Supplementary Material for "Deep Neural Networks in Fully Connected CRF for Image Labeling with Social Network Metadata"

Chengjiang Long Roddy Collins Eran Swears Anthony Hoogs
Kitware Inc.
28 Corporate Drive, Clifton Park, NY 12065
{chengjiang.long, roddy.collins, eran.swears, anthony.hoogs}@kitware.com

Abstract

This document provides more statistic information about the top frequently used words as tags among the positive training instances of the MIR-9K dataset to explain why our proposed DCRF is able to make better use of text information and why the DCRF-Text outperforms the DCRF-Tag, which are not included in the main paper due to space limitations.

1. Observation

To analyze why our proposed DCRF-Text s able to outperform DCRF-Tag, we visualize the top 20 words in decreasing order of frequence occurring among positive training examples, as shown in Figure 1 and Figure 2. We also provide the top 100 words in decreasing order of frequence occurring among positive training examples from Figure 3 to Figure 26. As we can observe, the top frequently co-occurring words such as "love", "photo", "great", "group", "added", "nice" *et. al* convey little information relative to any of the prediction 24 categories. This oberservation demonstrates the limitation of using the high occurring words as tags to extract informative information from text. Instead, we treat all the texts including title, description and comments information associated with an image as an unorganized incoherent sentence or a bag of words, and then we train a text-level CNN to explore the underlying information embedded in text. This fact may be able to explain why our proposed DCRF-Text is able to achieve better performance than DCRF-Tag. Again, such a text-level CNN enables our proposed DCRF is able to make better use of information to boost the quality of image labeling.

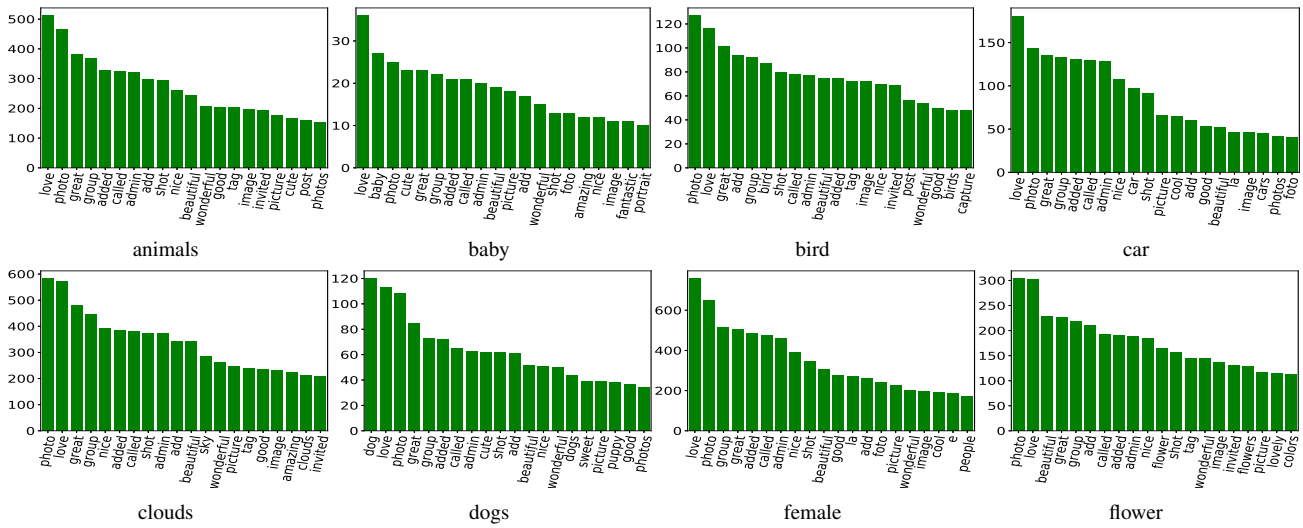

Figure 1: The visualization of the top 20 words frequently occurring among the positive training instances for each of the 1st-8th categories on the MIR9k dataset.

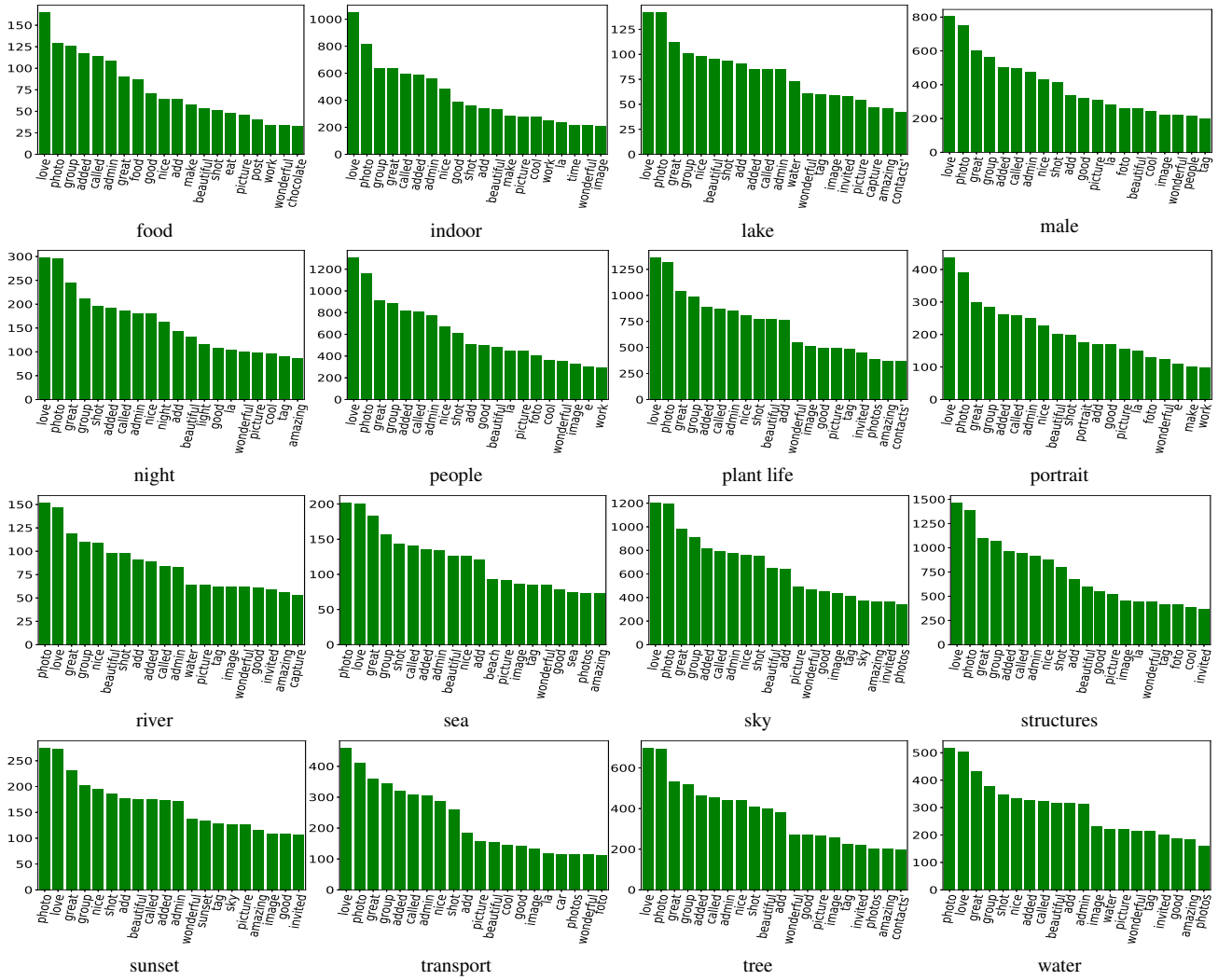

Figure 2: The visualization of top 20 words frequently occurring among the positive training instances for each of the 9th-24th categories on the MIR9k dataset.

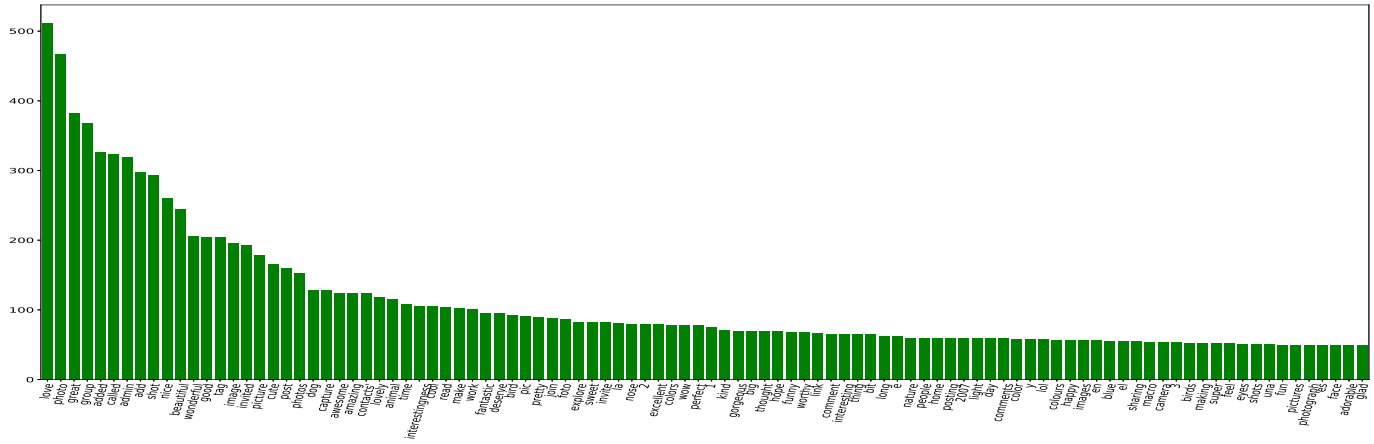

Figure 3: The visualization of top 100 words frequently occurring among the positive training instances for the 1st category (animal) on the MIR9k dataset.

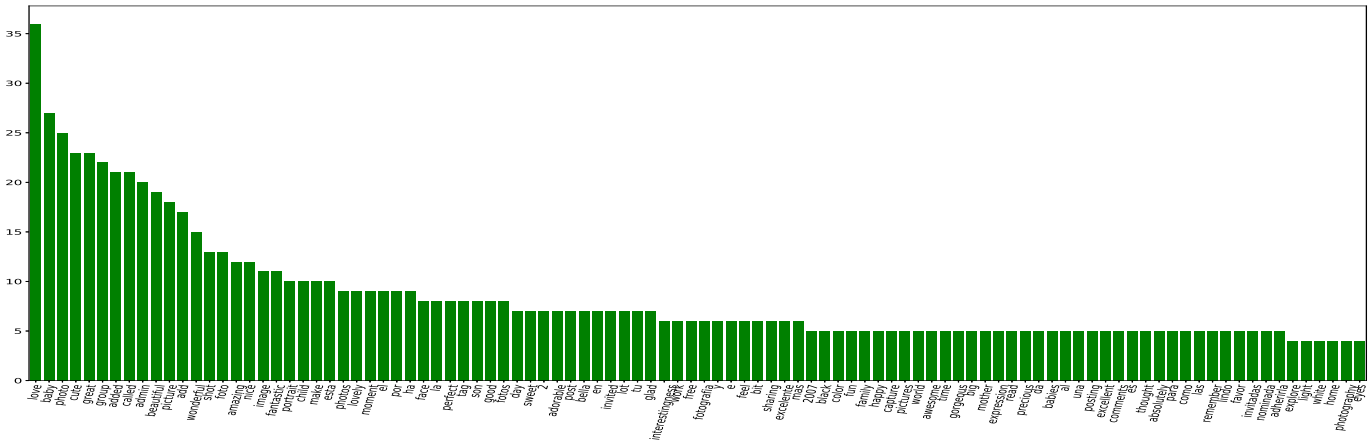

Figure 4: The visualization of top 100 words frequently occurring among the positive training instances for the 2nd category (baby) on the MIR9k dataset.

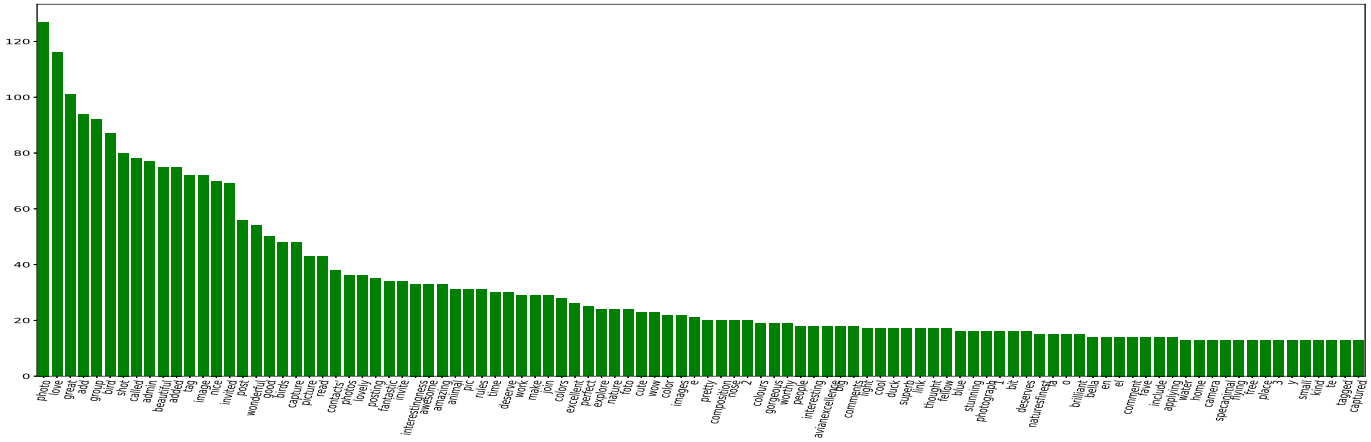

Figure 5: The visualization of top 100 words frequently occurring among the positive training instances for the 3rd category (bird) on the MIR9k dataset.

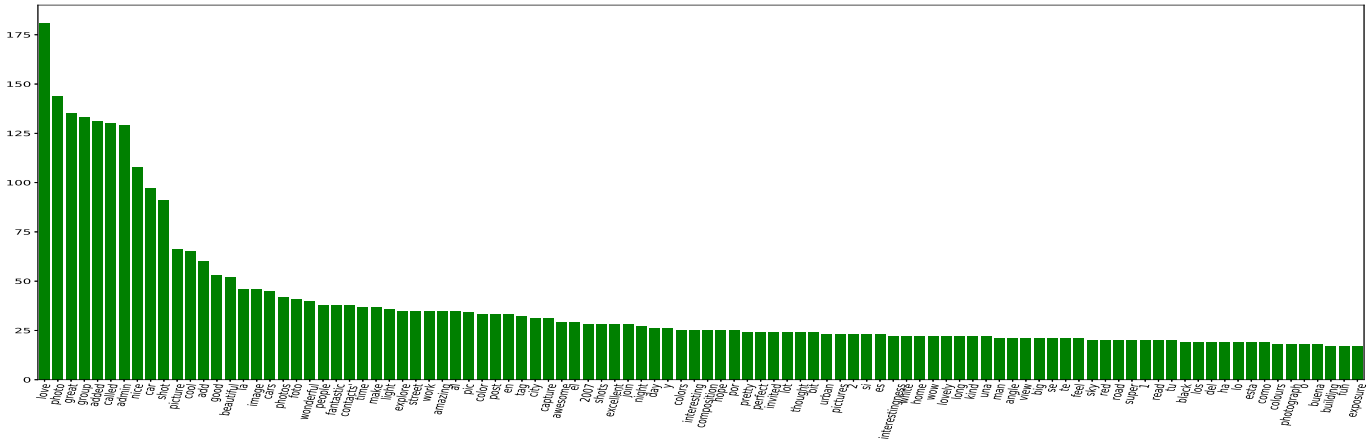

Figure 6: The visualization of top 100 words frequently occurring among the positive training instances for the 4th category (car) on the MIR9k dataset.

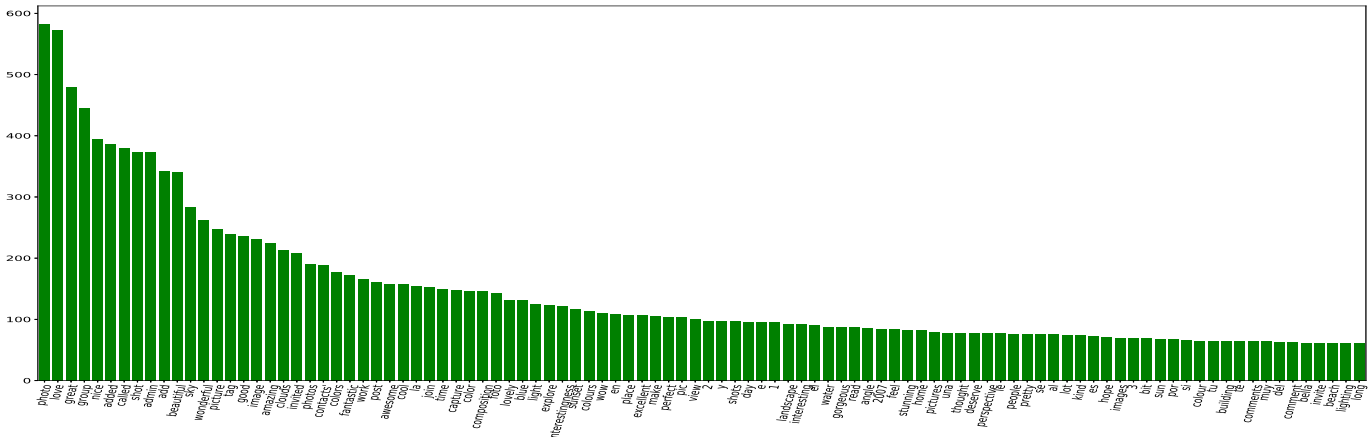

Figure 7: The visualization of top 100 words frequently occurring among the positive training instances for the 5th category (clouds) on the MIR9k dataset.

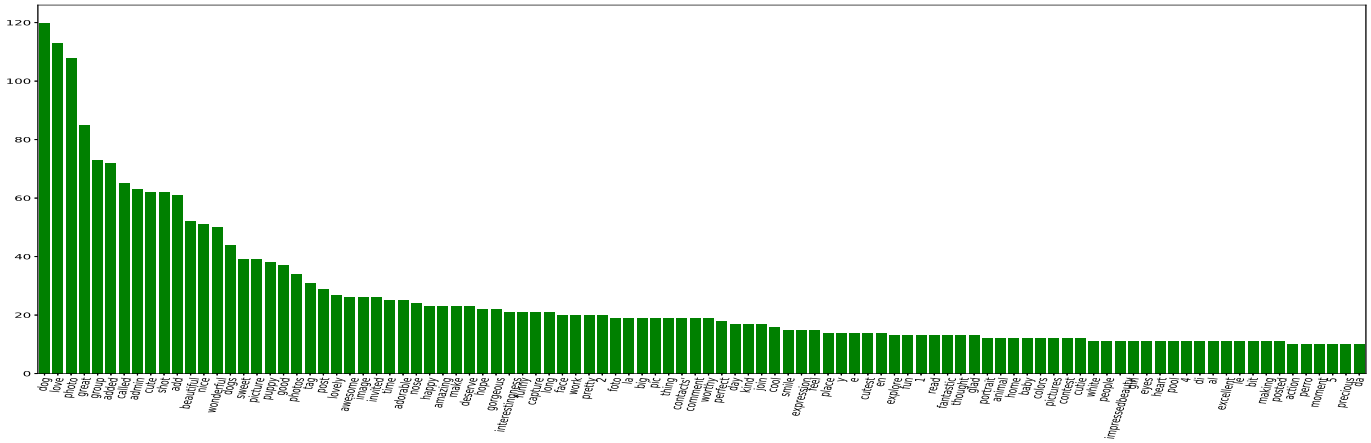

Figure 8: The visualization of top 100 words frequently occurring among the positive training instances forthe 6th category (dog) on the MIR9k dataset.

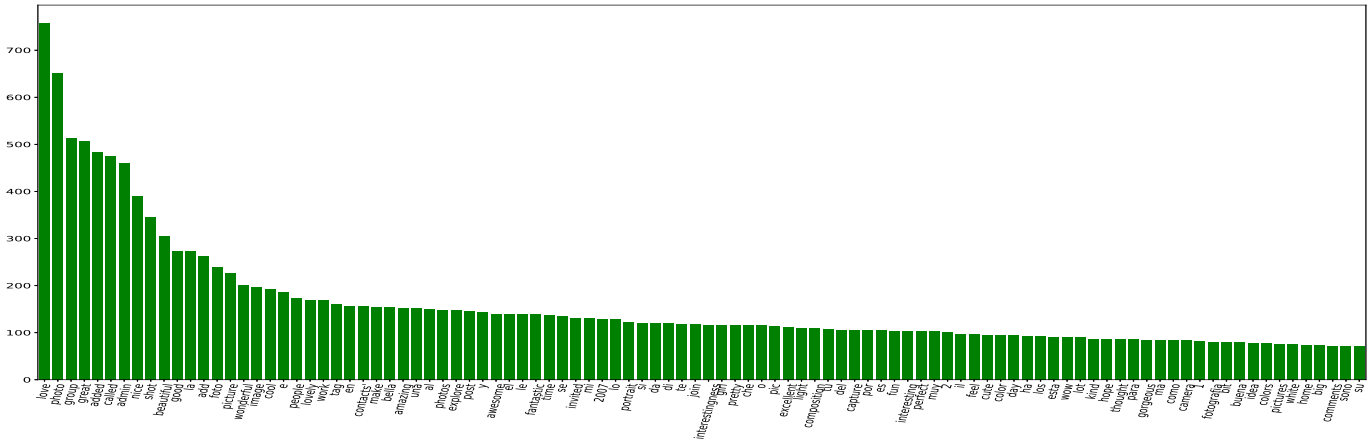

Figure 9: The visualization of top 100 words frequently occurring among the positive training instances for the 7th category (female) on the MIR9k dataset.

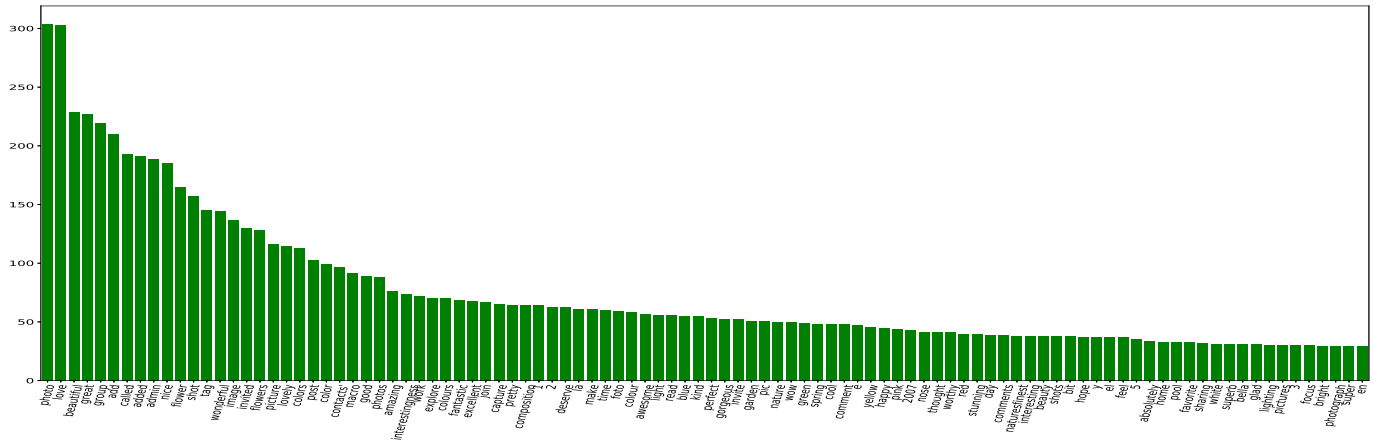

Figure 10: The visualization of top 100 words frequently occurring among the positive training instances for the 8th category (flower) on the MIR9k dataset.

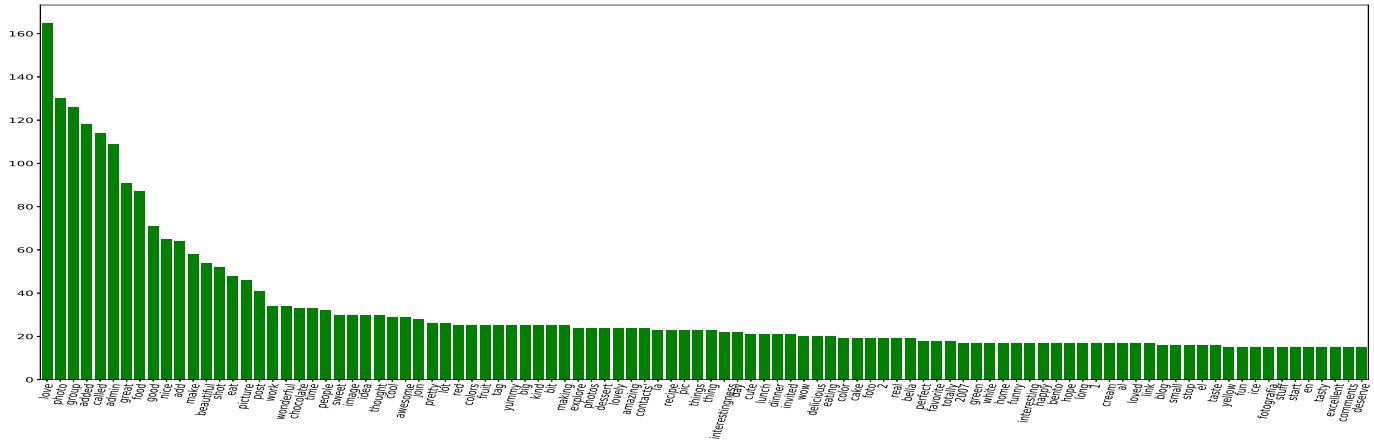

Figure 11: The visualization of top 100 words frequently occurring among the positive training instances for the 9th category (food) on the MIR9k dataset.

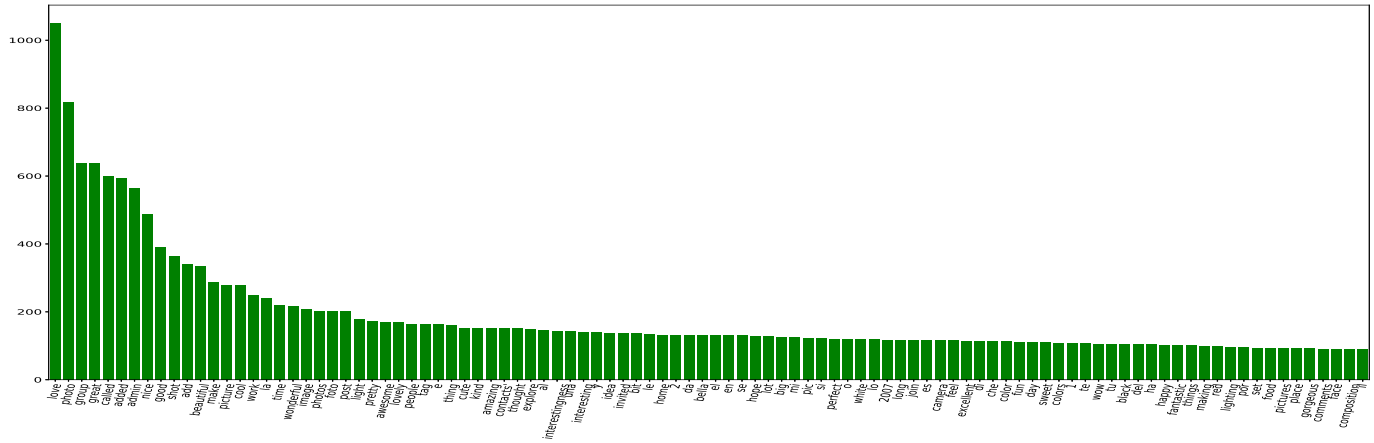

Figure 12: The visualization of top 100 words frequently occurring among the positive training instances for the 10th category (indoor) on the MIR9k dataset.

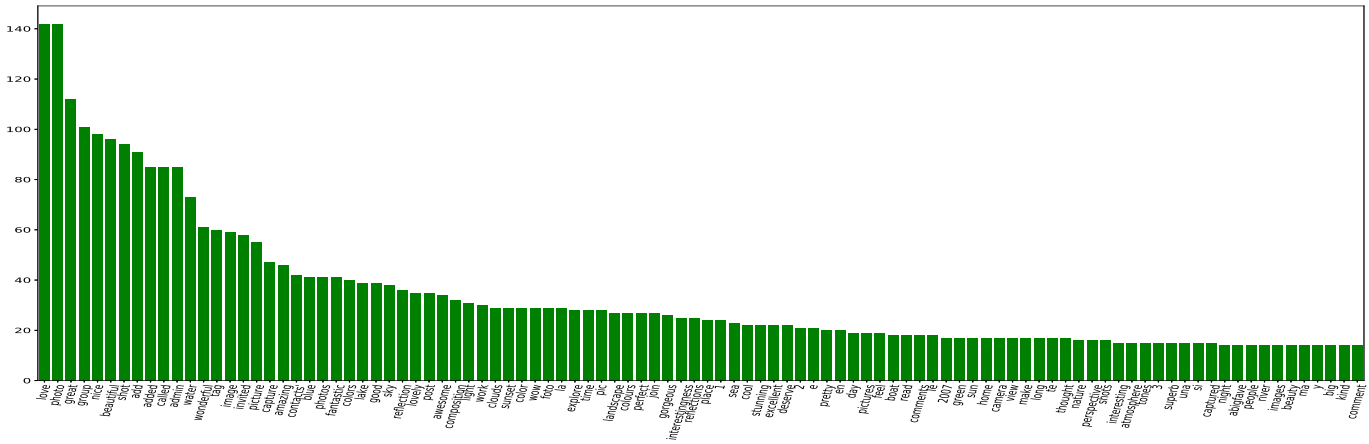

Figure 13: The visualization of top 100 words frequently occurring among the positive training instances for the 11th category (lake) on the MIR9k dataset.

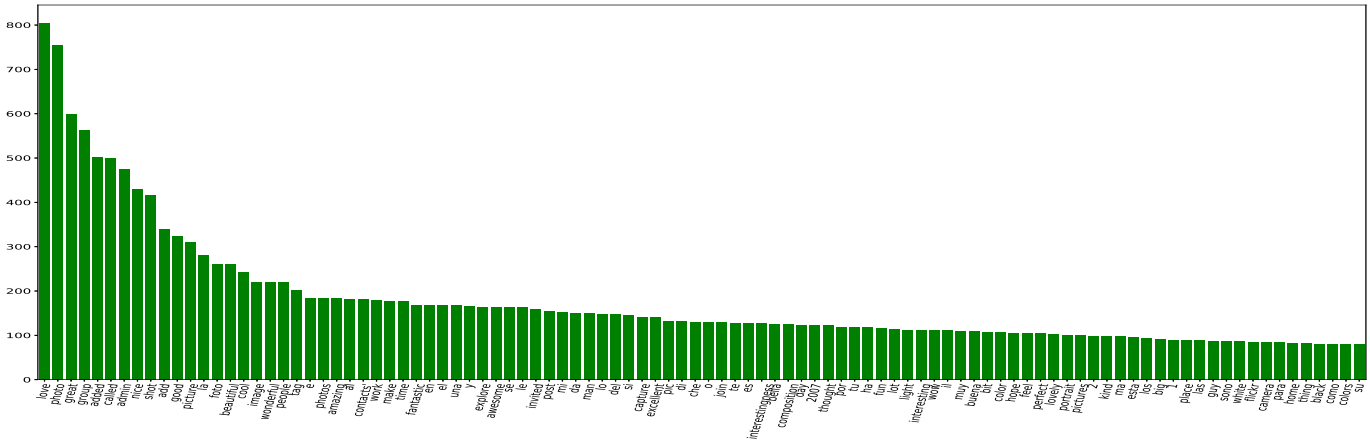

Figure 14: The visualization of top 100 words frequently occurring among the positive training instances for the 12th category (male) on the MIR9k dataset.

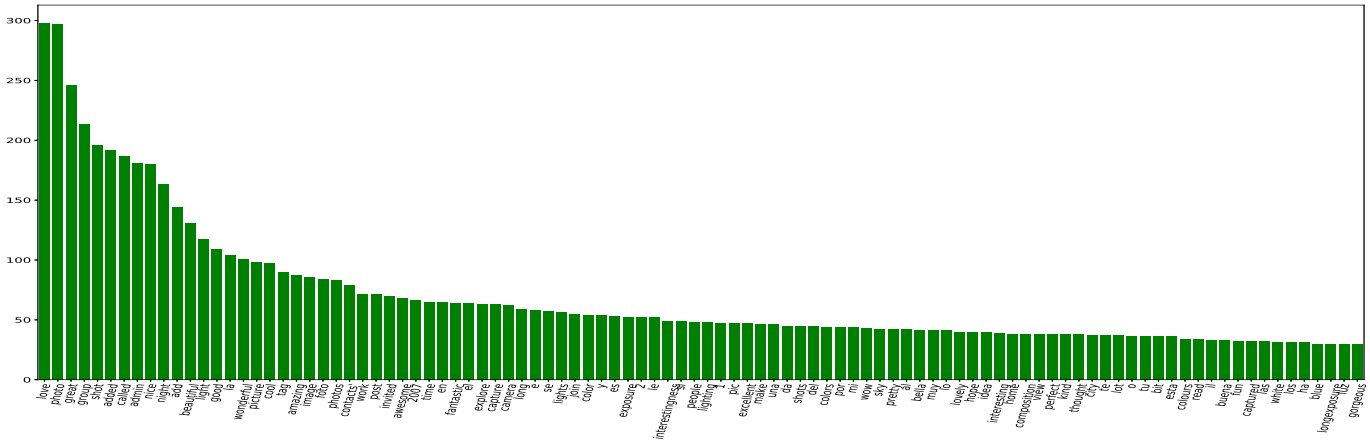

Figure 15: The visualization of top 100 words frequently occurring among the positive training instances for the 13th category (night) on the MIR9k dataset.

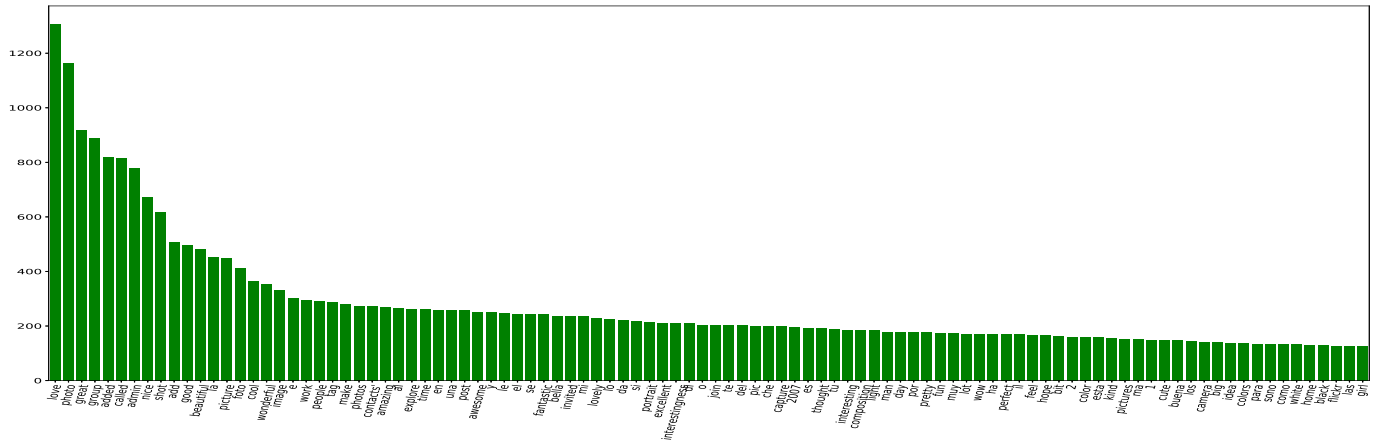

Figure 16: The visualization of top 100 words frequently occurring among the positive training instances for the 14th category (people) on the MIR9k dataset.

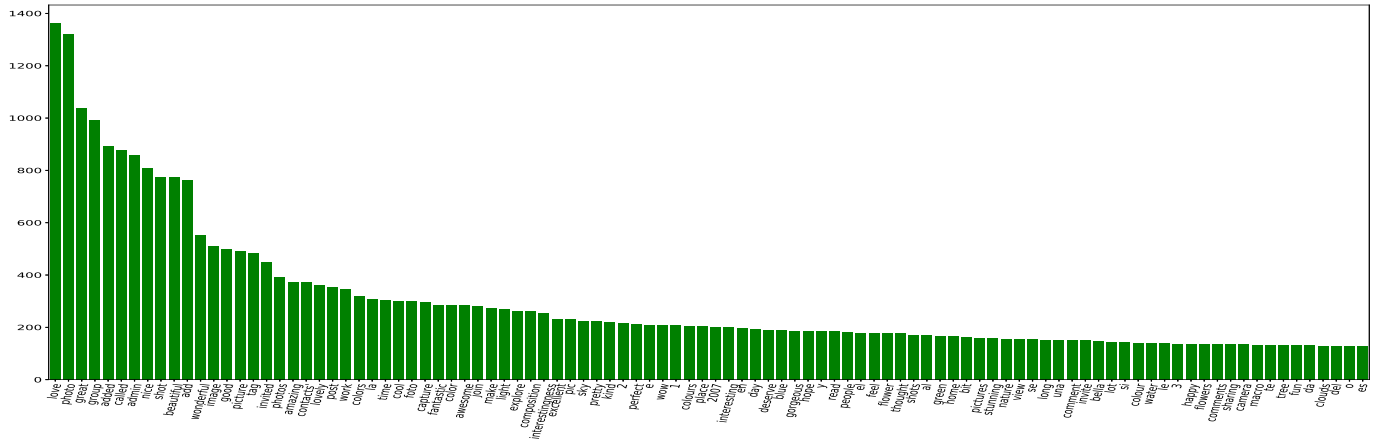

Figure 17: The visualization of top 100 words frequently occurring among the positive training instances for the 15th category (plant life) on the MIR9k dataset.

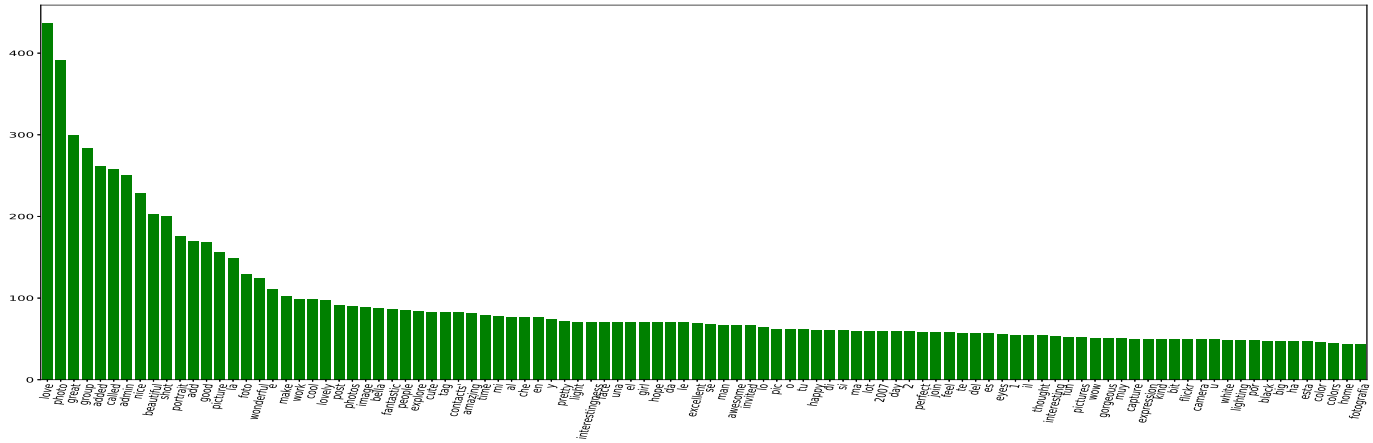

Figure 18: The visualization of top 100 words frequently occurring among the positive training instances for the 16th category (portrait) on the MIR9k dataset.

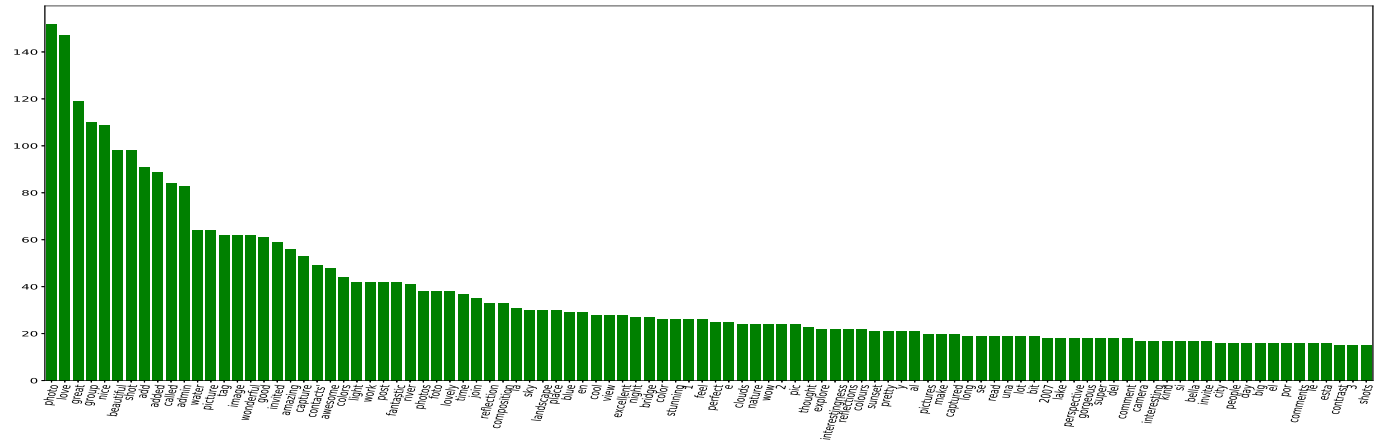

Figure 19: The visualization of top 100 words frequently occurring among the positive training instances for the 17th category (river) on the MIR9k dataset.

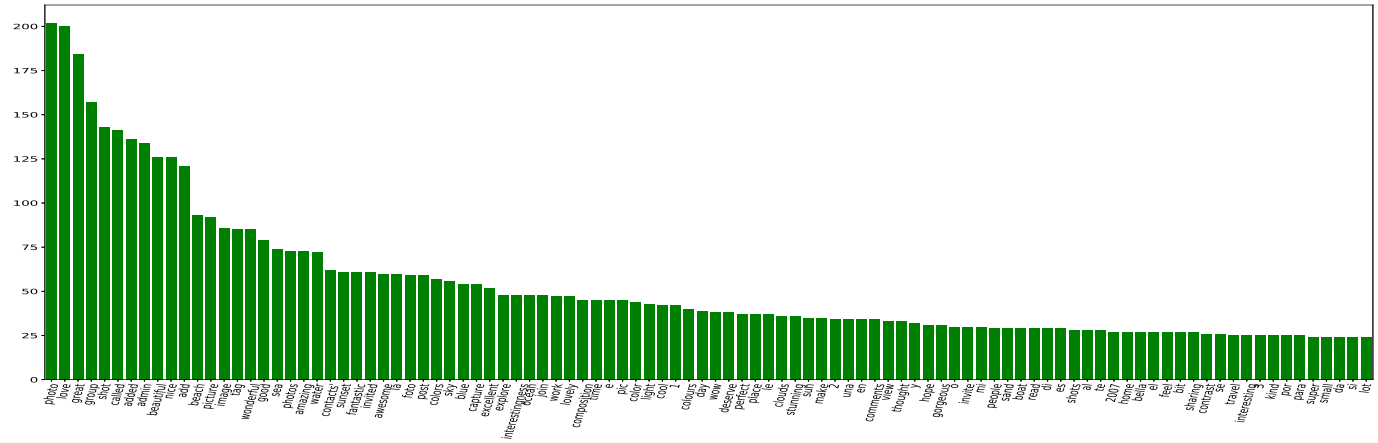

Figure 20: The visualization of top 100 words frequently occurring among the positive training instances for the 18th category (sea) on the MIR9k dataset.

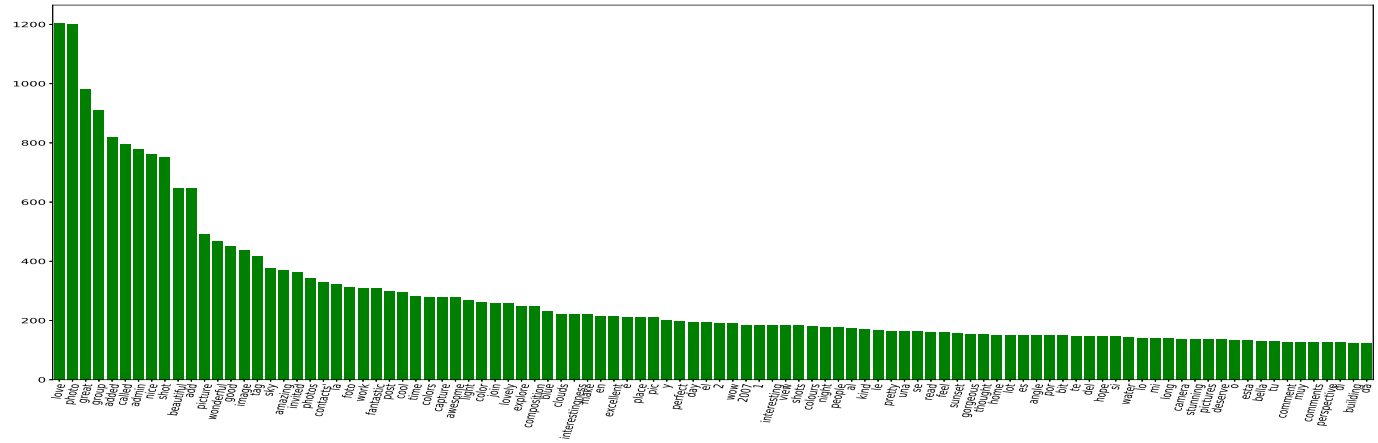

Figure 21: The visualization of top 100 words frequently occurring among the positive training instances for the 19th category (sky) on the MIR9k dataset.

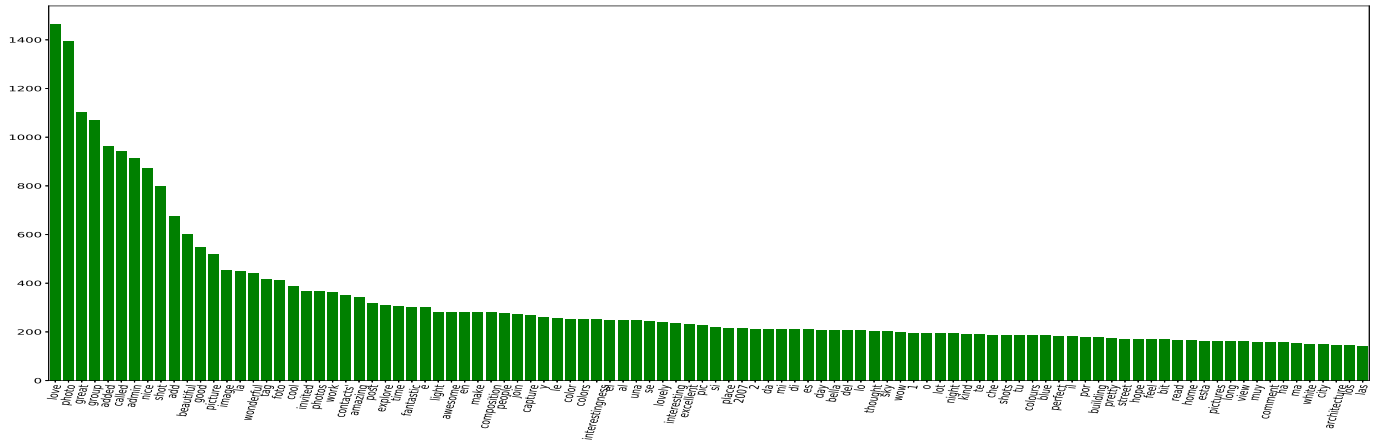

Figure 22: The visualization of top 100 words frequently occurring among the positive training instances for the 20th category (structures) on the MIR9k dataset.

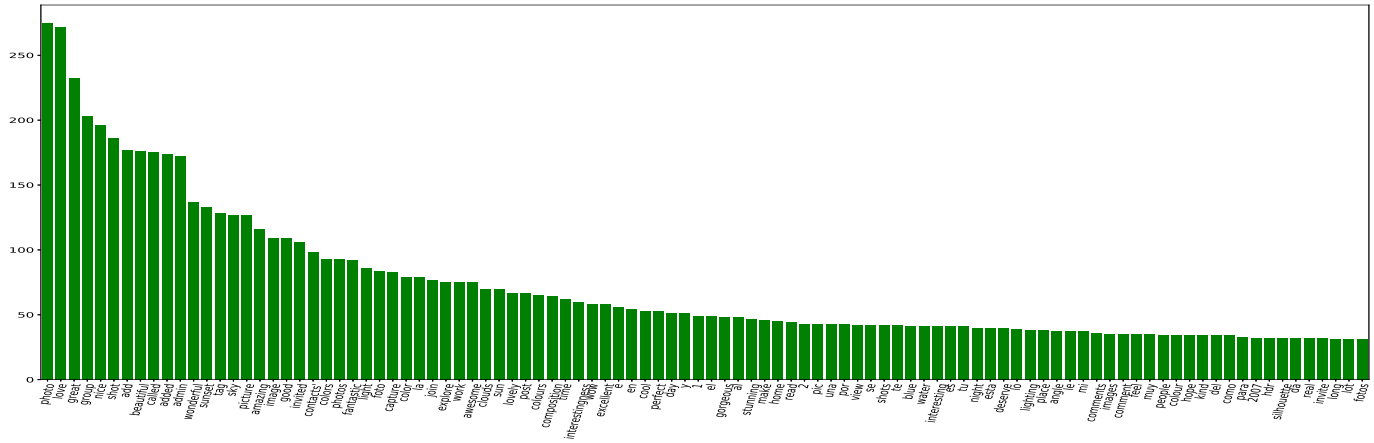

Figure 23: The visualization of top 100 words frequently occurring among the positive training instances for the 21th category (sunset) on the MIR9k dataset.

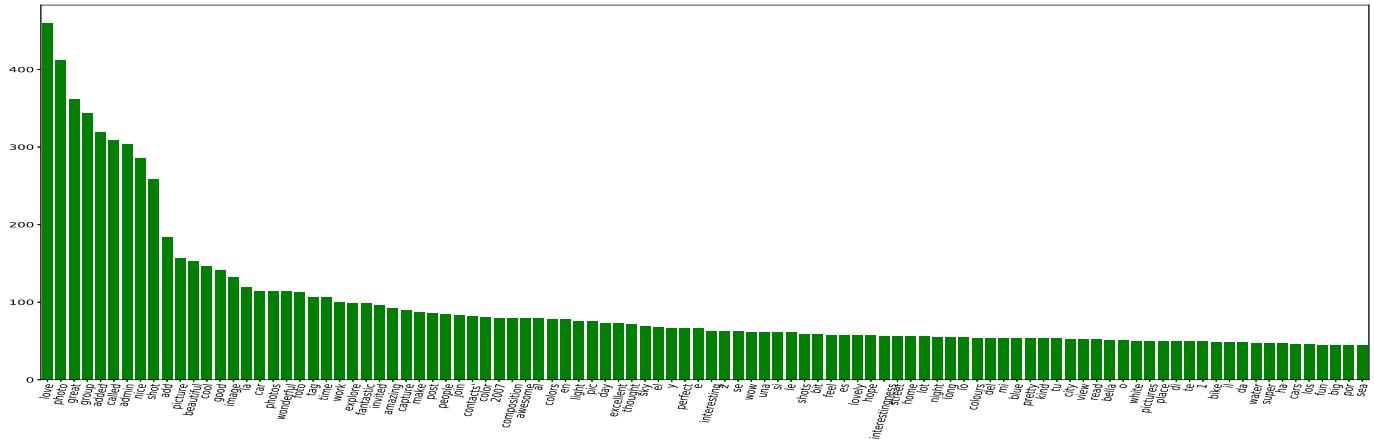

Figure 24: The visualization of top 100 words frequently occurring among the positive training instances for the 22th category (transport) on the MIR9k dataset.

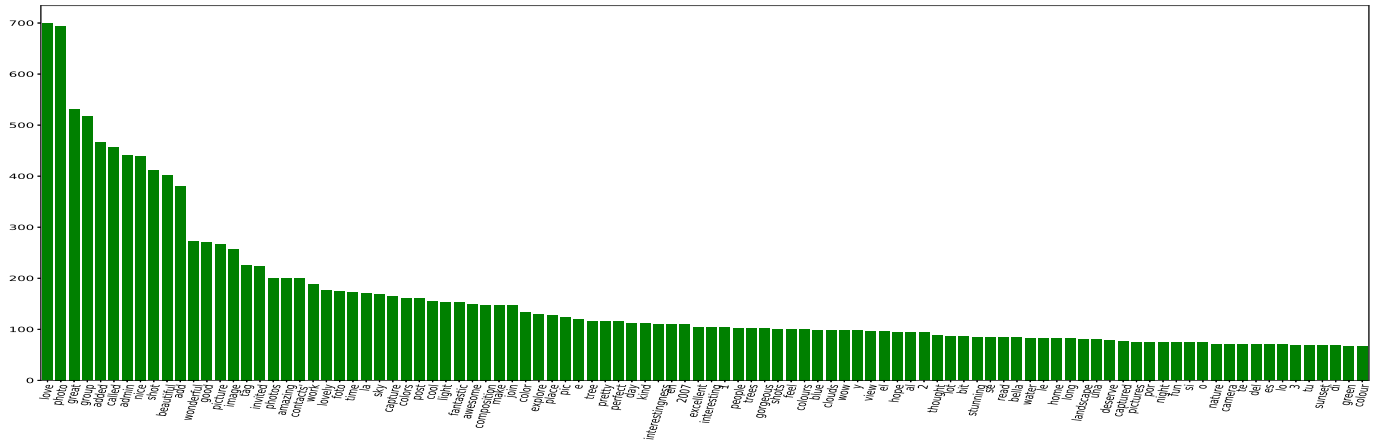

Figure 25: The visualization of top 100 words frequently occurring among the positive training instances for the 23th category (tree) on the MIR9k dataset.

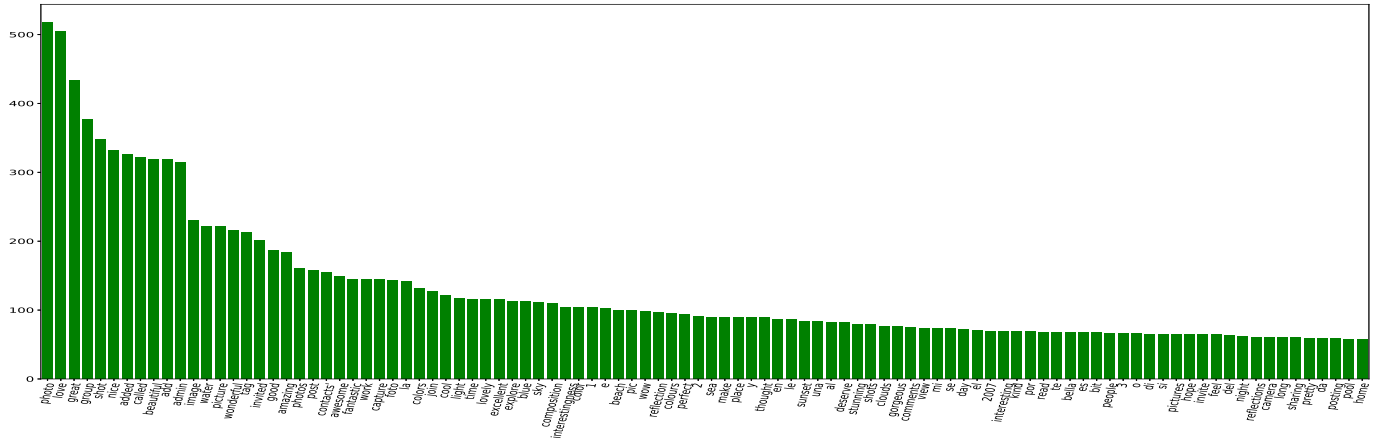

Figure 26: The visualization of top 100 words frequently occurring among the positive training instances for the 24th category (water) on the MIR9k dataset.