Mining Interesting Relationships between Cat Breeds using Deep Representations Learnt by a CNN

SAMARJIT KARMAKAR, HIMADYUTI BHANJA, DILISHA JAIN, and PRATHAM TANGRI,
National Institute of Technology, Warangal, India

Deriving relationships between objects has been far studied in data mining. We take this problem further by deriving interesting relationships between cat breeds using the deep representations learnt by a convolutional neural network (CNN). Our model is able to derive accurate relationships using agglomerative clustering between 12 diverse cat breeds. Further, we have visualized the various important features learnt by the CNN for a specific cat breed image.

Additional Key Words and Phrases: Convolutional Neural Networks, Computer Vision, Classification, Clustering, Deep Representations

ACM Reference Format:
Samarjit Karmakar, Himadyuti Bhanja, Dilisha Jain, and Pratham Tangri. . Mining Interesting Relationships between Cat Breeds using Deep Representations Learnt by a CNN. 1, 1 (April), 2 pages.

Fig. 1

1 INTRODUCTION

Making sense of an image is a task we humans are extremely good at. Our brain has evolved over millions of years, becoming an expert in making sense out of images, finding similarities, finding weird differences, etc. The same task is extremely difficult for a machine to perform as a machine operates on low level features such as pixel information.

Recent trends in deep learning has made it possible to perform tasks which may have been deemed impossible for a machine, such as object recognition, speech to text, machine translation, etc. We have taken up as similar task of making sense out of visual imagery.

In this project, we derive interesting relationships between cat breeds using the feature maps derived from a single convolutional architecture and using the learned features for agglomerative clustering.

2 OUR MODEL

We use a RowCNN based classifier network to classify the cat breed images into 12 classes. The classifier network is able to learn a mapping from breed images to their corresponding labels (one among 12 breeds).

We use the deep representations learnt by the above network to perform agglomerative clustering of 12 distinct cat breed images. The feature vector used for clustering is the feature activations obtained from an intermediate layer of the CNN, for each breed image.

Authors' address: Samarjit Karmakar; Himadyuti Bhanja; Dilisha Jain; Pratham Tangri, National Institute of Technology, Warangal, India.

XXXX-XXXX//4-ART
https://doi.org/

2.1 RowCNN based Classifier Network

Let the cat image be represented by $X \in \mathcal{R}^{x \times y}$, where x represents the horizontal axis and y represents the vertical axis. We apply convolution with multiple variable filter sizes $w_i \in \mathcal{R}^{h_i \times h_i}$, where h_i represents the dimension of the i^{th} filter, $i \in [1, k]$.

$$c_i = f(w_i * X + b) \tag{1}$$

Here, $b \in \mathcal{R}$ represents bias and f represents the ReLU non-linearity function. Multiple such operations are performed with variable windows sizes and concatenated into one single feature vector c, which is passed through fully connected layers for classification.

$$c = c_1 \oplus c_2 \oplus \ldots \oplus c_k \tag{2}$$

The idea is to capture breed relevant features, which could be a mixture of low, medium and high-level features. We take variable filter to capture the varying pixel axes and hence, capturing a hierarchy of breed relevant features. The concatenated vector c represents the whole feature vector breed relevant features of the cat image X.

The concatenated feature vector is passed to an agglomerative clustering algorithm to find out various relationships between the cat breeds.

The clustering is done using various linkage measures:

- Single Linkage: The smallest distance between medoids m_i, m_j in the two breeds.
- Complete Linkage: The largest distance between medoids m_i, m_j in the two breeds.
- Complete Linkage: The average distance between medoids m_i, m_j in the two breeds.
- Ward Linkage: The distance between medoids m_i, m_j in the two breeds, computed by Ward's method.

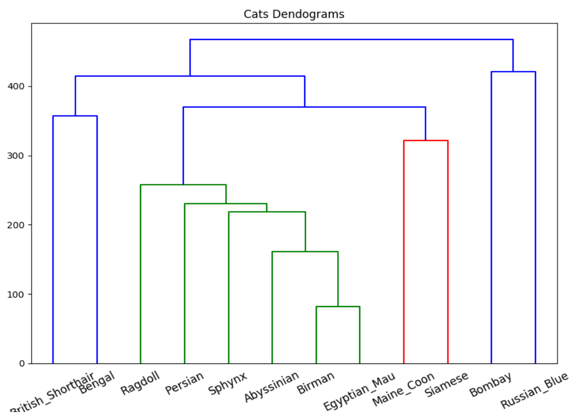

Fig. 4. Relationships using complete linkage measure.

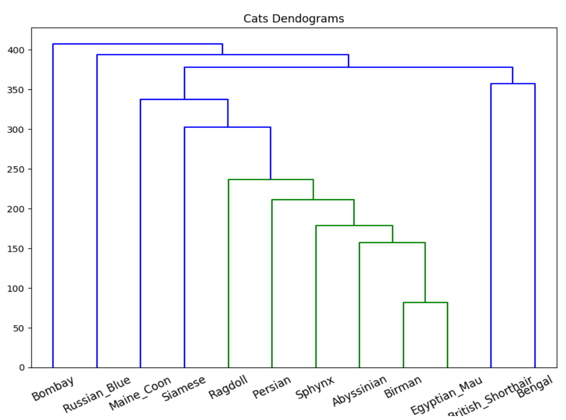

Fig. 5. Relationships using average linkage measure.

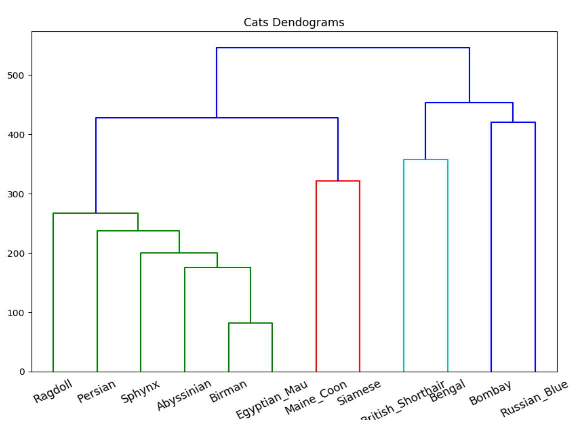

Fig. 6. Relationships using ward linkage measure.

Fig. 2

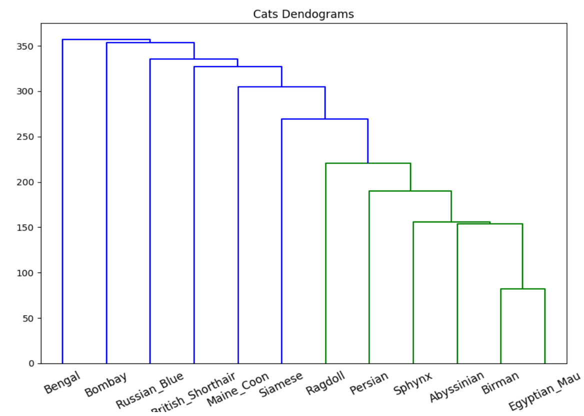

Fig. 3. Relationships using single linkage measure.

3 EXPERIMENTS AND RESULTS

The classifier network is tested on a dataset of cat images. The network is able to classify the cat images to their respective breeds with an accuracy of roughly 80.3%.

The relationships derived from agglomerative clustering are shown in Figure 3, 4, 5 and 6.

A few visualizations from the intermediate layers of the classifier network are shown in Figure 1 and 2.

The models have been implemented on the PyTorch deep learning framework and trained on a single Nvidia 1070 ti GPU.

4 CONCLUSION

Mining interesting relationships between cat breeds is an interesting data mining task. The important features relevant to breed identification are embedded in the intermediate layers of a neural network. Our model successfully derives interesting relationships between cat breeds. The applications of such a framework are huge, some of which may be deriving relationships between images, texts or audio using neural networks. Our framework works on a subset of such tasks proving the validity of such a model.