UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
A Neural Network Model of Complementary Learning Systems: Pattern Separation and Completion for Continual Learning

Permalink
https://escholarship.org/uc/item/6084m9rc

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 47(0)

Authors
Jun, James
Marupudi, Vijay
Shah, Raj Sanjay
et al.

Publication Date
2025

Copyright Information
This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

A Neural Network Model of Complementary Learning Systems: Pattern Separation and Completion for Continual Learning

James P Jun, Vijay Marupudi, Raj Sanjay Shah, Sashank Varma
{jjun44, vijaymarupudi, rajsanjayshah, varma}@gatech.edu
Georgia Institute of Technology

Abstract

Learning new information without forgetting prior knowledge is central to human intelligence. In contrast, neural network models suffer from catastrophic forgetting: a significant degradation in performance on previously learned tasks when acquiring new information. The Complementary Learning Systems (CLS) theory offers an explanation for this human ability, proposing that the brain has distinct systems for pattern separation (encoding distinct memories) and pattern completion (retrieving complete memories from partial cues). To capture these complementary functions, we leverage the representational generalization capabilities of variational autoencoders (VAEs) and the robust memory storage properties of Modern Hopfield networks (MHNs), combining them into a neurally plausible continual learning model. We evaluate this model on the Split-MNIST task, a popular continual learning benchmark, and achieve close to state-of-the-art accuracy (~90%), substantially reducing forgetting. Representational analyses empirically confirm the functional dissociation: the VAE underwrites pattern completion, while the MHN drives pattern separation. By capturing pattern separation and completion in scalable architectures, our work provides a functional template for modeling memory consolidation, generalization, and continual learning in both biological and artificial systems.

Keywords: catastrophic forgetting; catastrophic interference; continual learning; complementary learning systems.

Introduction

While neural networks excel at learning new tasks, they often suffer from *catastrophic forgetting* (CF), where performance on previously learned tasks drops precipitously when learning a new task (McCloskey & Cohen, 1989). In this regard, they differ from humans, for whom *continual learning* (CL) is natural (Mitchell et al., 2018; Parisi et al., 2019). In artificial systems, CL is typically supported through direct rehearsal (Gandhi et al., 2024; Rolnick et al., 2019), involving the replay of stored examples, or through synthetic rehearsal where generative models recreate prior experiences. Among synthetic rehearsal approaches, *generative replay* (GR), which interleaves synthetic samples of previous tasks with new training data, has proven to be highly effective in mitigating catastrophic forgetting (van de Ven et al., 2020, 2022). GR strategies draw direct inspiration from the CLS theory of the mammalian brain, which posits distinct but interacting memory systems within the neocortex and hippocampus (Kumaran et al., 2016; Mcclelland et al., 1995; O'Reilly et al., 2014).

Over time, the pursuit of state-of-the-art performance on CL tasks has detached GR approaches from their neural inspiration, sacrificing neural plausibility for engineering gains. Here, we describe a model architecture guided by the functional principles of the CLS framework. Rather than replicating the *structures* of the hippocampus, as computational cognitive neuroscientists have previously done (Mcclelland et al., 1995; Norman & O'Reilly, 2003; Schapiro et al., 2017), we abstract to the *functions* performed by those structures in concert with the cortex: *pattern separation* and *pattern completion* (Bakker et al., 2008; Larocque et al., 2013). Building on the biologically plausible architecture proposed by Spens & Burgess (2024), we extend it to a continual learning application by integrating the hippocampal (Modern Hopfield Network) and neocortical (variational auto-encoder) components. In consolidating these into a unified architecture (VAE+MHN model), we advance both performance and neuro-plausibility of this framework by enabling simulation of dynamic neocortical-hippocampal interactions during memory encoding and retrieval. We evaluate the model on the standard Split-MNIST benchmark and systematically track both pattern separation and pattern completion across continual learning, using quantitative representational analyses, Euclidean distances for separation, and structural similarity index measure (SSIM) scores for completion, to validate functional dissociations. Finally, we demonstrate that the model achieves high continual learning performance similar to non-continual learning baselines while preserving clear functional roles aligned with CLS principles.

The CLS Framework

In humans, the CLS framework directly addresses the question at the heart of CL: How does the brain learn new information while retaining older memories (Kumaran et al., 2016; Mcclelland et al., 1995; O'Reilly et al., 2014)? CLS posits that this is achieved through two complementary systems: (1) the hippocampus, which encodes episodic memories or the detailed snapshots of ongoing experiences, and (2) the neocortex, which gradually abstracts statistical regularities to form semantic memories. The hippocampus primarily supports *pattern separation*, encoding similar experiences as distinct representations to minimize interference. To facilitate consolidation, it actively *replays* episodic memories over time, training the slower-learning neocortex. The neocortex, in turn, specializes in *pattern completion*: reconstructing complete memories from partial or degraded inputs. Figure 1A illustrates these core structures and their roles within the CLS framework.

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society* ©2025 the author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

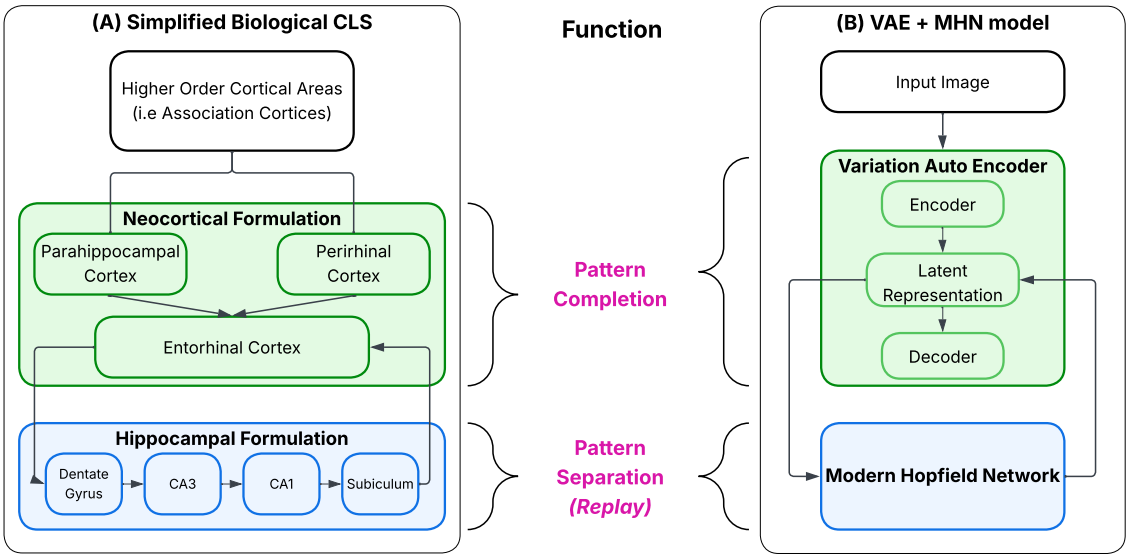

Figure 1: (A) Schematic of the CLS framework, showing the division of memory functions between the hippocampus, which supports pattern separation, and the neocortex, which supports pattern completion. (B) Computational instantiation of these functions in the VAE+MHN model: the MHN enables pattern separation (blue) by storing distinct episodic representations, while the VAE enables pattern completion (green) by abstracting generalized representations from replayed memories. The colors highlight the direct functional mapping between biological structures and their computational counterparts.

Neuroscience-Inspired Generative Replay Methods for CL

Early approaches to CL in artificial systems often relied on *explicit replay*, wherein models directly stored and replayed examples from prior tasks Rolnick et al. (2019). While effective, explicit replay poses serious limitations: it often requires storing large amounts of past data Gandhi et al. (2024), scales poorly with the number of tasks Krawczyk & Gepperth (2024), and lacks biological plausibility, as brains do not store exact copies of past experiences. To address these shortcomings, machine learning researchers turned to *synthetic replay* methods inspired by the CLS framework and, in particular, the hippocampal mechanism of memory replay. Standard GR implements this strategy by training a separate model to generate prior task samples alongside new task data. More recently, Brain-Inspired Replay (BI-R) refines this approach by generating compressed latent representations rather than raw inputs, further improving scalability and achieving strong performance on continual learning benchmarks such as Split-MNIST (van de Ven et al., 2020, 2022).

Through these advances, synthetic replay methods have come to diverge from core principles of biological memory systems. For instance, BI-R saves a separate copy of the generator after each task, an approach that is computationally inefficient and biologically implausible. Recent work by Spens & Burgess (2024) increases biological plausibility by pairing a VAE with an MHN (Ramsauer et al., 2021), where episodic memory replay from the MHN supports gradual learning in the VAE (Figure 1B). Replay from this network serves as a form of "teacher-student learning" (Hinton et al., 2015) to train the VAE. The VAE, in turn, gradually extracts relevant features from replayed samples from the MHN to learn a generalized version of the images presented. However, while their VAE+MHN model demonstrates pattern completion on a static MNIST task, these researchers did not test for pattern separation nor did they evaluate performance under continual, sequential learning settings like Split-MNIST.

In contrast, our model departs from prior work in three key ways: (1) adaption to continual learning to model human learning without storing separate, task-specific models, thereby preserving biological plausibility; (2) implementing latent-space replay instead of raw image replay to align with scalability and neuro-plausibility; and (3) introducing functional analyses to track pattern separation and completion over time, enabling direct evaluation of CLS principles for continual learning.

The Current Study

In this work, we extend the VAE+MHN model to explicitly operationalize both CLS functions: pattern separation and pattern completion, under a continual learning setting. We evaluate the model on the Split-MNIST benchmark, a popular class-incremental task in the machine learning literature. Our results show that the model achieves continual learning performance close to an upper baseline model trained non-continually, without the need for explicit task labels or stored replay buffers. This work demonstrates that biologically inspired memory mechanisms can support continual learning in neural networks.

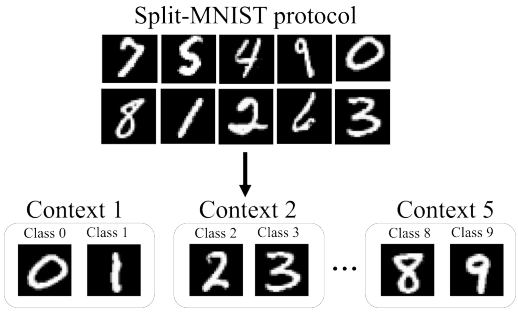

Figure 2: Class-incremental continual learning setup on the MNIST dataset (Split-MNIST Protocol).

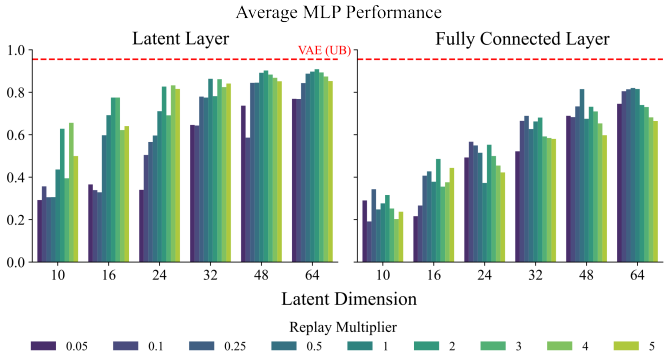

Figure 3: Different parameters were evaluated using Split-MNIST to determine the optimal model architecture. Each experiment was repeated 5 times; the average results are plotted. Upper baseline (VAE UB) performance is shown in red.

Method

The VAE+MHN model was trained on the Split-MNIST task (Figure 2), a benchmark commonly used in continual learning research. In the class-incremental setting, the MNIST dataset (60,000 training and 10,000 test images of handwritten digits, 0–9) is divided into five sequential binary classification tasks, each involving approximately 12,000 images from two classes (e.g., digits 4 and 5). The model experiences each new task *incrementally*, encountering only two new classes at each stage. This setup simulates continual learning by requiring the model to acquire new classes while retaining prior classes.

To understand how architectural hyperparameters impact memory retention and generalization, we systematically explored several model parameters. First, we tested if pattern completion affects eventual memory consolidation by comparing the latent layer (which supports pattern completion) or the fully connected layer (which performs a transformation but lacks generative properties) as inputs to the MHN. We also varied the dimensionality of the latent layer and the replay rate (i.e. ratio of replayed samples from older contexts to new context samples) during training and evaluated its impact on image reconstruction.

Training followed a generative replay procedure aligned with CLS principles. During the first context 1 (0 vs. 1), the model learned from real images without replay, as no prior memories existed. At the end of stage 1 training, 5% of the images (~600 samples) were randomly selected, and their latent representations from the VAE were stored in the MHN. During each subsequent stage, the model was trained on two types of samples: (1) from the current stage and (2) MHN generated samples representing those from prior stages. To produce replay samples, random noise vectors were used as retrieval cues. The MHN performed recurrent settling by iteratively updating its state to minimize an internal energy function (Ramsauer et al., 2021). The settled MHN state was "replayed" (i.e., decoded through the VAE's decoder), generating synthetic replay samples corresponding to previously learned classes. These replayed samples were interleaved with the new class data during training with the goal of mitigating forgetting.

As the generative model operates in an unsupervised manner, evaluating its performance requires judging whether the images it generates are valid numbers. To do so, we trained a separate MLP classifier on MNIST (achieving 97.58% test accuracy) and used it to classify images that were reconstructed by the model (demonstrated in Fig 3). Performance was assessed against the corresponding class label $0-9$.

All VAE models were trained for 10 epochs with a batch size of 32 and a fixed learning rate of 1E-3 with images normalized between 0 and 1 and shuffled to prevent order bias. A VAE loss function combining binary cross-entropy (for reconstruction accuracy) and Kullback-Leibler divergence (to regularize the latent space), with losses from current and (when applicable) replayed images, was summed and backpropagated to update model weights. For the MHN, a large inverse temperature parameter (β) was selected to ensure individual memory recall.

Results

Model performance

Model performance varied substantially with different hyperparameter settings (see Figure 3). Storing the latent layer in the MHN yielded better image reconstruction, particularly as latent dimensionality increased. Replay rate showed a quadratic effect: both low and high ratios impaired performance, with optimal results near a 1:1 replay ratio. We selected the model that stored latent layer representations with a dimensionality of 48 and used a 1:1 replay ratio for reasons of parsimony. This model achieved a test accuracy of 89.71%, which is just below the most accurate model (91.01%). All further results reported in this paper refer to this selected model. The performance of this model trained on Split-MNIST and three useful comparison models are reported in Table 1. To contextualize performance, we compared the selected VAE+MHN model to three benchmarks (Table 1). The upper baseline (UB) model, which was a VAE trained on all 10 MNIST classes simultaneously, reached 95.55% accuracy with no CF, as expected. The lower baseline (LB) was an untrained VAE. The control model, a standard VAE trained sequentially without GR or MHN, exhibited severe CF, achieved test accuracy nearly 30%

less than the UB. In contrast, the VAE+MHN model successfully mitigated CF, achieving test accuracy within 6% of the UB.

VAE+MHN Performance in Sequential Learning

Table 1: The test accuracy of the VAE + MHN model in the CL setting approached the performance of the upper baseline (UB). Offline: Offline learning; CIL: Class Incremental Learning.

Model	Training paradigm	Performance (via MLP)
VAE (UB)	Offline	95.55%
VAE (LB)	Untrained	8.92%
VAE (Control)	CIL	67.75%
VAE+MHN	CIL	**89.71%**

Pattern Separation vs Pattern Completion

Neuroscience studies have mapped the roles of the hippocampus and neocortex in pattern separation and pattern completion, respectively (Kumaran et al., 2016; Mcclelland et al., 1995; Norman & O'Reilly, 2003; O'Reilly et al., 2014; Schapiro et al., 2017). The current study explored how these functions emerge in the VAE+MHN model during continual learning. Guided by the CLS framework, we hypothesized that the MHN would primarily align with hippocampal processing and the pattern separation function. In contrast, the VAE's latent space would align with neocortical processing and the pattern completion function. We analyzed the latent space of the VAE and MHN components of the VAE+MHN model for evidence of these predictions. As references, we also considered the latent spaces of the UB VAE and Control VAE models. The former was trained on all classes simultaneously and should therefore have learned an optimized representation of the relationships between all classes. By contrast, the latter model was trained in a CL setting and experienced catastrophic forgetting, and should have learned an impoverished and incomplete representation.

Pattern separation We quantified intra-class pattern separation. To do this, a set of 128 held-out images per MNIST class was passed through the VAE+MHN (CL) model trained on SplitMNIST, an UB VAE classically trained on MNIST, and a Control VAE trained on Split-MNIST without replay. Representations were extracted at both the VAE latent layer and MHN (if present). MHN representations were extracted by passing latent layer activations of the test image as retrieval cues. All pairwise Euclidean distances within each class were computed and averaged. The larger this distance, the more distinct the representations, and thus the greater the pattern separation.

Pattern separation on the VAE's latent layer (i.e., neocortex) and the MHN (i.e., hippocampus) is shown in Figure 4. As predicted by the CLS, the MHN has learned the more differentiated representation necessary for supporting the pattern separation function, with a large Euclidean distance between

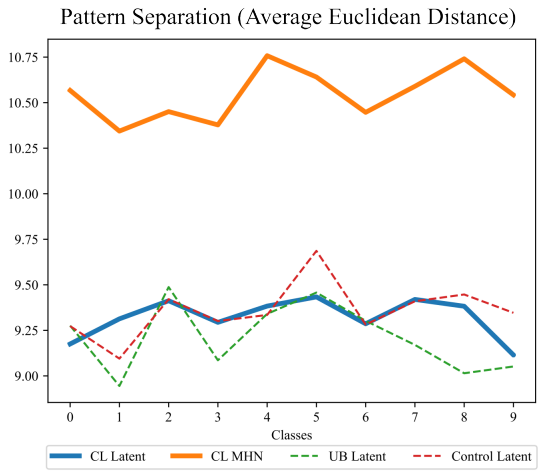

Figure 4: Averaged intra-class Euclidean distance across layers/models is shown. Consistent with CLS, representations within the MHN showing greater pattern separation (i.e., greater Euclidean distance between samples within a class).

the representations of images of the same class. By contrast, the VAE's latent layer shows smaller Euclidean distance between representations of images of the same class, i.e., they are more similar to (and less differentiated from) each other.

This difference is further illustrated in Figure 7, which shows the t-distributed stochastic neighbor embedding (t-SNE) of images of the same class, color coded, as the 5 contexts are learned. The high dimensionality of the activation vector representations is reduced to visualize their underlying structure. As the VAE+MHN model sequentially learns new classes, the MHN representations are more sparsely scattered across the space, reducing overlap between images of the same class and supporting pattern separation. By contrast, the VAE's latent layer representations reveal clear clusters by class. This supports pattern completion and is antithetical to pattern separation (although, as we will see next, these clustered representations support pattern completion and generalization). Thus, the different representations of the two components align with CLS expectations.

To assess statistical significance, we compared intra-class Euclidean distances of the VAE+MHN model's MHN representation and the VAE latent layer representations from the same model, the UB VAE, and the Control VAE. For each MNIST class, we computed all pairwise Euclidean distances between layer representations within that class and performed a two-sample t-test comparing thee MHN representation to each comparator model. This resulted in 30 total comparisons (10 classes x 3 layer pairs), for which we applied Bonferroni correction to control for family-wise error. In all cases, the MHN representation exhibited significantly greater intra-class distance (Bonferroni-corrected $p < 0.001$ after correction), indicating the MHN representation shows enhanced pattern separation within the representational space.

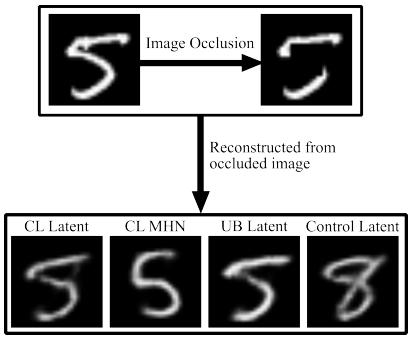

Figure 5: Held-out test images are perturbed (e.g., center occluded) and input into the VAE+MHN model, the UB VAE model, and the Control VAE model. Perturbed images are encoded and then decoded to reconstruct the original image for use in quantifying pattern completion.

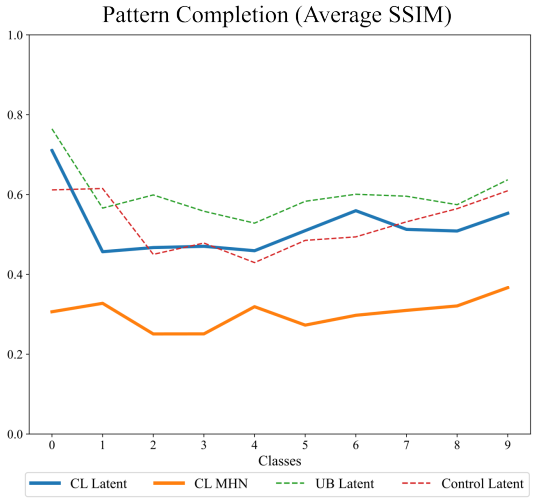

Figure 6: Structural Similarity Index Measure (SSIM) scores across models between reconstructed occluded stimuli and original, unperturbed stimuli are shown. Consistent with CLS, VAE latent layer representations support pattern completion.

Pattern completion To measure pattern completion, we again analyzed the VAE latent layer and MHN representations. We perturbed test images via center occlusion by adding a 10x10 pixel black mask. Each perturbed image was processed by the VAE+MHN model until the latent layer. First, the latent representation was directly decoded to reconstruct an unperturbed image. Second, the latent representation was projected to the MHN to retrieve a hippocampal representation, which was then decoded to reconstruct an unperturbed image. This process is depicted in Figure 5. For comparison, the perturbed image was also presented to the UB and Control models and reconstructed. The similarity of each of the four reconstructed images to the original, unperturbed image was then computed using the Structural Similarity Index Measure (SSIM) (Wang et al., 2004). This is a perceptual measure that captures the structure, contrast, and luminance similarity between a reference image and a reconstructed image. The greater the SSIM, the better the reconstruction, and thus the more effective the pattern completion.

The results are shown in Figure 6. For each of the 10 classes, reconstruction from the VAE+MHN model's VAE latent layer representations resulted in greater similarity to the original, unperturbed image than reconstruction from the MHN representations. In fact, reconstruction from the VAE's latent layer resulted in a performance close to that of the UB.

Referring to Figure 7, we see that the VAE's latent layer representation groups images of the same class. This clustering supports the pattern completion function. The VAE's latent layer representation efficiently consolidates class-level features despite perturbation. By contrast, the MHN prioritizes differentiation over consolidation, making it less suited for the reconstruction of perturbed images. This again reinforces the complementary roles of these components within the VAE+MHN model, consistent with the CLS framework.

We conducted statistical comparisons of intra-class SSIM values across models layers, using the VAE+MHN model's MHN representation as the baseline. Specifically, we compared the SSIM values between the MHN of this model's VAE latent layer representation and the VAE latent layer representations of the UB and Control models. For each MNIST class, we performed a two-sample t-test comparing the MHN's SSIM values against each comparator. Bonferroni correction was applied across all 30 comparisons. In all cases, the MHN representation exhibited lower SSIM values (Bonferroni-corrected $p < 0.001$ after correction), indicating poorer reconstruction consistent with its role in differentiating similar inputs in representational space.

Discussion

We evaluated whether the Spens & Burgess (2024) model, inspired by the CLS framework (Kumaran et al., 2016; Mcclelland et al., 1995; Norman & O'Reilly, 2003; O'Reilly et al., 2014; Schapiro et al., 2017), is able to overcome CF and demonstrate CL. We trained this VAE+MHN model (and several baseline models) on Split-MNIST, a popular class-incremental continual learning benchmark. We applied the concept of teacher-student learning using a fast, hippocampal learner as a "teacher" (MHN) of previous classes to a slow, neocortical "student" (VAE). The model incrementally learned all of the classes and showed high test accuracy at the end of training (89.71%), within range of the upper baseline VAE trained on all classes simultaneously (95.55%). Critically, a representational analysis showed that the MHN component (i.e., the hippocampus analog) supports pattern separation and the VAE's latent layer (i.e., the neocortex analog) supports pattern completion. These findings align with the CLS theory.

Unlike ML models, which focus on the engineering goal of overcoming CF and achieving CL (e.g., van de Ven et al. (2022)), the VAE+MHN model places a higher priority on fidelity to the memory architecture of the mammalian brain. Whilst not a 1:1 mapping, MHNs appear to capture key qualities of the hippocampus, particularly in modeling associative memory via CA3 pyramidal neurons and recurrent connec-

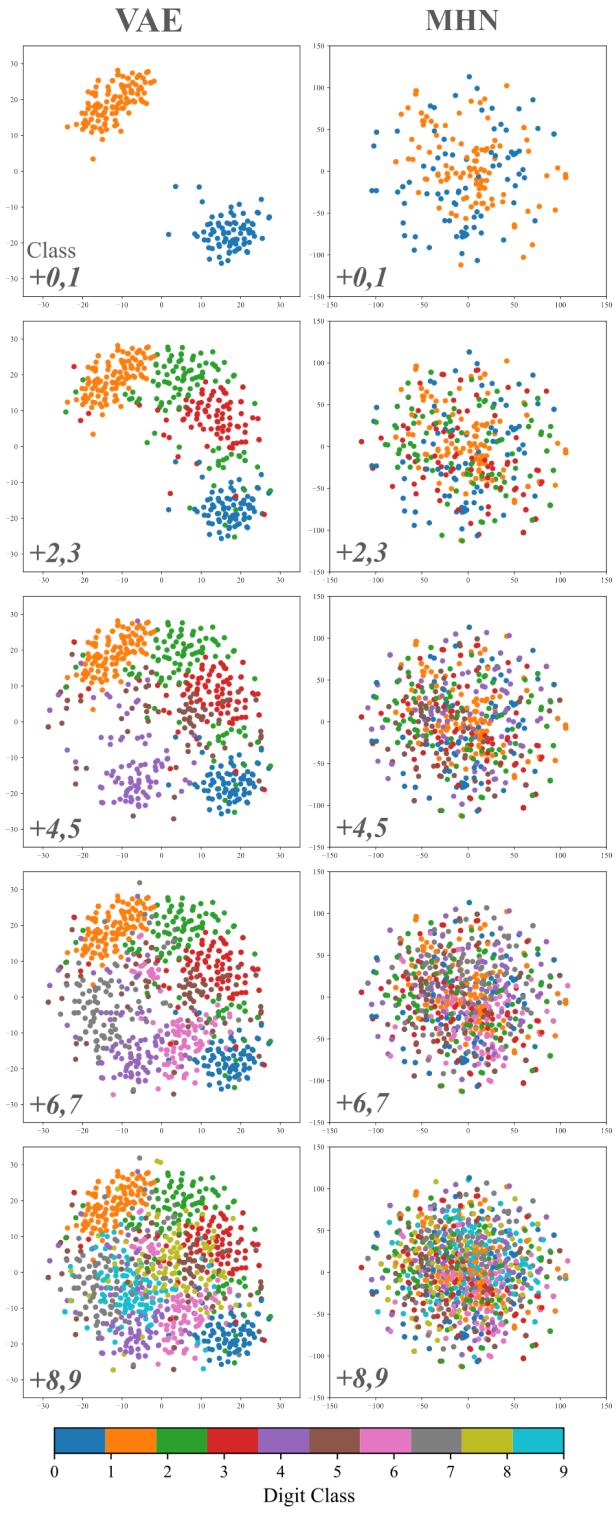

Figure 7: Visualization of t-SNE projections of latent representations from the VAE latent layer (left) and MHN (right) for classes 0-9. Each panel depicts the latent space as the number of classes increases during CL. The VAE latent layer exhibits clustering of the images of each class, supporting its role in pattern completion. By contrast, the MHN shows sparser and more distributed representations of images within each class, supporting its role in pattern separation.

tivity (Krotov & Hopfield, 2021). Our findings support their neural plausibility in showing that they capture the pattern separation function attributed to the hippocampus's CA3 subfield. Also critical from the perspective of neural plausibility is that the MHN is capable of being incrementally updated during learning. This captures the high-frequency interaction between the hippocampus and neocortex during learning, and differentiates the VAE+MHN model from ML approaches that are also inspired by the CLS framework. For example, BI-R saves a separate copy of the generative model (used to generate replayed samples from old tasks) after learning each new task, which is not neurally plausible (van de Ven et al., 2020, 2022). That said, the VAE+MHN model greatly simplifies the biological interactions within the brain compared to more biologically inspired CLS models (Kumaran et al., 2016; Mcclelland et al., 1995; Norman & O'Reilly, 2003; Schapiro et al., 2017), which model the individual subfields of the hippocampus and their connectivity to each other. On the other hand, these models are difficult to scale from laboratory tasks to more demanding ML tasks. The VAE+MHN model sits between these two extremes and provides an opportunity to bridge insights from both computational neuroscience and scalable model architectures.

Our findings should be considered in the context of several limitations. From the cognitive neuroscience perspective, the VAE+MHN model's design simplifies many complex interactions between and within the neocortex and hippocampus of the mammalian brain. Many other structures shape memory, such as the amygdala, prefrontal cortex, cerebellum, basal ganglia, etc. More generally, the model does not fully capture the dynamic and interdependent nature of neural processing.

Our analyses suggest complementary roles for the VAE and MHN in supporting pattern separation and completion, opening up multiple avenues for future exploration. One promising direction is to investigate the neural plausibility of this model beyond basic memory functions. The medial temporal lobe also supports higher-order cognitive processes such as transitive inference, paired associative inference, and acquired equivalence paradigms (Gross & Greene, 2007; Kumaran & McClelland, 2012; Wendelken & Bunge, 2010). While simpler, more neurally aligned models have explored the implementation of these functions (Schapiro et al., 2017), it remains an open question whether biologically grounded mechanisms like those in the VAE+MHN architecture can extend to such forms of reasoning. For example, when trained on "A is greater than B" and "B is greater than C", will the model understand that "A is greater than C", and if so, which components will underwrite that understanding? Our findings confirm that the MHN and VAE components implement dissociable mechanisms of pattern separation and completion, respectively, core functions implicated in memory consolidation. Investigating whether these mechanisms can also support higher-order processes like transitive inference offers a promising direction for understanding how memory structures contribute to reasoning and generalization.

References

Bakker, A., Kirwan, C. B., Miller, M., & Stark, C. E. (2008, 3). Pattern separation in the human hippocampal ca3 and dentate gyrus. *Science*, *319*, 1640-1642. doi: 10.1126/science.1152882

Gandhi, A., Shah, R. S., Marupudi, V., & Varma, S. (2024). Natural mitigation of catastrophic interference: Continual learning in power-law learning environments. *arXiv preprint arXiv:2401.10393*.

Gross, W. L., & Greene, A. J. (2007). Analogical inference: The role of awareness in abstract learning. *Memory*, *15*(8), 838–844. doi: 10.1080/09658210701715469

Hinton, G., Vinyals, O., Dean, J., Hinton, G., Vinyals, O., & Dean, J. (2015). Distilling the knowledge in a neural network. *arXiv*, arXiv:1503.02531. Retrieved from https://ui.adsabs.harvard.edu/abs/2015arXiv150302531H/abstract doi: 10.48550/ARXIV.1503.02531

Krawczyk, A., & Gepperth, A. (2024). An analysis of best-practice strategies for replay and rehearsal in continual learning. In *Proceedings of the ieee/cvf conference on computer vision and pattern recognition* (pp. 4196–4204).

Krotov, D., & Hopfield, J. J. (2021). Large associative memory problem in neurobiology and machine learning. In *International conference on learning representations.* Retrieved from https://openreview.net/forum?id=X4y_10OX-hX

Kumaran, D., Hassabis, D., & McClelland, J. L. (2016, 7). What learning systems do intelligent agents need? complementary learning systems theory updated. *Trends in Cognitive Sciences*, *20*, 512-534. doi: 10.1016/j.tics.2016.05.004

Kumaran, D., & McClelland, J. L. (2012). Generalization through the recurrent interaction of episodic memories: a model of the hippocampal system. *Psychological Review*, *119*(3), 573–616. Retrieved from https://doi.org/10.1037/a0028681 doi: 10.1037/a0028681

Larocque, K. F., Smith, M. E., Carr, V. A., Witthoft, N., Grill-Spector, K., & Wagner, A. D. (2013, 3). Global similarity and pattern separation in the human medial temporal lobe predict subsequent memory. *Journal of Neuroscience*, *33*, 5466-5474. doi: 10.1523/JNEUROSCI.4293-12.2013

Mcclelland, J. L., Mcnaughton, B. L., & O'reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*, 419-457.

McCloskey, M., & Cohen, N. J. (1989, 1). Catastrophic interference in connectionist networks: The sequential learning problem. *Psychology of Learning and Motivation - Advances in Research and Theory*, *24*, 109-165. doi: 10.1016/S0079-7421(08)60536-8

Mitchell, T., Cohen, W., Hruschka, E., Talukdar, P., Yang, B., Betteridge, J., ... Welling, J. (2018, 5). Never-ending learning. *Communications of the ACM*, *61*, 103-115. doi: 10.1145/3191513

Norman, K. A., & O'Reilly, R. C. (2003, 10). Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. *Psychological Review*, *110*, 611-646. doi: 10.1037/0033-295X.110.4.611

O'Reilly, R. C., Bhattacharyya, R., Howard, M. D., & Ketz, N. (2014). Complementary learning systems. *Cognitive Science*, *38*, 1229-1248. doi: 10.1111/j.1551-6709.2011.01214.x

Parisi, G. I., Kemker, R., Part, J. L., Kanan, C., & Wermter, S. (2019, 5). Continual lifelong learning with neural networks: A review. *Neural Networks*, *113*, 54-71. doi: 10.1016/j.neunet.2019.01.012

Ramsauer, H., Schäfl, B., Lehner, J., Seidl, P., Widrich, M., Adler, T., ... Hochreiter, S. (2021). *Hopfield networks is all you need.* Retrieved from https://arxiv.org/abs/2008.02217

Rolnick, D., Ahuja, A., Schwarz, J., Lillicrap, T., & Wayne, G. (2019). Experience replay for continual learning. *Advances in neural information processing systems*, *32*.

Schapiro, A. C., Turk-Browne, N. B., Botvinick, M. M., & Norman, K. A. (2017, 1). Complementary learning systems within the hippocampus: A neural network modelling approach to reconciling episodic memory with statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*. doi: 10.1098/rstb.2016.0049

Spens, E., & Burgess, N. (2024, 3). A generative model of memory construction and consolidation. *Nature Human Behaviour*, *8*, 526-543. doi: 10.1038/s41562-023-01799-z

van de Ven, G. M., Siegelmann, H. T., & Tolias, A. S. (2020, 12). Brain-inspired replay for continual learning with artificial neural networks. *Nature Communications*, *11*. doi: 10.1038/s41467-020-17866-2

van de Ven, G. M., Tuytelaars, T., & Tolias, A. S. (2022, 12). Three types of incremental learning. *Nature Machine Intelligence*, *4*, 1185-1197. doi: 10.1038/s42256-022-00568-3

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004, 4). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, *13*, 600-612. doi: 10.1109/TIP.2003.819861

Wendelken, C., & Bunge, S. A. (2010, May). Transitive inference: distinct contributions of rostrolateral prefrontal cortex and the hippocampus. *Journal of Cognitive Neuroscience*, *22*(5), 837–847. doi: 10.1162/jocn.2009.21226