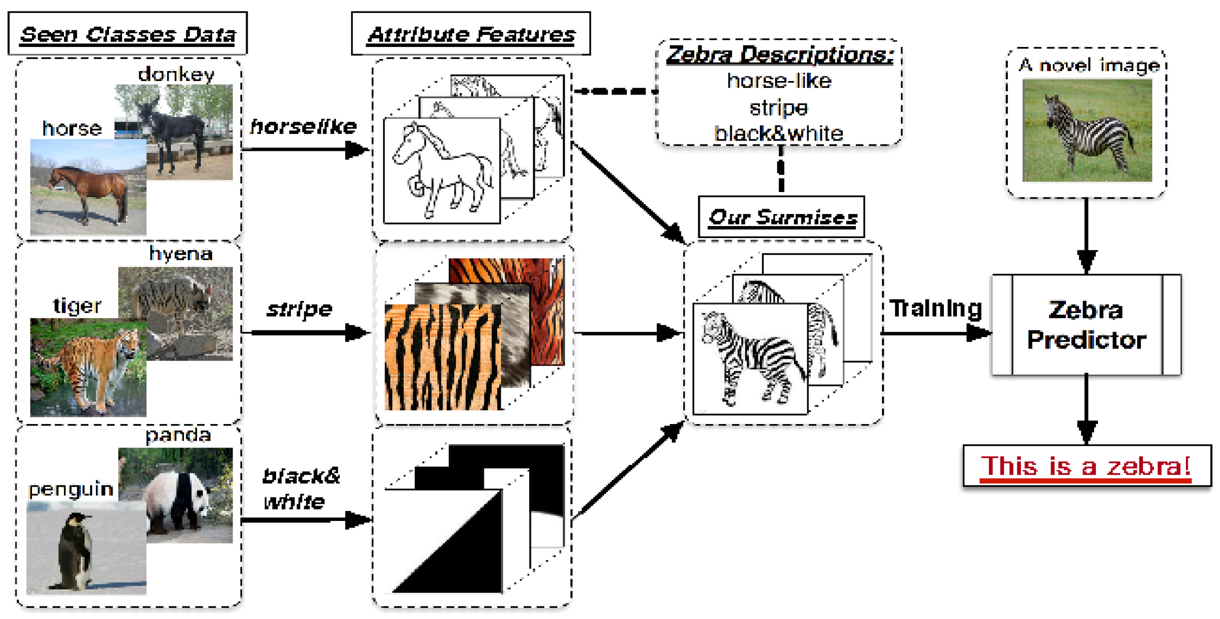
Attribute Vs Classification accuracy in Zero-shot Learning

Akash Ramesh Reddy, Gokulakrishnan Swaminathan, Mayur Kishor Deo

School of Industrial Engineering, Purdue University.

Problem Statement

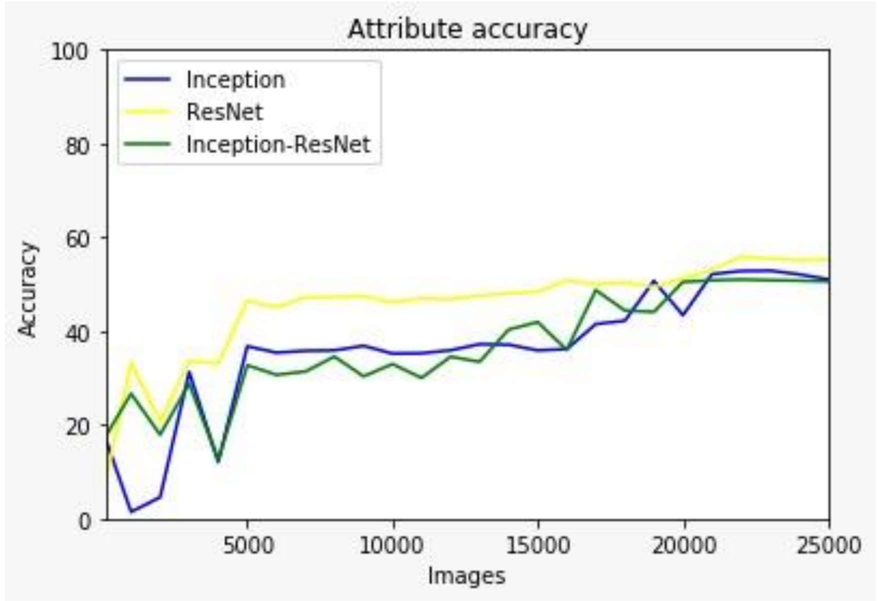
- Many deep learning methods focus on classifying instances whose classes have already been seen by the model during training.
- Zero-shot learning (ZSL) is a promising learning method wherein instances classified during testing are unlabeled.



- ZSL consists in recognizing new categories of instances without training examples, by providing high-level description of the new categories and using them with the previously extracted features that relate them to categories previously learned by the model.

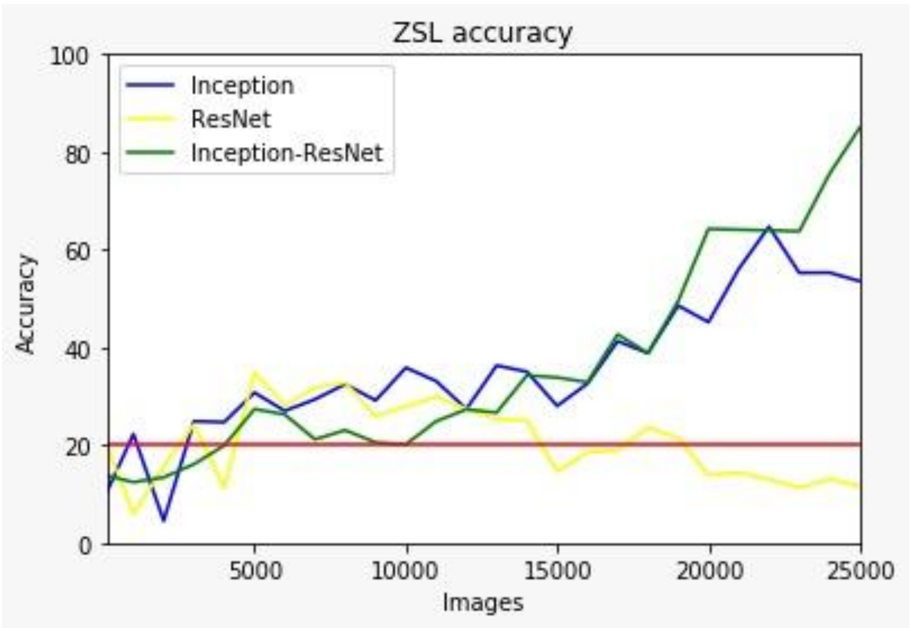
Results

Attribute accuracy obtained from the various architectures is plotted in the graph below:



- Attribute accuracy is highest for ResNet architecture.
- Inception follows ResNet with lesser attribute accuracy.
- Attribute accuracy is lowest for Inception-ResNet.

ZSL accuracy obtained from the various architectures is plotted in the graph below:



- ZSl accuracy is highest for Inception-ResNet architecture.
- Inception architecture follows closely to the Inception-ResNet model.
- ZSL accuracy is the lowest for ResNet .

Acknowledgements

We would like to thank Professor Juan P. Wachs and our student mentor Mr. Naveen Madapana for providing the incredible opportunity and guidance throughout the course of the project.

Approach and/or Methodology

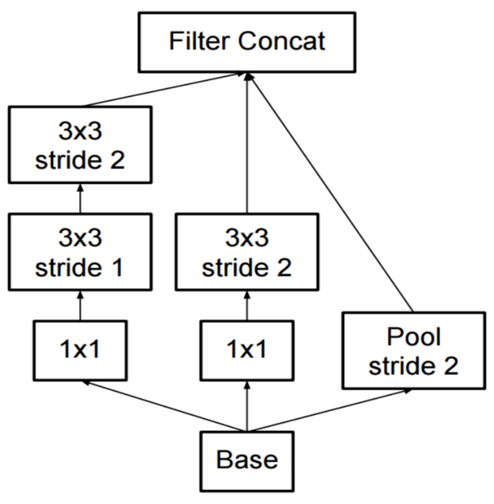
- Dataset chosen for our model is the Animals with Attributes (AwA) dataset which contains 50 classes of animal images.
- Of the 50 classes, 40 are used for training the model and 10 classes are used for testing.
- Along with the dataset, we obtain a predicate matrix [50 x 85] which contains the information of attributes of all the classes.

Predicate Matrix:

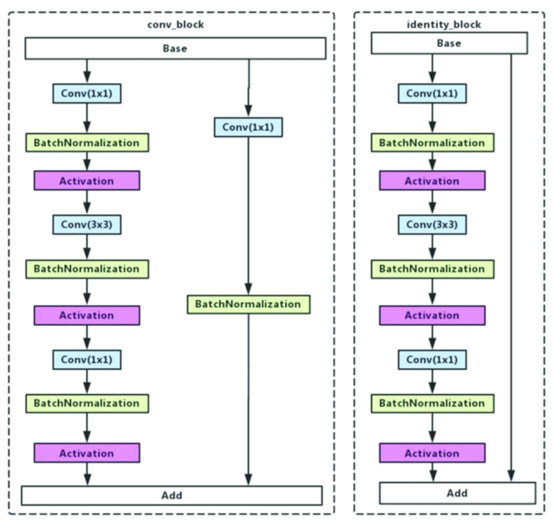
Black-0 (attribute not possessed by class). White-1(attribute possessed by class).

Back-end functions are used to extract layer outputs from model.

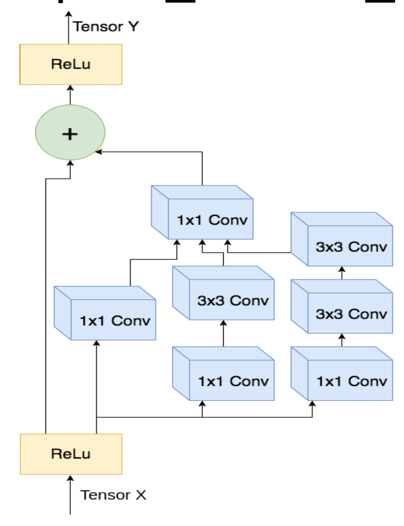
Inception_v3



ResNet_50 Architecture



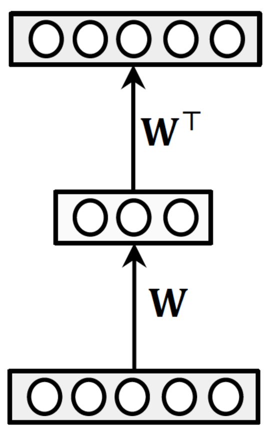
Inception_ResNet_v2



Classification-Algorithm

```
function W = SAE(X,S,lambda)
  % SAE - Semantic AutoEncoder
  % Input:
  %    X: dxN data matrix.
  %    S: kxN semantic matrix.
  %    lambda: regularisation parameter.
  %
  % Return:
  %    W: kxd projection matrix.

  A = S*S';
  B = lambda*X*X';
  C = (1+lambda)*S*X';
  W = sylvester(A,B,C);
end
```



Future Work

With the output obtained from the model and the features input by the user, we can create new images by de-convolution approaches.