TERMINAL ASSIGNMENT-BASED ASSESSMENT REPORT

Data Mining and Machine Learning – II

Abdul Azadh Abdul Saleem
Student ID: x18203621@student.ncirl.ie
M.Sc. in Data Analytics
School of Computing
National College of Ireland
Dublin, Ireland.

Supervisor: Noel Cosgrave

Q1) CASE STUDY

Object Classification from Photographic Image

INTRODUCTION:

Numerous researches on classification of different objects from the entire image which has multiple object classes present in it, has been conducted over years until now. This technique has been evolved over years and have been applied on numerous different applications depending on the purpose and the scenarios the methodology the implement of the technique has evolved. Parallelly several researches have been conducted to develop an efficient and cost-effective technique or network, that has the ability classify the objects from the collective environment, with higher accuracy. And these researches have also extended in the field of developing the better dataset of millions of classes of objects present in it, which could help the network to be trained to handle different situation. The traditional techniques of object detection proposed before has the ability to detect the objects which are in the first layer of the image (i.e. the objects present in the image without any overlap), whereas the model proposed here is focused on enhancing the ability of the model to detect and classify the objects which are present in the background of the subject of the image as well.

ImageNet is an image database that has about fourteen million images and about twenty-two thousand categories which can help the model classification model to perform with better than utilizing the other dataset extracted from various sites. ImageNet used to conduct the challenge called Large Scale Visual Recognition Challenge (ILSVRC), for comparing the algorithms that could perform better with object detection and classification. In the challenge conducted in 2015, it was reported that the Residual Network (ResNet) has outperformed the other networks in extracting the features of the objects and classifying the objects in the image (He *et al.*, 2016). The technique of object classification proposed in this report will be utilizing the deep network of ResNet 50 which provides the 49 layers of convolution and activation the output from this network is pooled in order to obtain the rounded binary values of the classes of classification.

PROPOSED APPROACH OF MULTI-SCALE OBJECT CLASSIFICATION:

From the research conducted on the previous methods employed to detect and classify the object in the images, the approach presented by (Yuan *et al.*, 2020). The researcher has approached a novel Convolutional Neural Network (CNN) model which is called as Gate-CNN network, this model resulted in performing the object classification in multi-layers of the image better than the previous models developed before. Since the aim of this proposal is to provide the robust approach of object classification model, this model can perform better in classifying almost most of the objects in the image provided to the model. Knowledge Database Discovery (KDD) methodology has been adapted to propose the implementation of this model for object classification as shown in the Fig.1 shown below.

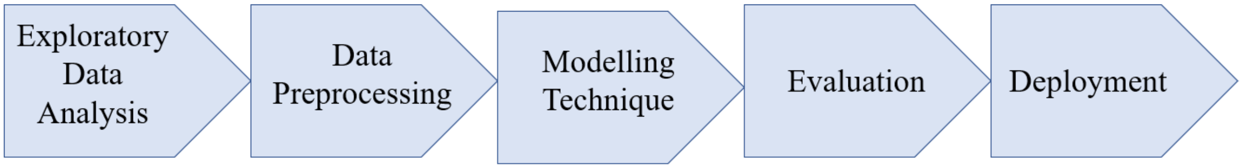

Fig.1 Proposed Methodology for Implementation of Object Classification

A) EXPLORATORY DATA ANALYSIS:

Exploring the content of the data provides some knowledge about the nature of the dataset and its internal division and the variables types, which can help the researcher to identify the discrepancies, that can be avoided to train the model. The dataset chosen here to train the feature selection layers of this model is the ImageNet dataset which has almost all the classes that could possibly present in the photographic image along with its subclasses. The dataset has the thumbnails of the images which are quality controlled and human annotated to define the class of the objects in the images.

B) DATA PREPROCESSING:

When the images from the dataset is imported, the images will be converted in numerical array of different dimensions. The input images that is preferred as input for this model is should be in RGB format. The images in the dataset may have different pixel dimensions, which has to standardized by converting them into the pixel ratio of 224*224 which makes the data fit to train the ResNet50 networks and further normalized by dividing the image array by 255.

C) OBJECT CLASSIFICATION MODEL:

The proposed design for Gated Convolutional Neural Network has been shown in the Fig.2 below. Utilization of Gated-CNN provides the advantage to the model by filtering the noises from the output of the network. The process of classification of the object takes place as follows.

1. FEATURE SELECTION:

When the preprocessed images will be provided as the input to the model, as the first step the selection of the features from the images will be done with the help of 50 layers network of ResNet50 network, which has been trained with the weight of ImageNet dataset. ResNet50 is the deep neural network that has 50 layers, with first layer of the network as the input layer which also the zero-padding which enables this layers to add the zero to the image matrix, if it doesn't has the required size. Further the other layers of the network contain 49 layers with each layer having the function of convolution and activation in it, which helps the model to perform better in training and classifying the objects. The output from these layers will be average pooled and flattened into different classes of object detected by these layers in this image. The layers of multi-scale feature that is obtained from this step will be provided to the successive step of feature extraction for further processing. Every layer in this process provides the different semantic information about the content in the image.

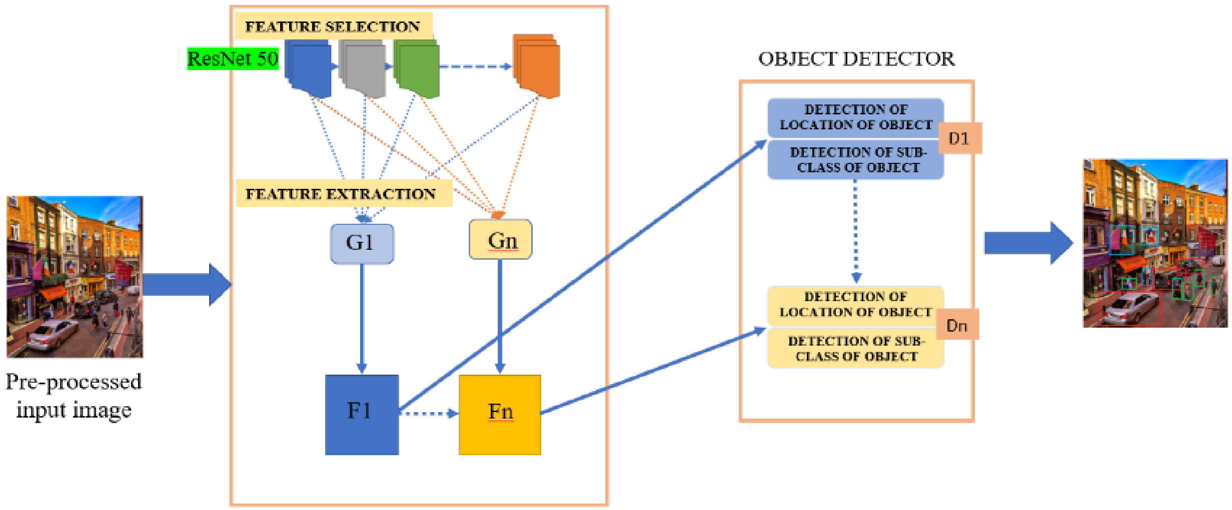

Fig.2 Proposed Convolution Neural Network for Object Classification.

2. FEATURE EXTRACTION:

This Feature Extraction layer will perform the operation of Gate process which is mentioned as G and F in Fig.2. The process of combining and processing multi-scale information obtained from the previous layer is will be explained with the help of Fig.3 below.

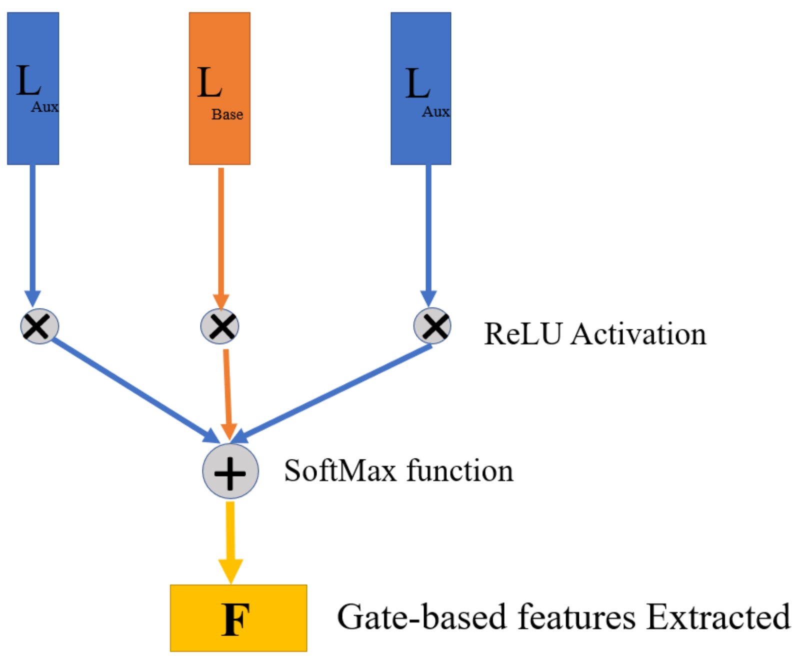

Fig.3 Process flow of Gate-based feature selection.

Among the layers of multi-scale features obtained, one layer will be considered as the Base layer which has the primary information of the image and the other layers are considered Auxiliary layers which has the information that can support the primary information. The information from the layers will be compared with the semantic information in the base with the help of Sigmoid function which helps to predict the probability of the classes. After this process, the information from each layer is passed to the ReLU activation layer which reduces the linearity of different classes of the objects in the processed data. The information of each layer from the ReLU activation layer will be combined for predicting the class of the object features in the image data with the help of SoftMax function.

3. OBJECT DETECTOR:

The features extracted from the previous layers will be fed to the detector layer which performs two operations. One operation is to perform the localization of the object by drawing a boundary box around the individual object in the image. And the other operation is to categorize the object under particular class depending on the probability of the class obtained. The main advantage of deploying this model is that it provides the information including the subclasses of the object detected from the image.

HYPERPARAMETER OPTMIZATION:

The classification model proposed here will be optimized in order to enhance the performance of the model and to reduce the computation time of running the classification model. (Zhang, 2019) has proposed that Adam optimizer which is usually used to optimize the classification model to update the weights of the network based on the training data for every iteration, does not perform well in case of Deep Neural Network. The author also reported in some cases this optimizer underperforms the pervious Stochastic Gradient Descent optimizer. The paper proposes the utilization of ND-Adam optimizer, which is improved version of Adam optimizer, can perform better in deep neural network. This optimizer adapts the learning rate depending the individual input vector, replacing the method of adapting the learning rate depending on the individual weights. And the loss function here will be considered will be ‘categorical_crossentropy’ loss which computes the loss between predicted object and classes with the ground-truth information of the object class.

D) EVALUATION:

Evaluation of the classification model is an important part of the implementation where the performance of the model can be evaluated, before deploying the model. The performance of the model can be calculated by using the evaluation metrics like Confusion Matrix, Accuracy, Recall, Precision, AUC-ROC curve, F1-score and Kappa-coefficient value. The formula for calculating such performance metrics is provided below.

1. **CONFUSION MATRIX -** Confusion matrix is plotted between the actual class of the objects in the ground-truth image data, with the class of the object predicted by the classification model proposed. The confusion matrix plot proposed below defines the performance of classification between two class of the object.

	ACTUAL CLASS	
	TRUE	**FALSE**
POSITIVE	**TP** (TRUE POSITIVE)	**FP** (FLASE POSITIVE)
NEGATIVE	**TN** (TRUE NEGATIVE)	**FN** (FLASE NEGATIVE)

PREDICTED CLASS

Fig.4 Confusion Matrix of Classification.

Since we have number of classes of object differing in each image let us consider 'n' be the number of class in the image and 'n x n' will be the dimension of the confusion matrix plotted for such classification. TP represents the actual class in the image is the same as the class predicted by the model, FP represents the class of object in the ground truth image is false and the prediction shows differently, TN represents the prediction of class is different from the class in the actual data, and FN shows the values which are actually negative and there are correctly classified by the model. With the help of information obtained from the confusion matrix the below metrics can be performed.

2. ACCURACY – Accuracy of the classification model can be calculated with the help of formula below.

$$\boldsymbol{Accuracy} = \frac{(\boldsymbol{TP} + \boldsymbol{TN})}{(\boldsymbol{TP} + \boldsymbol{TN} + \boldsymbol{FP} + \boldsymbol{FN})}$$

Accuracy is one of the main metrics that is helps to analyze the performance of the model in performing the classification in correctly predicting the classes of the object.

3. RECALL – Recall defines the sensitivity of the model in performing the classification which can be calculated by the formula below. This recall values extends between 0 to 1

$$\boldsymbol{Recall} = \frac{\boldsymbol{TP}}{(\boldsymbol{TP} + \boldsymbol{FP})}$$

4. **PRECISION** – Precision defines the specificity of the model, which can be calculated by the formula below. This precision values extends between 0 and 1, where the model with both precision and recall value as 1 will have the complete performance in classifying the class of the objects correctly.

$$\boldsymbol{Precision} = \frac{\boldsymbol{TP}}{(\boldsymbol{TP} + \boldsymbol{FN})}$$

5. F1-SCORE – F1-Score of the classification model shows the test accuracy of the model, which is nearly same as the accuracy value of the model. This metrics can be classified with the formula below. This F1-Score value extends between 0 and 1, where the value near to 1 represents the model has full test accuracy in classifying the objects.

$$\boldsymbol{F1} - \boldsymbol{Score} = \boldsymbol{2} * \frac{(\boldsymbol{Precision} * \boldsymbol{Recall})}{(\boldsymbol{Precision} + \boldsymbol{Recall})}$$

6. **COHEN'S KAPPA COEFFICIENT** – This metrics is used to measure the robustness or reliability of the model in performing classification. This also provides the details of relationship between the categorical classes in the model. Pe – represents the probability of getting the desired output and the Po – represents the observed outcome. K value should be nearly 1 which represent the model is robust in performing the classification.

$$k = \frac{p_0 - p_e}{1 - p_e}$$

E) DEPLOYMENT:

After evaluating the performance of the classification model, the model will be deployed for classification of the objects from the photographic images in real-time. If the performance of the model does not satisfy the desired performance the model will be again trained with the different parameter and the then evaluated again. This process is repeated until considerable performance in observed from the model.

SCALABILITY ISSUES:

The purpose of the model is to detect and classify the objects in the photographic image, which makes the model to get trained with a large set of image data. But in this proposed approach the model is trained with the predefined weights of the image from the ImageNet data which can reduce the complexity of computational run-time of the model. Further utilization of packages from TensorFlow module helps to perform the Neural Network model with better accuracy and reduced computational time. Further Google Colaboratory (Colab), which is similar to Jupyter Notebook in google drive helps to perform such task better, as this can also provide GPU till 12GB for such huge task, with free of cost.

ETHICAL IMPLICATIONS:

It is important for the researcher to conduct the research ethically without indulging or exposing any sensitive information utilized for the development of the project. Here this case the model will be trained with the weight of the images in the huge dataset of ImageNet which was openly available in their official site for research and education purpose. As a part of testing data no images will be considered which has sensitive information like exposing an individual or identity of the person.

REFERENCE:

He, K. *et al.* (2016) 'Deep residual learning for image recognition', in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 770–778. doi: 10.1109/CVPR.2016.90.

Yuan, J. *et al.* (2020) 'Gated CNN: Integrating multi-scale feature layers for object detection', *Pattern Recognition*. Elsevier Ltd, 105, p. 107131. doi: 10.1016/j.patcog.2019.107131.

Zhang, Z. (2019) 'Improved Adam Optimizer for Deep Neural Networks', in *2018 IEEE/ACM 26th International Symposium on Quality of Service, IWQoS 2018*. Institute of Electrical and Electronics Engineers Inc. doi: 10.1109/IWQoS.2018.8624183.

Q2) PAPER REVIEW

AUTOENCODER NODE SALIENCY – SELECTING RELEVANT LATENT REPRESENTATIONS

STRUCTURE TITLE:

Name of the article was framed precisely, and the project's intent can also be easily interpreted by the title name.

ABSTRACT:

The abstract of the paper clearly explains the key benefits of the paper and the background of the problem which the research paper is going to address has been briefly summarized. And, the author contribution for this research and the future scope of this research project along with major findings of this study is explained in this section with clear explanation. However, the outcome of the experiments results and major findings in the conducted research is missing in this section.

INTRODUCTION:

In this section author has explained how the research project is going to solve the research problem in a broad way and historical narrative of the previous research field is precisely stated. Introduction is written in a logically order starting with the background of the research and the importance of conducting the research and how the research question is solved and framing the research question and objective are mentioned clearly in this section. Reference for the saliency node method is given less importance and the autoencoders is referenced heavily in this section. Overall Introduction section contains all the information related to the experiment design and method, also points out the assumption which the experiments is going to be conducted. More material related to the autoencoders and the underlying theory is explained in detail which can be included in the methodology section and can be minimized in this section. Length of introduction section has more than 2 paragraphs which is bit lengthy and can be reduced.

GRAPHICAL ABSTRACT:

Author did not used any graphical abstract other than the screenshot of the validation and evaluation result of the experiments. Adding the pictorial design of the experiments will gives a better understanding for the readers in a single glance. The figures which shows the results of the experiments and the comparison tables are presented in a clear way different type of charts including bar, line and tables are used to explain the result which is obtained from both the dataset. Adding the visual representation of the process will give a better understanding for this research.

METHODOLOGY:

Before explaining the methodology, author has given the brief overview on Autoencoders and the experiments deign and the materials used to build the experiments is clearly explained also the previous work where the autoencoders used to build the learning task from the hidden nodes are reviewed. After the review of the autoencoders proposed node saliency method which is a novel method build in this research is explained step by step of the whole process and also the mathematical formula to compute the hidden nodes are also explained in a clear way. Methodology part is ordered in a precise way that the future researchers can follow the steps and implement the same research project without any issues and the author has followed the logical sequence of ordering the methodology heading for better understanding starting from dataset collection till the evaluation section. Two different type of dataset is used in this research and for the both dataset author has provided the adequate explanation of the dataset and provided the link from where the dataset is gathered in this section and also the dataset is chosen in an appropriate way and related to conduct this experiment. Autoencoders node saliency methods are a novel method is used in this experiment and the author has given a in depth explanation about the methods and included the references of the previous research conducted in this area and cited in a clear and understandable way. Visual representation can be added in this section as the novel methods is explained in theory and the researches may not find the sufficient information in a single glance. The main concern in this methodology section is while describing the measurement which was used in the research author has missed to give the measurement for both the dataset in a detailed way only the parameter setting and result obtained by the autoencoders are explained in detail and the measurement which the output of the autoencoders are not explained . Overall, by adding the visual representation of the novel node saliency methods and adding the measurements of the experiments conducted in both the dataset will be helpful for the future researches to replicate the same work with less bias.

RESULT:

Author has discussed all the key findings of the experiments he conducted on two different datasets in a precise way. New findings of the research have been explained with the charts and the experimental values are highlighted in the tables. For each dataset author had explained the step wise result progression and the explanation of the results from the figure mentioned in the report. In this research totally 4 steps are taken for each dataset to produce the result and for each section the explanation of the result and the discussion of the findings is stated. All the process is ordered in a logic way and sufficient information related to the experiments are clearly stated. Appropriate analysis is carried out which is related to the experiments and the research objective in the experiments conducted in each step the result of the research objective is carried out with good explanation. All the result drawn from this experiment are carried out and explained with correct statistics, from the figure the conclusion is drawn and used to describe the experiment result in this section.

Overall, result section contains all the sufficient information which are relevant to this research with sufficient information for the future researchers to understand this research. Also, both statistics and model comparison are described by the author in step by step and arranged in a logical way with a clear understanding of figure for each result section. Adding the references for evaluation metric based on which the experiments result is conducted can further be enhanced in this result section.

CONCLUSION / DISCUSSION:

All claims the author has concluded in this section are supported by the result. In this section the Novel autoencoder saliency method used in two different dataset results are explained and also author has explained about the finding he didn't get when conducting the research and how he deal when the uncertainties is encountered while conducting the research is also explained in the discussion section. Also, how the previous research helped to build this model and also in conducting the experiments in different ways is mention in discussion section with the result agree with the previous research. Finally, in the discussion section contribution of the conducted experiments which is related to the knowledge in the field is discussed for the future researchers to carry out this experiment. In the conclusion section author has clearly restated the purpose of conducting the research and the research objective of the project and the work done in this research and also the key findings of both the experiments conducted have been discussed and the comparison of both the method and the model which performed better is stated in this section. overall, author was so honest in criticizing the experiments which was conducted with sufficient information on the improvement and modification that is done which conducting the experiment is stated this section. Future work proposal part is not included in the result section which can be added so that the future research workers can improve the model or enhance the conducted experiments in this project.

LANGUAGE:

Author has used the simple language and reading the report and understanding the content is easy and not complex. However, there are few minor grammatical errors in the conclusion/discussion section which must be revisited by the author to do the minor correction. In this research to explain the experiments result author has used bar and line chart and in the line chart the x axis labels are not clear and are slightly overlapped reducing the size of the line chart will give the better understanding of the figure. Other than this all the axis labels and the size, color of the charts is consistently maintained.

PREVIOUS RESEARCH:

The references used in this section are properly cited by the author and top-rated references paper are chosen and referenced in this section. All the relevant reference paper which is related to the research objective is cited precisely. All the previous research paper which are used in writing the related work is referenced appropriately with accurate references. IEEE reference styling was used by the author to reference all the relevant papers used in this research with proper referencing style.