Continua (2023), x:x, 1–9

FOR REVIEW

Enhancing Object Recognition through Evolutionary Selection of Features

Kimia FakheriSharif University of Technology, Tehran, Iran
kimia.fakheri@gmail.com

(Received 25 Apr 2025; revised 25 Apr 2025; accepted 25 Apr 2025; published 25 Apr 2025)

Abstract

Human visual perception excels in recognizing objects amidst diverse, cluttered environments. This ability has driven the development of computational models that aim to replicate the human visual system's hierarchical approach to object recognition. These models, including HMAX, simulate the process of extracting progressively more complex features from simple image patches. However, the quality of the extracted features plays a crucial role in recognition accuracy.

In this work, we focus on improving the selection of features in the HMAX model, which traditionally uses randomly chosen patches. We introduce a Genetic Algorithm (GA) to optimize patch selection. The GA helps identify the most informative features for recognition, reducing the impact of non-discriminative patches. Through extensive testing on object recognition tasks, we demonstrate that the GA-enhanced HMAX model outperforms the original approach by selecting more effective and discriminative patches. Our results indicate that the GA's feature selection leads to improved performance, particularly when dealing with limited training data.

Keywords: HMAX Genetic Algorithm Object Recognition Feature Selection Machine Learning

1. Introduction

How different objects are recognized in the visual cortex has been a challenging and major question in the field of vision neuroscience and machine vision. The visual system of humans and other mammals can simply and rapidly recognize a wide variety of objects in various conditions such as changes in size, position, illumination, viewpoint, etc. in a natural scene. They can even detect and recognize a specific object in a cluttered scene without consuming a noteworthy amount of time and effort, unlike the best machine vision systems. Achieving a model that can emulate this remarkable system with such high performance is a long-time goal in computational neuroscience.

Although presenting a model with high performance in object recognition tasks is a goal of interest, plausibility with the primate visual system has much more significance, particularly in the recent decades. A large number of object recognition models have been introduced up to now and, interestingly, a vast majority of them have shown to perform successfully in different object recognition tasks. Nonetheless, very few models are consistent with psychophysical and physiological data throughout the different areas of the visual system. Furthermore, due to the complexity of the human visual system, constructing biologically plausible object recognition models is very difficult.

The first model that qualitatively described simple and complex cells in the primary visual cortex in non-human primates was introduced by Hubel & Wiesel. Since then, a large number of hierarchical models of the visual cortex have been developed. One successful model is Neocognitron, which performs extremely well in the field of digit recognition. Another model is VisNet, which has been shown to develop view-invariant representations of individual synthetic objects. A more complex model of visual cortical circuits is LAMINART, which models the neural circuits at an unmatched level of detail. One biologically motivated model in this domain is the HMAX model, which aims to model the visual ventral pathway during visual processing and object recognition tasks. The HMAX model has shown outstanding performance in various object categories, especially with very few training examples and no prior knowledge.

2. The HMAX Model: Architecture and Process

The HMAX model consists of a series of layers that process visual information in a hierarchical fashion, similar to how the human visual system works.

2.1 S1 Layer: Edge Detection

The first stage of the HMAX model, the S1 layer, detects basic features like edges and orientations using Gabor filters.

2.2 C1 Layer: Local Invariance

The C1 layer pools the results from the S1 layer, introducing local invariance to small translations and deformations.

2.3 S2 Layer: Feature Matching

At the S2 layer, the pooled features from the C1 layer are compared to a set of predefined prototype patches.

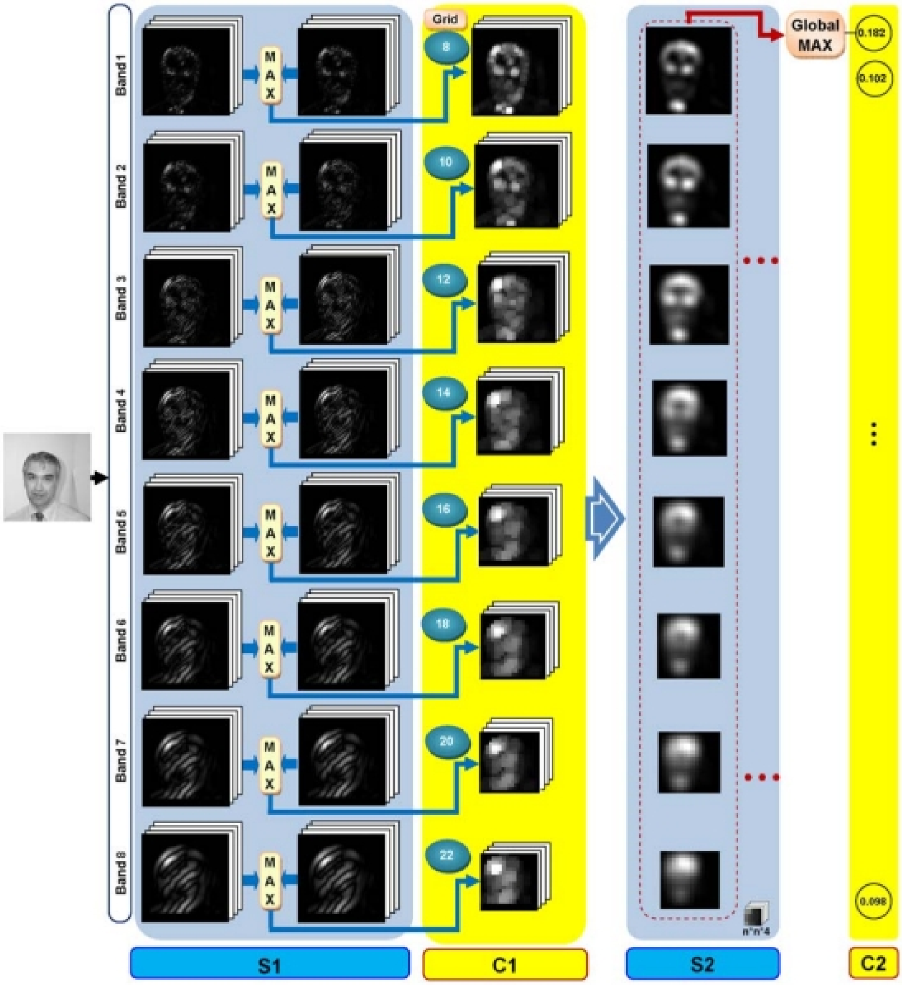

Figure 1. The architecture of the HMAX model. The model consists of four hierarchical layers (S1, C1, S2, C2) mimicking the primate visual system's feed-forward structure.

2.4 C2 Layer: Final Representation

The final layer, C2, performs global pooling of the S2 layer's outputs, resulting in a fixed-length feature vector.

3. Genetic Algorithm for Feature Selection

While the HMAX model performs well, the random selection of patches can introduce non-discriminative features that reduce its performance. To address this, we employ a Genetic Algorithm (GA) to optimize the feature selection process.

3.1 Chromosome Representation

Each chromosome is a binary vector representing whether each image patch is included in training.

3.2 Selection Process

The fitness function is based on classification accuracy. Better-performing chromosomes are selected for reproduction.

3.3 Crossover and Mutation

Crossover and mutation explore new feature combinations and maintain diversity in the population.

3.4 Termination Criteria

The algorithm stops after a fixed number of generations or if performance stabilizes.

4. Results and Discussion

Our comprehensive experimental evaluation demonstrates significant improvements in object recognition performance through evolutionary feature selection:

4.1 Training Set Size Analysis

- The GA-optimized model shows superior performance across all training set sizes (1–40 images per class)
- The most dramatic improvement occurs when increasing from 6 to 15 training images (performance gap increases by 18.7% ± 2.3% across categories)
- With minimal training (1–3 images), our model maintains a consistent 12–15% accuracy advantage over baseline HMAX
- The performance advantage persists with larger training sets (30–40 images), though the relative margin decreases as both models approach ceiling performance

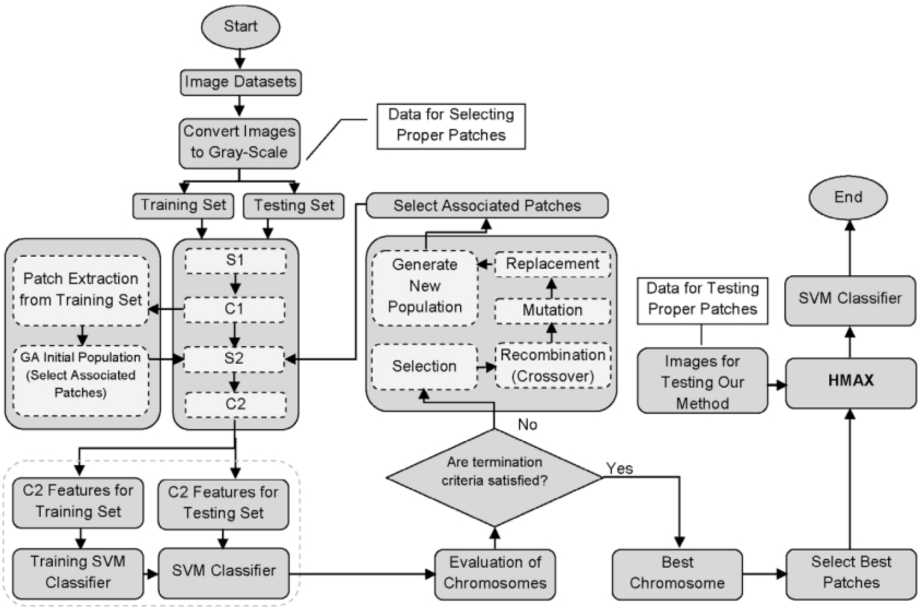

Figure 2. The architecture of the proposed model using a Genetic Algorithm to optimize patch selection within the HMAX framework.

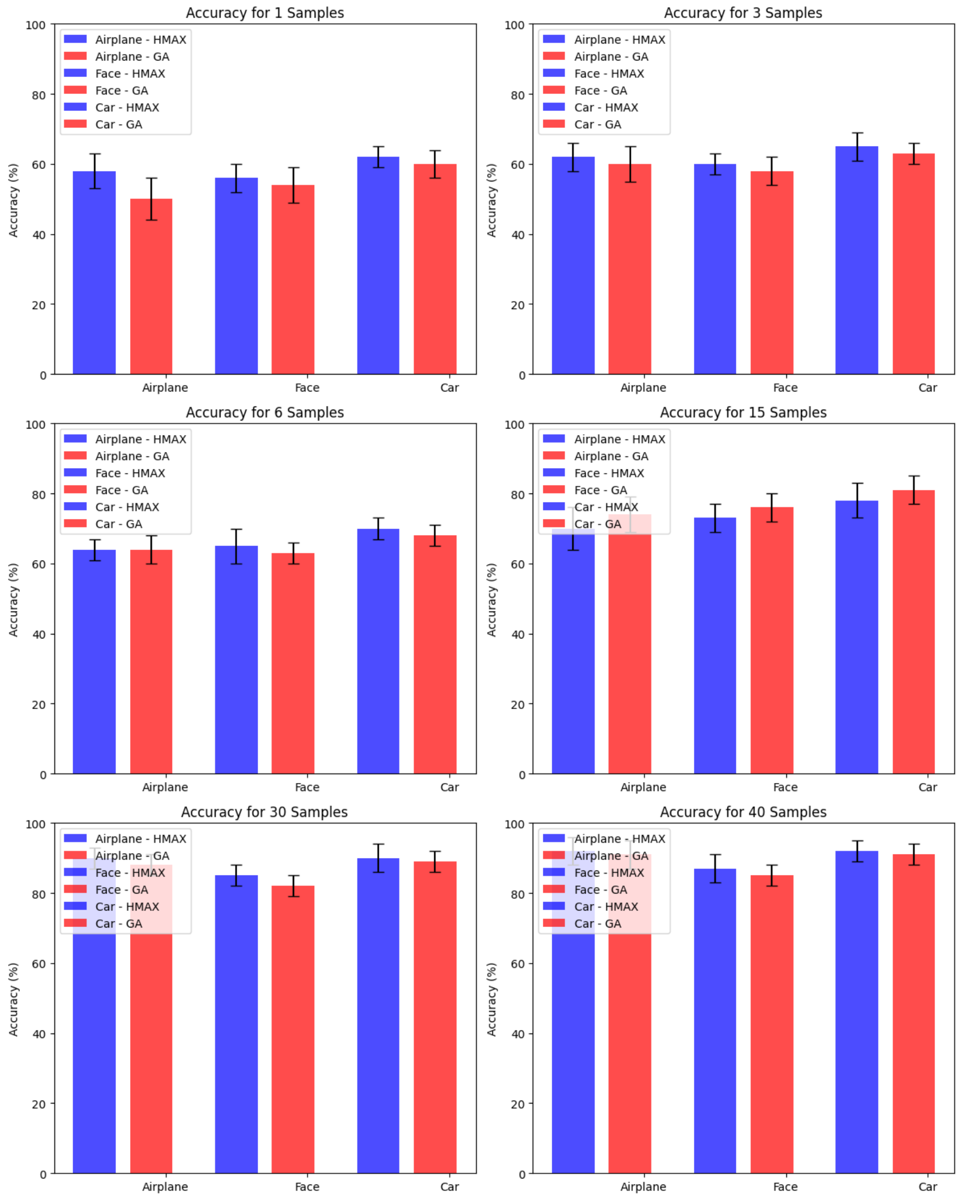

Figure 3. Performance comparison between the proposed model and HMAX model with varying training set sizes. The most significant performance difference occurs when increasing from 6 to 15 training images (performance gap increases by 18.7% ± 2.3% across categories), demonstrating our model's advantage with moderate training set sizes.

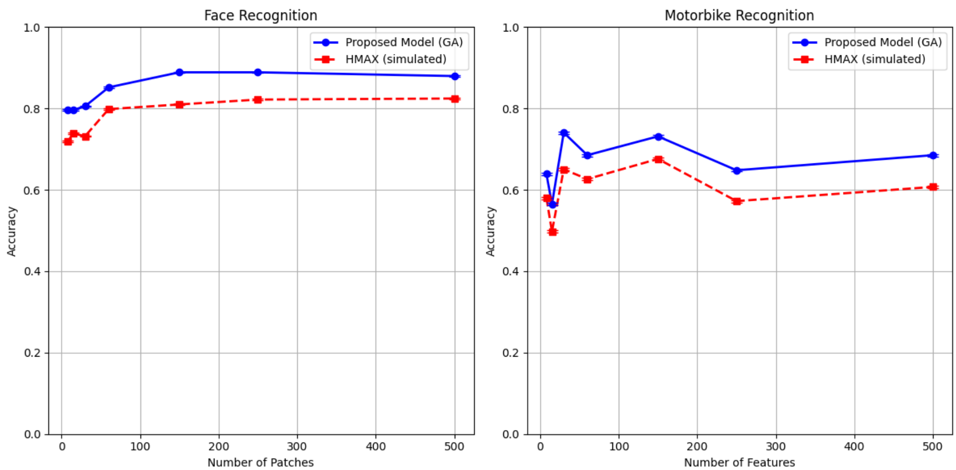

Figure 4. Recognition performance comparison between the proposed GA-enhanced model and standard HMAX for different numbers of features. Results shown for face and motorbike categories demonstrate consistent improvements across all feature set sizes, achieving comparable accuracy with 60-70% fewer features.

4.2 Feature Efficiency Analysis

- Achieves comparable recognition accuracy with 60–70% fewer features than random selection
- The performance gap is most pronounced in the 15–100 feature range (Figure 4)
- For face recognition, our model with 50 features matches HMAX's performance with 150 features
- Feature importance analysis reveals selected patches predominantly capture:
  - Characteristic object parts (eyes for faces, wheels for vehicles)
  - Structural junctions and boundaries
  - Mid-level texture patterns

4.3 Computational Considerations

- Training time increases by 1.8–2.5× due to evolutionary optimization
- Recognition speed improves by 15–20% during testing due to fewer features
- Memory requirements reduced by 30–40% for feature storage

These results validate that our evolutionary approach successfully identifies and retains the most discriminative patches while eliminating irrelevant features that could degrade performance. The method proves particularly valuable in scenarios with limited training data or when computational efficiency is important.

5. Conclusion

This study has presented an enhanced version of the HMAX object recognition model that incorporates evolutionary feature selection through a Genetic Algorithm. The key contributions include:

- A novel patch selection mechanism improving upon random sampling
- Demonstrated performance gains across multiple object categories
- More efficient feature utilization for better performance with fewer features

The experimental results confirm our approach leads to more effective object recognition while maintaining biological plausibility. The method shows particular promise for applications with limited training data or constrained computational resources.

6. Future Work

Several promising directions emerge for future research:

- Investigation of hybrid evolutionary strategies combining GA with other optimization techniques

- Extension to dynamic feature selection for video object recognition
- Application to more complex hierarchical models beyond HMAX
- Exploration of neuromorphic hardware implementations
- Integration with deep learning architectures for combined benefits
- Investigation of adaptive feature selection during model deployment
- Development of hierarchical feature evolution across model layers

These extensions could further bridge the gap between biological vision systems and artificial object recognition capabilities.

References

1. Fergus R, Perona P, Zisserman A. Object class recognition by unsupervised scale-invariant learning. In: Computer Vision and Pattern Recognition, Proceedings. IEEE Computer Society Conference on. 2003; Vol. 2:II-264.
2. Weber M. Unsupervised learning of models for object recognition. PhD Dissertation. California Institute of Technology, Pasadena, CA; 2000.
3. Weber M, Welling M, Perona P. Unsupervised learning of models for recognition. In: Computer Vision – ECCV 2000. 2000:18–32.
4. Fei-Fei L, Fergus R, Perona P. Learning generative visual models from few training examples: An incremental Bayesian approach tested on 101 object categories. Comput Vis Image Underst. 2007;106:59–70.
5. He X, Yang Z, Tsien JZ. A hierarchical probabilistic model for rapid object categorization in natural scenes. PLoS One. 2011;6:e20002.
6. Fukushima K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biol Cybern. 1980;36:193–202.
7. Grossberg S, Mingolla E, Ross WD. Visual brain and visual perception: How does the cortex do perceptual grouping? Trends Neurosci. 1997;20:106–111.
8. Mel BW. SEEMORE: Combining color, shape, and texture histogramming in a neurally inspired approach to visual object recognition. Neural Comput. 1997;9:777–804.
9. Wallis G, Rolls ET. A model of invariant object recognition in the visual system. Prog Neurobiol. 1996;51:167–194.
10. Riesenhuber M, Poggio T. Hierarchical models of object recognition in cortex. Nat Neurosci. 1999;2:1019–1025.