Detective StarLung: Model V1.1 Update

Implementing Data Augmentation to Enhance Generalization & Reduce Inter-Class Confusion

The Challenge: Overcoming Baseline Limitations

V1.0 Baseline Recap

Low Baseline (Attempt 1)

34% accuracy at 256x256 res, 10 EPOCHS

High Baseline (Attempt 2)

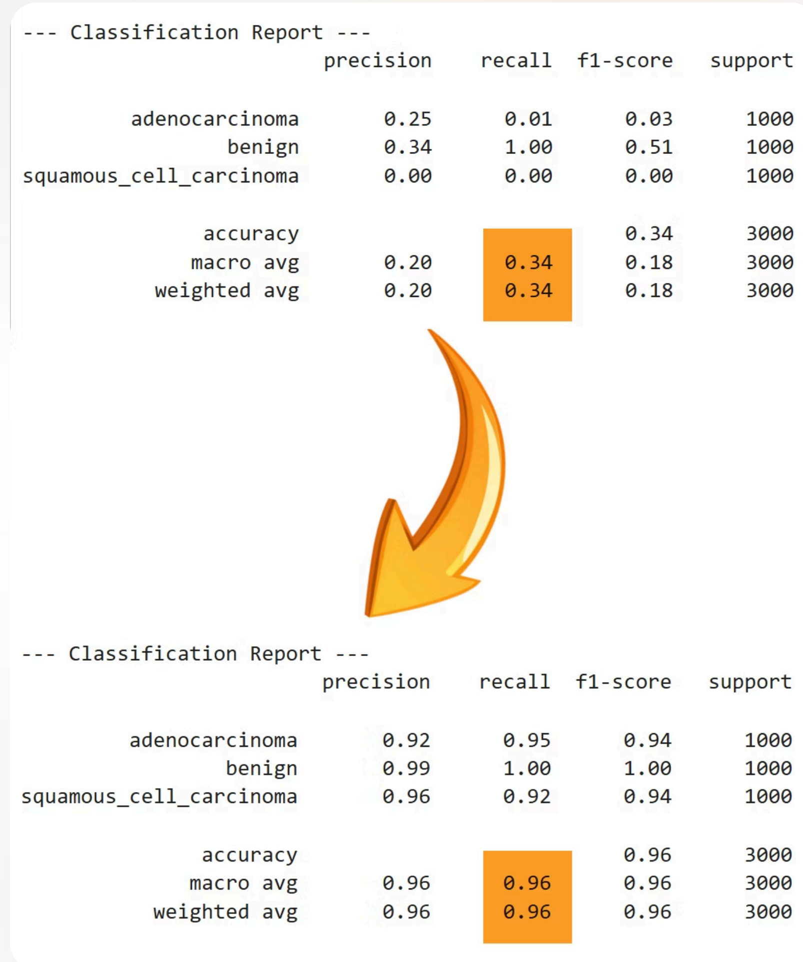
96% accuracy at 128x128 res, 5 EPOCHS

The Problem: Inter-Class Confusion

Previous models struggled to distinguish between:

- Adenocarcinoma
- Squamous Cell Carcinoma

Solution: V1.1 introduces **Data Augmentation** for robust feature learning.



--- Classification Report ---

	precision	recall	f1-score	support
adenocarcinoma	0.25	0.01	0.03	1000
benign	0.34	1.00	0.51	1000
squamous_cell_carcinoma	0.00	0.00	0.00	1000
accuracy			0.34	3000
macro avg	0.20	0.34	0.18	3000
weighted avg	0.20	0.34	0.18	3000

--- Classification Report ---

	precision	recall	f1-score	support
adenocarcinoma	0.92	0.95	0.94	1000
benign	0.99	1.00	1.00	1000
squamous_cell_carcinoma	0.96	0.92	0.94	1000
accuracy			0.96	3000
macro avg	0.96	0.96	0.96	3000
weighted avg	0.96	0.96	0.96	3000

Data Augmentation: Building Robust Models

Generating new training samples on-the-fly to prevent memorization.

Rotation

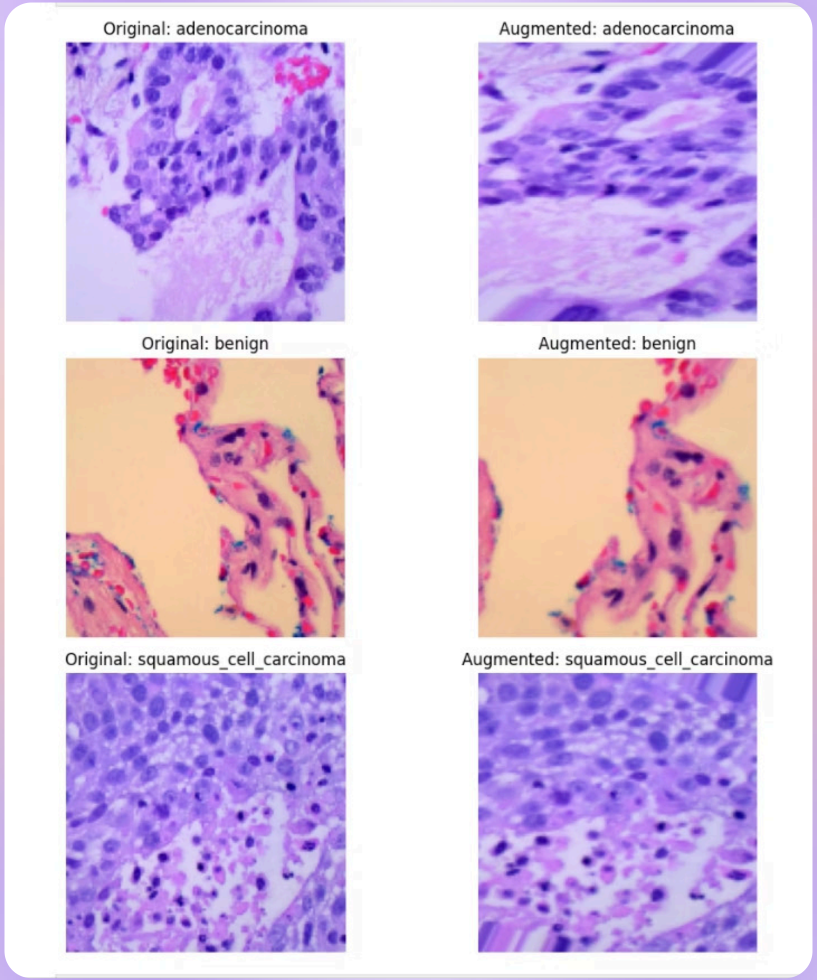
`rotation_range=40`

Zoom Range

`zoom_range=0.4` (mimics scanning distances)

Horizontal Flip

`horizontal_flip=True` (ensures lung invariance)

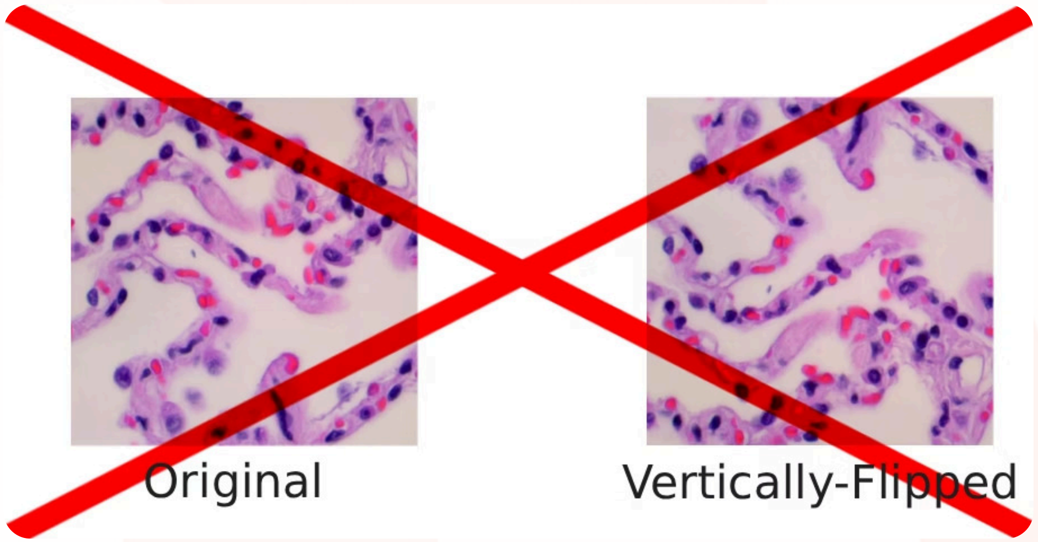


Data Augmentation Limits: Why Less is better

The decision to limit data augmentations for Detective StarLung V1.1 was strategic. In medical imaging—specifically CT scans—"more" is not always "better," as some standard augmentations can actually destroy the medical validity of the image.

Vertical Flip: Why Not?

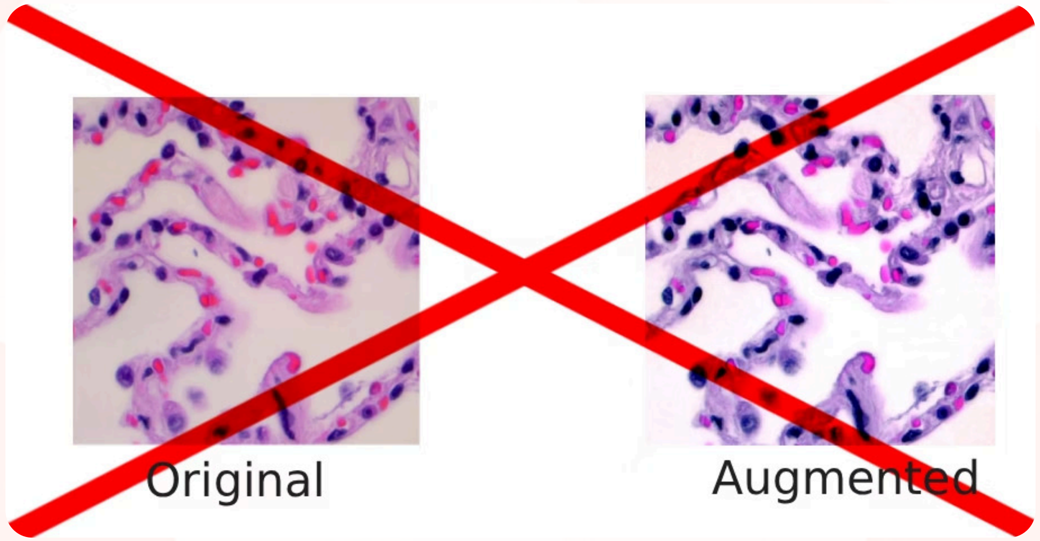
A human lung is never upside down in a scanner. If you teach the model to recognize an upside-down lung, you are training it on a biological impossibility.



Color & Light Adjustments: Why Not?

By looking at the samples, those are H&E (Hematoxylin [purple & blue] and Eosin [pink]) stained histopathology:

- If we apply hue and saturation adjustments, we break this chemical code.
- We need the model to trust that **Purple = DNA** and **Pink = Protein**. Heavy jitter destroys that trust.



```
# TRAINING
from tensorflow.keras.callbacks import EarlyStopping
early_stop_condition = EarlyStopping(monitor='val_loss', patience=3, restore_best_weights=True)
history = model.fit(
    train_generator,
    validation_data=validation_generator,
    epochs=EPOCHS,
    callbacks=[early_stop_condition]
)
model.save('../models/draft_model.keras')
```

Optimized Training Configuration (V1.1)

1

Model Architecture

Maintained successful CNN structure: Conv2D → MaxPooling → Dense

2

Increased Epochs

EPOCHS=40 for complex augmented patterns

3

Early Stopping

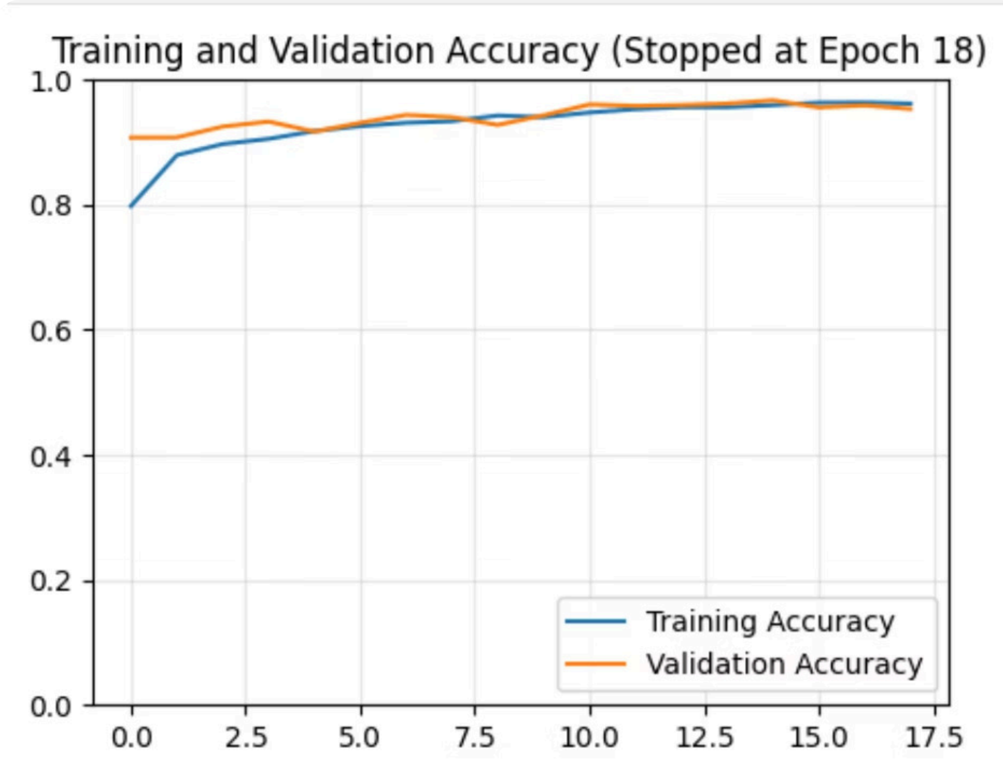
Monitored Validation Loss

4

Patience & Restoration

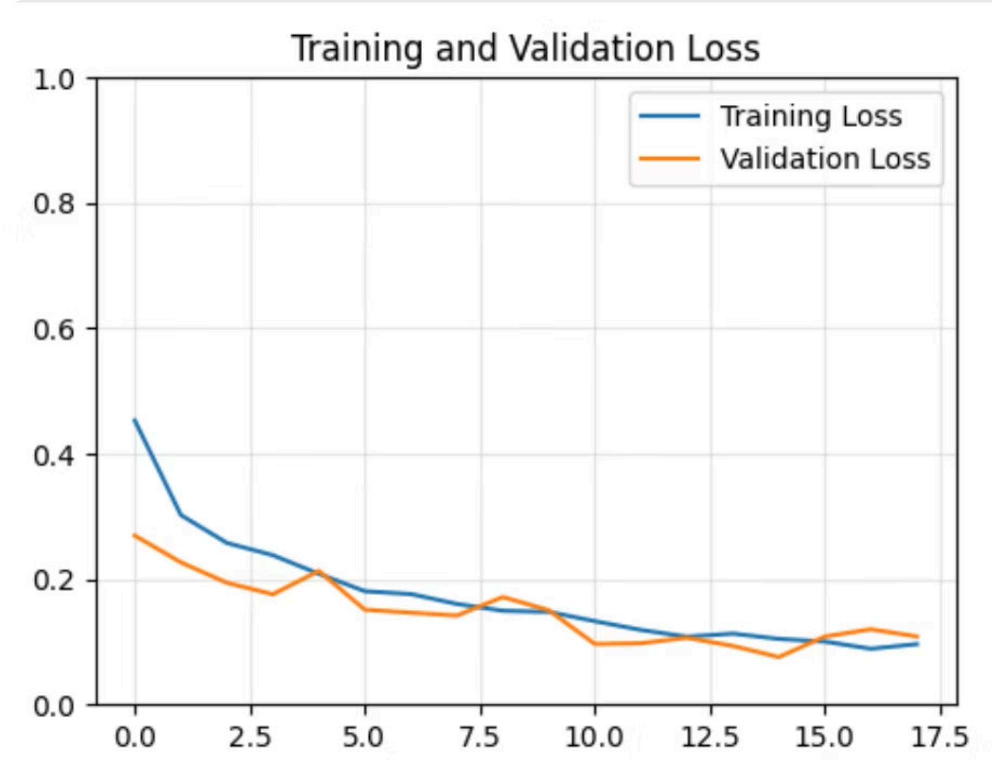
3 epochs patience; `restore_best_weights=True` to prevent overfitting

Performance Curves: Training Dynamics & Convergence



Training/Validation Accuracy

No Overfitting:: The training and validation curves converge beautifully, with validation accuracy rising from 0.9 to ~0.97 and the training curve catching up, until the final epoch—an indication of near-perfect generalization.



Training/Validation Loss

No Divergence: The Orange line (Validation) hugs the Blue line (Training) tightly until the very end.

These curves demonstrate better learning without overfitting, a significant improvement over V1.0.

Early stopping kicked in at exactly the right moment **(Epoch 18)**, preventing the lines from separating.

Confusion Matrix: Enhanced Distinction

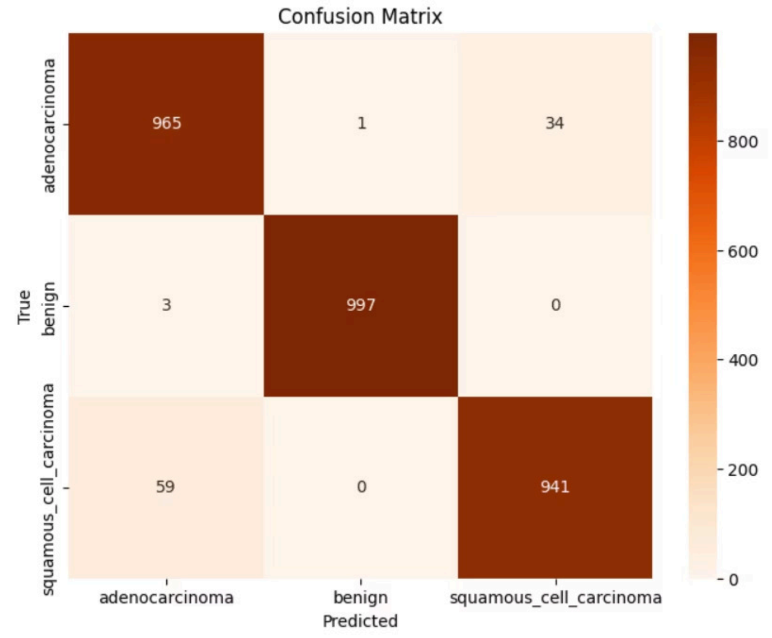
Post-augmentation, the confusion matrix reveals significantly improved performance, especially in differentiating tricky classes.

The Before (V1.0): The model was confused between the two cancer types

- It mistook **81** Squamous cases for Adenocarcinoma.
- It mistook **41** Adenocarcinoma cases for Squamous.

The After (V1.1 - Augmented):

- Squamous mistakes dropped to **59** (Huge improvement).
- Adenocarcinoma mistakes dropped to **34**.
- **Why?** The rotation and zooming forced the model to learn the *texture* of the cells, not just the shape of the lung.



Conclusion: A Robust, Generalized Model

The transition from Baseline (V1.0) to Augmented (V1.1) is complete, yielding a highly effective diagnostic tool.

Reduced Confusion

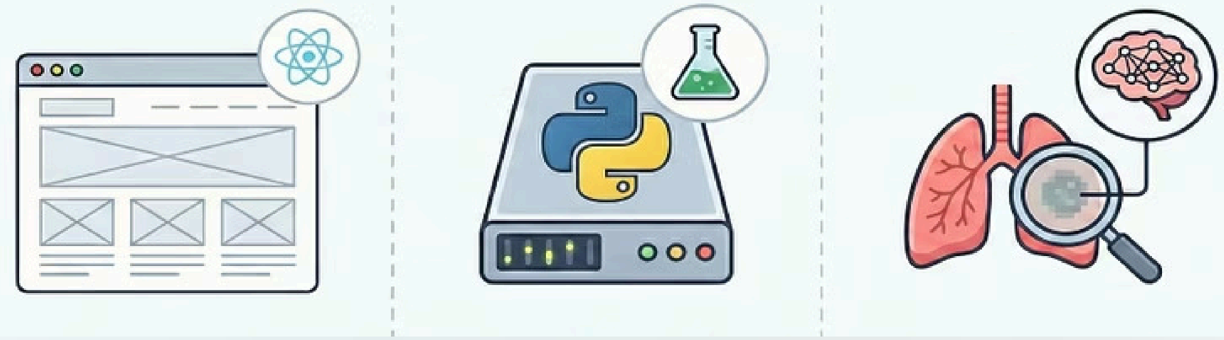
Better distinction between similar cancer types.

Increased Robustness

Less sensitive to image variations.

Ready for Deployment

V1.1 is ready for the next phase.



Next Steps: Real-World Application

1. **Phase 2: Live Demonstration Preparation**

   Deploying the model for a real-world, interactive demonstration.

2. **Frontend Development**

   Utilizing a React App for a responsive and intuitive user interface.

3. **Backend Integration**

   Python Flask to manage routes and send images to the AI model for prediction.

4. **Predicted Image Feedback**

   Model sends back the classified image with predicted cancer type.