96 Annu. Meeting. [Online]. Available: http://home.process.com/Intranets/wp2.htp [↑](#footnote-ref-8)
10. **G. N. Lester and J. H. Nelson, “History of Circuit Breaker Standards,” unpublished. *Presented at the IEEE/PES*** *General Meeting*, 24 July 2008. [Online]. Available IEEE/PES Switchgear Committee web site: http://www.ewh.ieee.org/soc/pes/switchgear/Presentations/2008CBtutorial/speaker1paper.pdf [↑](#footnote-ref-9)
11. Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” *IEEE Transl. J. Magn. Japan*, vol. 2, pp. 740–741, Aug. 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982]. [↑](#footnote-ref-10)
12. E. H. Miller, “A note on reflector arrays,” *IEEE Trans. Antennas Propagat.,* in press. [↑](#footnote-ref-11)
13. J. L. Alqueres and J. C. Praca, “The Brazilian power system and the challenge of the Amazon transmission,” in *Proc. 1991 IEEE Power Engineering Society Transmission and Distribution Conf.*, pp. 315–320. [↑](#footnote-ref-12)
14. S. Hwang, “Frequency domain system identification of helicopter rotor dynamics incorporating models with time periodic coefficients,” Ph.D. dissertation, Dept. Aerosp. Eng., Univ. Maryland, College Park, 1997. [↑](#footnote-ref-13)
15. *IEEE Guide for Application of Shunt Power Capacitors*, IEEE Std. 1036-2010, Sep. 2010. [↑](#footnote-ref-14)
16. G. Brandli and M. Dick, “Alternating current fed power supply,” U.S. Patent 4 084 217, Nov. 4, 1978. [↑](#footnote-ref-15)