Dog Breed Identification

Presented by
Yue Bi

Outline

- Dataset
- Model
- Experiment Results

Outline

- Dataset
- Model
- Experiment Results

Figure. Stanford Dogs Dataset

Figure. Stanford Dogs Dataset

Image id	Breed
000bec180eb18c7604dcecc8fe0dba07	boston_bull
001513dfcb2ffafc82cccf4d8bbaba97	dingo
001cdf01b096e06d78e9e5112d419397	pekinese
00214f311d5d2247d5dfe4fe24b2303d	bluetick
0021f9ceb3235effd7fcde7f7538ed62	golden_retriever
002211c81b498ef88e1b40b9abf84e1d	bedlington_terrier
00290d3e1fdd27226ba27a8ce248ce85	bedlington_terrier
002a283a315af96eaea0e28e7163b21b	borzoi
003df8b8a8b05244b1d920bb6cf451f9	basenji
0042188c895a2f14ef64a918ed9c7b64	scottish_deerhound

Table. Stanford Dogs Dataset (labels)

Dataset Overview

- Dataset: 10222 images
- Breed: 120
- Pick top 25 frequent breed to classify (2668 images)
- Randomly pick 80% of the images as training set and other 20% as validation set

Order	Breed	Order	Breed
1	scottish_deerhound	14	cairn
2	maltese_dog	15	beagle
3	afghan_hound	16	japanese_spaniel
4	entlebucher	17	australian_terrier
5	bernese_mountain_dog	18	blenheim_spaniel
6	shih-tzu	19	miniature_pinscher
7	great_pyrenees	20	irish_wolfhound
8	pomeranian	21	lakeland_terrier
9	basenji	22	saluki
10	samoyed	23	papillon
11	airedale	24	whippet
12	tibetan_terrier	25	siberian_husky
13	leonberg		

Table. Top 25 frequent breed (labels)

Outline

- Dataset
- Model
- Experiment Results

VGG16* Model

- Prepossessing
- Convolution Parameters
- Activation Function
- Pooling
- Fully-Connected Layers

- **Visual Geometry Group, University of Oxford*

Prepossessing Steps

- Data Augmentation: horizontal flip (double dataset size)

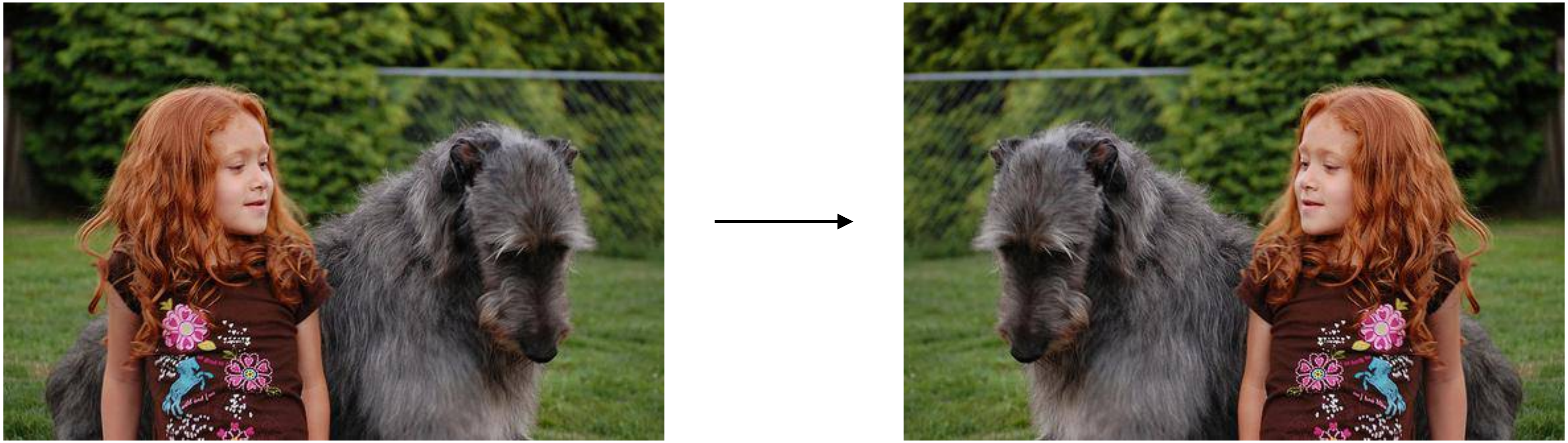


Figure. horizontal flip

Prepossessing Steps

- Fix image size to 224×224

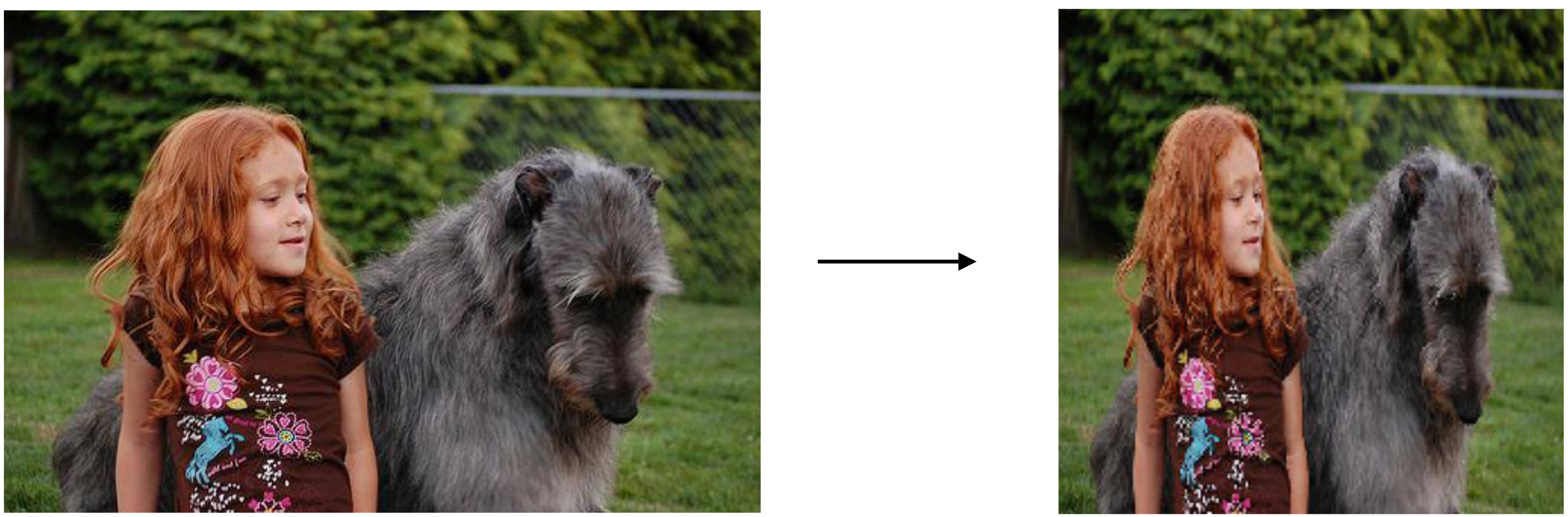
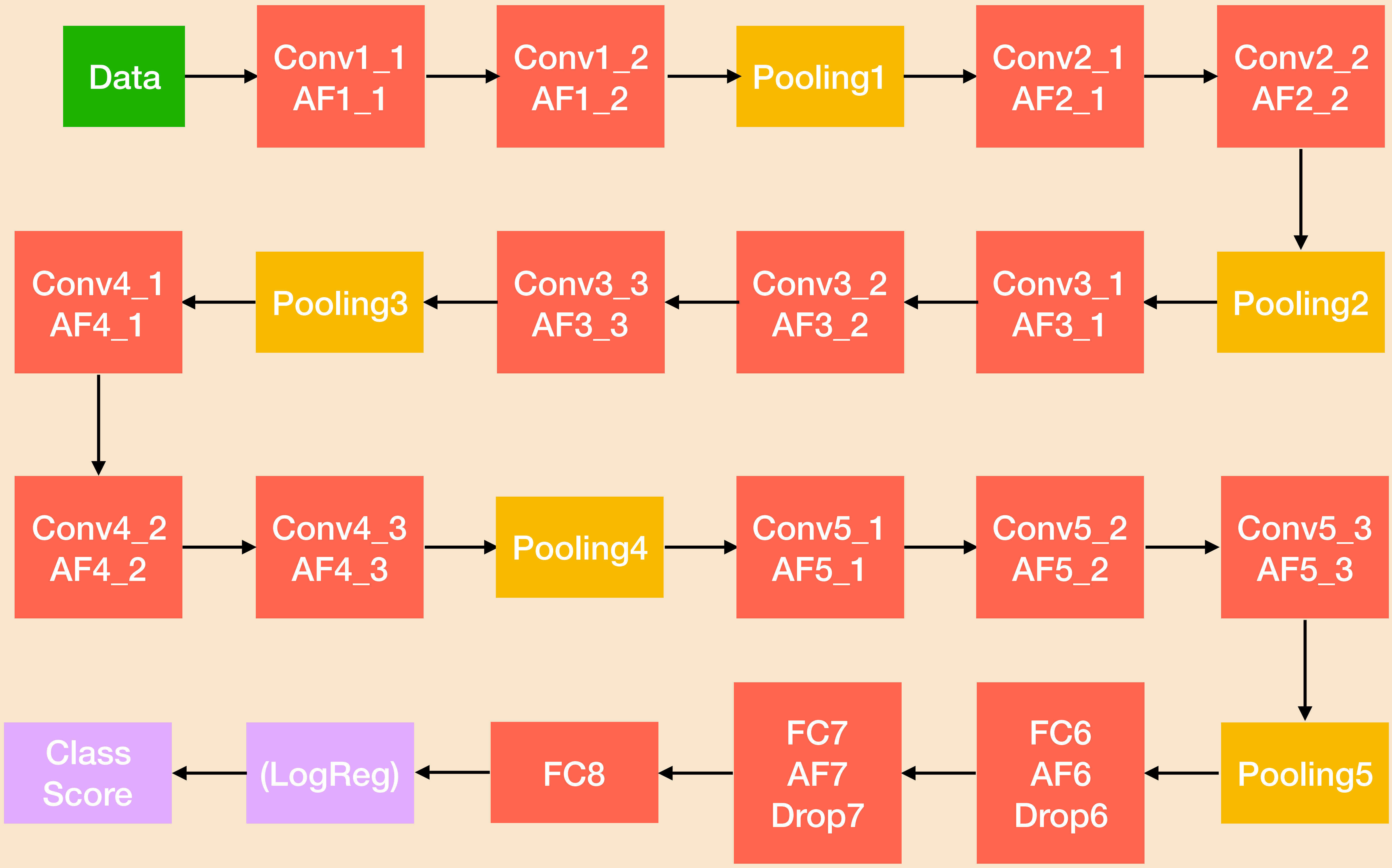


Figure. fix image size



- *Netscope (VGG16)

Convolution Parameters

- Kernel Size: $3 \times 3 \times 3$

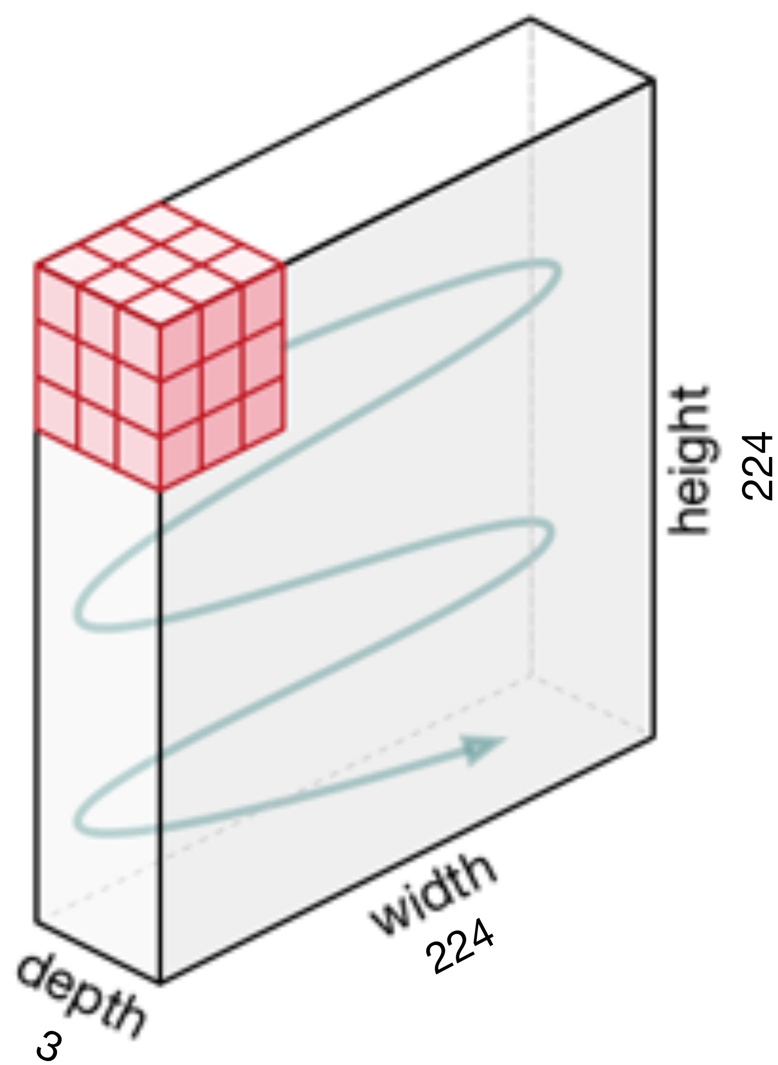


Figure. convolution kernel size

Convolution Parameters

- Stride: 1 pixel

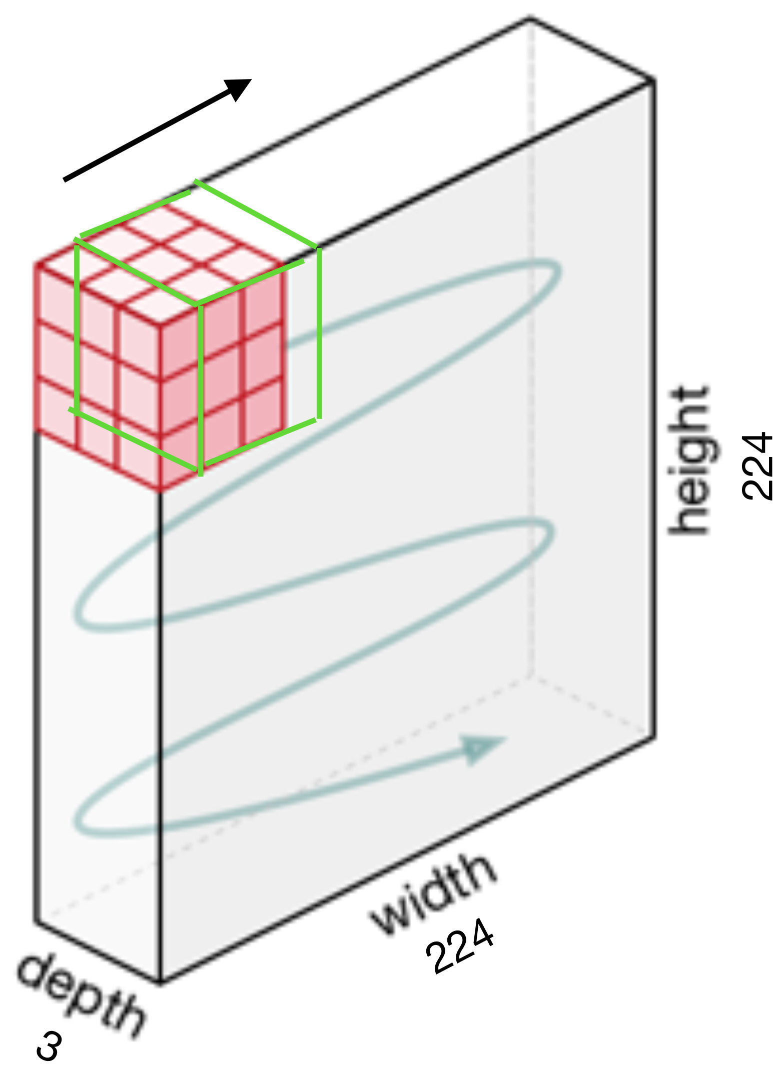


Figure. convolution stride

Activation Function

- ReLU

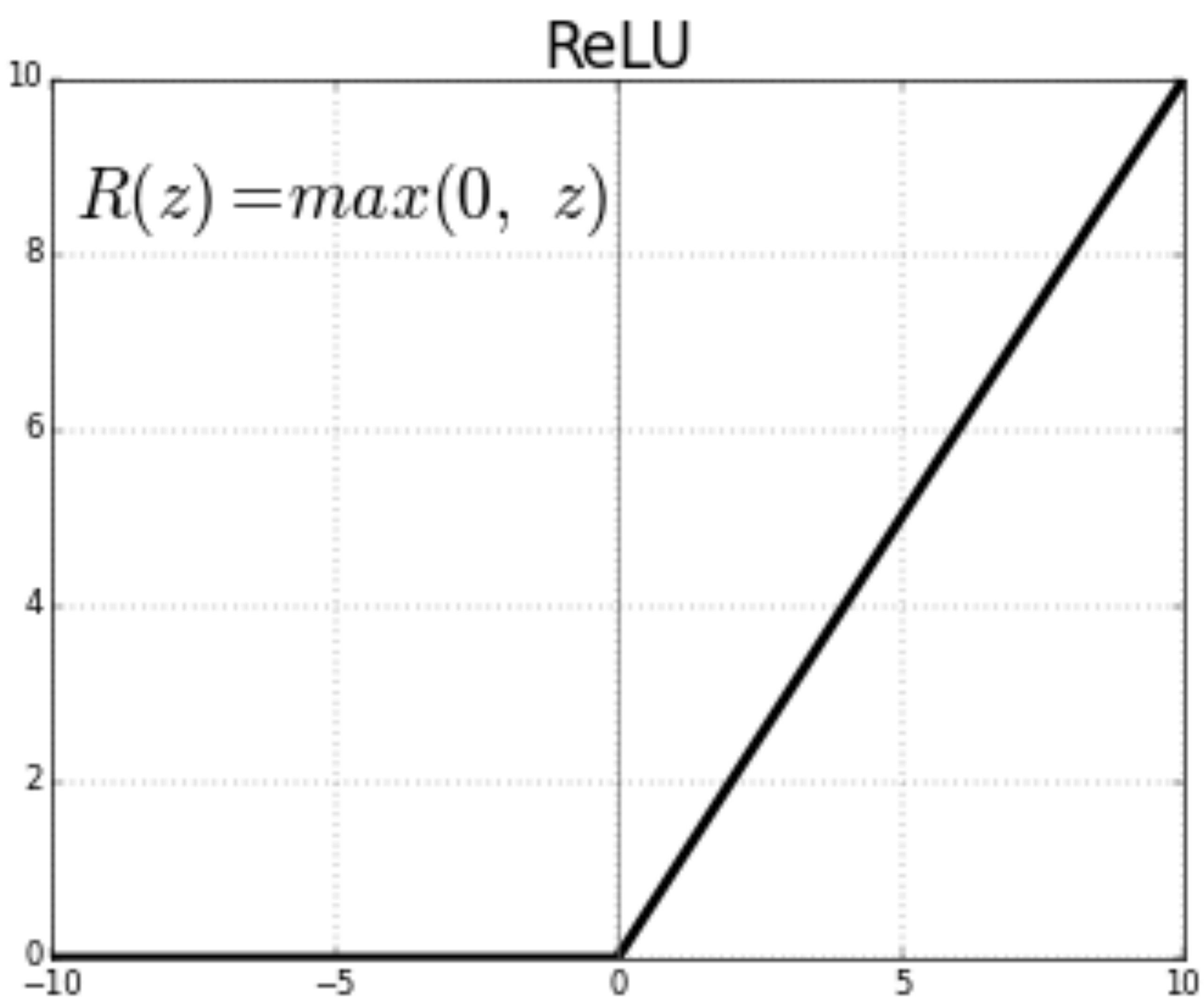


Figure. ReLU function

Pooling

- Max-pooling over a 2×2 pixel window with stride 2.

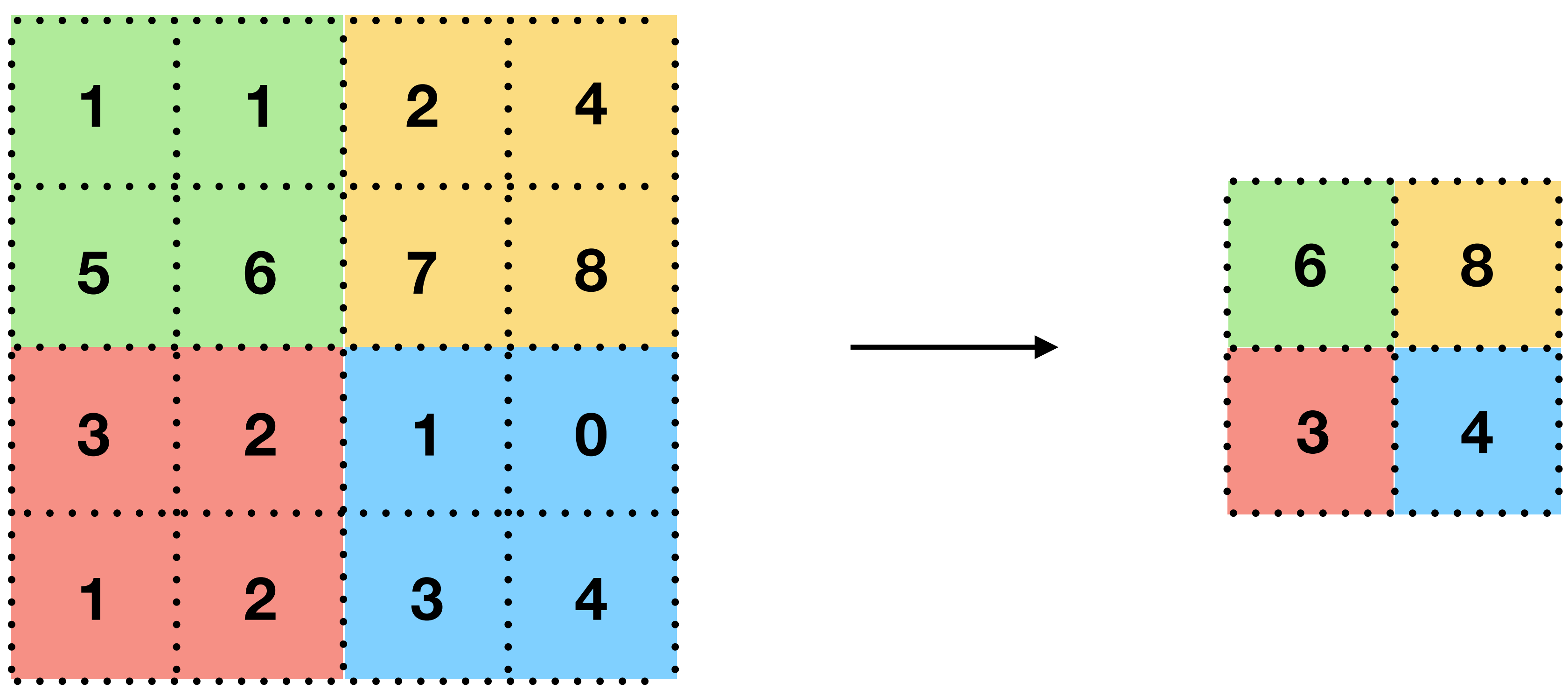


Figure. max-pooling example

Pooling

- Average-pooling over a 2×2 pixel window with stride 2.

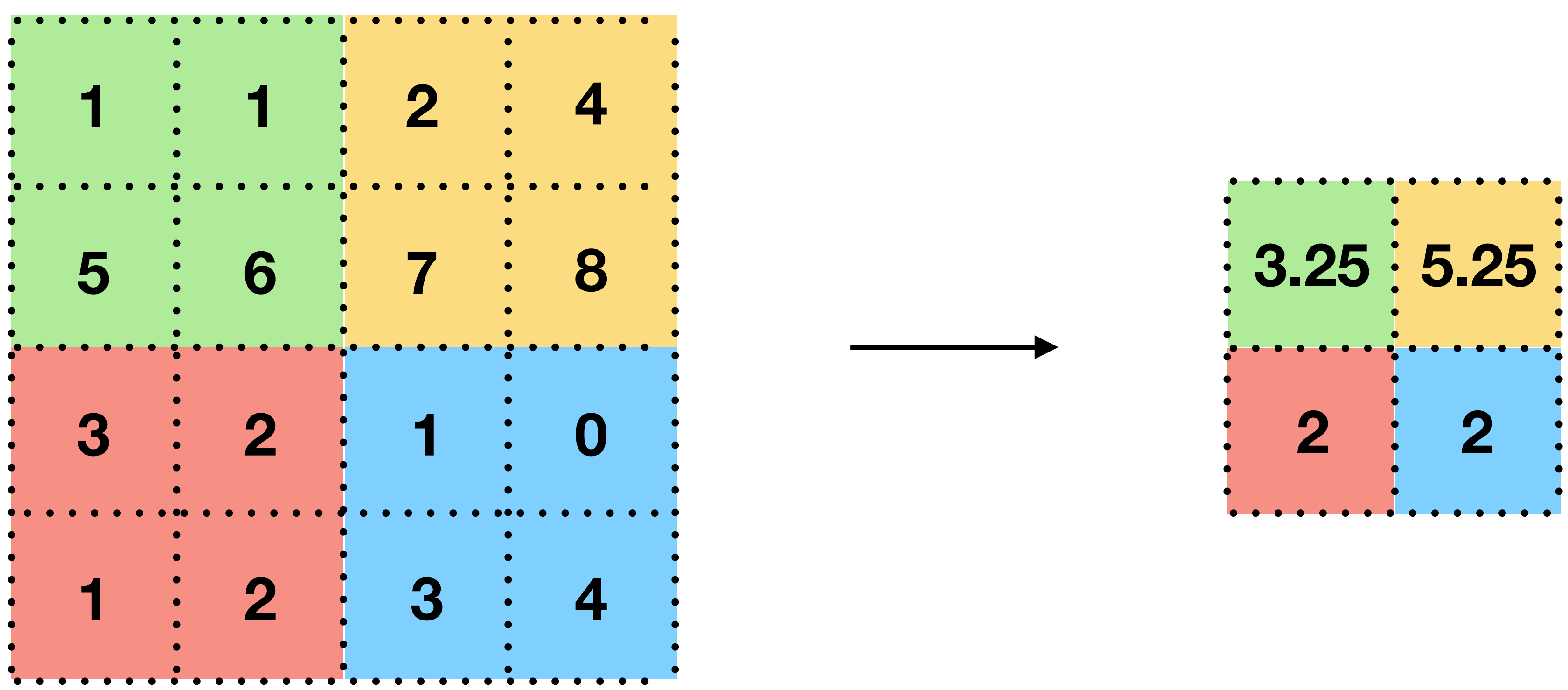


Figure. average-pooling example

Fully-connected Layers

- A stack of convolutional layers is followed by three Fully-Connected layers:

- The first two have 4096 channels each.

- The third contains 1000 channels.

Dropout

- Dropout regularization for the first two fully-connected layers to reduce overfitting.
- Dropout ratio is 0.5.

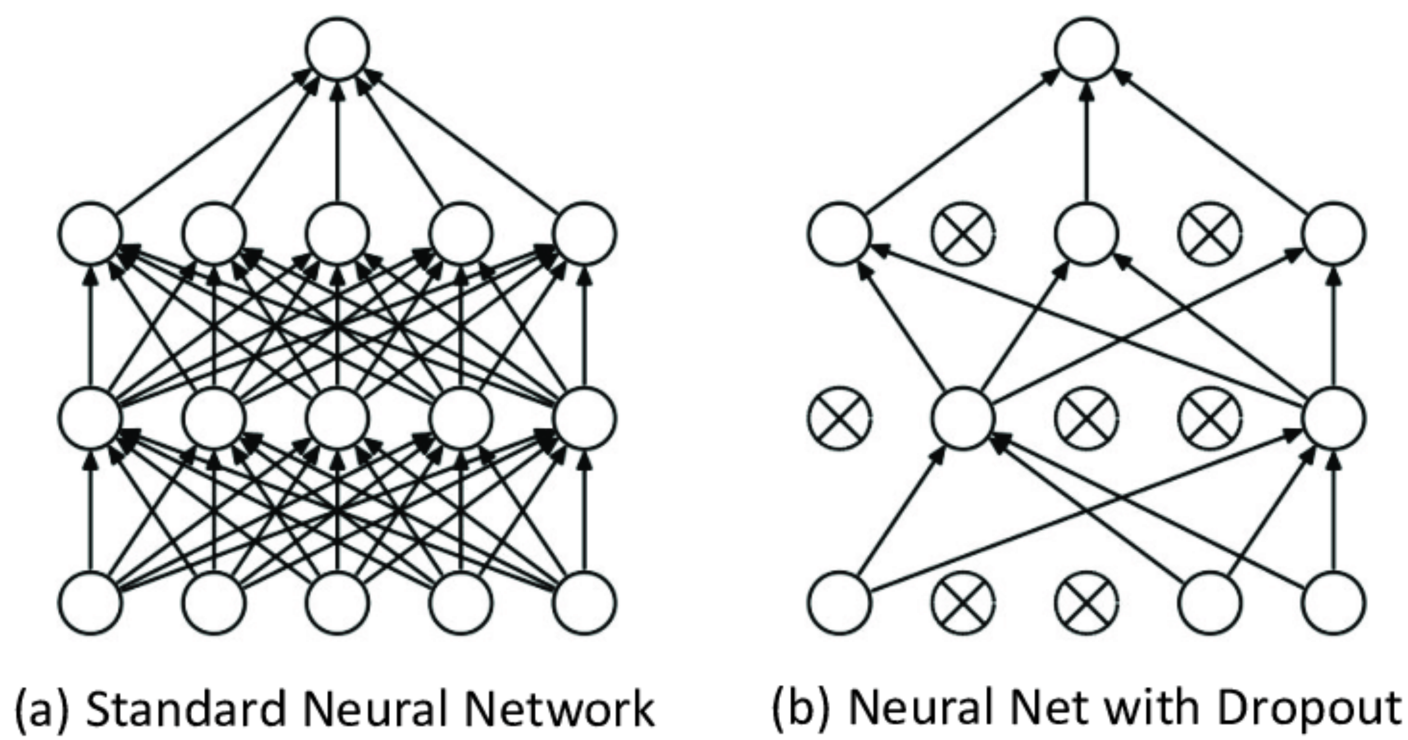


Figure. dropout example

Multinomial Logistic Regression

- Use Multinomial Logistic Regression as the last layer of the network (like softmax)

$$Pr(Y_i = 1) = \frac{e^{\beta_1 X_i}}{\sum_{k=1}^{25} e^{\beta_k X_i}} \qquad Pr(Y_i = 2) = \frac{e^{\beta_2 X_i}}{\sum_{k=1}^{25} e^{\beta_k X_i}} \qquad \cdots \qquad Pr(Y_i = 25) = \frac{e^{\beta_{25} X_i}}{\sum_{k=1}^{25} e^{\beta_k X_i}}$$

$$L(b_k) = \sum_{i=1}^{n} \left(\frac{e^{X_i b_k}}{1 + e^{X_i b_k}}\right)^{Y_i} \left(\frac{1}{1 + e^{X_i b_k}}\right)^{1-Y_i}$$

$$\hat{\beta}_k = argmax_{\,b_k}\ L(b_k)$$

Outline

- Dataset
- Model
- Experiment Results

Results

Max-pooling		Average-pooling	
Accuracy	0.874	Accuracy	0.866

Table. Results of 25 classes
(Training set: 4228 images, Test set: 1108 images)

Results of Different Number of Classes

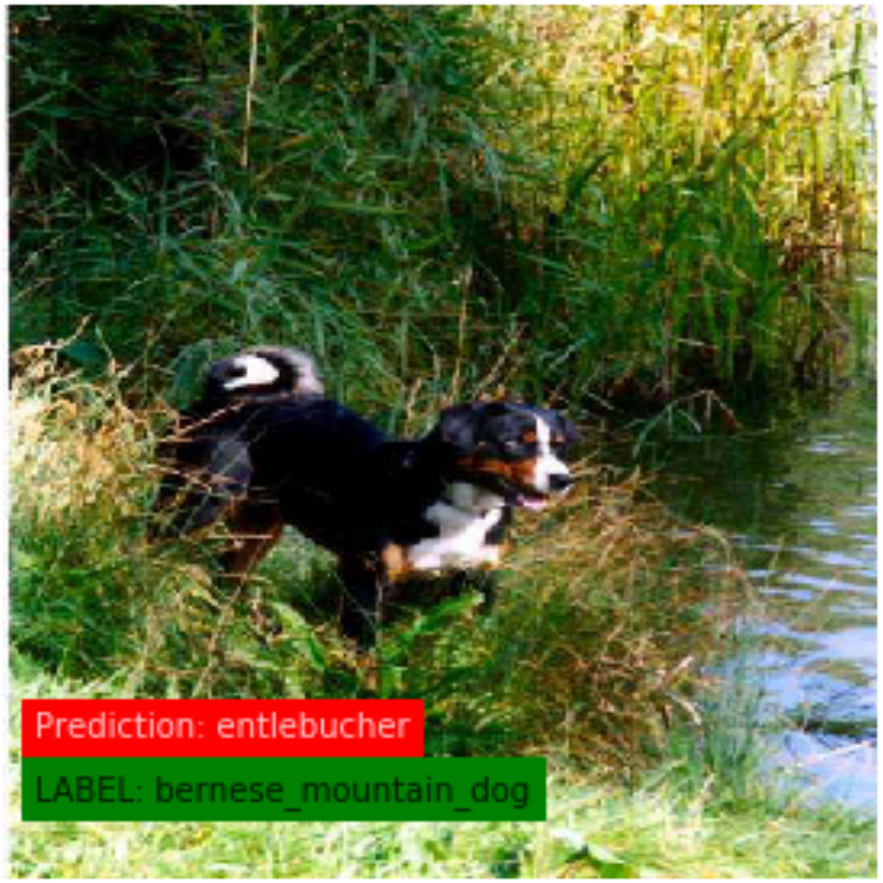
# of class	25	12	6
Accuracy	0.874	0.921	0.946
Training set size	4228	2126	1104

Table. Results of different number of classes (Max-pooling)

Show Errors

Show Errors

Show Errors



Future Work

- Detect mixture of breeds
- Help determine the breed of dog from shelter easily

Thank you !

References

- [1]: Karen Simonyan, Andrew Zisserman. “Very deep convolutional networks for large-scale image recognition.” ICLR, 2015.
- [2]: https://www.kaggle.com/gaborfodor/keras-pretrained-models
- [3]: https://keras.io/applications/#vgg16
- [4]: https://en.wikipedia.org/wiki/Multinomial_logistic_regression