เทคนิคการเขียนบทความวิจัยระดับนานาชาติทางด้านคอมพิวเตอร์

โอฬาริก สุรินต๊ะ

คณะวิทยาการสารสนเทศ
มหาวิทยาลัยมหาสารคาม

โอฬาริก สุรินต๊ะ

ภาควิชาเทคโนโลยีสารสนเทศ
คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม

Multi-agent Intelligent Simulation Laboratory (MISL)

E-mail: olarik.s@msu.ac.th

Research Fields

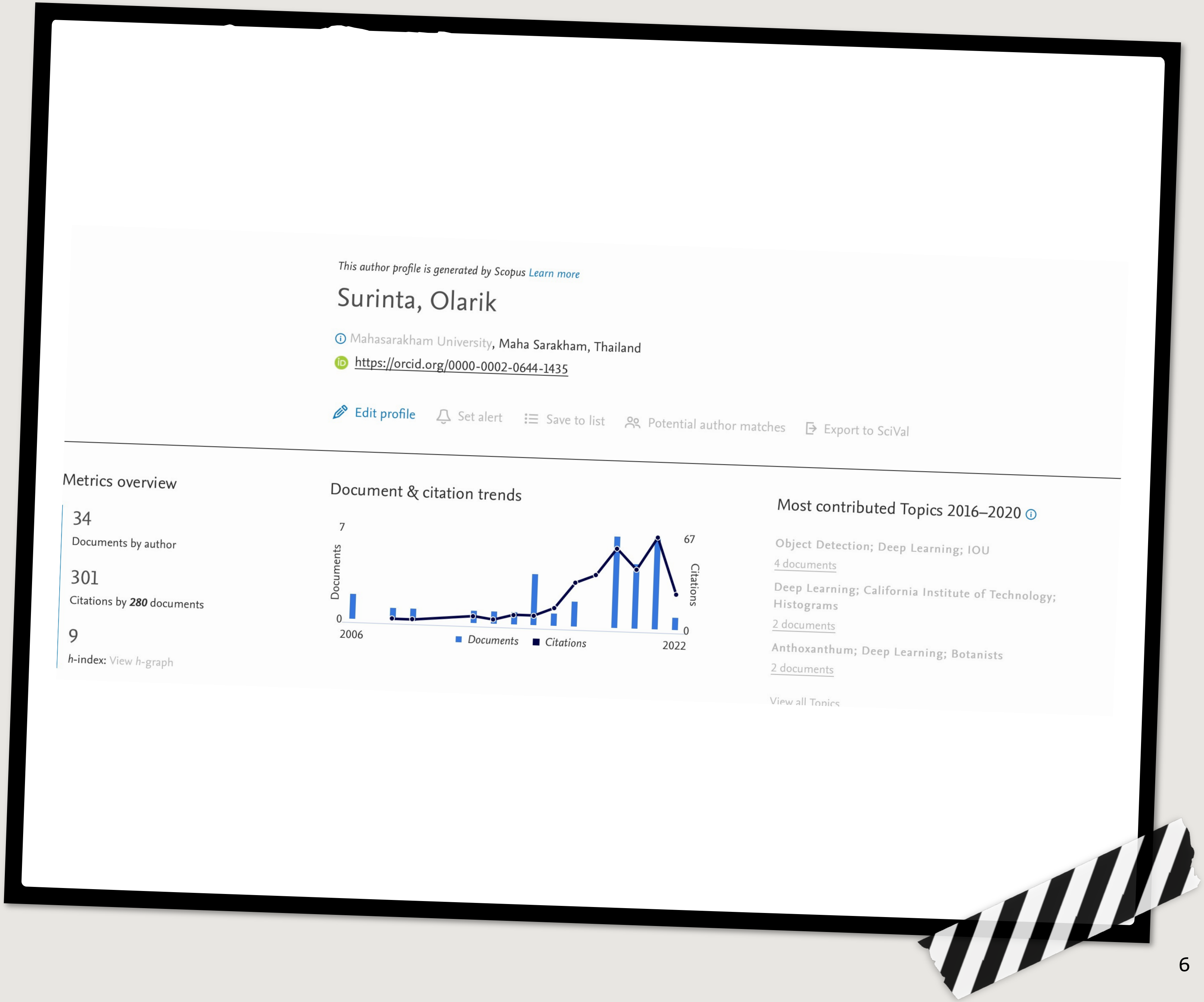
- Deep learning, Machine learning, Pattern recognition, Computer vision & Image processing, Handwritten word & character recognition, Document analysis, Word spotting, etc.

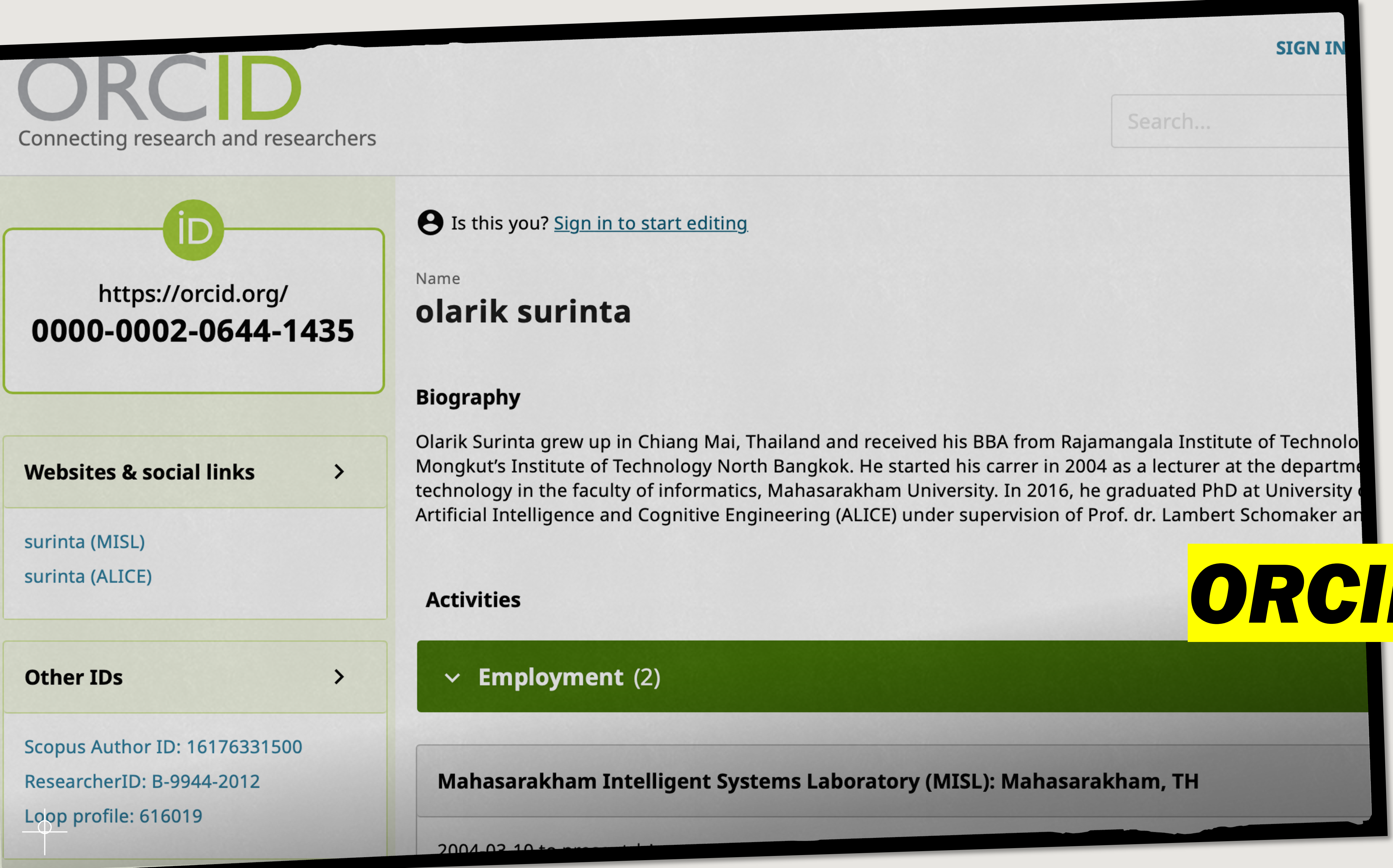
Feature Research

- Gonwirat, S. and **Surinta, O.** (2022). CycleAugment: Efficient Data Augmentation Strategy for Handwritten Text Recognition in Historical Document Images, Engineering and Applied Science Research, 49(4), 505-520. (Scopus Q4)
- Enkvetchakul, P. and **Surinta, O.** (2022). Effective Data Augmentation and Training Techniques for Improving Deep Learning in Plant Leaf Disease Recognition. Applied Science and Engineering Progress, 15(3), 1-12. (Scopus Q4)
- Singkhornart, T. and **Surinta, O.** (2022). Multi-Language Video Subtitle Recognition with Convolutional Neural Network and Long Short-Term Memory Networks, ICIC Express Letters, 16(6), 647-655. (Scopus Q3)
- Gonwirat, S. and **Surinta, O.** (2021). Optimal Weighted Parameters of Ensemble CNNs Based on a Differential Evolution Algorithm for Enhancing Pornographic Image Classification, Engineering and Applied Science Research, 48(5), 560-569. (Scopus Q3)
- Phiphitphatphaisit, S. and **Surinta, O.** (2021). Deep Feature Extraction Technique Based on Conv1D and LSTM Network for Food Image Recognition, Engineering and Applied Science Research, 48(5), 581-592. (Scopus Q3)

Publication Profile

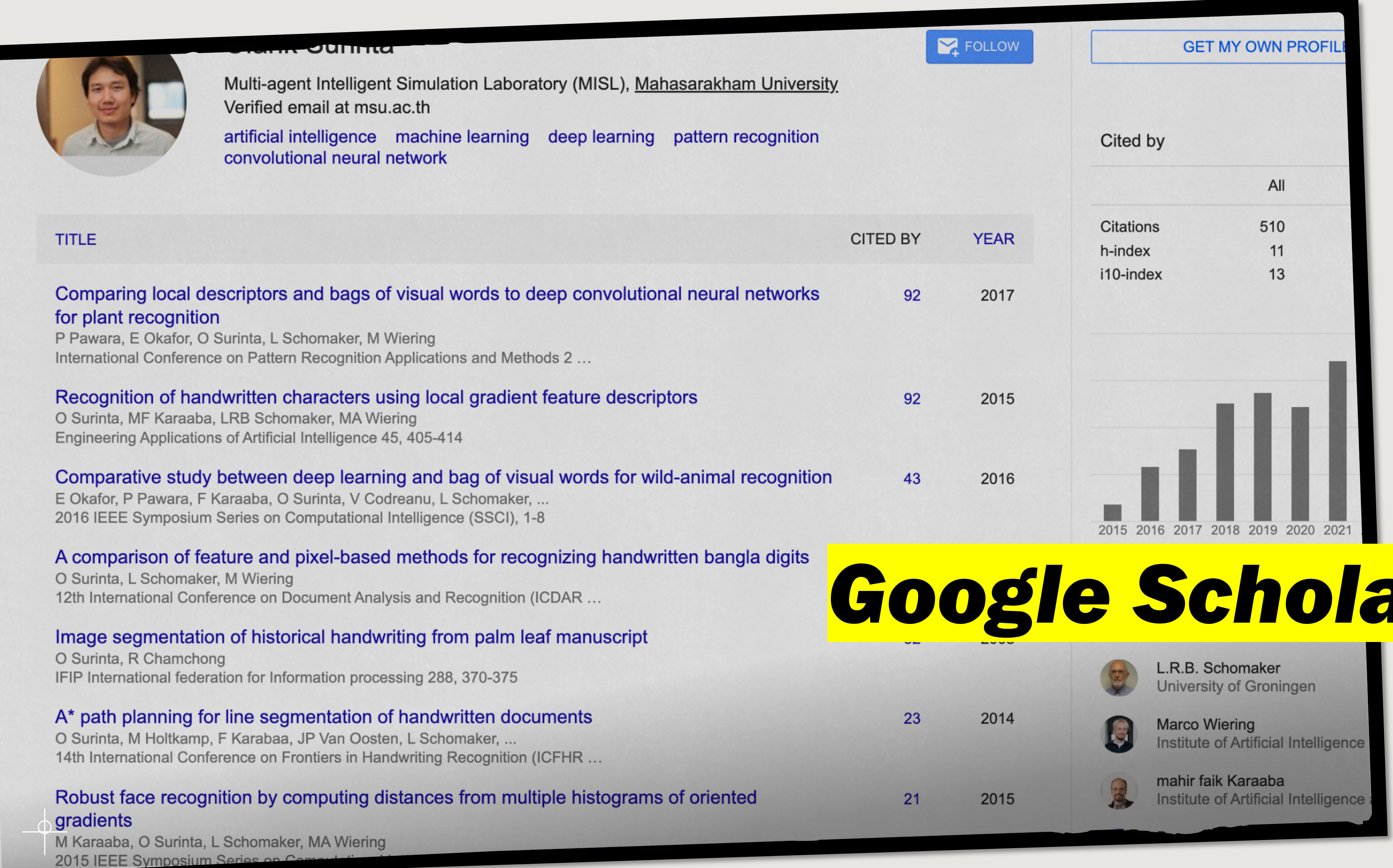
Scopus


This author profile is generated by Scopus Learn more
Surinta, Olarik
Mahasarakham University, Maha Sarakham, Thailand
https://orcid.org/0000-0002-0644-1435
Edit profile
Set alert
Save to list
Potential author matches
Export to SciVal
Metrics overview
34
Documents by author
301
Citations by 280 documents
9
h-index: View h-graph
Document & citation trends
Documents
Citations
7
0
67
0
2006
2022
Documents
Citations
Most contributed Topics 2016–2020
Object Detection; Deep Learning; IOU
4 documents
Deep Learning; California Institute of Technology; Histograms
2 documents
Anthoxanthum; Deep Learning; Botanists
2 documents



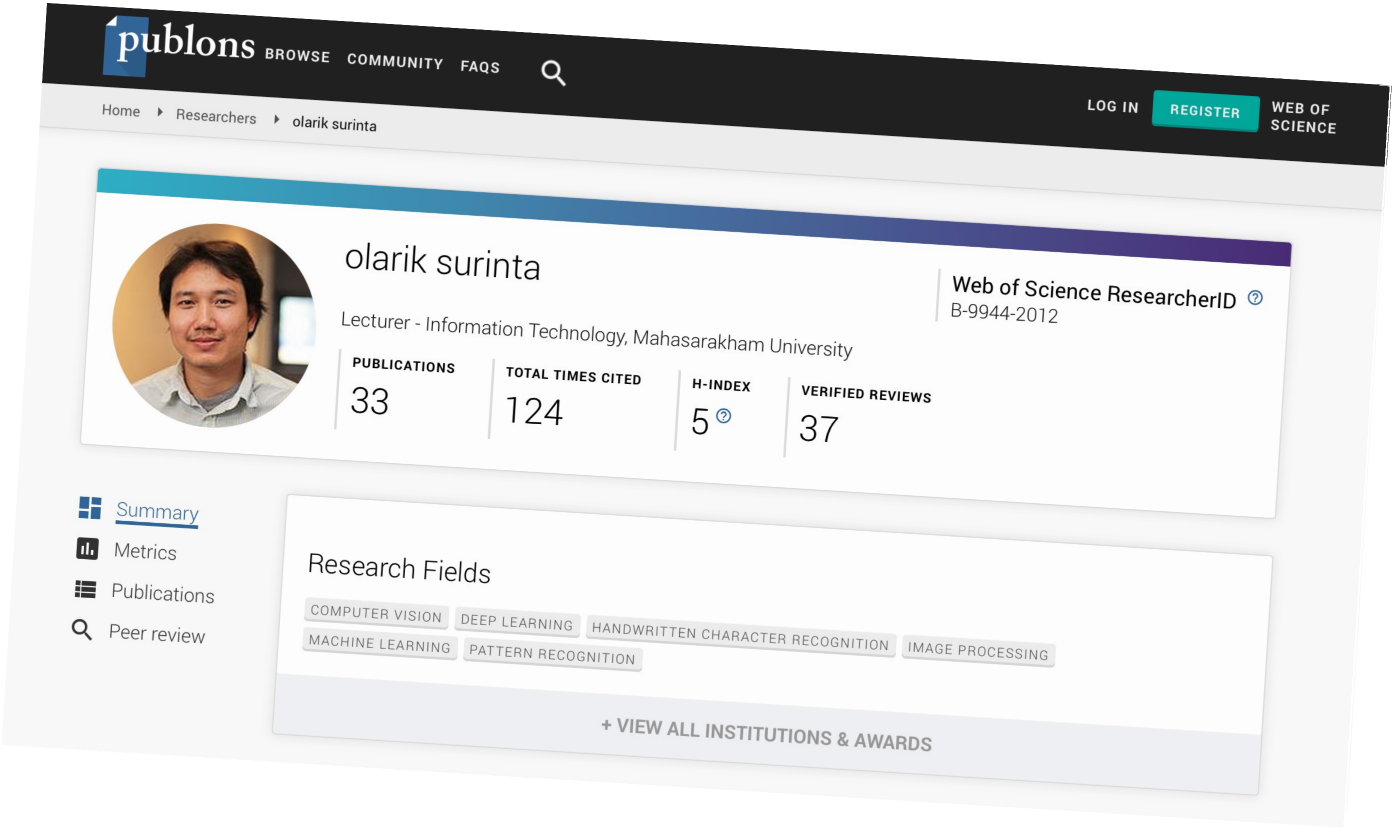
ORCID
Connecting research and researchers
SIGN IN
Search...
https://orcid.org/
0000-0002-0644-1435
Is this you? Sign in to start editing
Name
olarik surinta
Biography
Olarik Surinta grew up in Chiang Mai, Thailand and received his BBA from Rajamangala Institute of Technolo
Mongkut's Institute of Technology North Bangkok. He started his carrer in 2004 as a lecturer at the departme
technology in the faculty of informatics, Mahasarakham University. In 2016, he graduated PhD at University
Artificial Intelligence and Cognitive Engineering (ALICE) under supervision of Prof. dr. Lambert Schomaker an
Websites & social links
surinta (MISL)
surinta (ALICE)
Activities
Employment (2)
Other IDs
Scopus Author ID: 16176331500
ResearcherID: B-9944-2012
Loop profile: 616019
Mahasarakham Intelligent Systems Laboratory (MISL): Mahasarakham, TH

ORCID



Google Scholar

Publons



publons
BROWSE
COMMUNITY
FAQS
LOG IN
REGISTER
WEB OF SCIENCE
Home
Researchers
olarik surinta
olarik surinta
Lecturer - Information Technology, Mahasarakham University
Web of Science ResearcherID
B-9944-2012
PUBLICATIONS
33
TOTAL TIMES CITED
124
H-INDEX
5
VERIFIED REVIEWS
37
Summary
Metrics
Publications
Peer review
Research Fields
COMPUTER VISION
DEEP LEARNING
HANDWRITTEN CHARACTER RECOGNITION
IMAGE PROCESSING
MACHINE LEARNING
PATTERN RECOGNITION
+ VIEW ALL INSTITUTIONS & AWARDS

What are you looking for?

- Are you looking for publish a manuscript in the …
  - National / International Journal, called **Article**
  - National / International Conference, called **Proceeding** or Conference paper

วารสารวิชาทางวิชาการที่อยู่ในฐานข้อมูลตามประกาศ ก.พ.อ.

- **ฐานข้อมูลนานาชาติ**

+ ERIC (https://eric.ed.gov)

+ MathSciNet (http://www.ams.org/mathscinet)

+ Pubmed (http://www.ncbi.nlm.hih.gov/pubmed)

+ Scopus (http://scopus.com)

+ Web of Science (เฉพาะในฐานข้อมูล SCIE, SSCI และ AHCI เท่านั้น) (http://mjl.clarivate.com)

+ JSTOR (https://www.jstor.org)

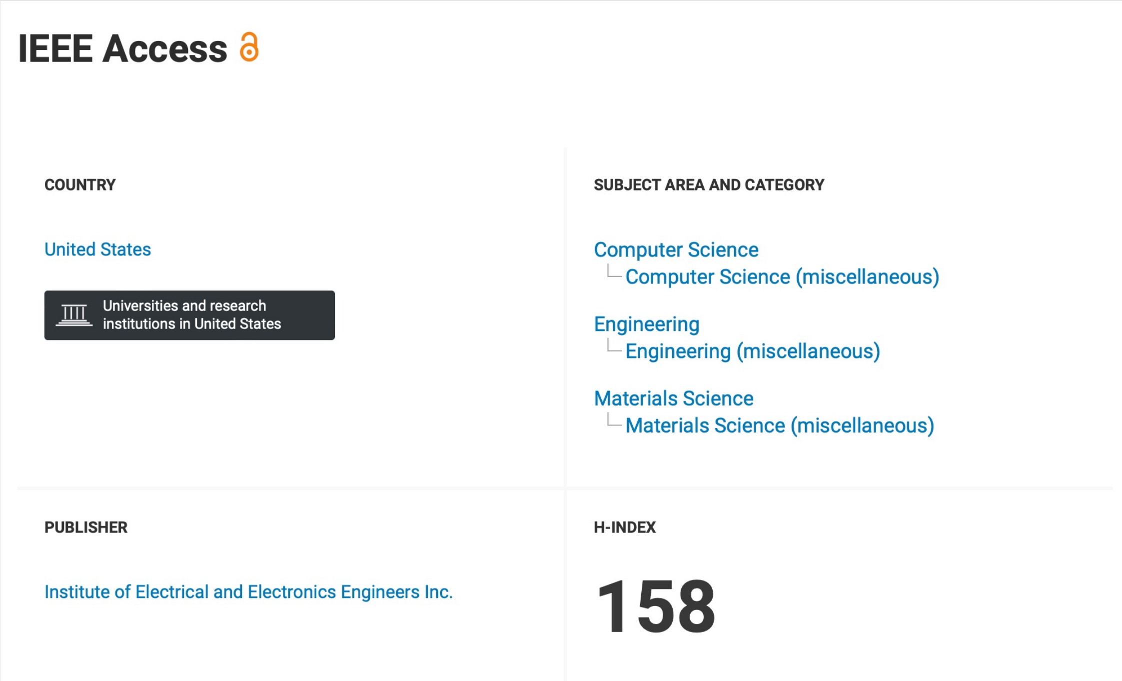
+ Project Muse (https://muse.jhu.edu)

+ รายชื่อวารสาร https://drive.google.com/file/d/1R-cWRgDTgUmMUO7Mc8H5yAVVA_FFxY8_/view

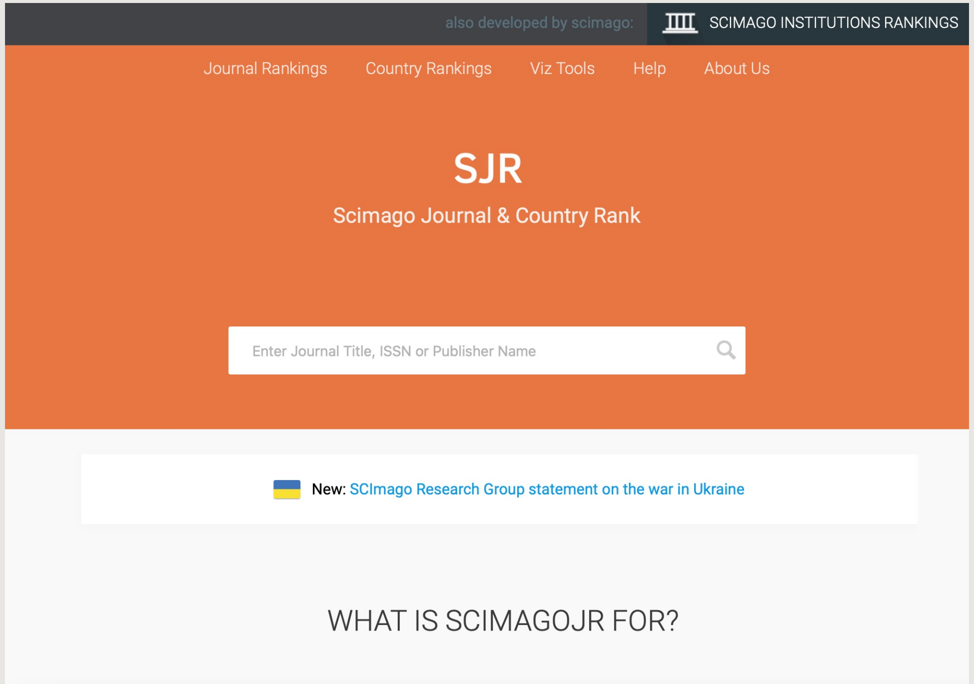
วารสารวิชาทางวิชาการที่อยู่ในฐานข้อมูลตาม ประกาศ ก.พ.อ.

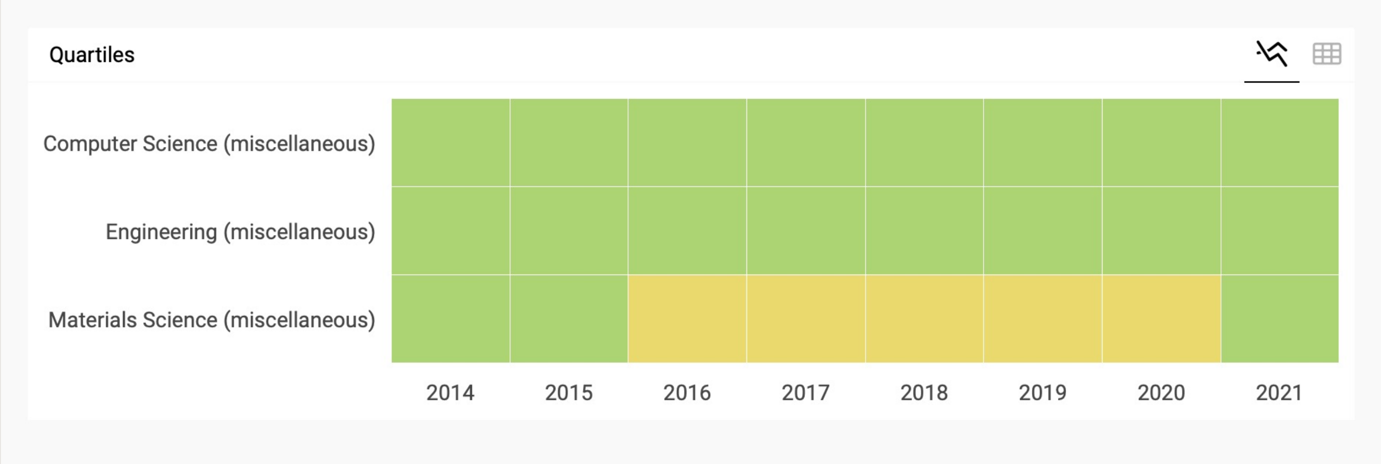
- **ฐานข้อมูลวารสารไทย**
  + ศูนย์ดัชนีการอ้างอิงวารสารไทย (Thai Journal Citation Index - TCI) – เฉพาะวารสารที่มีชื่ออยู่ในกลุ่มที่ 1 และ 2
    - กลุ่มชื่อวารสารที่อยู่ในฐาน TCI

สิ่งที่ต้องพิจารณาหากต้องการเผยแพร่บทความในวารสารวิชาการระดับนานาชาติ



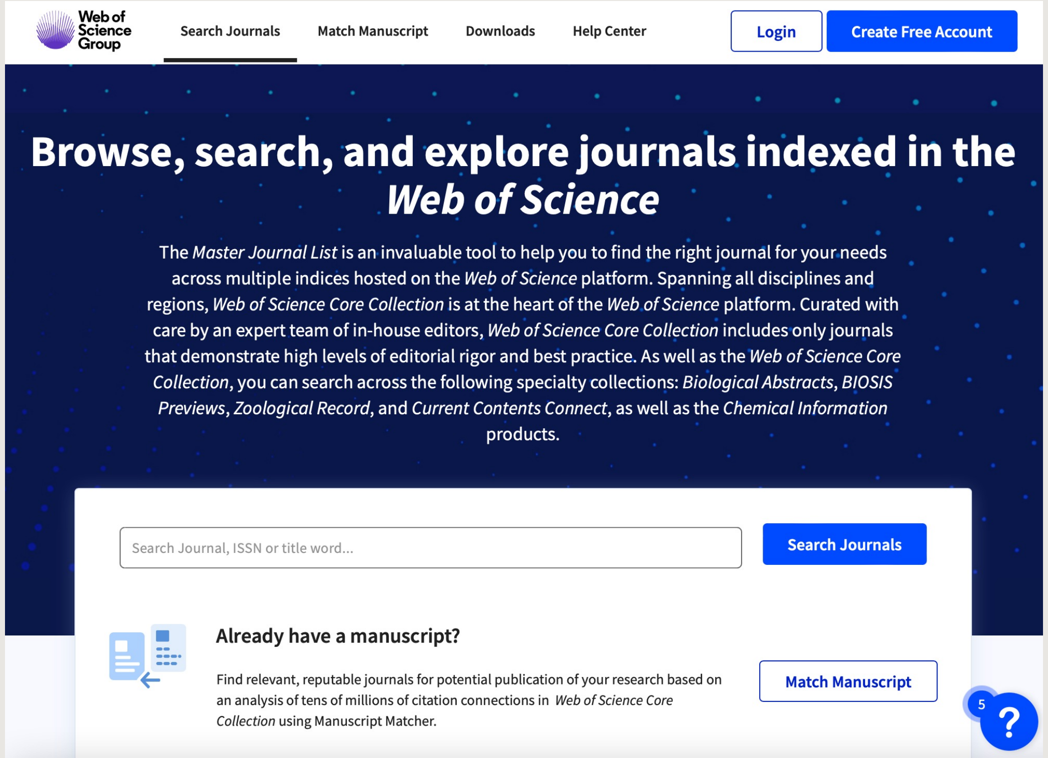
- **<u>Scopus</u>** – H-index and Quartile



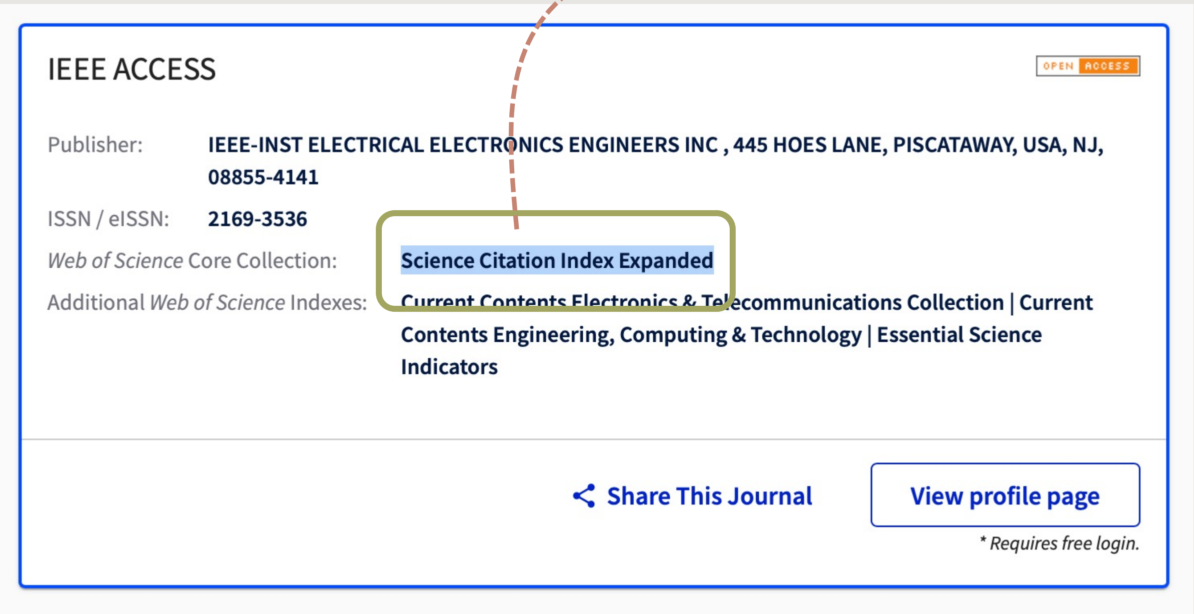


สิ่งที่ต้องพิจารณาหากต้องการเผยแพร่บทความในวารสารวิชาการระดับนานาชาติ

- <u>Web of Science</u> – impact factor and Quartile



SCIE



สิ่งที่ต้องพิจารณาหากต้องการเผยแพร่บทความ ในวารสารวิชาการระดับนานาชาติ

ESCI

<u>*ยังไม่อยู่ในฐาน ISI*</u>

IEEE CONTROL SYSTEMS LETTERS

Publisher: **IEEE-INST ELECTRICAL ELECTRONICS ENGINEERS INC , 445 HOES LANE, PISCATAWAY, USA, NJ, 08855-4141**

ISSN / eISSN: **2475-1456**

Web of Science Core Collection: **Emerging Sources Citation Index**

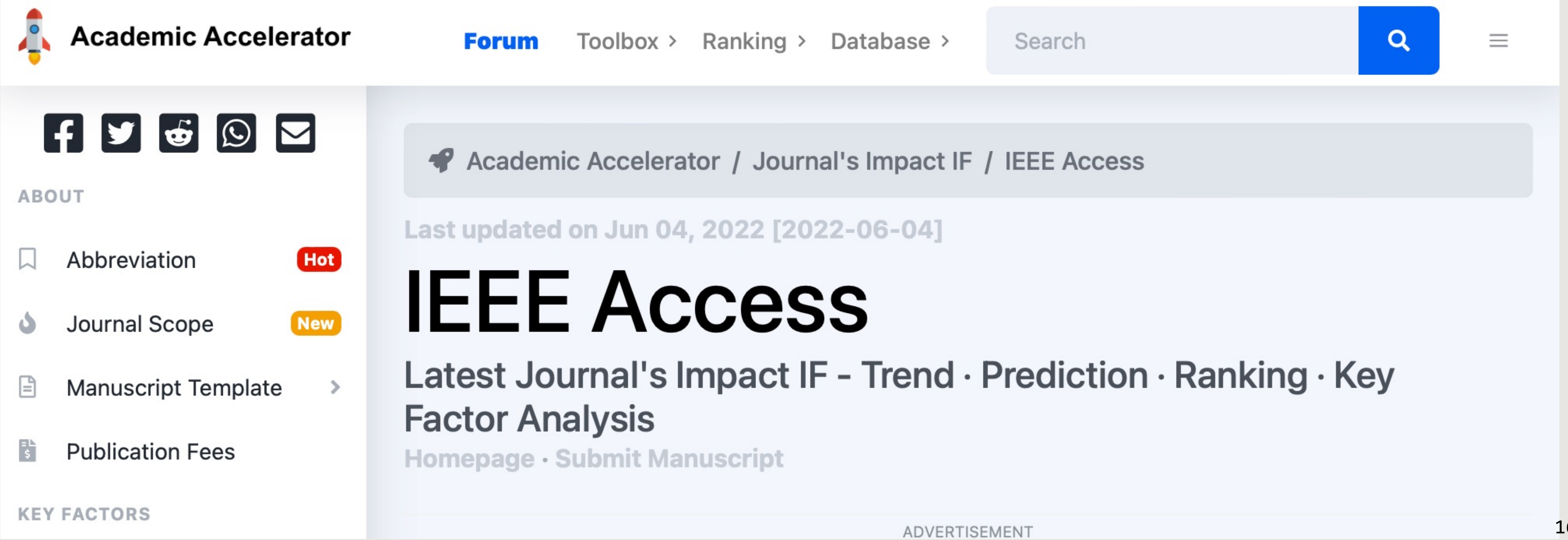
Share This Journal

View profile page

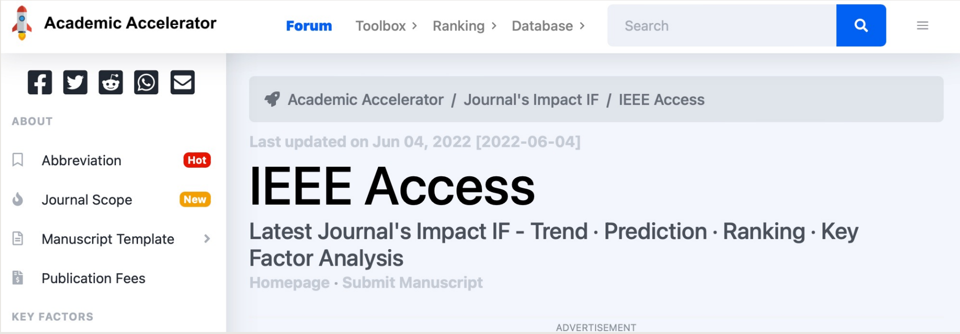
** Requires free login.*

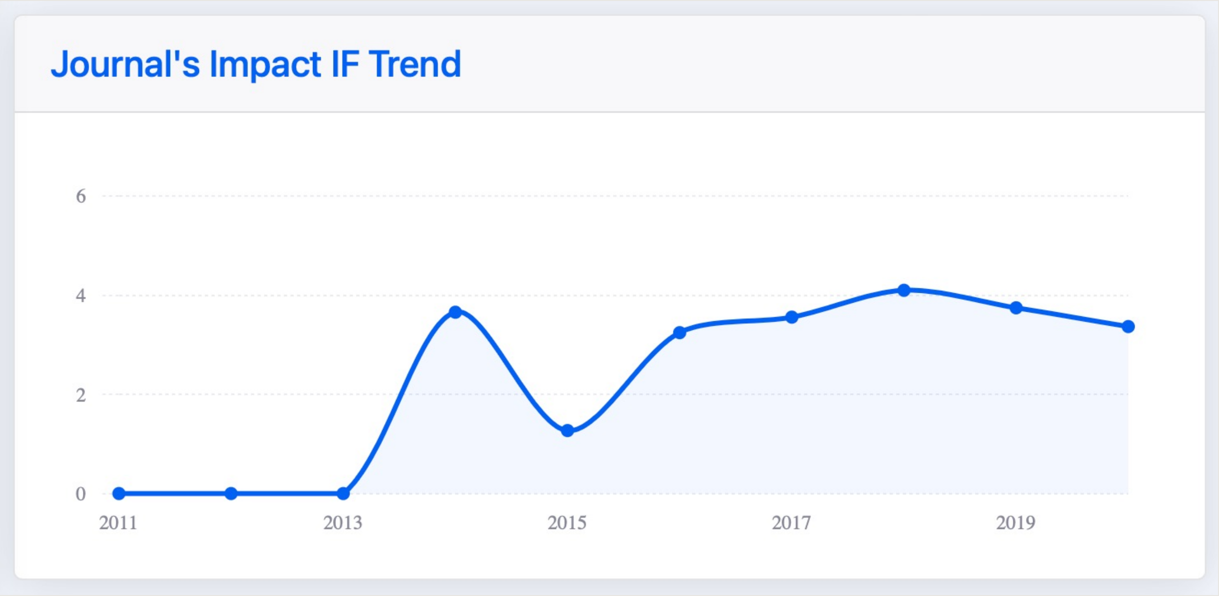
สิ่งที่ต้องพิจารณาหากต้องการเผยแพร่บทความ ในวารสารวิชาการระดับนานาชาติ

- Impact factor



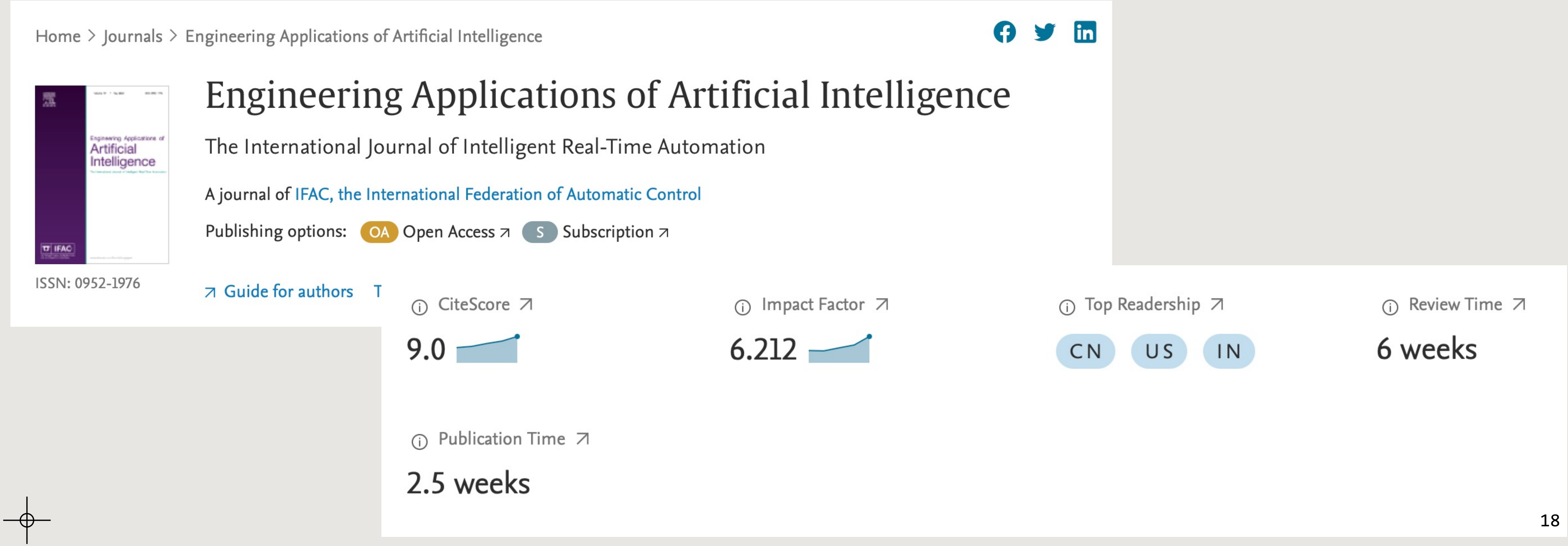
สิ่งที่ต้องพิจารณาหากต้องการเผยแพร่บทความ ในวารสารวิชาการระดับนานาชาติ





สิ่งที่ต้องพิจารณาหากต้องการเผยแพร่บทความในวารสารวิชาการระดับนานาชาติ

- Impact factor



How to calculate "impact factor"

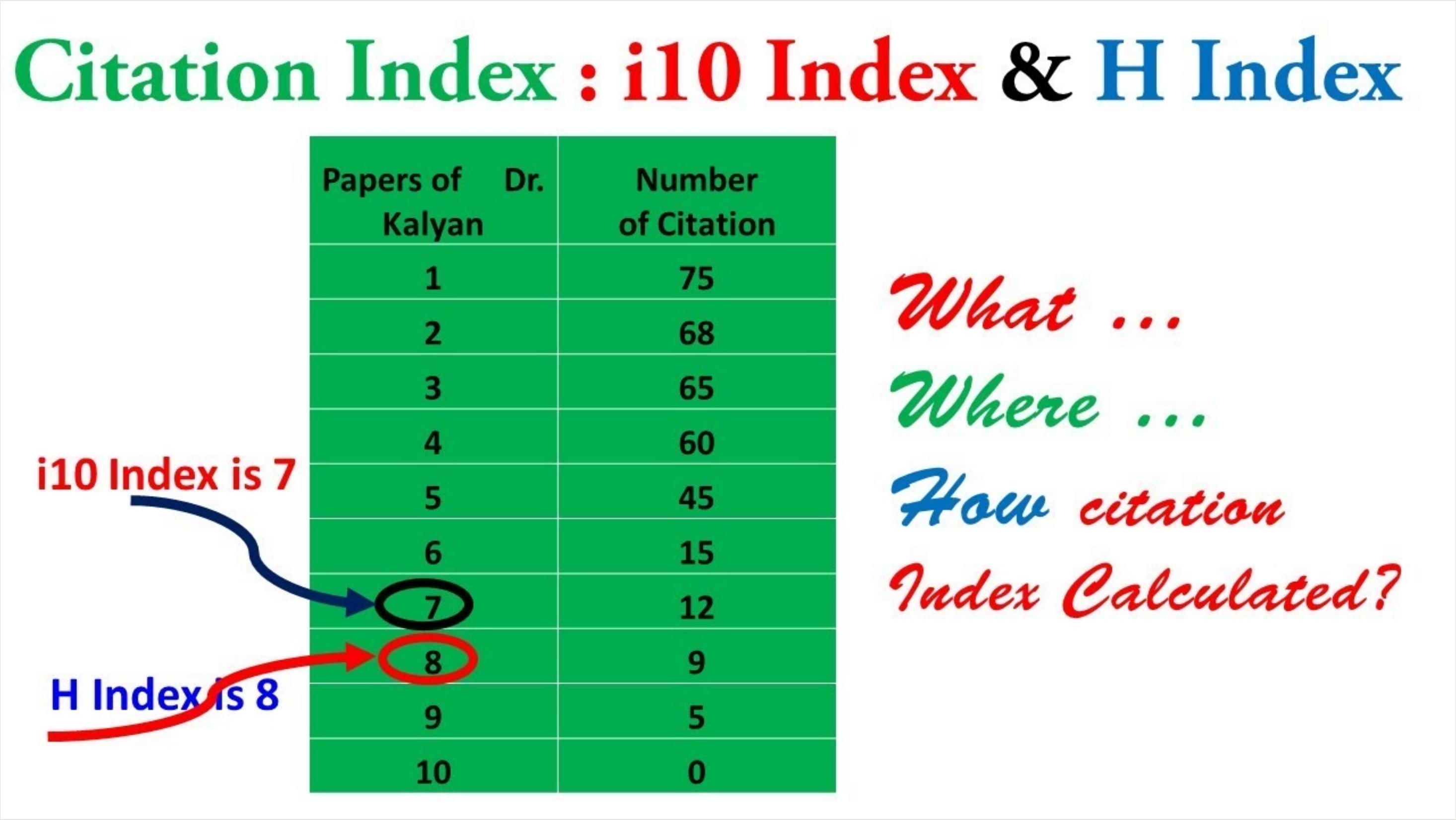
Journal Impact Factor

Cites in 2008 to items published in:	2007 =263	Number of items published in:	2007 =55
	2006 =374		2006 =49
	Sum: 637		Sum: 104

$$\text{Calculation} = \frac{\text{Cites to recent items}}{\text{Number of recent items}} \quad \frac{637}{104} = \mathbf{6.125}$$

https://osu.libguides.com/c.php?g=110226&p=714742

How to calculate "H-index"



Papers of Dr. Kalyan	Number of Citation
1	75
2	68
3	65
4	60
5	45
6	15
7	12
8	9
9	5
10	0

https://www.youtube.com/watch?v=hRPgg36m3ss

สิ่งที่ต้องตรวจสอบก่อนที่จะส่งบทความไปให้วารสารพิจารณา

IEEE Access

Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

ความเร็วในการเผยแพร่

สิ่งที่ต้องตรวจสอบก่อนที่จะส่งบทความไปให้วารสารพิจารณา

Scope

Learn More About IEEE *Access*

A Multidisciplinary Open Access Journal | A Scope Covering all IEEE Fields of Interest | Bibliometrics | Indexing Information | No Page Limits

A Multidisciplinary Open Access Journal

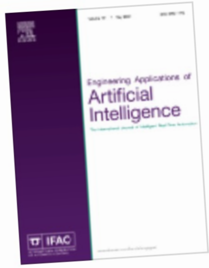
IEEE *Access* is a multidisciplinary, all-electronic archival journal, continuously presenting the results of original research or development across all IEEE fields of interest. Supported by article processing charges (APC), its hallmarks are a rapid peer review and publication process of 4 to 6 weeks, with open access to all readers.

A Scope Covering all IEEE Fields of Interest

IEEE *Access* publishes articles that are of high interest to readers: original, technically correct, and clearly presented. The scope of this journal comprises all IEEE fields of interest, emphasizing applications-oriented and interdisciplinary articles.

IEEE *Access* hosts IEEE Society/Council Sections as well as Special Sections, which give researchers the opportunity to publish their work in dedicated topical sections.

Scope



ISSN: 0952-1976

Engineering Applications of Artificial Intelligence

The International Journal of Intelligent Real-Time Automation

A journal of IFAC, the International Federation of Automatic Control

Publishing options: OA Open Access ↗ S Subscription ↗

↗ Guide for authors   Track your paper ⌵   Order journal ⌵

Artificial Intelligence (AI) is playing a major role in the fourth industrial revolution and we are seeing a lot of evolution in various machine learning methodologies. **AI techniques** are widely used by the practicing engineer to solve a whole range of hitherto intractable problems. This journal provides an international forum for rapid publication of work describing the practical application of **AI methods** in all branches of **engineering**. Submitted papers should report some novel aspects of AI used for a real world engineering application and also validated using some public data sets for easy replicability of the research results.

สิ่งที่ต้องตรวจสอบก่อนที่จะส่งบทความไปให้วารสารพิจารณา

สิ่งที่ต้องตรวจสอบก่อนที่จะส่งบทความไปให้วารสารพิจารณา

Focal points of the journal include, but are not limited to innovative applications of:

- Internet–of–things and cyber-physical systems
- Intelligent transportation systems & smart vehicles
- Big data analytics, understanding complex networks
- Neural networks, fuzzy systems, neuro-fuzzy systems
- Deep learning and real world applications
- Self-organizing, emerging or bio-inspired system
- Global optimization, Meta-heuristics and their applications: Evolutionary Algorithms, swarm intelligence, nature and biologically inspired meta-heuristics, etc.
- Architectures, algorithms and techniques for distributed AI systems, including multi-agent based control and holonic control
- Decision-support systems
- Aspects of reasoning: abductive, case-based, model-based, non-monotonic, incomplete, progressive and approximate reasoning
- Applications of chaos theory and fractals
- Real-time intelligent automation, and their associated supporting methodologies and techniques, including control theory and industrial informatics

Scope

Engineering Applications of Artificial Intelligence publishes:

- Survey papers/tutorials
- Contributed papers — detailed expositions of new research or applications
- Case studies or software reviews — evaluative and descriptive reviews of existing available AI software systems, discussing the experience gained and lessons learnt from using or developing AI systems for engineering applications
- IFAC EAAI Forum — problems arising from engineering practice, needing to be solved by somebody; solutions to problems discussed in this forum or elsewhere; critiques of a position or claim found in the literature

The Editors of *Engineering Applications of Artificial Intelligence* wish to inform authors that this journal will *not* publish papers that propose "novel" metaphor-based metaheuristics, unless the authors:

1. present their method using the normal, standard optimization terminology;
2. show that the new method brings useful and novel concepts to the field;
3. motivate the use of the metaphor on a sound, scientific basis;
4. present a fair comparison with other state-of-the-art methods using state-of-the-art practices for benchmarking algorithms.

For more details on the International Federation of Automatic Control (IFAC), visit their home page at http://www.ifac-control.org

สิ่งที่ต้องตรวจสอบก่อนที่จะส่งบทความไปให้วารสารพิจารณา

Guide for Authors

Explore the following resources that will guide you from article submission to online publication.

- Preparing Your Article
- Submission Guidelines
- Stages of Peer Review
- Post Acceptance Guide

Browse journals > Engineering Applications of Artificial Intelligence > Guide for authors

Guide for Authors

Download Guide for Authors in PDF

Aims and scope +

- Your Paper Your Way

INTRODUCTION

- Types of paper
- Submission checklist

BEFORE YOU BEGIN

- Ethics in publishing
- Declaration of competing interest
- Submission declaration and verification
- Use of inclusive language
- Open access
- Submission
- Submit your article

PREPARATION

- Queries
- NEW SUBMISSIONS
- Peer review
- REVISED SUBMISSIONS
- Article structure
- Essential title page
- Artwork
- Tables
- References
- Video
- Data visualization
- Supplementary material
- Research data

AFTER ACCEPTANCE

- Online proof correction

Guide for Authors

สิ่งที่ต้องตรวจสอบก่อนที่จะส่งบทความไปให้วารสารพิจารณา

Indexing

Indexing Information

- Current Contents/Engineering, Computing and Technology Edition (Clarivate Analytics)
- Directory of Open Access Journals (DOAJ)
- EBSCOhost
- Ei Compendex
- Google Scholar
- IET Inspec
- Journal Citation Reports/Science Edition (Clarivate Analytics)
- Science Citation Index Expanded (Clarivate Analytics)
- Scopus
- Web of Science (Clarivate Analytics)

IEEE Access®

Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Visit journal homepage >

Submit your paper >

Open access option >

Guide for authors >

Track your paper >

Order journal >

View articles >

Browse journals > Engineering Applications of Artificial Intelligence > Abstracting and Indexing

Abstracting and Indexing

- Cambridge Scientific Abstracts
- Current Contents - Engineering, Computing & Technology
- EBSCOhost
- Engineering Index Monthly
- Ei Compendex
- Gale Academic OneFile
- Gale Infotrac Custom
- INSPEC
- Science Citation Index Expanded
- PubMed/Medline
- Scopus
- Web of Science
- TEMA - Technik und Management
- Zentralblatt MATH

Article Processing Charge (APC)

Open access options

Engineering Applications of Artificial Intelligence offers authors two choices to publish their research:

Gold open access	Subscription
Articles are freely available to both subscribers and the wider public with permitted reuse.	Articles are made available to subscribers as well as developing countries and patient groups through our access programs.
An open access publication fee is payable by authors or their research funder.	No open access publication fee.

Gold Open access publication fee

Gold open access publication fee

To provide gold open access, this journal has a publication fee (Article Publishing Charge, APC) which needs to be met by the authors or their research funders for each article published open access. This ensures your article will be immediately and permanently free to access by everyone.

The Article Publishing Charge for this journal is USD 2,680, excluding taxes.

Limit on the Number of Pages and APC

No Page Limits

The article processing charge (APC) is $1,850 per article (plus applicable local taxes) without page limits. Corresponding authors who are IEEE members receive a 5% discount; IEEE members who are also IEEE Society members receive a 20% discount (the 5% IEEE member discount plus an additional 15% discount for Society membership). Please note that this discount does not apply to Student or Graduate Student Members. More information can be found on the IEEE Author Center.

In addition, your institution's library may have established a deposit account with IEEE to pay APCs on your behalf. Please check with your institution's library to see if a deposit account is available.

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Acceptance Rate

At a Glance

- **Journal:** IEEE *Access*
- **Format:** Open Access
- **Frequency:** Continuous
- **Submission to Publication:** 4-6 weeks (typical)
- **Topics:** All topics in IEEE
- **Average Acceptance Rate:** 30%
- **Model:** Binary Peer Review
- **Article Processing Charge:** US $1,850

Learn More

Submission

Submission

Our online submission system guides you stepwise through the process of entering your article details and uploading your files. The system converts your article files to a single PDF file used in the peer-review process. Editable files (e.g., Word, LaTeX) are required to typeset your article for final publication. All correspondence, including notification of the Editor's decision and requests for revision, is sent by e-mail.

Note that authors should include figures, tables and captions embedded directly within the text of the paper submitted for review. This will assist with the reviewing process. Source files are still required for final publication and will be required should your paper be accepted for publication.

Submission

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Submission Guidelines

Submission Checklist | Manuscript Types Acceptable for Peer Review

Submission Checklist

1. Manuscript should be prepared in a double column, single-spaced format using a required IEEE *Access* template. A Word or LaTex file **and** a PDF file **are both required** upon submission. Content on each file must match exactly. File sizes should not exceed 40MB. Download IEEE *Access* Templates for Microsoft Word and LaTex.
2. Author lists should be carefully considered before submission. For more information on what constitutes an author, please click here. (Changes to author list post acceptance are not allowed.)
3. Corresponding authors are required to have an ORCID ID associated with their account. For more information on ORCID, click here.
4. Short biographies are required for ALL authors. As per the required templates, biographies of all authors should be added to the article below the references.
5. All authors should be listed on both the source file (Word of LaTex file) and the manuscript PDF file. When submitting your files ensure the submission system successfully extracts the full author list from your files.
6. The article should be thoroughly reviewed for proper grammar before being submitted. Articles with poor grammar will be immediately rejected. If needed, IEEE *Access* offers Paperpal Preflight to assist authors in checking their manuscript for grammar issues prior to submission. Check your manuscript on Paperpal Preflight by clicking here.
7. All research works should be carefully referenced. More information to avoid plagiarism is listed below.

Cover letter

Dear Madam/Sir,

We are pleased to submit our manuscript titled 'Multi-layer Adaptive Spatial-Temporal Feature Fusion Network for Efficient Food Image Recognition' for consideration as an original article in the journal ''. The manuscript we submit is the result of our work on food image recognition. The manuscript submitted has not been previously published and is also not under consideration for publication in the same or substantially similar form in any other peer-reviewed academic journal or elsewhere.

Our paper presents a food image recognition system using adaptive spatial-temporal feature fusion network, called ASTFF-Net. The ASTFF-Net focus on fusion between spatial and temporal feature using ResNet50 and LSTM. However, we used convolutional 1D (Conv1D) block to fit the features before fed into the LSTM network. We have experimented on four different adaptive spatial-temporal feature fusion networks (ASTFF-NetB1 to B4) on four benchmark food image datasets. The experimental results show that the most accurate network for food image recognition is ASTFF-NetB3 and it also significantly outperformed the existing methods.

All authors listed have contributed sufficiently to the article and are therefore qualified to be listed as authors. To the best of our knowledge, no conflict of interest exists. If you have any further questions, please do not hesitate to contact us. Thank you very much for your consideration.

With kind regards

Sirawan Phiphitphatphaisit & Olarik Surinta
Address: Department of Information Technology
Faculty of Informatics, Mahasarakham University
Kam Riang, Kantarawichai, Maha Sakham, 44150, Thailand

Email: olarik.s@msu.ac.th

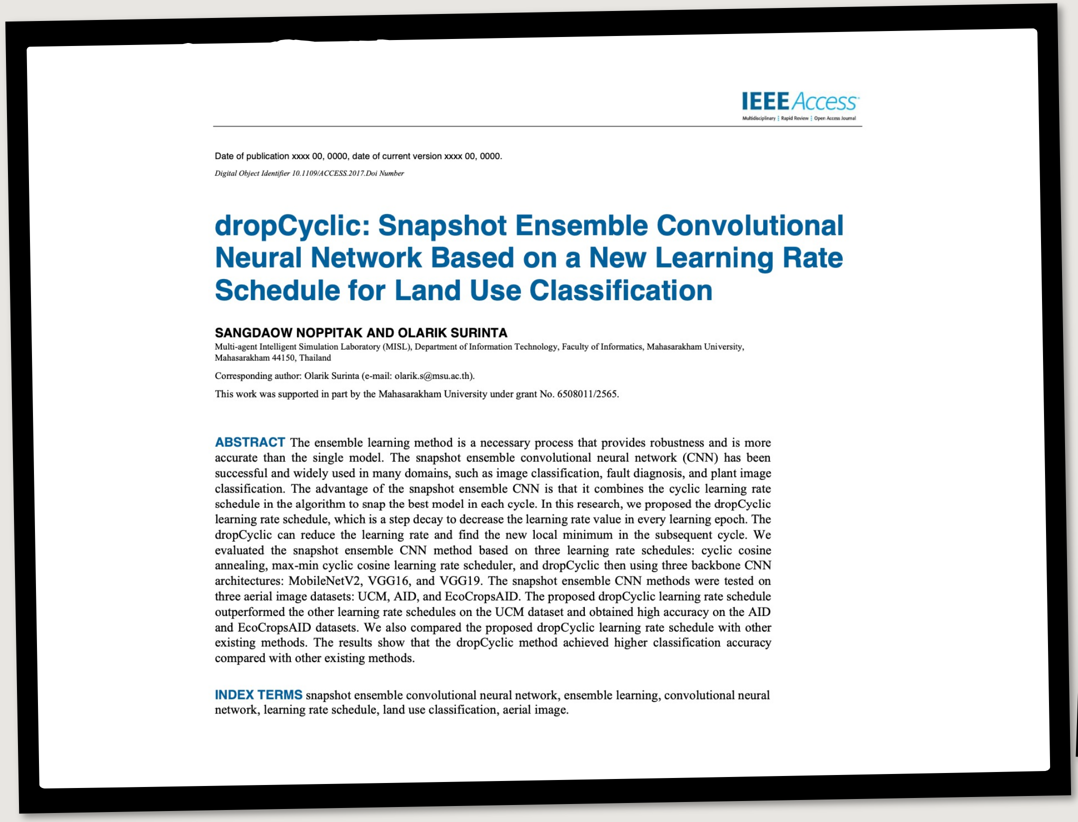
Preparing for submission

- Cover letter

Preparing for submission

- **Manuscript**

  + Proofreading



IEEE Access

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.Doi Number

dropCyclic: Snapshot Ensemble Convolutional Neural Network Based on a New Learning Rate Schedule for Land Use Classification

SANGDAOW NOPPITAK AND OLARIK SURINTA

Multi-agent Intelligent Simulation Laboratory (MISL), Department of Information Technology, Faculty of Informatics, Mahasarakham University, Mahasarakham 44150, Thailand

Corresponding author: Olarik Surinta (e-mail: olarik.s@msu.ac.th).

This work was supported in part by the Mahasarakham University under grant No. 6508011/2565.

ABSTRACT The ensemble learning method is a necessary process that provides robustness and is more accurate than the single model. The snapshot ensemble convolutional neural network (CNN) has been successful and widely used in many domains, such as image classification, fault diagnosis, and plant image classification. The advantage of the snapshot ensemble CNN is that it combines the cyclic learning rate schedule in the algorithm to snap the best model in each cycle. In this research, we proposed the dropCyclic learning rate schedule, which is a step decay to decrease the learning rate value in every learning epoch. The dropCyclic can reduce the learning rate and find the new local minimum in the subsequent cycle. We evaluated the snapshot ensemble CNN method based on three learning rate schedules: cyclic cosine annealing, max-min cyclic cosine learning rate scheduler, and dropCyclic then using three backbone CNN architectures: MobileNetV2, VGG16, and VGG19. The snapshot ensemble CNN methods were tested on three aerial image datasets: UCM, AID, and EcoCropsAID. The proposed dropCyclic learning rate schedule outperformed the other learning rate schedules on the UCM dataset and obtained high accuracy on the AID and EcoCropsAID datasets. We also compared the proposed dropCyclic learning rate schedule with other existing methods. The results show that the dropCyclic method achieved higher classification accuracy compared with other existing methods.

INDEX TERMS snapshot ensemble convolutional neural network, ensemble learning, convolutional neural network, learning rate schedule, land use classification, aerial image.

Engineering Applications of Artificial Intelligence

Multi-Layer Adaptive Spatial-Temporal Feature Fusion Network for Efficient Food Image Recognition

--Manuscript Draft--

Manuscript Number:	EAAI-22-169
Article Type:	Research paper
Keywords:	Food Image Recognition; Deep Feature Extraction Method; long short-term memory; Convolutional Neural Network; Spatial and Temporal Features; Adaptive Feature Fusion
Corresponding Author:	olarik surinta, Ph.D. Mahasarakham University Mahasarakham, THAILAND
First Author:	Sirawan Phiphitphatphaisit
Order of Authors:	Sirawan Phiphitphatphaisit
	olarik surinta, Ph.D.
Abstract:	Various deep learning methods have been proposed to address the challenge of food image classification. This researchproposed a robust adaptive spatial-temporal feature fusion network, called ASTFF-Net, to enhance the performance of the food image recognition system. The architecture of ASTFF-Net is divided into three parts; spatial feature extraction network, temporal feature extraction network, and adaptive feature fusion network. In the first part, we extracted the spatial features using the ResNet50 and then minimized the size of the parameters using the reduction operation. Further, the convolutional 1D (Conv1D) block was applied to fit the features into the recurrent neural networks. In the second part, the spatial features from the first part were given to the long short-term memory (LSTM) that allows learning various patterns from sequence features. In the final part, the spatial features from the first part and temporal features from the second part were concatenated and assigned to the Conv1D, followed by the softmax layer. The advantage of ASTFF-Net is that the proposed network can prevent overfitting problems due to the attachment of the global average pooling and dropout layers. These layers decreased the number of network parameters and dropped the number of connections between layers, respectively. In the experiments, we evaluated four different adaptive feature fusion networks (ASTFF-NetB1 to B4) on four benchmark food image datasets; Food11, UEC Food-100, UEC Food-256, and ETH Food-101. As a result, the proposed ASTFF-NetB3 achieved the best performance on four benchmark food image datasets. It also significantly outperformed the existing methods.

Manuscript types

IEEE Access

Multidisciplinary : Rapid Review : Open Access Journal

Manuscript Types Acceptable for Peer Review

Research Article - This is a classic research article that has a hypothesis, investigation, solution, model, physical experiment and/or simulation and a result that is of value to the community within that area of expertise.

Topical review - This is a review of an emerging area within the journal's scope that performs a technical and critical review of other articles. Calculations are performed and conclusions are drawn on the strengths and weaknesses. The conclusion can also discuss future challenges.

Theory - This is a scholarly article that uses mathematical methods to develop new theoretical results of importance to the field.

Survey - A survey article analyzes, summarizes, systemizes, and presents fresh conclusions from a large number of recently published scholarly articles.

Perspective - This category of article is an in-depth viewpoint article intended to bring together a big picture in a fast-evolving landscape of technological development. It will typically be a topic where there is community uncertainty and/or disagreement. Since IEEE *Access* does not have a page limit, publishing perspective articles will allow more technical discussion. This category of article will typically be written by a leading authority in an area.

Applied research - This article describes challenges and practical solutions for topics within the journal's scope. Quantitative results for validation of the approach are expected.

Negative result - This is a non-trivial theoretical or experimental negative or null result that does not support a hypothesis. Provided that the research question posed is meaningful and the study is rigorously conducted, this type of article has value to the engineering community.

Methods - This article will report the development of new or improved fabrication or manufacturing technique, or a new experimental, measurement or mathematical technique. Applied research articles focus on practical systems, while here the focus is on methods.

Publication Process

Reviewer Recommendation and Comments

Recommendation
✓ No Recommendation
Accept
Minor Revision
Major Revision
Reject

Overall Manuscript Rating (1-100):

Cancel | Save & S... | ...viewer Attachments | Proof & Print | Proceed

Reviewer Comments to Author

Reviewer Confidential Comments to Editor

Is there a financial or other conflict of interest between your work and that of the authors?

YES __ NO __

Please give a frank account of the strengths and weaknesses of the article:

Review Questions

*Does the title of this paper clearly and sufficiently reflect its contents?

*Are the keywords and abstracts/summary informative?

*Are the references relevant and up-to-date?

*Please rate the overall academic value and contribution of this manuscript.

Please select a response

*Are you willing to review the revision of this manuscript?

Referee Report Form

Please mark the appropriate space by typing an 'X' or N/A (not appropriate).

1. Scope

__Of general relevance

__Very relevant in the field of......................

__Very specialised

2. Information contained

__New techniques/theory

__New application of known concepts

__Valuable confirmation of known techniques

__Repetition of known material

__Too theoretical

3. Conclusions drawn

__Adequate

__Not justified

__Suffer from major omissions

__Suffer from loose generalisations

4. Title

__Adequately descriptive

__Should be changed

5. Abstract

__Clear and adequate

__Should be rewritten

__Missing

Referee Report Form

6. Language

__Grammatically good

__Needs revision

7. Presentation and style

__Adequate

__Too brief for clarity

__Too comprehensive, must be shortened

__Contains irrelevant material

__Arrangement unsuitable; Could be better subdivided

8. Illustrations

__Number and quality adequate

__Fig(s) ...... may be omitted

__A figure is desirable to illustrate

....

__Quality of prints/drawings inadequate

9 Tables

__Adequate

__Should be rearranged to represent data more clearly

__Table ...... may be omitted

10 Abbreviations, formulae, units

__Conform to acceptable norms

__Do not conform with accepted standards; should be changed

__Should be explained

11 Literature references

__Adequate

__Inadequate

__................. cannot be located

Referee Report Form

12 The paper is graded as

__Excellent

__Good

__Acceptable

__Sound, but dull

__Confirmatory

__Without obvious significance

__Weak

__Too speculative

__Too preliminary

__Outside this journal's scope

Referee Report Form

Please enter here any further information that is necessary before a decision can be made:

The below checklist covers the points made in the above review form, as well as other aspects which would need to be considered when making your assessment.

1. Scope

Is there an immediate "appeal" to a practising industrial engineer?

Is the title explicit, attractive and interesting?

Is the abstract clear and to the point, stressing both the specific application and the generic aspects of the work?

2.

Does the Introduction clearly state the application area?

Referee Report Form

3.

Is there real evidence of the practical industrial benefits of the technologies/methodologies introduced (e.g., where it was applied, and what improvements resulted)?

Does the Conclusion state these clearly?

4.

Are there generic aspects which make the work applicable beyond a narrow range of applications?

Are these clearly brought out in the paper?

5.

Is the paper correct technically?

6.

Is there some aspect, either in theory or application, which is new or innovative?

Referee Report Form

7.

Is the paper intelligible, but non-trivial, to a practising professional engineer in the field of intended application?

8. Is the paper intelligible, and of some relevance, to practising professional engineers in other fields?

9.

Is the paper easy to read, i.e.,

· Is it to the point?

· Is it grammatically and semantically simple and correct?

· Are the figures, graphs, etc., clear explicit and properly labelled?

· Are the mathematics essential? (Enough detail should be given so that numerical examples can be reproduced exactly, but mathematical proofs should be referenced, rather than spelt out in tedious detail.)

· Are the references complete, and relatively easy to obtain?

· Is the length appropriate? (Most papers will tend to be between 5 and 8 pages in length, but shorter or longer papers are acceptable if their lengths are appropriate to their contents.)

Editor Decision

- Accept
- Accept with minor Revision
- Accept with major Revision
- Reject

Reviewer Recommendation

Overall Recommendation

* Overall Recommendation

- ○ Accept in present form
- ○ Accept after minor revision (corrections to minor methodological errors and text editing)
- ○ Reconsider after major revision (control missing in some experiments)
- ○ Reject (article has serious flaws, additional experiments needed, research not conducted correctly)

Accepted

[Applied Science and Engineering Progress] Your Manuscript Submission No. 3810

Applied Science and Engineering Progress KMUTNB <asep@op.kmutnb.ac.th> Tue, 1 Dec 2020, 15:27

to me

Dear Dr. Olarik Surinta

It is a pleasure to accept your manuscript "Effective data augmentation and training techniques for improving deep learning in plant leaf disease recognition" for publication in Applied Science and Engineering Progress (ASEP) (E-ISSN: 2673-0421).

In order for Applied Science and Engineering Progress to proceed with publication of your article, you must complete a Copyright Form and email back to me (junjiraporn.t@op.kmutnb.ac.th). Under the agreement, you retain copyright to your work and grant an exclusive license to Applied Science and Engineering Progress to publish the article. Please note that without a completed agreement, we are unable to proceed with publication of your article.

Thank you for your fine contribution. We look forward to your continued contributions to the Journal.

Regards,

Editor: ASEP
Asst. Prof. Dr. Nawaporn Wisitpongphan

Response to the Reviewers

Reference No.: Access-2022-06942
Response to Reviewers

Dear Editors-in-Chief,

Thank you for giving us the opportunity to submit a revised manuscript "**dropCyclic: Snapshot Ensemble Convolutional Neural Network Based on a New Learning Rate Schedule for Land Use Classification**" for publication in the SCIE indexed journals. We appreciate the time and effort that you and the reviewers dedicated to providing feedback on our manuscript and are grateful for the insightful comments on and valuable improvements to our paper. We have incorporated the suggestions made by the reviewers. Those changes are highlighted within the manuscript. Please see below, for a point-by-point response to the reviewers' comments and concerns. In the manuscript, we use the yellow highlighter to highlight the change according to the comments.

Response to the Reviewers

ขอบคุณ Reviewer

Reviewer 3

1. The manuscript presents a novel learning rate schedule for snapshot ensemble CNN models. The work is evaluated on benchmark datasets for land use classification. The related work section is nicely written with detailed description of specific methods.

Author response: We would like to thank the reviewers for their thoughtful comments and efforts toward improving our manuscript.

2. In Section 3, most content is existing work and only subsection A.3 is the proposed dropCyclic schedule. Sections should be re-arranged for better clarity of the proposed work and the existing work.

Author response: Thank you for your suggestion. We re-arranged the manuscript in Section III, as follows. A. Snapshot Ensemble methods and B. Cosine Cyclic Learning Rate Schedule. However, in Subsection B, we presented the topic as follows 1) cyclic cosine annealing, 2) max-min cosine cyclic learning rate scheduler, and 3) proposed drop cyclic cosine learning rate schedule, because we would like to show the order of the first, second, and third improvement equation of the snapshot ensemble methods.

Response to the Reviewers

ตอบทุกข้อคำถามที่ Reviewer ถาม โดยตอบให้ละเอียด

<u>**** ตอบทุกข้อ ****</u>

Reviewer comment and response to reviewer form
Applied Science and Engineering Progress
Article number: <u>3810</u>

Topic	Comments and responses		
	Reviewer Comment (For reviewer)	**Response to comment (For author)**	**Note**
Reviewer 2			
	This paper presented The effectiveness of training techniques and data augmentation using deep learning for plant leaf disease recognition. It is very interesting, provides a reliable method, provides a good literature. It is very interesting, provides a proposed method, provides a good literature. I found the article to be particularly well-implemented. The main contribution is good enough to present in this journal.	We want to express our appreciation to the reviewer for your valuable comments on the manuscript, "*The effectiveness of training techniques and data augmentation using deep learning for plant leaf disease recognition*" for publication in the *Journal of Applied Science and Engineering Progress.* However, we want to improve the title to be appropriate for the article. The new title is "**Effective data augmentation and training techniques for improving deep learning in plant leaf disease recognition**."	
Reviewer 1			
	The strong points of this paper are well written and the structure is easy to follow. The results are sound and reasonable with sufficient explanations. The leaf disease dataset, which the authors created, is interesting as it contains the number of classes higher than the iCassava2019 dataset. The online and mixed training with data augmentation techniques strategy is an interesting concept. The authors provided enough details of their methodologies and the iCassava2019 dataset is public. Thus, this work can be reproducible to the certain degree.	We want to express our appreciation to the reviewer for your valuable comments, which have helped us improve our article. Thank you for allowing us to submit a revised version of the manuscript to the Journal of Applied Science and Engineering Progress. We would like to improve the title to be appropriate for the article. The new title is "**Effective data augmentation and training techniques for improving deep learning in plant leaf disease recognition**." We have included most of the suggestions given by the reviewer. Those changes are highlighted in the manuscript.	
	The authors also found that the "brightness	Thank you for pointing this out.	

Response to the Reviewers

Reviewer จะให้ความสำคัญกับเรื่อง **Novelty** มาก

Topic	Comments and responses		
	Reviewer Comment (For reviewer)	**Response to comment (For author)**	**Note**
	The main weak point of this paper is novelty. The main conclusion and contribution about “The performance of the deep learning method is improved when combining data augmentation techniques” “transfer learning shows a better result than training data from scratch” is not quite new. It can be found in other works such as “Data augmentation for plant classification” (Pawara et al).	Thank you for the suggestion. As you said, it can be found in the data augmentation experiments from Pawara et al. (2017), "Data Augmentation for Plant Classification." In their experiments, the illumination result (adding random values between 10 and 80 to the R, G, and B channels) showed that the performance was relatively high, with 99.42% with a fine-tuned GoogleNet and 98.46% with a fine-tuned AlexNet. For our experiments, however, the brightness technique gave low accuracy on the plant leaf disease dataset (see Table 5). It obtained 90.77% with fine-tuning MobileNetV2 and only 63.08 with scratch MobileNetV2. When experiments with the NASNetMobile, it obtained only 66.92% and 89.23% with scratch and fine-tuning, respectively. This because the brightness technique directly affected diseases by occasionally fading out the white spots and the disease spots on the leaves. In our experiments, zoom, rotation, and shift techniques showed better performance. We added two data augmentation techniques; Cutout and Mixup. The result of these two augmentation techniques does not affect high accuracy. Also, we concentrated on combining data augmentation techniques and training techniques to enhance the performance of the deep learning method; offline, online, and mixed methods. As a result, However, we examined to address the issue of plant leaf disease recognition. We obtained more than 95% on the plant leaf dataset and obtained around 83% on the iCassave 2019 dataset.	
	The sentence “The objective was to to compare	Thank you for your suggestion. We agree with the	

Response to the Reviewers

อย่าลืมขอบคุณ Reviewer

คำชมของ Reviewer

	There is a minor issue in table 1. "Details of the leaf disease dataset, which consists of 13 types of plant diseases and the number of images of leaf diseases as each type of plant disease". I think there are 12 types of plant diseases instead of 13 because one is healthy class.	We agree with the comments and have changed according to the suggestion. We changed caption of Table 1 to "Table 1: Details of the leaf disease dataset (consists of 13 types; 12 types of plant diseases and one type of healthy) and the number of images of leaf diseases as each type of plant disease."	
	In conclusion, this is good work and well written. The results are sound and reasonable. It would be even more interesting if the authors can point out the uniqueness of this dataset, for example, finding some data augmentation techniques that are inappropriate for plant disease recognition and providing explanations.	Thank you for pointing this out. We added this sentence to the conclusion section. "On the contrary, the brightness technique that generated a plant leaf image by adding high-intensity values affected the plant leaf disease images by changing the white spots and the disease spots on the plant leaves. Hence, it is inappropriate for plant leaf disease recognition."	

Copyright

ทุกวารสารจะให้ผู้แต่งเซ็น Copyright

Applied Science and Engineering Progress (2022) Vol.15

Paper Number: **3810** Title of Paper: Effective Data Augmentation and Training Techniques for Improving Deep Learning in Plant Leaf Disease Recognition

Language of Paper: English

Area of the work (Choose in the areas suggests in home-page or other) ____________________________

Copyright Transfer Form

The main author: Olarik Surinta

On behalf of the co-authors: Prem Enkvetchakul

of the Work: Effective Data Augmentation and Training Techniques for Improving Deep Learning in Plant Leaf Disease Recognition was accepted to be published in Applied Science and Engineering Progress, Volume 15, Number x, (2022) undersign hereby assigning all copyright rights in and to the above work to the Applied Science and Engineering Progress. The undersigned hereby represents and warrants that:

(1) He/she holds the copyrights of the work.

(2) The work is original and was not published previously.

(3) He/she can sign this transfer for joint author(s).

In return for these rights, the Applied Science and Engineering Progress recognizes the retained rights:

(1) The author(s) is/are entitled of using a substantial part of the article in future works.

(2)The rights revert to the author(s) if the article is not published.

Main Author Signature: ____________________________________ Date: ________________________________

( )

Acceptance Letter

บางวารสารจะมี
Acceptance Letter
แนบเป็นไฟล์มาให้

แต่ส่วนใหญ่จะแค่ส่งอีเมล

Applied Science and Engineering Progress

King Mongkut's University of Technology North Bangkok

Acceptance letter

2 December 2020

Paper Name: Effective Data Augmentation and Training Techniques for Improving Deep Learning in Plant Leaf Disease Recognition

Authors: Prem Enkvetchakul and Olarik Surinta

Dear Author,

I am very pleased to inform you that your paper is accepted for Applied Science and Engineering Progress, ISSN: 2672-9156. The paper is arranged to be published in Vol.15, No.2, April–June, 2022.

Best regards,

Prof. Dr.-Ing. habil. Suchart Siengchin

Editor-in-Chief

Proof your manuscript

Applied Science and Engineering Progress KMUTNB <asep@op.kmutnb.ac.th> Mon, 18 Jan 2021, 16:35

to me

Dear Dr. Olarik Surinta

We are pleased to inform you that your paper in title “Effective Data Augmentation and Training Techniques for Improving Deep Learning in Plant Leaf Disease Recognition” is nearing publication. You can help us facilitate quick and accurate publication by reviewing the proofs. Please keep in mind the following:

- Only errors introduced during production process or that directly compromise the scientific integrity of the paper may be corrected.
- Any changes that contradict journal style will not be made.
- Any changes to scientific content (including figures) will require editorial review and approval.

Please check the author names and affiliations very carefully to ensure correct spelling and correct sequence.

Please submit your corrections within 5 working days **before 25 January 2021**. Without your response to these queries, we will not be able to continue with the processing of your article for Online Publication.

If you have any correction, please color-highlight where it is and send the file to junjiraporn.t@op.kmutnb.ac.th.

Yours sincerely,
Applied Science and Engineering Progress

Revised Manuscript

First round: *Revise and Resubmission*

[ecti-cit] Editor Decision [Revise and Resubmission] External Inbox ×

THAIJO

Suphakant Phimoltares via Thai Journals Online (ThaiJO) <admin@tci-thaijo.org> Tue, 18 Aug 2020, 12:35

to [redacted], me

Dear [redacted] Olarik Surinta:

We have reached a decision regarding your submission to ECTI Transactions on Computer and Information Technology (ECTI-CIT), [redacted] ' has been completed.

Our decision is to: "Revise and Resubmission"

Based on the recommendation of the reviewers, the manuscript needs some revision based on the reviewers' comments and/or suggestions and then please resubmit the revised version within the next 3 weeks for further process.

On behalf of the Editor-in-Chief, Prof.Dr.Prabhas Chongstitvatana and Prof.Dr.Chidchanok Lursinsap, we thank you for considering the ECTI Transactions on CIT. Please contact us (E-mail: editorcit@ecti-thailand.org) for any questions or enquires.

Suphakant Phimoltares
Chulalongkorn University
suphakant@gmail.com

Reviewer A:
Recommendation: Revisions Required

Reviewer B:
Recommendation: Decline Submission

Problem

Reviewer C:
Recommendation: Revisions Required

Reject / has not been accepted

[ecti-cit] Editor Decision External Inbox ×

Suphakant Phimoltares via Thai Journals Online (ThaiJO) <admin@tci-thaijo.org> Sat, 3 Oct 2020, 22:08

THAIJO

to , me

Dear , Olarik Surinta:

We have reached a decision regarding your submission to ECTI Transactions on Computer and Information Technology (ECTI-CIT), ".

We regret to inform you that your paper has not been accepted for publishing at ECTI-CIT Transactions. The overall quality of submissions was extremely high, and we have had to decline a number of very good papers. Comments and/or suggestions from the reviewers are given below.

Thank you for your submission, and we hope you renew your interest in publishing your future work with ECTI Transactions on CIT.

Yours,
Prof.Dr. Prabhas Chongstitvatana

Prof.Dr.Chidchanok Lursinsap

Editor-in-Chief

Reject / has not been accepted

Reviewer A:

Recommendation: Resubmit Elsewhere

, Olarik Surinta:

lecision regarding your submission to ECTI Transactions on Computer an

".

ou that your paper has not been accepted for publishing at ECTI-CIT Tra
and we have had to decline a number of very good papers. Comments an

ubmission, and we hope you renew your interest in publishing your future

ongstitvatana

Reviewer B:

Recommendation: Accept Submission

Title

suitable

Abstract

suitable

Introduction and Objective

suitable

Methodology

suitable

Result and discussion

suitable

Conclusion

suitable

Other comments (Please provide constructive comments for manuscript improvment)

This research has been revised in a suitable direction.

Author

First Author

* Corresponding Author

CONCURRENCY AND COMPUTATION: PRACTICE AND EXPERIENCE
Concurrency Computat.: Pract. Exper. (2014)
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/cpe.3413

SPECIAL ISSUE PAPER

Evaluating automatically parallelized versions of the support vector machine

Valeriu Codreanu[1,5,*,†] Bob Dröge[2], David Williams[1], Burhan Yasar[6],
Po Yang[4], Baoquan Liu[4], Feng Dong[4], Olarik Surinta[3], Lambert R.B. Schomaker[3],
Jos B.T.M. Roerdink[1] and Marco A. Wiering[3]

[1]*Johann Bernoulli Institute for Mathematics and Computer Science, University of Groningen, Groningen, The Netherlands*
[2]*Donald Smits Centrum voor Informatie Technologie, University of Groningen, Groningen, The Netherlands*
[3]*Institute of Artificial Intelligence and Cognitive Engineering, University of Groningen, Groningen, The Netherlands*
[4]*Department of Computer Science and Technology, University of Bedfordshire, Bedford, UK*
[5]*Electronic Systems Group, Eindhoven University of Technology, Eindhoven, The Netherlands*
[6]*Rotasoft Inc., Ankara, Turkey*

Author

First Author and *Corresponding Author

Engineering Applications of Artificial Intelligence 45 (2015) 405–414

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

ELSEVIER

CrossMark

Recognition of handwritten characters using local gradient feature descriptors

Olarik Surinta *[,1], Mahir F. Karaaba [1], Lambert R.B. Schomaker [1], Marco A. Wiering [1]

Institute of Artificial Intelligence and Cognitive Engineering (ALICE), University of Groningen, PO Box 407, 9700 AK Groningen, The Netherlands

Supervisor

Engineering and Applied Science Research 2022;49(4):505-520 Research Article

EASR

Engineering and Applied Science Research

https://www.tci-thaijo.org/index.php/easr/index

Published by the Faculty of Engineering, Khon Kaen University, Thailand

CycleAugment: Efficient data augmentation strategy for handwritten text recognition in historical document images

Sarayut Gonwirat and Olarik Surinta*

Multi-agent Intelligent Simulation Laboratory (MISL), Department of Information Technology, Faculty of Informatics, Mahasarakham University, Mahasarakham 44150, Thailand

Received 24 November 2021
Revised 4 February 2022
Accepted 25 February 2022

First Author

*Corresponding Author

Author

Timeline

Engineering and Applied Science Research 2022;49(4):505-520 Research Article

EASR

Engineering and Applied Science Research

https://www.tci-thaijo.org/index.php/easr/index

Published by the Faculty of Engineering, Khon Kaen University, Thailand

CycleAugment: Efficient data augmentation strategy for handwritten text recognition in historical document images

Sarayut Gonwirat and Olarik Surinta*

Multi-agent Intelligent Simulation Laboratory (MISL), Department of Information Technology, Faculty of Informatics, Mahasarakham University, Mahasarakham 44150, Thailand

Received 24 November 2021
Revised 4 February 2022
Accepted 25 February 2022

3 เดือน

Timeline

CONCURRENCY AND COMPUTATION: PRACTICE AND EXPERIENCE
Concurrency Computat.: Pract. Exper. (2014)
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/cpe.3413

SPECIAL ISSUE PAPER

Evaluating automatically parallelized versions of the support vector machine

Valeriu Codreanu[1,5,*,†], Bob Dröge[2], David Williams[1], Burhan Yasar[6], Po Yang[4], Baoquan Liu[4], Feng Dong[4], Olarik Surinta[3], Lambert R.B. Schomaker[3], Jos B.T.M. Roerdink[1] and Marco A. Wiering[3]

[1]*Johann Bernoulli Institute for Mathematics and Computer Science, University of Groningen, Groningen, The Netherlands*
[2]*Donald Smits Centrum voor Informatie Technologie, University of Groningen, Groningen, The Netherlands*
[3]*Institute of Artificial Intelligence and Cognitive Engineering, University of Groningen, Groningen, The Netherlands*
[4]*Department of Computer Science and Technology, University of Bedfordshire, Bedford, UK*
[5]*Electronic Systems Group, Eindhoven University of Technology, Eindhoven, The Netherlands*
[6]*Rotasoft Inc., Ankara, Turkey*

SUMMA

The support vector machine (SVM) is a supervised le
data. It is a very popular technique in machine learnin
such as image classification, protein classification, and
tional complexity of the kernelized version of the algorit
examples. To tackle this high computational complexit
that converts a gradient-ascent based training algorithn
unit (GPU) implementation. We compare our GPU-base
CPU implementation, a highly optimized GPU-LibSV
OpenACC implementation. The results on different ha
an important speed-up for the current approach when compared to the CPU and OpenACC versions. Furthermore, our solution is almost as fast and sometimes even faster than the highly optimized CUBLAS-based GPU-LibSVM implementation, without sacrificing the algorithm's accuracy. Copyright © 2014 John Wiley & Sons, Ltd.

Received 30 December 2013; Revised 5 June 2014; Accepted 9 September 2014

9 เดือน

KEY WORDS: GPU; automatic parallelization; handwritten digit recognition; machine learning; support vector machine

Engineering Applications of Artificial Intelligence 45 (2015) 405–414

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

ELSEVIER

Recognition of handwritten characters using local gradient feature descriptors

CrossMark

Olarik Surinta [*,1], Mahir F. Karaaba [1], Lambert R.B. Schomaker [1], Marco A. Wiering [1]

Institute of Artificial Intelligence and Cognitive Engineering (ALICE), University of Groningen, PO Box 407, 9700 AK Groningen, The Netherlands

A R T I C L E   I N F O

Article history:
Received 14 April 2015
Received in revised form
5 June 2015
Accepted 23 July 2015
Available online 10 August 2015

Keywords:
Handwritten character recognition
Feature extraction
Local gradient feature descriptor
Support vector machine
k-nearest neighbors

3 เดือน

Timeline

Affiliation - the institution where you conducted the research

CONCURRENCY AND COMPUTATION: PRACTICE AND EXPERIENCE
Concurrency Computat.: Pract. Exper. (2014)
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/cpe.3413

SPECIAL ISSUE PAPER

Evaluating automatically parallelized versions of the support vector machine

Valeriu Codreanu[1,5,*,†], Bob Dröge[2], David Williams[1], Burhan Yasar[6], Po Yang[4], Baoquan Liu[4], Feng Dong[4], Olarik Surinta[3], Lambert R.B. Schomaker[3], Jos B.T.M. Roerdink[1] and Marco A. Wiering[3]

[1] *Johann Bernoulli Institute for Mathematics and Computer Science, University of Groningen, Groningen, The Netherlands*
[2] *Donald Smits Centrum voor Informatie Technologie, University of Groningen, Groningen, The Netherlands*
[3] *Institute of Artificial Intelligence and Cognitive Engineering, University of Groningen, Groningen, The Netherlands*
[4] *Department of Computer Science and Technology, University of Bedfordshire, Bedford, UK*
[5] *Electronic Systems Group, Eindhoven University of Technology, Eindhoven, The Netherlands*
[6] *Rotasoft Inc., Ankara, Turkey*

Affiliation

Engineering Applications of Artificial Intelligence 45 (2015) 405–414

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Recognition of handwritten characters using local gradient feature descriptors

Olarik Surinta [*,1], Mahir F. Karaaba [1], Lambert R.B. Schomaker [1], Marco A. Wiering [1]

Institute of Artificial Intelligence and Cognitive Engineering (ALICE), University of Groningen, PO Box 407, 9700 AK Groningen, The Netherlands

Affiliation

Engineering and Applied Science Research 2021;48(5):560-569 Research Article

EASR

Engineering and Applied Science Research

https://www.tci-thaijo.org/index.php/easr/index

Published by the Faculty of Engineering, Khon Kaen University, Thailand

Optimal weighted parameters of ensemble convolutional neural networks based on a differential evolution algorithm for enhancing pornographic image classification

Sarayut Gonwirat[1)] and Olarik Surinta*[2)]

[1)]Department of Computer and Automation Engineering, Faculty of Engineering and Industrial Technology, Kalasin University, Kalasin 46000, Thailand
[2)]Department of Information Technology, Faculty of Informatics, Mahasarakham University, Maha Sarakham 44150 Thailand

Received 18 December 2020
Revised 14 February 2021
Accepted 17 February 2021

Abstract

Abstracts are typically 150–250 words

- Research Problem and Objectives
- Methods
- Results
- Conclusion

Result

Objective

Proposed method

Conclusion

ABSTRACT

In this paper we propose to use local gradient feature descriptors, namely the scale invariant feature transform keypoint descriptor and the histogram of oriented gradients, for handwritten character recognition. The local gradient feature descriptors are used to extract feature vectors from the handwritten images, which are then presented to a machine learning algorithm to do the actual classification. As classifiers, the *k*-nearest neighbor and the support vector machine algorithms are used. We have evaluated these feature descriptors and classifiers on three different language scripts, namely Thai, Bangla, and Latin, consisting of both handwritten characters and digits. The results show that the local gradient feature descriptors significantly outperform directly using pixel intensities from the images. When the proposed feature descriptors are combined with the support vector machine, very high accuracies are obtained on the Thai handwritten datasets (character and digit), the Latin handwritten datasets (character and digit), and the Bangla handwritten digit dataset.

© 2015 Elsevier Ltd. All rights reserved.

Keywords – at least 3 keywords (Actually 3-5)

Abstract

Use of ensemble convolutional neural networks (CNNs) has become a more robust strategy to improve image classification performance. However, the success of the ensemble method depends on appropriately selecting the optimal weighted parameters. This paper aims to automatically optimize the weighted parameters using the differential evolution (DE) algorithm. The DE algorithm is applied to the weighted parameters and then assigning the optimal weighted to the ensemble method and stacked ensemble method. For the ensemble method, the weighted average ensemble method is applied. For the stacked ensemble method, we use the support vector machine for the second-level classifier. In the experiments, firstly, we experimented with discovering the baseline CNN models and found the best models on the pornographic image dataset were NASNetLarge with an accuracy of 93.63%. Additionally, three CNN models, including EfficientNetB1, InceptionResNetV2, and MobileNetV2, also obtained an accuracy above 92%. Secondly, we generated two ensemble CNN frameworks; the ensemble learning method, called Ensemble-CNN and the stacked ensemble learning method, called StackedEnsemble-CNN. In the framework, we optimized the weighted parameter using the DE algorithm with six mutation strategies containing rand/1, rand/2, best/1, best/2, current to best/1, and random to best/1. Therefore, the optimal weighted was given to classify using ensemble and stacked ensemble methods. The result showed that the Ensemble-3CNN and StackedEnsemble-3CNN, when optimized using the best/2 mutation strategy, surpassed other mutation strategies with an accuracy of 96.83%. The results indicated that we could create the learning method framework with only 3 CNN models, including NASNetLarge, EfficientNetB1, and InceptionResNetV2.

Keywords: Pornographic image classification, Differential evolution algorithm, Mutation strategy, Convolutional neural networks, Ensemble convolutional neural networks, Stacked ensemble learning method, Ensemble learning method

Abstract and Keywords

Research problem

Method

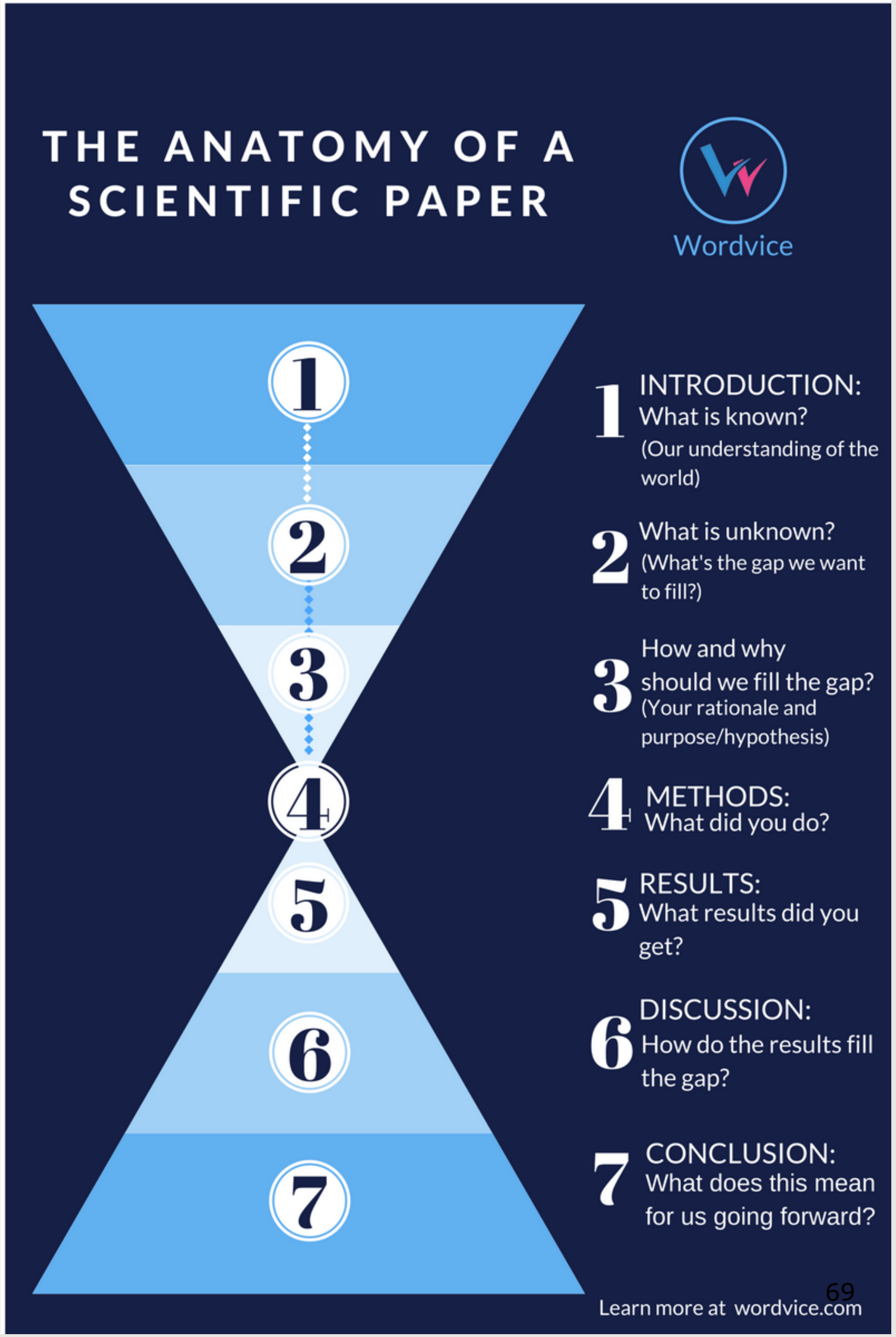
Result

Abstract

Plant disease is the most common problem in agriculture. Usually, the symptoms appear on leaves of the plants which allow farmers to diagnose and prevent the disease from spreading to other areas. An accurate and consistent plant disease recognition system can help prevent the spread of diseases and save maintenance costs. In this research, we present a plant leaf disease recognition system using two deep convolutional neural networks (CNNs); MobileNetV2 and NasNetMobile. These CNN architectures are designed to be suitable for smartphones due to the models being small. We have experimented on training techniques; online, offline, and mixed training techniques on two plant leaf diseases. As for data augmentation techniques, we found that the combination of rotation, shift, and zoom techniques significantly increases the performance of the CNN architectures. The experimental results show that the most accurate algorithm for plant leaf disease recognition is NASNetMobile architecture using transfer learning. Additionally, the most accurate result is obtained when combining the offline training technique with data augmentation techniques.

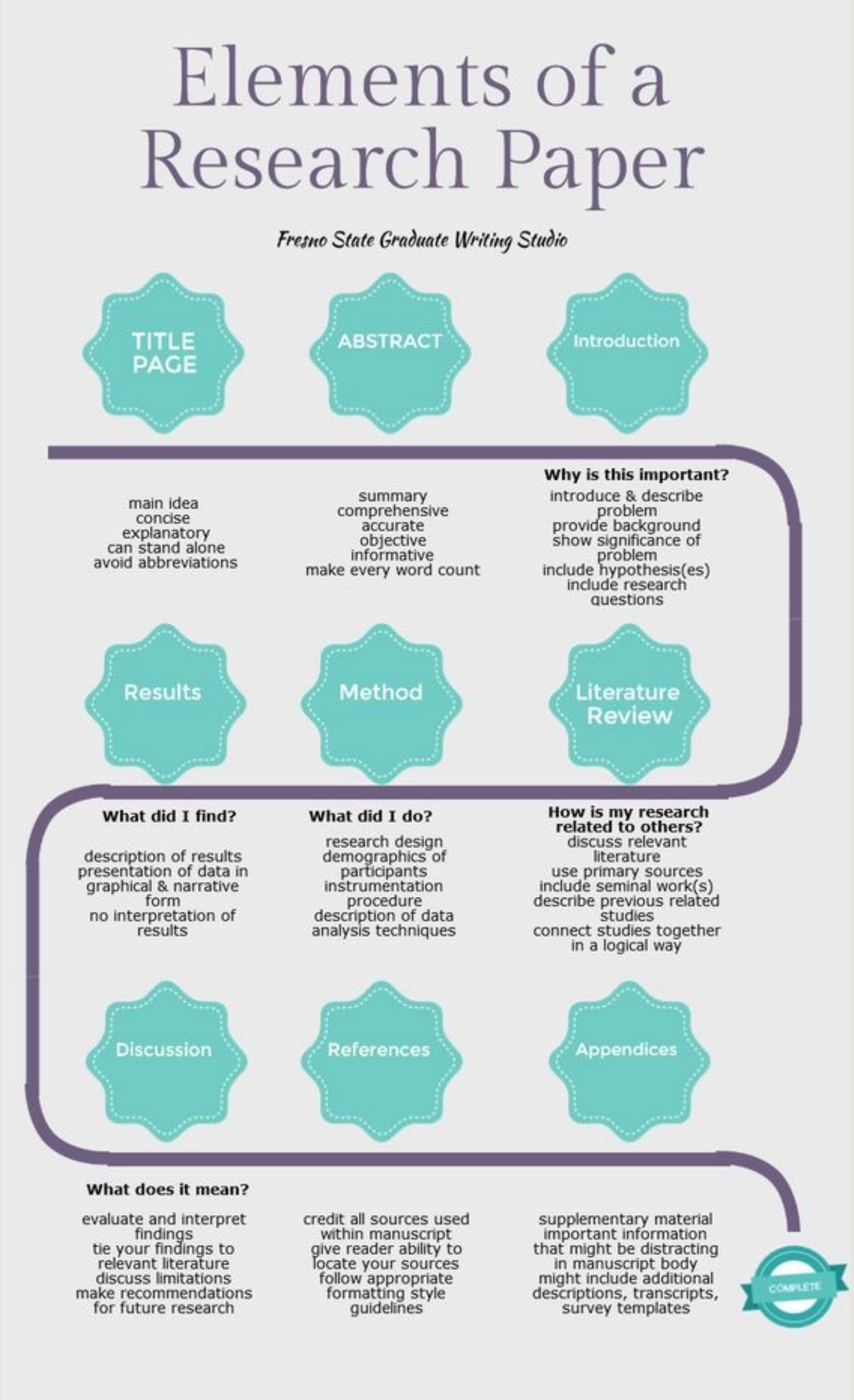
Keywords: Plant leaf disease recognition, Deep learning, Convolutional neural networks, Transfer learning, Data augmentation

Anatomy of a scientific paper



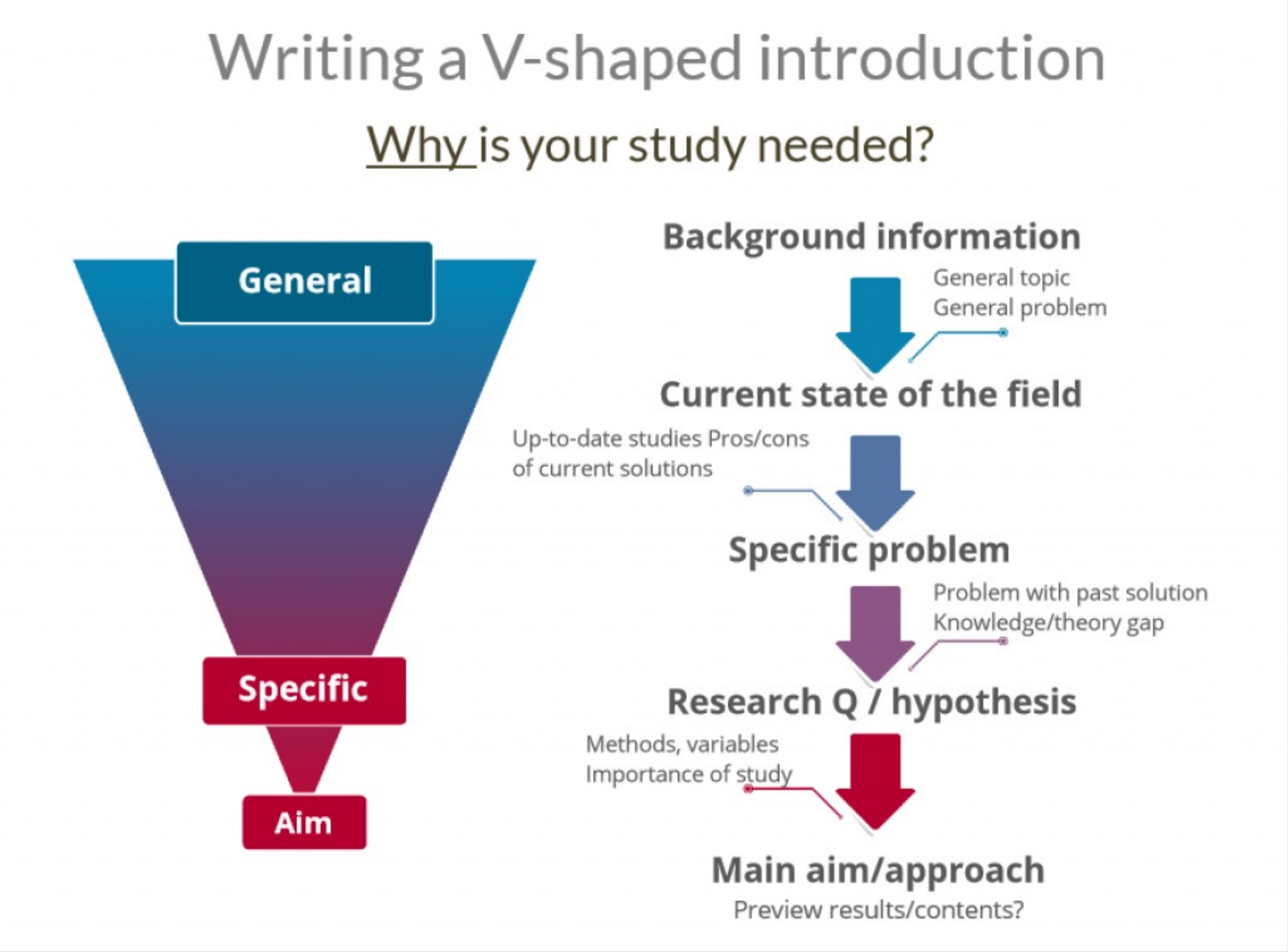
https://blog.wordvice.com/how-to-draft-a-compelling-introduction-for-your-journal-article/

Elements of a Research Paper



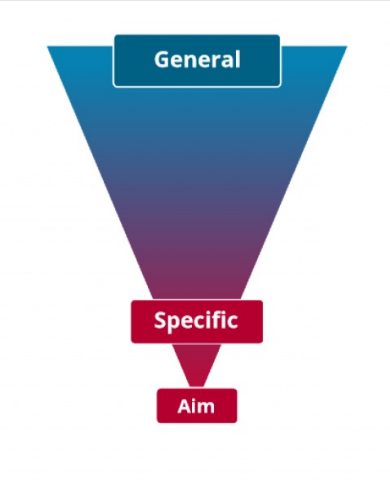
https://learning.edanz.com/writing-the-introduction-v-for-victory/

Introduction Section



url: researchgate

Introduction



Background information

Specific

1. Introduction

Overweight and obesity are the most significant factors for chronic diseases such as diabetes and cardiovascular diseases. The easiest way to avoid chronic diseases is to monitor and control people's dietary behavior. The advancement of artificial intelligence might help people to monitor and estimate daily calorie intake. Hence, food recognition systems are the most straightforward solution. Many systems can recognize several foods based on images. However, when people take a photograph several food characteristics (e.g. the shape and decoration of food, brightness adjustment, and non-food objects, called noise food images) are sent to the system to compute and predict the food type and calorific content. These issues can be a cause of weaknesses of food imaging systems. Computer vision and machine learning techniques are proposed to address the problems mentioned above. Many researchers employ computer vision techniques to generate hand-crafted visual features and send robust features to the novel machine learning techniques, such as support vector machine (SVM), multilayer perceptron (MLP), random forest, and Naive Bayes techniques [1-3] to classify objects [4, 5].

Furthermore, many studies have extracted the robust features, called the spatial features, using convolution neural network (CNN) architectures. The greatest benefit of this technique is that we can extract robust features with various CNN architectures. The robust features, however, are sent to be classified using traditional machine learning techniques. Additionally, the CNN architecture combined with a long short-term memory (LSTM) network has been applied for classification tasks. Nevertheless, a few researchers have invented CNN architectures and LSTM networks for food image recognition. In this research, we focus on improving the accuracy performance of the food image recognition based on CNN architectures and LSTM networks.

Introduction

Background information

Specific

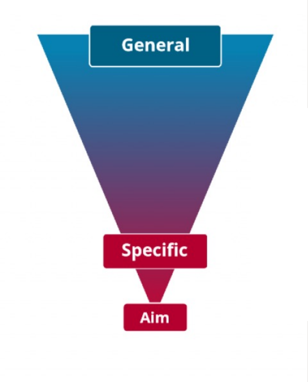
1. Introduction

The offline text recognition system is a vision-based application that automates extracting information from handwritten and printed manuscripts and transforms images into digitally readable text that is editable and comfortable to store and retrieve. Earlier research focused on character recognition which recognized isolated characters [1-4]. However, a few studies concentrated on the recognition of handwritten text. This is because it takes more effort to segment handwritten text into individual characters [5-7]. Due to messy handwriting, various writing styles, and cursive texts, as shown in Figure 1, it is difficult to solve by segmenting characters and then recognizing them by traditional optical character recognition (OCR). Character sequence learning is more suitable for word recognition [8, 9]. Hence, an effective feature-based sliding window and sequence learning methods are applied to recognize each character and then transcript to words [10-12]. However, handwritten text recognition (HTR) methods mainly focus on word recognition and have become a more prominent research domain nowadays.

Deep learning methods have become the principal method in various computer vision applications, such as object detection, object recognition, speech recognition, and natural language processing. Further, convolutional neural network (CNN) architectures, one of the deep learning methods, are widely proposed for feature extraction and image classification. CNN is also proposed to address the challenge of word recognition [13-15]. In addition, CNN and recurrent neural networks (RNNs), which are the famous sequence learner architectures, were proposed to recognize both printed and handwritten words [14, 15] and achieved a high accuracy performance. Consequently, state-of-the-art in handwritten character recognition is a combination of CNN and RNN, called convolutional recurrent network (CRNN). The CRNN also proposed solving problems in many text recognition fields such as scene text and video subtitle recognition.

Moreover, handwritten text recognition has been applied in many languages, such as English, Chinese, Arabic, Indian, and Amharic [13, 14, 16-19]. Particularly, historical manuscripts contain cursive writing, noisy background, and differing word spelling from an ancient and insufficient lexicon for transcription. The challenge of Thai handwritten character recognition is that the Thai language does not have an exact rule to split the sentences and no space between words. For explicit prediction, it is demanding to segment sentences into tokenized words.

Contribution / Aim



This paper aims to experiment with subtitle recognition that transforms the subtitle text image into text format. We propose CNN and LSTM architectures for recognition of Thai and English video subtitle images. The contributions of this paper can be summarized as follows.

- We propose the CNN and LSTM architectures, namely CNN-LSTM architecture for text line recognition. For the CNN architecture, we modify the VGGs architectures and then compare the experimental results with the method proposed by Chamchong et al. [1]. The experimental results show that our CNN-LSTM architecture obtains a lower character-level error value than Chamchong et al. [1].
- This paper aims to provide the new standard Thai and English languages video subtitle dataset for subtitle text recognition. The video subtitle dataset contains 4,224 images and includes 157 characters.

The main contribution of this paper is to present the new data augmentation strategy, namely CycleAugment. The proposed data augmentation strategy mainly focuses on minimizing the validation loss and avoiding overfitting. We achieve our goal with a simplistic strategy and implementation. Our research is motivated by Huang et al. [20], who proposed the cyclic cosine annealing method that calculated the learning rate in every epoch and then started the new learning rate at the beginning of a new cycle.

Furthermore, training the CRNN model usually allows choosing only to train the CRNN model with or without applying data augmentation techniques. We offer the CycleAugment strategy that provides the ability to train the CRNN model with and without applying data augmentation techniques simultaneously. Importantly, our CycleAugment strategy confirms that it can handle every CRNN architecture.

We evaluate the efficiency of the CycleAugment strategy on several CRNN architectures for handwritten word recognition on Thai archive manuscripts. To show the importance of the CycleAugment strategy, we compared it to the original data augmentation strategy. The results showed that the CycleAugment strategy significantly decreased the character error rate (CER). The CycleAugment strategy achieved the CER value of 5.43 and the original data augmentation strategy obtained the CER value of 7.31 on the Thai archive manuscript.

Contribution / Aim

The significant contributions of this research are summarized in the following:

1. We propose the deep learning framework that combines state-of-the-art ResNet50, which is the convolutional neural network (CNN) and long short-term memory (LSTM) network, called ResNet50+Conv1D-LSTM network. This framework can extract robust features that are spatial and temporal features, from the food images. Mixed data augmentation techniques are also involved while training the CNN model. The data augmentation technique insignificantly increases the performance of food image recognition.

2. In these experiments, LSTM and Conv1D-LSTM networks were proposed to create robust temporal features. For the Conv1D network, various layers were combined, including zero padding, batch normalization, Convolution 1D, ReLU, batch normalization, dropout, and average pooling layers. In the training scheme, batch size, which was the number of training examples, were applied as 16, 32, and 64. The LSTM network results showed that a batch size of 32 provided a better result than batch sizes of 16 and 64.

Contribution

Introduction + Literature Review

1. **Introduction.** Video media has been published on various channels such as YouTube, Facebook, and Instagram. It gives the audience friendly options to choose and watch freely. Nowadays, video subtitles have been added to the videos to make them accessible to a broad audience, including foreigners and the hearing impaired. Significantly, adding subtitles increases the audience watching the video content and the video creator also received increased revenue from more video views. Some examples of the video subtitles are shown in Figure 1.

In recent years, deep learning methods, such as convolutional neural network (CNN) and long short-term memory (LSTM), have been proposed to address text and character recognition. Chamchong et al. [1] proposed a hybrid deep neural network combined with CNN and recurrent neural network (RNN). They designed hybrid deep neural networks with a tiny CNN weight layer. It included two CNN weight layers and each layer consists of 16 feature maps. The last CNN layer was combined with two layers of a bidirectional gated recurrent unit (Bi-GRU). It was called the 3CNN+BiGRU network. They trained the models using the connectionist temporal classification (CTC) loss function on Thai ancient manuscripts. The result showed that the tiny 3CNN+BiGRU obtained the character-level error rate (CER) value of 11.9%. Yan and Xu [2] proposed using the residual network (ResNet) architecture, Bi-GRU, and CTC to recognize Chinese and English subtitle texts in video images. It obtained an accuracy of 92.3% on the ICDAR2003 and 89.2% on the ICDAR2013.

Gan et al. [3] proposed a 1D-CNN and temporal convolutional recurrent network, called 1D-TCRN, to recognize in-air handwritten Chinese text on a large-scale IAHCT dataset. In this model, the two 1D residual convolution blocks were applied. These two blocks were connected as sequence layers. Hence, this architecture was then connected with LSTM and CTC layers. This network could recognize 2,565 characters of Chinese handwritten text. In [4], the subtitle left/right boundary detection discovered text window regions, called the CNN ensemble algorithm. First, a sliding window slides through the text windows computing the deep features using the CNN architecture and sending these features to classify as text or not-text using a support vector machine algorithm. The CNN ensemble algorithm can determine the text region and recognize the characters at the same time.

Zhang et al. [5] invented a scale-aware hierarchical attention network, called SaHAN, for scene text recognition. This network included two schemes: encoder and decoder. For the encoder, a deep pyramid convolutional recurrent neural network was proposed to create the multi-scale features. The smallest features were then converted to 1D vectors to learn semantic information in the bi-directional LSTM (Bi-LSTM). For the decoder, the semantic information and multi-scale features were transferred to the hierarchical attention decoder. It included two stages in the hierarchical attention decoder: 1D and 2D. Hence, the output of the 1D attention decoder was trained by the GRU and predicted the sequence label.

Introduction + Related Work

1. Introduction

Handwritten character recognition systems have several important applications, such as zip-code recognition, writer identification for e.g. forensic research, searching in historical manuscripts, and others. For such applications, the system should be able to recognize handwritten characters written on many different kinds of documents, such as contemporary or historical manuscripts. The aim is to let the system to automatically extract and recognize the characters that are embedded in the manuscript. The quality of the manuscript is one of the factors that can improve the recognition accuracy (Gupta et al., 2011). It is essential to deal with the different problems that occur in the manuscripts, such as distortions in a character image and the background noise that can appear during the scanning process. The aim of our work is to develop new algorithms that can obtain a high recognition accuracy.

Obtaining high recognition accuracies on handwritten character datasets is a challenging problem, for which many different solutions have been proposed. Although on the standard MNIST dataset extremely high accuracies have been obtained (Meier, 2011), there are many other datasets consisting of less examples and which can be

* Corresponding author. Tel.: +31 50 363 9464.
E-mail addresses: o.surinta@rug.nl (O. Surinta),
m.f.karaaba@rug.nl (M.F. Karaaba), l.r.b.schomaker@rug.nl (L.R.B. Schomaker),
m.a.wiering@rug.nl (M.A. Wiering).
URL: http://www.ai.rug.nl/~mrolarik (O. Surinta).
[1] ALICE http://www.rug.nl/research/alice

considered more difficult. These data
different writing styles of the same
persons (with differences in age, gend
writing devices, and difficulties due to b
from the printer (Surinta et al., 2012).

In this paper we emphasize the imp
complex handwritten Thai, Bangla, and
handwritten characters and digits are h
shapes, strokes, curls, and concavities
samples of the handwritten characters
the handwritten images shown in this
resolution for illustration purposes. Due
use of pixel intensities may not work
sometimes little overlap between two
ing the same character. Therefore, in t
feature extraction techniques which
ments, but still provide discriminative
tion of the handwritten characters. Th
that we will make use of have also bee
recognition, namely the scale invaria
descriptor (Lowe, 2004) and the hist
(HOG) (Dalal and Triggs, 2005). This
these local gradient feature descripto
handwritten characters and digits lead
system. High recognition performances
ging handwritten datasets even with a s
nearest neighbor method, and very hig
obtained when using a support vector

Related work: In previous studies, the raw image (IMG) method, which directly copies the intensities of the pixels of the ink trace (Surinta et al., 2013), has often been used as the feature extraction method. It extracts a high dimensional feature vector that depends on the size of the input image.

In recent years, deep learning architectures (Hinton et al., 2006; Schmidhuber, 2015) have been effectively used for handwritten digit recognition. Most of the studies have focused on the benchmark MNIST dataset (LeCun and Cortes, 1998) and achieved high accuracies such as higher than 98% or 99%. The MNIST dataset consists of isolated handwritten digits with size of 28×28 pixel resolution and contains 60,000 training images and 10,000 test images. In Hinton et al. (2006), a greedy training algorithm is proposed for constructing a multilayer network architecture which relies on the restricted Boltzmann machine, called deep belief networks (DBN). The performance obtained from the DBN with three hidden layers (500–500–2000 hidden units) on the MNIST dataset was 98.75%. This accuracy is higher than obtained with a multi-layer perceptron and a support vector machine (SVM).

Outline

Paper Outline. This paper is organized as follows. Section 2 briefly explains deep learning researches in food image recognition systems and describes the different deep learning techniques. Section 3 describes the proposed approach for the food image recognition system. In Section 4, the experimental settings and the results of the deep learning methods are presented. The conclusion and directions for future work are given in Section 5.

The remainder of this paper is organized as follows. The related work is briefly described in Section 2. Section 3 deeply explains the proposed CRNN architecture and proposed CycleAugment strategy. Section 4, present the Thai historical document dataset, training strategy, and experimental evaluation. The discussion is presented in Section 5. Finally, the last section gives the conclusion and future direction

Literature Review & Related Work

2. Related work

In this section, we survey the HTR task based on deep learning techniques. We also study the transfer learning and data augmentation techniques that improve the performance of deep learning.

2.1 Handwritten text recognition

Text recognition systems have been proposed for several applications, such as scene text recognition [21-24], video subtitle recognition [17, 25], and handwritten text recognition in many languages [14, 16, 19]. Currently, most of the proposed HTR methods are based on the CNNs and RNNs architectures.

For HTR, Abdurahman et al. [16] proposed a convolutional recurrent neural network architecture, called AHWR-Net, to recognize Amharic words. The AHWR-Net architecture was divided into feature extraction, sequence modeling, and classification. First, they created a CNN model and compared their proposed CNN model with state-of-the-art CNN models: DenseNet-121, ResNet-50, and VGG-19. These CNN models were also proposed to extract the feature from the Amharic word images. Second, the RNN architecture was proposed as the sequence model to train spatial features extracted from the previous step. Finally, the probability distributions, which was the output of the RNN method, were classified using a connectionist temporal classification algorithm (CTC). In addition, Butt et al. [19] built a robust Arabic text recognition system using the CNN-RNN attention model from natural scene images. Their Arabic text recognition system addressed the challenge of working with texts in different sizes, fonts, colors, orientation, and brightness.

Furthermore, Ameryan & Schomaker [14] proposed a high-performance word classification using homogeneous CNN and long short-term memory (LSTM) networks. First, for the CNN model, they created five CNN layers. Each CNN layer contained a convolutional layer, normalization method, nonlinear rectified linear unit (ReLU), and max-pooling layer. Second, for the LSTM, bidirectional-LSTMs with three layers were used. Third, the CTC decoding was attached as the output of their network. Finally, invented ensemble system was proposed, which included five networks. The outputs of each network were sent to vote using the plurality vote method.

Table 1 Performance evaluation of classification results on the food datasets using deep learning techniques.

Datasets	Architectures	Accuracy	References
UEC-FOOD100 [19]	DeepFood	76.30	Liu et al. [16]
	InceptionV3	81.45	Hassannejad et al. [15]
	WISeR	89.58	Martinel et al. [20]
UEC-FOOD256 [21]	DeepFood	54.70	Liu et al [16]
	GoogLeNet	63.16	Bolanos and Radeva [22]
	InceptionV3	76.17	Hassannejad et al. [15]
	WISeR	83.15	Martinel et al. [20]
Food-101 [23]	Inception	77.40	Lie et al. [16]
	GoogLeNet	79.20	Bolanos and Radeva [22]
	InceptionV3	88.28	Hassannejad et al. [15]
	Ensemble Net	72.12	Pandey et al. [17]
	CNNs Fusion	86.71	Aguilar et al. [18]
	ResNet152	64.98	McAllister et al. [2]
	WISeR	90.27	Martinel et al. [20]

Comparison existing methods

Citation / In-text Citation

IEEE style

- Gonwirat & Surinta **[31]** trained CNN models (including Inception-ResNetV2 and VGG19 architectures) based on two training methods, scratch learning and transfer learning.

APA style

- The quality of the manuscript is one of the factors that can improve the recognition accuracy **(Gupta et al., 2011)**.

APA style

- In Hinton et al. **(2006)**, a greedy training algorithm is proposed for constructing a multilayer network architecture which relies on the restricted Boltzmann machine, called deep belief networks (DBN).

Method

2. Local gradient feature descriptors

To study the effectiveness of local gradient feature descriptors for handwritten character recognition, we compare two existing feature extraction techniques, namely the histogram of oriented gradients and the scale invariant feature transform keypoint descriptor. Moreover, these local gradient feature descriptors are compared to the IMG method. The IMG method uses the raw pixel intensities of the handwritten images and is a simple and widely used method. In this study, the handwritten images are resized to two pixel resolutions, 28×28 and 36×36, so that for the IMG method 784 and 1296 feature values are computed, respectively.

3. Proposed approach for the food image recognition system

This section explains the framework of food image recognition. Two main architectures, convolutional neural network (CNN) and long short-term memory (LSTM) network, are proposed to extract the robust features from the food images. Hence, the robust spatial and temporal features are extracted from state-ofthe-art ResNet architecture and the LSTM network. The temporal features extracted from the LSTM network are transformed into a probability distribution using the softmax function.

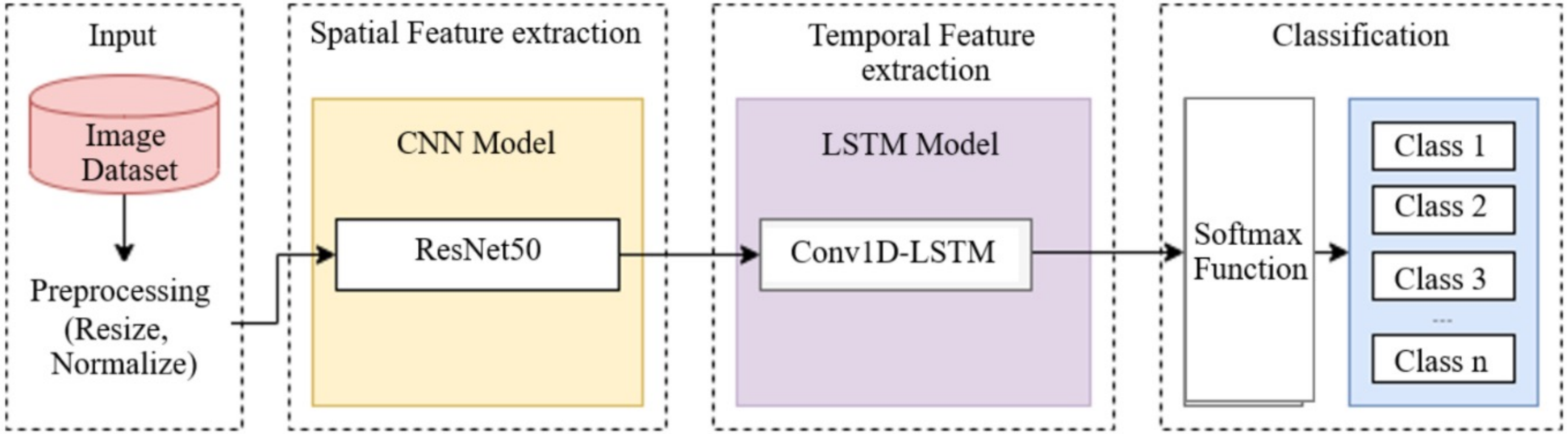


Figure 1 Architecture of our proposed framework for food image classification

Methods

Framework

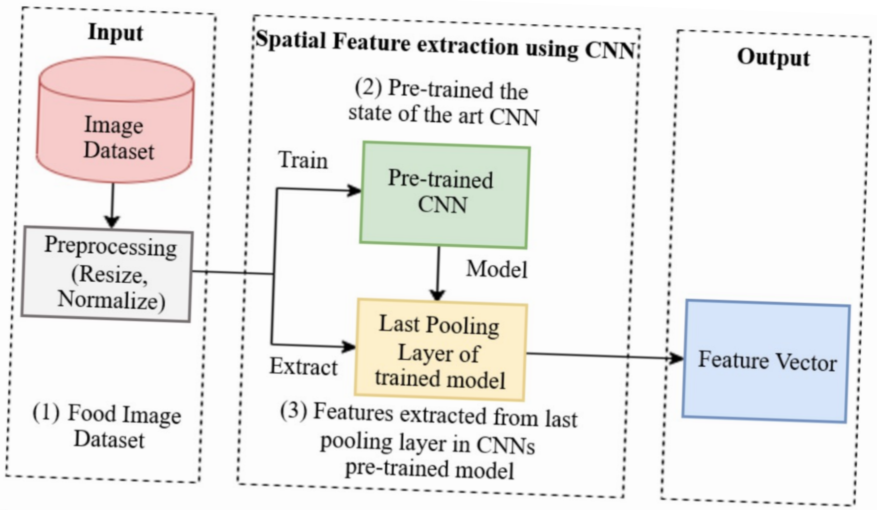


Figure 2 Diagram of the deep feature extraction technique. (1) food images are fed to the pre-processing step to resize and normalize. In the spatial feature extraction process, (2) food images are trained using state-of-the-art CNN architectures to find weights with low validation loss. Then, (3) the spatial features of the food images are extracted according to the best CNN model.

Method

3. The convolutional recurrent neural network

In this section, we present the convolutional recurrent neural network (CRNN) framework for Thai handwritten text recognition of historical document images with a new data augmentation strategy. Firstly, convolutional neural networks (CNNs) are described. Secondly, two recurrent neural networks (RNNs) (long short-term memory and gated recurrent unit) are briefly detailed. Thirdly, detail of the connectionist temporal classification (CTC) decoding is presented for the evaluation metric. Finally, the proposed cyclical data augmentation strategy, namely CycleAugment, is presented. The proposed framework is explained as follows.

3.1 Overview of the CRNN architecture

The CRNN network is illustrated in Figure 2. The CRNN has only one input. Our framework also supports both images of a group of words and short words as for the input. In the CNN architecture, we propose eight different CNN architectures to find the best base CNN model. For the RNN network, we propose two layers of bidirectional RNN networks and connect them to the CNN architecture. Hence, the outputs of the bidirectional RNN network are then classified using the softmax function. The output of the CRNN is a matrix containing character probabilities for each time step. Further, the CTC decoding is attached at the last layer to decode the probability of characters to make the final text output. Our framework can predict a maximum of 94 members in total, including characters, numbers, and blanks (space). The configurations of all CRNN architectures are shown in Table 1.

Furthermore, we propose the cyclical data augmentation strategy (CycleAugment). The CycleAugment strategy provides the CRNN model to train handwritten text images concurrently with and without applying data augmentation techniques. CycleAugment is a powerful strategy for obtaining various local optimal loss values in each cycle until they reach a minimum value at the end of training.

Framework

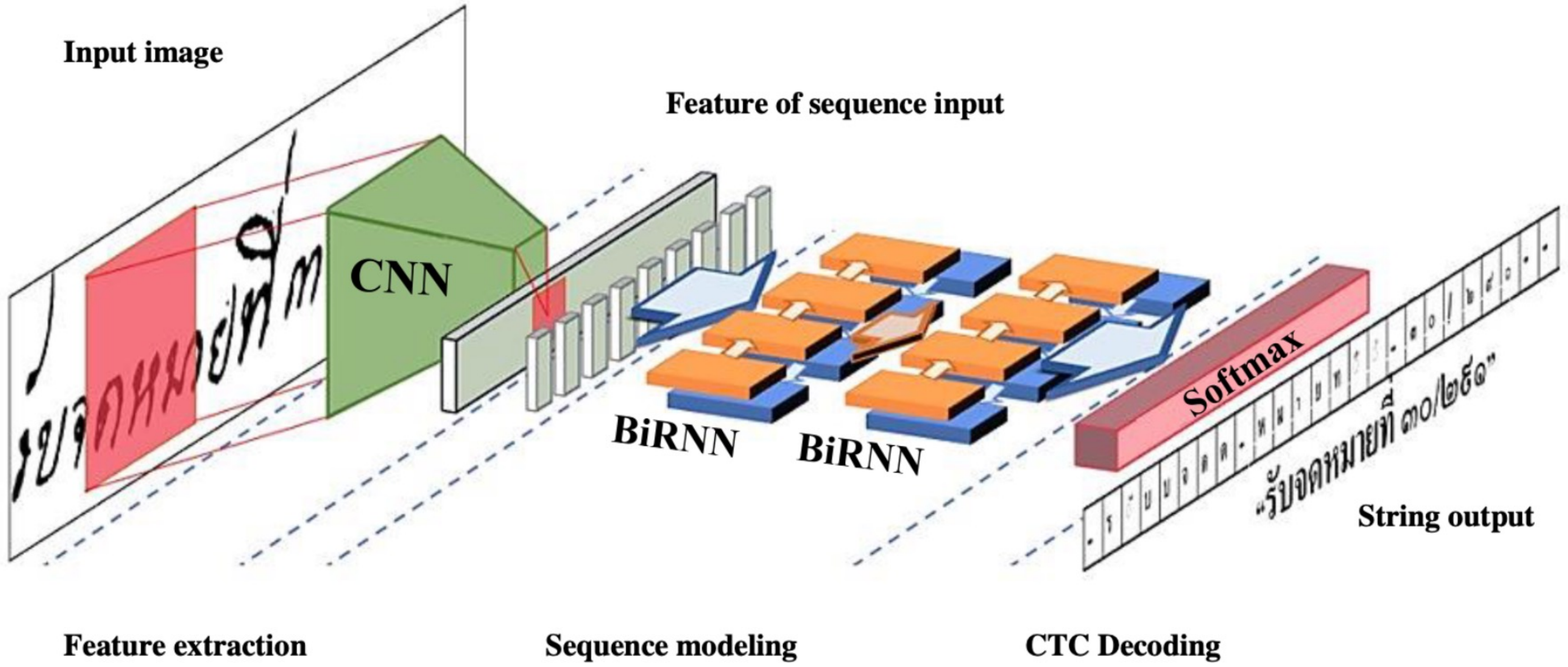


Figure 2 Overview framework of convolutional recurrent neural networks

Details

Table1 Configuration details of CRNN architectures

CCNet	mCCNet-64	mCCNet-512	mVGG16	mVGG19	mResNet50	mDenseNet-121	mMobileNet-V2	mEfficientNet-B1
6 weighted layers	7 weighted layers	7 weighted layers	14 weighted layers	16 weighted layers	26 weighted layers	43 weighted layers	23 weighted layers	29 weighted layers
Input image (64, 504, 1)			Input image (64, 504, 3)					
Conv3-16 Maxpool2-s2	Conv3-16 Maxpool2-s2	Conv3-16 Maxpool2-s2	Conv3-16 Conv3-16 Maxpool2-s2	Conv3-64 Conv3-64 Maxpool2-s2	Conv7-64 Maxpool3-s2	Conv7-64 Maxpool3-s2	Conv3-32-s2 DwConv3-32	Conv3-32-s2 DwConv3-32
Conv3-32 Maxpool2-s2	Conv3-32 Maxpool2-s2	Conv3-32 Maxpool2-s2	Conv3-128 Conv3-128 Maxpool2-s2	Conv3-128 Conv3-128 Maxpool2-s2	$\begin{bmatrix} \text{Conv1} - 64 \\ \text{Conv3} - 64 \\ \text{Conv1} - 256 \end{bmatrix}$ x3	$\begin{bmatrix} \text{Conv1} - 128 \\ \text{Conv3} - 32 \end{bmatrix}$ x6 Conv1-128 Avgpool2-s2	Conv1-16 Conv1-96 DwConv3-96	Conv1-16 Conv1-96 DwConv3-96 SE
Conv3-32 Maxpool2-s2	Conv3-32 Maxpool2-s2	Conv3-32 Maxpool2-s2	Conv3-256 Conv3-256 Conv3-256 Maxpool2-s2	Conv3-256 Conv3-256 Conv3-256 Conv3-256 Maxpool2-s2	$\begin{bmatrix} \text{Conv1} - 128 \\ \text{Conv3} - 128 \\ \text{Conv1} - 512 \end{bmatrix}$ x4	$\begin{bmatrix} \text{Conv1} - 128 \\ \text{Conv3} - 32 \end{bmatrix}$ x12 Conv1-512	$\begin{bmatrix} \text{Conv1} - 24 \\ \text{Conv1} - 144 \\ \text{DwConv3} - 144 \end{bmatrix}$ x2	$\begin{bmatrix} \text{Conv1} - 24 \\ \text{Conv1} - 144 \\ \text{DwConv3} - 144 \\ \text{SE} \end{bmatrix}$ x2
-	-	-	Conv3-512 Conv3-512 Conv3-512	Conv3-512 Conv3-512 Conv3-512 Conv3-512	-	-	$\begin{bmatrix} \text{Conv1} - 32 \\ \text{Conv1} - 192 \\ \text{DwConv3} - 192 \end{bmatrix}$ x3	$\begin{bmatrix} \text{Conv1} - 40 \\ \text{Conv1} - 240 \\ \text{DwConv5} - 240 \\ \text{SE} \end{bmatrix}$ x2
-	Conv1-64	Conv1-512	Conv1x1-512					
Global average pooling								
Bidirectional RNN-(N)								
Bidirectional RNN-(N)								
FC, Softmax (94)								
CTC decoding								

Detail

Table 1. Our proposed CNN-LSTM architecture

Stage	Operators	Resolution	Channels	Layers
1	Conv 3×3	32×379	64	2
2	Max-pooling 2×2			1
3	Conv 3×3	16×189	128	2
4	Max-pooling 2×2			
5	Conv 3×3	8×94	256	3
6	Max-pooling 2×2			
7	Conv 3×3	4×94	512	3
8	Max-pooling 2×1			
9	Bi-LSTM	94	256	2
10	Dense & Softmax Function	157		1
11	CTC Loss Function			

Material and Method

3. **Video Subtitle Dataset.** The video subtitle images used in this experiment were collected from 24 videos shared on Facebook and Youtube. The subtitle text included Thai and English languages, including Thai characters, Roman characters, Thai numerals, Arabic numerals, and special characters with 157 characters in total, as shown in Table 2.

FIGURE 5. Examples of subtitle text images and labels

Material and Method

4. Handwritten character datasets

We evaluate the different handwritten character recognition methods on three isolated handwritten script datasets belonging to three languages (Thai, Bangla, and Latin), all of which are composed of handwritten characters and digits. The original handwritten scripts in the datasets are not normalized to a fixed-size image and therefore are in numerous pixel space resolutions. Furthermore, we have manually collected a new Thai handwritten script dataset that contains 24,045 character images in total from various writers. The details of the Thai handwritten dataset are described in Section 4.1.

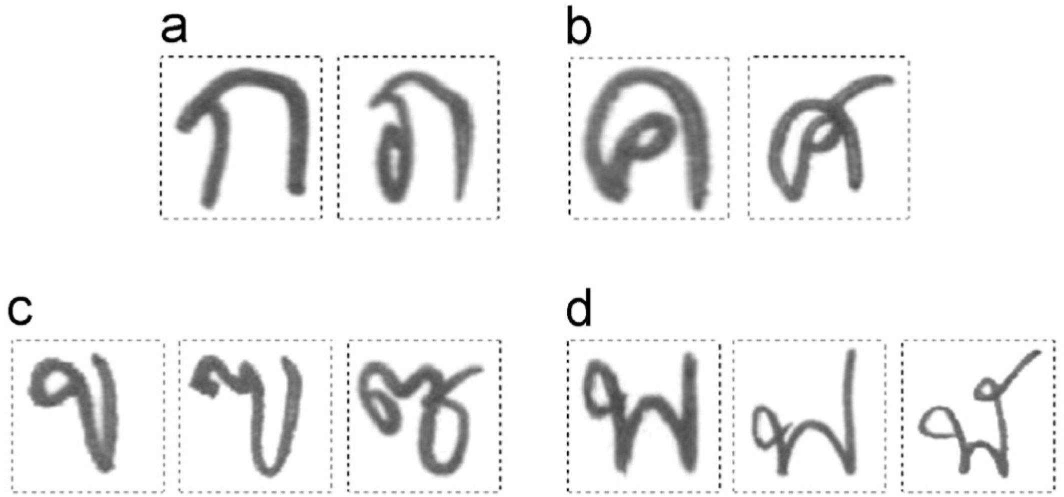


Fig. 4. Illustration of the relation between different Thai characters. In (a) and (b), the second character is constructed by slightly changing the first (different) character. In (c) and (d) the third is created by a modification of the second character, which is a modification of the first character.

Example

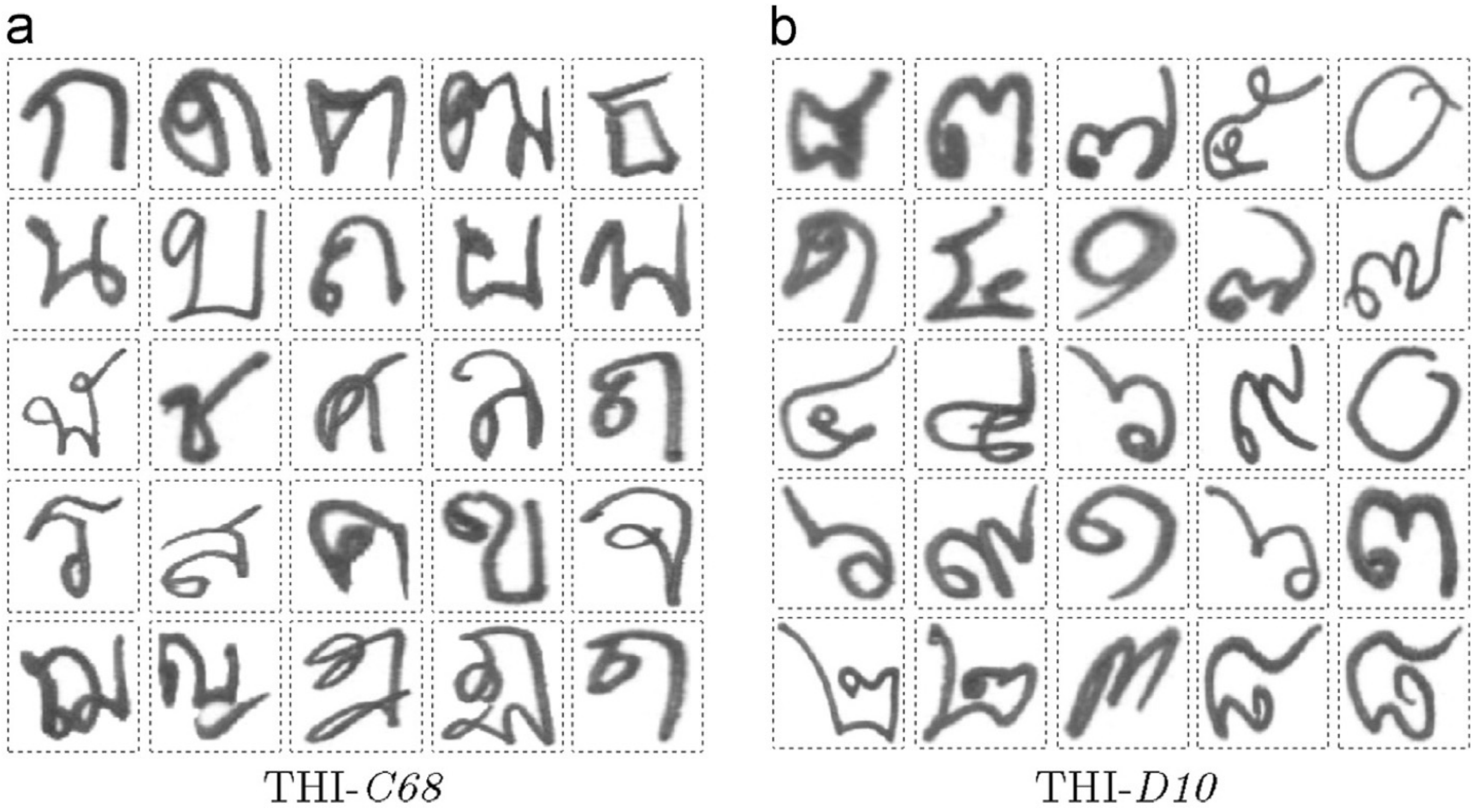


Fig. 6. Illustration of the Thai handwritten images. (a) Thai handwritten characters, and (b) Thai handwritten digits.

Overview of the datasets

Table 1
Overview of the handwritten datasets.

Dataset	No. of classes	Train	Test
THI-*C*68	68	13,130	1360
THI-*D*10	10	8055	1500
BANG-*C*45	45	4627	900
BANG-*D*10	10	9161	1500
LATIN-*C*25	25	26,329	11,287
LATIN-*D*10	10	1637	880

Experimental Results

- Experimental Setup
- Evaluation Metric
- Experiments
- Comparison with existing methods

Experimental Setup

4.2 Experimental setup

We implement the proposed framework with the TensorFlow platform. All experiments were performed on a Linux operating system with Intel(R) Core(TM) i7-4790 Processor 3.6GHz, 16GB DDR4 RAM. As explained in Section 3, we first used pre-trained models of six CNN architectures; VGG16, VGG19, ResNet50, DenseNet201, MobileNetV1, and MobileNetV2, to train and extract the spatial feature from food images. All CNNs were trained using the stochastic gradient descent (SGD) optimizer, rectified linear unit (ReLU) for activation function, and learning rate between 0.01 to 0.0001. Second, the spatial features were then sent to Conv1D-LSTM and LSTM networks to extract temporal features. In the LSTM network, the fraction of the units was employed to drop the linear transformation of the inputs. The initial weights were randomly selected by using a Gaussian distribution where the mean is zero.

We decided to train only 100 epochs to avoid overfitting when training the model. Figure 10 shows loss values while training the Conv1D-LSTM and LSTM model. According to loss values, better loss values were obtained after epoch 50 when they became stable values until epoch 100.

Evaluation Metric

4.3 Evaluation metrics

The evaluation metrics used for food image recognition were accuracy and F1-score. We used the accuracy score to evaluate the performance of the deep learning models on the test set and used the F1-score to examine the individual accuracy of each class. The accuracy and the F1-score were computed by Equations 2 and 3.

$$accuracy = \frac{TP_k + TN_k}{TP_k + TN_k + FP_k + FN_k} \tag{2}$$

$$F1 - score = 2 \times \frac{\left(\frac{TP_k}{TP_k + FN_k} \times \frac{TP_k}{TP_k + FP_k}\right)}{\left(\frac{TP_k}{TP_k + FN_k} + \frac{TP_k}{TP_k + FP_k}\right)} \tag{3}$$

Where TP_k called true positives, is the number of correctly classified images from class k.
FP_k called false positives, is the number of misclassified images from class k.
TN_k called true negatives, is the number of correctly classified image that does not belong to class k.
FN_k called false negatives is the number of misclassified images belong to class k.

4.3 Quantitative evaluation

In this section, we evaluate CRNN architectures on the Thai archive manuscript dataset using character-level error rate (CER) as the evaluation metric. We also compare nine state-of-the-art CRNN models regarding the number of parameters and training time. Moreover, we evaluate the new data augmentation strategy (CycleAugment) and compare our CycleAugment strategy with the original data augmentation strategy. Both strategies apply data augmentation techniques based on transformation techniques, including random shifting, rotation, and shearing. In addition, we evaluate the CRNN models that train from scratch and use the transfer learning technique to understand wherewith the transfer learning technique affects the CRNN models.

The performance of the handwritten text recognition was evaluated based on the CER. CER was calculated as the minimal Levenshtein distance, which is the number of single-character modifications that change the predictive text from the ground truth transcription of the word [46]. There are three operations of the CER metric: insertion, deletion, and substitution. The CER is calculated by the following Equation:

$$CER = \frac{I + S + D}{N} \qquad (20)$$

where I is the number of character insertions, S is the number of character substitutions, D is the number of character deletions, and N is the total number of characters in the target text.

Evaluation Method

Experiments

- ควรแบ่งเป็น Section ให้ขัดเจน
- แต่ละ Section ควรมีตาราง (Table) ประกอบ

4.4 Experiments with deep learning methods

In the experiments with deep learning methods, we first trained the Food-101 dataset using a pre-trained model of six state-of-the-art CNNs; VGG16, VGG19, MobileNetV1, MobileNetV2, ResNet50, and DenseNet201. Second, we proposed the deep feature method to extract the spatial feature from the last pooling layer of each CNN. The deep feature method extracted a high dimension of the spatial feature. The number of spatial features is reported in Table 4. It can be seen that ResNet50 provided 99,176 features. On the other hand, VGG16 produced only 25,088 features. Finally, we trained the high dimension of the spatial features using Conv1D-LSTM and LSTM networks.

Table 4 Illustration of the number of spatial features extract from different CNN architectures and size of each model

Deep feature methods	No. of parameters	No. of features
VGG16	14.7M	25,088
VGG19	20M	25,088
ResNet50	23.5M	99,176
DenseNet201	18.3M	94,080
MobileNetV1	3.2M	50,176
MobileNetV2	2.2M	62,720

Experiments

ตารางผลการทดลองควรมีอย่างน้อย 4 ตาราง

4.4 Performance of different combination of CRNNs

To evaluate the performance of CRNN architectures, we resized all images to 64x496 pixels and used them as the input to the CRNN architectures. We trained all the CRNN models using the Keras framework with TensorFlow backend and trained on Google cloud with NVIDIA Tesla P100 GPU with 16GB of RAM.

For the training process, we divided the Thai archive manuscript dataset with the ratio of 70:10:20 for training, validation, and test, respectively. The nine CRNN networks (see Table 1) were combined with two types of BiRNNs: BiLSTM and BiGRU. The number of RNN sizes with 128, 256, and 512 neurons was evaluated.

The CRNN networks were trained with the following parameters: 200 epochs, batch size of 32, Adam optimizer with learning rate of 0.001, the first- and second-moment estimate values of 0.9 and 0.999, and epsilon of 1e-07.

Table 3 Comparison of the parameters and computational time between different backbones CNNs and RNN sizes

Models	No. of parameters			Training time (hh:mm)			Character error rate (%)		
	RNN sizes								
	128	256	128	128	256	128	128	256	128
CCNet-BiGRU [26]	0.49M	1.75M	6.62M	00:26	00:27	00:30	13.32	14.54	14.43
CCNet-BiLSTM [26]	0.64M	2.30M	8.78M	00:26	00:27	00:34	14.54	14.81	15.29
mCCNet-64-BiGRU	0.50M	1.75M	6.62M	00:26	00:27	00:33	15.19	16.23	16.48
mCCNet-64-BiLSTM	0.64M	2.31M	8.78M	00:26	00:27	00:30	16.09	16.64	14.36
mCCNet-512-BiGRU	0.87M	2.47M	8.03M	00:26	00:27	00:30	14.48	16.15	15.70
mCCNet-512-BiLSTM	1.13M	3.26M	10.65M	00:26	00:26	00:33	14.26	12.69	11.35

Comparison with Existing Methods

Benchmark dataset

Reference within the text to Table, Figure

ทุกตารางและทุกรูปที่ปรากฏในบทความจะต้องกล่าวถึง

4.5 Comparison between ResNet50+Conv1D-LSTM network and previous methods

We made extensive comparisons between our ResNet50+Conv1D-LSTM network and existing state-of-the-art CNN architectures. The experimental results showed that our network performed better than all CNN architectures. The accuracy of 90.87% was obtained from the ResNet50+Conv1D-LSTM, while, the performance of the state-of-the-art WISeR architecture was 90.27% accuracy. The comparative results between the existing CNN architectures and our proposed architecture on the Food-101 dataset are shown in Table 7.

Table 7 Recognition performance on the Food-101 dataset when compared with different deep learning techniques.

Architectures	No. of training images per class	Accuracy	References
ResNet152	750	64.98	McAllister et al. [2]
EnsembleNet	750	72.12	Pandey et al. [17]
Modified MobileNetV1	400	72.59	Phiphiphatphaisit & Surinta [38]
DeepFood	750	77.40	Liu et al. [16]
GoogLeNet	750	79.20	Bolanos & Radeva [22]
CNNs Fusion	750	86.71	Aguilar et al. [18]
InceptionV3	750	88.28	Hassannejad et al. [15]
WISeR	750	90.27	Martinel et al. [20]
ResNet50+Conv1D-LSTM	750	90.87	Our proposed

Our Method

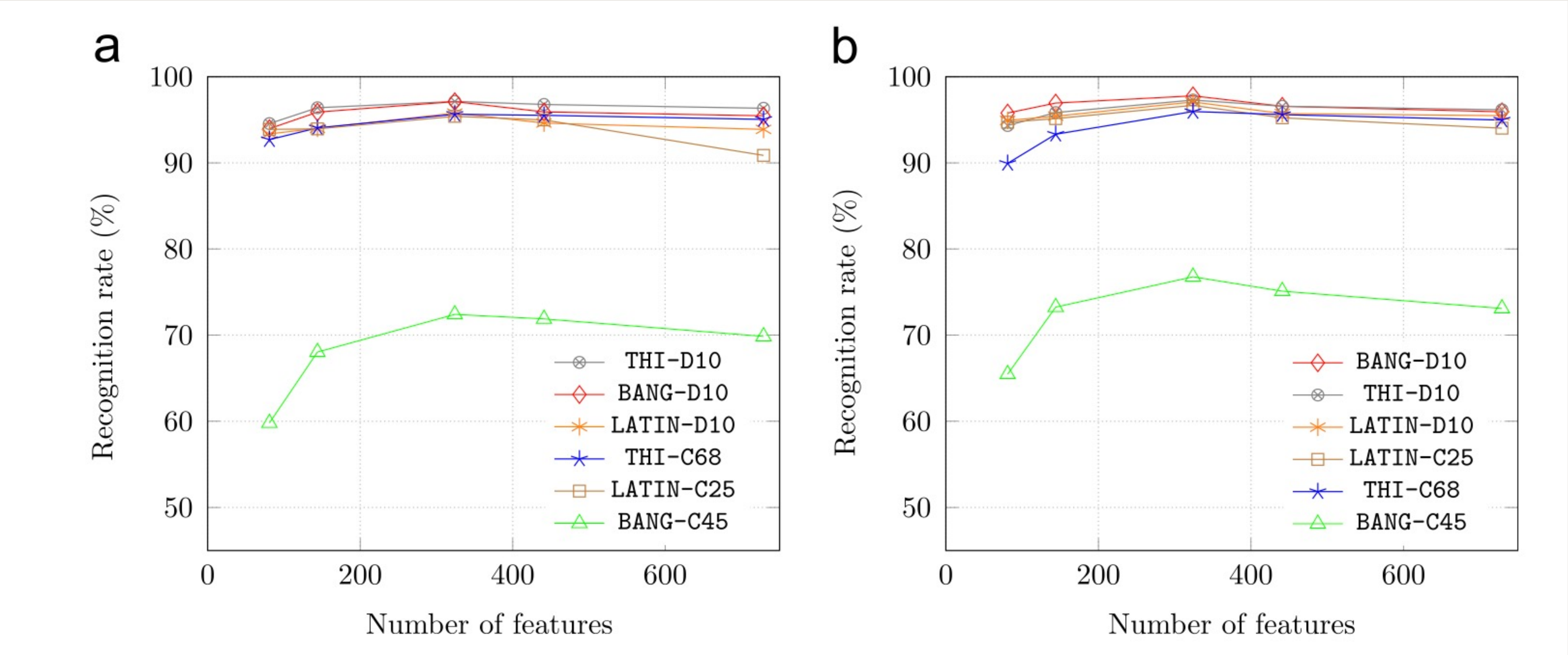
Comparison with Existing Methods

ทำได้สองรูปแบบ

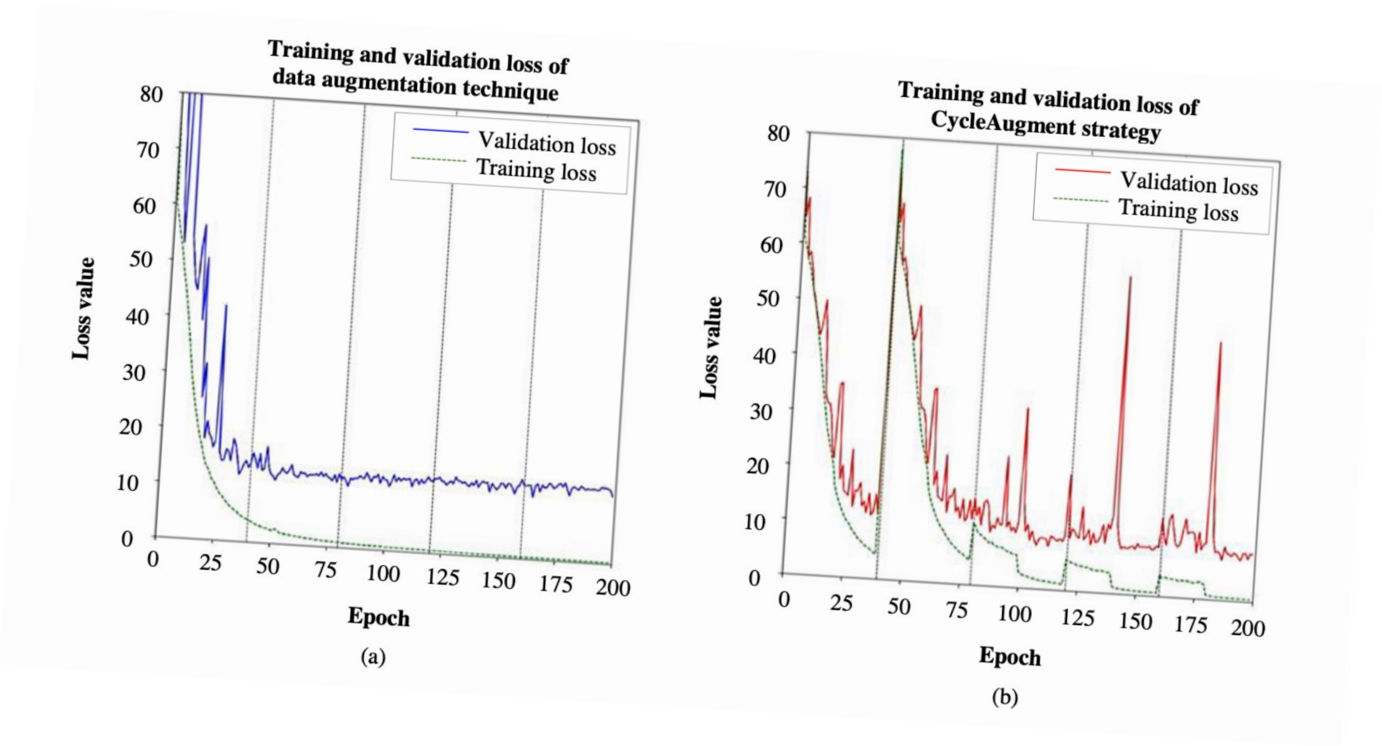
- กรณีที่ใช้ชุดข้อมูลมาตรฐาน (Benchmark dataset)
  - สามารถดูผลการทดลองจากบทความอื่นและนำมาเขียนเปรียบเทียบได้เลย
- กรณีที่เปรียบเทียบกับวิธีอื่น ๆ กับชุดข้อมูลที่เก็บรวบรวมมาเอง *– ต้องทดลองวิธีอื่น ๆ กับชุดข้อมูลที่เก็บรวบรวมขึ้นมาเอง*

Visualization

หากเป็นไปได้อย่าให้ข้อมูลซ้ำกับที่แสดงในตาราง



Visualization



(a) (b)

Visualization

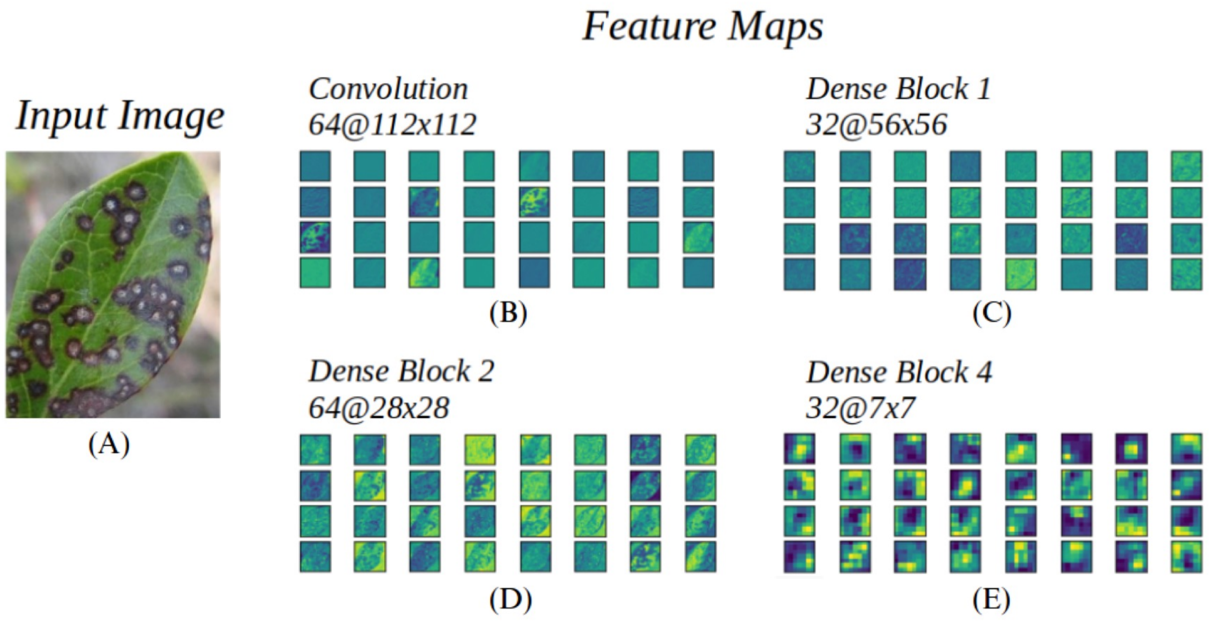
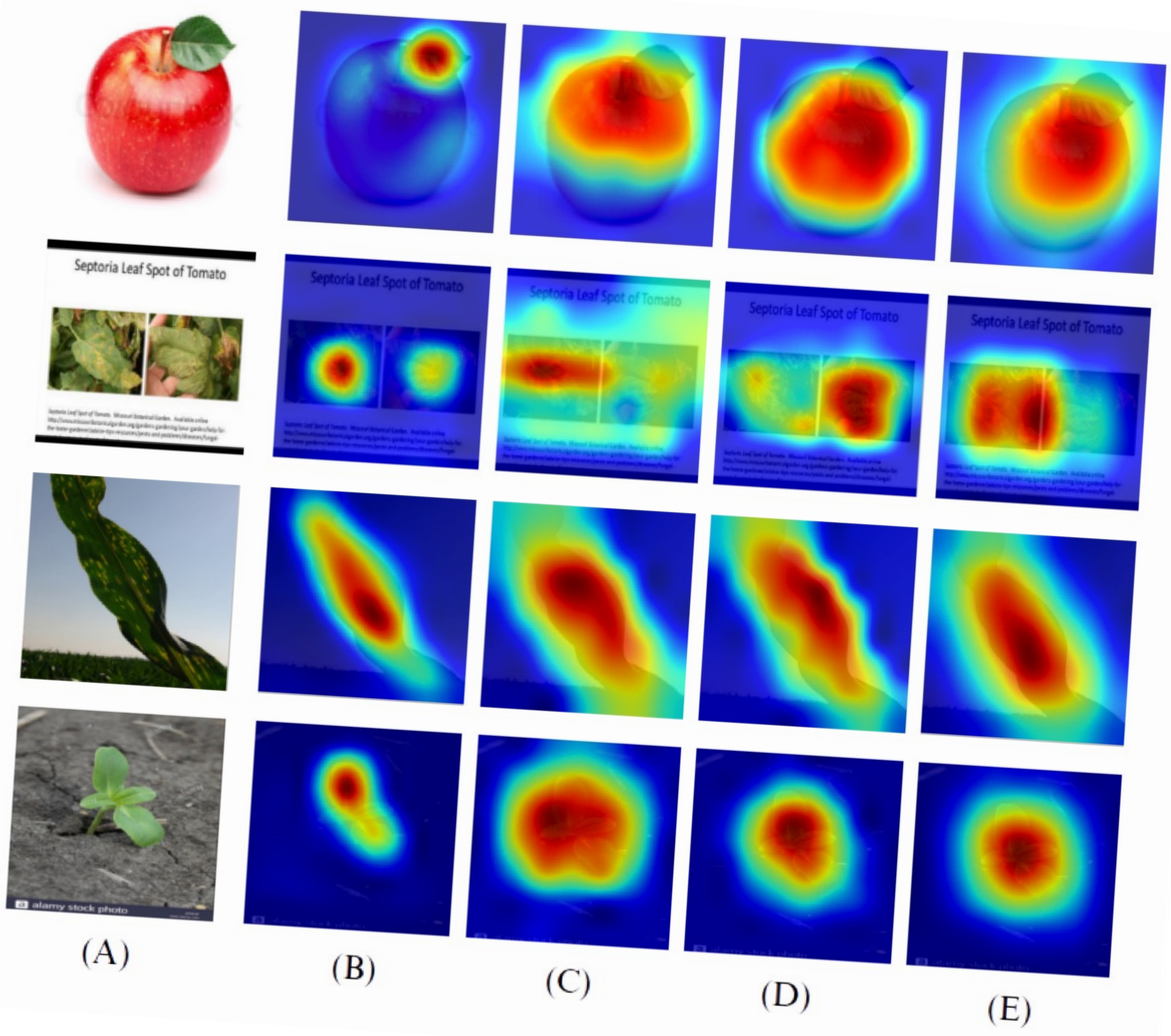


FIGURE 4. The feature maps visualization of the DenseNet201 architecture: (A) Input image, (B) first convolution layer, the output from (C) dense block layer 1, (D) dense block layer 2, and (E) dense block layer 4

Visualization



ทุกบทความควรมี Discussion

5. Discussion

5.1 CycleAugment strategy

It is known that deep learning requires data augmentation techniques to improve performance and avoid overfitting problems. To create the robust CRNN model, we then applied the data augmentation technique. The experimental results showed that the data augmentation techniques did not always confirm the best performance. Consequently, the CRNN model will find only the global minima value when training the CRNN with the original data augmentation strategy. The training loss never again increases, as shown in Figure 6(a). Indeed, it increases the chance of encountering overfitting problems.

We then proposed the new cyclical learning method, namely the CycleAugment strategy. The proposed strategy can effectively improve the performance of the handwritten text recognition by escaping the trapping in global minima and overfitting problems. The CycleAugment strategy increases the chances of discovering local minima in each cycle by switching between two training states with and without applying data augmentation while training the CRNN model, as shown in Figure 6(b). The CRNN model adapted to the local minima because the weight of the CRNN architecture is adjusted using a high error gradient value obtained from variation of the input images.

Discussion

5.2 Effectiveness of transfer learning technique

We have learned from many studies that the transfer learning technique consistently performed better than scratch learning [31, 32, 34]. Therefore, we evaluated the performance of the scratch and transfer learning, as shown in Table 5 and 6. The experimental results were quite surprising in that the transfer learning performance did not significantly outperform the scratch learning. However, in the CRNN architecture, we discovered that the transfer learning did not show outstanding results because the number of transfer parameters from the pre-trained CNN model was more limited than the parameters in the RNN architecture. We have to train the RNN model with a huge number of parameters that did not transfer from the pre-trained model. The parameters of the RNN architecture are larger, approximately four times more than the CNN architecture.

Discussion

Conclusion

คล้ายๆ กับ Abstract --- ***ห้าม Copy จาก Abstract*
--- ให้เขียนใหม่ทุกครั้ง

กล่าวถึง Method

กล่าวถึง ผลการทดลอง

5. Conclusions

This study proposed the ResNet50+Conv1D-LSTM network for accurate food image recognition. First, our network took advantage of extracting the robust spatial feature using a state-of-the-art convolutional neural network (CNN), called ResNet50 architecture. Second, we used the robust feature as input data for the Conv1D combined with the long short-term memory (LSTM) network, namely Conv1D-LSTM. The primary function of the Conv1D-LSTM network was to extract a temporal feature. Finally, the softmax function was employed to transforms the output of the Conv1D-LSTM into a probability distribution.

In the experiments, we evaluated six CNNs; VGG16, VGG19, ResNet50, DenseNet201, MobileNetV1, and MobileNetV2 to extract the feature, then classify with Conv1D-LSTM and LSTM networks on the Food101 dataset. The results showed that the ResNet50 combined with the Conv1D-LSTM network, called ResNet+Conv1D-LSTM network, provided the best performance (see Table 5). Additionally, we experimented with mixed data augmentation techniques; rotation, width shift, height shift, horizontal flip, shear, and zoom. The result of the data augmentation also insignificantly increased accuracy by 0.27%. Our experiments presented better results than previous work (see Table 7). The best result of the ResNet+Conv1D-LSTM obtained 90.87% on the Food-101 dataset.

In future work, we will experiment on increasing the performance of the food image recognition. We will consider other novel data augmentation techniques, which could be more efficient in the noise food images. Also, the ensemble and parallel networks will be involved in future work.

ต้องมี Future work
เสมอ

Acknowledgments

7. Acknowledgments

This research was funded under the Royal Golden Jubilee Ph.D. Program by the Thailand Research Fund (Grant No. PHD/0210/2561).

- *ไม่ต้องกล่าวขอบคุณพ่อแม่ ผู้มีพระคุณ ครูอาจารย์ในอดีต*
- *กล่าวถึงแหล่งให้ทุน*
- *กล่าวถึงคนสนับสนุน เช่น เครื่องมือวิจัย*

References

- หากจำนวนหน้าไม่จำกัด จำนวนของ Reference ควรมีอย่างน้อย 30 งาน
- หากจำกัดจำนวนหน้า เช่น ไม่เกิน 8 หน้า Reference ก็ประมาณ 12-15 งาน
- ควรมี Up-to-date อย่างน้อย 3-5 งาน
- สามารถ Reference บทความเก่า ๆ ได้ ในกรณีที่อ้างอิงถึงผู้คิดค้นทฤษฎี

Thank you for your attention