SvedbergOpen
DISSEMINATION OF KNOWLEDGE
International Journal of Artificial Intelligence and Machine Learning
Publisher's Home Page: https://www.svedbergopen.com/
Research Paper Open Access

Scalable Pattern Recognition In High-Dimensional Spaces Using Deep Learning Architectures

Rajashri CK[1], Shanthi R[2], Adars U[3], Vijayshree Khugshal[4], Vimal Bibhu[5], Dr. S. Balaji[6], Kiran Ingale[7], Dr. Ravi Kumar Sharma[8]

[1] Assistant Professor, Department of Computer Science, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, India. Email: rajashrick@maher.ac.in
[2] Assistant Professor, Department of Mathematics, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, India. Email: shanthir@maher.ac.in
[3] Research Scholar, Department of Electronics, VMKV Engineering College, Salem, Tamil Nadu, India. ORCID: 0009-0001-5905-1924
[4] Assistant Professor, Department of Computer Application, Dev Bhoomi Uttarakhand University, Dehradun, Uttarakhand, India. Email: socse.vijayshree@dbuu.ac.in ORCID: 0009-0002-0277-2154
[5] Professor, Department of Computer Science & Engineering, Noida International University, Greater Noida, Uttar Pradesh, India. Email: vimal.bhibu@niu.edu.in
[6] Professor, Department of Computer Science and Engineering, Panimalar Engineering College, Chennai, Tamil Nadu, India. Email: balajiit@gmail.com ORCID: 0000-0003-3208-1637
[7] Department of Electronics and Telecommunication Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra 411037, India. Email: kiran.ingale@vit.edu
[8] M.M. Institute of Computer Technology & Business Management, Maharishi Markandeshwar (Deemed to be University), Mullana – 133207, Haryana, India.

Abstract

Many problems in science and engineering can be formulated as geometric pattern recognition tasks in high-dimensional spaces. However, traditional convolution-based approaches are often unable to represent long-term dependencies and scale efficiently with increasing dimensionality. To address these drawbacks, this research proposes a scalable pattern recognition framework based on deep learning (DL) for high-dimensional data analysis. The proposed approach utilizes a High-Dimensional Dataset containing geometric pattern records extracted from image-based visual pattern data. During the preprocessing stage, noise handling is identified and handled appropriately, followed by Min-Max normalization to scale feature values and improve model convergence. Subsequently, Principal Component Analysis (PCA) is used for feature extraction and dimensionality reduction purposes. The proposed approach is first evaluated on synthetic datasets involving linear subspace detection in high-dimensional spaces demonstrating its capability to effectively model complex geometric structures. Furthermore, the Ebola Optimization Search-driven Adaptive Vision Transformer (EOS-Adaptive VT) uses AVT to learn long-range spatial and contextual dependencies in high-dimensional pattern data, and EOS adaptively explores and exploits the solution space, including 3D geometric registration under rigid transformations and image correspondence estimation. Experimental results shows that the proposed framework consistently performs with 97.24% accuracy, 97.02% precision, and 96.84% recall using Python 3.11. The findings highlight the effectiveness of transformer-based architectures for scalable pattern recognition, particularly in handling complex and high-dimensional datasets. Overall, this research shows the advancement of efficient, scalable, and high-performance pattern recognition systems with applications in computer vision, robotics, and high-dimensional scientific data analysis.

Keywords: Scalable pattern recognition, High-dimensional spaces, Deep learning, Geometric pattern analysis, 3D geometric registration, Image correspondence estimation, High-dimensional data analysis.

1. Introduction

Classification of high-dimensional data is an extensively used technique in scientific computing to organize data effectively. However, current techniques suffer from scalability problems, and reduced accuracy with large feature spaces [1]. The identification of patterns in physical systems is important for gaining insights into dynamical nucleation phenomena and self-organizing structures within non-equilibrium settings. Nevertheless, present-day methodologies encounter difficulties in accounting for subtle changes in time dynamics, structure, and pattern recognition [2]. Biological processes have developed highly sophisticated pattern recognition capabilities that enable survival within varied environments by contextually matching visual and behavioral stimuli. The problems faced include complexities associated with modeling dynamic transformations of patterns in contextual settings [3]. Network system intrusion detection plays an important role in the identification of cyber attacks based on detecting abnormal patterns of traffic. However, currently available techniques fail to address the issues such as data imbalance, high dimensionality, and evolving attacks [4]. The high-dimensional data analysis techniques are popular in data mining and knowledge discovery because it analyzes high-volume and complicated data sets with many redundant and correlated attributes. The problems that were encountered with it include issues with dimensionality reduction, information loss, and decreased interpretability [5]. Learning and prediction in high-dimensional space are crucial in modeling complex relations in large structured data with significant dependence on locally distributed features. The existing challenges are computation, generalization issues, and learning spatially distributed relations [6]. Fast searching and indexing in high-dimensional spaces are essential for information retrieval applications that need to perform similarity searches efficiently on a large-scale dataset. Current problems in this field include the curse of dimensionality, high computational costs, and the speed versus accuracy dilemma [7].

1.1 Problem Statement: The increasing presence of high-dimensional data within various scientific and engineering domains creates serious problems for scalable pattern recognition because traditional convolution-based methods usually find it difficult to capture long-range dependencies. To address these drawbacks, this research proposes a scalable Deep Learning (DL) framework that integrates Ebola Optimization Search (EOS)-driven Adaptive Vision Transformer (AVT) used for efficient high-dimensional feature learning, global-local relationship modeling, and robust pattern recognition.

1.2 Research Organization: Section 1 provides the introduction to scalable pattern recognition, including background, motivations, aims, and challenges. Section 2 summarizes previous research on transformer models. In Section 3, the developed EOS-AVT is introduced and formulated. Section 4 explains the experimental settings, including data sets and implementation. The analysis of the results is discussed in section 5. The discussion on the results, benefits and limitations is conducted in section 6. Finally, section 7 establishes the research and suggests further paths for future research.

2. Related Work

The current pattern recognition techniques have shown great results in terms of analyzing data with many dimensions and learning geometric features is summarized in Table 1. However, there are certain shortcomings associated with long-range dependencies and scalability when dealing with higher dimensions.

Table 1: Summary of current methods on high-dimensional pattern recognition approaches.

Ref	Objective	Method	Results	Limitations
[8]	Improved scalable similarity search in high-dimensional datasets for data science applications	The Graph-based Efficient Large-scale Point Indexing System (ELPIS) was developed for efficient nearest-neighbor retrieval.	The model demonstrated improved scalability and faster query response time compared to traditional graph indexing methods.	The approach was sensitive to graph construction parameters and showed reduced efficiency in extremely sparse datasets.
[9]	Detection of anomalies in high-dimensional spaces	Deep Hypersphere Network (DHN) was	The model achieved improved anomaly detection performance	The model struggled with interpretability and required careful

	with improved robustness.	employed to detect anomalies.	over baseline DL models.	tuning of hypersphere radius parameters.
[10]	Focused on retinal disease classification with improved explainability.	An Explainable Fully Dense Fusion Neural Network (E-FDFNN) r was used to enhance feature fusion.	The model showed better classification performance compared to traditional methods.	The architecture was computationally expensive and less suitable for real-time deployment.
[11]	Improved satellite image classification accuracy using deep learning.	A Vision Transformer (ViT)-based DL Model was used for remote sensing image classification.	The model achieved high classification performance and outperformed baselines.	The model required large-scale training data and high computational resources.
[12]	Developed lightweight models for prediction and classification.	A Lightweight Transformer Edge Intelligence Model (LTEIM) was introduced, using a compressed Transformer Encoder architecture optimized for edge devices.	The model showed improved efficiency and competitive predictive performance on industrial datasets.	Limited generalization was observed across different industrial operating conditions.
[13]	Focused on feature extraction and prognostics in machinery health monitoring.	A Transformer Self-Attention Transfer Network (TSATN) was used for high-dimensional time-series data.	The model improved feature representation and prognostic accuracy compared to traditional approaches.	The model performance decreased under noisy sensor conditions and required large labeled datasets.
[14]	Improved high-dimensional classification using ensemble transformer architectures.	Transformer-based Attention-guided Deep Ensemble Network (Trans-ADENet) is used for ensemble DL classification problems.	The method achieved strong classification performance across benchmark datasets.	The ensemble structure increased computational complexity and training time.
[15]	Improved classification accuracy and reduced computational complexity in high-dimensional medical datasets.	Maximum Pattern Recognition (MPR) was used for identifying non-linear dependencies between the attributes and also filtered out irrelevant attributes.	The experimental results conducted using medical databases showed good improvement in terms of performance measures.	Performance varied based on the tuning of the parameters used for the evolutionary algorithm.

2.1 Research Gap

The current state-of-the-art high-dimensional pattern recognition approaches have various problems, like improper modeling of dependencies between features that are far apart, poor incorporation of local and global inter-relationships between features, high computational cost, and poor scalability as the number of dimensions increases [8–14]. To solve these problems, the present approach uses AVT by incorporating EOS-based optimization, re-attention, and adaptive feature learning to accomplish better feature representation and geometric relationship modeling.

3. Proposed Methodology

A scalable framework for pattern recognition in high dimensions was created, which can effectively learn features, analyze geometric patterns. The scalable framework developed in this research is based on preprocessing high-dimensional data sets with data normalization, feature scaling, and noise reduction is shown in Figure 1. Next, it involves the use of EOS in conjunction with AVT to perform feature representation learning, and parameter optimization, and achieve higher levels of recognition accuracy.

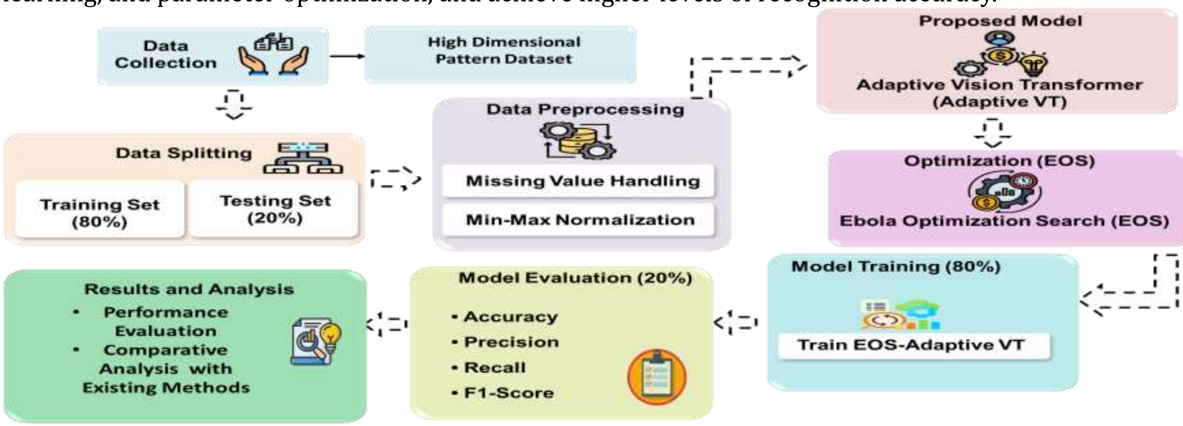

Figure 1: Overall Workflow of the EOS-AVT Framework for High-Dimensional Pattern Recognition

3.1 Data Collection

The High-Dimensional Pattern Dataset consists of 13,000 records, and it has 57 columns related to the structure of geometric pattern shapes, noise properties, and pattern behavior at the high-dimensional level. This dataset is composed of various numerical and structural features related to geometric patterns, both linear and nonlinear, subspace relationships, correspondence behavior, and transformation properties. In experimental evaluation, the data set is partitioned into the 80:20 ratio, which involves using 80 percent for training and 20 percent for testing. Source: https://www.kaggle.com/datasets/colabsss/high-dimensional-pattern-dataset

3.2 Data Pre-processing

This is done on High-Dimensional Pattern Dataset to make sure that there is high-quality data used in pattern recognition. The process begins by identifying any noise, Outlier handling in the dataset, and this is done to ensure that there is no data distortion in the database while minimizing bias. The next step after handling the missing values is Min-Max normalization for scaling down all features.

➢ **Noise Handling and Outlier handling for Robust High-Dimensional Feature Representation**

Management of noise and outliers improves the quality of feature spaces through the elimination of distortions that may arise due to noises in the high-dimensional data, and thus making the process of learning features easier. This will increase the capabilities of EOS-AVT to learn geometric complexities, together with building resilient models.

➢ **Min-Max Normalization for Feature Scaling**

Following data pre-processing on missing values, the obtained feature vectors were subjected to feature scaling to ensure that each feature plays an equally significant part in building the model. The various features differ in numerical values and are scaled down through Min-Max normalization to the range of 0 - 1.

$$\psi_{p,q} = \left[\frac{F_{p,q} - min(F_q)}{max(F_q) - min(F_q)} \times N\right] \tag{1}$$

In equation (1) below, $\psi_{p,}$ q is the normalized and quantized value of the q^{th} $the\ feature\ corresponding\ to\ the$ p^{th} sample. $F_{p,q}$ stands for the intitial value of the q^{th} feature for the p^{th} sample. $min(F_q)$ and $max(F_q)$ are the minimum and maximum values of the q^{th} feature across all samples. Where N is the scaling factor that defines the number of quantization levels. However, the output of the Min-Max normalization process was a set of normalized feature vectors that improved feature consistency and enhanced the effectiveness of subsequent feature learning and pattern recognition processes.

3.3 Principal Component Analysis (PCA) for High-Dimensional Feature Extraction

PCA comprises a normalized high-dimensional matrix of features wherein each sample consists of a number of inter-related features from the dataset. It is used to convert the feature space into a lower- dimension orthogonal space while retaining maximum variance within the data set.

$$\boldsymbol{A} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} \tag{2}$$

$$\boldsymbol{G}_{(p,p)} = \begin{bmatrix} g_{1,1} & \cdots & g_{1,p} \\ \vdots & \ddots & \vdots \\ g_{p,1} & \cdots & g_{p,p} \end{bmatrix} \quad (3)$$

$$\sigma_{g,r} = \frac{1}{M-1}\sum_{t=1}^{M}(a_{t,r} - \mu_r)(a_{t,s} - \mu_s) \quad (4)$$

$$det(G - \lambda I) = 0 \quad (5)$$

Equation (2-5), A denotes the high-dimensional input feature matrix. $a_1, a_2, \ldots, a_p$ represents individual feature vectors extracted from raw data, and p denotes the total number of features in the input. Where $\boldsymbol{G}_{(p,p)}$ represents the covariance matrix capturing feature dependencies. The diagonal elements $g_{1,1}, g_{p,p}$ represent the variances of individual features, while off-diagonal elements represent correlations among features. $\sigma_{g,}$ r denotes the covariance between feature dimensions g and r. M is the sample number, where t is the sample index, a_t and at are feature values, and μ_r *and* μ_s represent the corresponding mean values. G denotes the covariance matrix computed from high-dimensional feature representations, λ represents the eigenvalues derived from the covariance matrix, and I denotes the identity matrix used to ensure stability in the eigenvalue decomposition process during feature space transformation and dimensional structure analysis. However, the applications of PCA reduced dimensionality of the data while increasing computational efficiency and enhances effectiveness of further analysis.

3.4 EOS-Driven AVT for Scalable High-Dimensional Pattern Recognition

The proposed methodology combines the AVT and EOS algorithms to enable scalable and robust high-dimensional pattern recognition. The AVT captures both global dependencies re-attention, and an enhanced feed-forward mechanism where EOS, refines model parameters and feature embeddings by balancing exploration.

3.4.1 AVT for High-Dimensional Pattern Recognition

The Vision Transformer is used to extract relationships among the input tokens through the self-attention mechanism. The Vision Transformer suffers from limited local dependency modeling and reduced attention diversity, as multiple attention heads may generate highly similar attention maps during training. To address the drawback, the AVT includes the Re-Attention improves attention map diversity by allowing the attention heads to interact, and the Feed-Forward Network improves feature extraction.

The AVT includes high-dimensional representations of features from the input data. It is used to capture the global dependencies and re-attention, which helps capture attention in different forms; and a feed-forward network with depth wise convolution to capture local dependencies effectively.

- **Re-Attention**

The Re-Attention mechanism is employed to enhance the diversity of attention maps by enabling interactions among multiple attention heads, thereby reducing attention collapse and improving global feature relationship learning.

$$\text{Re-Attention}(A, B, C) = \text{Norm}\left(\Theta^T\left(\text{softmax}\left(\frac{AB^T}{\sqrt{m}}\right)\right)\right)C \quad (6)$$

In equation (6), A denotes the query matrix representing high-dimensional feature embeddings. B denotes the key matrix used to measure relationships among feature tokens. C denotes the value matrix containing contextual feature information. m represents the dimensionality of the feature embedding space used for attention normalization. Θ denotes the learnable attention transformation matrix. Θ^T represents the transpose of the attention transformation matrix. AB^T denotes the similarity matrix computed between query and key feature representations. B^T represents the transpose of the key matrix.

Feed-Forward

The Feed-Forward Network is employed to further transform and refine the extracted feature representations after the attention stage. This allows the model to be able to better capture complex structures, hence allowing for improved pattern recognition.

$$T_f = \text{Seq2Img}(T), T_f \in \mathbb{R}^{p\times q\times c} \quad (7)$$

Equation (7) shows T denotes the sequence of feature tokens generated from the input high-dimensional data. Seq2Img represents the transformation function that converts token sequences into structured feature maps. T_f denotes the intermediate feature map representation obtained from the token sequence. $\mathbb{R}$ represents the real-valued feature space of dimension $p \times q \times c$, dimensional feature tensor, where p and q orrespond to the height and width of the feature map. c indicates the number of feature channels.

$$M_f = g(g(T_f \circledast P_1) \circledast P_d) \circledast P_2 \quad (8)$$

In equation (8), M_f denotes the refined feature map containing enhanced feature representations. g represents the nonlinear activation function. T_f denotes the intermediate feature map obtained from the token sequence. $\circledast$ denotes the feature transformation operation and learnable projection matrices. P_1 denotes the first learnable projection matrix used to transform and expand feature representations. P_d represents the intermediate feature refinement matrix responsible for enhancing feature interactions and extracting meaningful patterns. P_2 denotes the final learnable projection matrix.

$$S = \text{Img2Seq}(M_f) \quad (9)$$

In equation (9), S denotes the output sequence of refined feature tokens generated after feature enhancement. $\text{Img2Seq}(\cdot)$ represents the transformation function that converts a feature map into a sequence of feature tokens. M_f denotes the refined feature map containing enhanced local and global feature information. $\text{Img2Seq}(M_f)$ represents the process of reshaping the refined feature map. However, the contextual feature embeddings capture both local and global information from the input data.

3.4.2 EOS for High-Dimensional Feature Optimization in Scalable Pattern Recognition

After obtaining refined feature representations from the AVT, EOS involves control variables such as the neighborhood threshold, short move rate, long move rate, displacement scaling factor, and randomness factors. This technique is adopted as an optimization strategy to boost the performance of the developed scalable pattern recognition system for high-dimensional settings.

$$X_i^{t+1} = X_i^t + p\,V(X) \quad (10)$$

Equation (10) shows X_i^{t+1} denotes the updated position, where i^{th} individual at iteration $t+1$. X_i^t denotes the current position of the i^{th} individual at iteration t. p represents the displacement scaling factor. $V(X)$ denotes the movement rate of the individual.

$$V(X) = srate \times rand(0,1) + V(Indbest) \quad (11)$$

$$V(Y) = lrate \times rand(0,1) + V(Indbest) \quad (12)$$

In equation (11-12) $V(X)$ denotes the feature-space movement rate of a candidate solution during the optimization process. $V(Y)$ denotes the movement rate of an alternative candidate solution group used to enhance search diversity. $V(\text{Ind}_{best})$ represents the movement rate in the high-dimensional search space. $srate$ denotes the local search rate that controls fine-grained exploration around promising feature representations. $lrate$ represents the high-dimensional feature space. $\text{rand}(0,1)$ represents the uniformly distributed random number. However, outcome generated by EOS includes a set of solution vectors with the optimized value for model parameters and feature embedding to obtain maximum optimization of the fitness function in pattern recognition.

The framework uses AVT-based feature extraction and then EOS-based optimization to generate high-dimensional embeddings based on global self-attention, Re-Attention, and feed-forward processing, along with feature aggregation using local-global features. EOS is subsequential employed to optimize both feature embeddings and model parameters by exploiting its exploration and exploitation capabilities to generate globally optimal solutions.

4. Results

The proposed scalable pattern recognition framework for high-dimensional space was successfully implemented by the EOS-AVT. Experiments were carried out on a high-performance workstation fitted 32 GB of DDR5 RAM, with an Intel Core i7-14700K processor, and an NVIDIA RTX 4070 Graphics Processing Unit (GPU) with 12 GB of VRAM. Implementation of the proposed pattern recognition framework made use of the following technologies: Python 3.11, TensorFlow 2.16, Keras 3.0, NumPy, Pandas, Scikit-learn, and Matplotlib. This section offers the findings of Evaluation Metrics, is illustrated in Table 2, Experimental Analysis, 5-fold cross-validation, Comparative Performance, and Computational Analysis.

Table 2: Performance Metrics Used for Evaluating the Proposed EOS-AVT

Metrics	Explanation
Accuracy (%)	Measures the percentage of correct classifications made by the classifier out of all the samples.

Precision (%)	Calculates the percentage of positive predictions made correctly out of all positive predictions.
Recall (%)	Calculates the percentage of positive class values out of all the samples that are predicted correctly by the model.
F1-Score (%)	Measures using the Harmonic Mean between Precision and Recall for a more balanced measure.
Training Time (s)	Estimates the time taken to train the proposed EOS-AVT model using the training data set.
Inference Time (ms)	defines the time taken by the trained model to make predictions for test inputs.

4.1 Experimental analysis

This analysis is used to examine the pairwise relationship between high-dimensional geometric features and the learned scoring functions. This facilitates the understanding of feature dependencies, clustering behaviors, and feature separation, is shown in Figure 2.

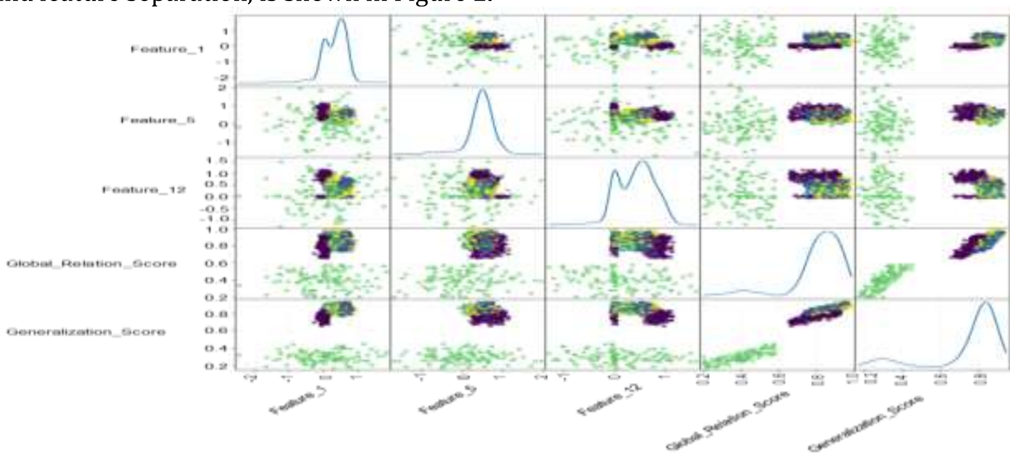

Figure 2: Pairwise relationships of high-dimensional geometric features and scores

The different levels of relationship between the features, which means that a highly related combination is a sign of high dependency is presented in Figure 2. Feature 1, Feature 5, Feature 12, Global Relation Score, and Generalization Score have different patterns, implying that there is stable representation learning in the proposed model due to consistent variability.

The distribution of the values for each variable using the density plot. The figure is used to examine the distribution of data, variability between variables, and possible skewness of the distribution prior to modelling is illustrated is figure 3.

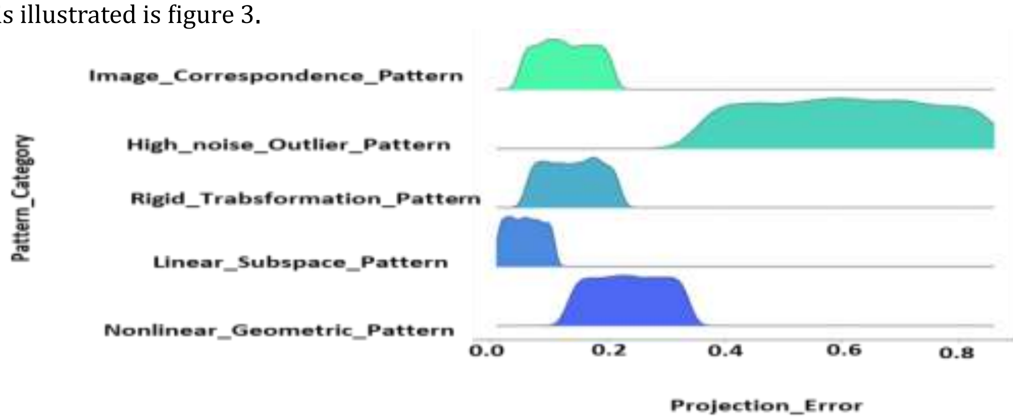

Figure 3: Density-Based Visualization of Feature Distributions

The horizontal density curves correspond to the probability distribution of a particular feature. The peaks correspond to the area where there is high concentration of values for the particular feature, while the range covered by each curve reflects the level of variation within the feature.

4.2 5-fold cross-validation

The 5-fold cross-validation consists of splitting the data into five equal sets. Four of the sets are used for training, and one set is used for testing at all times. The procedure is carried out five times, and eventually, the results' mean and standard deviations are determined after the fifth run, as presented in Table 3.

Table 3: 5-fold Cross-validation performance of the proposed EOS-AVT model

Fold	Training Time (s)	Inference Time (ms)	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
1	42.3	3.8	96.8	96.5	96.3	96.4
2	41.7	3.6	97.2	97.0	96.8	96.9
3	40.9	3.5	97.5	97.3	97.1	97.2
4	42.8	3.9	96.9	96.7	96.6	96.6
5	40.2	3.4	97.8	97.6	97.4	97.5
Average (Mean ± SD)	41.58 ± 0.95	3.64 ± 0.18	97.24 ± 0.39	97.02 ± 0.41	96.84 ± 0.42	96.92 ± 0.40
95% CI	[40.62, 42.54]	[3.46, 3.82]	[96.86, 97.62]	[96.62, 97.42]	[96.44, 97.24]	[96.53, 97.31]

4.3 Comparative Performance and Computational Analysis

The comparative analysis of the proposed EOS-driven AVT is performed by evaluating its performance against existing methods, including the Geometric Algebra Graph Attention Network (GGAN) [16] as the baseline approach. This experiment involves training and validating different models on the same high-dimensional dataset with identical data partitioning and experimental settings. The comparative results in terms of pattern recognition accuracy and geometric feature learning are summarized in Table 4. The proposed method demonstrates better performance compared to GGAN, achieving 97.8% accuracy, 97.02% precision, 96.84% recall, and 96.92% F1-score, indicating highly effective high-dimensional pattern recognition capability. In addition, a training time of 41.58 seconds and an inference time of 3.64 ms/sample demonstrate strong computational efficiency and scalability for geometric pattern recognition tasks.

Table 4: Comparative Performance and Computational Analysis

Method	Training Time (s)	Inference Time (ms/sample)	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Graph Geometric Algebra Network (GGAN)	48.20	4.20	94.12	93.85	93.76	93.80
EOS-AVT [Proposed]	41.58	3.64	97.24	97.02	96.84	96.92

5. Discussion

The developed EOS-AVT framework for large-scale pattern recognition has shown excellent results in providing precise, resilient, and explainable decisions. Previous approaches that utilized the Geometric Algebra Graph Attention Network (GGAN) [16] suffered from certain problems related to scalability issues, the capacity of feature extraction, and adequate consideration of complex geometrical dependencies. Furthermore, previous models lacked mechanisms for global/local feature learning, optimizing the model's parameters, and enriching the model with attention diversity capabilities. However, the presented EOS-AVT framework manages to solve these issues by introducing more expressive feature embedding, optimizing high-dimensional features, and enhancing pattern distinction.

6. Conclusion

The EOS-AVT framework enable scalable, accurate, and robust pattern recognition in high-dimensional spaces, ensuring by efficient feature learning and strong generalization capability. It utilizes a high-dimensional pattern dataset with preprocessing steps including normalization, feature scaling, and missing-value handling, while

EOS- AVT enhances representation quality and predictive performance. The proposed framework demonstrates strong results in precision (97.02%), recall (96.84%), accuracy (97.24%), F1-score (96.92%), training time (41.58), and inference time (3.64). However, the framework was sensitive to hyperparameter tuning and variations in data distribution, while future work will focus on real-time high-dimensional pattern monitoring, cross-domain adaptability, and scalable deployment of transformer-based optimization models in dynamic intelligent systems.

Reference

[1] Cai, X., Chan, R.H., Xie, X. and Zeng, T., 2024. An Efficient and Versatile Variational Method for High-Dimensional Data Classification. Journal of Scientific Computing, 100(3), p.81. https://doi.org/10.1007/s10915-024-02644-9

[2] Evans, C.G., O'Brien, J., Winfree, E. and Murugan, A., 2024. Pattern recognition in the nucleation kinetics of non-equilibrium self-assembly. Nature, 625(7995), pp.500-507. https://doi.org/10.1038/s41586-023-06890-z

[3] Woo, T., Liang, X., Evans, D.A., Fernandez, O., Kretschmer, F., Reiter, S. and Laurent, G., 2023. The dynamics of pattern matching in camouflaging cuttlefish. Nature, 619(7968), pp.122-128. https://doi.org/10.1038/s41586-023-06259-2

[4] Gul, S., Arshad, S., Saeed, S.M.U., Akram, A. and Azam, M.A., 2024. WGAN-DL-IDS: an efficient framework for an intrusion detection system using WGAN, random forest, and deep learning approaches. Computers, 14(1), p.4. https://doi.org/10.3390/computers14010004

[5] Huang, Z., Kabán, A. and Reeve, H., 2024. Efficient learning with projected histograms: Z. Huang et al. Data Mining and Knowledge Discovery, 38(6), pp.3948-4000. https://doi.org/10.1007/s10618-024-01063-6

[6] Stöckl, C., Yang, Y. and Maass, W., 2024. Local prediction-learning in high-dimensional spaces enables neural networks to plan. Nature Communications, 15(1), p.2344. https://doi.org/10.1038/s41467-024-46586-0

[7] Zhao, X., Tian, Y., Huang, K., Zheng, B. and Zhou, X., 2023. Towards efficient index construction and approximate nearest neighbor search in high-dimensional spaces. Proceedings of the VLDB Endowment, 16(8), pp.1979-1991. https://doi.org/10.14778/3594512.3594527

[8] Azizi, I., Echihabi, K. and Palpanas, T., 2023. Elpis: Graph-based similarity search for scalable data science. Proceedings of the VLDB Endowment, 16(6), pp.1548-1559. https://doi.org/10.14778/3583140.3583166

[9] Zheng, J., Li, J., Liu, C., Wang, J., Li, J. and Liu, H., 2022. Anomaly detection for high-dimensional space using a deep hypersphere fused with a probability approach. Complex & Intelligent Systems, 8(5), pp.4205-4220. https://doi.org/10.1007/s40747-022-00695-9

[10] Kayadibi, I. and Güraksın, G.E., 2023. An explainable, fully dense fusion neural network with a deep support vector machine for retinal disease determination. International Journal of Computational Intelligence Systems, 16(1), p.28. https://doi.org/10.1007/s44196-023-00210-z

[11] Adegun, A., Viriri, S. and Tapamo, J.R., 2024. Automated classification of remote sensing satellite images using deep learning-based vision transformer: A. Adegun et al. Applied Intelligence, 54(24), pp.13018-13037. https://doi.org/10.1007/s10489-024-05818-y

[12] Wang, L., Li, Y., Liu, H. and Liu, T., 2025. A lightweight transformer edge intelligence model for rule prediction classification. Sensors, 25(13), p.4224. https://doi.org/10.3390/s25134224

[13] Sun, S., Peng, T. and Huang, H., 2023. Machinery prognostics and high-dimensional data feature extraction based on a transformer self-attention transfer network. Sensors, 23(22), p.9190. https://doi.org/10.3390/s23229190

[14] Bhimineni, V.C. and Senapati, R., 2025. Trans-ADENet: Transformer-based Attention-guided Deep Ensemble Network for high-dimensional data classification. Array, 29 (2026), p.100659. https://doi.org/10.1016/j.array.2025.100659

[15] Nematzadeh, H., García-Nieto, J., Aldana-Montes, J.F. and Navas-Delgado, I., 2024. Pattern recognition frequency-based feature selection with multi-objective discrete evolution strategy for high-dimensional medical datasets. Expert Systems with Applications, 249, (2024) p.123521. https://doi.org/10.1016/j.eswa.2024.123521

[16] Zhong, J. and Cao, W., 2025. Graph Geometric Algebra networks for graph representation learning. Scientific Reports, 15(1), p.170. https://doi.org/10.1038/s41598-024-84483-0