

Evaluation of Optimizers in Pre-trained Neural Networks: Case Study on the Classification of Pathologies in Thyroid Scans Using a Reduced and Unbalanced Dataset

Benicio Grossling-Vallejos, Julio César Mello-Román, José Luis Vázquez Noguera, Horacio Legal-Ayala & Ronald Rivas Coluchi

Abstract

This study comparatively evaluates the performance of three pre-trained convolutional neural network architectures (ResNet50, InceptionV3, and DenseNet169) in the multiclass classification of thyroid scintigraphy images, using a reduced and unbalanced dataset. The models were trained in two experimental configurations and three optimizers (SGDM, AdamW, and RMSprop). Evaluation metrics included precision, precision, recall, F1-score and ROC-AUC curves per class. The results indicate that DenseNet169 achieved the best overall performance, with higher F1-scores (with RMSprop) and more balanced ROC curves between classes (with AdamW). InceptionV3 was found to offer intermediate performance, but with greater variability between classes, while ResNet50 showed lower discriminatory power, particularly in the less represented classes. The results highlight the usefulness of the DenseNet169 model for classification tasks with limited data availability and its use as a diagnostic support tool in nuclear medicine.