Code Crunch project

🩺 GastroVision — Automated Gastrointestinal Disease Classification

By: *Gaurav Kumar Dhakad*

Platform: Kaggle (TensorFlow + Keras)

Dataset: GastroVision – Gastrointestinal Disease Detection

🧠 Problem Statement

Gastrointestinal (GI) diseases are among the most common medical issues worldwide.

Early diagnosis using **endoscopic imaging** can save lives, but manual interpretation is slow and error-prone.

The goal of this project was to build a **Deep Learning model** that can automatically classify GI endoscopy images into **27 disease and anatomical landmark categories**, providing faster and more reliable diagnostics assistance.

🎯 Objectives

1. Load and preprocess the GastroVision dataset efficiently (without manual splitting).
2. Build a **Convolutional Neural Network (CNN)** capable of multi-class classification.
3. Train the model using **GPU acceleration on Kaggle**.
4. Evaluate the model using accuracy, loss, and validation metrics.

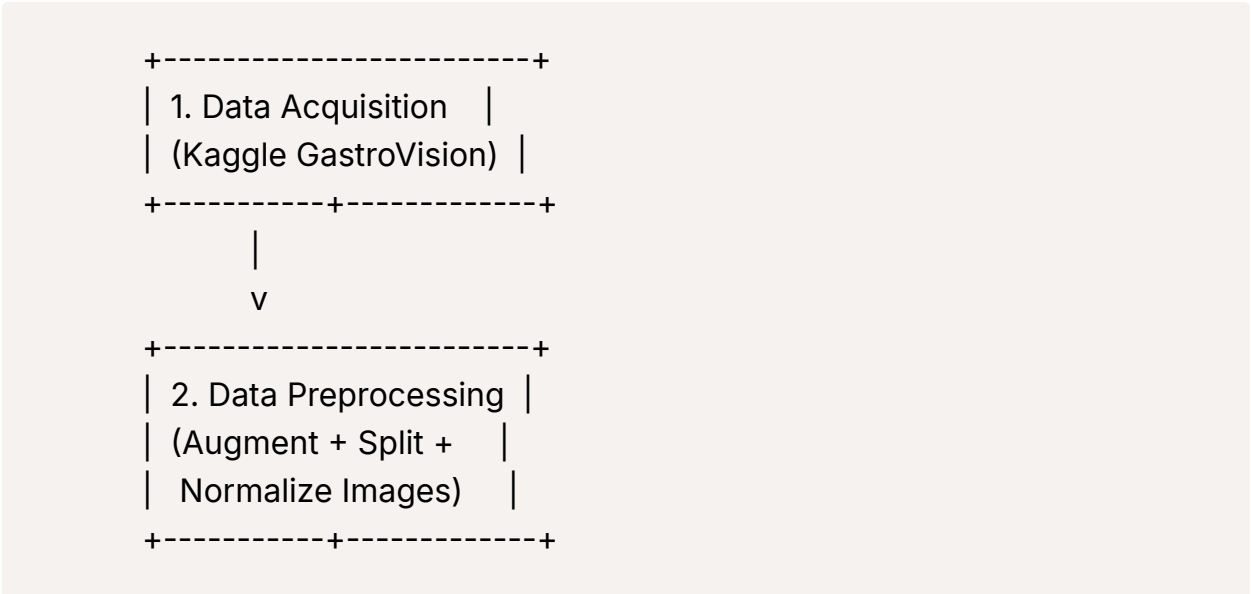
📦 Dataset Overview

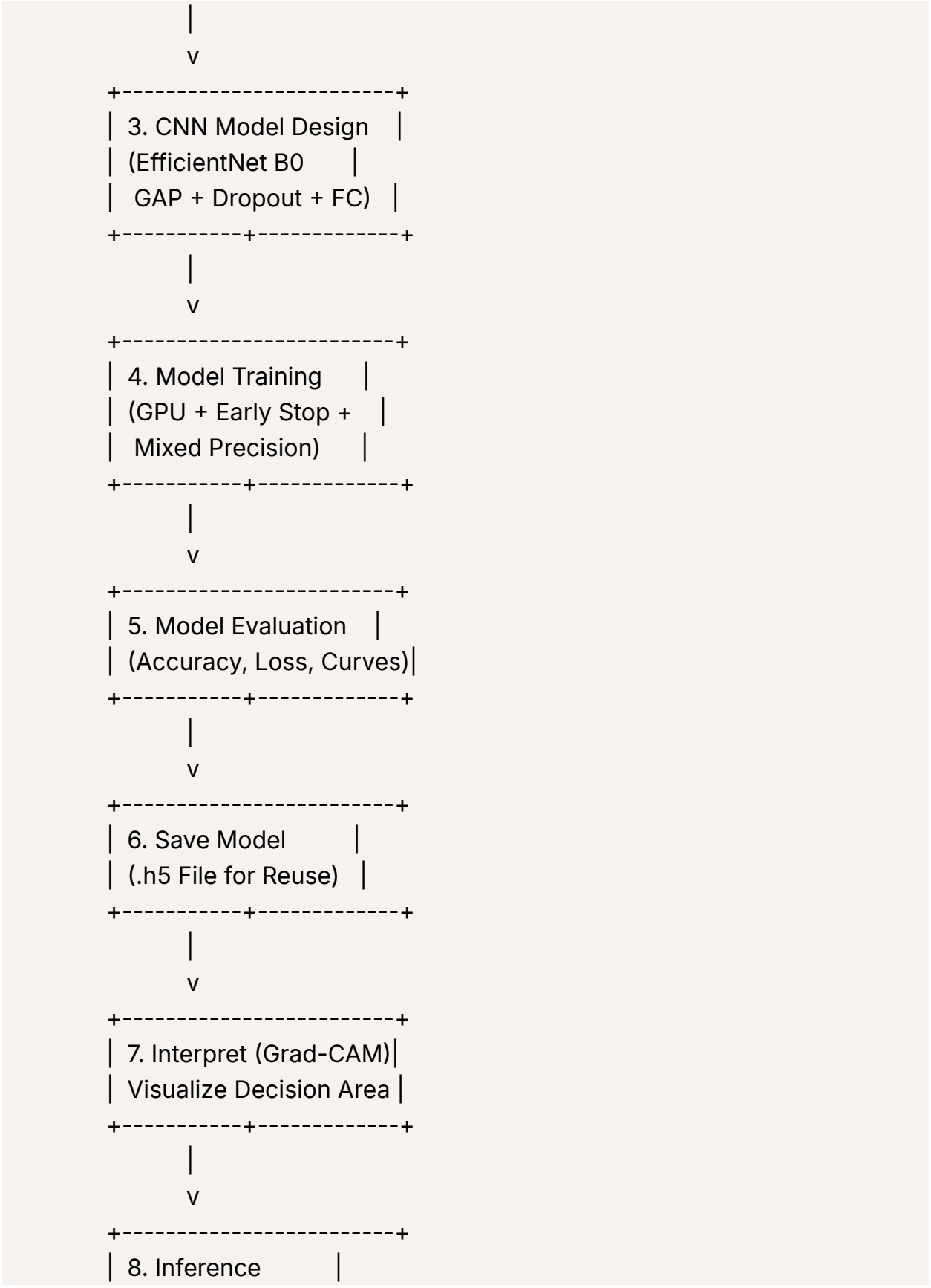
- **Dataset name:** GastroVision – Gastrointestinal Disease Detection
- **Source:** Kaggle (uploaded by Orvile)
- **Total images:** ~27,000
- **Classes:** 27 (diseases and anatomical landmarks)
- **Structure:**

```
Gastrovision/
  ├── Angiectasia/
  ├── Barretts esophagus/
  ├── Blood in lumen/
  ├── Colon polyps/
  ├── Colorectal cancer/
  ├── ... (27 folders total)
  └── Ulcer/
```

Each folder represents one class.

Model Workflow




3. CNN Model Design
(EfficientNet B0
GAP + Dropout + FC)
4. Model Training
(GPU + Early Stop +
Mixed Precision)
5. Model Evaluation
(Accuracy, Loss, Curves)
6. Save Model
(.h5 File for Reuse)
7. Interpret (Grad-CAM)
Visualize Decision Area
8. Inference

```
        | Predict on new images |
        +-----------------------+
```

Why i have use EfficientNet B0?

What is EfficientNet?

EfficientNet is a **family of convolutional neural networks** developed by **Google AI (2019)**.

Unlike older CNNs that scaled only one dimension (depth or width), EfficientNet uses **compound scaling** — a method that scales *depth*, *width*, and *resolution* **together** in a balanced way.

Why EfficientNetB0 Fits *GastroVision*

Challenge	Why EfficientNetB0 Solves It
Medium-size dataset (~27K images)	Pretrained EfficientNet generalizes well, even with fewer medical images.
High intra-class similarity (many GI images look alike)	EfficientNet learns fine-grained visual patterns efficiently.
Need for speed on Kaggle GPU	B0 is lightweight (5M params), runs fast even on free GPUs.
Limited compute resources	Lower FLOPs → faster inference, lower memory footprint.
Transfer learning ready	Pretrained on ImageNet — can transfer low-level features (edges, textures).

So, **EfficientNetB0 = sweet spot** of **accuracy, speed, and efficiency**.

EfficientNetB0 achieves similar or better accuracy with **8× fewer parameters** and **6× faster training**. It's more parameter-efficient and better suited for limited GPUs. The model uses modern architectural optimizations like compound scaling and squeeze-excitation, making it ideal for medium-scale medical datasets like GastroVision.

Visual summary of our model

```
Input Image (224×224×3)
      ↓
EfficientNetB0 (Feature Extractor)
      ↓
Global Average Pooling
      ↓
Dropout(0.4)
      ↓
Dense(27, activation='softmax')
      ↓
Predicted GI Disease Label
```

Model Performance Summary

Metric	Value (%)	Description
Accuracy	52%	The model correctly classified 52% of all validation images.
Precision	21.60%	On average, 21.60% of the model's positive predictions were correct.
Recall	20.82%	The model correctly detected 20.82% of all disease cases.
F1-Score	18.78%	