

# Computer Vision for Coffee Bean Roast Classification

GTL Workshop – Artificial Intelligence

January 23, 2026

RoastAI: Diego Aguiar, Mariana Sambucetti, Catalina Vaz Martins.

# Quick Question...

The number of coffee shops in big cities is rising – but is this always a good thing?

Share:



If you live in or visit a big city anywhere in the world, chances are you will be spoiled for choice when it comes to finding a specialty coffee shop. Over the last two decades, we have seen a proliferation of cafés and roasters worldwide as international markets have embraced third wave coffee culture.

In capital cities, this can mean thousands of options to decide between, including larger chains and independent outlets. For consumers, endless choice can seem appealing. But for café operators, an increasingly competitive market can be difficult to navigate.

To find out whether there is such a thing as "too many" coffee shops in bigger cities, I spoke to Jonathan Rubinstein, founder and CEO of Joe Coffee Company in New York City.

*You may also like our article on whether offline coffee shops are a dream or a reality.*



Which countries drink the most coffee in the world? Global consumption grows 20% in a decade

The new centre of



dynamic growth in coffee consumption over the last

SOCIETY

## Coffee, in Uruguay, is second only to mate.

Coffee is the third most consumed beverage in the world, behind water and tea.



by Manuel Coccaro · April 14, 2024



# The Coffee Journey

---



# What and how

---

Most people enjoy a good cup of coffee,

**AND** roasting is a key step that creates its flavors and aromas.

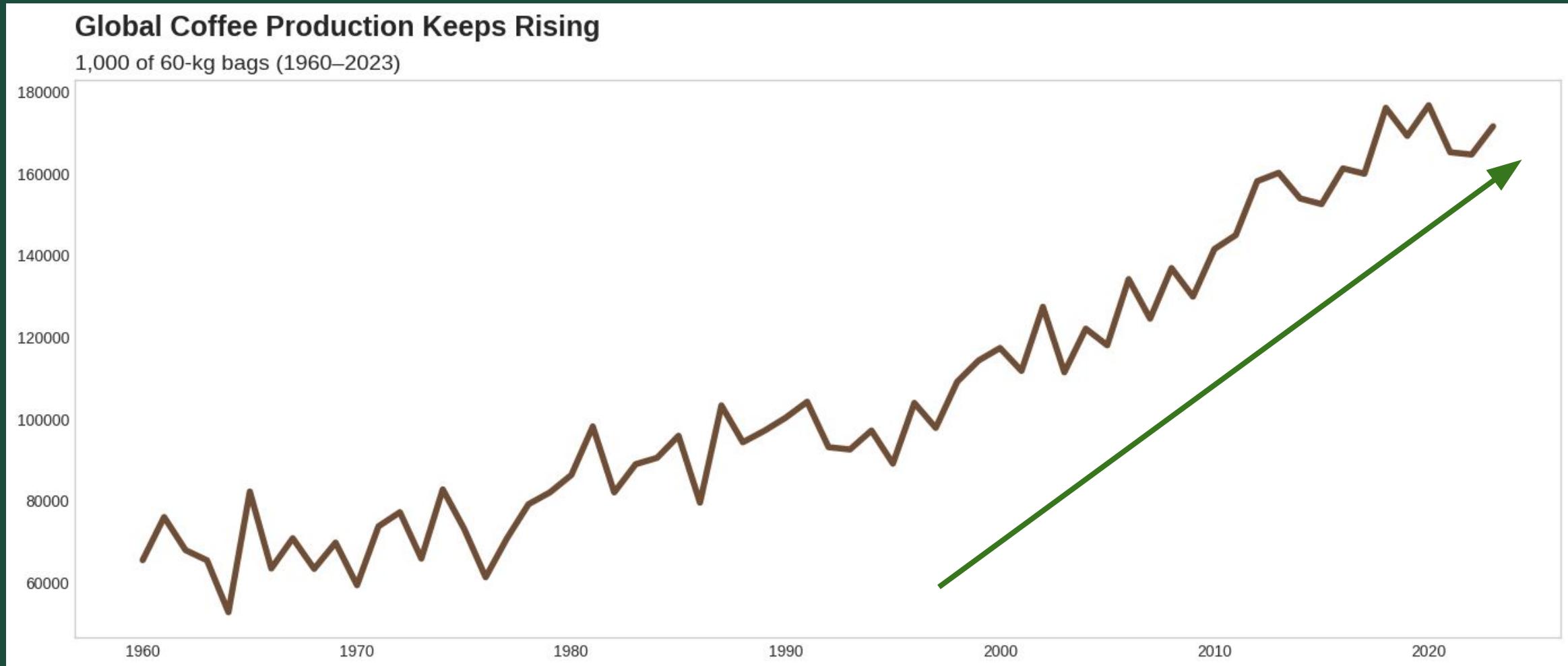
**BUT** identifying the correct roast level is difficult for humans, causing errors and extra costs.

**THEREFORE**, we are developing an automatic classification system to improve accuracy  
and efficiency.

**OBJECTIVE** develop an image-based machine learning  
model that classifies a coffee bean into four distinct roast  
levels to support automated quality control.

# Coffee is a Global Growing Market

---

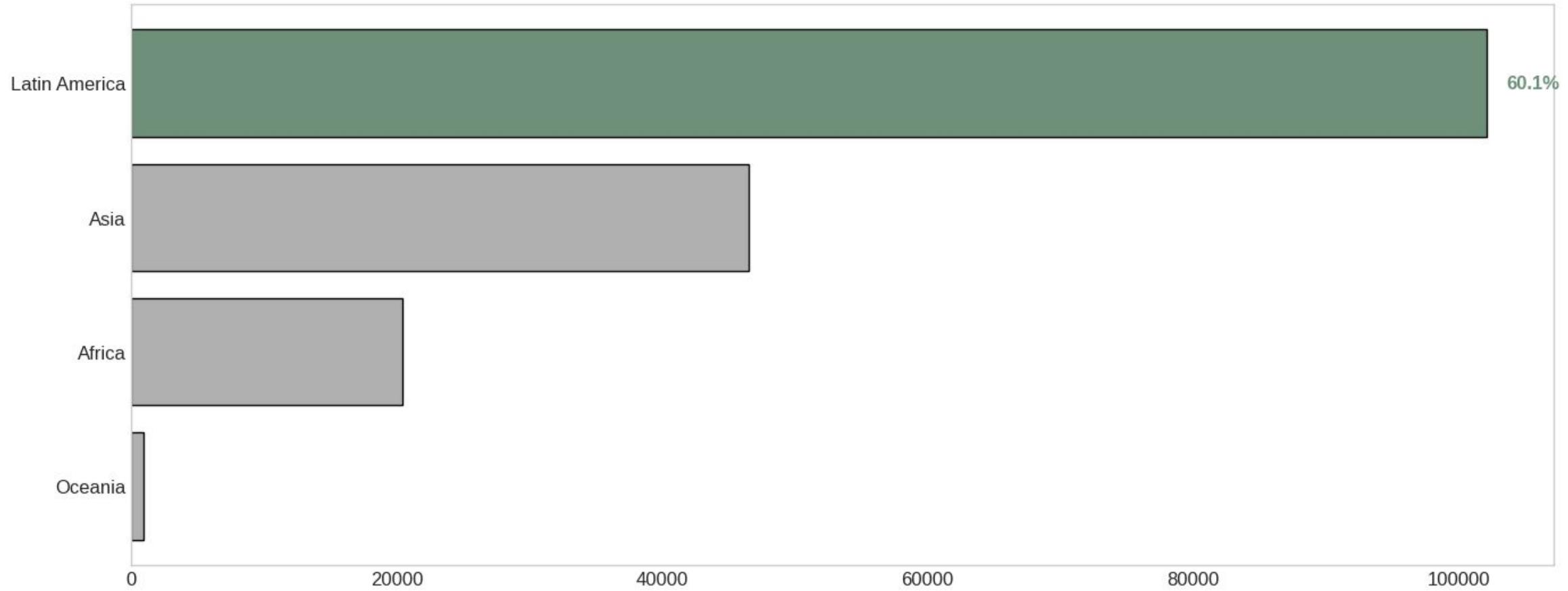


# South America is a Key Coffee Producer

---

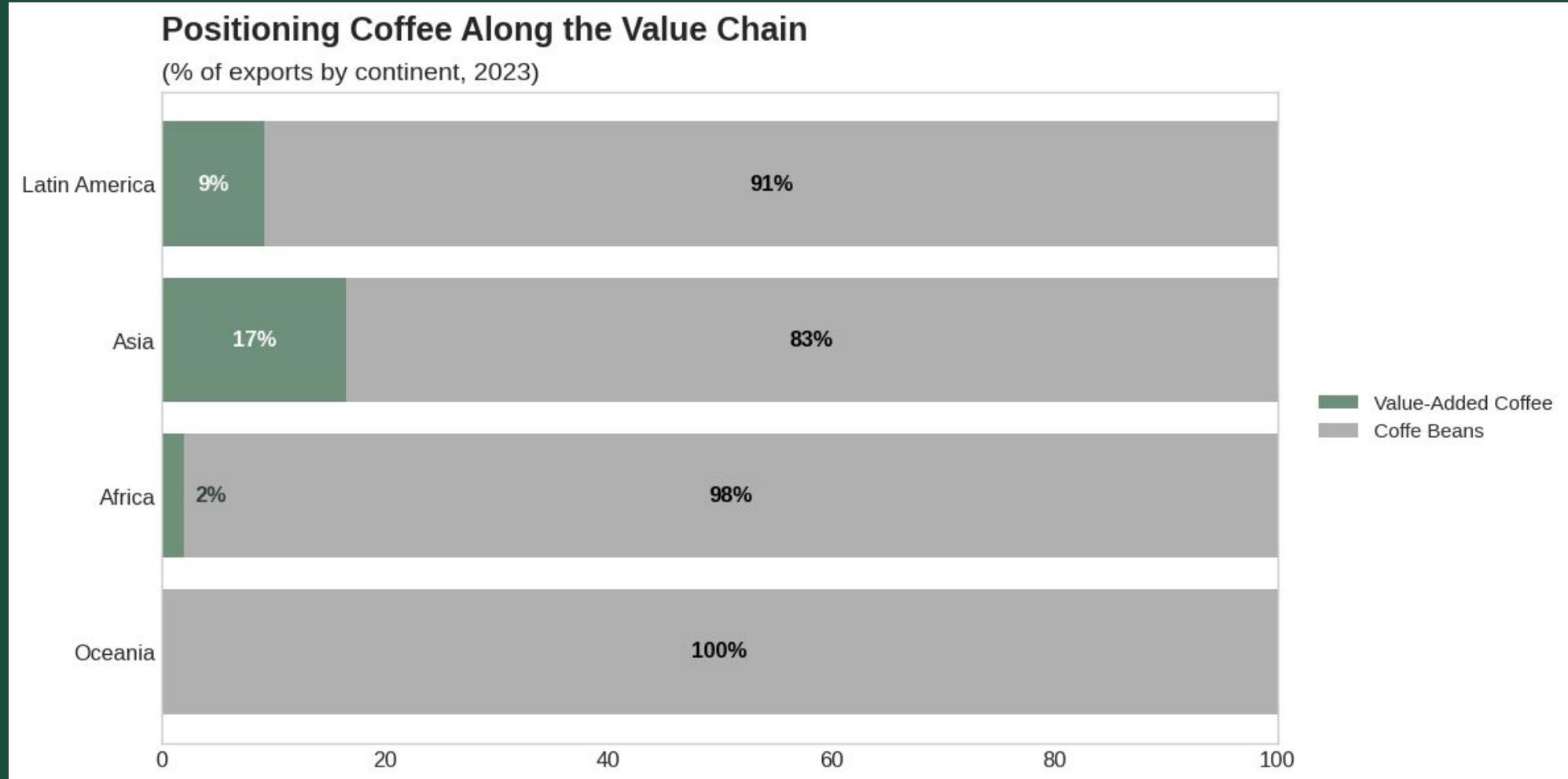
## Production Power Is Concentrated in Latin America

1,000 of 60-kg bags (2023)



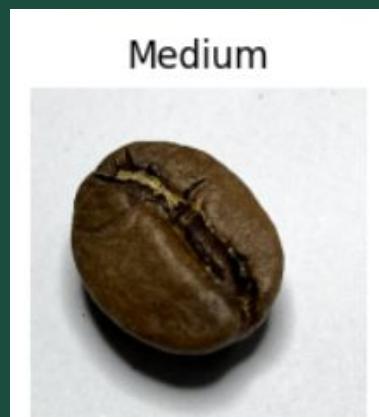
# Quality Differentiation Drives Value

---



# Roasting Levels

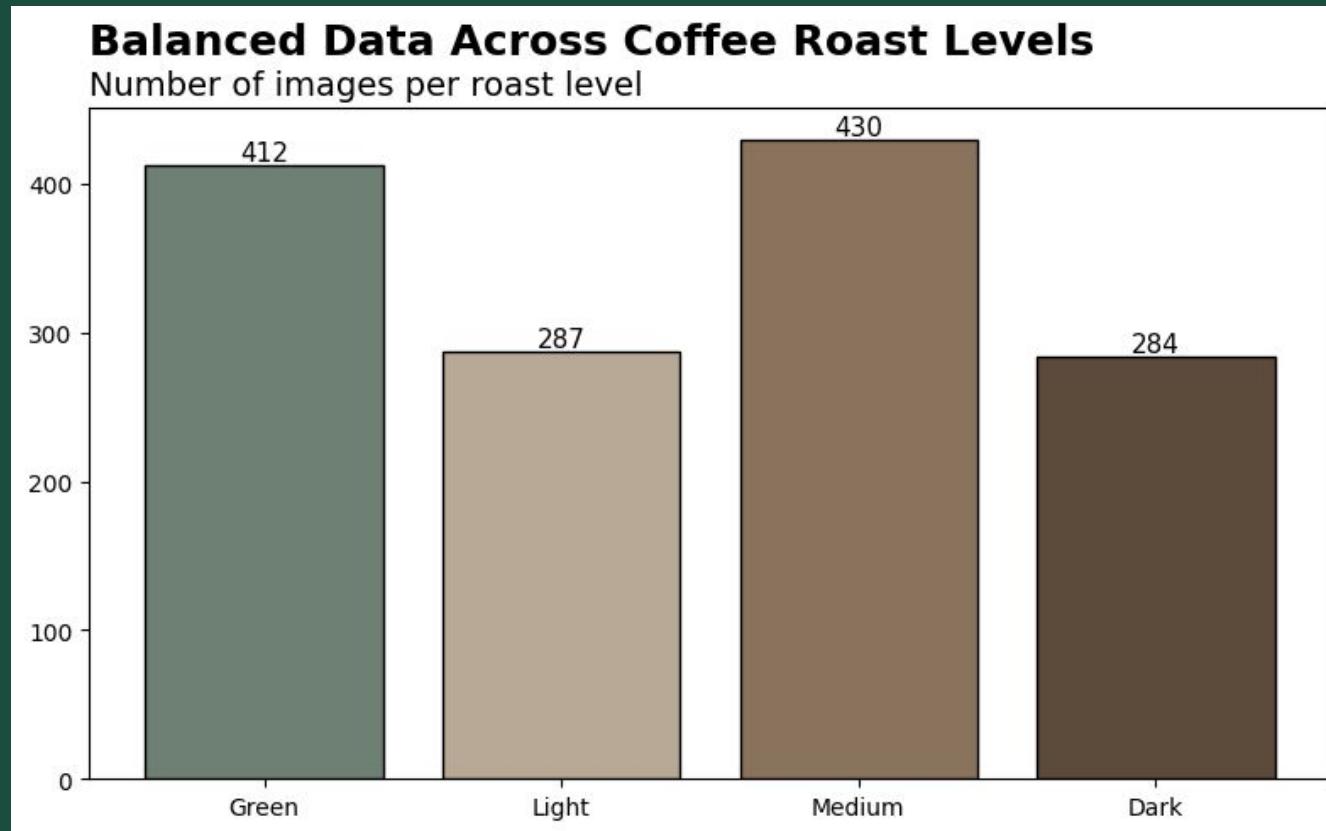
---



- **Green:** Raw beans with green color and no roasting development
- **Light:** Light brown color, higher acidity, origin flavors preserved
- **Medium:** Balanced flavor, moderate acidity, fuller body
- **Dark:** Dark brown color, bold flavor, low acidity, oily surface

# Our beans images

---



- 1,768 (1,413 Training) coffee bean images across 4 roast levels
- Data sources:
  - Public dataset
  - Additional images collected by our team
- Impact
  - Improved robustness and generalization
  - Better alignment with real-world usage

# Teaching the Model How to See

---



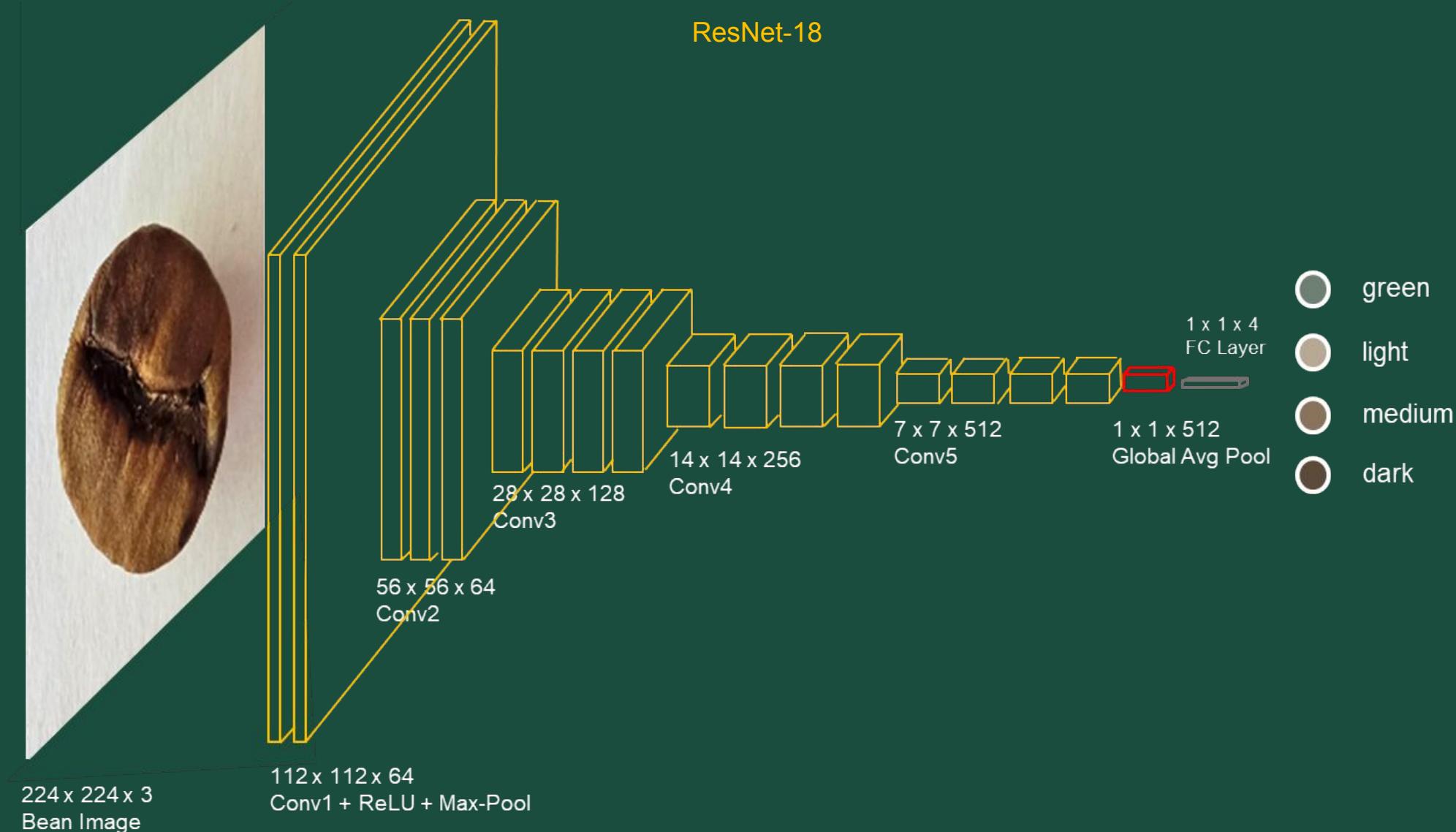
What we have

- Raw RGB images of coffee beans
- No handcrafted features
- No roast-specific rules

How we prepare the model input

- Resized to  $224 \times 224$
- Normalized with ImageNet statistics
- Data Augmentation

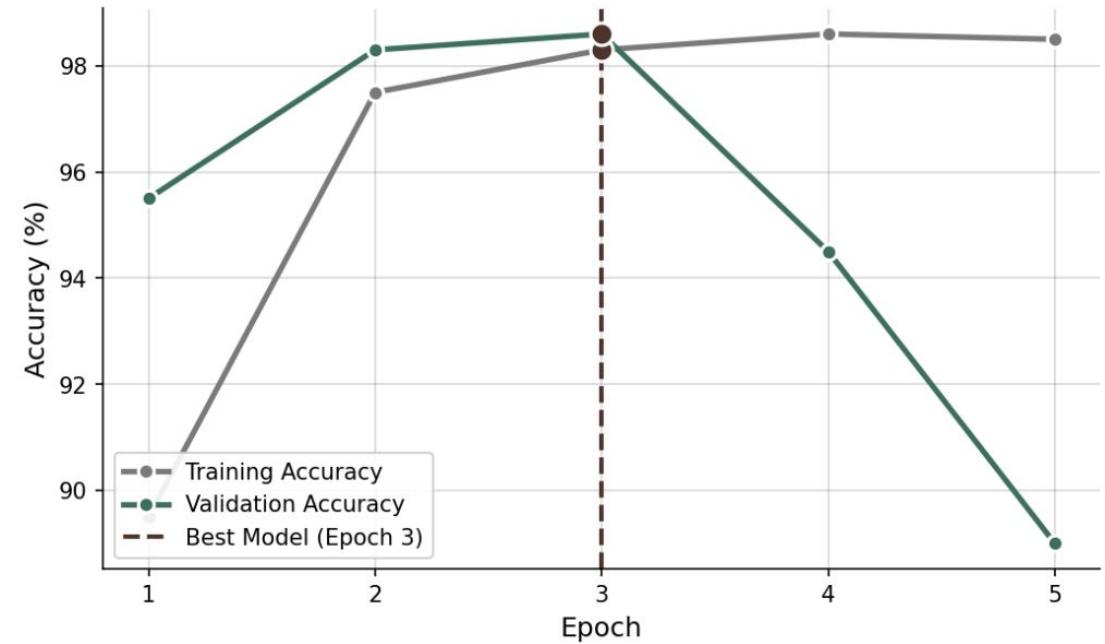
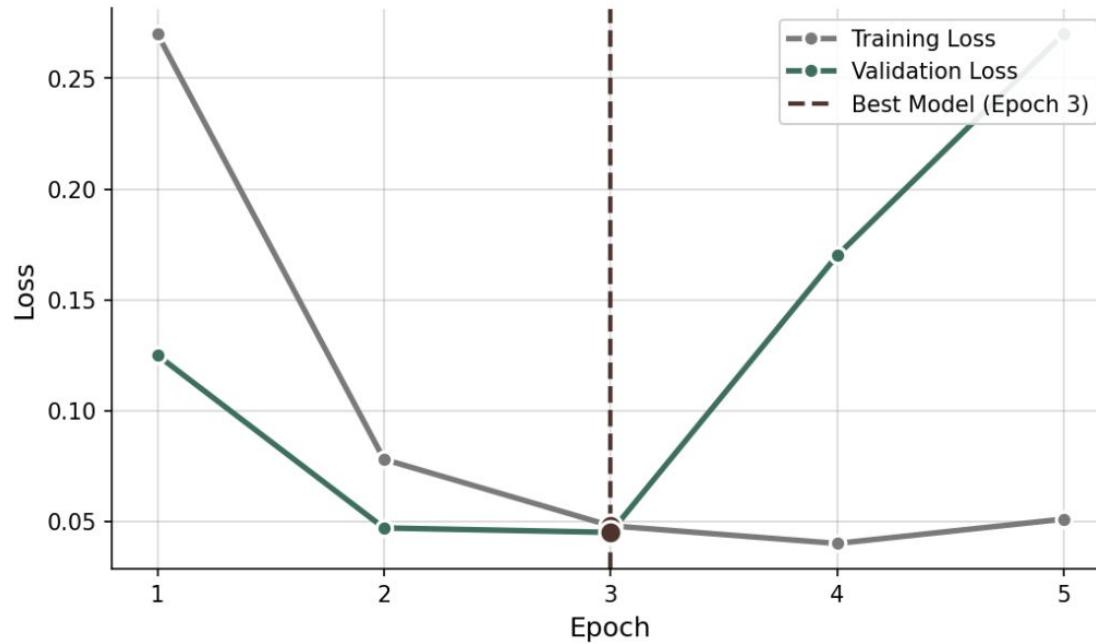
# Choosing and Training the Model



# Hyperparameter Tuning & Training Strategy

## Detecting Overfitting with Early Stopping at Epoch 3

Validation loss rises after epoch 3 while training loss continues to drop, signaling overfitting

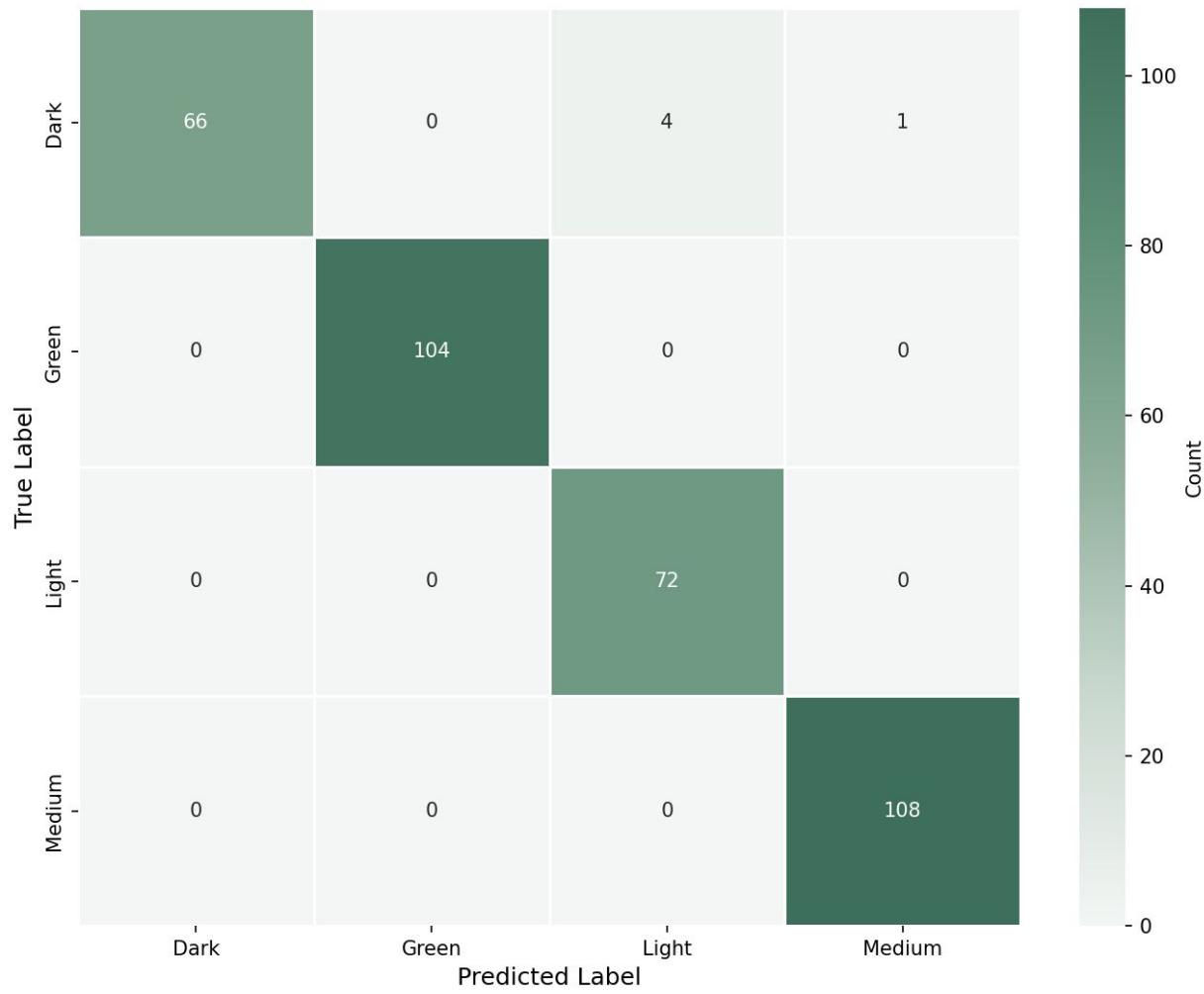


- Optimizer: Adam, lr = 1e-4
- Max epochs: 15 with Early Stopping (patience=2, min\_delta=0.001)
- Stopped at epoch 5. Best model: epoch 3 (lowest val loss)

# Measuring and Proving Model Performance

## Classifying Coffee Roast Levels with High Accuracy

Model achieves 98.6% accuracy across 4 roast categories (n=355)



- Test set size: N = 355 (20% holdout)
- Per-class recall (to show robustness):
  - Dark: 0.93
  - Green: 1.00
  - Light: 1.00
  - Medium: 1.00
- Overall accuracy: 0.99
- Macro F1: 0.98

# Conclusion

---

- ML model can classify coffee roast levels
- It is a scalable alternative to manual inspection
- Strong potential for quality control
- Limitations:
  - Images captured under controlled conditions
  - Model may need adaptation for different camera setups

# Next Steps

---

- Develop a mobile application to capture and upload coffee bean images
- Integrate the trained model for real-time roast level prediction
- Deploy the system for pilot testing in real roasting environments