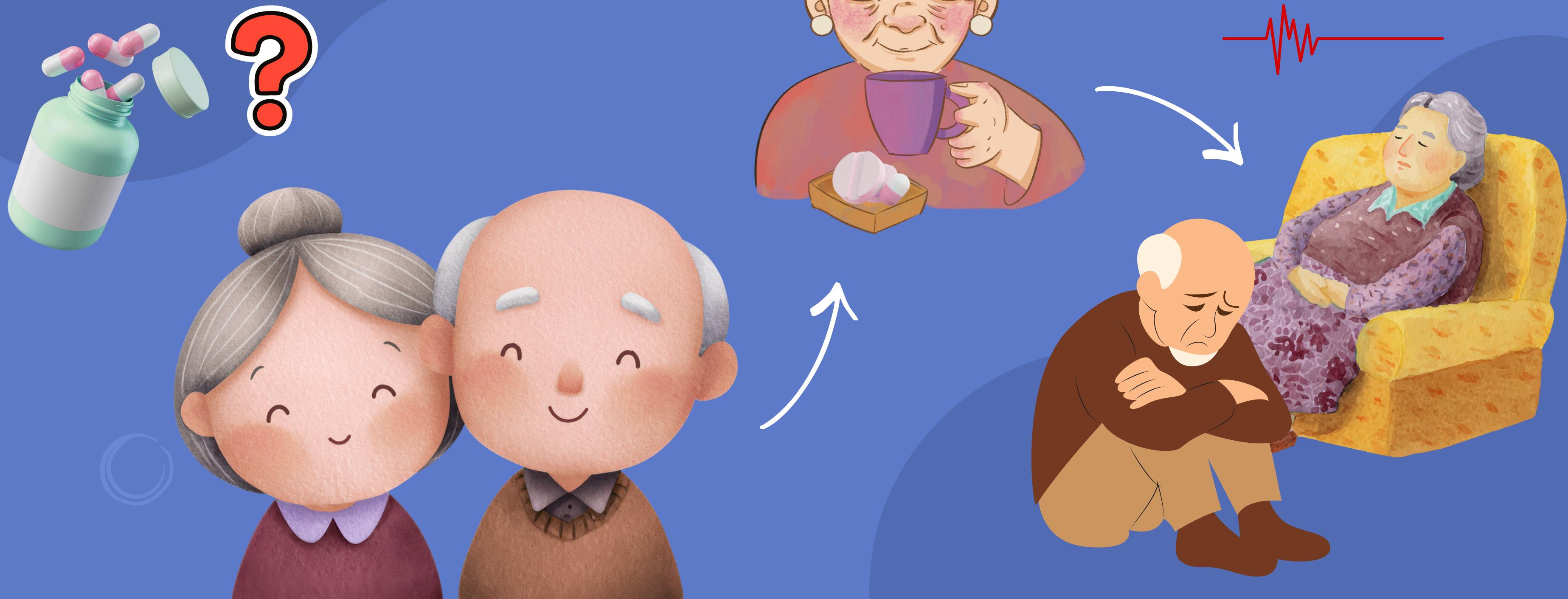




Pharmaceutical Drugs Classification

Airis Alesya Binti Shafarin | FCS13225

Problem Statement



The screenshot shows the Kaggle dataset page for 'Pharmaceutical Drugs and Vitamins Syn'. At the top, there's a search bar, 'Sign In' and 'Register' buttons, and a 'Download' button. Below the header, the dataset title 'Pharmaceutical Drugs and Vitamins Syn' is displayed with a count of 159 files. There are links for 'Data Card', 'Code (24)', 'Discussion (0)', and 'Suggestions (0)'. The main area is titled 'Data Combined (10 directories)' and shows a grid of ten folder icons. The first five are labeled: 'Alaxan 1000 files', 'Bactidol 1000 files', 'Bioflu 1000 files', 'Biogesic 1000 files', and 'DayZinc 1000 files'. The next three are partially visible. To the right is a 'Data Explorer' sidebar showing a hierarchical tree structure of the dataset. The root node 'Drug Vision' contains a 'Data Combined' folder which further branches into ten categories: Alaxan, Bactidol, Bioflu, Biogesic, DayZinc, Decolgen, Fish Oil, Kremil S, Medicol, and Neozep. Below the tree is a 'Summary' section indicating 20.0k files.

Data Overview

The dataset includes images of popular pharmaceutical drugs and vitamins in the Philippines. Currently, there are ten available classes of pill images.



Objectives & Key Questions

- To build an image-based pill classification model using deep learning.
- To analyze characteristics of pill images, ensuring the performance of the trained model.
- How accurately can the trained model classify the pills into their correct categories?
- Which pill category is the most frequently misclassified?



Methodology



Data Collection



Data Quality
Check



Data
Augmentation



Model Building,
Compiling & Training



Model
Evaluation

EDA Key Findings

The plot shows that the categories are balanced. This helps prevent the model from becoming biased towards a majority class.



Sample Image Visualization



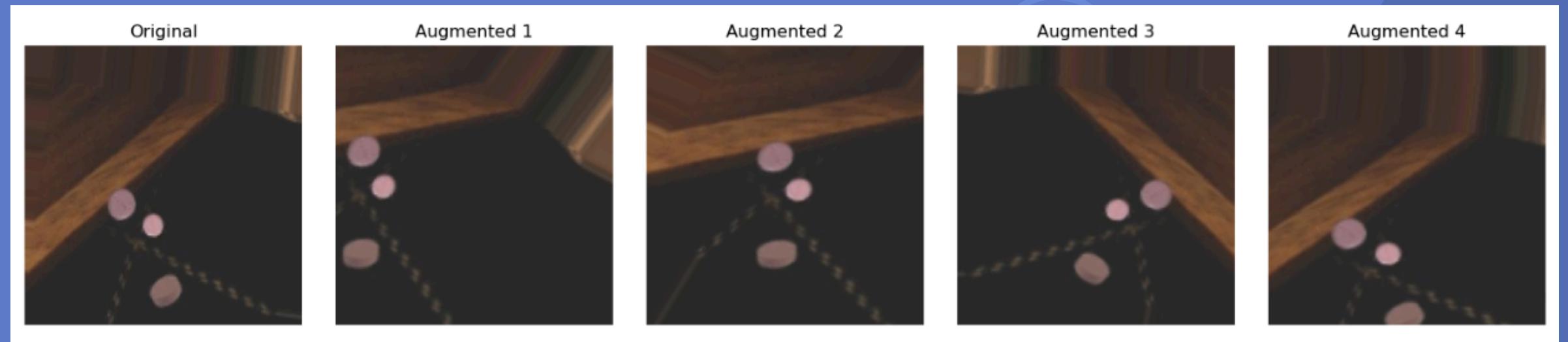
Before



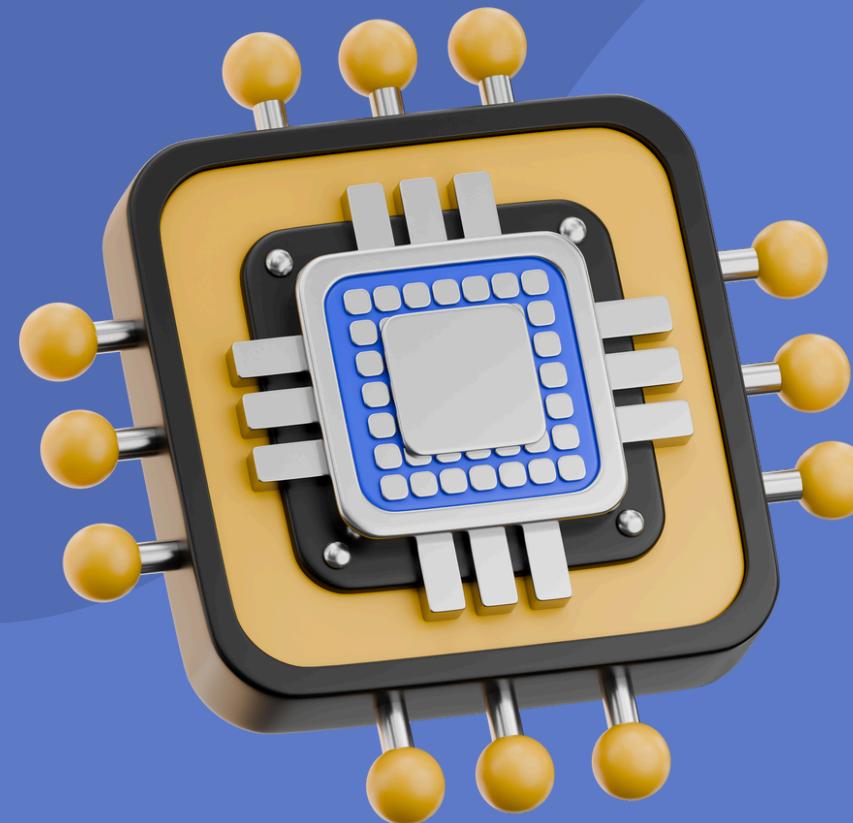
After

```
image_quality(samp_image_path)
```

The image is not blurry, Laplacian operator variance: 123.91
{'sharpness': 123.91, 'resolution': 330.23}



Modeling Approach



CNNs model can detect features regardless of where they appear in the image.



80% of the images were used for training, while 20% were used for validation.



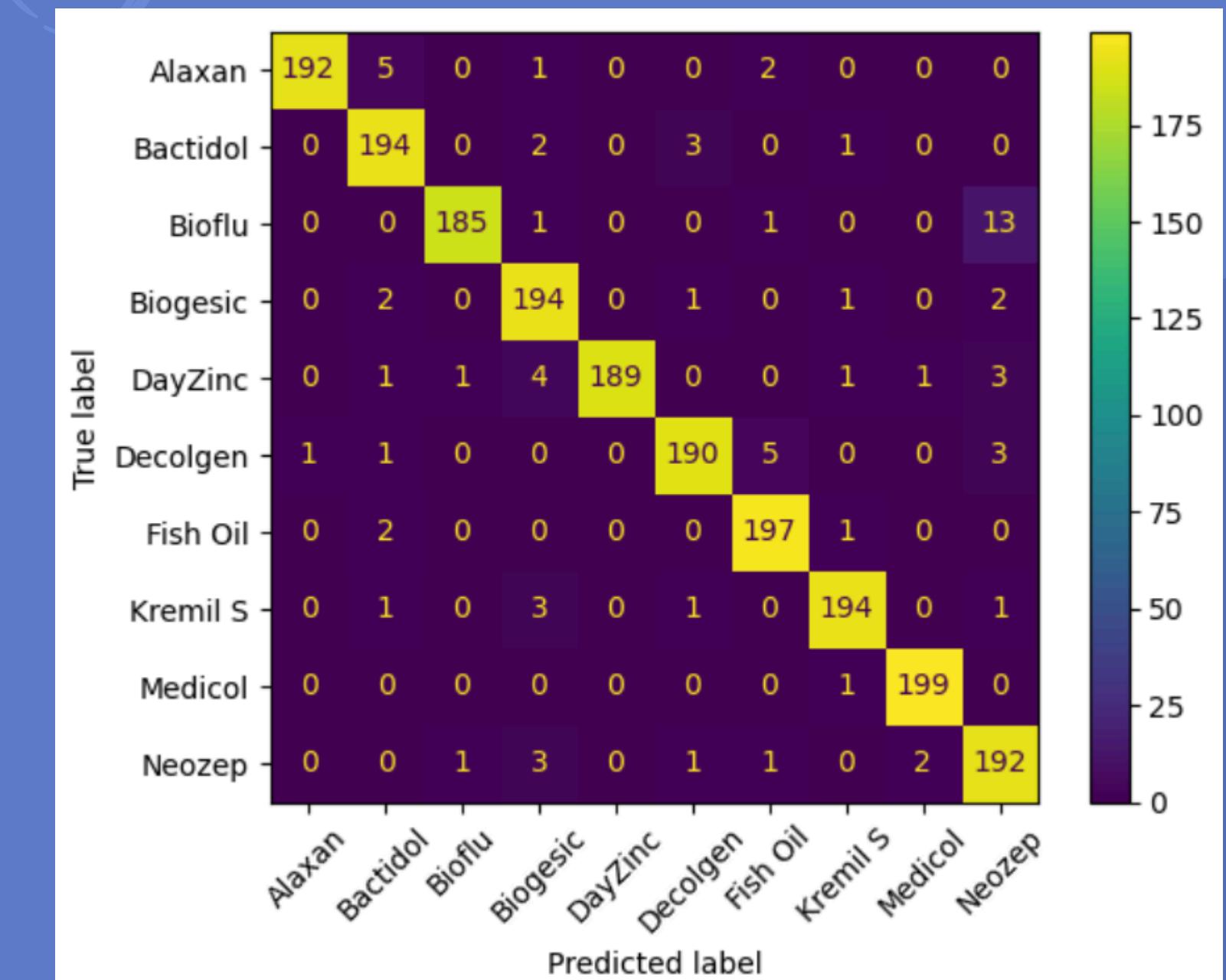
Images were rescaled and resized. Data augmentation was used to reduce overfitting.

Results & Evaluation

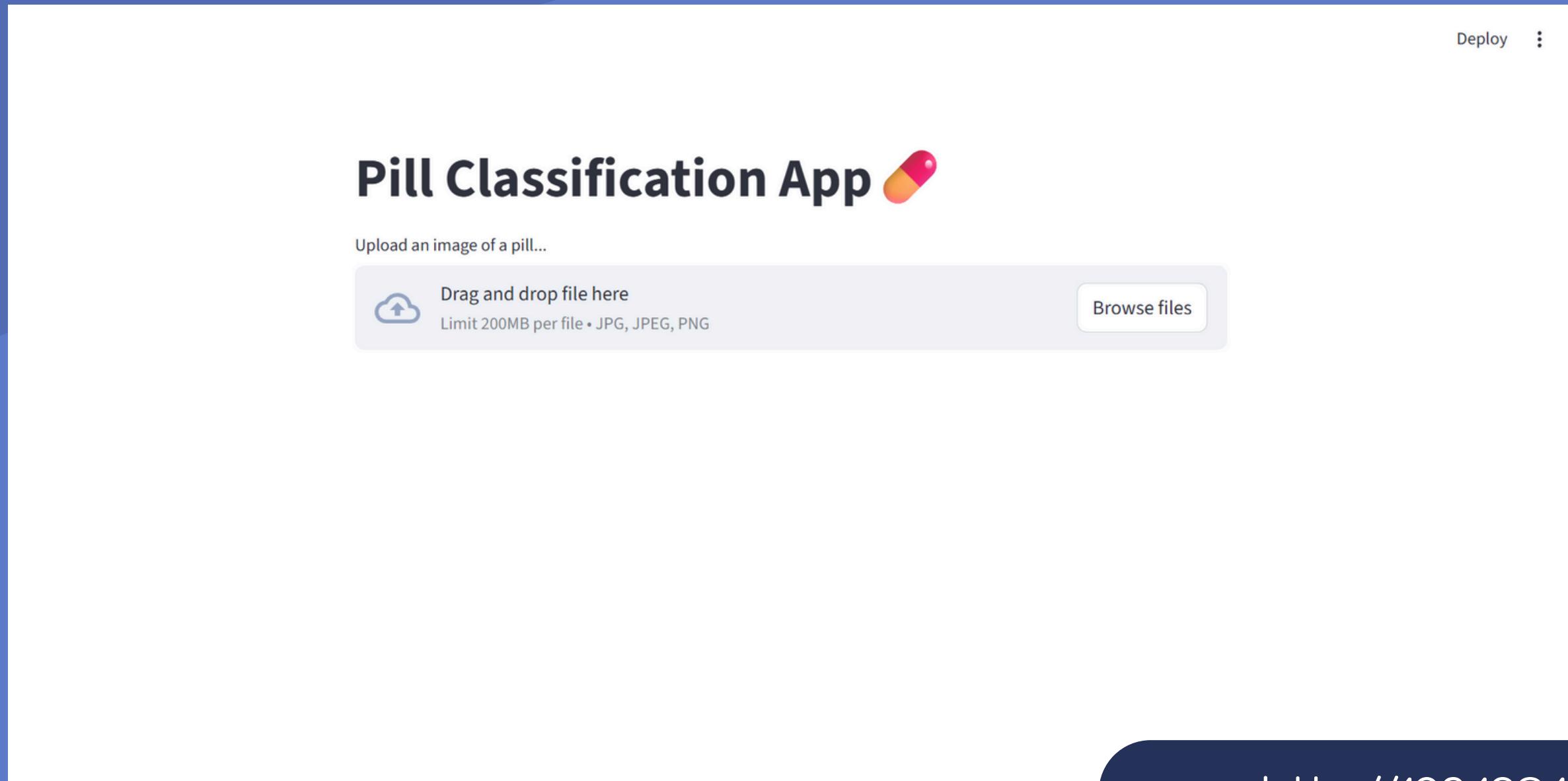
```
# accuracy score
model = tf.keras.models.load_model('pill.keras')
t_lost,t_acc = model.evaluate(val_generator, verbose=1)

63/63 ----- 10s 146ms/step - accuracy: 0.9630 - loss: 0.1358
```

With an overall accuracy of 96.3%, the model can accurately predict most pills from images. This reduces the risk of mistaking pills that can eventually lead to overdosing.

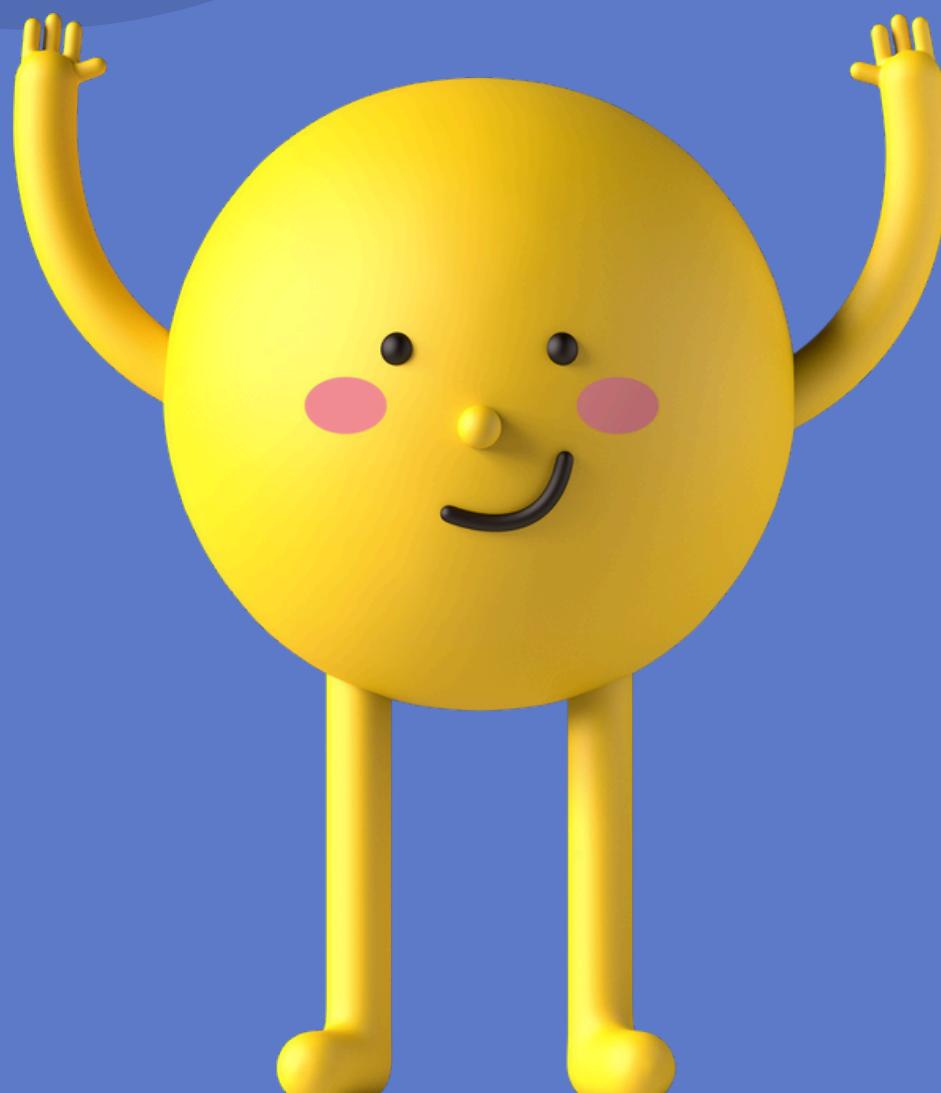


Project Demo



<http://192.168.100.150:8501>

Measure of Success



I targeted at least 90% accuracy score
and I achieved it!



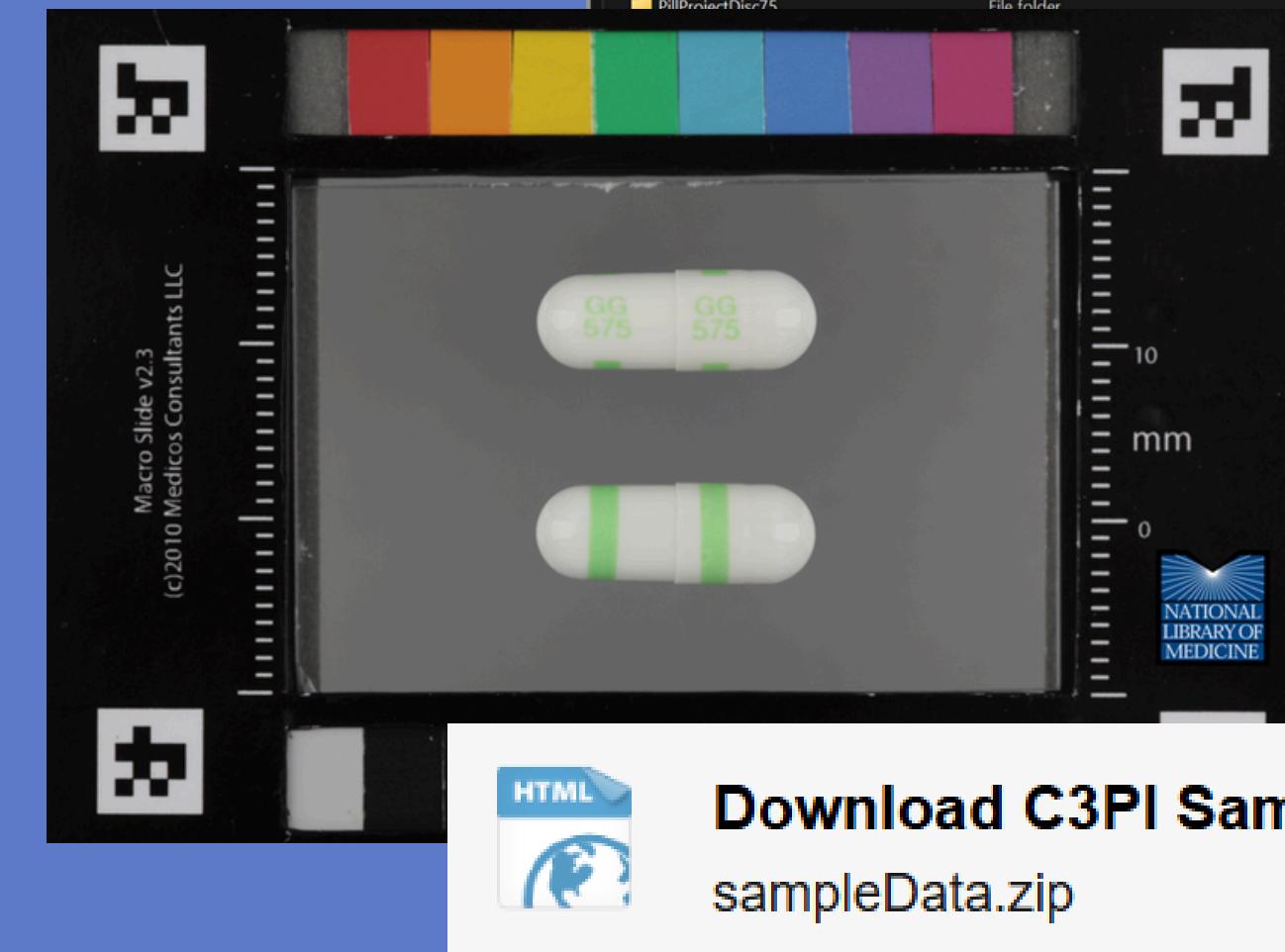
```
# accuracy score
model = tf.keras.models.load_model('pill.keras')
t_lost,t_acc = model.evaluate(val_generator, verbose=1)

63/63 ━━━━━━━━━━━━ 10s 146ms/step - accuracy: 0.9630 - loss: 0.1358
```

Limitations & Future Works



The data that I used instead.



[Download C3PI Sample Data \(493MB zip file\)](#) 🔥
[sampleData.zip](#)

Sample data that contains
images of a **single** pill.

Name	Type	Compressed size	Password pr...	Size	Ratio	Date modified
PillProjectDisc9	File folder					19/4/2019 11:36 AM
PillProjectDisc18	File folder					19/4/2019 11:35 AM
PillProjectDisc28	File folder					19/4/2019 11:32 AM
PillProjectDisc31	File folder					19/4/2019 11:35 AM
PillProjectDisc37	File folder					19/4/2019 11:34 AM
PillProjectDisc62	File folder					19/4/2019 11:35 AM
PillProjectDisc63	File folder					19/4/2019 11:36 AM
PillProjectDisc64	File folder					19/4/2019 11:34 AM
PillProjectDisc75	File folder					19/4/2019 11:36 AM

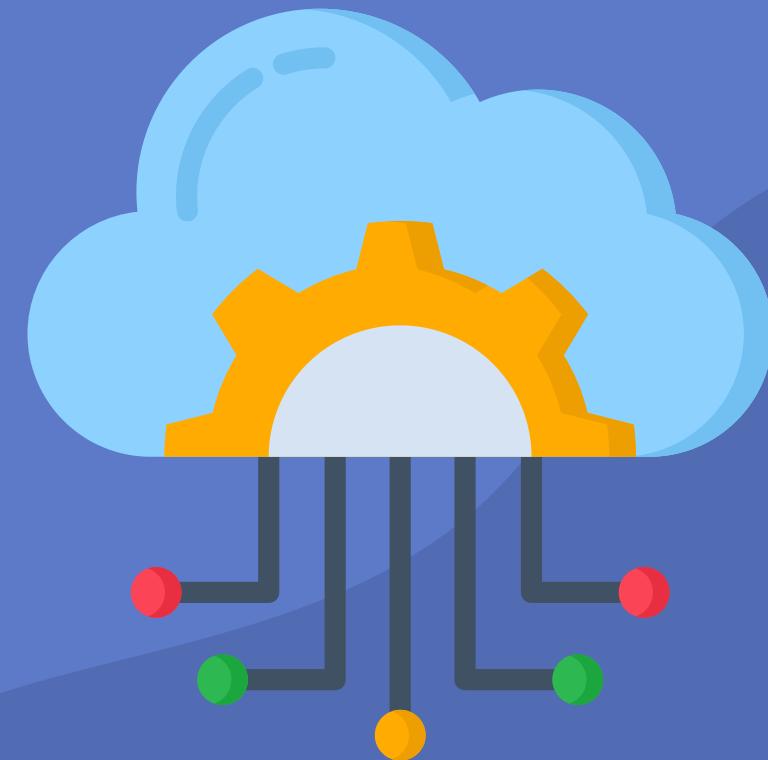
Tech Stack



Python language



Pandas, NumPy, Scikit-learn,
TensorFlow, Matplotlib,
OpenCV, Pillow



GitHub, Streamlit



Thank you for listening

Any questions?