

Deep learning based qualitative analysis for better performance in medical diagnosis

Muthupriya Shankaran (mush8206)

Department of Computer
and Systems Sciences

*Empirical Research Methodology for Computer and Systems
Sciences*

Stockholm University

Autumn term 2021

Supervisor: **Name**

Swedish title: XXX – not mandatory



Abstract

The paper is about the deep learning based qualitative analysis to show how effectively deep learning model performed to improve medical diagnosis. It uses inductive approach. Among several of medical diagnosis, this study adopted a convolutional neural network (CNN) for covid-19 testing and respiratory study and examined the performance with computed tomography (CT) pretrained model testing. The proposed transfer learning approach proves to be more successful than the approaches described in literature. In qualitative analysis, it identifies themes, patterns and relationships from data collection. Thus this approach is a satisfactory performance by our model.

Keywords: Deep learning (DL), convolutional neural network (CNN), Computed tomography (CT), medical diagnosis, qualitative analysis.

Table of Contents

1 Qualitative Research Plan.....	0
1.1 Research question.....	1
1.2 Sampling.....	1
1.3 Place of study.....	1
1.4 Time plan.....	1
1.5 Plan for Data collection.....	2
1.5.1 Kind of data needed.....	2
1.5.2 Data guide.....	3
References.....	4
Appendix A -An Appendix.....	5

List of Figures

List of Tables

Table 1: Gantt chart for time plan (Duration is calculated on weekly basis).....	2
--	---

1 Qualitative research plan

Qualitative research plan which gives the clear and detailed foresight before the research begins. The main considerations of planning this qualitative search will be outlined particularly with the question and background of the research area, in qualitative research, code system was developed gradually and collaboratively. Later, it categorize, identify the themes patterns and relationships hence these would perform better in diagnostic process using deep learning technique and in CT imaging. Moreover, Data collection will act as a same source for this report. Once the data is collected, thematic analysis followed by qualitative approach has been conducted. Then survey, which would help to identify patterns and themes to achieve the expected performance. Before survey, study plan is conducted in Gantt chart to know the specific timeline duration to deliver the report.

1.1 Research question

The research question for this paper is "How effectively deep learning model can support medical diagnosis?". The RQ was extracted from the literature review. The research explored how diagnostic process achieved better using deep learning technique. Qualitative research of deep learning has been carried out in order to obtain culturally specific information about the values and better outcomes of medical diagnosis.

1.2 Sampling

The sampling frame was formed by those ($n = 96$) who gave their consent to be contacted during their baseline interview for the randomized control trial (RCT). The study sample comprised respondents from severity measurement and control groups purposively selected to include those eligible for the following resources: mild, common, high severity and none. Sampling continued until no new themes emerged from the interviews until data 'saturation' was reached. Sometimes, due to missing of information from data or lack of information for instance while interviewing if one of the participants have not stated his/her problem properly during interview then sampling issues may occur and create distortions in the results and lead a user to draw incorrect conclusions. Hence to overcome this issue simple random sampling technique are used or they can be prevented if the analysts select subsets or samples of data to represent the whole population effectively.

1.3 Place of study

In geographical area, research on COVID-19 is being effectively carried out in various areas such as in Brazil belonged to the World Health Survey, Latin American countries and the Caribbean are few more region where this study has been conducted. And in locality, the place of study would be in medical laboratory or clinical laboratory where tests are carried out on clinical specimens to obtain information about the health of a patient to aid in diagnosis, treatment, and prevention of disease.

1.4 Time plan

Time plan plays a major role when it comes to research plan. For this Qualitative Research method from below Gantt chart we may come to know the time duration that how long period would take to deliver the report. Planning of time scheduling is important before any research begins to start the survey. Following (Table 1) are the steps can be planned on weekly basis

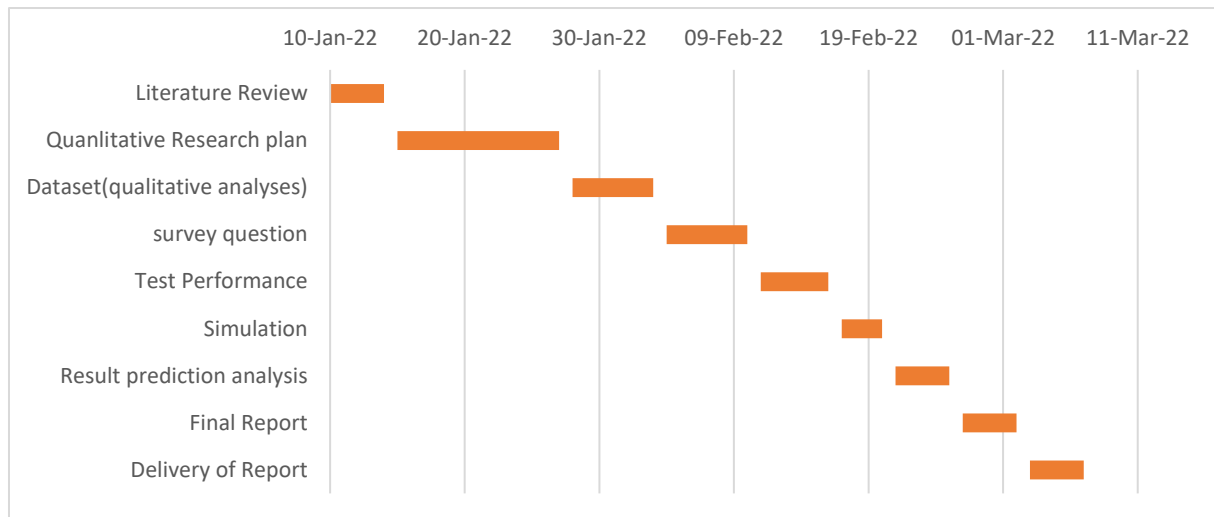


Table 1: Gantt chart for time plan (Duration is calculated on weekly basis).

1.5 Plan for data collection

In plan for data collection, to seek better answer for our research question, Interviews method are used which is most effective for qualitative research. Interviews questions are usually open ended questions and provide an indepth information about dataset. The most common type of interview method used in qualitative research and the healthcare context is semistructured interview. Semistructured interviews are mainly used to explore participants thoughts, feelings and to delve deeply in to personal and sometimes sensitive issues. Semistructured interviews can be conducted in multiple ways such as face to face, telephone, text/email, individual, group, brief, in-depth. To answer our research question we will focus on semistructured interviews within qualitative research-focus groups and in-depth interviews.

1.5.1 Kind of data needed

Interview based semistructured method are followed to collect the data from healthcare. With the help of medical care premises related to research question data are collected.

The source of input for our research question for the images in COVID*CT-2A are as follows:

- China National Center for Bioinformation (CNCB) (China)[1]
- National Institutes of Health Intramural Targeted Anti-COVID-19 (ITAC) Program (countries unknown) [2]
- Negin Radiology Medical Center (Iran) [3]
- Open access online collaborative radiology resource (Radiopaedia) (countries unknown) [4]

Thus the interviews will be used as a base to design survey for qualitative method.

1.5.2 Data guide

The data will be analysed by using thematic analysis. The thematic analysis provides theoretical freedom and flexibility which makes it a foundational method for qualitative analysis. Below are the following steps to answer research question through qualitative analysis method.

Survey Question: How do you avoid bias in medical diagnosis?

Data Extraction & Methods:

The purpose of establishing this validation set is to investigate, for instance, whether adding weak validation (i.e., findings without using RT-PCR and radiological tests) training data would boost the performance of the model. In another way, coronavirus pneumonia (NCP) infected by SARS-CoV-2, common pneumonia (CP), and normal controls are applied some modifications to images from the database to facilitate our models. Specifically, as the potential contrast in the background of the images may result in biases in the models.

Identifying themes, patterns and relationship:

Although performance indicators are useful for model evaluation, they fail to explain the decision-making behavior of the network. In this regard, we employed visualization technique of DL based gradient descent is applied to explore the areas of concern for the models in COVID-19 testing, to better understand which characteristics of CT images are key for diagnostic accuracy, and thus aid clinical decision-making. In most cases, the performance of the model is the same as would be expected for typical human visual cognition. This is particularly true for common pneumonia, as the model successfully focus on the disease areas, and display the affected regions of lungs. The radiologist further can apply color visualization approach using gradient descent for making efficient and confident decision. In addition, COVID-19 diagnosis include examination of clinical characteristics and the use of computed tomography (CT) imaging. Owing to its high sensitivity, CT imaging has been proposed as an essential substitute tool for COVID-19 screening, which is especially effective as a complementary method to RT-PCR, and by way of CT imaging was able to perform rapid prediction compared to RT-PCR.

Discussion:

The performance of the model with limited data and found that the model still perform satisfactorily. This shows that our model is still applicable with a limited data, which is characteristic of the real situation, where large and diverse datasets may not be readily available. Finally, we explored the relevant mechanism of COVID-19 testing using gradient descent visualization technique to make the proposed deep learning model more interpretable and explainable. The experiments demonstrate that our models are effective in COVID-19 testing. In future, we will pay attention to the evaluation of the severity of COVID-19 and attempt to discover more valuable information from CT images to combat the pandemic. Although the system has good performance on public datasets, the work is still at theoretical research stage, and the models has not been validated in actual clinical routine. Therefore, we will test our system in the clinical routine and communicate with physicians to understand how they use it and their opinions about the models. Thus, we can further improve the models in our future work.

References

- [1].Zhang, K. *et al.* Clinically applicable AI system for accurate diagnosis, quantitative measurements, and prognosis of COVID-19 pneumonia using computed tomography. *Cell* 181, 1423–1433 (2020).
- [2]. Harmon, S. A. *et al.* Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. *Nat. Commun.* 11, 1–7 (2020).
- [3]. Rahimzadeh, M., Attar, A. & Sakhaei, S. M. A Fully Automated Deep Learning-based Network For Detecting COVID-19 from a New And Large Lung CT Scan Dataset. *medRxiv* 2020.06.08.20121541. <https://doi.org/10.1101/2020.06.08.20121541> (2020).
- [4] COVID-19 | Radiology Reference Article Radiopaedia.org. <https://radiopaedia.org/articles/covid-19-4>.
- [5] Panwar, H. *et al.* A deep learning and grad-CAM based color visualization approach for fast detection of COVID-19 cases using chest X-ray and CT-Scan images. *Chaos Solitons Fractals* **140**, 110190. <https://doi.org/10.1016/j.chaos.2020.110190> (2020).
- [6] Wu, F. *et al.* A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265–269 (2020).

Appendix A – An Appendix

Assessment criteria