

Research Article

# Towards a Human-Like AGI Architecture: General Intelligence Framework (GIF)

Jiran Kurian Puliyanmakkal and Rohini V

Department of Computer Science, Christ (Deemed to be University), Bangalore, India

Article history

Received: 28-10-2024

Revised: 24-01-2025

Accepted: 14-02-2025

Corresponding Author:

Jiran Kurian Puliyanmakkal

Department of Computer Science,

Christ (Deemed to be University),

Bangalore, India

Email:

jiran.kurian@res.christuniversity.in

**Abstract:** Artificial Intelligence (AI) has achieved significant breakthroughs but remains limited by its specialization and inability to generalize across domains, unlike human cognition. Current models such as Large Language Models (LLMs) and Multimodal Large Language Models (MLLMs) excel at specific tasks but struggle with real-time adaptability and cross-domain generalization. This paper introduces the General Intelligence Framework (GIF), an approach designed to bridge this gap by mimicking human-like cognitive processes. By integrating Deep Learning (DL), Spiking Neural Networks (SNNs), and neuromorphic hardware, the framework fosters Real-Time Learning (RTL) and adaptability. The proposed framework holds potential for industries like robotics, healthcare, education, astronomy, defense, autonomous systems, etc., where flexible, adaptive AI is critical. We hypothesize that the framework will enable AI systems to handle unforeseen inputs and tasks without requiring extensive retraining, representing a step toward achieving Artificial General Intelligence (AGI).

**Keywords:** Artificial General Intelligence (AGI), General Intelligence Framework (GIF), Spiking Neural Networks (SNNs), Neuromorphic Computing, Real-Time Learning (RTL)

## Introduction

AI has brought transformative changes across various industries, including healthcare, robotics, education, business analytics, etc.,. While AI has demonstrated remarkable success in narrow applications like healthcare diagnostics and autonomous driving, a critical limitation persists AI systems cannot autonomously apply learned knowledge to new tasks or adapt to unforeseen scenarios (Job *et al.*, 2023). For instance, an autonomous vehicle trained to navigate urban environments may struggle to adapt when tasked with off-road driving, a scenario requiring rapid generalization. Similarly, a medical AI trained on a specific dataset may perform poorly when diagnosing diseases with variations not included in its training data. This rigidity contrasts sharply with human intelligence, which excels at drawing from past experiences to solve novel and diverse problems with minimal effort.

Previous efforts to develop AGI have struggled to bridge the gap between narrow AI and true general intelligence, particularly in systems that require adaptability across domains without significant retraining. Prominent models, such as Google's Gemini and OpenAI's GPT-4, have made strides in multimodal learning, yet they remain constrained by architectural

rigidity, limiting their ability to generalize knowledge and adapt in real-time.

While contemporary AI models like LLMs and MLLMs have advanced in handling multiple forms of input—such as text, audio, images, and videos—they remain fundamentally restricted by their rigid architectures (Zhang *et al.*, 2024). These models are optimized for specific tasks and struggle when confronted with novel input types, such as sensory data from previously unknown devices. Adapting these models to handle new forms of input often requires a complete overhaul of their architecture, limiting their ability to function in real-world, dynamic environments where flexibility and adaptability are essential.

A key reason for this limitation is the reliance of current AI models on vast amounts of domain-specific data. These systems generate outputs based on learned statistical patterns within the data, rather than by generalizing knowledge in the way humans do. Humans can rapidly adapt to new situations by relating them to past experiences, thereby learning efficiently with minimal data. For example, if a human sees a ball being thrown, they can instinctively predict its trajectory based on their prior understanding of physics and motion. Similarly, when witnessing a cannon fire for the first time, a person can quickly estimate the distance the

cannonball will travel by drawing on their knowledge of similar patterns. This ability to connect new experiences with existing knowledge allows humans to generalize learning across different domains.

In contrast, LLMs and MLLMs lack this human-like capability. Their understanding is confined to the patterns in the data they were trained on and they struggle to autonomously validate new information by connecting it to past experiences. This dependency often results in errors, overfitting, or hallucinations when these models encounter inputs outside their training data (Birhane *et al.*, 2023; Job *et al.*, 2023). This dependency on vast amounts of data and lack of experiential validation highlights a significant limitation in their current architecture (Naveed *et al.*, 2023; Ullah *et al.*, 2024). Moreover, as recent studies have shown, AI models experience performance saturation, where continual optimization yields diminishing returns. These models tend to become increasingly specialized and overfitted to specific benchmarks, limiting their ability to generalize knowledge across domains—one of the core requirements for achieving true AGI (Ott *et al.*, 2022).

To address these limitations, the AI research community must shift towards developing systems that mirror the adaptability and versatility of human intelligence. Unlike current AI systems, humans excel at synthesizing information from multiple sensory inputs and drawing upon past experiences to make informed decisions across various contexts (Rachel & MIT, 2024). This capability underscores the need for AI systems that can gather, analyze, and synthesize diverse types of information, recognize patterns, and apply knowledge from previous tasks to new situations without requiring re-architecture.

AGI represents the next frontier in AI research. AGI systems must autonomously adapt to new challenges, generalizing knowledge across a broad range of tasks. AGI aims to replicate the human brain's versatility and cognitive abilities, enabling systems to learn from diverse experiences, integrate various forms of input, and apply that knowledge to solve complex problems.

This study extends the work outlined in “A Deliberation on the Stages of Artificial Intelligence” (Kurian & Rohini, 2021), focusing on the Stage of General Intelligence. We introduce a novel GIF designed to bridge the gap between narrow AI and true general intelligence. This framework integrates DL and Brain-Inspired Learning, incorporating SNNs that mimic the brain's adaptability and RTL capabilities. The modules are designed to generalize knowledge across domains, adapt to new tasks, and incorporate feedback in real-time, much like the human brain learns from experience.

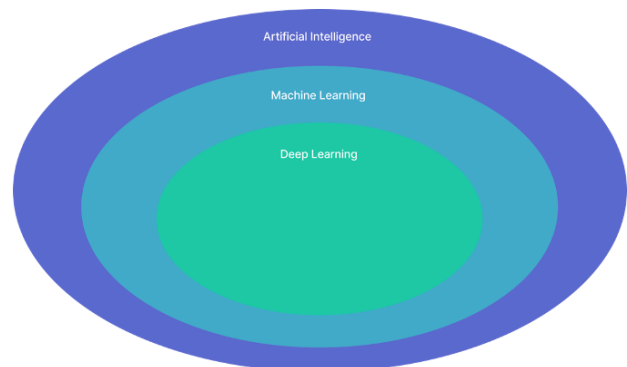
Despite significant advancements, current AI systems like Large Language Models (LLMs) and Multimodal Large Language Models (MLLMs) often exhibit rigidity,

struggling to adapt to new or unexpected tasks without extensive retraining. This study proposes the General Intelligence Framework (GIF) to address these limitations by integrating Real-Time Learning (RTL), Spiking Neural Networks (SNNs), and neuromorphic hardware for continuous, human-like adaptability. Our aim is to bridge the gap between narrow AI applications and Artificial General Intelligence (AGI), providing a modular design that seamlessly incorporates novel inputs and outputs while minimizing reconfiguration.

**Unique Contributions of the GIF.** The proposed framework introduces three primary innovations: (1) Real-Time Learning (RTL): GIF updates its internal parameters immediately upon receiving new inputs, akin to how biological neurons respond to stimuli in real time; (2) Spiking Neural Networks (SNNs): A biologically inspired architecture that efficiently processes temporal data while facilitating rapid adaptation to novel tasks; and (3) Modular Input-Output Design: A flexible architecture that can integrate additional sensory or output modules—such as cameras, LiDAR, or robotic actuators—without extensive re-architecting. Together, these innovations aim to achieve a level of adaptability and generalization closer to human cognitive processes than current AI paradigms.

### *Hierarchical Representation of Artificial Intelligence*

AI is a broad field that encompasses a wide range of technologies and approaches. At its core, AI refers to machines capable of performing tasks that typically require human intelligence, such as learning, reasoning, problem-solving, and understanding natural language. Within this expansive field, two primary subfields—Machine Learning (ML) and DL—have gained prominence due to their significant contributions to the advancement of AI. These concepts are often visualized as a set of nested ellipses, where AI forms the outermost layer, ML resides within AI and DL lies at the heart of ML. This hierarchical representation reflects the growing sophistication and specialization of AI technologies as we move from broader to more specific methods Figure (1).



**Fig. 1:** Hierarchical representation of artificial intelligence

## Artificial Intelligence

Initially, AI was the overarching concept involving building systems capable of mimicking human intelligence. It encompasses a broad range of applications, including rule-based systems, decision trees, robotics, expert systems, and more. AI has been applied in various fields, such as autonomous vehicles, healthcare diagnostics, financial fraud detection, and customer service chatbots.

A classic example of AI in action was the rule-based expert system. These systems are programmed with a set of rules that allow them to make decisions based on specific inputs. For instance, in a medical diagnostic system, AI can analyze a patient's symptoms and cross-reference them with a database of medical knowledge to suggest possible diagnoses.

However, while rule-based AI systems can perform specific tasks, they are rigid and lack the ability to learn from new data. This limitation paved the way for ML, which takes AI a step further by enabling systems to learn from experience and improve over time.

## Machine Learning (ML)

ML is a subset of AI focused on enabling systems to learn from data rather than relying solely on pre-programmed rules. ML algorithms analyze patterns in data, allowing the system to make predictions or decisions without explicit human intervention. The more data the system is exposed to, the better it becomes at identifying patterns and making informed predictions. ML has seen widespread use in recommendation systems, such as Netflix or Amazon's recommendation engines, as well as in autonomous vehicles, where it helps systems recognize objects like pedestrians or stop signs.

ML provides significant advantages over traditional rule-based AI by enabling systems to adapt and improve through experience. However, the effectiveness of ML is often limited by the quality and quantity of data available for training. As ML algorithms become more complex, they require more sophisticated methods for learning, leading to the emergence of DL.

## Deep Learning (DL)

DL is a subset of ML inspired by the structure and function of the human brain. DL models, known as neural networks, consist of layers of artificial neurons that work together to process and learn from large amounts of data. These models excel at recognizing patterns in unstructured data, such as images, audio, and text, making them particularly effective in tasks like image recognition, speech recognition, and NLP.

One of the key features of DL is its ability to automatically extract features from raw data. For example, in image recognition tasks, early layers of the

neural network might learn to recognize simple features like edges and textures, while deeper layers learn more complex features like shapes and objects. This ability to "learn features" makes DL particularly powerful for tasks that involve complex and high-dimensional data.

While DL models are highly effective, they also come with challenges. These models require massive amounts of data and computational resources to train and their complexity can lead to overfitting or performance saturation. Additionally, despite their ability to recognize patterns, DL models still struggle with tasks that require more generalized, human-like reasoning and adaptability, especially when faced with novel inputs or scenarios outside their training data.

## Literature Review

Recent advancements in LLMs, such as OpenAI's GPT-4, have brought significant breakthroughs in Natural Language Processing (NLP) and multimodal tasks. However, despite their impressive performance on specialized tasks, LLMs face key limitations that hinder the pursuit of true AGI. While these models excel at recognizing patterns within vast datasets, they struggle to exhibit the adaptability required for general intelligence.

One of the critical shortcomings of LLMs is their reliance on memorization and statistical associations rather than cognitive reasoning. For instance, while GPT-4 can effectively predict the next word in a sentence based on learned patterns, it struggles when tasked with novel or counterfactual scenarios. The paper "Reasoning Skills of Large Language Models Are Often Overestimated" underscores this limitation, demonstrating that these models often fall short when required to engage in complex reasoning or when presented with unfamiliar tasks (Rachel & MIT, 2024).

Recent work on Large Language Models (LLMs) demonstrates diminishing returns when scaling model size beyond certain thresholds. For instance, after a certain point, increasing the number of training tokens only improves accuracy by a marginal 1–2%, despite substantially higher computational and data requirements (Ott *et al.*, 2022). Moreover, these models often struggle when faced with out-of-distribution inputs; error rates can rise dramatically—sometimes by 10% or more—when the prompts deviate from familiar training contexts (Birhane *et al.*, 2023). Such limitations underscore the need for architectures that can adapt in real time without relying solely on memorized patterns.

In contrast, research on SNNs suggests that spike-based processing can reduce energy consumption by 15–20% compared to standard deep learning models in robotics applications (Madrenas *et al.*, 2023). By focusing on event-driven spikes rather than continuous activations, SNNs are inherently more efficient for time-sensitive tasks, forming the backbone of the GIF's real-time adaptability.

Beyond their reliance on vast datasets, LLMs also face challenges in maintaining contextual understanding over extended interactions. As detailed in “The Working Limitations of Large Language Models,” these models tend to lose coherence in long conversations, limiting their utility in tasks that require sustained reasoning and adaptability (Job *et al.*, 2023).

These challenges highlight the fundamental gap between the capabilities of current AI models and the requirements for achieving true AGI. The reliance on large datasets and the difficulty in generalizing knowledge across different domains remain significant barriers to progress.

### *Alternative AGI Approaches*

The pursuit of AGI has inspired a range of methodologies, each attempting to overcome the fundamental limitations of existing AI systems. While the GIF introduces real-time adaptability and neuromorphic processing through SNNs, several other approaches provide important context for the evolution of AGI. Here we compare some of these alternatives, highlighting their strengths and limitations.

#### *Reinforcement Learning and MuZero*

One prominent approach in AGI research is MuZero, a model developed by DeepMind that employs reinforcement learning to master games such as Go, Chess, and Shogi without prior knowledge of the game rules. MuZero’s ability to learn and plan within an environment demonstrates remarkable flexibility in handling complex tasks (Schrittwieser *et al.*, 2020). However, MuZero’s proficiency is confined to specific, structured domains, requiring substantial tuning for optimal performance in each new task or environment. This specialization limits its generalization capabilities, which is a critical goal in AGI.

In contrast, the GIF emphasizes autonomous adaptability across a wide range of tasks. By leveraging real-time experiential learning through SNNs, GIF aims to dynamically adjust to novel inputs and tasks without the need for extensive reconfiguration, offering a more generalized approach to learning that MuZero lacks.

#### *Neuro-Symbolic AI*

Neuro-Symbolic AI represents a hybrid approach that attempts to combine the strengths of symbolic reasoning and neural networks. This model merges logic-based systems, which excel at structured reasoning, with the statistical pattern recognition capabilities of DL models (Brachman & Levesque, 2022). By blending these techniques, Neuro-Symbolic AI aims to achieve more generalizable reasoning and improved domain flexibility.

While Neuro-Symbolic AI offers a promising middle ground, its real-time adaptability remains limited. In scenarios requiring dynamic adjustment to new data or

environments, Neuro-Symbolic models often lag due to their reliance on predefined logical structures. By comparison, the GIF leverages neuromorphic computing and real-time sensory feedback, which allow it to process novel inputs as they occur and adjust its behavior in real time without requiring re-engineering. This capability makes GIF more suited for real-world applications where adaptability is key.

#### *Transformers and LLMs*

Transformers, such as those used in models like GPT-4, have revolutionized NLP by enabling more effective handling of sequential data and long-term dependencies. These models have been extended into multimodal architectures, which process not only text but also images, audio, and other inputs (Vaswani *et al.*, 2017). While these models represent a significant step forward, their dependence on large, domain-specific datasets and performance saturation pose substantial barriers to achieving AGI.

One of the critical limitations of transformers and LLMs is their need for task-specific fine-tuning when introduced to novel environments. Despite multimodal learning advancements, these systems often struggle to generalize across domains without significant retraining. The GIF addresses this limitation by employing a modular architecture that allows it to seamlessly incorporate new sensory inputs without altering its core architecture. This modularity, combined with real-time adaptability, ensures that the GIF can handle diverse tasks with greater efficiency and flexibility than current transformer-based models.

#### *Spiking Neural Networks (SNNs) and Neuromorphic Computing*

SNNs, an integral part of the GIF, offer a distinct advantage over traditional DL models by mimicking the spike-based communication in biological neurons. SNNs are particularly well-suited for processing temporal data and making real-time decisions, such as those required in robotics or autonomous systems. Additionally, when implemented on neuromorphic hardware, SNNs can perform tasks with remarkable energy efficiency, making them ideal for low-power, real-time applications.

While traditional models like transformers require vast computational resources and extensive training, SNNs, through neuromorphic hardware, can operate more efficiently by processing data in real time. This shift towards energy-efficient, real-time adaptability sets the GIF apart from other AGI approaches.

#### *Challenges in Scaling Traditional Deep Learning Architectures*

Traditional DL architectures, particularly transformers, have been the foundation of many recent AI advancements. However, these models face

significant challenges, particularly related to scalability. As transformers grow in complexity and size, their computational resource requirements escalate, leading to diminishing returns in performance improvements a phenomenon known as performance saturation (Ott *et al.*, 2022). This limitation presents a challenge for models such as GPT-4, which struggle with generalization when processing novel data outside of their training set.

Neuromorphic computing, employed in the GIF, offers a more energy-efficient and scalable alternative. By mimicking the spiking behavior of biological neurons, neuromorphic hardware enables the framework to process data in real-time with lower power consumption. The GIF's integration of SNNs is a crucial differentiator, allowing it to overcome the limitations of traditional DL models by enhancing real-time adaptability while maintaining efficiency.

### *Multimodal Learning in AGI*

Recent advancements in MLMs, such as Google's Gemini, have demonstrated the ability to integrate diverse forms of input, including text, audio, and images, to handle more complex real-world tasks (Zhang *et al.*, 2024). These multimodal systems provide broader applicability, yet they still rely heavily on predefined architectures that are optimized for handling specific types of inputs.

The GIF builds on these ideas but takes a step further by enabling the seamless integration of new, previously unseen sensory inputs. Unlike multimodal models, which require task-specific adjustments, the modular input-output system in the GIF allows it to dynamically process and synthesize any form of data without reconfiguration. This adaptability makes the framework more suitable for AGI, where flexibility and RTL across diverse environments are crucial for success.

One of the promising avenues is the development of SNNs, which emulate the spiking behavior of biological neurons and offer potential advantages in energy efficiency and real-time processing. The paper "Advancements in Algorithms and Neuromorphic Hardware for Spiking Neural Networks" discusses how neuromorphic hardware can support the implementation of SNNs, providing a more efficient alternative to traditional neural networks (Javanshir *et al.*, 2022). However, despite these advancements, SNNs still face challenges in achieving the same level of accuracy and scalability as traditional LLMs, particularly in complex tasks.

The integration of AI-native memory systems with LLMs has also been proposed as a solution to enhance long-term reasoning and decision-making capabilities. The paper "AI-native Memory: A Pathway from LLMs Towards AGI" introduces a novel architecture that combines memory modules with LLMs, allowing them to store and retrieve important conclusions from past

interactions. While this approach shows promise in improving the generalization capabilities of LLMs, it also raises concerns about the complexity and feasibility of implementing such systems at scale (Shang *et al.*, 2024).

### *Need for Introducing the General Intelligence Framework*

Given the limitations of existing AI models, there is a clear need for a new framework that can overcome these challenges and pave the way for more adaptable and generalizable AI systems. The proposed GIF aims to integrate diverse learning methods, such as DL, brain-inspired learning, and real-time experiential learning, to create a more holistic AI system capable of generalizing across domains and adapting to novel inputs.

One of the key components of this framework is the incorporation of neuromorphic computing, which offers the potential to enhance the adaptability and efficiency of AI systems. The paper "Neuromorphic Computing Facilitates Deep Brain-Machine Fusion for High-Performance Neuroprosthesis" highlights the advantages of neuromorphic computing in creating more biologically plausible models that can interact seamlessly with human cognitive processes (Qi *et al.*, 2023). This approach is particularly relevant for developing AI systems that can learn and adapt in real-time, much like the human brain.

RTL, a concept still in its early stages, is another critical component of the GIF. Although research on RTL in AI is limited, the potential benefits of this approach are significant, particularly in dynamic environments where adaptability and quick decision-making are crucial. We will be explaining this in detail in our follow-up paper called "Deep Understanding".

In summary, the current landscape of AI research highlights both the remarkable advancements and the significant limitations of existing models like LLMs. While approaches such as multimodal integration, neuromorphic computing, and AI-native memory systems offer promising solutions, there is a clear need for a more comprehensive framework that can address the challenges of generalization, adaptability, and RTL. The proposed GIF aims to bridge this gap by integrating diverse learning methods and cognitive science principles, paving the way for the next generation of AI systems capable of achieving true general intelligence.

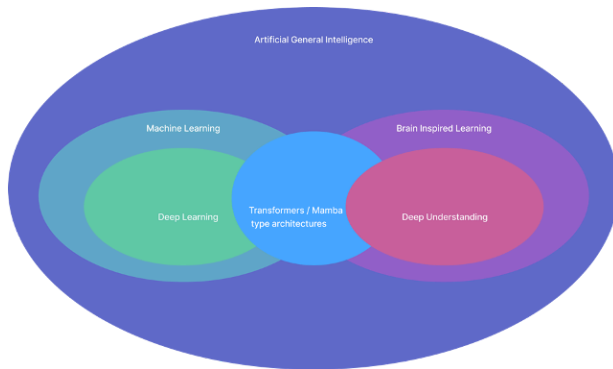
The following sections will delve deeper into the proposed framework, outlining its key components and the potential impact on the future of AI research and development. In comparison to existing hierarchical models, such as those used in the Gemini family and GPT-4o, which primarily focus on enhancing multimodality and efficiency, the proposed framework extends these ideas by incorporating RTL and brain-inspired processing, offering a more robust approach to achieving AGI.

## Introduction to the General Intelligence Framework

### Transition to Hierarchical Representation for Artificial General Intelligence

AGI refers to an AI system that can learn and apply new skills across diverse tasks without relying on prior specialized training. It involves generalizing knowledge to handle novel problems, demonstrating a level of reasoning and abstraction similar to human intelligence (ARC, 2024). As we strive to bridge the gap between Artificial Narrow Intelligence and AGI, it is essential to rethink the current hierarchical representation of AI, incorporating new concepts and developing frameworks that enable flexibility, adaptability, and cross-domain learning.

This section introduces a novel representation, emphasizing a hierarchical representation for AGI that integrates ML, Brain-Inspired Learning, and a novel concept called Deep Understanding (DU). Additionally, this framework incorporates neuromorphic computing, enabling AGI systems to mimic human cognitive functions more closely while also addressing the unique processing capabilities required for RTL and adaptation Figure (2).



**Fig. 2:** Hierarchical representation for artificial general intelligence

This new hierarchical representation for AGI redefines the organization of AI components, reflecting a shift towards more generalizable and adaptable systems. Unlike the traditional hierarchy where ML and DL are the key pillars of AI, the new structure expands the framework to better mirror the human brain's complexity and functionality. This transition to a new hierarchical representation for AGI involves the following components:

#### Artificial General Intelligence as the Outermost Layer

- AGI represents the broad field that encompasses all intelligent systems. It includes both classical AI approaches (rule-based, expert systems) and modern techniques like ML and brain-inspired models

- In this new hierarchy, AGI is the outermost ellipse, capturing the general capabilities and scope of artificial systems that perform tasks requiring intelligence

#### Machine Learning Layer

ML: This core component inside AI focuses on statistical and algorithmic learning from data. It includes DL, which is nested within ML, representing the deep neural networks that learn complex patterns and features from large datasets. DL models are powerful but have limitations in generalization, often requiring vast amounts of training data and task-specific fine-tuning.

#### The Transition Layer

Situated between ML and Brain-Inspired Learning is the Transition Layer, an essential part of the AGI hierarchical representation. Architectures such as Transformers and Mamba serve as key enablers of this transition. In the ML domain, Transformers have revolutionized language models by allowing the processing of sequential data and understanding long-term dependencies. These models represent a shift towards more adaptable architectures that can handle varied tasks, aligning with the goals of AGI.

#### Brain-Inspired Learning (BIL) Layer

- The second core component is Brain-Inspired Learning (BIL). This field draws from neuroscience to build AI models that mimic the human brain's adaptability and cognitive processes. Unlike ML, BIL is not purely algorithmic but also incorporates neuromorphic computing a hardware paradigm inspired by the brain's structure and function. This hardware aspect is crucial, as it allows AI systems to process information in real-time and in a more energy-efficient manner
- Deep Understanding: Inside Brain-Inspired Learning lies the concept of DU, a novel idea introduced in this study. In contrast to DL, which primarily focuses on pattern recognition, DU emphasizes SNNs and the incorporation of both learned knowledge and real-time sensory inputs. DU specifically addresses real-world experiences and learns by connecting newly acquired data with previously understood patterns. The hardware component of neuromorphic computing enables this process by mirroring the brain's neurons and synapses. The focus here is not just on learning but on building an understanding that can generalize across different contexts and domains

#### Defining Real-Time Learning (RTL) and Deep Understanding (DU)

In this study, Real-Time Learning (RTL) refers to an AI system's ability to update its internal parameters—

such as synaptic weights in SNNs—on a millisecond-to-second timescale, without pausing for a dedicated training phase. This approach mirrors the immediate synaptic plasticity observed in biological neurons (Park and Choi, 2024).

Deep Understanding (DU) builds upon RTL by not only recognizing patterns but also forming semantic connections between new inputs and previously learned experiences. Unlike traditional deep learning, which can overfit static training data, DU emphasizes dynamic, context-rich processing. In this sense, the system “understands” new stimuli by continuously relating them to prior knowledge, enabling more robust generalization and adaptability.

### General Intelligence Framework

The proposed framework for AGI is designed to overcome the limitations of current AI systems by mimicking the adaptability and comprehensive learning capabilities of the human brain. This framework aims to enable AI systems to learn from diverse types of data and apply their knowledge across various domains. Current AI systems, like GPT-4o, demonstrate significant improvements in multimodal reasoning but still require extensive fine-tuning for new tasks. In contrast, the proposed AGI framework is designed to adapt to new tasks autonomously, leveraging RTL and SNNs to mimic the flexibility and adaptability of the human brain.

For instance, in the paper “LM-Nav: Robotic Navigation with Large Pre-Trained Language, Vision and Action Models” (Shah *et al.*, 2022), a robot is able to perform multiple tasks, such as comprehending high-level textual instructions for navigation within a real-world environment using visual observation. This is achieved by integrating different models: A Language Model, a Vision and Language Model, and a Visual Navigation Model. However, even though LM-Nav can handle different tasks like comprehending high-level instructions or understanding its environment, it cannot be easily retrained or adapted to perform entirely new tasks, such as acting as a personal assistant. This lack of adaptability underscores the need for a GIF.

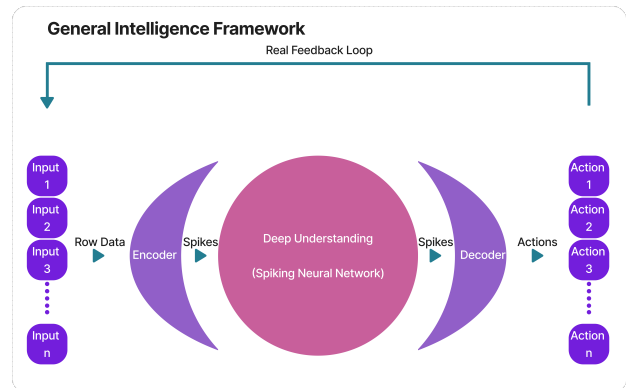
A GIF should incorporate several key characteristics to handle diverse tasks and adapt to new environments effectively, for instance:

- The system should understand the environment it is integrated into and adapt accordingly. This includes processing various types of sensory inputs and executing the necessary actions based on these inputs. For example, in the context of a self-driving car, the system should learn and adapt to the car’s functionalities, such as acceleration, steering, and braking, based on sensory inputs like camera feeds, LIDAR data, and GPS signals.
- The system should memorize its experiences and apply this knowledge when learning new tasks. For

example, if the system learns to ride a bicycle, it should be able to transfer that knowledge when learning to ride a motorbike, applying its understanding of balance and braking in a new context.

- The framework must incorporate RTL, allowing the system to learn from its own experiences as they happen. This concept will be explored further in subsequent sections.

The framework draws inspiration from how the human brain operates, particularly its ability to adapt to new situations, including the integration of prosthetic devices. Recent advancements in fields such as neuroprosthetics, neuroplasticity, and Brain-Machine Interfaces (BMIs) have provided valuable insights into developing a generic framework for building AGI. Unlike current AI models, which often require re-architecting to handle new tasks or inputs, the human brain can interact with new devices and sensory inputs without significant modifications. This adaptability is due to the brain’s ability to convert all types of sensory inputs (e.g., visual, auditory, tactile) into electrical signals, which neurons can process.



**Fig. 3:** General Intelligence Framework (GIF). Inputs are labeled as (A) Input 1 (Visual Input), (B) Input 2 (Auditory Input), (C) Input 3 (Sensor Data), etc. The arrow on the top indicates the continuous feedback loop for Real-Time Learning. Outputs specify whether they are physical actions (e.g., robotic arm movement, vehicle steering) or cognitive outputs (e.g., textual decisions, alerts). Fonts have been enlarged for clarity

For example, in the field of prosthetic technology, the brain is able to interact with new types of devices without modifying its neurons. Instead, it simply sends and receives signals that the brain can adapt to, regardless of the device type. The proposed GIF replicates this process. Inputs, regardless of type, are converted into a standardized format that the core processing module understands, much like how the brain converts sensory inputs into electrical signals for processing. A paper called Neuromorphic computing facilitates deep brain-machine fusion for high-performance neuroprosthesis discusses how

neuromorphic computing models mimic biological nervous systems to achieve deeper integration between the brain and prosthetic devices (Qi *et al.*, 2023).

In order to qualify an AI model as compatible with the GIF, the system should be adaptable enough to accept any type of input and provide any type of output (actions) without re-engineering the core architecture. The input and outputs can be attached or detached as needed, similar to how a module (library) in software development functions Figure (3).

The GIF is designed to bridge the gap between narrow AI systems and AGI by enabling RTL, adaptability, and generalization across diverse domains. This section provides a detailed breakdown of the GIF's core components, demonstrating how each contributes to achieving a flexible, adaptive AI system that mimics human cognition.

#### *Input Module*

The input module gathers data from the environment, encompassing all types of inputs from various sensors (visual, auditory, tactile, etc.). This could include binary data, sound waves, electromagnetic signals, or any novel sensory inputs that may emerge in the future. The system gathers inputs based on its sensory devices. For example, if the sensory device only includes a camera module, the input will initially be limited to visual data. The GIF is built around a modular architecture where input modules can be attached or detached dynamically, depending on the sensory devices available. For example, in an autonomous vehicle, the framework could integrate new sensors (e.g., a LIDAR module) on the fly, without re-engineering the core system. The encoded sensory inputs are processed by the core decision-making module, which adapts to the new data stream automatically, allowing the system to respond effectively to novel stimuli

#### *Encoder Module*

The encoders act as the bridge between raw sensory inputs and the core processing unit of the framework. Inspired by SNNs, the encoders convert sensory inputs into spikes or neural representations that mimic how the brain processes electrical signals. This encoding process identifies patterns within the input data and adapts to the signals, preparing them for deep analysis. When a new sensor is attached (e.g., a proximity sensor), the system instantiates a specialized encoder module that maps raw signals into the spike-based representation. A brief auto-calibration phase aligns incoming spike patterns with existing SNN synaptic weights. During this phase, the system leverages Real-Time Learning to minimize discrepancies between known sensor patterns and those of the newly introduced modality. If significant deviations persist—such as unusual frequency ranges or signal amplitudes—the GIF refines the relevant

connections on the fly, ensuring seamless integration without a full retraining cycle

#### *Core Processing Module (Deep Understanding)*

- At the heart of the GIF is the DU Module, which is built on SNNs. SNNs mimic the time-dependent behavior of biological neurons, making them uniquely suited for real-time processing and decision-making.
- How SNNs Work: Unlike traditional artificial neural networks, which use continuous activation functions, SNNs process information in discrete spikes of activity. These spikes are triggered when the neuron's membrane potential reaches a certain threshold. This spike-based mechanism allows SNNs to handle temporal dependencies more effectively than conventional neural networks, making them ideal for tasks that require real-time responses, such as navigating dynamic environments or controlling robotic systems.
- Several studies solidify the foundation of the DU Module. Two key studies highlight that SNNs are highly effective for real-time sensory processing due to their ability to handle time-series data and respond to events as they happen. This capability is particularly beneficial in applications like robotics and autonomous vehicles, where real-time decision-making is crucial. Research in this area demonstrates that when SNNs are implemented on specialized hardware, they can efficiently process complex sensory inputs with low power consumption, making them ideal for real-time applications in edge computing and neuromorphic sensors (Madrenas *et al.*, 2023; Park and Choi, 2024)
- Another study shows that SNNs excel in adaptive systems requiring RTL and adjustment, such as Brain-Computer Interfaces (BCIs) and adaptive control systems. These networks leverage synaptic plasticity mechanisms, similar to those in biological systems, allowing them to dynamically adjust their responses based on ongoing inputs and experiences. This adaptability makes SNNs well-suited for applications in cognitive neuroscience, rehabilitation, and adaptive signal processing (Madrenas *et al.*, 2023; Drigas and Sideraki, 2024). SNNs emulate the behavior of biological neurons, where neurons communicate using spikes, i.e., discrete events occurring at specific points in time. This time-based functionality makes SNNs ideal for handling real-time, continuous data. SNNs are capable of learning and making decisions in real-time, thanks to their ability to process temporal dependencies and adapt to new stimuli dynamically. Neuromorphic chips like Intel's Loihi use SNNs to enable low-latency and energy-efficient processing, making them especially suitable for AGI systems that require real-time adaptability.



- These studies lay the groundwork for the DU Module. This module serves as the brain of the GIF. Built on SNNs and enhanced by neuromorphic computing, this module not only recognizes patterns but also understands and connects them with prior experiences. It analyzes the encoded inputs and relates them to existing knowledge, continuously linking new information with past experiences in a way that mirrors the cognitive processes of the human brain. The ability to continuously connect new and past information is crucial for effective decision-making in dynamic and real-time environments, as it allows the system to adapt and make informed decisions based on both learned experiences and new inputs

#### *Decoder Module*

Once the DU module processes the input, the decoders convert the output spikes back into actionable signals. These outputs could be physical actions (such as moving a robotic arm or adjusting a vehicle's steering) or symbolic outputs (such as making a decision or providing a recommendation in a software system). Similar to the Encoders, users initially specify the expected output, allowing the system to convert the spikes into the desired output signals. As the system accumulates experience, it will be capable of determining the appropriate kind of output required for different scenarios without user intervention.

#### *Action Module*

Actions represent the physical or computational outputs generated by the GIF system. The Inputs module constantly monitors the environment for changes and the effects of its actions. This feedback provides new inputs, which are processed again, creating a cycle of learning and adaptation in real-time.

#### *Real-Time Learning (RTL) Module*

- RTL is a core feature of the framework. This module enables the AGI system to learn from its own experiences, much like a human learns from interacting with the environment (ARC, 2024). For example, if the system is tasked with moving an object (like a book) to a specific position, the sensory feedback from the task informs the system about the required pressure and time to complete the task. Over time, this learning is generalized and applied to new tasks, such as moving objects of different weights or navigating complex environments. This ability to transfer learning from one domain to another is what sets AGI apart from narrow AI. This is achieved with the help of the RTL which were introduced in the above-mentioned papers (Madrenas *et al.*, 2023; Park and Choi, 2024; Drigas and Sideraki, 2024). For example, in a self-driving car scenario, the RTL module would allow the AGI to adapt to new driving environments and

unexpected obstacles, much like how a human driver would learn from experience. This contrasts with traditional models, which require specific retraining to handle new conditions.

- To further enhance real-time processing capabilities, the GIF will be designed to run on neuromorphic hardware such as Intel's Loihi or IBM's TrueNorth chips. These chips mimic the brain's energy-efficient architecture, enabling the framework to perform complex, spike-based computations with lower power consumption and reduced latency compared to traditional hardware. Neuromorphic hardware accelerates the processing of SNNs, ensuring that the system can respond to dynamic inputs with minimal delay.

Numerous studies in modular neural network design show that adding or swapping sensor modalities can be achieved with minimal retraining when employing spike-based encodings (Madrenas *et al.*, 2023). In GIF, each new input device—ranging from simple cameras to specialized sensors like LiDAR—connects via a dedicated encoder module that converts signals into spikes. Because the core SNNs only process standardized spike events, the system inherently supports “plug-and-play” integration. Over time, the GIF refines synaptic connections via Real-Time Learning, requiring no fundamental re-engineering of the core structure. This design mirrors biological systems, where novel sensory inputs eventually map onto existing neural pathways through synaptic plasticity.

Recent models like Gemini (Gemini Team Google, 2023) and GPT-4o demonstrate the initial stages of integrating multiple sensory inputs to achieve similar goals. However, the proposed GIF takes this a step further by enabling systems to understand and adapt to new inputs and tasks without re-architecting the core structure by implementing RTL. For example, in a self-driving car, the system would gather information from various sensors (e.g., cameras, LIDAR, accelerometers, and GPS). This information is then decoded into signals that the core processing module can interpret. The Core Decision-Making Module identifies patterns in these signals and cross-references them with stored knowledge in the memory functions. Based on this analysis, the system makes informed decisions, such as reducing speed or changing direction, which are then executed by the car's control systems.

The most important aspect of this framework is that the inputs and the actions could be attached or detached as per the requirements and the model will be capable of adapting to those changes similarly to humans. The same framework can be applied across different systems, from simple devices like model cars to complex robots, highlighting the generality and scalability of the GIF.

#### *Hypothetical Validation and Expected Outcomes*

While the full implementation of the GIF is beyond the immediate scope of this study, several thought

experiments and hypothetical validations are proposed to illustrate the framework's capabilities. These experiments are designed to highlight the GIF's adaptability, RTL, energy efficiency, and generalization across tasks. Below, we outline the key experiments and the expected outcomes.

#### *Real-Time Learning with Spiking Neural Networks (SNNs)*

In this experiment, the RTL capability of the GIF will be tested using a robotic arm tasked with interacting with objects of varying sizes, weights, and textures. The goal is to validate the framework's ability to dynamically adjust its actions based on sensory feedback, without the need for retraining:

- **Experiment Setup:** A robotic arm, equipped with tactile and visual sensors, will be placed in a simulated environment where it must pick up, move, and manipulate objects of different shapes and weights. The objects will vary in texture (e.g., smooth, rough) and fragility (e.g., fragile glassware, solid metal blocks)
- **Procedure:** The arm will receive real-time sensory feedback through its sensors, allowing it to adjust its grip strength and movement speed dynamically. As the robotic arm interacts with each object, the SNNs within the GIF will process time-sensitive data from the sensors and adapt the arm's actions accordingly.
- **Expected Outcome:** The robotic arm should demonstrate the ability to autonomously adjust its behavior as it encounters new objects. For example, it should reduce its grip strength when handling fragile items and increase it when lifting heavier objects. The experiment will validate that the SNNs can facilitate real-time adaptability without requiring manual retraining. Furthermore, this experiment will demonstrate the system's ability to apply experiential learning, refining its actions based on previous encounters with similar objects.

#### *Energy Efficiency Comparison: SNNs vs. Traditional Neural Networks*

The second experiment is designed to compare the energy efficiency of SNNs within the GIF to traditional DL models, such as Convolutional Neural Networks (CNNs), in a task that requires real-time adaptability:

- **Experiment Setup:** A robotic agent, equipped with both a CNN and an SNN-based model, will be tasked with navigating an obstacle course that features dynamic changes, such as the introduction of new barriers or shifts in the environment. The agent will complete the course in two scenarios: one where it is controlled by a CNN model and the other where it is controlled by SNNs implemented on neuromorphic hardware.

- **Procedure:** In the first scenario, the CNN-controlled agent will navigate the course, but will require retraining to handle changes in the obstacle layout. In the second scenario, the SNN-controlled agent will adapt to the changing environment in real-time, using sensory feedback to modify its path without the need for retraining.
- **Expected Outcome:** The SNN-controlled agent is expected to demonstrate significantly lower energy consumption compared to the CNN-controlled agent. This is due to the energy-efficient design of neuromorphic hardware, which processes spikes more efficiently than the continuous activations required by CNNs. Additionally, the SNN-controlled agent should adapt to the new obstacle layout faster than the CNN model, validating the efficiency of SNNs in real-time adaptability and energy consumption.

#### *Neuromorphic Hardware Integration for Real-Time Processing*

To showcase the real-time processing capabilities of the GIF, this experiment will simulate the integration of the framework with neuromorphic hardware such as Intel's Loihi or IBM's TrueNorth chips. The goal is to test the framework's ability to handle continuous data streams in real-time, such as video or LIDAR input from a self-driving vehicle:

- **Experiment Setup:** A self-driving vehicle simulator equipped with sensors (e.g., cameras, LIDAR) will process continuous streams of sensory data. The vehicle will navigate a complex environment that includes moving obstacles, varying weather conditions, and fluctuating road textures.
- **Procedure:** The GIF, running on neuromorphic hardware, will process the sensory data in real-time and make decisions regarding steering, speed, and obstacle avoidance. The experiment will measure the latency in decision-making, as well as the power consumption of the hardware.
- **Expected Outcome:** The neuromorphic hardware should enable the GIF to process continuous sensory data with minimal latency, allowing the vehicle to respond to changes in the environment almost instantaneously. Additionally, the system is expected to consume significantly less power compared to traditional GPU-based systems, validating the energy-efficient processing capabilities of the GIF when integrated with neuromorphic hardware.

#### *Generalization and Task Transferability*

The ability of the GIF to generalize knowledge across domains is a key aspect of achieving AGI. In this experiment, the framework will be tested on its capacity to transfer learned knowledge from one task to a different, yet related, task without requiring retraining.

- **Experiment Setup:** Two robotic tasks will be designed. In the first task, a robot will learn to stack various objects, ranging from lightweight plastic blocks to heavier metal cylinders. The second task will involve a new set of objects with different dimensions and weights, requiring the robot to adjust its stacking technique accordingly.
- **Procedure:** The robot will first be trained on the initial task, where it will learn how to handle and stack objects based on their size and weight. After mastering this task, it will immediately transition to the second task, where it will encounter different objects with unique properties.
- **Expected Outcome:** The robot should be able to generalize its knowledge from the first task and apply it to the second task without retraining. For example, if the robot learned that heavier objects require more precise placement, it should carry over this knowledge when dealing with the new set of objects. This experiment will validate the GIF's ability to transfer learning across tasks, a key capability for achieving AGI.

#### *Real-World Application Scenarios*

To further illustrate the potential of the GIF, hypothetical real-world scenarios will be proposed, focusing on industries such as healthcare and autonomous systems. Two examples are presented:

- **Adaptive Medical Robotics:** A robotic arm used in surgery could adapt to various surgical tools and tissue types in real-time. As the arm receives sensory feedback, it adjusts its actions to match the resistance or fragility of the tissues it is handling. The GIF would enable the system to switch seamlessly between tasks, such as cutting, stitching, and applying pressure, without requiring manual intervention.
- **Expected Outcome:** The robotic arm would perform tasks autonomously, adjusting in real time to changes in the surgical environment. This showcases the GIF's ability to handle critical, high-stakes tasks where adaptability and precision are paramount.
- **Autonomous Vehicle Navigation:** A self-driving car equipped with the GIF could transition from paved roads to off-road conditions without requiring retraining. The system would adjust steering, speed, and braking based on real-time feedback from the car's sensors, adapting to changing terrains on the fly.
- **Expected Outcome:** The vehicle would autonomously navigate through new terrains, adjusting its driving behavior in response to sensory feedback. This demonstrates the GIF's ability to adapt to novel environments without pre-programmed solutions.

#### *Key Innovations of the General Intelligence Framework*

The GIF represents a novel approach to achieving true general intelligence by integrating real-time adaptability, modularity, and energy-efficient learning mechanisms. Several aspects of the GIF set it apart from existing AI models, including traditional DL systems, transformers, and LLMs.

##### *Real-Time Adaptability*

One of the most significant innovations of the GIF is its real-time adaptability. Traditional AI models often require task-specific reconfiguration or extensive retraining when encountering new environments or inputs. For instance, LLMs and CNNs are limited by their dependence on vast amounts of domain-specific data and their inability to generalize across tasks without significant modification.

The GIF, by incorporating SNNs and real-time sensory feedback, enables systems to autonomously adapt to novel situations without requiring retraining. This mirrors the human brain's ability to adjust to new stimuli by leveraging previous knowledge and integrating real-time feedback. This adaptability is critical in dynamic, real-world environments such as autonomous driving, robotics, and healthcare.

##### *Modularity and Input Flexibility*

The modular architecture of the GIF allows it to integrate and process diverse forms of input (e.g., text, image, audio, or even sensor data) without the need for architecture-specific configurations. This is a significant improvement over existing multimodal models such as Google's Gemini or GPT-4, which require task-specific architectures optimized for particular input types.

In contrast, the GIF can dynamically process any new input by employing a modular input-output system, allowing seamless integration of new sensory modalities. This is particularly valuable in fields where flexibility is required, such as robotic systems adapting to new hardware configurations or autonomous vehicles processing novel environmental data.

##### *Neuromorphic Hardware Compatibility*

Another core innovation of the GIF is its compatibility with neuromorphic hardware platforms, such as Intel's Loihi or IBM's TrueNorth chips. While traditional DL systems are resource-intensive and require significant computational power, neuromorphic computing offers a more energy-efficient alternative that mimics the behavior of biological neurons.

The integration of SNNs into neuromorphic hardware allows the GIF to perform real-time, energy-efficient processing. This makes it ideal for edge computing

applications where low power consumption is essential, such as wearable devices, IoT systems, or autonomous drones. Unlike conventional neural networks, which become computationally expensive as they scale, the GIF’s hardware compatibility ensures it remains scalable without excessive energy consumption.

### Overcoming the Scalability Challenges of Deep Learning

Traditional DL architectures, particularly transformers and LLMs, face scalability challenges as their computational requirements grow exponentially

with model size. The GIF addresses the scalability challenges faced by traditional DL models such as transformers, which suffer from performance saturation as their computational requirements grow exponentially with model size. Recent studies have shown that the computational cost of models like GPT-4 increases quadratically with model size, yet the improvements in generalization are marginal beyond a certain point. In contrast, SNNs, supported by neuromorphic hardware, scale efficiently without requiring such resource-heavy configurations, making them more suitable for real-time, low-power environments Table (1).

**Table 1:** General Intelligence Framework (GIF) vs Traditional AI; a Footnote: The table summarizes how the General Intelligence Framework (GIF) surpasses traditional AI models

Feature	GIF	Deep Learning Models	LLMs (GPT-4)	Multimodal Models (Gemini)
Real-Time Learning	Yes (via SNNs)	No (requires batch learning)	No (fine-tuning required)	Limited (task-specific)
Modularity of Input/Output	Fully modular; integrates new inputs seamlessly	Limited to predefined architectures	Limited to predefined tasks	Predefined architectures
Adaptability	High (adapts without retraining)	Low (task-specific training required)	Low (domain-specific tuning)	Medium (multimodal but static)
Energy Efficiency	High (optimized for neuromorphic hardware)	Low (high computational cost)	Low (computationally expensive)	Low (computationally expensive)
Scalability	Efficient (energy-efficient, scalable)	Poor (performance saturation)	Poor (performance saturation)	Moderate

### Generalization Across Domains

Unlike narrow AI systems, which excel in specialized tasks, the GIF aims to bridge the gap between narrow AI and AGI by generalizing knowledge across domains. While existing LLMs and multimodal models, like GPT-4, are optimized for specific benchmarks, they lack the ability to autonomously adapt to new tasks without manual intervention or fine-tuning.

In the GIF, generalization across domains is achieved by leveraging RTL and experiential memory, allowing the system to relate new inputs to past experiences. This is akin to how humans apply prior knowledge to novel situations. For example, a system trained to navigate a simple environment could apply its learned navigation strategies to a more complex environment without additional retraining, mimicking human cognitive flexibility. By contrast, systems like MuZero, though highly effective within specific domains, struggle to generalize beyond their training environment.

### Limitations

We acknowledge that the GIF remains a theoretical proposal at this stage. To validate the framework, our immediate next step is to develop a small-scale exoplanet identification system prototype utilizing a neuromorphic chip or stimulated chip. Specifically, we plan to measure (1) Identifying exoplanets under varying noises, (2) Energy consumption compared to a conventional CNN-based system, and (3) Response times under domain shifts (e.g., new extra-terrestrial bodies). These metrics

will provide early quantitative evidence of the GIF’s feasibility. Additionally, we intend to test domain-transfer scenarios—such as transitioning from identifying exoplanets to simple navigation tasks—without re-architecting the core SNN modules.

### Conclusion

In this study, we introduced the General Intelligence Framework (GIF) as a theoretical approach to achieving more flexible, adaptive AI. By uniting SNNs, Real-Time Learning, and neuromorphic hardware in a modular architecture, GIF aims to overcome the rigidities of current AI systems. While our examples and experiments remain hypothetical, future work will focus on implementing small-scale prototypes—such as a neuromorphic robotic arm—to gather empirical data on energy efficiency, real-time adaptability, and cross-domain transfer. We believe GIF has the potential to significantly advance the state of AI by offering a path toward domain-general learning, but its true impact will depend on thorough experimental validation in real-world scenarios.

Key Contributions of this study include:

- By leveraging SNNs and RTL, the GIF enables systems to autonomously adjust to novel inputs and tasks, mimicking the adaptability seen in human cognition.
- The GIF’s modular input-output system allows for the dynamic integration of new sensory inputs and outputs, providing unparalleled flexibility in diverse applications. This modularity ensures that the

framework can operate across various domains without requiring extensive reconfiguration.

- The framework's integration with neuromorphic hardware enables highly efficient processing, reducing power consumption while maintaining high performance in real-time applications. This makes it an ideal solution for autonomous systems and edge computing environments.
- One of the most critical aspects of the GIF is its ability to generalize knowledge across domains. By utilizing real-time sensory feedback and experiential learning, the framework can transfer learned skills from one task to another without requiring retraining, positioning it as a crucial step toward achieving AGI.

Future Work will focus on validating the framework through real-world experiments, including the implementation of the GIF in neuromorphic hardware environments and testing its adaptability in dynamic, real-time systems such as autonomous vehicles and medical robotics. Further research will also investigate expanding the framework's ability to handle increasingly complex tasks, incorporating advanced forms of RTL and exploring its scalability in industrial-scale applications.

While the immediate implementation of the framework is limited to theoretical validation, the potential applications of the GIF are vast. By addressing the shortcomings of current AI architectures and introducing a biologically inspired approach to adaptability, the GIF offers a path forward for the next generation of AI systems capable of autonomous decision-making, dynamic learning, and efficient real-time performance in various real-world environments.

The GIF represents a significant step forward in the journey toward AGI. By addressing the core challenges of real-time adaptability, energy efficiency, and task generalization, the GIF has the potential to reshape how AI systems are designed and deployed across industries. As future research continues to refine the framework and validate it in practical settings, the GIF could become a cornerstone in the development of more flexible, adaptable, and intelligent AI systems.

## Acknowledgement

The authors acknowledge the idea and suggestions given by Dr. Tulasi, Dr. Vaidhehi V, Dr. Blesson, Research Advisory Committee Members.

## Funding Information

The authors have not received any financial support or funding to report.

## Author's Contributions

**Jiran Kurian Puliyanmakkal:** Developed the Framework.

**Rohini V:** Designed the research plan and organized the study.

## Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

## References

- ARC. (2024). ARC-AGI Design. *ARC Prize*.  
<https://arcprize.org/arc>
- Birhane, A., Kasirzadeh, A., Leslie, D., & Wachter, S. (2023). Science in the Age of Large Language Models. *Nature Reviews Physics*, 5(5), 277-280.  
<https://doi.org/10.1038/s42254-023-00581-4>
- Brachman, R. J., & Levesque, H. J. (2022). Common Sense in Humans. *Machines like Us: Toward AI with Common Sense*, 9-24.  
<https://doi.org/10.7551/mitpress/14299.003.0004>
- Drigas, A., & Sideraki, A. (2024). Brain Neuroplasticity Leveraging Virtual Reality and Brain-Computer Interface Technologies. *Sensors*, 24(17), 5725.  
<https://doi.org/10.3390/s24175725>
- Gemini Team Google. (2023). Gemini: A Family of Highly Capable Multimodal Models. *ArXiv Preprint ArXiv:2312.11805*.  
<https://doi.org/https://doi.org/10.48550/arXiv.2312.11805>
- Ifrah, S. (2024). GPT-4o, DALL-E, and Whisper. *Getting Started with Azure OpenAI*, 217-250.  
[https://doi.org/10.1007/979-8-8688-0599-8\\_6](https://doi.org/10.1007/979-8-8688-0599-8_6)
- Javanshir, A., Nguyen, T. T., Mahmud, M. A. P., & Kouzani, A. Z. (2022). Advancements in Algorithms and Neuromorphic Hardware for Spiking Neural Networks. *Neural Computation*, 34(6), 1289-1328.  
[https://doi.org/10.1162/neco\\_a\\_01499](https://doi.org/10.1162/neco_a_01499)
- Job, M. B., Reeves, M., & Adam. (2023). The Working Limitations of Large Language Models. *MIT Sloan Manag. Rev. URL*.  
<https://sloanreview.mit.edu/article/the-working-limitations-of-large-language-models/>
- Kurian, J., & Rohini, V. (2021). A Deliberation on the Stages of Artificial Intelligence. *SCRS Conference Proceedings on Intelligent Systems, SCRS*, 1-11.  
<https://doi.org/10.52458/978-93-91842-08-6-1>
- Madrenas, J., Vallejo-Mancero, B., Oltra-Oltra, J. À., Zapata, M., Cosp-Vilella, J., Calatayud, R., Moriya, S., & Sato, S. (2023). Real-Time Adaptive Physical Sensor Processing with SNN Hardware. *Artificial Neural Networks and Machine Learning -- ICANN 2023*, 423-434.  
[https://doi.org/10.1007/978-3-031-44192-9\\_34](https://doi.org/10.1007/978-3-031-44192-9_34)
- Naveed, H., Khan, A. U., Qiu, S., Saqib, M., Anwar, S., Usman, M., Akhtar, N., Barnes, N., & Mian, A. (2025). A Comprehensive Overview of Large Language Models. *ACM Transactions on Intelligent Systems and Technology*.  
<https://doi.org/10.1145/3744746>

- Ott, S., Barbosa-Silva, A., Blagec, K., Brauner, J., & Samwald, M. (2022). Mapping Global Dynamics of Benchmark Creation and Saturation in Artificial Intelligence. *Nature Communications*, 13(1), 6793. <https://doi.org/10.1038/s41467-022-34591-0>
- Park, S. S., & Choi, Y.-S. (2024). Spiking Neural Networks for Physiological and Speech Signals: A Review. *Biomedical Engineering Letters*, 14(5), 943-954. <https://doi.org/10.1007/s13534-024-00404-0>
- Qi, Y., Chen, J., & Wang, Y. (2023). Neuromorphic Computing Facilitates Deep Brain-Machine Fusion for High-Performance Neuroprosthesis. *Frontiers in Neuroscience*, 17, 1153985. <https://doi.org/10.3389/fnins.2023.1153985>
- Rachel, G., & MIT, C. (2024). Reasoning Skills of Large Language Models are Often Overestimated. *MIT News Mass. Inst. Technol. URL*. <https://news.mit.edu/2024/reasoning-skills-large-language-models-often-overestimated-0711>
- Schrittwieser, J., Antonoglou, I., Hubert, T., Simonyan, K., Sifre, L., Schmitt, S., Guez, A., Lockhart, E., Hassabis, D., Graepel, T., Lillicrap, T., & Silver, D. (2020). Mastering Atari, Go, Chess and Shogi by Planning With a Learned Model. *Nature*, 588(7839), 604-609. <https://doi.org/10.1038/s41586-020-03051-4>
- Shah, D., Osiński, B., Ichter, B., & Levine, S. (2022). *LM-Nav: Robotic Navigation with Large Pre-Trained Models of Language, Vision, and Action*. <https://sites.google.com/view/lmnav>
- Shang, J., Zheng, Z., Wei, J., Ying, X., Tao, F., & Team, M. (2024). AI-Native Memory: A Pathway from LLMs Towards AGI. *ArXiv:2406.18312*. <https://doi.org/https://doi.org/10.48550/arXiv.2406.18312>
- Ullah, E., Parwani, A., Baig, M. M., & Singh, R. (2024). Challenges and Barriers of Using Large Language Models (LLM) Such as ChatGPT for Diagnostic Medicine with a Focus on Digital Pathology - A Recent Scoping Review. *Diagnostic Pathology*, 19(1), 43. <https://doi.org/10.1186/s13000-024-01464-7>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., & Kaiser, Ł. (2017). Attention Is All You Need. *Advances in Neural Information Processing Systems* 30, 30, 5998-6008. <https://doi.org/10.48550/arXiv.1706.03762>
- Zhang, D., Yu, Y., Dong, J., Li, C., Su, D., Chu, C., & Yu, D. (2024). MM-LLMs: Recent Advances in MultiModal Large Language Models. *Findings of the Association for Computational Linguistics ACL 2024*, 12401-12430. <https://doi.org/10.18653/v1/2024.findings-acl.738>