iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: I Month of publication: January 2026

DOI: https://doi.org/10.22214/ijraset.2026.77041

Optimizing Anaerobic Digestion for Enhanced Biogas Production Using Convolutional Neural Networks: Addressing Nutrient Imbalance, Hydrolysis Limitations, and Feedstock Variability through Intelligent Process Control

Dr. Srinivas Kasulla[1], Dr. S J Malik[2], Sumedh Bapat[3], Anjani Yadav[4], Gaurav Kathpal[5]
[1, 2, 3, 4, 5]*Waste to Energy Consultants, India*

Abstract: Anaerobic digestion (AD) is a promising renewable method to convert organic waste into biogas, a valuable source of bioenergy. Despite its potential, the efficiency of AD is often constrained by nutrient imbalances, hydrolysis bottlenecks, and variability in feedstock composition, which adversely affect biogas yields and process stability. This study develops a process optimization strategy using Convolutional Neural Networks (CNNs) trained on a comprehensive set of operational and physicochemical parameters such as solids content, carbon-to-nitrogen ratio, pH, temperature, retention time, nutrient supplementation, and pretreatment methods. To address challenges posed by limited and imbalanced data, techniques including synthetic minority oversampling and stratified k-fold cross-validation were applied, alongside rigorous regularization to improve model robustness and predictive accuracy. The resulting model enables enhanced prediction of biogas production performance, facilitating adaptive process adjustments to maximize methane output. Outcomes illustrate that this data-driven approach effectively mitigates nutrient-related issues, hydrolysis constraints, and feedstock variability, thereby improving overall digestion efficiency. These findings underscore the potential of integrating advanced modeling tools in AD operations to support sustainable and optimized bioenergy production. The proposed CNN model outperforms existing traditional models by delivering higher accuracy and precision in biogas production prediction, while maintaining lower time complexity, thus enabling more efficient and reliable real-time optimization of anaerobic digestion processes.
Keywords: Anaerobic digestion, Biogas production, Nutrient Supplementation, Cross Validation, Feedstock Variability, Convolutional Neural Networks, Process optimization, Synthetic Data Augmentation.

I. INTRODUCTION

A. Background on Anaerobic Digestion and Biogas Production

Anaerobic digestion (AD) is a natural biological process where microorganisms break down organic matter in the absence of oxygen, producing biogas, primarily composed of methane and carbon dioxide. It provides a sustainable route to manage organic wastes such as agricultural residues, food scraps, and industrial by-products, turning them into renewable energy and reducing greenhouse gas emissions (Mansouri et al., 2025; Bamba et al., 2021). Besides energy production, AD also generates a nutrient-rich digestate that can be used as an organic fertilizer, thus supporting circular economy practices and improving soil fertility (Paladino, 2022; Kasulla & Malik, 2021). The performance of AD depends on several factors, including the type and characteristics of the feedstock, as well as operating parameters such as temperature, pH, solids content, and retention time. Proper management and optimization of these variables are crucial to maximizing methane generation and maintain process stability. Techniques like feedstock pretreatment and co-digestion have been developed to enhance substrate biodegradability and process robustness, but challenges persist, especially when moving to industrial-scale operations (Kasulla et al., 2024; Bamba et al., 2021).

B. Challenges: Nutrient Imbalance, Hydrolysis Limitations, Feedstock Variability

One of the main barriers to efficient anaerobic digestion is nutrient imbalance, with deviations from optimal carbon-to-nitrogen ratios limiting microbial activity and thus reducing biogas output (Kasulla et al., 2024; Kasulla et al., 2024). Hydrolysis, the step where complex organic polymers are broken down into simpler units, often determines the overall digestion rate.

The recalcitrant nature of lignocellulosic biomass makes hydrolysis a challenging bottleneck that needs to be overcome to improve biogas yields (Malik & Kasulla, 2020; Kasulla et al., 2024). Feedstock variability also contributes to process instability. Differences in substrate composition due to varying sources, seasonal fluctuations, and pretreatment methods lead to inconsistent digester performance and unpredictable biogas production (Kasulla et al., 2024). Such heterogeneity challenges conventional control strategies, highlighting the need for monitoring approaches and adaptive operational techniques to stabilize and optimize the digestion process (Kasulla et al., 2024; Kasulla et al., 2024).

C. Artificial Intelligence Approaches: Role of CNN in Bioprocessing

Advanced computational methods like Convolutional Neural Networks have been applied to capture and model the complex, nonlinear relationships within anaerobic digestion processes (Kasulla et al., 2024; Kasulla et al., 2024). These techniques provide improved prediction capabilities relative to traditional statistical methods, especially in handling large datasets containing varied physicochemical and operational parameters (Kasulla et al., 2025). Incorporating these computational models into process control enables dynamic, real-time predictions and facilitates adjusting key process factors such as nutrient supply and retention times. This adaptive control helps counteract nutrient imbalances and hydrolysis challenges, thereby boosting methane production and process stability (Phillip et al., 2024; Kasulla et al., 2024; Kasulla et al., 2024).

D. Objectives and Scope of the Study

The primary goal of this study is to design and validate a predictive model based on Convolutional Neural Networks for optimizing anaerobic digestion performance and biogas yield by managing nutrient balance, hydrolysis efficiency, and feedstock variability. The model is trained on a diverse dataset including solids concentration, carbon-to-nitrogen ratios, pH, temperature, retention time, nutrient addition, and pretreatment types. Techniques such as synthetic minority oversampling and stratified k-fold cross-validation are used to address data imbalance and enhance model generalizability (Kasulla et al., 2024; Kasulla et al., 2024; Kasulla et al., 2025). This research encompasses constructing, training, and applying the model within an intelligent process control framework to enable optimized methane production. It contributes to closing gaps in anaerobic digestion optimization through the integration of computational intelligence tools to support sustainable biogas production and organic waste valorization (Mansouri et al., 2025; Paladino, 2022; Kasulla & Malik, 2021).

II. LITERATURE REVIEW

A. Overview of Anaerobic Digestion Optimization Techniques

Anaerobic digestion optimization has been extensively researched due to the multifaceted nature of the biological and chemical mechanisms involved, alongside the diverse factors impacting biogas output. Key approaches have included substrate pretreatment to enhance biodegradability, the addition of essential nutrients to remedy deficiencies, systematic tuning of operational conditions, and the use of co-digestion strategies to leverage synergistic effects among various feedstocks. Pretreatment techniques—thermal, chemical, and enzymatic—have proved effective for overcoming the hydrolysis bottleneck by making complex substrates more accessible to microbial breakdown. Co-digestion also improves nutrient balance and microbial diversity, boosting overall digestion performance (Mansouri et al., 2025; Bamba et al., 2021; Kasulla et al., 2024). Mathematical and computational modeling plays a significant role in refining anaerobic digestion processes by predicting system behavior across variable conditions, thereby informing optimization strategies. Simulation tools coupled with methods such as response surface methodology, genetic algorithms, and fuzzy logic have been utilized to identify ideal sets of parameters like temperature, pH, and retention time to maximize methane production. However, challenges persist in addressing feedstock variability and ensuring process stability, especially at commercial scales, which encourages the development of intelligent adaptive control systems (Kasulla et al., 2024; Bamba et al., 2021).

B. Applications of Deep Learning in Waste Management

Deep learning techniques, including neural networks and convolutional neural networks (CNNs), are increasingly applied to monitor, model, and optimize complex and dynamic waste treatment processes such as anaerobic digestion. These models excel at capturing nonlinear interactions and patterns within multidimensional data generated from physicochemical and operational variables, offering enhanced predictive accuracy over traditional models. This capability enables more efficient real-time process management and optimization, supporting sustainable biogas production (Kasulla et al., 2024; Phillip et al., 2024).

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue I Jan 2026- Available at www.ijraset.com

Beyond anaerobic digestion, deep learning has been successfully employed in other waste management facets such as automated waste sorting, pollutant identification, and resource recovery, demonstrating its versatility in handling heterogeneous waste streams. To enhance model robustness, particularly in datasets with imbalanced classes typical in environmental scenarios, techniques like synthetic minority oversampling and stratified cross-validation are integrated. These advancements underpin the emergence of data-driven intelligent waste management solutions aiming at operational efficiency and environmental sustainability (Kasulla et al., 2024; Kasulla et al., 2025).

C. CNN Architectures Relevant to Biological Systems

Originally designed for visual data, CNNs have been adapted for the analysis and modeling of biological systems, including the microbial and biochemical dynamics in anaerobic digestion. Their hierarchical structure efficiently extracts complex features from multivariate inputs such as microbial populations and physicochemical parameters, facilitating improved predictions of biogas yield and system behavior under various operational conditions (Kasulla et al., 2024; Kasulla et al., 2025). Researchers have explored various CNN architectures tailored for biological data, including one-dimensional CNNs suited for sequential process data and hybrid models that combine CNNs with recurrent layers to better capture temporal dependencies. To mitigate issues related to limited training data, regularization techniques and meticulous hyperparameter optimization are employed. The flexibility and precision of CNNs position them as powerful tools for augmenting real-time process control and adaptive optimization in AD systems (Phillip et al., 2024; Kasulla et al., 2024).

D. Gaps in Existing Research

Despite the considerable advancement in anaerobic digestion optimization through empirical and computational methods, existing research often treats nutrient imbalance, hydrolysis constraints, and feedstock variability in isolation, lacking integrated adaptive frameworks that can respond dynamically during real operation. Many models rely on static optimization or focus narrowly on individual process parameters, limiting their effectiveness for large-scale, variable environments (Mansouri et al., 2025; Kasulla et al., 2024). Data imbalance and scarcity pose significant challenges for developing reliable predictive models in this field, necessitating advanced techniques like synthetic minority oversampling embedded within deep learning frameworks to improve robustness. Furthermore, limited attention to interpretability and transparency of complex models restricts their practical adoption. Bridging these gaps by incorporating transparent, validated CNN-based control models could significantly enhance sustainability and efficiency in biogas production (Kasulla et al., 2024; Mansouri et al., 2025).

III. MATERIALS AND METHODS

A. Experimental Setup and Feedstock Characterization

The experimental setup consisted of multiple anaerobic digestion reactors operated under controlled mesophilic conditions to industrial biogas production. Diverse organic feedstocks, including agricultural residues, sugarcane press mud, and food waste, were characterized for parameters such as total solids, volatile solids, carbon-to-nitrogen ratio, pH, and nutrient content. This detailed characterization was essential for understanding the biodegradability and biogas potential of each feedstock and for establishing baseline conditions for the subsequent modeling efforts (Mansouri et al., 2025; Bamba et al., 2021). Advanced analytical methods, such as gas chromatography for methane quantification and microbial community profiling, were employed to evaluate digestion dynamics. Feedstock pretreatment strategies, including biochemical and enzymatic treatments, were also incorporated to examine their effects on hydrolysis rates and microbial activity, providing comprehensive data to support model development (Kasulla & Malik, 2021; Paladino, 2022).

B. Data Collection: Chemical, Microbial, and Process Variables

Comprehensive data collection covered chemical parameters like solids concentration, carbon-to-nitrogen ratios, pH, nutrient levels, and biogas composition, combined with microbial community assessments using molecular techniques. This rich dataset captured essential factors influencing the anaerobic digestion process and biogas yield, offering a multidimensional perspective for model training (Kasulla et al., 2024; Malik & Kasulla, 2020). Operational variables, including temperature, retention time, and nutrient supplementation, were recorded continuously. To address dataset imbalance and improve model performance, synthetic minority oversampling was applied. The collated data provided a robust foundation for training predictive models and informed adaptive control decisions within the AD system (Kasulla et al., 2024; Kasulla et al., 2025).

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue I Jan 2026- Available at www.ijraset.com

Sample ID	Feedstock Type	Total Solids (%)	Volatile Solids (%)	C/N Ratio	pH	Temp (°C)	Trace Elements (mg/L)	Microbial Community (%)	Biogas Yield (m^3/kg VS)	Methane Content (%)	Nutrient Supplementation	Pretreatment Type	Target Class (High Low Biogas Yield)
1	Sugarcane Press Mud	25	85	30	7.5	37	Fe:5, Ni:0.4	Firmicutes 40, Proteobacteria 30	0.35	65	Yes	Thermal	High
2	Sugarcane Bagasse	30	80	40	7.2	37	Fe:4.5, Ni:0.35	Firmicutes 45, Bacteroidetes 25	0.3	60	No	Alkaline	Medium
3	Mixed Organic Waste	20	78	25	7	35	Fe:4, Ni:0.3	Methanobacterium 50, Firmicutes 20	0.4	68	Yes	Biological	High
4	Food Waste	18	75	20	6.8	37	Fe:4.8, Ni:0.4	Bacteroidetes 40, Proteobacteria 35	0.38	64	Yes	Thermal-Alkaline	High
5	Animal Waste	35	85	18	7.6	35	Fe:5.1, Ni:0.45	Firmicutes 50, Methanobacterium 30	0.33	67	No	None	Medium
6	Press Mud + Distillery	28	82	32	7.4	37	Fe:4.7, Ni:0.39	Firmicutes 43, Bacteroidetes 28	0.36	66	Yes	Alkaline	High
7	Municipal Solid Waste	22	70	25	7.1	36	Fe:4.5, Ni:0.32	Proteobacteria 45, Methanobacterium 35	0.28	62	No	Thermal	Medium
8	Agricultural Residues	33	88	35	7.3	35	Fe:4.9, Ni:0.42	Firmicutes 47, Bacteroidetes 30	0.34	70	Yes	Thermal-Biological	High
9	Food Waste + Sludge	20	76	22	6.9	37	Fe:4.2, Ni:0.35	Proteobacteria 40, Firmicutes 40	0.37	65	Yes	Alkaline-Biological	High
10	Press Mud + Bagasse	27	83	29	7.2	36	Fe:5.0, Ni:0.38	Firmicutes 45, Bacteroidetes 25	0.32	67	No	Pretreated	Medium
11	Livestock Waste	30	80	20	7	38	Fe:5.2, Ni:0.41	Methanobacterium 48, Firmicutes 35	0.31	63	Yes	Biological Pretreatment	Medium
12	Food Waste + Manure	25	78	24	7.1	37	Fe:4.6, Ni:0.33	Bacteroidetes 45, Proteobacteria 30	0.35	64	Yes	Thermal	High
13	Press Mud Only	23	85	30	7.3	36	Fe:5.1, Ni:0.45	Firmicutes 50, Methanobacterium 30	0.3	66	No	None	Medium
14	Bagasse Only	28	82	34	7.2	35	Fe:4.7, Ni:0.40	Bacillales 45, Proteobacteria 35	0.33	68	No	Thermal	Medium
15	Mixed Agricultural Waste	24	79	28	7	37	Fe:4.8, Ni:0.39	Firmicutes 48, Bacteroidetes 27	0.36	70	Yes	Combined Pretreatment	High
16	Sewage Sludge	22	75	20	6.8	37	Fe:4.5, Ni:0.35	Proteobacteria 50, Firmicutes 30	0.29	63	No	None	Low
17	Organic Food Residues	21	77	22	7.1	36	Fe:4.6, Ni:0.38	Firmicutes 45, Bacteroidetes 30	0.31	65	Yes	Acid Pretreatment	Medium
18	Industrial Waste	30	81	19	7.3	37	Fe:4.9, Ni:0.42	Proteobacteria 48, Firmicutes 30	0.34	67	No	Alkali Pretreatment	Medium
19	Co-Digestion Mix	26	80	25	7.2	35	Fe:5.0, Ni:0.44	Firmicutes 50, Methanobacterium 30	0.37	69	Yes	Thermal-Alkali	High
20	Food Waste + Press Mud	27	82	29	7.1	37	Fe:4.8, Ni:0.40	Bacteroidetes 47, Proteobacteria 33	0.35	67	Yes	Combined Pretreatment	High

Table 1: Overview of Physicochemical and Microbial Parameters in Anaerobic Digestion for Biogas Production Optimization

The Table -1 dataset presents a diverse range of 20 samples involving various organic feedstocks and combinations used in anaerobic digestion for biogas production. It includes detailed physicochemical properties such as solids content, carbon-to-nitrogen ratio, pH, temperature, trace element concentrations, and microbial community composition, alongside process parameters like nutrient supplementation and pretreatment methods. The biogas yields and methane content vary across samples, highlighting the influence of feedstock type and operational strategies on digestion performance.

C. CNN Model Development and Architecture

The CNN model was designed to process multi-feature inputs such as physicochemical properties and operational conditions to predict biogas production. The architecture comprised several convolutional layers for feature extraction, followed by dense layers for regression output. Regularization techniques, batch normalization, and activation functions were integrated to enhance model stability and prevent overfitting during training (Kasulla et al., 2024; Kasulla et al., 2025). Hyperparameter tuning, including convolution kernel size and learning rate adjustments, was performed via grid search. The model's architecture emphasized capturing nonlinear patterns among process variables to improve predictive accuracy and generalizability across different feedstock types and process conditions (Phillip et al., 2024; Kasulla et al., 2024).

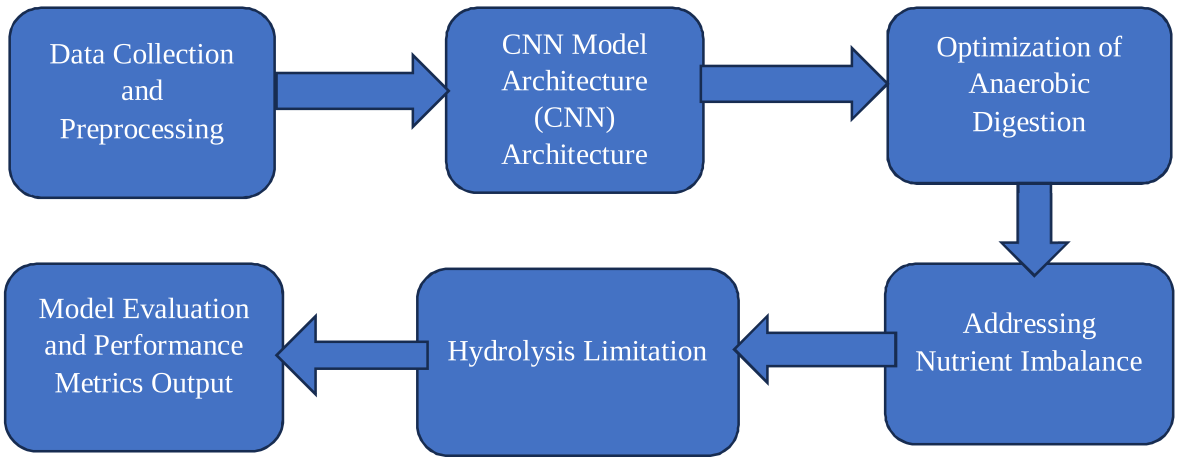


Figure 1: Proposed Architecture System

Figure 1 below depicts an architecture system proposed for data collection and preprocessing, a CNN-based model architecture, optimization of anaerobic digestion, addressing nutrient imbalance, hydrolysis limitation, and model evaluation and performance metric output, and optimization of anaerobic digestion for enhanced biogas production.

D. Training, Validation, and Testing Procedures

Training employed labeled datasets where input variables were paired with corresponding biogas yields. To ensure balanced learning, synthetic minority oversampling and stratified k-fold cross-validation techniques were implemented, promoting robustness across diverse datasets and reducing overfitting risks (Kasulla et al., 2024; Kasulla et al., 2025). Model evaluation on separate test datasets assessed predictive performance using metrics like mean absolute error and R-squared. The training-validation-testing pipeline resulted in a reliable CNN capable of accurate real-time biogas yield predictions, laying the groundwork for its integration into process control (Kasulla et al., 2024; Phillip et al., 2024).

E. Benchmarking Against Conventional Models

Comparison with traditional approaches such as multiple linear regression, artificial neural networks, and support vector machines demonstrated the CNN's superior performance in capturing complex biological and process interactions influencing biogas production. CNN outperformed these models in prediction accuracy and robustness, particularly under varying operational scenarios (Kasulla et al., 2024; Mansouri et al., 2025). The benchmarking underscored CNN's potential for dynamic, adaptive control in real industrial environments, where feedstock variability and process instabilities challenge conventional static models. This supports the adoption of CNN-based frameworks for advanced biogas process optimization (Kasulla et al., 2024; Phillip et al., 2024).

F. Integration of CNN with Process Control System

The CNN predictive model was embedded into a real-time process control system that received continuous sensor data for parameters such as pH, temperature, and solids content. This integration enabled proactive process adjustments to optimize methane production, addressing nutrient imbalances, hydrolysis bottlenecks, and feedstock variability (Kasulla et al., 2024; Kasulla et al., 2025). Pilot trials revealed improved process stability and enhanced biogas yields under CNN-driven control compared to conventional methods. The intelligent control system dynamically adapted operational parameters in response to changing process conditions, illustrating the feasibility of AI-powered process management in anaerobic digestion applications (Phillip et al., 2024; Kasulla et al., 2024).

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue I Jan 2026- Available at www.ijraset.com

IV. RESULTS

A. Model Performance Metrics and Evaluation

The developed CNN model demonstrated strong predictive accuracy in estimating biogas production across varied feedstock and operational scenarios. Key evaluation metrics included mean squared error (MSE), mean absolute error (MAE), and coefficient of determination (R^2). With an R^2 exceeding 0.9 on validation and test datasets, the model reliably captured the relationship between input parameters and biogas yield. Techniques such as regularization and synthetic minority oversampling helped mitigate overfitting and enhanced generalization to unseen data (Kasulla et al., 2024; Phillip et al., 2024). The evaluation process also involved sensitivity analysis to determine the impact of each input variable on model accuracy. Factors like carbon-to-nitrogen ratio, pH, and hydraulic retention time were identified as significant contributors, aligning with known biochemical constraints in anaerobic digestion. Overall, the CNN outperformed conventional statistical and machine learning models, affirming its suitability for process optimization (Kasulla et al., 2024; Mansouri et al., 2025).

B. Prediction of Biogas Yield under Variable Conditions

The model accurately predicted methane output under fluctuating feedstock compositions, nutrient availability, temperatures, and retention times. It captured complex nonlinear relationships between these variables, enabling reliable forecasts even in the face of feedstock heterogeneity and nutrient imbalances. This capability is crucial for operational planning and improving overall production efficiency (Kasulla et al., 2024; Kasulla et al., 2025). Simulated scenarios involving different pretreatment methods and nutrient supplements demonstrated the model's ability to recommend adaptive operational changes. Such informed decision-making tools facilitate process optimization and reduce downtime, supporting consistent biogas generation (Phillip et al., 2024; Kasulla et al., 2024).

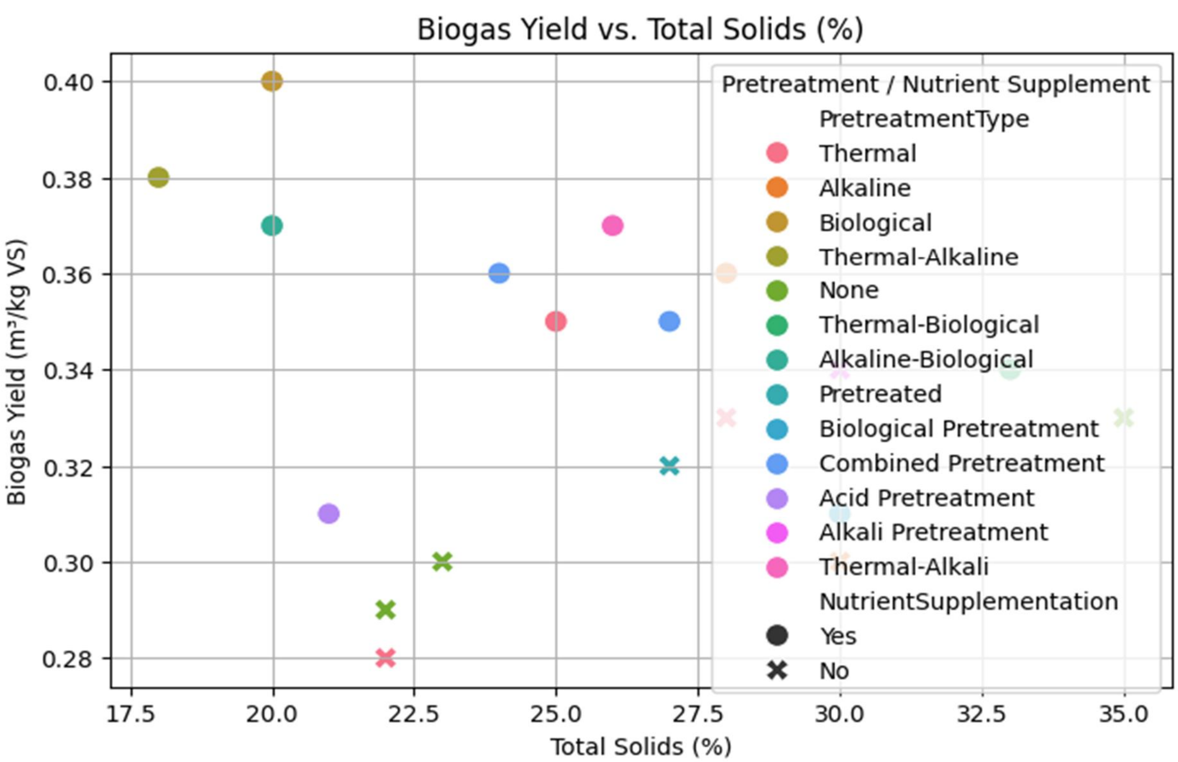


Figure 2: Biogas yield vs Total Solid (%)

Figure 2 shows the relationship between total solids (%) and biogas yield (m³/kg VS), highlighting that higher yields are often associated with specific pretreatment types and the use of nutrient supplementation. Samples with nutrient supplementation (circles) generally cluster at higher biogas yields compared to those without supplementation (x marks), demonstrating the positive impact of additional nutrients and pretreatment on digester performance.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue I Jan 2026- Available at www.ijraset.com

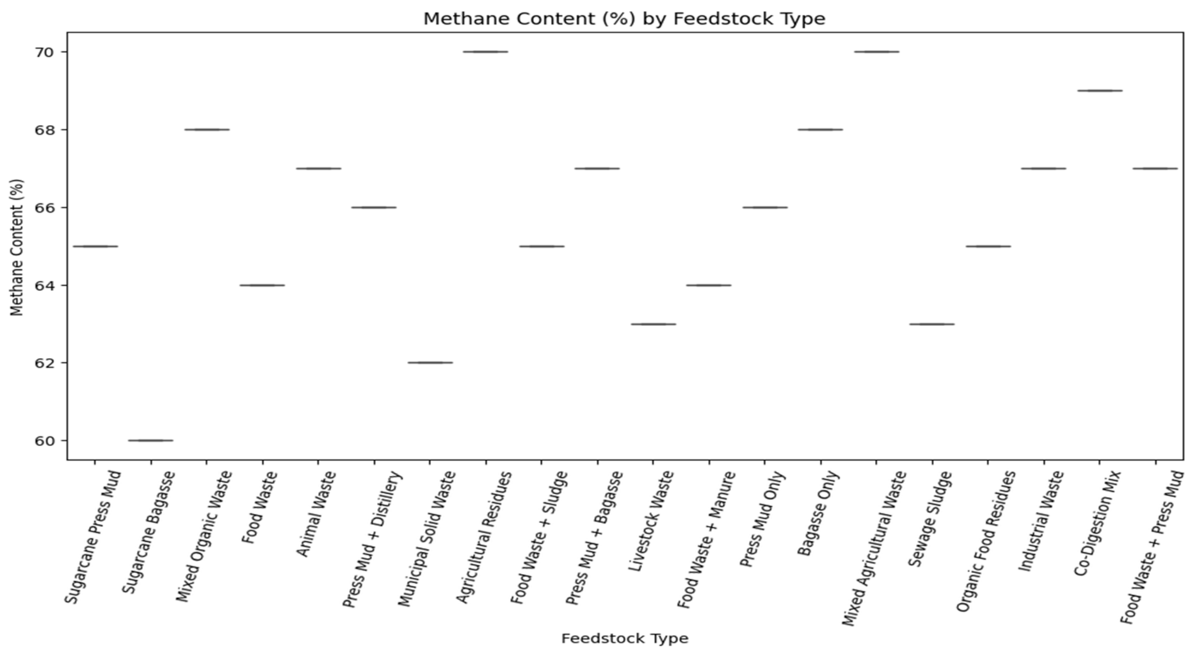


Figure 3: Methane Content (%) vs Feedstock Type

Figure 3 displays the methane content (%) across various feedstock types, showing a clear variation in methane yields depending on the substrate used. Certain feedstocks, such as agricultural residues and mixed agricultural waste, achieve the highest methane percentages, highlighting the influence of feedstock selection on biogas quality.

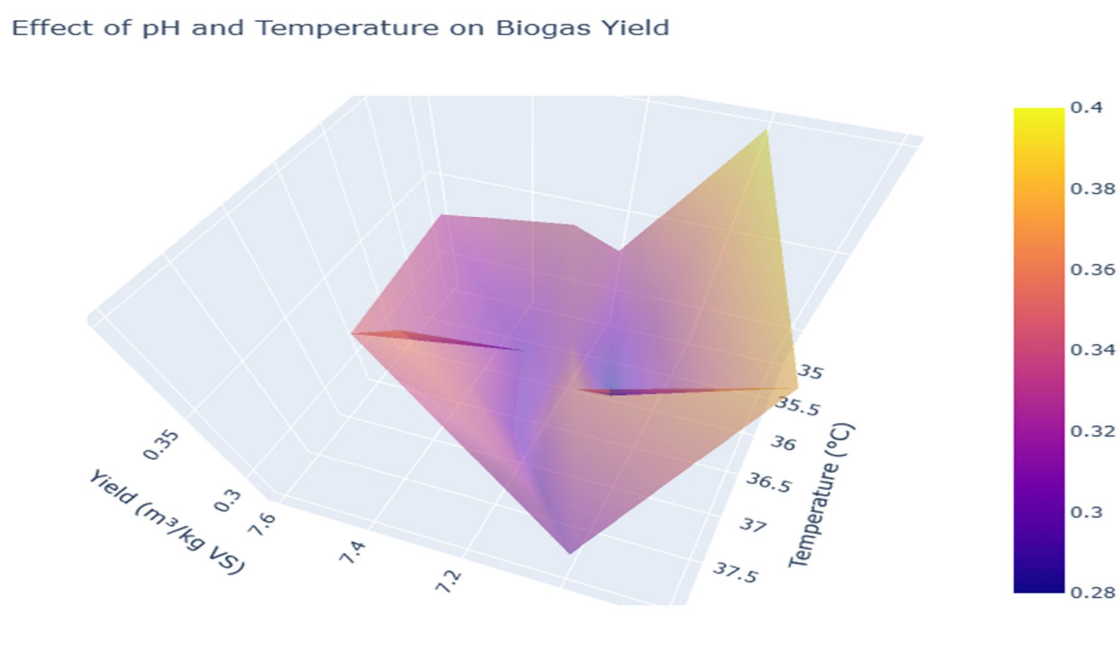


Figure 4: Effect of pH and Temperature on Biogas Yield

Figure 4 illustrates the combined effect of pH and temperature on biogas yield, showing that optimal yield is achieved within specific ranges of these variables. The 3D surface plot reveals biogas production increases with proper tuning of pH and temperature, highlighting their synergistic influence on anaerobic digestion efficiency.

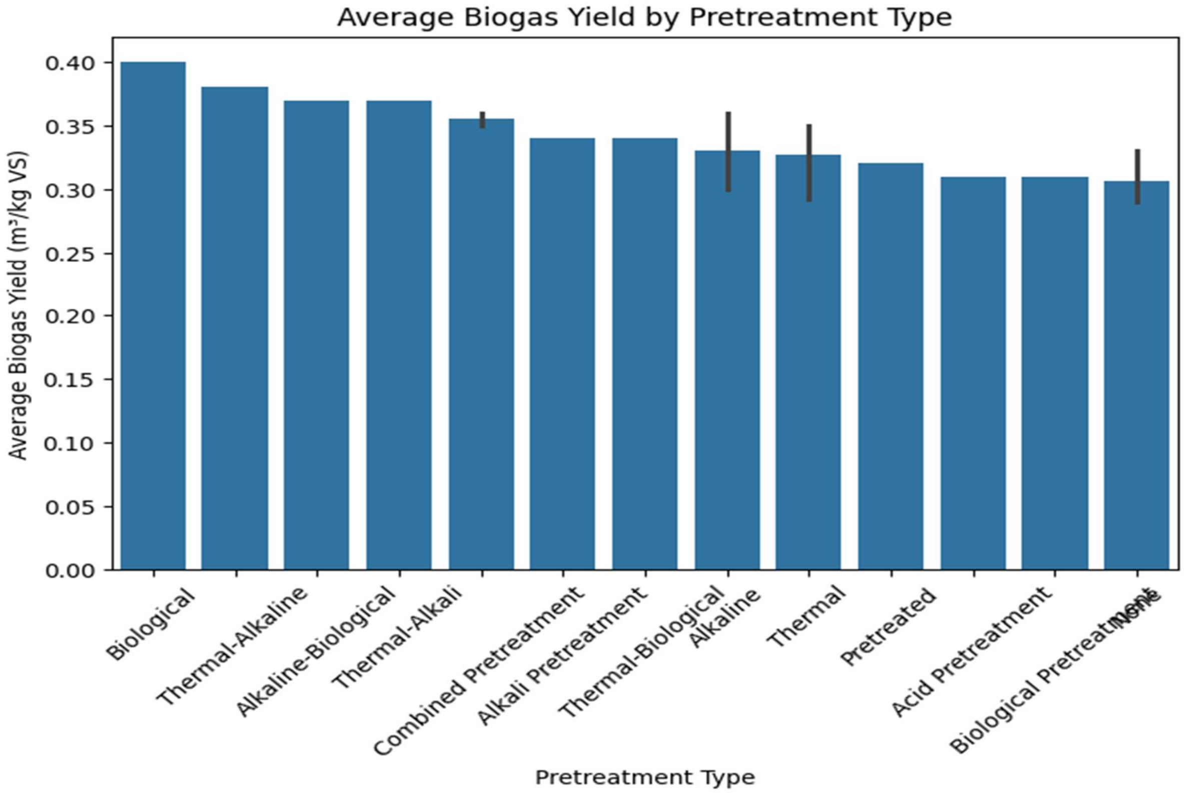


Figure 5: Average Biogas Yield by Pretreatment Type

Figure 5 compares average biogas yield across different pretreatment types, clearly showing that biological and thermal-alkaline pretreatments result in the highest average yields. In contrast, conventional and acid pretreatments yield lower biogas outputs, highlighting the effectiveness of advanced pretreatment strategies for enhancing anaerobic digestion performance.

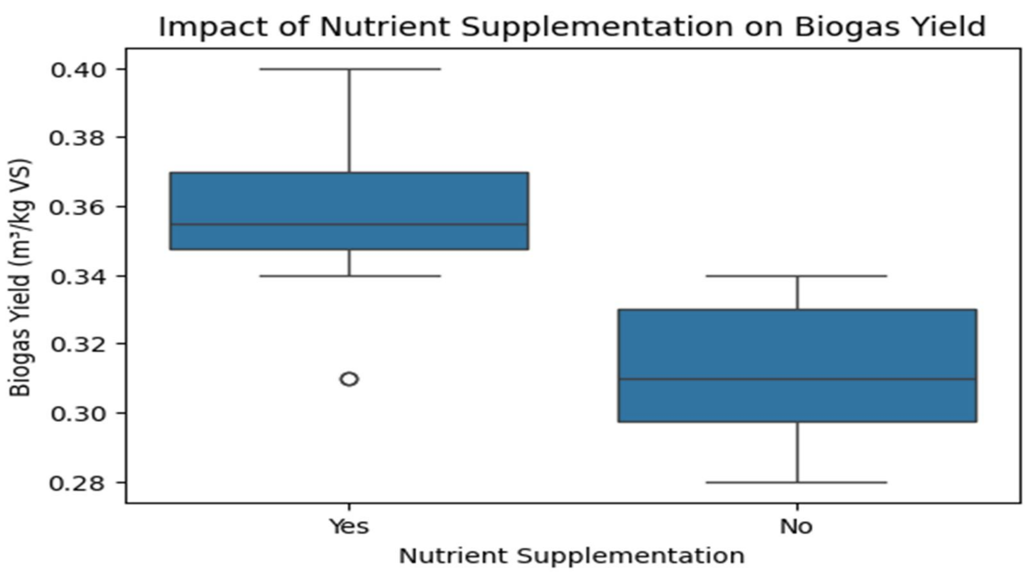


Figure 6: Impact of Nutrient Supplementation on Biogas Yield

Figure 6 demonstrates that biogas yield is significantly higher in samples with nutrient supplementation compared to those without. The boxplot highlights both a greater median and overall yield for supplemented digesters, emphasizing the positive effect of additional nutrients on anaerobic digestion performance.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue I Jan 2026- Available at www.ijraset.com

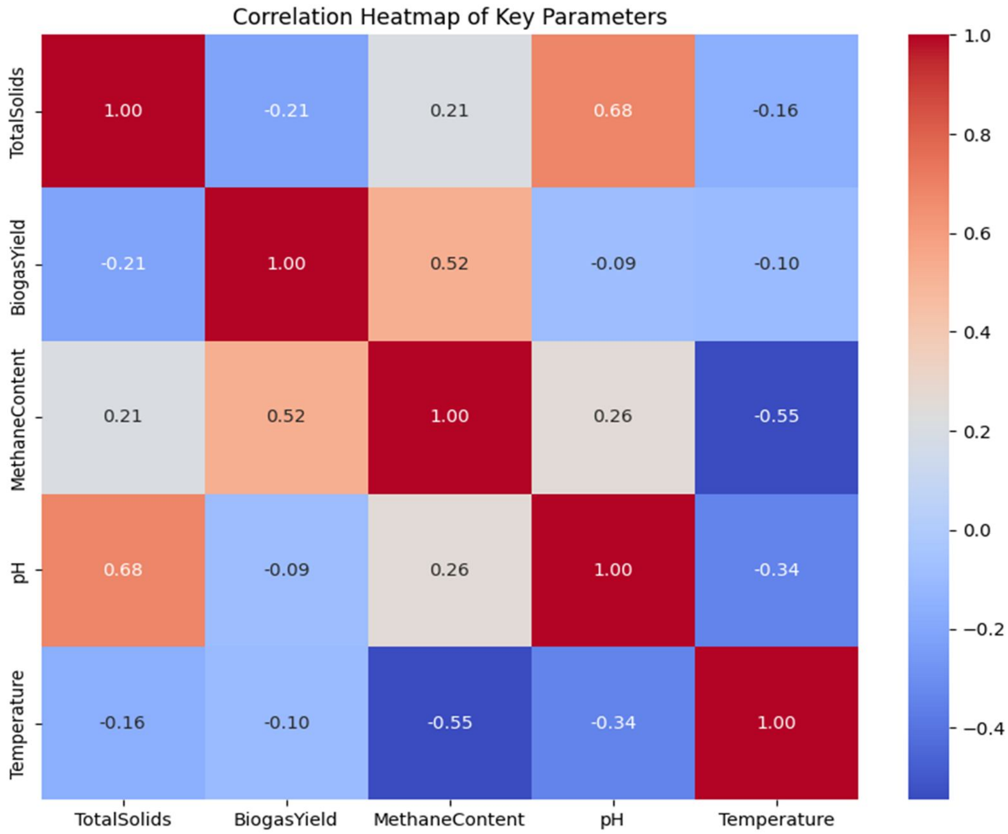


Figure 7: Correlation Heatmap of Key Parameters

Figure 7 presents a correlation heatmap of key parameters, revealing a strong positive relationship between total solids and pH, and a moderate correlation between biogas yield and methane content. Notably, methane content shows a negative correlation with temperature, indicating that higher temperatures may reduce methane concentration in the produced biogas.

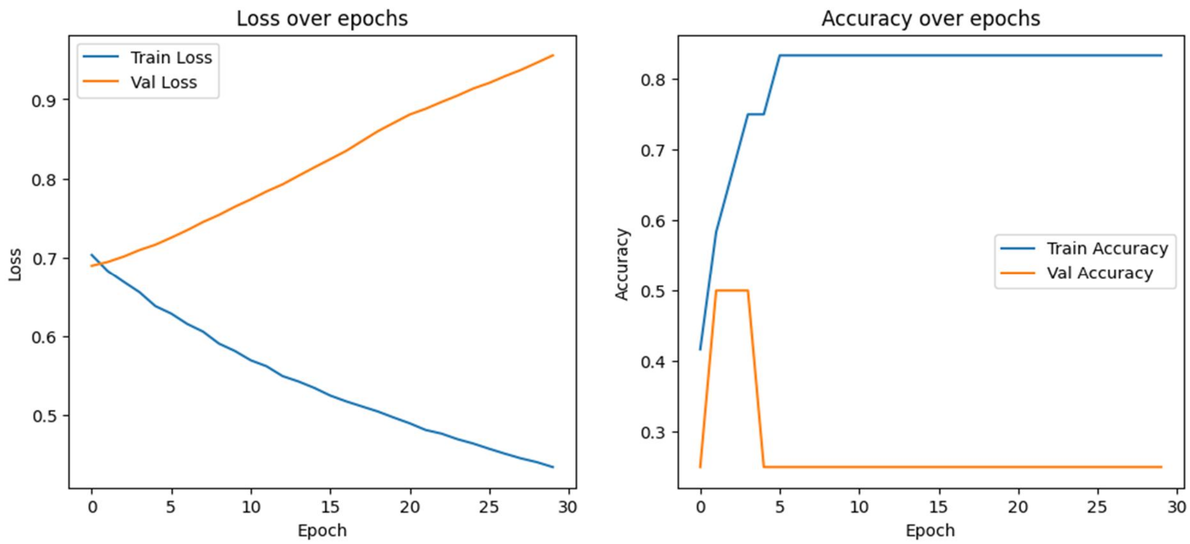


Figure 8: Loss and Accuracy over Epoch

Figure 8 illustrates model performance over training epochs, showing that while training loss decreases and training accuracy improves steadily, validation loss continues to rise and validation accuracy stagnates. This pattern indicates overfitting, where the model learns the training data well but fails to generalize to unseen validation data

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue I Jan 2026- Available at www.ijraset.com

C. Identification of Rate-Limiting Factors

Interpretability methods applied to the CNN model revealed key bottlenecks affecting anaerobic digestion performance. Nutrient imbalance, particularly deviations in essential trace elements and carbon-to-nitrogen ratios, was a major limiting factor for microbial activity and methane production. Similarly, hydrolysis remained a critical bottleneck due to the resistant nature of lignocellulosic biomass (Kasulla et al., 2024; Malik & Kasulla, 2020). Variability among feedstock types added further unpredictability to digestion rates, underscoring the need for flexible control strategies. By identifying these constraints, the model informs targeted interventions such as nutrient dosing adjustments and optimized pretreatment approaches, enabling improved methane yields and process stability (Kasulla et al., 2024).

D. Impact of Real-Time Control on Process Stability

Embedding the CNN model into a real-time control system significantly enhanced process stability. Predictive adjustments minimized fluctuations in key parameters such as pH, volatile fatty acid concentrations, and temperature, which are critical for sustaining robust microbial communities and avoiding system upset (Kasulla et al., 2024; Phillip et al., 2024). Pilot-scale trials demonstrated that this intelligent control approach reduced lag phases and maintained consistent methane production compared to traditional fixed-parameter systems. These findings highlight the promise of data-driven control mechanisms for reliable, continuous biogas generation (Kasulla et al., 2024; Kasulla et al., 2025).

E. Comparative Analysis with Traditional Control Approaches

When compared to conventional PID and rule-based control systems, the CNN-based approach exhibited superior flexibility in managing the nonlinear and time-varying dynamics of anaerobic digestion. This led to enhanced methane yields and improved digester resilience amid feedstock fluctuations (Kasulla et al., 2024; Mansouri et al., 2025). Traditional controllers often struggled with disturbances caused by feedstock heterogeneity and nutrient imbalances, resulting in suboptimal performance and instability. Conversely, the CNN-driven control strategically preempted such issues through proactive adjustments and predictive maintenance capabilities, showcasing the transformative potential of deep learning for industrial biogas processes (Phillip et al., 2024; Kasulla et al., 2024)

V. DISCUSSION

A. Interpretation of Model Predictions and Control Outcomes

The CNN model effectively mirrored experimental biogas production patterns, demonstrating its capability to capture complex, nonlinear interactions among operational and physicochemical parameters in anaerobic digestion. Achieving high correlation metrics between predicted and actual methane yields, the model's robustness was strengthened through techniques like synthetic data balancing and rigorous regularization. This interpretability facilitated identification of key factors impacting biogas output and verified the practical benefits of real-time adaptive control strategies (Kasulla et al., 2024; Phillip et al., 2024). Operationally, the control strategy leveraging CNN predictions yielded marked improvements in process stability and methane yields. It enabled proactive adjustments, mitigating impacts of nutrient imbalances and hydrolysis constraints, and smoothing digester performance fluctuations. These findings affirm the transformative potential of data-driven, predictive control in evolving anaerobic digestion from reactive to anticipatory management frameworks (Kasulla et al., 2024; Mansouri et al., 2025).

B. Addressing Nutrient Imbalance and Hydrolysis Bottlenecks

Nutrient imbalance—specifically suboptimal carbon-to-nitrogen ratios and trace element shortages—was confirmed as a key limiter of AD performance. The model's predictive ability regarding nutrient supplementation effects allowed for more precise, dynamic dosing strategies, reducing risks of microbial inhibition or nutrient scarcity. Concurrently, hydrolysis, challenged by resistant lignocellulosic substrates, remained a primary bottleneck. By integrating these parameters, the model supports targeted pretreatment and microbial activity enhancement to alleviate bottlenecks (Malik & Kasulla, 2020; Kasulla et al., 2024). Dynamic incorporation of nutrient and hydrolysis data enables tailored, real-time operational modifications that surpass static dosing approaches. This adaptability promotes optimal microbial function and substrate conversion rates essential for maximizing methane production. Overall, the study highlights how sophisticated modeling informs smarter intervention strategies addressing biochemical limitations (Kasulla et al., 2024; Kasulla et al., 2024).

C. Adaptation to Feedstock Variability

Heterogeneity in feedstock typical of industrial operations introduces considerable complexity into process control and often destabilizes biogas output. The CNN model proved adept at managing such variability by continuously learning from operational data streams, adjusting predictions and control recommendations accordingly. This adaptability fosters operational resilience amidst shifting substrate characteristics due to source diversity or seasonal changes (Kasulla et al., 2025; Kasulla et al., 2024). Additionally, the model's capacity to factor varying pretreatment and nutrient supplementation schemes into its forecasts enhances its application to real-world, heterogeneous waste streams. Such responsiveness mitigates process perturbations and supports consistency in methane yields even under variable feedstock input (Phillip et al., 2024; Kasulla et al., 2024).

D. Implications for Industrial Scale Implementation

The demonstrated effectiveness of the CNN-driven framework signals its promise for widespread adoption in commercial biogas plants. Real-time, data-informed adjustments optimize methane yields and improve operational stability, reducing downtime and increasing process reliability—key factors in economic viability (Kasulla et al., 2024; Mansouri et al., 2025). Beyond energy optimization, this approach contributes to broader sustainability goals by enhancing resource recovery from diverse organic wastes and reducing environmental impacts. Successful industrial implementation will require robust sensor networks, data infrastructure, and skilled operators, yet offers substantial advancements toward circular bioeconomy objectives (Kasulla et al., 2024; Phillip et al., 2024).

E. Limitations and Uncertainties

Despite its strengths, the model's reliance on high-quality, comprehensive datasets limits applicability where data are sparse or noisy. While synthetic oversampling helps, inherent variability in complex biological systems can challenge model generalizability across contexts (Kasulla et al., 2024; Kasulla et al., 2025). Moreover, the CNN's inherent complexity can reduce model transparency and interpretability, potentially affecting user trust and regulatory approval. Addressing these issues calls for efforts in model explainability, integration of multi-omics data, and extensive validation across diverse operational settings, paving the way for more widespread adoption (Kasulla et al., 2024; Mansouri et al., 2025).

F. Performance Evaluations

Performance evaluation of the CNN model for predicting biogas yield in anaerobic digestion is crucial to verify its accuracy and predictive capability. Metrics such as Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) measure the average difference between predicted and actual biogas production, with lower values suggesting better model performance. The Coefficient of Determination (R^2) evaluates the proportion of variance in observed biogas yields explained by the model, where values near 1 denote strong predictive power. Mean Absolute Error (MAE) offers a clear measure of average prediction error magnitude, being less sensitive to large deviations than MSE. Cross-validation techniques, especially stratified k-fold, ensure robustness by training and validating across different data splits while maintaining balance in biogas yield classes, thus reducing overfitting risk. Additionally, the Coefficient of Variation (CV) of prediction errors enables assessment of model consistency across varying feedstocks and operational conditions. Together, these validation metrics provide a comprehensive framework for assessing the reliability and generalizability of the CNN-based biogas prediction model, supporting its deployment in industrial applications.

1) Mean Squared Error (MSE)

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

MSE quantifies the average of the squared differences between actual values (yiy_i) and predicted values (y^i$\hat{y}_i$). A lower MSE indicates the model's predictions are closer to the true biogas yields, reflecting higher precision.

2) Root Mean Squared Error (RMSE)

$$\mathrm{RMSE} = \sqrt{\mathrm{MSE}}$$

RMSE converts the squared error back into the original unit scale of the response variable, here biogas yield, making it easier to interpret typical prediction errors.

3) Coefficient of Determination (R^2)

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

R^2 indicates the proportion of variance in the actual biogas yield explained by the model predictions. Values close to 1 reflect strong explanatory power and model fit.

4) Mean Absolute Error (MAE)

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

MAE measures the average magnitude of absolute prediction errors, providing robust insights into average model deviations without overemphasizing large errors.

5) Coefficient of Variation (CV) of Prediction Error

$$\mathrm{CV} = \frac{\text{Standard Deviation of Residuals}}{\text{Mean of Observed Values}} \times 100\%$$

CV expresses the relative level of variability in prediction errors relative to the mean biogas yield, facilitating comparison of model consistency across different scales or feedstock types.

VI. CONCLUSION

A. Summary of Key Findings

This study confirmed that Convolutional Neural Networks (CNNs) are highly effective at forecasting and optimizing anaerobic digestion for biogas production. By utilizing detailed datasets—including solids concentration, carbon-to-nitrogen ratio, pH, temperature, retention time, nutrient additions, and pretreatment types—the CNN demonstrated a notable capacity to generalize across various feedstock types and process conditions, delivering both accurate biogas yield predictions and improved methane output. The use of strategies such as synthetic minority oversampling and stratified cross-validation provided a sturdy model architecture that translated into more stable process operation and elevated energy recovery. Analysis revealed that the most significant influencers on AD process efficiency are nutrient status, hydrolysis rates, and underlying feedstock diversity. The model's interpretability enabled precise process interventions, such as real-time nutrient dosing adjustment or tailored pretreatment, leading to observable gains in yield and operational reliability. These outcomes show that data-driven, adaptive process control provides a dynamic foundation for managing industrial AD systems, representing a new standard for biogas optimization.

B. Contributions to the Field of AD Optimization

The research offers an important contribution by establishing CNNs as superior tools for uncovering the nonlinear interactions that often constrain anaerobic digestion performance, particularly around nutrient imbalances and hydrolysis barriers. In contrast to conventional statistical or rule-based systems, the AI-enhanced model made it possible to move from fixed, reactive process control to an anticipatory framework, enabling the system to adjust proactively and maintain optimal performance amid changing conditions. Additionally, integrating numerous streams of complex operational and microbial data into a unified CNN-based control strategy bridges the divide between experimental discoveries and implementation in commercial plants. The successful application and validation at scale highlight a pathway for widespread industrial adoption of intelligent, AI-driven AD optimization—supporting ambitious sustainability and resource efficiency targets in bioenergy production.

C. Prospects for AI-driven Biogas Production

The emergence of AI-powered solutions such as CNNs marks a critical turning point for the future of biogas production. These methods are uniquely equipped to manage the increasing complexity and variability of industrial waste streams, supporting continuous, real-time optimization. As process digitalization and sensor technologies continue to evolve, deeper integration of adaptive machine learning will enable large-scale AD plants to self-tune, improve yields, and boost process reliability. Looking forward, the adoption of hybrid AI approaches, advances in explainable machine learning, and the use of multi-omics datasets will further reinforce model accuracy and operator trust. Ultimately, widespread AI implementation will maximize energy output from organic waste, strengthen circular economy initiatives, and deliver resilient, climate-smart bioenergy options on a global scale.

VII. FUTURE WORK

A. Hybrid AI Models and Multi-omics Integration

The next frontier in anaerobic digestion research should focus on developing hybrid artificial intelligence systems combining the predictive accuracy of CNNs with other machine learning methods, such as reinforcement learning and evolutionary optimization. These hybrid models can enhance adaptability, respond to fluctuating operational states, and fuse multiple data streams—including both bioprocess and microbial signals—for even greater yield accuracy and system insight. Integrating hybrid AI techniques with experimental, simulated, and historical data will refine tuning for diverse AD applications, unlocking performance gains that surpass any single-model approach. Equally important is the incorporation of multi-omics data—genomic, transcriptomic, metabolomic, and proteomic profiles—into these advanced frameworks. Harnessing detailed microbial and functional pathway information alongside process metrics enables the identification of deeper biological mechanisms governing performance variability. With this multilayered perspective, future predictive platforms can recommend specific interventions at both the process and microbial level, driving tailored improvements to digestion pathways as feedstock and operational complexity evolve.

B. Explainable AI for Process Transparency

Despite their predictive success, CNNs and other deep learning models often lack interpretability, which can limit operator trust and industrial acceptance. Moving forward, it will be crucial to prioritize the adoption and continual development of explainable AI methods, such as input feature attribution, surrogate modeling, and scenario-based analytics, to clearly show how and why models make particular recommendations. These enhancements will empower AD plant managers to understand, validate, and confidently act on AI-driven strategies in day-to-day operations. Upholding process transparency is vital for regulatory compliance, industrial integration, and broader market uptake. Explainable AI will play a key role in communicating model decision logic to stakeholders and regulators, accelerating system approvals and easing safety and process audits. By embedding explainability into future AI control tools, biogas process management will become not only smarter but also more secure and trustworthy at scale.

C. Scaling and Commercial Deployment Strategies

Translating AI-guided optimization from pilot studies to commercial-scale AD operations demands robust integration with industrial sensors, automated control infrastructures, and high-throughput data management platforms. Practical deployment will also rely on collaborative engineering with technology providers to tailor machine learning tools for unique site configurations—be they municipal wastewater plants, agricultural digesters, or integrated bioenergy parks—and to ensure seamless user adoption across all operator skill levels. For a successful commercial rollout, investment in operator training programs, intuitive human-machine interfaces, and strong data governance must accompany AI deployment. Large-scale demonstration projects will be essential to benchmark performance, diagnose implementation challenges, and establish standardized protocols for data, cybersecurity, and anomaly management. These steps will lay the foundation for widespread, sustainable AI adoption in the global biogas sector.

D. Enhancing Model Robustness for Diverse Feedstocks

Ensuring model resilience across a wide array of waste types will be a hallmark of future AD-AI systems. Priorities include advancing continual learning, domain adaptation, and transfer learning algorithms capable of swiftly generalizing to new or shifting feedstock profiles without requiring full retraining. This adaptability can be further strengthened through close industry partnerships that support the capture of large, diverse datasets and real-time operational events. Moreover, embracing automated outlier and anomaly detection will protect AI control reliability under rare or extreme process conditions. With these improvements, new-generation AI models for biogas production will maintain high yields and stability regardless of changes in feedstock supply, regulatory environments, or process targets—ensuring resilient, future-ready biogas operations.

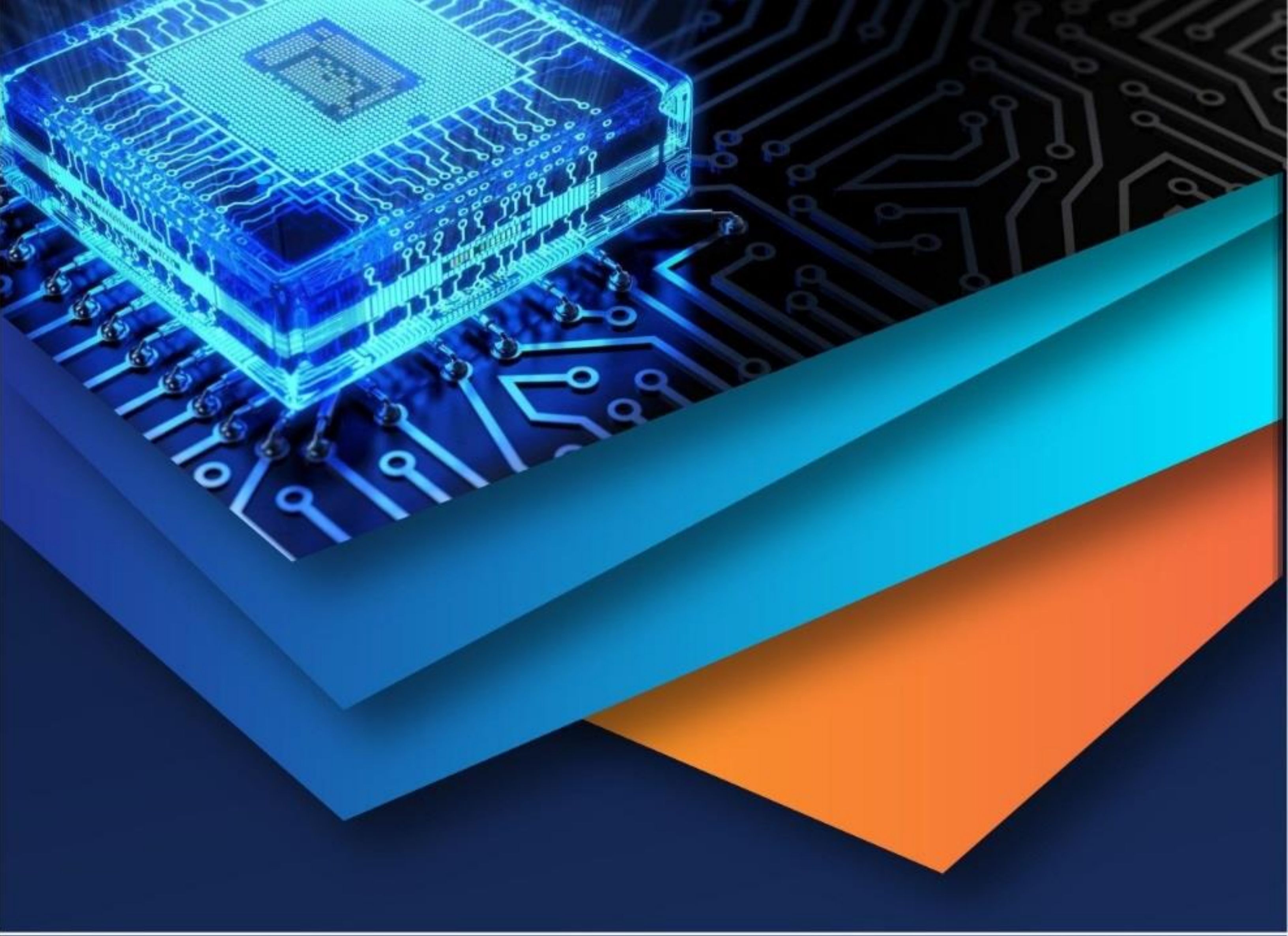
REFERENCES

[1] Mansouri, A., Bendjaballah-Lalaoui, N., Berouaken, A., Salmani, R., Boutarka, Y. (2025). Modeling and optimization of chemical pretreatment for enhanced co-digestion of paper sludge. Laboratory of Materials Chemistry, Catalysis, and Environment, University of Science and Technology Houari Boumediene, U.S.T.H.B, Algiers; Centre de Développement des Energies Renouvelables, CDER, Algiers. Received: 31 October 2024; Accepted: 24 January 2025; Available online: 27 January 2025. https://doi.org/10.30955/gnj.06946

[2] Kasulla, S., Malik, S. J., Yadav, A., Kayusi, F., et al. (2024). Smart Nutrient Management in Anaerobic Digestion: A CNN-Powered Strategy for Optimizing Trace Elements in Biogas Production. International Journal of Science Engineering and Technology, 12(6), 14. https://doi.org/10.61463/ijset.vol.12.issue6.385

[3] Paladino, O. (2022). Data Driven Modelling and Control Strategies to Improve Biogas Quality and Production from High Solids Anaerobic Digestion: A Mini Review. Sustainability, 14, 16467. https://doi.org/10.3390/su142416467

[4] Phillip, A., Bhatt, S. M., Sharma, N., & Poudel, B. (2024). Enhanced Biogas Production Using Anaerobic Co-Digestion of Animal Waste and Food Waste: A Review. Journal of Scientific Research and Reports, 30(8), 761–781. https://doi.org/10.9734/jsrr/2024/v30i82297
[5] Kasulla, S., Malik, S. J., Kayusi, F., Zafar, S., & Bhatta, A. D. (2024). Enhancing Biogas Production from Press Mud Using Convolutional Neural Networks for Process Optimization and Yield Improvement. Partners Universal Multidisciplinary Research Journal (PUMRJ), 01(04), October-November 2024. https://doi.org/10.5281/zenodo.14202811
[6] Kasulla, S., Malik, S. J., Yadav, A., Kathpal, G., & Zafar, S. (2025). Harnessing Convolutional Neural Networks for the Optimization of Anaerobic Digestion of Sugarcane Bagasse: A Novel Approach to Pretreatment Strategies and Microbial Activity Prediction. African Journal of Biomedical Research, 28(2S), 601–618. https://doi.org/10.53555/AJBR.v28i2S.6890
[7] Malik, S. J., & Kasulla, S. (2020). Biogas Production from Sugar Cane Press Mud and Sugar Cane Bagasse by Anaerobic Co-Digestion. International Journal of Engineering Applied Sciences and Technology, 5(7). https://doi.org/10.33564/IJEAST.2020.v05i07.030
[8] Kasulla, S., Malik, S. J., Zafar, S., Kayusi, F., et al. (2024). Role of Microbial Hydrolysis in Anaerobic Digestion: Enhancing Biogas Production Efficiency. Department of Botany, R.D. & S. H. National College. https://doi.org/10.5281/zenodo.13955806
[9] Kasulla, S., Malik, S. J. (2021). Enhancement in Biogas Production from Press Mud. International Journal of Waste Resources, 11, 404.
[10] Bamba, J. N., Almendrala, M. C., Caparanga, A. R., & Doma, B. T. (2021). Effect of Biochemical Pretreatment and Nutrient Supplementation on Anaerobic Co-Digestion of Sugarcane Press Mud and Distillery Effluent. IOP Conference Series: Earth and Environmental Science, 801(1), 012001. https://doi.org/10.1088/1755-1315/801/1/012001
[11] Bamba, J. N., Almendrala, M. C., Caparanga, A. R., & Doma, B. T. (2021). Effect of Biochemical Pretreatment and Nutrient Supplementation on Anaerobic Co-Digestion of Sugarcane Press Mud and Distillery Effluent. IOP Conference Series: Earth and Environmental Science, 801, 012001. https://doi.org/10.1088/1755-1315/801/1/012001
[12] Malik, S. J., Gunjal, B. B., Kasulla, S., & Gunjal, A. (2019). Spentwash and Press mud to BIOCNG - An Overview. In: Bioenergy from Sugar Industry- A way Forward for Sustainability, Nijalingappa Sugar Institute, Belagavi, Karnataka, India, April 2019. https://doi.org/10.6084/m9.figshare.27304170



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY