International Journal of
Computational Linguistics Research

DLINE JOURNALS

Print ISSN: 0976-416X
Online ISSN: 0976-4178

IJCLR 2026: 17 (1)
https://doi.org/10.6025/ijclr/2026/17/1/1-17

A Review of Viterbi Algorithm Applications in Pattern Recognition and Language Processing

Pit Pichappan
Digital Information Research Labs
Chennai. Tamil Nadu. India
pichappan@dirf.org

ABSTRACT

This work presents a comprehensive review of the Viterbi algorithm's applications in pattern recognition and natural language processing (NLP). Originally developed for decoding convolutional codes, the Viterbi algorithm has become a cornerstone technique for maximum a posteriori (MAP) sequence estimation in Hidden Markov Models (HMMs). The paper outlines its theoretical foundation as a dynamic programming method that efficiently identifies the most probable sequence of hidden states given observed data. It then explores numerous domain specific adaptations such as modified trellis structures, context aware decoding, beam search pruning, and finite state machine optimisations that enhance performance across diverse tasks.

These applications span distorted pattern recognition, named entity recognition (NER), stochastic grammar parsing, sign and gesture recognition, contextual text recognition, language testing, letter to phoneme conversion, and multilingual translation. In each domain, Viterbi based approaches consistently outperform traditional baselines by improving accuracy, reducing computational complexity (often from exponential or cubic to linear or quadratic time), and increasing robustness to noise and temporal distortions. The review also highlights recent innovations integrating Viterbi decoding with deep learning, parallel computing (e.g., GPU acceleration), and multimodal inputs.

A unified conceptual framework illustrates how multimodal inputs are processed through feature extraction and probabilistic modelling before Viterbi based decoding yields optimised decisions. Despite its maturity, the algorithm remains highly relevant due to its optimality, interpretability, and efficiency. However, challenges persist including scalability in large state spaces, integration with modern neural architectures, and the need for explainable, user centric decoding. The paper concludes that the Viterbi algorithm is not obsolete but evolving, with ongoing research poised to further extend its utility in intelligent, real time systems.

Keywords: Viterbi Algorithm, Hidden Markov Models, Pattern Recognition, Natural Language Processing, Sequence Decoding, Multilingual Translation

Received: 17 August 2025, Revised 25 November 2025, Accepted 8 December 2025

1. Introduction

Probabilistic modelling plays a central role in modern pattern recognition and language processing systems, where ambiguity, noise, and temporal dependencies are inherent. Among decoding techniques, the Viterbi algorithm remains one of the most widely used due to its optimality and computational efficiency. Originally introduced for convolutional code decoding, the algorithm has since been adapted to a wide range of stochastic models, particularly Hidden Markov Models (HMMs).

With the growth of data driven approaches in natural language processing (NLP), speech recognition, and multimodal interfaces, numerous variants of the Viterbi algorithm have been proposed to address domain-specific challenges. These include modified trellis structures, context aware decoding, and automaton-based optimizations. This review consolidates research efforts across these domains and presents them in a unified, systematic manner.

2. Potential of the Viterbi Algorithm

The Viterbi algorithm is a dynamic programming approach for finding the maximum a posteriori (MAP) state sequence in a stochastic model. In a hidden Markov model (HMM), the underlying finite state Markov chain is not directly observable; rather, it is inferred from an associated observation process. Our goal is to estimate the unobserved state sequence of this Markov chain. The most commonly used approach for this task is the maximum a posteriori (MAP) path estimator, which identifies the most probable sequence of hidden states given the observed data. This estimator can be computed efficiently using the Viterbi algorithm (Viterbi, 1967) a method widely employed in applications such as coding theory, inter symbol interference correction, and speech recognition. [1]

Much of the Bayesian literature on hidden Markov models (HMMs) see, for example, Cappé, and others [2, 3, 4, 5, 6] focuses on HMMs and Lember et al on Markov chain Monte Carlo (MCMC) methods [7, 8] When the objective is to estimate the underlying hidden state sequence corresponding to a given observation sequence, a task commonly referred to as *segmentation* (also known as *decoding* or *denoising*) various Gibbs sampling–based approaches are frequently employed, including techniques such as simulated annealing While decoding stochastic regular languages can be achieved in linear time, extensions to richer grammatical formalisms—such as stochastic context free or linear grammars often increase computational complexity. Modified Viterbi algorithms have been proposed to reduce these costs from cubic to quadratic time by constraining derivations or exploiting structural properties of the grammar.

Oskar Soop [9] demonstrated that, although the problem remains NP-hard, the use of tight upper and lower bounds, particularly when combined with m Viterbi approximations, can substantially reduce the number of candidate paths explored. This approach significantly outperforms exhaustive search in computational complexity and enables practical computation of Viterbi paths in TMMs of moderate size. Moreover, it enables early stopping during the search process, while still providing rigorous upper and lower bounds on the optimal path probability. Fishel [10] proposed a new TBD algorithm that combines dynamic programming with deep learning. It integrates a deep neural network (DNN) into the Viterbi algorithm to reduce computational complexity through state aware pruning and avoids the need for an explicit measurement model. The approach is validated using realistic Range Doppler radar data and shows improved detection of weak targets in noisy, cluttered environments. Hassan Razavi et al. [11] introduced a method to accelerate maximum a posteriori (MAP) trajectory estimation in continuous-time stochastic systems. By recasting the problem as an optimal control problem using the Onsager Machlup functional, the authors enable parallel in time computation via associative scan algorithms. This yields parallel versions of the Kalman Bucy filter and Rauch Tung Striebel smoother in linear Gaussian settings, extended to nonlinear models via Taylor expansions. GPU experiments confirm significant computational speedups without sacrificing accuracy.

3. Conceptual Framework for Viterbi-Based Systems

Viterbi-based systems rely on HMM elements: hidden states, transition probabilities between states, emission probabilities for observations from states, and initial state probabilities. These define the probabilistic model in which observations are the visible outputs of unobserved states. Systems apply Viterbi decoding in digital communications for error correction in noisy channels, speech recognition to map audio to word sequences, bioinformatics for gene prediction, and robotics for path localisation. [12 Variants include beam search for pruning low probability paths, logarithmic scaling to avoid underflow, and parallel implementations for speed; these adapt the framework to large scale or real time systems.[13]

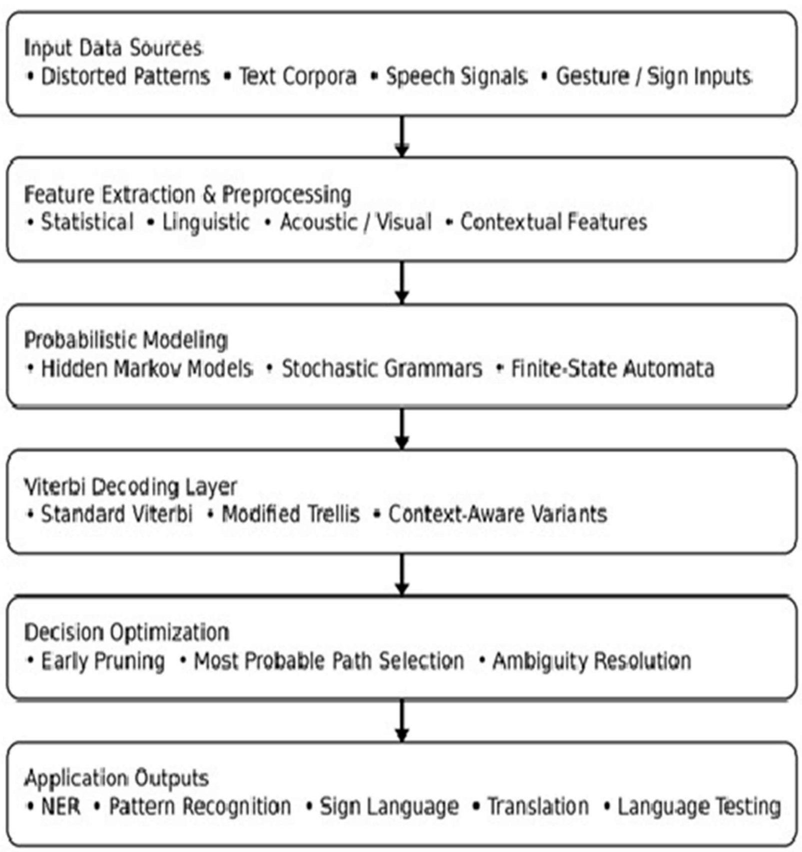

Figure 1. Conceptual framework illustrating the role of the Viterbi algorithm and its modified variants in pattern recognition and language processing systems

The framework highlights the flow from multimodal input data through feature extraction and probabilistic modelling to Viterbi-based decoding and optimised decision making. Modified trellis structures and context-aware decoding enable efficient pruning of irrelevant hypotheses, thereby improving accuracy and reducing computational complexity across applications such as named entity recognition, sign language recognition, language testing, and multilingual translation. The Viterbi algorithm combines the advantages of big data frameworks and hidden Markov models to solve the decoding problem for large-scale multidimensional data. [14]

Figure 1 presents a unified conceptual framework illustrating how Viterbi based systems operate across domains. The framework begins with multimodal inputs (text, speech, gestures, distorted patterns), followed by feature extraction and probabilistic modelling. The core decoding stage employs either standard or modified Viterbi algorithms, enabling early pruning of irrelevant hypotheses. The optimised decision layer produces application-specific outputs such as recognised entities, translated text, or decoded gestures.

This framework demonstrates that, despite differences in input modalities and end tasks, most applications share a common probabilistic decoding backbone centred on the Viterbi algorithm.

4. Viterbi Algorithm Applications Across Domains

We plan to provide a comparative overview of how the Viterbi algorithm (or its variants) has been applied across various application domains to improve performance relative to baseline methods. It includes and is structured around six key columns:

Application Domain	**Evaluation Metrics**	**Baseline Method**	**Viterbi-Based Improvement**	**Computational Complexity**	**Reported Performance Gains**
Distorted Pattern Recognition	Accuracy, Error Rate, Processing Time	Exhaustive pattern matching	Modified trellis Viterbi with early pruning	Reduced from exponential to near-linear	Higher recognition accuracy with significant reduction in decision time
Named Entity Recognition (NER)	Precision, Recall, F1-score	Rule-based / greedy decoding	Context-aware Viterbi sequence decoding	Linear in sequence length	Improved F1-score and stable boundary detection
Stochastic Grammar Parsing	Parsing accuracy, Likelihood score	CYK-style parsing	Extended Viterbi parsing	Reduced from $O(n^3)$ to $O(n^2)$	Faster parsing with comparable or improved accuracy
Sign Language Recognition	Recognition rate, Latency	Template matching	HMM + Viterbi decoding	Linear in frame sequence length	Robust temporal alignment and real-time recognition

Gesture Recognition	Accuracy, Robustness to noise	Frame-wise classifiers	Standard Viterbi	Linear	Improved robustness against temporal distortions
Contextual Text Recognition	Word accuracy, Context consistency	Independent symbol decoding	Context-sensitive Viterbi	Linear	Increased robustness under noisy or ambiguous input
Language Testing Systems	Scoring consistency, Processing speed	Manual / heuristic scoring	Viterbi-based decoding	Linear	Faster evaluation with consistent scoring outcomes
Letter-to-Phoneme Conversion	Phoneme accuracy, Runtime	Naïve Viterbi	FSM-optimized Viterbi	Linear with reduced constant factor	Large runtime reduction with minimal accuracy loss
Machine-Aided Translation	Parse accuracy, Disambiguation rate	Rule-based parsing	Viterbi-assisted disambig-uation	Polynomial	Improved syntactic ambiguity resolution
Multilingual Translation Systems	Translation accuracy, Response time	Single-modality translation	Viterbi-based multimodal decoding	Linear–polynomial	Improved translation flexibility and user adaptation

Table 1. Performance Metrics of Viterbi Algorithm Applications Across Domains

Table 1. summarises the key performance metrics used to evaluate Viterbi-based systems across multiple application domains, highlighting improvements in accuracy, robustness, and computational efficiency relative to baseline methods

4.1 Application Domain

This column lists various fields in which the Viterbi algorithm has been employed. The research on the performance of three Turbo decoding algorithms Soft Output Viterbi Algorithm (SOVA), Logarithmic Maximum A Posteriori (Log-MAP), and Maximum A Posteriori Probability (MAP) with a focus on their efficacy in ensuring reliable biomedical signal transmission for real-time telemedicine applications, contributes to enhancing the dependability and quality of telemedicine systems. [15] Viterbi decoders are extensively applied in communication systems, natural language processing (NLP), and other domains [16]. The Viterbi algorithm is a dynamic programming algorithm that solves fundamental problems in Hidden Markov Models and is used in real world applications such as big data and multimodal data [17]. Parts of Speech Tagging utilising a Modified Viterbi Algorithm performs exceptionally well. [18J. The errors in Micro Doppler Trajectory Estimation are highly reduced when the Viterbi Hough Joint Algorithm is used. [19] The curve extraction algorithm for multi-component m D signals based on the reassociation Viterbi algorithm is crucial for target recognition of space targets. [20] Leveraging the modelling capabilities of Hidden Markov Models (HMMs) and the decoding efficiency of the Viterbi algorithm (VA), this paper proposes a VA based scheme that successfully recognises ten common badminton strokes in real time, as demonstrated by experimental results. [21]. The Viterbi algorithm is adapted to extract reliability information within the turbo decoder, and when this soft information is incorporated back into the Viterbi decoding process, it yields a soft output version known as the Soft Output Viterbi Algorithm (SOVA). [22]

These applications include the following, but are not limited to:

- Distorted Pattern Recognition
- Named Entity Recognition (NER)
- Stochastic Grammar Parsing
- Sign Language Recognition
- Gesture Recognition
- Contextual Text Recognition
- Language Testing Systems
- Letter-to-Phoneme Conversion
- Machine-Aided Translation
- Multilingual Translation Systems

These domains include natural language processing, computer vision, speech processing, and educational technology. enables a wide range of data driven applications, such as speech recognition [22, 23, 25, 26]. For the speech recognition task considered, FLASH-BS Viterbi achieves a 69.5× speedup with only a 0.05% increase in error, demonstrating substantial gains in efficiency. [13]

4.2 Evaluation Metrics

The Viterbi algorithm is employed in speech and character recognition tasks, in which speech signals or characters are represented by hidden Markov models, as evidenced by numerous studies. The Viterbi algorithm is applied in digital communication systems and in speech and character recognition. [27]. The confidence score and vector as two indicators responsible for evaluating the Viterbi model internally before assessing the final results. [28] The Viterbi algorithm achieves an accuracy of 86.34%, surpassing the state of the art POS taggers for the Indian languages. For each domain, specific metrics are used to assess system performance.

[29] Examples include:

- Accuracy, Error Rate, Processing Time (for pattern recognition)
- Precision, Recall, F1-score (for NER)
- Parsing accuracy, Likelihood score (for grammar parsing)
- Recognition rate, Latency (for sign language)
- Word accuracy, Context consistency (for text recognition)
- Scoring consistency, Processing speed (for language testing)
- Phoneme accuracy, Runtime (for letter-to-phoneme conversion)

These metrics reflect both effectiveness (e.g., accuracy, F1-score) and efficiency (e.g., latency, runtime).

4.3 Baseline Method

This column identifies the traditional or non Viterbi approach used before applying the Viterbi algorithm. Aligning phrase pairs from comparable sentences and evaluating the utility of the extracted phrases by using them directly in the MT decoder are more effective. [30].The parsers in languages use hidden non Viterbi structures and can be relexicalized and adapted using unlabeled target language data. Examples include:

- Exhaustive pattern matching
- Rule-based or greedy decoding
- CYK-style parsing (a classic dynamic programming parser for context-free grammars)
- Template matching
- Frame-wise classifiers
- Independent symbol decoding
- Manual or heuristic scoring
- Naïve Viterbi (in one case, indicating further optimisation was possible)
- Rule-based parsing
- Single-modality translation

These baselines typically suffer from limitations such as high computational cost, poor handling of sequential dependencies, or limited contextual awareness.

4.4 Viterbi-Based Improvement

The Viterbi algorithm yields the best Conditional Random Field scores on social media texts [31]. For the Personalised Drug Prescriptions Decision Support System, the variants of the enhanced Viterbi algorithm play a central role. [32]

Describes the specific variant or adaptation of the Viterbi algorithm introduced in each domain:

- Modified trellis Viterbi with early pruning (for faster decisions in pattern recognition)
- Context-aware Viterbi sequence decoding (to capture dependencies between entities in NER)
- Extended Viterbi parsing (optimised for stochastic grammars)
- HMM + Viterbi decoding (standard in temporal signal modelling, like sign language)
- Standard Viterbi (applied to model gesture sequences)
- Context-sensitive Viterbi (to leverage linguistic context in text recognition)
- Viterbi-based decoding (for consistent automated scoring)
- FSM-optimised Viterbi (using finite-state machines to speed up phoneme conversion)
- Viterbi-assisted disambiguation (to resolve syntactic ambiguities)
- Viterbi-based multimodal decoding (integrating multiple input modalities in translation)

These improvements tailor the core Viterbi idea finding the most likely sequence of hidden states to domain-specific challenges.

4.5 Computational Complexity

Earlier reports show that the time complexity of the Viterbi-based approach differs from that of the baseline: high random access memory (RAM) usage and computational complexity limit its applicability in scenarios that demand real time reconfigurability. [33]

- Often reduced to linear in sequence length (e.g., $O(n)$)
- In grammar parsing, improved from $O(n^3)$ (CYK) to $O(n^2)$
- In pattern recognition, reduced from exponential to near-linear
- Some cases note reduced constant factors (same asymptotic class but faster in practice)

- A few applications report polynomial or linear polynomial complexity, reflecting more complex models(e.g., multimodal systems)

This highlights the Viterbi algorithm's role not just in improving accuracy but also in enabling scalable, real-time solutions. It is effective to detect the health issues accurately for each detection method [34]. Experiments on aggregated High Level Data Link Control (HDLC) protocol packets show that our method achieves performance close to the approximate theoretical upper bound while offering substantially lower computational complexity and enhanced usability. [35]. The MAP sequence estimation method effectively addresses multi-modality while substantially reducing computational and memory overhead. [36]

4.6 Reported Performance Gains

After the 2000s, various approaches have been proposed to address the main challenges of parallelizing the Viterbi algorithm [Imad Sassi]. Liu et al. analyse how to parallelise the three most important Hidden Markov Model (HMM) algorithms including the Viterbi algorithm for a GPU computing environment [37]. Sand et al. propose a C++ library that leverages linear algebra, multi-core processors, and SSE support to implement a parallel HMM [38]. Ahn et al. present a parallelisation technique using a graphics processing unit as a real-time modem solution for a WiMAX multiple-input multiple-output (MIMO) system [39]. Buthpitiya et al. [40] propose a GPU-based Viterbi implementation that outperforms a sequential CPU implementation [40]. In recent years, numerous studies have focused on parallel and/or distributed implementations of the Viterbi algorithm. Maleki et al. [41] introduce a novel method for parallelising a class of problems, including Viterbi, by exploiting algebraic properties. Hanif et al. describe the design of a parallel Viterbi algorithm based on matrix products, demonstrating that significant speedup can be achieved by applying GPUs to such parallel computing tasks [42].

We summarise the practical benefits observed as below:

- Higher recognition accuracy and faster decision time
- Improved F1-score and stable boundary detection in NER
- Faster parsing with comparable or better accuracy
- Robust temporal alignment and real time recognition in sign language
- Improved robustness against noise and temporal distortions in gestures
- Increased robustness under ambiguous inputs in text recognition
- Consistent scoring and faster evaluation in language testing
- Large runtime reduction with minimal accuracy loss in phoneme conversion
- Better ambiguity resolution in machine translation
- Enhanced flexibility and user adaptation in multilingual systems

These gains demonstrate that Viterbi-based methods offer a compelling balance of accuracy, efficiency, and robustness across domains involving sequential data and hidden state inference.

4.7 Derivation

The table illustrates the versatility and power of the Viterbi algorithm as a foundational technique for sequence modelling. By leveraging probabilistic models (often Hidden Markov Models) and dynamic programming, systems can make globally optimal predictions over sequences outperforming local or heuristic methods in both speed and accuracy.

5. Applications in Pattern Recognition and Language Processing

5.1 Distorted Pattern Recognition

In distorted pattern recognition, modified trellis structures incorporate statistical properties of distortion, allowing irrelevant pattern classes to be eliminated early. This human like decision strategy significantly reduces computational effort while improving recognition accuracy. The Markov Model Based Fusion Algorithm demonstrates strong overall performance in evaluating the quality of restored text based images [43]. The complex value neural network (CVNN) and RVNN equaliser connected with the Viterbi Viterbi algorithm have better BER performance with a lower error rate. [44]

5.2 Information Extraction and Named Entity Recognition

Information extraction systems aim to transform unstructured text into structured representations. Named Entity Recognition (NER), a key subtask, benefits from Viterbi decoding by modelling entity sequences probabilistically. Context aware Viterbi variants improve boundary detection and classification accuracy compared to greedy or rule based approaches. The Viterbi decoder based on label pair rules enables named entity recognition on a large volume of unknown data with a small amount of annotated data [45] manually annotated news corpus using the Viterbi decoding algorithm to assign the most probable named entity tags to words, enabling classification accuracy, which sets benchmarks for machine translation, question answering, summarisation, and information retrieval systems. [46]

5.3 Stochastic Grammar Parsing

Parsing stochastic grammars traditionally incurs high computational cost. Modified Viterbi algorithms enable efficient derivation of the most probable parse trees for one counter and linear grammars, achieving quadratic-time complexity while maintaining accuracy. Stochastic models are used to capture statistical correlations in the training data of large language models. [47]

5.4 Sign Language and Gesture Recognition

In sign and gesture recognition, temporal alignment is critical. HMM-based systems decoded using the Viterbi algorithm provide robust recognition of gesture sequences under temporal variability. Both word level and subunit level models have demonstrated real time feasibility. To assign polarity scores to the thesis or to entities within a phrase, in text analysis and analytics, machine learning (ML), and natural language processing (NLP) approaches are employed. This Sentiment Analysis using a POS tagger helps us derive a summary of the broader public's opinion on a specific topic. To this end, Chavali used the Viterbi algorithm, the Hidden Markov Model (HMM), and the Constraint based Viterbi algorithm for POS tagging. [48]

5.5 Contextual Text Recognition

Contextual text recognition systems employ modified Viterbi algorithms to integrate local observations with global constraints. While effective, the robustness of these systems with respect to varying source statistics remains underexplored. To recognise the whole word from a fixed lexicon, the Viterbi algorithm proved to yield highly accurate results. [49]

5.6 Language Testing Systems

Language testing and assessment systems employ Viterbi decoding to enable efficient and consistent scoring. Studies on washback effects suggest that algorithmic evaluation influences learning behaviours, highlighting the broader educational impact of such systems. The Viterbi algorithm is currently the most widely used and effective decoding method, offering high efficiency and speed within the system. [49]

5.7 Letter-to-Phoneme Conversion

In letter-to-phoneme conversion, naive implementations suffer from high runtime due to large correspondence sets. Optimised Viterbi decoding using deterministic finite-state automata significantly reduces execution time without sacrificing accuracy. The Tree constrained Pointer Generator (TCPGen) has shown promise in recognising context specific rare words, efficiently biasing them with a prefix tree. [50]. The tasks of microtext normalisation, sentence boundary disambiguation, part of speech tagging, text chunking, and lemmatisation in Computational syntactic processing employ Viterbi based syntactic analysis.

5.8 Machine-Aided and Multilingual Translation

Viterbi-assisted parsing helps resolve syntactic ambiguities in machine-aided translation, particularly for morphologically rich languages. In multilingual translation frameworks, Viterbi based decoding supports multimodal input (speech, text, printed text) and user adaptive output generation. The Hidden Markov Model (HMM) and the Viterbi algorithm have potential for training and decoding in multilingual translation. [51]. In machine translation, part of speech tagging is a crucial step, and the Hidden Markov Model achieves high accuracy in this task, leaving room for further improvement with enhanced datasets. [52]

6. Comparative Performance Trends

Across a wide range of application domains, Viterbi-based systems consistently demonstrate several fundamental performance advantages that explain their enduring relevance in sequential modelling tasks. First, they achieve improved accuracy through global sequence optimisation, as the Viterbi algorithm evaluates entire state sequences rather than making locally optimal or greedy decisions. This global optimisation property enables more reliable handling of long range dependencies, contextual constraints, and ambiguous observations, which is particularly important in tasks such as named entity recognition, grammar parsing, speech recognition, and machine translation.

Second, Viterbi based approaches offer reduced computational complexity through effective pruning strategies and optimised trellis structures. By eliminating low probability paths early in the decoding process and exploiting structural constraints such as finite-state automata, beam search, or context aware transitions, many applications report substantial reductions in runtime and memory usage. In several cases, computational complexity is reduced from exponential or cubic time to linear or quadratic time, enabling scalable and real-time deployment.

Third, these systems exhibit strong robustness to noise, distortion, and temporal variability. Because the Viterbi algorithm integrates probabilistic state transitions with observation likelihoods over time, it can smooth out local errors and maintain stable predictions even under noisy inputs, missing data, or temporal misalignments. This robustness is especially evident in gesture recognition, sign language interpretation, and distorted pattern recognition tasks.

Performance evaluation across studies commonly relies on metrics such as accuracy, precision, recall, F1-score, parsing likelihood, and runtime efficiency, reflecting both effectiveness and efficiency. Notably, optimised and parallelised Viterbi variants show particularly strong performance gains in time sensitive and large scale applications, where rapid decision making, low latency, and computational efficiency are critical, such as real time speech processing, multilingual translation systems, and high throughput information extraction pipelines.

7. Challenges and Open Research Issues

Despite its strengths, several challenges remain:

- Limited robustness analysis under varying source statistics, Scalability issues for extremely large state spaces, Integration with modern deep learning architectures, and the need for user centric and explainable decoding strategies.

Addressing these issues is critical for next generation Viterbi based systems.

8. Conclusion

The Viterbi algorithm remains a foundational and unifying technique for probabilistic sequence decoding across a wide range of pattern recognition and language processing applications. This review has demonstrated that, despite significant advances in machine learning and deep neural architectures, the core principles of Viterbi-based decoding remain highly relevant due to their optimality, interpretability, and computational efficiency. By systematically examining classical and modified variants of the algorithm, the study highlights how domain specific adaptations such as context-aware decoding, optimised trellis structures, finite-state constraints, and parallel implementations have enabled substantial improvements in both accuracy and runtime performance.

Across applications, including distorted pattern recognition, named entity recognition, stochastic grammar parsing, sign and gesture recognition, contextual text recognition, language testing, letter to phoneme conversion, and multilingual translation, Viterbi-based methods consistently outperform baseline or heuristic approaches. [53]. These gains are primarily attributable to the algorithm's ability to perform global optimisation across entire sequences, thereby handling uncertainty, temporal variability, and noisy observations. The comparative analysis further shows that many Viterbi-based solutions reduce computational complexity from exponential or cubic bounds to near-linear or quadratic time, making them suitable for real-time and large scale systems.

The proposed conceptual framework reinforces the idea that diverse application domains share a common probabilistic decoding backbone centred on the Viterbi algorithm. This unification clarifies how multimodal inputs, probabilistic modelling, and optimised decoding collectively contribute to robust decision making.

However, the review also identifies open challenges, including scalability to extremely large state spaces, robustness under changing data statistics, and seamless integration with modern deep learning models.

Overall, the findings confirm that the Viterbi algorithm is not merely a legacy technique but an evolving, adaptable methodology. Future research that combines Viterbi-based decoding with neural representations, parallel computing, and explainable AI paradigms is likely to further enhance its impact in next-generation intelligent and user centric systems.

This review presents a comprehensive, structured analysis of the Viterbi algorithm and its applications in pattern recognition and language processing. By unifying diverse studies under a common conceptual framework, the review highlights both the versatility and enduring relevance of Viterbi-based decoding. Continued research into robustness, scalability, and hybrid architectures is expected to further extend its applicability in intelligent systems.

References

[1] Caliebe, A., Rosler, U. (2002). Convergence of the maximum a posteriori path estimator in hidden Markov models, *IEEE Transactions on Information Theory*,. 48 (7) p. 1750-1758, July.

[2] Cappé, O., Moulines, E., Rydén, T. (2005). Inference in Hidden Markov Models. *Springer, New York*.

[3] Corander, J., Xiong, J., Cui, Y., Koski, T. (2013).Optimal Viterbi Bayesian predictive classification for data from finite alphabets. *J. Stat. Plan. Inference* 143 (2), 261–275.

[4] Koski, T. (2001). Hidden Markov models for Bioinformatics, Computational Biology Series, vol. 2. Kluwer Academic Publishers, *Dordrecht*.

[5] Marin, J., Robert, C. (2007). Bayesian Core: a Practical Approach to Computational Bayesian Statistics. Springer Texts in Statistics. *Springer, New York*.

[6] Rydén, T. (2008). EM versus Markov chain Monte Carlo for estimation of hidden Markov models: a computational perspective. *Bayesian Anal*. 3 (4) 659–688.

[7] Lember, J., Koloydenko, A. (2007). Adjusted Viterbi training: a proof of concept. *Probab. Eng. Inf. Sci.* 21 (3), 451–475.

[8] Besag, J. (2004). An introduction to Markov chain Monte Carlo methods. In: Mathematical Foundations of Speech and Language Processing, IMA Volumes in Mathematics and Applications, vol. 138, p. 247–270. *Springer, New York*.

[9] Soop, Oskar., Lember, uri., J. (2025). Branch and bound method for calculating Viterbi path in triplet Markov models, arXiv:2507.19338v1.

[10] Fishel, E., Tsarov, N., Tapiro, T., Nuri, I., Shlezinger, N. (2024). Learn to Track Before Detect via Neural Dynamic Programming, *In*: ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), *Seoul, Korea, Republic* of p. 9586-9590.

[11] Razavi, Hassan., Ángel, F., Fernández, García., Särkkä, Simo. (2025). Temporal parallelisation of continuous-time maximum a posteriori trajectory estimation. arXiv:2512.13319v1.

[12] https://www.tencentcloud.com/techpedia/120389. (2025). What is the application of the Viterbi algorithm in speech recognition decoding.

[13] Deng, Ziheng., Liu, Xue., Jiang, Jiantong., Li, Yankai. (2025). Qingxu Deng Xiaochun Yang FLASH Viterbi: Fast and Adaptive Viterbi Decoding for Modern Data Systems. arXiv:2510.19301v2 [cs.DC] 23 Oct 2025.

[14] Sassi, Imad., Reda, Oumaima., Anter, Samir., Zellou, Ahmed. (2022). Efficient parallel Viterbi algorithm for big data in a spark cloud computing environment, *Procedia Computer Science,* Volume 215, Pages 937-946.

[15] Kumar, Santhosh., K. B., Sujatha, B. R., Sushma, N., Venkateswara Rao Kolli. (2024). Smart and Sustainable Energy Systems: Enhanced Transmission of EEG Signals in Telemedicine Applications using Soft Output Viterbi Turbo Decoder, IOP Conf. Ser.: *Earth Environ. Sci.* 1375 012022.

[16] Bhattacharjya, Rajat., Maity, Biswadip., Dutt, Nikil. (2023). Locate: Low-Power Viterbi Decoder Exploration using Approximate Adders. GLSVLSI '23: *Proceedings of the Great Lakes Symposium on* VLSI 2023. P. 409 - 413.

[17] Bouguila, Nizar., Fan, Wentao., Amayri, Manar. (2022). Hidden Markov Models and Applications. *Springer*.

[18] Sabonsolin, J. E. A., Ambat, S. C., Arguson, A. C., Almeniana, F. C., Malasaga, E. V., Tejuco, H. J. (2025). Utilizing Modified Viterbi Algorithm for Religious Text: A Cebuano Part of Speech Tagging 2024 International Conference on IT Innovation and Knowledge Discovery (ITIKD), *Manama, Bahrain,* p. 1-6.

[19] Ding, Y., Liu, R., She, Y., Jin, B., Peng, Y. (2022). Micro Doppler Trajectory Estimation of Human Movers by Viterbi Hough Joint Algorithm, In:*IEEE Transactions on Geoscience and Remote Sensing,* vol. 60, p. 1-11, Art no. 5113111.

[20] Wu, J., Ai, X., Zheng, Y., Xu, Z., Zhao, F. (2024). Micro Doppler Curve Extraction Based on Reassociation Viterbi Algorithm, *In*: IEEE Transactions on Aerospace and Electronic Systems, vol. 60 (4) p. 4295-4309, Aug.

[21] Xiaofeng, Jiang., Xiude, Guo., Hui, Feng., Fangling, Ren. (2025). Discrete Dynamic Modeling Analysis of Badminton Games Based on Viterbi Algorithm in College Badminton Physical Education, *International Journal of High Speed Electronics and Systems*. 34 (2). p. 2440037.

[22] Zhao, G. (2024). Dissecting and Implementing SOYA Algorithm Variations in Convolutional and Turbo Code Decoding: An Analytical Approach, 4th Asia-Pacific Conference on Communications Technology and Computer Science (ACCTCS), Shenyang, China, 2024, p. 780-785.

[23] Ciaperoni, M., Gionis, A., Katsamanis, A., Karras, P. (2022). Sieve: A space efficient algorithm for viterbi decoding, *In:* Proceedings of the 2022 International Conference on Management of Data, ser. SIGMOD '22. New York, NY, USA: Association for Computing Machinery, p. 1136–1145. [Online]. Available: https://doi.org/10.1145/3514221.3526170.

[24] Ciaperoni, M. A., Katsamanis, A., Gionis, P., Karras, (2024). Beam-search sieve for low-memory speech recognition, *In*: Proc. Interspeech p. 272–276.

[25] Gales, M., Young, S. (2008). The application of hidden Markov models in speech recognition, Foundations and Trends, *Signal Processing*, vol. 1, no. 3, p. 195–304, 2008. [Online]. Available: http://dx.doi.org/10.1561/2000000004.

[26] Braun, H., Luitjens, J., Leary, R., Kaldewey, T., Povey, D. (2020). Gpu-accelerated viterbi exact lattice decoder for batched online and offline speech recognition, In:ICASSP 2020-2020 IEEE International Conference on Acoustics, *Speech and Signal Processing* (ICASSP), p. 7874–7878.

[27] Lou, H. L. (1995). Implementing the Viterbi algorithm, *IEEE Signal Processing Magazine*, vol. 12 (5) p. 42-52, Sept.

[28] Baklouti, Zied. (2023). Tweet Sentiment Extraction using Viterbi Algorithm with Transfer Learning. arXiv:2308.05973v1. https://doi.org/10.48550/arXiv.2308.05973.

[29] Phukan, R., Baruah, N., Sarma, S. K., Konwar, D. (2023). Parts of Speech Tagging for Unknown Words in Assamese using the Viterbi Algorithm. *Indian Journal of Science and Technology* 16 (SP2) 53-59.

[30] Hewavitharana, S., Vogel, S. (2016). Extracting parallel phrases from comparable data for machine translation. *Natural Language Engineering*. 22 (4): 549-573.

[31] Seth, R., Sharaff, A. (2024). BERT-CA Sentinels: Adapting BERT with Attention Mechanism for Satirical News Detection in Social Media Texts, 2nd World Conference on Communication Computing (WCONF), RAIPUR, India, p. 1-7.

[32] Bello, J. (2025). A Framework for Personalized Drug Prescriptions Decision Support System using Hybrid Techniques Vol. 14 No. 1. *Journal of Science and Logics in ICT Research.*

[33] Xian, Y., Li, K., Gao, K., Liu, J., Ma, S. A Complexity Reduced VMbPWM-Based All-Digital Transmitter with Multi-Bit Deduction Toward Higher In-Band SNR, *IEEE Transactions on Circuits and Systems I:* Regular Papers.

[34] Zar, W. T., Tun, H. M., Win, L. L. Y., Naing, Z. M. (2024). Innovative non-contact r-r intervals estimation using the Viterbi algorithm with Squared Branch Metric (VSBM). *Jurnal Pendidikan Teknologi Kejuruan*, 7 (1), 32-43.

[35] Li, Y., Yu, u H., Du, J. Zhang, X., Liu, G. (2024). Robust Delimitation of Aggregated Packets Based on Viterbi Decoding With Natural Redundancy in Protocols, In: IEEE Signal Processing Letters, vol. 31, p. 1970-1974.

[36] Min-Won, Seo., Solmaz, S., Kia. (2023). A Computationally Efficient Maximum A Posteriori Sequence Estimationvia *SteinVariational Inference*. https://doi.org/10.48550/arXiv.2312.08684.

[37] Chuan. (2009). Liu cuHMM: a CUDA implementation of hidden Markov model training and classification, *The Chronicle of Higher Education*, p. 1-13.

[38] Sand, Andreas., Christian, Pedersen., Mailund Thomas, N. S., et al. (2010). HMMlib: A C++ library for general hidden Markov models exploiting modern CPUs, *In*: 2010 Ninth International Workshop on Parallel and Distributed Methods in Verification, and Second International Workshop on High Performance Computational Systems Biology, *IEEE* p. 126-134.

[39] Chiyoung, Ahn., Kim, June., J. U., Jaehyuk, et al. (2011). Implementation of an SDR platform using GPU and its application to a 2× 2 MIMO WiMAX system, *Analog Integrated Circuits and Signal Processing,* 69 (2) p.107-117.

[40] Buthpitiya, Senaka., Lane, Ian., Chong, Jike. (2012). A parallel implementation of Viterbi training for acoustic models using graphics processing units, *Innovative Parallel Computing* (InPar), *IEEE* p. 1-10.

[41] Maleki, Saeed., Musuvathi, Madanlal., Mytkowicz, Todd. (2014). Parallelizing dynamic programming through rank convergence, *ACM SIGPLAN Notices*, 49 (8) p. 219-232.

[42] Kashif Hanif, Muhammad., Zimmermann, Karl-Heinz. (2017). Accelerating Viterbi algorithm on graphics processing units, *Computing*, 99 (11) p. 1105-1123.

[43] Wang, L. (2021). A Markov Model-Based Fusion Algorithm for Distorted Electronic Technology Archives. Computational Intelligence and Neuroscience, 2022 (1), 4202181.

[44] Xie, T., Sheng, Q., Yu, J. (2023). Deep Learning Equalizer Connected with Viterbi-Viterbi Algorithm for PAM D-Band Radio over Fiber Link. *Sensors*, 23, 9773.

[45] Liu, Q., Wang, J., Fu, J., Liu, J., Chen, B. (2024). Named Entity Recognition Method of Legal Instruments Based on Improved Few Shot Learning, *IEEE Access,* vol. 12, p. 157444-157454.

[46] Patil, N. V., Patil, A. S. (2026). Enhanced Named Entity Recognition in Marathi News Articles Using Machine Learning Approach. *In*: Bansal, J.C., Saha, S., Coello, C.A.C., Rathore, H. (eds) Advances in Data-Driven Computing and Intelligent Systems. ADCIS 2024. Lecture Notes in Networks and Systems, vol 1445. Springer, *Singapore*.

[47] Bender, E. M., Gebru, T., McMillan-Major, A., Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big *In*: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, FAccT '21, p. 610–623, New York, NY, USA, *Association for Computing Machinery*.

[48] Chavali, S. T., Kandavalli, C. T., S. R, S. T. M. (2022). Grammar Detection for Sentiment Analysis through Improved Viterbi Algorithm, *In*: 2022 International Conference on Advances in Computing, Communication

and Applied Informatics (ACCAI), Chennai, India, p. 1-6.

[49] Chandio, A. A., Asikuzzaman, M., Pickering, M. R., Leghari, M. Cursive Text Recognition in Natural Scene Images Using Deep Convolutional Recurrent Neural Network, *IEEE Access*, vol. 10, p. 10062-10078.

[50] Futami, H., Tsunoo, E., Kashiwagi, Y., Ogawa, H., Arora, S., Watanabe, S. (2024). Phoneme-Aware Encoding for Prefix Tree Based Contextual ASR, *In*: ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Seoul, Korea, Republic of, p. 10641-10645.

[51] Sen, V.S., Jamwal, S. S. (2024). Transcending Linguistic Boundaries: A Technical Exploration of the Evolution and Future Trajectory of Corpus-Based Machine Translation 2024. *J. Electrical Systems* 20-7s (2024) 2502-2509.

[52] Khedkar, Vijayshri., Shah, Pritesh. (2022). A Survey of Machine Translation and Parts of Speech Tagging for Indian Languages, *IJCSNS International Journal of Computer Science and Network Security,* VOL.22 No.4, April 245.

[53] Täckström, Oscar., McDonald, Ryan., Nivre, Joakim. (2013). Target Language Adaptation of Discriminative Transfer Parsers, *In*: Proceedings of NAACL-HLT 2013, ACM. pages 1061–1071.