Toward Neuro-Symbolic and Reservoir-Inspired Medical Imaging: A BD-CeNN Autoencoder with ASP Rule Mining for Robust and Explainable Interpretation of Grayscale and Color Images

Mugisha Muhire Jean-Pierre[§], Kyandoghere Kyamakya[*§], Christian Kutabuna Kubakisa[§],
Isaac Lukusa Kayembe[§], Divin Kayeye Kabeya[*], Samuel Muia Ngom[‡],
Witesyavwirwa Vianney Kambale[*‡], Hima Deepthi Vankayalapati[†], Selain Kasereka Kabunga[‡§]
[*]University of Klagenfurt, Institute of Smart Systems Technologies, Klagenfurt, Austria
[†]Faculty of English Engineering Studies, Technical University of Sofia, Sofia, Bulgaria
[‡]Faculty of Information and Communication Technology, Tshwane University of Technology, Pretoria, South Africa
[§]Faculté Polytechnique, Université de Kinshasa (UNIKIN), Kinshasa, DRC

***Abstract*—We introduce a novel neuro-symbolic AI framework for medical imaging that combines a Binary Discrete Cellular Neural Network (BD-CeNN)-based autoencoder, a reservoir computing (RC)-inspired feature refinement pipeline, and automatically mined Answer Set Programming (ASP) rules. Designed for interpreting grayscale and color medical images—such as CT, MRI, histopathology, and dermatology—the proposed hybrid architecture balances diagnostic accuracy, interpretability, and deployment feasibility. The BD-CeNN autoencoder transforms complex visual inputs into discrete, symbolic latent representations, then iteratively refined through a reservoir-style BD-CeNN stack. On top of this, ASP rules learned from annotated data enable logic-driven, transparent reasoning, producing diagnostic outputs that are both accurate and fully auditable. This framework suits resource-constrained clinical environments where trust, explainability, and adaptability are critical. The primary objective of this paper is to present the conceptual underpinnings of this innovative architecture, to highlight its scientific and practical potential, and to assess its feasibility for real-world deployment. While extensive experimentation and benchmarking are reserved for future work, this contribution lays the essential groundwork for a comprehensive, interpretable, and accessible AI paradigm in medical image analysis.**

***Index Terms*—Neuro-symbolic artificial intelligence, explainable medical imaging, Binary Discrete Cellular Neural Networks (BD-CeNN), reservoir computing, Answer Set Programming (ASP), multimodal symbolic encoding, edge-deployable diagnostic AI, symbolic feature extraction, temporal symbolic reasoning, interpretable clinical decision support.**

I. Introduction

The rapid evolution of artificial intelligence (AI) has significantly transformed medical imaging by enhancing diagnostic accuracy and automation. In particular, deep learning techniques—most notably convolutional neural networks (CNNs)—have achieved human-level or even superior performance in various tasks such as lesion detection, disease classification, and image segmentation across a wide range of modalities, including radiography, magnetic resonance imaging (MRI), computed tomography (CT), ultrasound, and digital pathology [1]–[3]. However, despite these advancements, the inherent opacity of such models continues to hinder their integration into clinical workflows, where interpretability, reliability, and clinician trust are essential for decision support [2], [3].

A. Limitations of Existing Methods

Despite their success, black-box deep learning models frequently lack the explanatory clarity required for clinical decision-making. Although post-hoc interpretability techniques such as Grad-CAM, SHAP, and LIME offer visual explanations, these methods often fall short in robustness, clinical reliability, and alignment with human reasoning processes [2], [4], [5]. Contemporary surveys emphasize that genuinely trustworthy AI systems must be intrinsically interpretable, clinically grounded, and designed with end-user needs in mind [2], [6]. Moreover, conventional static feedforward architectures are ill-equipped to replicate human-like cognitive processes. Clinicians often revisit regions of interest, integrate domain knowledge, and iteratively revise their interpretations—capabilities largely absent in standard deep learning systems [7], [8]. Another primary concern involves the high computational demands of state-of-the-art models. Many top-performing AI systems rely on vast datasets, powerful GPUs, and energy-intensive infrastructure—rendering them unsuitable for deployment in resource-constrained settings such as rural hospitals and mobile diagnostic units [9], [10].

B. Toward Neuro-Symbolic and Temporal Dynamics

We introduce a neuro-symbolic framework integrating (1) a BD-CeNN-based autoencoder to extract symbolic representations from grayscale and color medical images, (2) A reservoir-style BD-CeNN stack to capture temporal and contextual dynamics, and (3) an ASP (Answer Set Programming) engine to enable interpretable, rule-based diagnostic reasoning.

BD-CeNN Autoencoding: Binary Discrete Cellular Neural Networks (BD-CeNNs), which generalize the classical Cellular Neural Network paradigm, support local binary computation conducive to symbolic abstraction and efficient hardware implementation [11], [12]. These networks serve as interpretable autoencoders capable of aligning extracted features with clinically meaningful concepts while maintaining high computational efficiency. Their binary nature enhances compatibility with edge-computing devices and supports real-time deployment scenarios [13].

Reservoir-Style BD-CeNN: Inspired by the reservoir computing paradigm—previously applied in temporal domains such as EEG classification and radiographic sequence analysis—we adapt BD-CeNN layers to function as symbolic reservoirs. These stacked layers simulate cognitive reasoning cycles by embedding temporal dependencies and iterative pattern integration [14], [15]. This approach addresses a critical shortcoming in traditional deep architectures: the inability to emulate dynamic inference pathways essential for nuanced clinical interpretation.

ASP Rule-Based Reasoning: Answer Set Programming (ASP) is a form of declarative logic programming that has gained traction for its formal rigor and human-auditable output [16]. Our ASP module provides an interpretable reasoning layer atop neural feature extraction by learning symbolic rules from labeled datasets. This logic-based engine enables transparent diagnostic pathways and supports clinical audibility—essential features often missing in deep learning systems. While ASP has roots in expert systems, its integration with neuro-symbolic models for visual diagnostic tasks remains largely unexplored [17], [18].

C. Innovation and Scope

This triadic architecture systematically addresses four significant limitations: (1) *Inherent interpretability* The entire pipeline is designed for transparency, addressing critical demands for explainable AI in clinical contexts [2], [4], [19], (2) *Multimodal flexibility* It supports both grayscale and color modalities, enabling applications across radiology, dermatology, pathology, and beyond [1], [3], [11], (3) *Temporal reasoning capability* Through the symbolic reservoir stack, our system mimics iterative human reasoning for dynamic diagnostic contexts [14], [15], and (4) *Edge deployability* The lightweight nature of BD-CeNNs, combined with the logic-driven ASP layer, enables AI diagnostics in low-resource environments without sacrificing explainability or accuracy [12], [13], [20].

Our previous works [21], [22], [22]–[25] demonstrate a long-standing and diverse expertise in cellular neural networks (CeNNs): the authors have published a series of peer-reviewed studies—spanning 2011 to 2019—showing CeNNs as ultra-fast, flexible solvers for stiff ordinary and partial differential equations, real-time computational engineering, local traffic-signal control, time-series modeling and forecasting in transportation, and even raindrop detection for advanced driver-assistance systems. Collectively, these references [21], [22], [22]–[25] establish a solid, multidisciplinary foundation in designing CeNN-based algorithms that operate in real-time and address practical engineering problems.

II. A Critical Review of the Related State-of-the-Art

This section presents a comprehensive review of existing research, emphasizing scientific gaps this work addresses and highlighting innovation potential.

A. Explainable AI in Medical Imaging

Demand for explainable AI (XAI) in medical diagnostics increases steadily due to clinical safety, regulatory compliance, and trust implications [2], [4], [6]. Current XAI strategies fall into two categories: post-hoc methods (Grad-CAM, SHAP, LIME) that often lack fidelity and stability in complex medical settings [2], [5], [19], and self-explainable models that remain scarce in clinical workflows due to architectural limitations and performance trade-offs [7], [19]. Leading reviews stress that medical XAI must be truthful, clinically relevant, and resource-efficient [6], [26], yet few existing methods meet these criteria under operational constraints.

B. Hybrid and Neuro-Symbolic AI

Neuro-symbolic systems fuse sub-symbolic learning with symbolic reasoning, combining neural perception with logical inference [17], [27]. Frameworks like NeuroSymAD apply CNNs alongside ASP-based logic rules for Alzheimer's diagnosis [26], providing interpretable justifications aligned with clinical reasoning. However, most systems are limited to grayscale neuroimaging, rely on floating-point CNNs, and lack binary symbolic processing or color image integration, restricting generalizability across imaging domains.

C. Reservoir Computing in Medical Imaging

Reservoir computing (RC)—including Echo State Networks and Liquid State Machines—proves effective in temporal signal processing for EEG-based emotion recognition and seizure detection [14], [15]. Despite the potential for modeling temporal dependencies and iterative reasoning, RC has not been systematically applied to static medical images like CT or histopathology, presenting opportunities for dynamic feature refinement in image analysis.

D. Discrete Neural Networks and BD-CeNNs

Binary Discrete Cellular Neural Networks (BD-CeNNs) extend traditional CNNs through localized binary computations, facilitating high interpretability, low complexity, and edge deployability [11]–[13]. Used for image segmentation and denoising, their integration into reservoir architectures or neuro-symbolic frameworks for medical imaging remains unexplored. No current system fully leverages discretized symbolic encoding across grayscale and color modalities in unified pipelines.

E. ASP Rule Mining in Diagnostics

Answer Set Programming (ASP) supports logic-based reasoning with auditable decision paths, critical for trustworthy medical AI [16]–[18]. Prior diagnostic systems often required manually crafted rules, limiting scalability. Recent works auto-extract rules from CNN activations [26] but lack symbolic abstraction and discrete representations. There is no framework that currently auto-mines ASP rules from BD-CeNN symbolic encodings across diverse imaging types.

F. Scientific Gaps & Innovation Potential

Table I provides a comparative analysis of prevailing approaches, unresolved challenges, and innovations introduced by our framework. While progress exists in deep learning diagnostics, the field remains hampered by significant shortcomings affecting explainability, generalizability, and clinical integration. Our approach addresses six limitations, combining symbolic reasoning, discrete computation, dynamic inference, and lightweight architecture in unified neuro-symbolic systems.

1) Interpretability Gap: Current medical AI lacks inherently interpretable models, relying on post-hoc methods (Grad-CAM, SHAP, LIME) that frequently yield neither clinically validated nor robust explanations across samples [2], [5]. These black-box models offer no guarantees that explanations correspond to genuine causal factors, with outputs difficult to audit in medico-legal contexts. Our solution embeds interpretability architecturally using BD-CeNN autoencoders generating symbolic features interpreted via ASP-based logic reasoning, ensuring explanations are produced during inference rather than reverse-engineered [2], [6], [19].

2) Symbolic Integration: Most deep learning pipelines lack symbolic integration. Even neuro-symbolic approaches often isolate symbolic components to post-processing steps, detached from learned representations. Our model natively encodes symbolic structure through BD-CeNNs and maintains symbolic continuity by coupling with ASP-based inference, bridging neural perception with logic-driven decision-making closer to clinician reasoning [17], [27].

3) Temporal Inference: Diagnostic reasoning is iterative and temporal—clinicians revisit regions, adjust hypotheses, and incorporate contextual cues over time. Most AI systems process images in static, one-pass feedforward models, ignoring temporal interpretation aspects. Inspired by reservoir computing, our BD-CeNN reservoir layers preserve prior activations and evolve symbolic features across multiple inference cycles, mimicking temporal dynamics and recursive attention in expert clinical practice [14], [15].

4) Automated Rule Induction: Traditional ASP rule engines require manually defined rule sets, which are challenging to scale and maintain [16]–[18]. Recent approaches extract rules from CNN activations but lack symbolic fidelity or operate on fuzzy features. Our approach automates rule mining directly from discretized BD-CeNN symbolic features, creating meaningful ASP rules with strong semantic alignment to image content and clinical categories [18], [26], producing interpretable and data-driven explanations.

5) Multimodal Support: Medical imaging encompasses diverse modalities—grayscale (X-rays, MRIs) and color-rich images (dermatological scans, histopathology). Many XAI and neuro-symbolic systems are restricted to grayscale modalities, especially neurological imaging [26]. Our framework generalizes across grayscale and color inputs, supporting wider clinical applications and enhancing usability in mixed-imaging workflows across departments and specialties.

6) Resource-Efficiency and Edge Deployability: High-performing CNNs require significant computational resources, limiting adoption in rural clinics, point-of-care devices, and mobile units where compute capacity and internet access are limited [9], [10], [20]. BD-CeNNs operate with binary, localized computations, drastically reducing hardware requirements and power consumption [11], [13]. Coupled with lightweight ASP symbolic reasoning, the result suits edge deployment in resource-constrained environments.

G. Why This Matters

By embedding interpretability into core architecture, this framework overcomes post-hoc XAI brittleness. As an S-XAI model, it aligns with clinical demands for auditability, transparency, and efficiency [4], [6], [19]. BD-CeNN reservoirs offer novel approaches to iterative reasoning, simulating multi-pass diagnostic reasoning used by clinicians [8], [27]. The system generalizes to both grayscale and color imaging—contrasting with narrow existing XAI systems—and leverages binary symbolic encoding, optimizing speed, power consumption, and interpretability [11], [13], [20].

To our knowledge, this is the first framework integrating symbolic BD-CeNN autoencoding, temporal refinement through reservoir-like structures and ASP-based logic reasoning for diagnosis. This unified design lays the foundations for next-generation neuro-symbolic AI tailored for transparent, adaptive, and deployable medical diagnostics.

III. Our System Architecture

The proposed pipeline unfolds four stages (Figure 2), blending data-driven and logic-driven mechanisms for explainable medical image analysis.

A. Stage 1: BD-CeNN Autoencoder – Symbolic Latent Encoding

This stage transforms raw medical images into compact, symbolic, and interpretable representations using a Binary Discrete Cellular Neural Network (BD-CeNN)-based autoencoder. The system extracts clinically relevant features as logical, structured tokens that serve as the cognitive interface between machine perception and human-level reasoning.

1) From Visual Data to Symbolic Knowledge: In conceptual terms, clinicians traditionally analyze medical images: ”The lesion has spiculated margins and central necrosis” or ”This pattern resembles honeycombing seen in pulmonary fibrosis.” The BD-CeNN autoencoder addresses CNN limitations by

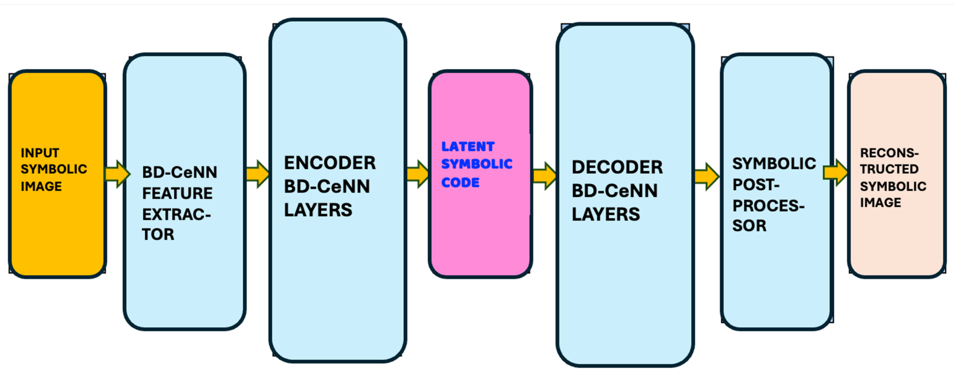
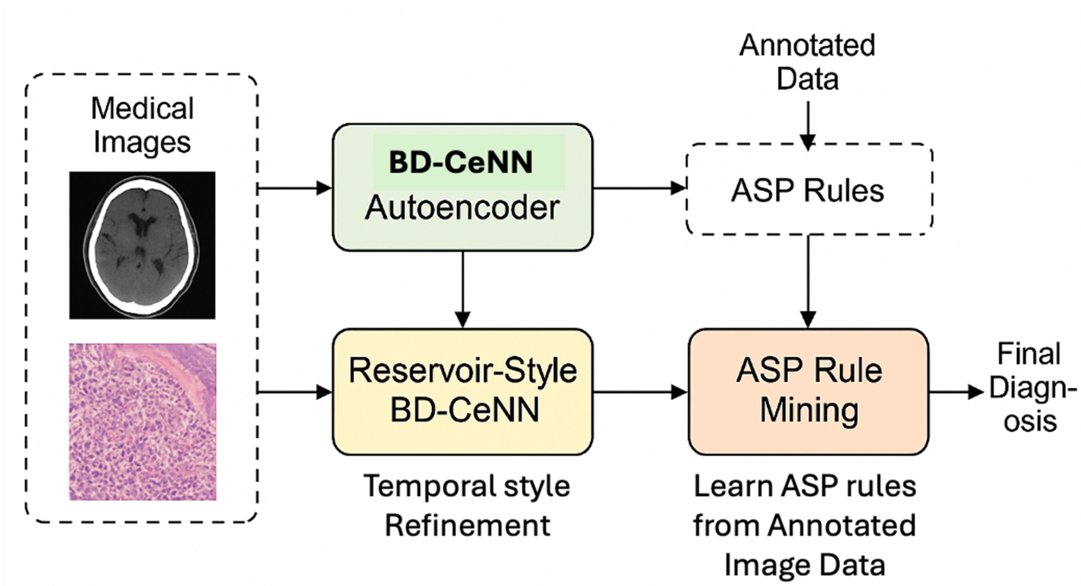
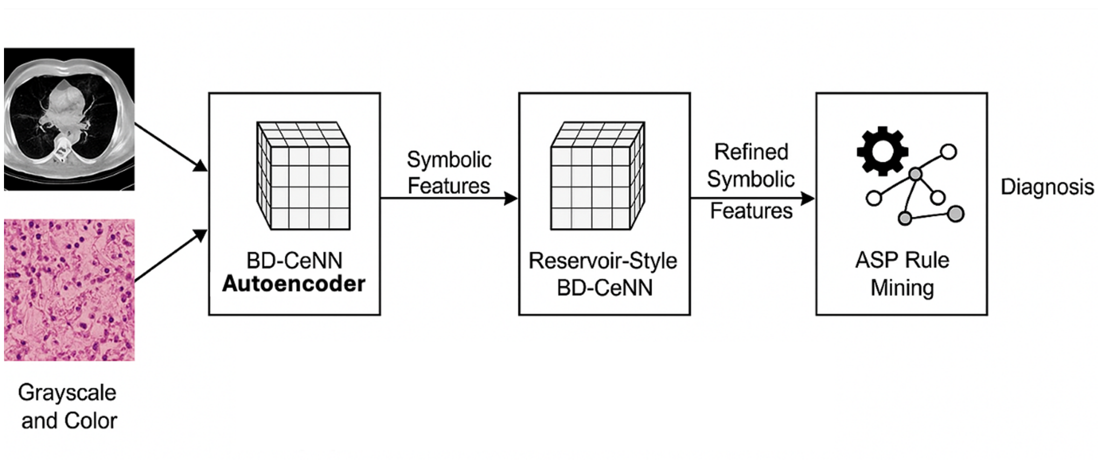


Fig. 1: BD-CeNN Autoencoder – visual representation of encoding/decoding symbolic features.



(a) Training Phase: BD-CeNN Autoencoder with ASP Rule Learning



(b) Inference Phase: Full Neuro-Symbolic Pipeline

Fig. 2: Full Neuro-Symbolic Pipeline – from input image to explainable output.

TABLE I: Comparison of Challenges, Existing Approaches, Unmet Needs, and Our Contributions

Challenge	Existing Approaches	Unmet Need	Our Contribution
Interpretability	Post-hoc XAI, self-explainable models [2], [6], [19]	Lack of intrinsic, auditable interpretability	BD-CeNN + ASP pipeline with built-in symbolic explanations
Symbolic Integration	Detached symbolic layers or none [17], [27]	Disconnection from learned features	Symbolic encoding and reasoning integrated throughout the pipeline
Temporal Inference	Static feedforward models; RC only in EEG/time series [14], [15]	No iterative reasoning for static images	BD-CeNN reservoir mimicking temporal diagnostic reasoning
Rule Induction	Manual rule design or fuzzy CNN-based rules [16], [18], [26]	No scalable, discrete rule mining from interpretable features	Automatic ASP rule generation from symbolic BD-CeNN features
Multimodal Support	Limited to grayscale or narrow domains [26]	Lack of generalization across imaging types	Full support for grayscale and color imaging
Edge Deployment	Deep CNNs: computationally intensive [9], [10], [20]	Inaccessible to low-resource or mobile clinical settings	Efficient, interpretable, and low-power model suitable for on-device deployment

directly translating visual data into symbolic assertions, capturing what is present, where it is located, and how it behaves, and organizing this information into discrete logic-compatible structures.

2) Preprocessing and Cross-Modal Harmonization: A harmonization process ensures modality-agnostic symbolic encoding: grayscale images (CT, MRI) focus on density gradients and anatomical location, while color images (histology, dermatology) encode hue and contrast into symbolic units. BD-CeNN layers normalize clinical visual cues into a unified symbolic vocabulary, enabling cross-modal logic.

3) BD-CeNN Architecture: BD-CeNN consists of a 2D cellular grid where each binary cell's next state depends on its current state, neighbors, and discrete update rules. This architecture provides spatial pattern emergence without convolutions, discrete semantics from local rules, and composability ideal for symbolic AI. Each layer acts as a symbolic transformation stage, progressively abstracting pixel-level data into structured patterns.

4) Encoding Pipeline: **Layer 1:** Detects contrasts, blobs, gradients, and boundary primitives, outputting binary maps like "edge present" and "contrast peak."

Layers 2-3: Aggregate patterns across neighborhoods, encoding regional features like "ring shape" and "bilateral symmetry."

Top Layer: Maps latent patterns to symbolic units using a pre-learned dictionary, outputting labels such as:

```
lesion_border = irregular, intensity_core = high,
↪ color_pattern = reticular.
```

The symbolic latent space is sparse (only meaningful concepts are activated), discrete (no ambiguous floating-point values), semantic (each symbol is meaningful and explainable), and relational (symbols participate in logic rules).

5) Clinical Example: CT Chest Scan: .

```
1. BD-CeNN layer 1: detects ring-shaped high-intensity
↪ regions
2. Layers 2-3: Characterize boundary as "spiculated,"
↪ location as "peripheral"
3. Final layer outputs: lesion_shape=irregular,
↪ density_core=hyperdense,
   boundary_type=spiculated, position=peripheral
4. ASP interprets: IF lesion_shape=irregular AND
↪ density_core=hyperdense
   AND boundary_type=spiculated THEN risk=high_malignancy
```

6) Key Benefits: The framework provides interpretability through semantic token mapping, composability enabling incremental concept addition without retraining, explainability via human-traceable inference chains, edge deployment through lightweight BD-CeNN logic, ontology compatibility with SNOMED/ICD coding systems, and interactive feedback loops for clinician-AI collaboration.

B. Stage 2: BD-CeNN Reservoir – Temporal Symbolic Refinement

The BD-CeNN reservoir refines symbolic features through multiple discrete time steps, emulating iterative human diagnostic reasoning. This Reservoir Computing-based architecture allows symbolic information to evolve temporally, creating a system that thinks repeatedly and contextually.

1) Architecture and Virtual Time Steps: The reservoir comprises stacked binary processing layers, each applying fixed local update rules to symbolic feature maps. Though medical images are static, the reservoir treats them as temporally unfolding symbolic events, with each layer performing incremental operations analogous to diagnostic iterations—first scanning for gross abnormalities, then focusing on finer details.

Key benefits include contextual awareness (later layers integrate broader patterns), symbolic consistency enforcement (resolving contradictions via propagation dynamics), noise suppression (stabilizing high-confidence patterns), and diagnostic focus emergence (converging on critical symbolic features).

2) Memory-Like Behavior: Using fixed transition rules with learning occurring only at later stages, the system enables fast deployment, better explainability, low computational cost, and neuromorphic hardware compatibility. Despite this, emergent memory behavior allows earlier symbolic decisions to influence later ones, supporting temporal integration.

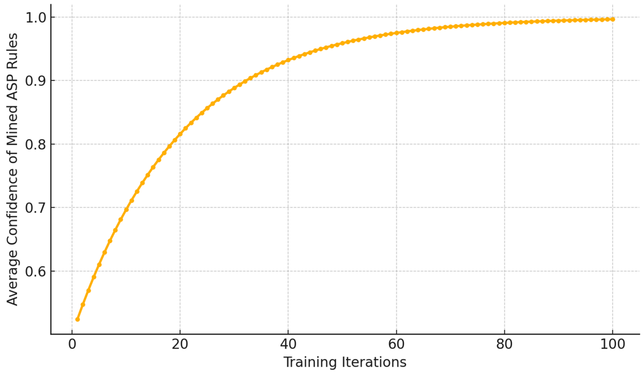


Fig. 3: ASP Rule Learning Curve showing confidence improvement over 100 training iterations, reflecting growing symbolic reasoning accuracy.

C. Stage 3: ASP Rule Mining – Symbolic Knowledge Extraction

This stage transitions from sub-symbolic processing to explicit symbolic reasoning using Answer Set Programming (ASP). Symbolic features are the foundation for constructing logical rules that enable transparent, auditable, and evolvable reasoning.

1) Symbolic Feature Space: Each image is reduced to structured symbolic representations, including morphological descriptors (size, border sharpness), textural properties (granularity, contrast), spatial relationships (central vs peripheral), and contextual metadata (patient age, anatomical zone). These interpretable descriptors form a semantically meaningful basis for reasoning.

2) Automated Rule Mining: An ASP-based algorithm uses supervised training data to search for consistent, minimal, discriminative rules connecting symbolic patterns to diagnoses. Example rules:

```
IF border_irregularity=true AND density=hyperdense AND
↪ region=upper_lobe
THEN diagnosis=suspicious_nodule

IF asymmetry=high AND lesion_color=dark_brown AND
↪ pattern=radial_streaks
THEN diagnosis=likely_melanoma
```

3) ASP Rule Execution: During inference, the ASP engine matches rules against symbolic configurations, generating diagnoses with traceable justification chains. This supports logical entailment, conflict resolution, explanation generation, and diagnostic traceability—vital for clinical validation.

Key advantages include transparency (human-readable reasoning chains), modifiability (expert rule editing without retraining), interdisciplinary adaptability (incorporating non-image data), auditability for regulatory compliance, medical pedagogy support, and patient personalization capabilities.

D. Stage 4: Explainable Output + Symbolic Feedback

The final stage generates explainable, symbolic outputs that are diagnostically meaningful and fully auditable, translating symbolic representations into actionable conclusions through logic-based reasoning.

1) Structured Diagnostic Output: Diagnostic predictions reflect clinical assessments like ”Possible early-stage tumor detected in upper right quadrant” or ”Hyperpigmented region consistent with dermal inflammation.” Each diagnosis derives from symbolic features via ASP logical inferences, anchored in structured reasoning rather than opaque neural patterns.

2) Symbolic Reasoning Trace: The system provides transparent, step-by-step reasoning chains, such as:

1) IF lesion_size=large AND border_irregularity=true THEN suspicion_level=high
2) IF suspicion_level=high AND patient_age>50 THEN diagnosis=”possible malignant tumor”

This transparency enables clinician auditing and transforms the system from a black-box to a white-box diagnostic collaborator.

3) Bi-directional Symbolic Feedback: An optional feedback mechanism allows ASP reasoning output to influence upstream BD-CeNN reservoir processing, enabling state reconfiguration, iterative reasoning, and adaptive refinement. This creates a closed-loop system capable of dynamic reinterpretation and knowledge-guided evolution.

4) Clinical and Educational Impact: The framework addresses regulatory requirements (EU AI Act, FDA GMLP) through embedded interpretability, explicit symbolic representations, and logic-based reasoning paths. This enables clinician scrutiny, verification, and intervention while serving as a pedagogical tool for medical education by making AI's ”thought process” visible and traceable.

5) Customizability and Human-in-the-Loop: The symbolic architecture supports dynamic customization through direct reasoning layer intervention. Clinicians can modify rules, introduce new ones, or deactivate outdated patterns within the ASP engine without full retraining. This flexibility accommodates evolving medical knowledge, patient-specific nuances, and institutional standards while maintaining collaborative intelligence between AI and human expertise.

IV. Key Contributions

This work introduces a novel, unified neuro-symbolic framework combining discrete neural computation, temporal dynamics, and symbolic logic for medical image interpretation. It offers several key innovations that distinguish it from existing AI-based diagnostic systems, particularly in explainability, adaptability, and deployment feasibility. First, we propose the first BD-CeNN-based autoencoder specifically designed for symbolic encoding of medical images, capable of processing both grayscale (e.g., CT, X-ray, MRI) and color domains (e.g., histopathology, dermatology). Unlike conventional autoencoders that rely on continuous-valued latent representations, the BD-CeNN model encodes discrete symbolic states that are directly compatible with logic-based reasoning. This symbolic abstraction facilitates interpretable downstream processing while reducing computational complexity. Second, the framework introduces a reservoir computing–inspired stack of BD-CeNN layers that enables dynamic symbolic reasoning over image content. These layers operate sequentially in a pseudo-temporal fashion, refining and evolving symbolic feature maps as they propagate through the architecture. This design simulates a form of iterative, memory-based processing akin to the diagnostic reasoning used by clinicians—thus bridging the gap between static neural inference and temporally-aware, context-sensitive analysis. Third, the architecture seamlessly integrates perception and cognition, where the BD-CeNN-based feature extraction layers (perception) are tightly coupled with an ASP-based rule engine (cognition). This coupling enables symbolic features to be directly mapped onto formalized diagnostic logic, supporting transparent, auditable, and expert-verifiable decision-making. The process remains traceable from raw input to symbolic output, a rare characteristic in modern AI pipelines. Finally, the framework is optimized for real-world applicability in constrained environments. Its discrete, low-resource BD-CeNN components and symbolic inference engine are well-suited for edge deployment in settings where computational resources, stable connectivity, and technical support may be limited. This includes rural health centers, mobile diagnostic platforms, emergency care units, and regulatory-sensitive environments. The model's interpretable nature aligns it with global trends toward ethical, human-centric healthcare AI, supporting accountability and clinical adoption. Together, these contributions establish a strong foundation for advancing the field of explainable and accessible medical AI, particularly at the intersection of neuro-symbolic reasoning and practical healthcare deployment.

V. Real-World Impact

The proposed system offers substantial promise for healthcare implementation across clinical diagnostics, medical education, and public health. Its hybrid neuro-symbolic design ensures high diagnostic accuracy while delivering explainable, traceable, and efficient decision-making that meets modern demands for transparent medical AI.

Clinical Applications: In radiology, the system analyzes grayscale modalities (X-rays, CT, MRI) by extracting symbolic features like geometric boundaries, intensity gradients, and structural anomalies. These map to logic-based rules, generating interpretable diagnostic suggestions that clinicians can validate and trace through symbolic reasoning paths. This explainability builds trust, reduces diagnostic ambiguity, and enhances radiologist acceptance, particularly for subtle abnormalities or multi-region pathologies.

For dermatology and pathology, where accuracy depends on color variation, texture, and morphological patterns, the system provides a symbolic representation of color-specific and shape-sensitive features. It enables robust skin lesion classification (melanoma detection) and histopathological cell identification. It is especially valuable for early cancer detection, infections, and autoimmune diseases where minor visual deviations carry significant diagnostic implications.

Telemedicine and Resource-Constrained Environments: The framework excels in telemedicine by transmitting symbolic encodings rather than full-resolution images in bandwidth-constrained environments (rural clinics, field hospitals, mobile units). These compressed representations preserve critical diagnostic cues while dramatically reducing data size, enabling real-time AI-assisted diagnostics on mobile or embedded devices.

Medical Education: The system serves as an interactive training tool by exposing symbolic structures underlying diagnostic outputs and rule-based reasoning chains. Trainees understand what the AI concludes and why, promoting conceptual clarity, supporting diagnostic skill-building, and fostering accountable AI-assisted practice. This enables curriculum integration in radiology, digital pathology, and clinical decision-making courses.

Regulatory Compliance: The symbolic foundation supports clinical auditability and compliance with regulatory requirements, including the EU AI Act and FDA GMLP guidelines, aligning with emerging standards for human-centered, trustworthy medical AI.

The architecture provides a flexible, future-ready solution spanning primary care, specialty medicine, education, and telehealth—offering a scalable model for integrating explainable AI into healthcare delivery.

VI. Conclusion and Future Work

We present a reservoir-enhanced neuro-symbolic AI architecture fusing discrete symbolic encoding, dynamic feature propagation, and logic-based reasoning to optimize diagnostic performance and model interpretability jointly. The Binary Discrete Cellular Neural Network (BD-CeNN) core encodes grayscale and color medical images into evolving, compact symbolic representations refined through reservoir computing dynamics and interpreted via Answer Set Programming (ASP) rules automatically derived from expert-labeled datasets—enabling logic-grounded, auditable decision-making [11], [14], [16].

This architecture addresses critical medical AI limitations: First, it delivers inherent interpretability by embedding explainability within model structure, surpassing post-hoc techniques like Grad-CAM or SHAP that often lack reliability and clinical alignment [2], [4], [6]. Second, reservoir-inspired symbolic refinement enables context-aware, temporally structured feature evolution missing from conventional static CNN models [3], [14], [15]. Third, binary and symbolic design enables energy-efficient inference suitable for edge deployment in resource-constrained clinical environments [9], [10], [13], [20]. Finally, the architecture supports diverse imaging modalities—grayscale (MRI, CT) and color-rich domains (histopathology, dermatology)—overcoming the narrow specialization of current models [1], [3], [8].

Future Directions: Symbolic Feedback Integration: Developing feedback mechanisms where ASP-derived symbolic rules dynamically influence BD-CeNN reservoir state evolution. This bi-directional coupling would strengthen semantic alignment between logical inference and symbolic feature dynamics, enabling adaptive, context-sensitive reasoning [14], [15].

Multimodal Data Fusion: Extending the framework to incorporate non-visual clinical data, including structured EHR entries, sensor data, and physician annotations alongside image inputs. Multimodal fusion significantly enhances contextual richness and diagnostic accuracy in healthcare AI systems [3], [7], [28].

Clinical Validation Studies: Engaging in clinician-in-the-loop evaluations, usability testing, and pilot deployments to ensure real-world utility. These efforts align with a growing emphasis on human-centered, explainable AI design, ensuring systems are interpretable, trustworthy, and practically deployable [2], [6], [28].

This work contributes to developing transparent, resource-efficient, and semantically grounded AI systems for medical imaging. By uniting BD-CeNN symbolic encoding, reservoir computing, and ASP logical inference, it advances theoretical foundations and practical applicability of explainable neuro-symbolic AI in clinical environments.

References

[1] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. Van Der Laak, B. Van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical image analysis*, vol. 42, pp. 60–88, 2017.

[2] E. Tjoa and C. Guan, "A survey on explainable artificial intelligence (xai): Toward medical xai," *IEEE transactions on neural networks and learning systems*, vol. 32, no. 11, pp. 4793–4813, 2020.

[3] A. S. Lundervold and A. Lundervold, "An overview of deep learning in medical imaging focusing on mri," *Zeitschrift fuer medizinische Physik*, vol. 29, no. 2, pp. 102–127, 2019.

[4] D. Gunning, M. Stefik, J. Choi, T. Miller, S. Stumpf, and G.-Z. Yang, "Xai—explainable artificial intelligence," *Science robotics*, vol. 4, no. 37, p. eaay7120, 2019.

[5] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Advances in neural information processing systems*, vol. 30, 2017.

[6] D. Dubois, H. Prade, and S. Schockaert, "Generalized possibilistic logic: foundations and applications to qualitative reasoning about uncertainty," *Artificial Intelligence*, vol. 252, pp. 139–174, 2017.

[7] A. Holzinger, A. Carrington, and H. Müller, "Measuring the quality of explanations: the system causability scale (scs) comparing human and machine explanations," *KI-Künstliche Intelligenz*, vol. 34, no. 2, pp. 193–198, 2020.

[8] E. J. Topol, "High-performance medicine: the convergence of human and artificial intelligence," *Nature medicine*, vol. 25, no. 1, pp. 44–56, 2019.

[9] X. Wang, Z. Tang, J. Guo, T. Meng, C. Wang, T. Wang, and W. Jia, "Empowering edge intelligence: A comprehensive survey on on-device ai models," *ACM Computing Surveys*, vol. 57, no. 9, pp. 1–39, 2025.

[10] S. Han, H. Mao, and W. J. Dally, "Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding," *arXiv preprint arXiv:1510.00149*, 2015.

[11] G. Manganaro, P. Arena, and L. Fortuna, *Cellular neural networks: chaos, complexity and VLSI processing*, vol. 1. Springer Science & Business Media, 2012.

[12] M. Milanova and U. Büker, "Object recognition in image sequences with cellular neural networks," *Neurocomputing*, vol. 31, no. 1-4, pp. 125–141, 2000.

[13] H. Li, X. Liao, C. Li, H. Huang, and C. Li, "Edge detection of noisy images based on cellular neural networks," *Communications in Nonlinear Science and Numerical Simulation*, vol. 16, no. 9, pp. 3746–3759, 2011.

[14] H. Jaeger and H. Haas, "Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication," *science*, vol. 304, no. 5667, pp. 78–80, 2004.

[15] M. Lukoševičius and H. Jaeger, "Reservoir computing approaches to recurrent neural network training," *Computer science review*, vol. 3, no. 3, pp. 127–149, 2009.

[16] M. Gelfond and V. Lifschitz, "The stable semantics for logic programs," in *Proceedings of the 5th International Conference on Logic Programming*, pp. 1070–1080, 1988.

[17] G. Brewka, T. Eiter, and M. Truszczyński, "Answer set programming at a glance," *Communications of the ACM*, vol. 54, no. 12, pp. 92–103, 2011.

[18] R. Kaminski and T. Schaub, "On the foundations of grounding in answer set programming," *Theory and Practice of Logic Programming*, vol. 23, no. 6, pp. 1138–1197, 2023.

[19] A. Holzinger, G. Langs, H. Denk, K. Zatloukal, and H. Müller, "Causability and explainability of artificial intelligence in medicine," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 9, no. 4, p. e1312, 2019.

[20] V. Hayyolalam, M. Aloqaily, Ö. Özkasap, and M. Guizani, "Edge intelligence for empowering iot-based healthcare systems," *IEEE Wireless Communications*, vol. 28, no. 3, pp. 6–14, 2021.

[21] J. C. Chedjou and K. Kyamakya, "A universal concept based on cellular neural networks for ultrafast and flexible solving of differential equations," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 4, pp. 749–762, 2014.

[22] J. Chedjou and K. Kyamakya, "Cnn-based ultrafast solver of stiff odes and pdes for enabling realtime computational engineering," *COMPEL-The international journal for computation and mathematics in electrical and electronic engineering*, vol. 30, no. 4, pp. 1333–1349, 2011.

[23] J. C. Chedjou and K. Kyamakya, "Cellular neural networks based local traffic signals control at a junction/intersection," *IFAC Proceedings Volumes*, vol. 45, no. 4, pp. 80–85, 2012.

[24] J. C. Chedjou and K. Kyamakya, "Cellular neural networks based time-series approximation for real time systems' modeling-and-identification and behavior forecast in transportation: motivation, problem formulation, and some research avenues," in *Autonomous Systems: Developments and Trends*, pp. 241–254, Springer, 2012.

[25] F. Al Machot, M. Ali, A. Haj Mosa, C. Schwarzlmüller, M. Gutmann, and K. Kyamakya, "Real-time raindrop detection based on cellular neural networks for adas," *Journal of Real-Time Image Processing*, vol. 16, pp. 931–943, 2019.

[26] Y. He, Z. Wang, Y. Zhang, T. Dan, T. Chen, G. Wu, and A. Li, "Neurosymad: A neuro-symbolic framework for interpretable alzheimer's disease diagnosis," *arXiv preprint arXiv:2503.00510*, 2025.

[27] T. R. Besold, S. Bader, H. Bowman, P. Domingos, P. Hitzler, K.-U. Kühnberger, L. C. Lamb, P. M. V. Lima, L. de Penning, G. Pinkas, *et al.*, "Neural-symbolic learning and reasoning: A survey and interpretation 1," in *Neuro-symbolic artificial intelligence: The state of the art*, pp. 1–51, IOS press, 2021.

[28] I. D. Mienye, G. Obaido, N. Jere, E. Mienye, K. Aruleba, I. D. Emmanuel, and B. Ogbuokiri, "A survey of explainable artificial intelligence in healthcare: Concepts, applications, and challenges," *Informatics in Medicine Unlocked*, p. 101587, 2024.