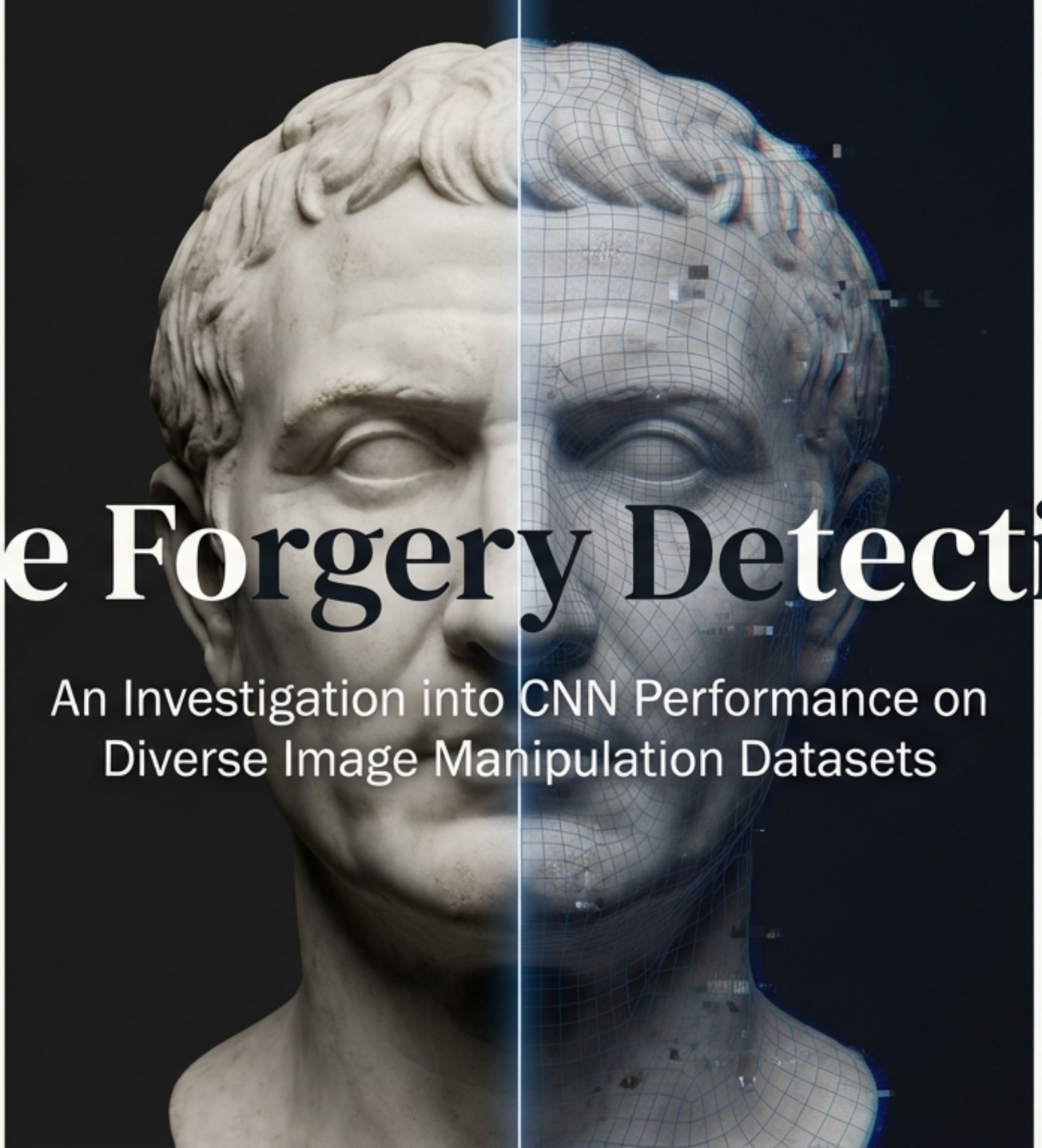


The Forgery Detective

An Investigation into CNN Performance on Diverse Image Manipulation Datasets

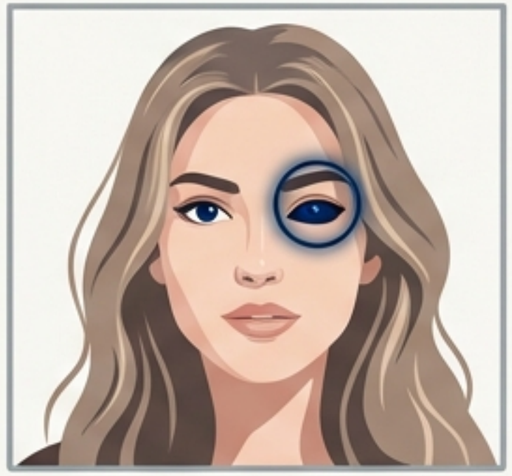
NotebookLM

The Challenge: A Crisis of Authenticity

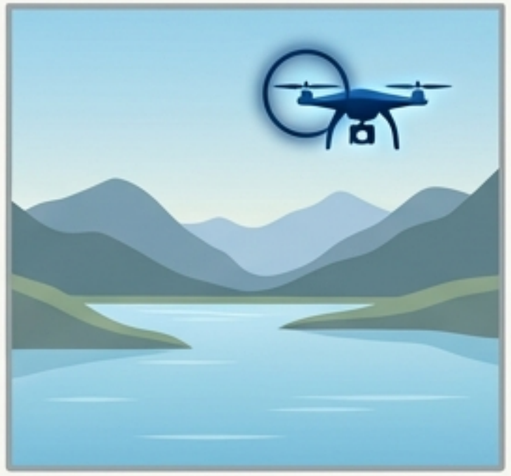
In an era of advanced digital tools, distinguishing authentic from manipulated images is a critical challenge. Our objective is to test a single, robust deep learning model's ability to act as a universal forgery detective.

Authentic Portrait

Authentic Landscape



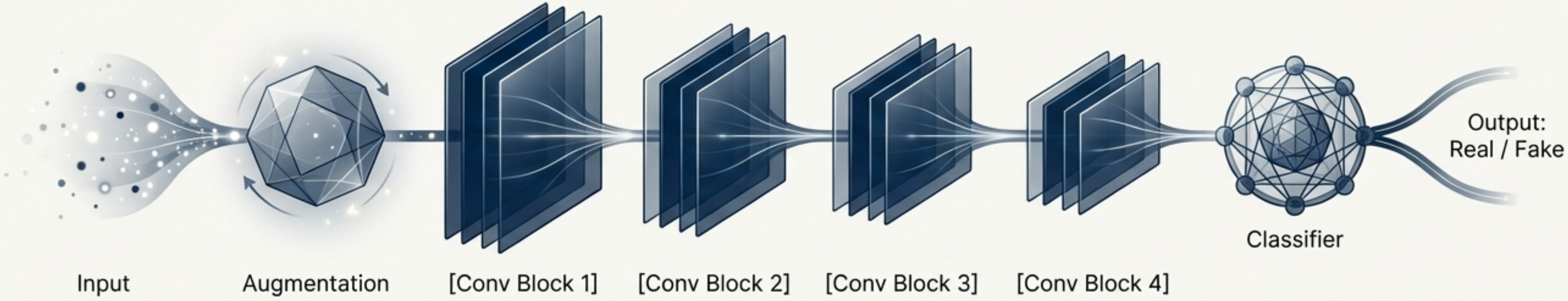
Manipulated Portrait



Manipulated Landscape

NotebookLM

The Tool: A Standardized CNN Architecture



We engineered a consistent CNN architecture with four convolutional blocks designed to learn image features. This **34-million-parameter** model serves as our single analytical tool across all experiments.

NotebookLM

The Proving Grounds: Four Distinct Datasets

To rigorously test our model, we used four separate datasets, each containing thousands of “real” and “fake” images. These datasets represent different types and qualities of image manipulation.

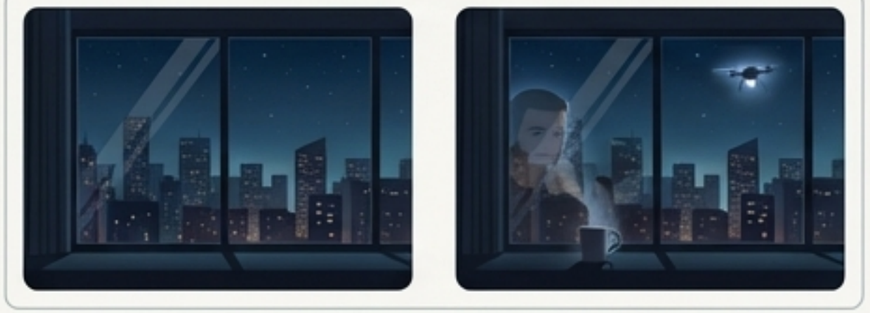
Dataset 1

Real Fake

Dataset 2

Real Fake

Dataset 3



Real Fake

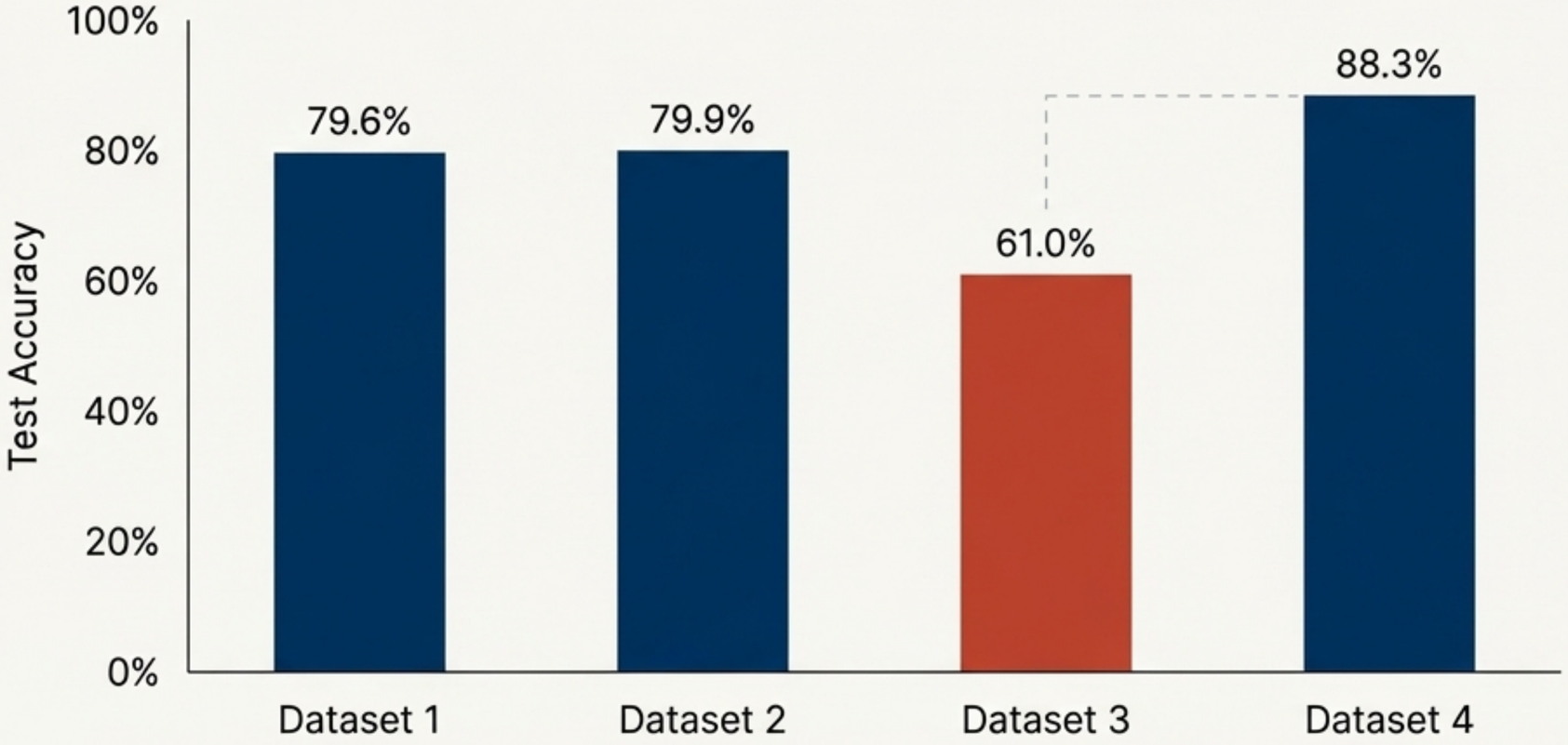
Dataset 4

Real Fake

NotebookLM

Initial Findings: A Story of Extreme Variance

Running the same model against each dataset yielded dramatically different outcomes. Performance ranged from an impressive 88.3% accuracy down to a surprising 61.0%, revealing high sensitivity to the data source.



NotebookLM

The High Achiever: Success on Dataset 4

The model achieved its peak performance on Dataset 4, reaching 88.3% test accuracy. The specific manipulations in this dataset aligned exceptionally well with the features our CNN learned to identify.

88.3%

Overall Test Accuracy

Precision (Fake)	Recall (Fake)
85%	93%

NotebookLM

The Anomaly: A Struggle with Dataset 3

Conversely, the model struggled significantly with Dataset 3, with accuracy falling to just **61.0%**. This indicates the forgeries in this collection possess artifacts that confused our model's feature detection.

61.0%

Overall Test Accuracy

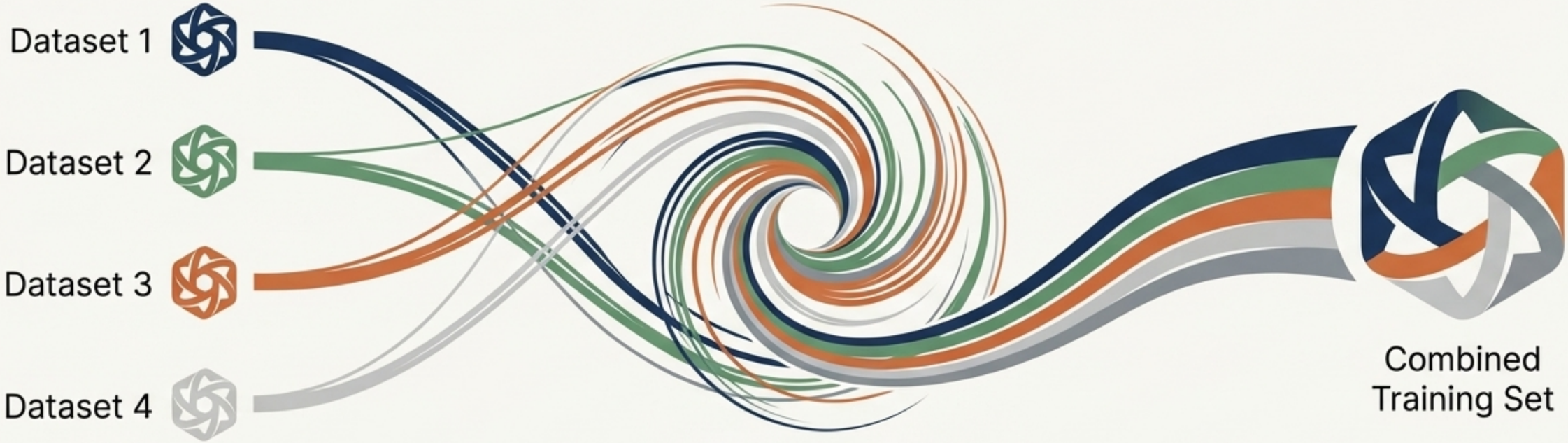
Precision (Fake)	Recall (Fake)
63%	54%

NotebookLM

New Hypothesis: Strength in Diversity

The performance variance suggests that specialization has its limits.

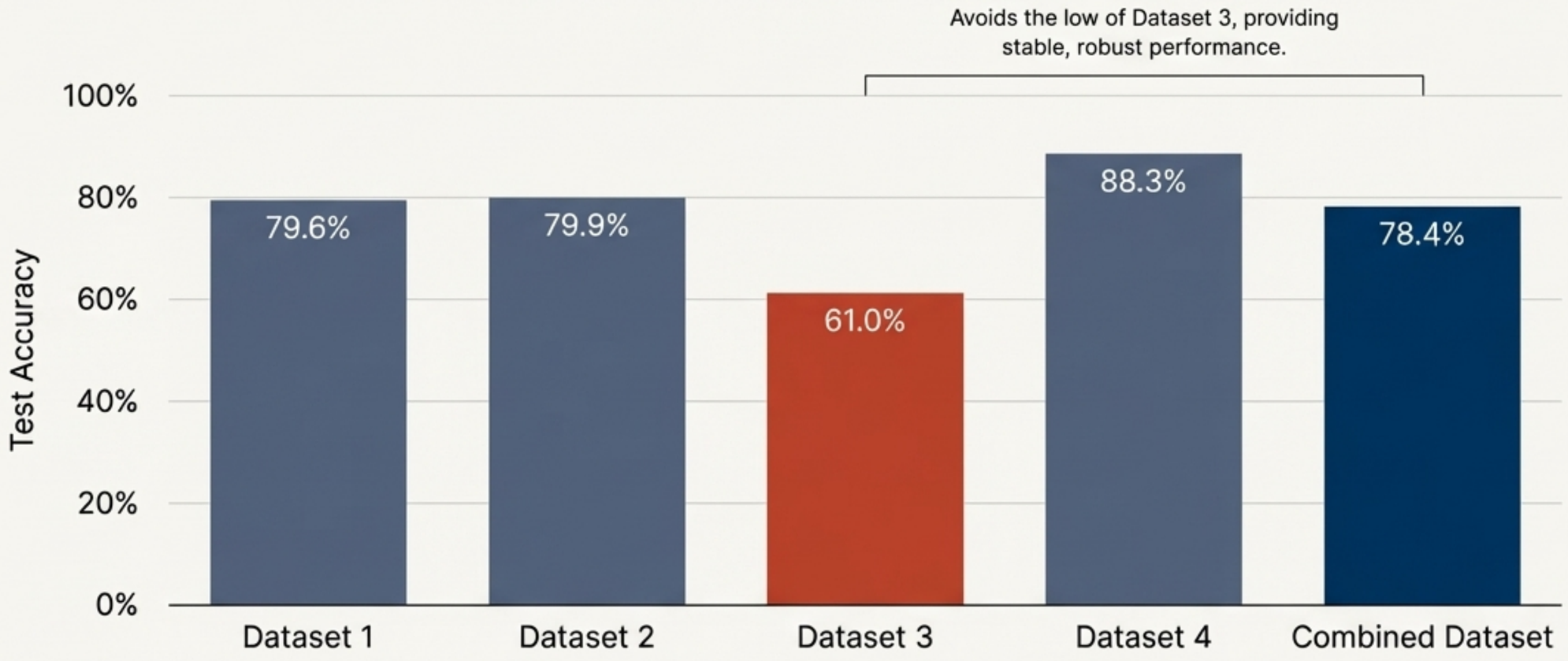
Our new hypothesis: could a model trained on all four datasets combined learn a more generalized and robust representation of forgeries?



NotebookLM

The Generalist: A Robust, Unified Model

When trained on the merged dataset, the model achieved a test accuracy of 78.4%. While not outperforming the top specialist, it created a reliable baseline, effectively navigating the complexities of all four data types.



NotebookLM

Conclusion: Data Is the Defining Factor

This investigation proves that for image forgery detection, the model architecture is only part of the story. The diversity and nature of the training data are the most critical factors in determining real-world performance and robustness.

NotebookLM