

# Animal Classification Using Deep Learning

<sup>1</sup>Mrs. Shweta Suryawanshi, <sup>2</sup>Miss. Vaishali Jogdande, <sup>3</sup>Miss. Ankita Mane

<sup>1</sup>Assistant Professor, Dept. of E&TC Engg., Dr. D. Y. Patil Institute of Engineering, Management and Research, Akurdi Pune, India

<sup>2,3,4</sup> Student, Dept. of E&TC Engg., Dr. D. Y. Patil Institute of Engineering, Management and Research, Akurdi Pune, India

## ABSTRACT

*Kaggle's Dogs vs. Cats contest is trying to solve the CAPTCHA challenge, which is based on the issue of distinguishing dog and cat images. It's easy for humans, but evidence suggests that the automatic separation of cats and dogs is particularly difficult. Many people have been working or working on building classifiers for machine learning to address this issue. A color-based classifier had an accuracy of 56.9% on the Asirra dataset. A SVM classifier achieved an accuracy of 82.7 percent based on a combination of color and texture characteristics. And in, they used the features of SIFT (Scale-Invariant Feature Transform) to train a classifier then finally got a 92.9 percent accuracy. We also want to solve this problem in our plan and achieve higher efficiency. We've tried various strategies. We tried Dense-SIFT features, combining DenseSIFT and color features, and features learned from CNN, for example. We also used SVMs on the learned features and finally achieved 94.00 percent of our best classification accuracy.*

**Keywords**—CNN, KAGGLE, DEEP LEARNING.

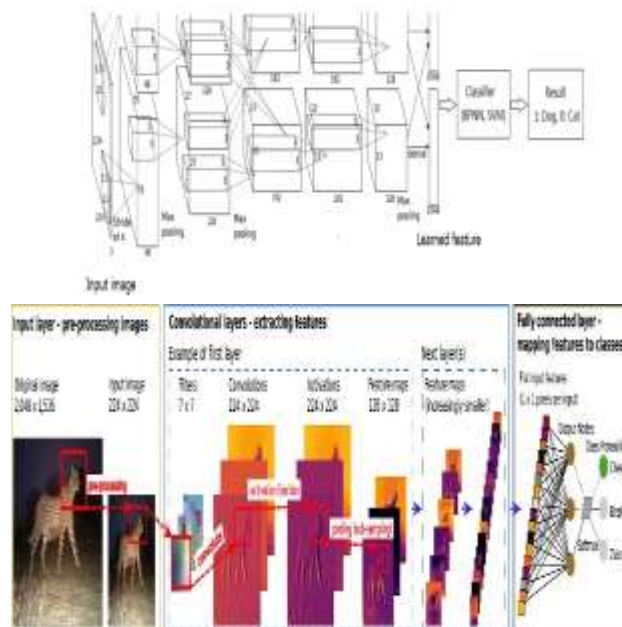
## 1. INTRODUCTION

One of the basic ideas of computer vision is the first intention to try to "understand the picture," which leads to a persistent rise in the need to comprehend the high-level meaning of objects concerning object recognition and identification of images. The area has grown rapidly by becoming a fundamental visual skill needed by computer vision systems. In a variety of fields, images have become omnipresent as so many people and systems extract vast amounts of information from images. Knowledge that can be critical in fields such as automation, schools, self-driving cars, tracking, or 3D model representation creation. While each of the above-mentioned applications differ by numerous factors, they share the common process of correctly annotating an image with one or a probability of labels correlating to a series of classes or categories. This method is known as object identification, and because of the emphasis on recognizing what an image is symbolic of, it has become an important research subject in the field. Detection and identification of objects dependent on image processing is a vast field of work focus. The motivation for this project is to build a system for the animal researchers and wildlife photographers to automatically detect and recognize wild animals and flowers. The detection and recognition of animals is an important area that has not been discussed quickly. However, this project is aimed at building a system that would help animal researchers and wildlife photographers to study animal behavior. Technology used in this research may be further modified for use in applications such as safety, purposes of monitoring, etc. Wildlife photography is considered one of photography's most challenging forms. It requires sound technical skills such as capturing properly. Photographers of wild life typically need a good field of expertise in crafting and a lot of patience. Some animals, for example, are hard to approach, so knowledge of the behavior of the animals is needed to predict their actions. Sometimes, until the exact time, photographers have to remain calm and quiet for many hours. Photographing certain species may require stalking skills or concealment using hide / blind. A great photograph of wild life is also the product of being at the right time in the right place. Because of the highly demanding nature of the work it costs too much and the same life in the company may be in risk. Animal researchers usually travel around the world to remote locations. Hostile conditions in the photojournalist's life are often the norm. Some are going to sit for hours and hours before they snap a shot worth selling. The photographers should be brave enough to stay comfortably and with great patience in a hostile environment until the animals appear. It always comes with fewer disturbances to the animal's natural behavior to have a perfect scene or result. Because of their high sensitivity, human presence can be easily identified by the dog. So photographers need to be prepared to face any critical jungle moment because we can't predict what's going to happen next. However, the DSLR cameras that are used in the industry are quite pricey and the shutter on-off process has drawbacks. There must therefore be proper recognition in such appliances. The higher the image quality, the more room it needs for memory. In wild photography, it also encourages proper recognition we have chosen to be conservative in this project and have limited our efforts to animals and flowers. This is a random selection because collecting a set of animal data to create a proper database is always a challenging part. Our purpose is to achieve the part of acceptance, so we used the plan for abstraction of characteristics, segmentation, and thresholding to achieve the probabilistic outcome. Many

descriptors such as Color format descriptor (CLD), Color structure descriptor (CSD), Edge histogram descriptor (EHD), Homogenous texture descriptor (HTD), Region form, Contour shape were used for the extraction of the function.

## 2. METHODOLOGY

In 2012, on the ImageNet 2012 classification benchmark, Krizhevsky and Hinton trained a CNN and achieved state-of-the-art performance. Our model is based on Krizhevsky's model. The original model has 8 layers and the last three layers are two fully connected layers (same as BP Neural Network's hidden layers) and output layers. The last 2 layers are cut off and the previous 6 layers are reused to extract features from images. Recent work has shown that features extracted from the activation of a CNN trained in a fully supervised fashion on a large, fixed set of object recognition tasks can be re-used for new generic tasks, which can vary significantly from the originally trained tasks. And a Deep Neural Network training needs a lot of experience and skill, and it takes a lot of time. That's why we re-used Krizhevsky's pre-trained network instead of creating our own deep architecture. Figure 6 shows the outline of responsibilities between two GPUs explicitly. One GPU runs at the top of the figure the layer-parts while the other runs at the bottom of the layer-parts. The input of the networks is 224x224 (RGB) and the number of neurons in the remaining layers of the networks is given by 55555 mm, 27 mm, 1313 mm, 13113 mm, 4096. Then we use the 4096 dimensional features to train a classifier and classify images as dog or cat. For more details on the network, we refer to it.



This model's architecture is illustrated by figure. First, input images are normalized into 224x224 (first normalize the short edge size to 224 and then select the center part of the images); then the first convolution layer filters the 224x224 input image with 96 kernels of size 11x11 with a 4-pixel step (this is the distance between neighboring neurons' receptive field centers in a kernel map). The second convolutionary layer takes the pooled output of the first convolutionary layer as its input and filters it with 256 size 5x5 kernels. The third, fourth, and fifth convolutionary layers are connected without any pooling layers interfering. The third convolution layer has 384 kernels of size 3x3, the fourth has 192 and 256 kernels of size 3x3. The fourth convolution layer. There are 4096 neurons in the fully connected layers. The 6-layer CNN converts an object into a 4096-dimensional function matrix. Then we trained the extracted features on the BP Neural Networks and SVMs with 5 fold cross-validation. After that, to classify the test set, we applied the learned models and evaluated their performance.

## 3. OBJECTIVES OF PROJECT

To identify the different animals and flowers

- **ADVANTAGES**
- 1. Real time system.
- 2. Low cost.
- 3. Easy to understand

#### • APPLICATIONS

1. Real time system.
2. Low cost .
3. Helpful for identifying different flower and animal.
4. Helpful for Wildlife Photographer.

#### 4. CONCLUSION

To improve the performance, we have a tendency to applied image segmentation approach to preprocess the info. However, because of poor segmentation result, we have a tendency to didn't reach any improvement. The most effective accuracy we have a tendency to got from the primary technique is just seventy one.47% (from AN SVM classifier). Seven to realize higher performance, we have a tendency to enforced our second approach, that could be a trainable model that applies the CNN to find out options. we have a tendency to additionally looked insight into what Deep Networks learned from pictures and explained why they sensible} good performance. the very best accuracy of this approach is ninety four.00% (from AN SVM classifier), that is additionally our greatest result and helps America rank ninth in ninety one groups within the Kaggle competition. In terms of classifiers, we have a tendency to primarily think of SVMs and BP Neural Networks, taking our high dimensional feature area into consideration. Numerous parameter settings were explored to boost classification accuracy on the take a look at dataset. for instance, for the BP Neural Networks, we have a tendency to tried completely different hidden layers and hidden units; for the SVMs, completely different kernel functions and C parameters were used.

#### 5. REFERENCES

- [1] Keqing Zhu, JieTian and Haining Huang, "Underwater object ImagesClassification Based on convolutional neural network," 2018 IEEE 3<sup>rd</sup> International Conference on Signal and Image Processing (ICSIP), pp.301-305,
- [2]Yichao Wu, Fei Yin and Chenglin Liu, "Improving handwritten Chinesetext recognition using neural network language models and convolutionalneural network shape models," Pattern Recognition, vol. 65,
- [3] SiyueXie and Haifeng Hu, "Facial expression recognition with FRCNN,"Electronics Letters, vol. 53, pp. 5-237, 2017.
- [4]N. Passalis and A. Tefas, "Spectral clustering using optimized bag-offeatures,"in Proceedings of the 9th Hellenic Conference on ArtificialIntelligence. ACM, 2016, pp. 1–9.
- [5]TajbakhshNima, Shin JaeY, Gurudu, Suryakanth R and et al, "Convolutional Neural Networks for Medical Image Analysis: Fine Tuning or Full Training?" IEEE Transactions on Medical Imaging, vol.35, pp. 1299-1312
- [6]Szegegy Christian, Vanhoucke Vincent, Ioffe Sergey and et al, "Rethinking the inception architecture for computer vision," Proceedings of the 2016 IEEE Conference on Computer Vision and PatternRecognition, pp. 2818-2826, 2016.
- [7]Kingma, D. P., & Lei Ba, J. (2015). Adam: A Method ofStochastic Optimization. arXiv:1412.6980 [Cs], (pp1–15).Retrieved from <https://arxiv.org/abs/1412.6980>
- [8]R. Khan, J. Van de Weijer, F. Shahbaz Khan, D. Muselet, C. Ducottet, and C. Barat, "Discriminative color descriptors," in Computer Visionand Pattern Recognition (CVPR), IEEE Conference on, 2013, pp.2866–2873.
- [9] A. Khosla, J. Xiao, A. Torralba, and A. Oliva, "Memorability of imagerregions," in Advances in Neural scientific discipline Systems,2012, pp.305–313.
- [10]P. Ye, J. Kumar, L. Kang, and D. Doermann, "Unsupervised feature learning framework for reference image quality assessment," in Computer Vision and Pattern Recognition (CVPR), IEEE Conferenceon, 2012, pp. 1098–1105.
- [11]Krizhevsky, A., Sutskever, I., & Geoffrey E., H. (2012).ImageNet Classification with Deep Convolutional Neural Networks. Advances in Neural information science Systems twenty five (NIPS2012), (pp19).<https://doi.org/10.1109/5.726791>
- [12] N. Silberman and R. Fergus. Indoor scene segmentation using a structured light sensor. In ICCV - 2011..
- [13] A. Krizhevsky. Learning multiple layers of features from tiny images. Master's thesis, Department of Computer Science, University of Toronto, 2009.
- [14] Deep Learning Face Attributes in the Wild, by Ziwei L., Ping L., Xiaogang W. & Xiaoou T. (2015) (Cited: 401).
- [15] Image Super-Resolution Using Deep Convolutional Networks, by Chao D., Chen C., Kaiming H. & Xiaoou T. (2014) (Cited: 591).