Explainable Fetal Ultrasound Classification with CNN and MLP Models

Dodda Venkatareddy, K.V. Narasimha Reddy, Yendluri Sowmya, Yarlagadda Madhavi, Shaik Chand asmi, Sireesha Moturi,
Dept Of CSE, *Narasaraopeta Engineering College*, Narasaraopeta, Palanadu, India
doddavenkatareddy@gmail.com, narasimhareddynec03@gmail.com, yendlurisowmya5@gmail.com,
yarlagaddamadhavi96@gmail.com, shaikchandasmi123@gmail.com, sireeshamoturi@gmail.com

***Abstract*—Artificial Intelligence has greatly influenced healthcare, most particularly in medical imaging. This paper represents a review in large form that classifies fetal ultrasound images with the use of convolutional neural networks and multi-Layer Perceptrons. While CNN is very good at spatial feature extraction in image classification, their lack of interpretability presents challenges toward applications in health. In this regard, we include methods of Explainable AI (XAI), more precisely Local Interpretable Model-Agnostic Explanations (LIME), for giving more transparency and confidence in the decision-making process of such models. The research here utilizes 12,400 fetal ultrasound images, which were classified under six anatomical structures. The CNN and MLP models showed very promising classification performances of 93.24% and 91.17%, respectively. LIME was implemented to interpret model predictions and to more clearly identify factors contributing to the classification. The results also show that explainability enhances not only trust in AI-based diagnostics but also model reliability in clinical settings.**

***Keywords*—Fetal ultrasound classification, Convolutional Neural Networks, Multi-Layer Perceptron, Explainable AI, Local Interpretable Model-Agnostic Explanations.**

I. Introduction

Over the years, artificial intelligence (AI) has taken a significant role in changing the healthcare scene, especially in medical imaging. Deep learning has been one of the many AI methods but has become the one that healthcare workers can use to make diagnoses and treatment decisions that are more accurate. The main thing that makes CNNs (convolutional neural networks) [2] so special is that they can learn from the images themselves to find the relevant features, thereby making them suitable for image classification tasks and object detection.

One popular method of prenatal care is fetal ultrasound imaging. Through others, it can provide the medical practitioners with as much as they need to know about the state of fetal development and ambler. The aforementioned pictures enable the attendants to evaluate the anatomy of the fetus, to supervise the growth, and to detect probable defects.

Relevant Pioneer of the fetal ultrasound image plane is the direct and most impactful pillar responsible for either a good or a bad health outcome of the mother or the fetus, so it is indeed a crucial point. Nevertheless, the complexity of ultrasound pictures and the need for a precise interpretation are the difficulties. [1]

While CNN is great at getting high classification accuracy, the decision-making processes can be intransparent, causing difficulties for health care takers to comprehend and have trust in the model's predictions. This lack of interpretability, along with the broad judgment AI tools make in a clinical setting, becomes a real concern in the domain where transparency and accountability are essential. systems and gain their trust by being clear to them. Methods such as LIME [6] are making it possible for scientists to observe the roles of various features in the given data and the model's predictions, thus pointing out the path between the little downfall and the readability. This project has the support of both CNNs and Multi-Layer Perceptrons (MLPs) for classifying the fetal ultrasound image planes. To the point that the classification is exact while using the spatial hierarchies and patterns, they describe the features of the method with spatial description, the following appropriate and recognized by the computer: a big picture being reused. But at the same time, one can argue that the growing number of e courses, especially ones hands-on with technology, could soon make the traditional education system obsolete.

”These newly created and innovatively designed online engineering courses that are hands-on and use the latest technology that is generally not available in traditional face toface classes need to be called the Future of Engineering Education.” By both the CNNs and the MLPs architectures, we are the best model to solve the problem, which is both we are able to achieve high classification accuracy. The chief goal of the investigation at hand is the production of a reliable and interpretable classification model for the ultrasound images of the fetus that will use both the CNN model and the MLP models. Intrinsically, the algorithm's capability has been promoted. The use of annotated fetal ultrasound images provides the AI assisted prenatal diagnostics process with more

reliability and, with that, will alleviate the all too common concern, hence addressing the need for transparency in AI healthcare.
Explainable AI techniques (XAI) have been devised as a means of making deep learning models more transparent by providing insights for the model's decisions. XAI is supposed to allow people to understand the inner workings of AI

II. LITERATURE REVIEW:

Befre Action that is most transformative for the deployment of pregnancy ultrasound category is the combination of convolutional networks and multilayer perceptrons with XAI tools and methods. Toward this end, we present a comprehensive list of relevant works and studies involving pretrained CNNs and MLPs in different AI applications to ultrasound images, as well as the general exploitation of XAI approaches in the deep learning context. Also, the possible pagination of the abstract with the keyword list is indicated in the main document. For example, Burgos-Artizzue et al. [7] mentioned automated maternal-fetal ultrasound abnormality classes that were generated using dense tissue segmentation. Their work highlighted the webcam, a CNN with an overall AUC of 1.0%, which significantly improved the performances in detecting different intrapelvic organs. Esteva et al. [3] in every alternative test copied the features of deep learning models in dermatology, reaching dermatologist-level classification of skin cancer. This work was the key point for the utilization of analogous techniques in different clinical imaging domains, such as fetal ultrasound analysis. The accomplishment of CNNs in identifying the diverse features of the picture in very fine detail has made them very popular in almost all of the medical imaging tasks. The applications include fetal biometry estimation as well as abnormality detection. CNNs have, in fact, become the most often used form of structure in the image-type tasks, but MLPs have still been recurrently utilized in clinical imaging applications. MLPs are distinguished for being both able to process and represent structured data in a proper manner and, therefore, can be an alternative approach to CNNs. For instance, MLPs have been used to analyze functions derived from CNNs, which leads to a more complete understanding of the information. Experiments have confirmed that the communication of MLPs with CNNs can be beneficial to the correct classification of data by both methods. Medical image interpretability and incorporation of Explainable AI (XAI) concepts into deep learning models are very important areas. LIME (local Interpretable model-agnostic explanations) can be referred to as one of the most recognized XAI approaches that have been used in many fields, including healthcare. The authors, Garreau and Luxburg [8], proposed a theoretical discussion on LIME in the medical field, the article of which paid attention to how the methods of LIME could help this department. Now and onwards, the same LIME article became the integral part of the discussion. Generally, XAI techniques are relatively new to fetal ultrasound imaging, but more and more people are gaining interest in this field. By using LIME researchers can pin down in an image where the most important points causing the model to decide like this are. Moreover, it is a must for the model to have the ability to provide logical decisions, as the latter tends to breed trust with healthcare professionals, who get the chance to acknowledge the basis of the prediction and hence make correct evaluations concerning the product of AI.

.

III. METHODOLOGY:

A. Dataset:

The dataset used to classify fetal ultrasound images consists of 12,400 images that are categorized into six main groups: AFD Fetal Abdomen, Fetal Femur, Fetal Thorax, Fetal Brain, Maternal Cervix, and Other. (https://zenodo.org/records/3904280). The csv dataset has basically 7 columns, i.e. patient num, Image name, Plane, Brain plane, Operator, US Machine and Train. and the images dataset contains the image name. All the images are a mix of RGB and Grayscale formats, and they have different resolutions and sizes. This complete dataset is not only helpful in classifying planes in the fetus but also it will function as a raw material for bringing out new and also confirming deep learning models in the context of prenatal healthcare.

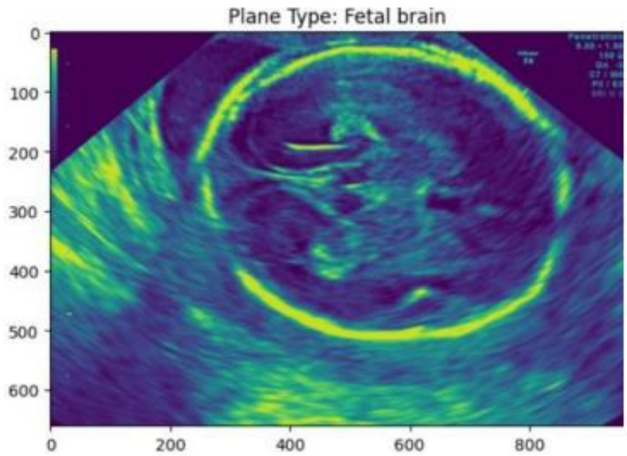

Fig1: Sample image

B. Data Preprocessing:

1) Resizing: All the pictures were changed to be the same measurement (224x224 pixels), which means that they are of the same size, and it is a must for the model's learning.

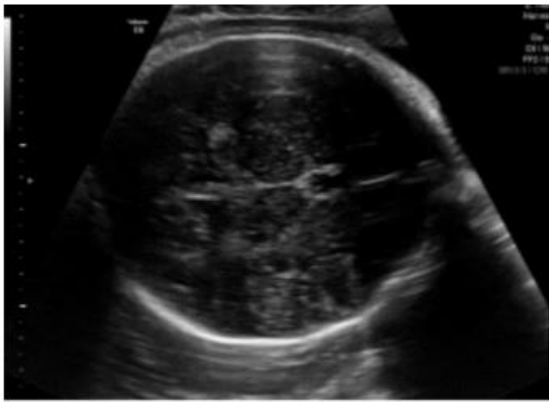

Fig2: Resized image

2.Normalization: Pixel values were reduced to a standard level of [0, 1]. This is a way of ensuring that the process is absolutely smooth and quick throughout the training of the model.

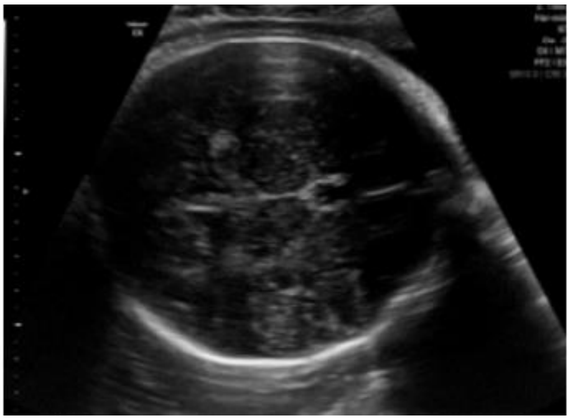

Fig3: Normalized Image

3) Feature Extraction: It is the code that makes use of the ResNet-50 model from torch vision. ResNet50 is a deep residual network honored for its strong performance and reasonable level of reliability in terms of image classification tasks. Functions such as the below mentioned code were performed for each image, passing the image into the ResNet50 model in order to extract its features. The model outputs the feature vector, which is then "printed" on the occasion of saving for further processing.

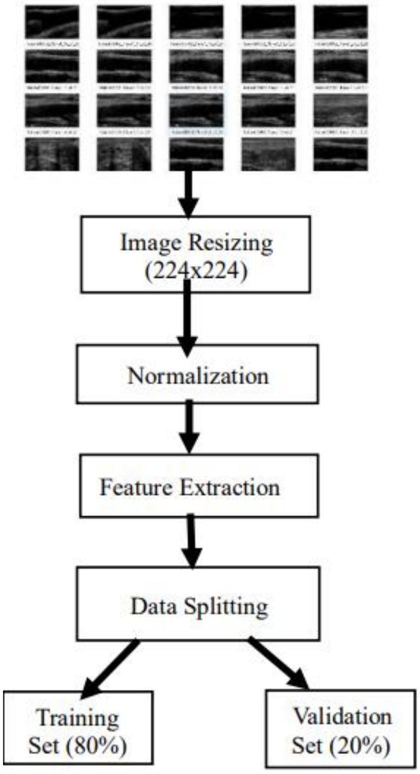

Fig 4: Data Preprocessing

C.Model Training:

Also, it is a design milestone to build a valid neural network that mixes CNN (Convolutional Neural Network) and MLP (Multilayer Perceptron)[7,8,9] models and manages the initial dataset. Data preparation is actually a process in which required procedures like normalization and transformation to tensors are implemented to ensure that the data entered is in a format model learning can use. Following the data loading process, we configure the two types of computational models: the first one for image data using CNN that obtains spatial features through the convolutional layers, and the second one, the so-called fully connected network that employs MLP for simpler input structures. [10,11] The first and the voice are the ones that are then being created, which is the initialization step. This approach keeps going unless the model achieves satisfactory production of the validation dataset.

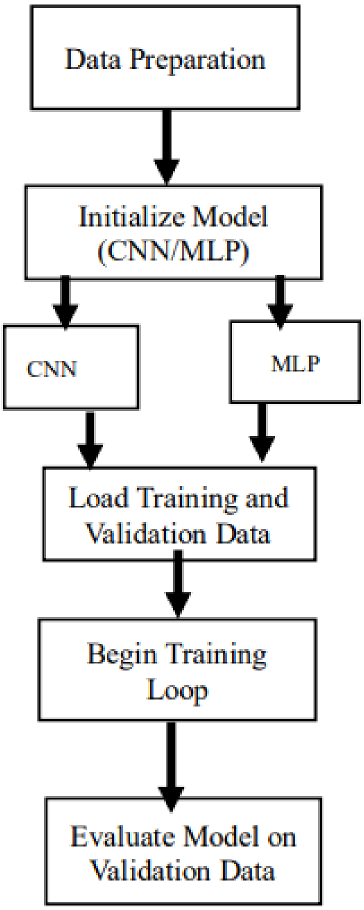

Fig 5:Model Training

1) Simple CNN Architecture:

• Convolutional Layers: conv1: Gets the initial input, which is given to the basic 64 channels, on applying a 3x3 filter to it with a 1 zero-pad use. [13] conv2: It acquires the 64 channels from the previous layer and outputs a 128-channel layer by using a filter of the same size as the other and a zero-pad.

• Pooling Layer: pool: performing max pooling with a 2x2kernel, which reduces the feature map's length and width. • Fully Connected Layers: fc1: This is a linear layer thattakes the flattened output from the sets of features of the convolutional layers and connects it to 512 nodes. fc2: It is the last linear layer that maps 512 nodes to the number of classes (num-classes). [12] • Activation and Dropout: ReLU activations are placed inthe convolutional and fully connected layers. There lies after the first fully connected ReLU activations and dropouts for regularization to prevent overfitting there.

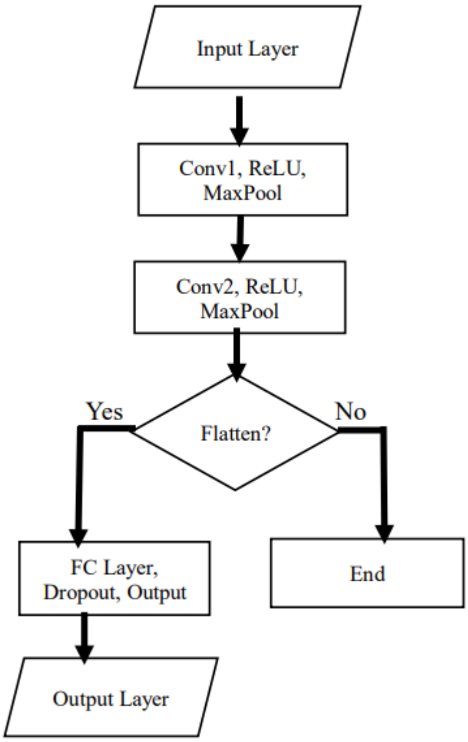

Fig 6: CNN model architecture

2) MLP Model Architecture And Explainability::

• Fully Connected Layers: fc1: The initial layer is as described. It's the very first linear layer that, when given the input features, maps the input features to 512 units. Fc 2: The second layer is further described. The number of units is decreased by the second linear layer, which is another feature of the MLP architectures. fc3: The very last layer is characterized by the following trait: The representation on the output layer that forms into the assigned categories (num-classes) is the last linear layer. • Activation and Dropout: Fully connected layers are separated by ReLu activations. ReLu activation layers are probably the best ones, and in between activation layers, they also do the job of breaking the symmetry too. Airplane.jpg are the users of the case of image classification, and my spatial data is the interpretation of it in 3-D with ReLu.

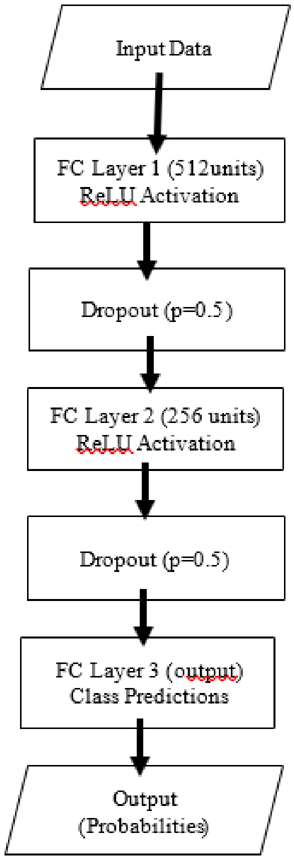

Fig 7: MLP model architecture

For this research, the Local Interpretable Modelagnostic Explanations (LIME) was employed to interpret forecasts of a highly advanced MultiLevel Perceptron (MLP) model executed using the PyTorch. The MLP model, the components of which are several fully connected layers with batch normalization and dropout, is one of the classification tasks. Generating modelunderstandable individual predictions with a simple and interpretable model is the primary idea of LIME. We can think of this procedure as building a simple model of a complex model to mimic the behavior of the complex model near a particular instance. The model is approximated by varying the input data along the features and monitoring the changes in the model's predictions. The explanation of the model's predictions is eventually based on the explanation of the specific feature. The diverse features and the relationship between them were noticed by the LIME model while generating the different versions of the input. Deploying LIME in this way is the mechanism by which the behavior of the model of the data can be examined in more detail from pred: production of choice. With the technique of LIME implementation, we were able to give indirect support to the whole process of generating interpretations through LIME. [14], [16] The visual presentations showed that the individual factors were of significant input to the classification, and the effect of the parameter change would be further encouraged; thereby, the reliability of the classification model was improved. This methodology not only helps to create a detailed central source model, but also facilitates crosschecking predictions and the existing domain knowledge. An example of this would be the MLP. When predicting an image as ”fetal femur” with 32% probability, LIME pointed out the top positive features, which were feature-73 gt 0.80 and feature14 gt 0.72, as the main sources of the variance, thus giving a lucid elucidation of the model’s prediction. The introduction of dizziness learning and interpretability supplied us with some salient information about the model’s performance as well as their decision-making transparency, a crucial foreshadower of clinical applications.

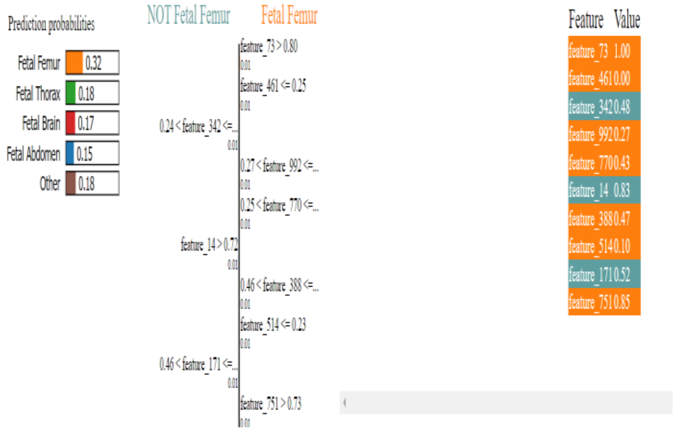

Fig 8:Lime Prediction

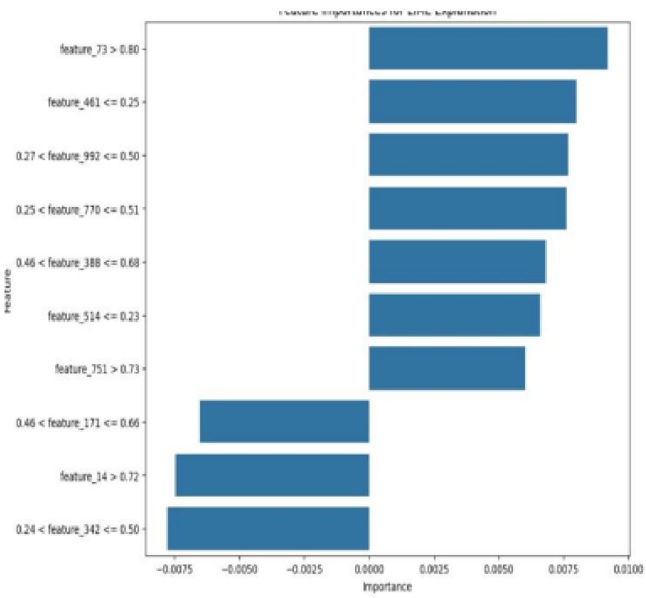

Fig 9: Lime Technique

IV. COMPARITIVE ANALYSIS AND DISCUSSION:

Comparative analysis of CNN and MLP clearly states the strengths and limitations of each method. The CNN’s unbeatable means of working are in part thanks to the privilege of its construction. It can extract and process the spatial and hierarchical relationships in the fetal ultrasound images. The increased precision that was reached by CNN proves its effectiveness in this classification mission and thus makes it favorable for applied domains of medicine. [15] MLP is not prominently known for a baseline model, but it was toppled by CNN, which is the promise of its practical asset of finding the best models for specific forms of data. A high rating of the MLP is a visual mean to show that sometimes the tasks with the more complex models can be the bigger ones. The required model has enough to be a classifier by CNN; thus, this model fits for this purpose better than others. Implications for Medical Imaging New research findings that have taken place this study will have a huge impact on a variety of uses of artificial intelligence in the medical imaging field. According to the result, various abilities like high accuracy, as well as the capability to be independent across all samples, are restrained in the CNN model, choosing to like a ray (”/” from the tools”). It will be easy for radiologists and medical professionals to have a diagnosis and analysis of fetal ultrasound images with the help of this particular technological tool.

Thus, the classification in the process of diagnosis can be automated by the use of an accurate model that could help in verifying the diagnosis of a patient and also quickly proceed to more effective treatment by the doctors. Also, this is where medical conditions can become difficult to understand since patients will have trouble figuring out how a machine knows better than them. For example, using ’Explanation AI (AX) methods, which are the following. [9], [14] ”CNN will help to unhide the models’ and make them see through, thus providing an opportunity to perceive everything that the model has gone through to reach a decision. The medical field is specifically the one that Explainable AI (XAI) cannot work without as the understanding of the explanation of the prediction by models is the key to doctors’ acceptance and trust. This is not only the support in the case of doctors but for patients also.

Perceptron (MLP) to determine the anatomical structures as well as classifying the fetal ultrasound images into different classes. The basic aim was to test the statistical reliability and efficacy of these models in the classification of images based on their anatomical structures and regions. Actually, the main reason the CNN model was chosen was that it is a model that can process and analyze images brilliantly. CNNs have such a way of design that convolutional layers can be used to get spatial hierarchies in images, which makes it really effective in image classification. The CNN model researchers used in the study was the one they had preprocessed.

V. RESULT

This experiment was performed by the application of a Convolutional Neural Network (CNN) and a Multi-Layer Perceptron (MLP) to determine the anatomical structures as well as classifying the fetal ultrasound images into different classes. The basic aim was to test the statistical reliability and efficacy of these models in the classification of images based on their anatomical structures and regions. Actually, the main reason the CNN model was chosen was that it is a model that can process and analyze images brilliantly. CNNs have such a way of design that convolutional layers can be used to get spatial hierarchies in images, which makes it really effective in image classification.

The CNN model researchers used in the study was the one they had preprocessed.

Table 1: CLASSIFICATION REPORT

Class	Precision	Recall	F1 score	Support
Fetal Abdomen	0.74	0.86	0.79	142
Fetal femur	0.98	0.98	0.98	610
Fetal thorax	0.72	0.9	0.8	213
Fetal Brain	0.85	0.94	0.9	358
Maternal Cervix	0.99	1.0	1.0	320
Other	0.95	0.82	0.88	837
Accuracy			0.91	2480
Macro Avg	0.87	0.92	0.81	2480
Weighted Avg	0.92	0.91	0.91	2480

on with neglected steps like normalization and augmentation, among other carefully implemented preprocessing steps, so that the model's learning capability and its generalizability could be enhanced. The model trained using CNN underwent a hard training and validation process. To be specific, the accuracy of the model at the training stage was 93.24% and at the validation stage was 91.17%. This means that the CNN model was capable of acquiring ultrasound images' features with high effectiveness and utilizing the same capacity to sort out new images that have not been seen before with a high level of accuracy. The fact that the differences between training and validation accuracy are so small implies that the model is finetuned and thus doesn't tend to overfit, which is an often seen phenomenon in deep learning models. Instead, the MLP model, a distributed type of neural network, was also brought to serve as a gauge for a better comparison with a CNN. The convolving of the filter's weights with the input matrix produces a new feature map. Nonetheless, the MLP model, although selecting some of the possible responses, did not win against the CNN in the aspect of accuracy and overall performance. Moreover, the application of these layers means MLP could not fully match the intricacy of the set. MLPs, on the other hand, consider the image as a one dimensional line of pixels, which limits their capability to capture all of the spatial information the image inherently possesses.

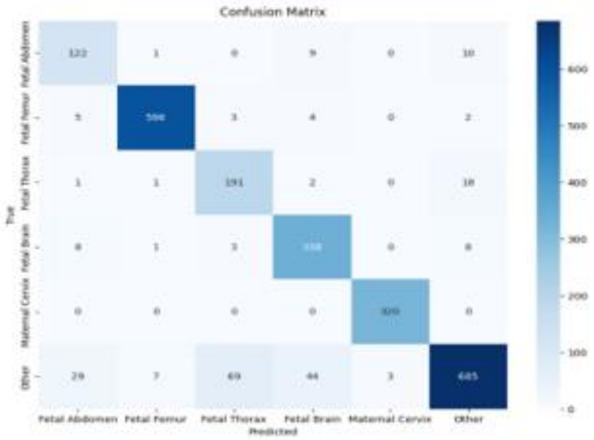

Fig 10:Confusion Matrix

VI. CONCLUSION :

Our study saw us carrying out the two models that we had designed that is the aftertaste of technology (CNN) and the Multilayer Perceptron (MLP) model aside for the classification of the fetus's ultra-sound images. The CNN model, with its capability to seize spatial hierarchies in images, was used to automatically morph the features and distinguish the images by the six superordinate classes which were the maternal cervix, fetal brain, fetal femur, fetal abdomen, and other. Additionally, we also utilized a MLP model for the process that the CNN model was applying but instead, it was doing it using fully connected layers. The MLP, on the other hand, was used to make the former approach by the feature vectors instead of image data, giving us an idea of the performance of the architecture on the task. Leveraging the LIME methodology, we obtained some explanations for the decisions that each model made which helped in achieving the interpretation so that humans can easily make sense of the models. Concerning the MLP model, LIME was able to diagnose the most critical features that altered its decision-making procedure.

REFERENCES

[1] Li, W. Cai, X. Wang, Y. Zhou, D. D. Feng and M. Chen, "Medical image classification with convolutional neural network," 2014 13th International Conference on Control Automation Robotics Vision (ICARCV), Singapore, 2014, pp. 844-848.IEEE

[2] Lin, T. Y., RoyChowdhury, A., Maji, S. (2015). "Bilinear CNN models for fine-grained visual recognition". In Proceedings of the IEEE international conference on computer vision (pp. 1449-1457)..

[3] esteva, a., kuprel, b., novoa, r. et al."dermatologist- level classification of skin cancer with deep neuralnetworks".nature542,(2017): 115–118.

[4] Driss, S. B., Soua, M., Kachouri, R., Akil, M. (2017, May). A comparison study between MLP and convolutional neural network models for character recognition. In Real-Time Image and Video Processing 2017 (Vol. 10223, pp. 32-42). SPIE.

[5] Chauhan, R., Ghanshala, K. K., Joshi, R. C. (2018, December). Convolutional neural network (CNN) for image detection and recognition.

In 2018 first international conference on secure cyber computing and communication (ICSCCC) (pp. 278-282). IEEE.

[6] N. Barr Kumarakulasinghe, T. Blomberg, J. Liu, A. Saraiva Leao and P. Papapetrou, "Evaluating Local Interpretable Model-Agnostic Explanations on Clinical Machine Learning Classification Models," 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS), Rochester, MN, USA, 2020, pp. 7-12.

[7] Burgos-Artizzu, X.P., Coronado-Gutierrez, D., Valenzuela-Alcaraz, B.´ et al. "Evaluation of deep convolutional neural networks for automatic classification of common maternal fetal ultrasound planes". Sci Rep 10, 10200 (2020).

[8] Garreau, D. amGarreau, D., & Luxburg, U. (2020, June). Explaining the explainer: A first theoretical analysis of LIME. In International conference on artificial intelligence and statistics (pp. 1287-1296). PMLR.

[9] Moturi, S., Srikanth Vemuru, D. S. (2020). Classification model for prediction of heart disease using correlation coefficient technique. International Journal, 9(2).

[10] Desai, M., Shah, M. (2021). An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and Convolutional neural network (CNN). Clinical eHealth, 4, 1-11.

[11] El Alani, O., Abraim, M., Ghennioui, H., Ghennioui, A., Ikenbi, I., Dahr, F. E. (2021). Short term solar irradiance forecasting using sky images based on a hybrid CNN–MLP model. Energy Reports, 7, 888- 900.

[12] Jagannadham, S. L., Nadh, K. L., Sireesha, M. (2021, November). Brain tumour detection using cnn. In 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC) (pp. 734-739). IEEE.

[13] Iswarya, B., Manimekalai, K. (2022). Drug discovery with XAI using deep learning. In Principles and Methods of Explainable Artificial Intelligence in Healthcare (pp. 131-149). IGI Global.

[14] Moturi, S., Vemuru, S., Tirumala Rao, S. N. (2022). Two phase parallel framework for weighted coalesce rule mining: a fast heart disease and breast cancer prediction paradigm. Biomedical Engineering: Applications, Basis and Communications, 34(03), 2250010.

[15] Van der Velden, B. H., Kuijf, H. J., Gilhuijs, K. G., Viergever, M. A. (2022). Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. Medical Image Analysis, 79, 102470.

[16] Vimbi, V., Shaffi, N., Mahmud, M. (2024). Interpreting artificial intelligence models: a systematic review on the application of LIME and SHAP in Alzheimer's disease detection. Brain Informatics, 11(1), 10.