

IntechOpen

IntechOpen Series
Artificial Intelligence, Volume 36

Emerging Technologies in Computational Cognitive Neuroscience

Edited by Vijayalakshmi Kakulapati



Emerging Technologies in Computational Cognitive Neuroscience

Edited by Vijayalakshmi Kakulapati

Published in London, United Kingdom

Emerging Technologies in Computational Cognitive Neuroscience

<http://dx.doi.org/10.5772/intechopen.1006448>

Edited by Vijayalakshmi Kakulapati

Contributors

Andrey Vinogradov, Aruna Varanasi, Emre Kapucu, Hiroshi Omori, Ismail Noori Mseer, Laura Ylä-Outinen, Md Jaffar Sadiq, Susanna Narkilanti, Vijayalakshmi Kakulapati

© The Editor(s) and the Author(s) 2025

The rights of the editor(s) and the author(s) have been asserted in accordance with the Copyright, Designs and Patents Act 1988. All rights to the book as a whole are reserved by INTECHOPEN LIMITED. The book as a whole (compilation) cannot be reproduced, distributed or used for commercial or non-commercial purposes without INTECHOPEN LIMITED's written permission. Enquiries concerning the use of the book should be directed to INTECHOPEN LIMITED rights and permissions department (permissions@intechopen.com)

Violations are liable to prosecution under the governing Copyright Law.



Individual chapters of this publication are distributed under the terms of the Creative Commons Attribution 4.0 License which permits commercial use, distribution and reproduction of the individual chapters, provided the original author(s) and source publication are appropriately acknowledged. If so indicated, certain images may not be included under the Creative Commons license. In such cases users will need to obtain permission from the license holder to reproduce the material. More details and guidelines concerning content reuse and adaptation can be found at <http://www.intechopen.com/copyright-policy.html>.

Notice

Statements and opinions expressed in the chapters are those of the individual contributors and not necessarily those of the editors or publisher. No responsibility is accepted for the accuracy of information contained in the published chapters. The publisher assumes no responsibility for any damage or injury to persons or property arising out of the use of any materials, instructions, methods or ideas contained in the book.

First published in London, United Kingdom, 2025 by IntechOpen

IntechOpen is the global imprint of INTECHOPEN LIMITED, registered in England and Wales, registration number: 11086078, 167-169 Great Portland Street, London, W1W 5PF, United Kingdom

For EU product safety concerns: IN TECH d.o.o., Prolaz Marije Krucifikse Kozulić 3, 51000 Rijeka, Croatia, info@intechopen.com or visit our website at intechopen.com.

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

Emerging Technologies in Computational Cognitive Neuroscience

Edited by Vijayalakshmi Kakulapati

p. cm.

This title is part of the Artificial Intelligence Book Series, Volume 36

Topic: Computational Neuroscience

Series Editor: Andries Engelbrecht

Topic Editor: Magnus Johnsson

Print ISBN 978-1-83634-962-4

Online ISBN 978-1-83634-961-7

eBook (PDF) ISBN 978-1-83634-963-1

ISSN 2633-1403

If disposing of this product, please recycle the paper responsibly.

IntechOpen Book Series
Artificial Intelligence
Volume 36

Aims and Scope of the Series

Artificial Intelligence (AI) is a rapidly developing multidisciplinary research area that aims to solve increasingly complex problems. In today's highly integrated world, AI promises to become a robust and powerful means for obtaining solutions to previously unsolvable problems. This Series is intended for researchers and students alike interested in this fascinating field and its many applications.

Meet the Series Editor



Andries Engelbrecht received the Masters and Ph.D. degrees in Computer Science from the University of Stellenbosch, South Africa, in 1994 and 1999 respectively. He is currently appointed as the Voigt Chair in Data Science in the Department of Industrial Engineering, with a joint appointment as Professor in the Computer Science Division, Stellenbosch University. Prior to his appointment at Stellenbosch University, he has been at the University of Pretoria, Department of Computer Science (1998-2018), where he was appointed as South Africa Research Chair in Artificial Intelligence (2007-2018), the head of the Department of Computer Science (2008-2017), and Director of the Institute for Big Data and Data Science (2017-2018). In addition to a number of research articles, he has written two books, *Computational Intelligence: An Introduction and Fundamentals of Computational Swarm Intelligence*.

Meet the Volume Editor



Dr. Vijayalakshmi Kakulapati is Professor and Associate Dean (R&D) at Sreenidhi Institute of Science and Technology (Autonomous), Hyderabad, India. She holds a Ph.D. in Computer Science and Engineering from JNTUH and a Post-Doctoral Fellowship in Artificial Super Intelligence. With over 27 years of academic and research experience, she has authored more than 250 research publications, including over 120 in Scopus, Web of Science, and SCI-indexed journals. She has edited or authored 13 books and contributed to more than 55 book chapters. Dr. Kakulapati holds nine patents and has received several national and international awards recognizing her contributions to research and innovation. Her expertise spans artificial intelligence, cognitive systems, data science, and intelligent computing.

Contents

Preface	XIII
Section 1	
Fundamental Aspects of Cognitive Computational Models in Neuroscience	1
Chapter 1	3
Introductory Chapter: Introduction to Cognitive Computational Models in Neuroscience <i>by Vijayalakshmi Kakulapati</i>	
Chapter 2	11
Impact of Cognitive Computational Models in Neuroscience <i>by Vijayalakshmi Kakulapati, Md Jaffar Sadiq and Aruna Varanasi</i>	
Chapter 3	27
Neurons on Microelectrode Arrays and <i>In Vitro</i> Electrophysiological Data Analysis <i>by Andrey Vinogradov, Laura Ylä-Outinen, Susanna Narkilahti and Emre Kapucu</i>	
Section 2	
Case Studies of Cognitive Computational Models in Neuroscience	55
Chapter 4	57
Perspective Chapter: What Image AI is Close to Human Senses? <i>by Hiroshi Omori</i>	
Chapter 5	83
E-Learning in Light of the Concept of Power <i>by Ismail Noori Mseer</i>	

Preface

Computational cognitive neuroscience is a field that combines biology and psychology to utilize neural network models to simulate cognitive functions and link biological processes to behavior. Researchers may test theories of cognition in ways that are not always attainable with only experiments in this discipline. We can use computer modeling to examine how networks of neurons learn, change, and occasionally fail, which is similar to how the brain functions normally and how it does not operate properly in people with neurological disorders. Students and researchers can now access this topic more easily thanks to advanced technologies; however, there is still some doubt. To be meaningful, computational models must be based on real-world data and have strict limits at multiple levels of analysis. The goal of this study is to examine the advantages and disadvantages of using computational modeling in cognitive neuroscience. It will provide both a theoretical basis and practical knowledge of how computational tools can help us understand the complex relationship between the brain and behavior.

Understanding how the brain works is one of the greatest challenges in science, as it is a complex and fascinating mechanism. Computational cognitive neuroscience aims to bridge the gap between biology and behavior by employing computer models to simulate and comprehend the brain's functioning. This book discusses neural networks designed to mimic the structure and operation of the brain. It focuses on aspects such as vision, attention, learning, memory, language, and cognitive control. It also examines how issues with brain systems may lead to cognitive difficulties in individuals with neurological disorders. As computational modeling tools improve, they enable us to create more precise simulations and gain a deeper understanding of how the brain functions. But there is still doubt about how biologically possible these models are. The book is for students, researchers, and professionals in neuroscience, psychology, artificial intelligence, and cognitive science.

Chapter 1: "Introductory Chapter: Introduction to Cognitive Computational Models in Neuroscience"

In neuroscience, cognitive computational models are crucial because they provide researchers with a means to explain, simulate, and predict brain functions, including language, learning, memory, attention, perception, decision-making, and consciousness. These models are based on concepts from various fields, including cognitive science, computer science, psychology, artificial intelligence, and neurobiology. They are now a key part of current neuroscience research. This chapter examines the evolution of cognitive modelling over time, from symbolic models based on logic and rule-based systems to more recent advancements in machine learning, particularly deep learning and neuro-symbolic AI. The chapter describes in detail how computer models can help us generate testable ideas, explain how the brain functions, and predict human behavior.

Chapter 2: “Impact of Cognitive Computational Models in Neuroscience”

Cognitive computational models have significantly enhanced our understanding of how the brain functions and its underlying structure. They have helped us learn more about cognitive processes, human behavior, cognitive diseases, and brain function. They also help us understand how people change over time by finding hidden influences. Cognitive science and computational neuroscience collaborate to investigate the brain's functioning. They use deep neural network models, machine learning, neuromorphic computing, and high-performance computing. This chapter discusses various CCMs and their potential benefits to healthcare workers.

Chapter 3: “Neurons on Microelectrode Arrays and *In Vitro* Electrophysiological Data Analysis”

This chapter describes how to employ Microelectrode array (MEA) technology in investigations of functional neural networks, with an emphasis on detecting spikes and bursts. It also discusses multicompartment MEAs for assessing network activity and connectivity, which helps us gain a deeper understanding of how brain networks function together *in vitro*. The chapter combines several analytic methodologies to provide a framework for understanding the complex dynamics of neural networks and their functional connections. This makes it a valuable resource for researchers who utilize MEA technology to investigate neuronal electrophysiology.

Chapter 4: “Perspective Chapter: What Image AI is Close to Human Senses?”

Using supervised learning, self-supervised learning, and CLIP, researchers have developed Computer Vision Models (CVMs) to examine how people perceive their world through images. They utilized three sets of photos, employing SoftMax regression, to identify a fake CVM that was most similar to the photos. The best CVM combinations changed significantly across picture sets, suggesting that they are quite subjective and reflect how people perceive things. The best CVM similarity was linked to how people are represented.

Chapter 5: “E-Learning in Light of the Concept of Power”

Access to technology, curriculum setup, social presence, and feedback are all factors that impact e-learning. Access to technology may create inequalities among children, and the way the curriculum is structured can reveal how power operates. The Community of Inquiry approach emphasizes the importance of teacher presence in facilitating effective learning. Not allowing pupils to discuss topics might make them less mentally and socially engaged. Because of this, e-learning requires a more open and effective teaching approach.

Vijayalakshmi Kakulapati
Sreenidhi Institute of Science and Technology,
Department of Information Technology,
Yamnampet, Ghatkesar, Hyderabad, Telangana, India

Section 1

Fundamental Aspects of
Cognitive Computational
Models in Neuroscience

Introductory Chapter: Introduction to Cognitive Computational Models in Neuroscience

Vijayalakshmi Kakulapati

1. Introduction

CCNs are particularly useful in neuroscience because they help us figure out how the brain thinks. They provide you with an organized approach to think about how the brain works and how to test its functions [1]. Researchers may utilize CCMs to do more than simply explain how the brain works; they can also make predictions about how neural circuits impact how we think and act. They employ math and algorithms to make theories of cognition more formal by putting together information from neuroimaging, electrophysiological, and behavioral testing.

CM has had a hard time figuring out how neural networks operate and learning from raw sensory information. Symbolic models like ACT-R and SOAR were the most frequent cognitive models in the past. DL and artificial neural networks are becoming trendy again. This has allowed contemporary models to get hierarchical representations from sensory information, which helps them make complex choices and think in an ordered fashion [2].

Investigators are integrating probabilistic reasoning and biological plausibility into cognitive models, such as Bayesian cognitive models, to build learning and inference processes that are extremely close to how individuals think. Deep generative and probabilistic graphical models are two examples of frameworks that use math to explain how the brain deals with uncertainty [3].

Neuro-symbolic models are a big step forward in the area. They combine organized illustrations of symbolic systems with logical reasoning from neural networks [4]. These hybrid models encompass the complexity of the human mind, including acquired patterns and organized reasoning. They are becoming more common in higher-order cognition.

To study behavior and structure at different levels, CCNs are very important. The Human Brain Project and the Allen Brain Observatory are two examples of large-scale brain simulations that provide us with a lot of neural data to use to test and improve these models. This shows how important it is to have models that describe how the brain works and how biological processes work [5].

Emerging trends in the field include the use of reinforcement learning to model adaptive AI systems, which have come a long way in several areas, such as behavior modeling, predictive coding models, and physiologically realistic learning algorithms.

These new developments are also having an impact on computational neuroscience, giving us new ways to simulate sequential cognition and memory [6]. There are still problems to solve, however, such as biological plausibility and interpretability. AI systems do a great job of mimicking human behavior, but their structures and learning principles are frequently quite different from those found in the brain. Also, validating models against real-world data is necessary to show that they are useful for science across species and tasks [7].

This chapter discusses cognitive computational modeling in neuroscience, including symbolic, connectionist, Bayesian, and hybrid methods. It talks about their brain bases, how they are used, and how modern models are used to deal with cognitive complexity. The chapter also talks about new trends and problems that are coming up in cognitive computational neuroscience.

2. About cognitive computational neuroscience

Cognitive computational neuroscience (CCN) is a discipline that brings together cognitive science, neuroscience, and computer modeling to learn how the brain makes cognitive activities happen. It wants to connect these domains by developing and testing biologically accurate computer models that show how the brain works and behaves while doing cognitive activities. CCN studies how the brain processes information to help in perception, memory, language, decision-making, and other cognitive skills [8]. Researchers create computer models based on biological brain systems to represent certain cognitive capabilities. They then test these models against data from studies of people and animals in the actual world. CCN study also helps make AI and ML systems that work more like real life. A CCN researcher may, for instance, create a neural network model of how the brain perceives things by using what they know about the anatomy of the visual cortex and how neurons respond to visual inputs.

CCNs show how the brain processes information using arithmetic. These models may be compared to how people act to learn more about how the brain works. The goal of these models is to help us better understand how people think and use that knowledge in fields including AI, brain-computer interfaces, social policy, education, and the treatment of cognitive disorders [9].

Theories of functional modules in cognitive psychology led to the creation of functional imaging methods such as electroencephalography, positron emission tomography, and fMRI. Inspired by cognitive psychology, cognitive neuroscience (CNS) came up with its version of the game “20 Questions” with nature, linking cognitive processes to brain areas. A brain map does not show how the brain works, but it does set limits for theories and might help us understand how the brain processes information [10, 11].

Cognitive science is important for computational neuroscience to understand how higher-level thinking works and how it affects behavior and cognition. We need AI, especially ML, to provide us with a theoretical and technical framework for modeling cognitive processes with physiologically realistic dynamic parts.

Computational modeling is an important technique in neuroscience because it lets researchers look at how neurons work on a biological scale [12]. It gives us a better understanding of how neurobiological features affect the characteristics of neural networks. Models may also be helpful at higher levels of study, such as showing how

the brain and behavior are connected. Some models look at behavioral phenomena at a middle level, incorporating information about the brain and how it changes over time. Some others, at the “anatomical systems” level, talk about how brain parts work based on their anatomical boundaries [13]. These models explore things that are not real, such the many sorts of representations that exist in the brain and how it learns or processes them.

Cognitive Computational Neuroscience (CCN) is a new discipline that brings together researchers from many professions to attempt to figure out how cognitive processes and brain systems operate together. Cognitive science models demonstrate how thoughts and actions are organized, whereas neuroscience focuses on the biological mechanisms that make these things happen. Computational models provide us with the formal tools we need to bring together different areas of study. This helps us model, guess, and evaluate cognitive and neurological phenomena in the real world.

3. Emerging technologies in cognitive computational neuroscience

Advanced technologies like neuroimaging, computational modeling, and brain-computer interfaces are being used to learn more about the brain and how it operates. Some of the most prominent areas of research include functional magnetic resonance imaging (fMRI), electroencephalography (EEG), magnetoencephalography (MEG), near-infrared spectroscopy (NIRS), transcranial magnetic stimulation (TMS), and positron emission tomography (PET). Researchers are also focusing on CM and ML, which includes encoding models, neural networks, ML algorithms, Bayesian Brain Modeling, and brain-computer interfaces that enable the brain to communicate with other systems [14].

CNS looks at how brain circuits handle information. Bayesian inference models, recurrent neural networks (RNNs), and reinforcement learning algorithms are some examples of mathematical models that may be used for this [15]. Researchers used deep convolutional neural networks (CNNs) to predict how neurons in the monkey visual cortex would act. This shows that there is a link between artificial and biological neural networks. Researchers have used recurrent neural network models to find out how prefrontal cortical networks retain information when doing memory tests. The CCN research uses AI’s reinforcement learning methods to copy how the brain learns from rewards and punishments. This is like how dopamine neurons and basal ganglia circuits operate in animals and humans [16].

AI algorithms are outstanding at predicting risks like depression, suicide, future drug addiction, and functional consequences. But they need to be validated to make sure they can be used in other situations and in the clinic. AI technologies save time in the health business by swiftly analyzing large volumes of data and anticipating treatments. They may also help create new theoretical models of how diseases work, find gene expression patterns that are connected to mental illnesses, and assist in predicting which medications will work, which can lead to the identification of new compounds that might be used to treat disorders. Cognitive neuroscientists may learn more about the brain and behavior, which helps both neuroscience and psychology. AI approaches may open new ways to intervene early, showing how AI and technology may help us learn more about behavior in science [17].

Functional neuroimaging employs neuroimaging technologies to look at how the brain works and how it relates to certain activities, stimuli, cognition, behaviors, or

neural processes. EEG, MEG, PET, fMRI, fNIRS, fTCDS, and MRS are some of the most common technologies. These technologies provide us with many ways of looking at complicated neurological processes [18]. As a result, multimodal imaging is becoming an area of interest for academic, therapeutic, commercial, and military purposes. In the future, these technologies may be used to gather information, teach people, screen for terrorism, and connect soldiers with machines.

Explainable AI is an important part of CM research because it makes models easier to understand and see through. Cognitive architectures like ACT-R and SOAR combine several models. Future research will concentrate on more complex ML methods. To make whole cognitive models that are easier to understand and see, multimodal modeling that includes behavioral, neurological, and physiological data is necessary [19].

HPC approaches may help with problems with computational performance in huge data sets and complicated models by employing platforms like JSC and supercomputers like Sequoia. They need more storage space, but they can conduct simulations that cut down on runtimes. Using big data and DL models together can make Alzheimer's disease diagnoses more than 90% accurate. As big data techniques provide bigger data pools, brain research will need quicker and more powerful computer approaches to be more complete.

Combining cognitive neuroscience, cognitive psychology, robotics, and AI gives us a full picture of how the brain works, which helps us make progress in AI, robotics, and healthcare.

4. Conclusion

Cognitive Computational Neuroscience is a new way to learn about the brain and cognition. CCN's goal is to uncover the neural basis for cognition by developing computational models guided by both cognitive theory and neuroscience evidence. This method might help not just with fundamental research but also with AI, neuropsychiatry, education, and how people and computers interact with one another. The discipline will only be successful if biological plausibility, computational rigor, and cognitive relevance continue to be combined.

5. Future enhancement

In the future, CNS will focus on creating encoding models to look at how brain representations change as people learn and grow, using more realistic tasks in neuroimaging studies to learn more about how the brain works in everyday life, combining knowledge from cognitive science, neuroscience, and AI to get a better picture of how the brain works, and making new and better neurotechnology for diagnosing, treating, and improving brain function. These new developments not only change the way cognitive neuroscience research is done, but they might also influence many other areas, such as mental health, neurology, education, and AI.


Author details

Vijayalakshmi Kakulapati

Sreenidhi Institute of Science and Technology, Hyderabad, Telangana, India

*Address all correspondence to: vldms@yahoo.com

IntechOpen

© 2025 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Kriegeskorte N, Douglas PK. Cognitive computational neuroscience. *Nature Neuroscience*. 2018;**21**(9):1148-1160. DOI: 10.1038/s41593-018-0210-5
- [2] Richards BA, Lillicrap TP, Beaudoin P, et al. A deep learning framework for neuroscience. *Nature Neuroscience*. 2019;**22**(11):1761-1770
- [3] Ma WJ, Jazayeri M. Neural coding of uncertainty and probability. *Annual Review of Neuroscience*. 2019;**42**:241-263
- [4] Besold TR, d'Avila Garcez AS, Bader S, Bowman H, Domingos P, Hitzler P, et al. Neuro-symbolic artificial intelligence: The state of the art. *Artificial Intelligence*. 2021;**295**:103459
- [5] Siegle JH, Jia X, Durand S, et al. Survey of spiking in the mouse visual system reveals functional hierarchy. *Nature*. 2021;**592**(7852):86-92
- [6] Schrimpf M, Blank I, Tuckute G, et al. The neural architecture of language: Integrative modeling converges on predictive processing. *Proceedings of the National Academy of Sciences*. 2021;**118**(45):e2105646118
- [7] Hassabis D, Kumaran D, Summerfield C, Botvinick M. Neuroscience-inspired artificial intelligence. *Neuron*. 2017;**95**(2):245-258
- [8] Naselaris T et al. Cognitive computational neuroscience: A new conference for an emerging discipline. *Trends in Cognitive Sciences*. 2018;**22**(5):365-367. DOI: 10.1016/j.tics.2018.02.008, ISSN 1364-6613
- [9] Cognitive Computational Neuroscience. The Centre for Human Brain Health. Available from: <https://www.birmingham.ac.uk/research/centres-institutes/human-brain-health/cognitive-computational-neuroscience>
- [10] Fodor JA. *Precis of the modularity of mind*. The Behavioral and Brain Sciences. 1985;**8**:1
- [11] Chklovskii DB, Koulakov AA. Maps in the brain: What can we learn from them? *Annual Review of Neuroscience*. 2004;**27**:369-392
- [12] Hampshire A, Chaudhry AM, Owen AM, Roberts AC. Dissociable roles for lateral orbitofrontal cortex and lateral prefrontal cortex during preference driven reversal learning. *NeuroImage*. 2012;**59**(4):4102-4112. ISSN 1053-8119. DOI: 10.1016/j.neuroimage.2011.10.072
- [13] Braver TS, Barch DM, Gray JR, Molfese DL, Snyder A. Anterior cingulate cortex and response conflict: effects of frequency, inhibition and errors. *Cereb Cortex*. Sep 2001;**11**(9):825-836. DOI: 10.1093/cercor/11.9.825
- [14] Mishra C. Application of technology in cognitive neuroscience and human behaviour. *Journal of Emerging Technologies and Innovative Research (JETIR)*. 2023;**10**(4):336-353, ISSN-2349-5162
- [15] Sutton RS, Barto AG. *Reinforcement Learning: An Introduction*. Cambridge, Massachusetts, London, England: The MIT Press; 2018
- [16] Mante V, Sussillo D, Shenoy KV, Newsome WT. Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature*. 2013;**503**(7474):78-84

[17] Lee EE, Torous J, De Choudhury M, Depp CA, Graham SA, Kim H-C, et al. Artificial intelligence for mental health care: Clinical applications, barriers, facilitators, and artificial wisdom. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*. 2021;**6**(9):856-864. DOI: 10.1016/j.bpsc.2021.02.001, ISSN 2451-9022

[18] National Research Council (US) Committee on Military and Intelligence Methodology for Emergent Neurophysiological and Cognitive/Neural Research in the Next Two Decades. *Emerging Cognitive Neuroscience and Related Technologies, 2, Current Cognitive Neuroscience Research and Technology: Selected Areas of Interest*. Washington (DC): National Academies Press (US); 2008. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK207951>

[19] Chen S, He Z, Han X, He X, Li R, Zhu H, et al. How big data and high-performance computing drive brain science. *Genomics, Proteomics & Bioinformatics*. 2019;**17**(4):381-392. DOI: 10.1016/j.gpb.2019.09.003

Impact of Cognitive Computational Models in Neuroscience

Vijayalakshmi Kakulapati, Md Jaffar Sadiq and Aruna Varanasi

Abstract

Cognitive computational models have significantly enhanced our understanding of the brain's structure and function, providing insights into cognitive processes, human behavior, cognitive illnesses, and brain function. They also aid in understanding lifetime development by identifying latent factors. Cognitive science and computational neuroscience work together to understand brain processing, with cognitive science studying how neurons perform basic cognitive functions and cognitive neuroscience creating models that break down cognition into functional components. Advancements in deep neural network models, machine learning, neuromorphic computing, high-performance computing, and machine learning enable cognitive computational neuroscience to simulate brain data arrangement and encode neural signals for brain-machine interfaces. This chapter explores different cognitive computational models (CCM) and their benefits, providing recommendations for healthcare professionals.

Keywords: computing, deep learning, machine learning, neurons, cognitions, technologies, healthcare, physicians, brain, data, signals

1. Introduction

Thinking, learning, and remembering are examples of cognitive processes that fall into declarative and procedural categories and are essential to human activity. Intelligence, a comprehensive cognitive capacity, encompasses thinking, problem-solving, and learning. Divergent individual capabilities might result in differing coping mechanisms under tough circumstances. Mild cognitive impairment manifests as a deterioration in memory and cognitive function associated with aging and dementia, perhaps including issues with memory, language, or judgment. Cognitive diseases often start insidiously but advance, impairing quality of life. Comprehending these illnesses and their manifestations is crucial [1].

The word “model” in the behavioral and brain sciences refers to a variety of models that are used to investigate information processing and brain function. Data-analysis methods delineate correlations among variables, whereas effective

connectivity and causal-interaction models deduce causal impacts. Box-and-arrow representations illustrate cognitive component functions, whereas word models offer ambiguous explanations. Oracle models depend on inaccessible information, whereas brain-computational models replicate task execution. Image-computable models forecast cerebral activity and behavioral reactions, whereas deep neural networks offer image-computable models for visual processing. Reinforcement learning models utilize authentic environmental input for their learning processes [2].

Cognitive computational models have considerably advanced our knowledge of the brain's structure and function, revealing information about cognitive processes, human behavior, cognitive diseases, and brain function. They facilitate the comprehension of lifelong development by recognizing hidden elements [3].

Through the conversion of hypotheses into implemented models and stimuli into numerical descriptors, computational modeling may help close the gap between neurobiology and cognition. This pertains to sophisticated deep learning models, neural encoding analysis techniques, and conventional instruments like as parsers, language models, and vector-space models. These models convert verbal assumptions into numerical representations or metrics that may be correlated with brain or psychological information [4].

Cognitive computational models have profoundly impacted neuroscience by enhancing our comprehension of the brain and its information processing mechanisms. They assist in delineating, quantifying, and validating intricate cognitive processes, forecasting behavior, and facilitating therapeutic applications. They amalgamate various neuroscientific data, formulate ideas on lifetime development, and simulate relationships among neurons and glial cells.

Computational modeling can connect cognition and neurobiology by converting hypotheses into executable models and stimuli into numerical representations. This pertains to sophisticated deep learning models, neural encoding analysis techniques, and conventional instruments such as parsers, language models, and vector-space models [4].

The present study is organized as follows: Section 2 provides a background study of computational cognitive neuroscience. Section 3 presents a relevant study in the field of computational cognitive neuroscience, and Section 4 gives detailed information about emerging technologies in computational cognitive neuroscience. Section 5 provides the overall discussion of the present study. Finally, Section 6 delves into concluding remarks, followed by future research prospects.

2. Background of the study

Cognitive science and cognitive neuroscience have advanced considerably since the era of Allen Newell, marked by the emergence of computational models and the necessity of artificial intelligence in the comprehension of cognition. During the 1980s, cognitive science achieved notable progress through symbolic cognitive architectures and neural networks that utilized human behavioral data [5, 6]. Nevertheless, computer hardware and machine learning have not yet progressed sufficiently to replicate cognitive processes in their entirety. The emergence of human functional brain imaging enabled cognitive neuroscience to align the information-processing modules of cognitive psychology with the brain; however, it was deficient in computational rigor. Cognitive science and neuroscience diverged in the 1990s [7].

The work in machine learning, artificial intelligence, cognitive science, and neuroscience has impacted each other's research. The research of McCulloch and Pitts on the brain was influenced by Turing's contributions to computation, which subsequently inspired Rosenblatt's development of the Perceptron. Neuroscience and psychology established the foundation for reinforcement learning, deep learning, and innovative external memory systems, such as the Neural Turing Machine [8].

The field of computational neuroscience emphasizes creating mathematical representations of fundamental cognitive processes such as sensory coding, normalization, working memory, evidence accumulation, decision-making processes, and motor control. The discipline is also evaluating intricate models to elucidate advanced sensory and cognitive brain representations (**Figure 1**) [10, 11].

Computational models have gained significance in human neuroimaging research, offering a framework for elucidating intricate brain-behavior correlations and forecasting cognitive, behavioral, and clinical results. This has fostered excitement about their capacity to augment our comprehension of cerebral activity and refine the forecasting of clinical problems [12]. Nonetheless, the constraints of neuroimaging have become more evident, particularly in studies concerning higher-order cognition and mental disorders [13, 14]. Swift breakthroughs in neurotechnology, including invasive recordings of human brain activity and real-time recordings through wearables, are vying for prominence in neuroimaging. These improvements provide unparalleled access to high-temporal and spatial resolution brain data and present novel opportunities for both fundamental and therapeutic research [15].

Various neuroscientific data from various research studies may be integrated by computational models, which can concentrate on factors such as learning rates and PEs. This convergence can be attained by online methods, controlled laboratory experiments, neuroimaging techniques, and invasive procedures. Comprehending neurobiological pathways via systematic study is essential for individualized mental wellness utilizing neuroscience. A multimodal approach is crucial for effectively utilizing many data sources and incorporating novel neuroimaging and behavioral techniques [16].

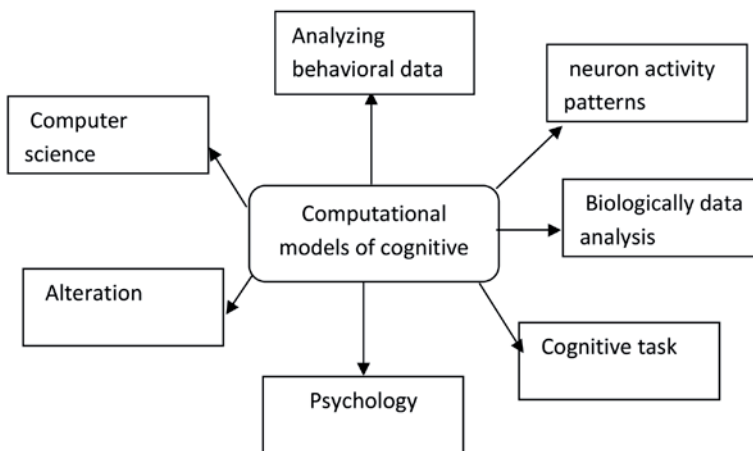


Figure 1.
Cognitive computational neuroscience [9].

3. Existing works

To produce human language and extract information from text input, Natural Language Processing (NLP) uses language models, which are computer systems. They forecast words or sequences depending on the prior or adjacent context. Examples include n-grams, probabilistic symbolic grammars, recurrent neural networks (RNNs), and transformer architectures. N-grams provide for a certain quantity of previous words but fail to represent hierarchical structures or long-range interdependence. Probabilistic context-free grammars openly represent hierarchical structures but do not use similarities across words and situations [17].

When analyzing Cognitive Networks (CCNs), encoding models are crucial because they use a weighted linear mixture of data taken from experimental inputs to predict brain activity. Models like the general linear model (GLM) utilize sensory properties as input and brain responses as output. Recent research has established models utilizing latent characteristics obtained from Artificial Neural Networks (ANN) [18]. Encoding models connect cerebral activity with artificial intelligence via a representational space, facilitating the evaluation of cognitive ideas [2].

The brain uses probability principles to evaluate data based on past knowledge, according to Bayesian cognitive theories [19]. In visual perception, retinal impulses represent external things, and to deduce these objects, we evaluate potential configurations and their explanatory adequacy for the image. Our preexisting ideas are encapsulated by a generative model that delineates the likelihood of each configuration and the probabilities with which a certain configuration would yield various retinal pictures. A Bayesian vision model may employ a generative model of the combined distribution of sensory input and environmental causes [20].

The China Brain Project, a component of the Innovation 2030 Major Science and Technology Projects, seeks to advance computational neuroscience to elucidate the neural foundations of cognition and neurological disorders. By mapping neural connections, the project will create effective algorithms and analytical tools to handle vast volumes of structural and functional data. The utilization of machine learning in analyzing brain abnormalities is crucial for the early detection of diseases and the assessment of therapy efficacy. The utilization of machine learning technologies for encoding and decoding neural signals will be essential in brain-machine interfaces [21].

Through the use of sophisticated models built using deep learning, neural encoding analysis techniques, and conventional tools, computational modeling may close the gap between neurobiology and cognition by converting inputs into numerical descriptors and hypotheses into implemented models [4].

In a study [22] on virtual or augmented reality digital communication frameworks underscores the difficulties in merging cognitive neuroscience with diverse technical fields. The efficacy of these systems is influenced by impaired perceptual and cognitive processing resulting from physiological and pathological aging causes, chiefly linked to neurotransmitter variations. The convergence of neuroscientific ideas with the Metaverse represents a potential multidisciplinary area.

Using mathematical frameworks such as partially-observable Markov decision processes (POMDP) and bilinear mappings, recent models have investigated the dynamically-gated working memory system, offering insights into its computational features [23, 24]. Dayan's research is on connecting fundamental gating processes for working memory updating with the system's capacity to execute tasks swiftly via vocal instructions. The research conducted by Ashby and colleagues demonstrates

that a PFC/BG system can elucidate significant discoveries in the categorization literature, hence providing more empirical opportunities for evaluating these theories [25]. Increasing empirical evidence from biological and cognitive neuroscience research substantiates the fundamental concepts behind these computer models [26].

Eye Movement analysis with Hidden Markov Models (EMHMM) has found prospective biomarkers for mental diseases such as sleeplessness, social anxiety, temperamental shyness, and stranger fear. Individuals with insomnia often misinterpret angry expressions as being afraid, owing to deficiencies in the eye area, indicating problems with attention management [27]. Eye movement flexibility predicts social anxiety symptoms in adults, but attention bias toward frightening faces correlates with temperamental shyness and stranger anxiety in infants [28]. Individuals experiencing their initial episode of schizophrenia have nose-centric eye movement patterns in gaze direction perception, signifying positive psychotic symptoms. Pain perception indicates that persons with unfavorable perceptions of illness-related situations have increased nose-focused eye movement patterns [29].

The viability of using online media mining to seed computer models of behavior modification for non-painful intentions (NPIs), such as mask wearing and social distancing, was shown by [30] using PVA simulations. They utilized two studies of COVID-19 Twitter datasets and predicted mask-wearing behavior in four US states. The first Twitter research concentrated on pro- and con-mask-wearing using hashtags, but the subsequent analysis offered a more nuanced examination of cognitive content utilizing Natural Language Processing (NLP) methodologies. The analyses were included in PVA models concerning attitudes and behaviors related to mask-wearing, and the predictions were juxtaposed with actual mask-wearing behaviors in those states throughout the 2020–2021 period. ACT-R theory, a cognitive architecture, serves as a computational realization of a cohesive theory of cognition, seeking to distill a consensus on the fundamental structures and processes of cognitive architectures [31].

PsyNeuLink is a platform created by Princeton University's Neuroscience of Cognitive Control Lab, providing a modeling environment for computational models utilized by neuroscientists, psychologists, and computer scientists to investigate brain functions, psychological behaviors, and their connection to artificial intelligence. It seeks to provide a centralized repository for model-sharing, enabling scientists to access optimal modeling tools for comprehending brain-based human intelligence [32].

4. About emerging technologies in neuroscience

Inspired by biological brain networks, the neural network model explains cognitive processes like visual object identification by combining information processing and various inputs. Sharp networks approximate any function, but deep networks address intricate tasks. Recurrent neural networks enhance engineering and capture repetitive signaling in the brain. Supervised learning techniques, including gradient descent, are employed for parameter estimation. Nonetheless, the use of this method by the brain remains contentious. Biologically realistic implementations of backpropagation and analogous supervised learning techniques may elucidate computational functions in the brain and propel developments in artificial intelligence.

Neuroscience includes diverse domains such as neural networks, brain-computer interfaces (BCIs), neuroinformatics, neuroimaging, cognitive neuroscience, neural

signal processing, neurodynamics, big data analytics, computational methods in neurobiology, neuroengineering, artificial intelligence in neuroscience, and mathematical simulation and evaluation of neurological diseases while tackling intricate issues in pattern recognition and signal processing.

The use of computational models to decipher intricate brain-behavior interactions and forecast cognitive, behavioral, and clinical consequences is growing in human neuroimaging research [12]. This has resulted in heightened interest in comprehending cerebral activity and forecasting clinical situations. Nonetheless, neuroimaging possesses limitations, particularly with higher-order cognition and mental illnesses [13]. Accelerated neuro-technological progress, including invasive brain activity recordings and wearable devices [33, 34], is vying for attention, offering access to high-temporospatial resolution neural data and ecologically relevant metrics. The issue persists regarding the continued significance of neuroimaging for behavioral studies of neuroscience.

Computational neuroscience is a swiftly advancing discipline that employs many models in neuroimaging and neuroscience research. These models may be either theory-driven or data-driven, with theory-driven models concentrating on elucidating the brain computations that underpin certain actions and cognitive processes [35]. Data-driven models, meanwhile, maintain a theory-agnostic approach, emphasizing data mining for practical objectives, such as discovering new patterns or forecasting clinical outcomes. Machine learning (ML) methodologies are pivotal to the AI revolution, employing both supervised and unsupervised learning approaches to attain practical results. These models emphasize empirical correlations within extensive datasets and present considerable therapeutic promise via biomarker discovery or the tailoring of treatment protocols [36, 37]. Hybrid models, which integrate diverse theoretical assumptions and empirical data, can be utilized alongside various imaging modalities, such as fMRI for theory-driven models that correlate cognitive functions with specific brain regions and circuits, and EEG for mechanistic, theory-driven, and predictive models that examine temporally sensitive phenomena [38–40]. MEG, due to its superior temporal and spatial resolution, is applicable for models necessitating accurate timing and localization of brain events, including predictive modeling [41].

Optogenetics is a technique that uses light to regulate and manipulate particular neurons in the brain, enabling researchers to investigate intricate neural circuits associated with many behaviors and brain processes. The fundamental aspect of optogenetics is the utilization of microbial opsins, including Channelrhodopsin-2 (ChR2) and Halorhodopsin (NpHR), which are light-responsive proteins derived from algae and bacteria. These proteins can be genetically modified and introduced into neurons using viral vectors, eliciting a response that can be documented and examined. Optogenetics is particularly intriguing as it may selectively target certain neuronal populations, such as those in the prefrontal cortex, which are implicated in decision-making and social behavior [42].

Systems that enable communication between the brain and external devices are known as brain-computer interfaces (BCIs). They record and analyze brain activity and convert it into commands that may be used to operate computers, prosthetic limbs, or other devices. There are two primary categories: non-invasive BCIs utilizing sensors placed on the scalp or ears, and invasive BCIs that involve the implantation of electrodes directly into the brain. A brain-computer interface (BCI) has prospective uses in the rehabilitation of paralysis, management of neurological disorders, and the gaming and entertainment sectors [43].

The CRISPR-Cas9 (Clustered Regularly Interspaced Short Palindromic Repeats) system relies on the guide RNA (sgRNA) to direct the Cas9 enzyme to the appropriate region of the genome for editing. It is a combination of CRISPR RNA (crRNA) and trans-activating CRISPR RNA (tracrRNA), enabling effective mammalian genome editing. The design of sgRNA, especially the gRNA domain, is essential for its effectiveness and precision. High-specificity sgRNAs are crucial for accurate genome editing, reducing off-target effects, and enhancing the efficacy of CRISPR-Cas9 technology in both research and therapeutic contexts [44].

The disordered electrical impulses in the brain of people with Parkinson's disease and other neurological disorders generate symptoms connected to movement. DBS disrupts these impulses, resulting in tremors and other motor symptoms. Neurosurgeons insert wires known as "leads" into the brain, which are linked to a compact neurostimulator. Consistent electric current pulses traverse the leads, and the neurostimulator is configured to provide an electrical signal that optimizes symptom management while minimizing adverse effects [45].

The National Parkinson's Foundation advocates for daily balance and deep brain stimulation (DBS) surgery for Parkinson's disease patients exhibiting symptoms that impede daily activities, experience mobility fluctuations due to PD medications, demonstrate a consistent positive response to medications, and require multiple combinations of PD medications under the supervision of a neurologist. Nevertheless, certain patients may be less appropriate candidates for DBS surgery owing to issues with balance, speech, cognitive disorientation, memory impairments, mental disorders, or heightened surgical risks.

Employing data-driven methodologies, AI systems can forecast risk variables such as substance misuse, depression, and suicide risk [46]. Nonetheless, validation is required for their generalizability and clinical applicability. AI technologies enhance efficiency in the healthcare business by forecasting treatment modalities and constructing innovative disease pathophysiology models. Machine learning may also identify gene expression patterns associated with psychiatric disorders. Artificial intelligence can assist in predicting pharmaceutical outcomes and identifying novel compounds with therapeutic potential. Nonetheless, AI encounters obstacles in healthcare systems due to the insufficient understanding of the molecular factors behind psychiatric disorders.

By gathering cognitive and behavioral data from people completing comparable tasks, XAI can build human strategy models. This knowledge may be juxtaposed with AI models' tactics, refining users' perceptions of AI decision-making processes. A Human Saliency Imitator forecasts human attention maps in image categorization by employing a deep learning model trained on human attention data [47]. A separate research study created a Human Attention-Guided XAI that optimizes the congruence of XAI saliency maps with human attention maps for object detection tasks. This approach can assist XAI in identifying possible adversarial assaults, including tiny alterations to an image or input that result in misclassification by AI [48].

The investigation of the brain's anatomical and functional interconnections. Graph theory is essential for examining brain connection patterns, offering insights into the communication and collaboration across brain areas. This understanding aids in comprehending cognitive processes and illnesses, enabling individualized therapies aimed at specific brain circuits.

Rapid advancements in brain-computer interfaces (BCIs) hold considerable promise for improving human communication and cognition. Recent advancements encompass implanted devices for those with paralysis and non-invasive

brain-computer interfaces utilizing electroencephalography and transcranial magnetic stimulation for brain-controlled prostheses. The necessity for further refining is to enhance the accessibility and user-friendliness of these technologies.

Neuroplasticity, the brain's extraordinary ability to remodel and alter its structure and function, is being investigated using brain stimulation methods such as transcranial direct current stimulation (tDCS) and transcranial magnetic stimulation (TMS). These strategies are being investigated to improve learning, memory, and rehabilitation in patients with neurological diseases. Nonetheless, ethical concerns and thorough study are essential for the secure and successful application of brain stimulation techniques [49].

The EEG signal, comprising many fundamental frequencies, represents cognitive functions and brain states across time. It is categorized into five sub-frequency ranges: delta, theta, alpha, beta, and gamma waveforms. Delta-band activity correlates with deep sleep, whereas theta waves are linked to awareness, attention, orientation, and working memory. Alpha waves are located in the occipital and posterior areas of the brain, whereas beta waves are mostly linked to awareness or alertness. Gamma frequencies are prevalent across cortical regions and are involved in processes including perception, attention, memory, consciousness, synaptic plasticity, and motor control. These frequencies facilitate the comprehension of brain activity and its influence on cognitive processes [50].

Invasive brain activity recordings and real-time wearable recordings are two examples of emerging neuro-technological technologies that are upending conventional imaging techniques by offering previously unheard-of access to high temporal resolution neural data and ecologically based measures [33, 51, 52]. These developments expand the methodological repertoire of the area, providing new opportunities for both fundamental and clinical study.

Blockchain technology and neuroscience may augment health and technological breakthroughs, especially in the treatment of neurological illnesses. The human brain, consisting of many neurons, generates a series of decentralized events that influence our conscious experiences, akin to blockchain technology. Each episode is safely archived and readily available, mitigating confusion with childhood recollections. Blockchain is a safe mechanism for storing and disseminating information, making it appropriate for tackling neurological issues and neurodegenerative diseases such as Parkinson's and Alzheimer's. A neuron-based consensus system may facilitate rapid and efficient blockchain operations, despite the analogy being complex and dependent on extensive data. The hash function in blockchain technology elucidates the human brain's sensitivity to slight input variations, therefore assuring data security and integrity within a network [53].

Cognitive augmentation, a technology using AI, virtual reality, and brain-computer interfaces, is rapidly advancing, improving cognitive capabilities such as attention, memory recall, and problem-solving, with applications in education, healthcare, and business, notwithstanding ethical issues. Cognitive augmentation is a technology-driven method that improves mental functions such as memory and problem-solving by using AI systems to interpret extensive data more effectively than humans. This collaboration seeks to provide a more efficient and effective method for managing intricate jobs. Nonetheless, ethical dilemmas emerge around accessibility and reliance on technology. As these advancements proliferate, access to sophisticated cognitive tools may be restricted to certain people, exacerbating existing inequities. Furthermore, dependence on technology for cognitive functions may result in a decline in people's inherent cognitive skills over time, prompting worries about autonomy and the ramifications of external technologies on critical thinking [54].

Motor skills are essential for learning and development, with fine motor skills associated with language acquisition in infancy and writing proficiency in older children. Gross motor abilities, including crawling and walking, are associated with overall well-being and scholastic achievements, particularly in mathematics. Excessive screen usage adversely affects fine motor skills and heightens the likelihood of obesity in preschool-aged children [55]. Attention control and executive function capabilities correlate with diverse educational results; television viewing adversely affects children's attention skills, but video games may enhance attention, task-switching, and multitasking abilities. A recent comprehensive analysis, however, revealed no significant effect of screen usage on executive functioning skills in young children [56]. Engaging educational applications focused on vocabulary and language acquisition may enhance linguistic results. Sleep serves as a crucial facilitator for learning, since blue light emitted from screens influences melatonin levels, impairing sleep quality and delaying its onset [57].

A technique called Cognitive IoT uses AI and IoT to enhance systems by analyzing vast amounts of data and coming to more intelligent conclusions. It entails the synthesis of data from the physical environment via sensors and actuators, then relaying it *via* the internet connection to the IoT infrastructure. Cloud computing oversees and evaluates this data, while artificial intelligence elements such as AI, natural language processing, and machine learning facilitate the system's intelligent adaptation and evolution. Edge computing analyzes data near its source, guaranteeing rapid and informative outcomes. Comprehensive security protocols safeguard devices and information, while an intuitive interface ensures seamless interactions between users and devices. Challenges include data overload, security issues, interaction with legacy systems, infrastructure expenses, the absence of standards, analytical complexity, talent scarcity, and fast technical advancements. The cognitive IoT market is projected to reach \$238.1 billion by 2032, representing a sevenfold rise within a decade [58].

Quantum cognition is a theoretical framework in neuroscience that employs quantum physics to explain how the brain works when it comes to things like memory, perception, and making decisions. It tries to use quantum principles to explain cognitive processes that do not make sense or are contradictory. This area is very important for solving research problems in bio-inspired soft computing and adaptive distributed computing systems. An algebraic approach to quantum physics using linear Hermitian operators creates a computer model of cognitive function and consciousness. The model includes decision functions based on neurodynamics and trigonometric wave functions that closely match the way neurons fire. This analytical technique shows how to use quantum mechanical processes to represent brain activation and other cognitive operations [59].

5. Discussion

The conception of computer models for neuroscience and neuroimaging research is being impacted by the philosophy of science [60]. There are two primary categories of models: theory-driven and data-driven. Theory-driven models elucidate the neural computations that underpin certain behavioral and cognitive processes, leveraging existing understanding of brain architecture, functionality, and cellular and molecular mechanisms. They generate testable predictions regarding neurological and behavioral effects, which may be objectively evaluated against observable data via simulations and model fitting. Differentiations in the models' design and parameters

may elucidate the neurological foundations of mental diseases. Data-driven models emphasize data mining for practical objectives, such as uncovering new patterns or forecasting clinical outcomes [61]. They employ machine learning methodologies to uncover empirical correlations within extensive datasets and provide therapeutic applicability via biomarker discovery or treatment protocol personalization [62]. Hybrid models, which blend theory and data, have been employed concurrently, combining diverse levels of theoretical assumptions and empirical evidence. Various imaging modalities are optimally utilized for informing computational models of differing scales, including fMRI for delineating cognitive activities to particular brain areas and circuits, EEG for enhanced temporal resolution, and MEG for accurate timing and localization of neural events [38, 39].

Various neuroscientific data, including online, controlled laboratory, and neuroimaging data, may be integrated using computational models that concentrate on factors such as PEs and learning rates. This methodology is essential for individualized mental wellness *via* neuroscience, necessitating a multimodal perspective to effectively use varied data sources and include innovative techniques [16].

Blockchain technology with neuroscience may improve health and technological advancements, especially in the treatment of neurological illnesses, resolving challenges within healthcare systems, and advancing human cognition across several sectors [63].

Digital technologies have a modest influence on learning and development, influenced by several circumstances. Nonetheless, not all consumers possess the knowledge to use technological equipment effectively. Policy recommendations might mitigate this gap by offering information on using digital technologies to give feedback and engage in discussions about content, even in the absence of technology. Cognitive IoT, integrating intelligent technology with cognitive computing, facilitates autonomous learning and decision-making across several domains, including urban environments, healthcare, and manufacturing. Notwithstanding problems such as data management and security, it has significant promise for intelligent settings.

6. Conclusion

Neuroscience is rapidly progressing with advancements in optogenetics, nanotechnology, and virtual reality, which are transforming research and therapy. These technologies manipulate neurons, provide pharmaceuticals, and enhance rehabilitation regimens. Neuroimaging, connectomics, machine learning, and artificial intelligence are driving progress in neuroscience research and clinical applications. Brain-computer interfaces and brain stimulation methods are elucidating cerebral enigmas and improving neurological conditions. Neurological research aims to understand the nervous system's workings and resolve problems affecting the body, brain, and nervous system. It recognizes impairments and mental diseases, detects, diagnoses, and prognoses using AI for medical data processing.

Brain-computer interfaces and non-invasive brain stimulation technologies are transforming mobility, communication, and brain-machine integration. AI is a potent instrument for advancing neurological research, facilitating the examination of intricate neurological data. With an emphasis on multidisciplinary research to address difficult issues and hasten their integration across sectors, the effort attempts to develop technologies that replicate human cognition and facilitate data-driven decision-making. Analytical approaches in computational neuroscience are essential

for domains such as spiking neural networks, neural field theories, tensor networks, quantum kernel learning, photonic quantum networks for intelligent network computing, and quantum state teleportation. A computational framework including many tiers of three-dimensional picture representation and behavioral manipulation is essential for the next era of neuroscience, with quantum computing possibly ushering in a new phase in the study of the brain.

7. Future enhancement


Through computational modeling, neuroscience may improve the treatment of neurological illnesses and close the gap between neurobiology and cognition. This encompasses deep learning models, neural encoding analysis techniques, and conventional instruments such as parsers, language models, and vector-space models. Future advancements will concentrate on multimodal analysis within computational cognitive neuroscience.

Author details

Vijayalakshmi Kakulapati*, Md Jaffar Sadiq and Aruna Varanasi
Sreenidhi Institute of Science and Technology, Hyderabad, Telangana, India

*Address all correspondence to: vldms@yahoo.com

IntechOpen

© 2025 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Moini J, LoGalbo A, Ahangari R. Cognitive functions. In: Moini J, LoGalbo A, Ahangari R, editors. *Foundations of the Mind, Brain, and Behavioral Relationships*. Academic Press; 2024. pp. 211-228. DOI: 10.1016/B978-0-323-95975-9.00002-0
- [2] Kriegeskorte N, Douglas PK. Cognitive computational neuroscience. *Nature Neuroscience*. 2018;**21**:1148-1160. DOI: 10.1038/s41593-018-0210-5
- [3] Available from: <https://www.emotiv.com/blogs/glossary/computational-neuroscience>.
- [4] Lopopolo A, Fedorenko E, Levy R, Rabovsky M. Cognitive computational neuroscience of language: Using computational models to investigate language processing in the brain. *Neurobiology of Language*. 2024;**5**(1):1-6. DOI: 10.1162/nol_e_00131
- [5] Simon HA, Newell A. Human problem solving: The state of the theory in 1970. *The American Psychologist*. 1971;**26**:145-159
- [6] Anderson JR. *The Architecture of Cognition*. Cambridge, MA, USA: Harvard University Press; 1983
- [7] McClelland JL, Rumelhart DE. *Parallel Distributed Processing*. Cambridge, MA, USA: MIT Press; 1987
- [8] Hebb DO. *The Organization of Behavior: A Neuropsychological Theory*. Wiley; 1949
- [9] Available from: <https://roberttlange.com/posts/2019/09/blog-post-7/>.
- [10] Kriegeskorte N. Deep neural networks: A new framework for modeling biological vision and brain information processing. *Annual Review of Vision Science*. 2015;**1**:417-446
- [11] Yamins DLK, DiCarlo JJ. Using goal-driven deep learning models to understand sensory cortex. *Nature Neuroscience*. 2016;**19**:356-365
- [12] Dolan RJ. Neuroimaging of cognition: Past, present, and future. *Neuron*. 2008;**60**:496-502
- [13] Gu X, Moran R, Srivastava D, Walton M. Harnessing the potential of cellular, animal, and computational neuroscience models for mental health. *Wellcome Trust*. 2023:19-23
- [14] Hauser TU, Skvortsova V, De Choudhury M, Koutsouleris N. The promise of a model-based psychiatry: Building computational models of mental ill health. *Lancet Digital Health*. 2022;**4**:e816-e828
- [15] Garcia-Ceja E, Riegler M, Nordgreen T, Jakobsen P, Oedegaard KJ, Tørresen J. Mental health monitoring with multimodal sensing and machine learning: A survey. *Pervasive and Mobile Computing*. 2018;**51**:1-26
- [16] Graeser M, Thieben F, Szwargulski P, Werner F, Gdaniec N, Boberg M, et al. Human-sized magnetic particle imaging for brain applications. *Nature Communications*. 2019;**10**:1936
- [17] Vaswani A. Attention is all you need. In: *Advances in Neural Information Processing Systems*, 31st Conference on Neural Information Processing Systems (NIPS 2017). Long Beach, CA, USA; 2017
- [18] Goldstein A, Zada Z, Buchnik E, et al. Shared computational principles for

language processing in humans and deep language models. *Nature Neuroscience*. 2022;**25**:369-380. DOI: 10.1038/s41593-022-01026-4

[19] Srikumar V, Kundu G, Roth D. On amortizing inference cost for structured prediction. In: *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*. 2012. pp. 1114-1124

[20] Bengio Y, Scellier B, Bilaniuk O, Sacramento J, Senn W. Feedforward initialization for fast inference of deep generative networks is biologically plausible. arXiv: 1606.01651. 2016. pp. 1-10

[21] Poo MM, Du JL, Ip NY, Xiong ZQ, Xu B, Tan T. China brain project: Basic neuroscience, brain diseases, and brain-inspired computing. *Neuron*. 2016;**92**(3):591-596. DOI: 10.1016/j.neuron.2016.10.050

[22] Shu-Chen L, Fitzek Frank HP, et al. Digitally embodied lifespan neurocognitive development and tactile internet: Transdisciplinary challenges and opportunities. *Frontiers in Human Neuroscience*. 2023;**17**:1-9. DOI: 10.3389/fnhum.2023.1116501

[23] Dayan P. Simple substrates for complex cognition. *Frontiers in Neuroscience*. 2008;**2**(2):255. DOI: 10.3389/neuro.01.031.2008

[24] Todd MT, Niv Y, Cohen JD. Learning to use working memory in partially observable environments through dopaminergic reinforcement. In: Koller D, editor. *Advances in Neural Information Processing Systems (NIPS)* 21. Curran Associates; 2008

[25] Ashby FG, Alfonso-Reese LA, Turken AU, Waldron EM. A

neuropsychological theory of multiple systems in category learning. *Psychological Review*. 1998;**105**:442. DOI: 10.1037/0033-295x.105.3.442

[26] Moustafa AA, Cohen MX, Sherman SJ, Frank MJ. A role for dopamine in temporal decision making and reward maximization in parkinsonism. *Journal of Neuroscience*. 2008;**28**(47):12294-12304. DOI: 10.1523/JNEUROSCI.3116-08.2008

[27] Zhang J, Chan AB, Lau EYY, Hsiao JH. Individuals with insomnia misrecognize angry faces as fearful faces while missing the eyes: An eye-tracking study. *Sleep*. 2019;**42**(2):zsy220. DOI: 10.1093/sleep/zsy220

[28] Wang L, Hsiao JH, Chan AB, Cheung J, Hung S, Au TK. On becoming socially anxious: Toddlers' attention bias for fearful faces. *Developmental Psychology*. 2023;**59**(2):353-363. DOI: 10.1037/dev0001472

[29] Chan FHF, Suen H, Chan AB, Hsiao JH, Barry TJ. The effects of attentional and interpretation biases on later pain outcomes among younger and older adults: A prospective study. *European Journal of Pain*. 2022;**26**(1):181-196. DOI: 10.1002/ejp.1853

[30] Pirolli P, Carley KM, Dalton A, Dorr BJ, Lebiere C, Martin MK, et al. Mining online social media to drive psychologically valid agent models of regional COVID-19 mask wearing. In: Thomson R, Hussain MN, Dancy C, Pyke A, editors. *Social, Cultural, and Behavioral Modeling*. Cham: Springer International Publishing; 2021. pp. 46-56

[31] Laird JE, Lebiere C, Rosenbloom PS. A standard model of the mind: Toward a common computational framework across artificial intelligence, cognitive

- science, neuroscience, and robotics. *AI Magazine*. 2017;**38**:13-26. DOI: 10.1609/aimag.v38i4.2744
- [32] Available from: <https://www.labxchange.org/library/items/lb:LabXchange:8744d3f3.html>:1
- [33] Paulk AC, Kfir Y, Khanna AR, Moustroph ML, Trautmann EM, Soper DJ, et al. Large-scale neural recordings with single-neuron resolution using Neuropixels probes in human cortex. *Nature Neuroscience*. 2022;**25**:252-263
- [34] Stangl M, Maoz SL, Suthana N. Mobile cognition: Imaging the human brain in the 'real world'. *Nature Reviews Neuroscience*. 2023;**24**:347-362
- [35] Pratte MS, Tong F. Integrating theoretical models with functional neuroimaging. *Journal of Mathematical Psychology*. 2017;**76**(80):93
- [36] Yip SW, Konova AB. Emerging topics in computational psychiatric research: Clarity through complexity? *Biological Psychiatry*. 2023;**93**:652-654
- [37] Linderman SW, Gershman SJ. Using computational theory to constrain statistical models of neural data. *Current Opinion in Neurobiology*. 2017;**46**(14):24
- [38] Türker B, Musat EM, Chabani E, Fonteix-Galet A, Maranci J-B, Wattiez N, et al. Behavioral and brain responses to verbal stimuli reveal transient periods of cognitive integration of the external world during sleep. *Nature Neuroscience*. 2023;**26**:1981-1993
- [39] Tang MF, Ford L, Arabzadeh E, Enns JT, Visser TAW, Mattingley JB. Neural dynamics of the attentional blink revealed by encoding orientation selectivity during rapid visual presentation. *Nature Communications*. 2020;**11**:434
- [40] Sun J, Sun Y, Shen A, Li Y, Gao X, Lu B. An ensemble learning model for continuous cognition assessment based on resting-state EEG. *NPJ Aging*. 2024;**10**:1-10
- [41] Hunt LT, Kolling N, Soltani A, Woolrich MW, Rushworth MF, Behrens TE. Mechanisms underlying cortical activity during value-guided choice. *Nature Neuroscience*. 2012;**15**:470-483
- [42] Available from: <https://www.tomorrow.bio/post/top-10-latest-technologies-in-neuroscience-2023-06-4603565337-futurism>.
- [43] Available from: <https://medium.com/@Shan3Y/brain-computer-interfaces-bridging-the-mind-machine-gap-43a1235eb403#:~:text=BCIs%20are%20systems%20that%20allow,%2C%20prosthetics%2C%20or%20other%20devices>.
- [44] Available from: <https://www.kbdna.com/crispr-cas9-revolutionizing-gene-editing-in-biomedical-research#:~:text=CRISPR%2DCas9%20is%20hailed%20as,based%20single%2Dcell%20molecular%20screens>.
- [45] Available from: <https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/deep-brain-stimulation>.
- [46] Lee EE. Artificial intelligence for mental health care: Clinical applications, barriers, facilitators, and artificial wisdom. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*. 2021;**6**(9):856-864. DOI: 10.1016/j.bpsc.2021.02.001
- [47] Yang Y, Zheng Y, Deng D, Yang Y, Huang Y, Hsiao JH, et al. HSI: Human saliency imitator for benchmarking saliency-based model explanations.

Proceedings of the Tenth AAAI Conference on Human Computation and Crowdsourcing. 2022;**10**:231-242

[48] Samek W, Binder A, Montavon G, Lapuschkin S, Muller K-R. Evaluating the visualization of what a deep neural network has learned. *IEEE Transactions on Neural Networks and Learning Systems*. 2016;**28**(11):2660-2673

[49] Available from: <https://timesofindia.indiatimes.com/blogs/voices/new-trends-and-technologies-in-neuroscience/>.

[50] Liu S, Wang L, Gao RX. Cognitive neuroscience and robotics: Advancements and future research directions. *Robotics and Computer-Integrated Manufacturing*. 2024;**85**:102610. DOI: 10.1016/j.rcim.2023.102610

[51] Ekstrom AD, Kahana MJ, Caplan JB, Fields TA, Isham EA, Newman EL, et al. Cellular networks underlying human spatial navigation. *Nature*. 2003;**425**:184-188

[52] Jacobs J, Weidemann CT, Miller JF, Solway A, Burke JF, Wei X-X, et al. Direct recordings of grid-like neuronal activity in human spatial navigation. *Nature Neuroscience*. 2013;**16**:1188-1190

[53] Available from: <https://www.linkedin.com/pulse/neuroscience-blockchain-technology-hamed-taherdoost>.

[54] Available from: <https://library.fiveable.me/key-terms/introduction-cognitive-science/cognitive-augmentation>.

[55] Li C, Cheng G, Sha T, Cheng W, Yan Y. The relationships between screen use and health indicators among infants, toddlers, and preschoolers:

A meta-analysis and systematic review. *International Journal of Environmental Research and Public Health*. 2020;**17**(19):1-19. DOI: 10.3390/ijerph17197324

[56] Lodge JM, Harrison WJ. The role of attention in learning in the digital age. *The Yale Journal of Biology and Medicine*. 2019;**92**(1):21-28

[57] Cheung CHM, Bedford R, Saez De Urabain IR, Karmiloff-Smith A, Smith TJ. Daily touchscreen use in infants and toddlers is associated with reduced sleep and delayed sleep onset. *Scientific Reports*. 2017;**7**(1). DOI: 10.1038/srep46104

[58] Available from: <https://www.neilsahota.com/cognitive-iot-integrating-cognitive-computing-with-iot/>

[59] Bagchi S. The quantum phenomena in computational model of neuro-cognition states: An analytical approach. *NeuroQuantology*. 2015;**13**(3):285-292

[60] Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science*. 1997;**275**:1593-1599

[61] Bishop CM. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Berlin, Heidelberg: Springer-Verlag; 2006

[62] Nielsen AN, Barch DM, Petersen SE, Schlaggar BL, Greene DJ. Machine learning with neuroimaging: Evaluating its applications in psychiatry. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*. 2020;**5**:791-798

[63] Taherdoost H. Neuroscience and blockchain. *Archives in Neurology and Neuroscience*. 2022;**12**. DOI: 10.33552/ANN.2022.12.000794

Neurons on Microelectrode Arrays and *In Vitro* Electrophysiological Data Analysis

Andrey Vinogradov, Laura Ylä-Outinen, Susanna Narkilahti and Emre Kapucu

Abstract

Microelectrode array (MEA) technology has been used in functional neuronal network studies for decades, enabling the analysis of extracellular neuronal activity in both *in vivo* and *in vitro*. The proper interpretation of MEA recordings relies on effective neuronal signal processing and analysis. A key challenge in the field is accurate detection of extracellular neuronal activity (spikes) and identifying their occurrence in dense temporal clusters known as bursts. These bursts can be recorded by individual electrodes (single-channel bursts) or across multiple MEA electrodes (network bursts), adding substantial complexity to the analysis. This chapter provides a comprehensive overview of MEA signal analysis techniques, including methods for spike detection, burst identification, and the assessment of functional connectivity within neuronal cultures. We also include the studies involving the multi-compartment MEAs in the context of network activity and connectivity assessment, since such MEA setups offer a spatiotemporal depth in neural network interactions *in vitro*, facilitating more thorough insights into neurodevelopment and disease modeling. By integrating various analysis methods, the chapter offers a framework for understanding the complex dynamics of neuronal networks and their functional connectivity, making it a valuable resource for researchers using MEA technology for analyzing neuronal electrophysiology.

Keywords: burst, EAP, extracellular action potential, extracellular electrophysiology, functional connectivity, MEA, microelectrode array, network burst, neuron, signal analysis, spike

1. Introduction

The neurons belong to electrically excitable cells, as their membrane is capable of rapid impulse conduction: a transient voltage perturbation—the action potential, also called a spike. The brain's remarkably complex function arises from the electrical activity of billions of neurons. The brain encodes information through the frequencies and timing patterns of action potentials generated by neuronal populations, which is rather difficult to study non-invasively [1–3].

The electrical properties of neuronal activity allow for distant acquisition through microelectrodes [4, 5]. However, accessing the micro- and mesoscale electrophysiological dynamics of the human brain *in vivo* presents significant challenges. On the other hand, these dynamics are effectively studied using animal models and *in vitro* human pluripotent stem cell-derived neuronal cultures by utilizing microelectrode arrays (MEAs) to record extracellular neuronal electrophysiological activity in both *in vivo* and *in vitro* settings [4–7].

Unlike intracellular recordings, MEAs capture activity from multiple neurons across several locations simultaneously [4, 5, 8–11], offering insights into network function and connectivity, which are crucial properties of neuronal networks. The obtained information about electrophysiological activity and functional interactions of the neurons serves as a proxy measure of synaptic transmissions [11]. Moreover, MEAs offer a non-invasive method to observe the electrical activity of the neuronal cells, thus allowing sequential or time-lapse follow-up of the same network. This approach is especially useful in developmental studies as well as mid- or long-term exposure experiments [3, 12, 13].

In vivo studies have the advantage of investigating neuronal networks' dynamics in their actual physiological environment; however, *in vitro* approaches offer a more controlled setting for analyzing interactions from the subcellular to the network level. Moreover, the *in vitro* application of MEAs is particularly valuable due to their non-invasive nature compared to the patch clamp technique. These capabilities facilitate the study of developmental processes, neurodegeneration, and long-term pharmacological effects [13–15]. A key advantage of *in vitro* functional research using MEAs is the ability to efficiently integrate pharmacological treatments, electrical stimulation, fluorescent labeling, and targeted optogenetic activation to enhance experimental outcomes [4, 11, 16].

MEA analysis of *in vitro* neuronal cultures, ranging in complexity from single to multiple cell types and from 2D to 3D network structures, is a powerful tool for studying neuronal dynamics by means of electrophysiology. Electrophysiological recordings of neuronal populations with MEAs rely on electrodes that vary in size, density, and spatial distribution. Standard *in vitro* MEAs consist of multiple electrodes embedded at the bottom of the culture wells, where neuronal cells are plated. This setup enables mesoscale electrophysiological recordings, allowing researchers to analyze neuronal activity at the population level. Each electrode detects superimposed signals from nearby neurons [4]. Two examples of the standard commercially available MEA plates produced by Axion Biosystems are shown in **Figure 1A**. Recently, so-called high-density (HD) MEAs have also been utilized widely. By reducing the electrode size and increasing the density, hence obtaining the ability to record thousands of electrodes, HD MEAs are advantageous for measuring subcellular neuronal activities. It is an important achievement toward microscale electrophysiology with the possibility to capture axonal signaling and spatially track the signal propagation [20, 21].

In vitro electrophysiological experiments require stable and controlled environmental conditions to ensure neuronal viability and reproducible signal acquisition. This necessity influences both the design and practical use of MEA systems. Notably, neuronal cultures must remain in an incubator to maintain stable temperature, humidity, and gas levels essential for their survival. Consequently, MEA plates are typically designed to be portable, while the remaining hardware components are integrated into an MEA headstage. The headstage is connected to a computer that runs MEA recording software. During recording sessions, the MEA plate is temporarily

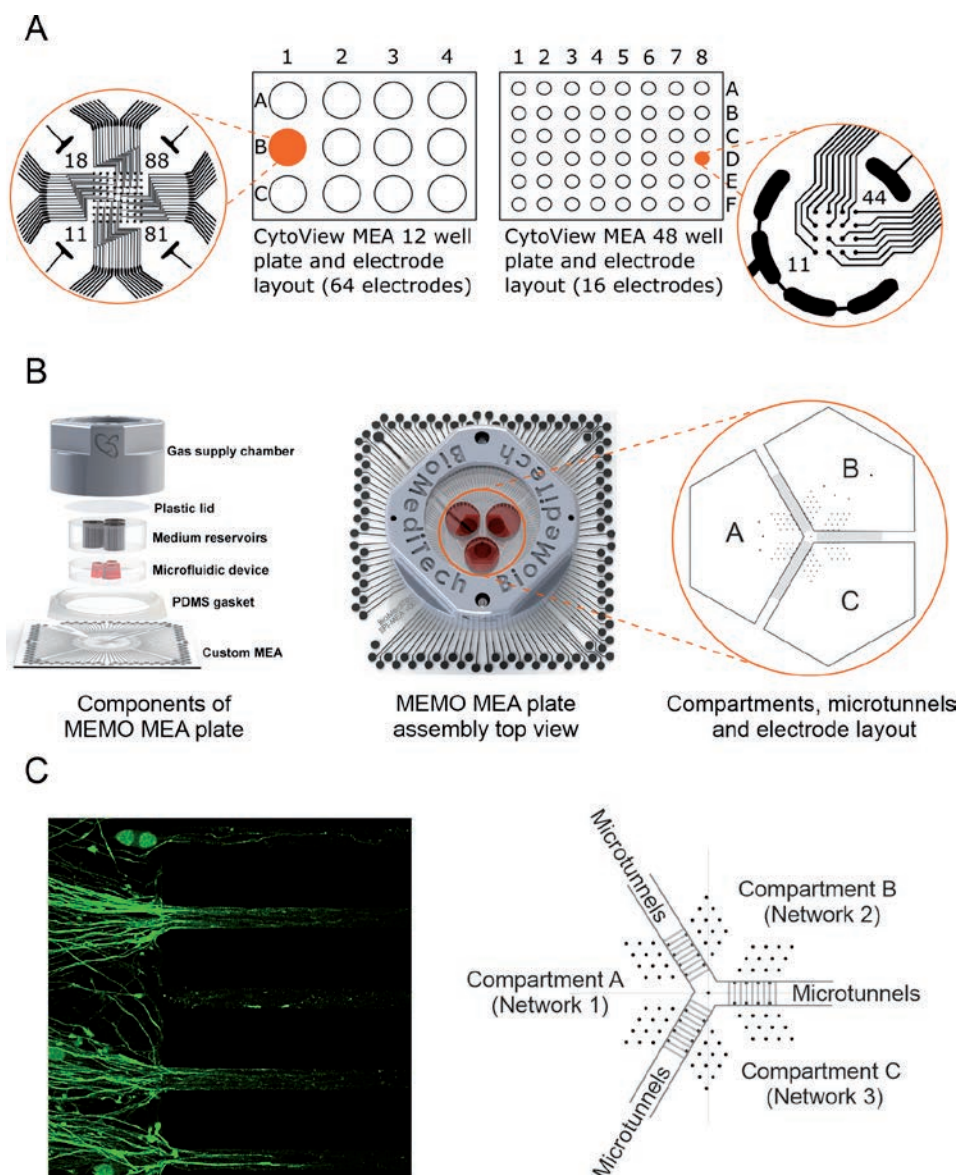


Figure 1.
 A. Schemes of well and electrode alignment of Axion MEA well plates. CytoView MEA 12 has 64 electrodes per well, and CytoView MEA 48 has 16 electrodes per well. Modified from Ref. [17]. B. Custom MEMO MEA plate components and scheme of its three compartments with microtunnels and electrode layout. Modified from Ref. [18]. C. (Left) Immunofluorescent images show neuronal axons penetrating the microtunnels. (Right) Zoomed schematic of a central part of the MEMO MEA plate. The black dots represent the electrodes, and the gray lines between compartments depict the microtunnels designed to enable axonal connections between three neuronal cultures. Each compartment has 24 electrodes. An additional eight electrodes per compartment are located near the entries of the microtunnels. Modified from Ref. [19].

positioned on the headstage [4, 22]. As an example, the components of the MEA2100-System produced by Multi Channel Systems MCS GmbH [23] are shown in **Figure 2A**. The proper temperature and gas environment can be maintained with integrated commercial solutions [24, 25] or by in-house developed designs [26, 27]. Some of the

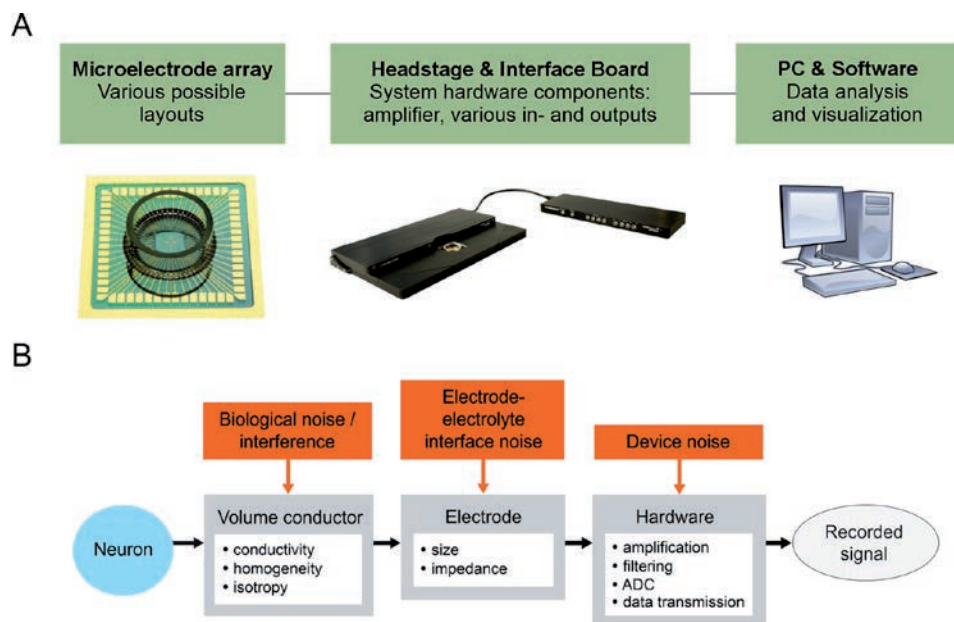


Figure 2.
A. Components of MEA2100-System. Modified from Ref. [23]. B. The block scheme illustrates the main contributors to the noise of the MEA system. Modified from Ref. [4].

most recent commercial solutions have the option to operate the system continuously in an incubator [28].

Novel culturing paradigms also impose another state-of-the-art subtype of *in vitro* MEA devices, multicompartmental MEAs. Such MEAs customized by different labs allow the study of the interactions between separated neuronal networks [18, 29–37], in some cases representing interconnectivities between different brain regions [36] or innervation to other cell types [38] as an advancement over cocultures located in the same compartments [39]. Some commercial solutions also exist for multi-compartmentalized MEAs [40]. An example of the custom three-compartment microfluidic MEMO MEA plate [18, 34, 41] is depicted in **Figure 1B–C**. The MEMO MEA plates are compatible with the MEA2100-System produced by Multi Channel Systems MCS GmbH.

This chapter focuses on the emerging use of *in vitro* neuronal cultures on MEAs and methods for analyzing their recorded signals. Precise analysis of recorded electrophysiological data is fundamental for research on neuronal development, brain injury, disease modeling, and pharmacological studies. Gaining insights into neuronal interactions will help deepen our understanding of neural circuit function in the brain.

2. Microelectrode array signals

When an electrode is positioned close enough to a spiking neuron, the recorded extracellular voltage signal typically exhibits a distinct and easily recognizable action potential signature. Since extracellular electrodes do not disrupt neural activity, these recordings offer an accurate and impartial representation of the activity within

intact neural circuits [11]. Electrophysiological research with *in vitro* MEAs typically involves culturing cells in MEA plates on top of the electrode grid to enable repeated electrophysiological recordings. These plates contain one or more wells with embedded electrodes at the bottom, onto which neuronal cells are seeded. The wells are filled with a cell maintenance medium that provides essential nutrients and chemicals to sustain cell viability [4]. This setup supports functional neuronal cultures in either 2D [7] or 3D [42, 43] configurations, where cells are distributed across the surface of the recording microelectrodes, which capture extracellular electric potential fluctuations [44]. Each microelectrode captures variations in the extracellular electric field within its local environment. Typically, these electrodes register activity from sources located within a range of up to 100 μm [22, 45]; however, the signal amplitude diminishes as the distance between the source and the electrode increases [46]. Additionally, an electrode's sensitivity is influenced by multiple factors, including the properties of the surrounding medium [47], the electrode material, and the characteristics of other hardware components of the MEA system [4].

The voltage signal detected by an MEA electrode results from the linear summation of contributions from current sources, with each contribution weighed inversely to its distance from the electrode. Calculating the voltage signal at the electrode necessitates two key assumptions: (1) neurons can be modeled as a spatial distribution of ion currents and sinks, approximated as monopolar point sources, and (2) the surrounding medium is homogeneous, isotropic, and unbounded. Consequently, the potential V_e (V) at the electrode location, influenced by N point sources, is determined using the following equation:

$$V_e = \frac{1}{4\pi\sigma} \sum_n^N \frac{I_n}{r_n} \quad (1)$$

where I_n is the amplitude (A) of a point source n , located at distance r (m) from the electrode; σ stands for the medium conductivity (S/m). The point source approximation for ion currents and sinks assumes that the volumes of their source regions are significantly smaller than their distances from the recording electrode, enabling them to be represented as points rather than complex shapes [4, 44, 48–50].

The recorded MEA signal incorporates multiple noise sources that degrade the recorded signal. Non-biological noise sources include $1/f$ or $1/f^2$ electrode-electrolyte interface noise and thermal noise. Moreover, digital recordings of the voltage signal and electrical stimulation require a connection between the MEA plate electrodes and various hardware, including amplifiers, filters, a digitizer, data transmission cables, a stimulator, and power lines. All these components of MEA systems contribute further to device noise. Signal distortions may also result from phase noise and lossy data compression methods [4, 51–53]. A block scheme of noise contributors in MEA recordings is shown in **Figure 2B**.

The biological components of the signal include the spikes and less prominent local field potential (LFP) fluctuations [5]. The LFP component contains hard-to-detect extracellular electric field changes, which originate from subthreshold transmembrane activities of neurons and distant spiking activity. The LFP resides in the lower part of the spectrum of the signal. However, a part of the LFP component is usually considered as biological noise. In such cases, the desired signal and the noise have overlapping spectral characteristics [4, 50, 52, 54–56]. Although the LFP component is often challenging to analyze, it can still provide valuable insights. *In vivo* MEA recordings from the human primary motor cortex [57] and monkey motor cortices [58, 59] have

demonstrated that the LFP carries rich movement-related information, which can be extracted even in the absence of detectable spike activity. Additionally, LFP analysis can reveal new insights into postsynaptic potentials (PSPs) in neuronal cultures [22].

In contrast, spikes buried in the MEA signal serve as precise indicators of neuronal action potentials. The initial step in processing a raw MEA signal typically involves pre-filtering to distinguish between the LFP and spiking components. The spike detection typically begins with the band-pass filtering within a frequency range of approximately 300–3000 Hz, whereas LFP analysis involves low-pass filtering with cutoff frequencies below 300 Hz [5, 56, 60–62].

Next, we will consider spike detection and consecutive analyses based on the spiking activity.

3. Microelectrode array signal analysis

Typically, the MEA signal analysis includes the detection of neuronal spikes in pre-filtered data, followed by the identification of spike-based features, such as firing rate (Section 3.1) or single-channel bursts and network bursts (NBs; Section 3.2). These analysis steps are visualized in **Figure 3**. Moreover, additional investigation of functional connectivity measures is often applied to map the functional topology of the networks [63–65] (Section 3.3). The functional connectivity evaluation methods may utilize both raw MEA data and the detected spikes [63–65]. Each analysis step is

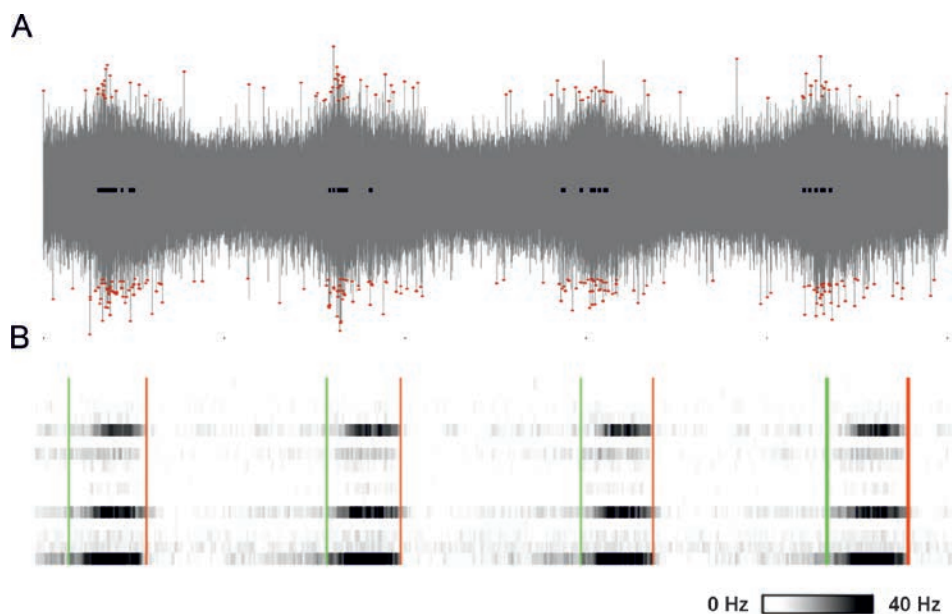


Figure 3.

A. Exemplary signal recorded by a single MEA electrode. The pre-filtered waveform is shown in gray; detected spikes are highlighted with red circles; and periods identified as single-channel bursts are marked by black lines superimposed on the middle of the waveform. B. The raster plot demonstrates a segment of data recorded from 16 electrodes simultaneously. The spike train from each electrode is represented as a row of colored bins in time. Each row corresponds to the spike train of a separate electrode. The rows from all electrodes are stacked on top of each other. The bin color encodes firing rate within the bin: white indicates 0 Hz, and black indicates 40 Hz. Green and red lines denote the start and end of detected NBs, respectively. The data were acquired with the Axion CytoView MEA 48 plate, which has 16 electrodes per well. Modified from Ref. [19].

associated with the derivation of numerous feature-specific output parameters, which are then used for scientific inference.

3.1 Spike detection and sorting

A single neuronal excitation, the extracellular action potential or the spike, is a key feature to extract from the pre-filtered signal of each electrode in an MEA [66], and it can be characterized by two main features: an amplitude peak and a biphasic shape [11]. These two characteristics are essential for designing spike detection algorithms. The segment of the signal induced by a spike in extracellular recordings typically lasts about 3 ms [11].

The typical shape of an extracellularly acquired action potential is considered biphasic. The action potential begins with a rapid influx of Na^+ ions, forming a current sink that generates a prominent negative deflection in the extracellular action potential. This is followed by a slower efflux of K^+ ions, creating a current source and leading to a smaller positive deflection [4]. However, the final signal shape captured by an electrode tends to depend on its position relative to the neuron's morphology: with large negative spikes recorded at the perisomatic area and positive spikes at the dendritic area, generated by balancing outward return current as the net membrane current summed over the entire neuron is always zero [4, 45, 67, 68]. Thus, it is rather difficult to conclude which polarity of the spike waveform should be considered characteristic.

One of the simplest spike detection methods involves thresholding based on the amplitude peaks: either negative or both negative and positive. Thresholds can be either static, defined by the user, or adaptive, computed from the data [69, 70]. Static thresholding struggles with the inherent variability in MEA data since noise levels can differ across channels and change over time. To overcome this limitation, adaptive thresholding techniques were introduced. The adaptive threshold is typically set in relation to the noise level, which is estimated from segments of the raw signal that are free of spikes [11]. The most common adaptive method, known as absolute amplitude-based thresholding (ABS), calculates the threshold by estimating the noise's standard deviation (SD), multiplied by a threshold calculation factor (TCF). This estimation uses the median of the absolute values of the signal to prevent excessively high thresholds [69]. The block diagram of the ABS spike detection method is shown in **Figure 4** below.



Figure 4.
Block diagram of the ABS algorithm.

The performance of this algorithm still decreases significantly when the signal-to-noise ratio (SNR) is low [46]. Despite this, the amplitude-based thresholding algorithms remain widely used in both *in vivo* and *in vitro* MEA recordings [60, 62, 71–76]. Another amplitude-based method is Precision Timing Spike Detection (PTSD), which detects spikes by comparing the difference between the local maximum and minimum (RMM) within a predefined time window denoted as a peak lifetime period (PLP). If the peak-to-peak amplitude exceeds the Spike Detection Differential Threshold (SDDT), the spike is identified. The SDDT is determined by multiplying the signal's SD by a specific TCF [77, 78]. This method captures both primary spike features: the transient amplitude peak and biphasic shape. The block diagram of the PTSD algorithm is provided in **Figure 5** below.

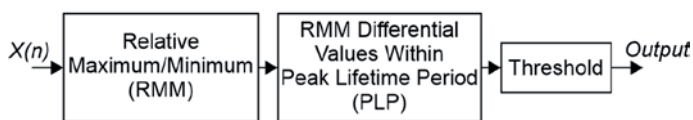


Figure 5.
Block diagram of the PTSD algorithm.

A different category of spike detection techniques focuses on transforming the signal into a specialized variable, known as the transformation variable or indicator signal, which is subsequently subjected to thresholding. This variable is designed to emphasize the key characteristics of the spikes and separate them from the noise. For instance, the Multiresolution Teager Energy Operator algorithm (MTEO) utilizes the Teager Energy Operator (TEO) concept, also known as Nonlinear Energy Operator (NEO), modified with a resolution parameter to detect various possible widths of action potential amplitude peaks. The discrete TEO is computed by finding the difference between the squared signal value at a given time point and the product of its adjacent samples. The algorithm highlights the transient amplitude changes in the signal. The corresponding equation is provided below:

$$\Psi \{x(n)\} = x^2(n) - x(n-k)x(n+k), \quad (2)$$

where $x(n)$ is the input signal, and k is the resolution parameter to reflect the width of the spikes to detect. The authors proposed utilizing multiple k values to generate several transformation variables, which were subsequently smoothed and processed through a maximum filter, which picked the maximum value from the three smoothed TEO outputs for every time instance, and thresholded it [54, 79]. The block diagram of the MTEO algorithm is presented in **Figure 6** below.

On the other hand, the transformation variable can be derived through various types of wavelet decomposition applied to the input signal. A wavelet is the probing function of finite duration, which can be described as a short-lived waveform or localized pulse [80, 81]. Wavelet basis functions typically exhibit biphasic, triphasic, or multiphasic waveforms with compact support, making them similar to action potentials. These functions are translated or slid over the signal to localize the time instants where the signal resembles the wavelet shape or its dilated versions in the time domain, though the exact waveform of the extracellular action potential is not known. The obtained output exhibits local maxima at points where the input signal closely matches the shape of the wavelet function [82, 83]. The wavelet decomposition is usually implemented with a pair

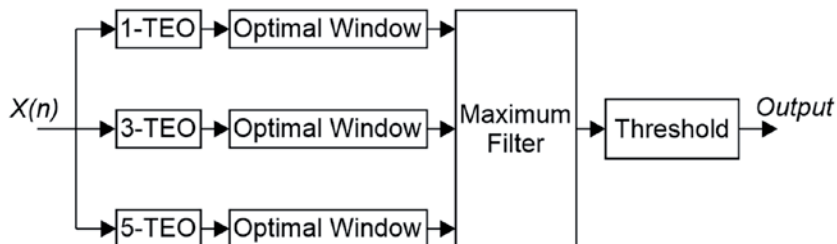


Figure 6.
Block diagram of the MTEO algorithm. Modified from Ref. [54].

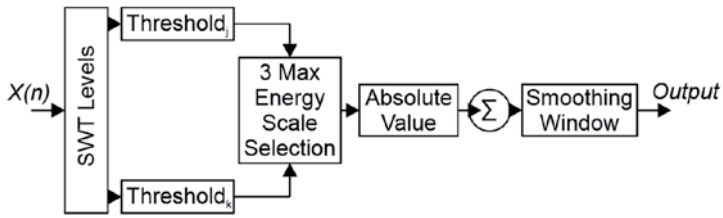


Figure 7.
 Block diagram of the SWT algorithm. Modified from Ref. [82].

of wavelets: a mother wavelet, also called the wavelet function, and a father wavelet, also called the scaling function. The mother wavelet is designed to capture a detailed (or high-frequency) profile of the input, while the father wavelet emphasizes the low-frequency approximation of the input signal. The resulting coefficients are denoted as detail coefficients and approximation coefficients, respectively. This multiresolution approach allows for efficient time-frequency analysis, particularly useful for non-stationary signals like MEA data [84]. A classical multilevel discrete wavelet transform (DWT) shifts to the next decomposition level, also called a scale, by downsampling the obtained coefficients and decomposing them with the same mother and father wavelets. Due to the downsampling, the classical DWT is not a time-invariant transform, which means that the DWT of a signal and the DWT of the same signal shifted in time do not differ in the shift of the coefficients only [46, 84]. A stationary wavelet transform is a time-invariant version of the DWT. Instead of downsampling the obtained coefficients, it moves to the next decomposition level by upsampling the mother and father wavelets [46, 82, 84].

Thus, in spike detection, wavelet-based methods are specifically designed to spotlight the distinctive shapes of spikes within an input signal. The previously published Stationary Wavelet Transform (SWT)-based spike detection algorithm involves denoising wavelet detail coefficients using MAD-based thresholding of each sub-band [82]. Since the signal's energy is concentrated within a limited number of coefficients while noise is distributed across many coefficients, such thresholding helps eliminate low-energy time-scale points across all scales [82]. The block diagram of the SWT algorithm is presented in **Figure 7** below.

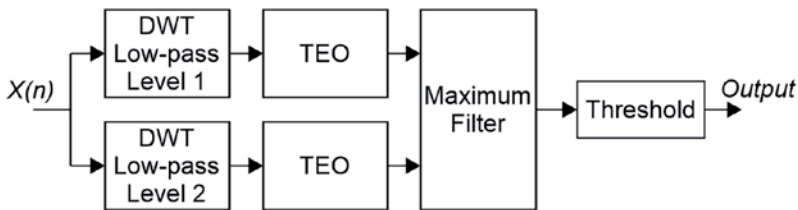


Figure 8.
 Block diagram of the WTEO algorithm. Adopted from Refs. [46, 83].

Another technique, the Wavelet-based Teager Energy Operator algorithm (WTEO), combines multilevel DWT using only the father wavelet with the application of the TEO to the resulting approximation coefficients of each sub-band [83]. The TEO outputs are combined with the maximum filter and then thresholded using an SD-based threshold [46, 83]. The block diagram of the WTEO algorithm is presented in **Figure 8** below.

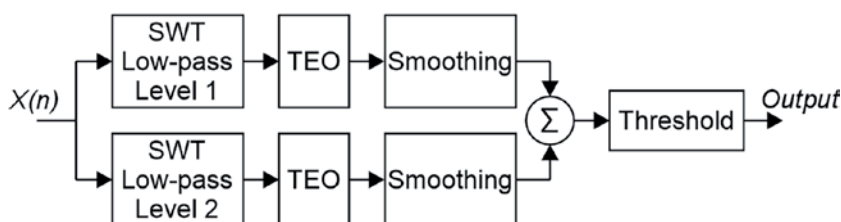


Figure 9.
Block diagram of the SWTTEO algorithm. Modified from Ref. [46].

Building on a similar principle, the Stationary Wavelet Transform-based Teager Energy Operator method (SWTTEO) utilizes the stationary wavelet transform instead of traditional wavelet decomposition [46]. The block diagram of the SWTTEO algorithm is presented in **Figure 9** below.

However, wavelet-based methods rely on the selection of a suitable wavelet basis—a pair of wavelets whose shape should roughly match that of the spikes, despite the actual spike morphology being unknown. Several alternative advanced methods were developed to avoid using such predefined spike shape characteristics. In those, the spike templates were derived from the MEA signal using higher-order statistical measures and subsequently convolved with the raw signal for spike detection [85, 86]. Alternatively, some recent approaches employed supervised machine learning (ML) techniques for spike detection, such as pre-trained convolutional neural networks (CNNs). These models were trained using a combination of hand-curated datasets and synthetic data, which significantly enhanced detection accuracy. Another strategy involves conducting a preliminary analysis to establish a “ground truth” using conventional spike detection methods, which is then used to train the model [87–89].

The spike detection algorithms are typically assessed on synthetic datasets crafted to mimic real MEA recordings. However, evaluations using different synthetic datasets often yield markedly different performance ratings for the same algorithms [46, 66, 86].

In certain experimental setups, the spike detection is followed by spike sorting to further analyze the detected spikes. For example, analyzing single-neuron activity requires accurately assigning each detected spike to its originating neuron [11]. This process involves extracting spike events as signal segments spanning over several tens of samples corresponding to the putative spike duration. These segments are then projected into a lower-dimensional space using techniques such as principal component analysis (PCA) or shape-based metrics such as peak-to-peak amplitude or spike width [90, 91]. Clustering algorithms are then applied to sort the spikes. Alternatively, the extracted waveforms can be decomposed using wavelet transforms, with selected wavelet coefficients serving as inputs for further clustering [69, 92]. The challenges imposed on spike sorting techniques arise from various features of MEA recordings, such as the interference of overlapping waveforms, spike shape changes during bursts, drift of neurons relative to their initial positions on the electrodes, and the unknown number of neurons within the electrode’s receptive field [11]. In conventional MEAs, many neurons usually contribute to spikes at the electrode, while in HD MEA setups, a single spike is clearly reflected in several electrode signals [11].

The general logic of the spike detection is rather consistent across standard MEAs, HD MEAs, and multicompartmental MEA setups. However, the choice of a method, its parameterization, and the interpretation of results should reflect the specific

characteristics of the MEA plate used, such as electrode size and density, cell type-related activity features, and other aspects of the experimental design.

3.2 Bursts

Bursts reflect network dynamics and are typically characterized by the simultaneous occurrence of frequent neuronal activity. Bursts imply compressed information transfer between neurons, consisting of a series of tightly packed single spikes. Bursts are recorded from single or several electrodes of an MEA and defined as single-channel bursts and NBs, respectively [11, 93, 94]. Although bursting is a key feature commonly used to evaluate neuronal network activity, a clear and universally accepted formal definition of a burst is still lacking in the field after more than 50 years of MEA research. As a result, burst detection strategies within the scientific community remain diverse and are often tailored to specific data at hand [11, 95, 96].

Bursts are believed to play a role in long-term potentiation and other processes related to synaptic transmission. This bursting activity contributes to a consistent and information-rich mode of neuronal communication [11, 93, 97–99]. *In vitro* bursts are considered to represent information packages similar to those observed *in vivo* neuronal networks [100]. Importantly, network bursts are characterized by their broader propagation within the network—they travel from cell to cell and across different regions of the neuronal network. The appearance of NBs in MEA data indicates the functional maturation of the recorded neuronal populations. NBs can be used to study the interconnectivity of the networks as well as connectivity velocity. Moreover, they represent the desired activity state for pharmacological experiments [18].

The lack of definition of bursts, together with unavailable ground truth MEA data, imposes serious constraints on the performance evaluation of the single-channel burst and NB detection algorithms. The efficiency assessment of the algorithms is usually achieved with synthetic data sets and visual performance estimation [11, 94, 99, 101].

3.2.1 Single-channel burst detection methods

Single-channel burst firing determines distinct developmental patterns in neuronal populations during MEA recordings [7, 98]. Numerous algorithms have been developed for detecting single-channel bursts [99], with many relying on interspike interval (ISI) characteristics within the recorded spike trains. The ISI-based approach may be further constrained with an additional criterion for a minimum number of spikes per burst to avoid detecting very brief single-channel bursts. The ISI threshold used for ISI-based detection can either be a fixed value [98] or be determined from the distribution of ISIs observed in the recorded signal [93, 101].

For example, the adaptive single-channel burst detection algorithm introduced by Pasquale and colleagues [93] constructed the logarithmic histogram of all ISIs of an input spike train. Then all peaks of the histogram were identified. The set of peaks was checked to contain the putative intra-burst peak: a maximum peak with an ISI value below the predefined intra-burst peak threshold. This peak was considered to represent the ISIs of intra-burst spikes. If no such peak was found, the data was marked as non-bursting. If the intra-burst peak was present, the algorithm evaluated the local minima between this peak and all other peaks that corresponded to longer ISI values. The evaluation criterion for the candidate minima was denoted as a *void* parameter, which reflected the quality of the peak separation and spanned from 0

when the peaks were not separated to 1 when the peaks were perfectly separated. The ISI value of the candidate minimum, which *void* parameter exceeded a predefined *void* threshold, was then selected as an ISI threshold, ISI_{th} , for the single-channel burst detection. If the previously mentioned intra-burst peak threshold appeared lower than the obtained ISI_{th} , both thresholds were used for detecting burst cores and extending their boundaries, correspondingly [93]. An example of the logarithmic histogram with highlighted peaks, intra-burst peak threshold, and derived ISI threshold ISI_{th} is depicted in **Figure 10A**.

Alternatively, the algorithm created by Kapucu and colleagues [101] computed a non-logarithmic histogram of ISIs and then derived the cumulative moving average (CMA) value for each bin. The maximum CMA value, CMA_m , was then used to identify the bin whose CMA value was closest to the product αCMA_m , where α denoted the tolerance parameter to identify the burst-related spikes. The ISI value x_t of that bin was then selected as an ISI threshold for the single-channel burst detection. The tolerance parameter α may vary from 0 to 1 [101]. The non-logarithmic ISI histogram with the overlaid CMA curve and the defined ISI threshold is depicted in **Figure 10B**. Both described algorithms process the ISI histogram to determine the ISI threshold that is assumed to separate the in-burst and out-burst firing regimes most efficiently. The data with the developed bursting activity tend to produce histograms

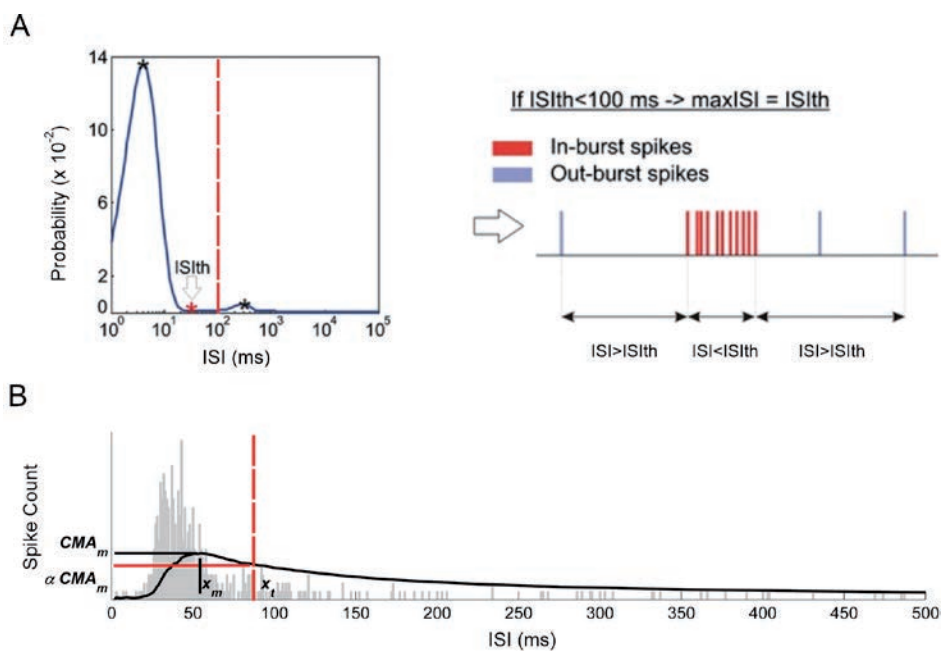


Figure 10. Two ISI-based single-channel burst detection methods. A. (Left) Logarithmic distribution of ISI with the red star-marked ISI threshold ISI_{th} that was selected at the local minima based on the peaks' separation criterion. The corresponding peaks are marked with the black star symbols. An intra-burst peak threshold of 100 ms is depicted with the red dashed line. The horizontal axis of the histogram is logarithmic. (Right) Detection outcome: an exemplary spike train with the red spikes classified as belonging to the single-channel burst and the blue spikes marked as the out-burst spikes. The spike train is presented as a point process in time. Reproduced from Ref. [93]. B. Definition of the ISI threshold using the maximum value of the CMA curve of the ISI histogram. The maximum value of CMA_m was reached at the bin with ISI value x_m , and the ISI threshold x_t was defined as the ISI value of the bin that corresponded to the αCMA_m value. The defined ISI threshold for burst detection is marked with the red dashed vertical line. The ISI histogram is depicted with gray bars. The axes of the histogram are linear. Reproduced from Ref. [101].

with an evident peak corresponding to shorter ISI values on the left side of the histogram.

On the other hand, there exists a family of so-called surprise-based methods for single-channel burst detection. These algorithms use statistical analysis to distinguish periods of bursting from baseline neuronal firing. Initially, a suitable assumption about the distribution of baseline spike firing times is developed. The bursting regimes are expected to produce evident deviation from this distribution, which is used to identify the single-channel bursts [96, 102, 103].

The spike timing-based detection of the single-channel bursts may, in principle, be applied to any MEA setup. However, the correct implementation of a particular algorithm relies on the general understanding of the origins of the recorded data. For example, with the standard commercial MEAs, it is impossible to infer directly whether the recorded single-channel bursting originates from one bursting neuron or from several active cells near the electrode [11]. Only the sorting of the in-burst spikes may provide an answer in this case.

3.2.2 Network burst detection methods

NB activity is commonly linked to network maturation, with its peak generally aligning with the period of highest synaptic density within a network [11, 93, 94, 104–106]. Numerous established methods for detecting NBs have been proposed [93, 94, 106–109]. Based on their underlying detection principles, these methods can be categorized into ISI-based approaches [93, 94, 108], population firing rate (PFR)-based approaches [106, 107], and hybrid methods [109]. PFR integrates the time-varying binned spike firing rates from all involved MEA electrodes into a common firing rate signal.

Some of the listed techniques extrapolate single-channel burst detection results to the network level [93, 108]. Pasquale and colleagues [93] extended their single-channel burst detection logic to the NB detection level. After the single-channel bursts were detected in the signals of all involved electrodes, they pooled the timings of all these bursts into the cumulative burst event train. The logarithmic inter-burst interval (IBI) histogram was constructed using that. The histogram was used to identify the IBI threshold and to detect NBs within the cumulative burst event train. Finally, the criterion for minimum participating channels was applied to the detected NBs to preserve only the NBs with contributions from multiple electrodes [93]. Similarly, Vällki and colleagues [108] identified the single-channel bursts among MEA electrodes using the algorithm by Kapucu and colleagues [101] described above and evaluated their synchronization. The total NB signal was obtained by summing bursting behaviors over channels, which was then thresholded in accordance with the minimum participating channels criterion [108].

Others treat the NB detection as an independent process [94, 106, 107, 109]. For example, Bakkum and colleagues [94] developed a logarithmic ISI_N histogram-based method. First, the spike trains from all involved MEA electrodes were combined and sorted into a joint spike train to calculate the logarithmic ISI_N histogram of ISIs between N consecutive spikes. Again, the histogram exhibiting two separate peaks was assumed to represent two different firing regimes: in this case, the spikes that belong to NBs and the spikes outside the network firing regime, respectively. Therefore, the threshold for detecting NBs was selected at the minimum point between these peaks in the histogram [94]. Hedrich and colleagues [107] presented the PFR average-based method. The spike trains from all involved MEA electrodes were summed into 5 ms time bins to calculate the

PFR, which was then smoothed by a Gaussian kernel with a 100 ms SD. The intervals with a PFR exceeding the slowly varying 1-second PFR average were selected as NB candidates. The candidates were accepted if they fit three criteria: their peak firing rate exceeded three SDs of the PFR, their peak firing rate exceeded 10% of the average of the top five peaks, and at least three electrodes contributed. At the last step, neighboring NBs with IBIs less than 200 ms were merged [107].

Matsuda and colleagues [109] combined ISI-thresholding with a form of PFR-thresholding. In their four-step method, the spike trains from all electrodes were pooled into a joint spike train, where the spikes separated by ISIs less than 4 ms were assigned to NB candidates. NBs with less than 20 spikes were discarded. Neighboring NBs with IBI less than 60 ms were merged. At the last step, only the NBs that accommodate more than 3000 spikes were considered valid [109]. The method involved a number of fixed parameters, which usually limit the adaptivity of the approach to highly variable MEA data.

The detection of NB activity provides an advanced measure of neuronal network functioning. **Figure 11** shows the gradual drop in NB counts in a kainic acid (KA)-exposed culture compared to its baseline activity. The data was acquired with an Axion CytoView MEA 48 plate, which has 16 electrodes per well.

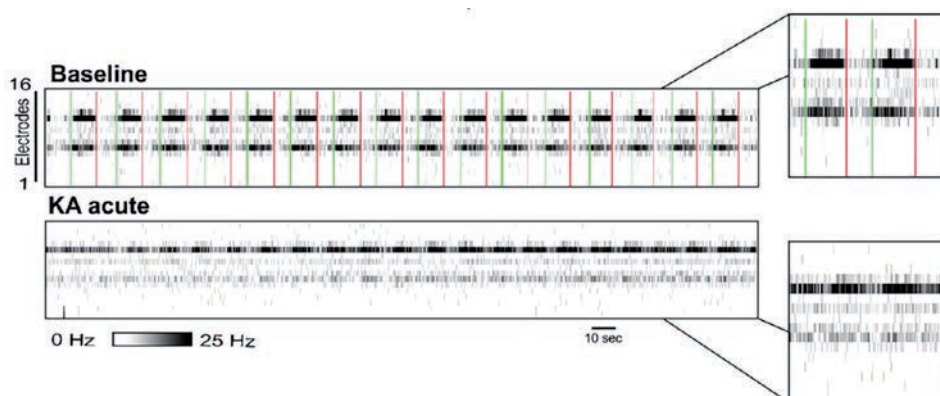


Figure 11. Detected NB counts at baseline and after KA exposure. The raster plots of KA-induced perturbation of NB activity. The upper plot shows baseline spiking activity in a well with detected NBs. The green and red vertical lines represent the starting and ending times of NBs, respectively. The lower plot demonstrates disruption of the NBs into spread tonic firing. No NBs were detected. Modified from Ref. [110].

3.2.3 Detection of network bursts in multi-compartment setups

Application of the NB detection methods to multi-compartment MEA setups possesses additional challenges. While the listed methods are well-documented for standard MEA setups, they are not directly applicable to the data acquired with multi-compartment MEAs. The architecture of multi-compartment MEAs involves multiple potential synchronization levels, requiring specialized analysis tools to fully exploit the data they generate. Without such tailored analysis techniques, the advantages of multi-compartment MEAs over conventional single-network setups may be lost. Overall, the signal analysis for such systems should account for both the spatial separation and the interactions between the interconnected networks to ensure accurate assessment [41].

Previous studies involving multi-compartment MEA data analysis often considered synchronous bursting as a key feature [30–33, 41]. Typically, these studies first identified local NBs within individual compartments using conventional methods, followed by evaluating the temporal alignment of these NBs within the entire compartmentalized network [30–32, 41]. For example, a recent work by Vinogradov and colleagues [41] defined three levels of synchronization within a three-compartment network of neurons (**Figure 1B–C**) recorded with the MEMO MEA plate. The initial detection of NBs was performed separately in each compartment using a custom algorithm developed based on a previously published method [94], which was extended with additional parameters to reflect local network synchrony [41]. The locally detected NBs were then evaluated for temporal alignment to capture their global, circuitry-level synchronization [41]. Beyond the localized NBs occurring within each of the three individual compartments of the device, two additional, more complex synchronization patterns were identified and analyzed. An intermediate circuitry burst (ICB) was characterized as simultaneous activity in two out of three compartments. Furthermore, a circuitry burst (CB) refers to synchronized bursting that encompasses the entire network across all three compartments [41]. The analysis demonstrated that a local network within a single compartment of a MEMO MEA plate could not only generate local NBs but also contribute to both ICBs and CBs observed at the global network levels, as illustrated by the network in compartment B (**Figure 12**).

3.3 Methods for functional connectivity analysis

Beyond analyzing synchronous bursting, various methods have been developed to assess functional connectivity and information propagation in neuronal cultures.

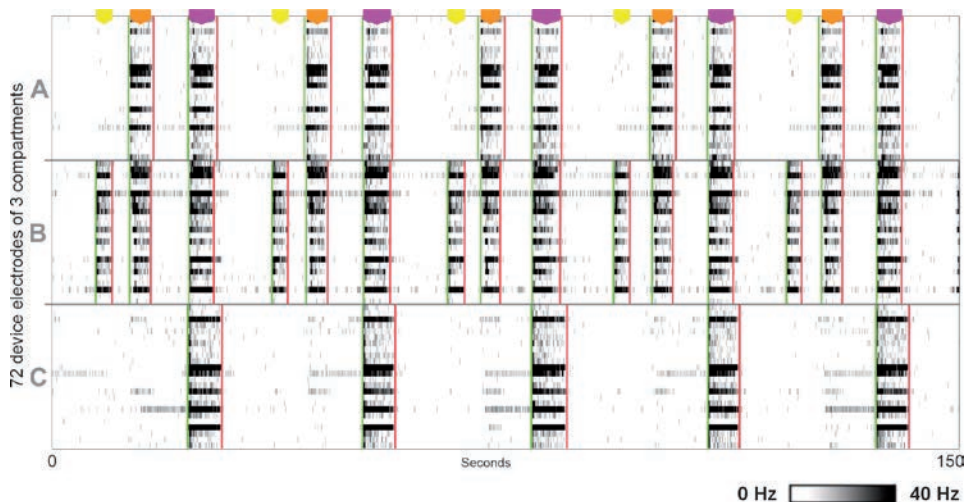


Figure 12. Examples of NB, ICB, and CB detection. The detections are visualized on a combined raster plot of spiking activity in three compartments within a 150-second segment. The data was recorded with the MEMO MEA plate. The green and red vertical lines represent the starting and ending times of detected local NBs, respectively. The purple polygons at the top depict the time intervals of detected CBs involving all three compartments; the orange polygons indicate ICBs with two-compartment synchronization; and the yellow polygons indicate local NBs in a single (B) compartment only. The length of the polygons corresponds to the duration of the corresponding synchronization patterns. Reproduced with modification from Ref. [41].

One such approach is the spike time tiling coefficient (STTC), which measures the proportion of spike co-occurrences between electrode pairs to determine pairwise correlations in their spiking activity [64, 111]. Similarly, functional connectivity maps can be constructed by analyzing correlations of simultaneous spikes across electrodes [35]. Cross-correlation-based methods have also been employed on compartment-wise firing rates within the network to infer the direction of activity propagation by evaluating the latency of peak correlations [30] or implemented straightforwardly between electrode pairs [31]. An additional alternative approach, known as correlated spectral entropy (CorSE), evaluates functional connectivity by analyzing temporal changes in the power spectra of raw MEA signals. The method involves segmenting the signals into windows, calculating spectral entropies for each window, and then determining connectivity through cross-correlation of the resulting entropy signals [63].

In a work by Hyvärinen and colleagues [112], the rat cortical and human pluripotent stem cell (hPSC)-derived cultures were compared in terms of functional connectivity development using both the CorSE and the STTC methods. The data was acquired with Axion CytoView MEA 12 plates with 64 electrodes per well. Both algorithms indicated earlier onset of functionally interconnected networks in the rodent data [112]. **Figure 13** provides the connectivity maps obtained using the CorSE

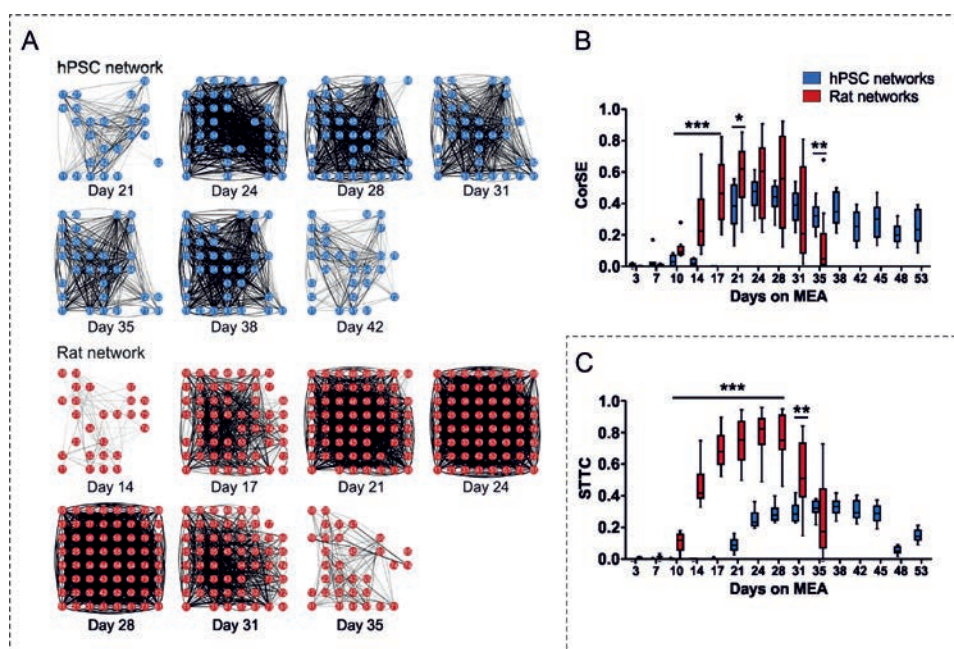


Figure 13.

*Functional connectivity development analysis. A. (Up) Functional connectivity maps from an hPSC-derived network over 10 minutes of recording between days 21 and 42 on MEA and (Down) from a rat cortical network between days 14 and 35 on MEA. Each map presents the electrode layout of the Axion CytoView MEA 12 plate as an 8×8 array with detected functional connections depicted as lines. An arbitrary CorSE value of 0.7 was used as the connectivity strength threshold for the line plotting. B. Functional connectivity development of the hPSC-derived and rat cortical networks described with the average CorSE values over time. C. The same as described by the average STTC values. The average values were calculated from all electrode pairs for each well. Both the hPSC-derived and rat cortical network data consist of $n = 12$ networks per group; data are shown as Tukey box plots. A Mann-Whitney U-test was performed to compare the two groups at each time point. Statistical significances are marked as * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$. Adapted from Ref. [112].*

algorithm and shows the dynamics of the average CorSE and STTC values for both culture types over the experimental timeline.

In addition to the correlation-based methods, various event-based and information theory-based approaches have been developed to estimate functional connectivity and causal interactions. Event Synchrony (ES), introduced by Quiroga and colleagues, quantifies synchronization based on the timing proximity of discrete events such as spikes. It assigns synchrony scores based on the coincidence of spike times within a local temporal window, providing an intuitive, symmetric measure of synchrony between two spike trains [113].

Mutual Information (MI) is a widely used metric that quantifies statistical dependencies between signals by comparing the joint probability distribution of events to their marginal distributions [65, 114]. MI has been particularly useful for assessing nonlinear relationships in spike train data [65, 115].

Extending the concept of MI, Transfer Entropy (TE) accounts for the directionality of information flow by incorporating time-lagged dependencies, thus estimating the extent to which past activity in one signal predicts future activity in another [116, 117]. Since TE is asymmetric and captures both linear and nonlinear interactions, making it suitable for analyzing directed functional connectivity [117].

Similarly, Granger causality [118], or its extended models for the frequency domain and multivariate signals [119], evaluates directional influences by relying on autoregressive modeling. It assumes that if a causal relationship exists, the past values of one time series will significantly contribute to the prediction of another. Several works in the field applied the Granger causality method for MEA data [31, 120, 121].

Each of these approaches complements classical correlation-based techniques and can enhance functional network mapping, particularly when directionality and information flow are of interest.

Functional connectivity analysis methods are particularly valuable in the context of multi-compartment MEA systems. Several illustrative examples are listed here. Le Feber and colleagues [30] applied cross-correlation analysis to compartment-specific firing rates within defined time bins, using the latency of peak correlations to infer the direction of activity propagation within a custom two-compartment MEA platform [30]. In another custom two-compartment MEA configuration, DeMarse and colleagues [31] estimated functional connectivity by computing cross-correlation and conditional Granger causality metrics from spike train data across electrode pairs [31]. Likewise, Van de Wijdeven and colleagues [35] generated functional connectivity maps by analyzing the correlations of simultaneous spikes detected across electrodes [35]. The CorSE method was used for evaluation of KA-exposure effects on functional connectivity in a three-compartment setup using previously mentioned MEMO MEA plates [41].

4. Conclusions

MEA technology remains a powerful tool for investigating the functional dynamics of neuronal networks. Precise analysis of electrophysiological signals is crucial for in vitro MEA studies focused on development, disease modeling, and drug testing. In this chapter, we introduced basic steps of neuronal data analysis and provided the corresponding analysis methods. This chapter has reviewed the core methods and current challenges in MEA signal analysis, including the accurate detection of spikes and the identification of bursts, both at single-channel and network levels, and functional connectivity mapping.

By presenting a comprehensive overview of established and emerging analysis strategies, this chapter aims to support researchers in making informed choices when designing experiments and interpreting MEA data. Integrating these analytical approaches enables a deeper understanding of neuronal network behavior and promotes the advancement of *in vitro* models in neuroscience research.

Acknowledgements

The work was supported by the Facility of Electrophysiological Measurements (Faculty of Medicine and Health Technology, Tampere University). The authors also thank Biocenter Finland for the support of the Electrophysiological Measurements facility. This work was supported by the Research Council of Finland and EU H2020-FETOpen (PRIME project).

Conflict of interest


The authors declare no conflict of interest.

Author details

Andrey Vinogradov, Laura Ylä-Outinen, Susanna Narkilahti* and Emre Kapucu
Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland

*Address all correspondence to: susanna.narkilahti@tuni.fi

IntechOpen

© 2025 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Cooper DC. The significance of action potential bursting in the brain reward circuit. *Neurochemistry International*. 2002;**41**:333-340. DOI: 10.1016/S0197-0186(02)00068-2
- [2] Emmenegger V, Obien MEJ, Franke F, Hierlemann A. Technologies to study action potential propagation with a focus on HD-MEAs. *Frontiers in Cellular Neuroscience*. 2019;**13**:159. DOI: 10.3389/fncel.2019.00159
- [3] Ylä-Outinen L, Tanskanen JMA, Kapucu FE, Hyysalo A, Hyttinen JAK, Narkilahti S. Advances in human stem cell-derived neuronal cell culturing and analysis. In: Chiappalone M, Pasquale V, Frega M, editors. *Vitro Neuronal Networks: From Culturing Methods to Neuro-Technological Applications*. Cham: Springer International Publishing; 2019. pp. 299-329. DOI: 10.1007/978-3-030-11135-9_13
- [4] Obien MEJ, Deligkaris K, Bullmann T, Bakkum DJ, Frey U. Revealing neuronal function through microelectrode array recordings. *Frontiers in Neuroscience*. 2015;**8**:423. DOI: 10.3389/fnins.2014.00423
- [5] Pelkonen A, Pistono C, Klecki P, Gómez-Budia M, Dougalis A, Konttinen H, et al. Functional characterization of human pluripotent stem cell-derived models of the brain with microelectrode arrays. *Cells*. 2022;**11**:106. DOI: 10.3390/cells11010106
- [6] Henze DA, Borhegyi Z, Csicsvari J, Mamiya A, Harris KD, Buzsáki G. Intracellular features predicted by extracellular recordings in the hippocampus *In vivo*. *Journal of Neurophysiology*. 2000;**84**:390-400. DOI: 10.1152/jn.2000.84.1.390
- [7] Heikkilä TJ, Ylä-Outinen L, Tanskanen JMA, Lappalainen RS, Skottman H, Suuronen R, et al. Human embryonic stem cell-derived neuronal cells form spontaneously active neuronal networks *in vitro*. *Experimental Neurology*. 2009;**218**:109-116. DOI: 10.1016/j.expneurol.2009.04.011
- [8] Thomas CA, Springer PA, Loeb GE, Berwald-Netter Y, Okun LM. A miniature microelectrode array to monitor the bioelectric activity of cultured cells. *Experimental Cell Research*. 1972;**74**:61-66. DOI: 10.1016/0014-4827(72)90481-8
- [9] Gross GW. Simultaneous single unit recording *in vitro* with a photoetched laser deinsulated gold multimicroelectrode surface. *IEEE Transactions on Biomedical Engineering*. 1979;**BME-26**:273-279. DOI: 10.1109/TBME.1979.326402
- [10] Pine J. Recording action potentials from cultured neurons with extracellular microcircuit electrodes. *Journal of Neuroscience Methods*. 1980;**2**:19-31. DOI: 10.1016/0165-0270(80)90042-4
- [11] Michela C, Valentina P, Monica F. *In Vitro Neuronal Networks: From Culturing Methods to Neuro-Technological Applications*. 1st ed. Cham: Springer International Publishing; 2019. DOI: 10.1007/978-3-030-11135-9
- [12] Shimba K, Sakai K, Iida S, Kotani K, Jimbo Y. Long-term developmental process of the human cortex revealed *In vitro* by axon-targeted recording using a microtunnel-augmented microelectrode Array. *IEEE Transactions on Biomedical Engineering*. 2019;**66**:2538-2545. DOI: 10.1109/TBME.2019.2891310

- [13] Frank CL, Brown JP, Wallace K, Mundy WR, Shafer TJ. From the cover: Developmental neurotoxicants disrupt activity in cortical networks on microelectrode arrays: Results of screening 86 compounds during neural network formation. *Toxicological Sciences*. 2017;**160**:121-135. DOI: 10.1093/toxsci/kfx169
- [14] Seidel D, Jahnke H-G, English B, Girard M, Robitzki AA. In vitro field potential monitoring on a multi-microelectrode array for the electrophysiological long-term screening of neural stem cell maturation. *Analyst*. 2017;**142**:1929-1937. DOI: 10.1039/C6AN02713J
- [15] Schulte S, Gries M, Christmann A, Schäfer K-H. Using multielectrode arrays to investigate neurodegenerative effects of the amyloid-beta peptide. *Bioelectronic Medicine*. 2021;**7**:15. DOI: 10.1186/s42234-021-00078-4
- [16] Zhang J, Laiwalla F, Kim JA, Urabe H, Van Wagenen R, Song Y-K, et al. Integrated device for optical stimulation and spatiotemporal electrical recording of neural activity in light-sensitized brain tissue. *Journal of Neural Engineering*. 2009;**6**:055007. DOI: 10.1088/1741-2560/6/5/055007
- [17] Kapucu FE, Vinogradov A, Hyvärinen T, Ylä-Outinen L, Narkilahti S. Comparative microelectrode array data of the functional development of hPSC-derived and rat neuronal networks. *Scientific Data*. 2022;**9**:120. DOI: 10.1038/s41597-022-01242-4
- [18] Pelkonen A, Mzezewa R, Sukki L, Ryyänen T, Kreutzer J, Hyvärinen T, et al. A modular brain-on-a-chip for modelling epileptic seizures with functionally connected human neuronal networks. *Biosensors and Bioelectronics*. 2020;**168**:112553. DOI: 10.1016/j.bios.2020.112553
- [19] Vinogradov A, Kapucu E, Narkilahti S. An analysis pipeline for signaling of human stem cell derived neurons in a microfluidics platform with an embedded microelectrode array. In: 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC); Oct 31 - Nov 4, 2021. Mexico: IEEE (Institute of Electrical and Electronics Engineers); 2021
- [20] Müller J, Ballini M, Livi P, Chen Y, Radivojevic M, Shadmani A, et al. High-resolution CMOS MEA platform to study neurons at subcellular, cellular, and network levels. *Lab on a Chip*. 2015;**15**:2767-2780. DOI: 10.1039/C5LC00133A
- [21] Radivojevic M, Franke F, Altermatt M, Müller J, Hierlemann A, Bakkum DJ. Tracking individual action potentials throughout mammalian axonal arbors. *eLife*. 2017;**6**:e30198. DOI: 10.7554/eLife.30198
- [22] Stett A, Egert U, Guenther E, Hofmann F, Meyer T, Nisch W, et al. Biological application of microelectrode arrays in drug discovery and basic research. *Analytical and Bioanalytical Chemistry*. 2003;**377**:486-495. DOI: 10.1007/s00216-003-2149-x
- [23] MEA2100-systems | www.multichannelsystems.com n.d. Available from: <https://www.multichannelsystems.com/products/mea2100-systems> [Accessed: April 22, 2025]
- [24] Maestro edge MEA and live-cell analysis system | Axion Biosystems n.d. Available from: <https://www.axionbiosystems.com/products/mea/maestro-edge> [Accessed: April 22, 2025]

- [25] MaxTwo HD multiwell microelectrode array. MaxWell Biosystems n.d. Available from: <https://www.mxwbio.com/products/maxtwo-multiwell-microelectrode-array/> [Accessed: April 22, 2025]
- [26] Kreutzer J, Ylä-Outinen L, Kärnä P, Kaarela T, Mikkonen J, Skottman H, et al. Structured PDMS chambers for enhanced human neuronal cell activity on MEA platforms. *Journal of Bionic Engineering*. 2012;**9**:1-10. DOI: 10.1016/S1672-6529(11)60091-7
- [27] Metsälä O, Kreutzer J, Högel H, Miikkulainen P, Kallio P, Jaakkola PM. Transportable system enabling multiple irradiation studies under simultaneous hypoxia in vitro. *Radiation Oncology*. 2018;**13**:220. DOI: 10.1186/s13014-018-1169-9
- [28] MEA2100-mini-systems | www.multichannelsystems.com n.d. Available from: <https://www.multichannelsystems.com/products/mea2100-mini-systems> [Accessed: April 22, 2025]
- [29] Taylor AM, Rhee SW, Tu CH, Cribbs DH, Cotman CW, Jeon NL. Microfluidic multicompartiment device for neuroscience research. *Langmuir*. 2003;**19**:1551-1556. DOI: 10.1021/la026417v
- [30] le Feber J, Postma W, de Weerd E, Weusthof M, Rutten WLC. Barbed channels enhance unidirectional connectivity between neuronal networks cultured on multi electrode arrays. *Frontiers in Neuroscience*. 2015;**9**:412. DOI: 10.3389/fnins.2015.00412
- [31] DeMarse TB, Pan L, Alagapan S, Brewer GJ, Wheeler BC. Feed-forward propagation of temporal and rate information between cortical populations during coherent activation in engineered In vitro networks. *Frontiers in Neural Circuits*. 2016;**10**:32. DOI: 10.3389/fncir.2016.00032
- [32] Gladkov A, Pigareva Y, Kutyna D, Kolpakov V, Bukatin A, Mukhina I, et al. Design of cultured neuron networks in vitro with predefined connectivity using asymmetric microfluidic channels. *Scientific Reports*. 2017;**7**:15625. DOI: 10.1038/s41598-017-15506-2
- [33] Forró C, Thompson-Steckel G, Weaver S, Weydert S, Ihle S, Dermutz H, et al. Modular microstructure design to build neuronal networks of defined functional connectivity. *Biosensors and Bioelectronics*. 2018;**122**:75-87. DOI: 10.1016/j.bios.2018.08.075
- [34] Kapucu FE, Tujula I, Kulta O, Sukki L, Rynänen T, Gram H, et al. Human tripartite cortical network model for temporal assessment of alpha-synuclein aggregation and propagation in Parkinson's disease. *Npj Parkinsons Diseases*. 2024;**10**:1-20. DOI: 10.1038/s41531-024-00750-x
- [35] van de Wijdeven R, Ramstad OH, Valderhaug VD, Köllensperger P, Sandvig A, Sandvig I, et al. A novel lab-on-chip platform enabling axotomy and neuromodulation in a multi-nodal network. *Biosensors and Bioelectronics*. 2019;**140**:111329. DOI: 10.1016/j.bios.2019.111329
- [36] Brofiga M, Pisano M, Tedesco M, Boccaccio A, Massobrio P. Functional inhibitory connections modulate the electrophysiological activity patterns of cortical-hippocampal ensembles. *Cerebral Cortex*. 2022;**32**:1866-1881. DOI: 10.1093/cercor/bhab318
- [37] Pigareva Y, Gladkov A, Kolpakov V, Bukatin A, Li S, Kazantsev VB, et al. Microfluidic Bi-layer platform to study functional interaction between Co-cultured neural networks with

unidirectional synaptic connectivity. *Micromachines*. 2023;**14**:835. DOI: 10.3390/mi14040835

[38] Duc P, Vignes M, Hugon G, Sebban A, Carnac G, Malyshev E, et al. Human neuromuscular junction on micro-structured microfluidic devices implemented with a custom micro electrode array (MEA). *Lab on a Chip*. 2021;**21**:4223-4236. DOI: 10.1039/D1LC00497B

[39] Takayama Y, Kushige H, Akagi Y, Suzuki Y, Kumagai Y, Kida YS. Selective induction of human autonomic neurons enables precise control of cardiomyocyte beating. *Scientific Reports*. 2020;**10**:9464. DOI: 10.1038/s41598-020-66303-3

[40] Gabriel-Segard T, Rontard J, Miny L, Dubuisson L, Batut A, Debis D, et al. Proof-of-concept human organ-on-chip study: First step of platform to assess neuro-immunological communication involved in inflammatory bowel diseases. *International Journal of Molecular Sciences*. 2023;**24**:10568. DOI: 10.3390/ijms241310568

[41] Vinogradov A, Kapucu EF, Narkilahti S. Exploring kainic acid-induced alterations in circular tripartite networks with advanced analysis tools. *eNeuro*. 2024;**11**(7):ENEURO.0035-24.2024. DOI: 10.1523/ENEURO.0035-24.2024

[42] Frega M, Tedesco M, Massobrio P, Pesce M, Martinoia S. Network dynamics of 3D engineered neuronal cultures: A new experimental model for in-vitro electrophysiology. *Scientific Reports*. 2014;**4**:5489. DOI: 10.1038/srep05489

[43] Smith I, Silveirinha V, Stein JL, de la Torre-Ubieta L, Farrimond JA, Williamson EM, et al. Human neural stem cell-derived cultures in

three-dimensional substrates form spontaneously functional neuronal networks. *Journal of Tissue Engineering and Regenerative Medicine*. 2017;**11**:1022-1033. DOI: 10.1002/term.2001

[44] Obien MEJ, Hierlemann A, Frey U. Factors affecting blind localization of a glass micropipette using a high-density microelectrode array. *Proceedings of IEEE Sensors*; Baltimore, MD, USA. 2013:1-4. DOI: 10.1109/ICSENS.2013.6688361

[45] Egert U, Heck D, Aertsen A. Two-dimensional monitoring of spiking networks in acute brain slices. *Experimental Brain Research*. 2002;**142**:268-274. DOI: 10.1007/s00221-001-0932-5

[46] Lieb F, Stark H-G, Thielemann C. A stationary wavelet transform and a time-frequency based spike detection algorithm for extracellular recorded data. *Journal of Neural Engineering*. 2017;**14**:036013. DOI: 10.1088/1741-2552/aa654b

[47] Pettersen KH, Einevoll GT. Amplitude variability and extracellular low-pass filtering of neuronal spikes. *Biophysical Journal*. 2008;**94**:784-802. DOI: 10.1529/biophysj.107.111179

[48] Nicholson C, Freeman JA. Theory of current source-density analysis and determination of conductivity tensor for anuran cerebellum. *Journal of Neurophysiology*. 1975;**38**:356-368. DOI: 10.1152/jn.1975.38.2.356

[49] Nunez PL, Srinivasan R. *Electric Fields of the Brain: The Neurophysics of EEG*. New York, USA: Oxford University Press; 2006. DOI: 10.1093/acprof:oso/9780195050387.001.0001

[50] Buzsáki G, Anastassiou CA, Koch C. The origin of extracellular fields

and currents-EEG, ECoG, LFP and spikes. *Nature Reviews Neuroscience*. 2012;**13**:407-420. DOI: 10.1038/nrn3241

[51] Mirfakhraei K, Horch K. Classification of action potentials in multiunit intrafascicular recordings using neural network pattern-recognition techniques. *IEEE Transactions on Biomedical Engineering*. 1994;**41**(1):89-91. DOI: 10.1109/10.277276

[52] Kim KH, Kim SJ. Neural spike sorting under nearly 0-dB signal-to-noise ratio using nonlinear energy operator and artificial neural-network classifier. *IEEE Transactions on Biomedical Engineering*. 2000;**47**(10):1406-1411. DOI: 10.1109/10.871415

[53] Jones IL, Livi P, Lewandowska MK, Fiscella M, Roscic B, Hierlemann A. The potential of microelectrode arrays and microelectronics for biomedical research and diagnostics. *Analytical and Bioanalytical Chemistry*. 2011;**399**:2313-2329. DOI: 10.1007/s00216-010-3968-1

[54] Choi JH, Jung HK, Kim T. A new action potential detector using the MTEO and its effects on spike sorting systems at low signal-to-noise ratios. *IEEE Transactions on Biomedical Engineering*. 2006;**53**(4):738-746. DOI: 10.1109/TBME.2006.870239

[55] Belitski A, Gretton A, Magri C, Murayama Y, Montemurro MA, Logothetis NK, et al. Low-frequency local field potentials and spikes in primary visual cortex convey independent visual information. *The Journal of Neuroscience*. 2008;**28**:5696-5709. DOI: 10.1523/JNEUROSCI.0009-08.2008

[56] Quian QR. What is the real shape of extracellular spikes? *Journal of Neuroscience Methods*.

2009;**177**:194-198. DOI: 10.1016/j.jneumeth.2008.09.033

[57] Lin J, Lai D, Wan Z, Feng L, Zhu J, Zhang J, et al. Representation and decoding of bilateral arm motor imagery using unilateral cerebral LFP signals. *Frontiers in Human Neuroscience*. 2023;**17**:1168017. DOI: 10.3389/fnhum.2023.1168017

[58] Heldman DA, Wang W, Chan SS, Moran DW. Local field potential spectral tuning in motor cortex during reaching. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. 2006;**14**:180-183. DOI: 10.1109/TNSRE.2006.875549

[59] Wang D, Zhang Q, Li Y, Wang Y, Zhu J, Zhang S, et al. Long-term decoding stability of local field potentials from silicon arrays in primate motor cortex during a 2D center out task. *Journal of Neural Engineering*. 2014;**11**:036009. DOI: 10.1088/1741-2560/11/3/036009

[60] Wang Y, Song Y, Dai Y, Li X, Xie J, Luo J, et al. The burst of electrophysiological signals in the suprachiasmatic nucleus of mouse during the arousal detected by microelectrode arrays. *Frontiers in Bioengineering and Biotechnology*. 2022;**10**:970726. DOI: 10.3389/fbioe.2022.970726

[61] Xu S, Deng Y, Luo J, Liu Y, He E, Yang Y, et al. A neural sensor with a nanocomposite interface for the study of spike characteristics of hippocampal neurons under learning training. *Biosensors (Basel)*. 2022;**12**:546. DOI: 10.3390/bios12070546

[62] Xu Z, Mo F, Yang G, Fan P, Wang Y, Lu B, et al. Grid cell remapping under three-dimensional object and social landmarks detected by implantable microelectrode arrays for the medial entorhinal cortex. *Microsystems*

- & Nanoengineering. 2022;**8**:1-11. DOI: 10.1038/s41378-022-00436-5
- [63] Kapucu FE, Vällki I, Mikkonen JE, Leone C, Lenk K, Tanskanen JMA, et al. Spectral entropy based neuronal network synchronization analysis based on microelectrode array measurements. *Frontiers in Computational Neuroscience*. 2016;**10**:112. DOI: 10.3389/fncom.2016.00112
- [64] Cutts CS, Eglén SJ. Detecting pairwise correlations in spike trains: An objective comparison of methods and application to the study of retinal waves. *The Journal of Neuroscience*. 2014;**34**:14288-14303. DOI: 10.1523/JNEUROSCI.2767-14.2014
- [65] Garofalo M, Nieuws T, Massobrio P, Martinoia S. Evaluation of the performance of information theory-based methods and cross-correlation to estimate the functional connectivity in cortical networks. *PLoS One*. 2009;**4**:e6482. DOI: 10.1371/journal.pone.0006482
- [66] Di Florio M, Iyer V, Rajhans A, Buccelli S, Chiappalone M. Model-based online implementation of spike detection algorithms for neuroengineering applications. In: *The 44th Annual International Conference of the IEEE Engineering in Medicine and Biology*; 11-15 July 2022. Glasgow, United Kingdom: IEEE (Institute of Electrical and Electronics Engineers); 2022. pp. 736-739. DOI: 10.1109/EMBC48229.2022.9871444
- [67] Gold C, Henze DA, Koch C, Buzsáki G. On the origin of the extracellular action potential waveform: A modeling study. *Journal of Neurophysiology*. 2006;**95**:3113-3128. DOI: 10.1152/jn.00979.2005
- [68] Someck S, Levi A, Sloin HE, Spivak L, Gattégno R, Stark E. Positive and biphasic extracellular waveforms correspond to return currents and axonal spikes. *Communications Biology*. 2023;**6**:1-15. DOI: 10.1038/s42003-023-05328-6
- [69] Quiroga RQ, Nadasdy Z, Ben-Shaul Y. Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. *Neural Computation*. 2004;**16**:1661-1687. DOI: 10.1162/089976604774201631
- [70] Wertenbroek R, Thoma Y, Mor FM, Grassi S, Heuschkel MO, Roux A, et al. SpikeOnChip: A custom embedded platform for neuronal activity recording and analysis. *IEEE Transactions on Biomedical Circuits and Systems*. 2021;**15**:743-755. DOI: 10.1109/TBCAS.2021.3097833
- [71] Liu M, Chen C, Gao K, Gao F, Qin C, Yuan Q, et al. Neuronal network-based biomimetic chip for long-term detection of olfactory dysfunction model in early-stage Alzheimer's disease. *Biosensors and Bioelectronics*. 2022;**216**:114619. DOI: 10.1016/j.bios.2022.114619
- [72] Chen R, Vakilha YS, Lassers SB, Tang WC, Brewer G. Hippocampal network axons respond to patterned theta burst stimulation with lower activity of initially higher spike train similarity from EC to DG and later similarity of axons from CA1 to EC. *Journal of Neural Engineering*. 2023;**20**:056004. DOI: 10.1088/1741-2552/acf68a
- [73] Yang Y, Deng Y, Xu S, Liu Y, Liang W, Zhang K, et al. PPy/SWCNTs-modified microelectrode Array for learning and memory model construction through electrical stimulation and detection of In vitro hippocampal neuronal network. *ACS Applied Bio Materials*. 2023;**6**:3414-3422. DOI: 10.1021/acsabm.3c00105

- [74] Zhang A, Mandeville ET, Xu L, Stary CM, Lo EH, Lieber CM. Ultraflexible endovascular probes for brain recording through micrometer-scale vasculature. *Science*. 2023;**381**:306-312. DOI: 10.1126/science.adh3916
- [75] Qin C, Yuan Q, Liu M, Zhuang L, Xu L, Wang P. Biohybrid tongue based on hypothalamic neuronal network-on-a-chip for real-time blood glucose sensing and assessment. *Biosensors and Bioelectronics*. 2024;**244**:115784. DOI: 10.1016/j.bios.2023.115784
- [76] Zhu Y, Jing L, Hu R, Mo F, Jia Q, Yang G, et al. High-throughput microelectrode arrays for precise functional localization of the globus pallidus internus. *Cyborg and Bionic Systems*. 2024;**5**:0123. DOI: 10.34133/cbsystems.0123
- [77] Maccione A, Gandolfo M, Massobrio P, Novellino A, Martinoia S, Chiappalone M. A novel algorithm for precise identification of spikes in extracellularly recorded neuronal signals. *Journal of Neuroscience Methods*. 2009;**177**:241-249. DOI: 10.1016/j.jneumeth.2008.09.026
- [78] Emery BA, Khanzada S, Hu X, Klütsch D, Amin H. Author spotlight: Advancing large-scale neural dynamics through HD-MEA technology. *Journal of Visualized Experiments (JoVE)*. 2024;**205**:e66473. DOI: 10.3791/66473
- [79] Choi JH, Kim T. Neural action potential detector using multi-resolution TEO. *Electronics Letters*. 2002;**38**:541-543. DOI: 10.1049/el:20020386
- [80] Mallat SG. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1989;**11**:674-693. DOI: 10.1109/34.192463
- [81] Daubechies I. *Ten Lectures on Wavelets*. Philadelphia, Pennsylvania, USA: SIAM (Society for Industrial and Applied Mathematics); 1992
- [82] Shalchyan V, Jensen W, Farina D. Spike detection and clustering with unsupervised wavelet optimization in extracellular neural recordings. *IEEE Transactions on Biomedical Engineering*. 2012;**59**:2576-2585. DOI: 10.1109/TBME.2012.2204991
- [83] Nabar N, Rajgopal K. A wavelet based Teager energy operator for spike detection in microelectrode array recordings. In: *TENCON 2009-2009 IEEE Region 10 Conference*; 23-26 January 2009. Singapore: IEEE (Institute of Electrical and Electronics Engineers); 2009. pp. 1-6. DOI: 10.1109/TENCON.2009.5395915
- [84] Semmlow JL, Griffel B. *Biosignal and Biomedical Image Processing*. 3rd ed. Boca Raton, Florida, USA: CRC Press; 2014
- [85] Shahid S, Walker J, Smith LS. A new spike detection algorithm for extracellular neural recordings. *IEEE Transactions on Biomedical Engineering*. 2010;**57**:853-866. DOI: 10.1109/TBME.2009.2026734
- [86] Natora M, Franke F, Obermayer K. Spike detection in extracellular recordings by hybrid blind beamforming. In: *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC'10*. Buenos Aires, Argentina: IEEE (Institute of Electrical and Electronics Engineers); 2010. DOI: 10.1109/IEMBS.2010.5626483
- [87] Lee JH, Carlson DE, Shokri Razaghi H, Yao W, Goetz GA, Hagen E, et al. YASS: Yet another spike sorter. In: *31st Conference on Neural Information Processing Systems (Advances in Neural*

- Information Processing Systems 30, NIPS 2017). Long Beach, California, USA: Curran Associates, Inc.; 2017
- [88] Rácz M, Liber C, Németh E, Fiáth R, Rokai J, Harmati I, et al. Spike detection and sorting with deep learning. *Journal of Neural Engineering*. 2020;**17**:016038. DOI: 10.1088/1741-2552/ab4896
- [89] Molen T v d, Lim M, Bartram J, Cheng Z, Robbins A, Parks DF, et al. RT-Sort: An action potential propagation-based algorithm for real time spike detection and sorting with millisecond latencies. *PLOS One*. 2024;**19**(12):e0312438. DOI: 10.1101/2024.04.08.588620
- [90] Abeles M, Goldstein MH. Multispike train analysis. *Proceedings of the IEEE*. 1977;**65**:762-773. DOI: 10.1109/PROC.1977.10559
- [91] Lewicki MS. A review of methods for spike sorting: The detection and classification of neural action potentials. *Network*. 1998;**9**:R53-R78
- [92] Chaure FJ, Rey HG, Quian QR. A novel and fully automatic spike-sorting implementation with variable number of features. *Journal of Neurophysiology*. 2018;**120**:1859-1871. DOI: 10.1152/jn.00339.2018
- [93] Pasquale V, Martinoia S, Chiappalone M. A self-adapting approach for the detection of bursts and network bursts in neuronal cultures. *Journal of Computational Neuroscience*. 2010;**29**:213-229. DOI: 10.1007/s10827-009-0175-1
- [94] Bakkum D, Radivojevic M, Frey U, Franke F, Hierlemann A, Takahashi H. Parameters for burst detection. *Frontiers in Computational Neuroscience*. 2014;**7**:193. DOI: 10.3389/fncom.2013.00193
- [95] Cocatre-Zilgien JH, Delcomyn F. Identification of bursts in spike trains. *Journal of Neuroscience Methods*. 1992;**41**:19-30. DOI: 10.1016/0165-0270(92)90120-3
- [96] Gourévitch B, Eggermont JJ. A nonparametric approach for detection of bursts in spike trains. *Journal of Neuroscience Methods*. 2007;**160**:349-358. DOI: 10.1016/j.jneumeth.2006.09.024
- [97] Lisman JE. Bursts as a unit of neural information: Making unreliable synapses reliable. *Trends in Neurosciences*. 1997;**20**:38-43. DOI: 10.1016/S0166-2236(96)10070-9
- [98] Chiappalone M, Novellino A, Vajda I, Vato A, Martinoia S, van Pelt J. Burst detection algorithms for the analysis of spatio-temporal patterns in cortical networks of neurons. *Neurocomputing*. 2005;**65-66**:653-662. DOI: 10.1016/j.neucom.2004.10.094
- [99] Cotterill E, Charlesworth P, Thomas CW, Paulsen O, Eglen SJ. A comparison of computational methods for detecting bursts in neuronal spike trains and their application to human stem cell-derived neuronal networks. *Journal of Neurophysiology*. 2016;**116**:306-321. DOI: 10.1152/jn.00093.2016
- [100] Madhavan R, Chao ZC, Potter SM. Plasticity of recurring spatiotemporal activity patterns in cortical networks. *Physical Biology*. 2007;**4**:181-193. DOI: 10.1088/1478-3975/4/3/005
- [101] Kapucu FE, Tanskanen J, Mikkonen J, Ylä-Outinen L, Narkilahti S, Hyttinen J. Burst analysis tool for developing neuronal networks exhibiting highly varying action potential dynamics. *Frontiers in Computational Neuroscience*. 2012;**6**:38. DOI: 10.3389/fncom.2012.00038

- [102] Legendy CR, Salcman M. Bursts and recurrences of bursts in the spike trains of spontaneously active striate cortex neurons. *Journal of Neurophysiology*. 1985;53:926-939. DOI: 10.1152/jn.1985.53.4.926
- [103] Ko D, Wilson CJ, Lobb CJ, Paladini CA. Detection of bursts and pauses in spike trains. *Journal of Neuroscience Methods*. 2012;211:145-158. DOI: 10.1016/j.jneumeth.2012.08.013
- [104] Kamioka H, Maeda E, Jimbo Y, Robinson HPC, Kawana A. Spontaneous periodic synchronized bursting during formation of mature patterns of connections in cortical cultures. *Neuroscience Letters*. 1996;206:109-112. DOI: 10.1016/S0304-3940(96)12448-4
- [105] Van Pelt J, Corner MA, Wolters PS, Rutten WLC, Ramakers GJA. Longterm stability and developmental changes in spontaneous network burst firing patterns in dissociated rat cerebral cortex cell cultures on multielectrode arrays. *Neuroscience Letters*. 2004;361:86-89. DOI: 10.1016/j.neulet.2003.12.062
- [106] Wagenaar DA, Pine J, Potter SM. An extremely rich repertoire of bursting patterns during the development of cortical cultures. *BMC Neuroscience*. 2006;7:11. DOI: 10.1186/1471-2202-7-11
- [107] Hedrich UBS, Liautard C, Kirschenbaum D, Pofahl M, Lavigne J, Liu Y, et al. Impaired action potential initiation in GABAergic interneurons causes hyperexcitable networks in an epileptic mouse model carrying a human NaV1.1 mutation. *The Journal of Neuroscience*. 2014;34:14874-14889. DOI: 10.1523/JNEUROSCI.0721-14.2014
- [108] Välkki IA, Lenk K, Mikkonen JE, Kapucu FE, Hyttinen JAK. Network-wide adaptive burst detection depicts neuronal activity with improved accuracy. *Frontiers in Computational Neuroscience*. 2017;11:40. DOI: 10.3389/fncom.2017.00040
- [109] Matsuda N, Odawara A, Katoh H, Okuyama N, Yokoi R, Suzuki I. Detection of synchronized burst firing in cultured human induced pluripotent stem cell-derived neurons using a 4-step method. *Biochemical and Biophysical Research Communications*. 2018;497:612-618. DOI: 10.1016/j.bbrc.2018.02.117
- [110] Mzezewa R, Lotila J, Kiiski H, Vinogradov A, Kapucu FE, Peltola J, et al. A kainic acid-induced seizure model in human pluripotent stem cell-derived cortical neurons for studying the role of IL-6 in the functional activity. *Stem Cell Research*. 2022;60:102665. DOI: 10.1016/j.scr.2022.102665
- [111] Gelfman S, Wang Q, Lu Y-F, Hall D, Bostick CD, Dhindsa R, et al. meaRtools: An R package for the analysis of neuronal networks recorded on microelectrode arrays. *PLoS Computational Biology*. 2018;14:e1006506. DOI: 10.1371/journal.pcbi.1006506
- [112] Hyvärinen T, Hyysalo A, Kapucu FE, Aarnos L, Vinogradov A, Eglen SJ, et al. Functional characterization of human pluripotent stem cell-derived cortical networks differentiated on laminin-521 substrate: Comparison to rat cortical cultures. *Scientific Reports*. 2019;9:17125. DOI: 10.1038/s41598-019-53647-8
- [113] Quian Quiroga R, Kreuz T, Grassberger P. Event synchronization: A simple and fast method to measure synchronicity and time delay patterns. *Physical Review E*. 2002;66:041904. DOI: 10.1103/PhysRevE.66.041904
- [114] Gray RM. *Entropy and Information Theory*. Boston, MA: Springer US; 2011. DOI: 10.1007/978-1-4419-7970-4

[115] Salazar RF, Dotson NM, Bressler SL, Gray CM. Content-specific fronto-parietal synchronization during visual working memory. *Science*. 2012;**338**(6110):1097-1100. DOI: 10.1126/science.1224000

[116] Schreiber T. Measuring information transfer. *Physical Review Letters*. 2000;**85**:461-464. DOI: 10.1103/PhysRevLett.85.461

[117] Ito S, Hansen ME, Heiland R, Lumsdaine A, Litke AM, Beggs JM. Extending transfer entropy improves identification of effective connectivity in a spiking cortical network model. *PLoS One*. 2011;**6**:e27431. DOI: 10.1371/journal.pone.0027431

[118] Granger CWJ. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*. 1969;**37**:424-438. DOI: 10.2307/1912791

[119] Baccalá LA, Sameshima K. Partial directed coherence: A new concept in neural structure determination. *Biological Cybernetics*. 2001;**84**:463-474. DOI: 10.1007/PL00007990

[120] Kispersky T, Gutierrez GJ, Marder E. Functional connectivity in a rhythmic inhibitory circuit using Granger causality. *Neural Systems & Circuits*. 2011;**1**:9. DOI: 10.1186/2042-1001-1-9

[121] Cadotte AJ, DeMarse TB, He P, Ding M. Causal measures of structure and plasticity in simulated and living neural networks. *PLoS One*. 2008;**3**:e3355. DOI: 10.1371/journal.pone.0003355

Section 2

Case Studies of Cognitive
Computational Models
in Neuroscience

Perspective Chapter: What Image AI is Close to Human Senses?

Hiroshi Omori

Abstract

The development of Image AI has been remarkable. In addition to Computer Vision Models (CVMs) that use supervised learning on large-scale data such as ImageNet-21 K, CVMs that use self-supervised learning and Contrastive Language Image Pre-training (CLIP) have been developed recently. We have been researching human environmental perception using photos. We measured the visual similarity by having many participants manually classify similar photos. We had three photo sets: 100 garden landscapes, 242 cityscapes, and 200 student life photos. We investigated how closely 26 types of pre-trained CVMs matched the human sense. We also used SoftMax regression to select a synthetic CVM from these CVMs that best correlated with the visual similarity. The optimal CVM combinations varied significantly across photo sets. For garden landscapes, which only have garden photos, semantic segmentation and self-supervised CVM were effective, while for student life photos, which have a wide variety of photos, supervised CVM was effective. For cityscapes, which have intermediate variations, self-supervised CVM was effective. MDS was used to examine in detail how the optimal CVM was like human perception and how it differed. For garden landscapes, humans and the CVM agreed on the judgment of garden size but differed on the judgment of garden style. In the case of cityscapes and student life photos, the recognition of patterns in the photos was roughly consistent between humans and the CVM, but the CVM made more detailed classifications. It was also suggested that some of the differences between the two stems from human representation. Although some differences remain, we found that by combining CVMs effectively, it is possible to construct a CVM that is quite close to human senses.

Keywords: visual similarity, computer vision model (CVM), self-supervised learning, softmax regression, multidimensional scaling (MDS), procrustes transformation, human representation

1. Introduction

We have conducted research [1–9] on collecting photos of a certain object and analyzing visual recognition of that object. The bottleneck in this research is the quantification and tagging of the photos. When researchers tag manually, problems with tagging bias arise. We have performed objective quantification using collective intelligence. That is, the similarity between photos was measured by the percentage of people

who perceived the two photos as “similar.” We called this similarity “visual similarity”. However, this method cannot measure the visual similarity for many photos, and an image analysis method that generates a similarity, like visual similarity, is needed.

In recent years, deep learning technology for images has developed rapidly and is now generally recognized as “Image AI”. By the way, Image AI includes a broader concept, such as image generation in addition to image recognition. In this article, we will deal with the relationship between computer image recognition and human image recognition, so we will not use the term Image AI from now on but will use the term Computer Vision Model (CVM).

For image classification CVMs, it has been common to perform supervised learning using large amounts of labeled image data such as ImageNet-21 K [10], which was an image set of 14,197,122 images labeled into 21,841 classes. CVMs with various architectures have been developed as supervised learning CVMs. In addition to Convolutional Neural Networks (CNNs), such as VGG16 [11] and ResNet [12], Vision Transformer (ViT) [13] has been released, which applies the self-attention mechanism [14], a standard in Natural Language Processing (NLP), to image processing.

However, there were problems with this method, such as the enormous cost of creating a large amount of labeled image data and the weak ability to recognize images that were not included in the training data. For this reason, a learning method that does not require labeling was required.

In fact, in the NLP, Self-Supervised Learning, which does not require labeling, such as *Large Language Models* (LLMs), for example, ChatGPT [15], is common for upstream tasks, and annotation, such as labeling, is performed in more specific downstream tasks. Its core technology is *Masked Language Modeling* (MLM), a learning method that masks any word in a sentence and predicts the masked word from the surrounding context. *Bidirectional Encoder Representation from Transformers* (BERT) [16] is a well known model that implements MLM.

Masked Image Modeling (MIM) [17] applies MLM to images. It masks parts of an image divided into patches and predicts them from the surrounding unmasked image. Examples of CVMs that implement MIM include *Bidirectional Encoder representation from Image Transformers* (BEiT) [18]. In addition, *Masked AutoEncoder* (MAE) [19] is a method like MIM. While MIM predicts pixels and features of a masked image, MAE predicts the features of the entire original image from an unmasked image. ConvNeXtV2 [20] implements MAE.

Self-supervised learning for images using data augmentation is also well known. *Simple Framework for Contrastive Learning of Visual Representations* (SimCLR) [21] performs data augmentation such as rotation, color conversion, and random cropping cutting out a part of the image. It performs contrastive learning to train an image encoder so that embedding vectors of positive example pairs that have been transformed from the same image are close to each other, and embedding vectors of negative example pairs that are combinations of different images are far apart.

Swapping Assignments between multiple views of the same image (SwAV) [22] is an online clustering method to train CNNs without annotations. By mining clusters invariant to data augmentations, SwAV learns semantic representations. *Self-supervised* (SEER) [23] is constructed by applying SwAV to large-scale data. SEER used 1B random, public, and non-EU (GDPR circumvention) Instagram images as pre-training. Recently, a super-large model (RG-10B) [24] with 10B parameters has been released.

Contrastive Language Image Pre-training (CLIP) [25] is also a learning method that does not require labeling. CLIP is based on the idea that words used in the text

surrounding an image on the Internet are likely to be highly related to the image, and so it simultaneously learns the embedding of images and words based on the affinity between images and words. Large-scale Artificial Intelligence Open Network (LAION)-5B [26] is a huge dataset consisting of approximately 5.85 billion image-text pairs processed by CLIP. LAION-5B is also known for being used to train the diffusion model [27], an image generation AI. Image encoders of LAION2B-en (English language consisting of 2.3B pairs) can be used for image classification.

Inspired by the rapid development of image deep learning in recent years, we began to study how well visual similarity can be approximated by pre-trained CVMs [28–32]. By combining semantic segmentation CNNs pre-trained on ADE20K [33], which consists of 20 K images colored with 150 semantics, such as sky, grass, building, and so on, CNNs and ViT pre-trained on ImageNet-21 K, and the image encoders of LAION-400 M with CLIP processing [24], we found that the synthesized CVM could well approximate the visual similarity for multiple photo sets [30–32]. However, in the past few years, many CVMs have been released one after another, including CVMs that have undergone self-supervised learning in Keras and Hugging Face. A total of 26 CVMs are covered in this report. Details are provided in Chapter 3. We believe that this report covers most of the current state of the art in CVMs for image recognition and image embedding.

We also devised SoftMax regression [34] to correctly find a CVM that mimics visual similarity from many CVMs and explain it in Chapter 4. The regression coefficients of the SoftMax regression give the contribution ratio of each CVM when the visual similarity is approximated by the similarity from CVMs. By combining CVMs using these regression coefficients, we can synthesize a CVM that is closest to human sense. In Chapter 5, we present the optimal combination of CVMs for the three photo sets discussed in this paper. For the synthesized CVM that is closest to human perception, we compared human and CVM MDS (Multi-Dimensional Scaling) and added a detailed discussion of the similarities and differences between the two in Chapter 6.

There were no reports that examined in detail the relations between human senses and the latest CVM senses. Since one of the goals of image AI is thought to be to imitate humans, this report is believed to have made a significant contribution to that goal.

2. Photo sets used and visual similarity measurement

2.1 Garden landscapes

In 2000 and 2001, we visited 184 randomly selected houses using detailed house maps and other sources and took numerous photographs of the gardens [1–5]. Within photos of each garden, we chose the one that seemed to best represent the overall atmosphere. Among these photos, 100 garden landscapes (40 in Aoba-Ku, Yokohama, Japan, 18 in Reading, UK, 7 in Edinburgh, UK, 15 in Turin, Italy, and 20 in Freising, Germany) were selected in consideration of landscape patterns and regionals. Twenty students measured the visual similarity between photos.

2.2 Kawagoe cityscapes

In 2010 and 2011, we asked local university students in Kawagoe, Saitama prefecture, Japan, to take “favorite” photos in the city [8]. 242 photos were collected from 172 students. Of these, 77 students measured the inter-photo visual similarity.

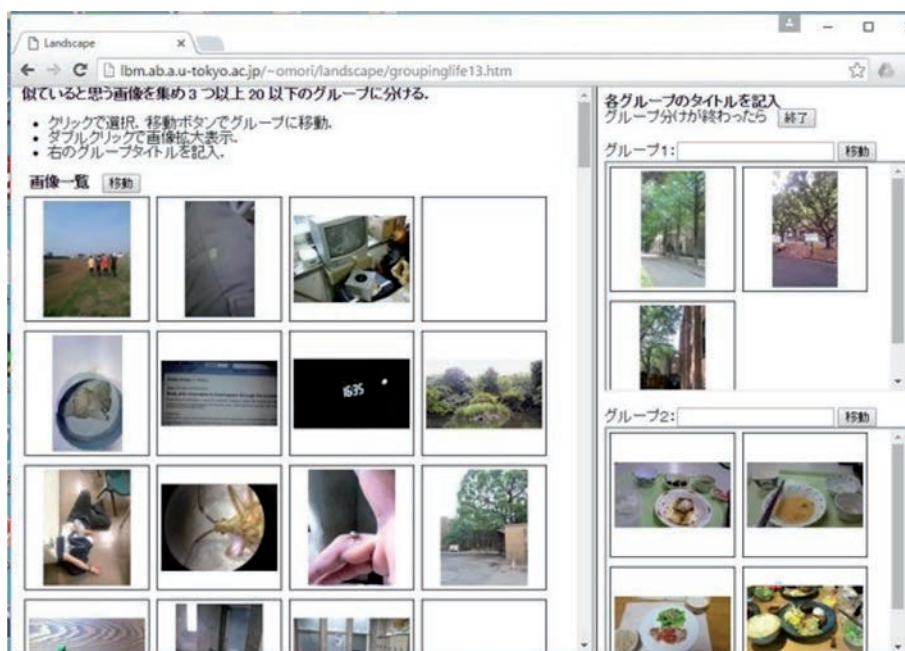


Figure 1.
The website to classify photos with similar ones [35].

	Photo1	Photo2	Photo3	Photo4
Photo1	36	1	14	2
Photo2	1	36	2	28
Photo3	14	2	36	3
Photo4	2	28	3	36



Figure 2.
An example of the similarity matrix.

2.3 Student life photos

We asked students of The University of Tokyo to take photos that they found “interesting” in their lives from 2012 to 2014 [9]. 200 photos were collected from 71 students. Of these, 36 students measured the visual similarity.

2.4 Differences among photo sets

In garden landscapes, there were only photos of the garden, with very little variation. In Kawagoe cityscapes, there was a variety of photos, such as distant views, streets, buildings, and so on. In student life photos, there were various types of photos, such as fields, buildings, indoors, foods, and so on. It is also the purpose of this paper to investigate the effect of different resolutions in photo sets on the relationship between the visual and the CVM similarity.

2.5 How to measure visual similarity

For the garden landscape study, students were given 100 postcard-sized photos of garden landscapes and asked to group them by similar photos. For Kawagoe cityscapes and student life photos, we asked the students to visit the website [35] shown in **Figure 1** and group similar photos on the web. On the left side, there was a box listing the photos, and on the right side, there were 20 boxes arranged vertically for storing similar photos. Students could select a photo by clicking on it and then move between photos by clicking the button above the box.

As an example of the visual similarity, **Figure 2** shows the measurement results for four photos from a set of student life photos. Of the 36 students, only one student felt that photos 1 and 2 were similar, so the similarity was measured as $1/36$, and 14 students felt that photos 1 and 3 were similar, so the similarity was measured as $14/36$. In this way, the visual similarity was measured as the proportion of people who felt that two photos were similar. The overall visual similarity for a set of n photos is represented by a vector y consisting of $n(n - 1)/2$ elements.

3. CVMs used and calculation of similarity between photos

The following pre-trained CVM parameters could be taken from Keras [36], GluonCV [37], or Hugging Face [38].

3.1 Supervised learning CVMs

3.1.1 Image classification by CNNs

We used CVMs pre-trained on ImageNet-21 K and fine-tuned on ImageNet-1 K, which is a subset of ImageNet-21 K and consists of an image set of 1.2 million images labeled into 1000 classes. For a given photo, we used an image feature (embedding) vector, the previous stage of classifying into 1000 classes. The CVM similarity between photos was calculated using the cosine similarity between these image feature vectors.

EfficientNet [39] was a CVM that balanced and optimized all dimensions of depth, width, and resolution using a simple, yet highly effective compound coefficient. We used six CVMs and called “Effi”.

Big Transfer (BiT) [40] was a CNN based on the ResNet architecture [12]. It was a powerful CVM for transferring learning to downstream tasks even with small amounts of data. We used nine CVMs and called “BiT”.

ConvNeXt [41] was a CNN based on the ResNet architecture, which improved the convolutional layer, optimized the placement of normalization layers and activation

functions, and incorporated the concept of the Transformer model. We used 15 CVMs called “Conv”.

3.1.2 Image classification by ViTs

We used ViT models pre-trained on ImageNet-21 K and fine-tuned on ImageNet-1 K. The similarity between image embeddings was calculated by the cosine similarity.

Vanilla ViT [13] applied the standard architecture of the Transformer [14], which was famous in the NLP, to image processing. It divided the input image into fixed-size patches, treated them as tokens, and learned features using a self-attention mechanism. We used four CVMs and called “ViT”.

FlexiViT [42] randomized patch sizes during training to obtain weights that performed well across a wide range of patch sizes. We used two CVMs and called “Flexi”.

Swin (Shifted *windows*) Transformer [43] processed images hierarchically to learn features at different resolutions and shifted the window incrementally to efficiently integrate information across different patches. We used 12 CVMs and called “Swin”.

Data-efficient image Transformers (DeiT) [44] was a ViT that leveraged knowledge distillation from CNN. The teacher model (CNN) outputted a probability distribution that represented class confidence for the input data. The student model (ViT) was learned using both the teacher model output (soft labels) and the actual label (the hard label). We used 10 CVMs and called “DeiT”.

3.1.3 Image classification by hybrids of CNN and ViT

We used hybrid models pre-trained on ImageNet-21 K and fine-tuned on ImageNet-1 K. The similarity between image embeddings was calculated by the cosine similarity.

Convolution and self-Attention (CoAtNet) [45] combined convolutional layers and a self-attention mechanism to effectively extract local and global features. We used eight CVMs and called “CoAt”.

EdgeNeXt [46] was a hybrid CNN and ViT that featured a lightweight and efficient model design. It was adapted for real-time processing in edge environments such as mobile devices and sensors. We used four CVMs and called “Edge”.

Multi-axis Vision Transformer (MaxViT) [47] divided an image into patches and efficiently extracted local and global features. The MaxViT Block used a hybrid structure that combined a convolutional network and a Transformer Block. We used 15 CVMs and called “MViT”.

3.1.4 Image classification by other CVMs

MLP Mixer [48] did not use convolutional layers or attention mechanisms but consisted only of fully connected layers. It pre-trained on ImageNet-21 K and fine-tuned on ImageNet-1 K. The similarity between image embeddings was calculated by the cosine similarity. We used six CVMs and called “Mixer”.

3.1.5 Multiple classification

After removing classes with few intra-class photos from ImageNet-21 K, there were 12,358,688 photos tagged into 11,220 classes (ImageNet-21 K processed). These classes

were organized into 11 layers from most abstract to least abstract, and multiple classifications of the photos were performed [49]. We used four CVMs and called “I21KP”.

It outputted classification probability distributions for each layer. The distance between two photos was calculated by taking the weighted average of the Hellinger distance between the probability distributions for each layer, calculated by the number of classes in the layer, and the similarity was calculated as $1 - \text{distance}$. The reason for using the Hellinger distance was explained in Refs. [28, 29].

3.1.6 Semantic segmentation

We used six CNNs [50–52] pre-trained on ADE20K [33]. These CVMs were called “Seg”. In addition, we also used four CVMs from SegFormer [53], which combined a Transformer encoder and a lightweight multi-layer perceptron (MLP) decoder to perform simple, efficient, and powerful semantic segmentation. It called “SegFor”.

In semantic segmentation, a photo was classified into 150 classes (sky, wall, etc.) for each pixel, resulting in a pixel count distribution for 150 classes of the entire photo. The distance between the two photos was calculated using the Hellinger distance between pixel count distributions, and the similarity was calculated using the $1 - \text{distance}$.

3.2 Self-supervised learning CVMs

3.2.1 Image classification using MIM or MAE

CVMs were typically pre-trained on ImageNet-21 K and fine-tuned on ImageNet-1 K. In this case, only ImageNet-21 K images were used for pre-training, and pre-training was done using MIM or MAE without using label information. Subsequent fine-tuning on labeled data could achieve high classification performance optimized for a specific task. The similarity between image embeddings was calculated by the cosine similarity.

BEiT [18] used MIM to divide the input image into patches and mask parts of them. The unmasked image was mapped to discrete tokens by discrete Variational AutoEncoders (VAE). Using the Transformer, it learned the context of the whole image from the tokens in the unmasked parts, predicted the tokens in the masked parts from this context, and learned the visual features of the whole image. We used nine CVMs and called “BEiT”.

BEiT v2 [54] used vector-quantized knowledge distillation to train the tokenizer, which discretizes a continuous semantic space to compact codes. It introduced a patch aggregation strategy which associates discrete image patches to enhance global semantic representation. We used eight CVMs and called “BEiT2”.

Swin Transformer V2 [55] introduced MIM as self-supervised learning to learn contextual relationships within the image and introduced data augmentation to increase the training data. We used 10 CVMs and called “Swin2”.

ConvNeXtV2 [20] divided an image into small patches and randomly masked some of the patches. The unmasked patches were input to the encoder, which extracted features using a convolutional layer. The output from the encoder was input to the decoder, which predicted the masked patches. The decoder reconstructed the masked patches and learned to approximate the original image. We used 15 CVMs and called “Conv2”.

3.2.2 Image classification using data augmentation

We used eight SEER [23, 24] models including RG-10B and called “SEER”.

self-Distillation with NO labels (DINO) [56] used unlabeled data to simultaneously train a teacher model and a student model by self-supervised learning, where the student model learned to mimic the output of the teacher model. It was pre-trained using a large and diverse dataset (LVD-142 M). This dataset comprised 142 million curated images, sourced from various datasets such as ImageNet, Google Landmarks, Mapillary street-level sequences (SLS), and Food-101, among others. The teacher model was input with a global crop that covered most of the image, allowing it to learn broad context and capture overall features. The student model was input with local crops in addition to the global crop, allowing it to learn finer details. Additionally, the parameters of the teacher model were updated using the exponential moving average (EMA) of the parameters of the student model, which allowed the teacher model to slowly track the evolution of the student model. We used three CVMs and called “DINO”.

3.3 CLIP related CVMs

3.3.1 Image encoder of CLIP

We could use image encoders of CLIP [25] to obtain an image embedding. The similarity between image embeddings was calculated by the cosine similarity. We used six image encoders of CLIP-LAION-400 M and called “CLIP”. There were two types of image encoders: ViT-based and ConvNeXt-based in CLIP-LAION-2B. We used 19 ViT encoders and called “CLIPV”. We used eleven ConvNeXt encoders and called “CLIPC”.

3.3.2 CVMs using CLIP

Explore the limits of Visual representation at scAle (EVA) [57] masked a part of the input image and extracted features from the unmasked part using CLIP. It learned to predict CLIP features, not the pixels in the masked part. This allowed it to learn both geometric and semantic features. We used seven CVMs and called “EVA”.

EVA02 [58] improved accuracy by introducing improvements, which are used in NLP, to the architecture of the Transformer from the image side. In addition, by co-training MIM and CLIP, it achieved high accuracy with less parameter data than EVA-01. We used 11 CVMs and called “EVA02”.

META Transformer [59] was an integrated model that uses CLIP as a text encoder and learned semantic embeddings of text and images, allowing it to process information from different modalities, including images, text, audio, infrared, and tabular data. We used two CVMs pre-trained on LAION-2B and called “META”.

4. Comparison of visual similarity with CVM similarity

4.1 SoftMax regression

Each of the m pre-trained CVM_i generates an inter-photo similarity vector x_i of length $n(n - 1)/2$ for a set of n photos. Consider a regression model that finds the combination of vectors x_i with the highest correlation for a visual similarity vector y . When performing a normal multiple regression of y on x_i , the regression coefficients may become negative, making interpretation difficult. Therefore, we

performed regression with the constraint that the regression coefficients be non-negative such as,

$$y = b_0 + \exp(b_1) x_1 + \dots + \exp(b_m) x_m \quad (1)$$

Then, the final coefficient a_i of x_i is calculated using the SoftMax function, as

$$a_i = \exp(b_i) / \sum_j \exp(b_j) \quad (2)$$

In this paper, we used 26 types of pre-trained CVMs, including “Effi”, “BiT” and so on. Also, multiple CVMs were used for each CVM, for example, six for “Effi”. Therefore, we used the SoftMax regression to find the optimal combination of CVMs for each CVM as a preliminary round. Then, we used the SoftMax regression to find the optimal combination of CVMs across all 26 types of CVMs, as the final round.

4.2 Comparison by the two MDS coordinates

The correlation between the inter-photo visual similarity and the final CVM similarity indicates how similar both are overall. The MDS arrangement is considered a low-dimensional representation of the similarity structure. Here, 3D MDS was used.

Since the MDS represents the relative positional relationship between photos, it can be freely inverted and rotated. Therefore, to compare the 3D MDS based on visual similarity with the 3D MDS based on the CVM, it is necessary to align the MDS axis of the CVM with the MDS axis of the visual similarity using the procrustes transformation and to align the scales as well.

The overall similarity between the 3D MDS coordinates from the visual similarity and the transformed 3D MDS coordinates from the final CVM could be measured with the trace of the correlation matrix between both MDS coordinates. If the two match perfectly, the correlation matrix would be the identity matrix and the trace would be three.

Furthermore, by overlaying the MDS from visual similarity and the MDS from CVM, the difference in perception between the two for each photo could be quantified and visualized. In other words, photos with a large distance between the two coordinates are thought to give different cognition between humans and the CVM.

5. Results

5.1 Result of the preliminary round

As an example of the preliminary round, the results of EVA02 are shown in **Table 1**. There were 11 CVMs of EVA02 available from Hugging Face, so these were denoted from M1 to M11. Specifically, M1 was “tiny_patch14-224.mim_in22k”, M2 was “base_patch14_448.mim_in22k_ft-in1k”, M11 was “large_patch14_448.mim_m38m_ft_in22k-in1k”, and so on.

CVMs with large regression coefficients in Table I indicated valid CVMs and CVMs with regression coefficients of 0 were invalid CVMs. The synthesized CVM for “EVA02” was obtained by combining the individual CVMs with these regression

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	Corr.
Garden	0.150	0.000	0.000	0.000	0.243	0.024	0.161	0.173	0.088	0.162	0.000	0.557
Kawagoe	0.001	0.000	0.000	0.000	0.000	0.000	0.172	0.128	0.000	0.292	0.407	0.669
Student life	0.134	0.052	0.000	0.001	0.158	0.000	0.061	0.000	0.345	0.000	0.259	0.623

Table 1.
Regression coefficients for “EVAo2”. A preliminary round.

		Supervised learning											
		Effi	BiT	Conv	ViT	Flexi	Swin	DeiT	CoAt	Edge	MViT	Mixer	I21KP
	Garden	0.394	0.483	0.479	0.459	0.330	0.467	0.457	0.432	0.414	0.489	0.469	0.294
	Kawagoe	0.581	0.538	0.628	0.583	0.558	0.609	0.608	0.604	0.570	0.657	0.594	0.469
	Student life	0.549	0.567	0.677	0.662	0.680	0.711	0.699	0.728	0.677	0.692	0.714	0.621
		Self-supervised learning						CLIP related					
Seg	Segfor	BEiT	BEiT2	Swin2	Conv2	SEER	DINO	CLIP	CLIPV	CLIPC	EVA	EVA02	Meta
0.489	0.501	0.399	0.478	0.487	0.503	0.481	0.427	0.383	0.518	0.522	0.505	0.557	0.471
0.378	0.368	0.642	0.613	0.620	0.649	0.514	0.695	0.472	0.628	0.637	0.609	0.669	0.568
0.448	0.515	0.712	0.630	0.669	0.685	0.657	0.572	0.620	0.646	0.666	0.618	0.623	0.522

Table 2.
 Correlation between visual similarity and synthesized CVM similarity.

		Effi	BiT	Conv	ViT	Flexi	Swin	DeiT	CoAt	Edge	MViT	Mixer	I21KP
Garden		0.000	0.002	0.000	0.021	0.000	0.000	0.081	0.000	0.000	0.000	0.000	0.000
Kawagoe		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.155	0.000	0.000
Student life		0.000	0.000	0.000	0.000	0.048	0.004	0.006	0.225	0.000	0.000	0.114	0.216
Seg	Segfor	BEiT	BEiT2	Swin2	Conv2	SEER	DINO	CLIP	CLIPV	CLIPC	EVA	EVA02	Meta
0.165	0.152	0.000	0.000	0.000	0.000	0.000	0.029	0.000	0.000	0.000	0.000	0.545	0.004
0.000	0.000	0.000	0.000	0.000	0.134	0.000	0.485	0.000	0.000	0.000	0.000	0.225	0.000
0.001	0.030	0.098	0.000	0.000	0.000	0.066	0.000	0.191	0.000	0.000	0.000	0.000	0.000

Table 3.
Regression coefficients for all types of CVMs, the final round.

coefficients. “Corr.” meant the correlation between the similarity based on this synthesized CVM and the visual similarity. We found that effective CVM combinations varied significantly depending on the photo set. Furthermore, “EVA02” was highly correlated with Kawagoe cityscapes, indicating that the sense of “EVA02” to Kawagoe cityscapes was like that of humans.

Table 2 shows the correlation between the visual similarity and the similarity by the synthesized CVM for the 26 types of CVMs discussed in this paper. We found that the highly correlated CVMs differed depending on the photo set. For garden landscapes, the highest correlations were with “EVA02” (0.557), “CLIPC” (0.522), and “CLIPV” (0.518), in that order. For Kawagoe cityscapes, the highest correlations were with “DINO” (0.695), “EVA02” (0.669), and “Conv2” (0.649), in that order. For student life photos, the highest correlations were with “CoAt” (0.728), “Mixer” (0.714), and “BEiT” (0.712), in that order.

It is noted that since there were several CVMs that add pre-trained models from Keras in addition to models from Hugging Face, the results in **Table 2** differed slightly those in the previous report [34].

5.2 Results of the final round

Table 3 shows the regression coefficients for the final round. We found that the effective CVMs varied significantly depending on the photo set. For garden landscapes, there are three main effective CVMs, such as “Seg” (0.165) and “SegFor” (0.152) from semantic segmentation, and “EVA02” (0.545) from CLIP related. For Kawagoe cityscapes, there were four main CVMs, such as “MViT” (0.155) from supervised learning, “Conv2” (0.134), “DINO” (0.485) from self-supervised learning, and “EVA02” (0.225) from CLIP related. For student life photos, there were four main CVMs, such as “CoAt” (0.225), “Mixer” (0.114), “I21KP” (0.216) from supervised learning, and “CLIP” (0.191) from CLIP related.

For garden landscapes, semantic segmentation, and self-supervised learning (“EVA02” was a CVM that combines CLIP and MIM) were effective, while for Kawagoe landscapes, the self-supervised learning CVMs were almost always effective. In contrast, for student life photos, the self-supervised learning CVMs were not very effective, and the supervised learning CVMs were often effective.

Table 4 shows correlations between the visual similarity and the similarity by the final synthesized CVM and the traces of the correlation matrix between 3D MDS coordinates. **Table 4** also lists the results from the previous report [31]. This time, 26 types of CVM were used, whereas the previous report used only five types of CVM. They were “CNN”, “ViT”, “I21KP”, “Seg”, and “CLIP”. Of these, “CNNs” were CVMs pre-trained on ImageNet-1 K and were not used this time. The synthesized CVM was

	This result				Previous result [31]			
	Corr.	MDS1	MDS2	MDS3	Corr.	MDS1	MDS2	MDS3
Garden	0.619	0.926	1.677	2.197	0.526	0.917	1.544	1.894
Kawagoe	0.716	0.936	1.880	2.672	0.603	0.918	1.734	2.496
Student life	0.778	0.947	1.881	2.756	0.740	0.953	1.880	2.758

Table 4.
 Comparison of correlations and traces with previous results.

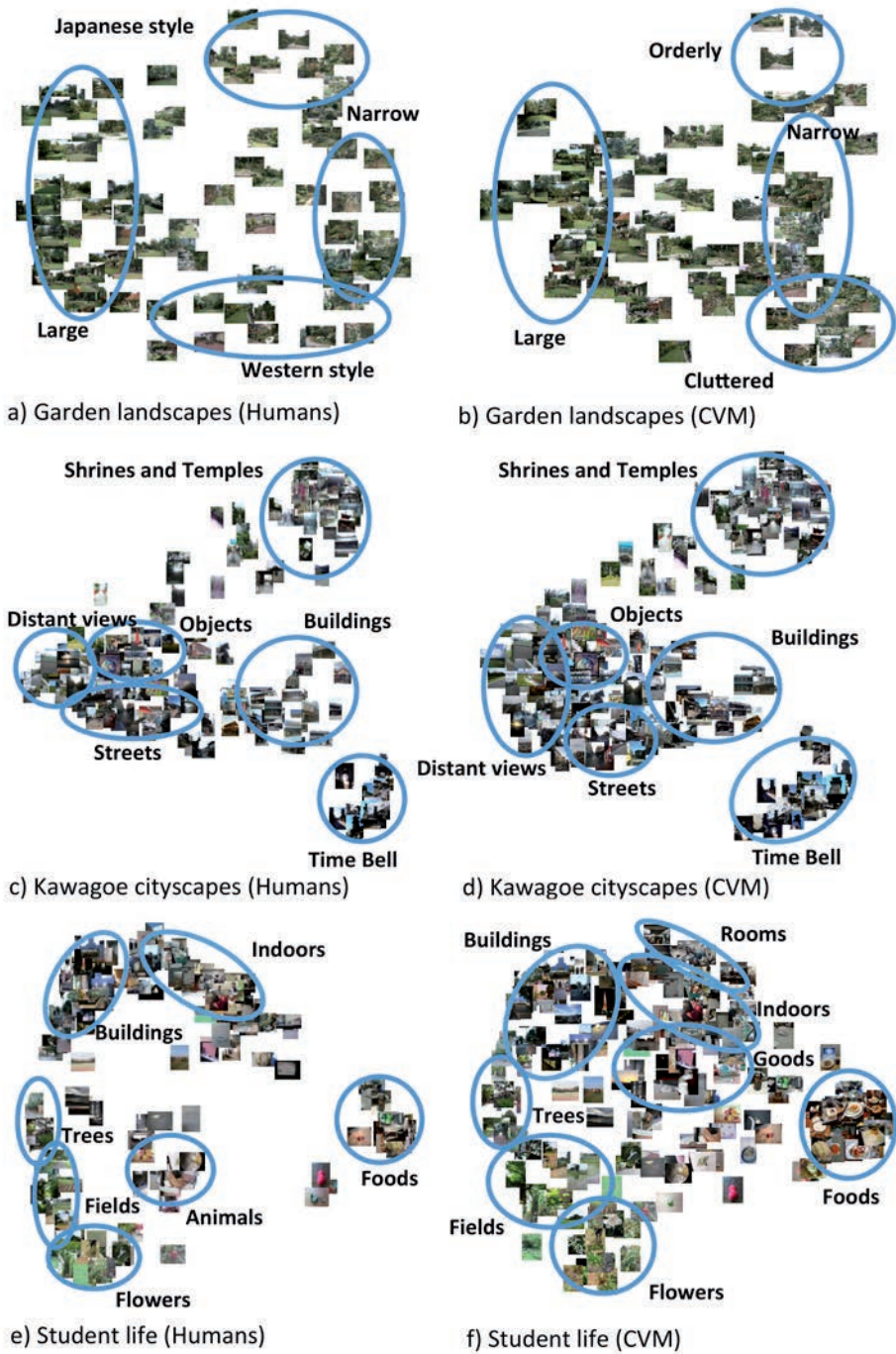


Figure 3. MDS placements by humans and CVM.

the average CVM similarity in the preliminary round, and the CVM with the highest correlation with visual similarity among the average, the top two, and the top three of each CVM in the final round was used as the final CVM.

Looking at **Table 4**, for garden landscapes and Kawagoe cityscapes, the correlation and trace values increased, and the results this time were better, but for student life photos there was not much difference from the previous results. In any case, these correlation coefficients represent the current state of similarity between CVM sensations and human sensations.

5.3 Comparison by MDS photo placements

Figure 3 shows 2D MDS photo placements for three photo sets. **Figure 3a** and **b** show MDS placements from humans and the final CVM for garden landscapes, respectively. **Figure 3c** and **d** show MDS placements from humans and the final CVM for Kawagoe cityscapes, respectively. **Figure 3e** and **f** show MDS placements from humans and the final CVM for student life photos, respectively.

Looking at **Figure 3a**, there were relatively large gardens covered with lawns on the left side and relatively narrow gardens on the right side. There were Japanese-style gardens with rounded pruning on the upper side and Western-style gardens with relatively many flowers or structures on the lower side. Thus, the first axis represented the size of the garden. The second axis seemed to represent the difference in the garden style: Japanese and Western. On the contrary, though the first axis of **Figure 3b** represented the same as that of **Figure 3a**, the second axis seemed different. That is, the upper side seemed to represent orderly and the lower side seemed cluttered gardens. This reflects the fact that, in the trace values in **Table 4**, while the MDS1 value for garden landscapes was high at 0.926, the MDS2 value was only 1.677 (an increment of 0.751), and the correlation between humans and CVM sense on the second axis was not so high.

Looking at **Figure 3c**, it looked like there were several clusters. There was a cluster of Kawagoe's symbol tower, "Time Bell", on the bottom right, "Shrines and Temples" on the upper right, "Distant views" on the left end, "Streets" on the bottom left, "Objects" above "Streets", and "Buildings" on the bottom center. Similar clusters were also seen in **Figure 3d**, and both placements were very similar. This was reflected by the 2D trace value for Kawagoe cityscapes in **Table 4**, which was 1.88, close to 2. However, looking more closely "Distant views" was split into two clusters, "Natural views" consisting of river scenery and "Distant views" consisting of cityscape views, and "Buildings" was also split into two clusters, "Wooden Bldg." and "Stone Bldg."; with the CVM providing more detailed classifications. This would be discussed in more detail in the next section. That is, the humans CVM MDS agglomerated, while the CVM MDS clusters appeared to be more spread out overall.

Looking at **Figure 3e**, it also looked like there were several clusters. There was a cluster of "Foods" on the far right, "Flowers" on the lower left, "Fields" on the middle-lower left, "Trees" on the middle left, "Buildings" on the upper left, "Indoors" on the upper center, and "Animals" near the center. On the other hand, though similar clusters were also seen in **Figure 3f**, the CVM MDS clusters appeared to be more spread out like Kawagoe cityscape case, the "Indoors" cluster spilt into three, that

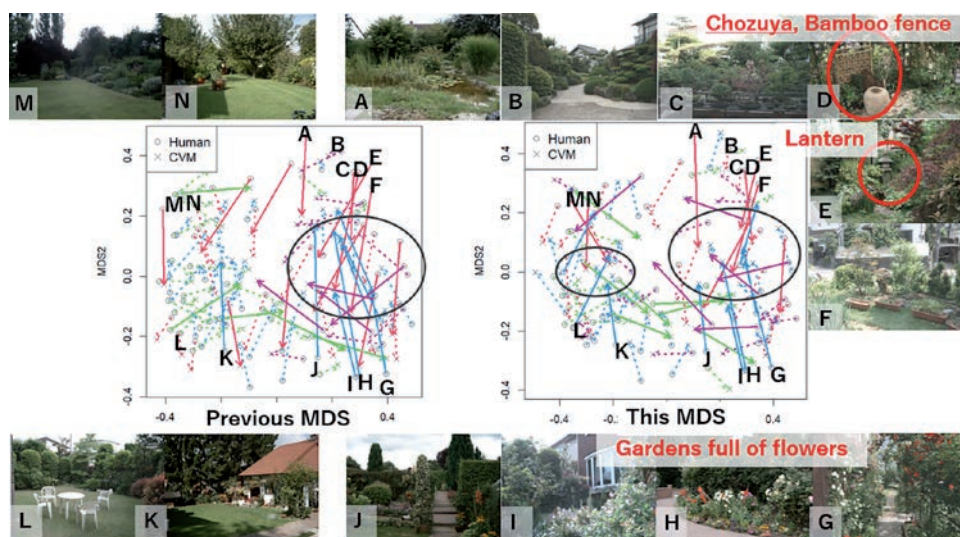


Figure 4. Superimposed MDS by humans and CVM for garden landscapes.

is, “Rooms”, “Indoors” and “Goods”, and the “Animals” cluster seemed disappear. Overall, the MDS placements of both were quite similar, which corresponded to the 2D trace value of 1.881 for student life photos in **Table 4**.

5.4 Photos that moved greatly MDS coordinates of both

Next, we examine how the impressions of individual photographs differ between humans and CVMs. This could be seen by superimposing the 2D human MDS and CVM MDS shown in **Figure 4** for garden landscapes, **Figure 5** for Kawagoe cityscapes and **Figure 6** for student life photos. The human MDS coordinates to photos were shown as “circles”, those of the CVM MDS as “crosses”.

Arrows were drawn from coordinates of the human MDS to those of the CVM MDS. Green indicated that the direction of the arrows was from 315° to 45, blue from 45 to 135, purple from 135 to 225, and red from 225 to 315. Among the coordinate movements in the MDS, the top 30 long arrows for garden landscapes and the top 50 for Kawagoe landscapes and student life are shown as solid lines, while the other short arrows are drawn as dotted lines.

5.4.1 In the case of garden landscapes

Looking at **Table 4**, in comparison with the previous results for garden landscapes, the current results showed that the optimal CVM is closer to human senses, but there is still a large difference. To examine this point, **Figure 4** shows each of the superimposed MDS results side by side. The left side shows the previous results, and the right side shows the current results. In addition, 14 photos of distinct gardens, from A to N, are placed around this figure. Six gardens (A to F) look Japanese style. In particular, the water basin (chozuya) and bamboo fence (D) and the lantern (E) are objects unique to Japanese gardens. On the other hand, four gardens (G to J) look Western style and are full of flowers.

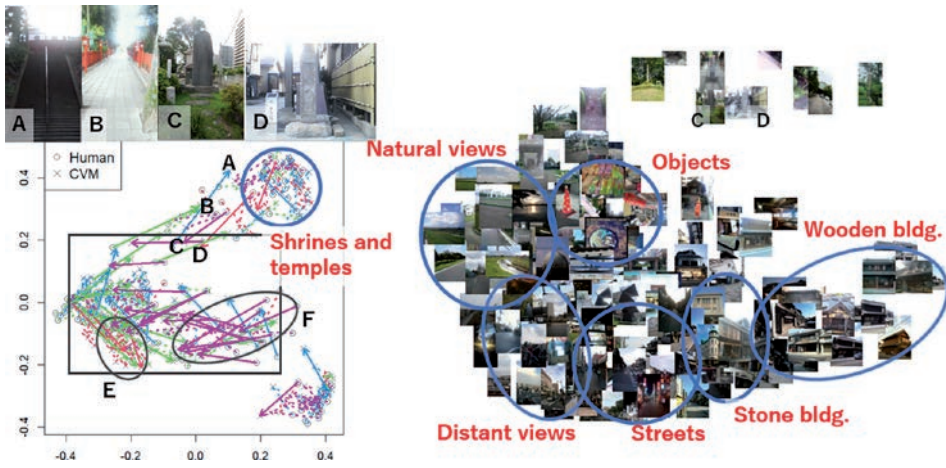


Figure 5.
 Superimposed MDS by humans and CVM for Kawagoe townscapes.

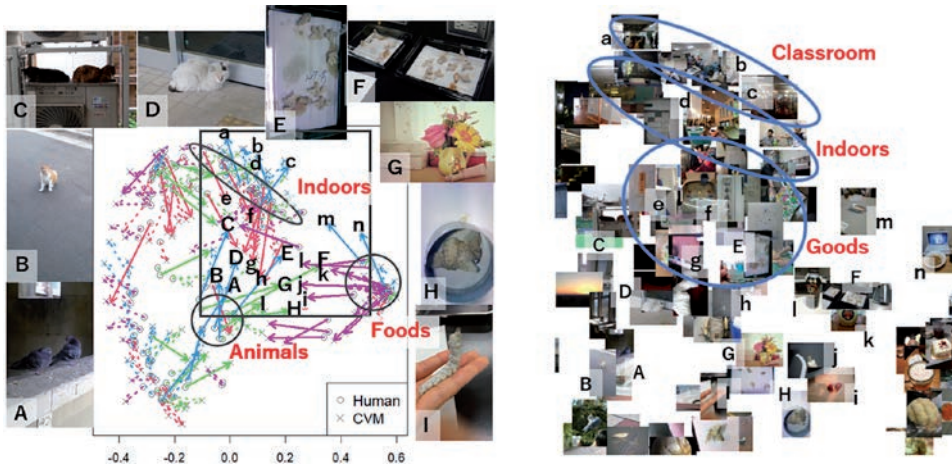


Figure 6.
 Superimposed MDS by humans and CVM for student life photos.

There were more vertical movements than horizontal movements in the photos that moved significantly in both cases. It might reflect that the CVM MDS was less like the humans MDS on the second axis than on the first axis. Overall, the previous results in the left figure have longer solid arrows, indicating a larger difference between human and CVM senses. There was a big difference between A, D to F and G to J to the human eyes, but when viewed from the CVM this time, there was no difference. Surprisingly, the previous CVM reversed the evaluation. As a result, in the CVM MDS, the expression Japanese-Western on the second axis disappeared and was changed to orderly-cluttered as shown in **Figure 3b**.

Next, let us look at the large garden on the left. M and N, which were located at the top in the human MDS, moved downward in the CVM MDS, while K and J, which were located at the bottom, moved upward, and there was almost no difference among M, N, K, and J. No artifacts could be seen in M and N, but artifacts were noticeable in K and J. This difference seemed large to the human eye, but it seemed

that the CVM did not perceive it as different. On the other hand, the previous CVM on the left side did not show any large difference between human and CVM senses. For this reason, as could be seen in **Figure 3b**, it was thought that the vertical variation was large in the narrow garden on the right, but the vertical variation shrank in the large garden on the left.

5.4.2 In the case of Kawagoe cityscapes

As shown in **Table 4**, the optimal CVM for Kawagoe cityscapes was quite close to human perception. This could be seen in **Figure 3c** and **3d**, where the cluster arrangements for humans and CVM were similar. However, clusters other than “The Time Bell” and “Shrines and Temples” were subtly different between humans and CVM, and this was shown in the superimposed MDS on the left in **Figure 5**, and in more detail in the enlarged view on the right.

In addition, photos A to D, which showed interesting movements in relation to the “Shrines and Temples” cluster, were placed at the top of the left figure. Photos of stone monuments such as A and B were located near the “Shrines and Temples” cluster in the human MDS, but in the CVM MDS, they shifted to the lower left. On the other hand, photos with compositions such as C and D shifted to the upper right in the CVM MDS, closer to the “Shrines and Temples” cluster.

What is striking here is that the photos that were included in the “Building” cluster in the human MDS all moved to the left and became attached to the “Streets” cluster in the CVM MDS then split into “Wooden” and “Stone buildings”. Furthermore, part of the “Distant views” cluster in the human MDS shifted to the lower right and formed a “Distant views” cluster adjacent to the “Streets” cluster, while the remainder became the “Natural views” cluster.

5.4.3 In the case of student life photos

As shown in **Table 4**, the optimal CVM for student life photos was quite close to human perception. This could be seen in **Figure 3e** and **3f**, where the cluster arrangements for humans and CVM were similar. However, the “Indoors” cluster in human MDS split into three in CVM MDS, and the “Animals” cluster disappeared, then it showed in detail in **Figure 6**.

In the “Animals” cluster, photos that showed large movements from humans to the CVM were shown from A to I around the left figure. Moving upward from the “Animals” cluster were photos of pigeons (A) and cats (B to D) and moving to the upper right were adult silkworms (E, F, and H), and a larval silkworm (I). These photos appeared to be quite like the human eye, but CVM seemed to see them as different from each other. The CVM seemed unable to understand the photos of silkworms lined up on pads (E and F), and thought they resembled a coffee can (I), so placed them near to it.

In the “Indoors” cluster, photos that showed large movements were shown from a to h. The “Classrooms” cluster in the right figure of **Figure 6** corresponding to photos from a to c could be seen seeping out of the human “Indoors” cluster to the upper left as shown in the left figure of **Figure 6**. Also, the “Goods” cluster corresponding to photos from e to h could be seen seeping out of the human “Indoors” cluster to the downright.

The “Foods” cluster in the human MDS of **Figure 3e** enlarged in the CVM MDS of **Figure 3f**, and this could be seen in more detail in **Figure 6**. The photos that moved largely from the “Foods” cluster were i to n in **Figure 6**. Looking at these photos, they seemed a little too subtle to be called “Foods”.

6. Discussion

Recent progress in deep image learning has been remarkable. We were interested in how similar the CVM similarity becomes to the visual similarity and investigated. We were surprised to see that many CVMs have been published even with a quick look at Keras and Hugging Face. As the number of candidate CVMs has increased, we developed SoftMax regression to easily find important CVMs. Even if new CVMs are announced in the future, we can quickly introduce their results to mimic human similarity perception by using SoftMax regression.

By the way, when we limited our search to CVMs that use huge training data, 26 types of CVMs were extracted at this time. We believe there are still other excellent CVMs, but this is almost the limit of our capabilities, so please forgive us. Many CVMs have been released for each individual CVM, and we have used a total of 214 CVMs in this research. We would like to express our gratitude once again to the researchers who developed these CVMs.

We were surprised to see that the effective CVMs varied greatly depending on the photo set. This is thought to reflect the fact that humans' sense of similarity differs depending on the photo. Therefore, it is necessary to carefully consider the subject of the research is what kind of photos. It may be necessary to randomly select about 100 photos from the ones you have, measure the visual similarity as shown in this paper, and select important CVMs using SoftMax regression.

This time, we used image encoders based on a huge amount of data called CLIP-LAION-2B. As shown in **Table 2**, "CLIPV" and "CLIPC" showed higher correlations with the visual similarity than the previous "CLIP", and a large effect was observed. However, the effectiveness measured by the regression coefficients was not observed. The reason is unclear, and this is a future issue. We also had high hopes for "SEER", but unfortunately the correlation with the visual similarity was not very high, and little effectiveness was observed.

Table 4 shows that the correlation increased significantly in garden landscapes and Kawagoe cityscapes, but not so much in student life photos. This was thought to be due to the big effect of the self-supervised learning CVMs, such as "EVA02" in garden landscapes, and "DINO" and "EVA02" in Kawagoe cityscapes. On the other hand, in the case of student life photos, sufficient correlation had already been achieved using the previous supervised learning CVMs, and the effect of the self-supervised learning CVMs was not so great.

The supervised learning CVM recognized the "objects" that appeared in a photo. Since student life photos were made up of a wide variety of images, such as foods, interiors, buildings, and landscapes, it was believed that the differences in the "objects" that appeared in the photos were the determining factor in the inter-photo visual similarity. For this reason, the supervised learning CVMs were effective.

On the other hand, since the visual similarity between photos of Kawagoe cityscapes was thought to be determined by the overall impression of the streetscape, rather than by the "objects", it was thought that the self-supervised learning CVMs would be effective. By the way, in the previous result [31], the clusters of the "Time Bell" and "Shrines and Temples" in the CVM MDS were more widely spread than the clusters in the human MDS, as shown in **Figure 7**. However, the clusters of the CVM MDS in **Figure 3d** were slightly more widely spread than the clusters in **Figure 3c**, with no large difference. The fact that the clusters by the CVM were not so widely spread was also thought to be due to the self-supervised learning CVM.

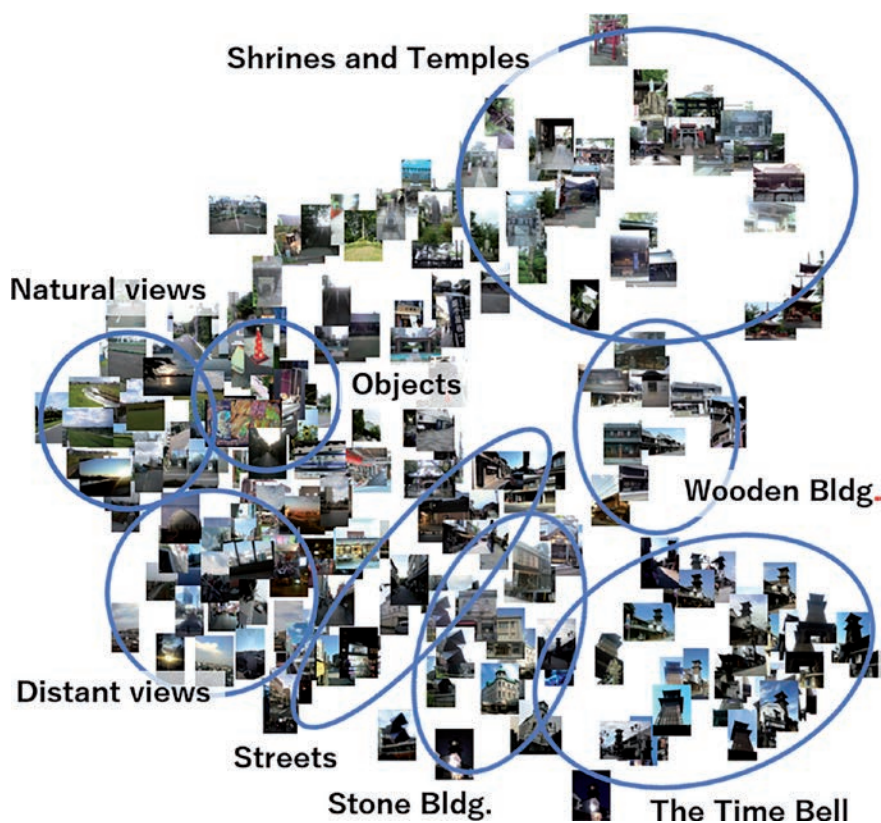


Figure 7.
Previous MDS for Kawagoe cityscapes.

In addition, the effectiveness of the semantic segmentation in garden landscapes was thought to be high since the size of the garden was related to the size of the semantic region called “grass”. In fact, when PSPNet [50] was used, which was the most effective among the six types of CVM used in “Seg” and the four types used in “SegFor”, the correlation between the CVM MDS1 and the percentage of “grass” pixels was 0.888, which revealed a relationship in which the more “grass” there was, the larger the garden appeared to be. Furthermore, the correlation between the percentage of “tree” pixels and the CVM MDS2 was 0.634, which revealed a tendency for gardens with fewer trees to appear cluttered.

Furthermore, looking at **Figure 4**, we can see that the human eye is quite different from that of the CVM. The human eye may sense a Japanese style from the objects in a garden, or a Western style if there are many flowers, but the CVM did not have this sense. This human sense can be called representation, and this representation may be unique to Japanese people. We also found that the presence or absence of artificial objects such as a table and chair at L in **Figure 4** did not have much impact on the CVM’s sense this time. This may be because CVMs that are effective for garden landscapes were not often selected from supervised learning CVMs, as shown in **Table 3**.

To support this, the left figure of **Figure 4** shows the results of a garden scene from our previous report [31], in which a CVM combining semantic segmentation and supervised learning was used. This result shows how photos that moved significantly from the human MDS to the CVM MDS in the current results moved in the previous

results. Looking at the right figure of **Figure 4**, the movement in the narrow garden on the right was like the previous time, but the movement in the large garden on the left was different. The movement of K was similar, but the movements of L, M, and N were smaller in the previous time. In other words, L, M, and N were not similar in the previous CVM, which is thought to indicate that the presence or absence of artificial objects influenced the sensation of CVM.

There are also significant differences in cluster formation between the human MDS and the CVM MDS. The clusters in **Figure 3c** and **e** are agglomerated, while the clusters in **Figure 3d** and **f** are more spread out. Humans classify photos based on some kind of criteria, which is thought to make it easier for clusters to form. On the other hand, the CVM does not have classification criteria, so it is thought that the separation of clusters can be ambiguous.

This is most evident in the “Indoors” cluster in **Figure 3e** and **Figure 6**. In student life photos, photos of food, buildings, plants, etc. can be relatively easily categorized according to the objects they contain, but indoor photos contain too many different objects, making it difficult to establish classification criteria. For this reason, we believe that the “Other” cluster was created to collect photos that could not be classified, which resulted in the formation of the “Indoors” cluster. On the other hand, CVM calculates the similarity between photos according to the objects they contain, so the “Indoors” cluster did not consolidate into one, and as a result it spread widely and appeared to be separated into the “Classrooms”, “Indoors”, and “Goods” clusters.

A similar phenomenon was observed in the “Buildings” and “Distant views” clusters in **Figure 3c** and **d**, and **Figure 5**. It seems that human senses separated “The Time Bell” and “Shrines and Temples” into clusters, separated single buildings into the “Buildings” cluster, and then gathered landscapes with the sky to form the “Distant Views” cluster. On the other hand, CVM judges the similarity of the photos themselves, so it separated “Buildings” into “Wooden” and “Stone” and approached “Streets” which shows multiple buildings. It also separated “Distant views” into “Natural views” which shows nature such as rivers, and “Distant Views” which shows cityscapes, and since many buildings are shown here, it was attached to “Streets”. Compared to the previous CVM, the individual clusters in **Figure 7** were spread out and the separation was unclear, but a cluster with the same composition as the cluster in **Figure 3d** was formed, so it was found that the scenery seen by CVMs was quite similar.

The “Foods” cluster in **Figure 3e** and **f** and **Figure 6** also seems to have this tendency. When humans look at many photos, they classify all photos that seem related to “Foods” as “Foods”, which results in the formation of a clear “Foods” cluster that is isolated from other clusters. On the other hand, CVM calculates the similarity between photos based on what is shown, so in **Figure 6** i to n cannot be said to be 100% “Foods”, and so they are thought to have been separated from the “Foods” cluster.

The same was true of the “Time Bell” cluster in **Figure 3c** and **d**. The “Time Bell” is the symbolic tower of Kawagoe City, and is clearly different from other Kawagoe cityscapes, so a clear and separate “Time Bell” cluster was formed. The shape of “Time Bell” is significantly different from other buildings, so a relatively clear and separate cluster was formed in the CVM MDS, but the cluster was somewhat broadened due to differences in the angle of the photo, etc. This cluster was much clearer and more isolated than the previous “Time Bell” cluster in **Figure 7**, clearly showing that the self-supervised learning CVM came much closer to human perception for photos like Kawagoe cityscapes.

Next, let us look at the disappearance of the “Animals” cluster in **Figure 3e**. Looking at **Figure 6**, the photos of A to I, which were in the original “Animals” cluster,

diffused from the right to the top in the CVM MDS, and the cluster disappeared. Looking at the photo arrangement in **Figure 6**, the photos around these photos are more diverse in the CVM MDS. Humans sense that living things and non-living things are different, but this is thought to be because the CVM does not have such a sense. This sense of living and non-living things may also be said to be a representation of human beings.

7. Conclusion


We used three photo sets with different characteristics to investigate how similar human senses are to the CVM senses and to investigate in what ways they differ. Overall, we found that human senses can be almost imitated by the CVM. Specifically, in garden landscapes with only garden photos, the percentage of “grass” and “tree” by semantic segmentation was effective in imitating human senses. In addition, it was shown that the self-supervised learning CVM could be used to approximate human senses quite closely in Kawagoe cityscapes. In student life photos, which consist of various types of photos, it was shown that the supervised learning CVM was close to the human senses. However, the taste of Japanese style and Western style in garden landscapes is a representation of Japanese people and cannot be imitated by the current CVM. In addition, the human representation of living and non-living things cannot be imitated by the current CVM.

Author details

Hiroshi Omori
Graduate School of Agricultural and Life Sciences and Faculty of Agriculture,
Tokyo, Japan

*Address all correspondence to: omori@ut-biomet.org

IntechOpen

© 2025 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Omori H, Adachi M, Hotta S, Yamashita M, Hanyu K, Kishino H, et al. Universality and specialty of the viewpoints to private gardens in Japan (in Japanese). *Japanese Journal of Biometrics*. 2001;**21**(2):53-72. DOI: 10.5691/jjb.21.53
- [2] Hanyu K, Yamashita M, Adachi M, Omori H, Hotta S, Kishino H, et al. Perception and inference of private garden: Influence of nationality and education. *MERA*. 2002;**15**:1-10. DOI: 10.20786/mera.8.1_1
- [3] Omori H, Hanyu K, Yamashita M, Watanabe S, Saio K. On the country image of private garden landscape (in Japanese). *Journal of the Japanese Institute of Landscape Architecture*. 2005;**68**(5):849-854. DOI: 10.5632/jila.68.849
- [4] Yamashita M, Hanyu K, Omori H. Comparison between English and Japanese in perception of private garden-views: By means of nationality discrimination. *Journal of Architecture and Planning (Transactions of AIJ)*. 2006;**60**7:93-99. DOI: 10.3130/aije.71.93_3
- [5] Hanyu K, Yamashita M, Omori H. Environmental inference on discrimination between Japanese and English private gardens. *MERA*. 2007;**10**(2):1-10. DOI: 10.20786/mera.10.2_1
- [6] Omori H, Hanyu K, Yamashita M, Saio K. Development of a landscape survey and analysis method using collective intelligence. *Landscape Research Japan Online*. 2011;**3**:39-42. DOI: 10.5632/jilaonline.3.39
- [7] Omori H, Hanyu K, Yamashita M. Slot map: Town map and collective intelligence of image about town integrating snapshot-location-texts. *Journal of Architecture and Planning (Transactions of AIJ)*. 2013;**78**(683):159-166. DOI: 10.3130/aija.78.159
- [8] Omori H, Hanyu K, Yamashita M. SALoT MAP as collective intelligence for the impression of Kawagoe landscape by local college students (in Japanese). *Journal of the Japanese Institute of Landscape Architecture*. 2014;**68**(5):849-854. DOI: 10.5632/jila.77.491
- [9] Omori H, Hanyu K, Yamashita M. SALoT MAP as collective intelligence on environment recognition: Visualization of the impression of student life (in Japanese). *International Journal of Affective Engineering*. 2017;**16**(1):171-179. DOI: 10.5057/jjske. TJSKE-D-16-00066
- [10] Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L. ImageNet: A large-scale hierarchical image database. *CVPR09*. 2009. pp. 248-255. DOI: 10.1109/CVPR.2009.5206848
- [11] Simonyan K, Zisserman A. Very deep convolution networks for large-scale image recognition. *arXiv [Internet]*. 2014:1-14. Available from: Available from: <https://arxiv.org/abs/1409.1556> [Accessed: February 28, 2025]
- [12] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *arXiv [Internet]*. 2015. Available from: <https://arxiv.org/abs/1512.03385>. [Accessed: February 28, 2025]
- [13] Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, et al. Image is worth 16x16 words: Transformers for image recognition at scale. *ICLR*. 2021:1-22. DOI: 10.48550/arXiv.2010.11929

- [14] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. arXiv [Internet]. Available from: <https://arxiv.org/abs/1706.03762>. 2017 [Accessed: February 28, 2025]
- [15] ChatGPT. Available from: <https://chatgpt.com/> [Accessed: January 25, 2025]
- [16] Devlin J, Chang MW, Lee K, Toutanova K. BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv [Internet]. 2018. Available from: <https://arxiv.org/abs/1810.04805> [Accessed: February 28, 2025]
- [17] Xie Z, Zhang Z, Cao Y, Lin Y, Bao J, Yao Z, et al. SimMIM: A simple framework for masked image modeling. arXiv [Internet]. 2021. Available from: <https://arxiv.org/abs/2111.09886> [Accessed: February 28, 2025]
- [18] Bao H, Dong L, Piao S, Wei F. BEiT: BERT pre-training of image transformers. arXiv [Internet]. 2021. Available from: <https://arxiv.org/abs/2106.08254> [Accessed: February 28, 2025]
- [19] He K, Chen X, Xie S, Li Y, Dollár P, Girshick R. Masked autoencoders are scalable vision learners. arXiv [Internet]. 2021. Available from: <https://arxiv.org/abs/2111.06377> [Accessed: February 28, 2025]
- [20] Woo S, Debnath S, Hu R, Chen X, Liu Z, Kweon IS, et al. ConvNeXt V2: Co-designing and scaling ConvNets with masked autoencoders. arXiv [Internet]. 2023. Available from: <https://arxiv.org/abs/2301.00808> [Accessed: February 28, 2025]
- [21] Chen T, Kornblith S, Norouzi M, Hinton G. A simple framework for contrastive learning of visual representations. arXiv [Internet]. 2020. Available from: <https://arxiv.org/abs/2002.05709> [Accessed: February 28, 2025]
- [22] Caron M, Misra I, Mairal J, Goyal P, Bojanowski P, Joulin A. Unsupervised learning of visual features by contrasting cluster assignments. arXiv [Internet]. 2020. Available from: <https://arxiv.org/abs/2006.09882> [Accessed: February 28, 2025]
- [23] Goyal P, Caron M, Lefaudeaux B, Xu M, Wang P, Pai V, et al. Self-supervised pretraining of visual features in the wild. arXiv [Internet]. 2021. Available from: <https://arxiv.org/abs/2103.01988> [Accessed: February 28, 2025]
- [24] Goyal P, Duval Q, Seessel I, Caron M, Misra I, Sagun L, et al. Vision models are more robust and fair when pretrained on uncurated images without supervision. arXiv [Internet]. 2022. Available from: <https://arxiv.org/abs/2202.08360> [Accessed: February 28, 2025]
- [25] Radford A, Kim JW, Hallacy C, Ramesh A, Goh G, Agarwal S, et al. Learning transferable visual models from natural language supervision. PMLR. 2021;**139**:8748-8763. DOI: 10.48550/arXiv.2103.00020
- [26] Schuhmann C, Beaumont R, Vencu R, Gordon C, Wightman R, Cherti M, et al. LAION-5B: An open large-scale dataset for training next generation image-text models. NIPS'22. 2022;**1833**:25278-25294. DOI: 10.48550/arXiv.2210.08402
- [27] Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models. NIPS'20. 2020;**574**:6840-6851. DOI: 10.48550/arXiv.2006.11239

- [28] Omori H, Hanyu K, Yamashita M, Shimada S. Prediction of the visual similarity between photos using several pre-trained CNNs. SCIS&ISIS. 2018:168-173. DOI: 10.1109/SCIS-ISIS.2018.00036
- [29] Omori H, Hanyu K, Shimada S. Prediction of the visual similarity between photos making use of deep learning: Image classification and semantic segmentation (in Japanese). International Journal of Affective Engineering. 2021;20(1):83-90. DOI: 10.5057/jjske.TJSKE-D-20-00028
- [30] Omori H, Hanyu K. Predict inter-photo visual similarity via pre-trained computer vision models. ACAI'22. 2022. pp. 1-9. DOI: 10.1145/3579654.3579769
- [31] Omori H, Hanyu K. How similar is image cognition in humans and computer visions? SCIS&ISIS. 2022. pp. 1-6. DOI: 10.1109/SCISISIS55246.2022.10002046
- [32] Omori H, Hanyu K. What is the difference between image cognition in humans and computer vision? (in Japanese). International Journal of Affective Engineering. 2024;23(2):107-117. DOI: 10.5057/jjske.TJSKE-D-23-00036
- [33] ADE20K database. 2025. Available from: <https://ade20k.csail.mit.edu/>
- [34] Omori H, Hanyu K. How close has computer vision become to human image cognition? SCIS&ISIS. 2024. pp. 1-6. DOI: 10.1109/SCISISIS61014.2024.10760175
- [35] Omori H. Grouping Cites. 2014. Available from: <http://lbn.ab.a.u-tokyo.ac.jp/~omori/groupinglife.html>
- [36] Keras Team. Keras: Deep Learning for Humans. Keras [Internet]. Available from: <https://keras.io/> [Accessed: January 25, 2025]
- [37] GluonCV Team. GluonCV: Computer Vision Toolkit. GluonCV [Internet]. Available from: <https://cv.gluon.ai/> [Accessed: January 25, 2025]
- [38] Hugging Face. Hugging Face: The AI Community Building the Future. Hugging Face [Internet]. Available from: <https://huggingface.co/> [Accessed: January 25, 2025]
- [39] Tan M, Le QV. EfficientNet: Rethinking model scaling for convolutional neural networks. arXiv [Internet]. 2019. Available from: <https://arxiv.org/abs/1905.11946> [Accessed: February 28, 2025]
- [40] Kolesnikov A, Beyer L, Zhai X, Puigcerver J, Yung J, Gelly S, et al. Big transfer (BiT): General visual representation learning. ECCV. 2020:491-507. DOI: 10.1007/978-3-030-58558-7_29
- [41] Liu Z, Mao H, Wu CY, Feichtenhofer C, Darrell T, Xie S. A ConvNet for the 2020s. CVPR. 2022:11976-11986. DOI: 10.48550/arXiv.2201.03545
- [42] Beyer L, Izmailov P, Kolesnikov A, Caron M, Kornblith S, Zhai X, et al. FlexiViT: One model for all patch sizes. CVPR. 2023:14496-14506. DOI: 10.48550/arXiv.2212.08013
- [43] Liu Z, Lin Y, Cao Y, Hu H, Wei Y, Zhang Z, et al. Swin transformer: Hierarchical vision transformer using shifted windows. ICCV. 2021:10012-10022. DOI: 10.48550/arXiv.2103.14030
- [44] Touvron H, Cord M, Douze M, Massa F, Sablayrolles A, Jegou H. Training data-efficient image transformers and distillation through attention. PMLR. 2021;139:0347-10357. DOI: 10.48550/arXiv.2012.12877

- [45] Dai Z, Liu H, Le QV, Tan M. CoAtNet: Marrying convolution and attention for all data sizes. arXiv [Internet]. 2021. Available from: <https://arxiv.org/abs/2106.04803> [Accessed: February 28, 2025]
- [46] Maaz M, Shaker A, Cholakkal H, Khan S, Zamir SW, Anwer RM, et al. EdgeNeXt: Efficiently amalgamated CNN-transformer architecture for Mobile vision applications. arXiv [Internet]. 2022. Available from: <https://arxiv.org/abs/2206.10589> [Accessed: February 28, 2025]
- [47] Tu Z, Talebi H, Zhang H, Yang F, Milanfar P, Bovik A, et al. MaxViT: Multi-axis vision transformer. arXiv [Internet]. 2022. Available from: <https://arxiv.org/abs/2204.01697> [Accessed: February 28, 2025]
- [48] Tolstikhin I, Houlsby N, Kolesnikov A, Beyer L, Zhai X, Unterthiner T, et al. MLP-mixer: An all-MLP architecture for vision. arXiv [Internet]. 2021. Available from: <https://arxiv.org/abs/2105.01601> [Accessed: February 28, 2025]
- [49] Ridnik T, Baruch EB, Noy A, Manor LZ. ImageNet-21K pretraining for the masses. NeurIPS. 2021. DOI: 10.48550/arXiv.2104.10972
- [50] Zhao H, Shi J, Qi X, Wang X, Jia J. Pyramid scene parsing network. CVPR. 2017:2881-2890. DOI: 10.48550/arXiv.1612.01105
- [51] Chen LC, Papandreou G, Schroffand FH, Adam H. Rethinking atrous convolution for semantic image segmentation. arXiv [Internet]. 2017. Available from: <https://arxiv.org/abs/1706.05587> [Accessed: February 28, 2025]
- [52] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. CVPR. 2015:3431-3440. DOI: 10.48550/arXiv.1411.4038
- [53] Xie E, Wang W, Yu Z, Anandkumar A, Alvarez JM, Luo P. SegFormer: Simple and efficient design for semantic segmentation with transformers. arXiv:2105.15203. 2021. DOI: 10.48550/arXiv.2105.15203
- [54] Peng Z, Dong L, Bao H, Ye Q, Wei F. BEiT v2: Masked image modeling with vector-quantized visual tokenizers. arXiv [Internet]. 2021. Available from: <https://arxiv.org/abs/2208.06366> [Accessed: February 28, 2025]
- [55] Liu Z, Hu H, Lin Y, Yao Z, Xie Z, Wei Y, et al. Swin transformer V2: Scaling up capacity and resolution. CVPR. 2022:12009-12019. DOI: 10.48550/arXiv.2111.09883
- [56] Caron M, Touvron H, Misra I, Jegou H, Mairal J, Bojanowski P, et al. Emerging properties in self-supervised vision transformers. ICCV. 2021:9650-9660. DOI: 10.48550/arXiv.2104.14294
- [57] Fang Y, Wang W, Xie B, Sun Q, Wu L, Wang X, et al. EVA: Exploring the limits of masked visual representation learning at scale. arXiv:2211.07636. 2022. DOI: 10.48550/arXiv.2211.07636
- [58] Fang Y, Sun Q, Wang X, Huang T, Wang X, Cao Y. EVA-02: A visual representation for neon genesis. arXiv:2303.11331. 2023. DOI: 10.1016/j.imavis.2024.105171
- [59] Zhang Y, Gong K, Zhang K, Li H, Qiao Y, Ouyang W, et al. Meta-transformer: A unified framework for multimodal learning. arXiv:2307.10802. 2023. DOI: 10.48550/arXiv.2307.10802

E-Learning in Light of the Concept of Power

Ismail Noori Mseer

Abstract

From the point of view of the analysis based on the aspect of power dynamics, e-learning reveals a number of critical factors influencing both the educational experience and broader implications for learners and educators. These include access to technology, curriculum design, social presence, and the role of feedback and assessment. Access to technology forms one of the very basic aspects of e-learning, which relates directly to the power structures. Online learning requires dependable access to the internet and appropriate devices. Such is a source of inequality among students hailing from different socio-economic backgrounds, geographic locations, and all manner of digital literacy. For instance, underprivileged students lack adequate resources to engage in effective e-learning and thus accelerate the existing inequalities in education. The concept of “virtual presence” underlines how technology access influences the experience of learning and either enables or constrains students’ possibilities to meaningfully engage with the learning material and peers. Curriculum design also reflects major discourses of power in e-learning. Choices on what is taught, how it is presented, and who is allowed to contribute or kept outside of such decisions are indicative of deeper structures of power at work within educational institutions. Often, the development of curricula does not take into account student input, which may be the solution for the disconnection that exists between what is taught and what really needs to or may interest a learner. This non-decision-making power can marginalize the voices of students and limit their active and deep engagement with the material. The Community of Inquiry framework stresses that teaching presence plays a critical role in fostering meaningful learning experiences; hence, if students are excluded from curriculum discussions, their cognitive and social presence is reduced.

Keywords: E-learning, power, technologies of self, commentary, taxonomy, biopolitical field, subjugated knowledges, cognitive discourse

1. Introduction

E-learning forms the basis of new power dynamics between instructors and learners through the incorporation of technology. E-learning bridges the gap in geography by allowing quality education to reach out to students in remote areas and thus helping in the empowerment of those citizens who otherwise were at a disadvantage due to their location or socio-economic background. Such access allows learners to engage

with educational content to which they otherwise did not have access [1]; this acts to redistribute power from traditional institutions to the individual learners who now can choose their own learning paths and resources. E-learning platforms often use adaptive technologies that permit individual learning to take place at an individual pace and according to individual style [2]. By putting the learners in control of their educational journeys, personalization empowers them with a sense of agency that is rarely experienced in a traditional classroom setting. In an e-learning environment, educators become more like facilitators or guides rather than the most important provider of knowledge [3]. This role adjustment will allow the teacher to spend more time on student learning processes rather than mere delivery of content, hence allowing for better engagement and understanding on the part of students.

Technology has made the learning materials more interactive with mechanisms for instant feedback. On the flip side, over-reliance on technology could also be the cause of inequality because students who do not have either the relevant internet or hardware would always lag behind and hence present a constant challenge in assuring equal opportunities for e-learning resources [4]. Data obtained through e-learning technologies creates insights on students' performance and behavior, empowering educators and learners alike. Educators can use this information to identify learning gaps, hence modify their teaching strategies accordingly [5], while at the same time allowing learners to monitor their progress and take ownership of their educational outcomes. E-learning gives a facelift to the power dynamics in educational contexts through increased access, personalized learning, changed roles by the educator, and the use of technology for enhancement in learning and feedback [6]. While it affords the opportunity for democratization across different strata, it equally demands consistent struggle for equal access to make such realities possible [7] for all learners.

Cognitive analysis in e-learning is one of the critical components whose keys to understanding lie in how learners process information, deal with cognitive load, and interact with digital content. It thus plays a very vital role in formulating effective e-learning experiences to improve learning outcomes. The theory of Cognitive Load Theory (CLT) focuses on the need to balance intrinsic load, related to material complexity, with extraneous load, related to the presentation of information so as to optimize learning. Effective instructional design minimizes extraneous cognitive load while maintaining intrinsic load at an appropriate level relative to a learner's current level of expertise. These are the main factors that influence the effectiveness of interactive learning media with regard to cognitive engagement and learning outcomes: the degree of interactivity should be optimized so as not to overwhelm learners with information. Immersion and Realism: Immersive technologies in learning, such as VR, have a motivational appeal that reinforces experiential processes of learning. On the other hand, a very high degree of immersion may be associated with an increased cognitive load that may detract from learning if it is not properly moderated. A moderate dose of disfluency introduced will stimulate deeper cognitive processing for better understanding and retention. Triggers presenting positive emotions in learners increase engagement and motivation to learn, hence motivating deeper cognitive processing and information retention. The assessments need to be aligned with learning objectives and cognitive processing types triggered by the instructional design so that the cognitive load invested at the time of learning translates into measurable educational benefits.

E-learning spaces are negotiated through the interactions of instructors and learners. These include: E-learning space reconfigures instructors from being

transmitted directly as an authority figure into a facilitator, in which direct student involvement and ownership of learning can be promoted that improve engagement. When students are given a choice in learning, for example, this increases motivation and engagement in pursuing ownership of their learning processes. Effective facilitation strategies provide inclusive dialogical spaces that balance power dynamics to ensure diverse participation in the engagement and learning outcome. Sometimes, access to technology and digital literacy widens inequities in engagement and learning outcomes. Inequalities within technology access and disparities in digital literacy create inequities within learning outcomes. The inclusion of formative assessments that allow for self-assessment-reflection and peer feedback may empower students by giving students more voices in their evaluation process. The integration of cognitive analysis with an understanding of power can enable educators to be more sensitive to the creation of inclusivity and more empowering online learning environments through collaboration and increased student engagement.

Even though e-learning is flexible and accessible, it introduces a number of considerations with respect to power: multilayered issues entailing motivation and access, digital literacy, and shifting roles of educators and learners. One of the major issues is that students tend to lack motivation when participating in an e-learning environment [8]. The absence of face-to-face interaction and a lack of immediate accountability can create less motivation and engagement because such students lose focus over some period of time. This lack of motivation is joined by feelings of isolation, given the fact that students do not get to participate in the social interactions that take place during a conventional classroom environment, and could affect their mental health and academic performance. Technical issues are a significant part of the problems that come with e-learning [9]. Poor connectivity or a lack of access to relevant devices to learn is often frustrating for students, as the process is mostly crippled and always leads to irritation among students. For example, this is particularly a problem when there is a “homework gap” because students come from low-income families or from rural areas where they have limited access to essential technology to complete assignments, thereby exacerbating disparities in education.

Further complicating all of these challenges is the added layer of power dynamics in e-learning. In other words, traditionally, the instructor has held the dominant position—that is, the only one with knowledge—which can effectively create a power differential that inhibits student agency. In a more student-centered approach, power is distributed to students through greater control over learning, which fosters community and engagement [10]. This shift in approach also means that educators need to get accustomed to new facilitator roles, which is not an easy task for those who are used to the traditional ways of teaching. Apart from that, digital literacy is another main factor that determines the success of e-learning. Good digital literacy enables a student to progress and function well in the depth of online platforms and learning resources with great efficiency [11]. Yet inequalities in digital literacy create unequal opportunities for participation that are dominated by more tech-savvy students over other less-proficient students.

Finally, there are the psychological effects that cannot be left unconsidered, related to remote learning: feelings of loneliness and increased stress. On the other hand, this might be reflected in the reduced level of motivation and engagement in class activities due to the exclusion of social interaction typical of the classical classroom environment [12]. E-learning in the context of power relationships is multilevel and quite intricate. In effect, overcoming these challenges would definitely necessitate concerted efforts from institutions of learning with regard to ensuring equal access

to technology, providing support for their students and educators, as well as promoting digital literacy in order to finally create effective and inclusive online learning environments.

2. Research methodology

The cultural context can be very important in the analysis of power relations in e-learning, as there are cultures that have different norms concerning authority, collaboration, and style of communication. E-learning research methodology is a sensitive field, more so on power issues and cultural consideration aspects, hence deepening the understanding of how the cultural norms and their power structures interplay in the shaping of learning environments [13]. Understanding Power Dynamics in e-learning shall be realized by knowing that interaction exists between instructor authority and student autonomy [14]. In most cases, instructors are granted enormous authority in traditional settings, which usually triggers power imbalance. This is further exacerbated in e-learning whereby the instructor can easily assume the position of the central authority, adopting a “sage on the stage” [15] mentality, which might reduce student participation and empowerment. Adoption of an empowerment framework may counteract some of these imbalances by fostering mutual responsibility for learning processes between instructors and students. Cultural context forms the core of any process or relationship encountered within e-learning [16]. Authority is not viewed equally across all cultures, and these cultural differences greatly affect how individuals relate to one another [17], make decisions, and exchange information. For instance, high power distance cultures, including many Asian and Middle Eastern cultures, deeply respect status and authority, which can result in very formal communications and decision-making approaches that are top-down [18]. On the other hand, low power distance cultures, such as the United States and Northern Europe, foster egalitarianism and encourage open communication, thus creating collaborative learning environments.

Cultural contexts influence collaboration styles in e-learning. For instance, collectivist cultures place a high value on group cohesion and consensus and are thus more likely to bear reflective evaluators who are inordinately concerned with the group’s cohesion versus the need for self-expression [19]. Individualistic cultures might provide more active initiators interested in personal, reflective contributions and assertiveness in discussions. The communication style is shaped by cultural differences, learning preferences, and the nature of the online interaction. High-context cultures rely heavily on non-verbal cues and context of communication and very often would favor indirect communication and hierarchical relationships [20], whereas low-context cultures depend on direct communication and egalitarian interactions. The difference also has an impact on the way learners will approach e-learning platforms: for example, students coming from high-context cultures may appreciate more guided approaches and explicit instructions from instructors whenever necessary [8]. Understanding the power relations in e-learning requires knowledge of the cultural context. The way perceptions of authority, collaboration style, and communication norms interact determines how learners will interact with educational content and with one another. E-learning methodologies need to be flexible in meeting such cultural differences as a means of fostering increased engagement and effectiveness across diverse learner populations [21]. It is through such factors that educators can be

set to establish e-learning environments that are not only more inclusive but also more productive, culturally sensitive, and equally offer opportunities for learning.

2.1 E-learning through the concept of technologies of self

E-learning, based on the concept of technologies of self, necessitates emphasizing learners' powers of self-management in education. This utilizes various digital tools and methodologies that allow personalized learning, targeting the individualizing of each person's education toward their specific needs and preferences [22]. These technologies lead to domination in cognitive guidance within a few key areas: The technologies of self allow learners to set personal goals, monitor their progress, and adjust their strategies in response to feedback. This level of autonomy creates ownership of the process that is motivating and engaging. Research is indicating that students who practice self-regulated learning are obtaining higher academic performance and satisfaction in an online environment. Technology integration within e-learning courses enhances cognitive presence in that learners are afforded reflection and discourse to construct meaning. Discussion forums and collaborative platforms foster a deeper level of cognitive engagement that supports more critical thinking and problem-solving skills [23]. Motivation and Engagement: E-learning environments utilizing technologies of self create increased motivation of the learners because one can select learning paths and use resources that best suit individual interests; hence, involvement with the content is deeper. Technologies of the self are directly fed back to learners through built-in digital tools of quizzes and analytics dashboards.

Real-time feedback allows the learner to diagnose their level of understanding from time to time and make necessary changes, something that is integral for any learning to take place. Technologies promoting social interaction foster the development of a learning community. High social presence will build closer collaboration, including peer support, essential in online contexts for motivation and engagement. In other words, technologies of self through e-learning significantly influence cognitive guidance in developing self-regulation, cognitive presence, motivation, immediate feedback, and social connectedness. Each of these factors leads to the transformative power of technology in shaping an efficient e-learning environment that focuses on learner autonomy while guaranteeing robust cognitive engagement.

2.2 E-learning through the concept of commentary

E-learning through the concept of commentary greatly powered the production of digital narratives by affording the digital specialist the ability to insert meaning and context into their storytelling. This approach uses interactive and multimedia features to facilitate learning in a more engaging and meaningful manner. The integration of different media formats is one aspect of this process that facilitates an enrichment of the narrative. Digital storytelling is not just an affair involving text; instead, it incorporates visuals, audio, and interactive components that might help in stimulating a range of senses. This multimodal focus enables learning more easily because learners can relate to the content at deeper levels owing to the fact that they would be able to connect with the content through different channels, which in turn makes their understanding and retention of the same easier. For instance, video content supported with written text would allow the learners to visualize ideas and at the same time process the information aurally, which, in turn, can also promote Richard Mayer's Cognitive Theory of Multimedia Learning [24], dual-channel processing

leads to better understanding and recall. The other key issue is related to commenting on the role of interaction between the narrator and the audience themselves. Where the storyline in traditional narration is set and often told from one perspective, in digital storytelling, many voices and perspectives may emerge. Interactivity enhances active involvement of learners in processes of learning, either through feedback or by creating alternative viewpoints that reshape the story. Engagement of this nature not only makes learning more dynamic but also fosters a sense of community among the learners, with the sharing of experiences and insights related to the content.

Moreover, commentary allows for emotional involvement with the text. Stories of emotions are easier to remember and interpret. When learners are involved in the narrative—through characters that they can relate to or situations that may seem gripping—they are bound to put more effort into grasping what is before them. This emotional interplay allows for better retention of information since students would likely remember information that appeals to them on a personal level. Digital narratives can be flexible [25], letting students individualize their learning experiences. Commented stories by educators can show relevance to a wide range of learning needs and styles. Students from a whole range of abilities can engage with the content in ways that resonate with their particular experiences—be it visualized through aids for ocular learners or audio for more auditory learners. This is important in today's diverse educational settings, where one-size-fits-all solutions mostly break down. Commentary as e-learning in digital narratives develops a media-based narration with opportunities for engagement, interaction that allows different views, emotion to help retention, and personalization of learning experiences. Together, these provide a more meaningful educational experience—one that does not just inform but also inspires the learner.

2.3 E-learning according to the concept of taxonomy

E-learning through the concept of taxonomy involves arranging and structuring so that learning experience may be enhanced and will allow scattering at different sites. Taxonomy here means a well-structured classification that navigates educators and learners through complex information landscapes. Among the major frameworks used to understand taxonomy in e-learning, Bloom's Taxonomy categorizes higher-order cognitive skills by naming them from basic knowledge recall to higher-level thinking skills involving creation and analysis. This framework can be applied in digital learning environments that enable instructors to design courses in accordance with specific learning outcomes. For instance, with Bloom's Digital Taxonomy, instructors can use technology effectively to provide an environment that elicits much engagement and thinking from students [26]. In this version of the update, one would emphasize the use of digital tools at each of the cognitive levels, be it doing simulations for application or using collaborative platforms for creation. Clearly specifying learning objectives within the framework of the taxonomy will enable educators to develop focused assessments that will, in turn, effectively measure student understandings at multiple cognitive levels. This structured approach ensures that learners do not only absorb information but also apply and analyze it meaningfully. Taxonomies will, in a way, help in categorizing content into manageable sets so that learners can have easier access to relevant content. It may further assist in effective content delivery on various digital platforms to ensure that learners receive the right information at the right time. For example, a well-developed taxonomy allows resources to be collated about particular topics or skills and hence enhances learning experience.

Taxonomies can also redefine relationships through the collaboration of learners. Students can discuss and share, using social media and collaboration tools within a taxonomy [27], their learning outcomes. Interaction in this manner not only enriches the learning but also creates community among learners. Well-defined taxonomy helps in the personalization of learning experiences, thus allowing educators to customize content with regard to assessment to meet each learner's needs. This is highly required in the e-learning environment wherein different students share different levels of pre-knowledge concerning a subject matter and, therefore, different learning styles. A taxonomy in e-learning will provide data on performance and learners' engagement. The instructor can use this information to fine-tune teaching strategies and course design to improve learning outcomes [28]. In other words, taxonomy for e-learning not only systematizes the educational content but also shapes anew the relationship between learners and knowledge through advanced learning objectives, better distribution of the content, collaboration, personalization of learning routes, and data analytics. It is by exploiting these latter aspects effectively that educators are able to provide for a much more interactive and effective online learning experience that caters to many different needs of various learners.

2.4 E-learning through the concept of the biopolitical field

E-learning, when viewed from a biopolitical perspective, seems to have shown that pedagogical practices are inherently linked with the struggle of powers constituting both knowledge and social norms. The concept of biopolitics is an analytical category developed by the French intellectual Michel Foucault and denotes how a government administered by controlled subjects orchestrates life control [29]. Biopolitics can be applied in a number of ways concerning education. Another important aspect of biopolitics in e-learning is its possibility to fabricate and devise the entire set of facts and events. For instance, analytics might become used to measure the level of student participation, performance, or even their mood on e-learning platforms. Data gathered from such efforts allows educational facilities to profile the learner in ways that might be material for pedagogical strategy and institutional policy. For example, pattern analysis identifying students at risk could trigger interventions specific to the performance in question. This is biopolitics in action: knowledge of individuals' lives applied to govern those lives in some respect [30]. Indeed, e-learning environments have a greater potential to regularize societal norms by integrating certain values and ideologies into their curricula. In these, specific content is presented based on broader ideas as dictated by government or institutional bodies. This could, therefore, be understood as one aspect of a biopolitical control of knowledge, where certain forms of knowing and practice are systematically enabled and developed, while others may be moved to the periphery. For example, the use of standardized testing in e-learning systems determines not only what is learnt but also how students conceptualize success and failure.

Biopolitical influence goes beyond the levels of individual learning to collective pedagogical outcomes evident in e-learning. For Foucault, power operates not only through hierarchies but also through networks [31], so that e-learning will enable new social formations of learners. The online communities coalescing around learning could subvert conventional educational structures through the encouragement of alternative types of knowledge generation. Yet these communities are also open to surveillance and regulation by the educational authorities anxious to reassert their control over the learning environment. This would be manifest in not only developing

countries but also in a variety of marginalized communities. Biopolitical principles toward e-learning show very different implications [28] where, on one hand, e-learning democratizes access to education, yet, on the other hand, may further exacerbate existing inequalities due to a lack of equal access to technology. The presence of the digital divide helps to further emphasize that in developing biopolitical strategies, the socio-economic factors of who benefits from any given e-learning initiative are paramount. In other words, e-learning, as part of a biopolitical field, serves to illustrate that aspects such as educational practices are never independent of the power relationships that create knowledge and social norms [32]. E-learning environments, through data analysis of learners and curation of content for institutional aims, epitomize Foucault's notion of biopower, whereby knowledge is a function of governance. It is in grasping these processes that educators and policymakers will be in a position to ensure equity and efficiency in learning environments that enable rather than constrain learners.

3. E-learning through the use of subjugated knowledges

E-learning platforms are important drivers in the reshaping and redefinition of knowledge categories by incorporating subjugated knowledges—that is, those knowledge forms that have been marginalized or excluded from mainstream discourse. These platforms democratize education by offering access to a diversity of perspectives, especially those emanating from the margins, and hence challenging conventional knowledge hierarchies. It increases access by overcoming geographical constraints [33], thereby allowing different learners to access content to which they can relate due to their experiences. This is a question of inclusivity that helps in the recognition of subjugated knowledges—for instance, indigenous knowledge systems or feminist theories—that can challenge dominant narratives. Indigenization of knowledge and skills might also be effected through the incorporation into the curriculum of various epistemologies, such as local ecological knowledge or traditional healing practices. Such a move would validate such knowledge claims as central rather than peripheral and would contribute to a more holistic understanding of subjects like science. The insights of historically oppressed groups become significant knowledge in e-learning platforms that value such knowledge. Hence, this model of learning opens up the academy to the inclusion of alternative perspectives that enrich the educational landscape. Traditionally, independent knowledge is taken to be the objective truth. However, from the perspective of subjugated knowledges, e-learning demonstrates that all knowledge is contextual and interdependent between different forms of understanding. Direct contact in e-learning through, for example, interactive discussions and collaborative projects increases student interaction, which in turn strengthens social ties among students. In addition, this contact may help reduce feelings of loneliness, and it enables feedback that provides an opportunity for personalized learning experience. Integration of subjugated knowledge in e-learning promotes a culture of lifelong learning and adaptability [34]. The more diverse the perspectives with which learners engage, the more likely they are to seek out new knowledge on their own initiative, which in turn should encourage organizations to invest in continuous training and development initiatives. In a nutshell, e-learning platforms have a transformational role in the reordering of educational paradigms by underlining access and direct contact with voices from the margin. It reconstitutes knowledge categories through an affirmation of diversity in epistemologies, furthers

equity in education, and complicates the understanding of knowledge as a complex play of cultural narratives and experiences.

4. E-learning through the concept of cognitive discourse

E-learning through the concept of cognitive discourse authority merges scientific knowledge with social interaction to work out effective online learning environments. Cognitive science is thus at the root of this concept, which guides the design and implementation of e-learning experiences, putting an emphasis on information processing and construction of knowledge by learners. Authority of cognitive discourse in e-learning is informed by a number of significant theories and frameworks that govern it. These are: Cognitive Load Theory (CLT): This theory, as propounded by John Sweller [35], prescribes that instructional materials should minimize extraneous cognitive load and maximize intrinsic and germane loads. By this approach, learners shall be able to focus on the essential concepts without being overwhelmed, thus enhancing retention and understanding. Cognitive Flexibility Theory: The theory insists on exposing information to different contexts so that a deeper understanding and adaptability can be achieved. E-learning with case studies and simulations enhances knowledge transfer and critical thinking in multiple dimensions.

A presence based on cognition within the Community of Inquiry model, cognitive presence deals with constructing and confirming meaning by reflection and discourse. Discussion forums and peer feedback are part of an effective e-learning platform that enhances cognitive engagement. Interaction with Peer and Instructor: Interaction will develop the feeling of belongingness and community [36]. It enhances the motivation and understanding when working by sharing different perspectives and getting support from each other. A strong feeling of social presence reduces feelings of loneliness while studying in distance learning; an effective role of participation and cooperation leads to enhanced learning outcomes. Self-Regulation and Technological Interaction: Efforts toward self-regulated learning in combination with effective use of technology are necessary in bringing involvement in course content to a deeper level of meaning. Rich multimedia resources and mechanisms for immediate feedback support cognitive engagement. Cognitive discursive authority also means cognizing how individuals evaluate sources for credibility. This includes the role of emotions and contexts in perceptions of authority. The emotional appeal of information can have a crippling effect on perceptions of authority. One can easily understand how navigating through diverse sources today is incredibly overwhelming. In a nutshell, the cognitive discourses' authority in e-learning underlines the interdependence between scientific knowledge and social interaction. Educators develop effective e-learning experiences that prepare students for academic success and active citizenship in a scientifically informed society by promoting rich dialogs and addressing socio-emotional factors.

5. E-learning through knowledge transfer

E-learning through the concept of knowledge transfer, in fact, tries to bridge the theoretical apprehension of knowledge with practical application and thus the ability of learners for the application of that knowledge in a real-world context. This is achieved through several key strategies: Step-by-step learning involves smaller “bytes”

of learning about certain concepts or skills [37]. Most e-learning courses are divided into sections that are manageable, either focusing on a concept or skill. In this respect, learners will find it easier to learn and confidently apply theoretical inputs effectively on the job. **Community Learning: Engaging learners in activities of collaboration,** such as discussion and group projects, creates interaction with peers that is very critical for both knowledge retention and application. The aspect of interaction in a social context promotes sharing experiences and insights during learning that is relevant and appropriate in a daily working environment. **Practical Exercises and Assessments:** The integration of practice questions and real-life scenarios in the e-learning course [38] will reinforce knowledge and better equip the learner to face every challenge in life. The continuous assessments offer feedback, which is then applied to create continuous improvement in the application of knowledge. **Business Goal Alignment:** The alignment of the learning outcomes to business objectives means that the training will be pertinent to the roles of the learners. Such alignment gives them a reason to go through the material and use it effectively; this helps in associating the objectives of training with organizational success. **Ongoing Support and Follow-Up: Refresher courses with follow-up exercises** enhance long-term learning. Ongoing practice will allow learners to review their knowledge and apply it in a different context, thereby keeping it usable and executable for a long period following the initial training. In other words, e-learning strengthens practical ways of knowledge by using structured content delivery [39], stimulating community involvement, applying practical exercises, aligning training to business objectives, and giving continued support. All these elements combined amplify the effective means of application of knowledge by the learner within realistic situations—a dream to cause meaningful learning impactful right at the heart of reality.

5.1 E-learning, through the concept of semantic representations

Semantic representations have recently gained immense traction in e-learning platforms for the betterment of educational experiences by facilitating personalization, content discovery, and assessment methods. Semantic representation involves data with a structure that conveys meaning, hence the ability of the systems to process information as it is comprehended by humans. Semantic analysis tailors learning content for each individual learner by highlighting strengths and weaknesses. This shall enable platforms to recommend resources that better match learners' current understanding and goals, hence boosting engagement and retention [1]. Semantic technologies improve the packaging of content and how it is accessed by providing users with intuitive search capabilities. For example, if searched regarding “machine learning,” results would include related topics and resources about the very same [40], based on context through semantics rather than a word-for-word search result [1]. E-learning systems, with natural language processing or NLP embedded within them, may comprehend what students respond to correctly reflect their comprehension and application of ideas, unlike traditional ways of testing. **Adaptive Learning Pathways:** It has the ability to dynamically adapt the curriculum to learner progress with appropriate challenge levels without students being overwhelmed. **Collaboration and Knowledge Sharing:** Semantic representation allows for better collaboration by indicating relevant discussions, thereby showing learners how ideas are related; this leads to a fuller collaborative environment.

Semantic technologies enhance education through information organization in meaningful ways and access to it. They enhance knowledge construction by enabling

intelligent searches where realization of the context and relations between concepts is facilitated, thereby enabling deeper understanding due to bringing coherence into information presentation. Personal Learning Networks (PLNs): Semantic technologies automate content curation [41], relieving learners from spending more time searching for content than engaging with the content itself. This improves learning experiences in educational terms, as learners are better connected to resources relevant to their interests and goals. Standardization of course descriptions through shared ontologies, facilitated by semantic technologies, has come to make credit transfers and recognition of prior learning between institutions easier, thus enabling students to create personalized pathways of education. Linguistic and Rhetorical Tools in E-Learning: It encompasses terminological, linguistic, rhetorical, and conceptual tools to provide semantic understanding in e-learning. Such tools will be supportive in navigating online communication and knowledge construction. Terminological Tools: It has to do with ensuring shared vocabularies are employed appropriately to reduce ambiguities; hence, it provides a better learning atmosphere. Linguistic Tools: The employment of language structures and styles that befit easy communication promotes better comprehension [42]. Rhetorical Tools: Strategies to engage the learner effectively through ethos, pathos, and logos. Conceptual frameworks: Utilize organized structures like Bloom's Taxonomy so as to conceptualize abstract thinking. Semantic representations in e-learning approach formal and cognitive semantics from various perspectives so as to construct appropriate learning environments. Logical structures, understanding cognitive processes or using rhetorical strategies are useful ways educators could improve the learning experience [4].

5.2 E-learning by means of the concept of discourse units

The concept of discourse units shows e-learning from the perspective of how communication is constructed and facilitated in the digital environment. By discourse units, it means segments in the flow of communication which may be subjected to analytical scrutiny in terms of how well they have served their purpose in relaying information and bringing about interaction among participants. Therefore, effective digital communication and dilapidated digital communication for e-learning are important to distinguish between in order to definitely influence the learning process [43]. This means that effective digital communication within an e-learning environment should be clear, interactive, and engaging, using a lot of digital tools and platforms for better collaboration and meaningful exchange between the learners and instructors. Platforms providing support for real-time discussions—for example [44], forums or chat applications—are a good platform to get immediate feedback or to encourage active participation. This is consistent with research that emphasizes how important collaborative learning environments are for students to communicate effectively with their co-students and instructors. Digital technologies have also been found through research to potentially really enhance the communication and collaboration competencies of students when introduced thoughtfully into the processes of learning. On the other hand, impoverished digital communication involves an e-learning system that has a poor design or poor implementation of the same that poses a barrier to effective interactions. This may be actualized through inadequate directions, the lack of tools to support participation, or an inability to enable activities necessary for collaboration. Moreover, a few of the learning contexts might always involve disappointments among the learners, making them disinterested, hence compromising the experience.

For example, if a digital platform is not friendly in providing such features or resource materials necessary for supporting collaboration, students may fail to communicate effectively to avoid disputes or perhaps without the motivation to work collaboratively. The burden of the communication space could be brought on by dilapidated digital environments, leading to cognitive overload. Poorly organized content or an overabundance of information without any support or guidance on how it is to be navigated impairs learners' capability to process [45] and engage with the material. This lies on the underlying cognitive load theory, which espouses that effective learning occurs when the cognitive resources are optimally challenged without irrelevant distractions or an overload of those resources. In a nutshell, the effectiveness of e-learning by discourse units relies upon the quality of the digital communication. While efficient digital environments facilitate engagement and collaborative effort due to clearly structured and supportive tools, dilapidated ones result in barriers that discourage learning. The effects of such differences reach much further because they not only impact the immediate learning outcomes but also influence the attitude taken toward online education in general. It therefore goes without saying that a well-designed e-learning platform, respecting the spirit of good communication and aimed at an improvement in educational experiences within digital contexts, is an investment worth its weight in gold.

5.3 E-learning in light of the concept of discursive formation

E-learning, within the concept of discursive formation, means knowledge of how language and discourse shape and mold educational experiences and life realities. Stemmed from theories by Michel Foucault, it identifies ways through which practices of language contribute to the dynamics of power and discrimination in digital learning environments. Discursive Formation in E-Learning: This is the discursive formation concerning e-learning, which involves how the language and the social practices shape understanding and implementation pertaining to digital education. Thus, it stipulates that knowledge and the power dynamic play a role in the implementation of educational practice and perception within an e-learning environment. Identity Construction: One key dimension concerns how the identities of teachers and learners are produced [46]. Thus, for example, in technology-mediated contexts, teachers are made out to be technologically proficient and their identities are mediated through dominant discourses about what constitutes a "good teacher." Through this process, then, good teaching is reduced to an issue of technological proficiency at the expense of pedagogical content. Technology itself plays an essential role as a mediator of knowledge. Most of the e-learning platforms encourage such discourses around learning through autonomous study and flexibility. These discourses tend to obfuscate the issues of equity and access as not all the learners are provided with equal means to succeed in an online environment. Power relations between institutions and e-learning discourses are highly intertwined. It is through institutional policy that decisions about the use of technology in teaching and learning determine what knowledge is to be considered valuable.

The impact of this may be to homogenize education where local concerns or cultural differences are potentially dismissed. Linguistic practices play a significant role in developing the e-learning experience in regard to issues of inclusion and discrimination. Such practices include neutralization of linguistic content so as not to cause misunderstanding, cultural sensitivity, and investment in quality translation services. Discrimination in e-learning can manifest as implicit bias in assessment, disparity

in access, and in the dynamics of online discussion. The solutions that address these issues include cultural competence training, effective feedback mechanisms, and designing contents inclusively. Theoretical education models lay the foundation for the design and delivery of e-learning. ADDIE, Merrill's Principles, and Constructivist Theory [47] are some of the models that inform effective learning content design to meet the needs of a diverse group of learners. This means that through the notion of discursive formation, e-learning emphasizes how language shapes educational experiences. Thus, with an understanding of linguistic practices and their implications for discrimination and representation, educators can work toward constructing learning environments that are inclusive and equitable, where students feel empowered to engage with the course material meaningfully.

5.4 E-learning, through the concept of knowledge

E-learning, from a knowledge perspective, can actually be regarded as a direct embodiment of power. Such a perspective arises from a number of interlinked factors that reveal how knowledge acquisition and sharing in online environments might nurture or question already established structures of power. E-learning platforms provide people with unparalleled access to knowledge and learning resources, an unprecedented shift in educational landscapes. The most important geographical or time-based barriers fall away since learners can access online courses and materials at any place and any time, accommodating differing schedules and learning rates. Such democratization of education implies that more people can acquire quality training without being confronted with financial burdens, as e-learning reduces costs associated with traditional education. The saying "knowledge is power" by Francis Bacon [48] denotes that knowledge allows individuals and society to make informed decisions and wield power over nature. The implications this has upon various areas of life are gigantic and range from education and personal growth [9] to the very philosophy of e-learning, which actually rests on providing access, affordability, personalization, and interactivity to continuing professional development and sustainability [49]. Besides, most e-learning systems utilize adaptive technologies for content presentation to meet the learning needs of each individual. This implies that learners would always get instruction matched to their current skill level and learning style.

This personalization enhances engagement and retention as a means of constructing understandings and challenging established narratives by learners, hence manifesting a form of power through knowledge. E-learning democratizes education; that is, making it more accessible and equal for several kinds of groups. It means that the e-learning courses and materials are open to different categories of students that henceforth give them the opportunity to learn what was previously inaccessible to them. In this regard, there is global contact that would lead to cultural interchange and understanding—an ingredient needed in today's world. In brief, it is the power of e-learning: a medium that reshapes knowledge production, amplifies learner participation and interaction—in particular ethical dimensions—while democratizing information access, thereby empowering the learner; at the same time, it unleashes new dynamics of control and responsibility that ought to be exercised with care [50]. This interplay between knowledge and power in e-learning environments suggests that critical engagement with technologies should be encouraged to ensure that technologies are used as an instrument of empowerment rather than as an instrument of control.

5.5 E-learning, through the concept of tacit knowledge

E-learning platforms act as important facilitators in the transfer and acquisition of knowledge that is subjective and hard to articulate by nature. Examples of such knowledge are experience, skills, and insight flowing out of personal and professional lives. The digital era ushers in a number of ways that facilitate better sharing of this kind of knowledge: E-learning platforms are integrated with collaborative tools that enable learners to discuss and share experiences. Social interaction is very important for the transfer of the knowledge [51], since people will construct the knowledge through shared conversations and collaborative projects. Online social networks sometimes turn out to be more viable in this area compared to face-to-face interactions. E-learning environments make use of storytelling for the delivery of tacit knowledge. Personal narratives or case studies help set information in context and allow better assimilation. The methodology encapsulates the nuances within the tacit knowledge and emotionally appeals to the learners for better retention and comprehension. Guided experience: Platforms host guided experiences where learners observe experts, practice [52] skills under supervision, and get feedback. This can reimagine traditional models of apprenticeship for digital environments via video demonstration or interactive simulation.

Some e-learning systems track lessons learned from projects or experiences. This helps retain valuable insight and allows others to learn from it. The process ensures that crucial tacit knowledge is not lost in the instance of employees leaving or moving positions. E-learning fosters a continuous learning environment by creating opportunities for skill development and knowledge sharing [53], activities such as webinars and discussion forums. It helps in converting the tacit knowledge into more explicit forms. Encouraging learners to “show their work” means their thought processes are made visible in the process of completing their tasks so that others understand the underlying tacit knowledge that informs the decision-making. The digital age has also transformed the way in which tacit knowledge is elicited and shared across organizations. While digital means can enable the access to information [54], especially for the nuance of understanding that is usually lost when it comes to tacit knowledge, social media and collaborative networks allow for real-time communication and an ecology that enables an experience-based learning with natural spontaneity. AI technologies will increasingly emerge as bridging tools of tacit-explicit knowledge gaps through patterned data analysis and insights that augment human intuition. In other words, e-learning platforms have a spectacular opportunity in supporting the transfer of knowledge by offering collaboration tools, storytelling techniques, guided experiences, tracking of lessons, creating continuous learning environments, and work process visibility. These approaches further strengthen the learning ecosystems and ensure that valuable knowledge of a tacit nature is always shared for the success and innovation of an organization [1].

6. Conclusions

E-learning represents an important development in modern education; indeed, it has totally changed the balance of power between teacher and student. This is actually reflected in a number of important ways. E-learning widens access since learners can utilize educational resources anytime and anywhere, hence surmounting many of the geographical and financial barriers that usually hinder conventional education.

Flexibility especially helps non-traditional learners, such as employed professionals or those with family constraints, study in comfort at their own speed. E-learning platforms often employ adaptive learning technologies, catering to individual learning styles and preferences. This kind of personalization of learning has shifted the balance of power in favor of students by placing them at the center of their educational journeys, where students can now choose what and how they learn. This makes students active learners rather than passive recipients of knowledge, hence increasing effectiveness and engagement in learning.

Another important factor in how e-learning impacts power dynamics is cost-effectiveness. It diminishes costs related to tuition, commutes, and materials; thus, it raises affordability and opens its doors to more people. In this respect, such democratization of education enables different groups from various socio-economic backgrounds to seek knowledge without the financial burden imposed by traditional institutions. E-learning crosses borders, creating international learning communities that allow participants from very different walks of life to engage in greater cultural exchange and cooperation. Interconnectedness supports a bottom-up approach to knowledge sharing, challenging established power structures in education.


With great potential also comes a great set of challenges, including the digital divide problem. Not every student has equal access to fast and reliable internet or to modern digital devices, which creates a chance for further inequity in the education system. Some scholars argue that such inequality gaps should be curbed to ensure equal distribution of benefits from e-learning across populations. E-learning has huge implications for changing the dynamics of power in education through increased access, personalization of learning experiences, affordability, and creating global communities. All these benefits must, however, be seen within the context of unequal access to technology.

Author details

Ismail Noori Mseer
Scientific Advisor at Bayt Al-Hikma, Iraq, Baghdad

*Address all correspondence to: imseer60@gmail.com

IntechOpen

© 2025 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Clegg S, Hudson A, Steel J. The emperor's new clothes: Globalisation and e-learning in higher education. *British Journal of Sociology of Education*. 2003;**24**(1):39-53
- [2] Bliss L, Brooks SL, Huq C. Creating online education spaces to support equity, inclusion, belonging, and wellbeing. *John Marshall Law Journal*. 2020;**14**:1
- [3] Leach T. Paulo Freire: Dialogue, politics and relevance. *International Journal of Lifelong Education*. 1982;**1**(3):185-201
- [4] Whitworth A. Researching the cognitive cultures of e-learning. In: *The Sage Handbook of E-learning Research*. 2007. pp. 202-220
- [5] Gligorea I, Cioca M, Oancea R, Gorski AT, Gorski H, Tudorache P. Adaptive learning using artificial intelligence in e-learning: A literature review. *Education Sciences*. 2023;**13**(12):1216
- [6] Diana P, Catone MC. E-learning in an undergraduate course in research methods for the social sciences: Reflections on teaching. *Italian Journal of Sociology of Education*. 2016;**8**(2):110-142
- [7] Goodfellow R. From "Equal Access" to "Widening Participation": The Discourse of Equity in the Age of e-Learning. Vol. No. 37. Peter Lang; 2007. pp. 55-74
- [8] Aparicio M, Bacao F, Oliveira T. Cultural impacts on e-learning systems' success. *The Internet and Higher Education*. 2016;**31**:58-70
- [9] Shahmoradi L, Changizi V, Mehraeen E, Bashiri A, Jannat B, Hosseini M. The challenges of E-learning system: Higher educational institutions perspective. *Journal of Education and Health Promotion*. 2018;**7**(1):116
- [10] Aldowah H, Ghazal S, Muniandy B. Issues and challenges of using e-learning in a Yemeni public university. *Indian Journal of Science and Technology*. 2015;**8**(32):1-9
- [11] Dolenc K, Aberšek B. E-learning and teaching-methodological or technological problem. In: *Distance Learning in Applied Informatics: Conference Proceedings*. 2012. pp. 73-81
- [12] Cappel JJ, Hayen RL. Evaluating e-learning: A case study. *Journal of Computer Information Systems*. 2004;**44**(4):49-56
- [13] Downey S, Wentling RM, Wentling T, Wadsworth A. The relationship between national culture and the usability of an e-learning system. *Human Resource Development International*. 2005;**8**(1):47-64
- [14] Siemens G, Yurkiw S. The roles of the learner and the instructor in e-learning. *Preparing Learners for e-learning*. 2003:123-138
- [15] King A. From sage on the stage to guide on the side. *College Teaching*. 1993;**41**(1):30-35
- [16] Masoumi D. *Quality in E-Learning within a Cultural Context*. 2010
- [17] Edmundson A. The cultural adaptation process (CAP) model: Designing e-learning for another culture. In: *Globalized e-Learning Cultural Challenges*. IGI Global; 2007. pp. 267-290

- [18] Martinsons MG, Davison RM. Strategic decision making and support systems: Comparing American, Japanese and Chinese management. *Decision Support Systems*. 2007;**43**(1):284-300
- [19] Cools E, Evans C, Redmond JA. Using styles for more effective learning in multicultural and e-learning environments. *Multicultural Education & Technology Journal*. 2009;**3**(1):5-16
- [20] Slotte V, Tynjälä P. Communication and collaborative learning at work: Views expressed on a cross-cultural e-learning course. *International Journal on E-learning*. 2005;**4**(2):191-207
- [21] Zhu C, Valcke M, Schellens T. A cross-cultural study of online collaborative learning. *Multicultural Education & Technology Journal*. 2009;**3**(1):33-46
- [22] Kanuka H. Understanding e-learning technologies-in-practice. In: *The Theory and Practice of Online Learning*. 2008. p. 91
- [23] Osborne DM, Byrne JH, Massey DL, Johnston AN. Use of online asynchronous discussion boards to engage students, enhance critical thinking, and foster staff-student/student-student collaboration: A mixed method study. *Nurse Education Today*. 2018;**70**:40-46
- [24] Mayer RE, Moreno R. A cognitive theory of multimedia learning: Implications for design principles. *Journal of Educational Psychology*. 1998;**91**(2):358-368
- [25] Stackelberg PV. Creating digital narratives: The structure and design of stories told across multiple media [doctoral dissertation]. 2011
- [26] Alaghbary GS. Integrating technology with Bloom's revised taxonomy: Web 2.0-enabled learning designs for online learning. *Asian EFL Journal*. 2021;**28**(1):10-37
- [27] Kim MS. Developing a competency taxonomy for teacher design knowledge in technology-enhanced learning environments: A literature review. *Research and Practice in Technology Enhanced Learning*. 2019;**14**(1):18
- [28] Supraja S, Hartman K, Tatinati S, Khong AW. Toward the Automatic Labeling of Course Questions for Ensuring their Alignment with Learning Outcomes. *International Educational Data Mining Society*; 2017
- [29] Deacon R. Strategies of governance: Michel Foucault on power. *Theoria*. 1998;**45**(92):113-149
- [30] Marzocca O. *Biopolitics for Beginners: Knowledge of Life and Government of People*. Mimesis; 2021
- [31] Gulson KN, Webb PT. "Life" and education policy: Intervention, augmentation and computation. *Discourse: Studies in the Cultural Politics of Education*. 2018;**39**(2):276-291
- [32] Lindberg S. Politics of digital learning—Thinking education with Bernard Stiegler. In: *Bernard Stiegler and the Philosophy of Education*. Routledge; 2021. pp. 59-71
- [33] Parchoma G. Toward diversity in researching teaching and technology philosophies-in-practice in e-learning communities. In: *Handbook of Research on Methods and Techniques for Studying Virtual Communities: Paradigms and Phenomena*. IGI Global; 2011. pp. 61-86
- [34] Remtulla KA. The knowledge-based economy and E-learning: Critical considerations for. *Convergence*. 2007;**40**(1-2):10

- [35] Lovell O, Sherrington T. Sweller's Cognitive Load Theory in Action. John Catt; 2020
- [36] Thacker I, Seyranian V, Madva A, Duong NT, Beardsley P. Social connectedness in physical isolation: Online teaching practices that support under-represented undergraduate students' feelings of belonging and engagement in STEM. *Education Sciences*. 2022;12(2):61
- [37] Willis BD. *Distance Education: Strategies and Tools*. Educational Technology; 1994
- [38] Havryliuk N, Osaulchuk O, Dovhan L, Bondar N. Implementation of E-learning as an integral part of the educational process. *Society Integration Education Proceedings of the International Scientific Conference*. 2020;4:449-459
- [39] Al-Samarraie H, Teo T, Abbas M. Can structured representation enhance students' thinking skills for better understanding of E-learning content? *Computers & Education*. 2013;69:463-473
- [40] Portugal I, Alencar P, Cowan D. The use of machine learning algorithms in recommender systems: A systematic review. *Expert Systems with Applications*. 2018;97:205-227
- [41] Jovanović J, Gašević D, Devedžić V. Dynamic assembly of personalized learning content on the semantic web. In: *European Semantic Web Conference*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2006. pp. 545-559
- [42] Rayson PE. *Matrix: A Statistical Method and Software Tool for Linguistic Analysis through Corpus Comparison*. United Kingdom: Lancaster University; 2003
- [43] Ravenscroft A. Communities, communication and cognitive change: Social processes and designing engaging e-learning discourse. In: *E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*. Association for the Advancement of Computing in Education (AACE); 2002. pp. 792-797
- [44] Lee SC, Song J, Ko EY, Park S, Kim J, Kim J. Solutionchat: Real-time moderator support for chat-based structured discussion. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 2020. pp. 1-12
- [45] Hsu TE. Effects of Learner Cognitive Styles and Metacognitive Tools on Information Acquisition Paths and Learning in Hyperspace Environments. 1994
- [46] Miller J. Teacher identity. *The Cambridge Guide to Second Language Teacher Education*. 2009;4(1):172-181
- [47] Sugrue B. Performance-based instructional design for E-learning. *Performance Improvement*. 2002;41(7):47-53
- [48] Zhimin C. Is knowledge power? *Contemporary Social Sciences*. 2023;8(2):38-51
- [49] Parry G. Making democrats: Education and democracy. In: *Democracy and Democratization*. Routledge; 2015. pp. 47-68
- [50] Hefzallah IM. *The New Educational Technologies and Learning: Empowering Teachers to Teach and Students to Learn in the Information Age*. Charles C Thomas Publisher; 2004
- [51] Hvorecký J, Šimúth J, Lipovská A. Ways of delivering tacit knowledge in e-learning. In: *2015 International*

Conference on Interactive Collaborative Learning (ICL). IEEE; 2015. pp. 523-526

[52] O'Malley C, Vavoula G, Glew JP, Taylor J, Sharples M, Lefrere P, et al. Guidelines for Learning/Teaching/Tutoring in a Mobile Environment. 2005

[53] Yilmaz R. Knowledge sharing behaviors in e-learning community: Exploring the role of academic self-efficacy and sense of community. *Computers in Human Behavior*. 2016;**63**:373-382

[54] Setia P, Setia P, Venkatesh V, Joglekar S. Leveraging digital technologies: How information quality leads to localized capabilities and customer service performance. *MIS Quarterly*. 2013:565-590

IntechOpen

IntechOpen Series
Artificial Intelligence, Volume 36

Emerging Technologies in Computational Cognitive Neuroscience

Edited by Vijayalakshmi Kakulapati



