IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 30 October 2023, accepted 9 November 2023, date of publication 13 November 2023, date of current version 16 November 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3332468

SURVEY

When, Where, and Which?: Navigating the Intersection of Computer Vision and Generative AI for Strategic Business Integration

MUHAMMAD HUSSAIN
Department of Computer Science, University of Huddersfield, HD1 3DH Huddersfield, U.K.
e-mail: m.hussain@hud.ac.uk

ABSTRACT In today's rapidly evolving digital landscape, Artificial Intelligence (AI) exerts a profound influence on our daily lives, from predictive text in emails to the ever-present virtual assistants like Alexa and Siri. This scholarly article embarks on a comprehensive exploration of the expansive world of Artificial Intelligence, with a keen focus on the domains of generative AI and computer vision. Our objective is to provide businesses with a nuanced and in-depth understanding of these critical AI subfields. By doing so, we empower organizations to make informed and strategic decisions regarding the adoption of generative AI and computer vision technologies. Our ultimate goal is to equip businesses with the knowledge and insights necessary to harness the potential of these AI domains effectively, driving innovation and bolstering their competitive edge in an increasingly technology-driven world.

INDEX TERMS Computer vision, generative AI, machine vision, business intelligence, automation.

I. INTRODUCTION

In our technology-driven era, Artificial Intelligence (AI) has transcended the realm of science fiction to become an integral part of daily life [1]. Familiar names like Gmail, Spotify, Alexa, and Siri epitomize AI's pervasive influence, seamlessly enhancing our experiences [2]. These AI-driven tools have quietly adapted to our preferences and habits over the years, making life a tad more convenient and enjoyable [3]. While the occasional uncanny ad targeting might raise eyebrows, we have become accustomed to AI's presence.

However, the recent surge in AI-related headlines extends beyond familiar territories, penetrating diverse domains, and garnering newfound attention [4]. AI is undergoing a renaissance, emerging as the "It" technology of our time, transcending the confines of business and technology sections to seize the limelight across the media spectrum [5].

A significant catalyst for this renewed fervour surrounding AI is the advent of OpenAI's ChatGPT in November 2022 [6]. ChatGPT, an acronym for Generative Pre-trained Transformer, introduced the public not to AI but to a specific facet of AI – generative AI [7].

The terminology surrounding AI, generative AI, and related concepts has grown increasingly convoluted within this expansive field [8]. To offer clarity amidst the complexity, this article embarks on a mission to elucidate the distinctions between AI, generative AI, and computer vision [9]. Our aim is to provide readers with a comprehensive understanding, presented in a manner both enlightening and engaging, surpassing the capabilities of virtual assistants like Alexa or Siri [10].

This scholarly article seeks to embark on a comprehensive journey through the realm of Artificial Intelligence, with a particular focus on the domains of generative AI and computer vision. Our endeavor aims to serve a critical purpose: to illuminate the significance of these AI subfields and address the essential need for businesses to comprehend their capabilities.

By doing so, we intend to provide organizations with a profound understanding of generative AI and computer vision, ultimately enabling them to make well-informed and strategic

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Wei.

decisions about the integration of these technologies. This exploration is grounded in the recognition that the adoption of generative AI and computer vision technologies is not merely an option but a compelling necessity, given their potential to drive innovation and enhance competitiveness in a world increasingly defined by technological advancements.

A. SURVEY OBJECTIVE

This research initiative is guided by the central objective of furnishing businesses with a comprehensive framework for strategically adopting AI technologies, with a specific focus on Generative AI and Computer Vision, tailored to their unique operational domains. The overarching goal is to facilitate informed decision-making processes within organizations, enabling them to harness the transformative potential of these AI subfields effectively. The survey seeks to accomplish the following specific academic objectives:

1) Clarify the Role of Generative AI and Computer Vision: The research begins by providing clear and precise explanations of Generative AI and Computer Vision within the context of AI technologies. This foundational understanding equips businesses with the knowledge necessary to make informed choices when considering the integration of Generative AI and Computer Vision into their operations.
2) Highlight the Business Relevance: Recognizing the significant influence of Generative AI and Computer Vision in contemporary business landscapes, this study emphasizes their pivotal roles across diverse sectors. Concrete examples illustrate how these AI subfields enhance user experiences and stimulate business value.
3) Explore Application Scenarios: Delving into the potential applications of Generative AI and Computer Vision, the research aims to demystify their nuanced distinctions and showcase their practical utility in various industries. By providing real-world use cases, this study helps businesses identify relevant scenarios for implementation.
4) Facilitate Informed Decision-Making: The core objective is to empower businesses with the knowledge and insights needed to make well-informed decisions regarding the adoption of Generative AI and Computer Vision. By carefully assessing their specific industry requirements, operational objectives, and technological capacities, organizations can make strategic choices that lead to innovation, operational efficiency, and enhanced competitiveness.
5) Provide Selection Guidance: To assist businesses in navigating the AI landscape effectively, this research offers guidance on selecting the most suitable Generative AI and Computer Vision solutions aligned with their domain-specific goals. This guidance encompasses a comprehensive evaluation of challenges, definition of objectives, data readiness, content generation needs, AI model selection, resource availability assessment, scalability planning, exploration of industry-specific solutions, pilot program implementation strategies, and vigilance regarding emerging trends in Generative AI and Computer Vision.

By addressing these academic objectives, this research aims to empower businesses with a solid foundation for making informed decisions about integrating Generative AI and Computer Vision technologies. The overarching aim is to enable organizations to leverage these AI subfields in a manner that seamlessly aligns with their unique operational characteristics, thereby fostering innovation and bolstering competitive resilience in an era increasingly shaped by advancements in AI.

B. ORGANIZATION OF PAPER

This academic review initiates by furnishing a thorough definition of AI and systematically investigates its multifaceted classifications. Subsequently, it provides an exhaustive and detailed exploration of generative AI, a subdomain within AI renowned for its unique ability to generate creative and novel content across diverse domains, including textual, visual, and auditory mediums. The review then transitions its focus towards an in-depth examination of computer vision, underscoring its extensive applicability and versatility across various industries, elucidating how it empowers machines to interpret and comprehend visual data from the physical world much like human visual perception.

Following this, the review exclusively delves into the critical process of selecting the most suitable AI solution for respective business applications. It expounds upon the essential considerations and methodologies involved in making an informed choice regarding AI implementation tailored to specific business objectives and requirements. Finally, the review culminates by offering a coherent and structured synthesis of the key points discussed throughout the paper. In doing so, it provides a conclusive and comprehensive evaluation of navigating the AI landscape, emphasizing its pivotal role in achieving optimal outcomes and business success in the contemporary technological landscape.

II. RESEARCH METHODOLOGY

A. SOURCE SELECTION AND ACQUISITION

The sources for this study were obtained from reputable academic databases, including Google Scholar, IEEE Xplore, and the ACM Digital Library. We conducted extensive searches using the keywords 'Computer vision,' 'Generative AI,' 'Computer vision for Business use cases,' and 'Generative AI for Business applications.' This approach allowed us to identify a broad range of sources that were potentially relevant to our research.

B. DATA COLLECTION AND SYNTHESIS

As a secondary research paper, our data collection process consisted of extracting and synthesizing information from the selected sources. We did not engage in primary data collection or quantitative research. Instead, we focused on summarizing and organizing the insights and findings presented in the literature to provide a cohesive and informative narrative.

III. WHAT IS AI?

The term "AI" encapsulates a vast and multifaceted domain of non-human intelligence, representing a significant departure from traditional computational paradigms [11]. It denotes the discipline dedicated to the development of computer systems imbued with the capability to execute tasks traditionally associated with human intelligence [12]. In concordance with the succinct articulation by Demis Hassabis, Co-Founder and CEO of DeepMind, AI can be succinctly characterized as "the science of making machines smart."

At its core, AI endows machines with the capacity to comprehend natural language, discern intricate patterns within data, make informed decisions, and accumulate knowledge from experiential encounters [13]. This emulation of human-like cognitive processes empowers machines to not only process and interpret information but also adapt to diverse contextual scenarios, iteratively enhancing their performance through continuous learning [14].

Within the expansive landscape of AI, several prominent techniques and methodologies emerge, with machine learning and deep learning occupying preeminent positions. Machine learning entails the meticulous training of algorithms on extensive datasets, facilitating their ability to extrapolate predictions or decisions from novel inputs [15]. In contrast, deep learning, a subdomain nestled within machine learning, harnesses intricate neural networks comprising interconnected layers, drawing inspiration from the intricate synaptic structure of the human brain [16]. These neural networks exhibit an innate proficiency in deciphering complex patterns within data, rendering them particularly well-suited for tasks such as image recognition [17].

The ubiquitous applicability of AI traverses a diverse spectrum of industries, encompassing domains as varied as manufacturing [18], retail [19], construction [20], healthcare [21], finance, transportation, and energy [22]. Within these sectors, AI assumes multifarious roles, endowing computational systems with the capacity to meticulously analyze prodigious datasets, execute arduous and repetitive tasks with unwavering precision, proffer tailored recommendations to end-users, and, notably, replicate human-like interactions through the deployment of chatbots [23] and virtual assistants [24].

Notwithstanding its historical inception, the recent surge of attention and enthusiasm directed towards AI has engendered a perceptible blurring of its definition and capabilities. Thus, to cultivate a more profound and discerning comprehension of AI's salient role within the contemporary business milieu, it becomes imperative to embark on a comprehensive exploration of the diverse categories of AI, each distinguished by its unique functions, competencies, and practical implications.

A. AI TAXONOMY: FROM REACTIVE TO HYPOTHETICAL SELF-AWARE ENTITIES

The domain of AI exhibits a rich tapestry of variegated typologies, each encapsulating distinct attributes and competencies, ranging from rudimentary predefined responses to the theoretical concept of sentient self-awareness. A thorough comprehension of these discrete AI classifications is imperative to fathom the expanse of functionalities that they proffer.

1) Artificial Narrow Intelligence (ANI): At the foundational echelon of AI classification exists ANI, signifying machines meticulously engineered to perform specific, delimited tasks with precision [25]. Although these machines excel within their designated domains, they lack the capacity for extrapolation and generalization. Emblematic of ANI is the historic triumph of IBM's Deep Blue, which, in 1997, etched its name in history by outwitting the reigning world chess champion, Garry Kasparov [26]. The triumph of Deep Blue hinged upon its judicious adherence to predefined algorithms and strategies, albeit devoid of adaptability or receptivity to novel experiential data.
2) Artificial General Intelligence (AGI): Ascending the gradient, AGI materializes as a focal point [27]. AGI systems aspire to mirror the cognitive proficiencies reminiscent of human faculties. Their remit encompasses an extensive gamut of cognitive tasks, effectively paralleling the spectrum of human cognitive processes. AGI, while largely aspirational, harbors the potential to orchestrate multifarious tasks, an accomplishment of transformative consequence.
3) Artificial Superintelligence (ASI): The apogee of AI acumen manifests through ASI, heralding the domain of AI systems endowed with capabilities that transcend the confines of human prowess [28]. ASI, however, remains entrenched in the realm of theoretical conjecture, revolving around the hypothetical supposition of AI entities that surpass human aptitude across diverse intellectual endeavours.

Delving further into the nuanced gradations of AI, four principal categories emerge as quintessential embodiments of AI's manifold capabilities:

1) REACTIVE MACHINES

Within this classification, reactive machines, typified by IBM's Deep Blue, operate within a rigid framework of predefined rules and algorithmic stratagems. They proficiently formulate decisions and prognostications predicated upon historical precedents but languish in the absence of cognitive adaptability or the capacity to assimilate newfound experiential data [29].

2) LIMITED MEMORY AI

In contradistinction to their reactive counterparts, limited memory AI systems possess the ability to refine their performance quotient commensurate with the accrual of data reservoirs [30]. This category of AI imbibes lessons from prior interactions, retaining and leveraging discernments to yield contextually pertinent responses. Conspicuous manifestations of limited memory AI include autonomous vehicular platforms, which harness historical data to navigate intricate traffic contingencies adeptly.

3) THEORY OF MIND AI

Theory of mind AI, while extant primarily within the realm of theoretical discourse, envisions AI entities endowed with the acumen to decipher the emotive substrata of human comportment [31]. Such AI systems would espouse the capacity to decipher intentions and envisage behavioral trajectories—a concept yet to crystallize into practicality due to its intrinsic intricacy.

4) SELF-AWARE AI

Functioning in alignment with theory of mind AI within the theoretical sphere, self-aware AI delineates the speculative concept of AI endowed with self-awareness [32]. Such theoretical AI entities would ostensibly possess idiosyncratic predilections, preferences, and an astute awareness of their emotional resonance with human interlocutors—an evocative notion whose concrete realization remains an elusive conjecture.

These delineations within the AI taxonomy serve as illuminative signposts in navigating the labyrinthine contours of artificial intelligence, elucidating the evolving trajectory of technology's role in the evolving human milieu.

IV. GENERATIVE AI: WHERE IMAGINATION MEETS ALGORITHM

Generative AI, an overarching concept encompassing a diverse array of AI systems, emerges as a transformative force in the realm of data-driven innovation [33]. At its core, generative AI represents the vanguard of technologies capable of generating multifaceted content, spanning textual narratives, visual artistry, code, and auditory compositions [34]. The essence of generative AI lies in its unique capacity to extrapolate from existing data reservoirs, metamorphosing raw information into novel creations.

In essence, generative AI functions as a creative alchemist, wielding complex algorithms to distill predictions and insights, catalyzing the genesis of entirely new data based on its pre-existing knowledge base. An emblematic instance of generative AI, ChatGPT [6], serves as a prime illustration of this paradigm, wherein it synthesizes coherent textual narratives drawing from an extensive corpus of training data. It is imperative to underscore, however, that generative AI operates within the confines of learned patterns and lacks genuine comprehension—a characteristic distinguishing it from human creativity.

In parallel to textual content generation, generative AI extends its purview to the realms of visual, auditory, and code generation. Models in this category undergo training on vast datasets, which imbues them with the aptitude to craft innovative visual artistry, encompassing an array of styles and genres. Likewise, generative AI systems demonstrate their mettle in the realm of auditory domain, producing audio recordings that traverse a gamut of tonalities and styles, resembling human conversation.

Figure 1 provides a visual framework for understanding the hierarchy of subfields within the expansive domain of Artificial Intelligence (AI). At the core of AI, it introduces two primary subfields: Generative AI and Computer Vision. Generative AI encompasses various facets, such as Text Generation, Image Generation, Voice Generation, Style Transfer, and Anomaly Detection, emphasizing its diverse content generation capabilities and its role in identifying anomalies in data. In contrast, Computer Vision encompasses Image Detection, Image Tracking, Image Reconstruction, Image Classification, Object Detection, and Image Segmentation, underlining its role in interpreting visual data, including object recognition, tracking, and image segmentation. This visual representation simplifies the complex landscape of Generative AI and Computer Vision within the broader field of AI.

Examples of AI-generated content encompass a wide spectrum, spanning from artistic creations, like the works discussed by [35], to the generation of synthetic human faces, as illustrated by [36], and even extending to scientific domains, such as the generation of magnetograms of the Sun, as demonstrated in the research conducted by [37]. Nonetheless, the proliferation of generative AI gives rise to a myriad of potential legal, ethical, and moral challenges. These encompass concerns related to copyright infringement in AI-generated artworks, a topic thoroughly examined by [35]. Additionally, generative AI engenders issues in academic settings, including academic dishonesty, cheating, and plagiarism, as acknowledged by [38]. Furthermore, considerations surrounding data privacy and security have been thoroughly explored by [39].

Table 1 presents the manifestation of Generative AI in several domains whilst Figure 2 presents popular architectures categorised based on input and output generation. The allure of generative AI is undeniably tantalizing, evoking astonishment through its capacity to seemingly conjure something out of nothing. Nevertheless, it is pivotal to discern that generative AI is bereft of the mystique of a magic wand; its output is intricately interwoven with the knowledge accrued during its training phase. The deployment of generative AI reverberates across multifarious domains. It unfolds as an indispensable tool for content generation, spanning the creation of social media posts, articles, images, and even code for website development. This engenders a manifold impact, precipitating temporal savings, catalyzing innovation by unburdening human resources from mundane undertakings, and fostering the flourishing of creativity.

Nonetheless, in navigating the terrain of generative AI, cognizance of its intrinsic limitations is imperative. The system's creative endeavors remain tethered to the scope and quality of its training data, invariably manifesting occasional

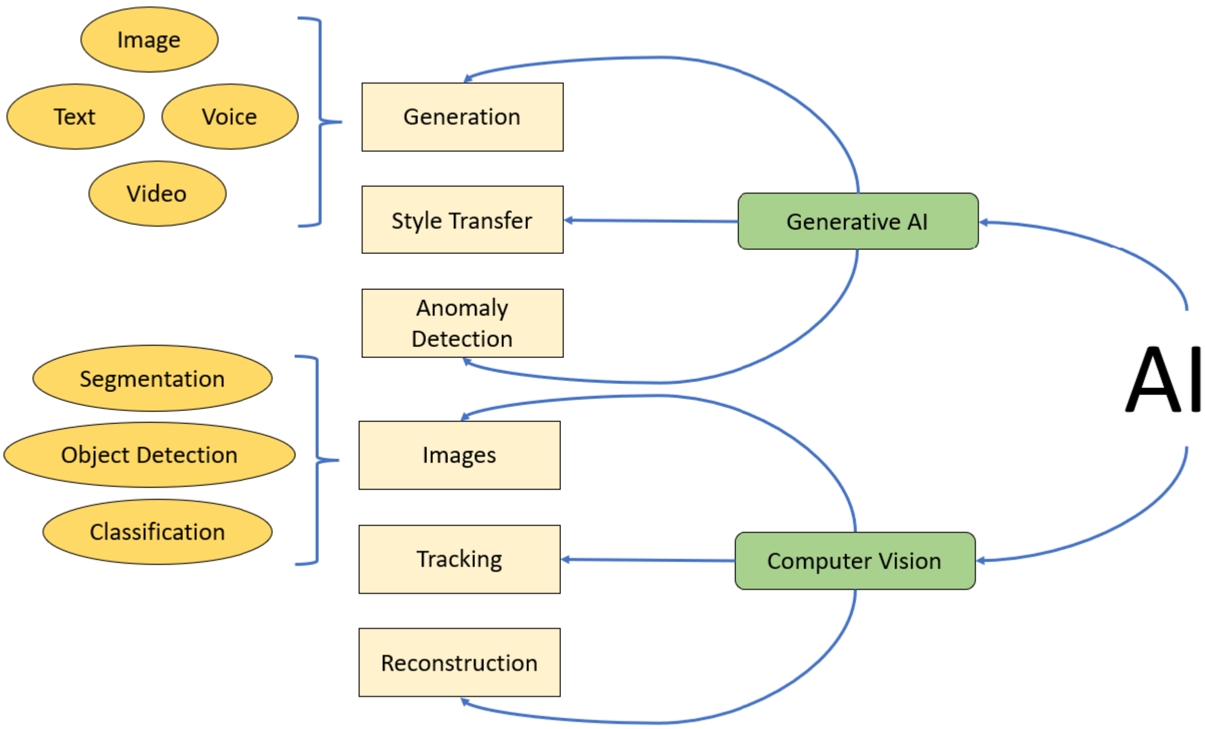

FIGURE 1. An abstract view of generative AI and computer vision.

inaccuracies, unpredictability, and incoherence in its output. Moreover, generative AI operates in the absence of real-time data, rendering it a static repository rather than a dynamic, context-aware entity. The meteoric ascent of generative AI to the forefront of technological prominence in November 2022, catalyzed by OpenAI's ChatGPT [6], has ignited a palpable buzz and widespread adoption. Its accessibility to the public sans intricate prerequisites, coupled with its utility and cost-effectiveness, have rendered it an indomitable presence within the AI landscape.

As the digital landscape metamorphoses, businesses seeking real-time visual insights and data-driven inferences within their operational domain veer toward the domain of computer vision—an AI variant endowed with the capability to discern and decipher visual data in real-time, engendering pivotal enhancements across diverse industries.

V. COMPUTER VISION: A GATEWAY TO VISUAL UNDERSTANDING

In the expansive domain of artificial intelligence (AI), computer vision stands as a transformative field strategically positioned at the confluence of AI and computer science [47]. Its overarching objective is to confer upon machines a remarkable cognitive ability: the aptitude to comprehend and decipher visual information akin to human perception. This paradigmatic shift empowers machines with cognitive faculties, endowing them with the capacity to perceive, comprehend, and derive meaningful insights from the extensive array of visual stimuli encountered in the physical world [48].

The foundational premise of computer vision hinges on the adept utilization of pre-existing camera technologies to accomplish swift detection and real-time processing of visual data. This encompasses a broad spectrum of visual elements, including but not confined to individuals, objects, and dynamic events. The innate immediacy of this feedback loop furnishes enterprises with a potent instrument, enabling them to respond promptly to unfolding scenarios, thereby culminating in substantial augmentations in both productivity and safety within their operational domains.

The pragmatic significance of computer vision becomes strikingly evident in its ability to address the burgeoning demand for immediate visual insights across diverse industries. This renders computer vision an indispensable asset for enterprises fervently in pursuit of optimizing and reinforcing their operational processes through the acquisition of real-time visual information.

Diverging from the purview of generative AI, which primarily centers on generating novel content derived from pre-existing data patterns, computer vision assumes the role of an additional sensory organ. It bestows entities with unparalleled access to real-time visual data, offering an expansive array of applications that span virtually every industry and sector. Consequently, computer vision emerges as a quintessential and transformative AI solution, uniquely poised to empower businesses with instantaneous visual information. This empowerment facilitates informed decision-making and process enhancements, serving as a catalyst across a wide array of operational scenarios.

TABLE 1. Applications of generative AI.

Domain	Application	Detailed Description	Specific Examples	References
Business	Marketing	Generative AI is employed to create highly personalized and context-aware advertisements. It leverages user data, preferences, and behavior to generate ad copy, visuals, and targeting strategies that are most likely to engage and convert potential customers. This application significantly enhances marketing efficiency and effectiveness.	Personalized ad campaigns, dynamic ad creatives, and targeted marketing based on user data.	[40]
Education	Learning Assistance	In the education sector, generative AI serves as a versatile learning assistant. It provides students with on-demand assistance, answering questions, explaining complex topics, and offering personalized learning resources. For educators, it helps in creating customized teaching materials, automating grading, and providing constructive feedback to students. This application enhances the overall educational experience.	AI-powered virtual tutors, personalized lesson plans, automated grading systems, and intelligent educational content generation.	[41]
Healthcare	Patient Interaction	Generative AI revolutionizes healthcare by improving patient interaction and support. Chatbots and virtual assistants powered by AI are capable of providing medical advice, answering health-related queries, and assisting in clinical diagnoses. These systems enhance access to healthcare information and services, especially in telehealth settings.	AI-driven healthcare chatbots, telemedicine platforms, and virtual health assistants.	[42]
Content Creation	Marketing	Generative AI creates tailored ad content, including headlines, body copy, images, and video scripts, for personalized and engaging marketing campaigns.	AI-generated ad creatives, personalized marketing content, and automated ad campaign management.	[43]
Content Creation	Journalism	Generative AI automates news production by analyzing data and generating data-driven news articles.	AI-generated news articles, data-driven reporting, and automated news writing.	[44]
Content Creation	Art	Generative AI aids artists, suggesting ideas, styles, and crafting visual or auditory art pieces, fostering novel artistic expression and creativity.	Collaborative art projects with AI, AI-generated artworks, and interactive artistic installations.	[45]
Gaming	Game Design	Generative AI automates game development, creating characters, plots, levels, and strategies, speeding up the process and adding innovation to gameplay.	AI-generated game content, procedurally generated game worlds, and automated game design tools.	[46]

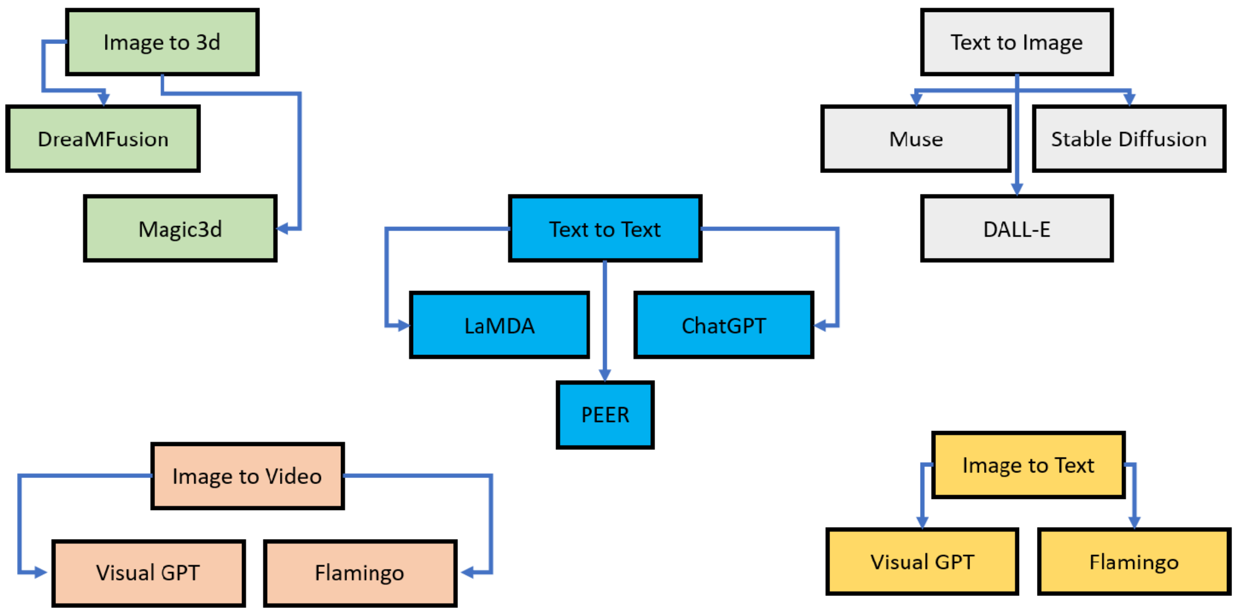

FIGURE 2. Application based generative AI models.

A. FACETS OF COMPUTER VISION

Within the domain of computer vision, three distinct categories come to the fore, each representing a specialized facet of visual understanding and processing: image classification [49], object detection [50], and segmentation [51]. These categories collectively constitute the foundation upon which computer vision systems operate, and they each serve unique and pivotal roles in various applications.

1) IMAGE CLASSIFICATION

Image classification, at its core, involves the categorization of visual inputs into predefined classes or labels. In this process, a computer vision system is trained to recognize and assign specific labels or categories to images based on their content. For instance, an image classification model might be trained to differentiate between defective and normal photovoltaic modules within a solar panel manufacturing complex [52]. This categorization is based on patterns and features extracted from the images, enabling the system to make accurate decisions regarding the class to which each image belongs. Image classification is fundamental in applications where understanding the content of an image in a general sense is paramount.

2) OBJECT DETECTION

Object detection represents a significant advancement beyond image classification. While image classification focuses on identifying and categorizing objects within an image, object detection goes further by not only identifying objects but also precisely locating their positions within the visual data. In essence, it answers the question, "Where are the objects in this image?" This capability is particularly practical and applicable in real-world scenarios where knowing the spatial arrangement of objects is crucial. For example, in autonomous driving systems, object detection enables a vehicle to not only recognize other vehicles on the road but also determine their precise locations and sizes. This spatial awareness is vital for making informed decisions and ensuring safety in several industries [53]. Object detection employs techniques like bounding boxes or more advanced methods like keypoint detection to achieve this precise localization of objects.

3) SEGMENTATION

Segmentation takes visual understanding to an even more granular level. Unlike image classification and object detection, which focus on objects as a whole, segmentation is concerned with identifying and delineating individual pixels or regions within an image. It aims to partition an image into semantically meaningful segments or masks, where each pixel is assigned to a specific object or background category [51]. This level of detail is immensely valuable in applications where pixel-level accuracy is required. For instance, in medical imaging, segmentation can be used to identify and outline tumors within MRI scans, enabling precise diagnosis and treatment planning. Similarly, in the field of image editing, segmentation allows for the selective manipulation of specific objects or regions within an image, providing powerful creative control.

In summary, these three categories within computer vision—image classification, object detection, and segmentation—represent a hierarchy of visual understanding, with each serving distinct purposes in a wide range of applications. While image classification provides a foundational understanding of image content, object detection adds the crucial element of spatial localization, and segmentation delves into pixel-level precision, enabling computers to interpret and interact with visual data in increasingly sophisticated ways. These capabilities continue to drive innovations across industries and are instrumental in shaping the future of AI and computer vision technologies.

Table 2 presents an overview of various computer vision-based applications along with relevant references, the models used, key features, and performance metrics for each application. These applications showcase the versatility and capabilities of computer vision in solving real-world challenges.

VI. PIONEERING THE TRANSFORMATIVE LANDSCAPE: COMPUTER VISION'S MULTIFARIOUS APPLICATIONS

The versatility intrinsic to computer vision transcends the realms of theoretical acumen, as its real-world applications span diverse sectors, galvanized by the relentless march of technology. This practical manifestation of artificial intelligence not only interprets visual data in real-time but also catalyzes process refinement, fosters informed decision-making, and bequeaths an elevated plane of efficiency. The transmutative potential of computer vision is ushering in a paradigm shift, reshaping the modus operandi of businesses across the spectrum.

A. RETAIL

In an arena characterized by cutthroat competition, where retailers straddle the dichotomy of online and brick-and-mortar sales, each competitive edge assumes pivotal importance. Computer vision emerges as a formidable ally in this dynamic milieu [69].

1) OCCUPANCY COUNTING

The precision of customer enumeration within retail spaces has long eluded conventional methodologies. Manual tallies or reliance on motion sensor cameras invariably culminate in inaccuracies, engendering lost revenue opportunities and eroding customer satisfaction [70]. Computer vision intervenes as the harbinger of rectification, orchestrating the automation of meticulous real-time occupancy tallies. Its acumen extends beyond mere enumeration, as it can discern between customers, staff members, and delivery personnel. This empowerment facilitates judicious staff allocation and expeditious customer assistance, engendering enhanced shopper experiences.

2) SPEED OF SERVICE

Protracted queues constitute a perennial scourge afflicting the retail landscape, with a staggering 86% of customers forsaking their purchases due to extended wait times, thereby translating into substantial revenue hemorrhages [71].

TABLE 2. **Computer vision-based applications.**

Reference	Application	Model Used	Key Features	Performance Metrics
[54]	Clothing Attribute Recognition	RCNN	Modified selective search for regional proposals, Inception-ResNet for feature extraction, Soft-NMS for object localization	Precision: 73.59%, Recall: 83.84%
[55]	Fabric Structure Defect Detection	CNN with Attention Mechanism	Vision-based tactile sensor, Frequency domain filtering	Accuracy: 99.77% with Attention Mechanism
[56]	Traffic Sign Detection	Faster RCNN	Surveillance video for drivers, traffic lights, and road signs, Challenge of covering only 5% of the image	mAP improved by around 12%
[57]	Car Damage Detection and Assessment	VGG16 and VGG19	Domain-specific pre-trained CNN models, Transfer learning, L2 regularization	VGG19 Accuracy: 95.22%, VGG16 Accuracy: 94.56%
[58]	Person Recognition	Semi-Supervised Faster RCNN	Long-distance video surveillance, High-end cameras, Transfer learning	False Positive Rate: 15.93%, True Positive Rate: 59.16%, False Negative Rate: 40.84% (Threshold: 0.0001)
[59]	Semiconductor Laser Chip Defect Detection	AlexNet, ZFNet, GoogLeNet	Comparison of deep learning models, GoogLeNet selected, Development of GoogLeNet-based system	Accuracy: 97%
[60]	Object Detection and Recognition	Faster RCNN	Improved model precision and detection loss, Saliency detection, Proposal generation, Bounding box regression	mAP: 74.6
[61]	Pavement Damage Detection	Pre-trained EfficientNet B4	Incorporation of thermal data, Augmented dataset with non-uniform illumination, camera noise	F1-score: 98.34%
[62]	Object Detection with Drone Cameras	RCNN	Embedded Hardware: Jetson TX/AGX Xavier and Intel Neural Compute Stick, Real-time object Detection	0.9 fps
[63]	Wind Turbine Surface Distress Analysis	UNet Deep Learning Framework	Automated system based on UNet, Training with UAV-acquired images	Accuracy: 96% in detecting micro-crack locations
[64]	Silk Industry Defect Detection	Faster RCNN	Effective results in various lighting conditions, Internal joint detection	Detection Accuracy: 98%
[65]	Weld Defects Detection and Classification	VGG16 and ResNet50	Fine-tuned transfer learning with VGG16 and ResNet50, Performance evaluation on augmented dataset	Accuracy: 90%
[66]	Pedestrian Detection	YOLO-R	Tertiary layers consisting introduced to integrate shallow and deep layer pedestrian features.	25 frames per second
[67]	Pallet Racking inspection	Custom CNN	Automated pallet racking inspection via a custom CNN architecture for lightweight inefrencing	above 90% accuracy
[68]	Photovoltaic Inspection	Custom CNN	Automated Micro-crack inspection in manufacturing facilities for solar panels. Augmentation pipeline is proposed for enhancing genrealisation of proposed lightweight CNN architecture	97% F1-score

Computer vision embarks on a remedial mission, meticulously tracking and monitoring patrons ensconced in checkout lines in real-time. It promptly disseminates alerts upon the breach of predefined thresholds for wait times or line lengths, enabling proactive interventions to obviate revenue losses.

3) FOOT TRAFFIC ANALYTICS

The comprehension of customer comportment within a retail establishment stands as a strategic imperative, yet this comprehension has remained elusive due to the dearth of real-time insights into customer movements and product engagements. Computer vision steps into this void, effecting comprehensive tracking of foot traffic, identification of congregation hotspots, and continuous monitoring of customer interactions with merchandise and services [72]. The resultant data empowers retailers to optimize product placements, make data-informed staffing determinations, and elevate the customer experience to uncharted zeniths, all unfolding in real-time.

4) FRICTION-LESS CHECKOUT

In an era permeated by labor scarcities, the exigency of staff optimization and process automation reigns supreme. Computer vision assumes the mantle of solution-provider by catalyzing friction-less checkouts [73]. It accomplishes this by identifying items traversing conveyor belts sans reliance on barcode scanning and subsequently automating the payment process. The outcome is an expeditious and hassle-free checkout experience for customers, concurrently mitigating the demand for staff interventions.

B. AUTOMOTIVE

The automotive industry stands as an epitome of innovation, with computer vision at the vanguard, steering transformative advancements.

1) AUTONOMOUS VEHICLES

In the ambitious quest for self-driving automobiles, computer vision emerges as a linchpin. It assumes the mantle of the ocular and cerebral faculties of autonomous vehicles, deftly deciphering the intricacies of the surrounding milieu and orchestrating real-time decisions that propel safe and reliable autonomous navigation [74]. By adeptly detecting and dissecting objects, pedestrians, road signage, and traffic dynamics, computer vision endows autonomous vehicles with the faculty of prudence. The seamless amalgamation of computer vision technology serves as the conduit for realizing the future replete with self-driving automotive marvels.

2) DRIVER MONITORING
Amplifying driver safety and attentiveness constitutes an overarching imperative in the automotive domain. Herein, computer vision ascends to the forefront by instituting real-time vigilance over driver comportment and attentiveness. Through the meticulous analysis of facial expressions, ocular movements, and sundry behavioral cues, computer vision systems proactively gauge driver alertness [75]. In instances where indicators of fatigue or distraction surface, computer vision intervenes as the sentinel, contributing significantly to accident prevention and fortifying the citadel of road safety.

C. AGRICULTURE
In the sphere of agriculture, computer vision orchestrates a transformative renaissance, redefining traditional farming practices and ushering in an era characterized by precision and efficiency [76].

1) CROP MONITORING
The utilization of drones and ground-based cameras fortified with computer vision technology has wrought a revolution in crop monitoring. These aerial and terrestrial systems function as vigilant sentinels, endowing farmers with real-time insights into crop health. This empowered perspective facilitates early disease detection, precise irrigation practices, and meticulously targeted fertilization regimens [77]. By scrutinizing plant condition and growth trajectories, computer vision equips farmers with the data-driven acumen necessary to make judicious decisions, ultimately culminating in the augmentation of crop yields.

2) PRECISION FARMING
Precision stands as the linchpin of contemporary agriculture, and it is within this paradigm that computer vision assumes a pivotal mantle. By harnessing the intricate algorithms intrinsic to computer vision, farmers can judiciously apportion resources—such as water, fertilizers, and pesticides. This scrupulously tailored approach ensures that each precinct of arable land receives precisely the requisite nourishment, thereby minimizing waste and maximizing crop yield [78]. The union of computer vision and precision farming stands as the quintessential embodiment of resource-efficient and sustainable agricultural practices.

D. MANUFACTURING
The manufacturing industry stands as another bastion where computer vision's value proposition shines brightly. This technological marvel finds fertile ground for implementation, offering multifarious benefits to augment the manufacturing process [79].

1) ANOMALY AND DEFECT DETECTION
Anomalies and defects within the manufacturing realm are inevitable, often culminating in substantial financial losses. Computer vision, equipped with its discerning eye, inspects products on assembly lines with meticulous precision [80]. It adeptly identifies defects, inconsistencies, or missing components. The real-time nature of this system ensures that issues are apprehended as soon as they arise, thereby averting slowdowns and augmenting return on investment.

2) PACKAGE AND LABEL DETECTION
In an era marked by labor shortages and stringent quality control requirements, reliance on manual labor for label and package detection is fraught with risks. Computer vision ascends as the savior in this context, automating package and label detection with unparalleled accuracy [81]. This automation not only reduces operational costs but also optimizes staff allocation, fostering efficiencies throughout the production line.

3) SAFETY MONITORING
The manufacturing sector is replete with complex machinery, characterized by moving parts, sharp edges, and scalding surfaces. This inherently dangerous environment necessitates stringent safety measures. Computer vision steps in as a proactive guardian, enhancing workplace safety by monitoring hazardous areas [82]. It also tracks compliance with personal protective equipment regulations and efficiently manages machine usage in real-time. Immediate alerts pertaining to safety concerns empower factory floor supervisors to preclude accidents before they transpire.

4) VOLUMETRIC SPACE DETECTION FOR DISTRIBUTION AND TRUCKING
Inefficiencies within shipping and delivery processes plague the trucking industry, resulting in annual losses surpassing $27.5 billion. The dilemma of underutilized truck capacities exacerbates these challenges [83]. Computer vision comes to the fore as a game-changer, rendering real-time insights into truck capacity and usage. This data is tailor-made for businesses, providing immediate alerts concerning unused space and facilitating prudent material delivery capacity management. The result is an unequivocal boost to return on investment.

E. ELEVATING INDUSTRIAL EFFICIENCY
Within the realm of industrial operations, computer vision has emerged as an indispensable instrument, orchestrating a paradigm shift across multifarious domains.

1) ROBOTICS AND AUTOMATION
Computer vision assumes a pivotal role in orchestrating the choreography of robots as they navigate intricate tasks encompassing item picking, placement, assembly, as shown in Figure 3, and complex navigational maneuvers. By bestowing robots with real-time visual feedback and perception capabilities, computer vision empowers them to function seamlessly within dynamic and multifaceted

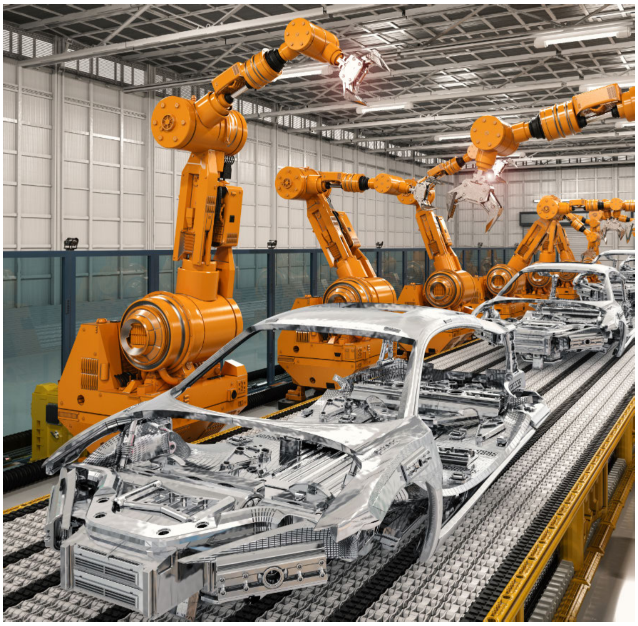

FIGURE 3. Automated car assembly.

environments [84]. This symbiotic interplay between robotics and computer vision culminates in heightened efficiency and precision, ushering in a new era of industrial applications.

2) SAFETY ENHANCEMENT

Safety stands as the bedrock of industrial settings, where stringent adherence to regulations is paramount. Computer vision emerges as a sentinel in this noble endeavor, steadfastly monitoring work environments in an uninterrupted fashion [85]. It not only ensures unwavering compliance with safety protocols but also stands vigilant in identifying nascent potential hazards. Through its proactive surveillance, computer vision initiates preemptive measures aimed at averting accidents. The real-time insights furnished by computer vision, therefore, constitute an invaluable asset in preserving and fortifying the security of the industrial ecosystem.

3) ENERGY EFFICIENCY

Energy consumption patterns within industrial facilities are often characterized by their intricacy and resource-intensive nature. Computer vision offers a judicious solution by subjecting these patterns to comprehensive analysis [86]. By meticulously identifying areas where energy utilization can be optimized, industrial operations embark on a trajectory marked by enhanced energy efficiency, for example the detection of Micro-cracks in Photovoltaic manufacturing facilities, as presented in Figure 4. The identification of opportune moments for energy savings becomes a streamlined and efficacious process with the proficient involvement of computer vision.

FIGURE 4. Photovoltaic micro-crack detection via computer vision.

VII. SELECTING THE OPTIMAL AI SOLUTION FOR YOUR BUSINESS

he decision to incorporate AI into business operations represents a significant stride towards innovation and operational efficiency. However, the prudent selection of the most fitting AI solution necessitates a meticulous evaluation of distinct needs and objectives. This section provides a framework to assist businesses in making a well-informed decision:

A. IDENTIFY KEY BUSINESS CHALLENGES

Commence your AI integration journey by meticulously identifying the specific challenges impeding your business. Whether these challenges pertain to customer service, production processes, or other facets of your operations, a lucid comprehension of these impediments will serve as the foundational compass for AI selection odyssey.

B. DEFINE YOUR OBJECTIVES

Endeavor to articulately delineate the precise objectives underpinning your decision to adopt AI technology. Do you aspire to curtail operational costs, amplify revenue generation, augment customer satisfaction, or refine your decision-making processes? The formulation of these objectives will furnish the metrics against which AI implementation can be gauged.

C. ASSESS REAL-TIME DATA NEEDS

Delve into the exigency of real-time visual insights within your business milieu. For industries characterized by dynamic and time-sensitive operational paradigms such as manufacturing, retail, and logistics, where expeditious decision-making is of the essence, the adoption of computer vision technology can proffer invaluable real-time data. The discernment of this exigency assumes paramount significance in the selection of the most apropos AI technology.

D. CONSIDER CONTENT GENERATION NEEDS

Conversely, undertake an evaluation of whether content generation, be it textual, visual, auditory, or other multimedia formats, constitutes a central requirement within your business framework. Generative AI stands as the paragon of efficacy in content creation. If content generation serves as a pivotal facet of your operational landscape, the incorporation of this technology is merited.

E. EVALUATE AVAILABLE RESOURCES

Methodically scrutinize the resources at your organization's disposal, encompassing budgetary considerations, technical

TABLE 3. Domain specific considersations.

Use Case	When to Use Generative AI	When to Use Computer Vision
E-commerce	• Personalized product recommendations based on user behavior and preferences. • Dynamic generation of product descriptions and marketing materials. • Content creation for virtual try-ons and product customization.	• Real-time product recognition for inventory management. • Visual search and object detection for product identification. • Monitoring and analyzing customer interactions with physical products in-store.
Manufacturing	• Design and generation of custom manufacturing prototypes. • Automating the creation of assembly instructions. • Virtual simulations and testing of product designs.	• Real-time quality control and defect detection on the production line. • Visual inspection of manufactured components for flaws or irregularities. • Monitoring equipment and machine performance through visual data.
Retail	• AI-powered chatbots for personalized customer assistance. • Creative content generation for marketing campaigns. • Virtual storefronts for online shopping experiences.	• Shelf monitoring for inventory tracking. • Facial recognition for customer identification and loyalty programs. • Foot traffic analysis and customer behavior insights.
Finance	• AI-generated financial reports and analysis. • Automated trading algorithms and predictive models. • Personalized financial advisory services.	• Fraud detection through real-time transaction analysis. • Visual data analysis for market trends and investment decisions. • Document scanning and data extraction for regulatory compliance.
Transportation	• Generating interactive and personalized travel itineraries. • AI-driven navigation and route optimization. • Creative content generation for travel advertisements.	• Real-time object detection for autonomous vehicles. • Traffic monitoring and congestion prediction. • Visual analysis for accident prevention and driver safety.
HealthTech	• AI-generated medical reports and patient summaries. • Personalized health advice and treatment plans. • Virtual medical assistants for remote patient monitoring.	• Medical image analysis for diagnosis and treatment planning. • Real-time monitoring of patient vital signs and anomalies. • Visual data for surgery assistance and precision medicine.

proficiency, and the existing technological infrastructure. The complexity of AI implementation can exhibit considerable variance, thereby necessitating an assurance of the requisite resources and support mechanisms to facilitate a seamless integration process. Contemplate whether the engagement of AI experts or collaboration with AI service providers is warranted.

F. PLAN FOR LONG-TERM SCALABILITY

Envisage the long-term implications of AI assimilation within your operational framework. How does AI dovetail with your business's overarching growth strategy? It is imperative that your chosen AI solution demonstrates scalability and adaptability to accommodate the evolving exigencies of your enterprise. Future-proofing your AI integration stands as an elemental prerogative to extract maximum dividends.

G. EXPLORE INDUSTRY-SPECIFIC SOLUTIONS

Embrace the prospect of industry-specific AI solutions calibrated to the idiosyncrasies of your sector. In the realm of healthcare, for instance, contemplate the exploration of AI applications tailored for patient diagnosis and treatment recommendations. Conduct thorough research into sector-specific use cases to unearth solutions in resonance with the particular requirements of your industry.

Table 3 provides an overview of when to consider Generative AI and when to opt for Computer Vision in various business use cases. It offers practical guidance for aligning AI strategies with specific needs and objectives respective of the domain.

Notably, Generative AI proves advantageous for content generation tasks, such as text, images, and videos, making it a valuable choice for creative industries, while Computer Vision excels in scenarios requiring visual analysis, object recognition, and image processing. Data availability is a crucial consideration, with Generative AI often necessitating large datasets for training, whereas Computer Vision can effectively process structured image data. Additionally, the complexity of the task and available expertise should be evaluated, as Generative AI may require deeper expertise in deep learning, while Computer Vision often benefits from established frameworks and libraries. Ultimately, the choice should align with desired outcomes, as Generative AI can generate novel content, while Computer Vision enhances visual understanding and recognition. These key takeaways from the table provide a concise summary for decision-makers seeking clarity in their AI implementation choices. By meticulously addressing these considerations and undertaking a holistic assessment of your business requisites, you are poised to execute a well-grounded choice in the selection of an AI solution that harmonizes seamlessly with your strategic objectives. The astute integration of AI technology bears the potential to instigate innovation, streamline operations, and amplify your competitive edge within the dynamic contours of the contemporary business milieu.

VIII. CONCLUSION: NAVIGATING THE AI LANDSCAPE FOR OPTIMAL BUSINESS IMPACT

As the business world undergoes a profound digital transformation, the rise of AI has captured the collective imagination. This surge in interest has been significantly fueled by the advent of generative AI, a technology capable of creating content and sparking innovative applications. Yet, amidst the excitement surrounding generative AI, another formidable AI variant, computer vision, remains equally noteworthy and transformative. Both AI technologies offer distinct advantages and hold the potential to revolutionize industries, but their applications and capabilities differ significantly.

Generative AI harnesses the power of machine learning to generate novel and creative content across various domains, from text and images to audio and more. It operates by analyzing vast datasets, learning intricate patterns, and producing content based on these learned insights. This technology has found success in applications such as content generation, creative projects, and natural language understanding. However, it is important to note that generative AI, despite its remarkable creative potential, lacks a genuine understanding of context and relies solely on patterns it has learned during training. It is reactive, offering responses based on past patterns but without memory of past interactions. Additionally, it is constrained by the quality and breadth of its training data, which can result in inaccuracies or unexpected outputs.

In contrast, computer vision is a field of AI dedicated to interpreting and comprehending visual information from the physical world, emulating human visual perception. It enables computers to detect, analyze, and understand visual data in real time, much like humans do. Computer vision can identify people, objects, and events, providing immediate insights into various processes and scenarios. This technology has proven invaluable across industries, enhancing safety, productivity, and efficiency. Unlike generative AI, computer vision does not generate new data or content; instead, it offers a critical extra set of eyes for businesses, providing real-time visual information for decision-making and process improvement. Its applications are diverse, spanning retail, manufacturing, restaurants, healthcare, automotive, agriculture, and more.

The choice between generative AI and computer vision should be driven by an organization's specific needs and objectives. Generative AI excels in creativity, content generation, and text-based applications. It is a valuable tool for businesses seeking to automate content creation, draft text, or engage in creative projects. On the other hand, computer vision shines when real-time visual insights are paramount. It offers immediate value by providing data-driven visual analysis, enhancing safety, optimizing processes, and improving decision-making across a wide range of industries. Whether monitoring occupancy in retail stores, identifying defects in manufacturing, or improving speed of service in restaurants, computer vision acts as a powerful ally in achieving operational excellence.

In this dynamic AI landscape, informed decision-making is crucial. Business leaders and decision-makers must thoroughly understand the capabilities, limitations, and applications of both generative AI and computer vision. By aligning AI adoption with specific organizational goals and challenges, businesses can harness these transformative technologies to drive innovation, streamline operations, and gain a competitive edge in a rapidly evolving digital era.

The strategic adoption of AI, whether generative or vision-based, has the potential to reshape industries, enhance customer experiences, and pave the way for a more efficient and intelligent future. As AI continues to mature and evolve, proactive businesses that navigate this landscape with clarity, purpose, and a deep understanding of their unique needs will emerge as leaders in the AI-driven revolution, shaping a brighter and more prosperous future for themselves and their stakeholders.

REFERENCES

[1] S. J. Russell, *Artificial Intelligence: A Modern Approach*. London, U.K.: Pearson, 2010.

[2] S. Russell and P. Norvig, "The history of artificial intelligence," in *Artificial Intelligence: A Modern Approach*, 4th ed. Hoboken, NJ, USA: Pearson, 2021, pp. 17–27.

[3] P. C. Bjork, "The master algorithm: How the quest for the ultimate learning machine will remake our world," *Perspect. Sci. Christian Faith*, vol. 68, no. 3, pp. 214–216, 2016.

[4] A. Kaplan and M. Haenlein, "Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence," *Bus. Horizons*, vol. 62, no. 1, pp. 15–25, Jan. 2019.

[5] D. Spencer, "Work in and beyond the second machine age: The politics of production and digital technologies," *Work, Employment Soc.*, vol. 31, no. 1, pp. 142–152, Feb. 2017.

[6] *Models GPT-3*, OpenAI, San Francisco, CA, USA, 2023.

[7] T. B. Brown et al., "Language models are few-shot learners," in *Proc. Adv. Neural Inf. Process. Sys.*, vol. 33, 2020, pp. 1877–1901.

[8] Z. Q. J. Lu, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2010.

[9] R. Szeliski, *Computer Vision: Algorithms and Applications*. Cham, Switzerland: Springer, 2022.

[10] H. Schütze, C. D. Manning, and P. Raghavan, *Introduction to Information Retrieval*, vol. 39. Cambridge, U.K.: Cambridge Univ. Press, 2008.

[11] S. Caner and F. Bhatti, "A conceptual framework on defining businesses strategy for artificial intelligence," *Contemp. Manage. Res.*, vol. 16, no. 3, pp. 175–206, Sep. 2020.

[12] D. B. Fogel, "Defining artificial intelligence," in *Machine Learning and the City: Applications in Architecture and Urban Design*, 2022, pp. 91–120.

[13] S. Samoili, M. L. Cobo, B. Delipetrev, F. Martinez-Plumed, E. Gomez, and G. De Prato, "AI watch. Defining artificial intelligence 2.0. towards an operational definition and taxonomy of AI for the AI landscape," JRC Res., Tech. Rep., JRC126426, 2021.

[14] R. Welsh, "Defining artificial intelligence," *SMPTE Motion Imag. J.*, vol. 128, no. 1, pp. 26–32, Jan. 2019.

[15] H. Wang, Z. Lei, X. Zhang, B. Zhou, and J. Peng, "Machine learning basics," *Deep Learn.*, pp. 98–164, 2016.

[16] X. Hao, G. Zhang, and S. Ma, "Deep learning," *Int. J. Semantic Comput.*, vol. 10, no. 3, pp. 417–439, 2016.

[17] M. Hussain, H. Al-Aqrabi, M. Munawar, and R. Hill, "Feature mapping for rice leaf defect detection based on a custom convolutional architecture," *Foods*, vol. 11, no. 23, p. 3914, Dec. 2022.

[18] B.-H. Li, B.-C. Hou, W.-T. Yu, X.-B. Lu, and C.-W. Yang, "Applications of artificial intelligence in intelligent manufacturing: A review," *Frontiers Inf. Technol. Electron. Eng.*, vol. 18, no. 1, pp. 86–96, 2017.

[19] P. K. Donepudi, "AI and machine learning in retail pharmacy: Systematic review of related literature," *ABC J. Adv. Res.*, vol. 7, no. 2, pp. 109–112, Nov. 2018.

[20] Y. Pan and L. Zhang, "Integrating BIM and AI for smart construction management: Current status and future directions," *Arch. Comput. Methods Eng.*, vol. 30, no. 2, pp. 1081–1110, Mar. 2023.

[21] M. Hussain, H. Al-Aqrabi, M. Munawar, R. Hill, and S. Parkinson, "Exudate regeneration for automated exudate detection in retinal fundus images," *IEEE Access*, vol. 11, pp. 83934–83945, 2022.

[22] M. Hussain, M. Dhimish, V. Holmes, and P. Mather, "Deployment of AI-based RBF network for photovoltaics fault detection procedure," *AIMS Electron. Electr. Eng.*, vol. 4, no. 1, pp. 1–18, 2020.

[23] N. P. Patel, D. R. Parikh, D. A. Patel, and R. R. Patel, "AI and web-based human-like interactive university chatbot (UNIBOT)," in *Proc. 3rd Int. Conf. Electron., Commun. Aerosp. Technol. (ICECA)*, Jun. 2019, pp. 148–150.

[24] F. Production Office, "Erratum: Motivation, social emotion, and the acceptance of artificial intelligence virtual assistants—Trust-Based mediating effects," *Frontiers Psychol.*, vol. 12, pp. 1–10, Oct. 2021.

[25] O. Kuusi and S. Heinonen, "Scenarios from artificial narrow intelligence to artificial general intelligence—Reviewing the results of the international work/technology 2050 study," *World Futures Rev.*, vol. 14, no. 1, pp. 65–79, Mar. 2022.

[26] W.-Y. Hsu and W.-Y. Lin, "Adaptive fusion of multi-scale YOLO for pedestrian detection," *IEEE Access*, vol. 9, pp. 110063–110073, 2021.

[27] T. Mahler, "Regulating artificial general intelligence (AGI)," in *Law and Artificial Intelligence: Regulating AI and Applying AI in Legal Practice*. Cham, Switzerland: Springer, 2022, pp. 521–540.

[28] K. S. Gill, "Artificial super intelligence: Beyond rhetoric," *AI Soc.*, vol. 31, no. 2, pp. 137–143, May 2016.

[29] A. Kar, A. Subash, and V. U. S. Rao, "Reactive artificial intelligence using big data in the era of precision medicine," *JAMA Surg.*, vol. 155, no. 7, p. 671, Jul. 2020.

[30] N. K. Upadhyay, H. Jiang, Z. Wang, S. Asapu, Q. Xia, and J. Joshua Yang, "Emerging memory devices for neuromorphic computing," *Adv. Mater. Technol.*, vol. 4, no. 4, Apr. 2019, Art. no. 1800589.

[31] M. M. Çelikok, T. Peltola, P. Daee, and S. Kaski, "Interactive AI with a theory of mind," 2019, *arXiv:1912.05284*.

[32] P. Baroni, F. Cerutti, D. Fogli, M. Giacomin, F. Gringoli, G. Guida, and P. Sullivan, "Self-aware effective identification and response to viral cyber threats," in *Proc. 13th Int. Conf. Cyber Conflict (CyCon)*, May 2021, pp. 353–370.

[33] H. Qiao, V. Liu, and L. Chilton, "Initial images: Using image prompts to improve subject representation in multimodal AI generated art," in *Proc. 14th Conf. Creativity Cognition*, Jun. 2022, pp. 15–28.

[34] S. Wang, L. Li, Y. Ding, and X. Yu, "One-shot talking face generation from single-speaker audio-visual correlation learning," in *Proc. AAAI Conf. Artif. Intell.*, vol. 36, 2022, pp. 2531–2539.

[35] J. L. Gillotte, "Copyright infringement in AI-generated artworks," *UC Davis L. Rev.*, vol. 53, p. 2655, Jan. 2019.

[36] L. Whittaker, T. C. Kietzmann, J. Kietzmann, and A. Dabirian, "'All around me are synthetic faces': The mad world of AI-generated media," *IT Prof.*, vol. 22, no. 5, pp. 90–99, Sep. 2020.

[37] J. Liu, Y. Wang, X. Huang, M. B. Korsós, Y. Jiang, Y. Wang, and R. Erdélyi, "Reliability of AI-generated magnetograms from only EUV images," *Nature Astron.*, vol. 5, no. 2, pp. 108–110, Feb. 2021.

[38] S. Awasthi, "Plagiarism and academic misconduct: A systematic review," *DESIDOC J. Library Inf. Technol.*, vol. 39, no. 2, 2019.

[39] K. Siau and W. Wang, "Artificial intelligence (AI) ethics: Ethics of AI and ethical AI," *J. Database Manage.*, vol. 31, no. 2, pp. 74–87, Apr. 2020.

[40] M. Chui, R. Roberts, and L. Yee, "Generative AI is here: How tools like ChatGPT could change your business," McKinsey, Chicago, IL, USA, Tech. Rep., 2022.

[41] E. Kasneci, K. Seßler, S. Küchemann, M. Bannert, D. Dementieva, F. Fischer, and U. Gasser, "ChatGPT for good? On opportunities and challenges of large language models for education," *Learn. Individual Differences*, vol. 103, Apr. 2023, Art. no. 102274.

[42] T. H. Kung, M. Cheatham, A. Medenilla, C. Sillos, L. De Leon, C. Elepaño, M. Madriaga, R. Aggabao, G. Diaz-Candido, J. Maningo, and V. Tseng, "Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models," *PLOS Digit. Health*, vol. 2, no. 2, Feb. 2023, Art. no. e0000198.

[43] L. Arango, S. P. Singaraju, and O. Niininen, "Consumer responses to AI-generated charitable giving ads," *J. Advertising*, vol. 52, no. 4, pp. 486–503, Aug. 2023.

[44] Y. Wong, S. Fan, Y. Guo, Z. Xu, K. Stephen, R. Sheoran, A. Bhamidipati, V. Barsopia, J. Liu, and M. Kankanhalli, "Compute to tell the tale: Goal-driven narrative generation," in *Proc. 30th ACM Int. Conf. Multimedia*, Oct. 2022, pp. 6875–6882.

[45] C. Guo, Y. Lu, Y. Dou, and F.-Y. Wang, "Can ChatGPT boost artistic creation: The need of imaginative intelligence for parallel art," *IEEE/CAA J. Autom. Sinica*, vol. 10, no. 4, pp. 835–838, Apr. 2023.

[46] G. Martínez-Arellano, R. Cant, and D. Woods, "Creating AI characters for fighting games using genetic programming," *IEEE Trans. Comput. Intell. AI Games*, vol. 9, no. 4, pp. 423–434, Dec. 2017.

[47] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *Proc. Int. Conf. Eng. Technol. (ICET)*, Aug. 2017, pp. 1–6.

[48] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed, and H. Arshad, "State-of-the-art in artificial neural network applications: A survey," *Heliyon*, vol. 4, no. 11, Nov. 2018, Art. no. e00938.

[49] A. Abu, N. H. Indra, A. H. A. Rahman, N. A. Sapiee, and I. Ahmad, "A study on image classification based on deep learning and TensorFlow," *Int. J. Eng. Res. Technol.*, vol. 12, no. 4, pp. 563–569, 2019.
[50] J. Du, "Understanding of object detection based on CNN family and YOLO," *J. Phys., Conf. Ser.*, vol. 1004, Apr. 2018, Art. no. 012029.
[51] R. Yang and Y. Yu, "Artificial convolutional neural network in object detection and semantic segmentation for medical imaging analysis," *Frontiers Oncol.*, vol. 11, Mar. 2021, Art. no. 638182.
[52] A. Zahid, M. Hussain, R. Hill, and H. Al-Aqrabi, "Lightweight convolutional network for automated photovoltaic defect detection," in *Proc. 9th Int. Conf. Inf. Technol. Trends (ITT)*, May 2023, pp. 133–138.
[53] B. Ataer Aydin, M. Hussain, R. Hill, and H. Al-Aqrabi, "Domain modelling for a lightweight convolutional network focused on automated exudate detection in retinal fundus images," in *Proc. 9th Int. Conf. Inf. Technol. Trends (ITT)*, May 2023, pp. 145–150.
[54] J. Xiang, T. Dong, R. Pan, and W. Gao, "Clothing attribute recognition based on RCNN framework using L-softmax loss," *IEEE Access*, vol. 8, pp. 48299–48313, 2020.
[55] B. Fang, X. Long, F. Sun, H. Liu, S. Zhang, and C. Fang, "Tactile-based fabric defect detection using convolutional neural network with attention mechanism," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–9, 2022.
[56] J. Cao, J. Zhang, and W. Huang, "Traffic sign detection and recognition using multi-scale fusion and prime sample attention," *IEEE Access*, vol. 9, pp. 3579–3591, 2021.
[57] P. M. Kyu and K. Woraratpanya, "Car damage detection and classification," in *Proc. 11th Int. Conf. Adv. Inf. Technol.*, Jul. 2020, pp. 1–6.
[58] H. Wei and N. Kehtarnavaz, "Semi-supervised faster RCNN-based person detection and load classification for far field video surveillance," *Mach. Learn. Knowl. Extraction*, vol. 1, no. 3, pp. 756–767, Jun. 2019.
[59] D. Hou, T. Liu, Y.-T. Pan, and J. Hou, "AI on edge device for laser chip defect detection," in *Proc. IEEE 9th Annu. Comput. Commun. Workshop Conf. (CCWC)*, Jan. 2019, pp. 247–251.
[60] V. K. Sharma and R. N. Mir, "Saliency guided faster-RCNN (SGFr-RCNN) model for object detection and recognition," *J. King Saud Univ.-Comput. Inf. Sci.*, vol. 34, no. 5, pp. 1687–1699, May 2022.
[61] C. Chen, S. Chandra, Y. Han, and H. Seo, "Deep learning-based thermal image analysis for pavement defect detection and classification considering complex pavement conditions," *Remote Sens.*, vol. 14, no. 1, p. 106, Dec. 2021.
[62] S. Hossain and D. Lee, "Deep learning-based real-time multiple-object detection and tracking from aerial imagery via a flying robot with GPU-based embedded devices," *Sensors*, vol. 19, no. 15, p. 3371, Jul. 2019.
[63] H. H. Shaheed and R. Aggarwal, "Wind turbine surface defect detection analysis from UAVs using U-Net architecture," in *Proc. Sci. Inf. Conf.* Cham, Switzerland: Springer, 2022, pp. 499–511.
[64] *A Faster-RCNN Based Chemical Fiber Paper Tube Defect Detection Method*. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8119386
[65] S. Kumaresan, K. S. J. Aultrin, S. S. Kumar, and M. D. Anand, "Deep learning-based weld defect classification using VGG16 transfer learning adaptive fine-tuning," *Int. J. Interact. Design Manuf. (IJIDeM)*, vol. 17, no. 6, pp. 2999–3010, Dec. 2023.
[66] W. Lan, J. Dang, Y. Wang, and S. Wang, "Pedestrian detection based on YOLO network model," in *Proc. IEEE Int. Conf. Mechatronics Autom. (ICMA)*, Aug. 2018, pp. 1547–1551.
[67] M. Hussain and R. Hill, "Custom lightweight convolutional neural network architecture for automated detection of damaged pallet racking in warehousing & distribution centers," *IEEE Access*, vol. 11, pp. 58879–58889, 2023.
[68] M. Hussain, H. Al-Aqrabi, and R. Hill, "PV-CrackNet architecture for filter induced augmentation and micro-cracks detection within a photovoltaic manufacturing facility," *Energies*, vol. 15, no. 22, p. 8667, Nov. 2022.
[69] B. Santra and D. P. Mukherjee, "A comprehensive survey on computer vision based approaches for automatic identification of products in retail store," *Image Vis. Comput.*, vol. 86, pp. 45–63, Jun. 2019.
[70] D. Beymer, "Person counting using stereo," in *Proc. Workshop Hum. Motion*, Dec. 2000, pp. 127–133.
[71] M. Hawkins, "Metaverse live shopping analytics: Retail data measurement tools, computer vision and deep learning algorithms, and decision intelligence and modeling," *J. Self-Governance Manage. Econ.*, vol. 10, no. 2, pp. 22–36, 2022.
[72] P. Cortez, L. M. Matos, P. J. Pereira, N. Santos, and D. Duque, "Forecasting store foot traffic using facial recognition, time series and support vector machines," in *Proc. Int. Conf. Eur. Transnational Educ.* Cham, Switzerland: Springer, 2017, pp. 267–276.
[73] N. James, "Automated checkout for stores: A computer vision approach," *Revista Gestão Inovação Tecnologias*, vol. 11, no. 3, pp. 1830–1841, Jun. 2021.
[74] M. Nagy and G. Lăzăroiu, "Computer vision algorithms, remote sensing data fusion techniques, and mapping and navigation tools in the industry 4.0-based Slovak automotive sector," *Mathematics*, vol. 10, no. 19, p. 3543, Sep. 2022.
[75] S. Battiato, S. Conoci, R. Leotta, A. Ortis, F. Rundo, and F. Trenta, "Benchmarking of computer vision algorithms for driver monitoring on automotive-grade devices," in *Proc. AEIT Int. Conf. Electr. Electron. Technol. Automot.*, Alessandro Ortis, France, Nov. 2020, pp. 1–6.
[76] J. F. S. Gomes and F. R. Leta, "Applications of computer vision techniques in the agriculture and food industry: A review," *Eur. Food Res. Technol.*, vol. 235, no. 6, pp. 989–1000, Dec. 2012.
[77] M. S. Uddin and J. C. Bansal, *Computer Vision and Machine Learning in Agriculture*. Cham, Switzerland: Springer, 2022.
[78] T. A. Shaikh, W. A. Mir, T. Rasool, and S. Sofi, "Machine learning for smart agriculture and precision farming: Towards making the fields talk," *Arch. Comput. Methods Eng.*, vol. 29, no. 7, pp. 4557–4597, Nov. 2022.
[79] N. Lyons, "Deep learning-based computer vision algorithms, immersive analytics and simulation software, and virtual reality modeling tools in digital twin-driven smart manufacturing," *Econ., Manage., Financial Markets*, vol. 17, no. 2, pp. 67–81, 2022.
[80] L. Scime and J. Beuth, "Anomaly detection and classification in a laser powder bed additive manufacturing process using a trained computer vision algorithm," *Additive Manuf.*, vol. 19, pp. 114–126, Jan. 2018.
[81] Q.-J. Zhao, P. Cao, and D.-W. Tu, "Toward intelligent manufacturing: Label characters marking and recognition method for steel products with machine vision," *Adv. Manuf.*, vol. 2, no. 1, pp. 3–12, Mar. 2014.
[82] S. Paneru and I. Jeelani, "Computer vision applications in construction: Current state, opportunities & challenges," *Autom. Construct.*, vol. 132, Dec. 2021, Art. no. 103940.
[83] K. Li, E. D. Miller, M. Chen, T. Kanade, L. E. Weiss, and P. G. Campbell, "Computer vision tracking of stemness," in *Proc. 5th IEEE Int. Symp. Biomed. Imag., Nano Macro*, May 2008, pp. 847–850.
[84] A. El-Komy, O. R Shahin, R. M. A. El-Aziz, and A. I. Taloba, "Integration of computer vision and natural language processing in multimedia robotics application," *Inf. Sci.*, vol. 7, no. 6, pp. 1–12, 2022.
[85] F. Safarov, F. Akhmedov, A. B. Abdusalomov, R. Nasimov, and Y. I. Cho, "Real-time deep learning-based drowsiness detection: Leveraging computer-vision and eye-blink analyses for enhanced road safety," *Sensors*, vol. 23, no. 14, p. 6459, Jul. 2023.
[86] Q. Lu and B. Murmann, "Improving the energy efficiency and robustness of TinyML computer vision using log-gradient input images," 2022, *arXiv:2203.02571*.

MUHAMMAD HUSSAIN received the B.Eng. degree in electrical and electronic engineering, the M.S. degree in the Internet of Things from the University of Huddersfield, Charlottesville, in 2019, and the Ph.D. degree in artificial intelligence for defect identification, in 2022. He is currently an Accomplished Researcher hailing from Dewsbury, West Yorkshire, U.K. His work contributes to optimizing PV systems' efficiency and reliability. He is equally passionate about machine vision, with a focus on lightweight architectures for edge device deployment in real-world production settings. Beyond fault detection, he explores AI interpretability, concentrating on developing explainable AI for medical and healthcare applications. His interdisciplinary approach underscores his commitment to ethical and impactful AI solutions. With his diverse expertise spanning AI, fault detection, machine vision, and interpretability, he aims to leave his mark on shaping the future of technology and its positive influence on society. His research interest includes fault detection, particularly microcracks on photovoltaic (PV) cells due to mechanical and thermal stress.