CTU Journal of Innovation and Sustainable Development
ISSN 2588-1418 | e-ISSN 2815-6412

DOI:10.22144/ctujoisd.2026.064

GFlowRule: Detailed-balance GFlowNets for diverse symbolic rule selection in neuro-symbolic image classification

Ma Truong Thanh[1], Thai Phu An[1], Dang Duy Khoa[1], Huynh Long[3], Dang Dai Tho[2], Pham Minh Tien[1], Nguyen Huu Hoa[1*], and Phan Ngoc Tuyen[4]

[1]*Department of Computer Science, Can Tho University, Viet Nam*

[2]*Viet Nam-Korea University of Information and Communications Technology, The University of Danang, Viet Nam*

[3]*Viet Nam Posts and Telecommunications Group, Viet Nam*

[4]*Faculty of Information Technology, Ca Mau College, Viet Nam*

**Corresponding author (nhhoa@ctu.edu.vn)*

Article info.

Received 14 Jun 2026
Revised 8 Sep 2026
Accepted 8 Oct 2026

Keywords

Convolutional neural network, cultural heritage image classification, detailed balance, GFlownet, neuro-symbolic learning, symbolic rule extraction

ABSTRACT

We propose a neuro-symbolic image-classification framework that combines CNN representations, Random-Forest-derived symbolic rules, and a Detailed-Balance GFlowNet for diverse rule-set selection. The GFlowNet optimizes a multi-component reward based on macro accuracy, correct and incorrect coverage, and inter-rule conflict, while the selected rules regularize CNN fine-tuning through either a fixed rule set or stochastic marginalization over $K = 32$ sampled rule sets. On a 12 class Vietnamese cultural-heritage dataset with 10,674 images, four backbones, and five random seeds, GFlowNet consistently obtains the highest matched-budget rule-set reward compared with Top-K Confidence, Greedy Coverage, and Random selection. Rule regularization generally matches or improves the CNN baseline, achieving its largest Macro-F1 gain of +0.73 percentage points on MobileNetV3-S. The learned policy also yields highly diverse rule sets, with a 100% unique-sample ratio over 5,000 draws. Overall, the results indicate that GFlowNet is effective at learning a diverse distribution of high-reward symbolic rule sets, while the downstream benefit of rule regularization remains dependent on the CNN backbone and regularization strategy.

1. INTRODUCTION

Automated recognition of cultural heritage imagery must satisfy two requirements at once: predictions need to be accurate, and the reasoning behind them needs to be auditable. Because misclassifications in this domain carry cultural and pedagogical weight, domain experts typically expect a system to justify a label rather than return it as an opaque output. Vietnamese intangible cultural heritage, spanning traditional costumes, ceremonial practices, and regional festivals, is a demanding benchmark: many categories share overlapping color palettes, ceremonial objects, and spatial layouts, so the discriminative signal often lies in subtle, localized differences rather than in global appearance.

Convolutional architectures perform strongly on visual recognition tasks (Simonyan & Zisserman, 2015), yet end-to-end training provides no mechanism for inspecting which visual cues drive a prediction, leaving it free to exploit incidental

correlations that can degrade robustness under distribution shift. Neuro-symbolic methods narrow this gap by pairing a neural feature extractor with an explicit symbolic layer, typically a set of human-readable rules distilled from the network's representation (Seo et al., 2021; Mao et al., 2025). Such rules follow naturally from a Random Forest (Breiman, 2001) fitted on CNN features, since every root-to-leaf path in a decision tree represents a conjunction of threshold conditions. In practice, this yields far more candidate rules than are useful: most are redundant, narrowly tied to a single majority class, or individually confident yet jointly conflicting. Reducing the pool with fixed support and confidence thresholds discards weak rules one at a time, without reasoning about the joint quality, coverage, or internal consistency of the surviving subset – the central difficulty this paper addresses: rule selection.

We treat rule selection as a sequential decision problem and solve it with a Generative Flow Network (Bengio et al., 2023), a probabilistic framework designed to sample structured objects in proportion to a reward rather than converge only to a single high-reward solution – useful when several rule subsets have comparable quality, since the objective is to model a distribution over solutions rather than identify one optimum. We train this policy with the Detailed Balance (DB) objective (Bengio et al., 2023), which imposes local flow-consistency constraints on transitions along selection trajectories. Beyond selecting a single rule set, we investigate whether conditioning CNN fine-tuning on multiple rule sets sampled from the learned GFlowNet distribution provides a different regularization signal from using a single elite solution. We therefore compare a static regularizer built from the best rule set found during training with stochastic rule-set marginalization, which resamples rule sets from the frozen GFlowNet policy at every optimization step. Finally, we package the trained pipeline into an interactive demonstrator so that the resulting symbolic evidence can be inspected directly by domain experts rather than only by the model's authors.

This paper makes four main contributions: (1) we introduce GFlowRule, a four-stage neuro-symbolic framework that combines CNN feature extraction, Random-Forest-derived symbolic rule generation, Detailed-Balance GFlowNet rule-set selection, and differentiable rule-regularized CNN fine-tuning; (2) we formulate rule-set selection with a multi-component reward that jointly accounts for macro-accuracy, correct and wrong coverage, and confidence-weighted inter-rule conflict, enabling the GFlowNet to model a distribution over high-quality rule subsets rather than a single deterministic solution; (3) we propose stochastic rule-set marginalization for symbolic CNN regularization, using Monte Carlo sampling from a frozen GFlowNet policy to expose the CNN to multiple rule subsets during fine-tuning; and (4) we develop an interactive rule-evidence explorer and conduct a multi-backbone evaluation on VN-ICH to examine rule-set quality, sampling diversity, computational cost, and downstream classification behavior.

2. BACKGROUND

2.1. GFlowNets and the Detailed Balance Objective

A Generative Flow Network (GFlowNet) is a probabilistic model that learns a stochastic policy for constructing objects sequentially, such that the marginal probability of sampling a complete object x is proportional to a non-negative reward $R(x)$ (Bengio et al., 2023).

Definition 1 (GFlowNet). Let $G = (S, A)$ be a directed acyclic graph whose nodes are states $s \in S$ and whose edges (s → s′) represent actions $a \in A$. A GFlowNet specifies a forward policy $P_F(s' \vee s)$ and a backward policy $P_B(s \vee s')$. A trajectory $\tau = (s_0, \dots, s_n)$ begins at a fixed initial state s_0 and terminates at $s_n \in S_T \subseteq S$, with the marginal probability of terminating at sn required to satisfy

$$P(\tau \ni s_n) \propto R(s_n) (1)$$

where R : ST → R≥0 is defined on the terminal states.

The Detailed Balance (DB) objective enforces Eq. (1) by requiring that, for every transition (s → s′), the flow entering s' from s equal the flow leaving s' back toward s:

$$F(s)P_F(s' \vee s) = F(s')P_B(s \vee s') (2)$$

where $F: S \to R_0$ is a state-flow function with boundary condition $F(x) = R(x)$ for all terminal x. In log-space, the log-flow residual of a transition is

$$\Delta_{DB}(s, s') = logF(s) + logP_F(s' \vee s) - logF(s') - logP_B(s \vee s') (3)$$

and the DB loss is the mean squared residual over sampled transitions, $L_{DB} = E_{(s,s')\sim P_{traj}}$

When $L_{DB} = 0$, Eq. (2) holds on every transition, which implies flow matching and, by a standard

flow-decomposition argument, the reward proportionality $P_F(x) = \frac{R(x)}{Z}$ for every terminal state x (Bengio et al., 2023). Consequently, any two terminal states with equal reward are sampled with equal probability, so a policy trained to $L_{DB} \approx 0$ does not collapse to a single mode when several rule subsets attain comparable reward, the property that motivates using a GFlowNet for rule selection, since class diversity among selected rules is an implicit requirement of the downstream regularizer.

In our implementation, $F(s)$, P_F, and P_B are each parameterized by a separate MLP with layer normalization, and the DB objective is provided by the torchgfn library's DBGFlowNet class, which evaluates L_{DB} over sampled trajectories.

2.2. Random Forest and Symbolic Rule Extraction

A Random Forest (Breiman, 2001) is an ensemble of T decision trees fitted on bootstrap samples of the training data with a random feature subset at every split. Let $h \in R^d$ be a CNN feature vector; each root-to-leaf path corresponds to a conjunction of axis-aligned threshold conditions (Breiman, 2001).

Definition 2 (Symbolic Rule). A symbolic rule ρ is a pair (C_ρ, y_ρ), where $C_\rho = \{c_1, \dots, c_k\}$ is a set of conditions and $y_\rho \in Y$ is the predicted class. Each condition $c_i = (j_i, \diamond_i, \theta_i)$ specifies a feature index j_i, a comparison operator $\diamond_i \in \{\leq, >\}$, and a threshold θ_i. A feature vector h satisfies ρ *if* $h_{j_i} \diamond_i \theta_i$ holds for every $c_i \in C_\rho$. The confidence of ρ on a dataset D is

$$conf(\rho, D) = \frac{|\{h \in D : hsatisfies\rho, y = y_\rho\}|}{|\{h \in D : hsatisfies\rho\}|} (4)$$

Enumerating every root-to-leaf path across all trees yields the raw rule set R_{raw}; rules whose validation-set support or confidence fall below fixed thresholds are discarded, leaving $R_{valid} \subseteq R_{raw}$.

2.3. Soft Rule Activation and Neuro-Symbolic Integration

To evaluate rules differentially on continuous features during CNN training, we use a sigmoid-based soft match: for $c = (j, \leq, \theta)$, $b_{\leq}(h_j, \theta) = \sigma\left(\tau(\theta - h_j)\right)$, and for $c = (j, >, \theta)$, $b\ (h_j, \theta) = \sigma\left(\tau(h_j - \theta)\right)$, where σ is the logistic sigmoid and τ > 0 a temperature that sharpens b toward a hard indicator as $\tau \to \infty$. For a rule ρ with k conditions, the degree of satisfaction is the soft conjunction

$$\phi_\rho(h) = \prod_{i=1}^{k} b_{\diamond_i}(h_{j_i}, \theta_i) (5)$$

Let $f_\theta = g_\theta \circ \psi$ be the CNN, with penultimate feature $h \in R^m$ and logits $z \in R^{Y\vee}$. Given a selected rule set R_{sel}, the rule-consistency penalty applies a confidence-weighted, label-smoothed cross-entropy at each rule's soft activation, averaged over the rule set:

$$L_{rule}(h, z) = \frac{1}{R_{sel} \vee \sum_{\rho \in R_{sel}} w_\rho \cdot \phi_\rho(h) \cdot l_{LS}(z, y_\rho)} (6)$$

where $w_\rho = conf(\rho)$ and l_{LS} is the label-smoothed negative log-likelihood with smoothing ε. Since ϕ_ρ and l_{LS} are both smooth, L_{rule} is differentiable in h wherever z is. The overall training objective is

$$L_{total} = L_{CE}(z, y) + \lambda L_{rule}(h, z) (7)$$

with λ > 0 controlling the strength of the symbolic constraint.

3. METHODOLOGY

The framework has four stages, summarized in Figure 1: CNN feature extraction (Stage 1), Random Forest rule generation and validation (Stage 2), GFlowNet-based rule selection under the DB objective (Stage 3), and rule-regularized CNN fine-tuning (Stage 4).

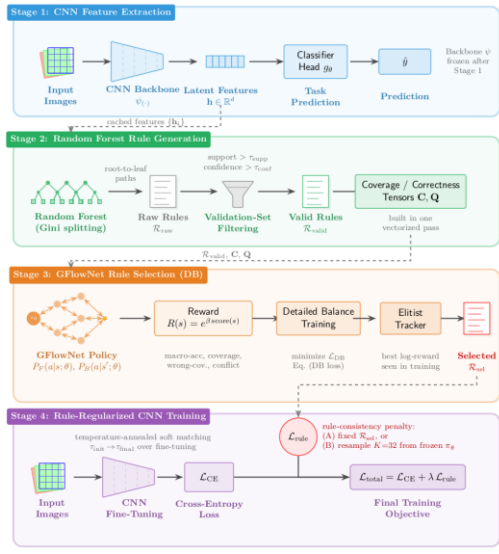


Figure 1. Overview of the proposed neuro-symbolic pipeline

Stage 1: CNN feature extraction. Each backbone is initialized from ImageNet pre-training and used as $f_\theta = g_\theta \circ \psi$, where ψ is the backbone and g_θ a linear classification head with |Y| = 12 outputs. Four backbones are evaluated: MobileNet V3 Small (d = 960) (Howard et al., 2019), AlexNet (d = 4,096) (Krizhevsky et al., 2012), ViT-B/32 (d = 768) (Dosovitskiy et al., 2021), and Swin-T (d = 768) (Liu et al., 2021). Each backbone is fine-tuned end-to-end by minimizing cross-entropy with Adam (Kingma & Ba, 2015) and differential learning rates for the backbone and head, using random resized cropping, horizontal flipping, and rotation, with early stopping (patience = 5) on validation loss. The backbone is then frozen and features $\{h_i\}$ are cached for the later stages.

Stage 2: Random Forest rule generation. A Random Forest with Gini-impurity splitting is fitted on the frozen training features. All root-to-leaf paths across every tree are enumerated into Rraw, with confidence computed via Eq. (4). Each candidate is then evaluated on the held-out validation set in a single vectorized pass: a rule is retained in Rvalid only if its fractional support exceeds 1% and its confidence exceeds 70%. This pass simultaneously builds a binary coverage tensor C and correctness tensor Q over the validation set, used throughout Stage 3.

Stage 3: GFlowNet rule selection. Selection is formulated as a GFlowNet over $s \in \{0,1\}^{R_{valid}\vee}$, a binary vector of chosen rules, starting from $s_0 = 0$. Each non-terminal action toggles one unselected rule on; a stop action terminates the trajectory. Forward masks block re-selection or exceeding the budget $K_{max} = 120$; backward masks allow removing any selected rule. P_F, P_B, F are each a two-hidden-layer MLP ($H = 256$ with layer normalization, trained with the DB objective via torchgfn's DBGFlowNet (Adam, η = 10−3, gradient clipping at norm 5.0, ReduceLROnPlateau on the validation DB residual).

The reward for a terminal state (selected rule set) is

$$R(s) = exp\big(\beta \cdot score(s)\big), \beta = 3.0 \quad (8)$$

with the four-component score

$$score(s) = w_{acc}A(s) + w_{cov}C^{+(s)} - w_{wrong}C^{-(s)} - w_{conflict}\gamma(s) \quad (9)$$

computed entirely from C, Q, with weights $w_{acc} = 1.0$, $w_{cov} = 0.5$, $w_{wrong} = 0.75$, $w_{conflict} = 0.1$. The reward weights were fixed across all backbones to preserve a common evaluation criterion. The macro-accuracy term receives the largest positive weight because predictive correctness is the primary objective. Correct coverage is assigned a positive but smaller weight to encourage broad sample coverage, while wrong coverage receives a stronger penalty to discourage rules that cover samples but frequently support incorrect predictions. The conflict term has a smaller weight because it acts as a secondary consistency regularizer rather than the primary measure of rule-set quality. Keeping these weights fixed across backbones also prevents backbone-specific reward tuning from introducing an additional source of variation into the comparison. These fixed weights define the designed rule-selection objective and may affect the resulting rule distribution. Macro-accuracy $A(s)$ is the accuracy of the rule set's confidence-weighted vote classifier, averaged per class; correct coverage $C^{+(s)}$ and wrong coverage $C^{-(s)}$ are the mean per-class fractions of samples covered and correctly / incorrectly predicted; and the conflict ratio is

$$\gamma(s) = \frac{1}{N_{val}} \sum_{j=1}^{N_{val}} \frac{V_j^{total} - V_j^{win}}{V_j^{total} + \varepsilon} \quad (10)$$

where V_j^{total} and V_j^{win} are the total and winning-class vote counts on sample j. The exponential in Eq. (8) maps the (possibly negative) score onto $R>0$, as GFlowNets require. An elitist tracker retains the highest-log-reward rule set observed across all sampled trajectories during training. After training, we restore the checkpoint with the best exponential moving average of the negative DB residual, draw 20 additional trajectories, and retain the better of the elitist record and these samples as R_{sel}. Two checkpoints are saved: the highest-EMA-reward converged checkpoint, used for the static penalty, and a diverse checkpoint satisfying the minimum diversity criterion, used as the frozen sampler for stochastic rule-set marginalization.

Stage 4: Rule-regularized CNN training. Fine-tuning resumes from the Stage 1 checkpoint under a rule-consistency regularizer, in two variants sharing the same training loop. Variant A (static) fixes R_{sel} from Stage 3 and applies Eq. (6) directly. Variant B (stochastic rule-set marginalization). Instead of conditioning on a single selected rule set, Variant B estimates the expected rule-consistency penalty under the frozen GFlowNet distribution,

$$L_{marg} = E_{s\sim\pi_\theta}[L_{rule}(H,Z;s)] \approx \frac{1}{K}\sum_{k=1}^{K} L_{rule}(H,Z;s_k) \quad (11)$$

where $K = 32$ rule sets are independently resampled at every optimization step from the frozen GFlowNet policy π_θ (requires_grad = False, evaluation mode), so gradients are propagated only through the CNN. By linearity of expectation, the Monte Carlo average is an unbiased estimator of the expected rule-consistency penalty under π_θ.

For stochastic rule-set marginalization, we fixed $K = 32$ samples per optimization step as a computational compromise between Monte Carlo estimation quality and training cost. Increasing K reduces the variance of the empirical estimate of the expected rule-consistency penalty, but requires additional rule-set sampling and penalty evaluation at each optimization step. In our implementation, the stochastic variant increases Stage 4 training time from approximately 25.5 min for the static variant to 36.6 min per run (Table 5). We therefore use $K = 32$ consistently across backbones rather than tuning K independently for each model. A systematic sensitivity analysis over K is left for future work.

For both variants, rule conditions are pre-compiled into fixed GPU buffers, allowing rule activations for a mini-batch to be evaluated using vectorized operations. The sigmoid temperature τ is held at $\tau_{init} = 2.0$ during a 2-epoch warm-up and is then geometrically annealed to $\tau_{final} = 12.0$ over the following 10 epochs. The complete temperature schedule therefore spans the first 12 epochs of Stage 4. The combined objective in Eq. (7), with $\lambda = 0.1$ and label smoothing $\varepsilon = 0.1$, is optimized using the same optimizer configuration as in Stage 1. The rule-regularization weight was fixed at $\lambda = 0.1$ across all backbones and selection strategies to maintain a common regularization scale and avoid backbone-specific tuning that could confound comparisons between rule-selection methods. The coefficient was used as a moderate penalty strength so that the symbolic term complements rather than dominates the supervised cross-entropy objective. We intentionally kept λ fixed across all experiments so that differences in downstream performance were not attributable to independently tuned regularization strengths. A systematic sensitivity analysis over λ is left for future work. Early stopping with patience 5 is applied based on validation accuracy.

4. RULE-SELECTION BASELINES

To isolate the GFlowNet's contribution, we compare it against three non-learned baselines on the same validated pool R_{valid}, scored with the same reward criterion (Eq. (9)) that the GFlowNet was trained to maximize. Since the GFlowNet decides autonomously how many rules to select (up to K_{max}), each heuristic is also run at the matched budget $k = |R_{sel}|$ equal to the count the GFlowNet chose in that run, our primary comparison, which controls for rule-set size.

Top-K Confidence ranks rules by validation-set confidence (Eq. (4)) and deterministically keeps the top-k; it favors per-rule precision but ignores class diversity and inter-rule conflict. Greedy Coverage adds rules one at a time to maximize newly covered validation samples,

$$\rho = argmax^{(t)}_{\rho\in R_{valid}\{R} \left|\{h_j \in S_t : h_j satisfies \rho\}\right| \quad (12)$$

where St is the set of samples not yet covered by the first t selected rules; this is computed with GPU-vectorized operations over C rather than a Python loop. Random Selection draws rules uniformly at random without replacement, with five seeds (0–4) to account for sampling variance.

5. EXPERIMENTAL RESULTS

5.1. Dataset and Experimental Setup

All experiments use VN-ICH, a Vietnamese Intangible Cultural Heritage image collection of 10,674 images across 12 classes. The dataset was compiled from two main image sources: (i) frames extracted from video recordings and (ii) photographs captured at real-world cultural festivals and events. During dataset preparation, basic image transformations provided by the employed image-processing/augmentation library were applied as standard preprocessing/augmentation operations. The resulting dataset was divided using stratified sampling into 7,466 training, 2,129 validation, and 1,079 test images, corresponding to a 70%/20%/10% split. (Table 1). All four backbones start from ImageNet pre-trained weights. Stage 1 fine-tuning runs up to 100 epochs with Adam (initial learning rate 10−4, batch size 64, early-stopping patience 5 on validation accuracy); Stage 4 restarts from the Stage 1 checkpoint with the same optimizer settings. Every experiment is repeated over five random seeds (42, 44, 46, 48, 50); we report means and standard deviations. All runs use an NVIDIA Tesla T4 GPU (16 GB) with 16-bit automatic mixed precision.

Table 1. Class-wise distribution of the VN-ICH dataset across the training, validation, and test splits

Class name	Train	Val	Test	Total
Cho Noi Cai Rang	650	185	94	929
Chua (Pagoda)	819	234	117	1170
Dieu Mua Truyen Thong	574	164	85	823
Don Ca Tai Tu	573	162	81	816
Dua Bo Bay Nui	581	166	84	831
Le Hoi Nghinh Ong	491	140	71	702
Nghe Dan Tre	651	186	94	931
Nghe Det Chieu	698	199	101	998
San Khau Du Ke Khmer	636	181	92	909
Ok Om Bok	620	177	89	886
Phong Tuc Cuoi	709	203	103	1015
Via Ba Chua Xu	464	132	68	664
Total	**7466**	**2129**	**1079**	**10674**

5.2. Rule-Set Quality: Comparison of Selection Strategies

Table 2 compares the four selection strategies on the validation-set reward score (Eq. (9)) at matched budget, averaged over five seeds. GFNet-DB attains the highest reward score on every backbone. Top-K Confidence scores lowest on three of the four backbones (MobileNetV3-S, Swin-T, ViT-B/32), because it concentrates on a narrow set of majority classes and leaves most classes with weak coverage ($C+ \approx 54\%–78\%$); AlexNet is the exception, where Greedy Coverage (1.161) instead falls below Top-K Confidence (1.208) because its weaker feature space causes max-coverage selection to pull in rules with high wrong-coverage (14.7%) and conflict (8.8%). Greedy Coverage and Random selection track GFlowNet more closely overall because both reach near-complete class coverage, but their shortfall comes from higher wrong-coverage and conflict. GFlowNet's advantage is that it balances all four reward components jointly rather than optimizing a single proxy in isolation.

Because the reward in Eq. (9) is the objective optimized by the GFlowNet, this experiment should be interpreted primarily as an assessment of whether the learned policy can identify high-scoring rule sets under the designed criterion, rather than as an independent measure of downstream generalization. The comparison remains informative because all heuristic baselines are evaluated under the same reward criterion and at the same matched rule-set budget, but it does not by itself establish a classification advantage.

5.3. Test Accuracy Comparison Across Backbones

Table 3 reports test-set accuracy and Macro-F1 for all six methods across four backbones (mean ± std over five seeds). The static methods apply the vectorized rule-consistency penalty using a rule set fixed after Stage 3, whereas GFNet-DB Marg. performs stochastic rule-set marginalization by resampling $K = 32$ rule sets per optimization step from the frozen GFlowNet policy.

Three observations stand out. First, stochastic rule-set marginalization matches or exceeds static GFNet-DB on three of the four backbones in Macro-F1, with gains of +0.60 pp on AlexNet and +0.14 pp on MobileNetV3-S relative to the static variant. The exception is ViT-B/32, where static GFNet-DB performs better (92.41% versus 92.01% Macro-F1). These results indicate that averaging the rule-consistency objective over multiple GFlowNet samples can benefit some backbone representations, but the effect is not universal. Second, MobileNetV3-S shows the most consistent lift: all five rule-regularized methods beat the CNN baseline on both metrics (ΔF1 from +0.34 to +0.73 pp), suggesting the rule penalty is a useful inductive bias when the baseline feature space is less saturated; on the strongest backbone (Swin-T), by contrast, every method stays within one standard deviation of the near-ceiling baseline. Third, Top-K Confidence, despite the weakest rule-set quality score, stays close to baseline accuracy on most backbones (ViT-B/32, at −0.72 pp, is the exception), because the cross-entropy term can partly compensate for a suboptimal penalty at moderate penalty weight – a decoupling between rule-set quality and downstream accuracy that Section 7 returns to.

Table 2. Rule-set quality on VN-ICH

Backbone	Method	Score↑	Macro-Acc (%)↑	$C+$ (%)↑	$C-$ (%)↓	Conflict (%)↓
AlexNet	GFNet-DB	1.221	88.6	93.9	10.4	5.0
	Greedy Coverage	1.161	85.3	100	14.7	8.8
	Random	1.165	86.0	93.1	12.5	5.0
	Top-*K* Conf.	1.208	92.9	58.2	0.9	0.1
MobileNet V3-S	GFNet-DB	1.303	92.3	94.6	7.0	4.7
	Greedy Coverage	1.268	90.0	100	10.0	7.2
	Random	1.249	98.8	93.9	9.1	4.7
	Top-*K* Conf.	1.184	91.6	56.4	1.1	0.1
Swin-T	GFNet-DB	1.419	96.8	98.7	3.2	2.5
	Greedy Coverage	1.394	95.4	100	4.6	3.2
	Random	1.393	95.6	98.9	4.4	2.4
	Top-*K* Conf.	1.325	93.3	78.4	0.0	0.0
ViT-B/32	GFNet-DB	1.344	93.7	96.5	5.9	2.3
	Greedy Coverage	1.305	91.5	100	8.5	4.0
	Random	1.316	92.5	96.6	7.2	2.5
	Top-*K* Conf.	1.018	74.9	54.1	0.1	0.0

Overall, these results indicate that rule-set quality and downstream classification performance should be treated as related but distinct evaluation dimensions. A higher validation reward does not necessarily imply a larger test-set accuracy gain, because the rule-consistency penalty interacts with the learned feature representation and the supervised cross-entropy objective. We therefore interpret the GFlowNet results primarily in terms of rule-set selection quality and diversity, while treating downstream classification improvements as backbone- and regularization-dependent empirical outcomes.

5.4. Test Accuracy Comparison Across Backbones

Table 3. Test-set performance on VN-ICH (%)

Backbone	Method	Acc (%)	Macro-F1 (%)	ΔAcc (pp)	ΔF1 (pp)
AlexNet	CNN Baseline	90.97±0.39	90.83±0.41	–	–
	GFNet-DB (static)	90.86±0.94	90.65±0.97	−0.11	−0.18
	GFNet-DB Marg.	**91.41±0.58**	**91.25±0.62**	+0.44	+0.42
	Greedy Coverage	90.71±1.00	90.53±0.99	−0.26	−0.30
	Random	90.92±0.86	90.76±0.88	−0.05	−0.07
	Top-*K* Confidence	90.99±0.72	90.79±0.76	+0.02	−0.04
MobileNet V3-S	CNN Baseline	94.29±0.59	94.20±0.63	–	–
	GFNet-DB (static)	94.85±0.28	94.79±0.30	+0.56	+0.59
	GFNet-DB Marg.	**95.00±0.39**	**94.93±0.42**	+0.71	+0.73
	Greedy Coverage	94.64±0.33	94.58±0.37	+0.35	+0.38
	Random	94.62±0.43	94.54±0.46	+0.33	+0.34
	Top-*K* Confidence	94.77±0.48	94.69±0.49	+0.48	+0.49
Swin-T	CNN Baseline	96.72±0.37	96.74±0.40	–	–
	GFNet-DB (static)	96.81±0.31	96.82±0.31	+0.09	+0.08
	GFNet-DB Marg.	96.79±0.05	96.82±0.08	+0.07	+0.08
	Greedy Coverage	**96.83±0.34**	**96.83±0.34**	+0.11	+0.09
	Random	96.64±0.48	96.65±0.50	−0.08	−0.09
	Top-*K* Confidence	96.62±0.31	96.61±0.30	−0.10	−0.13
ViT-B/32	CNN Baseline	92.25±0.54	92.16±0.55	–	–
	GFNet-DB (static)	**92.58±0.56**	**92.41±0.59**	+0.33	+0.25
	GFNet-DB Marg.	92.16±0.86	92.01±0.82	−0.09	−0.15
	Greedy Coverage	92.35±0.72	92.21±0.71	+0.10	+0.05
	Random	92.25±0.76	92.08±0.78	±0.00	−0.08
	Top-*K* Confidence	91.53±0.47	91.44±0.40	−0.72	−0.72

To assess whether the learned GFlowNet policy maintains diversity rather than collapsing to a small number of dominant solutions, we draw five independent batches of 1,000 rule sets from the converged checkpoint, yielding 5,000 samples per backbone. We compute diversity metrics for each seed; Table 4 reports the results for seed 42, while the remaining seeds show the same qualitative pattern. The unique-sample ratio is 100% for all four backbones, meaning that all 5,000 sampled rule sets are distinct. The empirical distribution also exhibits near-maximal normalized entropy, with a top-1 share of only 0.1%. These results indicate that the learned policy does not collapse to a small number of dominant rule sets and provides diverse samples for estimating the expected rule-consistency objective.

Unlike the reward comparison, this analysis evaluates a different property of the learned policy: whether probability mass concentrates on a small number of rule sets. The 100% unique-sample ratio, near-maximal normalized entropy, and 0.1% top-1 share provide evidence that the learned policy retains substantial diversity across sampled solutions rather than collapsing to a small set of dominant rule subsets.

5.5. Computational Overhead

Table 5 reports representative wall-clock training times for the main pipeline stages on VN-ICH, averaged over five seeds. GFlowNet training is the dominant additional cost, requiring approximately 53 min per run, whereas feature extraction and rule generation together require less than 2 min. Static GFNet-DB fine-tuning takes approximately 25.5 min, while stochastic rule-set marginalization requires approximately 36.6 min. The additional cost arises from repeatedly sampling and aggregating multiple rule sets at each optimization step, although rule-penalty evaluation itself is vectorized.

Table 4. GFlowNet policy quality for the converged checkpoint at seed 42

Backbone	Log-reward	Unique (%)	Entropy (%)	Top-1 share (%)
Alexnet	3.556	100.0	≈100.0	0.1
MobileNetV3-S	3.706	100.0	≈100.0	0.1
Swin-T	4.168	100.0	≈100.0	0.1
ViT-B/32	3.905	100.0	≈100.0	0.1

Table 5. Representative wall-clock training time per pipeline stage on VN-ICH, averaged over 5 seeds

Stage	Component	Time (min)
Stage 1	CNN baseline training	33.1 ± 4.7
Stage 2	Feature extraction	1.1 ± 0.1
Stage 2	Rule extraction (RF + filter)	0.5 ± 0.0
Stage 3	GFlowNet training	53.0 ± 2.2
Stage 4	GFNet-DB fine-tuning (static)	25.5 ± 1.6
Stage 4	GFNet-DB Marg. fine-tuning	36.6 ± 5.0
Stage 4	Greedy/Random/TopK + fine-tune	26.2 ± 2.1
Total (GFNet-DB static path)		≈113 min

6. APPLICATION

The reward, coverage, and entropy statistics of Sections 5.2–5.4 establish that the selected rule sets are accurate and diverse in aggregate, but a curator deciding whether to trust a single prediction cannot inspect a validation-set average. To make the symbolic account produced by the pipeline usable by non-technical stakeholders, we built a browser-based demonstrator on top of the trained pipeline, shown in Figure 2 and Figure 3.



Figure 2. Landing view of the rule-evidence explorer

The tool accepts a single uploaded image (Fig. 2) and routes it through three parallel inference branches corresponding to the variants considered in Stage 4: (i) the unregularized CNN baseline; (ii) the static rule-regularized model of Variant A, explained using its fixed rule set Rsel; and (iii) a multi-rule-set branch that aggregates evidence from multiple rule sets sampled from the frozen GFlowNet policy, following the stochastic marginalization formulation in Eq. (11). Presenting the three branches side by side (Fig. 3, panels 01–03) allows a curator to compare the baseline prediction with the corresponding static and distributional rule-based evidence.

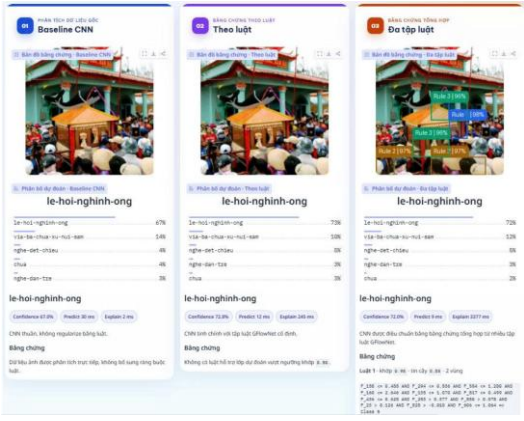


Figure 3. Side-by-side evidence panels for a single test image

7. DISCUSSION

The results reveal a systematic distinction between rule-set quality and downstream classification performance. GFlowNet consistently obtains the highest score under the designed rule-selection reward, but this advantage does not translate proportionally into test-set accuracy gains. This is expected because the reward evaluates the quality of symbolic rule subsets on the validation set, whereas downstream classification depends on how the resulting symbolic penalty interacts with the CNN representation and the supervised cross-entropy objective. The reward gap is also backbone-dependent, being widest for ViT-B/32 (1.344 versus 1.018 for Top-K Confidence) and narrower for AlexNet, suggesting that the effectiveness of axis-aligned symbolic rules depends on the geometry of the learned representation.

The stochastic-marginalization variant provides a more nuanced result. It matches or exceeds static GFNet-DB on three of the four backbones, but underperforms the static variant on ViT-B/32, indicating that averaging over the learned rule-set distribution is not universally beneficial. At the same time, the policy-quality analysis shows that the GFlowNet does not collapse to a small number of dominant solutions, supporting its intended role as a sampler over multiple competitive rule subsets rather than merely an optimizer for one elite solution. Taken together, the results suggest that the primary empirical benefit of GFlowRule lies in improving rule-set quality and diversity, whereas downstream classification gains depend on the backbone and its interaction with the symbolic regularizer.

Limitations. First, although the extracted rules are explicit and directly evaluable, their conditions are defined over latent feature dimensions (e.g., $h312 \geq 0.41$) rather than named visual concepts. They should therefore be interpreted as model-level symbolic rules rather than fully semantic human-interpretable concepts. Second, GFlowNet training adds approximately 53 min of overhead per run, and the multi-rule-set explanation itself costs seconds per image (Section 6), which may be prohibitive for rapid prototyping or high-throughput deployment. Third, the reward hyperparameters were fixed rather than tuned per backbone and therefore may not be optimal in every setting. Fourth, $\lambda = 0.1$ and $K = 32$ were fixed across experiments; although this provides a controlled comparison and avoids backbone-specific tuning, the absence of systematic sensitivity analyses leaves the effect of these hyperparameters insufficiently characterized. Finally, the rule-selection reward is the objective optimized by the GFlowNet, so higher reward should not be interpreted as independent evidence of improved downstream generalization. An independent cross-dataset or out-of-distribution evaluation would provide stronger evidence of whether the learned rule distribution transfers beyond the validation-set objective.

8. CONCLUSION

We presented GFlowRule, a neuro-symbolic framework that combines CNN representations, Random-Forest-derived symbolic rules, Detailed-Balance GFlowNet selection, and differentiable rule regularization. Across four backbones and five random seeds on the 12-class VN-ICH dataset, GFlowNet selection consistently achieved the highest matched-budget rule-set reward among Top-K Confidence, Greedy Coverage, and Random selection. The learned policy also generated highly diverse rule sets, with a 100% unique-sample ratio

and near-maximal normalized entropy over 5,000 samples, supporting its role as a distributional selector rather than a single-solution optimizer. Downstream classification effects were more modest and backbone-dependent: rule regularization improved or matched the CNN baseline in several settings but did not consistently provide gains across all backbones or regularization variants. These results suggest that the principal benefit demonstrated in the current study is improved selection and diverse sampling of high-quality symbolic rule sets, while the downstream classification benefit depends on the interaction between the selected rules and the underlying CNN representation. Future work will investigate sensitivity to the reward and regularization hyperparameters, connect latent rule conditions to named visual concepts, reduce sampling overhead, and evaluate GFlowRule on broader and independently collected cultural-heritage datasets.

CONFLICT OF INTEREST

There are no competing interests.

REFERENCES

Bengio, Y., Lahlou, S., Deleu, T., Hu, E. J., Tiwari, M., & Bengio, E. (2023). GFlowNet foundations. *Journal of Machine Learning Research, 24*(210), 1–55. https://jmlr.org/papers/v24/22-0364.html

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2021). An image is worth 16×16 words: Transformers for image recognition at scale. *International Conference on Learning Representations*. https://doi.org/10.48550/arXiv.2010.11929

Howard, A., Sandler, M., Chu, G., Chen, L.-C., Chen, B., Tan, M., Wang, W., Zhu, Y., Pang, R., Vasudevan, V., Le, Q. V., & Adam, H. (2019). Searching for MobileNetV3. *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 1314–1324). https://doi.org/10.1109/ICCV.2019.00140

Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. *International Conference on Learning Representations*. https://doi.org/10.48550/arXiv.1412.6980

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems, 25*, 1097–1105.

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., & Guo, B. (2021). Swin Transformer: Hierarchical vision Transformer using shifted windows. *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 10012–10022). https://doi.org/10.1109/ICCV48922.2021.00986

Mao, J., Tenenbaum, J. B., & Wu, J. (2025). *Neuro-symbolic concepts* [Preprint]. arXiv. https://arxiv.org/abs/2505.06191.

Seo, S., Arık, S. Ö., Yoon, J., Zhang, X., Sohn, K., & Pfister, T. (2021). Controlling neural networks with rule representations. *Advances in Neural Information Processing Systems, 34,* 11196-11207. https://doi.org/10.48550/arXiv.2106.07804

Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *International Conference on Learning Representations*. https://doi.org/10.48550/arXiv.1409.1556