Novelty Assessment Report

Paper: Vision Hopfield Memory Networks

PDF URL: https://openreview.net/pdf?id=BigoLH4PsK

Venue: ICLR 2026 Conference Submission

Year: 2026

Report Generated: 2025-12-30

Abstract

Recent vision and multimodal foundation backbones, such as Transformer families and state-space models like Mamba, have achieved remarkable progress, enabling unified modeling across images, text, and beyond. Despite their empirical success, these architectures remain far from the computational principles of the human brain, often demanding enormous amounts of training data while offering limited interpretability. In this work, we propose the Vision Hopfield Memory Network (V-HMN), a brain-inspired foundation backbone that integrates hierarchical memory mechanisms with iterative refinement updates. Specifically, V-HMN incorporates local Hopfield modules that provide associative memory dynamics at the image patch level, global Hopfield modules that function as episodic memory for contextual modulation, and a predictive-coding–inspired refinement rule for iterative error correction. By organizing these memory-based modules hierarchically, V-HMN captures both local and global dynamics in a unified framework. Memory retrieval exposes the relationship between inputs and stored patterns, making decisions more interpretable, while the reuse of stored patterns improves data efficiency. This brain-inspired design therefore enhances interpretability and data efficiency beyond existing self-attention- or state-space–based approaches. We conducted extensive experiments on public computer vision benchmarks, and V-HMN achieved competitive results against widely adopted backbone architectures, while offering better interpretability, higher data efficency, and stronger biological plausibility. These findings highlight the potential of V-HMN to serve as a next-generation vision foundation model, while also providing a generalizable blueprint for multimodal backbones in domains such as text and audio, thereby bridging brain-inspired computation with large-scale machine learning.

Core Task Landscape

This paper addresses: **Brain-Inspired Vision Backbone with Hierarchical Memory Mechanisms**

A total of **50 papers** were analyzed and organized into a taxonomy with **26 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Hierarchical Memory Architectures for Vision Systems**
- **Neuromorphic Hardware and In-Memory Computing for Vision**
- **Spiking Neural Networks for Visual Processing**
- **Attention and Recurrent Mechanisms in Vision Models**
- **Biologically Inspired Hierarchical Feature Extraction**
- **Memory-Augmented Learning for Recognition Tasks**
- **Hierarchical Temporal Memory Algorithms and Applications**
- **Computational Models of Visual Cortex Organization**
- **Neuroscience-Informed Architectures and Cognitive Models**

Complete Taxonomy Tree

- Brain-Inspired Vision Backbone with Hierarchical Memory Mechanisms Survey Taxonomy
- Hierarchical Memory Architectures for Vision Systems
  - Multi-Memory Integration Frameworks for Embodied Agents ★ (4 papers)
  - [0] Vision Hopfield Memory Networks (Anon et al., 2026) View paper
  - [1] RoboMemory: A Brain-inspired Multi-memory Agentic Framework for Lifelong Learning in Physical Embodied Systems (Cai, 2025) View paper
  - [2] RoboMemory: A Brain-inspired Multi-memory Agentic Framework for Interactive Environmental Learning in Physical Embodied Systems (Cai, 2025) View paper
  - [16] Neural Brain: A Neuroscience-inspired Framework for Embodied Agents (Liu Jian, 2025) View paper
  - Hippocampal-Inspired Memory Systems for Sequential Decision Making (1 papers)
  - [7] A scalable reinforcement learning framework inspired by hippocampal memory mechanisms for efficient contextual and sequential decision making (Hamed Poursiami, 2025) View paper
  - Working Memory Models for Visual Processing (3 papers)
  - [23] Working memory inspired hierarchical video decomposition with transformative representations (Qin Bin-jie, 2022) View paper
  - [43] Artificial Working Memory Constructed by Planar 2D Channel Memristors Enabling Brainâ□□Inspired Hierarchical Memory Systems (Xinglong Ji, 2021) View paper
  - [44] A Biologically Inspired Visual Working Memory for Deep Networks (Ethan Harris, 2022) View paper
  - Hierarchical Video Memory and Captioning Systems (1 papers)
  - [13] HiCMÂ²: Hierarchical Compact Memory Modeling for Dense Video Captioning (Minkuk Kim, 2025) View paper
- Neuromorphic Hardware and In-Memory Computing for Vision
  - Optoelectronic and Memristive Sensor Arrays (2 papers)

- [3] Full hardware implementation of neuromorphic visual system based on multimodal optoelectronic resistive memory arrays for versatile image processing (Guangdong Zhou, 2023) View paper
- [4] In-sensor image memorization and encoding via optical neurons for bio-stimulus domain reduction toward visual cognitive processing (Doeon Lee, 2022) View paper
- In-Memory Computing Architectures for Multimodal Processing (1 papers)
- [5] A brain-inspired hierarchical interactive in-memory computing system and its application in video sentiment analysis (Xiaoyue Ji, 2023) View paper
- Neuromorphic Circuits for Temporal Memory Algorithms (4 papers)
- [18] Hierarchical temporal memory based on spin-neurons and resistive memory for energy-efficient brain-inspired computing (Deliang Fan, 2015) View paper
- [19] Hierarchical temporal memory features with memristor logic circuits for pattern recognition (Olga Krestinskaya, 2017) View paper
- [24] Neuromorphic architecture for the hierarchical temporal memory (Abdullah M. Zyarah, 2019) View paper
- [25] Enhancing Biologically Inspired Hierarchical Temporal Memory with Hardware-Accelerated Reflex Memory (Pavia Bera, 2025) View paper

- Spiking Neural Networks for Visual Processing
  - Hierarchical Spiking Models for Image Classification (1 papers)
  - [6] Hierarchical Spiking-Based Model for Efficient Image Classification With Enhanced Feature Extraction and Encoding (Qi Xu, 2022) View paper
  - Unsupervised Spiking Networks for Optical Flow (1 papers)
  - [17] Unsupervised Learning of a Hierarchical Spiking Neural Network for Optical Flow Estimation: From Events to Global Motion Perception (Federico Paredes-VallÃ©s, 2018) View paper
- Attention and Recurrent Mechanisms in Vision Models
  - Transformer-Based Hierarchical Vision Architectures (2 papers)
  - [22] Brain-Inspired Stepwise Patch Merging for Vision Transformers (Yu YongHao, 2025) View paper
  - [34] A recurrent vision transformer shows signatures of primate visual attention (Jonathan Morgan, 2025) View paper
  - Recurrent and Feedback Networks for Visual Dynamics (3 papers)
  - [20] Deep recurrent neural network reveals a hierarchy of process memory during dynamic natural vision (Junxing Shi, 2018) View paper
  - [38] Brain-inspired Hierarchical Attention Recurrent CNN for Image Classification (Xinjing Song, 2022) View paper
  - [47] A Memory-Based Recurrent Neural Architecture for Chip Emulating Cortical Visual Processing (L Raffo, 1994) View paper
  - Attention-Based Saliency and Segmentation Models (4 papers)
  - [8] Lightweight salient object detection via hierarchical visual perception learning (Yun Liu, 2020) View paper
  - [10] Hierarchical Activation Dual Backbone Network for Weakly Supervised Semantic Segmentation (Congwei Zhang, 2024) View paper
  - [15] Seeing the Overlooked: Bio-Visual Inspired Weak Saliency Feedback Transformer for Person Re-identification (Changshuo Wang, 2025) View paper
  - [30] STM-SalNet: A Biologically-Inspired Spatial-Temporal Memory Network for Video Saliency Prediction (J Xu, 2025) View paper
- Biologically Inspired Hierarchical Feature Extraction
  - HMAX and Ventral Pathway Models (3 papers)
  - [35] Introducing memory and association mechanism into a biologically inspired visual model (Hong Qiao, 2013) View paper
  - [37] A brain-inspired network architecture for cost-efficient object recognition in shallow hierarchical neural networks. (Youngjin Park, 2021) View paper
  - [39] A biologically motivated visual memory architecture for online learning of objects (Stephan Kirstein, 2008) View paper
  - Multi-Level Interaction and Contour Detection (2 papers)
  - [9] Bio-inspired multi-level interactive contour detection network (Chuan Lin, 2023) View paper
  - [11] Bio-inspired smart vision sensor: toward a reconfigurable hardware modeling of the hierarchical processing in the brain (Pankaj Bhowmik, 2021) View paper
  - Hierarchical Models for Specialized Vision Tasks (3 papers)
  - [21] Vision-based topological mapping and navigation with self-organizing neural networks (Yue Hu, 2021) View paper
  - [27] A Brain-inspired Method for Occluded 3D Object Recognition (Zining Wan, 2024) View paper
  - [36] THIRDEYE: Cue-Aware Monocular Depth Estimation via Brain-Inspired Multi-Stage Fusion (Ioan, 2025) View paper
- Memory-Augmented Learning for Recognition Tasks
  - Graph-Based and Associative Memory Networks (2 papers)
  - [14] M-GCN: Brain-inspired memory graph convolutional network for multi-label image recognition (Xiao Yao, 2022) View paper
  - [29] Hierarchical associative memory based on oscillatory neural network (Yan Fang, 2013) View paper
  - Adaptive Memory for Few-Shot Learning (1 papers)
  - [26] Brain Inspired Adaptive Memory Dual-Net for Few-Shot Image Classification (Di Kexin, 2025) View paper
- Hierarchical Temporal Memory Algorithms and Applications
  - HTM for Object Recognition and Classification (1 papers)
  - [32] Object recognition using hierarchical temporal memory (FabiÃ¡n Fallas-Moya, 2018) View paper
  - HTM Feature Encoding and Spatial Pooling (1 papers)
  - [33] Feature extraction without learning in an analog spatial pooler memristive-CMOS circuit design of hierarchical temporal memory (Krestinskaya, 2018) View paper
- Computational Models of Visual Cortex Organization
  - Temporal Stability and Local Memory Models (1 papers)
  - [28] A model of the ventral visual system based on temporal stability and local memory (KÃ¶nig, 2006) View paper
  - Self-Organizing Hierarchical Memory Networks (1 papers)
  - [42] On the self-organization of a hierarchical memory for compositional object representation in the visual cortex (Evgueni Jitsev, 2010) View paper
  - Joint-Embedding and Predictive Architectures (1 papers)
  - [12] Understanding cortical computation through the lens of joint-embedding predictive architectures (AG Mohammadi, 2025) View paper

- Neuroscience-Informed Architectures and Cognitive Models
  - Molecular and Evolutionary Memory Architectures (1 papers)
  - [40] Hypernetworks: A molecular evolutionary architecture for cognitive learning and memory (Byoungâ□□Tak Zhang, 2008) View paper
  - Neural Decoding and Representational Dynamics (2 papers)
  - [31] Decoding Mouse Visual Tasks via Hierarchical Neural-Information Gradients (J Feng, 2025) View paper
  - [41] Convolutional neural networks uncover the dynamics of human visual memory representations over time. (Stas Kozak, 2024) View paper
  - Feature Binding and Attention in Working Memory (2 papers)
  - [46] A biologically inspired spiking neural P system in selective visual attention for efficient feature extraction from human motion (Esteban Anides, 2022) View paper
  - [49] Neural Architecture for Feature Binding in Visual Working Memory. (Sebastian Schneegans, 2017) View paper
  - Hierarchical Evolution and Localization Models (3 papers)
  - [45] Biologically Inspired Hierarchical Model for Feature Extraction and Localization (Liang Wu, 2022) View paper
  - [48] The Hierarchical Evolution in Human Vision Modeling. (D. Ballard, 2021) View paper
  - [50] Advances in Neural Information Processing Systems 13: Proceedings of the 2000 Conference (TK Leen, 2001) View paper

Narrative

Core task: brain-inspired vision backbone with hierarchical memory mechanisms. The field encompasses diverse approaches to integrating memory and biological principles into visual processing systems. At the broadest level, the taxonomy reveals several major branches: hierarchical memory architectures that organize multi-level storage for vision tasks, neuromorphic hardware implementations that exploit in-memory computing substrates, spiking neural networks that mimic temporal dynamics of biological neurons, attention and recurrent mechanisms that enable iterative refinement, biologically inspired feature extraction methods that mirror cortical organization, memory-augmented learning frameworks for recognition, hierarchical temporal memory (HTM) algorithms, computational models of visual cortex structure, and neuroscience-informed cognitive architectures. Some branches emphasize hardware efficiency and novel computing paradigms (Neuromorphic Visual Resistive[3], In-sensor Image Memorization[4]), while others focus on algorithmic innovations such as spiking dynamics (Hierarchical Spiking Classification[6]) or cortex-inspired hierarchies (Hierarchical Activation Backbone[10]). The interplay between these branches reflects ongoing efforts to balance biological fidelity, computational tractability, and practical performance.

Particularly active lines of work explore multi-memory integration frameworks for embodied agents, where systems must coordinate short-term sensorimotor memory with long-term episodic storage—exemplified by RoboMemory Lifelong[1] and RoboMemory Interactive[2], which address continual learning and interactive scenarios in robotics. Vision Hopfield Memory[0] sits within this cluster, emphasizing associative memory mechanisms that enable robust retrieval and pattern completion in visual backbones. Compared to the RoboMemory works that target embodied agent workflows, Vision Hopfield Memory[0] focuses more directly on the backbone architecture itself, leveraging Hopfield-style dynamics to create hierarchical memory layers. This contrasts with approaches like Neural Brain Framework[16], which integrates broader cognitive modeling, and with hardware-centric efforts such as Hierarchical Interactive In-memory[5] that prioritize physical substrate design. The central trade-off across these directions involves the granularity of memory organization, the degree of biological inspiration, and the balance between general-purpose learning and task-specific optimization.

Related Works in Same Category

The following **3 sibling papers** share the same taxonomy leaf node with the original paper:

1. RoboMemory: A Brain-inspired Multi-memory Agentic Framework for Lifelong Learning in Physical Embodied Systems

Authors: Cai, Honghao, Mingcong Lei, Honghao Cai, Binbin Que, et al. (30 authors total) | **Year/Venue**: 2025 | **URL**: View paper

Abstract

Embodied agents face persistent challenges in real-world environments, including partial observability, limited spatial reasoning, and high-latency multi-memory integration. We present RoboMemory, a brain-inspired framework that unifies Spatial, Temporal, Episodic, and Semantic memory under a parallelized architecture for efficient long-horizon planning and interactive environmental learning. A dynamic spatial knowledge graph (KG) ensures scalable and consistent memory updates, while a closed-lo...

Relationship Analysis

Both papers belong to the Multi-Memory Integration Frameworks for Embodied Agents category, focusing on unifying multiple memory types for embodied planning. While the original paper (V-HMN) proposes a vision backbone with local and global Hopfield memory modules for hierarchical image processing and classification, the candidate paper (RoboMemory) presents a complete embodied agent framework integrating spatial, temporal, episodic, and semantic memory modules for real-world robotic task execution and lifelong learning. The key difference is that V-HMN focuses on memory-based vision representation learning as a foundation model, whereas RoboMemory addresses end-to-end embodied agent planning with dynamic knowledge graphs and closed-loop control for physical robot deployment.

---

2. RoboMemory: A Brain-inspired Multi-memory Agentic Framework for Interactive Environmental Learning in Physical Embodied Systems

Authors: Cai, Honghao, Wu YiMou, Jiang Shao-han, Wang Ge, et al. (12 authors total) | **Year/Venue**: 2025 | **URL**: View paper

Abstract

Embodied agents face persistent challenges in real-world environments, including partial observability, limited spatial reasoning, and high-latency multi-memory integration. We present RoboMemory, a brain-inspired framework that unifies Spatial, Temporal, Episodic, and Semantic memory under a parallelized architecture for efficient long-horizon planning and interactive environmental learning. A dynamic spatial knowledge graph (KG) ensures scalable and consistent memory updates, while a closed-lo...

Relationship Analysis

Both papers belong to the Multi-Memory Integration Frameworks for Embodied Agents category, focusing on unifying multiple memory types for embodied planning. While the original paper (Vision Hopfield Memory Networks) proposes a brain-inspired vision backbone integrating local and global Hopfield memory modules with predictive-coding refinement for image classification tasks, the candidate paper (RoboMemory) presents a multi-memory agentic framework (Spatial, Temporal, Episodic, Semantic) specifically designed for interactive environmental learning in physical robotic systems. The key distinction is that the original paper focuses on hierarchical memory mechanisms within a vision backbone for perception tasks, whereas the candidate paper emphasizes memory-augmented planning and navigation for embodied agents in dynamic real-world environments.

---

3. Neural Brain: A Neuroscience-inspired Framework for Embodied Agents

Authors: Liu Jian, Jian Liu, Xiongtao Shi, Zhang Haitian, Thai Duy Nguyen, et al. (37 authors total) | **Year/Venue**: 2025 | **URL**: View paper

Abstract

The rapid evolution of artificial intelligence (AI) has shifted from static, data-driven models to dynamic systems capable of perceiving and interacting with real-world environments. Despite advancements in pattern recognition and symbolic reasoning, current AI systems, such as large language models, remain disembodied, unable to physically engage with the world. This limitation has driven the rise of embodied AI, where autonomous agents, such as humanoid robots, must navigate and manipulate uns...

Relationship Analysis

Both papers belong to the Multi-Memory Integration Frameworks for Embodied Agents category, focusing on unifying multiple memory types for embodied intelligence. The original paper (Vision Hopfield Memory Networks) proposes a vision backbone integrating local and global Hopfield memory modules with predictive-coding-inspired refinement for image classification tasks, while the candidate paper (Neural Brain) presents a broader neuroscience-inspired framework encompassing multimodal sensing, perception-cognition-action loops, neuroplasticity-based memory, and neuromorphic hardware for general embodied agents. The key difference is that the original paper focuses specifically on hierarchical memory architectures for vision backbones, whereas the candidate paper addresses a comprehensive system-level architecture for embodied agents across multiple modalities and cognitive functions.

Contributions Analysis

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: Vision Hopfield Memory Network (V-HMN) architecture

Description: The authors introduce V-HMN, a novel vision backbone that replaces conventional self-attention or convolution with hierarchical Hopfield-style associative memory modules. The architecture combines local memory for patch-level pattern completion and global memory for scene-level context, organized in a unified framework with iterative refinement.

This contribution was assessed against **8 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. STanhop: Sparse tandem hopfield model for memory-enhanced time series prediction

URL: View paper

Brief Assessment

Sparse Tandem Hopfield[53] focuses on time series prediction with temporal and cross-series memory modules, not hierarchical vision backbone architectures with local/global patch-level memory for image understanding.

2. iMixer: hierarchical Hopfield network implies an invertible, implicit and iterative MLP-Mixer

URL: View paper

Brief Assessment

Hierarchical Hopfield Mixer[54] focuses on deriving MLP-Mixer variants from hierarchical Hopfield networks for image classification, not on replacing self-attention with Hopfield-style associative memory modules as a vision backbone with local and global memory paths.

3. Entropy driven artificial neuronal networks and sensorial representation: A proposal

URL: View paper

Brief Assessment

Entropy Neuronal Networks[57] discusses Hopfield networks in a theoretical context for sensorial representation and precategorical refinement, but does not present a hierarchical vision backbone architecture with local/global memory modules for image classification tasks.

4. Hierarchical Hopfield Network Decomposition: A Spiked Covariance Framework for Latent Prototype Discovery

URL: View paper

Brief Assessment

Hierarchical Hopfield Decomposition[58] focuses on classical Hopfield networks for clustering and prototype extraction in a spiked covariance framework, not on building vision backbones with hierarchical memory modules for image classification.

5. Semantic enhancement and multi-level alignment network for cross-modal retrieval

URL: View paper

Brief Assessment

Semantic Multi-level Alignment[51] focuses on cross-modal retrieval between vision and text using semantic alignment networks, not on Hopfield-style associative memory modules for hierarchical vision backbones with iterative refinement.

6. Out-of-Distribution Nuclei Segmentation in Histology Imaging via Liquid Neural Networks with Modern Hopfield Layer

URL: View paper

Brief Assessment

Liquid Neural Hopfield[56] focuses on nuclei segmentation in histology imaging using liquid neural networks combined with Hopfield layers for out-of-distribution robustness. The original V-HMN is a general vision backbone for image classification that replaces self-attention with hierarchical Hopfield memory modules, whereas this candidate applies Hopfield layers as a stabilization mechanism within a domain-specific segmentation pipeline.

7. A universal abstraction for hierarchical hopfield networks

URL: View paper

Brief Assessment

Universal Hierarchical Hopfield[52] presents a general abstraction framework for building hierarchical Hopfield networks using hypergraphs with neuron layers and synapses, focusing on the mathematical formalism and software implementation (HAMux). The original paper proposes V-HMN as a specific vision backbone architecture with local/global memory modules for image classification tasks, which is a distinct application domain and architectural instantiation.

8. Energy-Based Learning and the Evolution of Hopfield Networks: From Boltzmann Machines to Transformer Attention Mechanisms

URL: View paper

Brief Assessment

Energy-Based Hopfield Evolution[55] provides a historical survey of energy-based learning and Hopfield networks, focusing on theoretical evolution from Boltzmann machines to transformer attention mechanisms. It does not present a hierarchical vision backbone architecture with local and global memory modules for iterative refinement as proposed in V-HMN.

Contribution 2: Predictive-coding–inspired iterative refinement mechanism

Description: The authors develop a lightweight refinement update rule where representations are gradually corrected toward memory-predicted prototypes through learnable error-correction steps. This mechanism provides an interpretable, brain-inspired alternative to purely feedforward processing.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. An approximation of the error backpropagation algorithm in a predictive coding network with local hebbian synaptic plasticity

URL: View paper

Brief Assessment

Predictive Coding Backpropagation[68] focuses on approximating error backpropagation through local Hebbian plasticity in supervised learning, not on iterative refinement for error correction in memory-based vision networks.

---

2. Neurocomputational Mechanisms of Sense of Agency: Literature Review for Integrating Predictive Coding and Adaptive Control in Humanâ Machine Interfaces

URL: View paper

Brief Assessment

Predictive Coding Agency[64] focuses on neuroscience models of sense of agency in human-machine interfaces, not on neural network architectures or iterative refinement mechanisms for computer vision tasks.

---

3. Associative memories via predictive coding

URL: View paper

Prior Art Analysis

Associative Predictive Memories[65] demonstrates that predictive coding can be used to implement iterative refinement for error correction in neural networks prior to the original paper. The candidate paper presents a generative predictive coding network where representations are iteratively updated through an inference phase that minimizes prediction errors via gradient descent on a global energy function. This iterative refinement mechanism, formalized through equations that update value nodes to reduce prediction errors between current representations and memory-predicted patterns, directly anticipates the original paper's contribution. Both papers use learnable parameters to control the strength of error-correction updates and frame the process as gradually correcting representations toward memory-based predictions.

Evidence

Evidence 1 - **Rationale**: Both papers describe using predictive coding to minimize errors through iterative updates. The candidate establishes the foundational mechanism of error minimization via gradient descent that the original paper builds upon. - **Original**: we propose the vision hopfield memory network (v-hmn), a brain-inspired foundation backbone that integrates hierarchical memory mechanisms with iterative refinement updates. specifically, v-hmn incorporates local hopfield modules that provide associative memory dynamics at the image patch level, glo... - **Candidate**: we now briefly recall predictive coding networks (pcns), and we introduce generative pcns, which are the underlying neural model for the novel ams introduced in the subsequent section. deep neural networks have a multi-layer structure, where each layer is formed by a vector of neurons [17]. while in ...

Evidence 2 - **Rationale**: The candidate paper explicitly describes an iterative refinement process where value nodes are gradually corrected to minimize prediction errors, which is the same core mechanism claimed as novel in the original paper. - **Original**: this mechanism can be viewed as a lightweight form of predictive-coding dynamics, where representations are gradually corrected toward memory-predicted patterns. in this way, the network gains an error-corrective feedback process that is absent in conventional feedforward models. - **Candidate**: inference: only the value nodes of the network are updated, while both the weight parameters and the memory vector are fixed. particularly, the value nodes are modified via gradient descent to minimize the global error of the network, expressed by the following energy function et: et = 1 2 ∑ i,l(ε l i...

Evidence 3 - **Rationale**: The candidate provides the mathematical formulation for iterative refinement with a learnable parameter (γ) controlling update strength, directly corresponding to the original paper's learnable β parameter for refinement. - **Original**: by organizing these memory-based modules hierarchically, v-hmn captures both local and global dynamics in a unified framework. memory retrieval exposes the relationship between inputs and stored patterns, making decisions more interpretable, while the reuse of stored patterns improves data efficienc... - **Candidate**: the process of minimizing et by modifying all xl i,t leads to the following changes in the value nodes: Δx l i,t = { γ· (-εl i,t + f′(xl i,t)∑nl-1 k=1 εl-1 k,t θl k,i,t) if 0 <l ≤ l 0 if l= 0 , (3) where γis the integration step, which is a constant determining by how much the activity changes in ea...

Evidence 4 - **Rationale**: The candidate demonstrates that their predictive coding network iteratively refines representations toward stored memory patterns through convergence to local minima, which is the same principle described in the original paper's refinement mechanism. - **Original**: this mechanism can be viewed as a lightweight form of predictive-coding dynamics, where representations are gradually corrected toward memory-predicted patterns. - **Candidate**: let ¯sbe a training data point, and m be the pcn considered above, already trained until convergence to generate ¯s. moreover, assume that m makes the total energy converge to zero at iteration t. at this point, the energy function defined on the value nodes has a local minimum ¯xin which the value n...

Evidence 5 - **Rationale**: Both papers position predictive coding as a biologically plausible mechanism for error correction. The candidate establishes this connection prior to the original paper's claim of introducing a brain-inspired refinement mechanism. - **Original**: in terms of biological plausibility, deep learning architectures and optimization methods are largely engineered for computational convenience rather than grounded in neuroscience. for example, backpropagation with gradient descent has no clear biological counterpart, and the conventional feedforwar... - **Candidate**: predictive coding (pc). this framework includes both a biologically plausible neural architecture and a learning algorithm [4] developed to emulate learning in the visual cortex. particularly, pc allows to describe multiple phenomena happening in the brain, using a single framework, such as free-en...

---

4. Towards the Training of Deeper Predictive Coding Neural Networks

URL: View paper

Brief Assessment

Deeper Predictive Coding[70] focuses on training deeper predictive coding networks through precision-weighted optimization and novel weight updates, not on developing iterative refinement as a general mechanism for vision models. The candidate addresses depth-scaling issues in predictive coding architectures rather than proposing iterative refinement as a brain-inspired alternative to feedforward processing.

5. Anopcn: Video anomaly detection via deep predictive coding network

URL: View paper

Brief Assessment

Predictive Coding Anomaly[61] focuses on video anomaly detection with frame prediction and error reconstruction, not general-purpose iterative refinement for neural network representations as in the original paper.

6. ActPC-Geom: Towards Scalable Online Neural-Symbolic Learning via Accelerating Active Predictive Coding with Information Geometry & Diverse Cognitive â¦

URL: View paper

Brief Assessment

Active Predictive Coding[63] focuses on reinforcement learning with neural generative coding circuits and optimal transport geometry, not on vision foundation models with memory-based iterative refinement for image classification.

7. Hybrid predictive coding: Inferring, fast and slow

URL: View paper

Brief Assessment

[Final Audit Failure] The model insisted on a refutation claim but failed to provide verifiable evidence after multiple retries. Marked as cannot_refute for safety. Please manually verify the candidate text.

8. Modelling Predictive Coding in the Primary Visual Cortex (V1): Layer 4 Receptive Field Properties in a Balanced Recurrent Spiking Neuronal Network

URL: View paper

Brief Assessment

Predictive Coding V1[69] focuses on spiking neural network models of V1 layer 4 with membrane potential dynamics encoding reconstruction errors, not on iterative refinement update rules for vision foundation models.

9. Neural elements for predictive coding

URL: View paper

Prior Art Analysis

Neural Predictive Elements[66] demonstrates that predictive-coding-inspired iterative refinement mechanisms for error correction in neural networks were proposed and analyzed prior to the original paper. The candidate paper presents a comprehensive framework where representations are iteratively corrected through prediction-error minimization, with explicit mathematical formulations showing how error signals drive refinement updates. The candidate describes how 'the model optimizes as error minimizes' through iterative processes, and provides detailed neural circuit implementations of error-correction dynamics. This establishes that the core concept of using predictive coding principles for iterative refinement in neural architectures predates the original work.

Evidence

Evidence 1 - **Rationale**: Both papers describe iterative refinement mechanisms based on predictive coding principles. The candidate establishes the foundational concept of iterative optimization through prediction-error minimization, which the original paper claims as novel. - **Original**: we propose the vision hopfield memory network (v-hmn), a brain-inspired foundation backbone that integrates hierarchical memory mechanisms with iterative refinement updates. specifically, v-hmn incorporates local hopfield modules that provide associative memory dynamics at the image patch level, glo... - **Candidate**: predictive coding theories of sensory brain function interpret the hierarchical construction of the cerebral cortex as a bayesian, generative model capable of predicting the sensory data consistent with any given percept. predictions are fed backward in the hierarchy and reciprocated by prediction e...

Evidence 2 - **Rationale**: The candidate paper explicitly discusses predictive coding as a framework for understanding error-corrective mechanisms in neural circuits, predating the original's claim of introducing this as a novel lightweight mechanism. - **Original**: this mechanism can be viewed as a lightweight form of predictive-coding dynamics, where representations are gradually corrected toward memory-predicted patterns. in this way, the network gains an error-corrective feedback process that is absent in conventional feedforward models. - **Candidate**: this simple conception, the hierarchical exchange of prediction and prediction error, confronts a rich cortical microcircuitry that is yet to be fully documented. this article presents the view that, in the current state of theory and practice, it is profitable to begin a two-way exchange: that predi...

Evidence 3 - **Rationale**: Both papers describe iterative update mechanisms where current representations are refined based on prediction errors. The candidate provides detailed circuit-level implementations of these iterative refinement dynamics. - **Original**: the central operation in both local and global modules is an iterative refinement rule. given a current representationzand a retrieved prototypem, the update is z(t+1) =z (t) +β(m-z (t)), whereβis a learnable update strength andtdenotes the refinement step. this mechanism can be viewed as apredictiv... - **Candidate**: the steps taken by the backward pathway leading to the suppression of error unit activity are discussed below in section 10; here, we attempt to trace it one step further, in the form of a subsequent inhibitory link from error unit back to expectation unit (figure 3 , pathway 8). this establishes a...

Evidence 4 - **Rationale**: The candidate explicitly describes the iterative optimization process through error minimization, which is the core principle the original paper claims to introduce as a novel predictive-coding-inspired mechanism. - **Original**: this mechanism can be viewed as a lightweight form of predictive-coding dynamics, where representations are gradually corrected toward memory-predicted patterns. - **Candidate**: thus begins an iterative process: the model optimizes as error minimizes. curiously enough our academic understanding of the brain may profit from a similar iterative strategy.

10. Tight stability, convergence, and robustness bounds for predictive coding networks

URL: View paper

Brief Assessment

Predictive Coding Stability[62] focuses on theoretical stability and convergence analysis of predictive coding networks through dynamical systems theory, not on iterative refinement mechanisms for error correction in vision models. The candidate analyzes mathematical properties of PC dynamics rather than proposing refinement architectures for neural networks.

Contribution 3: Class-balanced persistent memory banks with content-addressable retrieval

Description: The authors design explicit memory banks that store real sample embeddings in a class-balanced manner during training and remain frozen during inference. These banks enable content-addressable retrieval where stored prototypes act as reusable priors, improving data efficiency and interpretability.

This contribution was assessed against **2 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. AMDNet23: A combined deep Contour-based Convolutional Neural Network and Long Short Term Memory system to diagnose Age-related Macular Degeneration

URL: View paper

Brief Assessment

AMDNet Deep Learning[60] focuses on medical image classification using CNN-LSTM architectures for AMD diagnosis, not on class-balanced memory banks with content-addressable retrieval for vision models. The candidate does not address memory-based retrieval mechanisms or prototype storage systems.

2. Applied LLaMA: Systems, Methods, and Implementations

URL: View paper

Brief Assessment

Applied LLaMA[59] mentions content-addressable storage and class-balanced transitions only in passing, without describing a memory bank architecture for vision models or prototype-based retrieval mechanisms.

Appendix: Text Similarity Detection

No high-similarity text segments were detected across any compared papers.

References

- [0] Vision Hopfield Memory Networks View paper
- [1] RoboMemory: A Brain-inspired Multi-memory Agentic Framework for Lifelong Learning in Physical Embodied Systems View paper
- [2] RoboMemory: A Brain-inspired Multi-memory Agentic Framework for Interactive Environmental Learning in Physical Embodied Systems View paper
- [3] Full hardware implementation of neuromorphic visual system based on multimodal optoelectronic resistive memory arrays for versatile image processing View paper
- [4] In-sensor image memorization and encoding via optical neurons for bio-stimulus domain reduction toward visual cognitive processing View paper
- [5] A brain-inspired hierarchical interactive in-memory computing system and its application in video sentiment analysis View paper
- [6] Hierarchical Spiking-Based Model for Efficient Image Classification With Enhanced Feature Extraction and Encoding View paper
- [7] A scalable reinforcement learning framework inspired by hippocampal memory mechanisms for efficient contextual and sequential decision making View paper
- [8] Lightweight salient object detection via hierarchical visual perception learning View paper
- [9] Bio-inspired multi-level interactive contour detection network View paper
- [10] Hierarchical Activation Dual Backbone Network for Weakly Supervised Semantic Segmentation View paper
- [11] Bio-inspired smart vision sensor: toward a reconfigurable hardware modeling of the hierarchical processing in the brain View paper
- [12] Understanding cortical computation through the lens of joint-embedding predictive architectures View paper
- [13] HiCMÂ²: Hierarchical Compact Memory Modeling for Dense Video Captioning View paper
- [14] M-GCN: Brain-inspired memory graph convolutional network for multi-label image recognition View paper
- [15] Seeing the Overlooked: Bio-Visual Inspired Weak Saliency Feedback Transformer for Person Re-identification View paper
- [16] Neural Brain: A Neuroscience-inspired Framework for Embodied Agents View paper
- [17] Unsupervised Learning of a Hierarchical Spiking Neural Network for Optical Flow Estimation: From Events to Global Motion Perception View paper
- [18] Hierarchical temporal memory based on spin-neurons and resistive memory for energy-efficient brain-inspired computing View paper
- [19] Hierarchical temporal memory features with memristor logic circuits for pattern recognition View paper
- [20] Deep recurrent neural network reveals a hierarchy of process memory during dynamic natural vision View paper
- [21] Vision-based topological mapping and navigation with self-organizing neural networks View paper
- [22] Brain-Inspired Stepwise Patch Merging for Vision Transformers View paper
- [23] Working memory inspired hierarchical video decomposition with transformative representations View paper
- [24] Neuromorphic architecture for the hierarchical temporal memory View paper
- [25] Enhancing Biologically Inspired Hierarchical Temporal Memory with Hardware-Accelerated Reflex Memory View paper
- [26] Brain Inspired Adaptive Memory Dual-Net for Few-Shot Image Classification View paper
- [27] A Brain-inspired Method for Occluded 3D Object Recognition View paper
- [28] A model of the ventral visual system based on temporal stability and local memory View paper
- [29] Hierarchical associative memory based on oscillatory neural network View paper
- [30] STM-SalNet: A Biologically-Inspired Spatial-Temporal Memory Network for Video Saliency Prediction View paper
- [31] Decoding Mouse Visual Tasks via Hierarchical Neural-Information Gradients View paper
- [32] Object recognition using hierarchical temporal memory View paper
- [33] Feature extraction without learning in an analog spatial pooler memristive-CMOS circuit design of hierarchical temporal memory View paper
- [34] A recurrent vision transformer shows signatures of primate visual attention View paper
- [35] Introducing memory and association mechanism into a biologically inspired visual model View paper
- [36] THIRDEYE: Cue-Aware Monocular Depth Estimation via Brain-Inspired Multi-Stage Fusion View paper
- [37] A brain-inspired network architecture for cost-efficient object recognition in shallow hierarchical neural networks. View paper
- [38] Brain-inspired Hierarchical Attention Recurrent CNN for Image Classification View paper
- [39] A biologically motivated visual memory architecture for online learning of objects View paper
- [40] Hypernetworks: A molecular evolutionary architecture for cognitive learning and memory View paper

- [41] Convolutional neural networks uncover the dynamics of human visual memory representations over time. View paper
- [42] On the self-organization of a hierarchical memory for compositional object representation in the visual cortex View paper
- [43] Artificial Working Memory Constructed by Planar 2D Channel Memristors Enabling Brainâ□□Inspired Hierarchical Memory Systems View paper
- [44] A Biologically Inspired Visual Working Memory for Deep Networks View paper
- [45] Biologically Inspired Hierarchical Model for Feature Extraction and Localization View paper
- [46] A biologically inspired spiking neural P system in selective visual attention for efficient feature extraction from human motion View paper
- [47] A Memory-Based Recurrent Neural Architecture for Chip Emulating Cortical Visual Processing View paper
- [48] The Hierarchical Evolution in Human Vision Modeling. View paper
- [49] Neural Architecture for Feature Binding in Visual Working Memory. View paper
- [50] Advances in Neural Information Processing Systems 13: Proceedings of the 2000 Conference View paper
- [51] Semantic enhancement and multi-level alignment network for cross-modal retrieval View paper
- [52] A universal abstraction for hierarchical hopfield networks View paper
- [53] STanhop: Sparse tandem hopfield model for memory-enhanced time series prediction View paper
- [54] iMixer: hierarchical Hopfield network implies an invertible, implicit and iterative MLP-Mixer View paper
- [55] Energy-Based Learning and the Evolution of Hopfield Networks: From Boltzmann Machines to Transformer Attention Mechanisms View paper
- [56] Out-of-Distribution Nuclei Segmentation in Histology Imaging via Liquid Neural Networks with Modern Hopfield Layer View paper
- [57] Entropy driven artificial neuronal networks and sensorial representation: A proposal View paper
- [58] Hierarchical Hopfield Network Decomposition: A Spiked Covariance Framework for Latent Prototype Discovery View paper
- [59] Applied LLaMA: Systems, Methods, and Implementations View paper
- [60] AMDNet23: A combined deep Contour-based Convolutional Neural Network and Long Short Term Memory system to diagnose Age-related Macular Degeneration View paper
- [61] Anopcn: Video anomaly detection via deep predictive coding network View paper
- [62] Tight stability, convergence, and robustness bounds for predictive coding networks View paper
- [63] ActPC-Geom: Towards Scalable Online Neural-Symbolic Learning via Accelerating Active Predictive Coding with Information Geometry & Diverse Cognitive â□¦ View paper
- [64] Neurocomputational Mechanisms of Sense of Agency: Literature Review for Integrating Predictive Coding and Adaptive Control in Humanâ□□Machine Interfaces View paper
- [65] Associative memories via predictive coding View paper
- [66] Neural elements for predictive coding View paper
- [67] Hybrid predictive coding: Inferring, fast and slow View paper
- [68] An approximation of the error backpropagation algorithm in a predictive coding network with local hebbian synaptic plasticity View paper
- [69] Modelling Predictive Coding in the Primary Visual Cortex (V1): Layer 4 Receptive Field Properties in a Balanced Recurrent Spiking Neuronal Network View paper
- [70] Towards the Training of Deeper Predictive Coding Neural Networks View paper