Skin Cancer Classification using YOLOv8

1.Abstract

This project aims to classify skin lesions as **benign** or **malignant** using the **YOLOv8** object detection and classification framework. Skin cancer is among the most common types of cancer worldwide, and early and accurate diagnosis plays a critical role in successful treatment. Deep learning approaches like YOLO can provide fast and reliable assistance to dermatologists by automating lesion detection and classification from dermoscopic images.

This project follows a similar methodology to the previous blood cell classification project, with adjustments made for binary classification. The dataset used contains dermoscopic images labeled as either benign or malignant.

Data was preprocessed and divided into training, validation, and testing subsets to ensure accurate model evaluation. The YOLOv8 model was trained using PyTorch and the **Ultralytics** library with customized training parameters. Techniques like data augmentation and learning rate scheduling were applied to enhance model performance.

2.Dataset

- **Source**: Dataset:https://www.kaggle.com/datasets/fanconic/skin-cancer-malignant-vs-benign?resource=download&select=train
- **Classes**:
  - `benign` (non-cancerous)
  - `malignant` (cancerous)

Each image is annotated in YOLO format, and the dataset is structured into `train`, `val`, and `test` directories.

3.Code and Functions

Step	Description
Dataset Preparation(line9)	Randomly divides images into `train`, `val`, and `test` folders with appropriate class labels.
YOLO Installation(line50)	Installation of YOLOv8 and dependencies using `pip install ultralytics`.
Model Training(line58)	YOLOv8 is trained using custom dataset with binary classification setup.
Prediction(line89)	The trained model is run on unseen test images.
Visualization(line102)	Predicted class labels are overlaid on test images for performance analysis.

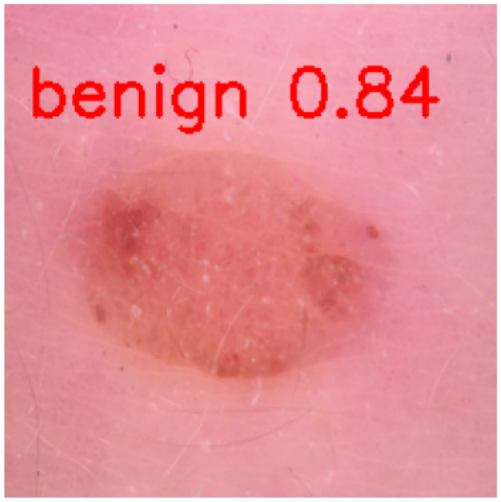
4.Results

4.i) Random Test Results for Skin Cancer Classification:

Below are randomly selected test results from the model's predictions. Each example shows the predicted lesion type (**benign** or **malignant**) alongside the actual type for comparison purposes. This section is intended to visually demonstrate the model's performance in distinguishing between cancerous and non-cancerous skin lesions.

Actual Type: Malignant

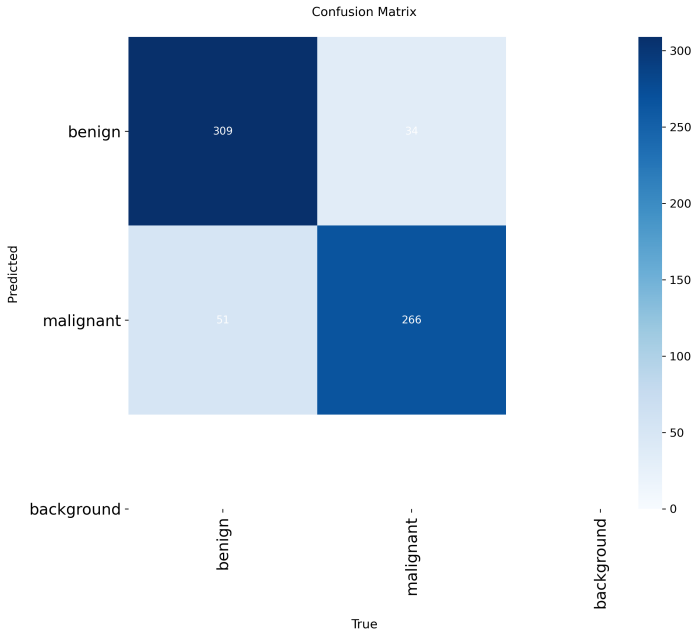
Fig.1: Skin Cancer Lesion Prediction



Actual Type: Benign

Fig.2: Skin cancer Lesion Prediction

4.ii) Confusion Matrix:



4.iii)Training Metrics:

Table 1: Training Results Table

Epoch	GPU Memory (GB)	Training Loss	Top-1 Accuracy	Top-5 Accuracy
1	5.81	0.4399	0.821	1.0
2	6.92	0.3960	0.824	1.0
3	6.95	0.3424	0.867	1.0
4	6.95	0.3182	0.859	1.0
5	6.95	0.2901	0.871	1.0

5. References and Acknowledgments:

• Dataset:https://www.kaggle.com/datasets/fanconic/skin-cancer-malignant-vs-benign?resource=download&select=train

• Thanks to the Ultralytics team for developing the YOLO framework.

Author:

Batuhan Kılıçarslan