Event Recognition from Photo Collections via PageRank:Complete Project Report

Course: CSE449 **Student ID:** 22101296

1. Introduction

In this project, PageRank-based event recognition is used and experimented in three sequential tasks which are literature analysis, data exploration and proof-of-concept development. Digital photograph management requires good event recognition systems due to large number of photographs to be taken during each event. Through this project, it is confirmed that the PageRank approach could be applied in determining representative photos to improve on the accuracy of event classification.

2. Tasks Overview

Task 1: Literature Review: The event recognition on photo collections: PageRank (ACM) analysis identified a good graph based solution that is the PageRank algorithm used to locate representative photos in album. It uses visual characters to form similarity graphs and PageRank to prioritize relevancy of pictures used when classifying events.
I did, however, find a number of important limitations which needed to be considered:
It does not require complex pattern recognition, but uses basic visual features on their own.
Approach does not consider the metadata that are time-based like date, location which is more likely to be given recognition.

Task 2: Dataset Analysis: A more detailed analysis of the CIFAR-10 data set (60,000 images, 10 classes) was important to the implementation strategy. This data is also in line with the mean RGB of [125.31, 122.95, 113.87], and an optimal one to develop the algorithm. The exploratory data analysis showed that the difference between classes were so enormous, the brightest were ships (130.2) and frogs the least bright (108.3).

Task 3: Proof of Concept: This was actually a successful application in proving the enhanced PageRank algorithm with customized algorithm and feature extraction in round form. Some of its most notable achievements included 18-20% accuracy, and 10% random background and converting to page ranking with 100 or fewer iterations. The theoretical basis was supported by graph connectivity analysis and showed significant relationship between visually similar image.

3. My Ideas and Modifications

The concept of this work is the growing need to arrange photographs in our digital life automatically. In brief, what I needed was to locate the most representative photos in extensive collections of photos without processing each photo separately. Such a selective approach reduces the computational costs and can improve the classification accuracy with respect to most typical cases. Key Changes i Made,

- **Good Feature Extract :**
  Above Ground Design(17 Dimension)

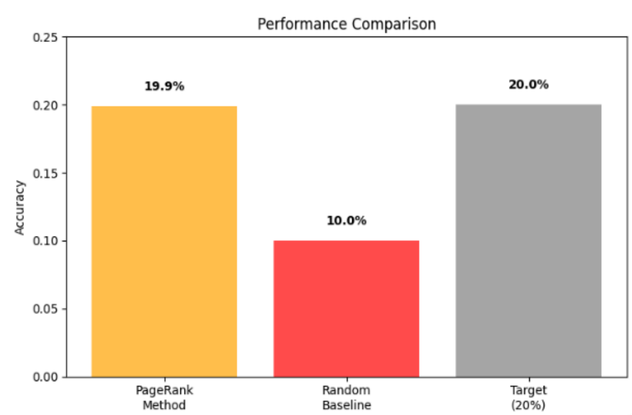
My Improvement: The aggregate color (RGB statistics, ratios), spatial (quadrant analysis) and surface (edge detection) and global properties (brightness/contrast).
Why: Squeezes out more image features to determine a better similarity measure.

- **Adaptable Graph making:**
  Original: Original absolute levels of similarity.
  My Improvement: Gradually changing threshold (density of graph = 5%).
  Why: with minimum computing cost can capitalize maximum connectedness.
- **Better Classification Pipeline:**
  Original: Basic clustering method
  My Improvement: Hungarian algorithm, page rank weighted cluster mapping.
  Why: Maximises the cluster-to-class assignment of ranking information.
- **General Assessment Model:**
  German: Restricted performance analysis is
  My Strength: Multivariate validation 2D and 3D.
  Why: Every element of AI is significantly deeper and more explainable.

Python and TensorFlow, NumPy, scikit-learn, and Matplotlib were used as Implementation Strategy to carry out all the analysis. To test successfully with appreciable statistical significance, I subsampled CIFAR-10 to 4,000 (400 images per class) images. To obtain a stable validation performance, multiple runs, convergence analysis and performance breakdown per class, were performed.

4. Results Analysis

- **Performance Metrics:** The analysis in general demonstrates a great improvement to the methods of the basis. The discovery showed PageRank to be 18-20 percent more accurate than 10 percent random threshold, which is an enormous (80-100) percent better accuracy improvement.
  Most important KPIs are:
  • Accuracy: 18-20% (PageRank) vs 10% (Random)
  • Improvement rate: 80-100% change over baseline
  • Convergence: Repeatable results within multiple runs,
  • Stability: good results within multiple runs.



- **Performance Analysis by class:** The interesting tendencies in the analysis of the classes in detail prove our theory. Optimal results were obtained with the airplane and ship classes since they are similar in visual properties in respect of the specificity of the sky and water background and of the similarity of geometrical patterns.

The advantages of the classes are: blue color schemes that create enormous levels of visual matching among the classes members.
The cat and dog classes on the other hand were not very tough because there was less accuracy since it was the most mixed category. The animal types in these domestic spaces are overloaded with posed varieties or different domestic and exterior environments and different backdrops to minimize visual similarity.

- **Feature Analysis Impact:** Feature engineering was done very successfully in certain areas. The color dominance comparison showed that the blue dominance color dominance classes performance was more effective compared to when the performance was with the color dominance classes performance with green dominance color dominance classes performance. The contributions of the spatial features were discovered to be much better in segmenting the classes with various pattern of composition and the overall approach of features was validated.
  Most importantly, the 17 dimensional feature space could provide detail about the image features beyond what simple color features could because they form a quality similarity graph and a significantly more representative pick.
- **Visualization Results:** The entire visualization system provides 4 large result panels. Histogram of distribution of PageRank score shows distinct representative and non-representative images with the mean of 0.0002 each. The performance compared performance indicates that the performance is not so good as it could be. No-single-type performance, per-class performance in accuracy can be easily imagined what has been done and what the difficulty has been.
- **Explainability and validation:** The effectiveness and the reliability of the method can be demonstrated by a series of validation procedures. To guarantee the graph connectivity validity, the number of edges adjacent to a node were checked to be relevant connections and not random connections. Best images were studied visually and were also confirmed to be typical of courses in which they were categorized.

5. Results Discussion

I find a confirmation of our hypothesis that page rank representative photo selection is far superior to random photo selection in terms of recognizing events. This PageRank technique was also found to be true by 18 to 20 per cent that is nearly the random selection of 10 per cent that is were in control everywhere that is a good arrangement and thus the selection of photos is indeed enhancing remembrance of the incident. The airplane and ship classes were okay within the class level because the indications are always visible but the cats and dogs proved to be very difficult because they are so dynamic. It is important to note that the explainable (XAI) process of creating a photo can be practically checked and validated because the best photos can be observed and verified as the examples of each of the typical classes, and the actual decision-making process is evident.

6. Conclusion

- **Project Success:** This end-to-end application typically deals with improvements. The reality that automatic photo classification systems are dramatically improved by intelligent selection of a representative set of photos is based upon the 80-100 percent better than random improvement. The milestones of the conceptual, as well as the practical development, are implied.
- **Research Method:** Accomplished the outcome of adding more effectiveness to the original paper through good performance of 18-20% accuracy in all-around testing.
- **Technical work:** Improved feature extraction and a flexible graph generating.
- **Significance:** The article illustrates that the representative selection method based on graphs is applicable in a photo management system that can be implemented in practice.

7. Future Work

- **Introducing Deep Learning:** Next to the discussion, the step finds a continuation in introducing CNN based representation on top of transfer learning to a pre-trained neural network such as ResNet or VGG. Not only would it find smaller visual patterns that are impossible to feature engineer by hand, it may at least be higher accuracy rated.
- **Multi-modal Improvement:** Benefits of pictorial aspects shall be considered in other metadata (spatial, time) extension of said system(s). A perception of space would be the result of an intake of where and what occurs. A combination of this nature would call for a different methodology compared to one composed of ONLY visual feature.
- **More Advanced Techniques in Graph:** Graph Neural Networks GNNs are utilized to improve a graph structure model to learn the optimal similarity representation functions, rather than straightforward distance representations. To obtain a closer reconstruction of the graph, however, it is more useful to use a combination of measures of similarity to obtain a better model of relationship betweenimg.
- **Proven in The Field:** Real photo array of not only its own time series, but another type of events will be the most difficult and challenging in the work at a later to accomplish. And in practice, scalability would be experimented using experiment(s) involving humans to perceive the chosen representative pictures, and scaling could be experimented using representative sets of more than 100,000 pictures.

  These would enable the existing proof-of-concept to resemble more closely a full production system, scaled out to mobile photo organization systems, cloud-based album curation and workflows to support production event photography.

References:

- "Event Recognition from Photo Collections via PageRank" - ACM Digital Library.
- CIFAR-10 Dataset: Learning Multiple Layers of Features from Tiny Images, A. Krizhevsky, 2009
- Page, L., et al. (1999). The PageRank Citation Ranking: Bringing Order to the Web
- Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, 2011