Unsupervised One-shot Learning of Both Specific Instances and Generalised Classes with a Hippocampal Architecture

Gideon Kowadlo[1] [0000-0001-6036-1180], Abdelrahman Ahmed[2,3] [0000-0001-9530-766X], and David Rawlinson[3] [0000-0001-9443-3840]

Cerenaut, Melbourne, Australia
info@cerenaut.ai
https://cerenaut.ai

Abstract. Established experimental procedures for one-shot machine learning do not test the ability to learn or remember specific instances of classes, a key feature of animal intelligence. Distinguishing specific instances is necessary for many real-world tasks, such as remembering which cup belongs to you. Generalisation within classes conflicts with the ability to separate instances of classes, making it difficult to achieve both capabilities within a single architecture. We propose an extension to the standard Omniglot classification-generalisation framework that additionally tests the ability to distinguish specific instances after one exposure and introduces noise and occlusion corruption. Learning is defined as an ability to classify as well as recall training samples. Complementary Learning Systems (CLS) is a popular model of mammalian brain regions believed to play a crucial role in learning from a single exposure to a stimulus. We created an artificial neural network implementation of CLS and applied it to the extended Omniglot benchmark. Our unsupervised model demonstrates comparable performance to existing supervised ANNs on the Omniglot classification task (requiring generalisation), without the need for domain-specific inductive biases. On the extended Omniglot instance-recognition task, the same model also demonstrates significantly better performance than a baseline nearest-neighbour approach, given partial occlusion and noise.

Keywords: CLS · Hippocampus · One-shot · Specifics · Instances · Unsupervised · Generalisation.

1 Introduction

One-shot learning has seen renewed interest in recent years. Many studies [12,27] are motivated by the apparent limitations of modern ML relative to animal-like learning [14]. An ability for one-shot learning alleviates the reliance on large labelled datasets in which samples are assumed to be i.i.d., implying an unchanging world. This is particularly relevant for autonomous real-world agents, where samples are necessarily highly correlated and typically unlabelled.

However, we believe that the standard approach - classification of general classes - does not go far enough. Learning specific instances is crucial for intelligent agents, and is something we take for granted. For example, identifying your own coffee cup from other cups, in addition to recognising that it belongs to the 'cup' category. More generally, it underpins memory for singular facts and an individual's own autobiographical history, important for future decision making.

At first glance, learning specific instances appears easy. An obvious starting point is nearest-neighbour lookup in a buffer of past observations. However, this approach may perform poorly given observational variation such as occlusion, and have trouble also generalising class recognition ability. Conversely, methods that can generalise would be unlikely to do well at learning specific instances, as they are conflicting capabilities.

Learning of specific instances is not to be confused with Instance-based Learning [24]. Such approaches store instances during training, and use them to classify test samples e.g. k-nearest neighbour, SVM's and RBFs. Our objective is learning a model of the instances, despite observational variation, while being able to distinguish even very similar instances from each other. We identify two important aspects of learning - classification and recall (generation) of concept.

Complementary Learning Systems (CLS) is a model of mammalian learning that describes the interplay between the neocortex and a region called the Hippocampal Formation (HF) [19,22]. CLS is believed to be crucial for fast learning and is recognised to be important for intelligent agents [10]. Our motivation is to expand the standard definition of one-shot learning, and to test if the CLS architecture can satisfy the requirements.

In this paper we propose a broader benchmark for one-shot learning that includes robust classification of specific instances given observational variance by introducing image corruption with occlusion and noise. We present an ANN implementation of CLS using an Artificial Hippocampal Algorithm (AHA) and apply it to the extended benchmark. The performance of the system is compared to two baselines, a simplified version of CLS that replaces the hippocampal model with a conventional ML model optimised for the task, and to the naive solution for learning specifics - a buffer with nearest neighbour lookup.

2 Background

2.1 One-shot Learning

Following seminal work by Li et al. [15,16], the area was re-invigorated by Lake et al. [11], who introduced a popular test that has become a standard benchmark [12]. It is a one-shot classification task on the Omniglot dataset of handwritten characters. Classification is posed as a matching task, where a given character must be matched with a character of the same class in a test set, see Section 4 for details. The framework was formalised by Vinyals et al. [27].

A typical approach is to pre-train a model on many classes and use learnt concepts to recognise new classes quickly from one or few examples. Often framed as meta-learning or "learning to learn", there are multiple implementations using neural networks that require external labels and supervised learning during pre-training, such as Siamese networks [8], matching networks [27], and prototypical networks [26]. Two notable Bayesian approaches, BPL [12] and RCN [2], achieve above and close to human level performance respectively. The superior performance of BPL may be partially explained by its use of prior knowledge about handwriting via stroke formation. RCN, by virtue of the design which is modeled on the visual cortex, is also specialised for this type of visual task. It is less clear how it could be applied to other datasets and problems where contour topology is less distinct or relevant. A comprehensive review is given in [13].

2.2 Complementary Learning Systems (CLS)

Complementary Learning Systems (CLS) is a standard framework for understanding the function of the HF [19,21,10]. CLS consists of two differentially specialised and complementary structures, neocortex and HF, shown in Figure 1a. In this framework, the neocortex is analogous to a conventional ML model, incrementally learning regularities across many observations, comprising a long-term memory (LTM). It forms overlapping and distributed representations that are effective for inference. In contrast, the HF rapidly learns distinct observations, forming sparser, non-overlapping and therefore non-interfering representations, functioning as a short term memory (STM). Recent memories from the HF are replayed to neocortex, re-instating the original activations resulting in consolidation as long-term memory (LTM). Patterns are replayed in an interleaved fashion, avoiding catastrophic forgetting. In addition, they can be replayed selectively according to salience. There have been numerous implementations of CLS [20,7,4,23,25] and Rolls et al. presented a similar model with greater neuroanatomical detail [22].

Overall, the HF functions as an autoassociative memory that can recall memories from partial cues. CLS describes the HF in terms of distinct functional units called subfields. Together they comprise a unification of pattern completion and pattern separation pathways. Reported implementations (see citations above) are expressed at the level of individual neurons replicating known biological plasticity and dynamics, and have not been applied to ML benchmarks.

3 Model

Our approach is to implement a CLS-style STM with an LTM (Figure 1a), with biological plausibility constraints [9] - all components are trained with local and immediate credit assignment. The LTM comprises a simple vision component suitable for image feature extraction. The STM is implemented with AHA, an Artificial Hippocampal Algorithm, which follows the subfield architecture of the HF described by CLS. There are two baselines for comparison. Firstly, the LTM alone is compared to performance of LTM+STM. It constitutes a naive solution to classifying specific instances (see Section 4). The second baseline, FastNN, is an alternate STM comprised of a standard ML component empirically optimised for the same tasks. We provide the code and configuration required to reproduce experiments in a GitHub repository[1].

3.1 Training and Testing Framework

With complementary systems, training and testing are non-standard, and are therefore explained here to provide context for the remainder of the paper.
Stage 1: Pre-train LTM: Train LTM on a training set over multiple epochs. The LTM learns incrementally about common features that can be used compositionally to represent unseen classes.
Stage 2: Evaluate LTM+STM: Evaluation is conducted with a disjoint evaluation set. The LTM does not learn during this stage. Training and Testing of STM occurs rapidly, allowing multiple internal cycles but only one exposure to an external stimulus. The STM is reset after each evaluation[2]. Evaluation consists of two steps performed in succession - Train (encoding) and Test (inference).

- *Train*: A small support set is presented once (referred to as 'study set' in CLS). STM modules are set to train mode to learn the samples.
- *Test*: A small query set is presented (referred to as 'recall' in CLS), STM modules are in inference mode. For each 'recall' sample, the system is expected to retrieve the corresponding sample from the 'study' set. If correct, it is considered to be 'recognised' - an AHA moment!

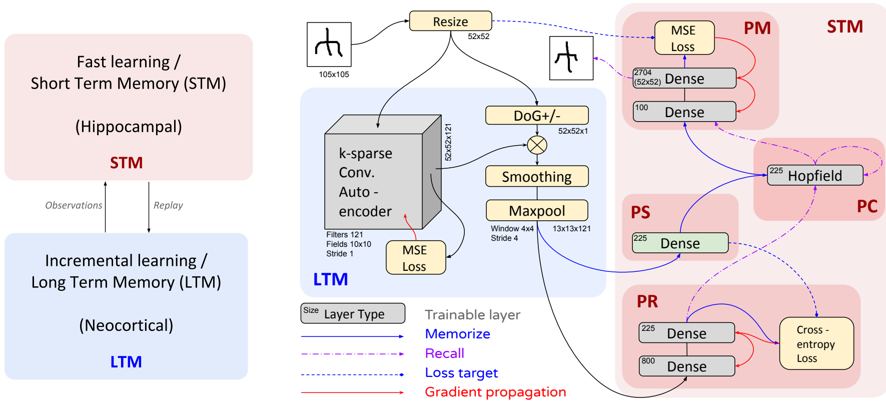
3.2 LTM - Vision Component

The role of the LTM is to process high-dimensional sensor input, and output relatively abstract visual features that can be used as compositional primitives. A single layer convolutional sparse autoencoder based on [17,18] provides the required embedding. However, in Omniglot there is a lot of empty background that is encoded with strong hidden layer activity. Lacking an attention mechanism, this detracts from compositionality of foreground features. To suppress encoding of the background, we added an 'Interest Filter' which loosely mimics known retinal processing (see below). Smoothing is applied to provide some tolerance to feature location and a final max-pooling stage to reduce dimensionality.

Interest Filter The retina possesses centre-surround inhibitory and excitatory cells that can be well approximated with a Difference of Gaussians (DoG) kernel [28]. Positive and negative DoG filters are used to enhance positive and negative intensity transitions. Local non-maxima suppression merges nearby features and a 'top-k' function creates a mask of the most significant features globally. Positive and negative masks are combined by summation giving a final 2D mask that is applied to all channels of the convolutional autoencoder output.

[1] https://github.com/Cerenaut/aha

[2] Adam Optimizer is reset and trainable parameters are re-initialised.



(a) (b)

Fig. 1. a) **CLS**: The STM learns and forgets rapidly. Salient memories are replayed to the LTM for incremental statistical learning. b) **System diagram:** Our implementation of CLS. Local credit assignment via shallow backpropagation is used throughout. The dense layer in PS (green) is initialised, but not trained.

3.3 STM - Artificial Hippocampal Algorithm (AHA)

AHA is our implementation of CLS. For greater details on CLS, the biological basis for AHA design choices, and in-depth implementation details, see [9]. The components and connectivity are shown in Figure 1b. LTM outputs sparse distributed overlapping patterns. The signal becomes sparser and more orthogonal through PS, minimising interference between patterns, resulting in distinct representations for similar inputs.

In train mode, PS patterns are encoded/memorised into PC, an autoassociative, content-addressable memory. They form a target, which PR learns to retrieve from LTM distributed representations. PM learns to map from the stored non-interfering patterns, to the originating sparse distributed patterns.

In test mode, PR retrieves the corresponding stored PC pattern using input from LTM, which is used to cue complete recall from PC. PS is not used as a cue, because even small input differences will result in orthogonal PS outputs. PC can retrieve a crisp, complete pattern, that in turn enables PM to recall the original observation. In future work this will be used for improved inference and consolidation of memories.

The use of PS for encoding and PR for recall is based on the Hippocampal model by Rolls [22,23]. The role of each subfield is detailed below:

PS - Pattern Separation PS is implemented with a single fully-connected Artificial Neural Network (ANN) layer with sparsity constraints and temporal inhibition. Sparsity is implemented as a 'top-k' ranking per sample, mimicking a local competitive process via inhibitory interneurons. Low k produces outputs with low overlap, but orthogonality is further improved by replicating the sparse connectivity observed in this pathway in the hippocampus [23]. A portion of the incoming connections are removed by setting weights to zero (similar to the sparsening technique of [1]). Additionally, after a neuron fires (i.e. it is amongst the top-k), it is temporarily inhibited, mimicking the refractory period observed in biological neurons. PS is initialised with uniformly distributed random weights and does not undergo any training.

PC - Pattern Completion PC is implemented with a Hopfield network [6]. Unlike a standard Hopfield network, there are separate input pathways for encoding (PS) and recall (PR). Output layers of PS and PR are the same size as PC. PS and PR output signals are conditioned from a continuous value [0,1] to a binary signed unit range [-1,1], chosen for better Hopfield performance.

PR - Pattern Retrieval PR is implemented with a 2-layer fully-connected ANN. In training, PS output is used as an internally generated label constituting self-supervised learning [3]. Usually in self-supervised learning, prior task knowledge is used to set a pre-conceived goal such as rotation, with the motivation of learning generalisable representations. In the case of AHA, no prior is required. The motivation is separability and as such, the use of orthogonal patterns as labels is very effective.

PM - Pattern Mapping PM is implemented with a 2-layer fully-connected ANN. In this study we trained it to reconstruct the input images rather than the LTM output, for easy assessment of recalled image quality and correctness.

AHA - Theory of Operation

Compositionality A central capability of AHA is the memorisation of new conjunctions of primitive concepts. The primitives can be composed in a vast number of new combinations, a feature of animal-like learning [14]. Memorisation of conjunctions of concepts is an aspect of episodic memory, as identified in the hippocampal computational modelling literature [7] (expanded in [9]).

Generalisation to subsequent observations of the new combination is achieved through unification of separation and completion (below). The scope of generalisation depends on the level of abstraction of the primitives.

Unifying Separation and Completion Separation and completion are conflicting capabilities requiring separate pathways. Unification is achieved through the collaboration of PS and PR. PS sets a target for PR and PC to learn, providing a common representational 'space'. This makes it possible to separate PC encoding and retrieval between the separation and completion pathways respectively. In this way, they don't conflict with each other, but each operate to their strengths.

3.4 Baseline STM - FastNN

FastNN is a 2-layer fully-connected ANN. Like AHA, the target for recall is the input image itself (rather than LTM encodings) for ease of analysis. It is 'fast' in that it also learns given only one external stimulus. We empirically optimised the learning rate, training iterations, number and size of hidden layers and the other hyperparameters.

4 Experiments

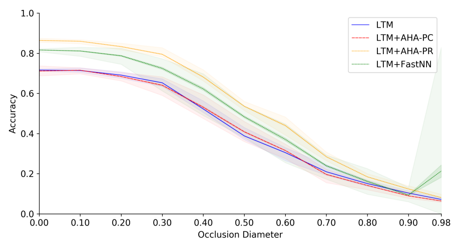
Omniglot Benchmark - One-shot classification task We tested our models on the one-shot classification test from [12]. Referring to the train/test framework (Section 3.1), first the LTM is pre-trained on a 'background' set of 30 alphabets. Then using a disjoint 'evaluation' set of 10 alphabets, a single 'train' character is presented. The task is to identify the matching character from 20 distinct 'test' characters from the same alphabet by a different writer. This is repeated for 20 characters comprising a single 'run'. The experiment consists of 20 runs in total. Accuracy is averaged over the 20 runs. Characters and alphabets were selected to maximise difficulty through confusion of similar characters [12].

The method used to determine the matching character varies between reported studies. In this work, we use minimum MSE of an internal representation. For LTM, we used the autoencoder encoding, for LTM+AHA, we report PR and PC and for LTM+FastNN the hidden layer encoding. None of these networks are explicitly trained to classify. In addition to accuracy, the quality of end-to-end retrieval of appropriate memories is assessed with MSE recall loss.

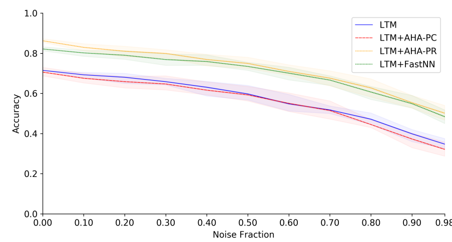
One-shot instance-classification task We extended the experiments with the one-shot instance-classification task. It is the same as one-shot classification, except that the 'train' character exemplar must be matched with the exact same exemplar amongst 20 'test' distractor exemplars of the *same* character class. Being the same character class, all the exemplars are very similar making separation difficult. In each run, the character class and exemplars are selected by randomly sampling without repeats from the 'evaluation' set.

Common Conditions In addition, we explored robustness by introducing image corruption to emulate realistic challenges in visual processing that could also apply to other sensory modalities. Noise emulates imperfect sensor capture. For example, in visual recognition, the target object might be dirty or lighting conditions changed. Occlusion emulates incomplete sensing e.g. due to obstruction by another object. Robust performance is a feature of animal-like learning that would confer practical benefits to machines, and is therefore important to explore [1]. Occlusion is achieved with randomly placed circles, completely contained within the image. Noise is introduced by replacing a proportion of the pixels with a random intensity value drawn from a uniform distribution.

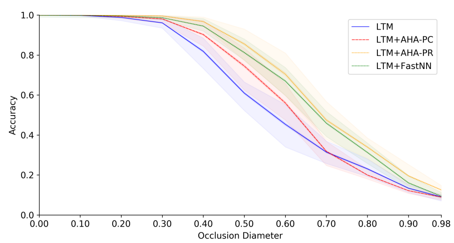
For both tests, instead of presenting 1 test character at a time, all 20 are presented simultaneously, made possible by the short term memory of CLS.



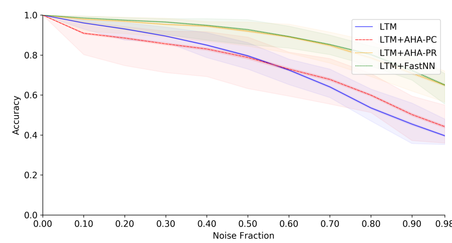
(a) One-shot classification with occlusion



(b) One-shot classification with noise



(c) One-shot instance-classification with occlusion



(d) One-shot instance-classification with noise

Fig. 2. Accuracy vs occlusion and noise. LTM+STM improves performance over baseline LTM. AHA STM is superior to the baseline FastNN STM. The effect is more pronounced for occlusion than noise. Occlusion diameter is expressed as a fraction of image side length, noise as a fraction of image area. The mean value is bold with medium shading for 1 standard deviation, light shading demarcates min/max values.

Noise and occlusion is increased from none, to almost complete corruption, in 10 increments. The highest level is capped at 98% corruption, to ensure some meaningful output. Every test is repeated with 10 random seeds.

In one-shot classification, strong generalisation is required as well as some pattern separation to distinguish similar character classes. The one-shot instance-classification task requires strong pattern separation, as well as some generalisation for robustness.

5 Results

5.1 Accuracy

One-shot classification results are shown in Figures 2a and 2b. LTM accuracy starts at 71.6%, without image corruption. Increasing noise affects all features equally and gradually, whereas occlusion increases the likelihood of suddenly removing important topological features. At very high occlusion, the character is mostly covered, leading to chance level performance. All signals follow the same overall trend as LTM.

Table 1. Comparison of algorithms for one-shot classification, without image corruption. LTM+AHA is competitive with state-of-the-art convolutional approaches whilst demonstrating a wider range of capabilities.

Algorithm	Accuracy (%)	Algorithm	Accuracy (%)
BPL [13]	96.7	**LTM+AHA**	**86.4**
Human [13]	*95.5*	Prototypical Net [13]	86.3
RCN [2]	92.7	**LTM+FastNN**	**81.9**
Simple Conv Net [13]	86.5	VHE [5]	81.3

With no noise or occlusion, PR at 86.4% has an advantage of almost 15% over LTM and is comparable to other ANN results for this task (Table 5.1). This advantage is maintained with moderate levels of occlusion. As extreme occlusion begins to cover most of the character, the accuracy of PR, PC and LTM converge. However, the advantage is maintained over all noise levels. PC accuracy was no better than LTM.

With no corruption, FastNN improves on LTM accuracy by 10.3%, 4.4% less than the AHA improvement. The advantage of AHA over FastNN is maintained over almost all levels of occlusion, and minor for all levels of noise.

For context, reported accuracy in the case of zero noise or occlusion is contrasted with other works in Table 5.1. Existing values are reproduced from [13].

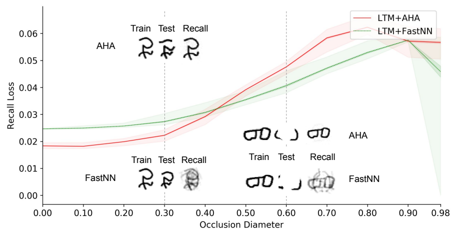
One-shot instance-classification results are shown in Figures 2c and 2d. LTM accuracy is perfect at low levels of image corruption, remaining almost perfect in the case of occlusion, until approximately one third of the image is affected. All signals follow the same trends observed for the one-shot classification task.

For AHA, PR accuracy remains extremely high, close to 100% until a 10% greater level of occlusion than for LTM i.e. addition of AHA increases tolerance to occlusion. The advantage over LTM increases with increasing corruption, fading away for occlusion but continuing to grow for noise.

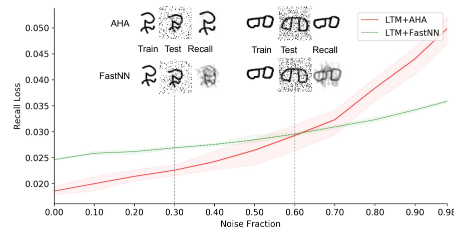
FastNN also improves on the baseline. It has worse accuracy than AHA for a given level of occlusion (less substantial than one-shot classification), and almost equal accuracy for varying levels of noise.

5.2 Recall

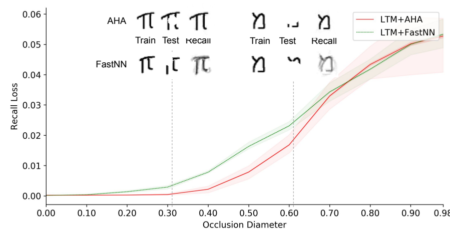
Recall-loss is shown in Figure 3. In the one-shot classification experiment, AHA demonstrates better performance than FastNN under moderate occlusion and noise. At higher levels of corruption, AHA may retrieve a high quality image of the wrong character, resulting in a higher loss than lower-quality images retrieved by FastNN. In the one-shot instance-classification experiment, this character confusion is less likely to occur and AHA is superior or equal to FastNN under all meaningful levels of image corruption. FastNN is qualitatively better for one-shot instance-classification than for one-shot classification, and almost as good as AHA, but recalls are typically an ‘average’ version of the character, rather than a specific instances.



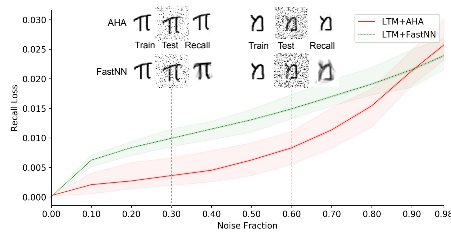
(a) One-shot classification with occlusion



(b) One-shot classification with noise



(c) One-shot instance-classification with occlusion



(d) One-shot instance-classification with noise

Fig. 3. Recall-loss vs occlusion/noise. AHA yields more specific and crisp recall images given moderate input corruption. With substantial corruption, AHA sometimes retrieves an accurate copy of the wrong character resulting in a higher loss. FastNN provides blurry or nonspecific recall in many conditions. Units and plot characteristics are the same as for Figure 2.

6 Discussion

The results demonstrate that CLS is an effective approach to one-shot classification of both specific instances and categories. At first sight, one-shot instance-classification is trivially solved by a nearest neighbour comparison, but the results show that this approach performs poorly given realistic levels of image corruption. In addition, the CLS-style architecture of AHA has an advantage over the simpler FastNN. For accuracy, this is most noticeable on one-shot classification and in the presence of occlusion, and it is significantly superior at recalling high quality images across the test conditions. To the authors best knowledge, this is the first application of CLS to a dataset derived from real-world observations featuring non-synthetic variation.

AHA was comparable to state-of-the-art approaches on the standard Omniglot Benchmark - a subset of our extended test. The reported approaches are optimised for one-shot classification without any image corruption suggesting that they are not suited for one-shot instance-classification. Referring to Table 5.1, BPL and RCN are significantly ahead of other methods, and similar to human performance. They have an advantage as they exploit domain specific priors as discussed in Section 2.1. The Simple Conv Net (CNN) represents a standard approach for deep learning. AHA is equally good despite being unsupervised (no external labels) and uses only local local credit assignment. Additionally, AHA demonstrates the broader range of capabilities discussed.

PR performs classification significantly better than PC. It partially fulfils the role of completion, as it learns to reproduce the target. PC fulfils a vital role for additional completion and sharpening for crisp recall. There is a small accuracy bias toward PR due to the fact that PR outputs a superposition of possible patterns, enhancing the chance of a correct match via MSE. In contrast, PC is designed to retrieve a single, sharp complete sample and in doing so is unable to hedge its bets.

The boundary between class and exemplar is continuous, subjective and may depend on the task. For example, you could define the character itself as a class, and the corrupted samples as exemplars. Or a Labrador dog: the class could be the animal type (dog), or the breed (Labrador). AHA demonstrates this flexibility to the task by accomplishing both one-shot classification and one-shot instance-classification. As per Section 3.3, AHA learns a conjunction of primitives, and then generalises over variations in that combination.

Acknowledgment

Thanks to Elkhonon Goldberg for enriching discussions on the hippocampal region and to Rotem Aharon for insights and analysis of Hopfield networks.

References

1. Ahmad, S., Scheinkman, L.: How Can We Be So Dense? The Benefits of Using Highly Sparse Representations. arXiv preprint arXiv:1903.11257 (2019)
2. George, D., Lehrach, W., Kansky, K., Lazaro-Gredilla, M., Laan, C., Marthi, B., Lou, X., Meng, Z., Liu, Y.: A Generative Vision Model that Trains with High Data Efficiency and breaks text-based CAPTCHAs. Science pp. 1–19 (2017)
3. Gidaris, S., Bursuc, A., Komodakis, N., Pérez, P., Cord, M.: Boosting Few-Shot Visual Learning with Self-Supervision. In: Proceedings of the IEEE International Conference on Computer Vision. (2019)
4. Greene, P., Howard, M., Bhattacharyya, R., Fellous, J.M.: Hippocampal anatomy supports the use of context in object recognition: A computational model. Computational Intelligence and Neuroscience **2013**(May) (2013)
5. Hewitt, L.B., Nye, M.I., Gane, A., Jaakkola, T., Tenenbaum, J.B.: The Variational Homoencoder: Learning to learn high capacity generative models from few examples. arXiv preprint arXiv:1807.08919 (jul 2018)
6. Hopfield, J.J.: Neural networks and physical systems with emergent collective computational abilities. Proceedings of the National Academy of Sciences **79**(8) (1982)
7. Ketz, N., Morkonda, S.G., O'Reilly, R.C.: Theta Coordinated Error-Driven Learning in the Hippocampus. PLoS Computational Biology **9**(6) (2013)
8. Koch, G., Zemel, R., Salakhutdinov, R.: Siamese Neural Networks for One-shot Image Recognition. In: Proceedings of the 32nd International Conference on Machine Learning (2015)

9. Kowadlo, G., Ahmed, A., Rawlinson, D.: AHA! an 'Artificial Hippocampal Algorithm' for Episodic Machine Learning. arXiv preprint arxiv:1909.10340 (2019)
10. Kumaran, D., Hassabis, D., McClelland, J.L.: What Learning Systems do Intelligent Agents Need? Complementary Learning Systems Theory Updated. Trends in Cognitive Sciences **20**(7), 512–534 (jul 2016)
11. Lake, B.M., Salakhutdinov, R., Gross, J., Tenenbaum, J.B.: One shot learning of simple visual concepts. In Proceedings of the 33rd Annual Conference of the Cognitive Science Society (2011)
12. Lake, B.M., Salakhutdinov, R., Tenenbaum, J.B.: Human-level concept learning through probabilistic program induction. Science **350**(6266), 1332–1338 (2015)
13. Lake, B.M., Salakhutdinov, R., Tenenbaum, J.B.: The Omniglot challenge: a 3-year progress report. Current Opinion in Behavioral Sciences **29**, 97–104 (2019)
14. Lake, B.M., Ullman, T.D., Tenenbaum, J.B., Gershman, S.J.: Building machines that learn and think like people. Behavioral and Brain Sciences **40**(2012), 1–58 (2017)
15. Li, F.F., Fergus, R., Perona, P.: A Bayesian approach to unsupervised one-shot learning of object categories. In: Proceedings Ninth IEEE International Conference on Computer Vision (2003)
16. Li, F.F., Fergus, R., Perona, P.: One-shot learning of object categories. IEEE Transactions on Pattern Analysis and Machine Intelligence (2006)
17. Makhzani, A., Frey, B.: K-Sparse Autoencoders. arXiv preprint arXiv:1312.5663 (2013)
18. Makhzani, A., Frey, B.J.: Winner-Take-All Autoencoders. In: Advances in Neural Information Processing Systems. pp. 2791–2799 (2015)
19. McClelland, J.L., McNaughton, B.L., O'Reilly, R.C.: 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**(3), 419–457 (1995)
20. Norman, K.A., O'Reilly, R.C.: Modeling Hippocampal and Neocortical Contributions to Recognition Memory: A Complementary-Learning-Systems Approach. Psychological Review **110**(4), 611–646 (2003)
21. O'Reilly, R.C., Bhattacharyya, R., Howard, M.D., Ketz, N.: Complementary learning systems. Cognitive Science **38**(6), 1229–1248 (2014)
22. Rolls, E.T.: A model of the operation of the hippocampus and entorhinal cortex in memory. International Journal of Neural Systems **6** (1995)
23. Rolls, E.T.: The mechanisms for pattern completion and pattern separation in the hippocampus. Frontiers in Systems Neuroscience **7**(October), 1–21 (2013)
24. Russell, S.J., Norvig, P.: Artificial intelligence: a modern approach. Prentice hall (2009)
25. Schapiro, A.C., Turk-Browne, N.B., Botvinick, M.M., Norman, K.A.: 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**(1711), 20160049 (2017)
26. Snell, J., Swersky, K., Zemel, R.S.: Prototypical Networks for Few-shot Learning. In: Advances in neural information processing systems. pp. 4077–4087 (2017)
27. Vinyals, O., Blundell, C., Lillicrap, T., Kavukcuoglu, K.: Matching Networks for One Shot Learning. In: Advances in neural information processing systems (2016)
28. Young, R.A.: The Gaussian derivative model for spatial vision: I. Retinal mechanisms. Spatial vision **2**(4), 273–93 (1987)