# CSE 584 – MACHINE LEARNING: TOOLS AND ALGORITHMS

## HOMEWORK-1

Divya Navuluri
PSU ID: 903661629

---

## PAPER 1: A Reinforced Active Learning Algorithm for Semantic Segmentation in Complex Imaging.

### 1. What problem does this paper try to solve?

*"A Reinforced Active Learning Algorithm for Semantic Segmentation in Complex Imaging"* highlights the two main problems in semantic segmentation:
(i) Mainly, the high costs and time consumption associated with pixel-wise labeling.
(ii) The datasets used for semantic segmentation tasks are unbalanced, as some classes are more frequently represented than others.

### 2. How does it solve the problem?

To solve the problem, the paper gave an overview of a new method to a reinforced active learning method that utilizes deep reinforcement learning algorithm. This method is an improvement of the Deep Q Learning procedure, adapted specifically for learning from images. This is done with the goal of training a computer to identify and select small yet crucial details from a pool of unlabeled images. These selected regions are considered more informative for the model which minimizes the need to label all the images while at the same time handling of class imbalance by asking for more labels to the underrepresented classes.

This method is effective in optimizing the efficiency because as the large quantity of images suggests it does not concentrate on the overall image set, which is common in conventional methodologies, but rather in the several measurable significant areas connected with images. This strategy enables the enhancement of the segmentation model with fewer labeled examples; thus, reduction of the annotation costs as well as compensating for the data imbalance problem by giving the under-represented classes appropriate representation.

### 3. A list of novelties/contributions?

- The authors introduced a new form of reinforced active learning technique for semantic segmentation problems which they cast in terms of a Markov Decision Process. This enables the learning of a strategy for a particular agent to select informative image regions from a particular unlabeled data pool.
- They proposed a batch-mode Deep Q-Network architecture that can find which batches of the regions should be labelled in each iteration. This makes the approach appropriate in large datasets since it does not take a lot of time to run.

- The method requires more labels from under- represented classes than baselines, which can address imbalanced classes concerns that exist in most semantic segmentation datasets.
- They evaluate the efficiency of the proposed approach on two challenging semantic segmentation problems – CamVid road scene segmentation and SUN RGB-D indoor scene segmentation.
- It does a better performance with the existing bounding box prediction methods as it proves better with the small or less object classes data set. This application establishes greater accuracy, mean Intersection over Union, and boundary F1 (BF) scores.
- As compared to active learning baselines, their approach necessitates the less amount of labeled data to provide equivalent performance.

**4. What do you think are the downsides of the work?**

I think the downsides of "*A Reinforced Active Learning Algorithm for Semantic Segmentation in Complex Imaging"* is - first of all, it is time-consuming, which is even more evident when working with big data because one has to train both segmentation and query networks. It highly depends on the quality of the initial labeled dataset which can be a constraint especially when there is scarcity of labeled data. Although the method has been designed to handle class imbalance to some extent, its effectiveness for the very rare classes may be compromised partly undermining its applicability in the different real-world conditions. Moreover, the performance gains are inconsistent and vary with different datasets and tasks, making it difficult to predict its effectiveness. The method also requires much human intervention in the process, which is rather time-consuming and costly, especially for large scale projects. Finally, there are various issues concerning generalization of learned active learning policy for different datasets.

## PAPER 2: New Semi-Supervised and Active Learning Combination Technique for Non-Intrusive Load Monitoring

**1. What problem does this paper try to solve i.e., its motivation?**

The paper *"New Semi-Supervised and Active Learning Combination Technique for Non-Intrusive Load Monitoring"* represents the issue of limited labeled data in the field of Non-Intrusive Load Monitoring applications. Non-Intrusive Load Monitoring aims to disaggregate a household's total energy consumption into the energy used by individual appliances without direct monitoring of each device. Traditional Non-Intrusive Load Monitoring techniques depend heavily on either supervised learning or unsupervised learning techniques, which require large amounts of labeled data, which is expensive and time consuming to acquire. On the other side, unlabeled data is under-utilized but is abundant. The motivation for this paper is to use both unlabeled and labeled data to improve the Non-Intrusive Load Monitoring performance and reduce the need for extensive manual labeling.

**2. How does it solve the problem?**

To solve the problem, semi-supervised learning is utilized to assist both labeled and unlabeled data. The active learning method is used to selectively request labels from users on the most informative samples. The method utilizes self-training as the semi-supervised method and entropy as the measure of certainty. The classifier selected the random forest due to its high accuracy compared to other algorithms that are experimented on. This combined approach is expected to reduce the number of times that the user is asked a query while at the same time every query that is asked is supposed to provide the maximum amount of information to the learner in the amount of effort that is required from the user.

**3. A list of novelties/contributions?**

There are several contributions to Non-Intrusive Load Monitoring:

- The paper proposed a combination of active learning and new semi-supervised methods for the implementation of Non-Intrusive Load Monitoring.
- Evaluates the effect of query timing on performance for large datasets of different complexities.
- Validates that semi-supervised learning outperforms supervised learning especially for small and complex datasets.
- Shows the combined technique and improves accuracy by up to 6% compared to semi-supervised learning alone.
- Advances the Non-Intrusive Load Monitoring by optimizing the use of limited labeled data, improving classification accuracy, and reducing the user query.

**4. What do you think are the downsides of the work?**

I think there are some potential downsides to consider in this paper- the authors in the research paper used only two phases of the BLUED dataset, thus not presenting a comprehensive view of possible real-world Non-Intrusive Load Monitoring scenarios. The approach still requires some level of user interaction, which might not be feasible in all applications. The performance improvement varies significantly between the two phases tested, suggesting that the method's effectiveness may be dataset dependent. In addition to that, this paper doesn't address the computational complexity of the proposed algorithm, which could be a concern for real-time applications. Lastly, the approach doesn't handle simultaneous events, which are common in real-world scenarios and could limit its practical applicability.

## PAPER 3: Captcha Recognition based on Multi-task Convolutional Neural Network and Active Learning

**1. What problem does this paper try to solve i.e., its motivation?**

The main motivation for "*Captcha Recognition based on Multi-task Convolutional Neural Network and Active Learning*" is to address the challenge of achieving high accuracy in CAPTCHA recognition with limited labeled training data. This becomes a very practical issue because collecting large volumes of labeled data for CAPTCHA's is not easy in real life situations. Current approaches involve the expensive and time-consuming process of manual labelling of CAPTCHA's, and while others have been shown to work well, they require large amounts of training data. The authors in the paper would like to propose an approach that can lead to a high recognition rate but with fewer labeled samples which will enable the recognition of CAPTCHA in practice.

**2. How does it solve the problem?**

To solve the problem, the authors introduced an end-to-end CAPTCHA recognition system using a multi-task learning convolutional neural network along with the active learning framework. Instead of segmenting each of the characters to find, Convolutional Neural Network takes in the whole CAPTCHA image as a feature. Several tasks are then used to identify each character individually but with the features that have been extracted in common. Active learning is used for the selection of the training samples which provide the most useful information for training purposes and to reduce the time, passive learning is used. Specifically, the method used uncertainty samples to identify high-uncertainty samples from the unlabeled pool and add them to the training set iteratively. This makes it possible for the model to generate high accuracy even with the limited number of labelled samples compared to the random selection. The method here involves the use of deep learning for the feature extraction and the classification of the data, which is a typical case since in practice, there are few labeled CAPTCHA data available.

**3. A list of novelties/contributions?**

- A framework recognizing CAPTCHA for solving without character segmentation by multiple work using the Convolutional Neural Network.
- Using the active learning strategy to increase the accuracy of a model with a smaller number of training samples.
- Some experiments just showed how advantageous active learning is particularly in sample selection compared to random selection.
- To demonstrate that, a high accuracy of 91.17% is achieved using a small training set of 1014 samples are sufficient.

**4. What do you think are the downsides of the work?**

I think there are some downsides of the work to be noted:

- The approach still depends on the need to label the first set of training data and label selection iterations of active learning.
- The success of the method might not be as effective when a completely different approach from the CAPTCHA is being used.
- The active learning process increases the time required to train the algorithm in comparison to the normal training process.
- The paper does not link the performance to other methods of CAPTCHA recognition.
- It is evident that from the analyzed dataset, there is no way of capturing different CAPTCHAs in the real environment.