Evolutive deep models for online learning on data streams with no storage

Y. Choi

2018.09.22

Introduction

- Deep learning community has been focused on the case of static datasets.
- There is a demand for real-time processing and analysis of continuously arriving huge amounts of data.
- These models often fail to match the required detection rates.

Contributions

1. Train online classifiers without storing the incoming data while being able to adapt to new classes.
2. Make quantitative evaluation of how the replacement of real data by generated data.

Algorithm

- Let S_i be the data whose label is i.

Require: $S = \bigcup_{i=1}^{\infty} S_i$: data stream, with i - class number
Require: n : number of already learned classes
Require: G_i : generative model for class i
Require: C_1^n : classification model for data from $\bigcup_{i=1}^{n} S_i$

$G_1 \Leftarrow$ initialize model
$n \leftarrow 1$
while are receiving samples from S **do**
  $d \Leftarrow$ get batch from S_j, j - current class
  if $j = n + 1$ **then**
    $n \leftarrow n + 1$
    $G_n, C_1^n \Leftarrow$ initialize models
    if $n > 2$ **then**
      $C_1^n \Leftarrow$ copy parameters from C_1^{n-1}
    end if
  end if
  $d^* \leftarrow \bigcup_{i=1..n}^{i \neq j} d_i^*$ generate synthetic data from $\{G_i\}$
  $C_1^n \Leftarrow$ train with $d \bigcup d^*$
  $G_j \Leftarrow$ train with d
end while

Algorithm 1: Online learning model, proposed in Sec. 3

Schematic representation of proposed appoach

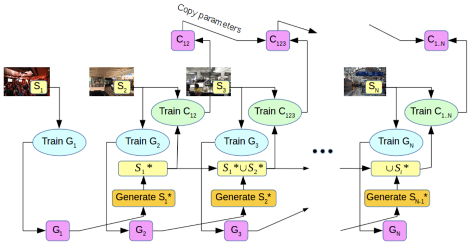


Fig. 1: Schematic representation of our online learning approach. Original data is presented to the model class by class. Each time new class of data appears we start training a new generator modeling that class. At the same time we train a classifier on the generated data from the previously learned classes and the original data from the new class that come from the stream.

Add labels

- Add a node to the out layer of the classifier.
- Initialize its connections with the previous layer.
- The rest of the network weights are copied from the previous state.

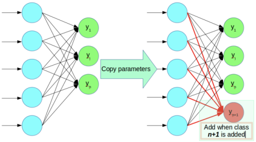


Fig. 2: Adding a node to the output layer and initializing the connections with the previous layer in the online learning scenario when new data class appears in the stream.

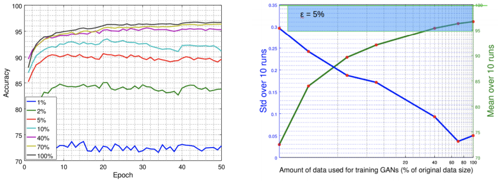
Questions

1. How well the generated data represents true data?
2. How much generated data is needed?

The measure of generalizability of generative model

- Let G be a generative model and D be a dataset.
- D^{val} is validation set of D
- μ is some similarity measure
- $\left|\mu(D \setminus D^{val}, D^{val}) - \mu(D^{gen}, D^{val})\right| < \epsilon$

The measure of generalizability of generative model



(a) (b)

Fig. 3: Results of the generalizability test on MNIST (see Sec. 4.1). (a) Classification accuracy for different GANs support sizes as a function of training time. Average over 10 runs; (b) Mean/std of the classification accuracies for different GANs support sizes over 10 runs after 50 training epochs for the generalizability tests. Blue box represents the area in which the generalization error does not exceed 5%

Representativity of generative models

- How much data do we need to represent the full richness of the information?