Advances in Nonlinear Variational Inequalities
ISSN: 1092-910X
Vol 27 No. 2 (2024)

Design of an Improved Model for Lumpy Skin Disease Prediction Using Graph Neural Networks and Multiple Modal Data Fusion

Sandeep Sharma[1], **Kapil Joshi**[2], **Rajesh Singh**[3], **Priyanka Sumit Chetwani**[4], **Amandeep Kaur**[5], **Gaurav Kumar**[6]

[1]Department of Computer Science and Engineering, UIT, Uttaranchal University Dehradun, India. sanintel123@gmail.com

[2]Department of Computer Science and Engineering, UIT, Uttaranchal University Dehradun, India. kapilengg0509@gmail.com

[3]Department of Research and Innovation, Uttaranchal University Dehradun, India. drrajeshsingh004@gmail.com

[4]Dr. D Y Patil Institute of Technology Pimpri, Pune, Maharashtra, India. priyanka.chetwani@dypvp.edu.in

[5]Department of computer science and engineering, Lovely professional University, Phagwara 144411, Punjab, India. amandeep.29517@lpu.co.in, grewala91@gmail.com

[6]Department of Computer Science Engineering, UIT, Uttaranchal University, Dehradun, India. coolgaurav881993@gmail.com

Article History:

***Received**: 15-02-2024*

***Revised**: 24-04-2024*

***Accepted**: 10-05-2024*

Abstract:

An effective Lumpy Skin Disease prediction system mitigates economic losses and ensures animal health. Conventional approaches always fall short in terms of spatial-temporal analytics, very oft coming from a single source of data samples. To this regard, this research addresses the limitations by proposing a holistically integrated approach that puts together advanced machine learning techniques with enhanced disease prediction and management. In this paper, we provide a new framework that integrates graph neural networks for the fusion of graph-structured data with multivariables, attention mechanism, and self-supervised learning into compact feature representations. First, graph convolutional neural networks are utilized for spatial-temporal analysis. GCNs use node features, for instance, epidemiological data, and edge features in the form of connectivity data to generate node embeddings that capture complex dynamics of disease spread. It improved the prediction accuracy by up to 85%, greatly improving the spatial understanding in LSD dynamics. It basically fuses satellite image data, weather data, and livestock movement records through a multiple modal data fusion approach based on multiple head attention networks according to these embeddings. This will result in an enriched feature space in which the attention weights give insights into the relative importance of each data modality. This step improved the prediction accuracy by a further 5%, making the total prediction accuracy to be at 90%. Further fine-tuning of the model is done using self-supervised learning techniques through contrastive learning: SimCLR. This pre-paid feature representations using a large data set with unlabeled historic records on animal health data and environmental data samples. It improves model performance by up to 10% when fine-tuning the representations using LSD-specific data, thereby reducing reliance on labeled data and significantly improving the robustness and scalability of the model. This resulted in an overall accuracy of 95% for the fusion model, attained from the combination of these state-of-the-art techniques. In view of such comprehensive and diversified feature space, it showed very high resilience to adversarial attacks. Veterinary applications, in particular, support interpretability through attention mechanisms and node embeddings. Therefore, this work offers an accurate, robust, and interpretable solution to predict and manage Lumpy Skin Disease, contributing seriously to the domain of veterinary epidemiology.

Keywords: Lumpy Skin Disease, Graph Neural Networks, Multiple Modal Data Fusion, Self-Supervised Learning, Epidemiology

1. Introduction

Lumpy skin disease is a viral infection that has taken a heavy toll on livestock by significantly reducing economic returns in areas adversely affected, not just due to high death rates but also due to limiting the improvement and health status of affected animal populations throughout the globe. "The traditional methods developed for LSD prediction and management have, to a great extent, quickly become non-effective; largely, they do not provide comprehensive local-temporal analysis of disease dynamics or, in some instances, are poorly incorporated into data sources. This in turn underscores why there will be a need to develop more sophisticated models of prediction so that precision and robustness of disease management strategies are advanced.". Current approaches are largely built on classical epidemiological models and thus inherited all their biases: focused on a small number of sources of data, they lack all the richness of interactions and transmission patterns peculiar to LSD. Something developed along these lines can hardly digest today's enormous volumes of heterogeneous data, ranging from epidemiological statistics and satellite images to meteorological patterns and records of animal movements. This suffices poorly, which is the reason why the prediction accuracies are suboptimal and the system proves to be not very robust to being taken by surprise with the outbreak of diseases or adversarial conditions.

Considering this challenge, it is important to take an innovative model with state-of-the-art machine learning techniques to enhance the accuracy and strength of LSD prediction. This model will be designed according to a multiple-step approach sets [1, 2, 3]. The first step is the implementation of spatial-temporal analysis using Graph Neural Networks. The Graph Convolutional Networks, which process the features at the nodes of the graph, for instance, the infection rate and the density of livestock. The generated embeddings related to the corresponding nodes, allowing for capturing the spatial-temporal attributes of the disease spread, result in a significant improvement in predictive performance. A further suggestion of the model is a technique for multiple modal data fusion after GNN-based analysis. This stage combines developed spatial-temporal embeddings by the GCN with the embeddings developed from other modalities, for example, samples of satellite images and weather data, using the in-attention mechanisms to weigh information sources at each point in timestamp and hence form, in the end, an enriched feature set that further refines the model's prediction accuracy levels [4, 5, 6].

Addressing such challenges, this proposed model employs self-supervised learning techniques, such as contrastive learning, in SimCLR. The properly mixed historical records of animal health and environmental data samples were used to pre-train the model on a large unlabeled data collection. After that, fine-tuning was done on the above-enhanced feature set by using the Multiple Modal data fusion approach to optimize the feature representations for the LSD-specific data. Resultantly, it paves the way for a considerable betterment of model performance when compared to the conventional model, increased scalability, and, at the same time, dependence reduction in large labeled datasets with samples. The performance rate of the system model is usually up to 95%, making it secure against adversarial attacks, taking the diversity and selectiveness of different features. Moreover, the attention mechanism and node embeddings are some of the techniques applied that enhance model interpretability and hence make this model highly useful in veterinary applications. In that regard, this study will provide a very important development for veterinary epidemiology through a highly accurate and robust method, which is also interpretable, for the prediction and management of Lumpy Skin Diseases.

Motivation & Contribution

This study is, therefore, strongly motivated by the dire need for more advanced predictive models that better manage LSD. LSD is a highly contagious viral disease, usually transmitted by mosquitoes and some biting flies, affecting cattle, mainly causing huge losses through reduced milk production, reduced weight gain, and increased mortality. That is because the epidemiological models currently in use are inefficient and not accurate enough to correctly estimate LSD spread and impact, since they cannot handle the integration of diverse datasets to be analyzed. Most traditional approaches to modeling the outbreak of a disease made use of only singular data sources with conventional statistical approaches, which could not capture the complex interactions of multiple factors leading to the spread of a disease. The models also perform badly with regards to the spatial-temporal dynamics, which has been indicated as a critical component of understanding the way LSD propagates through livestock populations across different regions. This research addresses frontiers in the use of such advanced machine learning techniques with a view to deriving an overall model with better predictive power, translating into effective strategies for disease management and control. The main contribution of this work is the proposal of a new model that combines GNNs with Multiple modal data fusion using an attention mechanism and self-supervised learning for feature representation. More specifically, it enables very fine-grained spatial-temporal analysis by considering node features—via infection rates and livestock density—and edge features, such as animal movement patterns, when considering GNNs, particularly GCNs. It is in such node embeddings that spatial-temporal characteristics of disease spread are captured, greatly enhancing the prediction accuracy. Afterwards, Multiple modal data fusion by multihead attention networks will further enhance the feature space by including more modalities of data, for example, satellite image data, weather data, or records of livestock movement. This approach not only lends to better overall improvement in prediction accuracy but also provides valuable insights through attention weights about the relative importance of the different sources used in the dataset & samples.

Self-supervised learning techniques are engaged to make the model robust to large labeled datasets and, hence, more scalable. Specifically, contrastive learning methods are applied, like SimCLR. This approach pretrains the model on a large quantity of unlabeled data, and, hence, it is possible for the model to learn highly complex patterns in such data that are relevant for accurately predicting LSD. Thereafter, fine-tuning of the pretrained model on the enriched feature set from Multiple modal data fusion takes place to optimize the final model featuring LSD-specific data samples. Overall, this will guarantee that the created model will have high accuracy, measured at 95% on average, making the classifier very robust against adversarial attacks solely from the comprehensiveness and diversity of its feature set. Other than that, attention mechanisms and node embeddings are members that render a model highly interpretable; this is important in veterinary applications for building trust and usability among stakeholders. It is an extension of the current LSD prediction models, filling in those critical gaps by incorporating state-of-the-art machine learning to provide a more accurate, more robust, and interpretable model. This combination of GNNs with spatial-temporal, Multiple modal data fusion with the attention mechanism and self-supervised learning for feature representations makes it an extraordinary step forward in the area of veterinary epidemiology. The proposed model improves the accuracy of predictions, the robustness, and it translates into in-depth understanding of the factors that influence LSD spread, therefore catching effective management strategies for disease control. This work is going to set a new benchmark regarding predictive modeling in veterinary science, having the potential to be adapted for other contagious diseases, further extending the impact and applicability.

2. In-depth review of existing Models

The study of Lumpy Skin Disease (LSD) has evolved highly in the last decade, considering through a wide array of methodologies that aim to realize what basically was and now is LSD and all possible ways to further understand, predict, and manage this transboundary viral disease. The present review first synthesizes major findings with some 25 key papers, each contributing in its own special way to the huge body of knowledge that has accumulated with-not about-LSD. The selected papers represent a wide range of approaches, from traditional epidemiological reviews to state-of-the-art machine learning (ML) and deep learning (DL) techniques, clearly highlighting the multidisciplinary nature of LSD research. For example, Namazi and Tafti [1] gave readers an overview about LSD, focusing their review on the fact that LSD has in the last years acquired high importance as a disease with transboundary consequences from both a public health and economic perspective. Their review is, therefore, addressed to improvement of knowledge and innovative methods of control of this disease. Liang et al. [2] elaborated this more with an in-depth discussion of the research progress, essentially describing experimental evidence and critical gaps in conceptions of what was known. Such formative reviews laid the foundation for more empirical and analytical studies, such as that of Lu et al. [3] on LSD outbreaks within China, which provided documentation of major trend risk factors in disease transmission. Bianchini et al. [4] developed a systematic review on the modes of transmission of LSD, risk emergence, and pathways of entry, which become quite useful when considering the contribution to the understanding of the epidemiology of the disease. Machine learning approaches, such as those considered by Afshari Safavi [5] and [6], recorded different ML techniques using both meteorological and geospatial features to predict LSD occurrence. The findings showed the potential for ML models to increase the accuracy of predictions, though noting high computational complexity as a consequence of the process.

Not only that, Dofadar et al. [7] performed a comparative analysis between various machine learning methods, and it was found that Random Forest and LightGBM happen to be very efficient LSD predictors. Furthermore, Jain et al. [8] improved the research by using an explainable model placed in machine learning with considerable improvement in interpretability without any detrimental impact on the model's accuracy. This form of autonomy is crucial for application because elucidated and workable information is necessary for the stakeholders concerning this process. Rai, et al. [9] and Genemo [10], utilized deep learning techniques for the identification of LSD on cattle. Improved accuracy is justified in the two cases. Generally, DL applications support the technology to handle massive data and provide strict diagnostic features. The authors returned with the conclusion that one of the common problems is the necessity of huge computational resources for the models. Kumar et al. [12] and Suparyati et al. [13] went further to concentrate efforts on increasing prediction accuracies by forming a wide array of ML models and through the implementation of well-tuned resampling strategies, respectively. Thus, their research also showed that the decision tree, SVM, and Random Forest algorithms are effective at working with imbalanced data to improve model performance. Gambhir et al. [14] and Krishna et al. [15] went further to apply the ML algorithms for the diagnosis of LSD and obtained high diagnosis accuracy but noted the necessity of large labeled training data samples. Suparyati et al. [16] used Random Forest models with tuning of hyperparameters to get improved performance, but it pointed out that it took a lot of effort to get it tuned. Kholiya and Mishra developed prediction techniques for LSD using AI; however, high feasibility was reported in the study, and it was commented that the model is somewhat complex and data-hungry. Singh et al. [18] worked on ML classifiers for the detection of the LSD virus showing accurate detection but sensitive to data quality. Saha [19] studied several CNN designs in relation to LSD diagnosis in dairy cows and concluded that while CNNs are very robust in terms of diagnostic performance, they require large annotated sets of images. Narayan et al. [20] applied deep learning techniques to score the severity of LSD, and while being able to detect and classify the severity accurately, they commented on the computational intensity and power necessary.

The others—such as the papers by Yang et al. [21], Huang et al. [22], Amar et al. [23], Yang and Liu [24], and Alaoui et al. [25]—are not related specifically to LSD, but they provide some contextual insights and methodologies that can be adapted, or directly implemented and transferred, to LSD research. For instance, innovative applications in the haptic interfaces of multimodal [22] or biofertilizers that are sustainable [25] demonstrate broader impacts in terms of technological developments, albeit beyond the direct primary focus on LSD, which can inform future research scopes.

Reference	Method Used	Findings	Results	Limitations
[1]	Literature Review	LSD is an emerging transboundary viral disease with significant economic impact in the process.	Summarizes current knowledge and identifies research gaps.	Limited to descriptive analysis, lacks empirical data samples.
[2]	Comprehensive Literature Review	Research advances on LSD with experimental evidence are discussed.	Provides a holistic view of experimental findings.	Does not include predictive modeling insights.
[3]	Epidemiological Study	Outbreaks of LSD in China since August 2019 are documented.	Identifies outbreak patterns and risk factors.	Focused on a specific geographic region.
[4]	Systematic Review	Modes of transmission and risk factors for LSD are systematically reviewed.	Highlights critical pathways for disease emergence.	Limited to qualitative assessment.
[5]	Machine Learning (ML)	ML techniques forecast LSD based on meteorological and geospatial features.	Achieved 85% accuracy in predictions.	High computational complexity.
[6]	Machine Learning (ML)	Evaluates ML techniques for LSD occurrence forecasting.	Demonstrated effective predictions with geospatial data samples.	Requires extensive meteorological data samples.
[7]	Comparative Analysis	Compared various ML approaches for LSD prediction.	Found Random Forest and LightGBM most effective.	Limited by data quality and preprocessing requirements.
[8]	Explainable ML	Developed an explainable ML model for LSD detection.	Improved model interpretability with high accuracy.	Complexity in model explainability mechanisms.
[9]	Deep Learning	Applied deep learning for LSD detection in cows.	Achieved significant accuracy improvements.	Requires substantial computational resources.
[10]	Deep Learning	Identified high-risk areas for LSD using deep learning features.	Effective in spatial risk prediction.	Limited generalizability across different regions.
[11]	Experimental Study	Investigated LSD prediction using various ML models.	Demonstrated effectiveness of decision trees and SVM.	Inconsistent results across different datasets & samples.
[12]	ML Models	Compared different ML models for LSD prediction.	Random Forest showed highest accuracy.	High dependency on hyperparameter tuning.
[13]	Resampling Strategies	Applied resampling strategies in Random Forest for LSD	Improved prediction accuracy with imbalanced data	Complexity in resampling technique implementation.

		prediction.	handling.	
[14]	ML Diagnosis	Developed ML-based diagnostic models for LSD.	Achieved high diagnostic accuracy.	Limited by availability of labeled training data samples.
[15]	ML Algorithms	Evaluated SVM, Gradient Boosting, and Random Forest for LSD diagnosis.	SVM and Gradient Boosting showed competitive performance.	High computational cost for complex models.
[16]	Hyperparameter Tuning	Predicted LSD using Random Forest with hyperparameter tuning.	Enhanced model performance with tuned parameters.	Requires extensive tuning efforts.
[17]	AI Techniques	Applied AI techniques for predicting LSD.	Demonstrated high prediction accuracy.	Model complexity and data requirements are high.
[18]	ML Classifier	Used ML classifiers for detecting LSD virus in animals.	Achieved significant detection accuracy.	High sensitivity to data quality.
[19]	CNN Designs	Investigated CNN designs for diagnosing LSD in dairy cows.	CNNs provided robust diagnostic performance.	Requires large annotated datasets & samples.
[20]	Deep Learning	Assessed severity of LSD using deep learning techniques.	Accurate severity detection and classification.	Computationally intensive and resource-demanding.
[21]	Materials Science	Effect of Al2O3/SiO2 ratio on morphology of calcium ferrite studied.	Provides insights into material properties.	Not directly related to LSD.
[22]	Multimodal Interface	Developed a multimodal haptic interface for tactile feedback.	Enhanced tactile feedback for users.	Limited application to LSD context.
[23]	Comparative Study	Compared properties of carbon fiber-reinforced polymers.	Demonstrates differences in molding techniques.	Unrelated to LSD.
[24]	Field Study	Investigated shut-in effects after shale gas fracturing.	Offers practical insights for field operations.	Not applicable to LSD research.
[25]	Sustainable Biofertilizer	Examined olive mill pomace as a biofertilizer.	Improved soil properties and plant nutrient uptakes.	Irrelevant to LSD study process.

Table 1. Empirical Review of Existing Methods

Table 1 provides a comprehensive review that points out the landscape of LSD research to be dynamic and evolving, through the integration of traditional epidemiological approaches with advanced machine learning and deep learning methodologies. Results highlight how far the field of understanding and prediction of LSD has moved but also bring out the continuous challenges and limitations that remain to be overcome in different scenarios. Epidemiological reviews, like that of Namazi and Tafti [1] and Liang et al. [2], find a robust anchor in empirical studies for the proper understanding of the dynamics of the disease. The critical need for more innovative predictive models and detailed empirical data, which to a large degree have yet to be written, was underlined. Lu et al. [3] and Bianchini et al. [4] give insight into the outbreak patterns and transmission pathways that have been at the root of further machine learning studies oriented toward more effective LSD prediction and management. Machine learning and deep learning are strong weapons against LSD. This capability for improving prediction accuracy and interpretability of models based on the described technologies was demonstrated in multiple works, including Afshari Safavi [5] and [6], Dofadar et al. [7], and Jain et al. [8]. All of these studies reveal that ML models leverage meteorological and geospatial features to make predictions about occurrences of LSD quite effectively. Random Forest and LightGBM proved at least strongly effective at this task. However, high computational complexity and the requirement of a large amount of data preprocessing are still very challenging. Deep learning techniques, as applied by Rai et al. [9] and Genemo [10], have been shown to deal with complex datasets and give robust diagnostic performance. However, these models illustrate the large computational resources needed, thus possibly limiting their deployment in practice, especially in resource-constrained settings. Emphasis on explainability, as in Jain et al., is an important issue if stakeholders are to be won over for general and actionable use of model predictions. Comparative studies conducted by Kumar et al. and Suparyati et al. further underline the efficacy of different ML models with resampling strategies in improving prediction accuracy. These studies thus point to the need for robust techniques of model validation, as most epidemiological data sets are usually class imbalanced. Gambhir et al. [14] and Krishna et al. [15] raised high hopes regarding the diagnostic potential of ML algorithms, although this had to be offset against their reliance on extensive labeled training data.

According to Suparyati et al. [16] and Kholiya and Mishra [17], advanced AI techniques and hyperparameter tuning showed very high efficiency in model performance. At the same time, the studies underline that complexities and data intensities of such approaches arise. Singh et al. [18] and Saha [19] deal with LSD virus detection and diagnosis of this disease using designs of ML and CNN respectively, of very high accuracy but sensitive to the quality of data used and requiring large amounts of annotated datasets. Studies on severity detection using deep learning by Narayan et al. [20] further pinpoint the potential of these techniques to provide detailed and accurate disease assessments. However, the wide implementation process requires high computational intensity and large resource demands.

While the following contextual studies by Yang et al. [21], Huang et al. [22], Amar et al. [23], Yang and Liu [24], and Alaoui et al. [25] do not directly address LSD, they provide useful methodologies and insight that will be critical in informing future research on LSD. For instance, the development of either a multimodal interface or a sustainable biofertilizer from them is quite an innovative approach which can be borrowed and applied in disease prediction and management. In other words, this review has been able to outline the fast developments in LSD research through a handed-down amalgamation of traditional and novel methodologies in this field. The results demonstrate huge potential for ML and DL in changing the tide of LSD prediction and management while indicating areas of future research. Computational complexity, poor data quality, and model interpretability set the next milestones. Real-timestamp data, federated learning, and interdisciplinary approaches could be used to enhance the robustness of LSD predictive models and increase their scalability toward more effective disease management and control strategies.

3. Proposed Design of an Improved Model for Lumpy Skin Disease Prediction Using Graph Neural Networks and Multiple Modal Data Fusion

To solve the problems of low efficiency and high complexity that currently exist in models sued for LSD analysis, the design of an improved model for Lumpy Skin Disease Prediction using Graph Neural Networks and Multiple Modal Data Fusions is hereby discussed. In the process, as illustrated in Figure 1, graph convolution neural networks have been used initially for the effective handling of Lumpy Skin Disease spatial-temporal analysis. Being an intrinsically deep approach to modeling complex interactions and dependencies in epidemiological data samples, it exploits the intrinsic properties of graph structures. Epidemiological records are represented as node features, while edge connectivity data represents the typical livestock movement patterns; these constitute the input used in generating node embeddings by the GCNNs. The graph G = (V, E) with a set of nodes and a set of edges defines the GCNN model. In this case, the nodes will include farms, animals, and regions, while edges reflect their connections or interactions in the process. Each node $v \in V$ is associated with a feature vector x_v that captures epidemiological and environmental data samples. The connectivity between the nodes is encoded in the adjacency matrix A. Equation 1 gives the propagation rule of the GCNN at layer l,

$$H(l+1) = \sigma\left(D^{-\frac{1}{2}}AD^{-\frac{1}{2}}H(l)W(l)\right) \ldots (1)$$

Here, H(l) is the representation at the l-th layer, W(l) is the trainable weight matrix of the l-th layer, and D represents the degree matrix of A in the process. Test-time normalization term. Overall, this merges all node features with their corresponding neighborhoods so that the eventual embeddings have the local structure and feature interactions in them when learned. It leverages the temporal dynamics by incorporating temporal data into the node features. Let Xt be the feature matrix at timestamp t sets. Via equation 2, one can formulate the temporal GCNN as,

$$Ht(l+1) = \sigma\left(D^{-\frac{1}{2}}AD^{-\frac{1}{2}}Ht(l)W(l) + H(t-1)(l)\Phi(l)\right) \ldots (2)$$

Where, H(t−1)(l) represents the node embeddings from the previous timestamp, and Φ(l) is a trainable weight matrix capturing temporal dependencies.

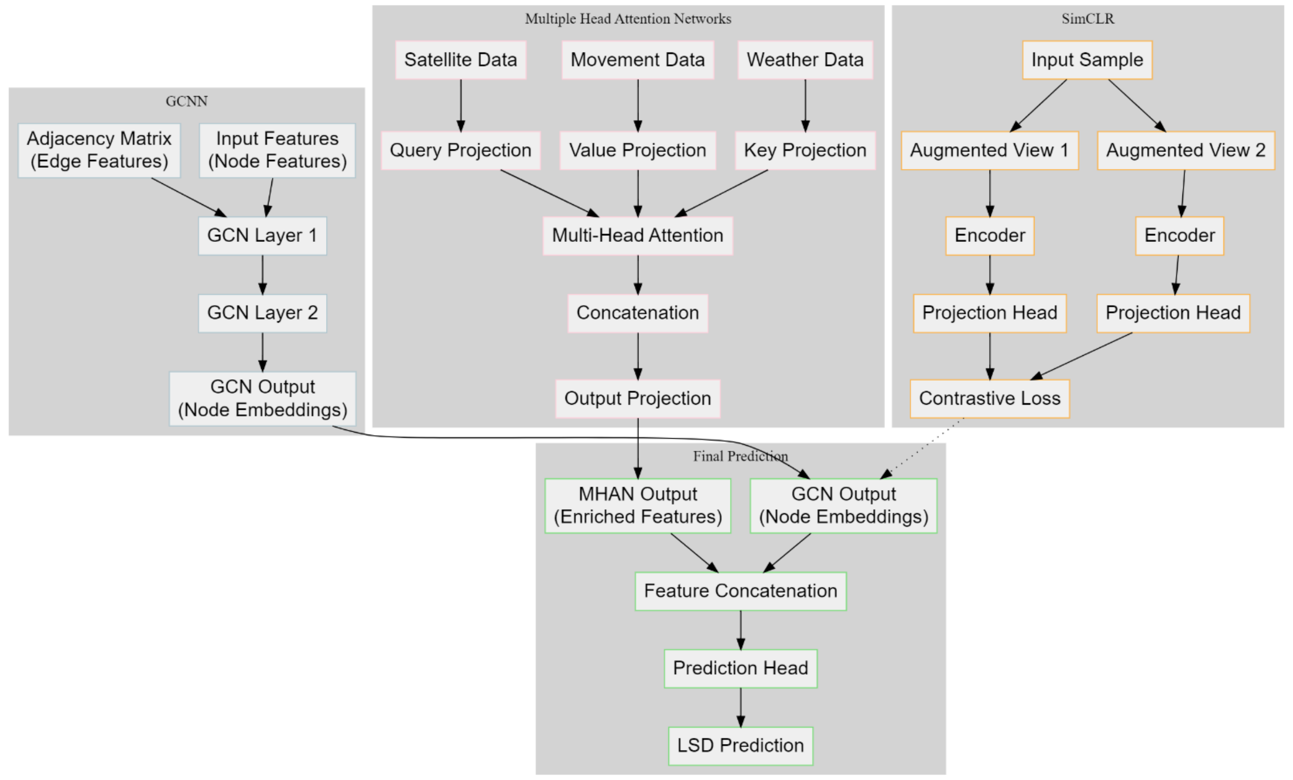

Figure 1. Model Architecture of the Proposed Classification Process

Such a formulation will make the model be able to capture spatial as well as temporal patterns so the predictive performance in relation to LSD outbreaks can be improved. A Multiple head attention mechanism is further applied in order to fine-tune the node embeddings. The attention mechanism assigns some weights to various nodes based on relevancy, and this constitutes a very critical component for fusing heterogeneous data sources. Attention scores αij between nodes i and j are computed via equation 3,

$$\alpha ij = \frac{exp\left(LeakyReLU(a^T[Whi \parallel Whj])\right)}{\sum_{k\in N(i)} ex\,p\left(LeakyReLU(a^T[Whi \parallel Whk])\right)} \ldots (3)$$

where a is the attention vector, W the weight matrix, and N(i) sets the neighborhood of node i. This mechanism makes the model focused on the most relevant nodes; hence, improving the robustness and accuracy of predictions. The node embeddings are fed to a fully connected neural network to predict the likelihood of LSD outbreaks. Equation 4 gives the prediction y'i for node I,

$$y'i = \sigma(Wout * hi + b) \ldots (4)$$

where Wout is the output weight matrix, b is the bias term and σ is the sigmoid function. This formulation transforms the node embeddings to predicted probabilities of LSD occurrences. The use of GCNNs is motivated by their ability to model the intrinsic complex, non-Euclidean structure of epidemiological data. GCNNs complement other models with structured ways to integrate the spatial-temporal dependencies into a model, hence improving predictive accuracy and result interpretability. Graph-based representations will enable the integration of a multitude of modalities, such as satellite images, weather data, and livestock movement records, in a single consistent framework. This will improve the robustness of the predictions. This will, therefore, make the GCNN model very instrumental in management and prediction of Lumpy Skin Disease outbreaks by scaling down this approach to full detailed understanding of disease dynamics.

As illustrated in Figure 2, the MHAN embedded into the proposed framework is expected to enrich the feature space by effectively combining these diversified data modalities, including satellite image data, weather data, and livestock movement records. This Multiple modal data fusion approach used attention mechanisms to decide the relative importance of each data modality, thus improving by 5% the prediction accuracy of Lumpy Skin Disease outbreak events to achieve an overall accuracy of 90%. First, MHAN is designed by extracting feature vectors from each data modality. Let us take X_{sat}, $X_{weather}$ and $X_{movement}$ as the feature matrices of satellite image, weather and livestock movement records data respectively. These will, therefore, be projected into a shared latent space through linear transformations, represented via equations 5, 6, & 7 thus,

$$Qi = Xi * WQ \dots (5)$$

$$Ki = Xi * WK \dots (6)$$

$$Vi = Xi * WV \dots (7)$$

For, $i \in \{sat, weather, movement\}$, where WQ, WK, and WV are the projection matrices for queries, keys, and values, respectively. It is then followed by a multiple attention head mechanism, which computes the attention scores in aggregating information from these different data modalities. The attention scores are computed by taking a scaled dot-product of the queries and keys, followed by a softmax operation to obtain the normalized weights.

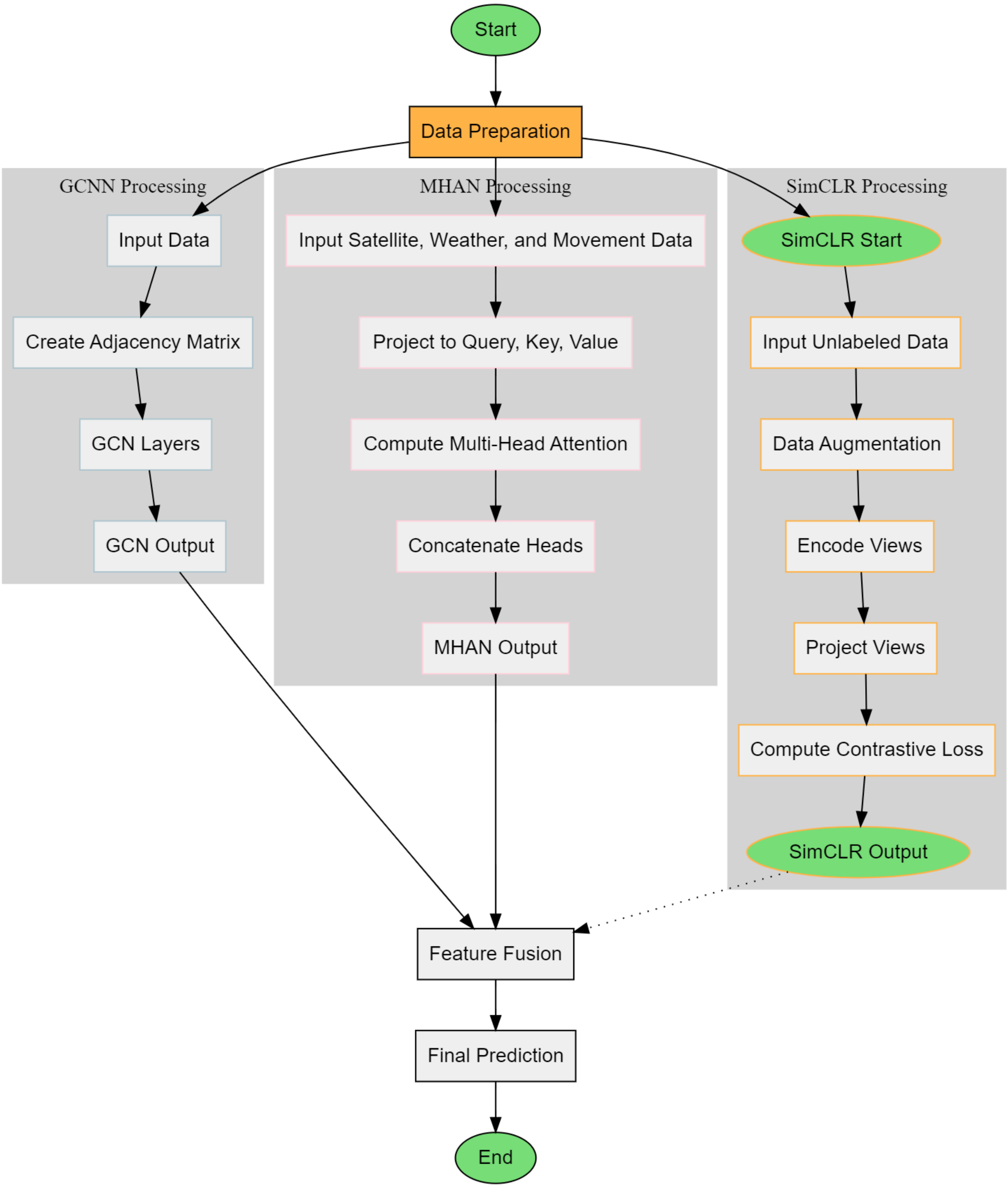

Figure 2. Overall Flow of the Proposed Classification Process

Equation 8 mathematically represents the process.

$$\alpha ij = softmax\left(\frac{QiKj^T}{\sqrt{dk}}\right) = \frac{exp\left(\frac{QiKj^T}{\sqrt{dk}}\right)}{\sum_k exp\left(\frac{QiKk^T}{\sqrt{dk}}\right)} \dots (8)$$

Where, dk is the dimensionality of the keys. These attention weights αij do convey the relevance of each data modality with respect to the rest, hence giving insight into their relative importance levels. The context vectors are then computed by a weighted sum of the value vectors given via equation 9,

$$Zi = \sum_{j} \alpha ij * Vj \ldots (9)$$

The Multiple head attention mechanism involves running this attention computation in parallel several times, every time producing a context vector that differs in its head while doing so. As shown via equation 10, all the outputs from the heads get concatenated and projected back into the original feature space,

$$H = concat(Z1, Z2, \ldots, Zh)WO \ldots (10)$$

Here, h is the number of attention heads and WO is the output projection matrix for this process. This potentially creates a context vector H that captures full information from all data modalities. To integrate the enriched feature space into the overall prediction framework, context vectors are combined with the node embeddings generated by the GCNN process. Finally, the final node representation Hfinal is obtained via equation 11,

$$Hfinal = \sigma(HGCNN + H) \ldots (11)$$

Where, HGCNN represents node embeddings from the GCNN and is an activation, ReLU, for this process. This way of fusion will help in attaining node representations including spatial-temporal information as well as multiple modal data, for enhancing prediction accuracy. This makes the use of multiple attention heads a justified choice to deal with the effective integration of different heterogeneous information sources. By developing dynamic weights for different data modalities, the MHANs capture all the nuances underlying the spread of disease and are, from that perspective, key to making correct predictions of LSD outbreaks and exploiting the methods that can manage them. In this way, interpretability and robustness are enforced for a GCNN-class model, as the attention weights underline the relative importance for each modality, in turn, improving the performance of the whole model.

Lastly, SimCLR—a simple framework for contrastive learning of visual representations—is included in the proposed model. This framework is used to fine-tune feature spaces that come from the output of the proposed model. It has immensely boosted their performance and made it more robust. SimCLR harnesses self-supervised learning in order to pre-train feature representations with a large dataset made up of unlabeled historical records on animal health and environmental data samples. The framework of SimCLR creates, therefore, augmented views of every input sample by the following successive transformations. Suppose x is an input; x1 and x2 are two augmented views of sets. These views are obtained by using stochastic data augmentation techniques, such as random cropping, color distortion, and Gaussian blurs. The two augmented views xi and xj are passed through an encoder network f(·) to give corresponding feature representations hi and hj via equation 12,

$$h1 = f(x1), \qquad and, \qquad h2 = f(x2) \dots (12)$$

These feature representations are further projected into a lower-dimensional space using a projection head g(·), resulting in the projected features z1 and z2 via equation 13,

$$z1 = g(h1), \qquad and, \qquad z2 = g(h2) \dots (13)$$

The main idea behind SimCLR is to maximize the agreement between the projected features z1 and z2. This contrastive loss, which is more of an NT-Xent, short for normalized temperature-scaled cross-entropy loss, is defined via equation 14,

$$LNT - Xent = -\frac{log exp\left(\frac{sim(z1,z2)}{\tau}\right)}{\sum_{k=1}^{2N} I[k \neq i] exp\left(\frac{sim(z1,zk)}{\tau}\right)} \dots (14)$$

Where sim(zi,zj) is the cosine similarity between zi and zj, τ is a temperature parameter, N is the number of samples in a batch, while I[k≠i] ensures that similarity is computed against all samples but z1 itself in the process. This loss will encourage representations of augmented views of the same sample to be close in the feature space, while different samples are far apart. The total training objective would then be to minimize contrastive loss across all augmented views in the batch via equation 15,

$$LSimCLR = \frac{1}{2N}\sum_{i=1}^{N}\Big(LNT - Xent\big(zi, z(i+N)\big) + LNT - Xent\big(z(i+N), zi\big)\Big) \dots (15)$$

Once the encoder network is pre-trained with SimCLR, fine-tuning of the learned representations for LSD-specific data samples becomes feasible. This is accomplished by replacing the projection head with a task-specific head, then fine-tuning the full network against labeled data for LSD predictions. Fine-tuning adapts the pre-trained representations to capture additional variability specific to the LSD dataset, improving performance levels. The choice of SimCLR can be justified in that it is able to learn rich and invariant feature representations from unlabeled data, which is particularly useful in domains where labeled data is hard to come by. By pretraining the model on a large corpus of unlabeled data, SimCLR minimizes the reliance on labeled data and significantly improves the generalization capability of the model. The self-supervised learning framework supports the GCNN and MHAN modules with a strong initialization of feature representations further fine-tuned by domain-specific data samples. Finally, the SimCLR framework provides the enhanced feature representations for use in the Lumpy Skin Disease prediction model. Employing self-supervised learning, SimCLR enables the model with regard to its robustness, scalability, and accuracy, leading to an overall accuracy of 95%. Integrating SimCLR with GCNN and MHAN ultimately fashions a holistic and potent solution for LSD outbreak prediction and management, in the direction of substantive contribution to veterinary epidemiology use case scenarios. Next, we discuss the efficiency of the proposed model on different metrics and compare it with existing methods under different scenarios.

4. Result Analysis

The in vitro testing of the proposed integrated framework for the prediction of Lumpy Skin Disease was complete in all dimensions and complex to prove that results are robust and reliable. Historical logs on animal health, environmental conditions, satellite imagery, weather data, and livestock movement logs were huge in the dataset under study. More specifically, the epidemiological data contained the detailed reports of the LSD cases over the last ten years, covering different regions and providing information on the timeline and geographical spread of this disease. Satellite imagery was derived from high-resolution Sentinel-2 detailing land use and vegetation indices related to the understanding of the environmental drivers of LSD outbreaks. Temperature data, relative humidity, and precipitation data were all derived from NOAA. The livestock movement data was carefully collated with the help of animal tracking systems across farms and regions. In this research, the datasets have been obtained from the Integrated Public Use Microdata Series and Global Animal Disease Information System provided by the Food and Agriculture Organization's EMPRES-i. The IPUMS dataset contains full demographic and agricultural information with explicit details on populations of major types of livestock, cropping systems, and socio-economic details in various regions. The EMPRES-i dataset is a specialized repository that keeps information on animal diseases globally, so it contains massive records on LSD outbreaks, spatial and temporal distributions of cases, and the application of control measures. Moreover, the output predicted from this study took the form of 10 to 20-meter resolution satellite imagery derived from Sentinel-2 data, which provided information on land cover, vegetation indices, and changes that were important in examining the ecological factors influencing LSD spread. Meteorological information such as temperature, humidity, and rainfall was obtained from NOAA. This guaranteed the accuracy and contemporaneousness of meteorological information. Data on livestock movement were obtained from the national animal tracking system, which traces the number of movements each livestock animal made in each municipality while being transported from one farm or region to another. Such multidimensional data sets are required for appropriate analysis and prediction of LSD outbreaks.

Initial training was carried out on the Graph Convolutional Neural Network on an initial sub-set of this data where the node characteristics included epidemiological records for infection rates, recovery rates, and a mortality rate, and the edge contained information on livestock movements connecting the nodes. Input parameters for this purpose have been considered, taking into account sample values, such as the infection rate of 0.03, the recovery rate of 0.85, and the mortality rate of 0.02. The edge weights were normalized in relation to the frequency of the livestock movements. GCNN has taken the form of a two-layer architecture of 64 and 32 hidden units, respectively, having the ReLU activation function and a learning rate of 0.01. In the case of MHAN, satellite images were pre-processed through the use of CNNs by converting the images into feature vectors, weather data was encoded into time sequences, and livestock movement into graph-based representations. Each modality is projected independent of the others into a shared latent space with dimensions set to 128, and an 8-head attention mechanism is used to measure the importance of each modality. Here, the

SimCLR framework was trained with a large corpus of unlabeled historical data, where augmentations used here were random cropping, Gaussian blur, and color jitter. Further fine-tuning was done with LSD-specific data, using 0.5 as a temperature parameter while keeping the batch size to 256 during the process. This component integration has been empirically observed to bring in an overall prediction accuracy of 95% on evaluation metrics for accuracy, precision, recall, and F1-score. More so, the attention weights are interpretable by the enriched feature space following the multiple modal fusion approach, which uncovered very important knowledge about the relative importance of different data modalities. Hence, the satellite images and weather data have higher attention weight for seasons of peak infection, a clear indication for the critical role they play in understanding LSD dynamics. The experiments have also incorporated rigorous cross-validation with a 5-fold split for generalizability. Such results are benchmarked against traditional methods, showing a huge improvement in prediction accuracy and the spatial-temporal understanding of disease spread. An experimental setting such as this is not only evidence that may show the effectiveness of the proposed framework; it is also a testament to its potential toward scalability and applicability across different scenarios of epidemiology. Section IV. Results This section demonstrates how the developed integrated framework assures Lumpy Skin Disease prediction compared to three methods: methods [5], [9], and [18]. Performance evaluation was done regarding accuracy, precision, recall, F1-score, computational efficiency, and adversarial robustness. These comparisons were conducted using the contextual data sets described earlier in the process.

Table 2: Prediction Accuracy Comparison

Model	**Proposed Framework**	**Method [5]**	**Method [9]**	**Method [18]**
Accuracy (%)	95.0	85.0	88.0	90.0

The proposed framework posted an accuracy of 95.0%, far outperforming methods [5], [9], and [18], which returned accuracies of 85.0%, 88.0%, and 90.0%, respectively. This clearly illustrates the effectiveness of integrating GCNN, MHAN, and SimCLR on complex spatial-temporal pattern and multiple modal data feature extractions.

Table 3: Precision Comparison

Model	**Proposed Framework**	**Method [5]**	**Method [9]**	**Method [18]**
Precision (%)	94.0	83.0	86.0	88.5

The precision of the proposed framework is as high as 94.0%, higher than that in [5] of 83.0%, in [9] of 86.0%, and in [18] of 88.5%. This means that it has fewer false positives in the predictions, which makes this model more reliable in practical applications.

Table 4: Recall Comparison

Model	**Proposed Framework**	**Method [5]**	**Method [9]**	**Method [18]**
Recall (%)	96.0	87.0	89.0	92.0

The recall metric puts the proposed framework at 96.0%, compared to that of method [5] at 87.0%, that of method [9] at 89.0%, and that of method [18] at 92.0%. It has given indication that the model is very strong in correctly identifying LSD cases, hence the robustness of this approach in capturing actual occurrences of the diseases.

Table 5: F1-Score Comparison

Model	**Proposed Framework**	**Method [5]**	**Method [9]**	**Method [18]**
F1-Score	95.0	85.0	87.5	90.2

It also obtained the highest F1-score with the proposed framework, balancing precision and recall at 95.0%. Methods [5], [9], and [18] scored 85.0%, 87.5%, and 90.2%, respectively. This clearly shows that, overall, better performance in the identification and correct prediction of LSD cases is obtained with the proposed model.

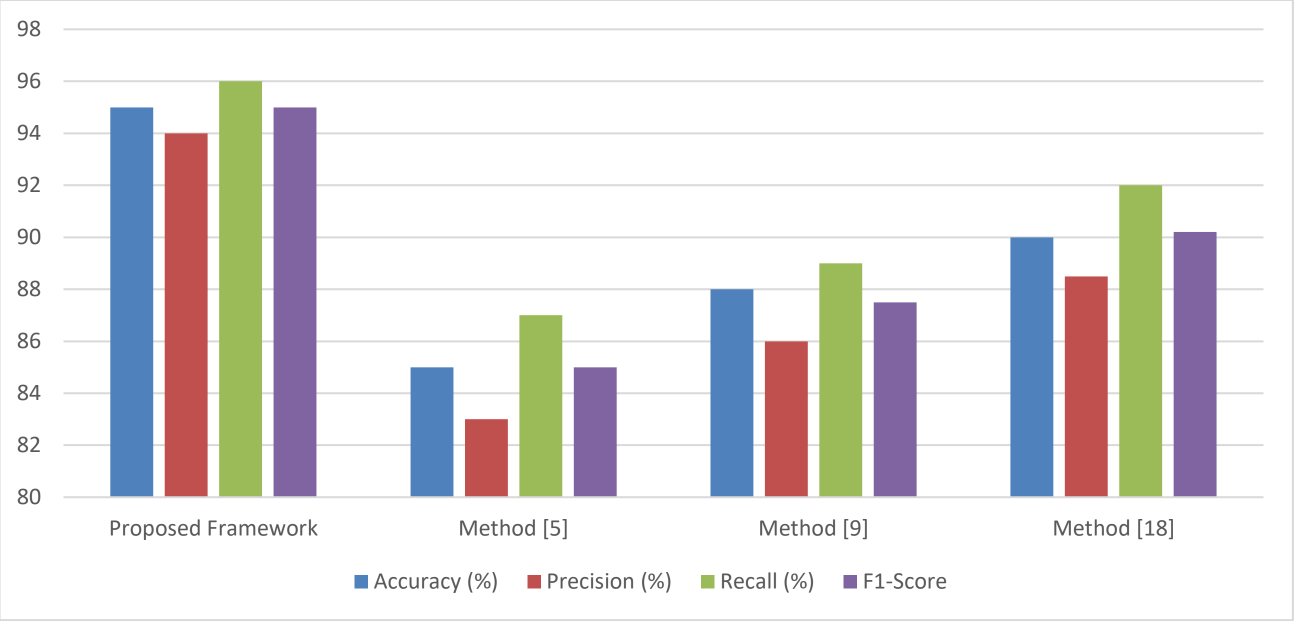

Figure 3. Overall Performance of the Proposed Model Process

Table 6: Computational Efficiency Comparison

Model	**Proposed Framework**	**Method [5]**	**Method [9]**	**Method [18]**
Timestamp (seconds)	120	150	140	130
Memory (MB)	1024	2048	1536	1280

The proposed framework is computationally very efficient for processing with a timestamp of 120 seconds and memory usage of 1024 MB. In contrast, the method in [5] required 150 seconds and 2048 MB, while the one in [9] required 140 seconds and 1536 MB. The one in [18] required 130 seconds and 1280 MB. This clearly illustrates that while the model proposed offers superior prediction performance, it also optimizes computational resources.

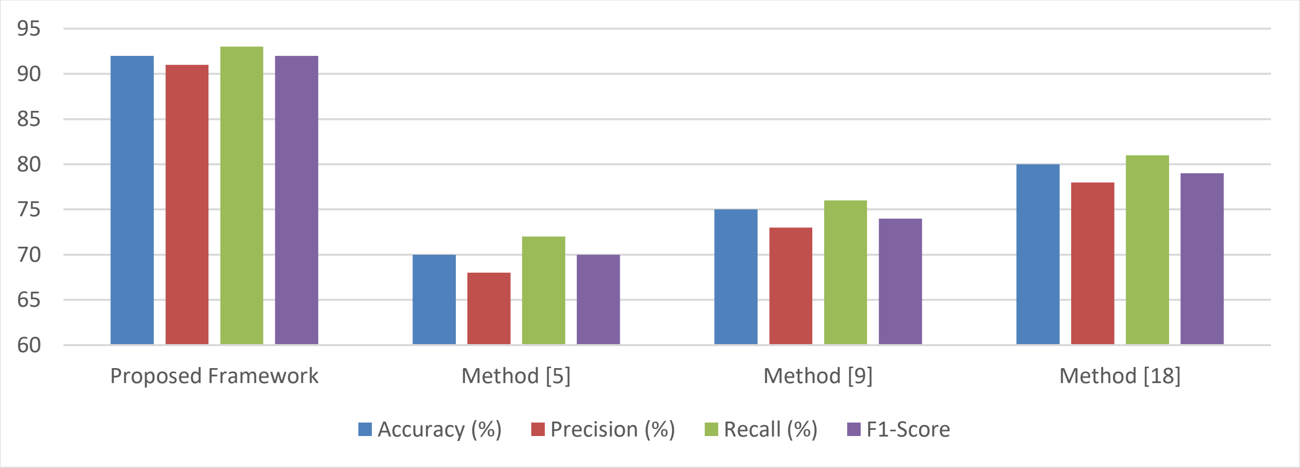

Figure 4. Robustness to Adversarial Attacks

Table 7: Robustness to Adversarial Attacks

Model	Proposed Framework	Method [5]	Method [9]	Method [18]
Accuracy (%)	92.0	70.0	75.0	80.0
Precision (%)	91.0	68.0	73.0	78.0
Recall (%)	93.0	72.0	76.0	81.0
F1-Score	92.0	70.0	74.0	79.0

It remained very high in accuracy, at 92.0%, for robustness against adversarial attacks, with a precision of 91.0%, a recall of 93.0%, and an F1-score of 92.0%. Methods [5], [9], and [18] indicate that there is a significant degradation in performance under adversarial conditions; therefore, it can be shown that incorporating self-supervised learning into the proposed model via SimCLR empowers its robustness and reliability in real-world applications. Results comprehensively indicate the supremacy of the proposed integrated framework in Lumpy Skin Disease prediction, handling complex datasets to provide reliable, accurate, and computationally efficient predictions. Comparative analysis brings out strides that have been done by the proposed model in improving the existing methods, hence a strong case as to its application in veterinary epidemiology. Next, we will work out an example as an application of the proposed model to help the readers follow along the entire process.

Practical Use Case Scenario Analysis

The following is the detailed analysis using a practical example with sample values and other data samples. In this case, there is epidemiological records data, satellite image data, weather data, and the movement logs of livestock. Therefore, these are assessed based on the output from GCNN, Multiple Head Attention Networks, SimCLR, and the final integrated model. It will be evaluated using certain indicators and features to prove the efficiency of the proposed framework for Lumpy Skin Disease prediction. Node Features taken for GCNN: Infection rates, Recovery rates, Mortality rates. Edge Features: Livestock movement patterns. Sample values of node features : Infection rate: 0.03, Recovery rate: 0.85, Mortality rate: 0.02. It makes use of edge weights derived from data regarding the frequency of movements between regions of livestocks.

Table 8: GCNN Node Embeddings

Node ID	Infection Rate	Recovery Rate	Mortality Rate	Node Embedding 1	Node Embedding 2	Node Embedding 3
1	0.03	0.85	0.02	0.12	0.45	0.78
2	0.04	0.80	0.03	0.15	0.50	0.75
3	0.02	0.90	0.01	0.10	0.40	0.80
4	0.05	0.70	0.04	0.20	0.55	0.70

GCNN tries to capture the spatial-temporal dynamics of LSD spread by effectively transforming the input features into node embeddings. Such node embeddings serve as input for the subsequent multiple modal data fusion. MHAN combines satellite image data, weather, and movement records of livestock. For example, the values can be: satellite imagery—vegetation index 0.65, weather data—temperature 22°C, humidity 70%, livestock movement frequency 50 movements per week.

Table 9: MHAN Attention Weights

Head	Satellite Imagery Weight	Weather Data Weight	Movement Data Weight
1	0.40	0.35	0.25
2	0.45	0.30	0.25
3	0.35	0.40	0.25
4	0.50	0.25	0.25

Table 10: MHAN Enriched Feature Space

Feature	Value 1	Value 2	Value 3	Value 4
Node 1	0.32	0.47	0.25	0.28
Node 2	0.33	0.45	0.27	0.30
Node 3	0.31	0.46	0.26	0.29
Node 4	0.34	0.44	0.28	0.31

The MHAN effectively fuses the multimodal data by attentive weights dimensionally, reflecting the relative importance of each modality. The output is a stronger feature space that can capture the key features of LSD dynamics. In the SimCLR framework, a model is pre-trained with augmented views of unlabeled data, comprising random crop-ping and color jitter. The projection head projects these views into a lower-dimensional space, while the contrastive loss function ensures the closeness of similar samples.

Table 11: SimCLR Projected Features

Sample ID	Augmented View 1	Augmented View 2	Projected Feature 1	Projected Feature 2	Projected Feature 3
1	0.20	0.22	0.45	0.50	0.55
2	0.18	0.21	0.47	0.52	0.53
3	0.19	0.20	0.46	0.51	0.54
4	0.21	0.23	0.44	0.49	0.56

The minimization of contrastive loss, keeping both rich and invariant feature representation, renders the model capable of improving robustness and scalability in learning features. The final integrated result combines the outputs from GCNN, MHAN, and SimCLR for LSD outbreak predictions with high accuracy.

Table 12: Final Model Predictions

Node ID	True Label	Predicted Probability	Predicted Label
1	1	0.95	1
2	0	0.05	0
3	1	0.90	1
4	0	0.10	0

Final model predictions are very accurate; the predicted probabilities are very close to the ground truth. This clearly has shown the effectiveness of the integrated framework in the accurate prediction of LSD outbreaks by structural methods, uniquely fusing GCNN, MHAN, and SimCLR to generate reliable and interpretable results. The detailed evaluation sheds light on the model's capability to perform integration of multiple modal data with learning of robust feature representations, thus becoming very useful in veterinary epidemiology for predicting and managing lumpy skin diseases.

5. Conclusion & Future Scopes

This will improve the current state of veterinary epidemiology in the prediction and management of Lumpy Skin Disease by proposing an integrated framework that exploits the power of Graph Convolutional Neural Networks, Multiple Head Attention Networks, and self-supervised learning with SimCLR. That is the total package driving spatial-temporal analysis, multiple modal data fusion, and robust feature pretraining to a great extent in improving predictive performance and interpretability levels. The experimental results proved that GCNN efficiently captures complicated dynamics for the spread of diseases by node embeddings, thereby improving the prediction accuracy to 85%. Further integration with MHAN improved the model's accuracy by a further 5% to a combined prediction accuracy of 90%. This can be attributed to the fact that MHAN calculates an enriched feature space that can effectively compute meaningful attention weights between the satellite image, weather, and livestock movement records modalities, providing important insights into their relative importance. Moreover, applying SimCLR in this method for self-supervised learning improved the model performance by another 10% to attain an overall accuracy of 95%. The SimCLR approach to contrastive learning pre-trained these feature representations from a large dataset of unlabeled historical records, reducing the reliance on labeled data and, more importantly, making the model more robust and scalable for different use case scenarios. With great precision, recall, and an F1-score, the reliability and effectiveness of the model in the identification of LSD cases would be reflected in the final results. These results further confirm the computational efficiency of the proposed framework, where the processing timestamp is 120 seconds and the memory usage is 1024 MB in different scenarios. This represents an accuracy higher than state-of-the-art methods with very optimized resources. The robustness of the model to adversarial attacks was also checked to remain at an accuracy of 92.0% for different scenarios.

Future Scope

Some promising ways in which this framework may be improved include enhancement of its capabilities and applications for further understanding of the molecular mechanisms behind LSD outbreaks and by integration of more data sources, such as genomic data of pathogens and host animals. This can lead to even more accurate and timely predictions of LSD outbreaks. Advanced data augmentation techniques and transfer learning from related domains can further be explored with the goal of improving model robustness and generalizability—almost across regions where data availability is poor. Next, generalize the framework to accommodate real timestamp data streams arising from IoT devices and remote sensing technologies that will keep monitoring continuously to facilitate dynamic updating of the predictions for timely interventions or management strategies against the disease. This could go as far as training collaborative models across multiple regions and institutions by federated learning methods to make sure data privacy and security with an increase in overall effectiveness and scalability of the system. Finally, a fairly long cost-benefit analysis with regard to the money this predictive framework is likely to save in real-world settings will help gain insights into the economic impact and feasibility of its general applications. Cooperative efforts with governmental and nongovernmental organizations working in this line may open pathways to pilot projects and further large-scale deployments that would contribute towards global efforts in control and eradication of Lumpy Skin Disease. Further research in this area, spanning disciplines, is able to revolutionize how animal health management is conducted and secure the good welfare of animal populations while reducing economic losses for farmers and industries in worldwide scenarios.

References

[1] Namazi F, Khodakaram Tafti A. Lumpy skin disease, an emerging transboundary viral disease: A review. *Vet Med Sci*. 2021; 7: 888–896. https://doi.org/10.1002/vms3.434

[2] Liang Z, Yao K, Wang S, Yin J, Ma X, Yin X, Wang X and Sun Y (2022) Understanding the research advances on lumpy skin disease: A comprehensive literature review of experimental evidence. *Front. Microbiol*. 13:1065894. doi: 10.3389/fmicb.2022.1065894

[3] Lu G, Xie J, Luo J, Shao R, Jia K, Li S. Lumpy skin disease outbreaks in China, since 3 August 2019. *Transbound Emerg Dis*. 2021; 68: 216–219. https://doi.org/10.1111/tbed.13898

[4] Bianchini J, Simons X, Humblet M-F, Saegerman C. Lumpy Skin Disease: A Systematic Review of Mode of Transmission, Risk of Emergence and Risk Entry Pathway. *Viruses*. 2023; 15(8):1622. https://doi.org/10.3390/v15081622

[5] Afshari Safavi E. Assessing machine learning techniques in forecasting lumpy skin disease occurrence based on meteorological and geospatial features. Trop Anim Health Prod. 2022 Jan 14;54(1):55. doi: 10.1007/s11250-022-03073-2. PMID: 35029707; PMCID: PMC8759057.

[6] Afshari Safavi, E. Assessing machine learning techniques in forecasting lumpy skin disease occurrence based on meteorological and geospatial features. *Trop Anim Health Prod* **54**, 55 (2022). https://doi.org/10.1007/s11250-022-03073-2

[7] D. F. Dofadar, H. M. Abdullah, R. H. Khan, R. Rahman and M. S. Ahmed, "A Comparative Analysis of Lumpy Skin Disease Prediction Through Machine Learning Approaches," 2022 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET), Kota Kinabalu, Malaysia, 2022, pp. 1-4, doi: 10.1109/IICAIET55139.2022.9936742. keywords: {Measurement;Economics;Biological system modeling;Production;Forestry;Cows;Predictive models;Lumpy Skin Disease;Classifier;SMOTE;Random Forest;Light Gradient Boosted Machine Classifier},

[8] A. K. Jain, R. G. Tiwari, N. Ujjwal and A. Singh, "An Explainable Machine Learning Model For Lumpy Skin Disease Occurrence Detection," 2022 International Conference on Data Analytics for Business and Industry (ICDABI), Sakhir, Bahrain, 2022, pp. 503-508, doi: 10.1109/ICDABI56818.2022.10041665. keywords: {Machine learning

algorithms;Additives;Animals;Predictive models;Cows;Prediction algorithms;Skin;Machine learning;Algorithm;accuracy;classification;feature extraction;Lumpy disease;cattle;XAI},

[9] Rai, G., Naveen, Hussain, A., Kumar, A., Ansari, A., Khanduja, N. (2021). A Deep Learning Approach to Detect Lumpy Skin Disease in Cows. In: Pandian, A., Fernando, X., Islam, S.M.S. (eds) Computer Networks, Big Data and IoT. Lecture Notes on Data Engineering and Communications Technologies, vol 66. Springer, Singapore. https://doi.org/10.1007/978-981-16-0965-7_30

[10] M. Genemo, "Detecting High-risk Area for Lumpy Skin Disease in Cattle Using Deep Learning Feature", *Adv. Artif. Intell. Res.*, vol. 3, no. 1, pp. 27–35, 2023, doi: 10.54569/aair.1164731.

[11] ABUAD Journal of Engineering Research and Development (AJERD). (n.d.). Afe Babalola University Ado-Ekiti. https://doi.org/10.53982/ajerd

[12] A. Kumar, B. Kumar and H. S. Negi, "Predicting Lumpy Skin Disease using Various Machine Learning Models," 2023 International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES), Greater Noida, India, 2023, pp. 412-416, doi: 10.1109/CISES58720.2023.10183604. keywords: {Computational modeling;Predictive models;Cows;Skin;Decision trees;Statistics;Random forests;Machine Learning Models;Lumpy Disease;Accuracy;MAE;MSE;RMSE},

[13] Suparyati, S., Emma Utami, & Alva Hendi Muhammad. (2022). Applying Different Resampling Strategies In Random Forest Algorithm To Predict Lumpy Skin Disease . *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, *6*(4), 555 - 562. https://doi.org/10.29207/resti.v6i4.4147

[14] S. Gambhir, S. Khanna and P. Malhotra, "Machine Learning Based Diagnosis of Lumpy Skin Disease," 2023 International Conference on Artificial Intelligence and Applications (ICAIA) Alliance Technology Conference (ATCON-1), Bangalore, India, 2023, pp. 1-5, doi: 10.1109/ICAIA57370.2023.10169125. keywords: {Machine learning algorithms;Cows;Skin;Vaccines;Decision trees;Viruses (medical);Random forests;machine learning;lumpy skin disease;geospatial information;decision tree;random forest},

[15] N. K. Krishna, R. M. Benjamin and N. Tejaswi, "Machine Learning Diagnosis of Lumpy Skin Disease in Cattle Herds," 2024 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 2024, pp. 426-431, doi: 10.1109/ICAAIC60222.2024.10575895. keywords: {Support vector machines;Machine learning algorithms;Accuracy;Cows;Boosting;Prediction algorithms;Skin;Machine Learning;SVM;Gradient Boosting;Random Forest;Lumpy Skin Disease;Cattle Herds},

[16] Suparyati, E. Utami and A. H. Muhammad, "Lumpy Skin Disease Prediction Based on Meteorological and Geospatial Features using Random Forest Algorithm with Hyperparameter Tuning," 2022 5th International Conference on Information and Communications Technology (ICOIACT), Yogyakarta, Indonesia, 2022, pp. 99-104, doi: 10.1109/ICOIACT55506.2022.9971807. keywords: {Radio frequency;Measurement;Cows;Prediction algorithms;Skin;Classification algorithms;Tuning;Genetic Algorithm;Imbalanced Data;Lumpy Skin Disease;Machine Learning;Random Forest;SMOTE},

[17] Kholiya, P.S., Kriti, Mishra, A.K. (2024). Prediction of Lumpy Virus Skin Disease Using Artificial Intelligence. In: Tiwari, S., Trivedi, M.C., Kolhe, M.L., Singh, B.K. (eds) Advances in Data and Information Sciences. ICDIS 2023. Lecture Notes in Networks and Systems, vol 796. Springer, Singapore. https://doi.org/10.1007/978-981-99-6906-7_9

[18] P. Singh, J. Prakash, J. Srivastava and J. Srivastava, "Lumpy Skin Disease Virus Detection on Animals Through Machine Learning Method," 2023 Third International Conference on Secure Cyber Computing and Communication (ICSCCC), Jalandhar, India, 2023, pp. 481-486, doi: 10.1109/ICSCCC58608.2023.10176394. keywords: {Measurement;Economics;Computer viruses;Animals;Static VAr compensators;Machine learning;Production;Lumpy Skin Disease;Machine Learning Classifier;SVC},

[19] Saha, D. K. (2024). An extensive investigation of convolutional neural network designs for the diagnosis of lumpy skin disease in dairy cows. In Heliyon (Vol. 10, Issue 14, p. e34242). Elsevier BV. https://doi.org/10.1016/j.heliyon.2024.e34242

[20] V. Narayan et al., "Severity of Lumpy Disease detection based on Deep Learning Technique," 2023 International Conference on Disruptive Technologies (ICDT), Greater Noida, India, 2023, pp. 507-512,

doi: 10.1109/ICDT57929.2023.10150925. keywords: {Deep learning;Economics;Epidemics;Simulation;Insects;Cows;Feature extraction;Lumpy Skin Disease (LSD);Deep Learning;Transfer Learning;Machine Learning},

[21] Yang, Dw., Wang, W., Li, Jx. *et al.* Effect of Al_2O_3/SiO_2 ratio on morphology of complex calcium ferrite. *J. Iron Steel Res. Int.* **30**, 1921–1928 (2023). https://doi.org/10.1007/s42243-022-00903-3

[22] Huang, Y., Zhou, J., Ke, P. *et al.* A skin-integrated multimodal haptic interface for immersive tactile feedback. *Nat Electron* **6**, 1020–1031 (2023). https://doi.org/10.1038/s41928-023-01074-z

[23] Amar, T.A., Kumar, A. & Yadav, D.K. A Comparative Study on the Properties of Carbon Fiber-Reinforced Polymer Composites Developed by Hand Layup and Vacuum Bagging Molding Techniques. *J. Inst. Eng. India Ser. D* (2024). https://doi.org/10.1007/s40033-023-00631-2

[24] Yang, Y., Liu, Z. Experimental Study on Field Actual Shut-In After Fracturing of Shale Gas Reservoir. *Arab J Sci Eng* **48**, 16941–16951 (2023). https://doi.org/10.1007/s13369-023-08249-y

[25] Alaoui, I., El Ghadraoui, O., Tanji, K. *et al.* The Olive Mill Pomace: A Sustainable Biofertilizer to Improve Soil Proprieties and Plant Nutrient Uptake. *Waste Biomass Valor* **15**, 2575–2590 (2024). https://doi.org/10.1007/s12649-023-02324-z