

ISSN : 2165-4069(Online)

ISSN : 2165-4050(Print)



IJARAI

International Journal of
Advanced Research in Artificial Intelligence

Volume 3 Issue 4

www.ijarai.thesai.org

A Publication of
The Science and Information Organization



INTERNATIONAL JOURNAL OF
ADVANCED RESEARCH IN ARTIFICIAL INTELLIGENCE



THE SCIENCE AND INFORMATION ORGANIZATION

www.thesai.org | info@thesai.org



Editorial Preface

From the Desk of Managing Editor...

"The question of whether computers can think is like the question of whether submarines can swim." — Edsger W. Dijkstra, the quote explains the power of Artificial Intelligence in computers with the changing landscape. The renaissance stimulated by the field of Artificial Intelligence is generating multiple formats and channels of creativity and innovation.

This journal is a special track on Artificial Intelligence by The Science and Information Organization and aims to be a leading forum for engineers, researchers and practitioners throughout the world.

The journal reports results achieved; proposals for new ways of looking at AI problems and include demonstrations of effectiveness. Papers describing existing technologies or algorithms integrating multiple systems are welcomed. IJARAI also invites papers on real life applications, which should describe the current scenarios, proposed solution, emphasize its novelty, and present an in-depth evaluation of the AI techniques being exploited. IJARAI focusses on quality and relevance in its publications.

In addition, IJARAI recognizes the importance of international influences on Artificial Intelligence and seeks international input in all aspects of the journal, including content, authorship of papers, readership, paper reviewers, and Editorial Board membership.

The success of authors and the journal is interdependent. While the Journal is in its initial phase, it is not only the Editor whose work is crucial to producing the journal. The editorial board members, the peer reviewers, scholars around the world who assess submissions, students, and institutions who generously give their expertise in factors small and large— their constant encouragement has helped a lot in the progress of the journal and shall help in future to earn credibility amongst all the reader members.

I add a personal thanks to the whole team that has catalysed so much, and I wish everyone who has been connected with the Journal the very best for the future.

Thank you for Sharing Wisdom!

Editor-in-Chief

IJARAI

Volume 3 Issue 4 April 2014

ISSN: 2165-4069(Online)

ISSN: 2165-4050(Print)

©2013 The Science and Information (SAI) Organization

Editorial Board

Peter Sapaty - Editor-in-Chief

National Academy of Sciences of Ukraine

Domains of Research: Artificial Intelligence

Alaa F. Sheta

Electronics Research Institute (ERI)

Domain of Research: Evolutionary Computation, System Identification, Automation and Control, Artificial Neural Networks, Fuzzy Logic, Image Processing, Software Reliability, Software Cost Estimation, Swarm Intelligence, Robotics

Antonio Dourado

University of Coimbra

Domain of Research: Computational Intelligence, Signal Processing, data mining for medical and industrial applications, and intelligent control.

David M W Powers

Flinders University

Domain of Research: Language Learning, Cognitive Science and Evolutionary Robotics, Unsupervised Learning, Evaluation, Human Factors, Natural Language Learning, Computational Psycholinguistics, Cognitive Neuroscience, Brain Computer Interface, Sensor Fusion, Model Fusion, Ensembles and Stacking, Self-organization of Ontologies, Sensory-Motor Perception and Reactivity, Feature Selection, Dimension Reduction, Information Retrieval, Information Visualization, Embodied Conversational Agents

Liming Luke Chen

University of Ulster

Domain of Research: Semantic and knowledge technologies, Artificial Intelligence

T. V. Prasad

Lingaya's University

Domain of Research: Bioinformatics, Natural Language Processing, Image Processing, Robotics, Knowledge Representation

Wichian Sittiprapaporn

Maharakham University

Domain of Research: Cognitive Neuroscience; Cognitive Science

Yaxin Bi

University of Ulster

Domains of Research: Ensemble Learning/Machine Learning, Multiple Classification Systems, Evidence Theory, Text Analytics and Sentiment Analysis

Reviewer Board Members

- **AKRAM BELGHITH**
University Of California, San Diego
- **ALAA F. SHETA**
Electronics Research Institute (ERI)
- **ALBERT ALEXANDER**
Kongu Engineering College
- **ALPA KALPESH RESHAMWALA**
NMIMS, MPSTME
- **AMIR HAJJAM EL HASSANI**
Université de Technologie de Belfort-Monbéliard
- **AMIT VERMA**
Department in Rayat & Bahra Engineering College, Mo
- **AMITAVA BISWAS**
Cisco Systems
- **ANTONIO DOURADO**
University of Coimbra
- **Alexandre Bouënard**
- **ASIM TOKGOZ**
Marmara University
- **B R SARATH KUMAR**
LENORA COLLEGE OF ENGINEERNG
- **BABATUNDE OPEOLUWA AKINKUNMI**
University of Ibadan
- **BESTOUN S.AHMED**
Universiti Sains Malaysia
- **BHANU PRASAD PINNAMANENI**
Rajalakshmi Engineering College; Matrix Vision GmbH
- **Badre Bossoufi**
University of Liege
- **CHIEN-PENG HO**
Information and Communications Research Laboratories, Industrial Technology Research Institute of Taiwan
- **DAVID M W POWERS**
Flinders University
- **Dewi Nasien**
Universiti Teknologi Malaysia
- **Dr.VUDA SREENIVASARAO**
School of Computing and Electrical Engineering,BAHIR DAR UNIVERSITY, BAHIR DAR,ETHIOPA
- **DIMITRIS CHRYSOSTOMOU**
Production and Management Engineering / Democritus University of Thrace
- **EHSAN MOHEBI**
University of Ballarat
- **FABIO MERCORIO**
University of Milan-Bicocca
- **FRANCESCO PERROTTA**
University of Macerata
- **FRANK IBIKUNLE**
Covenant University
- **GERARD DUMANCAS**
Oklahoma Medical Research Foundation
- **GORAKSH GARJE**
Pune Vidyarthi Griha's College of Engineering and Technology, Pune
- **GRIGORAS GHEORGHE**
"Gheorghe Asachi" Technical University of Iasi, Romania
- **GUANDONG XU**
Victoria University
- **HAIBO YU**
Shanghai Jiao Tong University
- **HARCO LESLIE HENDRIC SPITS WARNARS**
Budi LUhur university
- **IBRAHIM ADEPOJU ADEYANJU**
Ladoke Akintola University of Technology, Ogbomoso, Nigeria
- **IMRAN CHAUDHRY**
National University of Sciences & Technology, Islamabad
- **JABAR H YOUSIF**
Faculty of computing and Information Technology, Sohar University, Oman
- **JATINDERKUMAR R. SAINI**
S.P.College of Engineering, Gujarat
- **JOSÉ SANTOS REYES**
University of A Coruña (Spain)
- **KRASIMIR YORDZHEV**
South-West University, Faculty of Mathematics and Natural Sciences, Blagoevgrad, Bulgaria
- **KRISHNA PRASAD MIYAPURAM**
University of Trento

- **Le Li**
University of Waterloo
- **Leon Abdillah**
- Bina Darma University
- **LIMING LUKE CHEN**
University of Ulster
- **Ljubomir Jerinic**
University of Novi Sad, Faculty of Sciences,
Department of Mathematics and
Computer Science
- **M. REZA MASHINCHI**
- **MALACK OTERI**
jkuat
- **MAREK REFORMAT**
University of Alberta
- **MD. ZIA UR RAHMAN**
Narasaraopeta Engg. College,
Narasaraopeta
- **Mehdi Bahrami**
University of California, Merced
- **MOHAMED NAJEH LAKHOUA**
ESTI, University of Carthage
- **MOKHTAR BELDJEHEM**
University of Ottawa
- **MONJI KHERALLAH**
University of Sfax
- **Nidhi Arora**
M.C.A. Institute, Ganpat University
- **PARMINDER SINGH KANG**
De Montfort University, Leicester, UK
- **PETER SAPATY**
National Academy of Sciences of Ukraine
- **PRASUN CHAKRABARTI**
Sir Padampat Singhanian University
- **QIFENG QIAO**
University of Virginia
- **RAJESH KUMAR**
National University of Singapore
- **RASHAD AL-JAWFI**
Ibb University
- **REZA FAZEL-REZAI**
Electrical Engineering Department,
University of North Dakota
- **SAID GHONIEMY**
Taif University
- **Secui Calin**
IEEE Membership; IEEE Power & Energy
Society Membership; IEEE Computational
Intelligence Society Membership
- **Selem Charfi**
University of Valenciennes and Hainaut
Cambresis, France
- **SHAHABODDIN SHAMSHIRBAND**
University of Malaya
- **SIMON EWEDAFE**
Baze University
- **SUKUMAR SENTHILKUMAR**
Universiti Sains Malaysia
- **T C.MANJUNATH**
HKBK College of Engg
- **T V NARAYANA RAO**
Hyderabad Institute of Technology and
Management
- **T. V. PRASAD**
Lingaya's University
- **V BABY DEEPA**
- **VISHAL GOYAL**
- **VITUS S.W. LAM**
- **WEI ZHONG**
University of south Carolina Upstate
- **WICHIAN SITIPRAPAPORN**
Mahasarakham University
- **YAXIN BI**
University of Ulster
- **YUVAL COHEN**
The Open University of Israel
- **ZHAO ZHANG**
Deptment of EE, City University of Hong
Kong
- **ZHIGANG YIN**
Institute of Linguistics, Chinese Academy of
Social Sciences
- **ZNE-JUNG LEE**
Dept. of Information management, Huafan
University

CONTENTS

Paper 1: A Registration Method for Multimodal Medical Images Using Contours Mutual Information

Authors: Ying Qian, Meng Li, Qingjie Wei, Xuemei Ren

PAGE 1 – 5

Paper 2: Framework for Knowledge-Based Intelligent Clinical Decisionsupport to Predict Comorbidity

Authors: Ernest E. Onuiru, OludeleAwodele, Sunday A. Idowu

PAGE 6 – 17

Paper 3: Scale-Based Local Feature Selection for Scene Text Recognition

Authors: Boyu Zhang, JiaFeng Liu, XiangLong Tang

PAGE 18 – 23

A Registration Method for Multimodal Medical Images Using Contours Mutual Information

Ying Qian

The lab of Graphics and Multimedia
Chongqing University of Posts and Telecommunications
Chongqing, China

Meng Li, Qingjie Wei, Xuemei Ren

The lab of Graphics and Multimedia
Chongqing University of Posts and Telecommunications
Chongqing, China

Abstract—In recent years, mutual information has developed as a popular image registration measure especially in multimodality image registration. For different modality medical images, the contour of tissues or organs is similarity. In this paper, an effective new registration method of the multimodal medical images based on the contour mutual information is proposed. Firstly, get the contour through variational level set method. Secondly, within the contour pixels are assigned the same grayscale value, obtain two contour images. Finally, two contour images using mutual information as similarity measure for image registration. The experiment results show that the registration algorithm proposed in this paper can do more effectively and more accurately work than normalized mutual information and gradient mutual information. (*Abstract*)

Keywords—contour mutual information; mutual information; multimodal medical image; image registration; variational level set method

I. INTRODUCTION

In actual clinical diagnosis, a single modal of image usually unable to provide enough information to doctors. So we need to have the step of different modal image fusion, thus providing more information to doctor, and then they can understand the comprehensive information from pathological tissue or organs, make better medical diagnosis. Moreover, image registration is a premise key in image fusion, so the multimodal medical image registration becomes a research hotspot in the field of medical image processing. Medical image registration is the ascertaining process of spatial mapping between two images that differs in image acquisition time, image properties, or viewpoint and subsequently producing a result image that is informative. At present, medical image registration methods can be classified as feature-based and intensity-based [1]. In intensity-based registration, mutual information (MI) is one of the intensity based measure, and does not require the definition of landmarks or features such as surfaces [2]. It is an automatic method and it applies to a variety of modal image registration [3].

The multimodal image registration method based on mutual information of image gray intensity, which is no need for us to preprocess. In order to improve the accuracy of the registration, so many scholars made deep researches and improved the registration methods.

Studholme [4] proposed a method based on normalized mutual information. But the method ignores the spatial information. Plum [5] proposed to include spatial information by multiplying MI measure with a gradient term based on the magnitude and the orientation of the gradients. But gradient is sensitive to noise, thus the success rate become lower when images contains noise. Butz and Thiran [6] used mutual information in edge computing. But the edge information is sparse and the mutual information is sensitive to the number of sampling point, thus rendering appear local maximum and result in mismatching. Loeckx [2] proposed a method based on conditional mutual information, spread gray joint histogram from two-dimension to three-dimension. But the method has high computational cost. Ruecker [7] combined the mutual information with image's texture information which be described as gray level co-occurrence matrix. But this method assume that image's texture information have similarity when organ in different modal images.

This paper uses the characteristic that the contours of different modal medical images have similarity, proposed a registration method based on contour mutual information(CMI), extracts contours effectively from images which to be registered, and within the contour pixels are assigned the same gray-scale, That is, contour image. Finally, two contour images using mutual information as similarity measure for image registration. CMI efficiently solves problems of poor robustness and low accuracy in image registration.

II. MUTUAL INFORMATION

A. Definition

MI is selected as the similar measure. MI is an important concept of information theory which is used to represent the statistics correlation of two sets of data which has been widely used in the image registration. Starting from a reference image X and floating image Y with intensity bins x and y , the calculation of MI of two images X and Y is as follows:

$$I(X,Y)=H(X)+H(Y)-H(X,Y) \quad \square\square\square \quad (1)$$

$$H(X,Y)=-\sum_{x,y}p(x,y)\log(p(x,y)) \quad \text{and}$$

$H(X)=-\sum_x p(x)\log(p(x))$ the joint and marginal entropy of random variables X and Y .

B. Problems

Mutual information registration method based on gray value is a statistical method for registration, has nothing to do with the pixel location. Constructs a size for the test image A (256 * 256), pixels of each column from top to bottom, with gray value from 0 to 255, as shown in Fig.1(a), A rotate 180°, obtain test image B, as shown in Fig.1(b). A is the reference image, B is the floating image. Floating image 360° rotation around itself (from -180° to 180°), Observe the reference image and the floating image of the mutual information curve, Respectively, at 0° and 180°, the maximum mutual information appears, as shown in Fig.1(c).

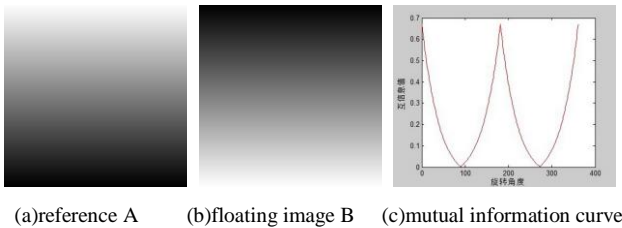


Fig. 1. Test images and mutual information curve.

The research direction of many researchers is how the spatial information of image is introduced to the mutual information to produce a new similarity measure. But the robust is not very good; the main reasons are the following:

- As shown in Fig.2(a), image with noise, weaken the correlation between images, easy to fall into local extreme value leads to wrong registration;
- As shown in Fig.2(b), if the image appears to be missing which make registration less effective, the registration effect will be worse;
- As shown in Fig.2(c), due to less spatial resolution and image information content, the image of low resolution images can lead to the problem of robust ;
- Images to be registered have less effective information, such as the top of the skull, as shown in Fig.2(d).

In this paper, analyzes the robustness of registration according to mutual information measure curve. In Fig.1, the abscissa of mutual information is the rotation angle (-25° to 25°), the vertical axis is the amount of the mutual information. As can be seen from the figure, there are have different degrees of local maxima in the four cases mentioned above.

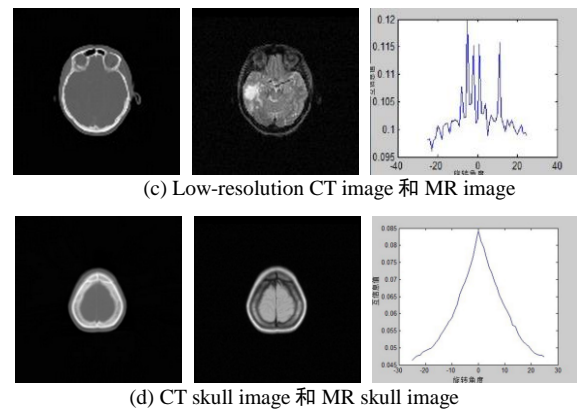
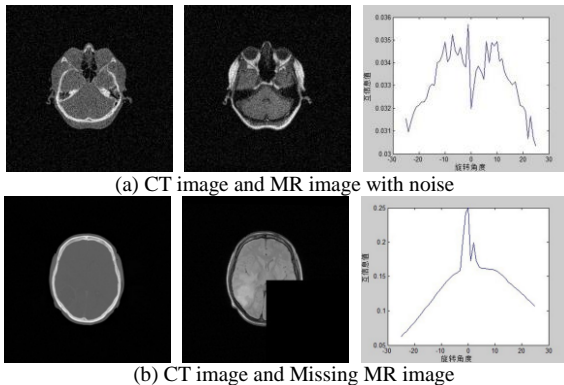


Fig. 2. Mutual information curves under different circumstances.

Although some of the improved algorithm considers the spatial information of the image, but have a major impact on the registration is still gray values, its spatial information is essentially the use of the gray information, because of the limitations are not widely used.

III. METHOD

Contour as a stable characteristic of image, when the image translating or rotating, which can keep better corresponding relation. So before get contour image, we should extract contours of image.

The contour extraction method will choose an operator to detect the image edge which based on the image edge features. Then we will connect the edge point into a closed curve to obtain the target contour.

Commonly used methods of edge detection include: Roberts operator, Sobel operator, Prewitt operator, etc. But these edge detection methods are sensitive to noise, and will enhance the noise in the edge detection. This paper mainly use feature of contour, so effectively getting contour of image is the key of CMI.

In this paper, we adopt a variational level set method proposed by Li [8], which have no need to reinitialize. The method bring penalty term in energy function, thus ensuring level set function keep become sign function in evolutionary process, and it avoids reinitialize in the process. After get contours of two images, filling contour unified gray value and then get contour image.

CMI has two advantages, firstly, ensures the strict corresponding relations exist between two contours images, add gray-scale correlation between images. Secondly, it used spatial information that contours in different modal images have similarity.

The following describes specific steps which about a way of image registration based on contour mutual information:

- 1) Two images are given to be registered, and they are reference image R and floating image F.
- 2) Get contours of two images by variational level set method, and then fill unified gray value to get contour image (R', F').

3) Mutual information be used as measure function, and Powell optimization algorithm execute optimization methodology of parameters for image R' and F' . The process of optimization uses the nearest neighbor interpolation algorithm.

4) Optimal spatial transform parameters reform the floating image F according to the Powell algorithm, obtain registration results.

IV. EXPERIMENT

In this paper, the experiment data are come from Retrospective Registration Evaluation Project of the Vanderbilt University, and the images resolution is 256×256 . The experiment 1 validates the robustness of CMI, the rest experiments were compared with Gradient Mutual Information (GMI, from reference [3]) and Normalized Mutual Information (NMI, from reference [4]).

A. Experiment 1 Validate the Robustness

According to mutual information curve from Fig.3, using image mutual information method can lead peak value become not so sharp. However, the method proposed in this paper has a sharp peak, and it is easy to detect the best position of registration. Meanwhile, it is not easy to appear the phenomenon of local maximum because of the smooth curve.

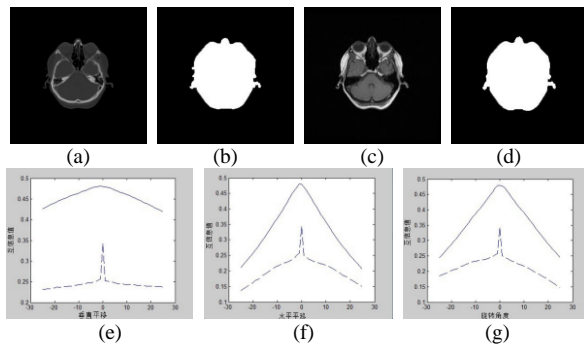


Fig. 3. The results of robustness of experiments.

(a) is CT original registration image and (b) is MR original registration image, (c) is CT contour image and (d) is MR contour image. (e) is mutual information curve based on image gray value and the method of paper to rotate the images (from -25° to 25°). (f) is mutual information curve based on image gray value and the method of paper for vertical translation (from -25 To 25 pixels). (g) is mutual information curve based on image gray value and the method of paper for horizontal translation. The solid line is the measure of the curve based on image gray value and the dashed line is the measure of the curve based on the contour mutual information.

In order to validate the robustness of the registration algorithm based on contour mutual information, using the method of this thesis to create mutual information curve, as Fig.4. Compared the curve of Fig.2, the mutual information curve created by CMI has good robustness.

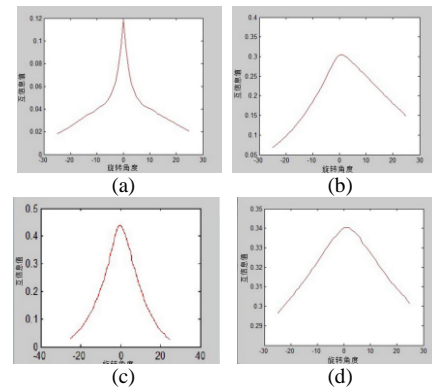


Fig. 4. Measure curve based on contour mutual information.

(a)(b)(c)(d) Corresponding to CMI of test image in Fig.1 mutual information measure curve.

B. Experiment 2 Validate the Accuracy

As shown in Fig.5, the edge images of CT are applied to the MR images, by this to detect the effect of registration. The bone tissue in MR images are close to the background gray-scale value, However, it have large different with the background gray-scale value of CT image, thus leading to registration errors which use the method of mutual information. As shown in Fig.5 (d); The gradient mutual information, due to considerate the information of space and orientation, the registration results have more advantages compare to the method which based on normalized mutual information; From the Fig.5 (e), there are many errors at edges, but effects obtained by CMI show that the image has been substantially aligned at edges after registration.

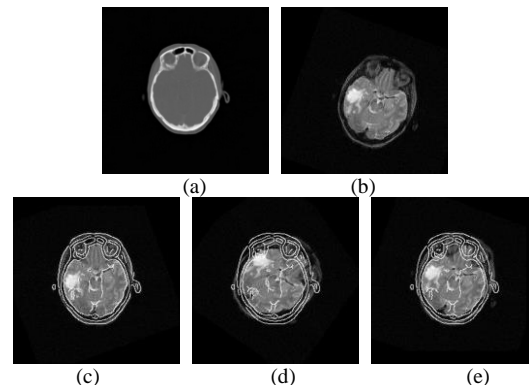


Fig. 5. Experimental results verify the accuracy of CMI.

C. Experiment 3 Validate the robustness when Image have Missing information

From Fig.6 (a), CT image are mainly reflects the skeletal structure, its gray-level changes small.

We mainly use edges information to register with MR image. When Image has large missing information, it will mismatch easier. From experimental results, CMI has obvious advantages compared with other two methods in gray-scale difference.

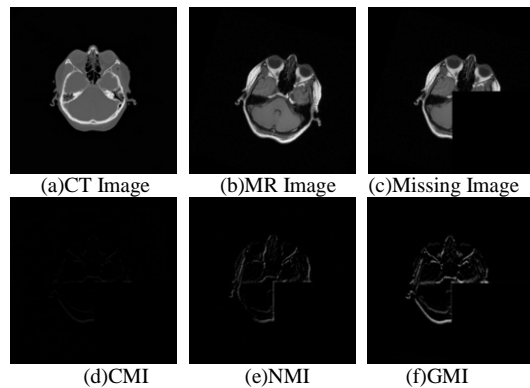


Fig. 6. The results on the robustness of image deletion.

(a)(b)(c)are CT images and MR image in the same layer from the head of a patient. (a) is original CT image (b) is original MR (T2-weighted) image, (c) have missing (Setting the region of the image's gray value to 0 to simulate the effect of information lacking). (d)(e)(f)corresponding to the result of the registration image obtained by CMI, NMI and GMI which subtract registration images.

D. Experiment 4 validate the robustness when image has noise

The gradient is sensitive to noise, when image have a little of noise, the registration method which based on gradient mutual information will easily appear local maximum, it will results in mismatching. As shown in Fig.7 (a), this paper choose three groups of head images in CT modal and MR modal (T1 weighted, T2 weighted, PD weighted), which have been registered and in the corresponding layer. As shown in Fig.7 (a), add noise in CT and MR images. Noised image of CT as a reference image, let the MR images horizontal, vertical position and angle transformation from $(-25,-25,-25^\circ)$ to $(25, 25, 25^\circ)$ get floating images, to get 50×3 groups of experimental subjects. Three parameters among them are the horizontal of translation (pixel), vertical translation (pixel), rotation angle. Results from three methods are shown in Table1. From data in the Table1we can see the registration of the CMI is more accurate than other two methods. In the CT image and PD image registration, the noise will make impact on gradient mutual information, thus making the effect of registration unsatisfactory.

TABLE I. THE EXPERIMENTAL RESULTS

Experimental object	Robustness verification method for noisy image		
	CMI	NMI	GMI
CT-T1	86%	80%	80%
CT-T2	86%	82%	78%
CT-PD	90%	84%	82%

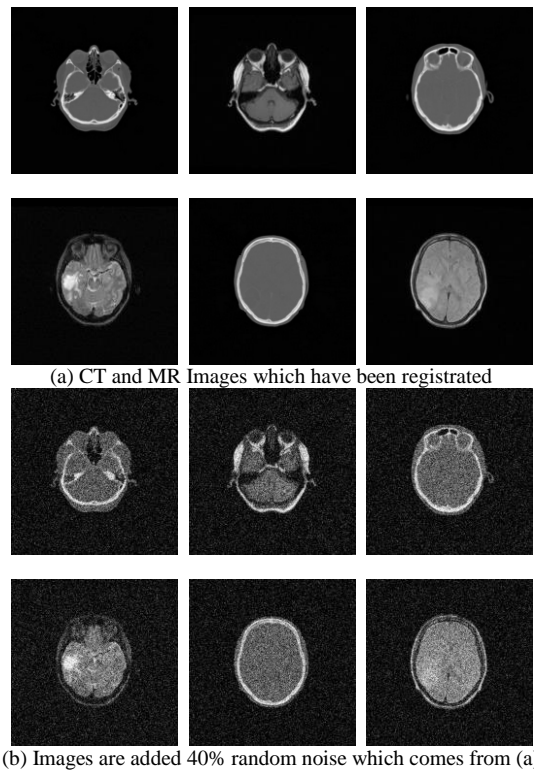


Fig. 7. Robustness verification method for noisy image.

V. CONCLUSIONS

Recently, the algorithm in multimodality medical image registration has poor robust and accuracy. This paper proposed a registration method based on contour mutual information. On the one hand, it uses a spatial characteristic, that the contours of tissue and organs in different modal have similarity. On the other hand, the method that fill the same gray values to the contour, which can add the correlation information among contour image and reduce the interference in image registration when image have missing information, thus improving precision of registration. The experiments show that the method of this paper can achieve more accurate registration results. But this paper has not considered nonlinear registration in multimodality medical image. We will make more research in further work.

REFERENCES

- [1] Zitova B, Flusser J, Image registration methods a survey, Image and vision computing, pp: 977-1000, 2003.
- [2] Loeckx D, Slagmolen P, Maes F, Vandermeulen D, Suetens P, Non-rigid image registration using conditional mutual information, IEEE transactions on medical imaging, pp:19-29, 2010.
- [3] Nigris D, Collins D, Arbel T, Multi-modal image registration based on gradient orientations of minimal uncertainty, IEEE transactions on medical imaging, pp:2343-2354, 2012.
- [4] Studholme C, Hill D L G, Hawkes D J, An overlap invariant entropy measure of 3D medical image alignment, Pattern recognition, pp: 71 -86, 1999.
- [5] Pluim J P, Maintz J B, Viergever M A, Image registration by maximization of combined mutual information and gradient information, IEEE transactions on medical imaging, pp: 809-814,2000

- [6] Butz T, Thiran J P, Affine registration with feature space mutual information, MICCAI 2001, LNCS 2208. 549-556.
- [7] Rueckert D, Clarkson M J, Hill D L G, Hawkes D J, Non-grid registration using higher order mutual information, Proceeding of SPIE Medical Imaging 2000 : Image Processing, San Diego, pp:438-447, 2000.
- [8] Li CM, Xu CY, Gui CF, et.al. Level set evolution without re-initialization: A new variational formulation. In: Schmid C, Soatto S, Tomasi C, eds. Proc. of the 2005 IEEE Computer Society Conference. on Computer Vision and Pattern Recognition. Washington: IEEE Computer Society Press, pp: 430 – 436, 2005.

Framework for Knowledge–Based Intelligent Clinical Decision support to Predict Comorbidity

Ernest E. Onuiri, OludeleAwodele and Sunday A. Idowu
Department of Computer Science, Babcock University,
Ilishan-Remo, Ogun State, Nigeria

Abstract—Research in medicine has shown that comorbidity is prevalent among chronic diseases. In ophthalmology, it is used to refer to the overlap of two or more ophthalmic disorders. The comorbidity of cataract and glaucoma has continued to gain increasing prominence in ophthalmology within the past few decades and poses a major concern to practitioners. The situation is made worse by the dearth in number of ophthalmologists in Nigeria vis-à-vis Sub-Saharan Africa, making it most inevitable that patients will find themselves more at the mercies of General Practitioners (GPs) who are not experts in this domain of interest. To stem the tide, we designed a framework that adopts a knowledge-based Clinical Decision Support System (CDSS) approach to deal with predicting ophthalmic comorbidity as well as the generation of patient-specific care plans at the point of care. This research which is within the domain of medical/healthcare informatics was carried out through an in-depth understanding of the intricacies associated with knowledge representation/preprocessing of relevant domain knowledge. Furthermore, we present the Comorbidity Ontological Framework for Intelligent Prediction (COFIP) in which Artificial Neural Network and Decision Trees, both being mechanisms of Artificial Intelligence (AI) was embedded into the framework to give it an intelligent (predictive and adaptive) capability. This framework provides the platform for a CDSS that is diagnostic, predictive and preventive. This is because the framework was designed to predict with satisfactory accuracy, the tendency of a patient with either of cataract or glaucoma to degenerate into a state comorbidity. Furthermore, because this framework is generic in outlook, it can be adapted for other chronic diseases of interest within the medical informatics research community.

Keywords—*Framework; Knowledge-based; Comorbidity; Clinical Decision Support System (CDSS)*

I. INTRODUCTION

In today's contemporary times, a trend that seems to be gaining a lot of ground is the integration of intelligent mechanisms in the development of applications to enable them make decisions and attempt to behave like humans. This is widespread in Expert Systems, a branch of AI which Professor Edward Feigenbaum of Stanford University defines as the use of knowledge and inference procedures to solve problems that are difficult enough to require significant human expertise for their solution [1]. Other areas gaining popularity include data mining, machine learning, neural networks, natural language processing, semantic web and so on. These different sub-fields of AI to a large extent make use of knowledge representation techniques. One of such techniques is ontology. "Ontology from the perspective of AI is a model that represents a set of

concepts within a specific domain as well as the relationships between those concepts" for the purpose of communicating knowledge between entities and how they inter-relate [2]. In addition, ontology describes a formal specification of a certain domain. It is constituted by a specific vocabulary used to describe a certain phenomenon as well as a set of explicit assumptions as to the intended meaning of the vocabulary [3]. It is the implementation of concepts like ontology in the development of applications such as Decision Support Systems (DSSs) that make them act intelligently.

A DSS is a computer application designed to aid decision makers with the task of decision making. Hence, it can be said that CDSS associates health observations with health knowledge in order to influence the health choices made by clinicians to improve healthcare. A CDSS can be said to be an active knowledge system, where two or more items of patient data are used to generate case-specific recommendation(s) [4]. This means that a CDSS is a DSS that uses knowledge management to achieve clinical advice for patient care based on some number of items of patient data. This goes a long way in easing the job of healthcare practitioners, especially in areas where the number of patients is overwhelming. In medical domains like ophthalmology, there is a dire need of CDSSs, given the few specialists in this area of medical practice in regions of the world like Africa.

Co morbidity is defined as "any distinct clinical entity that has co-existed or may occur during the clinical course of a patient who has the primary disease under study" [5]. Within ophthalmology, co morbidity is commonly used to refer to the co-existence of two or more ophthalmic disorders. Comorbidity between cataract and glaucoma disorders have gained increasing prominence in ophthalmology within the past few decades. In a survey carried out in Sweden, it was substantiated that as much as 36% of patients with cataract also had ocular comorbidities [6].

More so, it has been observed that chronic diseases are often associated with comorbidities. In view of the foregoing National Centre of Health Statistics reported that some of the reasons that explain this trend range from inadequate hospital resources, long waits in hospitals and inadequate medical practitioners [7]. Therefore, one can conclude that comorbidity can be referred to as a condition caused by the debilitating effects of a prevailing ailment. Cataract (Catt) is one of such ailment that can trigger comorbidities especially glaucoma. It is worthy of note that in some instances, glaucoma can also be the prevailing ailment that triggers cataract. Hence, the

comorbid condition of choice is that of cataract and Glaucoma (Gla) because both disorders are prevalent in sub-Saharan Africa.

Furthermore, in keeping with this established trend is the need for CDSSs of this nature to help GPs solve problems that are outside their knowledge-base/expertise (in this case, ophthalmology). Consequently, knowledge transition tools such as Evidence-Based Clinical Algorithms (EBCAs), which includes Clinical Practice Guidelines (CPGs) and Clinical Pathways (CPs), significantly go a long way in trying to reduce this care gap caused by the absence of an up-to-date knowledge [8]. CPGs are a function of a detailed and in-depth evaluation of scientific evidence about a specific medical condition/disease/procedure, designed for informed recommendations to aid clinicians in making decisions based on adequate evidence [9]. CPs on the other hand is used to implement the recommendations generated by the CPGs in actual clinical practice [10]. CPs also specifies the clinical processes as well as their workflow to implement the CPGs in a specific clinical setting. Consequently, a CPG entails *medical knowledge* whereas a CP entails *operational knowledge* about how to implement the CPG—i.e. the domain-specific protocols specifying the actual sequencing, decisions and scheduling of clinical tasks, as per the CPG, for the entire clinical course [11].

In view of the foregoing, a major challenge that arises is that which pertains to the alignment of multiple CPs of the comorbid diseases while conserving the integrity of medical knowledge and task pragmatics, and also ensuring patient safety. Hence, in this work, we modelled an intelligent/predictive CDSS for ophthalmic co-morbidity at the point of care. In addition, part of the emphasis was the examination and development of methods that use EBCAs to formalize, model, align and predict patient-specific clinical recommendations with a feedback mechanism, care coordination and decision support for ophthalmic comorbiditybased on substantial clinical evidence.

II. OBJECTIVES AND METHODOLOGIES

The main aim of this research was to design a predictive framework that adopts a knowledge-based CDSS approach to deal with ophthalmic comorbidity at the point of care. Therefore, preliminary methods for the formalization, modelling, alignment and prediction of EBCAs in comorbidities were determined and outlined. Focus was on the prevalent comorbidity of Cataract (Catt) and Glaucoma (Gla), because both are very rampant and common in sub-Saharan Africa and blacks in general. Accordingly, the objectives are to:

- 1) *Outline a stepwise description/exposition of how to build useful and valid knowledge-based CPs (guidelines/framework) through knowledge identification, acquisition, synthesis, formalization and alignment of relevant datasets especially those related to ophthalmic diseases of interest (cataract and glaucoma). This was achieved through an extensive review of literature leading to an in-depth understanding of the intricacies associated with knowledge representation/preprocessing of relevant*

domain knowledge. In doing this, the background natures of the ophthalmic diseases of interest are revealed.

- 2) *identify relevant classification/predictive algorithms to provide intelligence for the proposed intelligent framework that makes the system adaptive and consequently improves on the existing framework. This was carried out by investigating the appropriate/suitable AI mechanisms aimed at giving adaptability to the proposed framework. This was done by looking at results obtained when such mechanisms were deployed in existing frameworks.*

- 3) *Design a framework that is diagnostic, predictive and preventive. Hence, the system when operational zed, will predict the chances of a patient with either of cataract or glaucoma to degenerate into a state comorbidity. This led to the adoption and embedding of Artificial Neural Network (ANN – feed-forward back multi-layer propagation) and Decision Trees (DTs – C5.0 algorithm) which are tools of AI into the framework to give it an intelligent (predictive and adaptive) capability.*

III. ANALYSIS OF EXISTING AND RELATED FRAMEWORKS, THEIR FEATURES AND FINDINGS

This section introduces a discourse into the workings of existing and related frameworks especially COMET (Comorbidity Ontological Modeling & Execution) and PEDSS (Perinatal DSS).

A. COMET Framework

The system in Fig 1 is an ontology-based decision support framework called COMET for handling comorbidities by the alignment of ontologically modelled CPGs. It is built to formalize, model, align and operationalize the evidence-based clinical algorithms of co-morbid Chronic Heart Failure (CHF) and Atrial Fibrillation (AF) in order to provide evidence-based clinical recommendations, care coordination and decision support to GPs for effective management of CHF and AF.

Consequently, the framework addressed the following healthcare knowledge modelling issues:

- 1) *Modelling of healthcare knowledge, especially in terms of CPGs and CPs, to develop an ontology-based healthcare knowledge model for handling comorbid diseases.*
- 2) *Computerization of CPs to offer point-of-care decision support*
- 3) *Alignment of ontologically-modelled disease-specific CPs to handle comorbid diseases; and*
- 4) *The provision of computerized decision support for GPs, based on modelled CPGs and CPs, to assist them in handling chronic and comorbid diseases [12]*

Also an elaborate OWL-CP (Web Ontology Language – Clinical Pathway) ontology for comorbid CHF and AF was developed that can be executed to support the diagnosis and management of comorbid CHF and AF in a general practice setting. Hence, the COMET framework was implemented to handle three patient care scenarios:

- 1) Patient that has CHF
- 2) Patient that has AF and
- 3) Patient with comorbidity of both AF and CHF

COMET is accessible by web and is designed for GPs. It has been evaluated, both by simulated cases and by health professionals (GP and specialist), for its ability to handle single disease and comorbid care scenarios based on patient data and related constraints. The output at every phase was compared with the expected output as per single disease or comorbid management. Their results showed that the resultant sequence of plans and their outcomes are comparable to the CP knowledge [12].

B. Web-Based PEDSS Framework Using a Knowledge-Based Approach to Estimate Clinical Outcomes: Neonatal Mortality and Preterm Birth in Twins Pregnancies

This system whose architecture is shown in Fig 2 and a description of its components in Table 1, adopts an improved classification method that was derived to improve the prediction of two distinct medical problems non-invasively:

- 1) Neonatal mortality with information available at 10 minutes from birth and
- 2) Preterm birth in twin pregnancies before 22 weeks.

The framework was developed by assessing various data mining methods with the aim of improving the classification of neonatal mortality and preterm birth in twin pregnancies. The major analyzed models were DTs and hybrid ANN to see which produced better outcomes. Positive findings related to the DT mechanism showed that same method can be applied to many other multi-factorial medical problems to improve its classification. This is given the fact that most published risk estimation models attempt to meet clinically acceptable sensitivity and specificity, in which case successful identification of positive cases have been met with much difficulty. Also, with unnecessary variables adding noise and complexity to the problem, it reduces the likelihood of identifying positive cases. A major aim of this framework was to incorporate the advantages of DTs to create a system able to predict the two perinatal problems already mentioned at an earlier stage while maintaining high sensitivity, specificity, Positive Predictive Value (PPV) and Negative Predictive Value (NPV)[13].

Thus, the new approach provides several improvements to better predict medical problems as outlined below:

- 1) Pre-processed datasets run against C5.0 algorithm produced DTs superior to the DT-ANN hybrid method.
- 2) Two novel prediction models using DTs and hybrid ANN were evaluated. The DT prediction model had the highest performance outcome for predicting neonatal mortality (sensitivity = 62.24%, specificity = 99.95%, PPV = 72.34%, NPV = 99.92%) using information available within 10 minutes from birth, and preterm birth in twin pregnancies (sensitivity = 80.00%, specificity = 91.55%, PPV = 67.35%, NPV = 95.79%) before 22 weeks gestation. This was achieved using 5-by-2 cross validation. This indicates that the system is not

over trained and provides good generalization. [NPV = Negative Predictive Value PPV = Positive Predictive Value].

3) Creation of a neonatal mortality prediction system for newborn to be assessed with data available from the first 10 minutes from birth non-invasively with excellent discrimination, exceeding the results of current standard predictions.

4) Creation of a preterm birth prediction system for a high risk population (for women pregnant with twins) non-invasively before 22 weeks gestation with excellent discrimination, exceeding the results of current standard predictions

5) The previous neonatal prediction method only focused on newborns after admission to NICU. This is the first attempt at predicting neonatal mortality in a heterogeneous population with data available at 10 minutes from birth.

6) Several improvements were made compared to past models: For the neonatal mortality case, the prediction of neonatal mortality non-invasively was reduced to data available at 10 minutes from birth using only 13 attributes, whereas the previous models required up to 12 hours from birth using 3 variables derived from invasive methods.

7) A conceptual framework for a secure web-based Perinatal Decision Support System (PEDSS) was consequently developed (with components as seen in Table 1) to provide audience targeted information and risk prediction of neonatal mortality and preterm birth in twin pregnancies [13].

IV. THE PROPOSED FRAMEWORK [COMORBIDITY ONTOLOGICAL FRAMEWORK FOR INTELLIGENT PREDICTION (COFIP)]

Having analyzed the PEDSS and COMET frameworks, a description of COFIP is given in this section. The framework diagram is as represented in Fig 3.

A. Knowledge Representation/Preprocessing

This section is comprised of the knowledge identification, acquisition, synthesis, formalization and alignment layers. They are discussed below;

a) Knowledge Identification/Acquisition Layer: The cost and performance of an application depends directly on the quality of the knowledge acquired [14]. The purpose of this phase is to identify valid sources of relevant patient management knowledge as it pertains to two chronic disease conditions namely cataract and glaucoma. This is derivable from existing CPGs – a documentation that is predicated on evidence-based research and is thus a repository of knowledge aimed at providing guidance for decisions and criteria regarding diagnosis, management and treatment of specific disease conditions. The knowledge sources considered, not only entailed evidence-based recommendations but also specific tasks and procedures and their scheduling information. A number of knowledge sources are identified during this phase, including CPGs, institution specific drug management protocols, journal publications, and most

importantly domain experts (in this case a consultant ophthalmologist and an optometrist at the Babcock University Teaching Hospital, Ilishan-Remo, Ogun State, Nigeria).

b) Knowledge Synthesis Layer: The knowledge synthesis phase involves the acquisition of the clinically useful task-specific heuristics from the identified knowledge sources (such as the CPGs) through the processes of selection, interpretation and augmentation of the guideline statements, tacit knowledge and logic. Where necessary, the heuristics will be further decomposed into atomic tasks and then organized in such a way as to develop two (cataract and glaucoma) CPs packages containing clear and relevant evidence-based diagnostic and therapeutic plans for patient care management, especially by GPs. Knowledge synthesis is a process in which one builds concepts in cooperation with others. It provides the opportunity for one's hypothesis or assumption to be tested. Social intercourse is one of the most powerful media for verifying one's own ideas. As such, participants in the dialogue can engage in the mutual co-development of ideas [15].

c) Knowledge Formalization Layer: Written sources such as textbooks and technical treatises are often not precise enough for transformation into descriptive logic: there may be competing accounts of the same phenomena, overlapping taxonomies and standards, or outright contradictions [16]. Hence in the knowledge formalization layer, the fused knowledge from the previous layer is modelled and formalized in terms of a dedicated CP ontology to be developed using the Web Ontology Language (OWL). Ontology is the standard knowledge representation mechanism for the Semantic Web framework. The choice of OWL is predicated on the fact that it offers declarative semantics that allows us to associate natural language descriptions with formal statements, thereby allowing human and machine readability of knowledge and subsequent execution of the knowledge. In this phase, the comorbid clinical processes in the CP ontology is hierarchically decomposed into component tasks that are based on the available evidence for specific single disease and comorbid scheduling constraints. This will ensure the conceptualization of the domain into an unambiguous model, thereby determining all implicit constraints on the relationships between the domain concepts, particularly to assist the alignment of concepts in handling comorbidities.

d) Knowledge Alignment Layer: The knowledge alignment layer involves ontology alignment—i.e. alignment of discrete and ontologically defined care plans in response to single disease or comorbid preconditions. The alignment of comorbid CPs is achieved at knowledge modelling level by developing a unified ontological model that encompasses the combined knowledge of aligned CPs. Also, knowledge alignment is tackled at the ontology level, implying that all ontological constraints about knowledge consistency will be observed in the ontologically-modelled Cataract-Glaucoma CP that entails a network of specific classes and the relationship between them. This is indeed a complex activity given the fact

that the alignment of comorbid plans needed to take into account the medical correctness and clinical pragmatics of the resultant Cataract-Glaucoma CP.

B. Knowledge-Based Warehouse

The knowledge-based warehouse is the repository for all the relevant domain knowledge gotten from the knowledge stratified into precondition sets A, B and C by the knowledge representation/preprocessing section of the framework. This knowledge-base is also updated through the workings of the results classification/prediction algorithms section where the rules and prediction modules are housed. It is structured such that knowledge is represented in such a way as to promote an efficient system that gives results that tend towards what is obtainable in reality. Therefore explicit/domain knowledge is synergized with tacit knowledge leading to an optimized outcome that helps to inform a patient-specific care plan/recommendation.

The optimized outcome from the prediction algorithms prior to generating the CPs is also used to update the knowledge-base. When this happens, a similar problem can be taken care of without subjecting it to the ANN and DT algorithms, an instance of learning having taken place. This implies a smooth transition from the knowledge-based warehouse to the CP formulation. However, where the patient already exhibits a full-blown comorbid condition it is not subjected to the results classification/prediction algorithms since the essence of that section is to predict the percentage tendency for the emergence of comorbidity where one of both diseases has developed. The section is composed of sub-modules, namely:

- Pre-Condition Set A – Cataract Disease: this houses the knowledge set akin to the cataract condition and sets up a need to predict whether the patient is likely to develop glaucoma vis-à-vis the comorbid condition.
- Pre-Condition Set B – Glaucoma Disease: this contains the knowledge set akin to the glaucoma condition and sets up a need to predict whether the patient is likely to develop cataract vis-à-vis the comorbid condition.
- Pre-Condition Set C – Cataract-Glaucoma Comorbid Condition: this houses the knowledge set akin to the cataract-glaucoma comorbid condition that sets up a need to initiate patient treatment plan and management.

C. Results Classification/Prediction Algorithm

This section entails the mechanisms that make the framework adaptive. They include DTs and ANN which are both techniques of AI.

a) Artificial Neural Network: ANNs are powerful non-linear mapping structures and are especially useful for modelling relationships which are unknown. ANNs function similar to the human brain and can solve problems involving data that is complex, non-linear, imprecise and/or noisy [17]. The human brain is a collection of more than 10 billion interconnected neurons that is able to receive process and transmit data. The human brain also consists of a highly parallel computing structure to support computationally

demanding perceptual acts and control activities [18]. ANNs were developed as generalized mathematical models to represent the biological nervous system [18]. The ANN is trained to detect a pattern between the inputted data and the related output value from a dataset. After training the set, the ANN can be used to predict the result of a newly inputted data [17]. There are various types of ANNs including feed-forward, recurrent neural network and probabilistic network. The ANN structure used in this thesis is referred to as feedback oriented propagation multi-layer perceptron.

b) Decision Trees: Decision trees are favoured in the data mining community due to its highly interpretable structure, allowing business end users and analysts to understand the models, whereas neural networks are difficult to understand and interpret [19]. A decision tree consists of a root node, branch nodes and leaf nodes. The tree starts with a root node, is further split into branch nodes (each of the nodes represent a choice of various alternatives), and terminates with a leaf node which are un-split nodes (represents a decision) [20]. Classification of decision trees are conducted in two phases, including the tree building (top down) and tree pruning (bottom-up). Tree building is computationally intensive, and requires the tree to be recursively partitioned until all data items belong to the same class. Tree pruning is conducted to improve the prediction and classification of the algorithm and to minimize the effects of over-fitting, which may lead to misclassification errors [21]. There are a number of decision tree algorithms that exist including Classification and Regression Trees (CART), Iterative Dichotomiser 3 (ID3), C4.5 and C5.0. This thesis work uses C5.0 based decision tree algorithm which is an improvement over C4.5, which itself is an improvement over the earlier ID3 method.

c) Result Comparison and Optimization: The result comparison and optimization module is responsible for synergizing the outputs generated by the different classification/prediction algorithms i.e. the ANN and DT mechanisms so as to settle for an optimized output. The entire workings of the different modules in this unit are all geared towards finding a healthy association between the domain knowledge and tacit knowledge in order to make the overall system operation with some measure of expertise akin to a human expert. Tacit knowledge has to do with unwritten, unspoken, and hidden vast storehouse of knowledge held by an individual, based on the persons' emotions, experiences, insights, intuition, observations and internalized information. Tacit knowledge is integral to the entirety of a person's consciousness, is acquired largely through association with other people, and requires joint or shared activities to be imparted from one person to another. Like the submerged part of an iceberg it constitutes the bulk of what one knows, and forms the underlying framework that makes explicit knowledge possible. When the inputs from the knowledge-base is subjected to analysis by the ANN and DT mechanisms their outputs are compared and the outcome are optimized for the generation of rules with which predictions get carried out.

D. Intelligent Clinical Pathway Generator System

The intelligent clinical generator system unit is composed of rules that are a function of the results comparison/optimization module which is used to make predictions that gets to inform the generation of patient-specific care plans. It also contains a feedback mechanism.

a) Rules and Prediction: The rules module receives an optimized output upon which rules are generated and subjected to further coordinated analysis that becomes the yardstick for the prediction of the tendency for comorbidity. Once this is established the predicted values become the benchmark upon which generation of patient-specific care plans and recommendations take place.

b) Patient-Specific Care Plans/Recommendations: This system adopts patient-specific CPs/recommendations as against case-based CPs/recommendations because the modern patient wants to be treated as an individual person and not just as a statistic [22]. Patients want to know their own risk, not just a parameter regarding a class of people similar to them. This feature is highly enhanced through the deployment of ANNs which are able to reproduce the dynamical interaction of multiple factors simultaneously, allowing the study of complexity which is very important for a researcher interested in in-depth knowledge of a specific disease or to better understand the possible implications relative to strange associations among variables. This has to do with what is called "intelligent data mining". But on the other hand ANNs can also help medical doctors in making decisions under extreme uncertainty and to draw conclusions on individual basis and not as average trends.

c) Feedback Mechanism: In view of the fact that this framework serves as a platform for implementation and translation into a real system, a feedback mechanism is sacrosanct and must be included. More so, a typical developer may find it difficult to adapt a framework that is without a feedback system during implementation [23]. Hence, the proposed framework holds an extension to one. The schematic diagram in Fig 4 shows a typical feedback mechanism which is a sub-set of the patient-specific care plan generator system. The feedback mechanism is composed of functionalities that:

- 1) *Describe diagnosis:*
 - Inform patient about disease (you're diagnosed as suffering from...)
 - present supporting evidence for disease
- 2) *Alleviate patient fears:*
 - describe improvements
- 3) *Describe future prospects.*
- 4) *Describe disease triggers:*
 - present background information
 - List triggers mentioned by the patient.
 - List triggers mentioned by doctor.
 - Suggests methods of avoidance.

- 5) Describe drug prescription.
 - 6) Describe need for long term effort.
- d) Discussion of the Underlying Processes:

1) The patient information is fed into the system (See Fig 5). The patient information includes age, race, gender etc. medical history such as surgery, ocular disease, eye surgery, trauma etc. Questions about the patient's sight are also included. The patient's information is then subjected to the knowledge representation and preprocessing.

2) Knowledge representation/preprocessing has four layers: identification/acquisition, synthesis, formalization and alignment layers.

- Knowledge identification/acquisition layer identifies valid sources of relevant patient management knowledge as it pertains to two chronic disease conditions namely cataract and glaucoma. The knowledge is derived from existing CPGs.
- Knowledge synthesis layer involves the acquisition of the clinically useful task-specific heuristics from CPGs through the processes of selection, interpretation and augmentation of the guideline statements, tacit knowledge and logic.
- In the knowledge formalization layer (semantic layer) the synthesized knowledge will be modelled and formalized in terms of a dedicated CP ontology to be developed using the Web Ontology Language (OWL).
- Knowledge Alignment layer involves ontology alignment of discrete and ontologically defined care plans in response to single disease or comorbid preconditions.

3) The preprocessed results from the knowledge representation unit are passed into knowledge base warehouse for dynamic storage and updating. The knowledge-based warehouse is updated through the rules and prediction unit. Here, the preconditions are segmented and classified into one of cataract, glaucoma or comorbid precondition sets. If it is the comorbid condition the patient-specific care plan is generated straight away. Else, the disease identified (cataract or glaucoma) is subjected to the classification/prediction algorithm. The two algorithm used are ANN and DT. The two are used to complement each other's strengths and weaknesses as the case may be. The result from ANN and DT are compared for optimization and the best values are chosen and used to generate the rules.

4) The rules generated are made from the result of the optimization and they are used to update the precondition sets A and B (cataract and glaucoma) contents in the knowledge base warehouse. Once the condition is ascertained, the patient-specific care plan is generated.

V. CONCLUSIONS AND IMPLICATIONS

This research paper entails the design of an adaptive framework (COFIP) that is diagnostic, preventive and predictive. This is because COFIP was designed to predict

with satisfactory accuracy, the tendency of a patient with either cataract or glaucoma to degenerate into a state of comorbidity.

Furthermore, because this framework is generic in outlook, it can be adapted for other chronic diseases of interest within the medical informatics research community.

A. Recommendations and Future Work

Having built the predictive framework, full implementation will be carried out as follows:

1) Build useful and valid knowledge-based CPs (guidelines/framework) through the acquisition of relevant data sets especially those related to ophthalmological diseases (Catt-Gla).

2) Model the selected ophthalmological diseases such that the diagnostic and treatment concepts show inter-relationships in formal language in ways that curb every form of possible ambiguity. This is carried out in such a way as to ensure that the encoded knowledge and the underlying decision logic can be executed through computerized clinical decision support systems to provide patient-specific CPG-based recommendations.

3) Systematically align the designed model to handle ophthalmologic diseases without compromising the integrity of medical knowledge, care coordination and safety of patients.

4) Fully operationalize the designed framework that is diagnostic, predictive and preventive. Thus, the system when implemented will predict the chances of a patient with either of Catt or Gla to degenerate into a state comorbidity.

B. Summary

The coexistence of cataract and glaucoma accounts for alarming levels of visual impairment in our society vis-à-vis impairment of quality of life and hence increases the burden of illness, care plan, patient management and related concerns. A viable method to reduce the menace of blindness and other related conditions is to engage GPs in the management of cataract and glaucoma as well as its co-morbidities, because GPs are the first point of care for most patients. There are challenges in the diagnosis of glaucoma for instance given that many of its clinical features are time-specific and sometimes not obvious though present.

To make matters worse, concurrent presence of cataract complicates the management of either condition as the choice of treatment depends on individual factors of each disease as it manifests in the patient. EBCAs such as CPGs and CPs have the propensity to narrow this gap [8]. They can by assisting GPs to undertake complex diagnostic and management scenarios resulting from the comorbidity of cataract and glaucoma. In view of the foregoing we have designed an improved framework that enhances the classification/prediction of comorbidities that bring with it adaptability that helps to guide patients with one chronic disease to manage their condition so as to prevent the degeneration of their condition to full-blown comorbidities. COFIP was designed to improve on the existing COMET framework that is not adaptive.

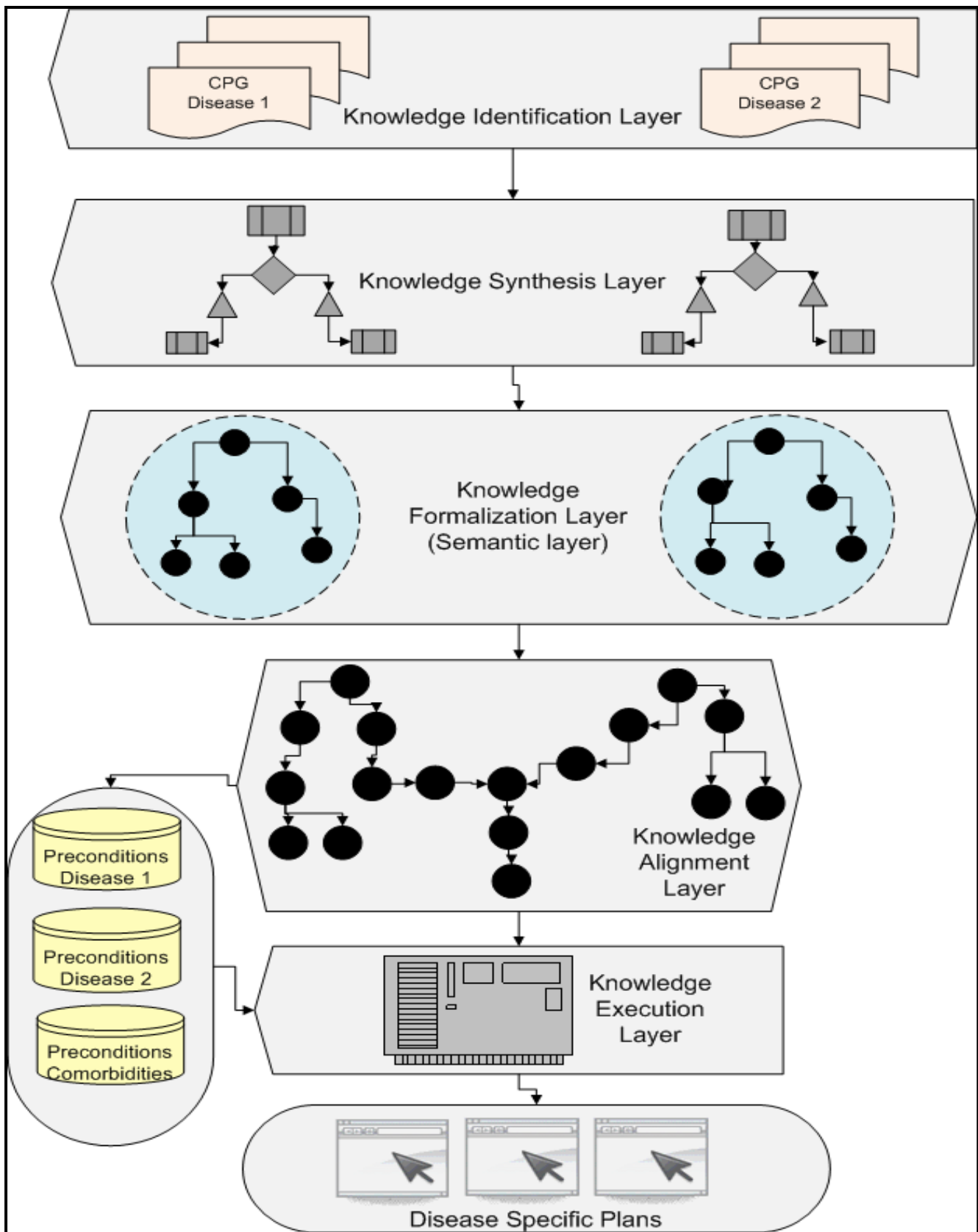


Fig. 1. COMET Framework [12]

TABLE I. COMPONENTS FOR A WEB-BASED PEDSS [13]

	Components	Description
1	Authentication Server	Authenticates users into the system
2	Content Management System (CMS)	The heart of the system used to display, search, and process the data, based upon the user request
3	Workflow Engine	Required to automate alerts, warning and actions
4	External Data Source	A repository of the patient, or user information
5	Directories	A database of user information, etc.
6	Other Web Servers	Other servers required to operate the PEDSS
7	ASP.Net, XML, HTML	The interface presented to the user

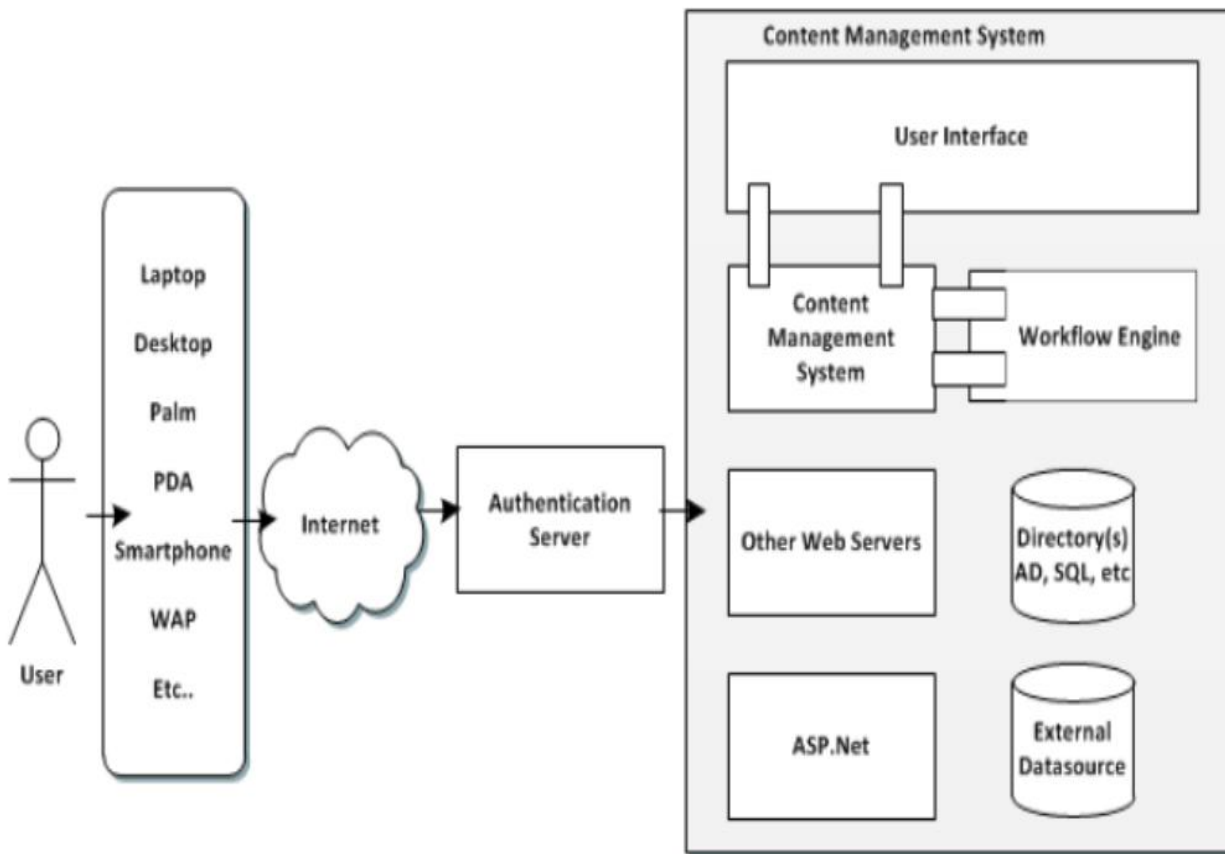


Fig. 2. System Architecture of Web-based PEDSS [13]

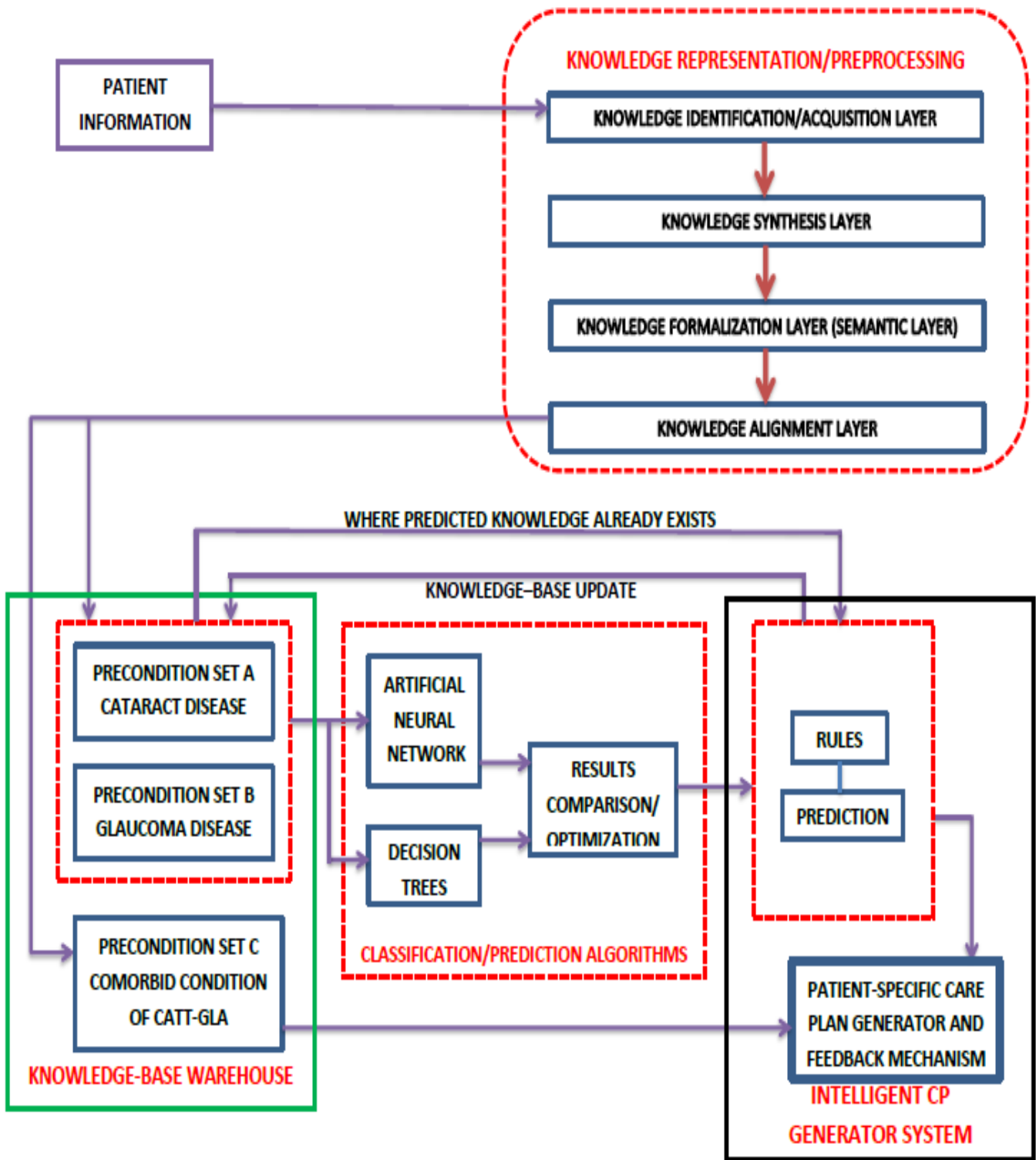


Fig. 3. The Proposed (COFIP) Framework

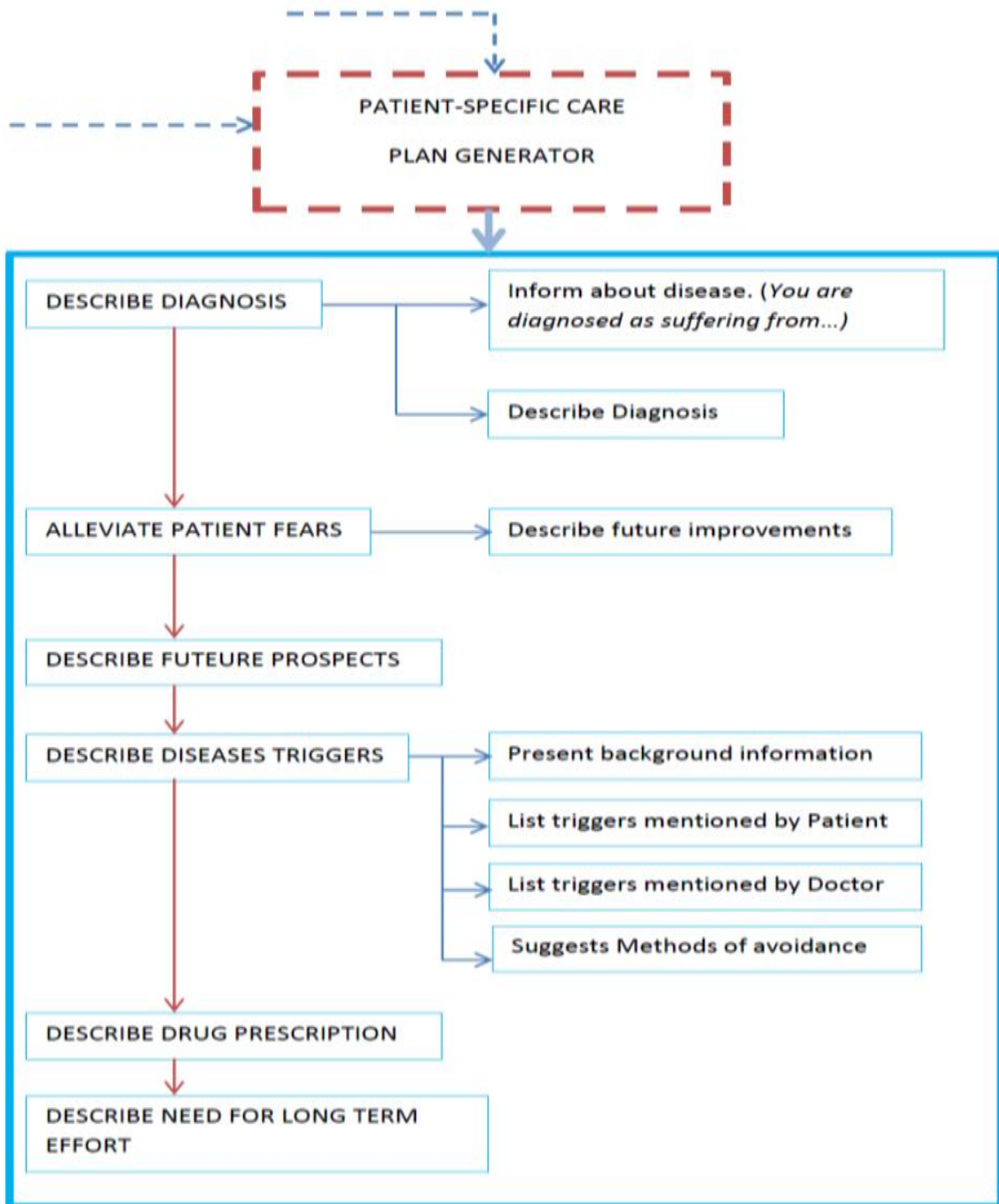


Fig. 4. Feedback Mechanism

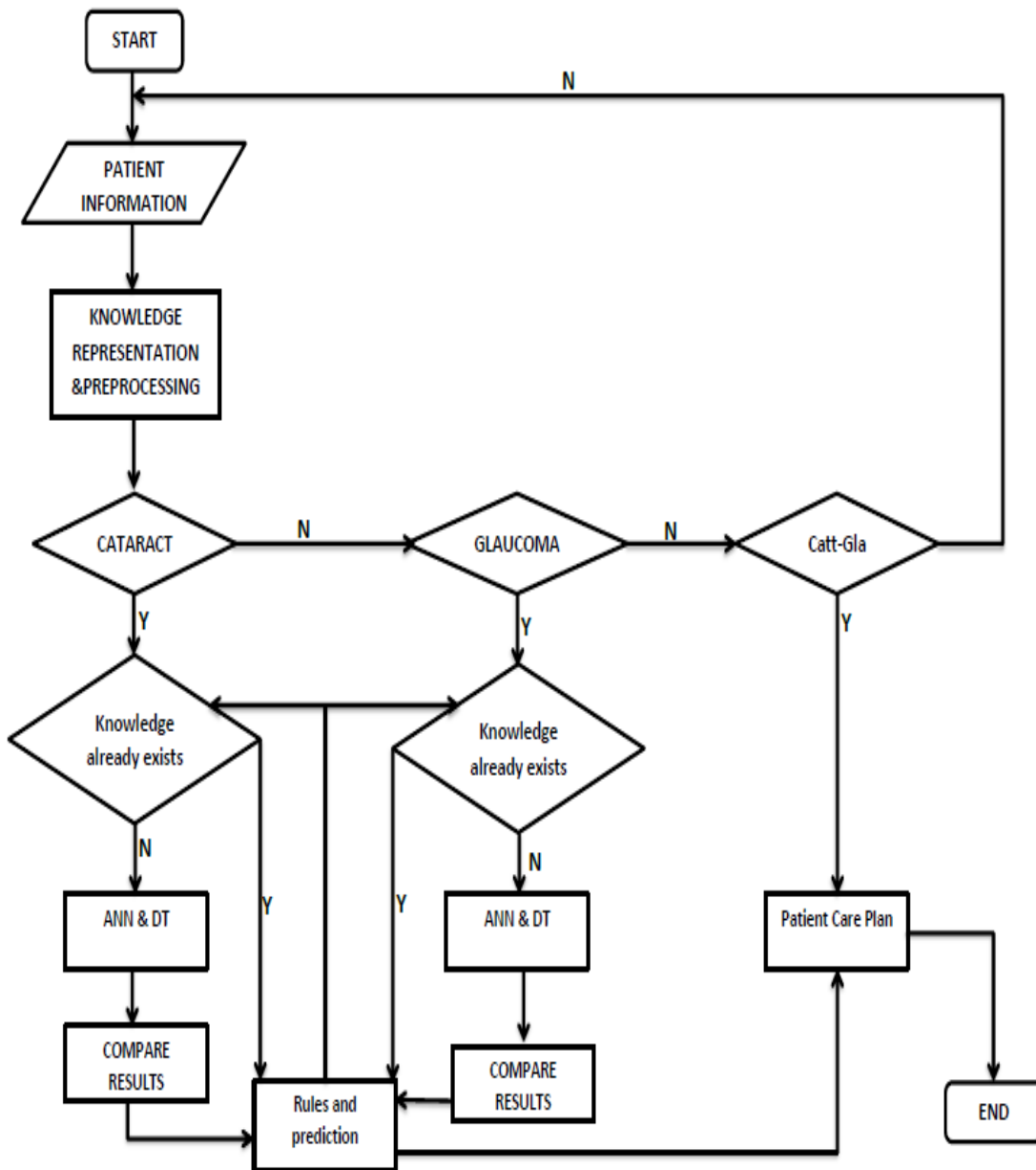


Fig. 5. Flowchart of COFIP

REFERENCES

[1] Angeli C. (2010). *Diagnostic Expert Systems: From Expert's Knowledge to Real-Time Systems*. Advanced Knowledge Based Systems: Model, Applications & Research. (Eds. Sajja&Akerkar), Vol. 1, pp 50 – 73, 2010

[2] Jones T. M. (2008). *Artificial Intelligence, a System Approach*.

[3] Fonseca, F. (2007). *The Double Role of Ontologies in Information Science Research*. Journal of the American Society for Information Science and Technology, 58(6), pp. 786-793.

[4] Chen J.Q & Lee S.M. (2002). *An Exploratory Cognitive DSS for Strategic Decision Making*. 2002 Elsevier Science B.V. All rights reserved. doi:10.1016/S0167-9236(02)00139-2

[5] Feinstein, A. R. (1970). *The Pre-therapeutic Classification of Comorbidity in Chronic Disease*. Journal of Chronic Diseases, 23, pp456-7.

[6] O'hEineachain R. (2005). *Glaucoma and Cataract surgery. Comorbidities: Patients With Cataracts Often Have Other Ocular Conditions Affecting Outcomes*.

[7] National Center of Health Statistics (2007). *N.C.H.S. Definitions*. <http://www.cdc.gov/nchs/datawh/nchsdefs/comorbidities.htm>. Retrieved November 12, 2008

[8] Brush, J.E, Radford, M.J. &Krumholz, H.M. (2005). *Integrating Clinical Practice Guidelines into Routine of Everyday Practice*. Crit Pathway Cardio, 4:161-167

[9] Woolf, S.H., Grol, R., Hutchinson, A., Eccles, M., &Grimshaw, J., (1999). *Clinical Guidelines: Potential Benefits, Limitations, and Harms of Clinical Guidelines*. BMJ, 318: 527-530

[10] Kitchiner, D.J. &Bundred, P.E. (1996). *Integrated Care Pathways*. Arch DisChild, 75: 166-168

[11] Pearson, S. D., Goulart-Fisher, D., & Lee, T.H. (1995). *Critical Pathways as a Strategy for Improving Care: Problems and Potential*. Annals of Internal Medicine, 123(12), 941-948

- [12] Abidi S.R. (2010). *A Knowledge Management Framework to Develop, Model, Align and Operationalize Clinical Pathways to Provide Decision Support for Comorbid Diseases*. <http://dalspace.library.dal.ca/handle/10222/13009>. Retrieved February 5, 2013.
- [13] Gunaratnam M. (2013). *A Web-Based Perinatal Decision Support System Framework Using a Knowledge-Based-Approach to Estimate Clinical Outcomes: Neonatal Mortality and Preterm Birth in Twins Pregnancies*. http://www.sce.carleton.ca/faculty/schwartz/RCTI/Seminar%20Day/Autonomous%20vehicles/theses/Thesis%20-%20Marry%20Gunaratnam%20_%20100550026.docx. Retrieved March 28, 2013.
- [14] Rhem A.J & Associates. (2002). *A framework for Knowledge Acquisition*. <http://www.ajrhem.com/framework.pdf>. Retrieved March 11, 2013.
- [15] Markova, I. & Foppa, K. (1990). *The Dynamic of Dialogue*. New York: Harvester Wheatsheaf.
- [16] Bowker, G. C., (2000). "Biodiversity Datadiversity." *Social Studies of Science* 30(5): 643-683.
- [17] Jha, G.K. (2011). *Artificial Neural Networks*. http://www.iasri.res.in/ebook/EB_SMAR/e-book_pdf%20files/Manual%20IV/3-ANN.pdf
- [18] Lisboa, P.J.G. (2011). *A Review of Evident of Health Benefit from Artificial Neural Networks in Medical Intervention*. *Neural Networks*, 2, 11-39.
- [19] Apte, C., et al. (2002). *Data Intensive Analytics for Predictive Modelling*. http://www.research.ibm.com/dar/papers/pdf/dar_overview_ibmjrd.pdf. Retrieved March 12, 2013.
- [20] Peng, W., Chen J. & Zhou H. (2006). *An Implementation of ID3 – Decision Tree Learning Algorithm*. University of New South Wales. Last accessed: December 17, 2011 from <http://web.arch.usyd.edu.au/~wpeng/DecisionTree2.pdf>.
- [21] Anyanwu, M.N., and Shiva, A.G., (2009). *Comparative Analysis of Serial Decision Tree Classification Algorithms*. *International Journal of Computer Science and Security*. 3, 230-237
- [22] Grossi E. (2011). *Artificial Neural Networks and Predictive Medicine: a Revolutionary Paradigm Shift*. http://cdn.intechopen.com/pdfs/14887/InTech-Artificial_neural_networks_and_predictive_medicine_a_revolutionary_paradigm_shift.pdf. Retrieved April 7, 2013.
- [23] Buchanan B.G, Carenini G, Mittal V.O. & Moore J.D. (1998). *Designing Computer-based frameworks that facilitate doctor-patient collaboration*. Elsevier, *Journal of Artificial Intelligence in Medicine*. <http://goo.gl/uGgxV>. Retrieved March 5, 2013.

ABOUT THE AUTHORS

Ernest E. Onuirih holds a Master of Science degree in Computer Information Systems from Babcock University, Ilishan-Remo, Ogun State, Nigeria. His interests include Medical informatics, Artificial intelligence, Bioinformatics and Networking.

Awodele Oludele is a professor of Computer Science at Babcock University, Ilishan-Remo, Ogun State, Nigeria. His interests include Artificial Intelligence, Modelling and Simulation and Green Computing

Sunday A. Idowu is a professor of Software Engineering and Information Systems at Babcock University, Ilishan-Remo, Ogun State, Nigeria. His interests include Software Engineering, Programming, Content Networking, and Web development.

Scale-Based Local Feature Selection for Scene Text Recognition

Boyu Zhang, Jia Feng Liu, Xiang Long Tang
School of Computer Science and Technology,
Harbin Institute of Technology,
Harbin 150001, China

Abstract—Scene text recognition has drawn increasing concerns from the OCR community in recent years. Among numerous methods that have been proposed, local feature based methods represented by bag-of-features (BoFs) show notable robustness and efficiency. However, as the existing detectors are based on assumptions about local saliency, a vast number of non-informative local features would be detected in the feature detection stage. In this paper, we propose to remove non-informative local features by integrating feature scales with stroke width information. Experiments taken both on synthetic data and real scene data show that the proposed feature selection method could effectively filter non-informative features and improve the recognition accuracy.

Keywords—Scene Text Recognition; Local Feature; Stroke Width

I. INTRODUCTION

In recent years, scene text recognition (STR) [1] technologies have got increasing concerns from OCR community and other related fields. Compared with surrounding text, scene text is more connected to image contents in most cases. Thus the rich semantic information contained in scene text often plays vital roles in a host of computer vision applications, including impaired people assist, visual land-mark robot navigation and intelligent traffic system.

Even numerous potential applications exist, the STR is still challenging due to the following disadvantages: (1) The scales of scene text, even in same sentences, vary a lot; (2) The shapes and styles may be different since scene text are specially designed to fit different requirements; (3) Scene images always contain illumination changes, viewpoints variations and other disadvantages such as a non-flatness surface; (4) In most cases, no context information is provided.

During the past decades, a number of methods are proposed in response to these disadvantages. The existing methods in STR area could be divided into two categories according to their basic ideas. One of which is to achieve accurate STR by developing traditional OCR methods. Most approaches under this idea contain three procedures, which are, text detection, segmentation and character recognition. For instance, Chen and Yuille [2] train strong classifier which contains multiple features by integrating weak classifiers with AdaBoost to extract text regions, then text are recognized by employing commercial OCR software. Coates et al. [3] apply scalable learning algorithm to feature extraction, text detector and classifier to produce high accurate STR system. Kai et al.

[4] designed an end-to-end system for scene text recognition, in which Random Fern [5] is utilized as raw character detector as well as classifier. Moreover, they proposed to improve the accuracy of STR by introducing pre-defined vocabulary.

Another idea is to treat scene text as objects. Thus researchers can transplant object recognition methods that are proposed mostly against image degradations and uncontrolled environments into STR area. For example, De Campos et al. [6] build up a STR framework by following classic BoF methods in which sample images are described by frequency histogram of local features. They also compare the effectiveness of different local descriptors by taking experiments one representative benchmark. Zheng et al. [7] recognize scene characters by matching detected SIFT [8] features between input samples and pre-build template images. Different from BoFs method that totally omits position information, they consider the relative position of local features by using MPLSH [9]. Diem and Sablatnig [10] build a historical document analysis system based on local descriptors and achieve a state-of-art accuracy for ancient character recognition.

Among these methods, the ones based on local features [6], [7], [10] show notable robustness and effectiveness, especially when in small sample size situations and situations containing image degradations [11]. They are more robust because they represent sample images using sets of local features and omitting other highly variable factors. It is obvious that their accuracy largely depends on the effectiveness of detected local features. However, even most local feature detectors assume that salient image patches are informative, the meanings of effective are different in different applications. Specific to our problem, not all detected saliency image patches reflect local structures of characters. Thus, for improving the accuracy, criteria are needed to filter features which are not related to the text.

In this paper, we focus on local feature based STR and propose a novel criterion which integrate stroke width information with local feature scales to remove non-informative local features and achieve higher accuracy. Our idea is based on the fact that text is constituted by strokes with specific width. Thus there should be an appropriate proportion between local feature scale and the corresponding stroke width if these features reflect local text structures such as corner and cross. Experiments taken on both natural and synthetic text images show that the proposed approach could effectively improve the accuracy of local feature based STR.

II. RELATED WORKS

Many techniques are developed for filtering redundancies and noises from original features set. In this paper, we make the specific consideration about methods based on codebook model. A classical codebook method includes local feature detection, codebook generation, quantization, and finally classification. Most efforts for feature selection are taken on codebook generation stage and code-word selection stages. In this section, we briefly introduce typical existing methods according to their categories and discuss differences between these methods and proposed method in the end.

A. Compact Codebook Generation

In codebook generation stage, the algorithm seeks for a group of code words (also referred as 'codebook'), which could describe the feature space effectively. A vast number of methods are proposed to generate effective codebook. For instance, Tuytelaars and Schmid [12] extract high-dimensional descriptors for sample images by partitioning feature space using lattices with regular sizes and then combine similar dimensions to make the descriptors more compact. The most widely applied idea is to get codebook utilizing unsurprised cluster algorithms such as K-means [13], which get the most descriptive k centers by minimizing the variance between k centers and the training data. Different from k -means that is dense sensitive, Jurie and Triggs [14] proposed a radius-based clustering which clusters all features within a fixed radius of similarity radius to one cluster.

B. Code-word Selection

Besides generating a compact codebook, a host of algorithms are proposed for picking the most effective subset from the original codebook. Code-word selection is equal to feature selection problem since sample images are represented by frequency histograms of code-words and each bin corresponding to a feature dimension. Distinguishing by whether class labels are given existing methods could be divided into supervised and unsupervised ones.

Supervised methods analyze the relationship between the class labels and code words and then pick more discriminate subset based on pre-defined criteria. Literature [14] gives a performance evaluation for three typical methods including MI [15], OR [16] and Linear SVM weights [17] on representative datasets. Moosmann et al. [18] proposed to build supervised indexing trees using an ERC-Forest that considers semantic labels as stopping tests. The work in [19] aims to find the Descriptive Visual Words (DVWs) and Descriptive Visual Phrases (DVPs) for each image category.

For unsupervised situations, Zhang et al. [20] proposed to pick out the most discriminative code words which lead to minimal fitting errors between data matrix and indicator matrix. Maximum variance selects features with the largest variances and unsupervised feature selection for PCA selects a subset of features that can best reconstruct other features. Laplacian score [21] selects features that preserve the local geometrical structure best. Q - α [22] measures the cluster coherence by analyzing the spectral properties of the affinity matrix.

C. Proposed Method

Different from the above methods, the proposed method in this paper filters non-informative features by performing a pre-selection based on analyzing both feature scale and stroke width information. Its advantage is that the algorithm effects before codebook generation stage and thus could avoid errors that occur in the following process. This means the proposed methods could be more effective when facing small sample size problems, which are common in STR and historical document analysis.

III. SCALE-BASED LOCAL FEATURE SELECTION

The fundamental assumption of designing most local feature detectors is that salient image patches are informative. In fact, the concepts of 'informative' are different in different situations. Specifically, in STR process, it is not promised each salient patches indeed reflects character structure. Thus criteria are needed to remove features that are not effective.

According to whether they are helpful for distinguishing different characters, we divide detected local features into informative and non-informative. Features belong to the first category always localize in character bounding-boxes and they are salient since they contain character structures such as corners and stroke crosses. In contrast, most features that belong to the second category are generated by cluttered background and noises, thus do not provide information for STR. It is worthwhile to emphasize that large local features that cover the majority of a character should be categorized into the second type since these features are not robust enough when numerous variations are included.

However, it is difficult to remove non-informative local features automatically as it is difficult to give a formally definition for non-informative features. The target can be achieved by training a binary classifier that could distinguish on-informative features from informative ones, however, a large number of training samples are needed to train such a classifier and the existence of varies fonts makes sample collecting rather difficult. Moreover, labeling all features manually is labor expensive and hardly objective. Another idea is to optimize learned codebook according to class label as we discussed in section II, which is under sophisticated mathematical model. These methods that select features by analyzing the relationship between code words and class labels also need large training dataset.

In this paper, we propose a novel local feature selection criterion that selects effective local features based on the ratio between character stroke width and local feature scale.

A. Feature Scale and Stroke Width

Our idea is based on the observation that it is impossible to write small character with wide strokes and large characters with thin strokes. Thus the ratio between character size s_c and stroke width w in the text area should keep within a reasonable range to ensure the character is recognizable. At the same time, for each detected local feature which reflects a local structure on character, its scale s_f should also be indirect

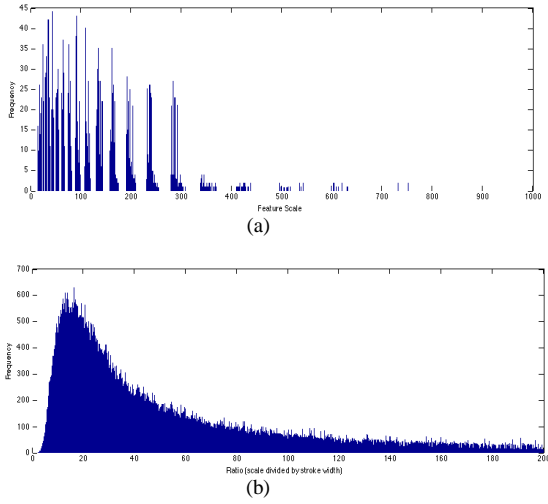


Fig. 1. Diagram (a) shows the frequency histogram of feature scales local features extracted form 'char 74k' dataset. (b) is the frequency histogram of ratio parameters that is calculate by feature scales divided by corresponding stroke width.

Proportion to character scale s_c abided by commonsense. This means that for a reasonable character, the scale of a representative local feature should have a stable ratio r with stroke width w . Based on this idea, we can filter non-effective features by checking whether the ratio r is in an interval $[r_{\min}, r_{\max}]$.

The reason we do not directly apply character size for feature selection is that local structures are directly instituted by strokes and thus the ratio between stroke width and feature scale is more stable than the ratio between character size and feature scale. Moreover, stroke width is more accurate than character size in two reasons. Firstly, the segmentation in scene images is difficult which would lead to inaccurate character size. Secondly, characters in the same size have different stroke width because of the existence of multi-font.

To prove this, we count the frequency histograms of the detected local features according to their feature scales and ratio parameters respectively. The definition of stroke width and the calculation of ratio parameters are described in detail in section IV. Fig 1(a) shows the frequency of local feature scale and Fig 1(b) gives the frequency of the ratio between feature scale and corresponding stroke width. We find that the ratio parameter depends on a uniform long-tail distribution which certify that a relationship exists between local feature scales and stroke width.

B. Scale-based Local Feature Selection

Typical local feature detectors such as SIFT and Multi-Scale Harris contain three stages. In the first stage, for each

pixel $I(i, j)$ in an image I , its local saliency H corresponding to scale s is evaluated by using measurement function F . By noting the neighborhood of point $I(i, j)$ as $r(i, j)$, we have:

$$H(i, j, s) = F(r(i, j, s)) \quad (1)$$

Then the algorithm searches local extreme through both spatial and scale space to find local maximums as candidate feature points, which we note as C . At last, a global thresholding process is taken on C abide by following equation:

$$L_{i,j} = \begin{cases} 1, & \text{if } H(i, j, s) > th_s \\ 0, & \text{else} \end{cases} \quad (2)$$

Where $L_{i,j}$ indicators whether pixel $r(i, j)$ is the center of an acceptable local feature and th_s is the threshold of feature saliency. Different from the above process considering the local saliency only, in our work, the relationship between the feature scale s and the stroke width w is also considered. Thus the probability that a local region is effective could be described as $P(H, s, w)$. According to Bayes formula, we have

$$P(H, s, w) = P(H | s, w)P(s, w) \quad (3)$$

Noticing that the calculation of local saliency H is independent to stroke width w , the probability $P(H, s, w)$ could be simplified into $P(H | s)$. Furthermore, in this paper, we describe the relationship $P(s, w)$ between s and w by a sign function of ratio r and use another sign function to describe $P(H | s)$, we get

$$L_{i,j} = P(H)P(r_{i,j}) \quad (4)$$

where

$$P(r) = \begin{cases} 1, & r \in [r_{\min}, r_{\max}] \\ 0, & \text{else} \end{cases} \quad (5)$$

and

$$P(H) = \begin{cases} 1, & H > th_s \\ 0, & \text{else} \end{cases} \quad (6)$$

Thus we could give the feature selection algorithm based on the above analysis. According to Algorithm 1, we can improve the accuracy and efficiency by removing non-informative local features. Section IV demonstrates the effect of the proposed algorithm.

Algorithm 1 Scale-based Local Feature Selection



Fig. 2. (a) and (b) shows character samples from Fnt and NS data, respectively. (c) shows Chinese word samples form CH data

IV. USING THE TEMPLATE

In this section, we verify the effect of the proposed scale-based feature selection algorithm with experiments on representative benchmarks. Section IV demonstrate the effect of the proposed algorithm.

A. Experiment Setup

1) *Experimental Data:* To prove that local features with proper scales are more effective, we conduct experiments on a representative benchmark which is referred as 'char74k' [6]. The 'char 74k' dataset contains both synthetic and natural samples. Synthetic samples include 52 classes of English characters (capital letters and lower case letters) and 10 classes of numbers (0~9). For each class, 1016 character samples are generated according to 256 different system fonts with 4 different styles. For natural samples, characters are cropped manually from scene images. Fig 2(a) and Fig 2(b) shows some typical samples of 'Fnt' data and 'NS' data in this benchmark. This dataset is selected for two reasons. Firstly, it contains typical scene character samples which are segmented manually and labeled in detail. Secondly, synthetic data could be used as baseline in our experiment since these samples certify accurate stroke width information and all detected local features are useful for character recognition. Moreover, we collect our own Chinese words dataset (the dataset will be referred as 'CH' in the following parts of this paper) beside the above benchmark using Internet searching engine according to 12 different key words. For each text image we get, accurate text regions are cropped and labeled manually. Examples of CH data are shown in Fig 2(c).

2) *Local Feature Detection:* We employ two typical detectors, which are, Hessian-affine and difference of Gaussian (DoG). According to the literature [6], the combination of DoG detector and SIFT descriptor performs much better than others.

3) *Stroke Width Extraction:* In this paper, stroke width information is extracted by utilizing stroke width transform [23]. For each pixel in a text image, if it is localized between two edges pixels with opposite gradient directions, its stroke width value is defined as the distance between these two edge pixels. If more than one pair of edge pixels are found, the stroke width value is set as the minimum one. On the contrary, stroke width value is set as infinite when the algorithm cannot find pixels like that. For more details about stroke width extraction, readers could refer to the original paper by Epstein et.al.[23]. Two factors should be considered for extracting

precise stroke width. The first one is the thresholds for edge detector (Canny here) should be selected very carefully since the precision of SWT heavily depends on the results of edge detection. The second is that the algorithm needs to know whether the character pixels are darker than the background or opposite. In practice, it is without any difficulties to assign parameters of edge detector for synthetic data as these images have high contrast (binary images, actually). Moreover, all synthetic samples have darker pixels compared to the background. For natural images, thresholds of Canny operator are assigned much lower by considering the image contrast and the contrast between text and background are assigned manually.

Based on detected local features and extracted stroke width value, we can calculate the ratio r for each local feature.

B. Character and Word Recognition

Text recognition is achieved based on classic bag-of-features framework, which is similar to literature [6]. In our experiments, 30 training samples and 15 testing samples are selected randomly for each class. Then local features are detected and described as mentioned above. The visual word vocabulary is generated by using k-means cluster algorithm, and the number P of visual words for each class is assigned equally (varies from 2 to 10 in the following experiments).

Finally, each sample is quantized into feature vector according to the vocabulary and thus each sample image is described by a $P \times C$ dimension vector where C is the number of classes. Support vector machine (SVM) with RBF kernel is chosen as classifier due to its effectiveness and representativeness and '1 VS all' strategy is employed to solve multi-class problem.

Besides, we perform recognition separately for numbers, lowercase letters and capital letters to avoid the influence of similar symbols such as 'o' and '0', 'p' and 'P'. Thus the accuracy for NS and Fnt data is calculated by using the weighted average according the following equation

$$Acc = \frac{C_c}{C} Acc_c + \frac{C_l}{C} Acc_l + \frac{C_n}{C} Acc_n$$

where C_c , C_l and C_n is the class number of capital letters, lowercase letters and numbers and $C = C_c + C_l + C_n$.

In the feature selection stage, a group of samples for each class are selected to find the best threshold for filtering especially large or small features. For each training process, we

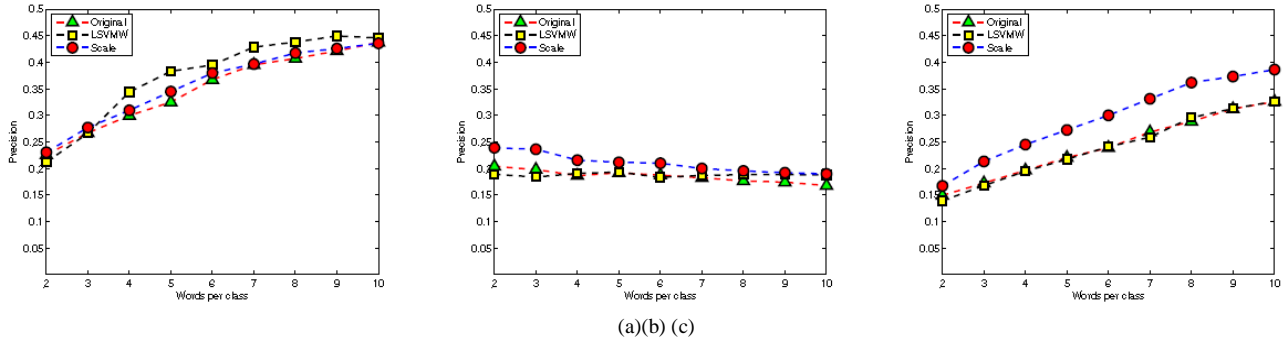


Fig. 3. Diagram (a) shows the recognition accuracy of LSVMW, proposed method (c) and original BoFs based method on Fnt data. Diagram (b) and (c) shows the corresponding results on NS and CH data, respectively.

TABLE I. IMPROVEMENTS BROUGHT BY SCALE-BASED LOCAL FEATURE SELECTION

Words per Class		2	3	4	5	6	7	8	9	10
Rate of Improvement (%)	Fnt	2.20	3.84	3.27	6.13	3.39	0.39	2.51	1.07	-0.42
	NS	17.17	19.34	15.61	10.31	11.94	9.61	10.59	10.19	13.04
	CH	12.05	23.61	24.70	23.64	25.56	23.43	25.31	19.58	18.57

remove features that have extremely large or small ratio parameter as percentage. The best filter threshold is found by employing grid search. For Fnt data, the algorithm search the best threshold from 1% to 10% for both large and small sides. The reason for limiting the searching range is that very few non-informative features are detected for Fnt data. The experimental results also show that the best thresholds in the neighborhood of 1% in most cases for Fnt data. We can find that the selection slightly improve the accuracy of Fnt data. Besides, the results of feature selection using linear SVM weight is also shown in the Fig 3. The results of MI and IG are not attached as LSVMW over-performs them. We can discover that both Scale-based feature selection and LSVMW-based method can improve the accuracy of Fntdata. However, the improvement of scale-based method is not very obvious and weaker then LSVMW-based one. The reason is that most detected local features are informative since no cluster background and noises are included in Fnt data. For the NS data which include more noises, scale-based feature selection strategy overruns both original data and LSVMW-based feature selection. The results show that the scale-based feature selection brings more benefits when a rare word number is used and the efforts of LSVMW is close to our method when the number of words increases. The reason is that when rare word number is used, the influence of noises is more obvious in that error code words will reduce the accuracy, and the proposed method is more effective for filtering non-informative features and avoiding the generation of error code words.

The accuracy for CH data is calculated under the same method. We can discover that the proposed scale-based feature selection method obviously overruns original and LSVMW-based method. This encouraging result further proves that we can filter non-informative local features by considering both feature scales and stroke width. To examine the improvement of the proposed method in greater detail, the recognition

accuracy of original data and filtered data is calculated. Moreover, the improvement brought by stroke width information is evaluated as follows: noting the recognition accuracy on original feature set as C_{ori} and C_{sel} as accuracy on selected feature set, the improvement C_{imp} can be evaluated as $C_{imp} = (C_{sel} - C_{ori}) / C_{ori}$. The results are shown in Table I. From Table I, we can see that supervised feature selection algorithms such as LSVMW are more effective for clean data and the proposed method is more effective when samples contain more noises and degradation such as NS data and CHdata.

V. CONCLUSION

In this paper, we proposed a new approach for filtering text-independent local features by considering both stroke width information and feature scales. The proposed approach is tested on representative benchmarks and the encouraging experimental results (a maximum improvement of 25.56% for CH data and 19.34% for natural data) prove the existence of relevancy between stroke width and feature scales. Different from traditional methods which need a group of training data, the proposed approach can effectively filter on-informative local features when only a few samples are used. Moreover, it is notable that the proposed approach is evidently effective for degraded images and small sample size situations. These two advantages ensure the proposed method could be widely applied in the fields such as historical document analysis and text-associate image retrieval. At the same time, we can find that there is much room for improvement in recognition rate for local feature based algorithms. Therefore, our future work include developing probability model which aims at increasing the accuracy of local feature based STR and building end-to-end scene text analysis system.

ACKNOWLEDGMENT

This work was supported by National Natural Science Foundation of China (grant No.61073128).

REFERENCES

- [1] K. Jung, K. In Kim, and A. K Jain, "Text information extraction in images and video: a survey," *Pattern recognition*, vol. 37, no. 5, pp. 977-997, 2004.
- [2] X. Chen and A. L. Yuille, "Detecting and reading text in natural scenes," in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, vol. 2. IEEE, 2004, pp. II-366.
- [3] A. Coates, B. Carpenter, C. Case, S. Satheesh, B. Suresh, T. Wang, D. J. Wu, and A. Y. Ng, "Text detection and character recognition in scene images with unsupervised feature learning," in *Document Analysis and Recognition (ICDAR), 2011 International Conference on*. IEEE, 2011, pp. 440-445.
- [4] K. Wang, B. Babenko, and S. Belongie, "End-to-end scene text recognition," in *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1457-1464.
- [5] A. Bosch, A. Zisserman, and X. Muoz, "Image classification using random forests and ferns," in *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*. IEEE, 2007, pp. 1-8.
- [6] T. de Campos, B. R. Babu, and M. Varma, "Character recognition in natural images," 2009.
- [7] Q. Zheng, K. Chen, Y. Zhou, C. Gu, and H. Guan, "Text localization and recognition in complex scenes using local features," in *Computer Vision-ACCV 2010*. Springer, 2011, pp. 121-132.
- [8] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91-110, 2004.
- [9] Q. Lv, W. Josephson, Z. Wang, M. Charikar, and K. Li, "Multi-probe lsh: efficient indexing for high-dimensional similarity search," in *Proceedings of the 33rd international conference on Very large databases. VLDB Endowment*, 2007, pp. 950-961.
- [10] M. Diem and R. Sablatnig, "Are characters objects?" in *Frontiers in Handwriting Recognition (ICFHR), 2010 International Conference on*. IEEE, 2010, pp. 565-570.
- [11] K. Das and Z. Nenadic, "An efficient discriminant-based solution for small sample size problem," *Pattern Recognition*, vol. 42, no. 5, pp. 857-866, 2009.
- [12] T. Tuytelaars and C. Schmid, "Vector quantizing feature space with a regular lattice," in *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*. IEEE, 2007, pp. 1-8.
- [13] D. Lee, S. Baek, and K. Sung, "Modified k-means algorithm for vector quantizer design," *Signal Processing Letters, IEEE*, vol. 4, no. 1, pp. 2-4, 1997.
- [14] F. Jurie and B. Triggs, "Creating efficient codebooks for visual recognition," in *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, vol. 1. IEEE, 2005, pp. 604-610.
- [15] M. Vidal-Naquet and S. Ullman, "Object recognition with informative features and linear classification," in *ICCV*, vol. 3, 2003, p. 281.
- [16] J. Brank, M. Grobelnik, N. Milic-Frayling, and D. Mladenic, "Interaction of feature selection methods and linear classification models," in *Workshop on Text Learning held at ICML, 2002*.
- [17] Y. W. Chang and C. J. Lin, "Feature ranking using linear svm," *Causation and Prediction Challenge Challenges in Machine Learning, Volume 2*, p. 47, 2008.
- [18] F. Moosmann, B. Triggs, F. Jurie et al., "Fast discriminative visual codebooks using randomized clustering forests," *Advances in Neural Information Processing Systems 19*, pp. 985-992, 2007.
- [19] S. Zhang, Q. Tian, G. Hua, Q. Huang, and W. Gao, "Generating descriptive visual words and visual phrases for large-scale image applications," *Image Processing, IEEE Transactions on*, vol. 20, no. 9, pp. 2664-2677, 2011.
- [20] L. Zhang, C. Chen, J. Bu, Z. Chen, S. Tan, and X. He, "Discriminative codeword selection for image representation," in *Proceedings of the international conference on Multimedia. ACM*, 2010, pp. 173-182.
- [21] X. He, D. Cai, and P. Niyogi, "Laplacian score for feature selection," in *Advances in neural information processing systems*, 2005, pp. 507-514.
- [22] L. Wolf and A. Shashua, "Feature selection for unsupervised and supervised inference: The emergence of sparsity in a weight-based approach," *The Journal of Machine Learning Research*, vol. 6, pp. 1855-1887, 2005.
- [23] B. Epshtein, E. Ofek, and Y. Wexler, "Detecting text in natural scenes with stroke width transform," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010, pp. 2963-2970.