

# DOG CLASSIFICATION INTO 120 BREEDS

GROUP10: HANG ZHANG, LEYAN ZHU, HAONAN SONG, ZAN DENG

{HAZ006, LEZHU, HASONG, Z1DENG}@UCSD.EDU

## Abstract

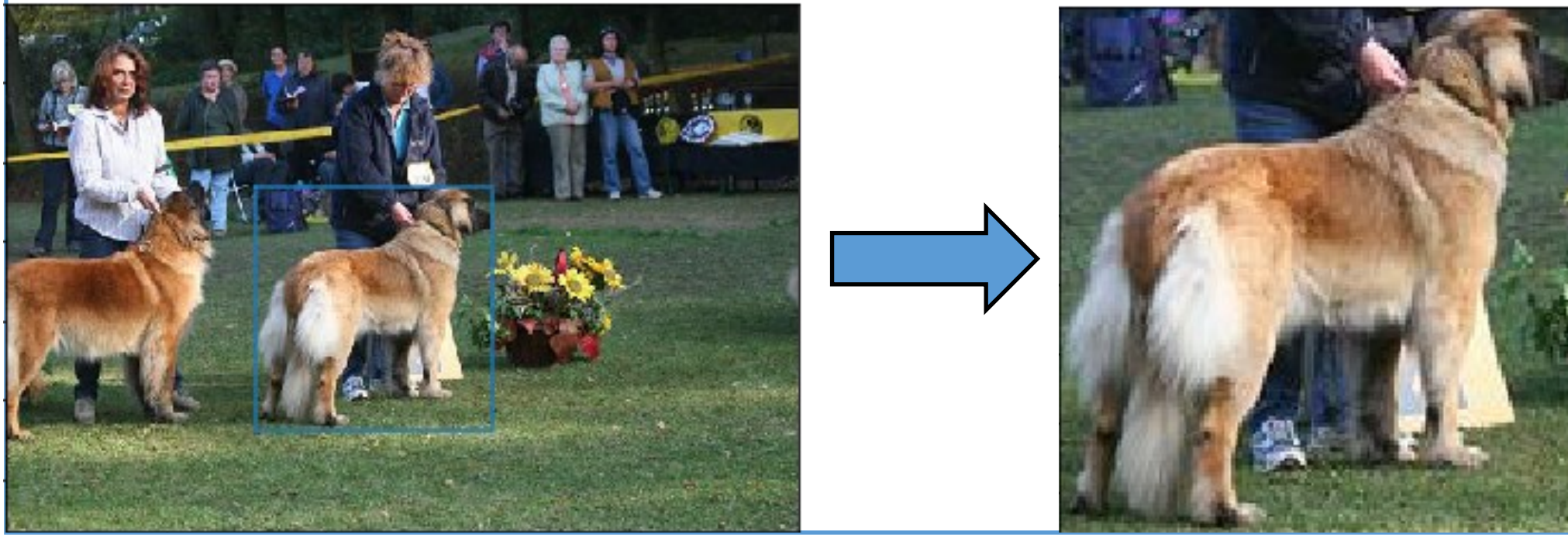
Dog classification is a challenging problem as certain dog breeds have near identical features or differ in color and age. Our project, based on the Stanford dog dataset, aims to classify the dogs into different breeds. We preprocess the dataset and use different models including Alexnet, VGG16, and Inception-v3 and achieve accuracy rate of 29.65%, 62.09%, and 79.28% separately.

## Dataset

The dataset comes from famous Stanford Dog Dataset[1][2][6] and consists of two parts as follow.  
Images: 20580 colorful images in 120 categories(120 breeds of dogs);  
Annotations: name, size and bounding box for each image.

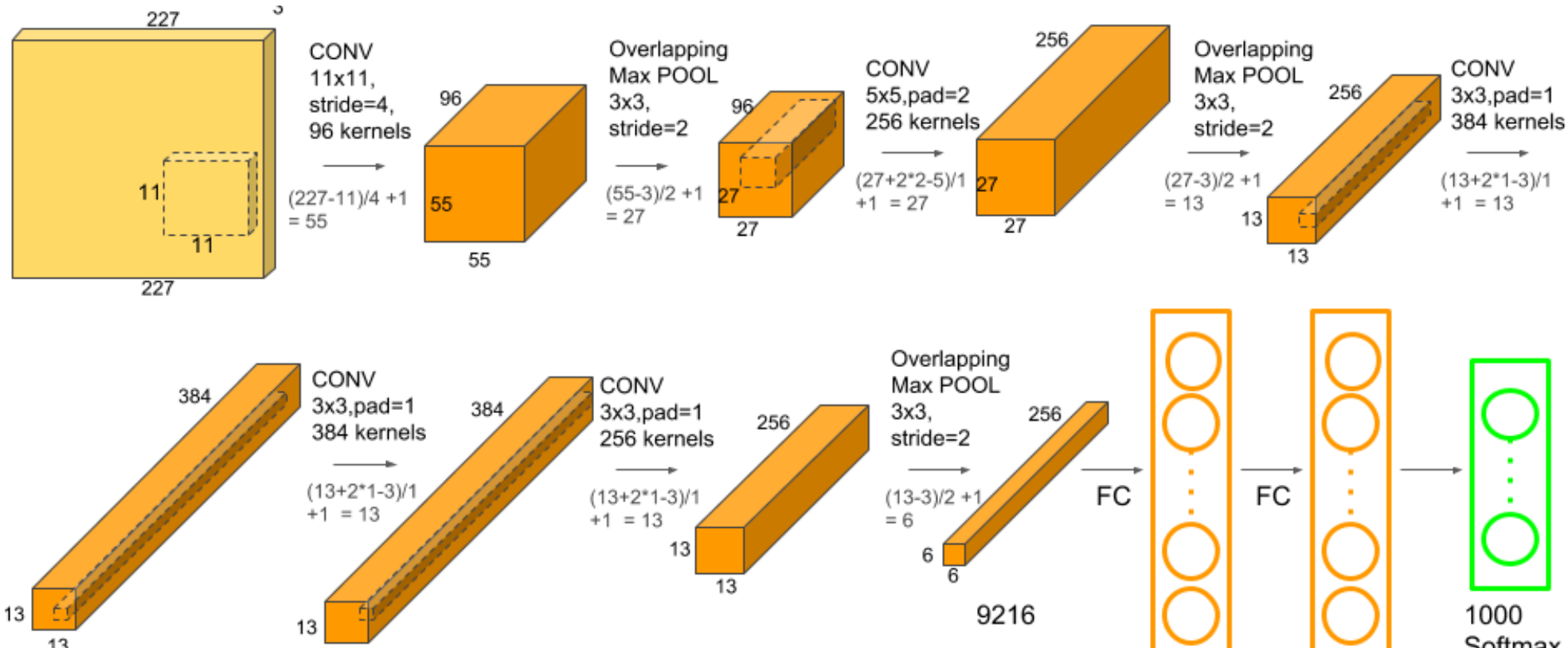
## Features

The pre-processing of dataset includes localization, data augmentation and normalization. We preprocessing the data with location feature with the help of Annotations. Since we use deep learning models, we do not need to extract specific feature by ourselves.

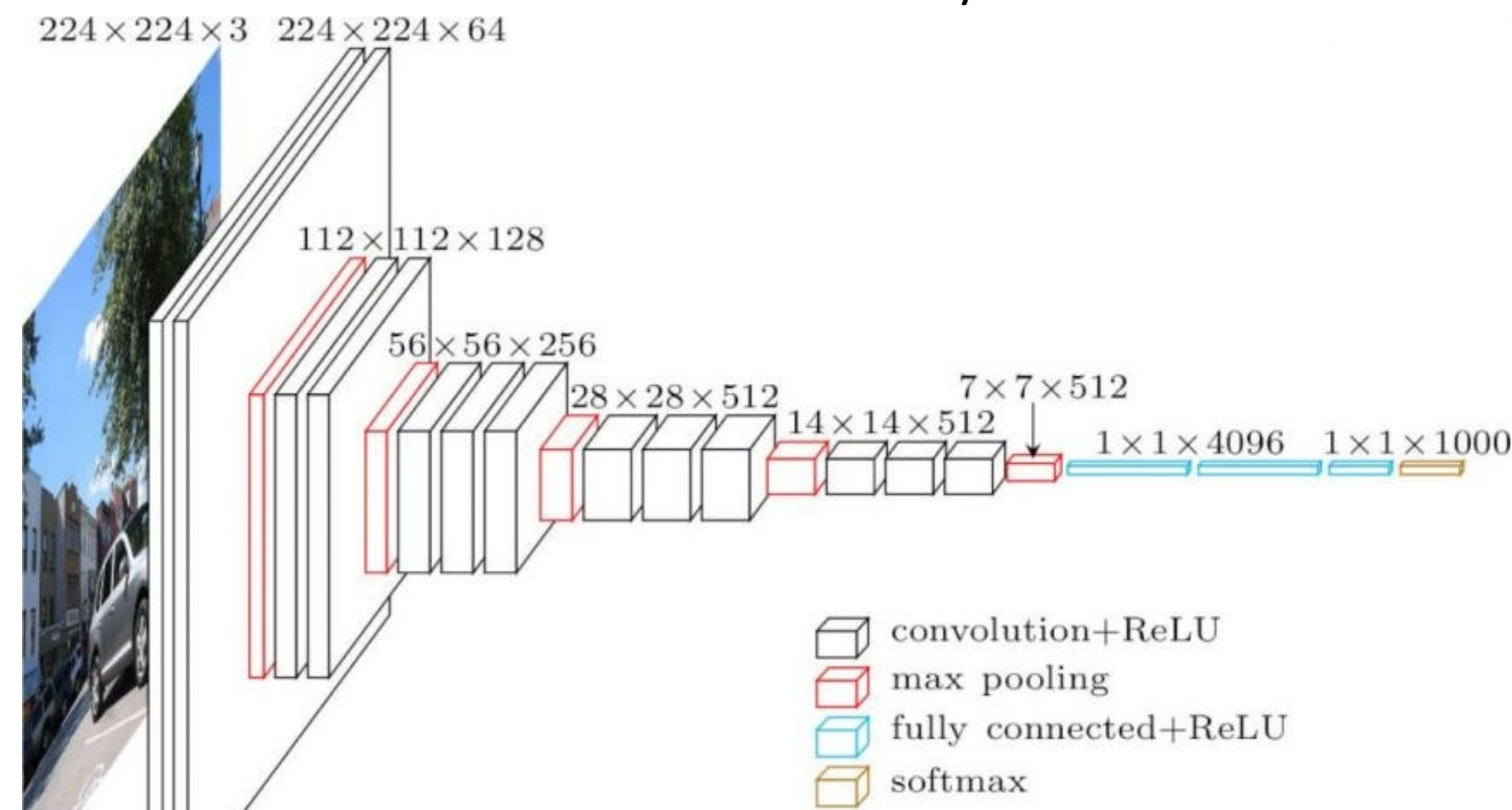


## Models

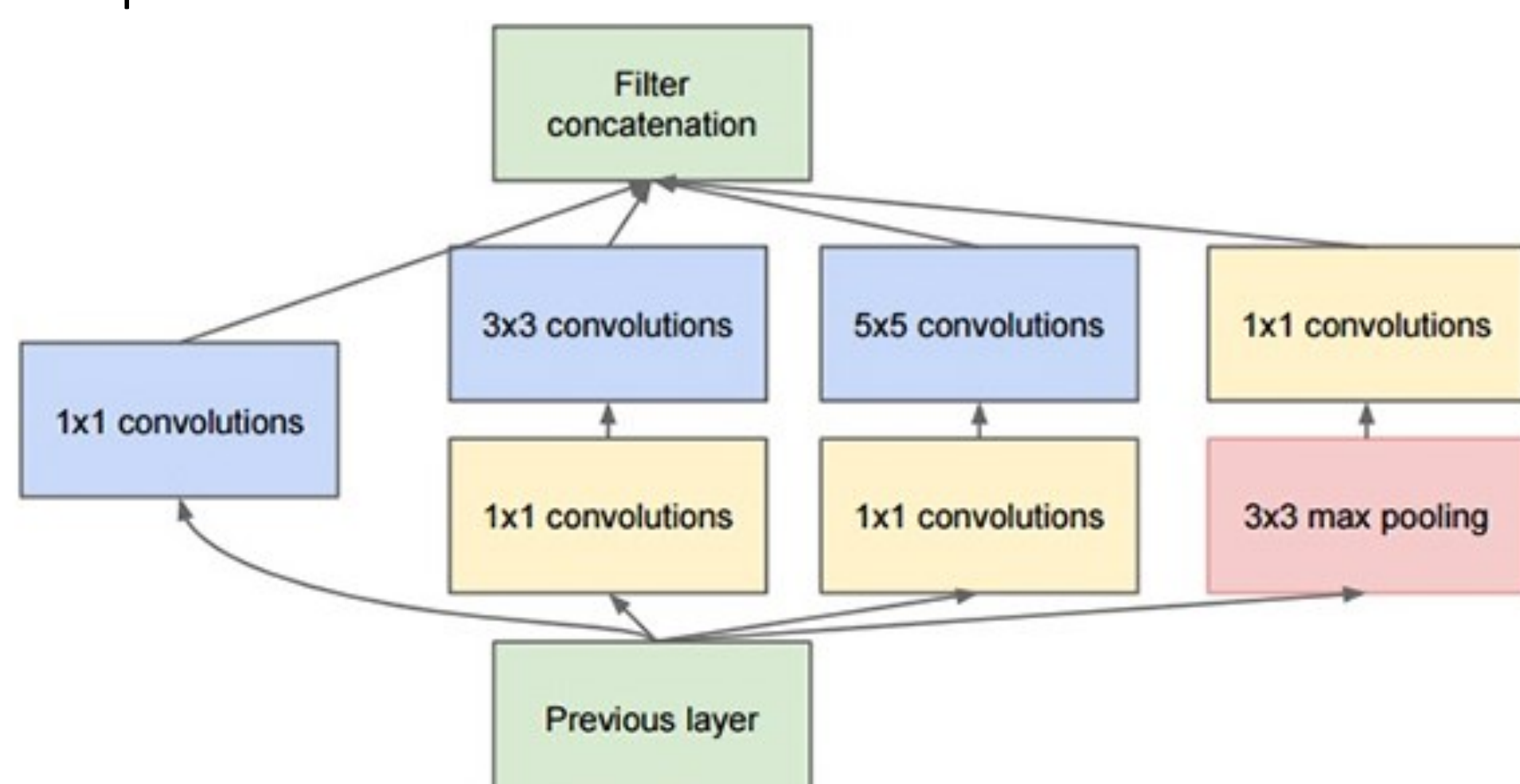
Alexnet[3] is a convolutional neural network designed by Alex Krizhevsky. It contains five convolution layers, some which follow with max-pooling layers, and three fully connected layers.



VGG16[4] is a convolutional neural network designed by Simonyan and Zisserman. It consists of 16 convolutional layers.

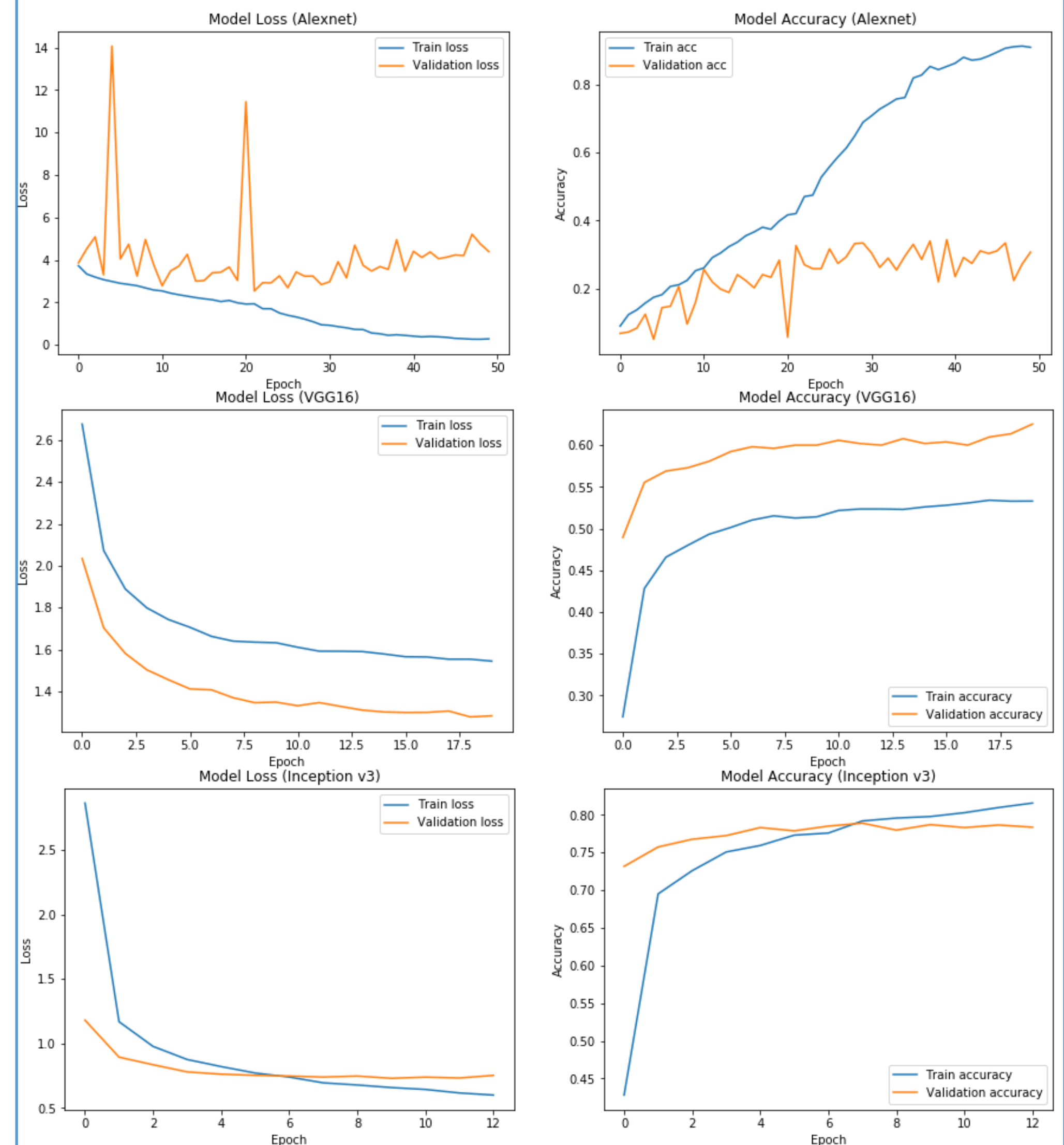


Inception-v3[5] implements an inception module, which is based on small convolutions. This unique module dramatically reduce the number of parameters.



## Results

Models	Alexnet	VGG16	Inception-v3
Train	93.25%	52.39%	93.22%
Test	29.65%	62.09%	90.69%

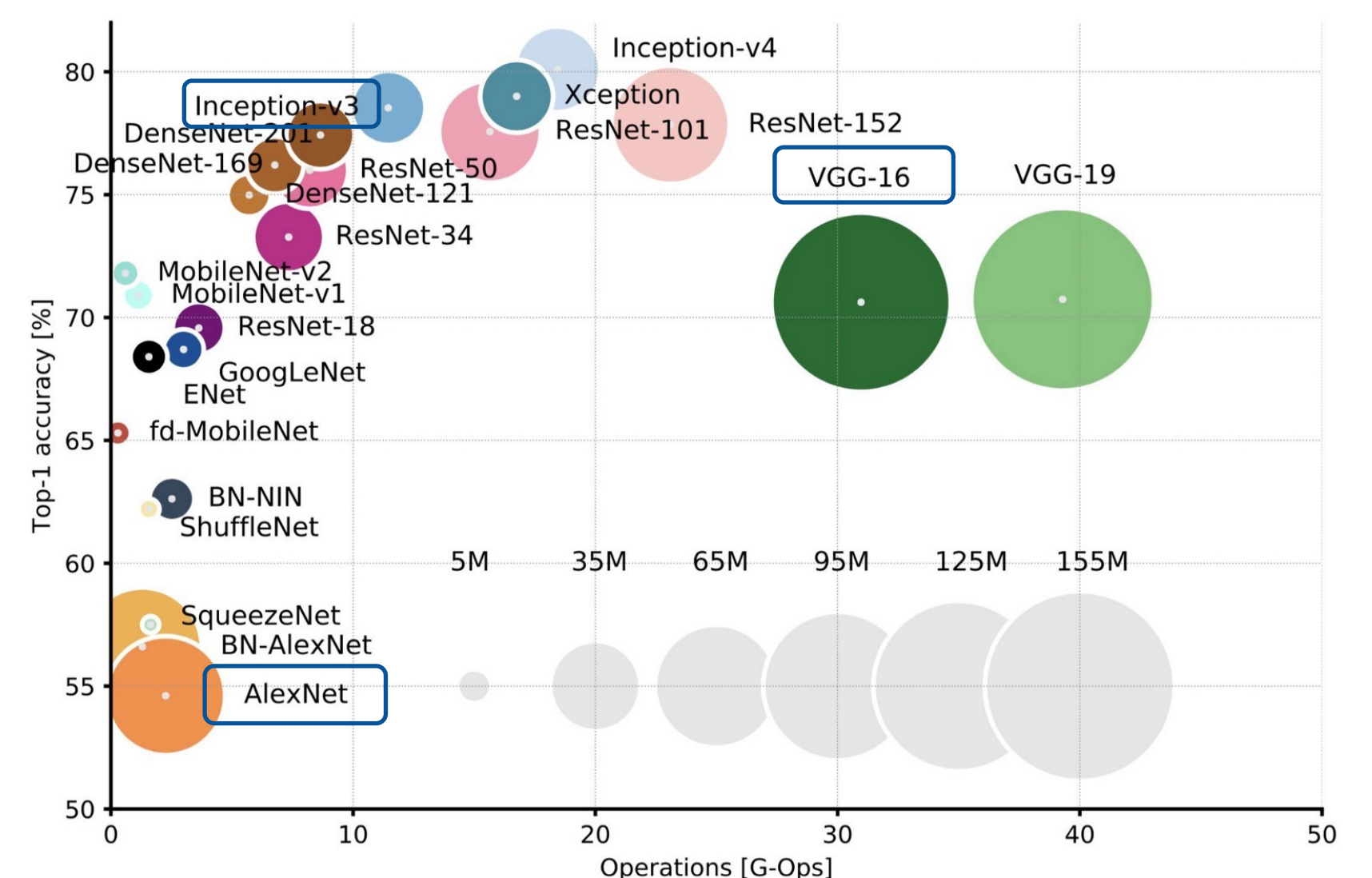


## Discussion

We have applied three models to classify 30 breeds of dogs and we find out that Inception-v3 model performs best. For Alexnet, it is comparably shallow and is overfitting, which leads to low accuracy of 30%. For VGG16, it is much better and we achieved a better performance on validation set as we use the method of dropout. For Inception-v3, it is a relatively latest model thus it achieved the best result.

## Future Work

For future work, we would like to apply new models like VGG19 and Inception-v4 to see if we can get a better result. Also, we would like to dive into the comparison of different models and find out why some work better while others are not.



## References

- <http://vision.stanford.edu/aditya86/ImageNetDogs/>
- J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li and L. Fei-Fei, ImageNet: A Large-Scale Hierarchical Image Database. IEEE Computer Vision and Pattern Recognition (CVPR), 2009.
- Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." *Advances in neural information processing systems*. 2012.
- Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." *arXiv preprint arXiv:1409.1556* (2014).
- Szegedy, Christian, et al. "Rethinking the inception architecture for computer vision." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
- Aditya Khosla, Nityananda Jayadevaprakash, Bangpeng Yao and Li Fei-Fei. Novel dataset for Fine-Grained Image Categorization. First Workshop on Fine-Grained Visual Categorization (FGVC), IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011.