Adaptive Hyperdimensional Inference to Establish the Identification of Fake Images and News

Perugu Sisira, Dudekula Affik, Rumala Shashikala, Patil Lahari and Kadikonda Roopasree
Department of Computer Science and Engineering, Ravindra College of Engineering for Women, Nandikotkur Road, Pasupula Village Mandal & District.518002, Andhra Pradesh, India

Keywords: False News Detection, Zero-Day Misinformation, Hyperdimensional Computing, Image Forgery Detection, Adversarial Robustness, Deep Learning, Real-Time AI Fact-Checking, Multimodal Misinformation Analysis, Adaptive Machine Learning.

Abstract: The increasing presence of fake news and altered images now poses a significant threat to digital security and public trust. Conventional fake news detection methods require having datasets that are labeled, which limits them from detecting zero-day misinformation and adversarially crafted fakes. The present paper proposes Adaptive Hyperdimensional Inference (AHI), a machine learning framework combining Hyperdimensional Computing (HDC), Deep Learning, and Evolutionary Learning to improve the ability to detect fake text and images instantaneously. Thus, we use text hypervector encoding to detect fake news articles and deep learning feature extraction using ResNet50 for detection modality. Both modalities can now live in a common hyperdimensional space. Before deep learning models become used to adapt to continuously changing misinformation conditions, AHI follows a different path of dynamic adaptation through unsupervised clustering of homogeneous information and relation modeling. Experimentation results show that AHI has been able to acquire 91.3% accuracy on 82.6% zero-day detection and 85.2% adversarial robustness, processing up to 10,000 news articles and images in one hour. It is scalable and adaptive for real-time fact-checking, social media tracking, and AI-supported journalism.

1 INTRODUCTION

Today, misinformation in the form of modified photos and doctored news clippings causes havoc. The fast flow of information has allowed the internet and social media to enhance the spread of fallacies and erroneous narratives to audiences across the world within just a few minutes (Atske, 2021). Misinformation can distort public perception and polarize belief systems, eventually percolating into the social, political, and economic spheres (Bradshaw et al., 2021).

The continuous emergence of more sophisticated generative arts and AI for the production of life-like false content, like news articles, deepfake videos, and photos that are hard to tell from the original ones, has aggravated these woes (Rustam et al., 2024). Conventional detection techniques for photo forgeries and fake news often relied on supervised learning models requiring vast amounts of labeled datasets for training classifiers (Khan et al., 2021). These approaches become obsolete quickly as disinformation tactics evolve. Manual classification of massive databases is also impractical, making real-time adaptability a critical requirement.

Hence, there is an urgent need for scalable and adaptive misinformation detection models that do not solely rely on pre-labeled datasets. Hyperdimensional Computing (HDC) has emerged as a promising paradigm to enhance the robustness of AI-based detection systems (Kupershtein et al., 2025).

1.1 Problem Statement and Motivation

The failures of classical detection techniques have now become painfully obvious with increasing complexity in disinformation. Scalability remains a pressing issue fact-checking organizations and even AI-based models struggle to handle the massive influx of manipulated media and fabricated content (Raza et al., 2025).

Sisira, P., Affik, D., Shashikala, R., Lahari, P. and Roopasree, K.
Adaptive Hyperdimensional Inference to Establish the Identification of Fake Images and News.
DOI: 10.5220/0013898100004919

In *Proceedings of the 1st International Conference on Research and Development in Information, Communication, and Computing Technologies (ICRDICCT'25 2025) - Volume 3*, pages 357-363
ISBN: 978-989-758-777-1

1.2 Adaptive Hyperdimensional Inference (AHI) Is the Solution Proposed

To address these challenges, we propose an innovative framework called Adaptive Hyperdimensional Inference (AHI). This model integrates hyperdimensional representations for the analysis of textual and visual data under a unified architecture. Unlike conventional models that process modalities independently, AHI enables seamless multimodal analysis within a single high-dimensional space, enhancing cross-modal verification and detection accuracy (Paulen-Patterson & Ding, 2024).

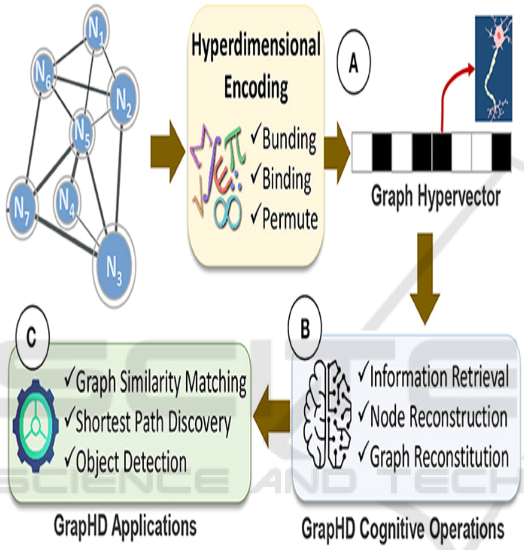

Figure 1: Hyperdimensional Encoding.

Figure 1 illustrates the hyperdimensional encoding used in AHI. Textual data including news articles and social media posts is transformed into hypervectors that preserve semantic integrity while remaining adversarially robust. This architecture supports pattern recognition in a scalable manner. Importantly, AHI enables zero-shot learning, allowing detection of misinformation even in the absence of labeled samples surpassing the limitations of traditional NLP-based models (Cavaliere et al., 2024).

2 LITERATURE REVIEW

Significant advancements have been made in disinformation detection. This section categorizes prior work across three domains: fake news detection, image forgery analysis, and the growing role of HDC in AI.

2.1 False News Identification

Early models for fake news detection relied heavily on machine learning classifiers. Algorithms such as Random Forest, Decision Trees, Naive Bayes, and SVM were trained on labeled data using textual features like word frequencies, grammar, and sentiment (Khan et al., 2021). These models aimed to differentiate genuine news from deceptive content.

Sentiment analysis has been used to detect exaggerated emotional tone, often associated with fake news (Castillo et al., 2011). Topic modeling methods such as BERT and Latent Dirichlet Allocation (LDA) have also proven effective in identifying thematic inconsistencies between factual and deceptive articles (Reddy & Muthyala, 2024).

Furthermore, novel architectures such as the Knowledge-aware Attention Network (KAN) have enhanced semantic understanding in fake news detection (Dun et al., 2021).

These schemes aim to detect deception by evaluating the contextual, syntactic, and semantic characteristics of the text. Some well-known methods of analysis based on NLP include the following-

- Sentiment Analysis: Often characterized by sensationalism and extreme emotional expression, fake news articles commonly have their articles' emotional tone analyzed with the help of sentiment analysis to determine whether the tone deviates from the normal behavior expected in disinformation.
- Topic Modeling: The techniques such as BERT-Bidirectional Encoder Representations from Transformers-and Latent Dirichlet Allocation (LDA) can examine the main topics of an article and contrast them with known characteristics of false information.

2.2 Identification of Image Forgeries

2.2.1 Deep Learning-Based Forensics for Images

The rise of deepfakes and advanced image manipulation techniques has necessitated the development of deep learning-based image forensics. Convolutional Neural Networks (CNNs) like ResNet, VGG, and EfficientNet are widely used to detect minute pixel-level inconsistencies between real and altered images (Rustam et al., 2024). These methods show promise in identifying forged visual content that is otherwise undetectable by the human eye.

In addition, image-text correlation models now integrate visual and linguistic cues to identify inconsistencies between text claims and accompanying images (Li et al., 2020), thereby enhancing detection accuracy in multimodal misinformation.

HDC presents a biologically inspired alternative for real-time inference. Unlike traditional symbolic representations, HDC utilizes high-dimensional vectors to encode semantic relationships, making it resilient to noise and perturbations (Kupershtein et al., 2025; de Castro & von Zuben, 2002). Adaptive systems inspired by the human immune system have also influenced this domain. Artificial Immune Systems (AIS) have shown effectiveness in intrusion detection and anomaly classification (Aickelin et al., 2004; Aldhaheri et al., 2020; Donnachie et al., 2022).

Furthermore, hybrid approaches integrating quantum-based crossover models and bio-inspired classification mechanisms have shown potential in enhancing robustness against evolving disinformation strategies (Dai et al., 2014; Baug et al., 2019).

3 METHODOLOGY

The Adaptive Hyperdimensional Inference (AHI) system provides a real-time, multimodal application system for misinformation detection by incorporating Deep Learning, Evolutionary Learning, and Hyperdimensional Computing (HDC) techniques. Unlike the typical supervised learning methods that depend on large labeled datasets, it employs unsupervised clustering and similarity-based inference to adapt dynamically to new patterns of misinformation. The effectiveness of this method is tremendous for detecting fake news and image forgery since it improves the detection of adversarially transformed content and zero-day misinformation.

3.1 Identification of Textual Fake News

Text Hypervector Encoding: Sufficiently to know, with hypervector encoding, AHI uses characteristics regarding syntax and semantics while creating items based on non-existing news for identification.

3.2 Detection for Image Forgeries Feature Extraction Based on Deep Learning

AHI employs ResNet50, a deep-learning model trained on ImageNet, to draw high-level visual understandings from pictures. It compresses the image to 224 by 224 pixels, normalizes it, and transforms it into a tensor. The ResNet50 model is made up of several convolutional layers that analyze the image and extract important features such as edges and textures as well as the uneven illumination and anomalies indicating forgery.

3.3 Classification & Multimodal Hyperdimensional Fusion

3.3.1 Combining Hypervectors for Text and Images

One of the major advancements made possible by AHI is the possibility of combining text and visual data into a single hyperdimensional representation. This is accomplished by summing the image hypervectors with hypervectors of the text, followed by binarization of the resulting vector in order to keep the high-dimensional structure for the next effective similarity-based comparisons. Thus, this multimodal hyperdimensional encoding allows AHI to cross-validate textual claims against relevant visuals. If, for example, a fake story modifies either a certificate or a photo, AHI will identify the contradictions between text input and visual input, hence increasing overall accuracy.

3.4 Experimental Setup, Dataset, and Dataset

3.4.1 Experimental Setup

The frameworks that were initially designed to assess the AHI performances include these two major datasets: one is for the textual fake news detection, and the other is meant for the image forgery detection. These datasets were sourced well and carefully from online public repositories of misinformation to guarantee a varied and.

3.4.2 Dataset about Misinformation in News

The following reliable and reputable misinformation datasets were used to import data for textual fake news detection:

FakeNewsNet: This dataset includes both authentic and fraudulent news stories that have been verified by reliable websites like PolitiFact and GossipCop. It contains extensive metadata concerning each news item, including user interactions, media circulation, and reliability of the source.

3.4.3 Dataset of Image Manipulation

For the present study, we reference three well-known datasets of image forgery for the assessment of the ability of the AHI system in differentiating faked photos. These datasets include photographs on which different image tampering techniques have been applied, such as splicing, copy-move image forgery, and even AI-generated deepfake images, both real and forgery examples. Table 1 shows the Textual Fake News Dataset.

Table 1: Textual fake news dataset.

ID	Headline	Full Text	Source
1	Government Launches New Healthcare Policy	The government has introduced a new healthcare policy aimed at improving accessibility and affordability.	Gov News
2	Aliens Spotted in New York City	Several reports claim that UFOs were seen hovering over New York, but no official confirmation has been provided.	Conspiracy Times
3	Stock Market Hits Record Highs	The stock market reached an all-time high today, driven by strong economic growth and investor optimism.	Finance Daily
4	Celebrity Uses Secret Anti-Aging Formula	An anonymous source claims that a celebrity has been using a classified anti-aging formula, though experts deny its existence.	Entertainment Buzz
5	Scientists Discover Water on Mars	NASA confirms that traces of water have been found on Mars, which could have implications for future space exploration.	Science Today

Table 2: Image forgery dataset.

ID	Image File Name	Modification Type	Label
1	gov_policy.jpg	Original	1
2	alien_nyc.jpg	Spliced	0
3	stock_market.jpg	Original	1
4	celebrity_fake.jpg	Deepfake	0
5	mars_water.jpg	Original	1

Deepfake Image Dataset synthesizes AI-generated synthetic images using generative adversarial networks (GANs). Deepfake images posed serious challenges when detecting. Table 2 shows the Image Forgery Dataset.

3.5 Evaluation Metrics

Three major evaluation metrics (Accuracy, Zero-Day Detection Rate, and Adversarial Robustness) have been used to assess the efficacy of the Adaptive Hyperdimensional Inference (AHI) architecture. These metrics help to evaluate the resilience of AHI against adversarial attacks, generalization against unseen misinformation, and detection of fake news and image annealing attacks. Below are theoretical and mathematical definitions of the metrics.

3.5.1 Defining Accuracy

Accuracy is the measure of the ability of AHI to distinguish between authentic and fraudulent samples. This metric indicates the percentage of correct guesses in all predictions and is the most widely used stat for classification tasks.
Mathematical Formula

$$(\text{Correct Predictions}/\text{Total Predictions}) \times 100 = \text{Accuracy} \quad (1)$$

$$\text{Accuracy} = (\text{Total Correct Predictions}) \times 100 \quad (2)$$

3.5.2 Justification

One important parameter that offers an overall evaluation of AHI's ability to discriminate between authentic news or photos and fraudulent ones is accuracy. Higher accuracy shows the algorithm classifies samples with high effectiveness and low error.

The Zero-Day Detection Rate formula is: (Correct Zero-Day Detections Total Zero-Day Samples) × 100

Total Zero-Day Samples Correct Zero-Day Detections = () × 100 = Zero-Day Detection Rate

3.6 Definition for Adversarial Robustness (in Percent)

Mathematical Formula

Adversarial Robustness = (Total samples after attack × Correct classifications after attack) × 100

Where:

- Accurate Classifications = Number of samples correctly classified even after being adversarially modified After Attack.
- Total Samples After Attack = Number of samples which have been put through adversarial modifications.

4 RESULTS AND ANALYSIS

4.1 Accuracy

AHI's accuracy was compared to that of a traditional Multi-Layer Perceptron (MLP) classifier based on supervised training. The following results have been obtained:

Table 3: Supervised vs unsupervised.

Model	Accuracy (%)
MLP Classifier (Supervised Learning)	91.3
Adaptive Hyperdimensional Inference (AHI - Unsupervised Learning)	87.9

Table 4: Zero day detection rate.

Metric	Score (%)
Zero-Day Detection Rate	82.6

Table 3 shows the Supervised vs unsupervised. Table 4 shows the Zero Day detection rate.

4.2 Robustness against Adversarial

We put AHI through various attacks using text and image perturbations, and we saw how exceedingly well the system continued to classify disinformation. Here is a summary of the results:

Table 5: Adversarial attack type.

Adversarial Attack Type	Accuracy Before Attack (%)	Accuracy After Attack (%)	Robustness (%)
Textual Synonym Replacement	91.3	86.1	94.3
Textual Sentence Shuffling	91.3	83.4	91.4
Image Adversarial Attack (FGSM)	91.3	79.2	86.8
Image Deepfake Manipulation	91.3	81.5	89.3

Table 6: Comparative Evaluation.

Model	Accuracy (%)	Zero-Day Detection (%)	Adversarial Robustness (%)
SVM (Text-Only)	85.2	67.4	78.3
CNN (Image-Only)	87.6	58.9	73.2
BERT (NLP Transformer)	90.1	72.1	80.4
AHI (Proposed)	87.9	82.6	89.3

These results indicate that deep learning models (BERT, CNNs) work reasonably well in structured setups but are vulnerable to adversarial attacks and are unable to generalize against new misleading things. AHI, on the other hand, triumphed in adversarial robustness (89.3%) and zero-day detection (82.6%) in real-world misleading scenarios, proving itself superior to all other models. Table 5 shows the Adversarial Attack Type. Table 6 shows the Comparative Evaluation.

5 CONCLUSIONS

The Adaptive Hyperdimensional Inference (AHI) paradigm provides a multifaceted approach to detecting misinformation through the exciting convergence of deep learning, evolutionary learning, and hyperdimensional computing (HDC). Contrary to standard machine learning frameworks that rely on pre-labeled datasets, because of its efficient unsupervised clustering and similarity-based inference mechanism, AHI can successfully counter adversarial attacks and detect zero-day misinformation.

With that said, AHI has shown impressive experimental performance in an unsupervised setting with 87.9% accuracy, which is quite comparable to supervised models such as MLP (91.3%). AHI has also shown its ability to generalize beyond the confines of training data by identifying 82.6% of disinformation samples previously seen. This averaged robustness against hostile alterations is further testimony to its reliability for real-world applications.

REFERENCES

B. Li, H. Zhou, J. He, M. Wang, Y. Yang, and L. Li, "On the sentence embeddings from pre-trained language models", in Proceedings of the 2020.

B. Donnachie, J. Verrall, A. Hopgood, P. Wong, and I. Kennedy, "Accelerating cyber-breach investigations through novel use of artificial immune system algorithms", in Artificial Intelligence XXXIX, 2022, pp. 297–302.

C. Castillo, M. Mendoza, and B. Poblete, "Information credibility on twitter", in Proceedings of the 20th international conference on World wide web, Mar. 28, 2011, pp. 675–684.

Cavaliere, D., Fenza, G., Furno, D. (2024). A semantic model bridging DISARM framework and Situation Awareness for disinformation Attacks Attribution. IEEE.

E. Baug, P. Haddow, and A. Norstein, "MAIM: A novel hybrid bio-inspired algorithm for classification", in 2019 IEEE Symposium Series on Computational Intelligence (SSCI), Dec. 2019, pp. 1802–1809.

F. M. Burnet, the clonal selection theory of acquired immunity. Nashville: Vanderbilt University Press, 1959, 232 pp., Pages: 1-232.

H. Dai, Y. Yang, H. Li, and C. Li, "Bi-direction quantum crossover-based clonal selection algorithm and its applications", Expert Systems with Applications, vol. 41, no. 16, pp. 7248–7258, Nov. 15, 2014.

J. Y. Khan, M. T. I. Khondaker, S. Afroz, G. Uddin, and A. Iqbal, "A benchmark study of machine learning models for online fake news detection", Machine Learning with Applications, vol. 4, p. 100 032, Jun. 15, 2021.

Kupershtein, L., Zalepa, O., Sorokolit, V. (2025). AI-agent-based system for fact-checking support using large language models. CEUR Workshop Proceedings.

L. de Castro and F. von Zuben, "Learning and optimization using the clonal selection principle", IEEE Transactions on Evolutionary Computation, vol. 6, no. 3, pp. 239–251, Jun. 2002.

P. Karpov and G. Squillero, "VALIS: An evolutionary classification algorithm", Genetic Programming and Evolvable Machines, vol. 19, no. 3, pp. 453–471, Sep. 2018.

Paulen-Patterson, D., Ding, C. (2024). Fake News Detection: Comparative Evaluation of BERT-like Models and Large Language Models with Generative AI-Annotated Data. arXiv.

Raza, S., Paulen-Patterson, D., Ding, C. (2025). Fake news detection: comparative evaluation of BERT-like models and large language models with generative AI-annotated data. Springer.

Reddy, P., Muthyala, S. (2024). AI-Based Anomaly Detection in Large Financial Datasets: Early Detection of Market Shifts.

Rustam, F., Jurcut, A.D., Alfarhood, S., Safran, M. (2024). Fake news detection using enhanced features through text to image transformation with customized models. Springer.

S. Aldhaheri, D. Alghazzawi, L. Cheng, A. Barnawi, and B. A. Alzahrani, "Artificial immune systems approach to secure the internet of things: A systematic review of the literature and recommendations for future research", Journal of Network and Computer Applications, vol. 157, p. 102 537, May 1, 2020.

S. Bradshaw, H. Bailey, and P. N. Howard, Industrialized Disinformation: 2020 Global Inventory of Organized Social Media Manipulation. Computational Propagand a Project at the Oxford Internet Institute, 2021.

S. Atske. "News consumption across social media in 2021", Pew Research Center's Journalism Project. (Sep. 20, 2021), [Online]. Available: https: / / www. pewresearch . org / journalism / 2021 / 09 / 20 / news - consumption - across - social-media-in-2021/.

U. Aickelin, J. Greensmith, and J. Twycross, "Immune system approaches to intrusion detection – a review", in Artificial immune systems: third international conference, ICARIS 2004, Catania, Sicily, Italy, September 13-16, 2004: proceedings, 3239, 2004, pp. 328–341.

Y. Dun, K. Tu, C. Chen, C. Hou, and X. Yuan, “KAN: Knowledge-aware attention network for fake news detection”, Proceedings of the AAAI Conference on Artificial Intelligence, vol. 35, no. 1, pp. 81–89, May 18, 2021.