I. J. Engineering and Manufacturing, **2025, 3, 1-17**
Published Online on June 8, 2025 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijem.2025.03.01

An Intelligent Deep Learning Model for Neonatal Asphyxia Identification Using Linear Graph Neural Networks Optimized with African Vultures Algorithm

Madhusundar Nelson
Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India
Email: madhusundarn9022.sse@saveetha.com
ORCID iD: https://orcid.org/0009-0005-9325-6623

Surendran Rajendran*
Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India
Email: surendranr.sse@saveetha.com
ORCID iD: https://orcid.org/0000-0002-6945-8461
*Corresponding Author

Received: 04 November, 2024; Revised: 03 January, 2025; Accepted: 27 January, 2025; Published: 08 June, 2025

Abstract: Birth marks the evolution from a fluid environment to one in which the baby breathes. Respiration difficulties after birth and in the first hours of life are common. Some babies will experience more severe respiration problems, and some of these will require medical care that only expert physicians can offer. In the first six hours of life, asphyxia is recognized. Neonatal asphyxia is a syndrome categorized by the deprivation of oxygen supply at birth and causes many morbidities and mortalities among newborns. Improved results depend on early detection and action. Using the African Vultures Algorithm that performs hyperparameter optimization expands accuracy and effectiveness in supporting the model. This integration improves the detection and sympathetic of neonatal asphyxia by using African Vultures Algorithm's capability to find fine results in high-dimensional places as well as the capability of LGNN to method structured clinical data. Herein, our model excels over present deep learning models and computer vision techniques for accurate detection of neonatal asphyxia. LGNN model achieved an accuracy of 98% in Dataset 1 and 96.5% in Dataset 2. F1-Score of the proposed model was achieved as 97.6% in Dataset 1 and 96.7% in Dataset 2. In Dataset 1 and Dataset 2, LGNN model as long as 97.9% and 97.1% precision, individually. In Dataset 1 and Dataset 2, the proposed (LGNN) provided 98.8% and 97.7% recall, one-to-one. Efficiently exploit the graph structure for feature aggregation and inference by employing a linear graph neural network architecture. Improved newborn outcomes are made possible by the suggested model, which is a potent diagnostic tool that also helps clinicians comprehend underlying pathophysiological causes.

Index Terms: Neonatal, Asphyxia, Linear Graph Neural Network (LGNN), Deep Learning, Computer Vision, African Vultures Algorithm

1. Introduction

Neonatal asphyxia which is characterized by the newborn receiving little oxygen at birth poses a serious risk to the health of newborns. Neonatal asphyxia continues to be a foremost source of illness and mortality in meanness of improvements in medical care, especially in environments with little resources. Mitigating the negative effects of asphyxia and enhancing newborn outcomes need prompt identification and management. Blood gas analysis, vital sign monitoring, and clinical evaluations are frequently used in the conventional methods of diagnosing newborn asphyxia. The sole viable treatment for "neonatal encephalopathy" is therapeutic hypothermia, which entails a body temperature reduction of 33.5 degrees Celsius within the first 72 hours of life.

Asphyxia versus hypoxia: Hypoxia is characterized by low oxygen delivery to tissues or low blood oxygen levels. It could be minor or really serious. While some biological tissues can withstand low oxygen levels over extended periods of time, others are more susceptible to injury. On the other hand, hypoxia occurs when your tissues are unable to absorb or utilize oxygen as well as they should, whilst asphyxia occurs when oxygen isn't reaching your airways.

Hypoxia can induce asphyxia, but not the other way around. Both hypoxia and asphyxia have an impact on how much oxygen your body gets. neonatal asphyxia, also known as birth asphyxia Pregnancy can sometimes result in an unborn kid receiving insufficient oxygen. This could occur, for instance, due to placental issues or a shortage of oxygen in the mother's blood [1,2]. If the labor was prolonged or there were issues with the umbilical cord, the baby might not get as much oxygen during birth. The duration of the baby's oxygen deprivation, the depth of their oxygenation, and the speed at which they receive medical attention all determine this is how dangerous.

Fig. 1. Sample image of Neonatal Asphyxia

Fig.1. A sample picture of asphyxia in neonates. Asphyxia can cause a variety of symptoms, some of which appear immediately and others which take time to manifest. These consist of: lips or face having a shade of red, purple, blue without intending to, urinate or defecate difficulties swallowing, breathlessness, Whether breathing quickly or deeply (hyperventilation). In severe cases, organ dysfunction or failure, including brain injury (hypoxic-ischemic encephalopathy), kidney failure, or damage to other vital organs. Although these methods yield useful data, they might not be sensitive enough or specific enough, which could cause delayed or incorrect diagnosis. The arising deep learning techniques these days have providing new directions for attractive the accuracy and performance of neonatal asphyxia identification. Neural networks have, in specific, demonstrated to be uniquely able to excerpt complex patterns from a variety of sources, counting structured clinical data and medical imaging. It remains challenging, in any case, to combine these two forms of data in a manner that fully imitates the complexity of newborn asphyxia [3,4].

The diagnosis of newborn asphyxia using the Linear Graph Neural Network, a precise deep learning architecture. Since heart rate and oxygen overload are two examples of physiological time series from neonates that are intrinsically linked, LGNNs are particularly good at processing graph-structured data. Here, it uses the aptitude of Linear Graph Neural Networks (LGNNs) to provide a novel deep learning framework for knowing neonatal hypoxia. This method goals to make use of the additional information that medical pictures provide, including CT or MRI scans of the newborn brain, along with structured clinical data, usually stored within CSV files. The dataset in this investigation contained of CSV files, for which neonatal brain MRI scan images were used to regulate the status of neonatal asphyxia. Using LGNNs, make one graph-based presentation of the data; graph node and edge features show indications based on clinical characteristics captured from CSV files. Also improve the discriminative capacity of the model by using advanced image processing techniques to augment the network with visual information taken from medical images, namely, MRI scans of the neonatal brain.

Different sources of data: Would collect the information for newborn asphyxia using organized clinical data, which is mainly stored in a CSV format, and medical imaging such as neonatal brain MRIs [5,6]. Fig.2. Example of neonatal asphyxia MRI. Uniform graph representation construction: Anticipate the graph representation to depict edges as being the representation of a relation between nodes, where nodes themselves are referred to clinical data that come from CSV files. In addition, we anticipate an extension of the graph representation by including relevant information via visualization from medical photographs. The area of our project on the development of a deep learning model will be a novel deep architecture constructed on top of Linear Graph Neural Networks (LGNNs). This graph representation has made it easier to deal and integrate the various forms of heterogeneous data such that the diagnosis of newborn asphyxia will be easy. Training and assessment on a wide range of datasets: We will use a large collection of picture and CSV data obtained from newborn patients with different levels of severe asphyxia to train the suggested deep learning model. We will undertake a systematic scrutiny of the model's routine to determine its precision, sensitivity, specificity, F1score and accuracy Furthermore, research aim is to offer a knowledge of the crucial elements that contribute to the model's decision-making process, which will improve physicians' comprehension of the pathophysiology of newborn asphyxia and facilitate customized treatment plans.

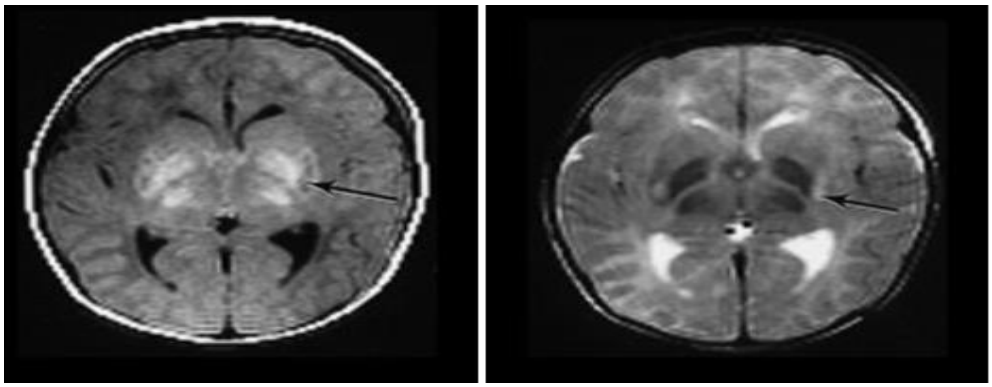

Fig. 2. Sample MRI image of Neonatal Asphyxia

Neonatal asphyxia is the condition that involves insufficient oxygen supply to the newborn at birth. It still represents a significant cause of morbidity and mortality, mainly in low-resource settings. Despite advancements in medical care, it remains hard to be diagnosed early and appropriately managed. Thus, it demands new approaches towards better clinical outcomes. The traditional approaches include blood gas analysis, monitoring of vital signs, and clinical assessment, which are not very sensitive and specific for early and timely diagnosis. What is needed still is a uniform framework of such heterogeneous data types, which could include structured clinical data and medical imaging, to diagnose neonatal asphyxia. Deep learning models are only partially able to exploit the relationship present in such data sources.

Contribution and Novelty: The proposed outline introduces a Linear Graph Neural Network (LGNN) for neonatal asphyxia diagnosis that exploits the strength of graph-structured data processing. The LGNN successfully merges physiological time-series data and medical imaging, which captures subtle patterns associated with asphyxia. Hyperparameter optimization is performed through the consumption of the African Vultures Algorithm that attracts the efficiency and accuracy of the model.

This model will be able to detect neonatal asphyxia more accurately than the existing methods and will classify infants who have experienced this pathology with precision, recall, and high F1-scores on several datasets. A holistic approach based on a fusion of image processing and graph-based learning offers insights into pathophysiology as a complement to this diagnosis tool that can pave the way toward developing more tailored therapies.

2. Recent Works

Graph Convolutional Neural Network (GCN), Graph Attention Network (GAT), and the primary break process are integrated with the graph neural network model ENode-GAT to improve the exactness of minor trial node categorization consuming the approach of outside reference of similar word nodes, With some classification benefits, the experimental findings show that the ENode-GAT model suggested in this study achieves 85.1% sorting accuracy on the Cora dataset and 85.3% cataloging accuracy on the Stock dataset by Zhang et al. An LSTM collector and biased pooling to create a unique graph neural network architecture for the speech-graph utilizing feature similarity. The IEMOCAP dataset yielded an unweighted accuracy of 65.39% and a subjective accuracy of 71.83%. Li et al. [1,2]. Early detection and accurate diagnosis of newborn hypoxia may lead to better results. The goal of recent research is to find indicators for diagnosis in newborns who could suffer brain injury Boskabadi H et al. Considering the growing capability of AI, talk about potential guidelines for new models and the future of neonatology, offering strategies for integrating AI into newborn intensive care facilities. The persistence of the learning was to appraise the potential investigative use of novel biomarkers for newborn asphyxia. Keles E et al. [3,4].

AI tools tested in neonatology contain dynamic sign nursing; risk social lamination (retinopathy of prematureness, abdominal puncture, jaundice), disease forecast (breathing suffering disorder, bronchopulmonary dysplasia, apnea of prematureness), and neurological investigative and predictive support (electroencephalograms, sleep phase arrangement, neuroimaging). Innovative image gratitude knowledges are especially helpful for early infection credit Chioma et al. A summary of brain damage associated with hypoxic-ischemic newborn encephalopathy. Predominant were the WM with periventricular sublocation lesions and the basal ganglia with thalamus damage. In the context of HIE, a thorough diagnosis of brain lesions may shed light on the mechanism and timing of occurrence Beck J et al. [5,6]. Evaluating brain damage in babies with HIE, magnetic resonance imaging (MRI), which incorporates diffusion-weighted imaging, has emerged as the gilt standard and has good prognostic value. Magnetic resonance spectroscopy has been demonstrated to be a dependable prognostic biomarker and offers supplementary metabolic information. More and more researches are being done to learn more about the causes and outcomes of brain injury through the use of advanced imaging modalities including arterial spin labelling and diffusion tensor imaging Parmentier CEJ et al.

Seven characteristics of the taxonomy image categories, data augmentation, features extraction, deep learning algorithm types, transfer learning, the collaborative of classifiers. Other researchers could use the provided taxonomy to organize their contributions to and activities related to their research. The proposed future path may enhance the

effectiveness and increase the number of applications for deep learning-assisted lung disease detection Kieu et al. [7,8]. LGNN allows nodes to iteratively gather and acquire the feature evidence of high-order neighbors by manipulative a high-order adjacency matrix broadcast mechanism. LGNN employs a straightforward direct mapping to preserve computational proficiency and produce the final node illustration once it has obtained the node illustration of the network topology Cao Shujuan et al. a novel fusion deep learning system by merging CNN, spatial transformer network (STN), data expansion, and VGG. Here,refer to this novel hybrid method as VGG Data STN with CNN (VDSNet). The dataset of NIH chest X-ray images that was gathered from the Kaggle fountain is subjected to the new prototypical. Both the dataset's full and sample versions are taken into account. VDSNet performs better than current procedures for both full and sample datasets in standings of several system of measurement such as validation accuracy, precision, recall, and F0.5 score. While vanilla grey, vanilla RGB, hybrid CNN and VGG, and modified capsule network have accuracy values of 67.8%, 69%, 69.5%, and 63.8%, respectively, VDSNet shows a justification accuracy of 73% for the complete dataset Subrato Bharati et al. [9,10].

3. Materials and Methods

In order to obtain and preprocess the csv-based data Synthetic-Infant-Health-Data dataset from Kaggle in order to generate a clever deep learning model for the identification of neonatal asphyxia. Electronic medical records (EMRs) are the data source. They include sequential information as well as physiological data (LVHreport, LowerBodyO2, RUQO2, CO2Report etc..). These procedures have to be followed in order to preprocess the Synthetic-Infant-Health-Data dataset in demand to make a deep learning prototypical to identify newborn asphyxia.

3.1 Improving the Quality of CSV Data

Managing Missing Values: As was already established, missing values can seriously affect how well a model performs. methods of imputation such as mean. The choice of outlier detection and removal method be contingent on numerous issues, counting the landscape of the data, the distribution of the facts, and the detailed necessities of the problem you're addressing. Here are some common methods for outlier detection and removal. Data points that advance extra prearranged quantity of standard deviations after the mean are measured outliers according to the standard deviation method. Z-Score Method: Determine each data point's z-score, which expresses how far off the mean it is by a given number of standard deviations. Outliers are data points that have a z-score greater than a predetermined level, such 3. This technique is comparable to the standard deviation method, but adjust the threshold according to how sensitive want to be to outliers.

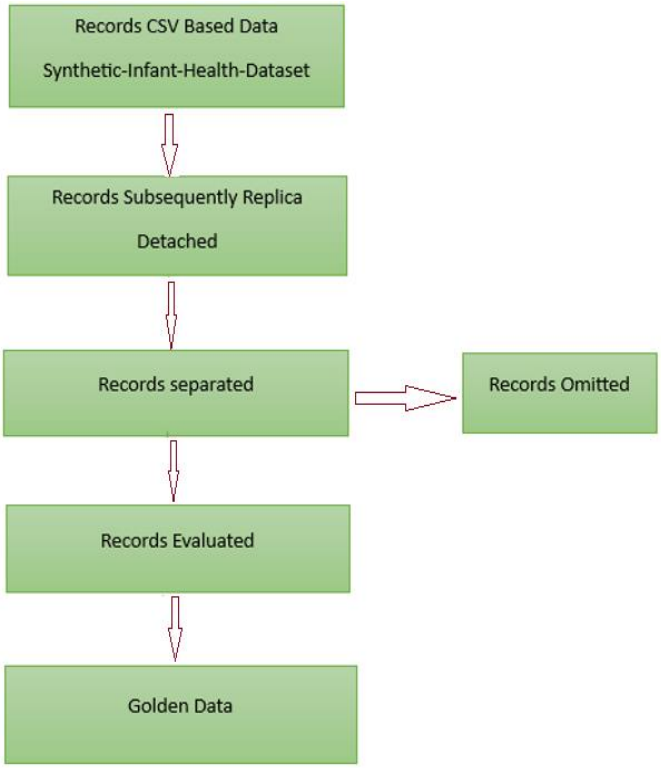

Fig. 3. Preprocessing steps of electronic medical records (EMRs)

Table 1. Characteristics of Neonatal as per medical records

Newborn ≥ 36 weeks		
Patient Characteristics	**All**	**Therapeutic Hypothermia**
	n (%) Mean ±SD	**n Collection (%) Mean ±SD**
Sex, male	278/520 (53.5)	255/479 (53.2)
Cannulation in delivery room	388/517 (75.1)	368/476 (77.3)
Birth openair a TH center	377/520 (72.5)	349/479 (72.9)
Delivery mode	511/520	470/479
Cesarean	310 (60.7)	289 (61.5)
Sentinel event	264/520 (50.8)	248/479 (51.8)
Birth weight, g	3172 ±537	3180 ±523
Apgar score at 5 min < 5	283/520 (54.4)	268/479 (55.9)
Apgar score at 10 min < 5	219/439 (49.9)	207/403 (51.4)
Vaginal, no instrumental removal	89 (17.4)	78 (16.6)
Vaginal, instrumental removal	122 (21.9)	103 (21.9)
First-hour breast-feed (mmol/L)	12.43 ±4.96	12.50 ±4.97
Hypoglycemia ≤ 24 h of life	36/435 (8.3)	35/400 (8.8)
First-hour base surplus (mmol/L)	11.81 ±6.64	11.93 ±6.64
First-hour pH	6.97 ±0.18	6.96 ±0.18
Glycemia at entry (mmol/L)	6.87 ±4.59	6.95 ±4.67
Seizures > 24 h of lifespan	99/514 (19.3)	91/473 (19.2)
Seizures ≤ 24 h of lifespan	145/514 (28.2)	127/473 (26.9)
Death throughout hospitalization	81/520 (15.6)	74/479 (15.4)

A grade assigned by Sarnat & Sarnat [11] to therapeutic hypothermia (TH). A glucose level of less than 2.2 mmol/L is considered hypoglycemia.

3.2 Obtaining and Preparing Image Data

Data collection: Get an MRI image dataset of newborns, both those with and without asphyxia diagnoses. Data Annotation: Label the MRI pictures to show whether or not neonatal asphyxia is present. Automated systems or medical professionals can perform this. Data preprocessing: To improve the diversity and size of the dataset, preprocess the MRI pictures through shrinking them to a uniform magnitude, normalizing pixel intensities, and maybe elevating the data.

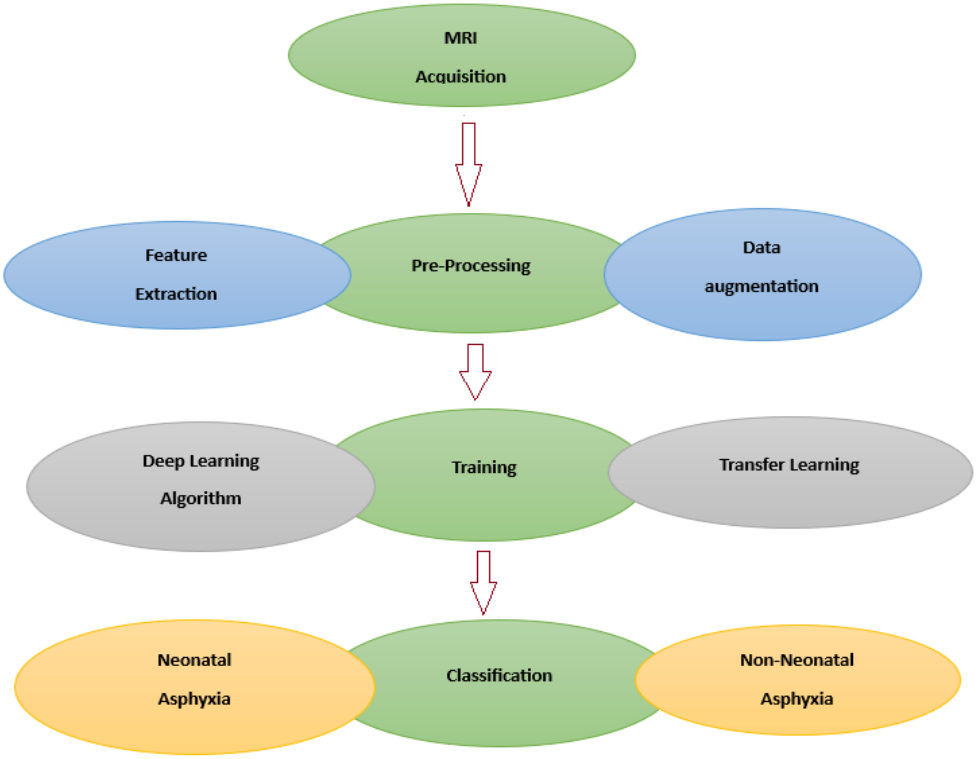

Fig. 4. Overview of MRI image based Neonatal Asphyxia

Data augmentation: To insincerely enlarge the size of the training dataset, MRI pictures can be subjected to augmentation techniques such rotation, flipping, scaling, and noise addition. This strengthens the deep learning model's capacity for generalization and increases its resistance to changes in the input data. Annotation and Labelling: Relevant information, such as the existence of anomalies or the anatomical structures of interest, must be annotated or labelled on MRI images in many deep learning applications. The supervised learning models are trained on this annotated data.

Table 2. Imaging findings of deviations for each MRI mode

Outline of Wound	Imaging mode	Imaging Discoveries	Period Frame of Deviations
BGT	SWI	Noticeable, low-intensity veins	Although there isn't much research available right now, protruding hypo-intense strains have been seen as primary as 18 hours later delivery.
	T1WI/T2WI	Irregular sign strength in the thalami, perirolandic cortex, and basal ganglia. lack of a PLIC with a typical high signal strength.	Anomalies that are initially insignificant or minor but progressively more noticeable by the second part of the first week after the insult. After a month, an MRI may reveal gliosis, capacity loss, cysts, and poor myelination in the perirolandic cortex and central grey substance.
	H-MRS	Reduced NAA and increased lactate in the thalami and basal ganglia.	In general, lactate rises.
	DWI	In impacted locations, there is a high motion strength on isotropous DWI with short ADC ideals.	Later the insult, irregularities highest in 3-5 days later. For infants receiving therapeutic hypothermia, pseudo-normalization happens next roughly 11–12 days, and for neonatal not receiving cooling, it happens after 6–8 days.
WM/WS	SWI	Noticeable hypo-intense veins and weak signal strength where the hemorrhagic lesions are located.	Although there isn't much research available right now, protuberant hypo-intense veins have been seen as initial as 18 hours afterwards the delivery. Hemorrhagic lesion sites have low signal intensity right away and may continue for several months.
	T1WI/T2WI	Irregular sign strength in the surrounding cortex and white substance of the cerebral artery crunch regions in children with severe disease. Loss of gray-white substance dissimilarity at the cortical level may be seen in T2WI.	Anomalies that are initially insignificant or minor but progressively more noticeable by the second part of the first week after the insult. After a month, cortical retreating, white substance size loss, cysts, and gliosis of the lens cortex and white substance can all be shown on an MRI.
	H-MRS	Lower NAA and higher breast-feed in the squeezed white matter.	In overall, breast-feed levels rise.
	DWI	Low ADC values and high signal strength on isotropic DWI in the impacted locations	After the insult, abnormalities peak 3-5 days later. For infants receiving therapeutic hypothermia, pseudo-normalization happens after roughly 11–12 days, and for neonatal not receiving cooling, it happens after 6–8 days.

Plan from earlier research conducted by Barkovich et al. Described by Penrice et al. and Barkovich et [13,14] al ADC: superficial dispersion coefficient; BGT: basal ganglia and thalami; DWI: diffusion-weighted imaging; H-MRS:proton magnetic resonance spectroscopy; MRI: magnetic quality imaging; NAA: N-acetylaspartate; PLIC:posterior limb of the inner container; SWI: susceptibility weighted imaging; T1WI: T1-weighted imaging; T2WI:T2-weighted imaging; WM/WS: white substance/watershed.

3.3 Linear Graph neural Network

This work suggests an LGNN neural network, whose topology is established as illustrated in, to address the issue of shallow linear networks' inadequate perception of high-order neighborhood information. Initially, the input graph is pre-processed using balanced normalization and feature normalization at the graph representation and preprocessing stage in order to produce a feature matrix and a normalized adjacency matrix. Secondly, the last node illustration data is gained through effective linear revolution mapping, and the high-order neighbor feature evidence of the node is repeatedly gathered by creating a high-order adjacency matrix dissemination process.

ADC image generation and ROI selection:

To accomplish every objective related to image processing, including feature extraction, ROI selection, ADC image synthesis, image selection, and visualization. Every ROI was chosen by hand.

$$\mathrm{ADC} = \sum_{i=1}^{n} \frac{\ln \frac{T_i}{T_0}}{b_i} \tag{1}$$

By creating a 3D region of interest (ROI) that included the asphyxia, the anomalous range of each 2D ADC picture portion remained identified. The pixel values inside the targeted area were then retrieved.

Feature extraction:

Eqs. and were utilized to compute the higher order instant values and the mean pixel value inside the ROI. V is the mean of the pixel values, f (Vi) is the possibility of the signal strength of a pixel, and Vi is the signal strength in the ith pixel. N is the whole amount of pixels within the ROI.

$$\text{Mean } = \frac{\sum_{i=1} V_i}{S} \tag{2}$$

$$n\text{th moment } = \sum_i (V_i - V)^n f(V_i) \tag{3}$$

$$O_{f,\delta}(p,m) = \sum_{x,y=1}^{n} \begin{cases} 1 \text{ if } f(w,s) = p \text{ and } f(s + \delta_{x,y} + \delta_y) = x \\ 0 \text{ otherwise} \end{cases} \tag{4}$$

From the produced GLCM matrices, the arithmetical texture properties of GLCM (mean, variance, energy, entropy, contrast, homogeneity, association, prominence, and shade values) were retrieved. Additionally, the GLCM features were retrieved using Equations 5 through 13.

GLCM mean

$$\mu_i = \sum_{k,l=0}^{N-1} i(U_{k,l})\mu_j = \sum_{k,l=0}^{N-1} j(U_{k,l}) \tag{5}$$

GLCM variance

$$\sigma_i^2 = \sum_{k,l=0}^{N-1} U_{k,l}(i - \mu_i)^2 \sigma_j^2 = \sum_{k,l=0}^{N-1} V_{k,l}(l - \mu_j)^2 \tag{6}$$

GLCM energy (Ey)

$$\text{Ey} = \sum_{k,l=0}^{N-1} U_{k,l}^2 \tag{7}$$

Entropy (Ety)

$$\text{Ety} = \sum_{k,l=0}^{N-1} U_{k,l}(-\ln U_{k,l}) \tag{8}$$

GLCM contrast (Cont)

$$\text{Cont} = \sum_{k,l=0}^{N-1} U_{k,l}(k - l)^2 \tag{9}$$

Homogeneity (Homoy)

$$\text{Homoy } = \sum_{k,l=0}^{N-1} \frac{U_{k,l}}{1+(k-l)^2} \tag{10}$$

Association (Assn)

$$\text{Assn} = \sum_{k,l=0}^{N-1} U_{k,l} \left[\frac{(k-\mu_k)(l-\mu_l)}{\sqrt{(\sigma_k^2)(\sigma_l^2)}} \right] \tag{11}$$

Cluster shade (Sde)

$$\text{Sde} = \sum_{k,l=0}^{N-1} \{k + l - \mu_k - \mu_l\}^3 U_{k,l} \tag{12}$$

Cluster prominence (Prome)

$$\text{Prome} = \sum_{k,l=0}^{N-1} \{k + l - \mu_k - \mu_l\}^4 U_{k,l} \tag{13}$$

3.3.1 Graph Representation

As an pointless graph G= (V, E, X), there are nodes vk ∈V and edges (Vk, Vl) ∈ E. LGNN uses the similar depth as GCN and GraphSage, i.e., a two-layer fully connected neural network construction, to address the issue that a single-channel low graph neural network cannot imprisonment high command constructions. Additionally, stop using any techniques or parts that are frequently found in neural network architectures, such as activation functions and density processes. The node's data is combined iteratively, and effective linear transformation mapping yields the final node representation data.

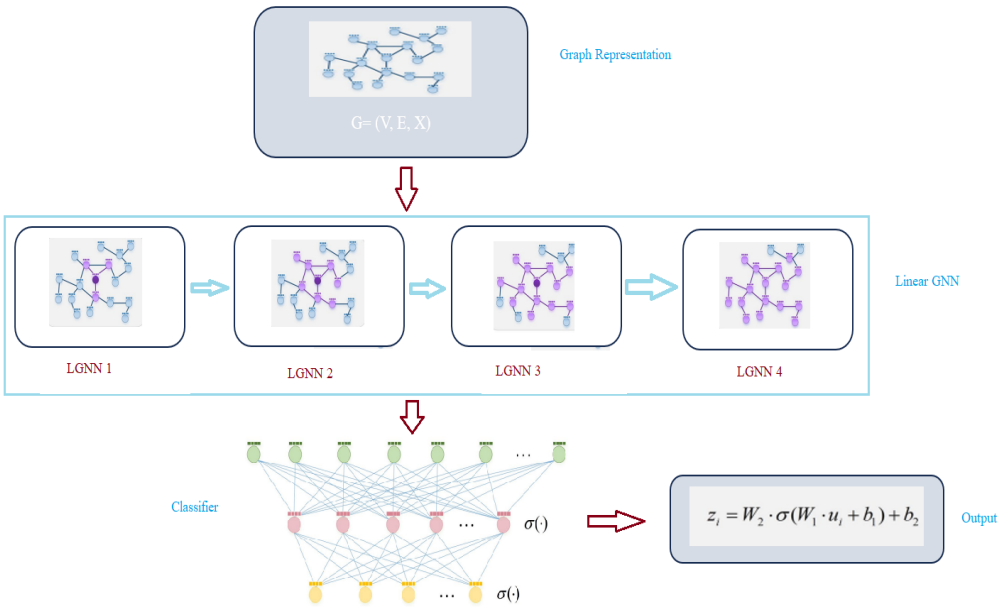

Fig. 5. Framework of a Linear Graph neural Network

For each node vi ∈V, there is a primary feature illustration xi ∈Rd, where d is the measurement of the feature. In the LGNN model, use the adjacency matrix B ∈RN* N to epitomize the regional anatomy of the graph, where Bij =1 represents that there are edges between nodes vi and v j, and Bij = 0 signifies that there are no edges amongst them. In directive to avoid the loss of data, regularly control the adjacency matrix:

$$B^{(sym)} = \mathrm{E}^{-1/2} \cdot \hat{B} \cdot \mathrm{E}^{-1/2} \tag{14}$$

Amongst them, $\hat{B}$ = B+ I characterize the self-ring matrix extra on the source of the adjacency matrix, E is the grade matrix.

Define it as taking into account the variation in node features, eliminating the impact of node degree, and achieving a additional consistent illustration.

$$\hat{\mathrm{Y}} = \frac{\mathrm{Y}-\mu}{\delta} \tag{15}$$

The normalized feature matrix is denoted by $\hat{\mathrm{Y}}$, while the mean and standard deviation of feature matrix Y are represented by μ and δ, respectively. Dissimilar node attributes have comparable scales because to the standardized approach, which enhances the LGNN model's performance and convergence rate.

Higher-Order neighbor Dissemination:

Towards better detention the intricate relationships among nodes in the network, we suggest a feature learning technique that utilizes high-order adjacency matrix circulation. To investigate the relationship between nodes at various orders, we develop a linear propagation approach grounded on the provided node feature matrix Y and standardized B (sym). The apprise equation of node features in k-order propagation is computed as follows.

$$\mathrm{m}_i^{(\mathrm{k})} = B^{(sym)} \cdot \hat{\mathrm{y}}_i \tag{16}$$

The k-th-order post-propagation node feature illustration is denoted by $\mathrm{m}_i^{(\mathrm{k})}$ in this instance. By using this communication method, the node features are able to improved depict the properties of nodes in high-order graded associations by capturing high-order neighbor information.

$$B^{(sym)(\mathrm{k})} = \prod_{i=1}^{k} B^{(sym)(\mathrm{i})} \tag{17}$$

A multi-scale feature fusion approach that enables the thorough utilization of circulation data of various orders. Specifically, we create a multi-scale node feature tensor by multiplying the node features of individually directive by a matching weight matrix and stacking the features of all orders in the route of the channel.

$$z_i = \mathrm{n}sf\left(\left[\mathrm{m}_i^{(1)} \cdot \mathrm{O}^{(1)} \parallel \mathrm{m}_i^{(2)} \cdot \mathrm{O}^{(2)}\right] \parallel \cdots \parallel \mathrm{m}_i^{(k)} \cdot \mathrm{O}^{(k)}\right) \tag{18}$$

nsf (.) signifies the multi-scale synthesis method, O (1) …O(k)signifies the weight matrix of the appropriate order throughout circulation, and || denotes the tensor edging technique. By using a multi-scale feature combination method, LGNN can better graph and represent graph structure by extracting ironic evidence from the propagation of various orders. We present a regularization technique to preserve the steady circulation of gradient and prevent the issue of gradient desertion during the propagation of the multi-order adjacency matrix:

$$\mathrm{Q}_i = \frac{\hat{y}_k - \frac{1}{\mathrm{n}}\sum_{l=1}^{\mathrm{n}} \hat{y}_k}{\sqrt{\frac{1}{\mathrm{n}}\sum_{l=1}^{\mathrm{n}} \left(\hat{y}_k - \frac{1}{\mathrm{n}}\sum_{l=1}^{\mathrm{n}} \hat{y}_k\right)^2 + \xi}} \tag{19}$$

A constant among them to prevent divisor zero is ξ. Furthermore, to further improve the expression capacity of features, we further map features to the semantic space using multi-layer perceptron's (MLPs):

$$\mathrm{z}_i = \mathrm{O}_2 \cdot \sigma(\mathrm{O}_1 \cdot \mathrm{u}_i + \mathrm{C}_1) + \mathrm{C}_2 \tag{20}$$

3.3.2 Training and Optimization

Two pieces make up the loss function for training a model. The negative log-likelihood loss function is the first and has the following mathematical expression:

$$\begin{aligned} \text{Loss} &= l1 + l2 \\ l1 &= -\frac{1}{\mathrm{P}}\sum_{i=1}^{\mathrm{P}} \log(\text{softmax}(\text{ lgits }_i)[\text{ labels }_i]) \end{aligned} \tag{21}$$

where labelsi is the actual labelled values and lgitsi is the model's output, a tensor comprehending the log chance standards of the anticipated upshots. The log prospect values are converted to probability values using the function softmax(.). By linking the anticipated probability dispersal with the genuine label, this loss function determines the discrepancy between the model estimate and the true label. The L2 regularization term, or second portion l2, is added by appending a penalty term to the loss function. In order to increase the model's size for generalization and avoid overfitting, the goal is to reduce the undesirable log-likelihood loss whereas limiting the model's masses over the regularization term. The negative log-likelihood loss role.

$$L(\theta) = -\sum_{i=1}^{\mathrm{p}} \log\left(\mathrm{S}_{\text{model}}(\mathrm{x}_i \mid \mathrm{y}_i; \theta)\right) \tag{22}$$

The suggested LGNNs are given numerous versions in this work, which are explained as follows: LGNNoriginal: The node's representation vector in each input feature is not multiplied by $\hat{B}$. After the node's features are directly entered, a two-layer fully associated network is assembled for drill. LGNN1: Following the multiplication of the node's feature matrix by $\hat{B}$ 1 for each input feature, a two-layer fully connected network is constructed for exercise.

$$\begin{cases} \mathrm{m}_i^{(1)} = B^{(sym)(1)} \cdot \hat{\mathrm{y}}_i \\ \mathrm{m}_i^{(2)} = B^{(sym)(2)} \cdot \hat{\mathrm{y}}_i \\ \mathrm{m}_i^{(3)} = B^{(sym)(3)} \cdot \hat{\mathrm{y}}_i \\ \mathrm{m}_i^{(4)} = B^{(sym)(4)} \cdot \hat{\mathrm{y}}_i \end{cases} \tag{23}$$

LGNN2: For every input feature, B^ 1 and B^ 2 reproduce the node's feature matrix that results in two feature matrices. These are then horizontally split with the rows being fed as input features to the algorithm, and a two-layered fully-connected network was constructed for training.

LGNN3: For every feature, the feature matrix of node gets multiplied by B^ 1, B 2 and B^ 3 and leads to three feature matrices. These three feature matrices are divided in parallel wherein every row serves as an input feature to this process. This in turn, configures a fully connected, two-layer network for training.

LGNN4: The node's feature matrix is increased by B^ 1, B^ 2, B^ 3 and B 4 each time the input features the outcome is four feature matrices, which are then split horizontally. Each row of the four feature matrices is then used as an involvement feature in the technique and a two-layered fully-connected network is built for exercise.

3.4 African Vultures optimization algorithm

The African Vultures Optimization Algorithm simulates the African vulture behavior in hunting and navigation. Use the explore(vulture) function to explore your search space while applying random perturbation to your hyperparameters that define the vulture. To assess how well the vulture's hyperparameters are doing, you would compute the objective function in the evaluate(vulture) function. The best-performing vultures are chosen using the select_best() method, and information sharing between vultures is facilitated via the share_information() function. Lastly, the population is updated via update_population() using the shared data and the chosen vultures.

$$F(i) = \begin{cases} \text{Superlative Vulture}_{\,1} \text{ if}_{\,i} = E_1 \\ \text{Superlative Vulture}_{\,2} \text{ if}_{\,i} = E_2 \end{cases} \tag{24}$$

The superlative result is chosen as the best vulture of the first collection and the second- superlative result as the superlative vulture of the second collection after the initial population is formed. The other solutions then use a change near the superlative solutions for the first and second groups. The suitability of all the solutions is calculated. Every fitness iteration involves recalculating the entire population.

$$\begin{aligned} C_1 &= \text{Superlative Vulture}_{\,1}(i) - \frac{\text{Superlative Vulture}_{\,1}(i) \times P(i)}{\text{Superlative Vulture}_{\,1}(i) - P(i)^2} \times F \\ C_2 &= \text{Superlative Vulture}_{\,2}(i) - \frac{\text{Superlative Vulture}_{\,2}(i) \times P(i)}{\text{SuperlativeVulture}_{\,2}(i) - P(i)^2} \times F \end{aligned} \tag{25}$$

Every vulture's path towards the food source is investigated. A number of vulture classes may collect everywhere a single food source owing to intense race for food and infrequent starvation. This vulture movement has been formulated using Equation (25).

$$P(i+1) = \frac{C_1 + C_2}{2} \tag{26}$$

The LFs were calculated by using

$$LF(x) = 0.01 \times \frac{u \times \sigma}{\frac{1}{|v|^{\beta}}}, \sigma = \left(\frac{\Gamma(1+\beta) \times \sin\left(\frac{\pi\beta}{2}\right)}{\Gamma(1+\beta 2) \times \beta \times 2\left(\frac{\beta-1}{2}\right)} \right)^{\frac{1}{\beta}} \tag{27}$$

In Eq. (27), d signifies the problem scopes, u and v are a chance number between 0 and 1, and β is a static and avoidance number of 1.5.

4. Experimental Results and Discussion

The experiment confirms the model's performance and flexibility by testing it on a publicly available data set. the Synthetic-Infant-Health-Data (Dataset 1). This is a dataset on infant health and artificially generated. It is supposed to look and behave like actual data but contains no real infant health data. Much information about the infants is found in the dataset, including age in days, birth diseases, body oxygen and carbon dioxide levels, and chest x-rays. This synthetic dataset will be helpful for research and analysis because it gives a real view of the newborn dataset without compromising on privacy or using actual personal data.

MRI of the Neonatal Brain - Mary A Rutherford files databases (Dataset 2) Nearly as momentous is the movement of diagnoses from adults to children, and from children to neonates. Magnetic resonance imaging offers a complete view of the extraordinary deviations that happen in the normal brain, which form the backdrop to the newborn's vast range of pathologies. Accuracy indicators for baseline neural networks and LGNNs when carrying out downstream tasks for node classification.

CNN, Deep Walk, GCN, GNN and LGNN methods are used for comparison with different performance metrics like accuracy, F1-Score, Precision, Recall were evaluated for two datasets like the Synthetic-Infant-Health-Data (Dataset 1) and MRI of the Neonatal Brain - Mary A Rutherford files databases (Dataset 2).

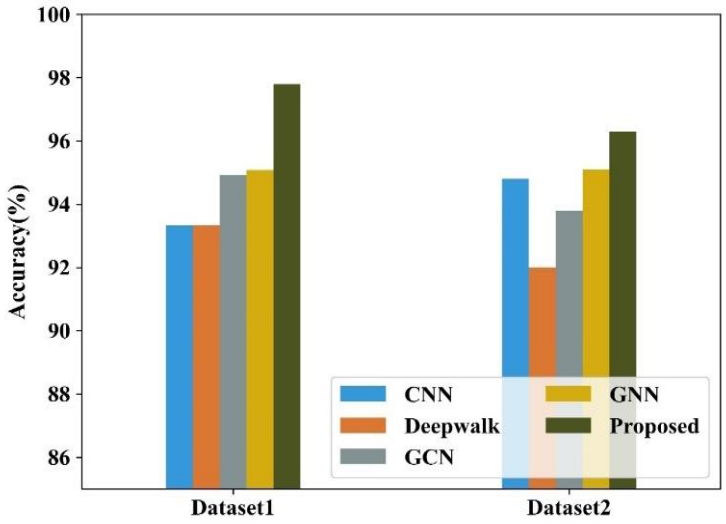

Fig. 6. Accuracy

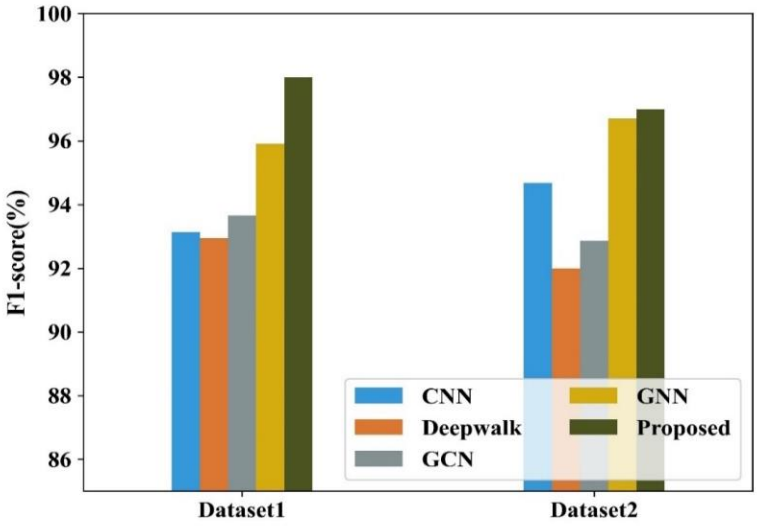

Fig. 7. F1-Score

A bar graph comparing the accuracy of various models for a task is shown in Fig.6. Accuracy. The graph's y-axis is labelled "Dataset," while the x-axis is labelled "Accuracy (%)." Dataset 1 and Dataset 2 are the two datasets that are displayed on the y-axis. The various models and their accuracy on the two datasets are broken down as follows:93% on Dataset 1 and 95% on Dataset 2 were accurately recorded by CNN; 94.5% on Dataset 1 and 95% on Dataset 2 were accurately recorded by GNN; 93% on Dataset 1 and 91.8% on Dataset 2 were accurately recorded by Deep walk; and the accuracy of 95% on Dataset 1 and 93.2% on Dataset 2 was proposed by GCN. The Proposed (LGNN) gave 98% accuracy in Dataset 1 and 96.5 % in Dataset 2.

A bar graph likening the F1-Score of various models for a task is shown in Fig.7. F1-Score. The graph's y-axis is labelled "Dataset," while the x-axis is labelled " F1-Score (%)." Dataset 1 and Dataset 2 are the two datasets that are displayed on the y-axis. The various models and their F1-Score on the two datasets are broken down as follows:93% on Dataset 1 and 95% on Dataset 2 were F1-Score logged by CNN; 95.5% on Dataset 1 and 97% on Dataset 2 were F1-Score logged by GNN; 93% on Dataset 1 and 92% on Dataset 2 were F1-Score documented by Deep walk; and the F1-Score of 95% on Dataset 1 and 93.4% on Dataset 2 was projected by GCN. The Proposed (LGNN) gave 97.6% F1-Score in Dataset 1 and 96.7 % in Dataset 2.

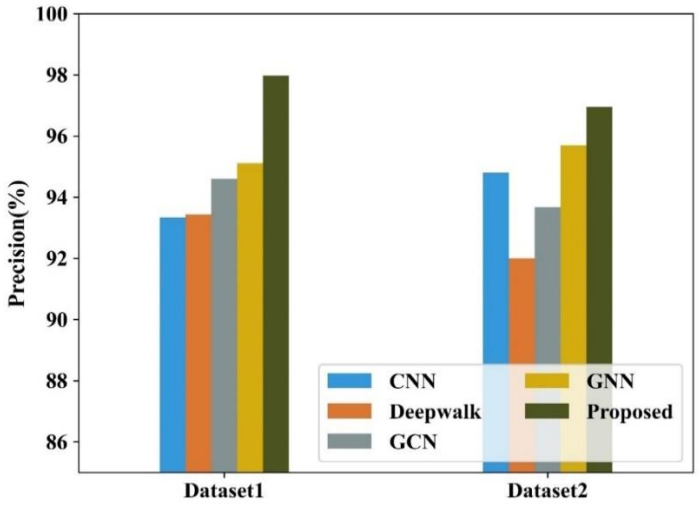

Fig. 8. Precision

A bar graph relating the Precision of various models for a task is shown in Fig.8. Precision. The graph's y-axis is labelled "Dataset," while the x-axis is labelled " Precision (%)." Dataset 1 and Dataset 2 are the two datasets that are exhibited on the y-axis. The various models and their Precision on the two datasets are broken down as follows:93% on Dataset 1 and 95% on Dataset 2 were Precision logged by CNN; 95.5% on Dataset 1 and 95.8% on Dataset 2 were Precision logged by GNN; 93.2% on Dataset 1 and 91.8% on Dataset 2 were Precision documented by Deep walk; and the Precision of 94.8% on Dataset 1 and 93.8% on Dataset 2 was projected by GCN. The Proposed (LGNN) gave 97.9% Precision in Dataset 1 and 97.1 % in Dataset 2.

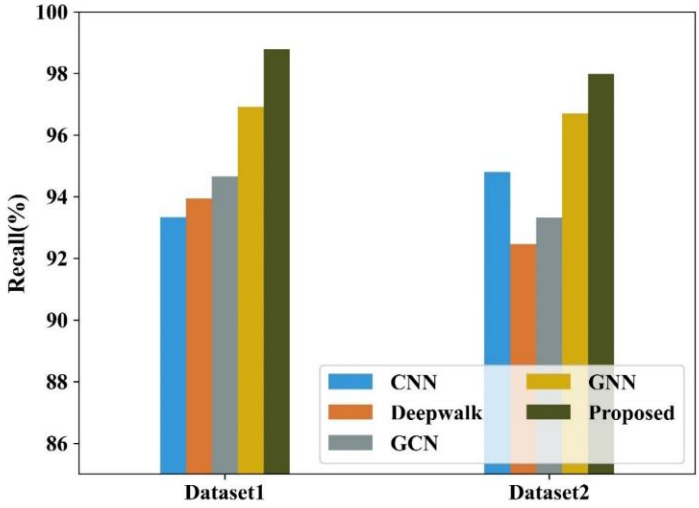

Fig. 9. Recall

A bar graph concerning the Recall of various models for a mission is shown in Fig.9. The graph's y-axis is categorized "Dataset," while the x-axis is labelled " Recall (%)." Dataset 1 and Dataset 2 are the two datasets that are revealed on the y-axis. The various models and their Recall on the two datasets are broken down as follows:93% on Dataset 1 and 95% on Dataset 2 were Recall charted by CNN; 95.5% on Dataset 1 and 95.8% on Dataset 2 were Recall logged by GNN; 93.2% on Dataset 1 and 92% on Dataset 2 were Recall recognized by Deep walk; and the Recall of 94.7% on Dataset 1 and 92.8% on Dataset 2 was projected by GCN. The Proposed (LGNN) gave 98.8% Recall in Dataset 1 and 97.7 % in Dataset 2.

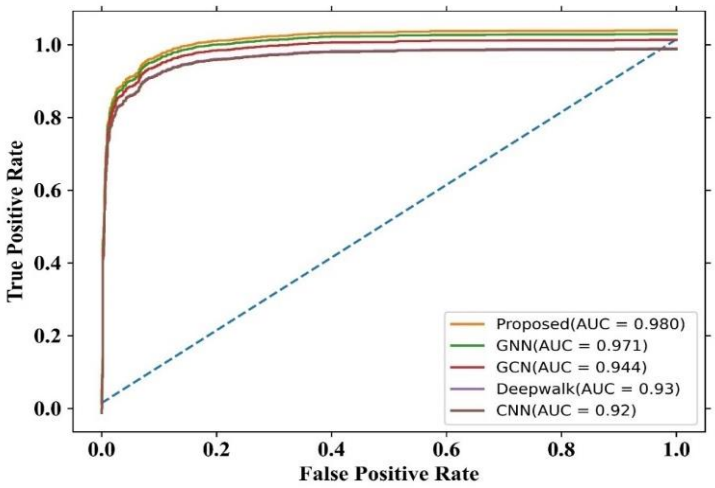

Fig. 10. ROC for Dataset 1

The ROC curve provides the following information regarding the models' performance on Dataset 1: The False Positive Rate (FPR) is plotted on the x-axis. The True Positive Rate (TPR), sometimes referred to as Recall, is shown on the y-axis. A conventional line in the upper left corner of the graph would characterize the faultless ROC curve. This means a model (FPR = 0, TPR = 1) that can distinguish well between positive and negative cases. The model in the image proposed here, LGNN has the highest AUC. The curve itself has this AUC value of 0.980. A higher AUC means greater general effectiveness of the model in the differentiation between positive and negative cases.

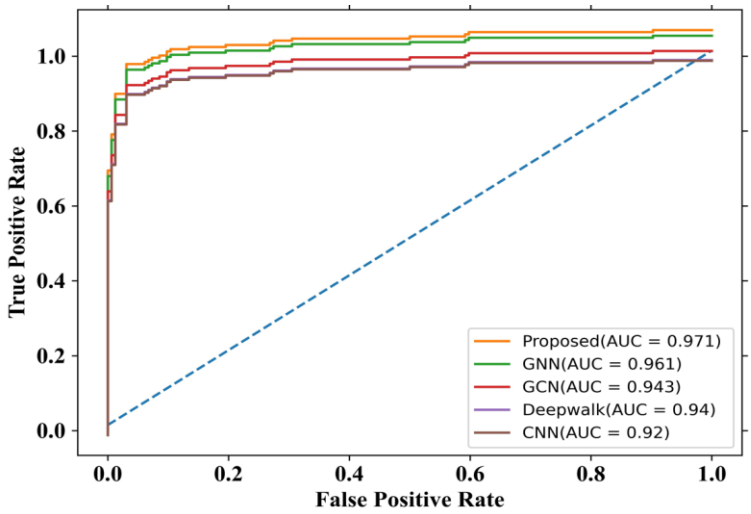

Fig. 11. ROC for Dataset 2

ROC Curve The following describes how the models performed on Dataset 2. The False Positive Rate (FPR) is plotted along the x-axis. The True Positive Rate (TPR), also known as Recall, is exposed along the y-axis. A conservative line in the upper left corner of the graph would denote the faultless ROC curve. This advocates a model (FPR = 0, TPR = 1) that can precisely distinguish between positive and negative situations. The proposed model (LGNN) in the picture has the highest AUC (Area Under the Curve). The curve itself has this AUC value of 0.971. Better overall success of the model in separating between positive and negative cases is indicated by a sophisticated AUC.

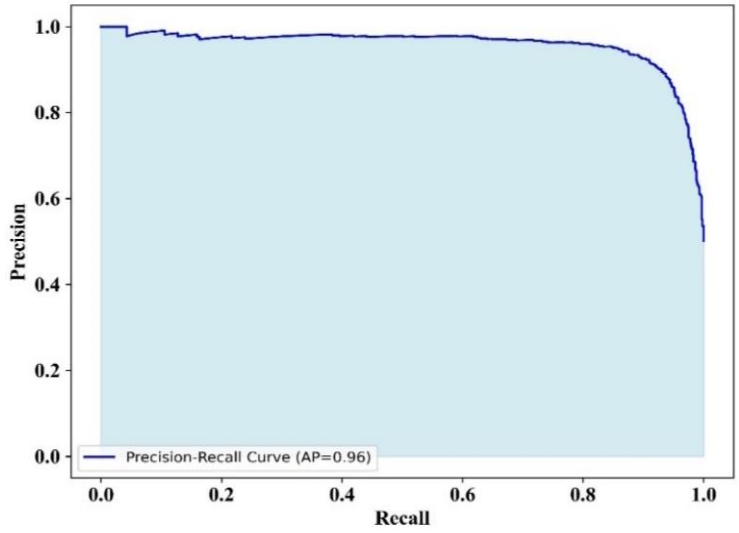

Fig. 12. Precision-Recall Curve for Data Set 1

Fig.12. represents a perfect precision-recall curve would be a traditional line in the top left corner of the graph (Recall = 1, Precision = 1), indicating a model that can perfectly separate amongst positive and negative cases. The text "1.0" on the top of the y-axis indicates a perfect precision of 1, and the text "0.0" on the left of the x-axis indicates a perfect recall of 1. Precision-Recall Curve for Data Set 1 AP=0.96.

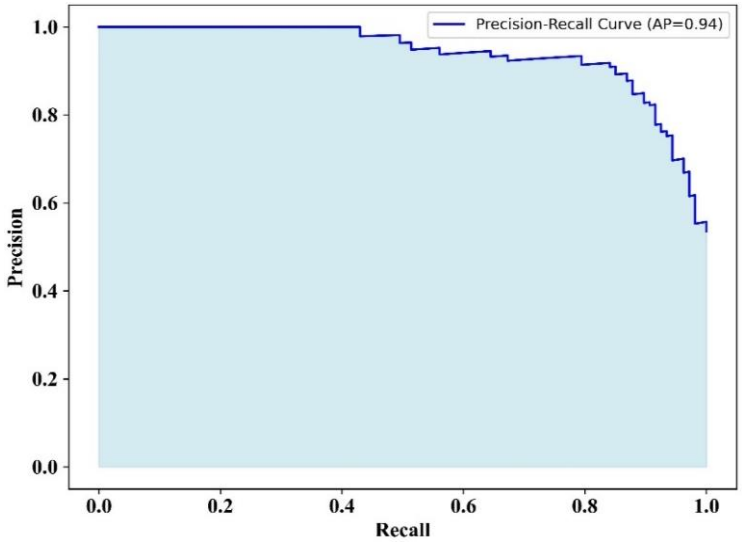

Fig. 13. Precision-Recall Curve for Data Set 2

Fig.13. signifies a faultless precision-recall curve would be a conventional line in the top left corner of the graph (Recall = 1, Precision = 1), indicating a model that can perfectly extricate amid positive and negative cases. The text "1.0" on the top of the y-axis indicates a perfect precision of 1, and the text "0.0" on the left of the x-axis indicates a perfect recall of 1. Precision-Recall Curve for Data Set 2 AP=0.94.

Fig.14. represents a matrix of confusion for Data Set 1. A table that shows you how well an algorithm performs is called a confusion matrix. This instance demonstrates how well a medical categorization model distinguishes between cases that are normal and those that are abnormal. This is how the confusion matrix is broken down. Rows: Show the actual tags (Normal or Abnormal) of the data points. Columns: Show the expected tags (Normal or Abnormal) for the data points. The numbers in the table indicate how many data points there are in each category: TP = 379 (True Positive). These are the data points that the typical appropriately branded as anomalous.TN = 312 (True Negative). These are the data points that the model accurately identified as typical. True Positive (FP): 12. These are the data points that the model mistakenly identified as abnormal when they were, in fact, normal. True Negative (FN): 16. These are the data points that the model misclassified as normal once, in fact, they were aberrant.

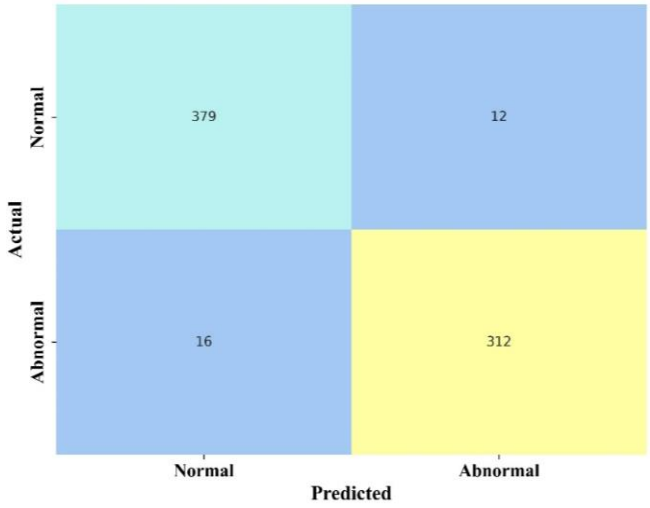

Fig. 14. Confusion Matrix for Data Set 1

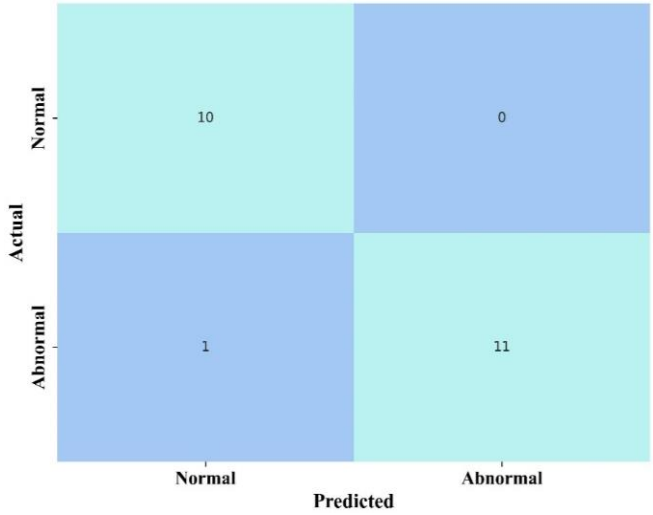

Fig. 15. Confusion Matrix for Data Set 2

Fig.15. a matrix of confusion for Data Set 2. A table that shows you how well an algorithm performs is called a confusion matrix. This instance demonstrates how well a medical categorization model distinguishes between cases that are normal and those that are abnormal. This is how the confusion matrix is broken down. Rows: Show the actual markers (Normal or Abnormal) of the data points. Columns: Show the expected tags (Normal or Abnormal) for the statistics points. Better performance is indicated by higher numbers (TP and TN) on the diagonal. The mainstream of the data points in a well-performing model will be accurately identified as either positive (TP) or negative (TN). Performance is poorer when values off the diagonal (FP and FN) are higher. These show how the model misclassifies certain cases as either positive or negative.

5. Conclusion

A new methodology in the diagnosis of neonatal asphyxia by using Linear Graph Neural Networks and African Vultures Algorithm for optimization of hyperparameters. This paper seeks to combine structured clinical data with medical imaging in a manner that can possibly overcome some of the inherent limitations of conventional diagnostic methods toward the identification of neonatal asphyxia. The model consequently seems to be efficient in such identification. Such interventions will result in improved clinical outcomes because such interventions are usually done earlier and more accurately in avoiding difficulties resulting from birth oxygen deprivation. The practice of the African Vultures Algorithm for hyperparameter optimization highlights finding optimal solutions in high-dimensional spaces, which is critical for attractive the routine of complex neural network architectures. The model's capacity to achievement the graph structure for feature aggregation and inference offers new insights into the submission of graph neural networks in healthcare, particularly for conditions that involve multifaceted data inputs. While the results look promising, further studies are required to ensure that the performance of the model will be robust over a more heterogeneous population and setting. Also, finding the entrenching of other deep learning techniques such as attention

mechanisms or transfer learning may even bring the model's precision and generalizability to an even higher level. Future work should proceed with a plan for real-time implementation and clinical trials in neonatal intensive care to test the utility of the model. The proposed LGNN-based model advances the state-of-the-art in neonatal asphyxia diagnosis and further offers a framework that may be adapted to other complex medical conditions. This can potentially enhance further understanding of pathophysiological processes, which, in turn, has great promise for personalized medicine and the better care of patients in the neonatology unit.

References

[1] Boskabadi H, Zakerihamidi M, Ghayour Mobarhan M, Bagheri F, Moradi A, Beiraghi Toosi M. Comparison of new biomarkers in the diagnosis of perinatal asphyxia. *Iran J Child Neurol.* 2023 Winter;17(1):99-110. doi: 10.22037/ijcn.v17i2.38561. Epub 2023 Jan 1. PMID: 36721830; PMCID: PMC9881829.1

[2] Keles E, Bagci U. The past, current, and future of neonatal intensive care units with artificial intelligence: a systematic review. *NPJ Digit Med.* 2023 Nov 27;6(1):220. doi: 10.1038/s41746-023-00941-5. PMID: 38012349; PMCID: PMC10682088.

[3] Chioma, R.; Sbordone, A.; Patti, M.L.; Perri, A.; Vento, G.; Nobile, S. Applications of Artificial Intelligence in Neonatology. *Appl. Sci.* 2023, 13, 3211. https://doi.org/10.3390/app13053211.

[4] Beck J, Loron G, Ancel PY, Alison M, Hertz Pannier L, Vo Van P, Debillon T, Bednarek N. An Updated Overview of MRI Injuries in Neonatal Encephalopathy: LyTONEPAL Cohort. *Children (Basel).* 2022 Apr 14;9(4):561. doi: 10.3390/children9040561. PMID: 35455605; PMCID: PMC9032533.

[5] Parmentier CEJ, de Vries LS, Groenendaal F. Magnetic Resonance Imaging in (Near-)Term Infants with Hypoxic-Ischemic *Encephalopathy. Diagnostics (Basel).* 2022 Mar 6;12(3):645. doi: 10.3390/diagnostics12030645. PMID: 35328199; PMCID: PMC8947468.

[6] Kieu, S.T.H.; Bade, A.; Hijazi, M.H.A.; Kolivand, H. A Survey of Deep Learning for Lung Disease Detection on Medical Images: State-of-the-Art, Taxonomy, Issues and Future Directions. J. *Imaging* 2020, 6, 131. https://doi.org/10.3390/jimaging6120131.

[7] Cao Shujuan, Wang Xiaoming, Ye Zhonglin, Li Mingyuan, Zhao Haixing, LGNN: a novel linear graph neural network algorithm, *Frontiers in Computational Neuroscience.*

[8] Subrato Bharati, Prajoy Podder, M. Rubaiyat Hossain Mondal, Hybrid deep learning for detecting lung diseases from X-ray images, *Informatics in Medicine Unlocked,* Volume 20,2020,100391,ISSN 2352-9148,https://doi.org/10.1016/j.imu.2020.100391.

[9] Sarnat, H.B.; Sarnat, M.S. Neonatal Encephalopathy Following Fetal Distress. A Clinical and Electroencephalographic Study. Arch. Neurol. 1976, 33, 696–705.

[10] Barkovich, A.; Miller, S.; Bartha, A.; Newton, N.; Hamrick, S.; Mukherjee, P.; Glenn, O.; Xu, D.; Partridge, J.; Ferriero, D.; et al. MR imaging, MR spectroscopy, and diffusion tensor imaging of sequential studies in neonates with encephalopathy. AJNR Am. J. *Neuroradiol.* 2006, 27, 533–547.

[11] Barkovich, A.J.; Baranski, K.; Vigneron, D.; Partridge, J.C.; Hallam, D.K.; Hajnal, B.L.; Ferriero, D.M. Proton MR spectroscopy for the evaluation of brain injury in asphyxiated, term neonates. *AJNR Am J Neuroradiol.* 1999, 20, 1399–1405.

[12] Penrice, J.; Lorek, A.; Cady, E.B.; Amess, P.N.; Wylezinska, M.; Cooper, C.E.; D'Souza, P.; Brown, G.C.; Kirkbride, V.; Edwards, A.D.; et al. Proton magnetic resonance spectroscopy of the brain during acute hypoxia-ischemia and delayed cerebral energy failure in the newborn piglet. *Pediatr. Res.* 1997, 41, 795–802.

[13] Zhang, Y.; Xu, Y.; Zhang, Y. A Graph Neural Network Node Classification Application Model with Enhanced Node Association. *Appl. Sci.* 2023, 13, 7150. https://doi.org/10.3390/app13127150.

[14] Li, Y., Wang, Y., Yang, X. et al. Speech emotion recognition based on Graph-LSTM neural network. J AUDIO SPEECH MUSIC PROC. 2023, 40 (2023). https://doi.org/10.1186/s13636-023-00303-9.

[15] M. Nelson, S. Rajendran, Y. Alotaibi, Vision graph neural network-based neonatal identification to avoid swapping and abduction, *AIMS Mathematics,* 8 (2023), 21554–21571. https://doi.org/10.3934/math.20231098

[16] T. Tamilvizhi, R. Surendran, K. Anbazhagan, K. Rajkumar, Quantum behaved particle swarm optimization-based deep transfer learning model for sugarcane leaf disease detection and classification, *Math. Probl. Eng.,* 2022 (2022), 3452413. https://doi.org/10.1155/2022/3452413

[17] https://www.kaggle.com/datasets/bhavkaur/synthetic-infant-health-dataset.

[18] https://www.mrineonatalbrain.com/

Authors' Profiles

Madhusundar Nelson is a dedicated researcher with a strong foundation in academia, industry, and revolutionary research. The past 2.5 years, I have been serving as a full-time Ph.D. Research Scholar at SIMATS, focusing on advanced fields such as Computer Vision, Deep Learning, and Graph Neural Networks.

Surendran. R. is working as a Professor in the Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India. His name is included in the 2024 World Top 2% Scientists List, published by Stanford University and Elsevier. He has 17 years of teaching experience including 5 years of teaching experience in Gulf Countries (Oman and Bahrain). He handled interesting roles and responsibilities including Teaching in various Institutions with more success factors like Research Faculty, Trainer of Trainer & Technical trainer for Placement, Coach for various program competition, Quality Improvement Committee Chair (ABET Accreditation), Course Equalization Committee chair and Head-Technology Development. He has published more than 175 articles in various international journals and conferences and 4 Patent published.

How to cite this paper: Madhusundar Nelson , Surendran Rajendran, "An Intelligent Deep Learning Model for Neonatal Asphyxia Identification Using Linear Graph Neural Networks Optimized with African Vultures Algorithm", International Journal of Engineering and Manufacturing (IJEM), Vol.15, No.3, pp. 1-17, 2025. DOI:10.5815/ijem.2025.03.01