Image Classification Using Graph Convolutional Networks

By

Achyuth Pothuganti, Phanindra Chowdary Borra, Venkata Sai Veyyagandla, Surya Prakash Bandam

1. Introduction

Image classification is a crucial task in computer vision with applications in environmental monitoring and urban planning. Traditional convolutional neural networks (CNNs) excel at feature extraction but fail to capture relational and structural dependencies in data. Graph Neural Networks (GNNs), especially Graph Convolutional Networks (GCNs) introduced by Kipf and Welling in 2017, address this limitation by processing graph-structured data. This project leverages GCNs for natural scene classification using graph representations derived from Delaunay Triangulation and SLIC-based super pixel segmentation. The Intel Image Classification dataset is used to evaluate the accuracy, efficiency, and interpretability of these methods.

2. Related Work

Graph Neural Networks are effective for analyzing non-Euclidean data like social networks, molecular graphs, and images. GCNs have been widely applied to tasks such as node classification, graph classification, and link prediction. Recent advancements demonstrate their utility in image-based tasks by converting pixel data into graph structures. Delaunay Triangulation and SLIC-based segmentation are prominent methods for graph construction, offering different trade-offs in granularity and efficiency. This project evaluates their performance on the Intel Image Classification dataset, a standard benchmark for image-based GCNs.

3. Methodology

3.1 Dataset

- **Source**: Intel Image Classification Dataset (Kaggle)
- **Images**: 24,000 (150x150 resolution)
- **Categories**: Buildings, Forest, Glacier, Mountain, Sea, and Street
- **Data Split**: 14,000 training, 3,000 testing, 7,000 prediction

3.2 Preprocessing

- **Image Resizing**: Standardized to 32x32 pixels.
- **Feature Extraction**:
    - **Delaunay**: RGB pixel values as node features.

- **SLIC**: Mean color, eccentricity, solidity, aspect ratio, and perimeter for superpixels.

3.3 Graph Construction

1. **Delaunay Triangulation**
   - **Nodes**: Pixels
   - **Edges**: Constructed using Delaunay triangulation.
   - **Node Features**: RGB values.
2. **SLIC Superpixel Segmentation**
   - **Nodes**: Superpixels generated by SLIC algorithm.
   - **Edges**: Region adjacency graph (RAG).
   - **Node Features**: Regional attributes like mean color and shape descriptors.

3.4 Model Architecture

- **GCN Layers**: Graph convolution layers with BatchNorm and ReLU activation.
- **Pooling**: Global mean pooling for graph-level feature aggregation.
- **Classification**: Fully connected layers with dropout for regularization.

Training Configuration:

- **Optimizer:** Adam
- **Loss Function:** CrossEntropy
- **Learning Rate:** 0.001
- **Batch Size:** 32
- **Epochs**: 50 (Delaunay), 100 (SLIC)

4. Results and Discussion

4.1 Results

Metric	Delaunay Triangulation	SLIC Segmentation
Preprocessing Speed	Faster	Slower
Training Speed	Slower	Faster
Test Accuracy	63.57%	73.0%
Strong Classes	Glacier, Street	Forest, Street
Challenges	Class overlap (Sea/Glacier)	Fewer overlaps

4.2 Comparative Analysis

Compared to CNNs like ResNet and VGG, which achieve higher accuracy (~85-90%) on image datasets, graph-based methods like GCNs provide greater interpretability by explicitly modeling relationships. While CNNs excel in feature extraction, they lack the ability to process graph-structured data. Alternative graph construction methods like k-Nearest Neighbors (simple but less interpretive) and fully connected graphs (informative but computationally expensive) also have limitations.

In this project, Delaunay Triangulation achieved 63.57% accuracy but struggled with scalability due to high computational costs. SLIC segmentation outperformed Delaunay with 73.0% accuracy by capturing regional features, enabling faster training despite slower preprocessing. However, SLIC still lags behind CNNs in accuracy, reflecting a trade-off between interpretability and performance.

4.3 Discussion of Limitations

1. **Delaunay Triangulation:**
    - Computationally expensive due to dense connectivity, making it unsuitable for high-resolution images.
    - Limited accuracy (63.57%) due to reliance on pixel-level features, failing to capture broader semantic information.
2. **SLIC Segmentation:**
    - Although it improves accuracy (73.0%), it still falls behind traditional CNNs (~85-90%) and requires time-intensive preprocessing.
    - Loses fine-grained details, leading to challenges in distinguishing visually similar classes (e.g., Sea vs. Glacier).

5. Conclusion and Future Work

5.1 Conclusion

This project demonstrates that GCNs are effective for image classification when combined with graph-based representations:

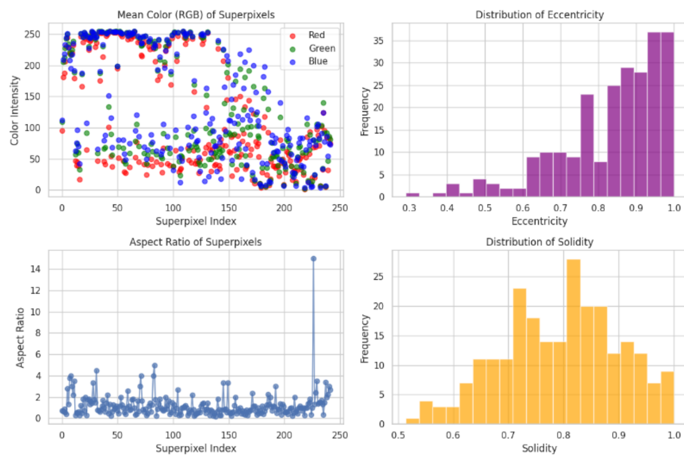
- **SLIC Segmentation** achieves significantly better accuracy (73.0% vs. 63.57%) and faster training than Delaunay Triangulation.
- Graph-based approaches enhance interpretability compared to pixel-based methods.

5.2 Future Work

1. **Hybrid Graph Construction**: Combine pixel-level and regional features for better representation.
2. **Hierarchical GCNs**: Use multi-scale graph structures for richer feature extraction.
3. **Optimized Preprocessing**: Reduce time complexity for large-scale datasets.
4. **Generalization**: Apply to multi-label and hierarchical classification tasks.

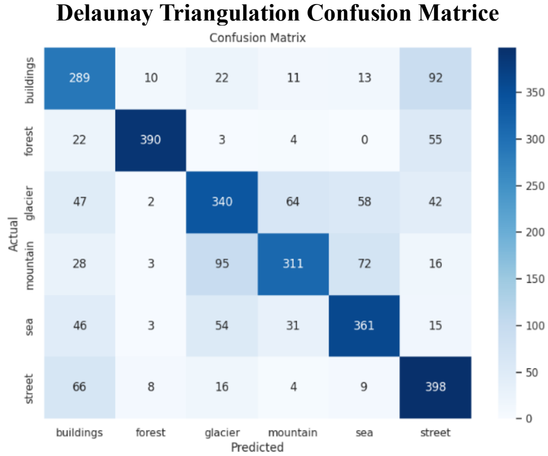
6. Screenshots and Visualizations

6.1 Data Visualizations

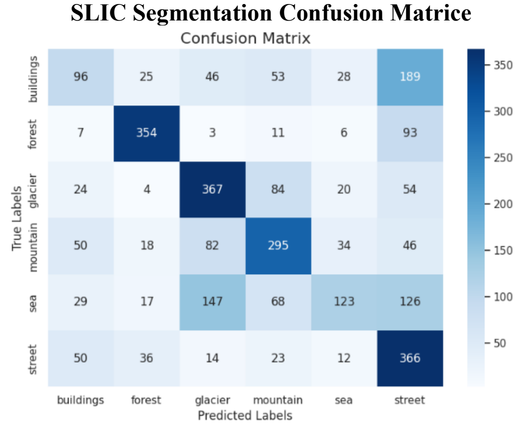


6.2 Model Performance

Delaunay Triangulation Confusion Matrice



SLIC Segmentation Confusion Matrice



7. References

1. Kipf, T. N., & Welling, M. (2017). "Semi-Supervised Classification with Graph Convolutional Networks."
2. Intel Image Classification Challenge dataset: [Kaggle Link](#).