A research group in the field of Health Informatics aimed to develop an Artificial Intelligence (AI)-based decision support system for the early detection and staging of Chronic Kidney Disease (CKD) using kidney ultrasound images and patient clinical records. The dataset was collected from three hospitals over eight months and consisted of only 920 patient samples, including 450 Healthy cases, 180 CKD Stage 1 cases, 120 CKD Stage 2 cases, 80 CKD Stage 3 cases, 50 CKD Stage 4 cases, and 40 CKD Stage 5 cases, resulting in a highly imbalanced dataset. Due to differences in ultrasound machines and acquisition protocols, many images suffered from low quality, noise, and inconsistent resolution. The researchers performed data cleaning, missing value handling, image resizing (224×224), normalization, and basic augmentation techniques such as rotation, flipping, and brightness adjustment. The dataset was divided into training (70%), validation (15%), and testing (15%) sets. Three deep learning models—Custom CNN, ResNet50, and MobileNetV2—were implemented and evaluated. The best-performing model, ResNet50, achieved only 67.8% accuracy, while the Custom CNN and MobileNetV2 achieved 61.4% and 65.2% accuracy, respectively. To increase transparency and physician trust, Explainable AI (XAI) techniques including Grad-CAM and SHAP were incorporated. However, the explanations revealed that the model frequently focused on irrelevant image regions and struggled to distinguish between advanced CKD stages due to limited and imbalanced training data. The researchers then deployed the ResNet50 model as a web-based healthcare application that allowed physicians to upload ultrasound images, receive CKD stage predictions, visualize Grad-CAM heatmaps, and generate patient reports. During pilot testing on 50 unseen patients, the system's accuracy decreased to approximately 60%, and doctors reported concerns regarding reliability, prediction consistency, and trustworthiness. Although the study successfully demonstrated an end-to-end AI-powered CKD screening framework with deployment and explainability features, several limitations remained, including insufficient data volume, severe class imbalance, low image quality, lack of multimodal data integration, inadequate generalization capability, limited clinical validation, and relatively poor diagnostic performance.

1. Draw a complete high-level workflow diagram representing the entire research process from data collection to deployment.

2. Identify and explain all major limitations of the study. And Propose technical and methodological improvements that could significantly enhance the system's performance and reliability.