

Penguin colony detection with deep learning.

Hieu Le

Computer Science Department, Stony Brook University
{hle}@cs.stonybrook.edu

1 2nd-May-2018 — Exp: Performance of the model as the amount of training data grows.

The experiments are conducted 3 times independently. For each trial, all data is randomly divided into 5 folds. The last chunk of data is kept untouched for testing. The first fold is used for training the first model. The first and second are used for training the second model. Likewise, the third model is trained on three chunk of data and the fourth model trained on four. Thus, from the first model to the fourth model, each is introduced with more training data. All other settings are the same for all models.

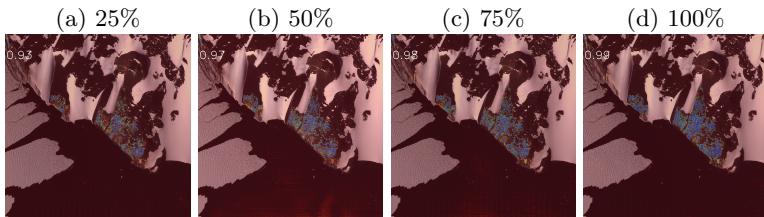


Fig. 1. Performance of our model with different amount of training data.

2 2nd-May-2018 — Exp: Varying dropout

Dropout is a technique used in deep learning to reduce overfitting. For each training iteration, a percentage of units in each layer are selected to be deactivated as dropout means "dropping out units". This prevent complex co-adaptation on training data, or naively memorizing the training data, enforcing each unit to independently learn the features.

To understand the effect of dropout, we conduct two set of experiment with dropout = 0.5 and 0.2 respectively.