Gastrointestinal Image Classification using MobileNetV2

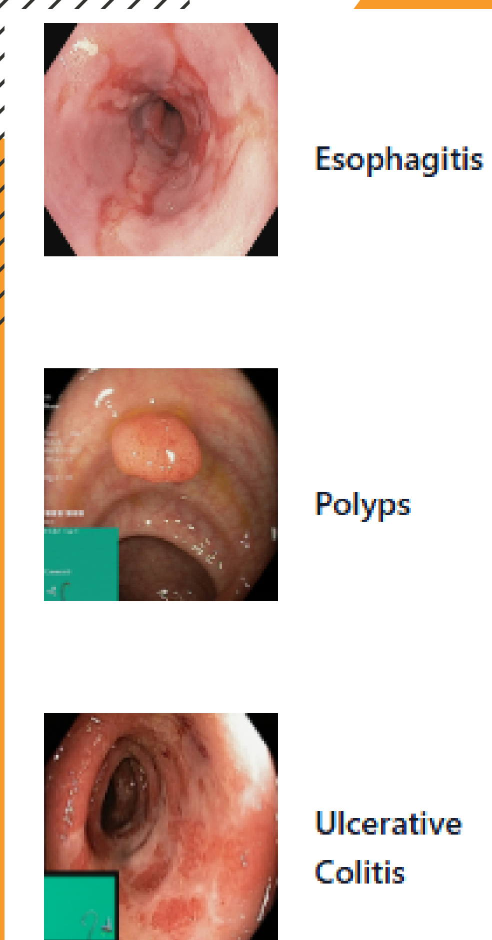
A PRESENTATION BY PABLO PEREZ

Introduction

The goal was to train a model that could distinguish between esophagitis, polyps, and ulcerative colitis with high accuracy, and ideally help in diagnostic support.

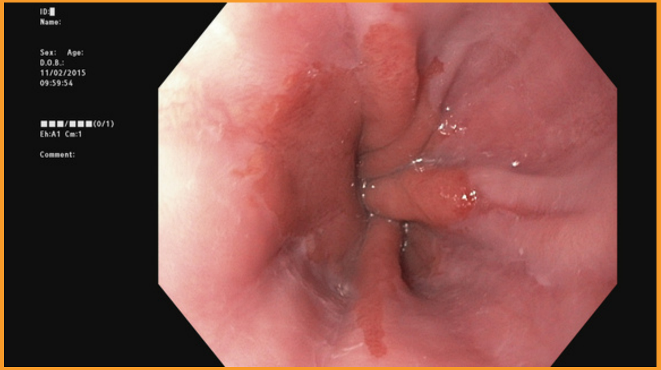
Objective: Classify endoscopic images into three GI conditions.

Dataset Overview



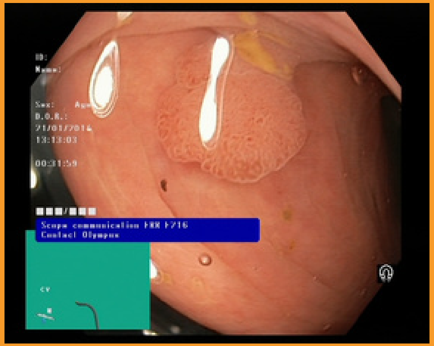
- The dataset used is the Kvasir v2 — it's a public medical dataset with a bunch of high-quality images. We focused on just three classes for this project: esophagitis, polyps, and ulcerative colitis.
- Each class had around 1000 images, so we were working with a total of about 3000 images.
- This gave us a total of 3000 high-resolution endoscopic images. The dataset was split into 80% training and 20% validation to ensure balanced and robust evaluation.

Data Problems



Black Boxes
esophagitis

green corner box
polyps



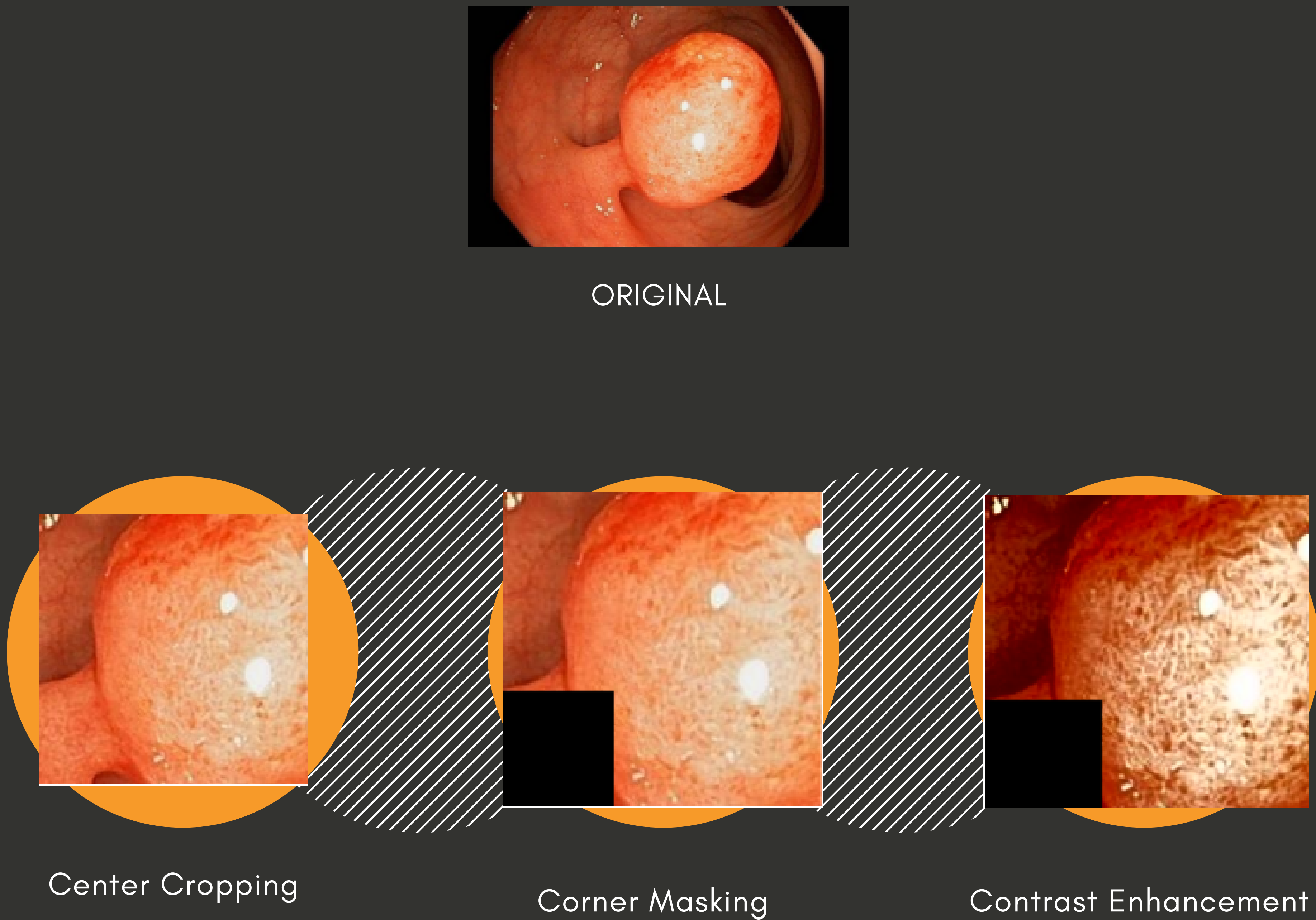
weird large blue boxes

Preprocessing Pipeline

Center Cropping: Standardizes input size to 224x224 pixels.

Corner Masking: Removes artifacts often found in the bottom-left region.

Contrast Enhancement: Uses histogram equalization in YCrCb color space to enhance detail.

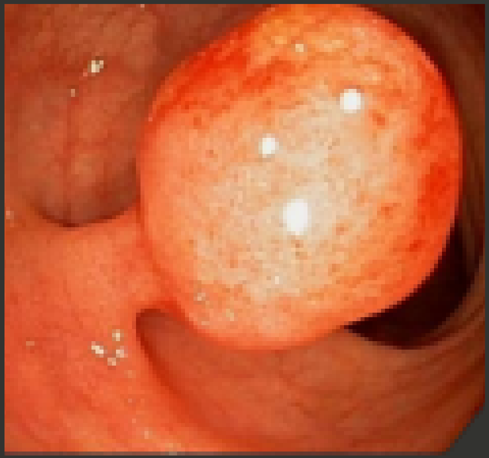

ORIGINAL
Center Cropping
Corner Masking
Contrast Enhancement

Data Augmentation

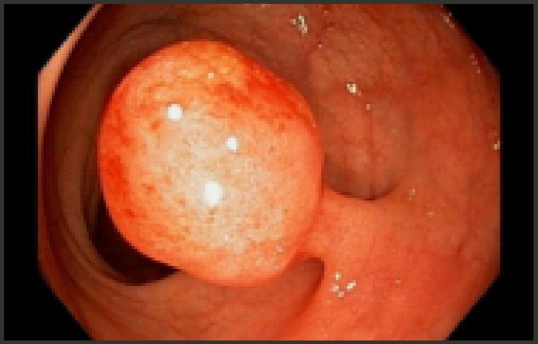
To improve generalization and address overfitting, I applied the following augmentations during training:

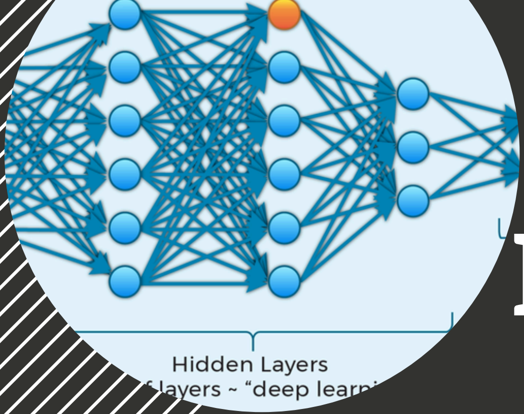
Random rotations, zooming, shearing

Width and height shifts



Horizontal flipping
This approach allowed the model to become more robust to variations in image acquisition.





Model Architecture

We employed MobileNetV2, a lightweight CNN pretrained on ImageNet. The architecture was adapted as follows:

- A global average pooling layer,
- A dense layer with 128 neurons,
- A dropout layer with a 0.5 rate for regularization,
- And a final dense layer with 3 output neurons and softmax activation for our three classes."

Training Strategy

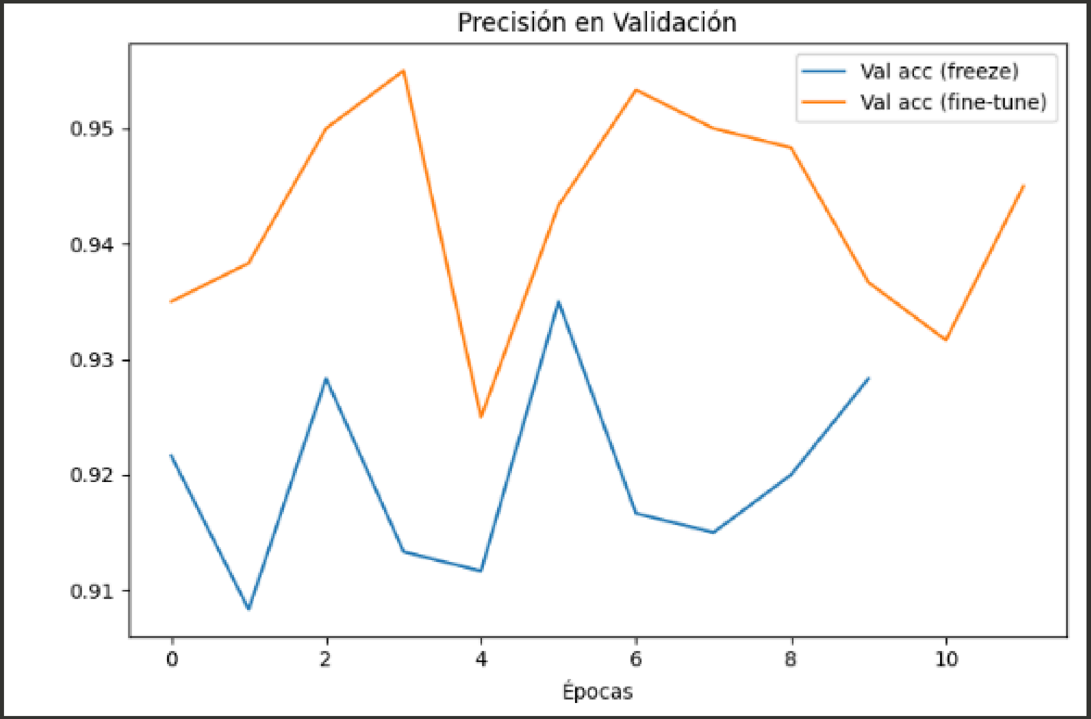
The training was conducted in two phases:

Phase 1: Freeze the base model and train the top layers for 10 epochs.
Phase 2: Unfreeze the last 30 layers and fine-tune the model for 20 more epochs using a lower learning rate.

Early stopping and model checkpointing were used to avoid overfitting and retain the best model.

Accuracy and Loss Curves

The learning curves showed consistent improvement:



Evaluation Results

- After loading the best-performing weights, the model was evaluated on the validation set:

Matriz de confusión:

```
[[195   1   4]
 [  5 180  15]
 [  0  11 189]]
```

Reporte de clasificación:

	precision	recall	f1-score	support
esophagitis	0.9750	0.9750	0.9750	200
polyps	0.9375	0.9000	0.9184	200
ulcerative-colitis	0.9087	0.9450	0.9265	200
accuracy			0.9400	600
macro avg	0.9404	0.9400	0.9399	600
weighted avg	0.9404	0.9400	0.9399	600

Evaluation Results

```
Clase: esophagitis
TP: 195, FP: 5, FN: 5, TN: 395
Precisión: 0.9750, Recall: 0.9750, Specificity: 0.9875
Accuracy: 0.9833, F1-score: 0.9750

Clase: polyps
TP: 180, FP: 12, FN: 20, TN: 388
Precisión: 0.9375, Recall: 0.9000, Specificity: 0.9700
Accuracy: 0.9467, F1-score: 0.9184

Clase: ulcerative-colitis
TP: 189, FP: 19, FN: 11, TN: 381
Precisión: 0.9087, Recall: 0.9450, Specificity: 0.9525
Accuracy: 0.9500, F1-score: 0.9265
```

Conclusion

MobileNetV2 proved to be a strong base model for medical image classification.
The custom preprocessing pipeline significantly improved image quality and model focus.
Fine-tuning the last layers **further enhanced model accuracy** and class recall.

Future Work:

- Explore other architectures like EfficientNet or Vision Transformers.
- Incorporate more GI conditions and expand dataset diversity.
- Test real-world deployment in clinical environments.

Thank You

WWW.NORIACCLINIC.COM