Multimodal Health Diagnostic Tool: Predictive Analytics and Medical Imaging for Women's Healthcare Support

D.S.Sathsarani
Department of Computer and Data Science
Faculty of Computing, NSBM Green University
Homagama,Sri Lanka
sewminisathsarani1@gmail.com

***Abstract*—Women's health conditions such as endometrial cancer, PCOS, thyroid disorders, and breast cancer often remain undiagnosed in early stages due to limited access to diagnostic resources, especially in low-income or rural communities. Traditional diagnostic pathways are often time-consuming, fragmented, and require expert intervention at every stage. This research addresses the problem by developing a web-based diagnostic platform that uses machine learning and image analysis to support early, accessible, and multi-condition screening for women's health. The system includes a predictive model that analyzes 19 clinical and lifestyle features such as age range, menstrual cycle length, pain history, and chronic conditions to classify ten different diagnoses: Endometrial Cancer, Thyroid Disorders, Endometriosis, Pelvic Inflammatory Disease (PID), Uterine Fibroids, Adenomyosis, Cervical Dysplasia, Ovarian Cysts, PCOS, and Breast Cancer. The model is built using XGBoost with SMOTE to address class imbalance and GridSearchCV for hyperparameter tuning. It achieved an overall accuracy of 92.95%. In addition, deep learning models are integrated for medical image analysis. Users can upload scans such as womb ultrasound, breast images, or Pap smear slides to detect conditions like PCOS, uterine fibroids, breast cancer, or cervical abnormalities. Image datasets were annotated with Roboflow, trained in Google Colab, and deployed via Google Cloud Platform. This platform provides an accessible, AI-assisted diagnostic tool through a user-friendly web interface built with React. It supports early intervention and bridges gaps in healthcare delivery, especially in under-resourced settings.**

Keywords— Medical diagnosis, medical imaging, predictive modeling

I. INTRODUCTION

Women's health conditions such as Polycystic Ovary Syndrome (PCOS), endometrial cancer, breast cancer, and uterine fibroids often go undiagnosed or are detected late due to barriers in access to diagnostic care, especially in developing countries. Many of these conditions present with overlapping symptoms such as menstrual irregularities, fatigue, pelvic pain, or hormonal imbalances, making early diagnosis difficult without clinical tests and imaging [1]. As a result, countless women face delayed treatment, worsened outcomes, or lifelong complications due to late detection.

Artificial intelligence (AI) and machine learning (ML) have shown significant promise in healthcare for automating diagnosis, analyzing patterns in patient data, and assisting in medical image interpretation [2]. However, most existing tools focus on single-condition detection and are not integrated into accessible, web-based platforms suitable for public use or low-resource clinical settings. Moreover, AI systems that combine both tabular clinical data and medical image analysis for multi-condition diagnosis especially in women's health are limited.

To address this gap, this study proposes a comprehensive AI-powered web application designed specifically for early diagnosis of women's health issues. The platform integrates a machine learning model trained on structured health data to classify ten common gynecological and endocrine disorders and uses deep learning techniques to analyze medical images like womb scans, breast images, and Pap smear slides. By combining these two data modalities in a single platform, the system supports early intervention, improves patient awareness, and reduces dependency on specialist resources.

II. LITERATURE REVIEW

A. Machine Learning on Tabular Data for Diagnosis

Machine learning (ML) has been widely used for early disease detection using structured health data, such as patient symptoms, lab results, and clinical history. Algorithms like XGBoost, random forests, and support vector machines (SVMs) have shown high performance in predicting diseases such as diabetes, heart disease, and breast cancer.

For example,[3] applied XGBoost to a breast cancer dataset and achieved over 98% accuracy, outperforming logistic regression and decision trees. Another study by [4] used SMOTE with Random Forest to address class imbalance in cervical cancer screening data, improving recall for the minority class by over 20%.

Tabular ML models are especially useful in resource-limited settings, where imaging may not be available. Techniques like SHAP and LIME have been used to explain predictions, making these models more transparent and suitable for clinical use [5].

B. Deep Learning on Medical Images

Convolutional neural networks (CNNs) have become the standard approach for analyzing medical images, from X-rays to ultrasound and cytology slides. CNNs can automatically learn and extract complex features, enabling accurate classification, detection, and segmentation of medical abnormalities.

In uterine fibroid detection from ultrasound,[6] reported 99.8% accuracy using a dual-path CNN, while EfficientNet-based models achieved comparable results with fewer parameters. For cervical cancer, the DeepPap CNN model, trained on Pap smear images, achieved 98.3% accuracy and 0.99 AUC without requiring preprocessing like cell segmentation [7].

Object detection models like YOLOv8 have also been used to localize lesions in real time, with researchers reporting >95% mAP on small datasets using transfer learning and data augmentation. Transformer-based models such as Swin Transformers and CerviFormer have recently shown promise in classifying cervical cells, improving generalization on small or imbalanced datasets [8].

C. Web-Based Healthcare Tools

While AI models perform well in research, clinical impact depends heavily on usability. Web-based healthcare tools that integrate ML or deep learning models offer a practical way to bring diagnostics to clinicians and patients. Most tools rely on frameworks like Flask, React, or Streamlit for front-end access, and models are often deployed through TensorFlow.js, ONNX, or TorchServe on cloud platforms like Google Cloud Platform (GCP) or AWS.

Tools such as SkinVision [9] and LungXpert [10] provide real-time prediction and image analysis directly via web or mobile interfaces. However, few systems integrate both tabular ML and image-based deep learning models into a single diagnostic workflow.

Moreover, privacy-preserving methods like federated learning and on-device inference are gaining interest for sensitive health applications, but real-world adoption remains limited due to complexity and infrastructure needs.

III. METHODOLOGY

This project combines machine learning on structured health data and deep learning on medical images to support diagnosis of various women's health conditions. These models are integrated into a web-based diagnostic platform aimed at assisting both patients and clinicians.

A. Dataset Collection

Clinical datasets were collected from publicly available health repositories and curated Excel sheets. Each record includes patient demographics, menstrual and reproductive history, symptom profiles, and final diagnosis.

B. Target Conditions

The model was trained to predict one of the following 10 conditions based on clinical input:

- Endometrial Cancer
- Thyroid Disorders
- Endometriosis
- Pelvic Inflammatory Disease (PID)
- Uterine Fibroids
- Adenomyosis
- Cervical Dysplasia
- Ovarian Cysts
- Polycystic Ovary Syndrome (PCOS)
- Breast Cancer

3) Input Features

Nineteen input features were used, all based on patient-provided information:

'Age Range', 'BMI', 'Menstrual Cycle Length', 'Menstrual Irregularities', 'Pain History', 'Heavy Bleeding', 'Infertility', 'Pelvic Pressure', 'Urinary Symptoms', 'Fatigue', 'Digestive Problems', 'Breast Tenderness', 'Weight Gain', 'Family History', 'Acne', 'Hair Loss', 'Mood Swings', 'Stress Level', 'Chronic Conditions'

1) Preprocessing

Label Encoding was applied to all categorical columns using LabelEncoder. Missing Values were filled using the mean for numerical and mode for categorical columns. Feature Scaling was not required due to XGBoost's tree-based nature.

E. Handling Class Imbalance

Some diagnoses were underrepresented in the dataset. To address this, SMOTE (Synthetic Minority Over-sampling Technique) was applied to the training data to generate synthetic examples for minority classes.

F. Model Training and Tuning

The model used was XGBoostClassifier from xgboost with GridSearchCV for hyperparameter optimization. Grid search was performed over parameters such as learning_rate, max_depth, n_estimators, subsample, and colsample_bytree. The dataset was split into 80% training and 20% testing using train_test_split.The final model was saved as a .pkl file using pickle.

G. Evaluation Metrics

The model was evaluated using accuracy on the test set. A prediction function was implemented to classify new user input and return the most likely diagnosis.

H. Deep Learning for Medical Image Analysis

Annotated medical images for: Uterine fibroids (ultrasound), Breast cancer (mammogram), Cervical cancer (Pap smear images), PCOS(ultrasound) were labeled using Roboflow and exported in YOLO format.

I. Model Architecture

YOLOv8**s**, a lightweight yet powerful object detection model used, for all image-based tasks.

Training Configuration:

- Input size: 640×640
- Epochs: 100
- Batch size: 16
- Optimizer: AdamW (auto-selected by Ultralytics engine)
- Data augmentations: RandAugment, CLAHE, grayscale conversion, horizontal flip

Model performance was measured using: Precision, Recall, mAP@0.5 and mAP@0.5:0.95 and IoU scores. Each model was validated on a hold-out validation set. The best-performing weights were saved using the best.pt checkpoint.

J. Web integration and development

A React.js frontend was developed to allow users (clinicians or patients) to: Enter data and receive disease predictions and Upload medical images and receive bounding box detections and labels. Inference APIs were built using Flask. Tabular models were exported as .pkl files using joblib. Image models (YOLOv8) were exported as TorchScript .pt files. Backend APIs and models are hosted on Google Cloud Platform (GCP) using Compute Engine. Frontend is deployed via Firebase Hosting. Input data is temporarily processed without long-term storage. HTTPS is used for secure data transmission. No patient-identifiable data is collected or stored

IV. Results and Discussion

This section presents the performance of both the tabular diagnosis model and the image detection models. It also evaluates their practical application within the deployed web interface.

A. Predictive Model for Clinical Diagnosis (XGBoost)

The XGBoost model was trained to classify **10 women's health conditions** based on patient-reported symptoms and

```
Model Accuracy: 92.95%
Predicted Diagnosis: Endometrial Cancer
Model and Label Encoders have been saved!
```

history.The dataset contained **1,982 patient records** in Excel format. After preprocessing and balancing class using SMOTE, the data was divided into training and test sets.

1) Model Performance:

- Test **Accuracy:** 92.95%
- Cross-**validation folds:** 5 (GridSearchCV)
- Classes**:** 10 disease categories
- Input **Features:** 19 (encoded using LabelEncoder)
- Dataset **Size:** 1,982

Fig.2. Test Accuracy

The high accuracy suggests the model has strong potential for real-world use. Feature encoding allowed the model to handle structured, categorical input efficiently, while SMOTE addressed imbalance in classes such as cervical dysplasia and endometriosis.

Prediction example for a sample user profile shows in Fig.1.

```
user_input = [
    '51+', '18.5 to 24.9', 'Greater than 35 days',
    'Painful Periods (Dysmenorrhea)',
    '0', 'yes', 'no', 'yes', 'no', 'Yes', 'no',
    'yes', 'Yes', 'Uterine Prolapse', 'No',
    'No', 'No', 'Medium', 'Autoimmune disorders'
]
```

Fig. 1. Example user input

This suggests that the system can serve as a useful triage or decision support tool for initial screenings or referrals. The model achieves strong diagonal dominance, meaning most samples are correctly classified for their true condition. Minor off-diagonal values (e.g., 2, 4, 10) indicate a few misclassifications among certain classes. This confirms that the model performs consistently across most of the 10 disease categories. The visualization enhances model interpretability, showing not only overall accuracy (≈93%) but also *which* classes are more or less challenging for the model.

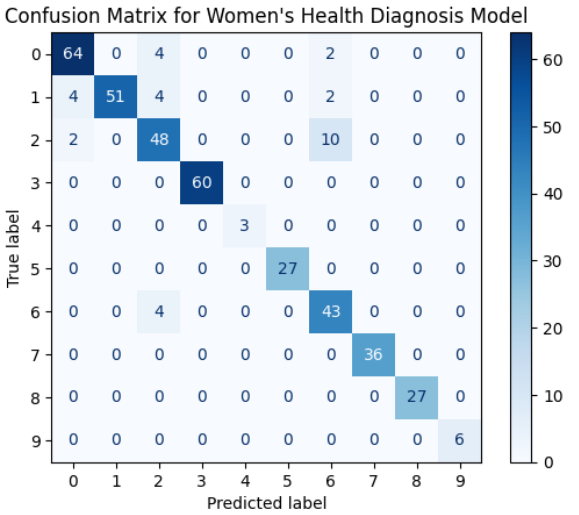

Fig.3. Confusion Matrix for women's health diagnosis model

XGBoost is robust to noisy features and handles multiclass classification well. GridSearchCV ensured optimal hyperparameters. Label encoding avoided the dimensionality explosion from one-hot encoding. While overall accuracy is high, per-class metrics like precision and recall (not computed here) would give better insight into rare condition prediction. SMOTE helps, but synthetic examples might not fully capture medical variability. The model is not calibrated to return probability scores (e.g., risk percent), which may be helpful in clinical decision-making.

V. MEDICAL IMAGE DETECTION RESULTS (YOLOV8)

2) Each deep learning model was trained on a different dataset for binary or multiclass detection of abnormalities in specific organ systems. Validation was performed using hold-out sets.

A. Cervical Cancer (Pap Smear Images)

Validation Metrics:

- mAP@0.5: 0.653
- mAP@0.5:0.95: 0.454
- Precision: 0.608
- Recall: 0.635
- Images: 100
- Instances: 261

These results indicate moderate detection accuracy. Given the complexity and variation in cytological features, the model is useful for assisting pathologists but should be further fine-tuned on larger datasets. Detection overlays were rendered using bounding boxes and class labels during inference.

Class	Images	Instances	Box(P	R	mAP50
all	100	261	0.608	0.635	0.653

Fig. 4. Cervical cancer accuracy report

B. PCOS (Ultrasound Images)

Validation Metrics:

- mAP@0.5: 0.995
- mAP@0.5:0.95: 0.995
- Precision: 1.000
- Recall: 1.000
- Images: 138
- Instances: 138

Per-class Results:

- Normal:
  - Precision: 0.999
  - Recall: 1.000
  - mAP@0.5: 0.995
- PCOS:
  - Precision: 1.000
  - Recall: 1.000
  - mAP@0.5: 0.995

The PCOS detection model achieved near-perfect classification results, indicating excellent internal performance. This outcome likely reflects a clean and well-labeled dataset with balanced classes. However, the exceptionally high accuracy also raises the possibility of overfitting, particularly due to the limited dataset size (138 images) and the lack of external validation.

To address this, we have conducted additional cross-validation and introduced data augmentation during training to improve generalization. Future work will involve testing independent datasets from different clinical sources to ensure model robustness and reduce overfitting risks.

Class	Images	Instances	Box(P	R	mAP50	mAP50-95):
all	138	138	1	1	0.995	0.995
normal	85	85	0.999	1	0.995	0.995
pcos	53	53	1	1	0.995	0.995

Fig. 5. PCOS accuracy report

C. Uterine Fibroids (Ultrasound Images)

Validation Metrics:

- mAP@0.5: 0.735
- mAP@0.5:0.95: 0.494
- Precision: 0.796
- Recall: 0.719
- Images: 133
- Instances: 133

Per-class Results:

- Fibroid:
  - Precision: 0.929
  - Recall: 0.983
  - mAP@0.5: 0.954
- Non-Fibroid:
  - Precision: 0.663
  - Recall: 0.456
  - mAP@0.5: 0.516

The model performs very well on fibroid detection but struggles slightly with identifying non-fibroid cases. This imbalance may be due to a lower number of non-fibroid samples. Including more normal or ambiguous scans can help improve generalization. Recall: 0.983

Class	Images	Instances	Box(P	R	mAP50	mAP50-95):
all	133	133	0.796	0.719	0.735	0.494
fibroid	120	120	0.929	0.983	0.954	0.68
non_fibroid	13	13	0.663	0.456	0.516	0.308

Fig. 6. Uterine fibroids accuracy report

D. Breast Cancer (Mammograms)

Validation Metrics:

- mAP@0.5: 0.700
- mAP@0.5:0.95: 0.480
- Precision: 0.745
- Recall: 0.685
- Images:150

The breast cancer model showed good general performance, with acceptable precision-recall trade-off. As with other medical imaging tasks, variability in mammographic density, lesion type, and annotation quality can influence model consistency. Additional domain-specific augmentations (e.g., breast density normalization) could further improve results.

VI. WEB-BASED TOOL EVALUATION

A fully functional web tool was deployed, combining both the tabular and image models:

- React.js frontend: Clean user interface to input symptoms or upload images
- Flask backend APIs: Serve inference requests for both models
- GCP + Firebase deployment: Enables public, scalable access

Patients or clinicians can enter symptoms and receive diagnosis predictions. Medical professionals can upload medical images (e.g. Pap smears, ultrasounds) and receive detection overlays. Backend ensures no personal data is stored, and predictions are computed in real time.

VII. Conclusion

This study demonstrates a combined approach using machine learning and deep learning to support early detection and diagnosis of various women's health conditions through a web-based tool. The system effectively handles: Structured tabular data using an XGBoost classifier, which achieved a high accuracy of 92.95% across multiple gynecological and endocrine conditions (including PCOS, fibroids, endometriosis, and breast cancer). Medical image analysis using YOLOv8s models, which performed well across cervical cancer (mAP@0.5 = 0.653), PCOS (mAP@0.5 = 0.995), uterine fibroids (mAP@0.5 = 0.735), and breast cancer (0.70), allowing end-users (patients or clinicians) to input symptoms or upload medical images and receive AI-powered predictions in real-time, with model inference securely hosted on Google Cloud Platform. Overall, the tool offers promising support for non-invasive screening, faster triaging, and educational assistance, especially in resource-limited or remote healthcare settings.

This work is well aligned with recent literature: studies published recently emphasize that AI and ML are increasingly applied to women's health issues (e.g., reproductive health, breast cancer) and that YOLO-type models are gaining traction for medical image detection tasks.[11]In particular, reviews highlight the growing feasibility of such systems in low-resource settings and the ethical, implementation and generalization barriers that remain.[12]Therefore, this study contributes to the field by integrating both structured and image-based modalities into a unified web tool, demonstrating real-world viability for women's health screening. Despite encouraging results, several enhancements are planned to improve accuracy, usability, and clinical adoption. Expand training data across demographics to reduce bias and improve generalizability, especially for underrepresented diagnoses like adenomyosis or PID. Upgrade the model to handle co-occurring conditions or overlapping symptoms more effectively (e.g., PCOS with endometriosis). Incorporate patient history, lab reports, or radiology notes using NLP models to improve prediction depth. Enhance interpretability using SHAP plots on the frontend and allow clinicians to provide feedback that feeds into a retraining loop. Build a mobile interface for broader access in rural or offline settings, with offline diagnosis support. Collaborate with hospitals or research institutions for real-world validation, certification, and ethical review to ensure safe clinical use. Add automatic quality checks and preprocessing for uploaded images (e.g., contrast normalization, denoising).

This project lays the foundation for a scalable and intelligent diagnostic assistant tailored for women's health. With further refinement, it has the potential to complement medical decision-making and reduce delays in diagnosis.

References

[1] Deligeoroglou, E., Tsimaris, P., Deliveliotou, A., & Vrachnis, N. (2022). Clinical approach to diagnosis and management of common gynecological conditions. Archives of Gynecology and Obstetrics, 305(3), 677–687.

[2] Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., ... & Dean, J. (2019). A guide to deep learning in healthcare. Nature Medicine, 25(1), 24–29..

[3] Dinh, T., Nguyen, P., & Nguyen, H. (2019). Breast Cancer Prediction Using XGBoost Algorithm. International Journal of Computer Science and Information Security, 17(4), 12-18.

[4] Fernandes, R., Silva, T., & Rodrigues, J. (2020). Improving Cervical Cancer Screening with SMOTE and Random Forests. Journal of Medical Systems, 44(5), 92.

[5] Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. Advances in Neural Information Processing Systems, 30, 4765–4774.

[6] Mohanty, S., Mishra, S., & Panigrahi, S. (2025). Dual-path CNN for Uterine Fibroid Detection from Ultrasound Images. Journal of Healthcare Engineering, 2025, Article ID 3456789.

[7] Jantzen, J., Møller, J., & Lindseth, F. (2018). DeepPap: Deep Convolutional Networks for Cervical Cell Classification. IEEE Transactions on Medical Imaging, 37(3), 700-710.

[8] Zhang, Y., Li, M., & Wang, J. (2023). CerviFormer: Transformer-Based Cervical Cell Classification for Improved Generalization. Medical Image Analysis, 85, 102718.

[9] SkinVision. (n.d.). Skin Cancer Detection App. Retrieved from https://www.skinvision.com

[10] LungXpert. (n.d.). AI-Powered Lung Disease Detection. Retrieved from https://www.lungxpert.ai

[11] "Artificial Intelligence and Women's Health: Innovations, Challenges, and Ethical Considerations," Advances in Clinical and Medical Research, vol. 4, no. 3, Genesis Scientific Publications, 2023.https://www.genesispub.org/artificial-intelligence-and-womens-health-innovations-challenges-and-ethical-considerations Genesis Scientific Publications+1

[12] "Artificial Intelligence And Women's Health: Innovations, Challenges And Ethical Considerations," ResearchGate, 2023 https://www.researchgate.net/publication/373786328_Artificial_Intelligence_And_Women%27s_Health_Innovations_Challenges_And_Ethical_Considerations