

Large-Scale Validation and Analysis of the Rejection Capacity in Entropic Associative Memory

Ángel Fernando Bórquez^a, Rafael Morales^b, Luis A. Pineda^c

^a Universidad de Sonora

^b Universidad de Guadalajara

^c Universidad Nacional Autónoma de México

Guadalajara, Jalisco, México

0000-0003-4347-6028

Abstract—The Entropic Associative Memory (EAM) is a cognitively inspired computational model that combines probabilistic recognition with an efficient pattern storage and retrieval architecture. Previous evaluations, limited to small domains of about ten classes and a few thousand instances per class, showed promising results in classification and reconstruction. This work extends those studies in two ways. First, we assess scalability using Google's *Quick, Draw!* dataset, which includes over 300 classes and 100,000 instances per class. Second, we test EAMs rejection capability by storing only half of the classes and evaluating performance on the full dataset. Results confirm EAMs adaptability to large-scale scenarios and its effective rejection of novel stimuli, underscoring its potential as a robust and explainable AI model.

Index Terms—Entropic Associative Memory, Quick Draw, scalability, open-set recognition, rejection.

I. INTRODUCTION

Research on memory models inspired by human cognition has grown rapidly within artificial intelligence (AI) [1]. These approaches aim to computationally reproduce the capacity of biological systems to store, associate, and retrieve past experiences. In this context, the Entropic Associative Memory (EAM), proposed by Pineda and Morales [2], constitutes a theoretical advance by offering a declarative, distributed, and efficient architecture for probabilistic recognition and flexible pattern retrieval.

Early studies with EAM focused on small datasets such as Fashion MNIST [3], where the model stored latent representations from a convolutional autoencoder and successfully retrieved them under noise or partial input. These results validated its effectiveness in closed-set recognition, where all test classes had been previously stored in memory. However, the original framework had three main limitations: (1) evaluations were restricted to small domains, hindering scalability analysis; (2) the model's ability to reject non-stored classes crucial in open-set recognition [4] was untested; and (3) the experimental setup was tightly coupled to fixed neural-network configurations.

This work addresses these limitations by (1) evaluating EAMs scalability on the large-scale Quick, Draw! dataset [5],

which includes over 300 classes and hundreds of thousands of examples per category, and (2) designing a rejection experiment to assess its performance with non-stored classes. These experiments extend our understanding of EAMs behavior in complex scenarios, providing empirical evidence of its robustness, generalization capacity, and potential as an alternative paradigm in explainable AI.

II. BACKGROUND

The Entropic Associative Memory (EAM), proposed by Pineda and Morales [2], is a cognitively inspired computational model designed to probabilistically and distributively store and retrieve patterns. Unlike conventional systems, EAM employs a single Associative Memory Register that concentrates all learned object representations in a matrix representation, enabling flexible recognition and retrieval while allowing the study of emergent behaviors such as generalization and synthetic image generation.

The model operates through three core mechanisms:

- **λ -register**: records patterns by reinforcing feature associations.
- **η -recognition**: determines whether an input stimulus matches a stored pattern, using tolerance and confidence thresholds.
- **β -retrieval**: reconstructs complete patterns from partial or noisy cues through probabilistic inference.

A distinctive feature of EAM is the transparency of its memory representation. Because all associations are explicitly stored in the memory matrix, its state can be directly inspected and visualized, offering clear insight into how features interact. This transparency allows tracing recognition or retrieval outcomes to their underlying associations, contrasting with the black-box nature of deep learning models and supporting EAMs potential as an explainable AI approach.

Previous experiments [2], [6] evaluated EAM using Fashion MNIST [3] and CIFAR-10 [7], where latent representations from a convolutional autoencoder were successfully recognized and reconstructed despite noise or incompleteness. Emergent phenomena such as association-chains transitions between stored representations were also observed. A later study employed the Quick, Draw! dataset with 1013 classes, obtaining promising results [8].

The work was conducted during the first author's summer stay at the University of Guadalajara, supported by a scholarship from the Programa Delfin.

Nonetheless, earlier frameworks showed key limitations: (1) experiments were confined to 1013 classes, preventing scalability analysis; (2) the ability to reject unseen patterns, essential for open-set recognition, was not tested; and (3) the design was tightly coupled to a fixed number of neural network classes, reducing flexibility.

These constraints motivate the present study, which aims to evaluate EAM in a more demanding context using Google's Quick, Draw! dataset [5], decouple the number of classes used in network training and memory loading, and analyze the models rejection performance for unseen classes.

III. METHODOLOGY

To extend the validation of the EAM model and analyze its behavior in more complex scenarios, an experimental methodology was designed comprising three phases: dataset migration and large-scale validation, rejection testing, and system refactoring for flexibility.

A. Dataset Migration and Large-Scale Validation

The first phase involved integrating the Quick, Draw! dataset [5], which required:

- **Data module adaptation:** modification of the data loader to handle an arbitrary subset of Quick, Draw! classes, ensuring balanced samples and variable class counts.
- **Neural network retraining:** the perceptual system—an autoencoder and a classifier based on [9]—was retrained from scratch using Quick, Draw! to obtain new latent feature representations.
- **Pipeline validation:** the full workflow (training, feature generation, memory storage, and recognition testing) was executed to confirm compatibility with the new domain.

B. Rejection Experiment Design

The second phase evaluated EAMs ability to reject unseen classes:

- **Partial memory loading:** although the perceptual system was trained with all classes, only half were stored in EAM.
- **Full-set evaluation:** the complete dataset, including non-stored classes, was used to assess explicit rejection behavior.
- **Metrics:** standard measures quantified the systems capacity to avoid false associations with unseen patterns.

C. Refactoring for Experimental Flexibility

The third phase improved modularity and scalability:

- **Class parameterization:** the number of classes became a runtime argument rather than a fixed value.
- **Parameter propagation:** this value was passed to the data module for dynamic subset selection.
- **Dynamic loading:** experiments could now run with variable class counts, supporting controlled complexity levels.

Together, these phases enabled EAMs migration to a large-scale domain, evaluation of its rejection capability for novel

stimuli, and a more flexible infrastructure for future experimentation.

IV. RESULTS

This section presents the performance of the autoencoder-classifier and the entropic associative memory as the number of classes varies. Out of the 345 classes in the *Quick, Draw!* dataset, a total of 64 were randomly selected for the experiments, and balanced to 113,613 instances each. The perceptual system was trained using 70% of this dataset. Experiments were run using the first 2, 4, 8, 16, 24, and 32 classes only, due to memory constraints. For each configuration, 20% of the corresponding dataset was used as the entropic associative memory's filling corpus, and 10% as the testing corpus for both the perceptual system and the memory. Due to time constraints, and memory limitations of the server, experiments with more classes were not performed. However, the selection of 64 classes is considered sufficient to observe a clear trend in the results, as their consistency suggests that they are not far from what would be obtained with cross-validation.

As in [2] and [9], the neural networks implementing the perceptual system, and autoencoder and a classifier, were trained jointly. Table I presents their performance on the test set for the different numbers of classes taken from the dataset. As can be observed, the classifier's performance tends to decrease as the number of classes increases, but it remained above 91% accuracy. The autoencoder's performance is slightly more stable, with a difference of less than 3% across all cases.

TABLE I
PERFORMANCE OF THE AUTOENCODER (ROOT MEAN SQUARE ERROR)
AND THE CLASSIFIER (ACCURACY) FOR DIFFERENT NUMBERS OF
CLASSES.

Number of Classes	Classifier Accuracy	Autoencoder RMSE
2	0.976	0.218
4	0.950	0.215
8	0.921	0.204
16	0.914	0.218
24	0.917	0.228
32	0.917	0.229

To facilitate a comparison of the results, Table II provides a summary of the key performance metrics across the two experimental setups; it shows the baseline classifier accuracy alongside the estimated peak accuracy of the EAM in both the recognition and rejection experiments.

A. Recognition Experiment Results

In this experiment, the memory stored all the available classes. The graphs included in Figure 1 show the performance of the entropic associative memory using precision and accuracy metrics. In all cases, the same number of memory columns (256) was kept, while the number of rows was varied in powers of two, from 2 up to 1024. Overall, it can be observed that the number of rows needed for the memory to perform well increases rather slowly with the number of classes stored in it.

TABLE II
SUMMARY OF PERFORMANCE ACROSS EXPERIMENTS.

Number of Classes	Classifier Accuracy	EAM Peak Accuracy (Recognition Exp.)	EAM Peak Accuracy (Rejection Exp.)
2	0.976	0.975	0.943
4	0.950	0.951	0.925
8	0.921	0.920	0.814
16	0.914	0.940	0.686
24	0.917	0.915	0.584
32	0.917	0.917	0.601

The performances with 4 or more rows were very similar to each other and comparable to those obtained by the classifier.

B. Rejection Experiment Results

In this experiment, only the first half of the classes were stored in the memory, but evaluation was performed against the complete set, using the same procedures and the same metrics as in the recognition experiment. The graphs in Figure 2 present the results. It can be observed that the memory performs very well with two and four classes, achieving results close to the corresponding ones in the recognition experiment. However, its performance decreases as the number of classes from the filling corpus (and the number of instances of those classes) stored in memory increases.

V. DISCUSSION

The experimental results allow for several relevant observations about the behavior of the Entropic Associative Memory (EAM) in more complex scenarios. First, as progressively increasing the number of classes, a general trend of performance stability in recognition, while decreasing performance in rejection, was observed. This behavior is expected, as the domain complexity and pattern variability grow with the number of categories and their instances—it is important to notice that the maximum memory size, with 256 columns and 1024 rows, is only of 0.5 MB, considering two bytes per cell. Second, the rejection experiment showed that the memory was capable of identifying patterns corresponding to classes registered within it and rejecting instances of unregistered ones. This result provides empirical validation for the hypothesis that EAM can operate in open-set recognition scenarios, differentiating between familiar and unfamiliar stimuli. However, cases of false acceptances for instances of unregistered classes were also detected as the number of classes increased. This suggests a need of adjusting threshold parameters (ι , κ , and ξ [2] were all set to their default value of zero for these experiments), and exploring complementary discrimination mechanisms.

Collectively, these findings demonstrate that EAM not only can adapt to more complex domains, but it also offers an effective mechanism for handling novel information. At the same time, they highlight future challenges related to optimizing its parameters and evaluating it in even larger and more heterogeneous domains. It is important to note, however, that these conclusions are based on experiments with only 64 of

the 345 available classes, and future work should aim to scale these tests to the entire dataset.

VI. CONCLUSIONS

This work extended the experimental framework of the Entropic Associative Memory (EAM) to evaluate its performance in more complex scenarios and analyze its rejection capacity for unregistered classes. The main contributions can be summarized as follows: (1) we migrated EAM to a wider and more diverse domain using *Quick, Draw!*, with results that confirm the model’s ability to scale in contexts with much more classes and greater variability in patterns; (2) a new experiment was designed and implemented, empirically showing how the EAM is capable of issuing a rejection of stimuli corresponding to unseen classes, validating its usefulness in open-set recognition problems; (3) the system was refactored to accept a variable number of classes and different memory sizes, which enables a more versatile environment for exploring more configurations and comparative analysis.

The findings demonstrate that the EAM is an explainable and robust model, capable of combining distributed storage with probabilistic retrieval and rejection mechanisms. These results reinforce its potential as an alternative paradigm within artificial intelligence, particularly in tasks where the ability to handle novel information is critical.

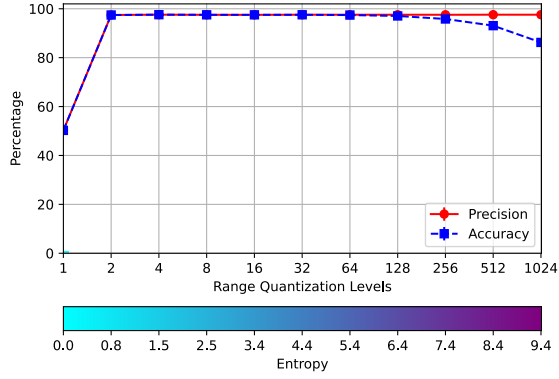
For future work, it is proposed to extend the experiments to even larger and more heterogeneous domains, to optimize the threshold parameters, and investigate the integration of EAM with more advanced deep learning architectures to strengthen both its accuracy and generalization capacity.

REPRODUCIBILITY

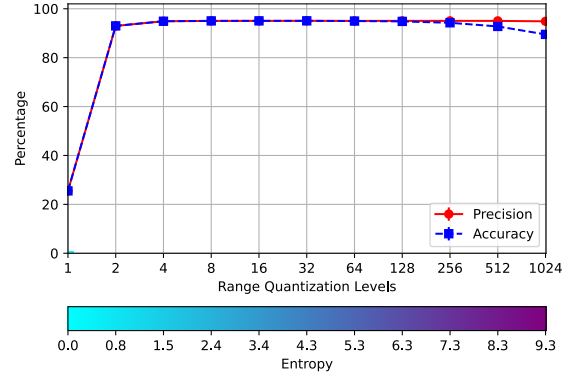
The source code, experimental data, and scripts required to reproduce the results presented in this paper are available in a public repository at <https://github.com/FernyBoy/quick.git>.

REFERENCES

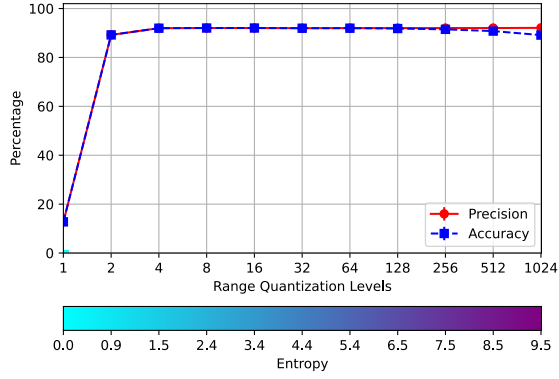
- [1] S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*. Pearson, Apr. 2020.
- [2] L. A. Pineda and R. Morales, “Imagery in the entropic associative memory,” *Scientific Reports*, vol. 13, no. 1, p. 9553, Jun. 2023.
- [3] Z. Research, “Fashion-MNIST,” Nov. 2022.
- [4] W. J. Scheirer, A. de Rezende Rocha, A. Sapkota, and T. E. Boulton, “Toward Open Set Recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 7, pp. 1757–1772, Jul. 2013.
- [5] G. C. Lab, “Quick, Draw! The Data,” <https://quickdraw.with-google.com/data>, Mar. 2025.
- [6] N. Hernández, R. Morales, and L. A. Pineda, “Remembering CIFAR-10 images with the entropic associative memory,” *Pattern Recognition*, vol. 172, p. 112639, Apr. 2026.
- [7] Alex Krizhevsky, “Learning Multiple Layers of Features from Tiny Images,” University of Toronto, Toronto, Canada, Tech. Rep., 2009.
- [8] J. R. González-Hernández, K. M. Figueroa-Mora, L. A. Pineda, and R. Morales-Gamboa, “Clasificador de bosquejos utilizando memoria asociativa entrópica pesada,” *Research in Computing Science*, vol. 153, no. 7, pp. 217–227, 2024.
- [9] R. Morales and L. A. Pineda, “The missing cue problem in hetero associative memory retrieval,” *Scientific Reports*, vol. 15, no. 1, p. 21850, Jul. 2025.



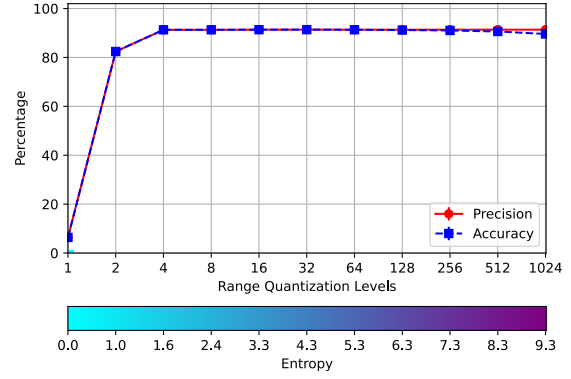
(a) 2 classes



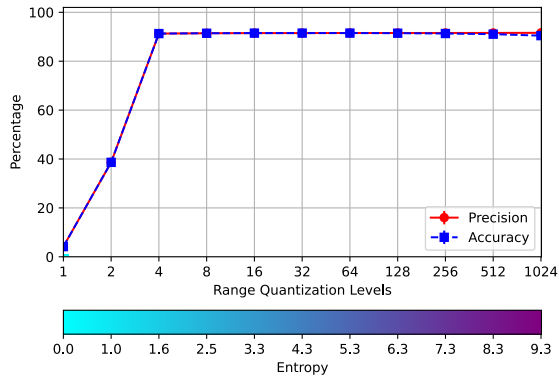
(b) 4 classes



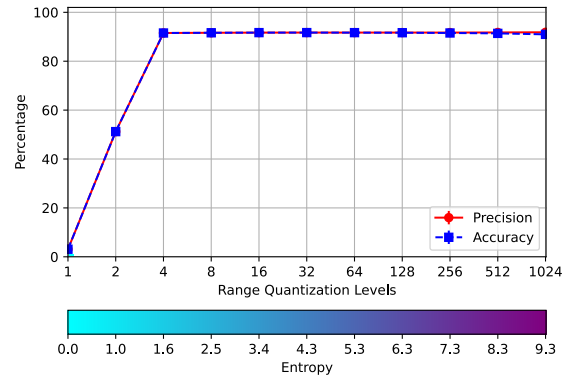
(c) 8 classes



(d) 16 classes

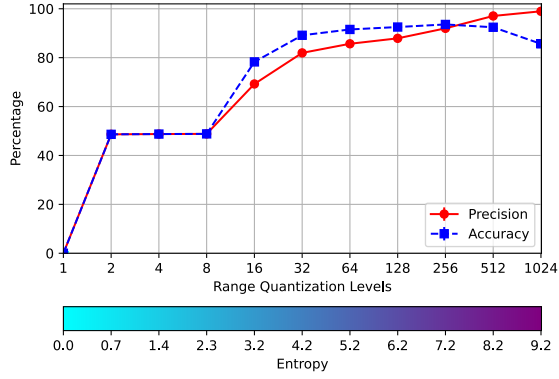


(e) 24 classes

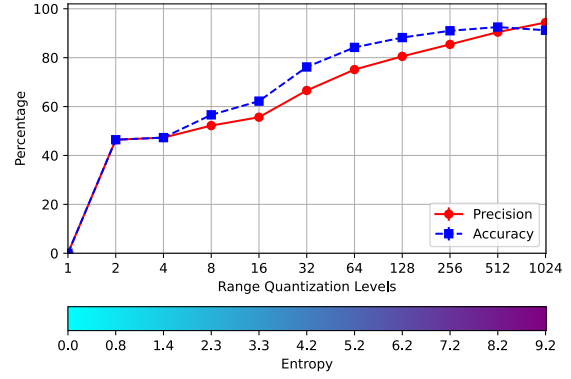


(f) 32 classes

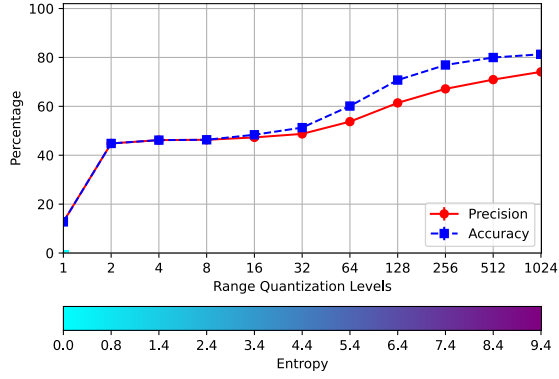
Fig. 1. Precision, accuracy, and entropy of the associative memory across the recognition experiment. The color bar at the bottom of each graph represents the entropy of the memory, where lower entropy (blue) indicates high certainty, and higher entropy (red) indicates high uncertainty.



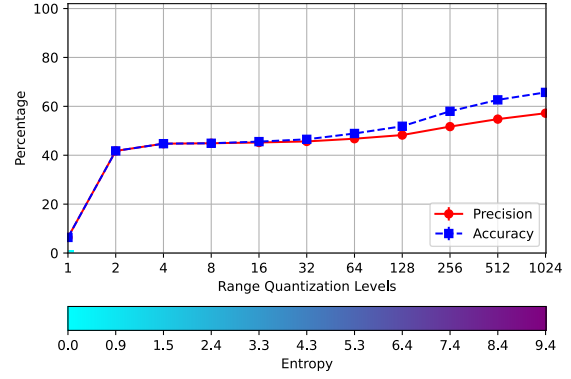
(a) 2 classes



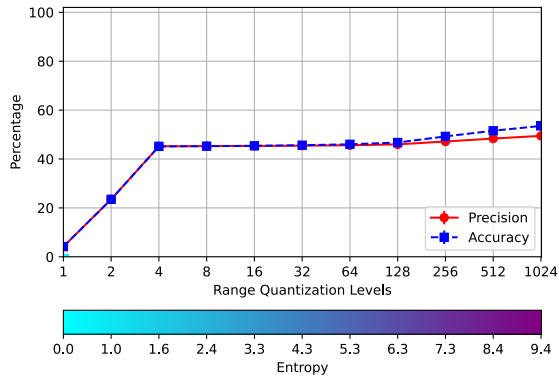
(b) 4 classes



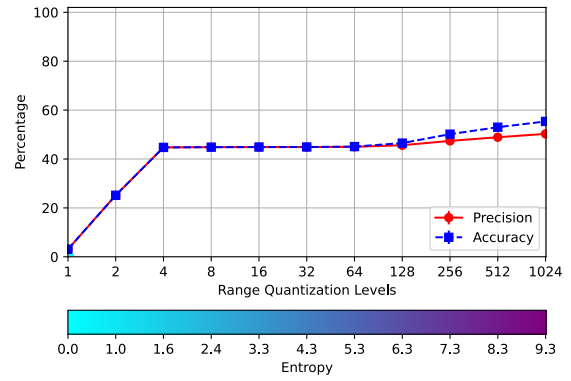
(c) 8 classes



(d) 16 classes



(e) 24 classes



(f) 32 classes

Fig. 2. Precision, accuracy, and entropy of the associative memory across the rejection experiment. The color bar at the bottom of each graph represents the entropy of the memory, where lower entropy (blue) indicates high certainty, and higher entropy (red) indicates high uncertainty.