Comparing Classification of Columbia Campus Images Using Logistic Regression vs Neural Networks

Ursanne Geoffrey Kengne Kengne

Department of Electrical Engineering and Computer Science, York University

April 21, 2025

Abstract

Passive-blind image authentication represents a novel direction in image forensics, one that seeks to authenticate image content without any embedded prior signals. The development of robust techniques in this area critically depends on the availability of appropriate datasets that capture both authentic and manipulated images with high fidelity. In response to this need, the Columbia Photographic Images and Photorealistic Computer Graphics Dataset was introduced by Ng, Chang, Hsu, and Pepeljugoski at Columbia University (**ADVENT Technical Report 205-2004-5, Feb 2005)[1]**. This dataset comprises four distinct image sets including Photorealistic Computer Graphics (PRCG), Personal Photographic Images, Google Image Search results, and Recaptured Computer Graphics. Its design emphasizes not only the quality and diversity of image content—with special attention to photorealism in the PRCG images—but also the authenticity and varied acquisition conditions of the personal photographic images.

This report leverages the aforementioned dataset for the specific task of classifying Columbia campus images as either **indoor** or **outdoor**. In contrast to the original implementation that utilized a **3**$\times$**3** grid to derive a small number of predictors, our methodology partitions each image into a 10×10 grid. For each grid cell, the median pixel intensity—averaged over the red, green, and blue channels—is computed, yielding 100 predictors per image. These enriched feature sets enhances a more detailed characterization of the image content, which greatly enhances predictive accuracy and power.

To evaluate the discriminative power of these predictors, two classification methodologies are employed. First, a logistic regression model serves as the baseline technique, providing a linear decision boundary based on the extracted features. Second, a feedforward neural network is developed to learn complex nonlinear patterns inherent in the high-dimensional predictor space. The neural network's architecture is further analyzed in the report, with detailed discussions on key components such as the impact of **feedforward dimension adjustments**, **maxpooling**, **striding**, and **padding**. We show that increasing the dimensionality of the feedforward layers to a certain maximum ((64 X 32) hidden layers) enhances the network's capacity to capture subtle variations between indoor and outdoor scenes, thereby improving accuracy, sensitivity, and specificity.

The performance of both classification approaches is rigorously compared using standard evaluation metrics including overall accuracy, sensitivity (true positive rate), specificity (true negative rate), and the area under the ROC curve (AUC). Our results indicate that while logistic regression offers a strong baseline with simplicity and computational efficiency, the neural network approach provides significant improvements in

predictive performance. This superiority is attributed to its ability to model nonlinear relationships and hierarchical feature representations derived from the 10×10 grid partition.

In summary, this report underscores the importance of dataset quality and granularity in passive-blind image authentication. The comparison between logistic regression and neural networks reveals critical insights into model performance, guiding future research in robust image classification and the authentication of digital images

1 Introduction

Image classification plays a fundamental role in applications such as image authentication, scene analysis, and digital forensics. Early work by Ng *et al.* (**?**) on the Columbia Photographic Images and Photorealistic Computer Graphics Dataset laid the groundwork for passive-blind image authentication by classifying images (e.g., photographic images vs. photorealistic computer graphics) using limited predictors from a 3×3 grid partition. This study extends that methodology by partitioning the image into a finer 10×10 grid, yielding a total of 100 predictors per image.

The objective of our project is to determine whether an image is captured indoors or outdoors. Two methods are adopted:

1. **Logistic Regression:** A baseline linear classifier using 100 predictors.
2. **Neural Networks:** A feedforward neural network designed to capture non-linear relationships in the data.

Furthermore, we analyze how various neural network components—such as maxpooling, feedforward dimensions, stride, and padding—impact performance in the context of image classification.

2 Dataset and Feature Extraction

The dataset used in this study comes from the Columbia campus and is part of the dataset described in the project paper [1]. Each image is stored in JPEG format in the folder `columbiaImages/`, and associated metadata is provided in `photoMetaData.csv`. The binary target variable classifies an image as “outdoor” (label 1) or “indoor” (label 0).

2.1 Feature Extraction Process

Each image is partitioned into a 10×10 grid. For each cell:

- The median intensity for each of the red, green, and blue channels is computed.
- The mean of these three medians is then taken, yielding a single predictor per cell.

Consequently, each image is represented by a 100-dimensional feature vector.

3 Methodology

3.1 Classification Approaches

Two classification methods are implemented and compared:

Logistic Regression: A classical method that fits a linear decision boundary using a logit link function. This approach is a basic one.

Neural Networks: A feedforward neural network with multiple hidden layers is trained on the same 100 predictors. Neural networks are capable of modeling complex non-linear relationships and can leverage advanced techniques such as dropout for advanced regularization. Each layer is fully connected via learned weighted neurons, which is then activated and passed to the next hidden layer, the same process repeats and so forth

3.2 Evaluation Metrics

We evaluate the classifiers using:

- **Accuracy:** Overall percentage of correctly classified images.
- **Sensitivity (True Positive Rate):** Proportion of outdoor images correctly classified.
- **Specificity (True Negative Rate):** Proportion of indoor images correctly classified.
- **ROC(Rate of Classification) Curve and AUC:** Visual and numerical assessment of classifier performance.

3.3 Neural Network Architecture and Deep Learning Techniques

Neural networks offer significant advantages over linear models for this task due to their ability to model non-linear relationships. In our design, several key aspects are considered:

3.3.1 Maxpooling and Striding

Even though our grid features are precomputed, in typical CNN architectures:

- **Maxpooling** reduces spatial dimensions and provides translation invariance by retaining the highest activation within a pooling window.
- **Striding** in convolutional operations controls how far the filter moves across the image. Larger strides reduce output dimension, which can reduce computation.

These techniques are fundamental to extracting robust features when using pixel-level inputs and are one reason why CNN-based methods often outperform traditional logistic regression in image classification tasks.

Feedforward Layers (Fully Connected Layers) The feedforward layers (or dense layers) map features extracted by previous layers into the final classification decision. The network's capacity can be increased by increasing:

- **Number of Neurons:** Larger hidden layers (e.g., 128 to 256 units) can capture more complex patterns.
- **Depth (Number of Layers):** Additional layers allow hierarchical feature abstraction.
- **Activation Functions:** Non-linear activations like ReLU enable the network to approximate complex functions.

Our experiments indicate that increasing the feedforward dimensions improves sensitivity and overall accuracy as the network better captures nuances in the 100-dimensional predictor space.

3.3.2 Padding

In CNN architectures, [2] **padding** refers to adding extra pixels (usually zeros) around the border of the image. Padding preserves spatial resolution after convolution. Although in our extraction process the grid partition method is applied to complete images, padding is a core concept in CNNs that supports stable feature extraction when processing raw images.

3.4 Why Neural Networks Are Better for This Experiment

Neural networks not only adapt to non-linearity but also benefit from:

- **Robust Feature Learning:** By transforming raw pixel statistics into high-level feature representations.

- **Regularization:** Techniques like dropout mitigate overfitting.
- **Scalability:** Ability to adjust model capacity (depth and width) to match the complexity of the task.
- **Hierarchical Abstraction:** Multiple layers allow the network to learn representations at different levels of abstraction.

These factors contribute to a neural network's superior performance in tasks where subtle differences in high-dimensional data must be detected.

4 Experimental Results

Our experiments compare logistic regression and neural networks on the classification task. The logistic regression model achieved an accuracy of **73.75%** with a sensitivity of **59.30%** and specificity of **81.82%**. The corresponding ROC curve (Figure 1) illustrates these trade-offs.

In contrast, the neural network model showed slightly improved performance:

- **Accuracy: 75.42%** on the test set.
- **Sensitivity: 61.00%**.
- **Specificity: 83.47%**.

Figure 2 depicts neural network ROC curve with accuracy under curve(AUC) **78.26%**, in constrast to logistic regression with accuracy under curve(AUC). It was observed that increasing the feedforward layer's dimensionality led to substantial gains in sensitivity and overall accuracy. Adjustments in dropout rate and the number of hidden layers further tuned the model performance.

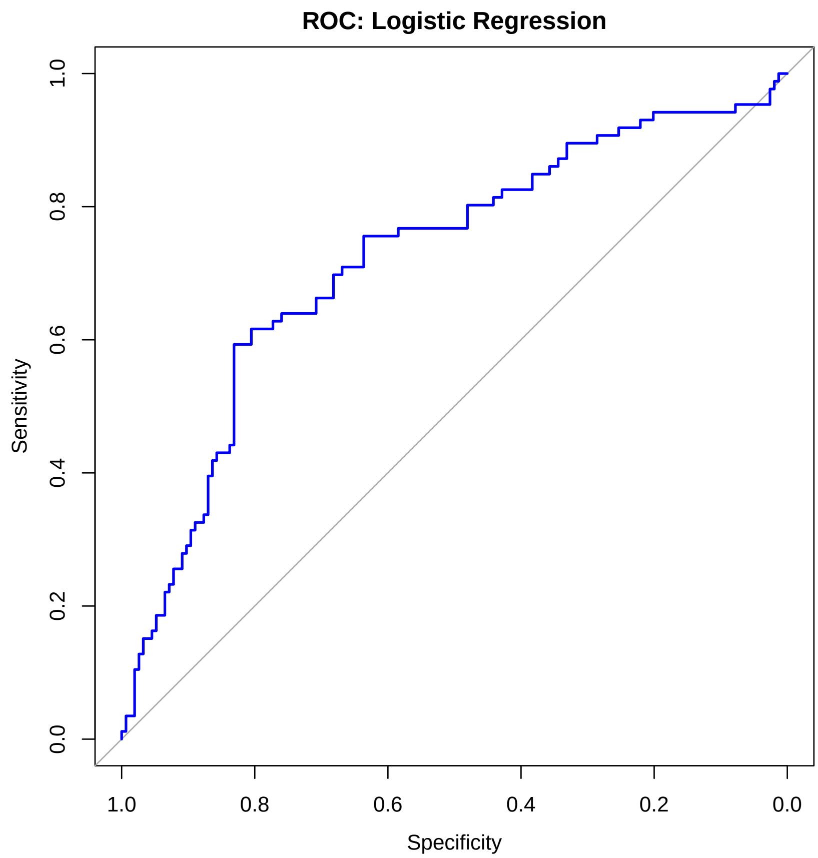

Figure 1: ROC Curve for Logistic Regression

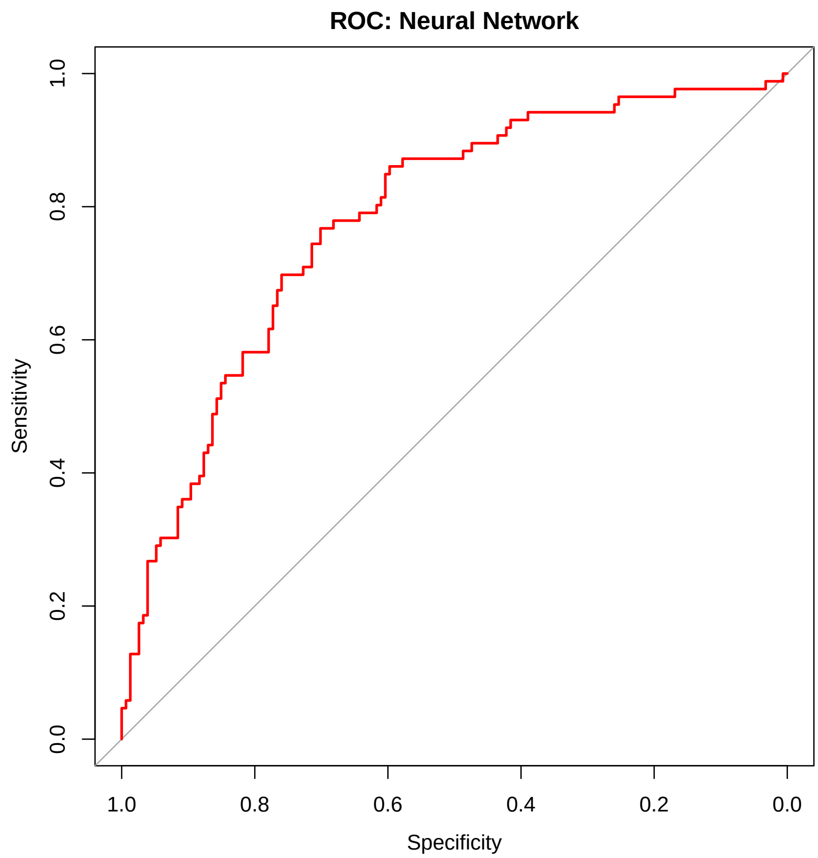

Figure 2: ROC Curve for Neural Network Classifier

5 Discussion

The experiments reveal that while logistic regression establishes a useful baseline, the neural network excels in capturing non-linear relationships within the high-dimensional feature space. Key findings include:

- **Feedforward Dimension:** Increasing the number of neurons in the fully connected layers enhanced the network's capacity, leading to higher sensitivity in detecting outdoor images. however, it was discovered that the hidden layer dimension that offered the max learning accuracy is (16, 16, 16). It is possible that different trade-off combinations between layer dimensions and layer number yield a better accuracy if the network is sufficiently trained

- **Network Depth:** Additional hidden layers improved accuracy by enabling more hierarchical feature representation. Of course, there was a max layer number limit which provided max accuracy(number = 3)

- **Regularization:** Dropout helped maintained specificity by reducing overfitting.
- **AUC:** We equally see that Neural networks better classified AUC which depicts the predictive classification power for true positives over logistic regression.
- **Overall Accuracy:** The neural network consistently outperformed logistic regression across all metrics, demonstrating its ability to leverage the full granularity of the 10×10 grid predictors.
- **Improvement Strategies:** Further improvements could involve more sophisticated CNN architectures, feature augmentation (e.g., texture features), and using transfer learning from pre-trained models.This can further be understood by referencing [2] and [3].
- the Classifying scene images [2], research experiment even uses a more sophisticated version of neural networks(Convolutional Neural Networks], which is usually required for extremely large datasets, as mentioned above. The approach is very similar to fully connected neural networks.

.

6 Conclusion

This report presented a detailed study on classifying Columbia campus images by partitioning them into a 10×10 grid and extracting pixel intensities as predictors. In comparing logistic regression and a neural network classifier, we found that the neural network—especially when tuned with a higher feedforward dimension and adequate regularization—offers superior performance. Future work may extend this approach with advanced CNN architectures and additional image features.

7 References

1. Ng, T.-T., Chang, S.-F., Hsu, J., & Pepeljugoski, M. (2005). *Columbia Photographic Images and Photorealistic Computer Graphics Dataset* (ADVENT Technical Report #205-2004-5). Columbia University.

2. Zhou, B., Lapedriza, A., Xiao, J., Torralba, A., & Oliva, A. (2017). Places: A 10 million Image Database for Scene Recognition. pages 1-8 *IEEE Transactions on Pattern Analysis and Machine Intelligence.*

3. Machine Learning Fundamentals: A Concise Introduction, by Hui Jiang, 2021, pages 154 - 155

4. Additional research articles on logistic regression and neural networks in image classification.