Exploring Research Area

Amartya Dutta

1 Research Area

My area of interest lies in Machine Learning in Computer Science. Specifically, it is Computer Vision along with Deep Learning. Deep learning is a subset of machine learning, which essentially uses neural networks which are a series of layers used for processing information from a given set of data. This idea of Neural Networks was inspired by the behavior of the human brain to understand and process information. Whereas Computer vision is another field of Artificial Intelligence (AI) that works on enabling computers to derive information from images, videos, and other inputs. While traditional Machine Learning deals with numerical data, Computer Vision deals with image data. The inspiration behind the emergence of this field was to enable Machines to see and visualize just like humans do.

Computer Vision started out as an independent field. However, with the rapid emergence of Deep Learning and the availability of more and more data, Computer Vision also saw a huge boom. Computer Vision tasks improved hugely with the use of Deep Learning models. My research interest lies in the intersection of these fields, with an aim to enhance existing Computer Vision benchmark scores such that applications built upon these can improve our life.

2 Open Research Problems

Computer Vision and Deep Learning are inseparable today. Hence, simply stating Computer Vision suffices. It is an extremely hot field in terms of research and publications, with a large number of papers making significant contributions to the field being published each year. Accordingly, there are several research problems that make researchers in the field scratch their heads and continue to make them ponder over it.

The first is Image Classification. Image Classification is the most basic problem in Computer Vision. The research problem is as simple as; if given an image and a set of possible classes, what class does the image belong to? For example, if the set of classes is dogs and cats, a model should be able to predict whether a specific image is either a cat or a dog. Even though several State of the Art models exists for this problem, most of them fail under different circumstances. The classification results hold only when the distribution of classes is uniform or if the images of the classes are not too hard to learn from. I came across the problem of classification while I was working with medical image classification where the class distribution was highly skewed and the classification results didn't perform well.

Object Detection is yet another challenge that persists in the field. Similar to Image Classification, a lot of literature exists on this problem, however, they fail under different circumstances. The aim behind working on these problems should be to beat the existing benchmark scores and make sure these methods are robust and don't fail under other circumstances. While I was working on the same medical dataset and wanted to localize and detect the contusions from the images,

most of the models failed. This is because there wasn't a stark contrast between the regions of interest and the background. Under such circumstances, most existing works fail to perform.

Another class of problem that continues to exist is Image Segmentation. The task of Image Segmentation is highly dependent on the shape of the object of interest. Furthermore, in order for the models to successfully predict segmentation masks for images, a lot of data has to be manually annotated to generate the training data. Therefore, this is a costly process and to avoid this, Weakly Supervised Segmentation is being considered. This is my current research problem statement and the aim is to effectively generate segmentation masks without the headache of needing someone to annotate data.

3   Description of Research Problem

My research problem in technical terms would be stated as **Weakly Supervised Segmentation** of objects. This means that instead of providing the model with a segmentation mask as a label for each sample of data during training, it will learn to predict the masks over a relatively loose label such as the class prediction of the image.

In layman's terms, given an image and the only information available is the class the image belongs to, one should be able to correctly identify the shape of the object belonging to that particular class, in that image.

4   Research Type

Applied research is a research methodology that creates practical solutions for specific problems while basic research is an approach that seeks to expand knowledge in a field of study. This means that applied research is more solution-driven while basic research is driven by the discovery of additional knowledge in the field.

My area of research lies in the field of Computer Vision, which is purely applied research. The main aim of working on research problems in Computer Vision is to improve existing baseline scores so that they are good enough to be put into practical use.

Computer Vision deals with image data, which is an array of numbers in mathematical representation. The aim is to find which set of numbers in that array has a relationship with each other and how their mutual presence has a correlation with the predicted output of the model. Since it deals with discovering the relationship between input features and the output using numerical methods, it falls under Quantitative Research. Qualitative Research on the other hand is purely about making observations and drawing conclusions through methods such as surveys and interviews.

5   Professors Working on Relevant Area

Computer Vision is an exciting field of research and there are lots of researchers working in this field, be it Industry or Academia. The professors at Virginia Tech. are no exceptions. There are some highly esteemed Faculties that have made notable contributions to the field.

Dr. Anuj Karpatne is an Assistant Professor in the Department of Computer Science at Virginia Tech. Even though his primary research focuses on Knowledge Guided Machine Learning, there is a lot of ongoing and exciting research work going on in the field of Computer Vision in his lab. I am currently a student researcher in his lab, working on the topic of Weakly Supervised Segmentation. Dr. Karpatne currently has 3716 citations and has an h-index of 22 as per Google Scholar. In the

last two years, Dr. Karpatne has published over 43 research papers of which around 10 were from the field of Computer Vision

Dr. Chandan K. Reddy is a Professor in the Department of Computer Science at Virginia Tech. One of his research areas lies in the intersection of Deep Learning and Healthcare, some of which also lies in the field of Computer Vision. Currently, as per his Google Scholar profile, he has 8936 citations and an h-index of 38. In the last two years, he has published around 40 papers of which 6 were on healthcare images and finding insights about them through Deep Learning.

Another professor who has some very exciting work on Computer Vision is Dr. Ismini Lourentzou. Dr. Lourentzou is an Assistant Professor in the Department of Computer Science at Virginia Tech. She has a total of 378 citations as of now and an h-index of 11. Over the last two years, she has published around 24 papers. Out of these, around 6 papers were relevant to my area of research, Computer Vision.