Paper Review - Assignment - Unit Test 2

1. Students are requested to select the research paper related to the course CS 365 Machine Learning and application in the field of respecitve discipline of the student strictly.
2. The paper should be selected from reputed conference or journal only. The journal is preferred.
3. Review the paper, analyze it thoroughly. Answer the following points.
4. One student from each department should coordinate the group of respective department students while selecting the paper to avoid the selection of same paper. It is student responsibility only student to avoid the duplication of paper for assignment.
5. If any students' papers found same, zero unit test assignment marks will be assgined and the second chance will not be given to the student.

* Required

Name of the student *

Bhagya Vinod Rana

Admission number of the student *

U19CS012

Deaprtment of the student *

Computer Science

Area of the paper selected *

Machine Learning & Computer Vision and Patt

Title of the paper *

Unbiased Teacher for Semi-Supervised Object

Is it conference paper? *

- (●) Yes
- ( ) No

Is it journal paper? *

- ( ) Yes
- (●) No

Attach a soft copy of the paper (file of paper uploaded should be name with Admission-number-Unit-Test-Assignment-Two.pdf) *

PDF U19CS012_Unit_... ×

What is the problem addressed in the research paper? *

Problem Definition: Our goal is to address obje

Applications of the problem addressed in the research paper. *

1.) The Classification of Content on the Internet
This may be applied to web page categorization. You can utilize a group of labelled web sites to forecast the labels of all the additional online pages you'll need.
A semi-supervised learning model is used by several search engines, including Google, to categorize and rank web pages in their search results. [Paper Link: http://www.cs.cmu.edu/~avrim/Papers/cotrain.pdf]
Google, in 2016 launched a new Semi-Supervised learning tool called Google Expander [https://ai.googleblog.com/2016/10/graph-powered-machine-learning-at-google.html]

2.) Speech Analysis
Since labeling of audio files is a very intensive task, Semi-Supervised learning is a very natural approach to solve this problem.

3.) Classification of Protein Sequences:
Protein sequences include a large number of amino acids, making it impossible to study each one and identify it as one of two types. With the help of semi-supervised learning, this activity may be accomplished quickly. All you need is a database of already sequenced proteins, and the model will take care of the rest.

Literature summary with key papers. For summary name the major algorithmic approaches reported in the literature for addressing the same problem. *

The Unbias Teacher, a work presented at ICLR 2021 by Facebook, examines the significant distinctions between semi-supervised object detection (SS-OD) and traditional semi-supervised tasks (such as image classification).

When the standard semi-supervised approach is applied to object detection, it is impacted by class imbalance and produces pseudo-labeling bias, which causes performance to deteriorate. The research argues that by combining the progressively updated Teacher network with Focal Loss, the state-of-the-art (SOTA) may be effectively enhanced, and the SOTA can be renewed with significant progress.

The overall practice method is probably the combination of STAC [1], Mean Teacher [2] and Focal Loss [3]. Through the progressive Teacher network and Focal Loss, the paper proved experimentally to reduce the negative effects of class imbalance in semi-supervised object detection.

[1] A Simple Semi-Supervised Learning Framework for Object Detection [https://arxiv.org/abs/2005.04757]
[2] Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results [https://arxiv.org/abs/1703.01780]
[3] Focal Loss for Dense Object Detection [https://arxiv.org/abs/1708.02002]

Logical discussion of the apporach for the problem addressed in the research paper. (Logical solution to the problem addressed) *

It's well known that more data = better quality models in deep learning (up to a certain limit obviously, but most of the time we don't have that much data). But getting labeled data is expensive. If you want to train a model to identify birds, you can set up a bunch of cameras to automatically take pictures of birds. That's relatively cheap.

Hiring people to label those pictures is expensive. What if you have a large number of pictures of birds, but only hire people to label a small subset of the pictures?

As it turned out, instead of just training the models on the labeled subset, you can pre-train the model on the entire training set, before fine-tuning it with the labeled subset, and you get better performance this way. That's semi-supervised learning. It saves you money.

Unbiased Instructor, a simple but efficient method for training a student and a steadily improving teacher in a mutually beneficial manner. Unbiased Teacher consistently outperformed state-of-the-art approaches on COCO-standard, COCO-additional, and VOC datasets, using a class-balance loss to downweight excessively confident pseudo-labels.

Technical solution description for the problem addressed in the research paper. *

The training process of Unbias Teacher is as follows:
1.) Use labeled data to train the Teacher network (and use Focal Loss). The paper calls this step burn-in.
2.) Copy the network parameters in two copies and use them as the Teacher and Student networks.
3.) Using data augmentation of different intensities, the pseudo-label will be obtained through the unlabeled data of weak Aug. through the Teacher network.
4.) The same unlabeled data passes strong Aug., and pseudo-label composes label data, and performs supervised learning on the Student network (using Focal Loss as well).
5.) Refer to the parameters of the Student network and update the parameters of the Teacher network with EMA (exponential moving average).

Neat Diagram for Visualizing the Whole Process - [https://gifyu.com/image/S2bjc]

Formulas and equations in support of the technical solution description with explaination should be handwritten on A4 white paper, signed and admission number should be written on top - left corner of the paper and should be uploaded as part of response to this point. *

PDF U19CS012_ML_P...

Simuation - data set selected, characteristic of the data set, or data set collected *

1.) Data Set selected - COCO dataset [Common Objects in Context]

Following three datasets are used:
1.) COCO-standard: Take a small part from MS-COCO's labeled data, and the rest are regarded as unlabeled data.
2.) COCO-addiction: Use all MS-COCO labeled data, and use COCO2017-unlabeled as unlabeled data.
3.) VOC: Use VOC07 trainval as labeled data, and use VOC12 trainval as unlabeled data.

On the whole, the experimental data of Unbias Student significantly surpassed other previous methods.

```
# For downloading images
wget http://images.cocodataset.org/zips/train2017.zip
wget http://images.cocodataset.org/zips/val2017.zip
# For downloading annotations
wget http://images.cocodataset.org/annotations/annotations_trainval2017.zip
```

2.) Characteristic of the data set
COCO is a large-scale object detection, segmentation, and captioning dataset.
COCO has several features:

- Object segmentation
- Recognition in context
- Superpixel stuff segmentation
- 330K images (>200K labeled)
- 1.5 million object instances
- 80 object categories
- 5 captions per image
- 250,000 people with keypoints

Simulation - tools used - programming or ready to use software *

Linux or macOS with Python ≥ 3.6, PyTorch ≥ 1

Objective parameter used for analysis *

1.) Focal Loss
Focal Loss was published in 2017. The core idea is to reduce the loss weight of simple samples. For object detection, the points falling on the background are numerous and easy to classify. The goal of Focal Loss is to avoid paying too much attention to these easy examples and to focus on hard examples.

Foal loss and cross entropy [https://ibb.co/sVkVN7P]

2.) EMA Update
Speaking of such a symmetrical EMA method, the first thing I think of is not Mean Teacher, but self-supervised learning MoCo (Momentum Contrast) [6] and BYOL (Bootstrap Your Own Latent) [7]. To put it simply, for a group of symmetric networks, use gradient decent training for one of them, and the other parameters follow the parameters of the first encode, but use momentum to update.

Well-known method in the field of self-supervised learning: BYOL [https://ibb.co/NN71264]

In fact, in addition to the method of self-supervised learning, EMA-related concepts are also very important in the field of deep learning. For example, the concepts of EMA are included in Adam optimizer and batch normalization.

Discussion on the simulation performed and results obtained. *

The experiment of the paper can be divided into two scenarios: limited labeled data and more unlabeled data

A.) Limited labeled data

The first part of the experiment is to train Faster-RCNN with limited labeled data. In the experimental results, Unbias Teacher has won a great victory in the proportion of various labeled data, and the less the labeled data, the more obvious the effect.

The performance of Unbias Teacher and other methods on COCO-standard [https://ibb.co/Tqt9RJm]

The paper attributed this effect to two points:
1.) The paper believes that updating the Teacher network through EMA will make the entire structure smoother and obtain a more accurate pseudo-label.
2.) EMA and Focal Loss alleviate class imbalance problems.

B.) More unlabeled data

The second part of the experiment is to use a sufficient amount of labeled data, accompanied by more unlabeled data, to see how much effect improvement can be obtained. Unbias Teacher is still strong in terms of data, showing all aspects of performance improvement in terms of final results and convergence speed.

The performance of Unbias Teacher and other methods on COCO-addictional [https://ibb.co/9tVmBMh]

The performance of Unbias Teacher and other methods on VOC [https://ibb.co/Rh4dVFN]

Under the design of EMA, the Teacher network can actually be regarded as an ensembled model of the Student network at different times. The paper believes that this can relieve the teacher network from generating bias (class imbalance) or noisy pseudo-label. As for Focal Loss, it is natural to improve class imbalance.

Conclusion (in your words not to be copied the same conclusion mentioned in the paper) *

Judging from the experimental data provided by the paper, the addition of EMA not only improves mAP, but also makes the training process more stable. The Focal Loss is a more significant impact on the final mAP, and even can be said to significantly reduce the model-Fitting over. On the other hand, the paper also uses the KL-divergence between the real label and the pseudo-label to indicate the degree of class imbalance. Also, the Focal Loss influence is more pronounced than EMA.

The effect of EMA Teacher and Focal Loss [https://i.ibb.co/884CNv4/The-effect-of-EMA-Teacher-and-Focal-Loss.png]

In addition, a very intuitive and powerful proof of the thesis is that the number of Pseudo-labels of Unbias Teacher has increased significantly, in other words, it reduces the state of class imbalance in the context.

Unbias Teacher can indeed predict more objects [https://i.ibb.co/qj9THb8/Unbias-Teacher-can-indeed-predict-more-objects.png]

On the whole, the core idea of Unbias Teacher is similar to the recent SOTA method Meta Pseudo Labels of image classification. Unbiased Teacher is a holistic framework that consists of a Teacher and a Student who work together to better each other.

We show that our approach eliminates pseudo-labeling bias owing to class imbalance and overfitting due to labelled data scarcity in the trials. Across numerous semi-supervised object identification datasets, our Unbiased Teacher performs admirably.

Your ideas to solve the same problem using any other techniques or approach learned in the subject with detail description. *

Judging from the results, Unbias Teacher has undoubtedly refreshed the SOTA of SS-OD, and the components used in it have also been proven effective through multiple studies.

Although it is a pity that there is no clear explanation of the relationship between EMA and alleviating class imbalance, and it is indeed as pointed out that there are insufficient innovations at the time of submission, but after all, it has opened the research road for all researchers to explore the details of SS-OD.

After all, in the past - The largest study of supervised in the field of computer vision is still in image classification. I believe that in the future, in addition to class imbalance, other key factors of semi-supervised in object detection will definitely be decrypted step by step.

Submit Clear form