Computer Vision for Embedded Systems

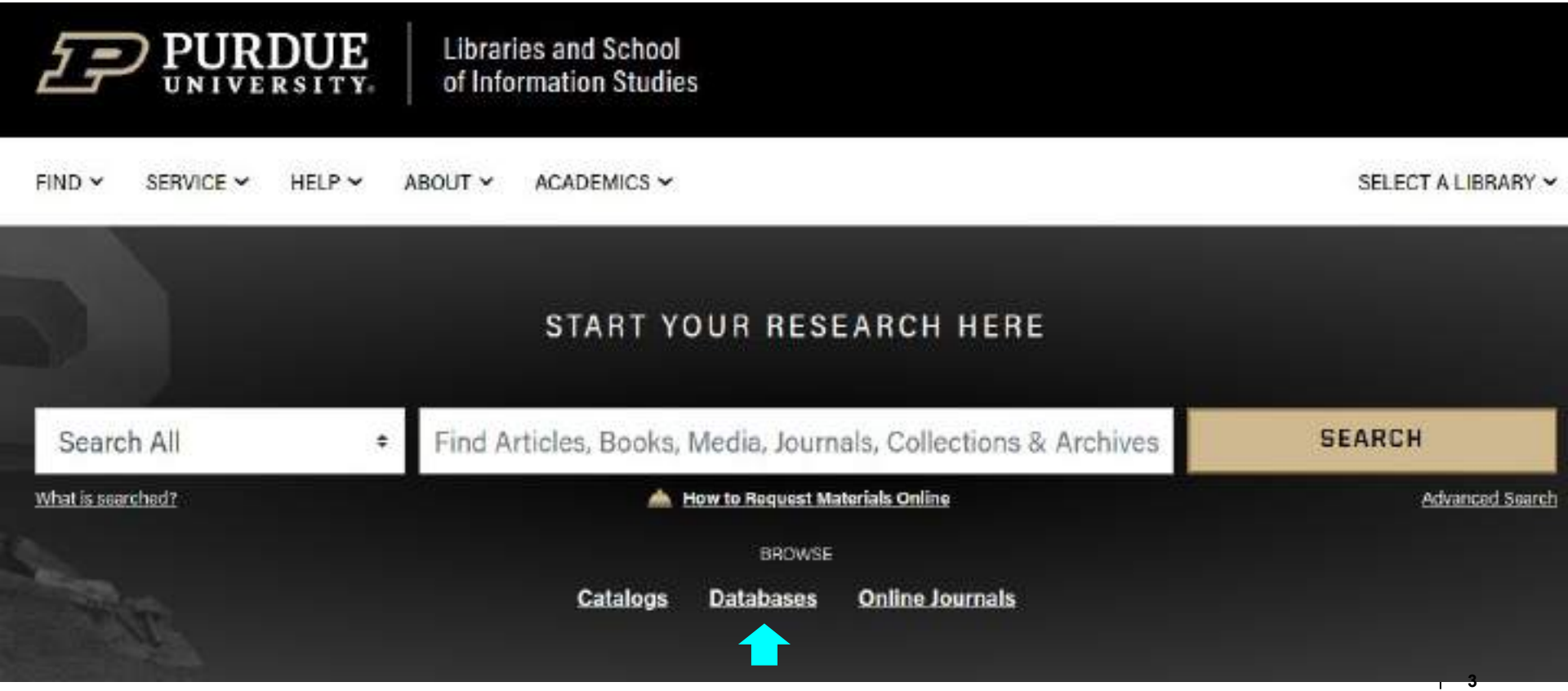
Yung-Hsiang Lu
Purdue University
yunglu@purdue.edu

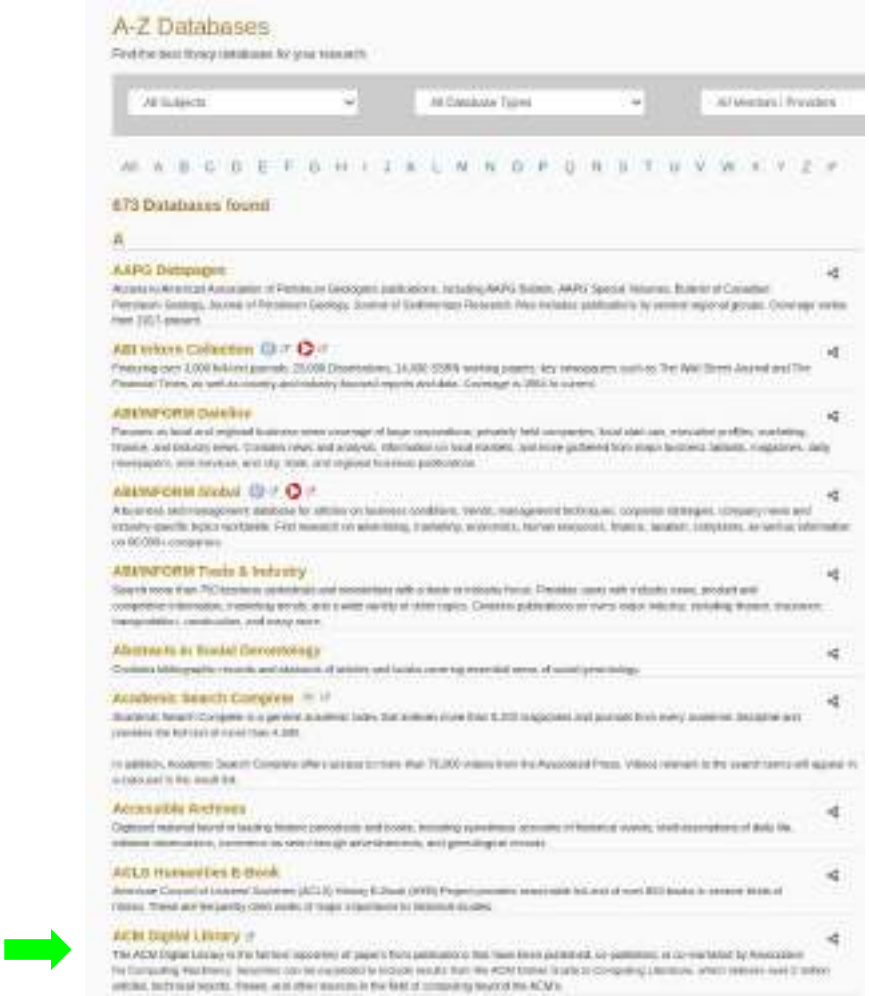
PURDUE UNIVERSITY

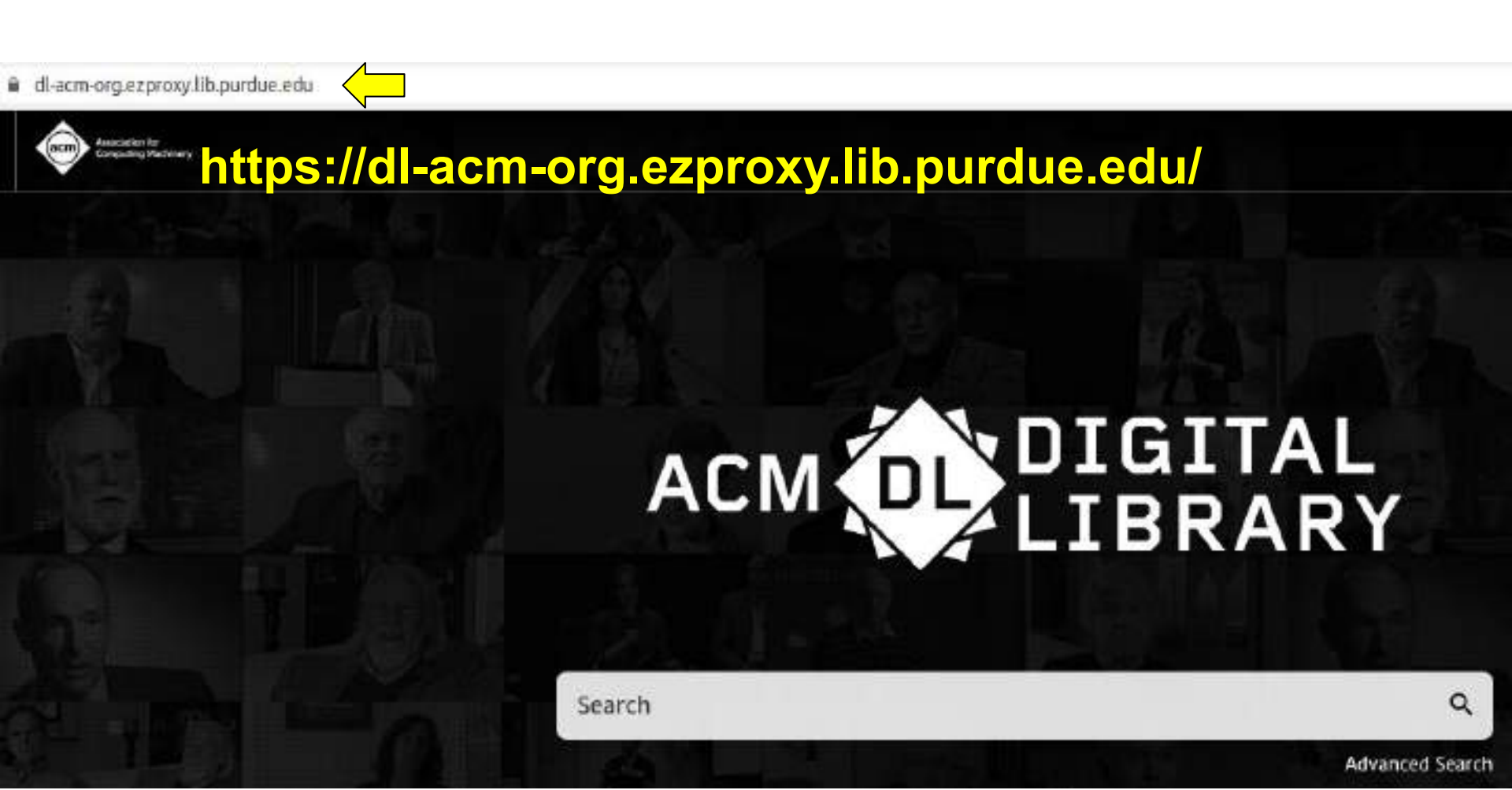
Sources of papers

- IEEExplore, https://ieeexplore.ieee.org/Xplore/home.jsp
- ACM Digital Library: https://dl.acm.org/
- Google scholar
- Computer Vision Foundation open access, https://openaccess.thecvf.com/menu
- arXiv: https://arxiv.org/

Download papers outside campus




A-Z Databases
673 Databases found
A
AAPG Datapages
ABI Inform Collection
ABI/INFORM Dateline
ABI/INFORM Global
ABI/INFORM Trade & Industry
Abstracts in Social Gerontology
Academic Search Complete
Accessible Archives
ACLS Humanities E-Book
ACM Digital Library

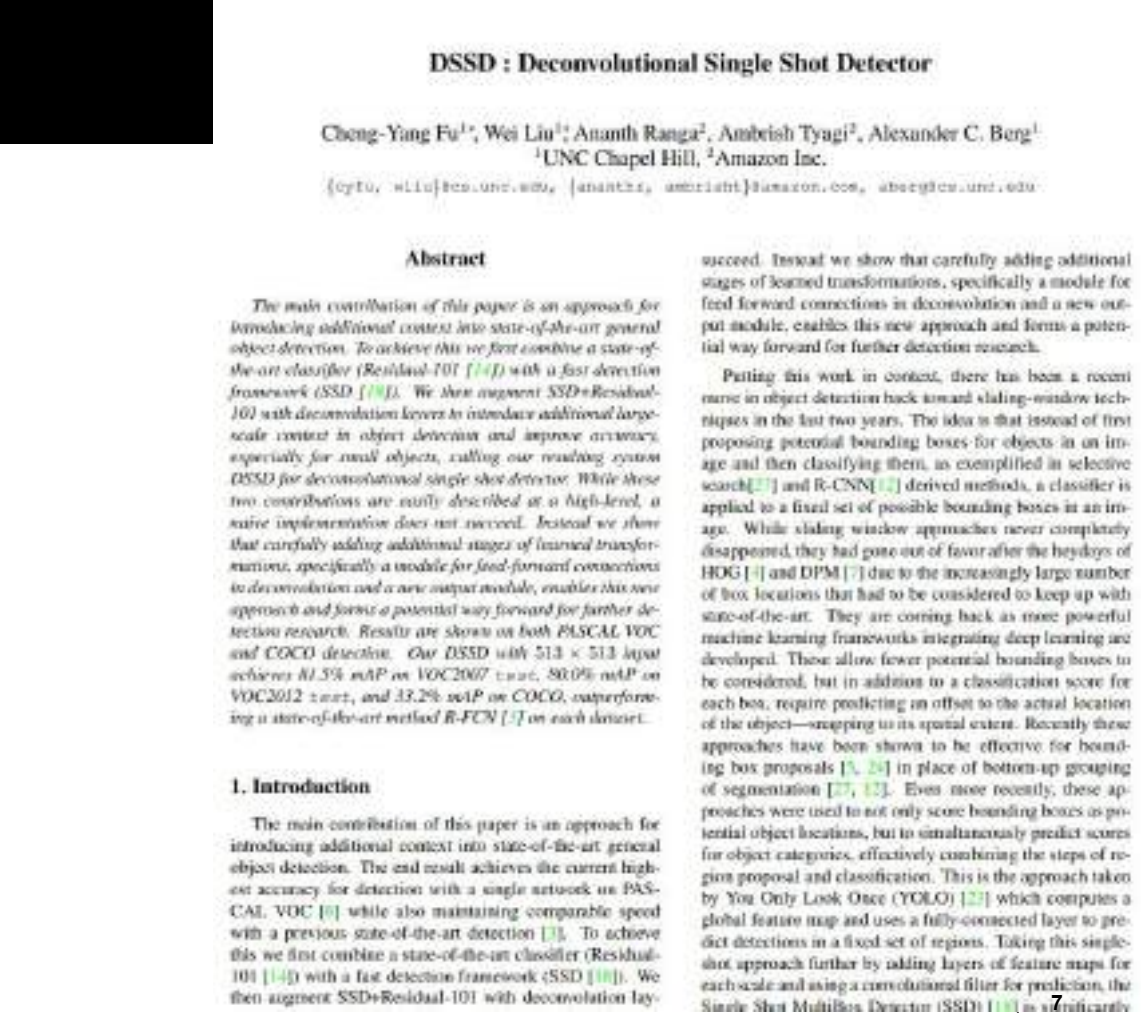

dl-acm-org.ezproxy.lib.purdue.edu
https://dl-acm-org.ezproxy.lib.purdue.edu/
ACM DL DIGITAL LIBRARY
Search
Advanced Search

Types of papers

Type	**Reader**	**Content**	**Authors**
Research	in the same field	New knowledge	Beginners to experts
Survey	wide	Summary of existing work	Experts
Tutorial	beginners in the field	How to start doing research	Beginners to Experts
Magazine	wide	Introduction	Usually experts
Vision	researchers in the field	Important problems to solve	Experts

Research Paper

- Title, Authors
- Abstract, Introduction
- Tight space
- (often) double column



DSSD : Deconvolutional Single Shot Detector

Cheng-Yang Fu[1]*, Wei Liu[1], Ananth Ranga[2], Ambrish Tyagi[2], Alexander C. Berg[1]
[1]UNC Chapel Hill, [2]Amazon Inc.
{cyfu, wliu}@cs.unc.edu, {ananthr, ambrisht}@amazon.com, aberg@cs.unc.edu

Abstract

The main contribution of this paper is an approach for introducing additional context into state-of-the-art general object detection. To achieve this we first combine a state-of-the-art classifier (Residual-101 [14]) with a fast detection framework (SSD [18]). We then augment SSD+Residual-101 with deconvolution layers to introduce additional large-scale context in object detection and improve accuracy, especially for small objects, calling our resulting system DSSD for deconvolutional single shot detector. While these two contributions are easily described at a high-level, a naive implementation does not succeed. Instead we show that carefully adding additional stages of learned transformations, specifically a module for feed-forward connections in deconvolution and a new output module, enables this new approach and forms a potential way forward for further detection research. Results are shown on both PASCAL VOC and COCO detection. Our DSSD with 513 × 513 input achieves 81.5% mAP on VOC2007 test, 80.0% mAP on VOC2012 test, and 33.2% mAP on COCO, outperforming a state-of-the-art method R-FCN [3] on each dataset.

1. Introduction

The main contribution of this paper is an approach for introducing additional context into state-of-the-art general object detection. The end result achieves the current highest accuracy for detection with a single network on PASCAL VOC [6] while also maintaining comparable speed with a previous state-of-the-art detection [3]. To achieve this we first combine a state-of-the-art classifier (Residual-101 [14]) with a fast detection framework (SSD [18]). We then augment SSD+Residual-101 with deconvolution layers to introduce additional large-scale context in object detection and improve accuracy, especially for small objects, calling our resulting system DSSD for deconvolutional single shot detector. While these two contributions are easily described at a high-level, a naive implementation does not succeed. Instead we show that carefully adding additional stages of learned transformations, specifically a module for feed forward connections in deconvolution and a new output module, enables this new approach and forms a potential way forward for further detection research.

Putting this work in context, there has been a recent move in object detection back toward sliding-window techniques in the last two years. The idea is that instead of first proposing potential bounding boxes for objects in an image and then classifying them, as exemplified in selective search[27] and R-CNN[12] derived methods, a classifier is applied to a fixed set of possible bounding boxes in an image. While sliding window approaches never completely disappeared, they had gone out of favor after the heydays of HOG [4] and DPM [7] due to the increasingly large number of box locations that had to be considered to keep up with state-of-the-art. They are coming back as more powerful machine learning frameworks integrating deep learning are developed. These allow fewer potential bounding boxes to be considered, but in addition to a classification score for each box, require predicting an offset to the actual location of the object—snapping to its spatial extent. Recently these approaches have been shown to be effective for bounding box proposals [5, 26] in place of bottom-up grouping of segmentation [27, 12]. Even more recently, these approaches were used to not only score bounding boxes as potential object locations, but to simultaneously predict scores for object categories, effectively combining the steps of region proposal and classification. This is the approach taken by You Only Look Once (YOLO) [23] which computes a global feature map and uses a fully-connected layer to predict detections in a fixed set of regions. Taking this single-shot approach further by adding layers of feature maps for each scale and using a convolutional filter for prediction, the Single Shot MultiBox Detector (SSD) [18] is significantly

Survey Paper (often more than 100 references)

IEEE Access

Received January 4, 2018, accepted February 8, 2018, date of publication February 19, 2018, date of current version March 28, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2807385

Threat of Adversarial Attacks on Deep Learning in Computer Vision: A Survey

NAVEED AKHTAR AND AJMAL MIAN

Department of Computer Science and Software Engineering, The University of Western Australia, Crawley, WA 6009, Australia

Corresponding author: Naveed Akhtar (naveed.akhtar@uwa.edu.au)

This work was supported in part by ARC under Grant DP 160101458.

[180] *Class Central. Deep Learning for Self-Driving Cars*. Accessed: Dec. 2017. [Online]. Available: https://www.class-central.com/mooc/8132/6-s094-deep-learning-for-self-driving-cars

[181] C. Middlehurst. (2015). *China Unveils World's First Facial Recognition ATM*. [Online]. Available: http://www.telegraph.co.uk/news/worldnews/asia/china/11643314/China-unveils-worlds-first-facial-recognition-ATM.html

[182] *About Face ID Advanced Technology*. Accessed: Dec. 2017. [Online]. Available: https://support.apple.com/en-au/HT208108

[183] M. Sharif, S. Bhagavatula, L. Bauer, and M. K. Reiter, "Accessorize to a crime: Real and stealthy attacks on state-of-the-art face recognition," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2016, pp. 1528–1540.

[184] R. Shin and D. Song, "JPEG-resistant adversarial images," in *Proc. Mach. Learn. Comput. Secur. Workshop*, 2017, pp. 1–6.

[185] W. Brendel and M. Bethge. (Apr. 2017). "Comment on 'Biologically inspired protection of deep networks from adversarial attacks.'" [Online]. Available: https://arxiv.org/abs/1704.01547

[186] N. Carlini and D. Wagner. (Nov. 2017). "MagNet and 'efficient defenses against adversarial attacks' are not robust to adversarial examples." [Online]. Available: https://arxiv.org/abs/1711.08478

Tutorial

Computer (Volume: 29, Issue: 3, Mar **1996**)

Theme Feature

Artificial Neural Networks: A Tutorial

Anil K. Jain
Michigan State University

Jianchang Mao
K.M. Mohiuddin
IBM Almaden Research Center

Numerous advances have been made in developing intelligent systems, some inspired by biological neural networks. Researchers from many scientific disciplines are designing artificial neural networks (ANNs) to solve a variety of problems in pattern recognition, prediction, optimization, associative memory, and control (see the "Challenging problems" sidebar).

Conventional approaches have been proposed for solving these problems. Although successful applications can be found in certain well-constrained environments, none is flexible enough to perform well outside its domain. ANNs provide exciting alternatives, and many applications could benefit from using them.[1-3]

This article is for those readers with little or no knowledge of ANNs to help them understand the other articles in this issue of *Computer*. We discuss the motivations behind the development of ANNs, describe the basic biological neuron and the artificial computational model, outline network architectures and learning processes, and present some of the most commonly used ANN models. We conclude with character recognition, a successful ANN application.

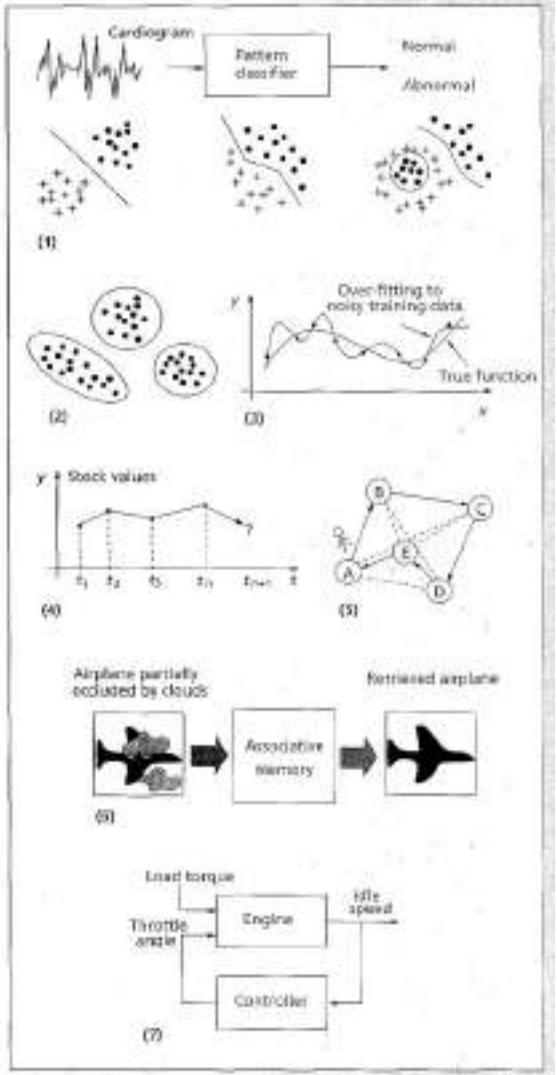


Table 1. Von Neumann computer versus biological neural system.

	Von Neumann computer	Biological neural system
Processor	Complex High speed One or a few	Simple Low speed A large number
Memory	Separate from a processor Localized Noncontent addressable	Integrated into processor Distributed Content addressable
Computing	Centralized Sequential Stored programs	Distributed Parallel Self-learning
Reliability	Very vulnerable	Robust
Expertise	Numerical and symbolic manipulations	Perceptual problems
Operating environment	Well-defined, well-constrained	Poorly defined, unconstrained

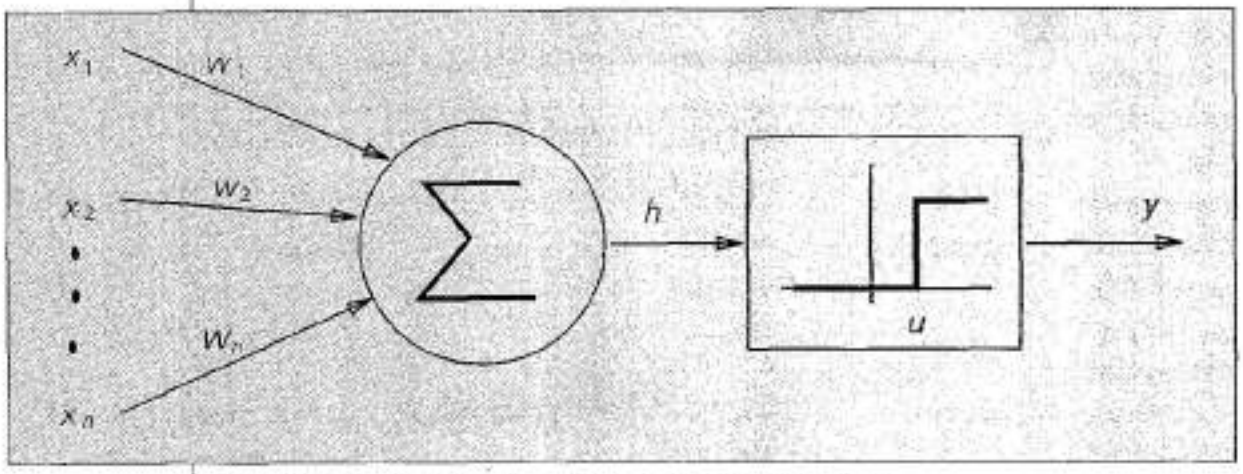


Figure 2. McCulloch-Pitts model of a neuron.

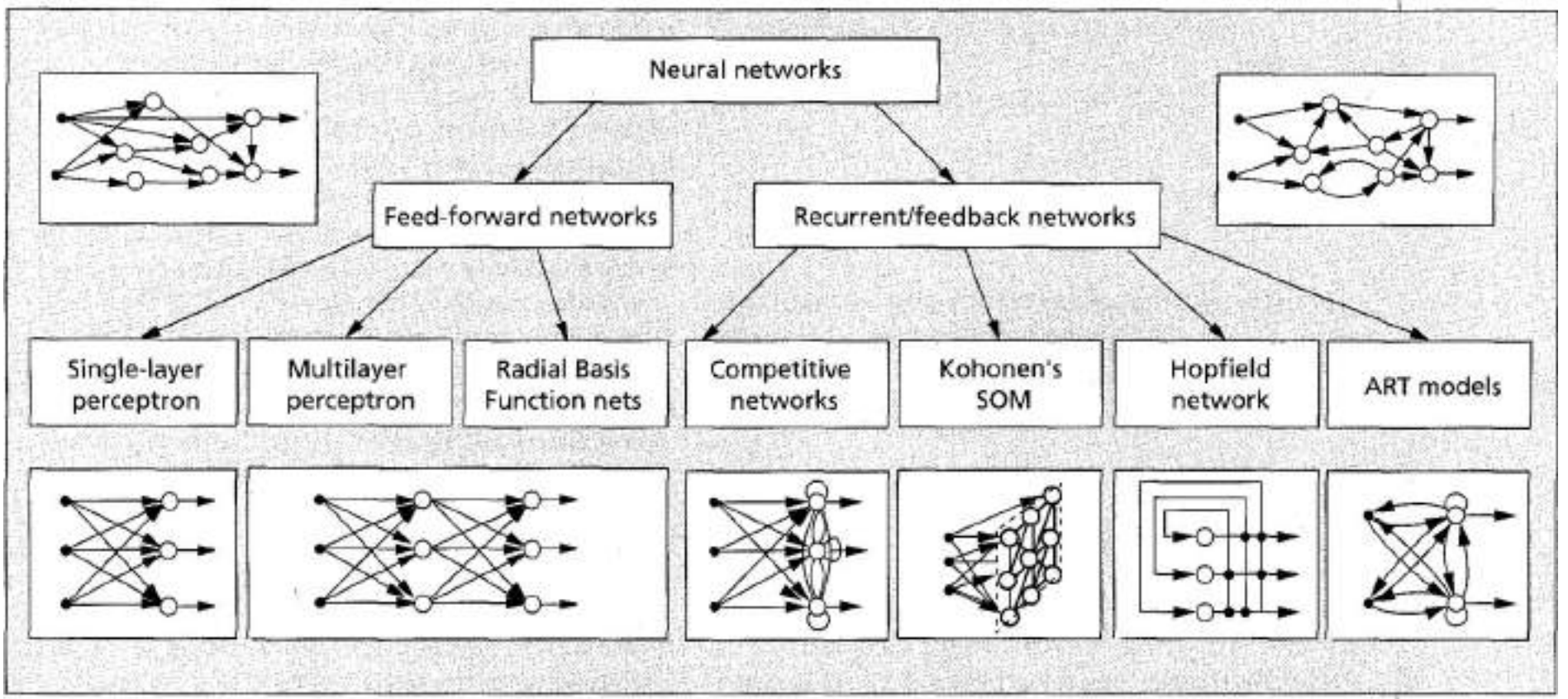
and has the desired asymptotic properties. The standard sigmoid function is the *logistic* function, defined by

$$g(x) = 1/(1 + \exp\{-\beta x\}),$$

where β is the slope parameter.

Network architectures

ANNs can be viewed as weighted directed graphs in which artificial neurons are nodes and directed edges (with weights) are connections between neuron outputs and neuron inputs.


Neural networks
Feed-forward networks
Recurrent/feedback networks
Single-layer perceptron
Multilayer perceptron
Radial Basis Function nets
Competitive networks
Kohonen's SOM
Hopfield network
ART models

Magazine paper

The Future of Human-in-the-Loop Cyber-Physical Systems

Gunar Schirner, Deniz Erdogmus, and Kaushik Chowdhury, *Northeastern University*
Taskin Padir, *Worcester Polytechnic Institute*

A prototyping platform and a design framework for rapid exploration of a novel human-in-the-loop application serves as an accelerator for new research into a broad class of systems that augment human interaction with the physical world.

Human-in-the-loop cyber-physical systems (HiLCPSs) comprise a challenging and promising class of applications with immense potential for impacting the daily lives of many people. As Figure 1 shows, a typical HiLCPS consists of a loop involving a human, an embedded system (the cyber component), and the physical environment. Basically, the embedded system augments a human's interaction with the physical world.

A HiLCPS infers the user's intent by measuring human cognitive activity through body and brain sensors. The embedded system in turn translates the intent into robot control signals to interact with the physical environment on the human's behalf via robotic actuators. Finally, the human closes the loop by observing the physical world interactions as input for making new decisions.

Examples of HiLCPSs include brain-computer interface (BCIs), controlled assistive robots,[1] and intelligent prostheses.

HiLCPS applications offer benefits in many realms—for example, the population of functionally locked-in individuals would benefit tremendously from such systems. Because these individuals cannot interact with the physical world through their own movement and speech, they often must rely heavily on support from caregivers to perform fundamental everyday tasks, such as eating and communicating. As the "Fundamental Autonomy for Functionally Locked-In Individuals" sidebar describes, a HiLCPS could aid in restoring some autonomy by offering alternative interfaces to the cyber-physical environment for interaction, communication, and control.

MULTIDISCIPLINARY CHALLENGES

Designing and implementing a HiLCPS poses tremendous challenges and is extremely time-consuming. Experts from many disciplines need to join forces to successfully solve these challenges.

Transparent interfaces

Traditional dedicated interfaces to the virtual world, such as the keyboard, mouse, and joystick, are less suitable for augmenting human interaction in the physical world. This environment requires transparent interfaces that use existing electrophysiological signals such as electroencephalography (EEG), electrocardiography (ECG), and electromyography (EMG), which measure electrical signals emitted by the brain, heart, and skeletal muscles, respectively. Additional auxiliary sensors

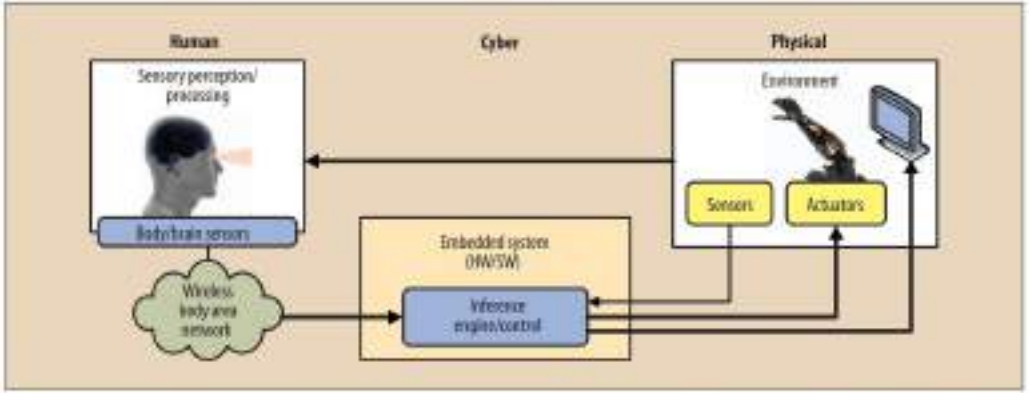


Figure 1. Human-in-the-loop cyber-physical system (HiLCPS). The loop consists of a human, an embedded system, and the physical environment.

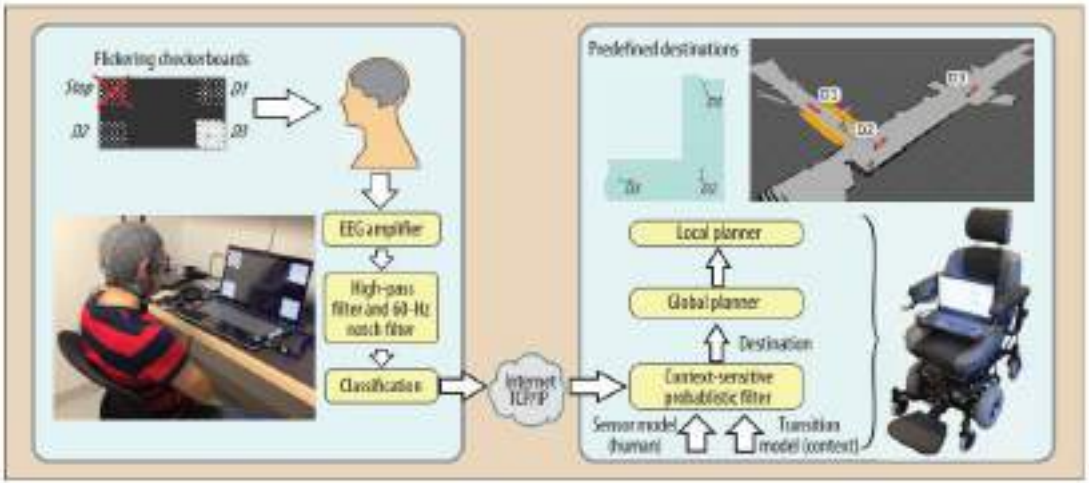


Figure 4. System architecture for the BCI-based control of an intelligent wheelchair as an example of a HiLCPS. (left) The semi-autonomous wheelchair receives brain signals from the user for a high-level activity; (right) it then executes the tasks of path planning, obstacle avoidance, and simultaneous localization and mapping.

Sidebar (common in magazine papers)

to monitor the respiratory rate, pulse oximetry, and skin resistance can help provide a more comprehensive view of the whole body.

The challenge with analog interfaces is in accurately detecting electrophysiological signals; electric potentials can be as low as in the mV range. Moreover, connecting a human to a host of wires to centrally gather these signals is not only impractical but also too restrictive. The optimum solution is to have a distributed sensor network with power-efficient and reliable communication, for example, through wireless body area networks (WBANs).

Human Intent Inference

HiLCPSs put high demands on intent inference algorithm design because input signals are inherently noisy. Intelligent sensor fusion can help compensate for inconsistent measurements of individual sensors and form a complete, coherent picture from the multimodal sensor input. One approach to deal with the interpretation of noisy signals is to take the physical world context into account, eliminating contextually impossible decisions such as actions that are not physically possible given the current state of robot control, or letters that do not make sense given the language used for typing interfaces. Because a HiLCPS continuously interacts with the physical environment, real-time intent inference is crucial for keeping up with the constantly changing environment.

Fundamental Autonomy for Functionally Locked-In Individuals

A human-in-the-loop cyber-physical system (HiLCPS) can offer assistive technology that helps to restore fundamental autonomy—self-feeding and communication—for people who are functionally locked-in due to various neurological or physical conditions. Depending on their clinical diagnosis and condition, these individuals might have full cognitive capabilities yet lack the ability to execute any motor actions that can generate movement or speech. Consequently, they rely heavily on caregivers to accomplish everyday tasks.

We are developing a HiLCPS that augments the neurophysiological capabilities of a functionally locked-in individual to facilitate self-feeding, communication, mobility, and digital access. As depicted in Figure A, we intend to build a brain–computer interface (BCI)–controlled wheelchair as a mobility platform, construct a robotic arm for self-feeding, and establish a communication interface. In addition to restoring the ability to meet basic needs, a HiLCPS can help to close the digital divide, making it possible for users to access the information and social resources that computers offer and contributing to a sense of self-fulfillment that is essential for a productive life.

Figure A. Restoring fundamental autonomy for functionally locked-in individuals through a HiLCPS.

Vision Paper

ACM SIGMOBILE Mobile Computing and Communications Review April 2011

Mobile Computing: the Next Decade

Mahadev Satyanarayanan
School of Computer Science, Carnegie Mellon University

In the inaugural issue of MC2R in April 1997 [24], I highlighted the seminal influence of mobility in computing. At that time, the goal of "information at your fingertips anywhere, anytime" *was only a dream. Today, through relentless pursuit of innovations in wireless technology, energy-efficient portable hardware and adaptive software, we have largely attained this goal. Ubiquitous email and Web access is a reality that is experienced by millions of users worldwide through their Blackberries, iPhones, iPads, Windows Phone devices, and Android-based devices. Mobile Web-based services and location-aware advertising opportunities have emerged, triggering large commercial investments. Mobile computing has arrived as a lucrative business proposition.*

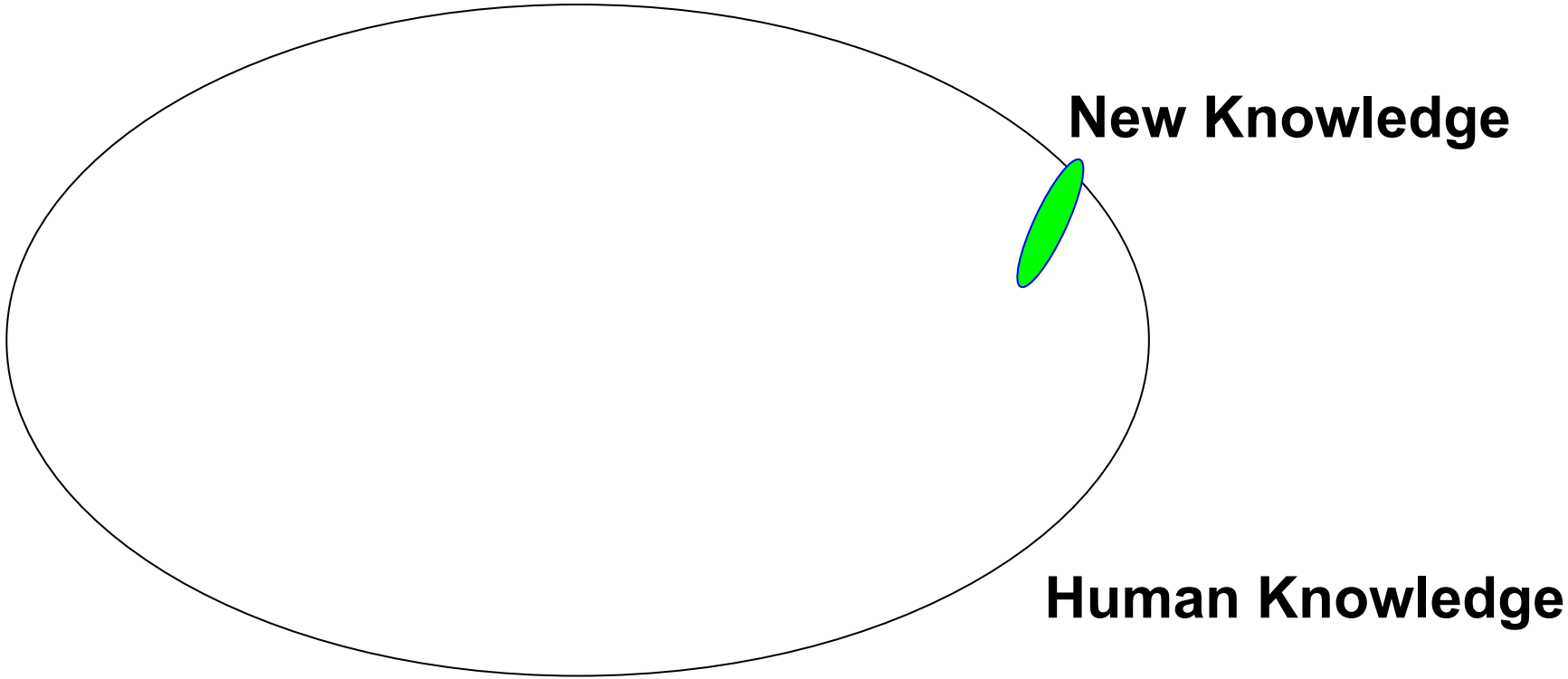
Looking ahead, what are the dreams that will inspire our future efforts in mobile computing? We begin this paper by considering some imaginary mobile computing scenarios from the future. We then extract the deep assumptions implicit in these scenarios, and use them to speculate on the future trajectory of mobile computing.

Write Research Papers

Yung-Hsiang Lu
Electrical and Computer Engineering

PURDUE UNIVERSITY

You have something to say (discovery)

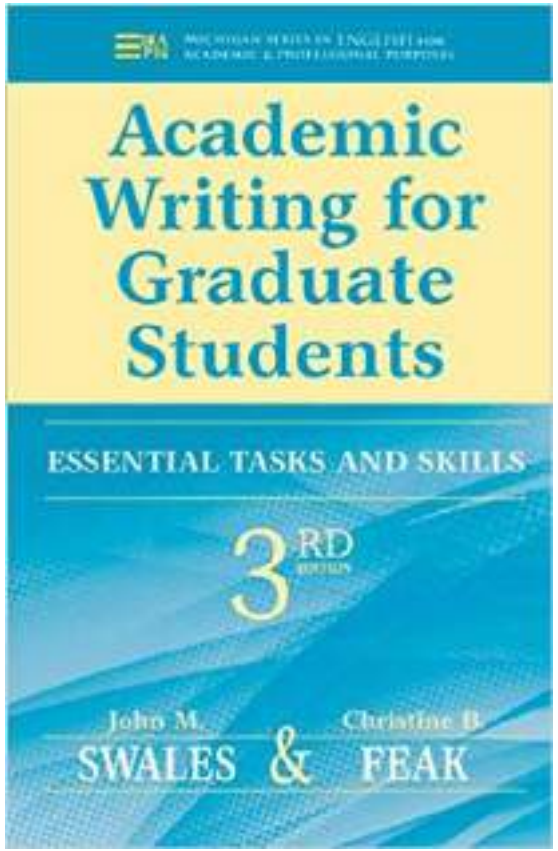


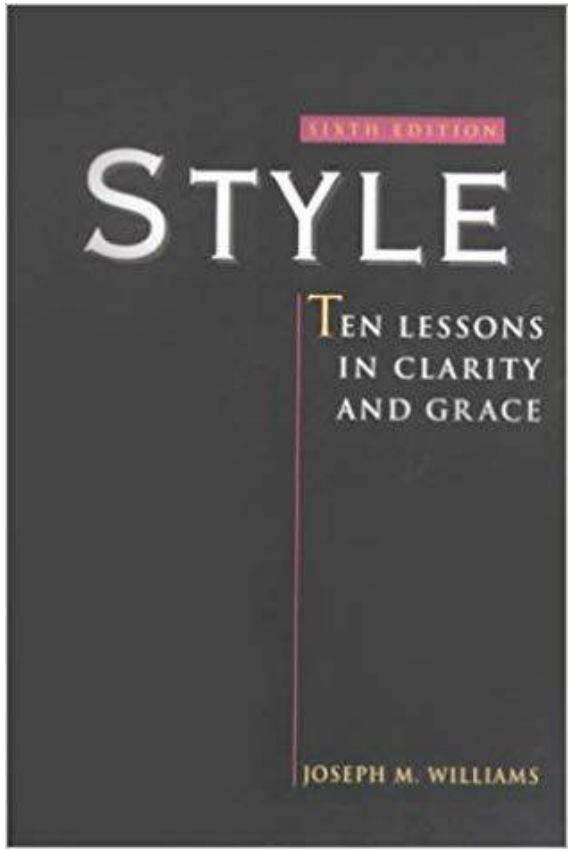
Why do you write research papers?

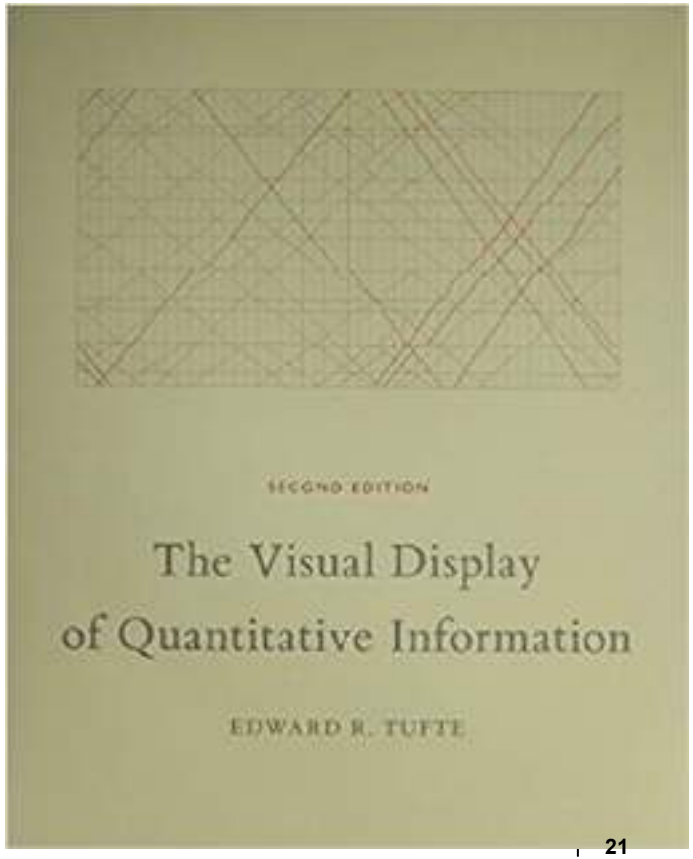
- You have something worth sharing (research findings)
- You are a researcher (or a doctoral student)
- You want to become special and famous
- “Publish or perish” in academia

If I do not want to get PhD, I do not need to know how to write papers ⇒ wrong.
If you invent new technologies, you may want to publish them (as papers or patents or both) so that (1) people understand your inventions (2) your inventions are protected.

Recommended Books







Why to Write Research Papers?

- You have something new to tell the research community
- ***New solution to an existing problem***
- New problem
- New understanding of a problem
- New formulation or interpretation

New to the research community, not new to you.

Anything that can be found in books or papers is not new.

Examples of Research Papers

- New ways to transmit wireless signals
- New methods to protect data from unauthorized access
- New evidence explaining dinosaur extinction
- New disease or new ways of transmitting the disease
- New materials for buildings or airplanes
- New safety features for vehicles
- New relationship between interest rates and economy

When should you start writing a paper?

- The earlier, the better (a least three months)
- Help you organize your thinking
- Guide you design and conduct experiments
- Formulate hypothesis to be validated or overthrown
- Focus your effort on relevant tasks

Reduce “deleted scenes” like movies

Start with "one sentence summary".

The summary guides the creation of the paper. Examples:

- We ***demonstrate*** that masks can effectively reduce the transmission of coronavirus.
- We ***create*** a vaccine for coronavirus; that is 94% effective in phase 3 trials.
- We ***propose*** a method to train machine learning 26% faster than the best existing method.

More examples of one-sentence summary

- We ***discover*** a security vulnerability among 62% of deployed computers.
- We ***build*** the world-first solar-powered airplane that can fly over 1,000 km.
- We ***design*** and ***implement*** a system that can search web pages on the Internet.
- We ***collect*** the data from five million people to analyze the relationship between wealth and longevity.

Understand the different strengths of words: prove, demonstrate, observe, hypothesize

How to write the one-sentence summary

- Select verbs: demonstrate, discover, design, build, create
- Select problems: transmission of diseases, cure of diseases, solar-powered plane, search the Internet
- Set metrics for success: 26% faster, 1000km, 62% deployed computers, 5 million people

Writing is a process of Iterative improvements

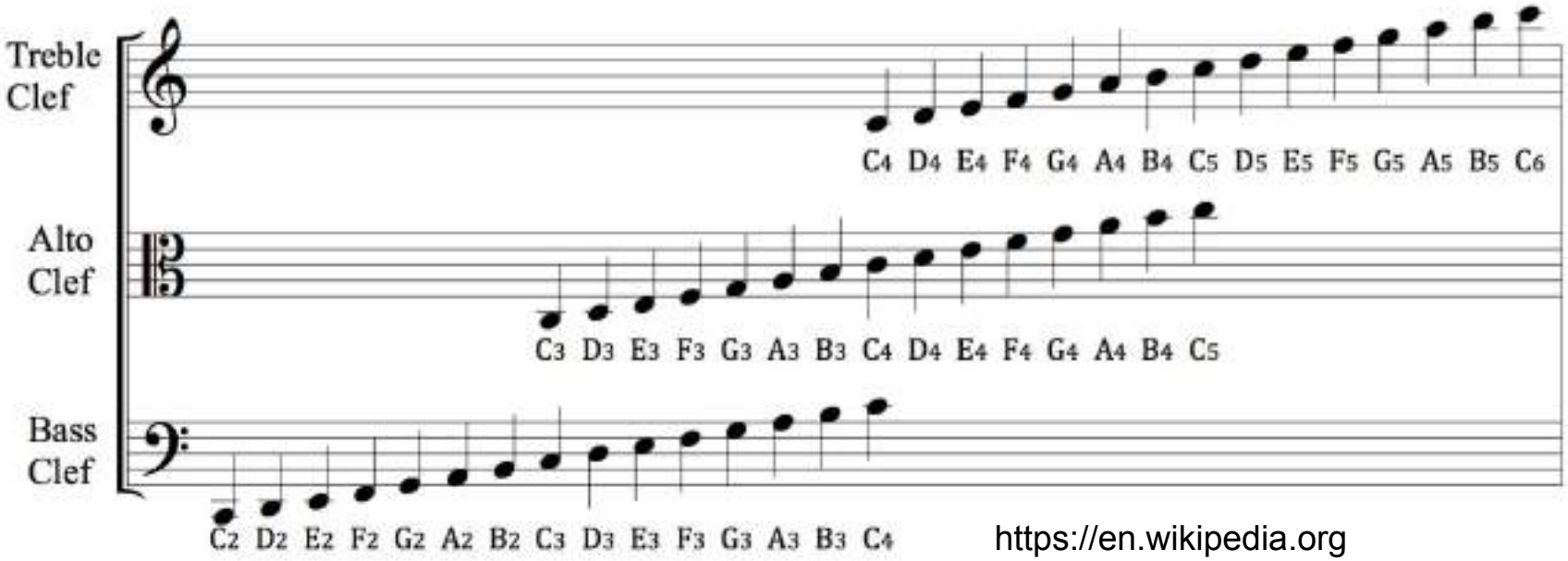
You will write and revise many times

Structure of Research Papers

Structure of a Research Paper

- Title, authors, affiliations
- Abstract (~150 words, ¼ page)
- Introduction (15% of paper)
- Background, Related Work (10% of paper)
- (optional) Description of the problem, examples (5-10%)
- (optional) Settings of the problems (5%)
- Your solution (30-40%, depending on areas)
- Evaluation and comparison (20-40%)
- (Optional) Discussion (5-10%)
- Conclusion (~150 words, ¼ page)
- References (10%, 2-3 x number of pages)

Creativity in content, not the formats.



https://en.wikipedia.org

Examples of Paper Titles

2019 CVPR

- Efficient Video Classification Using Fewer Frames
- Learning to Generate Synthetic Data via Compositing
- Weakly Supervised Person Re-Identification
- Guided Stereo Matching

2019 MobiSys

- Graphics-aware Power Governing for Mobile Devices
- Understanding and Detecting Overlay-based Android Malware at Market Scales
- Liquid Testing with Your Smartphone
- Are RFID Sensing Systems Ready for the Real World?

To Add or Not To Add? Project's Name

MURP: My Unknown Research Project

- Many people include the projects' names in the title.
- It is your decision, of course.
- I think that's bad because nobody knows what it means.
- If your project is famous, you don't need to publish.
- If your project is unknown, the name wastes space.

Authors

IEEE considers individuals who meet all of the following criteria to be authors:

1. Made a significant intellectual contribution to the theoretical development, system or experimental design, prototype development, and/or the analysis and interpretation of data associated with the work contained in the article.
2. Contributed to drafting the article or reviewing and/or revising it for intellectual content.
3. Approved the final version of the article as accepted for publication, including references.

Authors

Anyone listed as Author on an ACM manuscript submission must meet all the following criteria:

- they have made substantial intellectual contributions to some components of the original work described in the manuscript; and
- they have participated in drafting and/or revision of the manuscript and
- they are aware the manuscript has been submitted for publication; and
- they agree to be held accountable for any issues relating to correctness or integrity of the work.

Lightweight Multi-View 3D Pose Estimation through Camera-Disentangled Representation

Edoardo Remelli[1] Shangchen Han[2] Sina Honari[1] Pascal Fua[1] Robert Wang[2]
[1]CVLab, EPFL, Lausanne, Switzerland
[2]Facebook Reality Labs, Redmond, USA

Title

Authors

Affiliations

A Unified Optimization Framework for Low-Rank Inducing Penalties

Marcus Valtonen Örnhag[1] Carl Olsson[1,2]
[1]Centre for Mathematical Sciences
Lund University
[2]Department of Electrical Engineering
Chalmers University of Technology
{marcus.valtonen_ornhag, carl.olsson}@math.lth.se

Pixel Consensus Voting for Panoptic Segmentation

Haochen Wang
Carnegie Mellon University
whcw@cmu.edu

Ruotian Luo
TTI-Chicago
rluo@ttic.edu

Michael Maire
University of Chicago
mmaire@uchicago.edu

Greg Shakhnarovich
TTI-Chicago
greg@ttic.edu

Structure of a Research Paper

- Title, authors, affiliations
- **Abstract (~150 words, ¼ page)**
- Introduction (15% of paper)
- Background, Related Work (10% of paper)
- (optional) Description of the problem, examples (5-10%)
- (optional) Settings of the problems (5%)
- Your solution (30-40%, depending on areas)
- Evaluation and comparison (20-40%)
- (Optional) Discussion (5-10%)
- Conclusion (~150 words, ¼ page)
- References (10%, 2-3 x number of pages)

Abstract (12-15 sentences)

1. **Problem description (2 sentences)**
2. Existing work and deficiencies (2-3 sentences)
3. Your method and why it is better (3-5 sentences)
4. Evaluation methods (2-4 sentences)
5. Results and comparison (3 sentences)
6. (Optional) Implications and impacts

Motivation and Problem

- Motivation: Improve environmental sustainability
- Problem: Create degradable plastic

https://www.usatoday.com

- Motivation: Improve transportation safety
- Problem: Create vehicle communication

https://daseuropeanautohaus.com

Common Mistakes

- These are not problems: motivation, deficiencies in existing solutions, difficulty in existing solutions, history
- Eliminating a disease: too big for a paper
- Existing solution is too slow or expensive
- “I do not know how to solve it”
- “People have tried and failed”
- Some beginning researchers say, “There is no existing work. I am the first.”
- These people treat ignorance as innovation
- Before selecting a problem, read research papers

Abstract (12-15 sentences)

1. Problem description (2 sentences)
2. **Existing work and deficiencies (2-3 sentences)**
3. Your method and why it is better (3-5 sentences)
4. Evaluation methods (2-4 sentence)
5. Results and comparison (3 sentences)
6. (Optional) Implications and impacts

Four stages of doing research

1. Excitement due to ignorance
2. Frustration and disappointment with knowledge (after reading papers)
3. Experiments and failures
4. Improvement and innovation

Abstract (12-15 sentences)

1. Problem description (2 sentences)
2. Existing work and deficiencies (2-3 sentences)
3. **Your method and why it is better (3-5 sentences)**
4. Evaluation methods (2-4 sentence)
5. Results and comparison (3 sentences)
6. (Optional) Implications and impacts

Explain Your Method

1. What do you do?
2. How is it different?
3. Why is it better? (Conceptually)

If it is not better, do not waste time writing a paper

Some students say, “It is better by experiments.”

If you do not know why it is better, you have no paper.

Abstract (12-15 sentences)

1. Problem description (2 sentences)
2. Existing work and deficiencies (2-3 sentences)
3. Your method and why it is better (3-5 sentences)
4. **Evaluation methods (2-4 sentence)**
5. Results and comparison (3 sentences)
6. (Optional) Implications and impacts

Abstract (12-15 sentences)

1. Problem description (2 sentences)
2. Existing work and deficiencies (2-3 sentences)
3. Your method and why it is better (3-5 sentences)
4. Evaluation methods (2-4 sentence)
5. **Results and comparison (3 sentences)**
   Usually have numbers
6. (Optional) Implications and impacts

Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton, ImageNet Classification with Deep Convolutional Neural Networks Advances in Neural Information Processing Systems 25 (NIPS 2012)

we achieved top-1 and top-5 error rates of 37.5% and 17.0% … ILSVRC-2012 competition and achieved a winning top-5 test error rate of 15.3%, compared to 26.2%

Structure of (Computer System) Papers

- ***Title, authors, affiliations***
- ***Abstract (~150 words, ¼ page)***
- Introduction (15% of paper)
- Background, Related Work (10% of paper)
- (optional) Description of the paper, examples (5-10%)
- (optional) Settings of the problems (5%)
- Your solution (30-40%, depending on areas)
- Evaluation and comparison (20-40%)
- (Optional) Discussion (5-10%)
- Conclusion (~150 words, ¼ page)
- References (10%)

Concise

- Use as few words as possible, as many as necessary
- Shorter sentences are stronger
  - I have a dream
  - We choose to go to the Moon
  - Don’t follow the crowd, let the crowd follow you.



Use as few words as possible

For the purpose of creating a plan that aims to respond to events which may happen in the future, our team in this organization will have a meeting in which the team will talk, discuss, and write a plan on paper.
⇒ We will meet and write a plan.

Our team in this university has discovered a mathematical procedure which shows this description of the properties is correct in every possible scenario that can be constructed.

⇒ We prove this is true.

Structure of a Research Paper

- Title, authors, affiliations
- Abstract (~150 words, ¼ page)
- ***Introduction (15% of paper)***
- Background, Related Work (10% of paper)
- (optional) Description of the problem, examples (5-10%)
- (optional) Settings of the problems (5%)
- Your solution (30-40%, depending on areas)
- Evaluation and comparison (20-40%)
- (Optional) Discussion (5-10%)
- Conclusion (~150 words, ¼ page)
- References (10%, 2-3 x number of pages)

Abstract ⇒ Introduction

1. Problem description ⇒ First Paragraph
2. Existing work and deficiencies ⇒ Second Paragraph
3. Your method and why it is better ⇒ Third Paragraph
4. Evaluation methods + Comparison ⇒ Fourth Paragraph

One (***good***) figure is worth 1,000 words

Si Liu, Jiashi Feng, Csaba Domokos, Hui Xu, Junshi Huang, Zhenzhen Hu, and Shuicheng Yan, Fashion Parsing With Weak Color-Category Labels, IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 16, NO. 1, JANUARY 2014

Voice and Tense

- Active voice (we) is common in some areas but not others
- e.g. “We propose a solution”, “We present a method”, “We design a system”, “We prove”, “We demonstrate” …
- Present tense is common, even though the research has already been done.
- Avoid mixing present and past tenses in the same sentences, e.g., "We invented a solution and the solution uses machine learning."
- Mix “We propose a method”, “The proposed method”, “Our method”, “The new method”, “The improved method”

On Spectral Clustering: Analysis and an algorithm

Andrew Y. Ng
CS Division
U.C. Berkeley
ang@cs.berkeley.edu

Michael I. Jordan
CS Div. & Dept. of Stat.
U.C. Berkeley
jordan@cs.berkeley.edu

Yair Weiss
School of CS & Engi
The Hebrew Univ.
yweiss@cs.huji.ac.il

Abstract

Despite many empirical successes of *spectral clustering* methods—algorithms that cluster points using eigenvectors of matrices derived from the data—there are several unresolved issues. First, there are a wide variety of algorithms that use the eigenvectors in slightly different ways. Second, many of these algorithms have no proof that they will actually compute a reasonable clustering. In this paper, we present a simple spectral clustering algorithm that can be implemented using a few lines of Matlab. Using tools from matrix perturbation theory, we analyze the algorithm, and

DEEP COMPRESSION: COMPRESSING DEEP NEURAL NETWORKS WITH PRUNING, TRAINED QUANTIZATION AND HUFFMAN CODING

Song Han
Stanford University, Stanford, CA 94305, USA
songhan@stanford.edu

Huizi Mao
Tsinghua University, Beijing, 100084, China
mhz12@mails.tsinghua.edu.cn

William J. Dally
Stanford University, Stanford, CA 94305, USA
NVIDIA, Santa Clara, CA 95050, USA
dally@stanford.edu

ABSTRACT

Neural networks are both computationally intensive and memory intensive, making them difficult to deploy on embedded systems with limited hardware resources. To address this limitation, we introduce "deep compression", a three stage pipeline: pruning, trained quantization and Huffman coding, that work together to reduce the storage requirement of neural networks by 35× to 49× without affecting their accuracy. Our method first prunes the network by learning only the important connections. Next, we quantize the weights to enforce weight sharing, finally, we apply Huffman coding. After the first two steps we retrain the network to fine tune the remaining connections and the quantized centroids. Pruning, reduces the number of connections by 9× to 13×; Quantization then reduces the number of bits that represent each connection from 32 to 5. On the ImageNet dataset, our method reduced the storage required by AlexNet by 35×, from 240MB to 6.9MB, without loss of accuracy. Our method reduced the size of VGG-16 by 49× from 552MB to 11.3MB, again with no loss of accuracy. This allows fitting the model

introduce

apply

Why Do Internet Services Fail, and What Can Be Done About It?

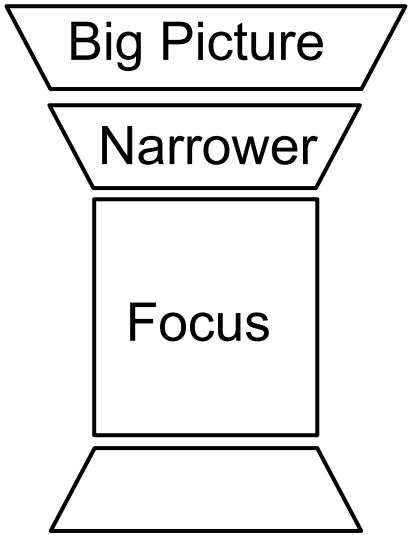
David Oppenheimer, Archana Ganapathi, and David A. Patterson, *University of California, Berkeley*

Abstract

In 1986 Jim Gray published his landmark study of the causes of failures of Tandem systems and the techniques Tandem used to prevent such failures See J. Gray. Why do computers stop and what can be done about it? Symposium on Reliability in Distributed Software and Database Systems, 1986.. Seventeen years later, Internet services have replaced fault-tolerant servers as the new kid on the 24x7-availability block. Using data from three large-scale Internet services, we analyzed the causes of their failures and the (potential) effectiveness of various techniques for preventing and mitigating service failure. We find that (1) operator error is the largest single cause of failures in two of the three services, (2) operator errors often take a long time to repair, (3) configuration errors are the largest category of operator errors, (4) failures in custom-written front-end software are significant, and (5) more extensive online testing and more thoroughly exposing and detecting component failures would reduce failure rates in at least one service. Qualitatively we find that improvement in the maintenance tools and systems used by service operations staff would decrease time to diagnose and repair problems.

Structure of Introduction

1. "Big picture", description of the problem
2. Related work and needed improvements
3. Your solution and why it is better
4. [More details about your solution]
5. Evaluation and comparison
6. Summary of contributions
7. [Implications]
8. [Structure of the paper]



Write Good Introduction

- Determine the audience's expected levels of expertise
- Use good figures to illustrate whenever possible
- Follow the "rules of 3": Each sentence has at most 30 words and 3 verbs. Each paragraph no more than ⅓ page
- Use simple English, even though content may be complex
- Do not require anyone to study grammar again
- Follow cohesion and coherence rules

Cohesion (connections of concepts)

Cohesion: Each sentence should be connected to the main concept(s) introduced in the previous sentence(s).

- Good: A→B, B→C, C→D ...
- Bad: A→B, C→D, E →F ...
- Avoid A→B, C→B, D→C ...
- Avoid A→B1, B2→C1, C2→D ...

Examples of cohesion

As computers are widely used in business transactions, security becomes increasingly important. Computer security usually relies on users providing passwords for authentication. Passwords need to meet certain rules to be considered secure. However, different organizations have different rules and create confusion among users.

Examples of cohesion

As computers are widely used in business transactions, security becomes increasingly important. Computer security usually relies on users providing passwords for authentication. Passwords need to meet certain rules to be considered secure. However, different organizations have different rules and create confusion among users.

Examples of cohesion

As computers are widely used in business transactions, security becomes increasingly important. Computer security usually relies on users providing passwords for authentication. Passwords need to meet certain rules to be considered secure. However, different organizations have different rules and create confusion among users.

Avoid similar but different expressions

As computers are widely used in business transactions, ***security*** becomes increasingly important. To use computers ***safely***, users need to enter ***passwords*** for authentication. Users typically answer two ***questions*** before entering a system.

⇒ Is security the same as safety? why two different words?

Autonomous robots can improve our everyday life.
Intelligent machines can work in dangerous areas.

⇒ Is autonomous robot the same as intelligent machine? why two different words?

Check Cohesion

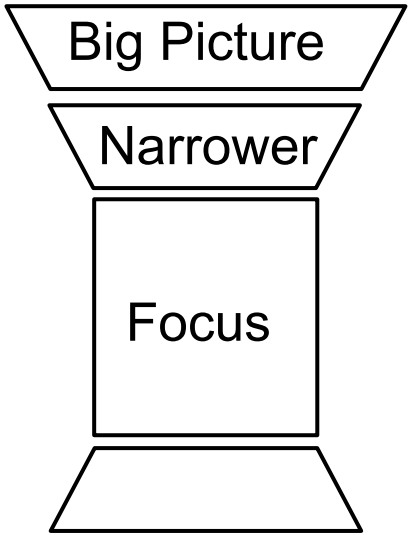
- Use color pens to mark the important concepts
- Check whether the concepts use the same words
- Examine the connections of concepts

Coherence (paragraph needs a focus)

- Coherence: the entire paragraph should have a focus. Otherwise A →B, B →C, C →D, ..., Y →Z
- Write down “one sentence summary” before writing each paragraph

Structure of Introduction

1. "Big picture", description of the problem
2. Related work and needed improvements
3. ***Your solution and why it is better***
4. [More details about your solution]
5. Evaluation and comparison
6. Summary of contributions
7. [Implications]
8. [Structure of the paper]



Third (+ Fourth) Paragraph in Introduction

- Start with “This paper” or “We” or “In this paper, we”
- Readers look for expressions like these
- Choose the right verbs: present, propose, investigate, demonstrate, develop, design, build, construct, prove, suggest, hypothesize
- Clearly explain: What you have done, How you do it, How you evaluate it, What results you get, Is it better than other methods? By how much?

When to write introduction?

- Some people suggest writing introduction last.
- You will need to revise the introduction many times.
- Need to think about your contributions early.
- Need to think about evaluation early.
- Need to think about the differences between your method and existing methods.
- You are already writing the introduction.

Structure of a Research Paper

- Title, authors, affiliations
- Abstract (~150 words, ¼ page)
- Introduction (15% of paper)
- ***Background, Related Work (10% of paper)***
- (optional) Description of the problem, examples (5-10%)
- (optional) Settings of the problems (5%)
- Your solution (30-40%, depending on areas)
- Evaluation and comparison (20-40%)
- (Optional) Discussion (5-10%)
- Conclusion (~150 words, ¼ page)
- References (10%, 2-3 x number of pages)

Related Work (not Works)

- Provide the context of your work
- Inform readers of the state of the art
- Explain the deficiencies of the existing work
- Describe the metrics for comparison
- Establish your credibility
- Position your own paper
- Cite the papers in the same venue you will submit to
- The number of references is approximately 2 x # pages

2. Related Work

Events-to-video reconstruction is a popular topic in the event camera literature. Early approaches did not reconstruct videos, but focused on the reconstruction of a single image from a large set of events collected by an event camera moving through a static scene. These works exploit the fact that every event provides one equation relating the intensity gradient and optic flow through brightness constancy [15]. Cook *et al.* [10] used bio-inspired, interconnected networks to simultaneously recover intensity images, optic flow, and angular velocity from an event camera performing small rotations. Kim *et al.* [17] developed an Extended Kalman Filter to reconstruct a 2D panoramic gradient image (later upgraded to a full intensity frame by 2D Poisson integration) from a rotating event camera, and later extended it to a 3D scene and 6 degrees-of-freedom (6DOF) camera motion [18] (albeit in a static scene only). Bardow *et al.* [3] proposed to estimate optic flow and inten-

Events-To-Video: Bringing Modern Computer Vision to Event Cameras

Henri Rebecq, Rene Ranftl, Vladlen Koltun, Davide Scaramuzza; Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 3857-3866

II. Related Work

In this work, we provide relevant background on previous work on video captioning, temporal attention mechanism and exploitation of local features.

Video Captioning: Up to now, some methods have been proposed for addressing the problems of video captioning, and these approaches have proved to be significant progress. These methods could be roughly classified into three types, depending on the manner in which the sentences are generated. (1) template-based method; (2) the method based on neural network.

Template-based method [13]–[17] firstly predicts semantic concepts or words (e.g., subjects, objects and verbs) by different classification methods, then employs a pre-defined sentence template to form them into a description. This method is intuitive, but need to deal with the complex data. Meanwhile, the limitation of sentence template cannot flexibly generate meaningful sentences [5].

With the development of deep learning, recent work [5]–[8], [18], [19] turn to employ neural network-based method to

C. Yan et al., "STAT: Spatial-Temporal Attention Mechanism for Video Captioning," in IEEE Transactions on Multimedia, vol. 22, no. 1, pp. 229-241, Jan. 2020, doi: 10.1109/TMM.2019.2924576.

Table I. Related Work on Mobile CBIR

Paper	Purpose	Feature Extracted	Search
[Sonobe et al. 2004]	recognize fish images	server	server
[Noda and Sonobe 2002]	identify flowers	server	server
[Ahmad and Gabbouj 2005]	show feasibility of architecture	server	server
[Yeh et al. 2004]	identify monuments	server	server
[He et al. 2008]	identify garments	server	server
[Chen et al. 2009]	compress features	client	server
[Rohs and Gfeller 2004]	identify color codes	client	client
[Yang et al. 2008a]	implement mobile CBIR	client	client
[Zhu et al. 2009]	retrieve images	client	clients
This article	save energy	client/server	client/server

Karthik Kumar, Yamini Nimmagadda, and Yung-Hsiang Lu. 2012. Energy Conservation for Image Retrieval on Mobile Systems. *ACM Trans. Embed. Comput. Syst.* 11, 3, Article 66 (September 2012), 22 pages. https://doi.org/10.1145/2345770.2345779

2 Background

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, Illia Polosukhin, Attention is All you Need, Advances in Neural Information Processing Systems 30 (NIPS 2017)

The goal of reducing sequential computation also forms the foundation of the Extended Neural GPU [20], ByteNet [15] and ConvS2S [8], all of which use convolutional neural networks as basic building block, computing hidden representations in parallel for all input and output positions. In these models, the number of operations required to relate signals from two arbitrary input or output positions grows in the distance between positions, linearly for ConvS2S and logarithmically for ByteNet. This makes it more difficult to learn dependencies between distant positions [11]. In the Transformer this is reduced to a constant number of operations, albeit at the cost of reduced effective resolution due to averaging attention-weighted positions, an effect we counteract with Multi-Head Attention as described in section 3.2.

Self-attention, sometimes called intra-attention is an attention mechanism relating different positions of a single sequence in order to compute a representation of the sequence. Self-attention has been used successfully in a variety of tasks including reading comprehension, abstractive summarization, textual entailment and learning task-independent sentence representations [4, 22, 23, 19].

End-to-end memory networks are based on a recurrent attention mechanism instead of sequence-aligned recurrence and have been shown to perform well on simple-language question answering and language modeling tasks [28].

Techniques	Description and Deficiencies
Distances Between Feature Vectors [41][42]	Use the distances between the centroids of feature vectors to find similarity. All groups of categories lying within a radius of each other are grouped. Difficult to determine the optimal radius.
Hierarchical Clustering [45]-[49]	Use the distances between the centroids of feature vectors to find similarity. The k closest categories are grouped at every stage. Fixing the value k produces poor accuracy and degenerated hierarchies.
HSV and Gabor Features [50]	Use texture and color information from categories to find similarities. All categories sharing the same color template or textures are grouped. Images in the same category may have different colors and textures.
Semantic Similarity[51]-[54]	Use semantic information from sources like WordNet to quantify the similarity. Categories that share more semantic details are grouped together. Semantic and visual similarities often do not correlate.
MNN-Tree (Proposed Method)	Use DNN's softmax output for a category, averaged over all input images belonging to other categories to find similarity. The softmax output is an indication of the confusion between categories. Categories that are frequently confused are grouped together.

Table 2. Different techniques to build hierarchies for image classification.

Description of Related Work

- State facts
- Acknowledge their contributions
- Take a positive tone (this method can be improved)
- Citing websites not always accepted

References

- Readers can study more details
- “References”, not “Prerequisites”
- Readers may check references first
- Must give strong positive impression

Authors

[27] Z. Zivkovic, "Improved adaptive gaussian mixture model for background subtraction," in *Proceedings of the International Conference on Pattern Recognition*, vol. 2, 2004, pp. 28–31.

[28] J. Shi and C. Tomasi, "Good features to track," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1994, pp. 593–600.

[29] J.-Y. Bouguet, "Pyramidal implementation of the affine lucas kanade feature tracker description of the algorithm," *Intel Corporation*, vol. 5, no. 1-10, p. 4, 2001.

[30] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, 2005, pp. 886–893.

Title

[27] Z. Zivkovic, "Improved adaptive gaussian mixture model for background subtraction," in *Proceedings of the International Conference on Pattern Recognition*, vol. 2, 2004, pp. 28–31.

[28] J. Shi and C. Tomasi, "Good features to track," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1994, pp. 593–600.

[29] J.-Y. Bouguet, "Pyramidal implementation of the affine lucas kanade feature tracker description of the algorithm," *Intel Corporation*, vol. 5, no. 1-10, p. 4, 2001.

[30] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, 2005, pp. 886–893.

Conference or Journal

[27] Z. Zivkovic, "Improved adaptive gaussian mixture model for background subtraction," in *Proceedings of the International Conference on Pattern Recognition*, vol. 2, 2004, pp. 28–31.

[28] J. Shi and C. Tomasi, "Good features to track," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1994, pp. 593–600.

[29] J.-Y. Bouguet, "Pyramidal implementation of the affine lucas kanade feature tracker description of the algorithm," *Intel Corporation*, vol. 5, no. 1-10, p. 4, 2001.

[30] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, 2005, pp. 886–893.

Year

[27] Z. Zivkovic, "Improved adaptive gaussian mixture model for background subtraction," in *Proceedings of the International Conference on Pattern Recognition*, vol. 2, 2004, pp. 28–31.

[28] J. Shi and C. Tomasi, "Good features to track," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1994, pp. 593–600.

[29] J.-Y. Bouguet, "Pyramidal implementation of the affine lucas kanade feature tracker description of the algorithm," *Intel Corporation*, vol. 5, no. 1-10, p. 4, 2001.

[30] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, 2005, pp. 886–893.

Pages

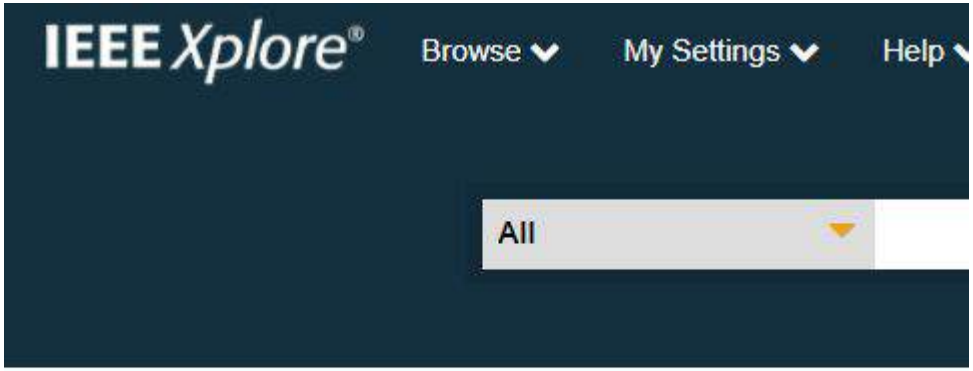
[27] Z. Zivkovic, "Improved adaptive gaussian mixture model for background subtraction," in *Proceedings of the International Conference on Pattern Recognition*, vol. 2, 2004, pp. 28–31.

[28] J. Shi and C. Tomasi, "Good features to track," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1994, pp. 593–600.

[29] J.-Y. Bouguet, "Pyramidal implementation of the affine lucas kanade feature tracker description of the algorithm," *Intel Corporation*, vol. 5, no. 1-10, p. 4, 2001.

[30] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, 2005, pp. 886–893.

Download Citations



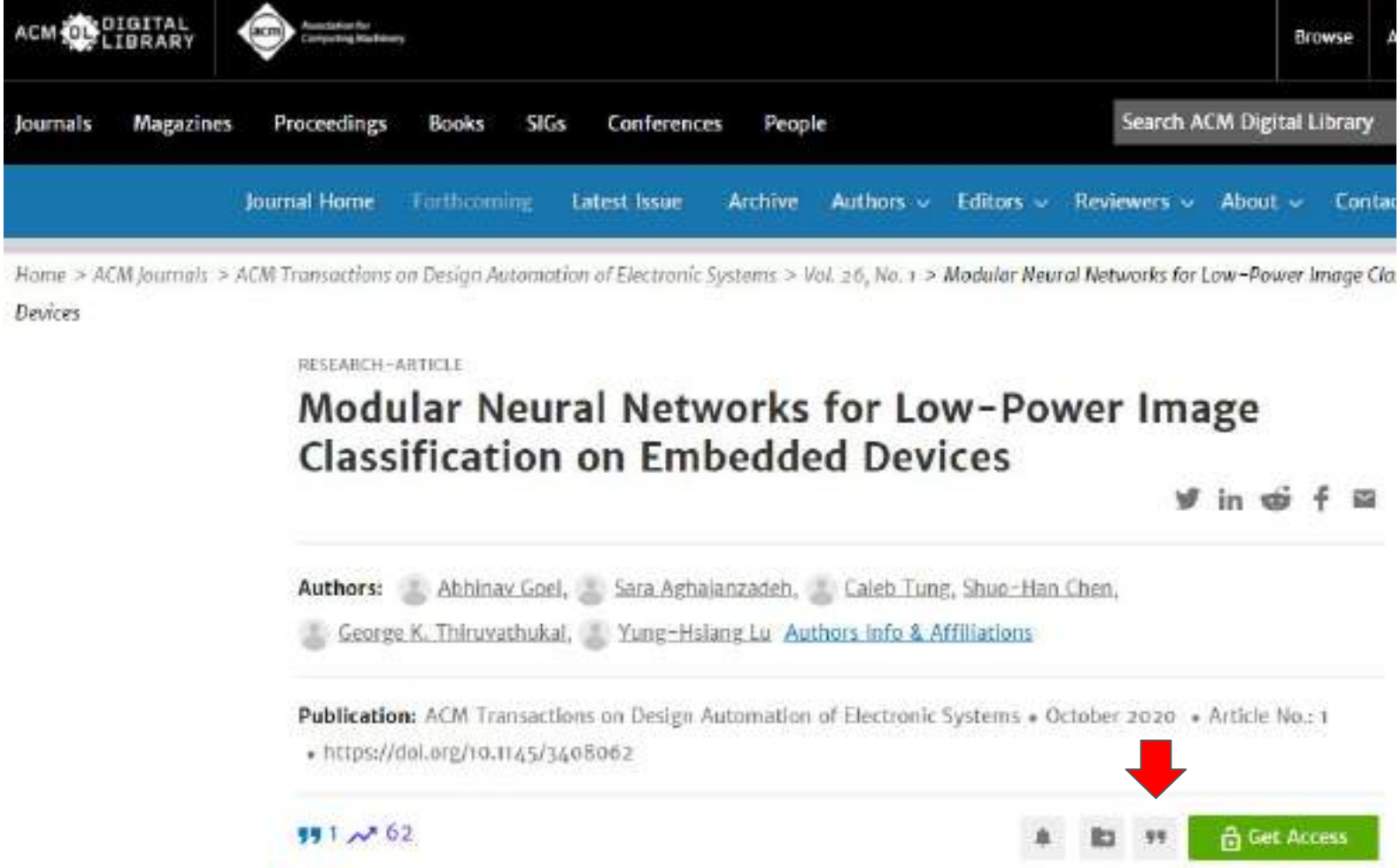
Journals & Magazines > Computer > Volume: 52 Issue: 10

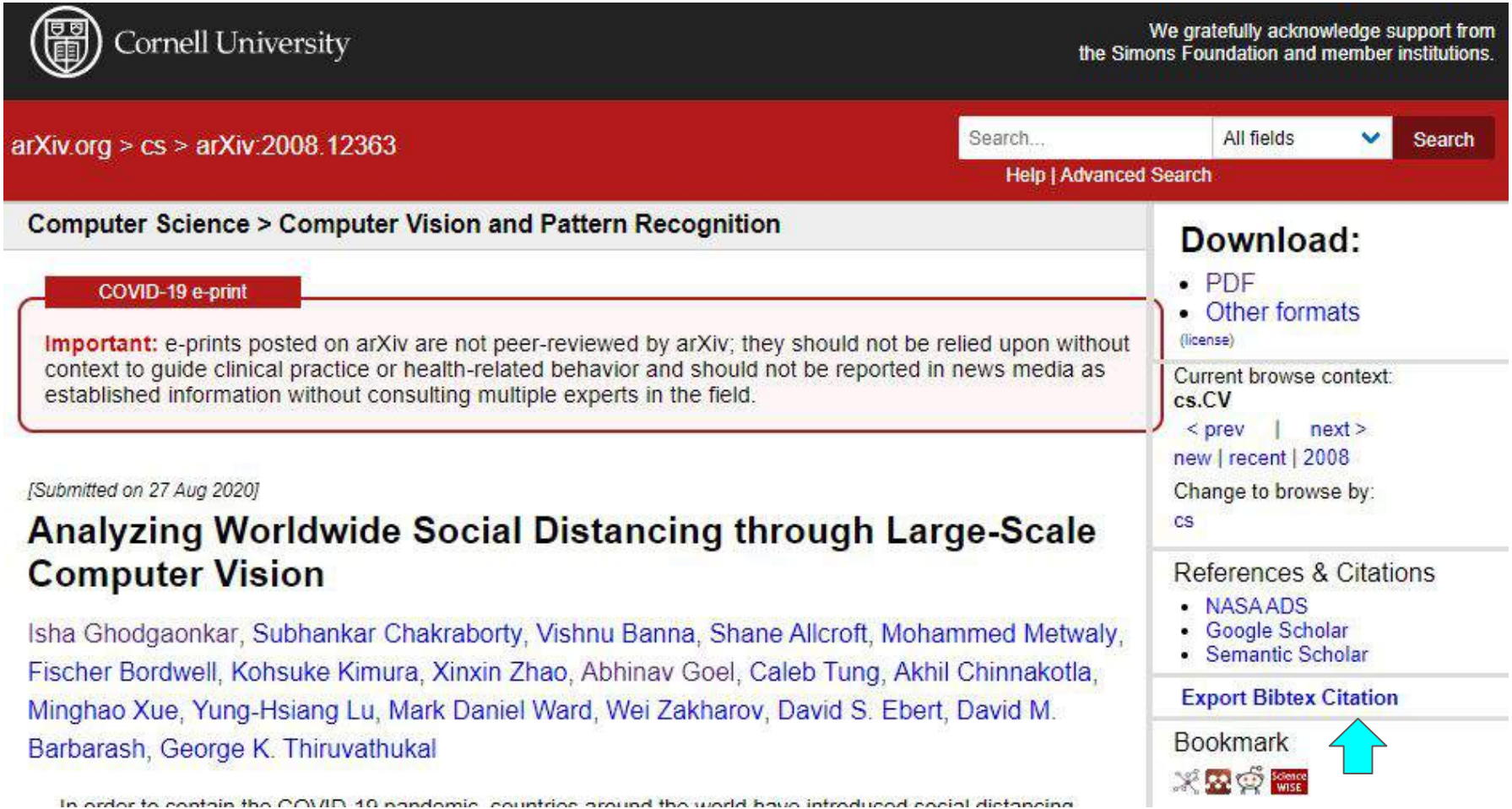
See the World Through Network Cameras

Publisher: IEEE

Cite This

PDF


ACM DIGITAL LIBRARY
Browse
Journals
Magazines
Proceedings
Books
SIGs
Conferences
People
Search ACM Digital Library
Journal Home
Forthcoming
Latest Issue
Archive
Authors
Editors
Reviewers
About
Home > ACM Journals > ACM Transactions on Design Automation of Electronic Systems > Vol. 26, No. 1 > Modular Neural Networks for Low-Power Image Cla
Devices
RESEARCH-ARTICLE
Modular Neural Networks for Low-Power Image Classification on Embedded Devices
Authors: Abhinav Goel, Sara Aghajanzadeh, Caleb Tung, Shuo-Han Chen, George K. Thiruvathukal, Yung-Hsiang Lu Authors Info & Affiliations
Publication: ACM Transactions on Design Automation of Electronic Systems • October 2020 • Article No.: 1 • https://doi.org/10.1145/3408062
1
62
Get Access


Cornell University
We gratefully acknowledge support from the Simons Foundation and member institutions.
arXiv.org > cs > arXiv:2008.12363
Search...
All fields
Search
Help | Advanced Search
Computer Science > Computer Vision and Pattern Recognition
COVID-19 e-print
Important: e-prints posted on arXiv are not peer-reviewed by arXiv; they should not be relied upon without context to guide clinical practice or health-related behavior and should not be reported in news media as established information without consulting multiple experts in the field.
[Submitted on 27 Aug 2020]
Analyzing Worldwide Social Distancing through Large-Scale Computer Vision
Isha Ghodgaonkar, Subhankar Chakraborty, Vishnu Banna, Shane Allcroft, Mohammed Metwaly, Fischer Bordwell, Kohsuke Kimura, Xinxin Zhao, Abhinav Goel, Caleb Tung, Akhil Chinnakotla, Minghao Xue, Yung-Hsiang Lu, Mark Daniel Ward, Wei Zakharov, David S. Ebert, David M. Barbarash, George K. Thiruvathukal
Download:
PDF
Other formats
(license)
Current browse context:
cs.CV
< prev | next >
new | recent | 2008
Change to browse by:
cs
References & Citations
NASA ADS
Google Scholar
Semantic Scholar
Export Bibtex Citation
Bookmark

Design, Automation, and Test in Europe pp 207-219 | Cite as

Quantitative Comparison of Power Management Algorithms

Authors Authors and affiliations

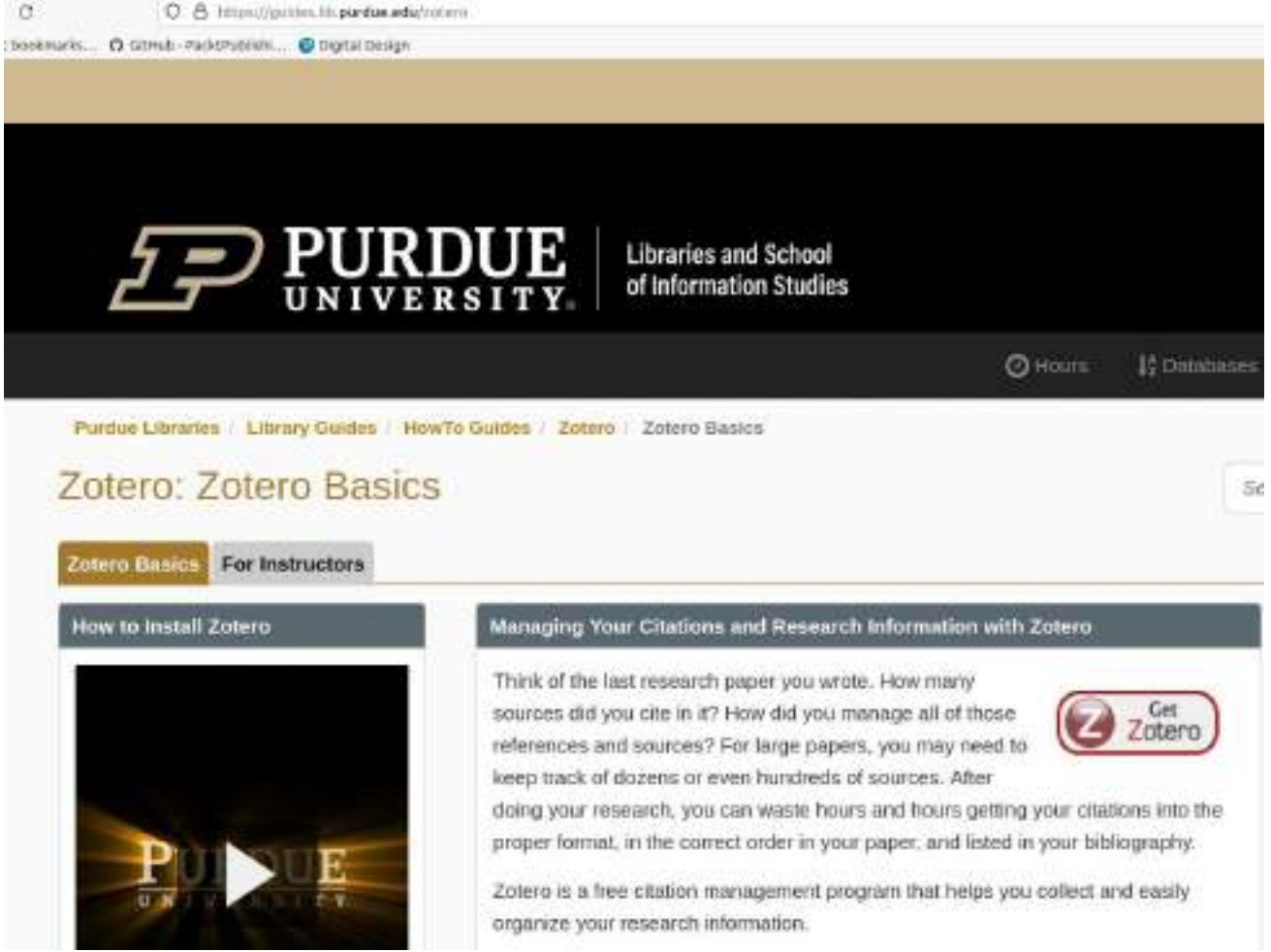
Yung-Hsiang Lu, Eui-Young Chung, Tajana Šimunić, Luca Benini, Giovanni De Micheli

Chapter

852 Downloads

Abstract

Dynamic power management saves power by shutting down idle devices. Several management algorithms have been proposed and demonstrated effective in certain applications. We quantitatively compare the power saving and performance impact of these algorithms on hard disks of a desktop and a notebook computers. This paper has three contributions. First, we build a framework in Windows NT to implement power managers running realistic workloads


PURDUE UNIVERSITY
Libraries and School of Information Studies
Hours
Databases
Purdue Libraries / Library Guides / HowTo Guides / Zotero / Zotero Basics
Zotero: Zotero Basics
Zotero Basics
For Instructors
How to Install Zotero
Managing Your Citations and Research Information with Zotero
Think of the last research paper you wrote. How many sources did you cite in it? How did you manage all of those references and sources? For large papers, you may need to keep track of dozens or even hundreds of sources. After doing your research, you can waste hours and hours getting your citations into the proper format, in the correct order in your paper, and listed in your bibliography.
Get Zotero
Zotero is a free citation management program that helps you collect and easily organize your research information.

Structure of a Research Paper

- Title, authors, affiliations
- Abstract (~150 words, ¼ page)
- Introduction (15% of paper)
- (optional) Description of the problem, examples (5-10%)
- (optional) Settings of the problems (5%)
- ***Your solution (30-40%, depending on areas)***
- Evaluation and comparison (20-40%)
- (Optional) Discussion (5-10%)
- Conclusion (~150 words, ¼ page)
- References (10%, 2-3 x number of pages)

Solution (or Method)

- Give a meaningful title (e.g., “Robots for Recycling” or “Incentive for Recycling”, not “The Idea” or “The Method”)
- Need to provide an overview of the section
- Divide the section into subsections
- Use figures to help readers understand

Advisable Learning for Self-Driving Vehicles by Internalizing Observation-to-Action Rules

Jinkyu Kim, Suhong Moon, Anna Rohrbach, Trevor Darrell, John Canny; Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 9661-9670

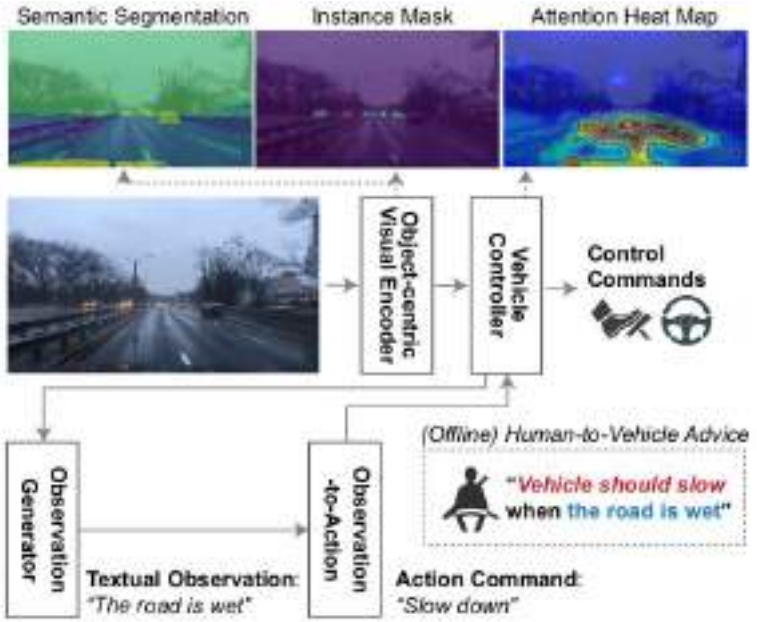


Figure 1: Our model consists of four main parts: (1) an object-centric visual encoder built upon a semantic segmentation model, (2) an observation generator, which generates textual observation about the scenes ("The road is wet"), (3) an observation-to-action module, which maps a visual scene description to a (high-level) action command ("Slow down"), and (4) a vehicle controller conditioned on the generated action command.

How to find better solutions?

- Reproduce results in papers
- List assumptions
- Test new data
- Compare and mix methods
- Generalize or specialize
- Use new hardware
- Read more papers
- Solve “real world” problems
- Research means re-search: If it is easy, it is not worth doing
- Attend many seminars, talk to many people

	Feature			
Method	1	2	3	4
1	✔	✔		
2	✔		✔	
3	✔	✔		✔
4			✔	✔
Yours	✔	✔	✔	✔