Masters Research Proposal - What are the effects of implementing a two-step system that includes a dimensionality reduction process for the detection of bone fractures

Alice Yang (597609)

School of Electrical and Information Engineering, University of the Witwatersrand, Private Bag 3, 2050, Johannesburg, South Africa

Abstract: The objective of this paper is to propose a two-step system for the detection of bone fractures in order to investigate the performance of the system compared to a single step system. The performance are defined by the speed of its execution, classification accuracy and error rate. The two-step system consists of a dimensionality reduction and a automated decision making process. The objective of the introduction of the dimensionality reduction process is to improve the performance of the automated decision making process during the training and decision making sessions.

Key words: Dimensionality Reduction, Neural Networks, Supervised Learning, Unsupervied Learning, Bone Fractures, Medical Images

1. Introduction

In medical the field there is an emphasis on the importance of medical diagnosis. In critical cases, an inaccurate diagnosis of a symptom can cost a patient's life. Many diagnosis are performed by trained medical doctors, however they are prone to making mistakes which can be influenced by external factors, such as fatigue and lack of training in the particular field. Automation of medical diagnosis was introduced in the early 1970s [1]. There are many algorithms available in diagnosing patients. The basic algorithm consists of "if... then..." statements. Although the algorithm is simple, the list of conditions in the medical field are endless and as a result the execution time to diagnosis a patient's symptoms is not realistic. There are many pattern recognition algorithms available which are categorized in two categories, namely, unsupervised and supervised. The reasoning behind making use of artificial intelligence is to reduce costs. In this proposal report, section 2 presents the background in the related topics, section 3 describes the problem, section 4 illustrates the proposal methodology, section 5 shows the time management for the proposed research and section 6 concludes the paper.

2. Background

There are many simple implementations of programs that are proposed to diagnosis common medical conditions. These implementations requires a vast majority of the medical knowledge to express the different descriptions of possible diseases to be stored in the program's database. Although the program is simple to implement, it has no method of measuring the severity of the patient's illness. Another downfall to using a database to diagnosis an illness, is that it can mean that the database can grow exponentially otherwise it would be outdated.

An alternative approach in medical diagnosis is to focus on one disease to ensure accuracy and quick response. The use of artificial intelligences can be focussed in one area. The focus of this paper is to detect bone fractures from X-Ray images. There are a selection of tools that can be applied in order to achieve the desired outcome for this investigation. These tools can be in the form of supervised learning, reinforcement, unsupervised learning or a combination of the two. Support Vector Machine and Neural Networks are common tools used in artificial intelligence (AI).

2.1 Literature Review

The review done by [2] for the automation of bone fracture detection indicated that there are various techniques in pattern recognition that comes in the form of unsupervised, reinforcement and supervised. The challenge found in classification are that there is an information overload and size and dimension are a concern. The problem with information overload is that it becomes to computationally expensive to find crucial information needed for the classification. The size and di-

mension of the information is a concern because large sizes of data contributes to the number of dimensionality, with high number of dimensions it presents a challenge in the classification process and in turn may require more computational resources.

2.1.1 *Unsupervised Learning*

Unsupervised learning is defined by the not needing to train the algorithm in order for it to classify data into different categories. This process allows the algorithm to learn more about the given data. There are two categories in unsupervised learning which are clustering and association. Popular unsupervised algorithms are k-means and Apriori algorithm. The clustering technique is an example of un-supervised learning, since the technique analyses a given set of data and groups the data that have similar characteristics together. One of the main advantages of clustering is that it is adaptable when given new data sets compared to classification. It is also useful for singling out useful features. There are various types of clustering methods, namely, partitioning, hierarchical, density-based, grid-based, model-based and constraint-based. Other unsupervised learning techniques includes Principal Component Analysis, Independent Component Analysis, Non-Negative Matrix Factorization and Singular Value Decomposition.

2.1.2 *Dimensionality Reduction*

The collection of digital data has increased drastically over the past decade, which led to datasets having high dimensionality. The high dimensionality found in datasets affected the performance of data processing algorithms. This is known as the curse of dimensionality [3]. Dimensionality reduction is a process targeted at reducing the dimensionality of the considered dataset by reducing the number of random variables. This process can be divided into two stages, namely Feature Selection and Feature Extraction. Both these stages are crucial for automating bone fracture detection. Dimensionality reduction techniques includes, Principal Component Analysis (PCA), Independent Component Analysis (ICA), Kernel Principal Component Analysis (KPCA), Graph-based kernel PCA, Linear Discriminant Analysis (LDA) and General Discriminant Analysis (GDA). The use of these techniques are used in the feature extraction process.

Principal Component Analysis (PCA) is a form of unsupervised learning and has the ability to compress data [4]. It is a widely used statical technique that reduces the dimensionality in large datasets, whilst preserving as much crucial information as possible [5].

Similar to PCA, Independent Component Analysis (ICA) is a technique which analyses data of multidimensional high order. It identifies the independent implicit information in order to remove redundant as well as extract independent information [6].

Kernel Principal Component Analysis (KPCA) is again similar to the regular PCA, however the difference lies in its application in which KPCA is applied in a none-linear manner. As a result, it creates a none-linear mapping scheme to maximize the variance within the considered data [7].

Graph-Based Kernel Principal Component Analysis are techniques that are prominent none-linear techniques that makes use of a cost function to construct a low-dimensional representation of the considered data. Such techniques includes Locally Linear Embedding (LLE), Hessian LLE, Laplacian Eigenmaps and Local Tangent Space Alignment (LTSA).

Linear Discriminant Analysis (LDA) is a method that finds linear combinations of features that can be characterized two or more classes. It is commonly used in statistics, pattern recognition and machine learning [8].

General Discriminant Analysis (GDA) is similar to LDA in terms of its objectives, in which it is to find crucial features in the considered data such that it lowers the dimensionality of the data. GDA operates with non-linear discriminants by which it makes use of a kernel function. Its underlying theory is similar to that of support vector machine (SVM).

2.1.3 *Supervised Learning*

The goal of supervised learning to to create a mapping function, such that when given new input data a prediction of the output data can be made. The mapping function is generated by with given data sets that consists of inputs with its corresponding out-

puts. Linear regression, random forest and SVM are popular supervised algorithms. Support Vector Machine (SVM) is a classifier mainly used for complex classification purposes. In order to separate new data into specified categories, SVM is trained with given labelled data making it supervised learning [9]. An implementation of SVM classifiers is done in [10] where the it is used to classify X-Ray images of the body into five categories, namely, head-neck, body, upper-limb and true-negative. Spatial Pyramid Histogram are used to decipher the medical images which is then used to train the Chi-Kernel based SVM. In general, SVM is used for binary classification [9].

Neural Networks in another common tool selection for AI, whereby it consists of three distinguished layers, the input, hidden and output layer. A general neural network consists of one or two hidden layers. The network is trained before use, which deems it as supervised classification or predictions. In [11], the technique implemented is back-propagation neural network (BPNN). The network is trained using a Supervised Delta Learning Rule. The result from the developed technique is a binary outcome in which it indicates whether the subject is normal or abnormal.

Studies of improving the common tools in AI have been done by combining the tools. There are two types of combinations, hybrid and non-hybrid multiple classifiers. Hybrid multiple classifiers combines different types of AI algorithms, whereas non-hybrid multiple classifiers only consists of one base algorithm and is replicated multiple times. A hybrid multiple classifier is done by [12], in which the algorithms used are K-Nearest Neighbour, Back Propagation Neural Network (BPNN) and SVM. According to the authors, the each of the AI tools is trained with a different set of data. The decision is based on a voting scheme that is called fusion selection. However, the paper presented in [12] performs binary results. In other words, it indicates whether a fracture is present or absent.

Another AI algorithm to be considered for medical diagnosis is deep learning. According to the author in [13], deep learning has the ability to combat complex problems whereas SVM would not be able to, since SVM works better with small data sets that have few outliers. However, the downfall of deep learning is its learning process, in which it can become tedious and computational demanding. Thus for small data sets, it is more plausible to use an off-shelf classifier such as SVM.

2.1.4 **Curse of Dimensionality** The main problematic factor affecting the performance in classification is the curse of dimensionality. The curse of dimensionality is a problem due to the sparsity of high dimensional spaces, in which an absurd amount of training data is needed in order to get low variance estimators [14].

Feature extraction with supervised learning algorithms may seem most desired compared to unsupervised learning algorithms, since supervised learning algorithms have more information about the problem and features. However, unsupervised methods do not suffer the curse of dimensionality as supervised learning does since it makes use of local measures to optimally estimate a single dimensional projection function [14].

2.1.5 **Combination Techniques** There have been studies done which combine the algorithms of supervised and unsupervised. The authors of [15] proposed a method which combines both supervised and unsupervised algorithms, name.

The authors of [16] designed a system in which it comprises of Principal Component Analysis(PCA) and Neural Network for the classification of EEG Signals. PCA is a form of unsupervised learning whilst neural network is supervised learning and requires training. The authors compared the correctness of the classification using Neural Network to using Principal Component Analysis with Neural Network. The final outcome from the experimentation is that the correctness of the classification of the Principal Component Analysis with Neural Network is better than Neural Network alone.

3. Problem Analysis

The modern world has presented the technology to predict human speech as well as digitally categorizing handwritten numbers. However, this technology has not been fully utilized in the medical field. There are many debates surrounding

the topic of whether machines are better than doctors. In many cases, the argument for doctors states that those who develop AI do not understand the complexity of medicine and those arguing for the machines say that not many understand how AI operates. However, at the end of the day there are still sick patients waiting to be treated with the right medication or medical care. In order to determine the type of medical care the patient needs a diagnosis must be performed. According to [17], where studies have been done around the United States, it was reported that doctors correctly diagnosed 55.3% for easier cases and 5.8% for more difficult cases. However, this is only one article and more investigation is needed to confirm these numbers. Even so, these numbers still indicate that doctors are not perfect and can mis-diagnosis patients. As a result of mis-diagnosis, the patient is subjected to stress or even life threatening situations should the disease be mis-treated. Therefore there is a large emphasis on accuracy and quick detection of diseases.

There are existing pattern recognition techniques available which have the ability to detect the presence of bone fracture. However, there are large amounts of datasets that is passed into the techniques. These datasets create dimensionality, the more datasets there are the higher the number of dimensionality there are. High number of dimensionalities affects the performance of the techniques. This is known as the curse of dimensionality.

4. Proposed Methodology

The purpose of this research is to develop a system which consists of two steps in order to determine the presence of a fracture with medical x-ray images as the system's input. The first step is a dimensionality reduction process and the second step is the automation of the decision making process. The objective of the dimensionality reduction process is to map the data from a high dimensionality to a low dimensionality without loosing any critical information needed for the decision making process.

The proposed approach adopted for this research is illustrated in Figure 1. The proposed system consists of an image processing technique, dimensionality reduction and an automated decision

making technique. The proposed system mainly focuses on the dimensionality reduction and automated decision making techniques. The objective of the proposed system is to determine the effects of using two steps compared to using one for the detection of bone fracture. The evaluation of the effects of the systems is done by observing the execution speed, classification accuracy and error rate.
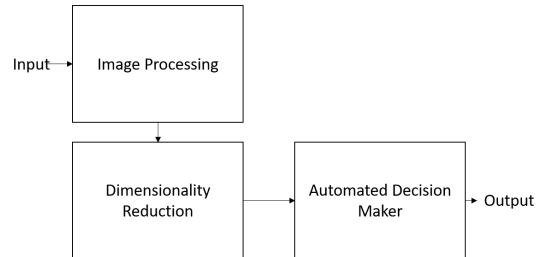

Figure 1 : Block Diagram illustrating an overview of the system

4.1 Feature Selection

Feature selection is performed through the process of image processing, whereby the Canny and Sobel techniques are utilized. The features selected are shown in Table 1.

Table 1 : Table showing the selected features and the dimensions of each feature.

Features	No. of Dimension	Description
gradient	2	hough line value, angle (in degrees)
lines	4	x_1, y_1 x_2, y_2
contours	N	line values that make up the contour

4.2 Dimensionality Reduction

The selected features illustrated in Table 1 have differing dimensions. The dimensions of the contour feature can vary depending on the type bone presented in the x-ray image. Therefore there is a high dimensionality in the dataset that represents one x-ray image. Although there is enough information from the selected features to represent an x-ray image, there is also redundant information present in the dataset. The automated decision making process has the ability to process the raw data, however in doing so it will hinder its over-

all performance. The PCA and ICA techniques are under consideration for the dimensionality reduction technique. Therefore it is subjected to change.

4.3 Automated Decision Making Techniques

The automation of bone fracture detection can be done by utilizing a neural network technique. The implementation of the neural network technique consists of two stages. The first stage is the training of the technique and the second is the classification of the bone fracture. The neural network technique has the ability to process the raw data from the selected features, however since the raw data has high dimensionality it has the potential to hinder the performance of both stages in the neural network technique. This will be observed during execution. The neural network technique in consideration is Back Propagation Neural Network, however this choice of technique is subjected to change.

5. Time Management

Table 2 presents the time management for the proposed research presented.

Table 2 : Table showing the time management of the implementation of a bone fracture detection two step system.

Date	Task
1 March	1st Draft Hand-In
2 March	Investigate best suited dimensionality reduction technique for selected features
8 March	Present 1st iteration of source code for the two step system
15 March	Present 2nd iteration with automated decision making process implemented
22 March	Submit Proposal

6. Conclusion

To conclude, the system proposed for the detection of bone fracture consists of two steps. The first step is dimensionality reduction and the second step is the automation of decision making. The objective of implementing a dimensionality reduction process is to improve the performance of the automated bone fracture detection. The performance is measured by speed, classification accuracy and error rate.

REFERENCES

[1] A. N. Ramesh, C. Kambhampati, J. R. T. Monson, and P. J. Drew. "Artificial intelligence in medicine." *Annals of the Royal College of Surgeons of England*, vol. 86, no. 5, pp. 334–338, 2004.

[2] S. K. Mahendran and S. S. Baboo. "Automatic Fracture Detection Using Classifiers- A Review." vol. 8, no. 6, pp. 340–345, 2011.

[3] I. K. F. Center, A. Scienti, and L. L. N. L. P. O. Box. "A survey of dimension reduction techniques." *Computing*, , no. 1, pp. 1–18, 2002.

[4] K. Mahajan, M. Vargantwar, and S. M. Rajput. "Classification of EEG using PCA, ICA and Neural Network." *International Journal of Engineering and Advanced Technology*, vol. 1, no. 1, pp. 80–83, 2011.

[5] I. T. Jolliffe, J. Cadima, and J. Cadima. "Principal component analysis : a review and recent developments Subject Areas : Author for correspondence :." 2016.

[6] Q. Cai and X. Tang. "A digital audio watermarking algorithm based on independent component analysis." In *2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, pp. 1053–1058. Oct 2016.

[7] D. Kapsoulis, K. Tsiakas, V. Asouti, and K. Giannakoglou. "The use of Kernel PCA in evolutionary optimization for computationally demanding engineering applications." In *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 1–8. Dec 2016.

[8] E. I. G. Nassara, E. Grall-Mas, and M. Kharouf. "Linear Discriminant Analysis for Large-Scale Data: Application on Text and Image Data." In *2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pp. 961–964. Dec 2016.

[9] P. Rebentrost, M. Mohseni, and S. Lloyd. "Quantum support vector machine for big data classification." *Physical Review Letters*, vol. 113, no. 3, pp. 1–5, 2014.

[10] S. Saha, A. Mahmud, A. A. Ali, and M. A. Amin. "Classifying digital X-ray images into different human body parts." In *2016 5th International Conference on Informatics, Electronics and Vision (ICIEV)*, pp. 67–71.

[11] S. Jyothi and K. Vanisree. "Congenital Heart Septum Defect Diagnosis on Chest X-Ray Features Using Neural Networks." In *2016 Second International Conference on Computational Intelligence Communication Technology (CICT)*, pp. 265–269.

[12] N. Umadevi and S. N. Geethalakshmi. "Multiple classification system for fracture detection in human bone x-ray images." In *2012 Third International Conference on Computing Communication Networking Technologies (ICCCNT)*, pp. 1–8.

[13] Y. LeCun, Y. Bengio, and G. Hinton. "Deep learning." *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[14] N. Intrator. "Feature Extraction Using an Unsupervised Neural Network." vol. 4, no. 1, pp. 98–107.

[15] V. E. Neagoe and R. M. Stoica. "A new neural network-based approach for automatic annotation of remote sensing imagery." In *2014 IEEE Geoscience and Remote Sensing Symposium*, pp. 1781–1784.

[16] R. Kottaimalai, M. P. Rajasekaran, V. Selvam, and B. Kannapiran. "EEG signal classification using Principal Component Analysis with Neural Network in Brain Computer Interface applications." In *2013 IEEE International Conference ON Emerging Trends in Computing, Communication and Nanotechnology (ICECCN)*, pp. 227–231.

[17] A. N. D. Meyer, V. L. Payne, D. W. Meeks, R. Rao, and H. Singh. "Physicians Diagnostic Accuracy, Confidence, and Resource Requests: A Vignette Study." vol. 173, no. 21, pp. 1952–1958.