ONE NETWORK FOR ALL

Image captioning, similar words, relevant images for text query and similar images for image query

Adithya Chowdary Boppana
aboppana@iu.edu

Sahaj Singh Maini
sahmaini@iu.edu

Siva Charan Mangavalli
simang@iu.edu

INTRODUCTION

- We try to exploit the fact that each layer of a neural network will hold important information that can be further used
- Hence, the same network can be leveraged for different applications

PREVIOUS WORK

- Previous work included using the generated captions for image retrieval instead of meta data[1]
- Use of skip thought vectors [4] and locality sensitive hashing for image retrieval

DATA

- The Flickr 8K dataset includes images obtained from the Flickr website and corresponding captions
- Training data contains 20K pairs and the test data consisted of 500 pairs

ARCHITECTURE

IMPLEMENTATION

- Adam optimizer was used for training with a learning rate of 0.001 for 10 epochs and sparse categorical cross entropy loss function was employed. It was trained and tested on Google Colab (Tesla K80)

IMAGE CAPTION

Caption for input image

A surfer be ski down a wave*

A dog run through the snow*

*Lemmatized text

SIMILAR WORDS

Similar words for a given word

Black
- White
- Gray
- Orange
- Yellow
- Brown

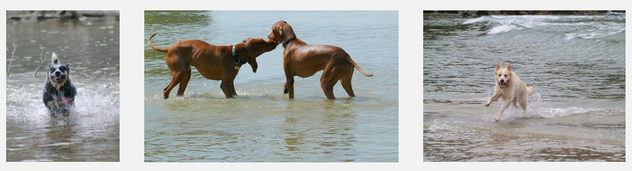
Soccer
- Tennis
- Baseball
- Basketball
- Football
- Rugby

RESULTS

Query

Retrieved Images

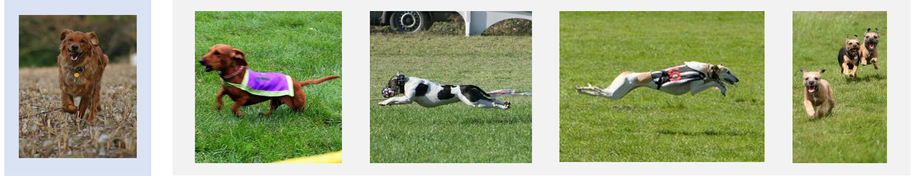
dog play in a beach

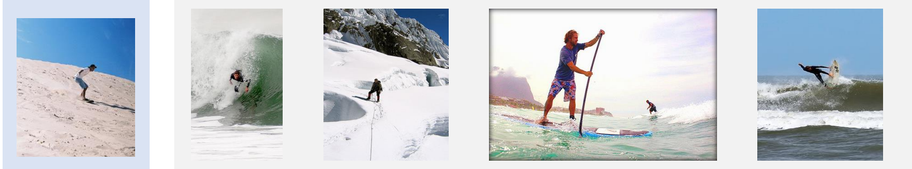


people play football

Query

Retrieved Images





TEXT QUERY

Relevant images for text query

Similar images for image query

IMAGE QUERY

CONCLUSION

- A single neural network can be used for multiple applications using the embeddings generated
- In future, Sent2Vec and LSH can be used for the same process in a more efficient and quicker way

REFERENCES

1. Oriol Vinyals, Alexander Toshev, Samy Bengio, Dumitru Erhan “Show and Tell: A Neural Image Caption Generator”
2. Tanya Piplani “DeepSeek: Content Based Image Search & Retrieval”
3. Shuang Bai “A Survey on Automatic Image Caption Generation”
4. Ryan Kiros, Yukun Zhu, Ruslan Salakhutdinov, Richard S. Zemel, Antonio Torralba, Raquel Urtasun, Sanja Fidler “Skip-Thought Vectors”
5. Flickr 8k https://forms.illinois.edu/sec/1713398
6. Google Conceptual Captions https://ai.google.com/research/ConceptualCaptions