Jurnal ELTIKOM :
Jurnal Teknik Elektro, Teknologi Informasi dan Komputer

Vol. 10, No. 1, June 2026, page. 24-36
ISSN 2598-3245 (Print), ISSN 2598-3288 (Online)
DOI: http://doi.org/10.31961/eltikom.v10i1.2098
Available online at http://eltikom.poliban.ac.id

DETECTING RESEARCH EVOLUTION AND TRENDS IN THE COMPUTER VISION DOMAIN USING TOPIC MODELING AND LARGE LANGUAGE MODELS

Setio Basuki[1*], Zamah Sari[1], Rizky Indrabayu[1], Masatoshi Tsuchiya[2]
[1)] Informatics Engineering, Universitas Muhammadiyah Malang, Malang, Indonesia
[2)] Department of Computer Science and Engineering, Toyohashi University of Technology, Toyohashi, Japan
e-mail: setio_basuki@umm.ac.id, zamahsari@umm.ac.id, bayyy17@webmail.umm.ac.id, tsuchiya@is.cs.tut.ac.jp

Received: 10 December 2025 – Revised: 15 March 2026 – Accepted: 16 March 2026

ABSTRACT
Research evolution and trends in computer vision (CV) are important for understanding the field's landscape. Trends show which topics are gaining attention, while evolution reveals how those topics change over time. Understanding both helps researchers gain insight into CV and anticipate emerging areas of focus. However, the rapid growth of publications makes such detection challenging. This paper aims to detect research evolution and trends in CV using topic modeling (TM) and large language model (LLM) techniques. The study applies TM and LLM approaches to papers from leading CV conferences, Computer Vision and Pattern Recognition (CVPR), International Conference on Computer Vision (ICCV), and Winter Conference on Applications of Computer Vision (WACV), published between 2013 and 2023, totaling more than 21,000 papers, using only abstracts and titles. The TM methods used are Latent Dirichlet Allocation (LDA) and Bidirectional Encoder Representations from Transformers for Topic Modeling (BERTopic), which generate keywords that represent topics. LLMs then refine these topics to support better analysis. The results show that research evolution and trends are easier to identify from abstracts than from titles, with BERTopic outperforming LDA in internal validity based on coherence metrics and external validity based on human judgment. In addition, the topics evolved from traditional image processing tasks in earlier years to a stronger focus on deep learning and, more recently, generative approaches. Integrating TM techniques with LLMs enhances the detection of evolving research themes and trends in CV. This approach provides a clearer understanding of the field's development and helps anticipate future directions.

Keywords: *BERTopic, computer vision, research evolution, large language model, topic modeling.*

I. INTRODUCTION

UNDERSTANDING the evolution, trends, and research opportunities in a field can provide a deeper understanding of the state of the art (SOTA), which is highly beneficial for peer review stakeholders [1]. For researchers, understanding SOTA helps determine research directions, identify gaps, and find the most appropriate current methods [2]. For reviewers, this knowledge is useful for assessing novelty, evaluating contributions offered, and providing feedback aligned with the latest developments [3]. For editors, SOTA helps maintain the reputation and scope of journals and attract more high-quality submissions (citation). The urgency of effective peer review through a deep understanding of SOTA can be addressed by a smart system for detecting research evolution and trends. Specifically, such a system should be able to identify evolution, trends, and research development opportunities, which can be categorized into four sub-questions: (i) popular topics, (ii) less favored topics, (iii) topics with stable popularity, and (iv) research development opportunities. Furthermore, the system needs to provide an interactive visualization of the detected trends so that users can understand them more easily.

At present, several studies have aimed to develop systems for detecting research evolution and trends. These studies use the topic modeling (TM) approach to identify topics that represent the development of a scientific field over time. This approach has been applied in various fields, such as informatics and other engineering areas [4],

[5], social sciences [6], [7], economics [8], [9], and medicine [10], [11]. Specifically in the CV-related field, trend detection has been applied in two studies: one [12] employed a dynamic topic modeling approach on large-scale CV papers, while another [13] implemented Latent Dirichlet Allocation (LDA) on papers related to the application of deep learning in the CV domain. Literature reviews indicate several research gaps, particularly (i) the absence of studies on the evolution of research topics focusing solely on computer vision, (ii) the lack of studies using papers from the most prestigious and specialized conferences in computer vision, (iii) the tendency of existing studies to focus on keyword extraction methods without capturing more recent approaches such as Large Language Models (LLMs), which limits interpretability, and (iv) the absence of studies comparing topic validity based on paper titles and abstracts. Together, these issues contribute to a limited and less comprehensive understanding of the evolution of topics in computer vision research.

This research aims to develop an automated model for detecting research evolution and trends in CV. Trends indicate the popularity of research topics, while evolution highlights changes in those trends over time. CV was selected because of its scientific prominence, broad range of applications, strong market demand, and substantial investment. The foundation for detecting the evolution of research topics is a collection of papers from three prestigious conferences in the field of computer vision: Computer Vision and Pattern Recognition (CVPR), International Conference on Computer Vision (ICCV), and Winter Conference on Applications of Computer Vision (WACV). Topics are extracted from the collection of papers by focusing on titles and abstracts, as these sections provide concise summaries of the papers' overall content. This study employs Latent Dirichlet Allocation (LDA) and Bidirectional Encoder Representations from Transformers for Topic Modeling (BERTopic) [14]. Furthermore, this paper leverages a recent technology, the Large Language Model, to refine the generated topics. Evaluation is conducted based on internal validity using coherence metrics and external validity through human judgment.

More specifically, this paper reports several findings: (i) BERTopic outperformed LDA on two evaluation criteria, namely coherence scores and manual or human judgment, based on the generated keywords, (ii) LDA was ineffective in detecting research evolution and trends because it mostly produced generic machine learning terms rather than CV-specific terms, (iii) BERTopic was effective in identifying research evolution and trends by producing creative, in-depth, and CV-specific keywords, (iv) the use of abstracts as the dataset provided more informative keywords representing evolution and trends than titles, (v) topics evolved over time, with early years focusing on typical image processing tasks, middle years adopting deep learning, and recent years emphasizing generative approaches, and (vi) the use of LLM helps refine the generated topics and makes them more readable.

II. Related Works

This section reviews existing work on detecting research trends in CV and other areas of computer science. The review focuses specifically on trend detection in the development of research areas themselves, rather than on their applications in other domains. For example, instead of examining trends in the use of artificial intelligence within the business domain, we focus on understanding how artificial intelligence has evolved over time as a field. Table 1 presents studies on research trend detection in computer science-related domains.

Several researchers have applied various topic modeling techniques to identify trends across different research areas in computer science. One of the most widely used methods is Latent Dirichlet Allocation (LDA). For instance, a study by [15] used LDA to explore trends in deep learning, identifying its core research themes and their evolution. Research by [13] applied LDA specifically to deep learning for computer vision, providing insights into how computer vision has advanced within the broader deep learning context. The study of [16] employed LDA in cybersecurity, revealing popular topics and emerging areas in that domain. Following this, studies by [17], [22] extended the application of trend detection in computer science. The usefulness of LDA is also reflected in studies focusing on specific domains, such as blockchain by [23], IoT by [20], [27], information security by [26], and gamification by [28]. Furthermore, the study by [21] used LDA to explore trends in software development, providing a broader view of evolving software engineering practices. The paper by [18] employed a structural topic model to identify trends in IoT security, offering context-sensitive insight into the challenges and advances within this domain. Similarly, the paper by [16] on the use of LDA in cybersecurity highlights its effectiveness in uncovering key research directions.

More advanced methods have emerged to capture the complexity of evolving research fields. For instance, dynamic topic modeling (DTM) was used by [24] to analyze computing scientific literature, enabling the study of temporal changes in academic topics. Similarly, the paper by [12] applied DTM to understand research progression within computer vision, providing valuable insight into its temporal dynamics. In addition, BERTopic, applied by [25] in the language model domain, has shown potential

TABLE 1
RELATED WORKS

Year	Research	Focused Task	Method
2019	[15]	Deep Learning	Latent Dirichlet Allocation
2019	[13]	Deep Learning for Computer Vision	Latent Dirichlet Allocation
2020	[16]	Cybersecurity	Latent Dirichlet Allocation
2020	[17]	Computer Science	Latent Dirichlet Allocation
2021	[18]	IoT Security	Structural Topic Model
2022	[19]	Sentiment Analysis	Latent Dirichlet Allocation
2022	[20]	IoT	Latent Dirichlet Allocation
2022	[21]	Software Development	Latent Dirichlet Allocation
2022	[22]	Computer Science	Graph based topic modelling
2022	[23]	Blockchain	Latent Dirichlet Allocation
2023	[24]	Computing Scientific Literature	Dynamic Topic Modeling
2023	[25]	Language Mode	BERTopic
2023	[26]	Information Security	Latent Dirichlet Allocation
2023	[27]	IoT	Latent Dirichlet Allocation
2023	[28]	Gamification	Latent Dirichlet Allocation
2023	[29]	Mobile Robot	Latent Dirichlet Allocation
2023	[12]	Computer Vision	Dynamic Topic Modeling
2024	[30]	Metaverse	Latent Dirichlet Allocation
2024	[31]	Academic Text	Large Language Models using SOLAR 10.7B

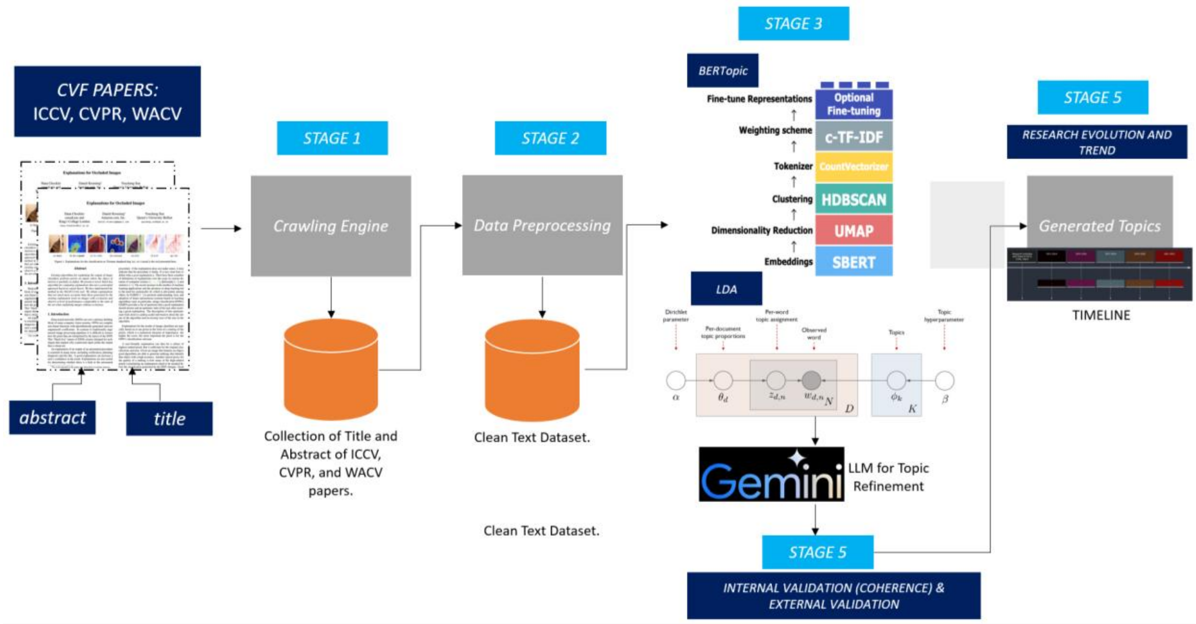


Figure 1. Proposed System for Research Trend Detection on CFV Conference Papers.

in handling nuanced and context-aware trend analyses. Its ability to extract semantically rich themes offers a significant advantage over traditional methods in detecting emerging topics. The rise of large language models (LLMs) has also introduced transformative approaches to trend detection. The paper by [31] used the SOLAR 10.7B LLM to analyze academic texts, enabling the discovery of complex and emerging patterns in academic writing. These models represent a significant advance over traditional topic modeling techniques.

III. RESEARCH METHOD

This paper develops a smart system to detect research evolution and trends in computer vision (CV) using topic modeling and a large language model. We use paper collections from three leading conferences under the Computer Vision Foundation (CVF): CVPR, ICCV, and WACV, from 2013 to 2023. The topic detection system analyzes the titles and abstracts of papers from each conference. We used two approaches to better understand the generated topics: (i) Bidirectional Encoder Representations from Transformers for Topic Modeling (BERTopic) [14] and (ii) Latent Dirichlet Allocation (LDA) [32]. The LLM, Gemini [33], is used to refine the generated topics for better visualization.

As shown in Figure 1, the proposed system consists of several stages. First, this paper collects titles and abstracts from ICCV, CVPR, and WACV. Second, data preprocessing is performed to remove noise

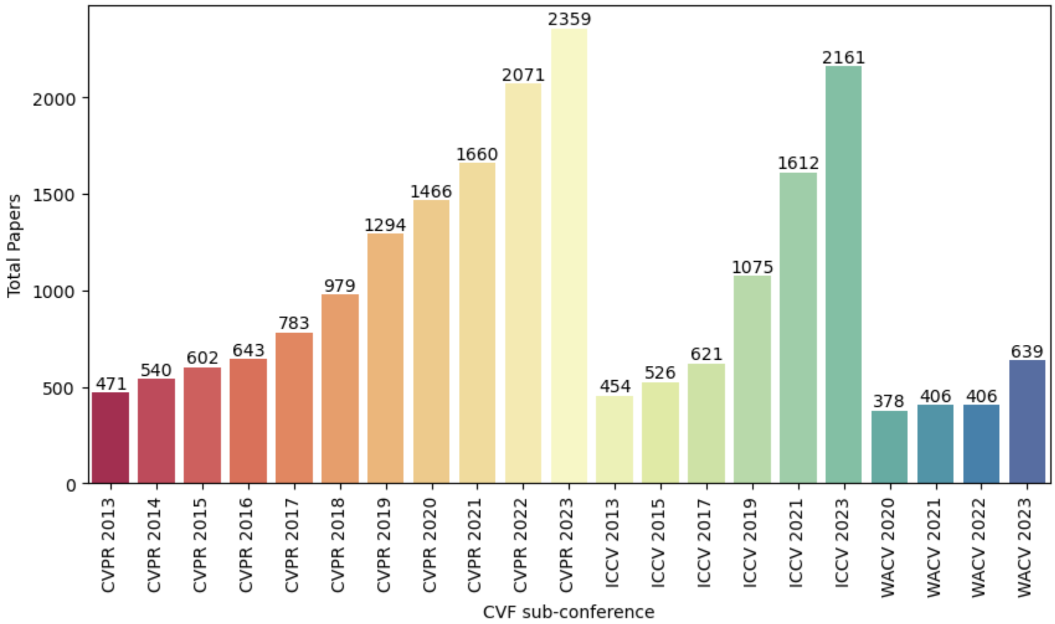


Figure 2. The Distribution of Computer Vision Foundation (CVF) Papers.

and fill in any missing content. Third, the main process detects research evolution and trends using TM and LLM. Finally, the analysis is conducted based on the generated keywords representing the topics through internal validation (coherence) and manual or human judgment.

A. Dataset

Both paper titles and abstracts are collected using crawling techniques from the CVF website. From the three conferences, approximately 21,146 papers were obtained, consisting of 21,146 titles and 21,146 abstracts. More specifically, the number of papers from each conference is as follows: CVPR, 12,868 papers; ICCV, 6,449 papers; and WACV, 1,829 papers. CVPR focuses on the use of machine learning for image and video understanding, while ICCV emphasizes three main areas: pattern recognition, object recognition, and image segmentation. WACV, in turn, focuses on adversarial learning, adversarial attacks, and generative models in images, videos, and 3D applications. The number of papers available for each conference varies, and these papers are openly accessible on the CVF website.

Next, data preprocessing is performed. The paper data obtained through the automatic scraping process contains formatting and cleanliness issues. At this stage, we clean tags such as HTML attributes ('\n', '\t', '\r'), remove stopwords ('in', 'on', 'at', etc.), and perform stemming and lemmatization to reduce words to their base forms ('equations' to 'equation', 'modules' to 'module', etc.). Through preprocessing, we aim to optimize the LDA and BERTopic algorithms in detecting the list of keyword topics discussed in the paper titles and abstracts. The paper distribution is shown in Figure 2.

B. Latent Dirichlet Allocation (LDA)

Latent Dirichlet Allocation (LDA) is a generative probabilistic model used for topic modeling, introduced in 2003, in which documents are treated as mixtures of topics, with each topic defined by a distribution over words. The main parameters of LDA include α (alpha), β (beta), θ (theta), ϕ (phi), w, and z. The α (alpha) parameter regulates the probability that multiple topics will appear in a document; higher α values produce a wider variety of topics, whereas lower values produce fewer, more dominant topics. The β (beta) parameter influences the lexical diversity of themes, with higher values producing broader vocabulary and lower values resulting in more specialized topics. The θ (theta) parameter denotes the distribution of topics within a document, indicating the probability of each topic's occurrence, and it is derived from a Dirichlet distribution characterized by α. The ϕ (phi) parameter denotes the distribution of words for each topic, indicating the probability of each word within that topic, parameterized by β. The w component denotes the actual observed words in the documents, which are assumed to be generated from topics according to the distribution ϕ. Finally, z represents the latent assignment for each word, indicating the topic that produced that term. These assignments originate from a multinomial distribution derived from the document's topic distribution, θ. Overall, these elements enable LDA to represent documents as combinations of topics, with each topic characterized by a distribution of words, thereby capturing complex relationships among documents, topics, and words.

C. BERTopic

BERTopic is a topic modeling technique that uses embedding technology (BERT) and class-based Term Frequency-Inverse Document Frequency (c-TF-IDF) to produce topic clusters. The use of word embeddings offers substantial advantages, including improved interpretability in topic identification, stronger semantic understanding of the generated keywords that represent topics, and greater coherence among keywords. The process begins with an input text corpus, which BERTopic converts into embedding representations through two steps: tokenization and embedding creation. This phase produces an embedding representation that captures the semantic relationships within the text. The next phase involves the use of Uniform Manifold Approximation and Projection (UMAP), which is a dimensionality reduction method [34]. This reduction is important because it preserves semantic relationships while reducing computational complexity. The clustering stage uses Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) [35], which produces topic clusters. The clustering results are then processed by c-TF-IDF to identify the most representative words or phrases in each cluster. The number of topics generated by BERTopic is variable because the HDBSCAN algorithm determines the most suitable number of topic clusters.

D. Gemini for Topic Refinement

Gemini 1.5, developed by Google, is an advanced multimodal large language model (LLM) capable of processing text, audio, video, and many other data types, while handling contexts of up to 10 million tokens [33]. Its foundation is based on a Sparse Mixture-of-Experts (MoE) architecture that selectively activates relevant subsets of model parameters for each task, enabling efficient and scalable computation. This architecture, together with training on diverse datasets and instruction tuning through human input, enables Gemini to produce coherent outputs across various tasks, including long-document analysis, video understanding, and low-resource language translation. A notable characteristic is its capacity for in-context learning, which allows it to adapt to new tasks during inference without explicit retraining. Using the Gemini API, we performed topic representation tasks with optimized LLM sampling techniques. We set the temperature to 0.8 to enhance nuance and creativity, while Top-P (0.95) and Top-K (64) limited the model to the most likely token probabilities. This combination ensured high-quality, coherent outputs, enabling precise, interpretable topic representations. Additionally, we used the "gemini-1.5-pro-latest" model. The Python program asks Gemini to generate short topic labels based on keywords using the following prompt: "Based on these [KEYWORDS], determine a concise and relevant label that accurately represents the overall topic in 5 words or fewer"

E. Experiment Scenario

The experiment to detect research evolution and trends was conducted using two scenarios. The first scenario applied detection to the title dataset, followed by the second scenario, which applied detection to the abstract dataset. Both scenarios used a cleaned dataset produced through preprocessing and transformation. The research trend detection experiment on the title and abstract dataset was conducted using LDA as the baseline method and BERTopic as the contemporary method. In each scenario, the generated keywords representing topics were evaluated through internal validation using coherence and external validation based on manual or human judgment [36]. For each manual evaluation scenario (1 CV researcher or expert), the assessment was based on key elements of research evolution and trends, such as the produced terms (generic or specific), CV-related terms (high or low), the depth of CV terms (deep or superficial), variation in CV-related terms (high or low), terms related to CV applications (high or low), the ability to capture recent updates (effective or ineffective), and the ability to detect research trends over the years (effective or ineffective). Following this, an analysis comparing the effectiveness of titles and abstracts was conducted. Note that LDA was designed to produce between 1 and 10 topics, while BERTopic was designed to generate a maximum of 10 topics. Finally, the generated topics were refined using the LLM Gemini for better interpretation. This refinement stage was intended to make the generated keywords more readable for humans by improving presentation and insight.

TABLE 2
COMPARISON OF KEYWORDS PRODUCED BY BERTOPIC AND LDA WHEN DETECTING RESEARCH TRENDS USING TITLES

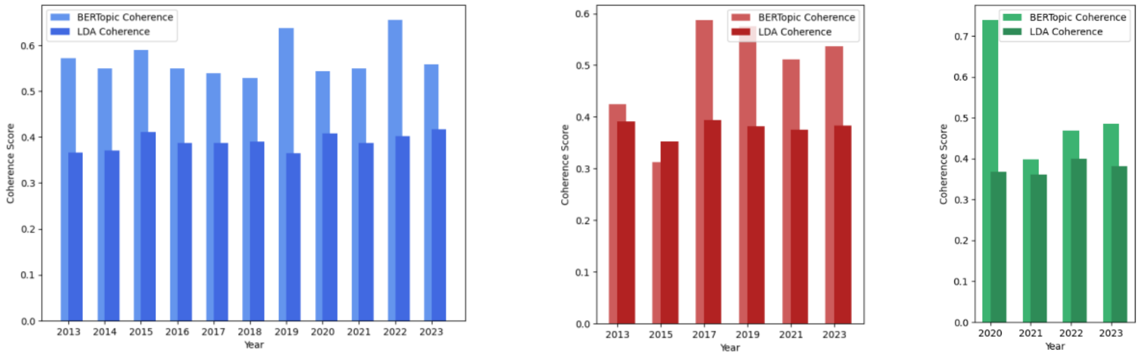
Conferences	BERTopic		LDA	
	Number of topics	Keywords	Number of topics	Keywords
CVPR 2013	1	tracking, object, visual, multiple, target, multi, pursuit, latent, pose, real	6	object, model, image, approach, detection, dataset, propose, method, paper, state
CVPR 2014	1	tracking, online, target, object, visual, multi, occlusion, robust, multiple, motion	8	method, image, model, state, proposed, approach, based, art, propose, data
CVPR 2015	2	tracking, target, tracker, multi, correlation, adaptive, foreground, reliable, object, multiple	9	image, feature, object, approach, state, art, video, dataset, descriptor, representation
CVPR 2016	7	network, neural, convolutional, deep, architecture, using, learning, pooling, vision, random	2	action, image, object, video, model, dataset, method, approach, frame, network
CVPR 2017	10	flow, optical, stereo, camera, based, slow, fast, interpolation, correspondence, accurate	9	image, network, data, method, object, dataset, learning, video, feature, cnn
CVPR 2018	5	visual, question, answering, attention, language, reasoning, generation, dialog, learning, grounding	5	image, model, approach, method, art, training, based, state, dataset, using
CVPR 2019	10	navigation, language, path, trajectory, vision, imitation, prediction, agent, lstm, visual	8	model, network, method, image, task, learning, propose, video, based, approach
CVPR 2020	10	reinforcement, learning, driving, simulation, embodied, environment, agent, meta, interactive, free	2	method, image, model, network, dataset, task, object, state, art, approach
CVPR 2021	8	action, temporal, recognition, activity, supervised, localization, weakly, video, spatio, actor	2	image, method, model, task, propose, supervised, network, training, approach, learning
CVPR 2022	10	federated, learning, heterogeneous, privacy, data, client, iid, heterogeneity, local, personalized	9	model, task, transformer, learning, video, image, vision, object, attention, training
CVPR 2023	7	stereo, view, depth, multi, estimation, disparity, ray, consistent, using, gated	5	training, model, image, task, pre, learning, video, method, dataset, text
ICCV 2013	1	action, recognition, event, video, complex, activity, detection, learning, interaction, feature	3	image, method, model, problem, data, propose, proposed, approach, algorithm, object
ICCV 2015	1	tracking, visual, online, multiple, filter, object, learning, multi, weighted, spatially	9	method, image, model, object, based, problem, paper, propose, approach, data
ICCV 2017	3	pose, estimation, human, hand, monocular, rgb, object, real, learning, robust	8	image, model, approach, state, network, task, object, art, end, feature
ICCV 2019	9	architecture, search, distillation, neural, network, knowledge, training, sequential, data, product	10	method, image, feature, network, dataset, state, learning, art, model, task
ICCV 2021	4	pose, human, estimation, hand, mesh, shape, motion, body, regression, using	3	image, learning, method, model, based, network, dataset, domain, data, representation
ICCV 2023	10	federated, learning, communication, global, personalized, vertical, client, personalization, iid, overlapping	9	model, task, training, learning, data, method, performance, image, propose, class
WACV 2020	7	retrieval, image, answering, question, content, text, exploring, search, modal, visual	7	image, network, method, detection, object, approach, learning, using, dataset, feature
WACV 2021	4	retrieval, text, visual, image, graph, identification, learning, teacher, student, table	6	method, dataset, network, class, model, object, image, feature, state, art
WACV 2022	2	video, action, temporal, supervised, recognition, segmentation, self, multi, audio, learning	1	image, method, model, learning, domain, approach, data, task, propose, training
WACV 2023	5	action, temporal, recognition, transformer, skeleton, network, spatio, contrastive, based, multi	8	method, feature, based, model, training, image, state, attention, art, dataset

IV. RESULTS

This section presents the analysis of the generated topics and their corresponding keywords produced by LDA and BERTopic. To enable a comparative analysis between titles and abstracts, this paper divides the conference period into three categories: early years (2013-2015), middle years (2016-2019), and recent years (2020-2023). This division allows a clearer understanding of how the research focus has shifted and developed over time.

A. Topic Detection Based on Title

Detecting research trends based on titles reveals several notable patterns. The effectiveness of BERTopic over LDA is also shown in Figure 3 by comparing the highest coherence scores produced by both methods. BERTopic demonstrates superiority in coherence scores over LDA across all conferences. This indicates that BERTopic performs better in generating more coherent and relevant topics across various CVF conferences. Furthermore, BERTopic's superiority can be seen in Figure 4, which illustrates the coherence scores obtained for topic numbers ranging from 1 to 10. However, in WACV,



(a) (b) (c)

Figure 3. Comparison of the Highest Coherence for all Conference Categories Based on (a) CVPR; (b) ICCV; (c) WACV

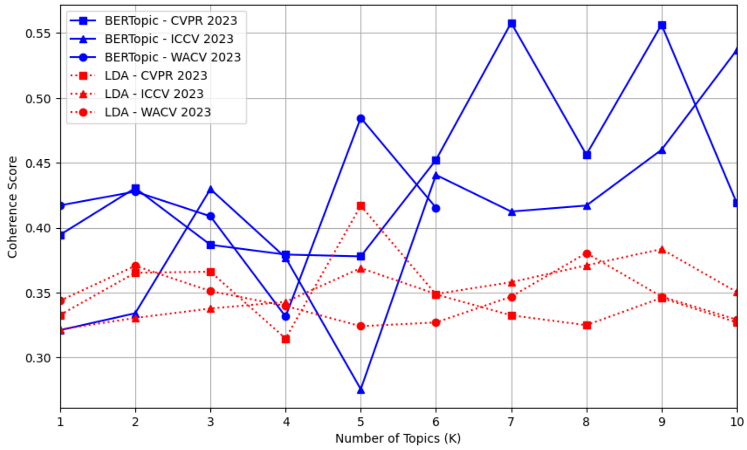
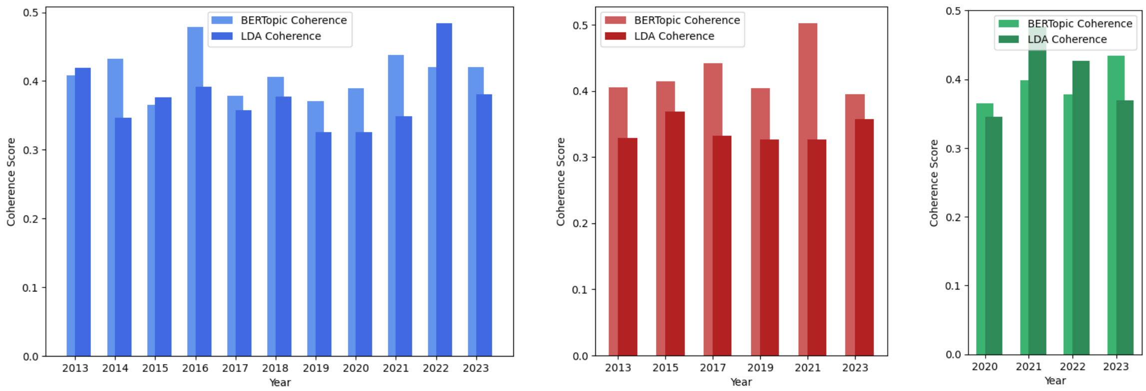


Figure 4. Comparison of Coherence Scores Between BERTopic and LDA Based on Titles.



(a) (b) (c)

Figure 5. Comparison of the Highest Coherence for All Conference Categories Based on (a) CVPR; (b) ICCV; (c) WACV

BERTopic stops generating topics when the number reaches 6 because of its reliance on vocabulary size and clustering methods. Specifically, the clustering algorithm (HDBSCAN) determined that 6 was the optimal number of dense clusters based on the document embeddings and the limited vocabulary size, as mentioned earlier.

Turning to the topics generated by both methods, the keywords produced by LDA tend to be more general, whereas BERTopic produces more specific keywords. For instance, BERTopic presents application-specific terms such as “multi-object tracking" through keywords like “multi”, “object”, and “tracking”, and “action recognition" through keywords like “action" and “recognition”. In contrast, the keywords produced by LDA tend to reflect broad CV terms such as “algorithm”, “method”, and “dataset”. Furthermore, the LDA keywords appear to cover a wider area of machine learning rather than focusing specifically on the CV domain. As a result, the evolution of topics in CVF papers can be recognized more easily through the keywords produced by BERTopic. Table 2 presents the topic detection

TABLE 3
COMPARISON OF KEYWORDS PRODUCED BY BERTOPIC AND LDA WHEN DETECTING RESEARCH TRENDS USING ABSTRACTS

Conferences	BERTopic Number of topics	BERTopic Keywords	LDA Number of topics	LDA Keywords
CVPR 2013	7	blur, image, noise, kernel, denoising, patch, deblurring, method, quality, deconvolution	1	model, learning, recognition, graph, human, video, based, activity, multi, cloud
CVPR 2014	8	image, blur, deblurring, method, denoising, algorithm, gradient, result, proposed, problem	3	image, reconstruction, object, scene, learning, domain, via, recognition, video, tracking
CVPR 2015	7	saliency, salient, dataset, fixation, object, image, visual, region, gaze, data	9	image, learning, visual, rank, video, metric, estimation, based, dataset, object
CVPR 2016	7	question, image, visual, dataset, video, language, sentence, description, text, answer	10	network, image, multi, object, visual, learning, neural, sparse, deep, convolutional
CVPR 2017	2	object, segmentation, semantic, detection, network, approach, instance, task, image, model	8	deep, learning, network, detection, object, video, multi, recognition, feature, visual
CVPR 2018	7	image, network, method, rain, quality, resolution, end, based, proposed, state	3	image, learning, using, depth, domain, single, network, estimation, adaptation, deep
CVPR 2019	7	image, method, blur, network, resolution, real, super, model, deep, result	5	learning, network, image, neural, convolutional, recognition, deep, graph, scene, feature
CVPR 2020	8	reflection, image, scene, mirror, light, transparent, method, surface, illuminant, normal	2	image, network, learning, detection, object, attention, feature, graph, adaptive, via
CVPR 2021	9	forgetting, learning, incremental, task, continual, class, new, catastrophic, data, old	9	learning, representation, image, supervised, video, action, via, self, recognition, point
CVPR 2022	9	client, federated, fl, local, model, data, learning, gradient, privacy, global	9	deepfake, image, light, generalization, implicit, animation, cloning, affine, vanishing, field
CVPR 2023	9	client, federated, fl, data, learning, model, local, global, privacy, server	3	image, learning, supervised, visual, detection, sketch, text, data, training, pre
ICCV 2013	5	saliency, gaze, salient, region, model, eye, image, method, object, detection	6	recognition, action, image, occlusion, estimation, prior, human, segmentation, pose, structure
ICCV 2015	5	video, concept, image, description, visual, sentence, language, phrase, word, text	1	learning, detection, image, object, deep, semantic, segmentation, video, model, shape
ICCV 2017	5	person, tracking, object, feature, identification, video, method, tracker, id, frame	7	object, detection, network, video, using, based, convolutional, image, deep, learning
ICCV 2019	8	face, facial, landmark, expression, image, structure, method, model, identity, loss	8	image, segmentation, network, video, object, learning, deep, attention, detection, shot
ICCV 2021	7	search, quantization, architecture, nas, neural, precision, bit, space, network, accuracy	2	learning, supervised, segmentation, point, cloud, image, object, detection, representation, multi
ICCV 2023	10	matrix, solver, camera, problem, bundle, flag, point, fundamental, adjustment, minimal	6	detection, object, learning, supervised, via, semi, training, dataset, e, transformer
WACV 2020	2	image, model, face, gan, method, network, approach, training, generative, propose	9	detection, object, visual, multi, scale, network, self, approach, tracking, detector
WACV 2021	2	point, method, object, camera, depth, network, map, cloud, visual, estimation	9	network, prediction, image, application, video, uncertainty, learning, fast, convolutional, neural
WACV 2022	5	detection, based, stereo, method, object, image, map, view, lane, fusion	3	supervised, shot, self, multi, semantic, learning, segmentation, video, image, using
WACV 2023	9	attack, adversarial, defense, certified, based, perturbation, domain, sample, model, training	6	detection, segmentation, learning, video, object, image, semantic, training, via, supervised

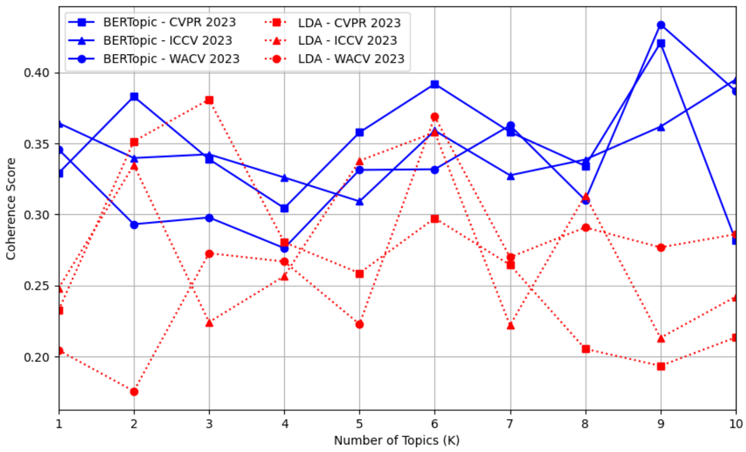


Figure 6. Comparison of the Coherence Scores Between BERTopic and LDA Based on Abstracts.

results based on titles. Note that the topics presented in the table are represented by the 10 keywords with the highest coherence scores.

In the early years, CVPR research focused on topics related to “object tracking and detection”, represented by keywords such as “tracking”, “object”, “visual”, and “multiple”. After that, the use of deep learning in computer vision began to emerge around 2016, indicated by terms such as “network”, “neural”, “convolutional”, “pooling”, and “deep”. In recent CVPR papers, the focus has shifted to more advanced learning techniques, such as reinforcement learning, reflected in keywords like “reinforcement” and “learning”, and attention mechanisms, highlighted by terms such as “attention” and “transformers”. Additionally, there is a stronger focus on privacy and data security, emphasized by terms such as “federated” and “privacy”.

From 2013 to 2015, ICCV papers showed terms similar to those found in CVPR, such as action recognition and object detection, through keywords like “action”, “recognition”, “object”, and “detection”. Interestingly, starting in 2017, BERTopic captures a trend in computer vision applications related to humans, indicated by keywords such as “object”, “pose”, “hand”, and “human”. In recent years, there has been a stronger focus on neural network architectures, supported by terms such as “neural”, “network”, “knowledge”, “distillation”, and “learning”. Similar to CVPR, recent ICCV conferences also include terms like “federated” and “learning”, which point to a shared focus on privacy and data security.

As the most recent conference, WACV shows more advanced topics than CVPR and ICCV. There is a strong emphasis on question answering, supported by terms like “question” and “answering”, which indicates the intersection between computer vision and natural language processing. This topic is further supported by keywords such as “retrieval”, “text”, and “visual”, which suggest work on multimodal learning and cross-domain applications. In recent years, there has also been a resurgence of action recognition, which was prominent in the early years of ICCV. This resurgence is marked by terms like “action” and “recognition”, along with the use of advanced models such as transformers, supported by keywords such as “transformer” and “attention”.

B. Topic Detection Based on Abstract

Experiments on detecting research trends using abstracts reveal several important findings. As shown in Figure 5, the coherence scores produced by BERTopic are generally higher than those produced by LDA across all conferences. Similar to the experiment conducted using titles, BERTopic performs better in generating coherent topics. Figure 6 further supports the superiority of BERTopic over LDA by showing changes in the coherence scores of both methods across 1 to 10 topics.

Table 3 presents the results of topic detection based on abstracts. It is worth noting that the quality of the generated keywords representing topics in LDA improves compared with those produced using titles. Although both algorithms demonstrate strengths in generating creative and domain-specific keywords, the keywords produced by BERTopic appear more creative. While LDA generated typical CV-specific terms such as “image”, “object”, “convolutional”, and “segmentation”, BERTopic generated not only CV-specific but also deeper and more contemporary keywords such as “catastrophic”, “illuminant”, “adversarial”, and “federated”.

The keyword analysis of CVPR using both methods shows a clear evolution in research focus over the years. In the early years (2013-2015), there was a strong emphasis on image processing to improve data quality, reflected in terms like “image”, “quality”, “deblurring”, “denoising”, “saliency”, and “reconstruction”. In the middle years (2016-2019), the focus shifted to more complex tasks such as question answering and image captioning, indicated by keywords like “image”, “text”, “question”, “answer”, and “description”. Additionally, the use of deep learning began to emerge, along with strong interest in image segmentation and object detection. In the most recent years (2020-2023), notable differences appeared between BERTopic and LDA. BERTopic tends to generate more advanced computer vision applications such as catastrophic forgetting, deepfake detection, and data privacy, indicated by keywords like “catastrophic”, “forgetting”, “deepfake”, “federated”, “learning”, and “fl”. In contrast, LDA still produces more common computer vision terms such as “segmentation”, “video”, and “image".

A closer look at ICCV in the early years using both LDA and BERTopic shows that the keywords were dominated by standard computer vision terms, such as “region”, “recognition”, and “visual”. From 2017 to 2019, BERTopic captured more discussion of human and object tracking, while LDA remained focused on generic computer vision terms with an emphasis on deep learning. In recent years (2021-2023), both methods expanded their focus to algorithm improvement, including neural architecture search (NAS) and transformers, supported by terms like “quantization”, “neural”, “architecture”, “search”, “NAS”, and “transformer”.

For WACV, themes such as object detection and image segmentation have remained consistently popular each year based on both methods. This is captured by LDA through keywords such as “object”, "detection", "image", and "segmentation". In the early years, there was a notable emphasis on Generative Adversarial Networks (GANs), indicated by BERTopic keywords like "generative" and "GAN". In recent years, the focus has shifted to adversarial robustness, reflected in keywords such as "attack", "adversarial", and "defense" produced by BERTopic, while LDA remains more constant with typical computer vision terms.

V. DISCUSSION

A. Analysis Comparing Title- and Abstract-Based Research Trend Detection

This paper proposes several elements to assess the keywords generated by LDA and BERTopic. Table 4 compares the keywords representing topics produced by both methods. In general, BERTopic's superiority in achieving higher coherence scores is consistent with its effectiveness in capturing research trends, as measured by several indicators. When applied to titles, LDA is ineffective. When applied to abstracts, LDA becomes more competitive by producing computer vision-related terms, although these terms remain at a surface level and show limited variation. Compared with the results generated from titles, LDA performs more effectively when combined with the abstract dataset. It is worth noting that BERTopic leverages clustering algorithms and embedding models, which allow a better understanding of vocabulary data or corpora than traditional methods. By contrast, LDA relies on probabilistic calculations, which allow it to run faster while still offering useful insight. Overall, BERTopic consistently shows superior performance compared with LDA, indicating higher-quality generated topics and greater precision, with fewer overly common keywords.

B. LLM-based Topic Representation

This section discusses how the LLM approach, particularly Gemini, is used for topic representation. Specifically, the LLM is to refine the generated keywords so they are easier to understand, provide clearer topic insight, and offer better presentation, rather than relying only on the keyword lists produced by LDA and BERTopic. Note that the use of the LLM is intended as a supplementary step for the generated keywords. It is preferable to visualize the topic produced by BERTopic, as LDA is less effective for detecting research evolution and trends (see Table 4).

The refined topics generated from titles are shown in Figure 7, while those generated from abstracts are shown in Figure 8. It can be seen that the same conference, in the same period, produces different refined topics. This difference is caused by the variation in the keywords generated from titles and abstracts, which Gemini then uses to propose refined topics. This difference is expected because titles are typically written in 10 to 20 words, whereas abstracts may contain up to 250 words. This variety of refined topics enriches the understanding of research evolution and trends in the CV domain. Even so, it still captures the general trend for each period. For instance, in the early years (2013-2015), the topics emphasized typical image processing tasks such as object detection and visual tracking. In the middle years (2016-2019), deep learning terms became more dominant, including reinforcement learning, deep image, and visual-text processing. After that, the recent trend (2020-2023) across all conferences shows a similar pattern centered on generative approach-related terms such as "transformer", "adversarial", and "neural", as well as machine learning security and privacy, including "federated learning", "defense", and "personalized". Because these trends are consistent with the analysis based on the generated keywords in the previous section, the use of the LLM makes the topics more human-readable.

C. Topic Detection in Previous Work

At present, several studies have applied various topic modeling techniques to identify trends across different research areas in computer science. Among these studies, only two [12], [13] focused on trend analysis in computer vision. These reviews indicate research gaps, particularly (i) the limited number of studies on the evolution of research topics focusing solely on computer vision, (ii) the lack of studies using papers from the most prestigious conferences in computer vision, (iii) the limited use of more recent approaches such as LLMs, which reduces interpretability, and (iv) the absence of studies comparing topic validity based on paper titles and abstracts. Together, these issues contribute to a limited and less comprehensive understanding of the evolution of topics in computer vision research. In addition, only a small number of existing studies address the temporal dimension of evolving research topics, which makes it difficult to gain insight into research evolution. Therefore, this paper offers a potential way to address these issues.

TABLE 4
COMPARISON OF TOPIC DETECTION ON TITLE AND ABSTRACT

Topic Modeling Methods	Research trend indicators	LDA	BERTopic
Title	Produced terms	Generic machine learning	CV-specific
	CV-related terms	Low	High
	The depth of CV terms	Surface	Depth
	CV-related terms variation	Low	High
	Terms-related to CV application	Low	High
	Capturing recent update	Ineffective	Effective
	Detecting research trend over years	Ineffective	Effective
Abstract	Produced terms	CV-specific	CV-specific
	CV-related terms	High	High
	The depth of CV terms	Surface	Depth
	CV-related terms variation	Low	High
	Terms-related to CV application	Low	High
	Capturing recent update	Quite Effective	Effective
	Detecting research trend over years	Quite Effective	Effective

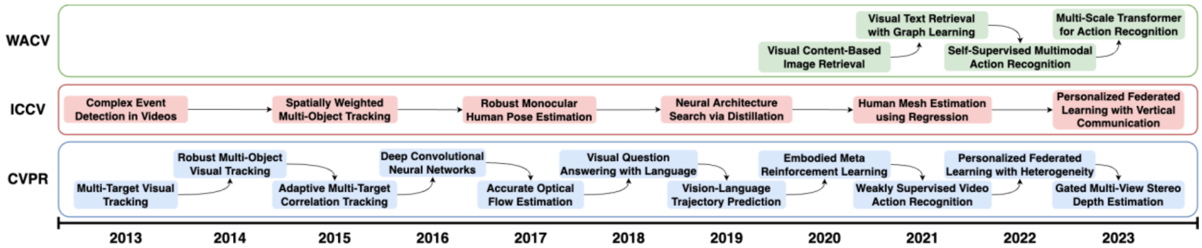


Figure 7. Refined Topics Representation Based on Title.

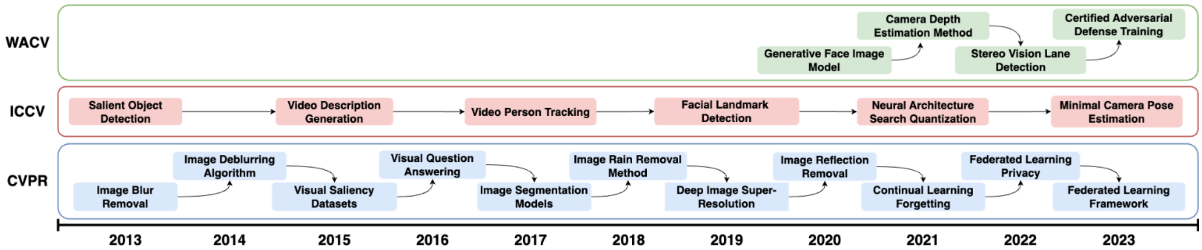


Figure 8. Refined Topics Representation Based on Abstract.

D. Dataset Research Limitation

Although the LLM in this paper successfully provides better representations of generated topics, several research limitations remain. When identifying research trends, researchers need more detailed information than research topics alone. More detailed information, such as algorithms, datasets, research problems, and performance, would improve researchers' understanding of the current SOTA in specific topics. Therefore, future work should examine how information extraction techniques can be combined more deeply with topic modeling approaches, especially BERTopic. In this way, it may become possible to automatically extract a dynamic leaderboard or identify the current SOTA of a specific topic. Moreover, the role of the LLM could become more substantial, not only in topic refinement but also in guiding the extraction process. Using an LLM for topic refinement also presents several challenges. The refinement process is sensitive to the prompt used. Furthermore, LLM output can be non-deterministic, meaning each run of the same prompt may produce different labels or topic representations. The resulting labels also risk losing context because the refinement process relies only on generated keywords rather than the full document. For that reason, careful attention is needed in several areas, such as providing an appropriate list of topic modeling keywords, setting LLM parameters consistently, such as temperature, and introducing a mechanism that provides context representative of the original document.

VI. CONCLUSION

This paper has identified the evolution and trends in the computer vision domain based on three leading CV conferences. The experiments show that the modern topic modeling method, BERTopic, outperformed the conventional LDA method in terms of internal validation (coherence) and external validation (manual or human judgment). BERTopic demonstrated its effectiveness in uncovering popular research topics (trends) and tracking topic changes over time (evolution). Furthermore, BERTopic produced consistent results when applied to both title and abstract datasets. Although the generated topics were represented by understandable keywords, the Gemini LLM improved the interpretability of these topics. The topics evolved as follows: in the early years, the focus was more on typical image processing tasks; in the middle years, it shifted to the adoption of deep learning; and in the most recent years, it emphasized generative approaches. In future work, we plan to explore information extraction techniques to provide more detailed information, such as algorithms, datasets, research problems, and performance. This additional detail will improve researchers' understanding of the current SOTA of a specific research topic beyond topic evolution alone. In addition, the use of LLMs may be more effective for extracting detailed information than for topic refinement alone.

Declaration of AI and AI assisted technologies in the writing process

During the preparation of this work the authors used Google Translate and Quilbot in order to translate the text into English and revise grammar error. After using this tool/service, the authors reviewed and edited the content as needed and takes full responsibility for the content of the publication.

CRediT Authorship Contribution Statement

Setio Basuki: Conceptualization, Data curation, Formal Analysis, Project administration, Resources, Software, Validation, Visualization, Writing – original draft, and Writing – review & editing. **Zamah Sari**: Conceptualization, Validation, Visualization, and Writing – review & editing. **Rizky Indrabayu**: Formal Analysis, Software, and Writing – review & editing. **Masatoshi Tsuchiya**: Conceptualization, Validation, and Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] K. Goffin, P. Åhlström, M. Bianchi, and A. Richtnér, "Perspective: State-of-the-Art: The Quality of Case Study Research in Innovation Management," *J. Prod. Innov. Manag.*, vol. 36, no. 5, pp. 586–615, Sep. 2019.

[2] E. S. Barry, J. Merkebu, and L. Varpio, "State-of-the-art literature review methodology: A six-step approach for knowledge synthesis," *Perspect. Med. Educ.*, vol. 11, no. 5, pp. 1–8, Sep. 2022.

[3] M. L. Weaver *et al.*, "The art of peer review: Guidelines to become a credible and constructive peer reviewer," *Semin. Vasc. Surg.*, vol. 35, no. 4, pp. 470–478, Dec. 2022.

[4] C.-H. Lee, C.-L. Liu, A. J. C. Trappey, J. P. T. Mo, and K. C. Desouza, "Understanding digital transformation in advanced manufacturing and engineering: A bibliometric analysis, topic modeling and research trend discovery," *Adv. Eng. Inform.*, vol. 50, p. 101428, Oct. 2021.

[5] F. Gao *et al.*, "Bibliometric analysis on tendency and topics of artificial intelligence over last decade," *Microsyst. Technol.*, vol. 27, no. 4, pp. 1545–1557, Apr. 2021.

[6] M. Mustak, J. Salminen, L. Plé, and J. Wirtz, "Artificial intelligence in marketing: Topic modeling, scientometric analysis, and research agenda," *J. Bus. Res.*, vol. 124, pp. 389–404, Jan. 2021.

[7] S. Boon-Itt and Y. Skunkan, "Public Perception of the COVID-19 Pandemic on Twitter: Sentiment Analysis and Topic Modeling Study," *JMIR Public Health Surveill.*, vol. 6, no. 4, p. e21978, Nov. 2020.

[8] T. Riepponen, M. Moilanen, and J. Simonen, "Themes of resilience in the economics literature: A topic modeling approach," *Reg. Sci. Policy Pract.*, vol. 15, no. 2, pp. 326–357, Apr. 2023.

[9] R. Agrawal, V. A. Wankhede, A. Kumar, S. Luthra, A. Majumdar, and Y. Kazancoglu, "An Exploratory State-of-the-Art Review of Artificial Intelligence Applications in Circular Economy using Structural Topic Modeling," *Oper. Manag. Res.*, vol. 15, no. 3–4, pp. 609–626, Dec. 2022.

[10] J. C. Lyu, E. L. Han, and G. K. Luli, "COVID-19 Vaccine–Related Discussion on Twitter: Topic Modeling and Sentiment Analysis," *J. Med. Internet Res.*, vol. 23, no. 6, p. e24435, Jun. 2021.

[11] H. Jang, E. Rempel, D. Roth, G. Carenini, and N. Z. Janjua, "Tracking COVID-19 Discourse on Twitter in North America: Infodemiology Study Using Topic Modeling and Aspect-Based Sentiment Analysis," *J. Med. Internet Res.*, vol. 23, no. 2, p. e25431, Feb. 2021.

[12] W. Lou and J. Meng, "The diversity of canonical and ubiquitous progress in computer vision: A dynamic topic modeling approach," *Inf. Process. Manag.*, vol. 60, no. 3, p. 103238, May 2023.

[13] A. A. S. Li, A. J. C. Trappey, C. V. Trappey, and C. Y. Fan, "E-discover State-of-the-art Research Trends of Deep Learning for Computer Vision," in *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, Bari, Italy: IEEE, Oct. 2019, pp. 1360–1365.

[14] M. Grootendorst, "BERTopic: Neural topic modeling with a class-based TF-IDF procedure," Mar. 11, 2022, *arXiv*: arXiv:2203.05794.

[15] F. Gurcan and S. Sevik, "Mapping the Research Landscape of Deep Learning from 2001 to 2019," in *2019 1st International Informatics and Software Engineering Conference (UBMYK)*, Ankara, Turkey: IEEE, Nov. 2019, pp. 1–5.

[16] M. R. Alagheband, A. Mashatan, and M. Zihayat, "Time-based Gap Analysis of Cybersecurity Trends in Academic and Digital Media," *ACM Trans. Manag. Inf. Syst.*, vol. 11, no. 4, pp. 1–20, Dec. 2020.

[17] N. T. Viet and A. G. Kravets, "Analyzing Recent Research Trends of Computer Science from Academic Open-access Digital Library," in *2019 8th International Conference System Modeling and Advancement in Research Trends (SMART)*, Moradabad, India: IEEE, Nov. 2019, pp. 31–36.

[18] J. Y. Lee and J. Lee, "Current Research Trends in IoT Security: A Systematic Mapping Study," *Mob. Inf. Syst.*, vol. 2021, pp. 1–25, Mar. 2021.

[19] S. Rouhani and F. Mozaffari, "Sentiment analysis researches story narrated by topic modeling approach," *Soc. Sci. Humanit. Open*, vol. 6, no. 1, p. 100309, 2022.

[20] M. Inaam Ul Haq, Q. Li, and J. Hou, "Analyzing the Research Trends of IoT Using Topic Modeling," *Comput. J.*, vol. 65, no. 10, pp. 2589–2609, 2022.

[21] G. Tokdemir and S. Uguz, "Research Trends in Agile Software Development," in *2022 International Conference on Emerging Trends in Computing and Engineering Applications (ETCEA)*, Karak, Jordan: IEEE, Nov. 2022, pp. 1–5.

[22] N. Tsapatsoulis, H. Partaourides, C. Christodoulou, and C. Djouvas, "Quo Vadis Computer Science? The topics of the influential papers during the period 2014-2021," in *2022 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech)*, Falerna, Italy: IEEE, Sep. 2022, pp. 1–8.

[23] C. Sharma, S. Sharma, and Sakshi, "Latent DIRICHLET allocation (LDA) based information modelling on BLOCKCHAIN technology: a review of trends and research patterns used in integration," *Multimed. Tools Appl.*, vol. 81, no. 25, pp. 36805–36831, Oct. 2022.

[24] R. V. Encinas Quille, J. M. Barros, M. B. Júnior, F. V. De Almeida, and P. L. P. Corrêa, "Detecting Favorite Topics in Computing Scientific Literature via Dynamic Topic Modeling," *IEEE Access*, vol. 11, pp. 41535–41545, 2023.

[25] W. Kang, Y. Kim, H. Kim, and J. Lee, "An Analysis of Research Trends on Language Model Using BERTopic," in *2023 Congress in Computer Science, Computer Engineering, & Applied Computing (CSCE)*, Las Vegas, NV, USA: IEEE, Jul. 2023, pp. 168–172.

[26] C.-D. Curiac and M. V. Micea, "Identifying Hot Information Security Topics Using LDA and Multivariate Mann-Kendall Test," *IEEE Access*, vol. 11, pp. 18374–18384, 2023.

[27] D. Choi, "Analysis of IoT Research Topics Using LDA Modeling," in *Big Data, Cloud Computing, and Data Science Engineering*, vol. 1075, R. Lee, Ed., in Studies in Computational Intelligence, vol. 1075., Cham: Springer International Publishing, 2023, pp. 153–163.

[28] A. Ayaz *et al.*, "Exploring Gamification Research Trends Using Topic Modeling," *IEEE Access*, vol. 11, pp. 119676–119692, 2023.

[29] S. Uğuz, B. Özyer, and A. Coşkunçay, "Research Trends in Mobile Robots: A Comparative Analysis within the Periods before and after the Industry 4.0 Revolution," in *2023 International Conference on Intelligent Computing and Control (IC&C)*, Wuhan, China: IEEE, Feb. 2023, pp. 43–47.

[30] H. Park, B. Ahn, and T. Kim, "An exploration of research trends on metaverse: topic modeling with latent dirichlet allocation," *Qual. Quant.*, vol. 59, no. 1, pp. 233–252, Feb. 2025.

[31] B. Gana, A. Leiva-Araos, H. Allende-Cid, and J. García, "Leveraging LLMs for Efficient Topic Reviews," *Appl. Sci.*, vol. 14, no. 17, p. 7675, Aug. 2024.

[32] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet Allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, 2003.

[33] Gemini Team *et al.*, "Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context," 2024, *arXiv*.

[34] L. McInnes, J. Healy, and J. Melville, "UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction," 2018, *arXiv*.

[35] R. J. G. B. Campello, D. Moulavi, and J. Sander, "Density-Based Clustering Based on Hierarchical Density Estimates," in *Advances in Knowledge Discovery and Data Mining*, vol. 7819, J. Pei, V. S. Tseng, L. Cao, H. Motoda, and G. Xu, Eds., in Lecture Notes in Computer Science, vol. 7819., Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 160–172.

[36] Jana Bernhard, Martin Teuffenbach, and Hajo G. Boomgaarden, "Topic Model Validation Methods and their Impact on Model Selection and Evaluation," *Comput. Commun. Res.*, vol. 5, no. 1, p. 1, Jan. 2023.