NAME: Angad Partap Singh

ID: M00912257

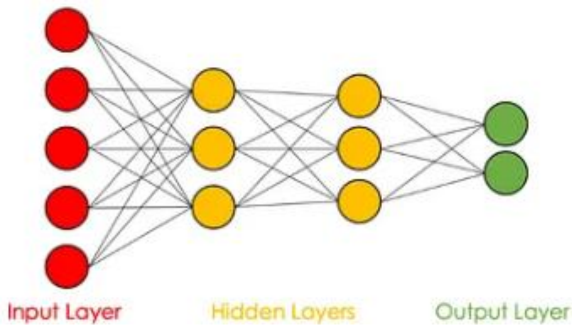
CST-4060-Visual Data Analysis- CW4: Essay on DEEP LEARNING

1. Abstract

A subfield of AI, as we address Deep Learning, has been able to attain a supreme position when it comes down to complex task handling and artistic results deliverable. Visual object detection, underwater imagery, speech recognition etc have all been impacted tremendously by DL. This essay explains the concept of Deep Learning and its applications. In the subsequent section, we'll study about how DL is applied in SONAR for detection of target underwater. Major applications discussed are medical imaging and Target detection. In the third section, underwater based targets are analysed and detected using different methods like CNN (Convolutional Neural Networks) and ATR (Automatic Target Recognition). Their operation and mathematics revolving the two methods have been discussed as well.

2. Deep Learning & Applications

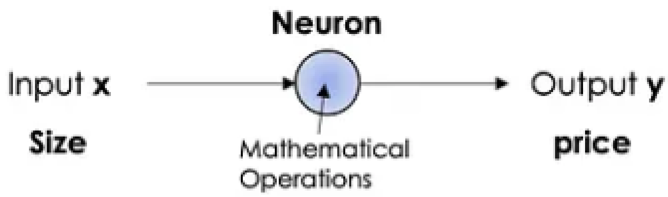
Deep Learning is a subset of Machine Learning, which is a further smaller set of AI. Artificial Intelligence focuses on the procedures of mimicking Human Behaviours. Depp Learning, part of ML can be thought of a Human Brian. The structure of Human Brain in terms of Depp learning is called Artificial Neural Network. It is an amazing technique to deal with unstructured data. As Machine Learning works upon the data provided to bifurcate between objects, Deep Learning picks the features basis Neural Network to get the job done with no Human Intervention.

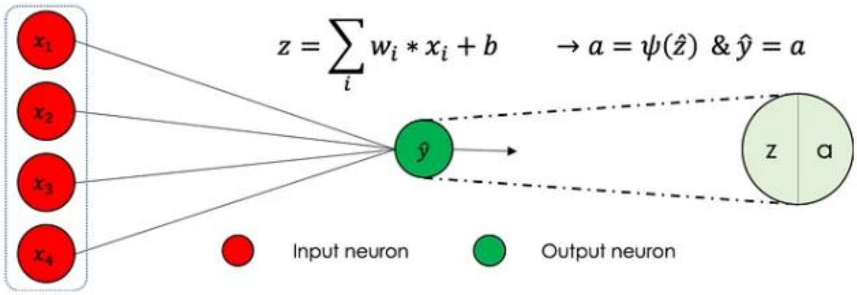


A basic Neural Network can be represented as:

- 5 neurons at Input layer
- 3 Neurons at Hidden layers.
- 2 at O/p layer.

The Neurons are the core entity of the Neural Networks. Data is fed to the Neurons at the Input Layer. This information is then transferred to Hidden Layers and the Output is given out at the Output Layer. The information is transferred through Connecting Channels between the Layers. These channels have values attached to them and are thus called Weighted Channels. Each neuron has Bias attached to it. The Bias is added to the weighted sum of inputs, which is further applied to a function. This created Activation Function. This Action Function results in whether a neuron gets activated or not. Every activated Neuron has some information for its subsequent layers. These weights and biases are adjusted to create a well-trained network. DL posses many applications in medical, image processing etc. One of them is detecting targets underwater, which is further explained in the upcoming sections. This target detection helps in the judgement of position along with feature of the raw images. Further, segmented images, and recognition is combined in one. Earlier, this work was done by artificial features, which included large datasets, exhausting segmenting procedures and then classification. But as the DL has come into picture, the time-consuming segmentation and recognition have tremendously reduced. Time complexity reduced, robustness increased along with greater accuracy are all the positive aspects that come handy with the introduction of deep learning into detection of targets and image processing.





Each Neuron has two main blocks:

1. Computing z using the inputs xi.

$$z = \sum_i w_i \star x_i + b$$

2. Computing z (y at o/p layer), using a.

$$a = \psi(z)$$

Medical Imaging: The solutions for medical image analysis and processing are all carried out using DL. In the recent few years, the pipelines and healthcare algorithms have been developed, that has accelerated the process of medical imaging. Being able to do MRI and CAT Scan, such that medical practitioners can take careful measurements. In the present scenario, this is carefully taken care by DL. This best of this is that this kind of work is uniform around the globe. Meaning medical industry now progresses to get better results. DL is reshaping the medical line by improving people's lives. The cancer cells and tumours are detected with the usage of DL. The new drugs, which take way lot of time for synthesizing, are now done at a faster rate. Apart from that, new and sophisticated equipment are also being invented on regular basis.

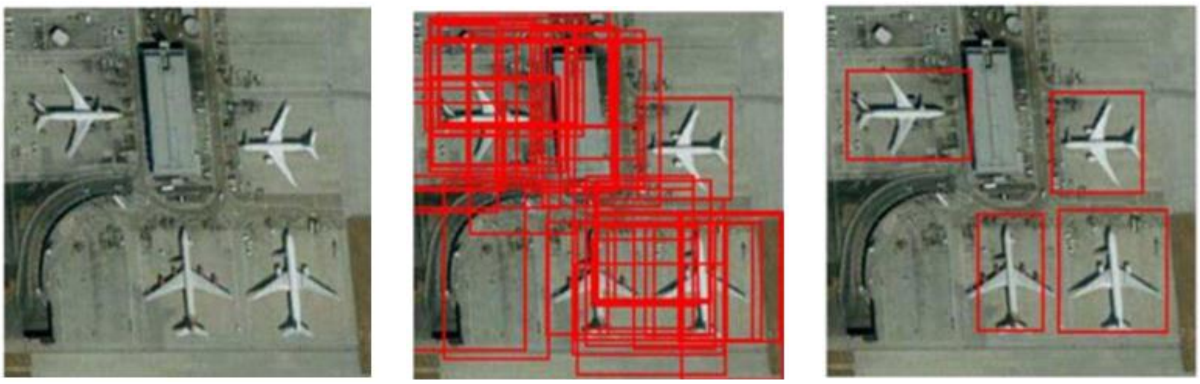
Reduction in unnecessary imaging, improvements in positioning, and characterization of the findings. In radiology, the Magnetic Resonance Imager has been made intelligent such that it recognizes and suggests modifications in a lesion to achieve its optimal characterization. Automated detection is under pipeline as it still faces huge research challenges and data. The long-term goal is to make the machine an independent image interpreter, however it requires huge data to derive knowledge and extract preferable outputs. For nuclear medical imaging, DL has also achieved the landmark to identify the number of doses specified for a patient, who needs to be treated with nuclear medicines.

There have been technical advancements to develop deeper neural architecture, in comparison to the traditional ANN (Artificial Neural Networks). These deep neural architectures have been trained in unsupervised manner with the application of Boltzmann machine. These processes were carried out to overcome the limitation of Deep Neural Networks (local minima establishment and overfitting). Thus, meaningful insights have been generated post analysing the large, uncategorical data, and meaningful predictions too. Further, the ensemble techniques, Naïve Bayesian models and SVM (Support Vector Machines), have always proven well regarding fields of genome segmentation and rRNA and mRNA sequencing.

Target Detection: With the usage of CNN (Convolutional Neural Networks) and Edge Boxes, DL has achieved the difficult task of classifying targets and non-target objects, in outer space (using satellites), and under water (SONAR). As the attributes in the contained in the images are prominent and concise, these are all included in the edge information of targets. This edge information is then used by EdgeBoxes to filter out the set target of proposals. On the other hand, CNN, a deep learning classifier, has a high learning capacity, and the capability of self-learning from the training data. Also, CNN is invariant when it comes down to minor rotations and slight shifts in the targets. This makes the performance and robustness of the proposed models quite optimum. Automatic detection of targets, such as oil tanks, deep-sea debris, artillery, etc. if of great significance for the armies of different nations and the navy forces. Using Adaboost Classifier, Zhang et al. [1] built an algorithm for oil-tanks detection. Yildiz et al. in [3] deployed SVM classifier for aircraft detection. Deep Belief Networks (DBN) and Convolutional Neural Network (CNN) are popular concepts widely used for Target recognition. CNN, as we already know, doesn't get affected by small rotations and shifts. DBN, a generative model, is pr-trained by Boltzmann Machine. Each layer is trained and then hyper tuned by Back-propagation, to act like a classifier. A block-level diagram for target detection is shown below.

This is a basic block-level diagram for an Automatic Target Detection. The basis being EdgeBoxes and CNN. The EdgeBoxes are used to generate object proposals in the first stage. Then Geometric filtering and checks are performed on the proposals, keeping high recall rates. CNN is then applied to the filtered proposals to extract the features and for classification purpose. The evaluation is then done on the larger dataset. This evaluation is done on a dataset with aircraft and patches of non-aircraft targets. The detection results are shown in the image below.

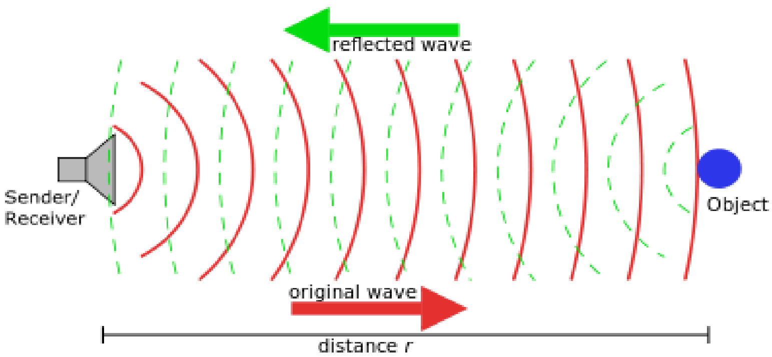
1. Raw Image **2. Objects Proposals w/ EdgeBoxes** **3. Target Detection**



Binarized Normed Gradients (BING) and selective search merges low-level pixels for object proposals. Ross Girshick et al. [17] proposed the usage of CNN along with Selective Search to achieve amazing results, EdgeBoxes depends upon the boundaries of the objects, thus making it a feature for target detection. The major issue that has been caused in target detection is the varying sizes of the objects. BING produces loosely fitting models, but Selective Search has been proven amazing for the general detection of targets. Although it is considerably slower than other models. This trade-off between quality and speed has been overshadowed by EdgeBoxes.

3. Deep Learning based Underwater Target Detection via SONAR

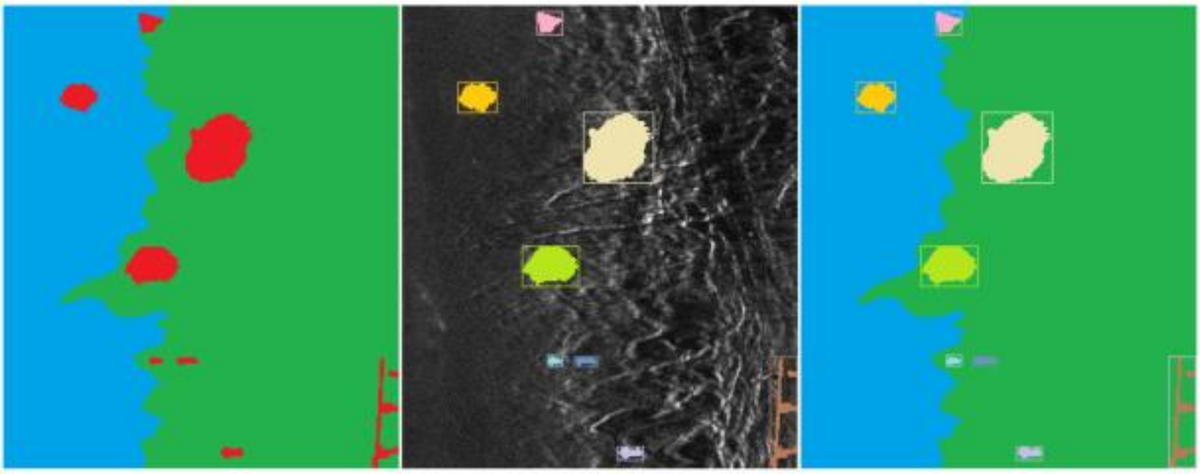
SONAR- viz. sound Navigation and Ranging. It's a technique (followed in submarine navigation) used basically to measure distances, communicate as well as detect objects underwater, or near seabed. Methods for communication in submarines, studies about underwater atmosphere etc are all part of SONAR image and target detection.



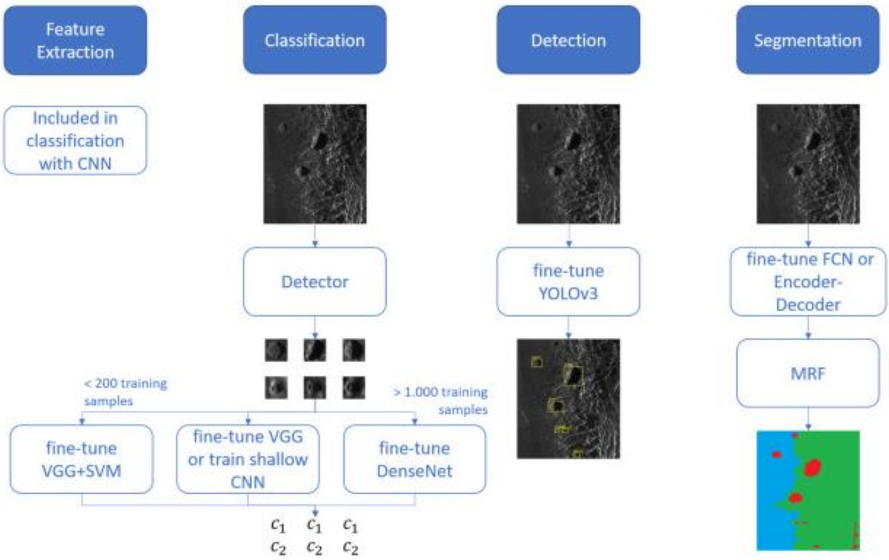
As the domains are evolving, this data-driven world is now taking help of deep learning and machine learning methods to gain information about underwater data using sounds.

The features are down sampled with Gaussian filtering, to make Gaussian image pyramid. The image I(x,y), and 2-D Gaussian function G(X,Y,d). The basic convolutional neural network has neurons and can be represented by the following. Multiple neurons are combined to create a model. The backpropagation is then applied to train the CNN. The more the epochs, the better the model. The weights are then adjusted to get a better output.

$$h_{w,b}(x) = f\left(W^T X\right) = f\left(\sum_{i=1}^{3} W_i x_i + b\right)$$



Target Detection is the most important feature in underwater object analysis. The Sonar images are mapped and classified using Automatic Target Recognition (ATR). The excelling in AlexNet, CNN have prevailed as state-of-the-art for tasks of visions in computer. The difference is compared between the raw images taken and the RGB pictures which are natural. The type of networks taken into consideration collates the data, and forms chunks of data such that the model trains from each chunk. The results from previous learning and new data produces amazing results.



The traditional ATR works as the features are calculated for classifier's training. These features depict the intensities of pixels and the shadowed/highlighted areas. This is turn helps in the improvement of the ATR processing chain. The methods like CNNs, AlexNet, Auto-encoders helps in the generation of the image from the classifiers output. These extracted features have been recently utilized with the ensemble methods of classification. Random Forest, Logistic regression, SVM etc have been proven to be beneficial in comparison of the classifiers working individually. Classical SVMs have been proven to be well to extract ROIs (Region of Interests), but conventional CNNs have worked well for the classification of the images.

Mathematically, there lies a decision boundary which separates the target from the unwanted noise. This can be written as follows:

$$\min_{w,b} \frac{1}{2}|w|^2 + C\sum_{i=1}^{n} \max\left(0, 1 - y_i(w^T x_i + b)\right)$$

W being the weights, x being the vector for features, b the bias. The solution can be found quadratically. Post the solution, the prediction points can be generated and tuned for future dataset/images. This is turn helps finetuning the detector. This detector, a subtask for processing does the detection of the boundary. A boundary box is calculated using the detector and the classification task is accomplished. The optimum outcome is evaluated based on high precision rate and recall. The overall working of the deep learning techniques can be visualised in the picture below.

4. Conclusion

The main work of this paper is of advantages of various algorithms w.r.t to feature extraction, preprocessing, classification, segmentation and then detection in SONAR. The limitations of underwater environment are realized by underwater acoustic information. The introduction of DL has surpassed the issues with traditional techniques. The comparison has been made b/w CNNs, SVMs, AlexNet etc. Usage of DL has made tremendous impact in underwater target detection.

5. References

Hohman, F. (2019). Visualization in Deep Learning. [online] Multiple Views: Visualization Research Explained. Available at: https://medium.com/multiple-views-visualization-research-explained/visualization-in-deep-learning-b29f0ec4f136.

Mebsout, I. (2020). Deep Learning's mathematics. [online] Medium. Available at: https://towardsdatascience.com/deep-learnings-mathematics-f52b3c4d2576.

Aggarwal, R., Sounderajah, V., Martin, G., Ting, D.S.W., Karthikesalingam, A., King, D., Ashrafian, H. and Darzi, A. (2021). Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. npj Digital Medicine, 4(1). doi:https://doi.org/10.1038/s41746-021-00438-z.

Steiniger, Y., Kraus, D. and Meisen, T. (2022). Survey on deep learning based computer vision for sonar imagery. Engineering Applications of Artificial Intelligence, 114, p.105157. doi:https://doi.org/10.1016/j.engappai.2022.105157.

Jin, A. and Zeng, X. (2023). A Novel Deep Learning Method for Underwater Target Recognition Based on Res-Dense Convolutional Neural Network with Attention Mechanism. Journal of Marine Science and Engineering, 11(1), p.69. doi:https://doi.org/10.3390/jmse11010069.

Yang, W., Fan, S., Xu, S., King, P., Kang, B. and Kim, E. (2019). Autonomous Underwater Vehicle Navigation Using Sonar Image Matching based on Convolutional Neural Network. IFAC-PapersOnLine, 52(21), pp.156–162. doi:https://doi.org/10.1016/j.ifacol.2019.12.300.

Sun, P. (2021). Real-life Applications of Deep Learning AI. [online] Blog For Data-Driven Business. Available at: https://www.intellspot.com/deep-learning-real-life-applications/.

Suganyadevi, S., Seethalakshmi, V. and Balasamy, K. (2021). A review on deep learning in medical image analysis. International Journal of Multimedia Information Retrieval, 11(1), pp.19–38. doi:https://doi.org/10.1007/s13735-021-00218-1.