

ABSTRACT

A COMPARATIVE STUDY OF IMBALANCE HANDLING TECHNIQUES (SMOTE, DATA AUGMENTATION, AND RANDOM OVERSAMPLING) IN KIDNEY CT SCAN CLASSIFICATION USING DENSENET-121 DEEP LEARNING MODEL

Renal diseases, including tumors, cysts, and stones, present significant health challenges and can progress to Chronic Kidney Disease (CKD) if not detected early. The use of CT scans enables detailed analysis of kidney conditions, yet traditional diagnostic processes are time-consuming and heavily reliant on radiological expertise, which can lead to delays. With advancements in artificial intelligence, deep learning models have shown promise in enhancing diagnostic accuracy by automating the analysis of medical images. However, one key challenge in utilizing deep learning for kidney disease detection is the class imbalance in medical imaging datasets, where certain conditions are underrepresented, leading to biased model predictions.

This study proposes a comparative analysis of various data imbalance handling techniques—Synthetic Minority Over-sampling Technique (SMOTE), data augmentation, and random oversampling—to evaluate their effectiveness in kidney CT scan classification using deep learning. The chosen deep learning architecture, DenseNet-121, will be trained and evaluated on a dataset with diverse kidney conditions, with experiments conducted on both balanced and imbalanced data. By implementing these balancing methods, this study aims to assess how each technique influences model performance in accurately classifying renal diseases, thereby providing insights into optimizing deep learning for medical diagnosis.

This research identified Random Oversampling as the most effective data-balancing technique for improving classification accuracy and reliability in imbalanced kidney CT scan datasets. Our findings demonstrated that Random Oversampling enhanced the overall accuracy and maintained a balanced precision-recall trade-off, proving it to be a robust method for handling class imbalance. These insights are expected to influence future medical imaging research and aid in developing diagnostic tools that improve early detection and treatment of kidney diseases. The success of Random Oversampling, along with the comparative analysis of SMOTE and Data Augmentation, suggests that careful selection and application of data balancing techniques can significantly enhance deep learning models in healthcare, addressing dataset imbalance challenges and improving clinical outcomes for patients with renal conditions.

Keywords: Renal disease, kidney CT scan, deep learning, data balancing, medical imaging, kidney tumor, cyst, stone classification