52 pages including cover

CSI Communications

Knowledge Digest for IT Community

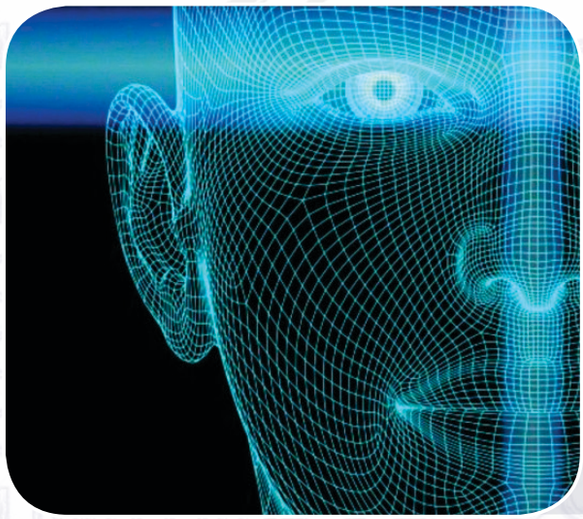
Volume No. 41 | Issue No. 11 | February 2018

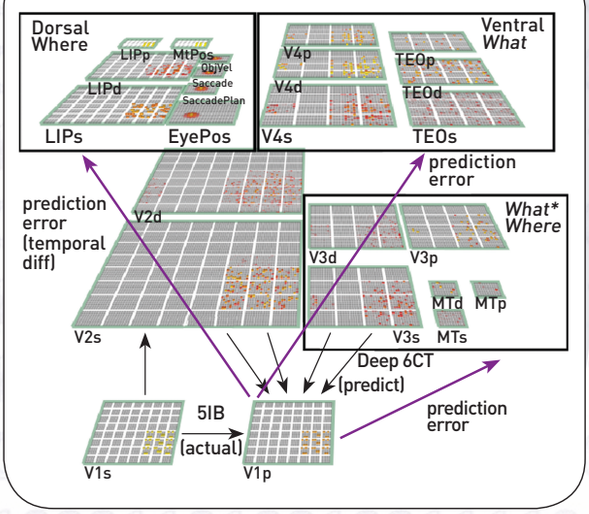
₹ 50/-

ISSN 0970-647X

SPECIAL RESEARCH ISSUE ON

PATTERN RECOGNITION





ARTICLE

www.csi-india.org

Know Your CSI

Executive Committee (2017-18/19)

President
Mr. Sanjay Mohapatra
D/204, Kanan Tower,
Patia Square, Bhubaneswar
Email : president@csi-india.org

Vice President
Mr. Gautam Mahapatra
Vailla No: 8, Maithri Enclave,
Near Tulsi Gardens, Yapral Kapra,
Hyderabad-500 062.
(E) vp@csi-india.org

Hon. Secretary
Prof. A. K. Nayak
Director, Indian Institute of Business Management, Budh Marg, Patna
(E) secretary@csi-india.org

Hon. Treasurer
Mr. Manas Ranjan Pattnaik
Plot No. N-24, 25 Chandaka Indl. Estate, Patia, KIIT, Bhubaneswar
(E) treasurer@csi-india.org

Immd. Past President
Dr. Anirban Basu
309, Ansal Forte, 16/2A,
Rupena Agrahara, Bangalore
Email : ipp@csi-india.org

Regional Vice-Presidents

Region-I
Mr. Arvind Sharma
3/294, Vishwas Khand,
Gomati Nagar, Lucknow-226010. UP
(E) rvp1@csi-india.org

Region-II
Mr. Devaprasanna Sinha
73B Ekdalia Road,
Kolkata - 700 019
(E) rvp2@csi-india.org

Region-III
Prof. Vipin Tyagi
Dept. of CSE, Jaypee University of Engg. and Tech., Guna - MP
(E) rvp3@csi-india.org

Region-IV
Mr. Hari Shankar Mishra
Command Care, Opp. Loreto Convent School, A. G. Office Road, Doranda,
Ranchi – 834002, Jharkhand
(E) rvp4@csi-india.org

Region-V
Mr. Vishwas Bondade
No. 774, 2nd Stage, Indiranagar,
Bangalore 560038
(E) rvp5@csi-india.org

Region-VI
Dr. Shirish S. Sane
Vice-Principal, K K Wagh Institute of Engg Education & Research,Nashik,
(E) rvp6@csi-india.org

Region-VII
Dr. M. Sundaresan
Professor and Head,
Dept. of IT, Bharathiar University,
Coimbatore - 641046, Tamil Nadu.
(E) rvp7@csi-india.org

Division Chairpersons

Division-I
Mr. Apoorva Agha
8, Katra Road, Allahabad, UP - 211002
(E) div1@csi-india.org

Division-II
Prof. P. Kalyanaraman
Plot No. 139, Vaibhav Nagar,
Phase I, Opp VIT Gate 3, Vellore – 632014.
(E) div2@csi-india.org

Division-III
Mr. Raju L. kanchibhotla
Aashirvad, 42/260/1/2,
Shramik Nagar,
Moulali Hyderabad-500 046, India
(E) div3@csi-india.org

Division-IV
Dr. Durgesh Kumar Mishra
H-123-B, Vigyan Nagar,
Annapurna Road, Indore
(E) div4@csi-india.org

Division-V
Dr. P. Kumar
Professor and Head
Dept. of Computer Science and Engineering,
Rajalakshmi Engineering College, Chennai – 602 105.
(E) div5@csi-india.org

Nomination Committee (2017-2018)

Prof. K. Subramanian
B 28,Tarang Apmts,
Plot 19, IP Extn,
Patparganj, Delhi - 110092
(E) ksmanian48@gmail.com

Dr. Brojo Kishore Mishra
Associate Professor, Dept. of IT,
C. V. Raman College Engineering,
Bhubaneshwar - 752054. India
(E) brojokishoremishra@gmail.com
brojokishoremishra@yahoo.com

Mr. Subimal Kundu
Flat No. 1A, Block - 7,
Space Town Housing Complex,
P.O. Airport, Kolkata – 700052
(E) subimalkundu@yahoo.com
kundu.subimal@gmail.com

- an individual.
- 2 are friends.
- 3 is company.
- more than 3 makes a society. The arrangement of these elements makes the letter 'C' connoting 'Computer Society of India'.
- the space inside the letter 'C' connotes an arrow - the feeding-in of information or receiving information from a computer.

Editors - CSI Communications

Prof. (Dr.) S S Agrawal
Chief Editor

Prof. Prashant R Nair
Editor

CSI COMMUNICATIONS

VOLUME NO. 41 • ISSUE NO. 11 • FEBRUARY 2018

Chief Editor
S S AGRAWAL
KIIT Group, Gurgaon

Editor
PRASHANT R NAIR
Amrita Vishwa Vidyapeetham, Coimbatore

Published by
AKSHAYA KUMAR NAYAK
Hony. Secretary
For Computer Society of India

Editorial Board:
Arun B Samaddar, NIT, Sikkim
Bhabani Shankar Prasad Mishra, KIIT University, Bhubanewar
Debajyoti Mukhopadhyay, MIT, Pune
J Yogapriya, Kongunadu Engg. College, Trichy
M Sasikumar, CDAC, Mumbai,
R Subburaj, SRM University, Chennai
R K Samanta, Siliguri Inst. of Tech., West Bengal
R N Behera, NIC, Bhubaneswar
Sudhakar A M, University of Mysore
Sunil Pandey, ITS, Ghaziabad
Shailesh K Srivastava, NIC, Patna
Vishal Mehrotra, TCS

Design, Print and Dispatch by
GP OFFSET PVT. LTD.

Contents

Articles

PLUS

Please note:
CSI Communications is published by Computer Society of India, a non-profit organization. Views and opinions expressed in the CSI Communications are those of individual authors, contributors and advertisers and they may differ from policies and official statements of CSI. These should not be construed as legal or professional advice. The CSI, the publisher, the editors and the contributors are not responsible for any decisions taken by readers on the basis of these views and opinions.

Although every care is being taken to ensure genuineness of the writings in this publication, CSI Communications does not attest to the originality of the respective authors' content.

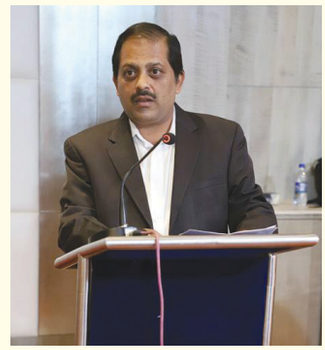
© 2012 CSI. All rights reserved.

Instructors are permitted to photocopy isolated articles for non-commercial classroom use without fee. For any other copying, reprint or republication, permission must be obtained in writing from the Society. Copying for other than personal use or internal reference, or of articles or columns not owned by the Society without explicit permission of the Society or the copyright owner is strictly prohibited.

Printed and Published by Akshaya Kumar Nayak on behalf of Computer Society of India, Printed GP Offset Pvt. Ltd. 269, 2nd Floor, A-2, Shah & Nahar Indl. Estate, Sitaram Jadhav Marg, Lower Parel, Mumbai 400 013 and Published from Samruddhi Venture Park, Unit No. 3, 4th Floor, Marol Industrial Area, Andheri (East), Mumbai 400 093. • Email : **hq@csi-india.org**
Editor: Prashant R. Nair

February 2018.indd 3 2/5/2018 5:55:29 PM

President's Message



From : President, Computer Society of India
Date : 01 February, 2018
Email : president@csi-india.org / Cell : (91) 9861010656

52nd Annual Convention of CSI @ Kolkata

We had a spectacular 52nd Annual Convention of our Computer Society of India (CSI) at Science City, Kolkata from 19 to 21 January, 2018. CSI shares a proud association and history with Kolkata; the first annual convention of CSI was held here at Indian Statistical Institute (ISI) in 1966. Subsequently, Kolkata hosted several CSI Annual Conventions in 1978, 1986, 1990 (Silver Jubilee), 1994, 2001, 2006 and 2012. I take this opportunity to congratulate the CSI Kolkata chapter for organizing this CSI annual convention in an exemplary manner. I think this is a record for Kolkata of hosting the convention 9 times.

We were honored by the presence of a distinguished keynote speaker in Prof. Mike Hinchey, President of International Federation for Information Processing (IFIP). IFIP is a global organization for ICT researchers and professionals to conduct research, develop standards and promote information sharing established in 1960 under the auspices of UNESCO. CSI represents India as a member society in IFIP. The technical sessions of the convention were very informative and lively with active knowledge exchange and networking among the delegates as well as researchers and students. The convention saw speakers from top universities, institutions and companies such as IBM, Tata Consultancy Services (TCS), Microsoft, Kansas State University USA, Nanyang Technological University (NTU) Singapore, IISc Bangalore, IIT Kharagpur, ISI Kolkata, Jadavapur University, Saha Institute of Nuclear Physics, Internet Corporation for Assigned Names and Numbers (ICANN), IFIP, CDAC etc. CSI is planning to bring the IFIP flagship events to India like the IFIP General Assembly and World Computing Congress in the future.

Another major highlight of the convention was the CSI Nihilent e-Governance awards being hosted for the 15th time. I take the opportunity to congratulate the CSI Special Interest Group (SIG) in e-Governance for anchoring this premier event of CSI, which has now become a benchmark in recognizing e-Governance initiatives in India true to the mission of Government of India, 'Digital India'. CSI recognized 2 outstanding women leaders of this country, Ms. Mamata Banerji, Chief Minister of West Bengal and Mrs. Vasundhara Raje, Chief Minister of Rajasthan for promoting e-Governance initiatives in their respective states. Ms. Mamata Banerji was represented by her Finance & Industry Minister, Mr. Amit Mitra and Mrs Raje was represented by the deputy speaker of Rajasthan, Mr. Rao Rajendra Singh. CSI is proud to confer these awards on these exceptional women leaders and honored by the presence of their representative senior leaders for the CSI Convention.

Please write your valuable ideas for growth of CSI at president@csi-india.org

With kind regards

Sanjay Mohapatra
President, CSI

February 2018.indd 4 2/5/2018 5:55:32 PM

Guest Editorial

Many real life problems suffer from uncertainty, imprecision, vagueness to name a few. Conventional computing paradigms often fall short of offering solutions to them. Even latest soft computing paradigms are not too robust to handle the situations. Hybrid soft computing is a paradigm which addresses these issues to a considerable extent.

Of late, there is enormous growth of research exploration of injecting elements of intelligence using efficient hybrid techniques. All these initiatives indicate that the individual soft computing techniques do not behave in conflicting manner rather behaves complimentary to one another, as has already been indicated. Hybrid intelligent systems stem from the synergistic integration of the different soft computing tools and techniques. The fusion of these techniques towards achieving enhanced performance and more robust solutions can be achieved through proper hybridization. In fact, recent reports reveal the inherent strength of such hybridization of computational methods.

Deep learning is an emerging approach for finding concise, slightly higher level representations of the inputs, and has been successfully applied to many practical learning problems, where the goal is to use large data to help on a given learning task. Deep learning is an emerging approach within the machine learning research community. Deep learning algorithms have been proposed in recent years to move machine learning systems towards the discovery of multiple levels of representation. Learning algorithms for deep architectures are centered on the learning of useful representations of data, which are better suited to the task at hand, and are organized in a hierarchy with multiple levels. There are several motivations for deep architectures: Brain inspiration (several areas of the brain are organized as a deep architecture); Cognitive arguments and engineering arguments (humans often organize ideas and concepts in a modular way, and at multiple levels.); Sharing of statistical strength for multi-task learning; Computational complexity. In fact, it was found recently that the features learnt in deep architectures resemble those observed in the areas V1 and V2 of visual cortex, and that they become more and more in variant to factors of variation in higher layers. Learning a hierarchy of features increases the ease and practicality of developing representations that are at once tailored to specific tasks, yet is able to borrow statistical strength from other related tasks. Finally, learning the feature representation can lead to higher-level (more abstract, more general) feature that are more robust to unanticipated sources of variance extant in real data. Typical examples of this learning technique in the soft computing paradigm are manifested in the learning in deep neural networks.

Pattern recognition and analysis have been a daunting task in the computer vision research community given the vast amount of uncertainty involved therein. Proper analysis of patterns and images plays a key role in many real life applications. Traditional applications include image processing, image mining, image inpainting, video surveillance, intelligent transportation systems to name a few. As an example, albeit ageing mitigates the glamour in human beings, wrinkles in face images can often be used for estimation of age progression in human beings. This can be further utilized for tracing unknown or missing persons. Images exhibit varied uncertainty and ambiguity of information and hence understanding an image scene is far from being a general procedure. The situation becomes even graver when the images become corrupt with noise artifacts. Image denoising, as one of the glaring problem in pattern recognition, can be described as the problem of mapping from a noisy image to a noise-free image. Various methods have been proposed for image denoising. One approach is linear or non-linear filtering methods which are a relatively simple approach based on smoothing, such as Median filtering which replace each pixel with the median of the value of a set of neighboring pixels, linear smoothing and wiener filtering. Other methods are based on wavelet or dictionary decompositions of the image. Wavelet decompositions transfer image signals to an alternative domain where they can be more easily separated from the noise, such as BLS-GSM. The dictionary-based method is to denoise by approximating the noisy patch using a sparse linear combination of atoms, including KSVD which is an iterative algorithm that learns a dictionary on the noisy image at hand, NLSC which is one of the best currently available denoising algorithms in terms of quality of the results, but requires long computation times.

Apart from the many challenges that a robot faces today, one of the biggest challenges still, is to increase the vision of the robot; making the robot able to see things, identify them, identify the hurdles, etc. Hurdle detection is one of the common tasks that have been done through image processing, by identifying and recognizing different types of objects in the image and then calculating the distance between robot and hurdles. Image analysis has a lot to contribute in this direction.

Another emerging area of image analysis and understanding is color image processing which includes processing of colored images in different color spaces. It also involves studying transmission, storage, and encoding of these color images. Pattern recognition involves study from image processing and from various other fields that includes machine learning (a branch of artificial intelligence). In pattern recognition, image processing is used for identifying the objects in an image and then machine learning is used to train the system for the change in pattern. Pattern recognition is used in computer aided diagnosis, recognition of handwriting, recognition of images etc.

In addition, a video is nothing but just the very fast movement of pictures. The quality of the video depends on the number of frames/pictures per minute and the quality of each frame being used. Video processing involves noise reduction, detail enhancement, motion detection, frame rate conversion, aspect ratio conversion, color space conversion etc.

Pose and gesture analysis plays an important role in various applications such as human-machine interaction, behavior analysis, video surveillance, annotation, search and retrieval, motion capture for the entertainment industry and interactive web-based applications. Current trends in research are based on the development of realtime video analysis algorithms mainly focusing on hand and head tracking and gesture

February 2018.indd 5 2/5/2018 5:55:32 PM

analysis. The main features for tracking and segmentation of hands and head are skin-color and motion. This provides robust and temporarily stable recognition results in relatively unconstrained scenarios. A faithful gesture recognition algorithm can be conducted with techniques from computer vision and image processing.

New instruments for imaging human brain activity, such as functional Magnetic Resonance Imaging (fMRI), offer a wonderful opportunity to study mechanisms in the brain. Several statistical machine learning algorithms have been developed to analyze fMRI data in order to learn to identify and track the cognitive processes that give rise to observed fMRI data. These fMRI data can be decoded faithfully using image analysis algorithms to understand as to which word a person is thinking about based only on the neural activity captured in their fMRI data.

Last but not the least the fields of remote sensing and medical image processing are also demanding arenas under consideration by the image processing research community.

This special issue on pattern recognition intends to bring together researchers to report the latest results or progress in the development of the hybrid soft computing paradigm for faithful pattern recognition. As such, the focus of this special issue is the methods of Computational Intelligence, with a focus on hybrid soft computing methods applied to pattern recognition research.

Siddhartha Bhattacharyya

Profile of Guest Editor

Dr. Siddhartha Bhattacharyya did his Bachelors in Physics, Bachelors in Optics and Optoelectronics and Masters in Optics and Optoelectronics from University of Calcutta, India in 1995, 1998 and 2000 respectively. He completed PhD in Computer Science and Engineering from Jadavpur University, India in 2008. He is the recipient of the University Gold Medal from the University of Calcutta for his Masters. He is the recipient of the coveted National Award **Adarsh Vidya Saraswati Rashtriya Puraskar** for excellence in education and research in 2016. He is the recipient of the **Distinguished HoD Award** and **Distinguished Professor Award** conferred by Computer Society of India, Mumbai Chapter, India in 2017. He is the recipient of the coveted **Bhartiya Shiksha Ratan Award** conferred by Economic Growth Foundation, New Delhi in 2017. He received the NACF-SCRA, India award for **Best Faculty for Research** in 2017. He received the **Honorary Doctorate Award (D. Litt.)** from The University of South America and the South East Asian Regional Computing Confederation (SEARCC) International Digital Award **ICT Educator of the Year** in 2017. He also received the **Rashtriya Shiksha Gaurav Puraskar** from Center for Education Growth and Research, India in 2017. He has been appointed as a ACM Distinguished Speaker for the period 2018-2020.

He is currently the Principal of RCC Institute of Information Technology, Kolkata, India. In addition, he is also serving as the Professor of Computer Application and Dean (Research and Development and Academic Affairs) of the institute. Prior to this, he was the Professor of Information Technology of RCC Institute of Information Technology, Kolkata, India. He served as the Head of the Department from March, 2014 to December, 2016. Prior to this, he was an Associate Professor of Information Technology of RCC Institute of Information Technology, Kolkata, India from 2011-2014. Before that, he served as an Assistant Professor in Computer Science and Information Technology of University Institute of Technology, The University of Burdwan, India from 2005-2011. He was a Lecturer in Information Technology of Kalyani Government Engineering College, India during 2001-2005. He is a co-author of 4 books and the co-editor of 8 books and has more than 185 research publications in international journals and conference proceedings to his credit. He has got a patent on intelligent colorimeter technology. He was the convener of the AICTE-IEEE National Conference on Computing and Communication Systems (CoCoSys-09) in 2009. He was the member of the Young Researchers' Committee of the WSC 2008 Online World Conference on Soft Computing in Industrial Applications. He has been the member of the organizing and technical program committees of several national and international conferences. He served as the Editor-In-Chief of International Journal of Ambient Computing and Intelligence (IJACI) published by IGI Global, Hershey, PA, USA from 17th July 2014 to 06th November 2014. He was the General Chair of the IEEE International Conference on Computational Intelligence and Communication Networks (ICCICN 2014) organized by the Department of Information Technology, RCC Institute of Information Technology, Kolkata in association with Machine Intelligence Research Labs, Gwalior and IEEE Young Professionals, Kolkata Section and held at Kolkata, India in 2014. He is the Associate Editor of International Journal of Pattern Recognition Research. He is the member of the editorial board of International Journal of Engineering, Science and Technology and ACCENTS Transactions on Information Security (ATIS). He is also the member of the editorial advisory board of HETC Journal of Computer Engineering and Applications. He is the Associate Editor of the International Journal of BioInfo Soft Computing since 2013. He is the Lead Guest Editor of the Special Issue on Hybrid Intelligent Techniques for Image Analysis and Understanding of Applied Soft Computing, Elsevier, B. V. He was the General Chair of the 2015 IEEE International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN 2015) organized by the Department of Information Technology, RCC Institute of Information Technology, Kolkata in association with IEEE Young Professionals, Kolkata Section

February 2018.indd 6 2/5/2018 5:55:33 PM

and held at Kolkata, India in 2015. He was the Lead Guest Editor of the Special Issue on Computational Intelligence and Communications in International Journal of Computers and Applications (IJCA); Publisher: Taylor & Francis, UK in 2016. He is the Editor of International Journal of Pattern Recognition Research since January 2016. He was the General Chair of the 2016 International Conference on Wireless Communications, Network Security and Signal Processing (WCNSSP2016) held during June 26-27, 2016 at Chiang Mai, Thailand. He was the General Chair of the 2016 IEEE Second International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN 2016) organized by the Department of Information Technology, RCC Institute of Information Technology, Kolkata jointly with IEEE Young Professionals, Kolkata Section and held at Kolkata, India during September 23-25, 2016. He is the member of the editorial board of Applied Soft Computing, Elsevier, B. V. He is serving as the Series Editor of the IGI Global Book Series **Advances in Information Quality and Management (AIQM)** from January 01, 2017. He is also Series Editor of the De Gruyter Book Series **Frontiers in Computational Intelligence (FCI)** from April 27, 2017.

His research interests include soft computing, pattern recognition, multimedia data processing, hybrid intelligence and quantum computing. Dr. Bhattacharyya is a fellow of Institute of Electronics and Telecommunication Engineers (IETE), India. He is also a senior member of Institute of Electrical and Electronics Engineers (IEEE), USA, International Institute of Engineering and Technology (IETI), Hong Kong and Association for Computing Machinery (ACM), USA. He is a member of Institution of Engineering and Technology (IET), UK, International Rough Set Society, International Association for Engineers (IAENG), Hong Kong, Computer Science Teachers Association (CSTA), USA, International Association of Academicians, Scholars, Scientists and Engineers (IAASSE), USA, Institute of Doctors Engineers and Scientists (IDES), India and The International Society of Service Innovation Professionals (ISSIP). He is a life member of Computer Society of India, Optical Society of India, Indian Society for Technical Education and Center for Education Growth and Research, India.

PUBLIC NOTICE FOR PROPERTY ON AUCTION

Sealed offers are invited by the Computer Society of India (CSI) bearing Regd. No. 41 of 1965, for sale of immovable property (72.1 Sq mtr.) at, **3rd floor, Nagesh Apartments, Nr. Panjim Municipal Market, Goa**, "AS IS WHERE IS BASIS" and Persons interested in purchasing the said property shall submit their offers latest by 25h February, 2018 to the **President, Computer Society of India, Unit No. 3, 4th Floor, Samrudhi Venture Park, MIDC, Andheri (E), Mumbai - 400093** together with a Demand Draft drawn in favour of the Computer Society of India, payable at Mumbai for 25% of the offer amount by way of earnest money security deposit.

The prospective purchasers may visit the property and inspect the same at their cost.

Sealed offers shall be opened on **25/03/2018 (Sunday) at 11a.m. at Computer Society of India, Unit No. 3, 4th Floor, Samrudhi Venture Park, MIDC, Andheri (E), Mumbai - 400 093**. Offerer's can remain present if desired.

The Computer Society of India, reserves their right to accept and/or reject all or any of the offers without assigning any reasons thereof. Sale is subject to sanction of Registrar of Societies, Hyderabad.

Sd/

Sanjay Mohapatra
President, CSI
5th February, 2018

CSI COMMUNICATION ANNOUNCEMENT

CSI Communications February 2018 Issue is a special research issue on Pattern Recognition. Because of shortage of space in this issue, we are not able place Chapter, Student Branch and Student Convention in this issue. These reports will appear in the next issue of CSI Communications. Inconvenience is regretted.

Sd/

Editor

February 2018.indd 7 2/5/2018 5:55:34 PM

How deeply we can recognize patterns today?

Anasua Sarkar
Assistant Professor, Computer Science & Engineering Departmwnt, Jadavpur University, E-mail: anasua.sarkar@cse.jdvu.ac.in , ashru2006@hotmail.com

1. Introduction

Arthur Samuel coined the term "**machine learning**" in **1959** [1]: "Machine Learning is a field of study that gives computers, the ability to learn without explicitly being programmed." He implements a Checker playing program to learn withr time. Initially it loses. But after some time, it follows all the board positions to win or loss and thus succeeds to be a good chess player than Samuel himself. Later on, **T.H. Davenport** defines in Analytics as an *excerpt from The Wall Street Journal* - "Humans can typically create one or two good models a week; machine learning can create thousands of models a week."

Pattern recognition is a subfield in machine learning which explores to recognize patterns and regularities in data. Pattern recognition is one of the most exciting technologies that one would have ever come across. We imagine to create our intelligence through explorations of pattern recognition algorithms on real-world data.

2. Why is Pattern Recognition important?

Resurging exploration in pattern recognition is due to the same features which make **data mining** and Bayesian analysis so well-known. In the era of big data, with boosting volumes and several varieties among data, computing approaches needs to be faster and more efficient.

Therefore, present scenario is suitable to quickly and automatically detect patterns and to form models to explore bigger, more complex, noisy data. It needs modern algorithms to provide more accurate results in very fast time and even on a very large scale. And applying these innovative pattern recognition approaches, an organization succeeds to identify profitable opportunities in its own filed and also can avoid unknown risks in today's world.

Because of new intelligent machine learning technologies, pattern recognition today has evolved in bigger dimensions than its previous explorations. It was started to detect patterns from noisy or simulated data, and emerges now to become the procedure to learn computers to solve with no need of programming to achieve certain computable solutions. AI today has even started its breakthrough to create artificial thinking through programming. The iterative aspect of pattern recognition has provided it a power to look into deeper view of data looking through the window of features, to learn by itself. This deep technology has gained fresh boosting in this big data era.

To produce an effective pattern recognition method, first stage is to produce noiseless data, then an advanced and fast algorithm needs to be chosen specific for the problem to solve. The automatic flow of data to develop the model needs to be built. Big data boom adds the requirement of scalability in all kinds of applications today. Lastly, to provide as best efficient solution as possible, ensemble approaches over several pattern recognition techniques need to be implemented for real-life applications.

While several pattern recognition algorithms have been explored in last decades, the advantage to automatically perform complex mathematical computations over big data as well as to simulate thinking and creation like human mind, is a recent development. Some popular and widely publicized examples of pattern recognition applications are:

- Self-driving Google car,
- Online recommendation offers from Amazon and Netflix,
- Machine learning approach with linguistic rules collects customers' review from Twitter
- Fraud detection approaches of PayPal.

3. How Data Mining, Machine Learning and Deep Learning Differs?

Real world of data are infinite to explore and analyze. However, we name them in different ways for our method of explorations on those data. All these approaches have similar power of extracting knowledge, patterns as well as associations from noisy data to make decisions. But each of these methods, has its own unique approach and ability.

- **Data Mining**

Data mining is a mining exploration which extracts knowledge hidden in the data. It may incorporate existing statistical approaches or innovative machine learning approaches. Data mining also exploits data storage and manipulation methods. Data mining applications are spreaded across different areas/applications utilizing concepts from statistical approaches, machine learning, language analysis, time-series prediction as well as in application areas.

- **Machine Learning**

Statistical models tries to learn from the data its structure, using theoretical distributions to fit to the data. But this methodology requires strong assumptions on data. Differently, machine learning tries to learn a model from the given data, which it can apply later on other test data, for prediction of values (regression) or labels (classifications). It provides validation errors on new test data and do not stop after proving null hypothesis. For its own iterative approach, it can learn automatically to detect patterns, using any of its supervised, unsupervised or semi-supervised algorithms.

- **Deep learning**

Deep learning is the self-taught learning method, which learns itself

February 2018.indd 8 2/5/2018 5:55:34 PM

from data representations. Recently this Deep structured learning technology has been spreaded across all fields of applications like - Computer visions, speech recognition methods, text analytics and video-game methods. Deep learning approaches may be of all three kinds - Supervised, Semi-supervised and Unsupervised.

Deep Neural Network (DNN) is an Artificial Neural Network which has several hidden layers lies between its two end layers - input and output layers. It is a feed-forward network where data flows from input layer to output layer without looping back. Deep Belief Network is another variation with back-propagation with connections between the layers but which do not have any connections between units within each layer. It gains the insight to re-create specific type of inputs by utilizing its back-propagation approach. Recurrent neural networks (RNNs) is the remarkable advancement of Pattern Recognition, where data can flow in any direction. This back-propagation method provides the power of self-taught learning to these neural networks, which are now evolving as the next boom in AI to imitate human brain and its activities in all application fields.

4. Pattern Recognition Applications

As we move towards the big data digital era, our artificial intelligence leaps and strides forward. Most recent invention in pattern recognition is the Deep Machine Learning. This innovative form of pattern recognition has become very popular in various professional applications - in marketing, medicine as well as web security. Pattern recognition makes it easier to predict without much human efforts. This approach only requires training data to learn better features needed to learn a given system. The only drawback is that pattern recognion sometimes is not efficient for non-convex problems or with non-differentiable discrete parameters. Pattern recognition methods are popular now in many real-life problems. Here vare some examples from those applications.

1) How Does Pattern Recognition Enhance Robotics?

In the basic definition of Machine Learning, it is a way in pattern recognition which distributes ability to computers to learn new knowledge without new programming. In depth, these methods can learn themselves and expand their lewrning over new test data from real life to test on.

Machine Learning was defined as to be *"based on algorithms that can learn from data without relying on rules-based programming. It came into its own as a scientific discipline in the late 1990s as steady advances in digitization and cheap computing power enabled data scientists to stop building finished models and instead train computers to do so."- in the article An Executive's Guide to Machine Learning*, from McKinsey Quarterly issue in 2015.

Machine learning has an ability to generate knowledge for collaborative robotics. Given human inputs by Demostration, deep programming can enhance robotic skillsets to collaborate properly. The name 'robot' was first coined by the Czech playwright Karel Capek (pronounced "chop'ek") from the Czech word for forced labor or serf [2]. Carnegie Mellon CS Department defines it as: *"Force through intelligence or where AI meets the real world." The steps to combine pattern recognition with robotics are-*

a) Learning proper grasp movements by human demonstration.
b) Planning new contact points on objects.
c) Optimizing and performing "reach-and-grasp" movements.

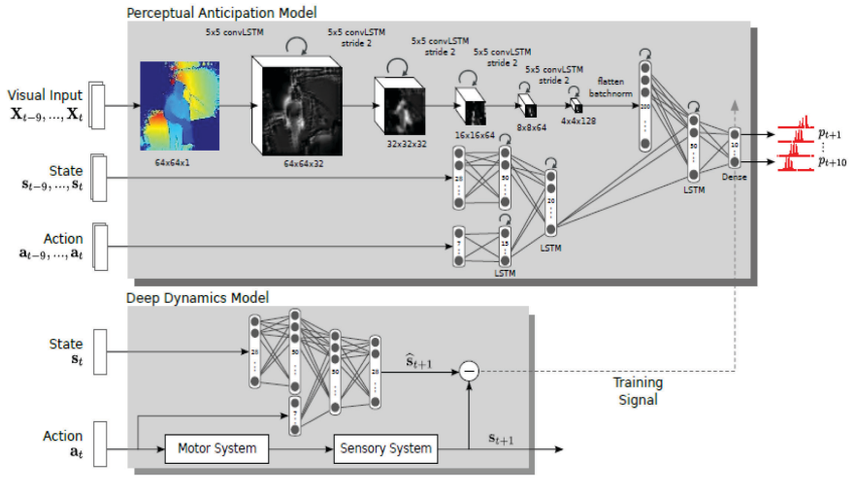


Fig. 1 : Architecture of Deep Predictive Model for perturbations [3].

Imitation Learning for robots is something like to do as infants. This approach is one advancement in utilizing robotics in construction, agriculture, security, rescue and other industries. In all these situations, manually programming of the robots to proceed is very difficult. Baxter robot anticipates physical contact by human [3]. It follows the architecture of Deep Predictive Model for perturbations. The Likelihoods of perturbation are predicted by Perceptual Anticipation Model after self-supervised training. The different predictors fire at different moments ahead of the actual physical perturbation, indicating real physical contacts.

2) Natural Language Processing (NLP)

NLP is today very popular in different disciplines. Machine learning approaches in natural language helps customer to obtain services in their mother languages. It can help to extract legalese being translated to help attorneys to sort through high volume of data to make the case.

a) Automatic Translation

Today, language translation in Google is very common. However, it can also translate text in any language from images. Google can make real time visual translation of 20 languages. The neural network first recognizes letters, then approximate dictionary look up is performed on words and then image

February 2018.indd 9 2/5/2018 5:55:35 PM

replaced with translated word.

b) **Handwriting recognition**

Most common utilizations of machine learning today in pattern recognition is for its capability to recognize different kinds of images. For example, US-Postal-Service utilizes machine learning approaches for handwriting recognition. It is very easy to exploit pattern recognition methods in some customized way for needed applications using SAS® Enterprise Miner™.[4].

3) **Artificial Musics and Drawing**

Magenta is the project of Google Brain team. It tries to find out one basic query: Can machines create compelling arts and music? [5] It is developed using an open-source infrastructure on TensorFlow to enhance the machine intelligence involved in music and art generations. They have developed approaches to learn to create art and music, generating impulsive and artistic contents by machines. This method is termed as "Experiments in Musical Intelligence (EMI)", where an Artificial Intelligence model decomposes works of human composers, then searches for common signatures among those earlier works, and then recombines them to create new compositions. Similar another approach is Performance RNN, which is one LSTM-based RNN -Recurrent Neural Network, which can model polyphonic music with expressive timing and dynamics. Sketch-RNN is another generative model for vector drawings[18].

4) **Marketing Personalization - Improve Customer Service**

Today we are familiar with personalized customer assistance, as they are provided by Amazon [6] or Netflix . Amazon provides you to recommend any product through its five stars rating with every item purchase. Based on these choices, Amazon generates "Your recommendations" page for its every registered user with their own personal preferences. Amazon deep learning recommendation algorithm uses item-to-item collaborative filtering.

Amazon's Destiny project [7] is their implementation of Deep Scalable Sparse Tensor Network Engine (DSSTNE) (pronounced "destiny"), which is a library created by Amazon to build Deep Learning models. They cliam it to be advanced than TensorFlow, because they enable users to distribute the deep-learning problem not only among different servers, but alos among different processors in each server. This system has the ability to predict efficiently using less data. DSSTNE was built with a number of features for production recommendation workloads:

i) Multi-GPU Scale: Training and prediction both phases use multiple GPUs.

ii) Large Layers: Model-parallel scaling enables larger networks.

iii) Sparse Data: DSSTNE is optimized to work faster sparse datasets, common in recommendation problems. Customized GPU kernels experiment sparse computation on GPUs.

The newly launched, Lumidatum proprietary machine learning platform [8] also helps its registered companies to find new revenue streams fueling the growth to its business. This machine learning project, uses big data science and statistical models to deploy predictive analytics, considering each and every interaction the company is making with its customers.

5) **Online Search**

Online search is the most common usage of pattern recognition in our everyday web life. Every time we search a new word in Google or other search engine today, their program machine learning program watches our responses. If we choose from their top results, it detects a success for its programming. However, if we go for a second choice page, then these search programs tries learn the mistake to deliver a better result next time.

For Google Brain project [9], in Google's mysterious X laboratory, they built a neural network with 16,000 processors using one billion connections. Then the network browses YouTube in search of cats. This "brain" simulation learned from 10 million randomly selected YouTube video thumbnails in three days. Later given 20,000 different items, it finally started recognizing cats in images using deep learning method. Using most common images in YouTube, this model obtained 81.7 % accuracy to detect human faces, 76.7 % accuracy to detect human body parts and 74.8 % accuracy to identify cats. This program is called "deep Q-network," or DQN, which can run on a regular PC.

"The idea is that instead of having teams of researchers trying to find out how to find edges, you instead throw a ton of data at the algorithm and you let the data speak and have the software automatically learn from the data," said Andrew Ng, who is the computer

```
In [73]: # randomly unconditionally generate 10 examples
         N = 10
         reconstructions = []
         for i in range(N):
           reconstructions.append([decode(temperature=0.5, draw_mode=False), [0, i]])

In [74]: stroke_grid = make_grid_svg(reconstructions)
         draw_strokes(stroke_grid)
```

Fig. 2 : Sketch-RNN to generate drawings.[18]

February 2018.indd 10 2/5/2018 5:55:35 PM

scientist at Stanford University involved in this self-taught learning algorithm project. They implement this Deepmind project on a large scale distributed infrastructure for neural network, called DistBelief. TensorFlow [10] is the next deep learning project platform, developed by Google after DistBelief, in 2011. This open source deep learning framework is very popular today for Convolutional neural network implementations and it is improved and generalized to use in user-chosen applications.

Google even has succeeded to make its Deepmind AI to learn itself the art of walking from object orientations [11]. It was never shown, how actually human walks. But using its own virtual sensors, it can decide how to move forward. Their four legged spider like model choices to jump, while two-legged humanoid model choses to walk. It can jump, keep balance, learns from mistake trials itself. This model even has learned itself a sideway walking style, which we even do not know.

Google uses a neural network they call "deep learning", to detect web spams in email. Their approach obtains data from different users and from different natural-language-processing machine learning models, they draw conclusions about the emails for spam detection. Google claims that that they are able to block 99.9 percent of spam to reach inboxes, while incorrectly classifyied legitimate e-mail as spam is only 0.05 percent.

6) Data Security

Data security has become vulnerable today, as malware has become very big problem. In 2014, Kaspersky Lab received 325,000 new malwares *each day*. But, new intelligence companies like Deep Instinct defines new malwares tend with similar code from previous ones with only 2 %-10% changes. Therefore, they can predict which files are malware with these 2–10% variations with greater efficiency. In other applications, machine learning models find out patterns in accessing cloud for data, and report anomalies to predict security breaches.

a) Fraud Detection

PayPal is using its own pattern recognition application to fight money laundering. PayPal's fraud-management options are like "detective-like methodology". They use Deep learning algorithms at granular levels [12]. It is like anomaly detection methods in pattern recognition. This algorithm can analyze around thousands of latent features like time signals, performers and geographic locations etc., to detect one fraud. It also is even able to detect "sub modus operandi," or different variants of the same scheme. To this approach, a sudden series of several small purchases at convenience stores predict a fraud.

PayPal also employs human detectives who are constantly analyzing patterns in the system for anomaly. With anomaly, the human detective makes up "a good story and a bad story" to explains two possible reasons for the transaction. Machine learning helps to learn and scale it up.

Trustev is the first online web system available today that fuses identity data with digital data in real time[13]. They examines all data for an online transaction or event to keep fraud out while keeping good customers.

7) Healthcare

Machine learning algorithms can detect more patterns than humans in healthcare applications. Computer assisted diagnosis (CAD) learns from early mammography scans of women and the method detects 52% of breast cancers as much as a year before their officially diagnosis. The industry Medecision developed a model that was able to detect eight variables to predict avoidable hospitalizations in diabetes patients. An expert with another data scientist generate regression models to find out relationships between independent variables which may lead to future events. in Medecision[14]. It trains the platform on a database of approximately 8 million patients. The analysis indicated that most of the avoidable hospital admissions were for upper respiratory infections that were complicated by diabetes but not caused by it.

Today, deep learning applications are trends poised to sweep through $6.5 trillion healthcare industry. Artesys has introduced a platform based on cloud and deep learning to which can revolutionize medical imaging. It exploits parallel GPU architecture to perform instant big data quantification and provides 4D Blood Flow post-processing. Imagia is another application, with its mission to detect and quantify cancer changes early. It supports efficient cancer drug development with new imaging biomarkers appropriate for modern clinical trial endpoints.

a) Assistive and Medical Tech

David L. Jaffe of Stanford defines Assistive robots to be devices which can process automatically its sensory information, and decide to perform tasks with disabled people. With these concepts, the Smart Tissue Autonomous Robot (STAR) was developed [15]. Utilizing machine learning and 3D sensing, this device was tested to stitch together pig intestines, which performed better than any human surgeon. STAR is a collaborative solution for delicate steps in surgeries.

8) Smart Cars

IBM recently surveyed that 74% is that we could drive smart cars on road by 2025. A smart car integrates into the Internet of Things, but can also learn knowledeg about its owner and its environment through sensors. This will learn to adjust automatically internal comfortable settings, like temperature, audio, seat position, etc. from driver and sensor reports. It will drive itself, and may check driver's emotion to avoid accidents. Some recent examples fo smart cars are - Google Driverless Car project, Autopilot –Tesla, IBM Watson IoT for Automotive and Aeromobil.

Google Chauffeur [16] is the program for Google Driveless car project. It uses a LIDAR system Velodyne 64- laser beam to generate a detailed 3D map of its environment. The car then generates maps combined with world, and apply different data models to navigate safely through roads. It has driven over 1 million miles for testing with only 14 minor accidents.

IBM uses cognitive computing to connect vehicles with its drivers. In IBM Watson IoT for Automotive initiative [17], connected vehicles obtain real-time information from their sensors about vehicles and in-vehicle data to understand the driving to improve the

February 2018.indd 11 2/5/2018 5:55:35 PM

in-car experience.

Tesla has already launched Autopilot, for driveless transport, while Aeromobil is the first official flying car which has been marketed.

5. Conclusions

Pattern Recognition is an incredible breakthrough in artificial intelligence. Although it recreates some dangerous banes like Blue whale video games, these innovative algorithms can lead us to an era of illuminated intelligent future life, which will be much smarter and easier than today.

6. References

[1] Samuel, Arthur L. (1959). "Some Studies in Machine Learning Using the Game of Checkers". *IBM Journal of Research and Development*

[2] Robots Defends Himself - Karl Capek, Lidove noviny, June 9, 1935.

[3] Indranil Sur and Heni Ben Amor , Robots that Anticipate Pain: Anticipating Physical Perturbations from Visual Cues through Deep Predictive Models, International Conference on Intelligent Robots and Systems (IROS) 2017 .

[4] https://www.sas.com/en_us/software/enterprise-miner.html

[5] https://magenta.tensorflow.org/welcome-to-magenta

[6] https://www.amazon.com/gp/help/customer/display.html?nodeId=13316081

[7] https://github.com/amzn/amazon-dsstne

[8] https://www.lumidatum.com/

[9] https://www.wired.com/2012/06/google-x-neural-network/

[10] https://www.tensorflow.org/

[11] https://deepmind.com/

[12] https://www.technologyreview.com/s/545631/how-paypal-boosts-security-with-artificial-intelligence/

[13] https://www.trustarc.com/

[14] https://www.medecision.com/

[15] S. Leonard, K. L. Wu, Y. Kim, A. Krieger and P. C. W. Kim, "Smart Tissue Anastomosis Robot (STAR): A Vision-Guided Robotics System for Laparoscopic Suturing," in *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 4, pp. 1305-1317, April 2014.

[16] https://www.google.com/selfdrivingcar/

[17] https://www.ibm.com/internet-of-things/iot-solutions/iot-automotive/

[18] https://magenta.tensorflow.org/sketch-rnn-demo

About the Author

Anasua Sarkar has completed her Ph D work on Bioinformatics in Jadavpur University, Kolkata, India in 2014. She was also awarded EMMA-EPU fellowship, 2011 to pursue her research work at LaBRI, University Bordeaux1, France. She is presently an Assistant Professor in Computer Science and Engineering Department, Jadavpur University, Kolkata. She has published 14 original research papers in peer-reviewed journals. She also a co-author in 16 book chapters and more than 15 conference papers. She is also reviewer in journals of Parallel and Distributed Computing (JPDC), Elsevier and IEEE SMCC-C. She is also a Student member, ISCB, 2010-2011, SMIEEE since 2014, member CSTA, ACM Chapter, ACM IGUCCS and IAENG. She has worked with INRIA MAGNOME group for 18 months in France for the expansion of Genolevures database for inclusion of PISO and ARAD species. Her research interests include Proteomics, Phylogenetics, Computational Biology, Pattern Recognition, Data Mining, Bioinformatics, Embedded, Parallel Systems.

OBITUARY

Computer Society of India mourns the sad demise of **Brig. (Retd.) Saranjit S. Sahney**, who left for Heavenly Abode on November 8, 2017. Brigadier Sahney was the person who initiated Chandigarh Chapter of Computer Society of India and was the founder Chairman. He worked very hard for the growth of the CSI Chandigarh Chapter. He had made a significant place in the hearts of everybody through his loving nature and management skills. The void so created with his leaving is impossible to fill and he will be remembered for the times to come.

CSI prays to the Almighty God to give peace to this pious soul and strength to all family members, friends and relatives to bear this irreparable loss.

Subhash Chander Jain
Chairman, CSI Chandigarh Chapter

February 2018.indd 12 2/5/2018 5:55:36 PM

Human Interpretability in Machine Learning Based Solutions

Pawan Kumar
Lovely Professional University, Phagwara, India. Email: pawan.11522@lpu.co.in

Majority of the machine learning models for classification or prediction tasks behave like a black-box. They are able to give a high predictive performance, but, they lack in terms of their outcomes being interpretable for a human expert i.e. why a particular outcome was given for a given instance. This interpretability is particularly important where decision-making is involved based on outcomes of machine learning models. This article, explores recent approaches proposed to confer interpretability to machine learning based solutions. There is a lot of scope for future work in the field of human interpretability as it has the potential to develop trust in machine learning models and facilitating deployment of machine learning based solutions.

Keywords : Machine Learning. Human Interpretability. Model-agnostic.

1. Introduction

Machine learning refers to algorithms that can make a computer learn from a large number of examples. The basic idea is to extract a formal statistical model from the given examples and using it to predict the value or class of the target variable for an unseen example [1]. If the value of target variable is known for each example, it is called supervised machine learning. Further, if the target variable is categorical, the task is known as classification and if the target variable is continuous, the task is known as regression. In this article, we are limiting our scope to classification or prediction task i.e. when the dependent variable is categorical in nature. For example, whether a given patient is having diabetes or not? Whether a student will graduate or not? If the possible outcomes are limited to two, it is called binary classification. If the possible outcomes are above two, it is termed as multi-class classification problem.

The primary objective of a machine learning model is to correctly classify an instance. However, in many problem-domains like medicine, the classification is followed by a critical decision making. For example, take the case of using machine learning to predict whether a tumour is malignant or benign. In this problem, both the type of misclassifications i.e. false positives as well as false negatives are hazardous. A false positive i.e. classifying a benign tumour as malignant will recommend a not-required chemotherapy. On the other hand, a false negative i.e. classifying a malignant tumour as benign will lead to no treatment at all and let the disease advance further. In such problem domains, to develop a trust in the machine learning model, it is desirable that the outcome of the model should be understandable to a human expert [2]. In other words, the human expert should be able to identify what features made the model predict a particular outcome for a given instance. So, a human interpretable machine learning model is basically one whose outcomes can be interpreted by a human expert.

2. Importance of human interpretability in machine learning

Trust: Particularly in problems that involve critical decision-making, if the outcomes of a machine learning model are understandable to a human domain expert, it helps building trust of the domain expert or stake holders in the model. While interpreting the model, if an association or causal relation is observed that is contradicting our so-far knowledge or common sense, it acts as an indicator of a potential problem in the data or the developed model [3].

New Insights: When we are applying machine learning in problem domains related to physical sciences e.g. understanding gene data, the input consists of a large number of features in the order of thousands or even more. Using complex machine learning models, we are able to achieve high prediction accuracy. However, if the outcomes are understandable to human experts in terms of relative contribution of features towards model outcome, new insights can be obtained about our understanding of the underlying physical processes. New correlations may be detected while interpreting that were not known so far. Interpretability can help mitigating gap between real life and machine learning objectives [3].

Right to Explanation Moreover, new regulations of European Union is giving stake-holders a right to interpretability or explanations of the outcomes if they are using or are affected by the machine learning algorithms.

3. Trade off in Accuracy and Interpretability

A few machine learning algorithms like decision trees and CART (Classification and Regression Trees)

February 2018.indd 13 2/5/2018 5:55:37 PM

are inherently interpretable. Similarly, linear models like SVM (Support Vector Machine) are easy to interpret. However, more complex models like random forests, tree ensembles, boosted trees, artificial neural networks are not interpretable. However, in complex problems where the number of features is very large and relationship between dependent variables and independent variables is highly non-linear, the simpler models like decision trees, CART etc do not give high prediction accuracy. So, machine learning engineers are making use of more complex models like Boosted trees, Tree ensembles and Artificial neural networks. These models have been able to learn complex relationships among data and hence to give high prediction accuracy. However, the disadvantage is that these models behave like a black-box i.e. they do not explain their outcomes. Hence, there seems to be a trade-off between interpretability and prediction accuracy of machine learning models. In other words, an attempt to increase the interpretability is usually at the cost of accuracy.

4. Approaches towards interpretability

Although there have been model specific (or algorithmic specific) interpretability attempts also, but it is desirable that there should be a model-agnostic approach to bring interpretability to machine learning based solutions. A model-agnostic approach is that which can explain or interpret classification or prediction outcomes of any classifier, be it Random forests or artificial neural networks. A clear advantage of model-agnosticism is that the choice of the model is abstracted or it does not affect the interpretability approach.

5. Open Research Questions

Although, a number of techniques have been proposed recently to address the issues related to interpretability, a lot of scope is there for future work. First, addressing trade-off between interpretability and prediction accuracy is of important concern to develop solutions that are human interpretable without compromising on accuracy. Second, there is a need to explore the proposed interpretability approaches in the medical, speech and video domains. Third, as interpretability is an issue usually in complex models due to large number of features and complex architecture of machine learning algorithms like deep neural networks, computational optimization approaches are worth exploration to facilitate real time solutions. Fourth, there is a lack of consensus about formal definitions of interpretability measures. It makes comparative analysis difficult. Fifth, there is another area is to develop evaluation metrics for interpretability using visual analytics. Sixth, the approaches proposed for interpretability in classification models can be explored for regression tasks. Also, it is important to enable machine learning models to interact with humans so that feedback mechanisms can be implemented.

6. Growing interest in interpretable models

In the recent years, there has been an increase in interest shown in development of human interpretable machine learning. Starting with model specific approaches, there is a shift towards developing model-agnostic approaches. Recent editions of the topmost conferences in machine learning and machine learning competitions have witnessed focus on having human interpretable machine learning solutions. For example, the two topmost conferences in the field of machine learning i.e. ICML (International Conference on Machine Learning) and NIPS (Neural Information Processing Systems) had hosted workshops on human interpretability in 2016 and 2017[9-11].

7. Conclusion

Most of the machine learning models behave like a black-box. Despite of being good at predictive performance, they lack interpretability. In problem

Table 1 : Approaches used for conferring human interpretability

Paper Title	Aim	Contribution
An Efficient Explanation of Individual Classifications using Game Theory [4]	A method for explaining outcome of any model i.e. model-agnostic	Outcomes are explained in terms of individual contributions of features using concepts of coalitional game theory
How to Explain Individual Classification Decisions [5]	Explain classification outcome by identifying most influential features	Providing local explanation vector that is basically an estimation of local gradients to understand instance-specific outcome
Making Tree Ensembles Interpretable [6]	Lack of interpretability in tree ensembles like random forest, boosted trees etc	To approximate the learned complex tree ensemble by a simple model. Idea is to use two models, 'P' for prediction and 'I' for interpreting
Using Visual Analytics to Interpret Predictive Machine Learning Models [7]	To achieve interpretability without losing on accuracy using visual analytics	Interpretability by representing input-output relationships of a machine learning classifier through visual analytics
Why Should I Trust You? Explaining the Predictions of Any Classifier [8]	Explain the prediction outcome of any classifier by learning an interpretable model locally around the prediction	LIME – an algorithm to interpret an instance outcome i.e. "Trusting the instance", SP-LIME – an algorithm for "Trusting the model"

February 2018.indd 14 2/5/2018 5:55:37 PM

EAIT 2018

Mr. Subir Lahiri, Chairman, CSI Kolkata chapter presenting memento to Prof. A.K. Nayak, Hon. Secretary, CSI at 5th International Conference on Emerging Applications of Information Technology (EAIT) 2018 at Kolkata

domains involving critical decision making based on outcomes of machine learning models, interpretability of outcomes by a human expert is of important concern. Interpretability of machine learning solutions can play an important role in trust building, getting new insights and facilitating deployment of machine learning based solutions. The challenges in this area include interpretability-accuracy trade-off, computational optimizations and lack of consensus about formal definition of interpretability and evaluation metrics.

8. References

[1] Pedro Domingos, "A Few Useful Things to Know about Machine Learning", Communications of the ACM, Volume 55 Issue 10, October 2012, Pages 78-87

[2] Kiri L. Wagstaff, "Machine Learning that Matters", Proceedings of the 29th International Conference on Machine Learning, Edinburgh, Scotland, UK, 2012

[3] Zachary C. Lipton, "The Mythos of Model Interpretability", 2016 ICML Workshop on Human Interpretability in Machine Learning (WHI 2016), New York, NY, USA

[4] Erik ˇStrumbelj, Igor Kononenko, "An Efficient Explanation of Individual Classifications using Game Theory", Journal of Machine Learning Research 11 (2010) 1-18

[5] David Baehrens, Timon Schroeter, Stefan Harmeling, "How to Explain Individual Classification Decisions", Journal of Machine Learning Research 11 (2010) 1803-1831

[6] Satoshi Hara, Kohei Hayashi, "Making Tree Ensembles Interpretable", 2016 ICML Workshop on Human Interpretability in Machine Learning (WHI 2016), New York, NY, USA.

[7] Josua Krause, Adam Perer, Enrico Bertini, "Using Visual Analytics to Interpret Predictive Machine Learning Models", 2016 ICML Workshop on Human Interpretability in Machine Learning (WHI 2016), New York, NY, USA

[8] Marco Tulio Ribeiro, Sameer Singh, Carlos Guestrin, "Why Should I Trust You? Explaining the Predictions of Any Classifier", KDD 2016, San Francisco, CA, USA, http://dx.doi.org/10.1145/2939672.2939778

[9] https://sites.google.com/view/whi2017/home

[10] https://sites.google.com/site/nips2016interpretml/

[11] http://www.filmnips.com/

■

About the Author

Pawan Kumar, (CSI Membership No. 1183942) Assistant Professor, School of Computer Applications, Lovely Professional University, Punjab

Research Interests: Machine Learning, Data Analytics and ICT Applications in health care

Experience: More than 14 years of experience in academics

Kind Attention: Prospective Contributors of CSI Communications

Please note that Cover Theme for **March 2018 issue is Sensors for Internet of Everything (IOET)**. Articles may be submitted in the categories such as: Cover Story, Research Front, Technical Trends, Security Corner and Article. Please send your contributions by 20th February, 2018.

The articles should be authored in as original text. Plagiarism is strictly prohibited.

Please note that CSI Communications is a magazine for members at large and not a research journal for publishing full-fledged research papers. Therefore, we expect articles written at the level of general audience of varied member categories. Equations and mathematical expressions within articles are not recommended and, if absolutely necessary, should be minimum. Include a brief biography of four to six lines, indicating CSI Membership no., for each author with high resolution author photograph.

Please send your article in MS-Word format to to Editor, **Prof. Prashant R. Nair** in the email ids **csic@csi-india.org** with cc to **prashant@amrita.edu**

(Issued on the behalf of Editorial Board CSI Communications)

Dr. S S Agrawal
Chief Editor

February 2018.indd 15 2/5/2018 5:55:37 PM

Implementation of Gesture Vocalizer that enables Full Duplex Communication between Deaf & Dumb and Normal People

S Hemalatha
Professor, Department of CS and Engg., Panimalar Inst. of Tech., Chennai, Tamilnadu
Email: pithemalatha@gmail.com

M Archana
Final Year, Dept. of CS and Engg., Panimalar Inst. of Tech., Chennai - Tamilnadu, Email: archanamurugan20@gmail.com

S Haripriya
Final Year, Department ofCS and Engg., Panimalar Inst. of Tech., Chennai, Tamilnadu
Email: haripriya2296@gmail.com

E Shiny Freeda
Final Year, Dept. of CS and Engg., Panimalar Inst. of Tech., Chennai - Tamilnadu, Email: e.shinyfreeda@gmail.com

In this world many people suffer from hearing loss (deaf) and speech loss (dumb) that might have occurred since birth or during their lifetime later. It is tedious for the deaf & dumb people to talk with the ordinary people. Peoples with sensory disabilities use their hands, fingers to communicate and have access to the world. Sign language is an expressive way for communication between normal and deaf-dumb people and it is the symbolic representation of words by using hands and fingers in different positions. So ordinary people like us learn the sign language for communication. The sign language of deaf and dumb is tedious to learn and it is not possible for everybody to learn that language. So every person cannot come and share their thoughts with these physically impaired people. Sign language has the major limitation as it is hard to understand by a normal people, only persons those who know the sign language can able to communicate. So it creates barrier in communication between the impaired and normal people. Hence there must a midway that would perform gesture detection, convert the gestures into speech format (gesture to voice translating module) and converts the speech of normal person to gesture (speech to image translating module) and it is displayed on display. So, the whole idea is to build an android application that enables two way communications between deaf-mute (speech-hearing impaired) person and a normal person. This application uses certain image processing techniques.

Keywords : Image processing, Gesture detection, Sign language.

1. Introduction

Humans know each other Humans grasp one another by transfer their ideas, thoughts, and experiences to the individuals around them. Though speech everybody will terribly convincingly transfer their thoughts and perceive one[1] another. Communication could be a method of exchanging ideas, thoughts, feelings and data in kind of verbal or non verbal message. For effective communication it's necessary that sender also as receiver will realize it. the massive reason behind the shortage of communication is that deaf individuals ar unable to pay attention and dumb individuals ar unable to talk.

The only means that of communication on the market to the deaf and dumb individuals is that the use of linguistic communication. Victimization linguistic communication they're restricted to their own world. This restriction prevents them from interacting with the outer world to share their feelings, inventive ideas and Potentials. only a fewfolks that don't seem to be themselves deaf and dumb ever learn to linguistic communication. This will increase the isolation of deaf and dumb individuals from the common society. So, deaf and dumb individuals use gesture to speak with traditional person. Gesture refers to any bodily motion or states significantly any hand motion or face motion (or) Gestures square measure numerous hand movement of specific form. Traditional individuals cannot perceive the particular which means of these gestures. Technology is a method to get rid of this hindrance and profit these individuals.

Hand gesture recognition provides Associate in Nursing intelligent and natural approach of human laptop interaction (HCI). [10]Hand gesture recognition is vicinity in computing and language technology that aims in process human gestures via mathematical algorithms. With gesture recognition it's potential for humans to act simply with machines while not the help of any mechanical devices. Hand

February 2018.indd 16 2/5/2018 5:55:38 PM

gesture is one in every of the foremost communicatory and most often used among a unique gestures.

Applications of hand gesture recognition square measure varied from linguistic communication to computer game. Gesture recognition is classed into 2 main classes i.e. vision primarily based and device based. Vision primarily based is capturing hand gesture victimization internet camera whereas, device primarily based is capturing gesture victimization flex device and hand gloves. [2]Some disadvantages of vision primarily based techniques like it includes advanced algorithms. Another challenge in image process includes variant lighting conditions, backgrounds and field read constraints and occlusion.

The aim of this paper is to gift associate degree application which will expeditiously convert linguistic communication gesture to speech. the sole thanks to enhance the communication between dumb-deaf folks and traditional folks is to acknowledge the linguistic communication and changing it to the corresponding voice signal. A digital camera is placed before of the physically impaired person. The physically impaired person would be actinghand gestures exploitation his fingers. once he makes the gestures, the digital camera[11] can capture the precisepositions of the fingers and perform image process to see the co-ordinates of the displacements. The co-ordinates captured are going to be mapped with the one antecedently hold on and consequently actual gesture are going to be captured. continued during this manner physically impaired person are going to be able to bear the wholespeech that he needs to speak. Afterward this gesture is going to be regenerate into voice in order that it'd be sounding to everybody. Additional the applying aims[3] to convert the speech signals from the traditional person to linguistic communication and also the equivalent gesture for that speech signal are going to be showed the impaired person. thus the 2 manner communication is feasible. This application may facilitate the visually impaired folks to speak with speech impaired folks. This project aims to bridge the gap by introducing a reasonable golem mobile within the communication path in order that the linguistic communication are often mechanically captured, detected and translated to voice for the advantage of the blind folks and other people World Health Organization can not perceive the linguistic communication.

2. Literature Survey

S. No.	Title	Author Name	Publication Journal/Date	Disadvantage	Advantage
1	Gesture-controlled user interfaces, what have we done and what's next?	Moniruzzaman Bhuiyan, Rich Picking	Proceedings of the Fifth Collaborative Research Symposium on Security, E-Learning, Internet and Networking (SEIN 2009), Darmstadt, Germany. 2009.	Some users are uncomfortable with the common input devices.	We present a review of the history of **Gesture controlled user interface (GCUI)**, and identify trends in technology, application and usability.
2	Hand Gesture Recognition Based on Karhunen-Loeve Transform	Joyeeta Singha, Karen Das	Mobile & Embedded Technology International Conference, pp. 365-371, 2013.	Continuously focusing on hand gestures by manual is very difficult.	we propose an easy approach to recognize different hand gestures based on KL Transform.
3	Approach to Hand Tracking and Gesture Recognition Based on Depth-Sensing Cameras and EMG Monitoring	Ondrej Kainz, František Jakab	Acta Informatica Pragensia, Vol:3, No:1, pp.104–112, 2014.	Nature of gesture can be not only static but also dynamic, so it is difficult to notice.	A new approach for hand tracking and gesture recognition based on the **Leap Motion device and surface electromyography (SEMG)** is presented.
4	Lessons Learned in Exploring the Leap MotionTM Sensor for Gesture-based Instrument Design	Jihyun Han, Nicolas Gold	Proceedings of the International Conference on New Interfaces for Musical Expression, pp. 371-374, July 2014.	The problem is that the methods are called on every frame so MIDI messages were being triggered for every consecutive frame where the fingers moved faster than a certain speed	It presents lessons learned from work-in-progress on the development of musical instruments and control applications using the **Leap MotionTM sensor**.
5	The Leap Motion controller: A view on sign language	Leigh Ellen Potter, Jake Araullo, Lewis Carter	Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration, pp. 175-178. ACM, 2013	We find less accuracy in the existing system.	This paper presents an early exploration of the suitability of the **Leap Motion controller** for Australian Sign Language (Auslan) recognition.

February 2018.indd 17 2/5/2018 5:55:38 PM

3. Methodology Used

Deaf and mute face the communication downside with everybody they move and therefore the key answer is that thesigning. Communication between the deaf and non-deaf has continually been a really cumbersome task. Finding ANknowledgeable interpreter for the day to day activities may be a tough task and is additionally unaffordable. Since communication is that the elementary side of human survival, varied [4] measures are taken to technically improve the benefit of communication for the deaf. This paper describes the system that overcomes the matter two-faced by the speech and hearing impaired. the most objective is to ascertain the communication method between Deaf & Dumb and therefore the traditional person. The modules includes

1. Mobile application
2. Voice to Image Conversion
3. Gesture to Speech Conversion

3.1 Mobile Application

One of the most popular handicaps is the deaf and dumb type, which prevent person from listening and talking. The number of deaf and dumb in the world continuously increasing and they are introverted closed society. In this paper, we introduce an integrated android application to blend uneducated Deaf-Dumb people within society, and help them to communicate with normal people. The mobile application is proposed to recognize the sign language. Android based mobile application is deployed in the deaf & dumb person for communication purpose[5]. Once the application is installed in the user's android mobile using .apk file, the user can open an android application. The user is asked to enter the mobile number and the logical address[11]. Then, the application takes the user to the section where he/she has to select either gesture input or voice input. Generally, the gesture input is provided by the impaired person and the voice input is given by normal people. This is the main model to obtain the gesture inputs and to process them. In this mobile application, camera is automatically initiated and captures the gesture input image provided by the deaf & dumb person. [9]Then the stored pre-recorded voices are generated as the output for input gestures. Normal person will speak through an android application, it recognizes the voice input which is converted into text and corresponding image is displayed to the Deaf & Dumb person. So this application can be implemented from the both the end.

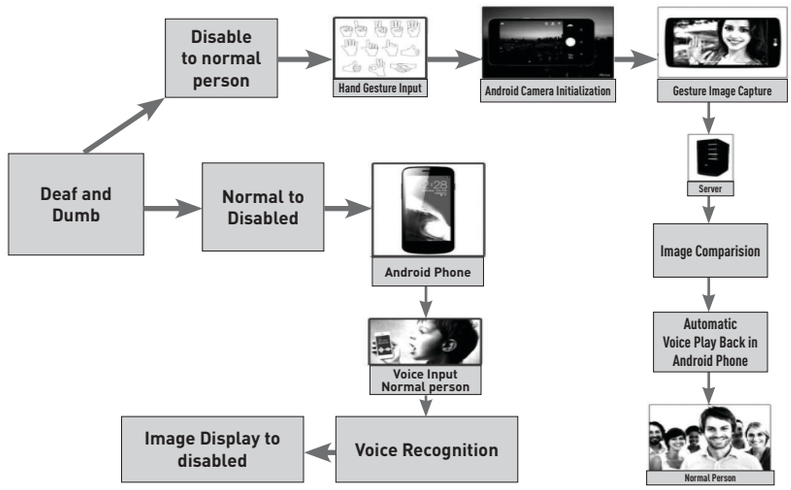


Fig 1 : Block Diagram

3.2 Voice to Image conversion

In this module, the translation of the voice input to the hand gesture images is illustrated. When the user chooses voice input, he/she will be asked to speak by clicking a specific button in an application. Once the user's voice is recognized using speech recognition functions, it is converted into text. If not recognized, application will trigger an automatic toast message. The converted text is taken as an input[12]. The text obtained is compared with the pre-defined set of texts. As the text matches, the corresponding and the equivalent image or the hand gesture frame is displayed as a message to the impaired user[6]. The hearing impaired person can easily understand what the normal person speaks. The conversion of voice to image is shown in the fig.2

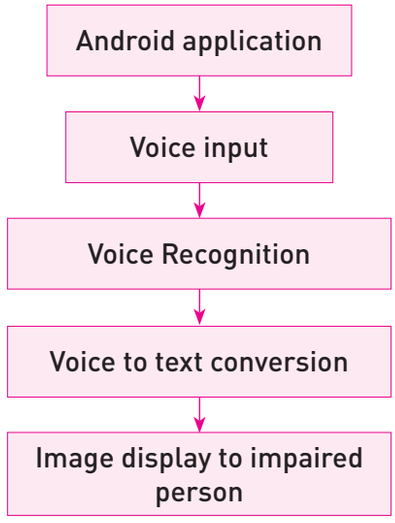


Fig 2 : Flow diagram for voice to image conversion

3.3 Gesture to Speech Conversion

In this module, the translation of the hand gesture input to the speech is explained. This module includes sub modules image acquisition, image preprocessing and gesture recognition. From the hand gesture image captured, the background is eliminated and converted into voice. The flow chart for the image to speech conversion is shown in the fig.3.

3.3.1 Image Acquisition

The first stage of any vision system is that the Image acquisition stage. Image acquisition is that the creation of photographic pictures, like of a physical scene or of the inside structure of associate degree object. The term is usually assumed to imply or embrace the process, compression, storage, printing, and show of such pictures. Image acquisition is that the method to

February 2018.indd 18 2/5/2018 5:55:39 PM

capture the hand gesture pictures that represents totally different signs. The resolution of varied image capturing devices might not be an equivalent. This ends up in totally different resolution of the captured pictures. For correct comparison of the options and to scale back the process effort required for process, all the pictures ought to be scaled to an identical size. so the pictures for coaching and testing square measure captured with net camera and info is made. This info consists of hand gesture. a picture acquisition humanoid camera is employed, then frames square measure send to the server that is followed by cutting that scale back the noise. The module in brief describes the schemes of capturing the image from humanoid device, image detection, and process the image to acknowledge the gestures.

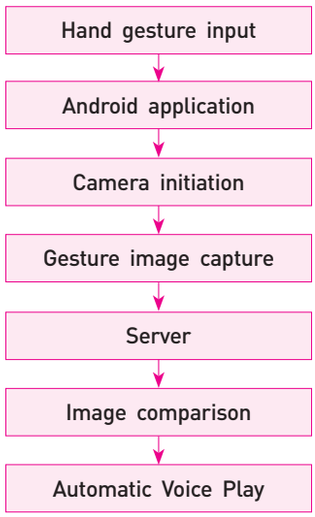


Fig. 3 : Flow diagram for gesture to voice conversion

3.3.2 Image preprocessing

A pre-processing is incredibly abundant needed task to be drained hand gesture recognition system. Pre-processing is applied {to pictures to pictures to photographs before we are able to extract options from hand images. Segmentation is that the classification of the input colored image into skin and non-skin pixels supported complexion data. Preprocessing additionally includes the background subtraction formula. the item or hand gesture targeted is taken into account as dark pixels and therefore the background is taken into account as light-weight pixels the sunshine pixels square measure eliminated and therefore the gesture is metameric as shown in fig four. Image to Sound Conversion

Input Image Detected Image Identified Gesture Identified Gesture Audio File Output Fig 4. Background is eliminated from image and regenerate to voice options of the metameric image may be extracted in several ways in which in step with specific application. under neath totally different scene conditions[7], the performance of totally different| of various feature detectors are going to be considerably different the character of the background, existence of alternative objects (occlusion), and illumination should be thought-about to work out what reasonably options may be with efficiency and faithfully detected.

3.3.3 Gesture Recognition

Gesture recognition may be a growing field of analysis among varied human laptop interactions; hand gesture recognition is incredibly fashionable for interacting between human and machines. it's nonverbal manner of communication and this analysis space is packed with innovative approaches. This module aims at recognizing basic hand gestures. the most options used square measure centre of mass within the hand, presence of thumb and variety of peaks within the hand gesture. that's the formula based mostly|is predicated|relies

on form based options by keeping in mind that form of human hand is same for all kith and kin except in some things. the popularityapproach utilized in this project is artificial neural network in Machine learning formula. for every gesture, 2 or a lot ofpictures ought to be keep within the info. the appliance is trained to acknowledge the pre-stored gestures underneath varied conditions like differing angles, orientation and illumination. a man-made Neural Network (ANN) is associate degree scientific discipline paradigm that's galvanized by the manner biological nervous systems, like the brain, method data. ANNs, like folks, learn by example. associate degree ANN is organized for a particular application, like pattern recognition or knowledge classification, through a learning method.

Module Implementation

The android application is installed in the user's mobile phone. Then the connection is made between the android application and the server i.e. system. The application is shown in Fig. 5.

Fig. 5 : Application Start Page

When an application gets started, user will enter the mobile number, logical address of the server system and submit the inputs. This is done to enable communication between an application and server. The input page is shown in Fig. 6.

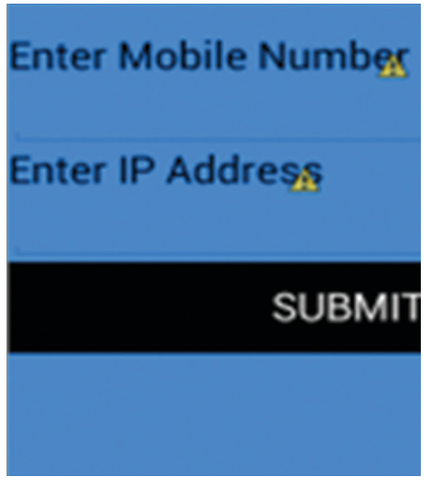


Fig. 6 : Input Page

When the inputs are submitted

February 2018.indd 19 2/5/2018 5:55:39 PM

and connection is made, an application moves to a page where choices are given. The first option is to give the hand gesture input and the next option is to give the voice input. The options are shown in Fig. 7.

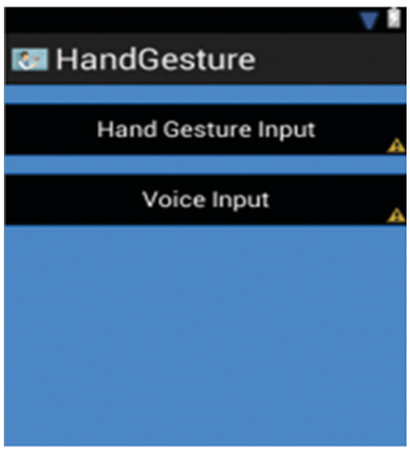


Fig. 7 : Options Page

If the user chooses to give voice as input it switches to a page where a speaker button is given as shown in fig.8. When the user presses the button, it will ask the user to speak. As the user conveys a word, the voice is recognized and converted to text and displayed in the text field above the speaker button. Then the text is compared and equivalent image file is displayed in image view field.

Fig.8 Voice Input and Image Output

If the user chooses to give gesture as an input, the front camera is initiated. The user will provide the gesture as an input as shown in fig.9. The given gesture is captured and it undergoes pre-processing i.e. resizing the images[8] to uniform size and subtracting its background. Next, the obtained gesture is compared with images in the training database. Finally the equivalent audio file is played for the given gesture input.

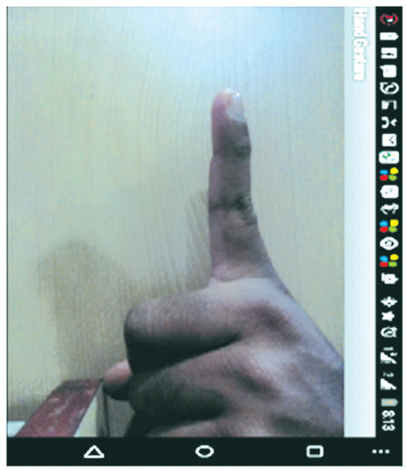


Fig. 9 : Gesture Input and Voice Output

For recognizing the gestures accurately this paper preferred the concept of machine learning and training the server. For one gesture, more than five images are stored in database for effective comparison and recognition. The training phase is shown in fig.10.

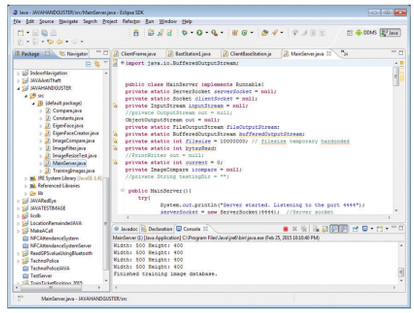


Fig. 10 : Training the Server

3. Result

In this paper, an android application that enables two way communication between deaf and dumb and normal people. The projected methodology interprets sign language gesture into speech and the speech of normal person is converted into text and corresponding hand gesture is displayed, so the communication between them can take place easily. The proposed method is tested on different gestures. It produces fairly stable and good results using Image Processing technique like background subtraction method. The huge variation in images also played a part in low accuracy rates. This includes different sizes of gestures in the images, different background of images, different orientations and angle of gestures, etc.

4. Conclusion

The main aim of our project is to provide a helping hand to the deaf and dumb people. We as normal human beings too find it difficult to communicate our feelings with them, this application will help us to share our feelings with them. This paper proposes a simple and efficient sign language recognition system to eradicate the communication gaps between the deaf and dumb people with normal people. Our method doesn't require any gloves and devices for detecting the hand movements, owing to cost of the system is reduced by great margin. This system can be used by people by spending less cost when compared to the existing system, since the latter requires additional devices to detect hand gestures. There are reasons for poor performance of testing data. Sign language is a useful tool to ease the communication between the deaf person and normal person. The application aims to lower the communication gap between deaf people and normal world, since it facilitates two way communications. Our intention in this project is that we make our end users, the deaf and dumb and the normal people happy and satisfied while using this application for communication.

Future work will address extraction of frame from live video feed using and applying frame filtering techniques such as blurring, RGB(Red, Green, Blue) to HSV (Hue, Saturation, Value) conversion and removing noise so that the image so obtained would make the gesture recognition simpler. After getting the gesture, the database is matched with the gesture and the corresponding speech is given as output.

References

[1] Bhuiyan, Moniruzzaman, and Rich

February 2018.indd 20 2/5/2018 5:55:40 PM

Picking, "Gesture-controlled user interfaces, what have we done and what's next.", Proceedings of the Fifth Collaborative Research Symposium on Security, E-Learning, Internet and Networking (SEIN 2009), Darmstadt, Germany. 2009.

[2] Singha, Joyeeta, and Karen Das, "Hand gesture recognition based on "Karhunen-Loeve transform.", Mobile & Embedded Technology International Conference, pp. 365-371, 2013.

[3] Kainz, Ondrej, and František Jakab, "Approach to Hand Tracking and Gesture Recognition Based on Depth-Sensing Cameras and EMG Monitoring.", Acta Informatica Pragensia, Vol:3, No:1, pp.104–112, 2014.

[4] Han, Jihyun, and Nicolas Gold, "Lessons Learned in Exploring the Leap Motion TM Sensor for Gesture-based Instrument Design", "Proceedings of the International Conference on New Interfaces for Musical Expression, pp. 371-374, July 2014.

[5] Potter, L. E., Araullo, J., & Carter, L. "The leap motion controller: a view on sign language", Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation,Application, Innovation, Collaboration, pp. 175-178. ACM, 2013

[6] Rachel Bainbridge,Joseph A. Paradiso, "Wireless Hand Gesture Capture Through Wearable Passive Tag Sensing", International Conference on Body Sensor Networks, pp. 200-204, IEEE 2011.

[7] Pragati Garg, Naveen Aggarwal and Sanjeev Sofat ,"Vision Based Hand Gesture Recognition", International Journal of Computer, Electrical, Automation, Control and Information Engineering Vol:3, No:1, pp. 186-191, 2009.

[8] Shraddha R. Ghorpade, Prof. Surendra K. Waghmare, "A Communication System for Deaf and Dumb People", International Journal of Science and Research (IJSR) Volume: 4, Issue: 9, pp. 1914-1917 September 2015.

[9] Koli P.B., Chaudhari Ashwini, Malkar Sonam, Pawale Kavita & Tayde Amrapali, "Image Processing Based Language Converter for Deaf and Dumb People", IOSR Journal of Electronics and Communication Engineering (IOSR-JECE), e-ISSN: 2278-2834, p-ISSN: 2278-8735 pp.25-30, 2015

[10] Kshirasagar Snehal P., Shaikh Mohammad Hussain, Malge Swati S., Gholap Shraddha S.,Mr. Swapnil Tambatkar, "Gesture vocalizer for deaf and dumb", International Journal Of Engineering Sciences & Research Technology, Volume: 5, Issue: 4, pp. 547-553, 2016.

[11] Sangeetha, K. and Barathi Krishna, L., "Gesture detection for deaf and dumb people", International Journal of Development Research, Vol. 4, Issue. 3, pp. 749-752, March, 2014.

[12] Shraddha R. Ghorpade, Surendra K.Waghamare, "Full duplex communication system for deaf and dumb people", International Journal of Emerging Technology and Advanced Engineering, Volume: 5, Issue: 5, pp. 224-227, 2015.

■

About the Authors

Dr. S. Hemalatha did her Bachelor degree in Computer Science and Engineering and Masters in Computer Science and Engineering University of Madras and Anna University Chennai, India in 2000, and 2004 respectively. She completed PhD in Computer Science and Engineering from Anna University, India in 2016. She has totally 17 years of experience in different engineering colleges and currently she is working as a Professor in Computer Science and Engineering in Panimalar Institute of Technology, Chennai, India. Shea has published 45 national and International journals and 30 papers in international conferences. She is the recipient of Distinguished Professor award from CSI Mumbai chapter in the year 2016, Professional achiever award from IEEE madras section in the year 2017. She has guided many UG and PG level projects and being a committee member for Research Scholars in Anna University. Her research areas are Network Security, Mobile Communication and Mobile Application Development. She is a member of CSI, IEEE and ISTE.

M. Archana is a final year Computer Science and Engineering student in Panimalar Institute of Technology. She is a member of Computer Society of India.

S. Haripriya is a final year Computer Science and Engineering student in Panimalar Institute of Technology. She is a member of Computer Society of India.

E. Shiny Freeda is a final year Computer Science and Engineering student in Panimalar Institute of Technology. She is a member of Computer Society of India.

February 2018.indd 21 2/5/2018 5:55:40 PM

Study on the Aiding Tools of for Effective Social Contact

S Hemalatha
Professor, Department of CS and Engg., Panimalar Inst. of Tech., Chennai, Tamilnadu
Email: pithemalatha@gmail.com

M Archana
Final Year, Dept. of CS and Engg., Panimalar Inst. of Tech., Chennai - Tamilnadu, Email: archanamurugan20@gmail.com

S Haripriya
Final Year, Department ofCS and Engg., Panimalar Inst. of Tech., Chennai, Tamilnadu
Email: haripriya2296@gmail.com

E Shiny Freeda
Final Year, Dept. of CS and Engg., Panimalar Inst. of Tech., Chennai - Tamilnadu, Email: e.shinyfreeda@gmail.com

This paper proposes survey on the assorted techniques utilized by Hearing-Speech impaired person to speak with the traditional person. during this world many of us suffer from hearing disorder (deaf) and speech loss (dumb) either from their birth or later. it's tedious for the deaf & dumb folks to speak with the standard folks. Peoples with sensory disabilities use their hands, fingers to speak and have access to the planet. language is AN communicative method for communication between traditional and deaf-dumb folks and it's the symbol of words by mistreatment hands and fingers in numerous positions. therefore normal folks like U.S.A. learn the language for the aim of communication. The language of deaf and dumb is tedious to find out and it's impractical for everybody to find out that language. therefore one and all cannot return and share their thoughts with these physically impaired folks. language has the key limitation because it is tough to know by a standard folks, solely persons those that understand the language will ready to communicate. therefore it creates barrier in communication between the impaired and traditional folks. This paper demonstrates numerous approaches for gesture recognition. It includes mistreatment hardware like good gloves with flex sensors, passive RFID tags, Depth sensing cameras and Leap motion atomic number 69 device. This paper additionally presents the algorithms used for gesture recognition likeskin detection, Karhunen-Loeve rework and chemist worth approaches. every methodology has its own blessings and downsides that has been in short delineated during this paper.

Keywords : Image process, Gesture recognition, signing.

1. Introduction

Humans grasp one another by transference their concepts, thoughts, and experiences to the folks around them. Through speech everybody will terribly convincingly transfer their thoughts and perceive one another. Communication may be a method of exchanging concepts, thoughts, feelings and data in sort of verbal or non verbal message. For effective communication it's necessary that sender moreover as receiver will realize it. The large reason behind the shortage of communication is that deaf folk's area unit unable to concentrate and dumb folk's area unit unable to talk.

The only suggests that of communication offered to the deaf and dumb folks is that the use of signing. Victimisation signing they're restricted to their own world. Only a few those that don't seem to be themselves deaf and dumb ever learn to signing. These limitations will increase the isolation of deaf and dumb folks from the people. So, deaf and dumb folks use gesture to speak with traditional person. Gesture refers to any bodily motion or states significantly any hand motion or faces motion (or) Gestures area unit numerous hand movement of specific form. Traditional folks cannot perceive the particular that means of these gestures. Technology is a method to get rid of this hindrance and profit these folks.

Hand gesture recognition provides AN intelligent and natural method of human pc interaction (HCI). Hand gesture recognition may be a section in engineering science and language technology that aims in describing human gestures via mathematical algorithms. With gesture recognition it permits the humans to move naturally with machines while not the assistance of any devices. Hand gesture is one among the foremost communicatory and most often used among a range of how. Applications of hand gesture recognition area unit completely different from signing to video game. Gesture recognition is classed into 2 main classes i.e. vision based mostly and sensing element based. Vision based mostly is capturing hand gesture

February 2018.indd 22 2/5/2018 5:55:41 PM

victimisation net camera whereas, sensing element based mostly is capturing gesture victimisation flex sensing element and hand gloves. Some disadvantages of vision based mostly techniques like it adopts complicated algorithms. Another issue in image process includes variable lighting conditions, backgrounds and field read constraints and occlusion (presence of objects in background). Interaction, as well as the assorted out there approaches is explained by Pragati Garg, Naveen Aggarwal and Sanjeev Sofat [7]. The concept of building a communication system that permits communications between speech-hearing impaired and a standard person is illustrated [8][9]. Detector primarily based techniques is employed for gesture detection [10][12]. Sangeetha, K. and Barathi avatar, L. planned the system uses Skin detection formula that detects the colouring of the users whereas giving gestures. The easy methodology is constructing a bar chart of properties. Another methodology is to produce coaching for the neural networks [11].

Voice to Image conversion Acoustic pre-processing, Extraction and Speech Recognition by Shraddha R. Ghorpade, Surendra UN agency square measure uncomfortable with input devices. GCUI appears acceptable for current and future omnipresent devices. within the early analysis gesture management or recognition method was terribly advanced due to the hardware necessities, however currently it's straightforward as straightforward vision technique is employed [1]. The system supported KL (Karhunen-Loeve) remodel to acknowledge hand gestures by Joyeeta Singha1, Tibeto-Burman language Das [2]. Ondrej Kainz, František Jakab conferred the analysis on typical approaches toward hand trailing and gesture recognition as given in [3]. The vision primarily based surface *diagnostic procedure is applicable to several areas of Human laptop Interaction (HCI) [3].*

Leap motion metallic element sensor/controller offers fine-grained gesture recognition and hand trailing that is explained in [4][5]. Married woman Bainbridge, Joseph A. Paradiso explained the new approach for capturing hand gestures that features wearable computing and therefore the associated Human laptop Interface [6]. A survey of Vision primarily based Hand Gesture Recognition ways for human laptop K. Waghamare is given in [12].] Moniruzzaman Bhuiyan and wealthy selecting bestowed paper Gesture-controlled user interfaces, what have we have a tendency to done and what's next? This paper presents trends within the technology of Gesture controlled computer program (GCUI).

GCUI provides realistic opportunities for specific applications and for users WHO square measure uncomfortable with input devices. This paper covers analysis concerning differing kinds of gestures, technology used and therefore the problems self-addressed. GCUI appears applicable for current and future omnipresent devices. Within the early analysis gesture management or recognition method was terribly advanced attributable to the hardware necessities, however currently it's straightforward as straightforward vision technique is employed. Initio gesture recognition was with tough technology like sensors and gloves etc. currently it's easier with digital camera and image process software system. The usually used input devices like keyboard and mouse give the means that of virtual input. Different choices like grasping virtual objects, hand, head or body gesture have become common. Hand movements vary person to person and it acts as command or interface to work machines.

The aim of this paper is to gift a survey on numerous strategies adopted for hand gesture recognition. Moniruzzaman Bhuiyan and wealthy choosing projected GCUI. GCUI provides realistic opportunities for specific applications and for users interaction, as well as the assorted out there approaches is explained by Pragati Garg, Naveen Aggarwal and Sanjeev Sofat [7]. The concept of building a communication system that permits communications between speech-hearing impaired and a standard person is illustrated [8][9]. detector primarily based techniques is employed for gesture detection [10][12]. Sangeetha, K. and Barathi avatar, L. planned the system uses Skin detection formula that detects the colouring of the users whereas giving gestures. the easy methodology is constructing a bar chart of properties. Another methodology is to produce coaching for the neural networks [11].

[1] Voice to Image conversion Acoustic pre-processing, Extraction and Speech Recognition by Shraddha R. Ghorpade, Surendra UN agency square measureuncomfortable with input devices. GCUI appears acceptable for current and future omnipresent devices. within the early analysis gesture management or recognition method was terribly advanced due to the hardware necessities, however currently it's straightforward as straightforward vision technique is employed [1]. The system supportedKL (Karhunen-Loeve) remodel to acknowledge hand gestures by Joyeeta Singha1, Tibeto-Burman language Das [2]. Ondrej Kainz, František Jakab conferred the analysis on typical approaches toward hand trailing and gesture recognition as given in [3]. The vision primarily based surface diagnostic procedure is applicable to several areas of Human laptop Interaction (HCI) [3]. Leap motion metallic element sensor/controller offers fine-grained gesture recognition and hand trailing that is explained in [4][5]. Married woman Bainbridge, Joseph A. Paradiso explained the new approach for capturing hand gestures that features wearable computing and therefore the associated Human laptop Interface [6]. A survey of Vision primarily based Hand Gesture Recognition ways for human laptop K. Waghamare is given in [12].] Moniruzzaman Bhuiyan and wealthy selecting bestowed a paper Gesture-controlled user interface, what have we have a tendency to done and what's next? This paper presents trends within the technology of Gesture controlled computer program (GCUI). GCUI provides realistic opportunities for specific

February 2018.indd 23 2/5/2018 5:55:41 PM

applications and for users WHO square measure uncomfortable with input devices. This paper covers analysis concerning differing kinds of gestures, technology used and therefore the problems self-addressed. GCUI appears applicable for current and future omnipresent devices. Within the early analysis gesture management or recognition method was terribly advanced attributable to the hardware necessities, however currently it's straightforward as straightforward vision technique is employed. ab initio gesture recognition was with tough technology like sensors and gloves etc. currently it's easier with digital camera and image process software system. The usually used input devices like keyboard and mouse give the means that of virtual input. Different choices like grasping virtual objects, hand, head or body gesture have become common. Hand movements vary person to person and it acts as command or interface to work machines.

[2] Joyeeta Singha1, Karen Das projected an inspiration handy Gesture Recognition that is predicated on Karhunen-Loeve remodel during this paper. This paper proposes a system supported KL (Karhunen-Loeve) remodel to acknowledge hand gestures. The system has 5 major steps. It includes skin filtering, palm cropping, edge detection, feature extraction and classification. ab initio the hand is detected mistreatment skin filtering wherever the RGB image is reborn to HSV image and palm cropping is performed to extract solely the palm portion of the hand. Define pictures of the palm is extracted mistreatment cagey Edge Detection technique. Options of the hand were extracted mistreatment KL remodel when palm extraction. Finally the input gesture was recognized mistreatment correct classifier. During this system, the accuracy rate was ninety six and this approach is straightforward to acknowledge totally different hand gestures. The most important blessings of this paper is it eliminates the correlative information, reduces dimensions keeping average sq. error minimum, and it offers wonderful cluster character when the remodel. The disadvantage with this paper is ceaselessly that specialize in hand gestures by manual is extremely tough.

[3] Ondrej Kainz, František Jakab illustrated associate Approach at hand chase and Gesture Recognition supported Depth-Sensing Cameras and electromyogram watching. During this paper, gesture recognition supported leap motion device and surface diagnostic technique (SEMG) is bestowed. The paper includes the procedure to method the depth image data and therefore the electrical activity created by skeletal muscles on forearm. This paper conjointly illustrates the analysis on typical approaches toward hand chase and gesture recognition. The summary of depth sensing cameras with specialize in leap motion device is enclosed. The vision primarily based surface diagnostic technique is applicable to several areas of Human pc Interaction (HCI). Visual image detection of hand movement and chase of diagnostic technique (EMG) signals is monitored. Classification of specific patterns for SEMG watching is completed by utilization of neural networks. To acknowledge gestures, the leap motion device is employed however there square measure some problems with continuous finger chase. This technique overcomes the shortcomings of visual primarily based system wherever chase of hand isn't continually recognized thanks toobstacles – e.g. hand coated by hand. Main downside once mistreatment this device is demand for a user to stayclose to the interaction zone, i.e. close to pc. Nature of gesture may be not solely static however conjointly dynamic, thus it's tough to note.

[4] Jihyun Han dynasty, Nicolas Gold conferred Lessons Learned in Exploring the Leap Motion metallic elementdevice for Gesture supported Instrument style. during this paper, leap motion metallic element device offers fine-grained gesture recognition and hand pursuit area unit explained. This paper presents management of musical instruments victimization the Leap Motion metallic element device. 2 variants of a virtual instrument were presented: Air-Keys and Air-Pads, each victimization the Leap Motion metallic element device for management. This paper has conferred shopper level gesture detection technology as suggests that for making and dominant instruments. The system uses device for augmenting instruments. The Microsoft Kinect metallic element device spawned several new applications in musical management. Instruments supported Leap Motion metallic element area unit rising in video. The new development as well as motion capture and glove based mostly system is reported during this paper. the arrival of gesture capture devices like Kinect metallic element had diode to more developments within the space. The system used visual and tactile feedback to help the entertainer in locating their hand position and provides ability to live latency. This paper evaluates platform from the attitude of instruments victimization touch-less action. The matteris that the strategies area unit referred to as on each frame thus messages area unit triggered for each consecutive frame once the fingers touched quicker than a particular speed.

[5] actress Ellen Potter, Jake Araullo, Lewis Carter projected a plan on leap motion controller in paper Leap Motion controller: A read on signing. This paper presents the employment of leap motion controller for recognizing Australian signing (AUSLAN). The controller is in a position to

February 2018.indd 24 2/5/2018 5:55:41 PM

supply correct pursuit of hand and finger movements. This detection loses accuracy once the hand moves into a foothold that obstructs controller's ability to look at, for example: once the hand rotates and is perpendicular to the controller and once the hands area unit brought along like finger to finger. Leap motion controller, alittle device which will be connected to the pc victimization USB. This paper includes the functionalities of the leap motion controller. This method principally aims at manufacturing associate application that acknowledges Auslan signs. This paper mainl focuses on evaluating the Leap Motion controller for its ability to acknowledge signs that area unit created within the field of read of the controller. The system gets trained to acknowledge and determine signs. this method incontestable a high level of recognition: ninety four for signers that had been trained with the system and eighty fifth for non-trained signers. the downside with this paper is that, the controller is used with important work for recognition of basic signs, but it's not applicable for advanced signs, particularly people who need important face or body contact. The Leap Motion controller has issue maintaining accuracy and fidelity of detection once the hands don't have direct line of sight with the controller.

[6] Wife Bainbridge, Joseph A. Paradiso explicit Wireless Hand Gesture Capture Through wearable Passive Tag Sensing. This paper illustrates the primary implementation of associate HCI finger gesture pursuit system engineeredfrom passive RFID tags. This paper explains a brand new approach for capturing hand gestures. Wearable computing and also the associated Human pc Interface should move aside today's keyboard, keypad, bit screen and differentlarge handheld interfaces used. The technologies enable user to specify the input through their fingers. This paper principally focuses on the work of passive RFID device tags to trace hand gestures. This method consists of associate ultra-high frequency (UHF) reader and little finger-worn tags with form of sensors to sight gestures. The tags area unit steam-powered by transmit RF energy. This paper primarily introduces the hardware and example for the computer program implementations like mouse scrolled by hand, a click by a mouse, coming into input by touching fingers and mapping management onto motion of hand, arm and fingers. the first physical goals of the system is to trace explicit hand movements that would be accustomed management a wearable pc or facilitate in interaction with present and different wearable devices. the system was absolutely practical, however as this can be associate early implementation, it had been still considerably restricted by power required by the tiny passive tags, antenna potency, and also the size of passive tags. Suggests that of scaling to lower power and smaller size area uniturged.

[7] Pragati Garg, Naveen Aggarwal and Sanjeev Sofat explained Vision based mostly Hand Gesture Recognition during this paper. The aim of this paper is to gift a survey of Vision based mostly Hand Gesture Recognition techniques for human pc interaction, as well as the assorted approaches. With present computing, user interaction approaches like keyboard, mouse and pen appears low. Use of hand as associate data input device is a hornytechnique for Human pc Interaction that has evolved from text-based interfaces to totally fledged multi-participant Virtual surroundings (VE) systems. This paper additionally illustrates the human-computer interaction of the future: A 3Dapplication wherever you'll move and rotate objects by the action of moving and rotating your hand while nottouching any data input device. the present approaches area unit categorised into 3D model {based|based mostly|primarily based mostly} approaches and look based approaches, explaining their benefits and shortcomings and distinctive the problems. Within the Vision based mostly strategies, solely a camera is employed while notthe employment of any further devices. This poses a difficult drawback because the systems got to be background invariant, lighting i.e. illumination insensitive, person and camera freelance to realize sensible performance. Moreover, such systems ought to be optimized so as to fulfill the wants, as well as accuracy. This paper is meant to illustrate the assorted open analysis problems similarly as act as a start line for anyone inquisitive about victimization hand gesture recognition in their interfaces. The paper highlights the assorted aspects of hand posture and gesture recognition technology. And it discusses the out there algorithms for hand posture and gesture recognition. The most disadvantages with this approach is presumptuous high distinction stationary backgrounds and close lighting conditions. Another issue long-faced is recognizing the temporal begin and finish points of purposeful gestures from continuous motion of the hand. This drawback is usually brought up as "gesture spotting".

[8] Shraddha R. Ghorpade, Prof. Surendra K. Waghmare provides a thought of Communication System for Deaf and Dumb folks. This paper presents the communication system for hearing-speech impaired folks. folks with such disabilities area unit finding it arduous to cope up with the quick growing technology. Deaf and dumb folks use signing for communication however they're restricted to their own world. Signing is associate communicative means for communication between traditional and dumb folks (information principally sent through the hand gesture). Only a few those that aren't themselves deaf and dumb ever learn to sign.

February 2018.indd 25 2/5/2018 5:55:41 PM

Technology is a way to get rid of this hindrance and profit these folks by facultative them to speak with others WHO don't perceive signing. So, there's a desire for translator to grasp what they speak and communicate with North American nation. The signing translation system coverts the traditional signing to speech or text and thence makes the communication between normal person and dumb folks easier. So, the concept is to create a system that allows communication between speech-hearing impaired and a standard person. This method includes image acquisition, pre-processing, feature extraction and classification procedures. This paper details Hand gesture recognition that provides associate intelligent and natural means of human pc interaction (HCI). Hand gesture recognition could be a section in applied science and language technology that aims in explaining human gestures via mathematical algorithms. With gesture recognition its potential for humans to act with machines whiles not the assistance of any devices. The downside with this paper is that, the character of the background, existence of different objects (occlusion), and illumination should be thought of.

[9] Koli P.B., Chaudhari Ashwini, Malkar Sonam , Pawale Kavita & Tayde Amrapali projected Image process based mostly Language device for Deaf and Dumb folks. This paper introduces a language device for deaf and dumb folks. Today it's a difficult for the deaf and dumb folks to speak with the traditional folks. As a result of the traditionalfolks don't recognize the signing utilized by the impaired folks. The signing is difficult to grasp and learn. Thus it'snot possible for the person to specific their thoughts with these deaf and dumb folks. Thus here could be a system during which a camera is placed before the impaired person. That person ought to wear the colored rings in his finger. If the person makes the gestures, it's captured by the camera. The image captured is RGB image. Within the next section it's born-again into binary image. In binary conversion RGB to gray so gray to binary image conversion takes place. Then by victimization the color recognition technique identifies the coordinates of the color. When this the coordinate's area unit mapped with the photographs that area unit already hold on within the info and actual image for that gesture is captured. Like this deaf and dumb folks can communicate. Limitations with this method area unit, the colored rings ought to be worn by the impaired person then solely he will build gestures. These aren't cheap by the poor folks. And also the methodology used here is advanced. The captured original image is born-again into binary image so mask image when thresholding.

[10] Kshirasagar Snehal P., Shaikh prophet Hussain, Malge Swati S., Gholap Shraddha S., Mr. Swapnil Tambatkar conferred a paper on Gesture Vocalizer For Deaf And Dumb. This paper discusses the means during which the poor folks can communicate with the traditional folks. Those folks use the signing (i.e.) gestures for communication. Here the gesture recognition is employed to spot the hand movements. Gesture recognition is classed into vision based mostly and device based. During this system device based mostly techniques area unit used. as a result of the vision based mostly techniques have drawbacks as advanced algorithms, lighting conditions and background. Device based mostly techniques have advantage as bigger quality. 2 sensors area unitused here like Flex sensors and measuring system that area unit mounted on gloves. Flex sensors square accustomed measure the positions and bending of the fingers. Measuring system is employed to livestatic and dynamic acceleration. during this system AVR microcontroller and speech synthesizer is employed. AVR microcontroller could be a low power C-MOS image is captured by the cameras and supported degree of bending of fingers the flex sensors resistance changes. The output are going to be analog values born-again into digital values. Then it's compared with the hold on prices and also the nearest value is chosen and displayed. The most disadvantage during this paper is, to implement this method Flex sensors, measuring system, AVR microcontroller and Speech synthesizer area unit required. Thus it's not cheap by the poor folks.

[11] Sangeetha, K. and Barathi avatar, L. explicit Gesture Detection For Deaf And Dumb folks. This paper introduces associate mechanical man application to cut back the barrier within the communication between deaf & dumb and traditional folks. This method uses the Skin detection algorithmic program that detects the skin color of the users whereas giving gestures. This method eliminates the employment of different devices like measuring system, gyroand gloves. Here the sign languages given by the impaired person area unit captured by the camera so the image is compared with the photographs hold on within the info and corresponding text or voice is generated. The skin detection techniques area unit used along side different strategies like Hough transforms or Manfred Eigen price based mostly approaches. This can be used for the pattern detection. Many strategies area unit accustomed produce the skin detector. The easy technique is constructing a bar graph of properties. Another technique is to supply coaching for the neural networks. The bar graph will store solely the properties of the pixels however the neural networks will predict the skin tone. Some tools are out there to classify the skin. Limitation with this paper is, during this system there's no predefined threshold set of skin

February 2018.indd 26 2/5/2018 5:55:41 PM

likelihood values. Thus it'll result in a tangle in skin detection. Another issue is usually the background is analogous to the skin color. It should be detected as skin.

[12] Shraddha R. Ghorpade, Surendra K. Waghamare introduced Full Duplex Communication System For Deaf & Dumb folks. This paper introduces associate electronic system referred to as artificial speaking mouth. The folks disabilities cannot able to cope up with the inheritable technologies. Particularly the deaf and dumb folks feel arduous to speakwith the traditional folks. They use signing to speak. However the signing isn't a effective means. Therefore thesolely means is convert the signing into voice signals. Thus the new technique referred to as artificial speaking mouth is projected. During this technique flex sensors plays a significant role. It modification the resistance supported the bending of the fingers. When the image is captured it's born-again into voice through 3 steps like Extraction technique, options extraction and Gestures classification. Within the extraction technique the photographs area unit solely divided if it's static gesture otherwise it's settled and caterpillar-tracked. This can be the method of dividing input image.

3. Proposed System

The main objective is to establish the communication process between Deaf & Dumb and the normal person. It is two way communications. The idea is to deploy an Android based application where by user will provide gestures & recognized by the server and the corresponding voice is initiated to communicate with the normal person. Normal person can speak out voice is recognized and corresponding image is displayed to the impaired person so that this application can be implemented from the both the end. Our application may use the camera of the phone to detect the sign shown by the deaf or dumb people and process it to either text or voice output which will help other people to understand the deaf and dumb people with ease. Our aim is to bridge the communication gap between the deaf and dumb people with the world (other people). The proposed procedure was implemented and tested with set of images. Image acquisition is the process to capture the hand gesture images which represents different signs. The resolution of image capturing devices will not be the same. This results in varying resolution of the captured images. For accurate comparison of the features of images and to reduce the effort needed for processing all the images should be scaled to a uniform size. Thus the images for training and testing are captured in a white background with web camera and database is created. The set of images of single person is used for training database; the sample database is taken. The hand gesture is pre-processed using background subtraction algorithm. The pre-processed gesture is taken as input for feature extraction and classification stage. Once the gesture is recognised; the equivalent gesture audio file is played at the output. Similarly, the voice input given by the end user is recognized and converted to text by calling built in functions. The text is then compared and corresponding image is displayed to the impaired person. Hence, the system enables two way communications between deaf and dumb and normal person.

4. Conclusion

The main aim of our project is to provide a helping hand to the deaf and dumb people. Normal human beings too find it difficult to communicate our feelings with them, this application will help us to share our views with them. This paper proposes review on various simple and efficient sign gesture recognition system to eradicate the communication gaps between the deaf and dumb people with normal people. Our method doesn't require any gloves and devices for detecting the hand movements, owing to this cost of the system is reduced by great margin. The main aim of this system is to increase the accuracy rate, efficiency, decrease the hardware equipments and make the system affordable to people. This paper adopts background subtraction algorithm. This system can be used by people by spending less cost when compared to the existing system, since the latter requires additional devices to detect hand gestures. There are reasons for poor performance of testing data. The huge variation in images also played a part in low accuracy rates. This includes different sizes, background, orientations and angle of gestures, etc.

References

[1] Bhuiyan, Moniruzzaman, and Rich Picking, "Gesture-controlled user interfaces, what have we done and what's next.", Proceedings of the Fifth Collaborative Research Symposium on Security, E-Learning, Internet and Networking (SEIN 2009), Darmstadt, Germany. 2009.

[2] Singha, Joyeeta, and Karen Das, "Hand gesture recognition based on "Karhunen-Loeve transform.", Mobile & Embedded Technology International Conference, pp. 365-371, 2013.

[3] Kainz, Ondrej, and František Jakab, "Approach to Hand Tracking and Gesture Recognition Based on Depth-Sensing Cameras and EMG Monitoring.", Acta Informatica Pragensia, Vol:3, No:1, pp.104–112, 2014.

[4] Han, Jihyun, and Nicolas Gold, "Lessons Learned in Exploring the Leap Motion TM Sensor for Gesture-based Instrument Design", "Proceedings of the International Conference on New Interfaces for Musical Expression, pp. 371-374, July 2014.

[5] Potter, L. E., Araullo, J., & Carter, L. "The leap motion controller: a view on sign language", Proceedings of the 25th Australian Computer-Human Interaction Conference:Augmentation, Application, Innovation, Collaboration, pp. 175-178. ACM, 2013

[6] Rachel Bainbridge,Joseph A. Paradiso, "Wireless Hand Gesture Capture Through Wearable Passive Tag Sensing", International Conference on Body Sensor Networks, pp. 200-204, IEEE 2011.

[7] Pragati Garg, Naveen Aggarwal and Sanjeev Sofat, "Vision Based Hand Gesture Recognition", International Journal of Computer, Electrical, Automation, Control and Information Engineering Vol:3, No:1, pp. 186-191, 2009.

[8] Shraddha R. Ghorpade, Prof. Surendra K. Waghmare,"A Communication System for Deaf and Dumb People", International Journal of Science and Research (IJSR) Volume: 4, Issue: 9, pp. 1914-1917 September 2015.

[9] Koli P.B., Chaudhari Ashwini, Malkar

February 2018.indd 27 2/5/2018 5:55:41 PM

Sonam, Pawale Kavita & Tayde Amrapali, "Image Processing Based Language Converter for Deaf and Dumb People", IOSR Journal of Electronics and Communication Engineering (IOSR-JECE), e-ISSN: 2278-2834, p-ISSN: 2278-8735 pp.25-30, 2015

[10] Kshirasagar Snehal P., Shaikh Mohammad Hussain, Malge Swati S., Gholap Shraddha S.,Mr. Swapnil Tambatkar, "Gesture vocalizer for deaf and dumb", International Journal Of Engineering Sciences & Research Technology, Volume: 5, Issue: 4, pp. 547-553, 2016.

[11] Sangeetha, K. And Barathi Krishna, L., "Gesture Detection for Deaf and Dumb People", International Journal of Development Research, Vol. 4, Issue, 3, pp. 749-752, March, 2014.

[12] Shraddha R. Ghorpade1, Surendra K.Waghamare, "Full duplex communication system for deaf and dumb people", International Journal of Emerging Technology and Advanced Engineering, Volume: 5, Issue: 5, pp. 224-227, 2015.

■

About the Authors

Dr. S.Hemalatha did her Bachelor degree in Computer Science and Engineering and Masters in Computer Science and Engineering University of Madras and Anna University Chennai , India in 2000, and 2004 respectively. She completed PhD in Computer Science and Engineering from Anna University, India in 2016. She has totally 17 years of experience in different engineering colleges and currently she is working as a Professor in Computer Science and Engineering in Panimalar Institute of Technology, Chennai, India. Shea has published 45 national and International journals and 30 papers in international conferences. She is the recipient of Distinguished Professor award from CSI Mumbai chapter in the year 2016, Professional achiever award from IEEE madras section in the year 2017. She has guided many UG and PG level projects and being a committee member for Research Scholars in Anna University. Her research areas are Network Security, Mobile Communication and Mobile Application Development. She is a member of CSI, IEEE and ISTE.

M. Archana is a final year Computer Science and Engineering student in Panimalar Institute of Technology. She is a member of Computer Society of India.

S. Haripriya is a final year Computer Science and Engineering student in Panimalar Institute of Technology. She is a member of Computer Society of India.

E. Shiny Freeda is a final year Computer Science and Engineering student in Panimalar Institute of Technology. She is a member of Computer Society of India.

ICSCAAIT-2018

International Conference on "Sustainable Computing and Advanced Artificial Intelligence Techniques (ICSCAAIT-2018)" held during during 24th & 25th January 2018 jointly organized by CSI Noida Chapter & LINGAYAS UNIVERSITY Delhi NCR . Prof. (Dr.)SK Kak Chief Guest(Chairman AICTE NWRO and Ex VC UPTU & CCS Univ Meerut) addressed the delegates during the inaugural ceremony on 24th January 2018.

February 2018.indd 28 2/5/2018 5:55:42 PM

A framework for monitoring healthcare of University students using Machine Learning and Data Analytics

Pawan Kumar
Lovely Professional University, India. Email: pawan.11522@lpu.co.in

Advances in computer technologies like data analytics and machine learning, is enabling us to improve systems or processes characterized by large volume of data. This work proposes a framework to monitor and analyse the quality of healthcare services being provided at educational institutes in India. Additionally, the proposed framework will result in an ever-growing database of medical records that can be useful for government, regulatory bodies and researchers in the healthcare industry. The larger objective is to have a collection of medical records of students at colleges or universities across India and subsequently extract medical knowledge from this database.

Keywords : Student health, predictive analytics, data mining, machine learning, ICT

1. Introduction

We are in a time when advancements in computing and internet technologies are affecting almost every aspect of our life. The techniques like machine learning, data analytics and big data are enabling maintaining and analysing huge volume of data and adding value to decision making. These techniques are particularly suitable where amount of information is large enough to go beyond human capabilities to analyse. According to a study, India has the largest student population in the world [1]. So, it is worth exploring how advancements in computer technologies can help in monitoring health of students and quality of healthcare services that are primarily meant for youth.

This paper proposes to develop a framework for compiling and analysing medical records of students' visit to healthcare facilities at their respective colleges or universities. The significance of such a system lies in the fact that it can help identify type and frequency of health issues being faced by students at colleges or universities across India. This can be useful for the government in framing healthcare related strategies.

2. Proposed System

Most of the colleges and universities own a facility to take care of health issues of their students and staff. Taking example of Lovely Professional University, Punjab, India, there is a full-fledged medical facility available named as Uni-Hospital. This hospital provides different facilities like OPD, Pathological tests and Physiotherapy also. The staff at the hospital has been provided a software module, where details are recorded for student visits [2]. Panjab University, Chandigarh, has Bhai Ghanaiya Ji Institute of Health that caters to the medical needs of around 50000 population students. It has facilities like clinical laboratory, ECG machine and X-ray machine [3]. DR B. R. Amebdkar NIT Jalandhar also has a dispensary facility equipped with a pathological lab [4]. Sant Longowal Institute, Sangrur also has a similar healthcare centre [5].

Although, most of the educational institutes provide basic healthcare facilities, however, at present, most of these maintain offline records of the patients. Also, the prescription suggested by the doctor is not recorded in a software facility. The idea is to maintain a database of records of students visiting healthcare facility at their respective college or university. The items of interest include student details like Gender, age, State, Problem and prescription given by the doctor. This will be an active database that will be live and growing every day. The proposed framework will provide an interface to every college to upload records of student visits at their respective health centres. On this database, data mining and machine learning techniques can be applied to extract hidden patterns. Also, the proposed system will be able to generate different kinds of reports as per the needs of the different stake holders.

Student Visit...Colleges Uploading data....Resulting in a big database... Machine Learning & Big Data....Medical Knowledge/Pattern extraction..... Decision Making.

The input to the system will be details of the students visiting healthcare centres at their respective colleges or universities. This data will be uploaded to a central server to develop an integrated database. This database will be growing on daily basis with new records being added to it. Using data analytics and machine learning algorithms, models will be

February 2018.indd 29 2/5/2018 5:55:43 PM

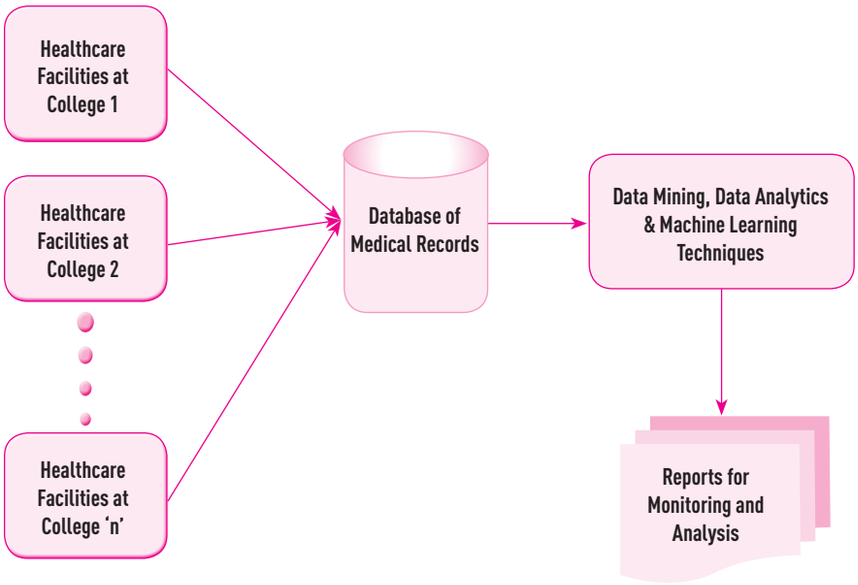


Fig. 1 : Block Diagram of working of the proposed system

developed to extract patterns of interest and subsequently help in healthcare prediction and corrective strategies.

3. Outcomes of the proposed system

The proposed system has the potential of offering value addition through the following outcomes:

3.1 Database of medical records:

The proposed system will gradually result into a large database of medical records. This database will be useful for different stakeholders be it university herself or government bodies. Using data mining and data analytics techniques, we can come to know about different types of health issues and their frequencies prevailing among college going youth in India. This data can be further analysed in term of gender, geographical regions or programmes of study.

3.2 Research Opportunities:

Being a live database of Indian subjects, that too from a specific population i.e. college going students, this database will be helpful for researchers from computing and medical field. They can apply their research hypothesis or algorithms on Indian subjects and hence publish results that can be of specific interest to India. This will be particularly useful for student-centric healthcare studies. Otherwise, they are to apply their algorithms on medical records from public hospitals that consist of records of all ages and of people with different life style and standard of living.

3.3 Extraction of Medical Knowledge:

One of the major concerns in healthcare research is that prescription advised by doctors to patients is in hard copy at most places. These prescriptions if made online and mapped to the response of recovery of the patient, will result in structured information in huge volume. The proposed system has potential to contribute in this direction as it records not only student details but also the prescription given by the doctors. Using artificial intelligence techniques, medical knowledge can be extracted from these records. This can provide new insights to decision makers from the medical regulatory bodies.

3.4 Foreseeing the future:

Applying predictive analytics the proposed system can forecast health issues that students may encounter due to seasonal or geographical reasons. Predictive analytics refers to the computing techniques that allow predicting the values of variables by learning a model from the existing examples. Knowing the type and frequency of health issues, efforts can be made to handle the likely diseases that this population may face when they transit from youth population to old population. Consequently, relevant awareness programmes can be started and medical precautionary practices be put in place.

3.5 Programme of study versus health:

Using data mining techniques, we can analyse is there any impact of programme of study on the type of health issues that a student is likely to encounter. For example, the frequency and type of health issues reported by engineering students may be significantly different from those reported by students studying social sciences or healthcare.

3.6 Scope for Self-learning algorithms:

Since this system is maintaining a live database, there is a scope of implementing machine learning algorithms that can update themselves regularly as new unseen data is fed into the database. This is likely to increase the performance and reliability of these algorithms.

3.7 For regulatory authorities:

This system will be able to generate reports that can be used by regulatory bodies like UGC or MCI or Health ministry for monitoring on quality of healthcare services being provided by the colleges or universities. Consequently, necessary instructions or corrective actions can be carried out.

3.8 Impact on the College or University:

Once online, the quality of healthcare facilities is likely to become an important parameter for students and their parents, in opting for a particular college at the time of admission. This is likely to encourage colleges to upgrade their healthcare infrastructure. Consequently, it will be beneficial to the students.

4. Expected challenges and possible remedies

Not all the colleges and universities across India can boast of good health care facilities. Such colleges will hesitate to be a part of this project. The management of these colleges must be instructed to upgrade their facilities within a timeline. Also, a provision of financial assistance can be made by

February 2018.indd 30 2/5/2018 5:55:43 PM

the government. Most of the healthcare facilities at educational institutes follow manual processes. So, the medical staff at the healthcare facilities will need training. The required training programs can be implemented through trained students.

5. References

[1] http://timesofindia.indiatimes.com/india/At-315-million-India-has-the-most-students-in-world/articleshow/37669667.cms

[2] http://www.lpu.in/student_services/healthcare.php

[3] http://bgjih.puchd.ac.in/

[4] http://www.nitj.ac.in/index.php/nitj_cinfo/index/21

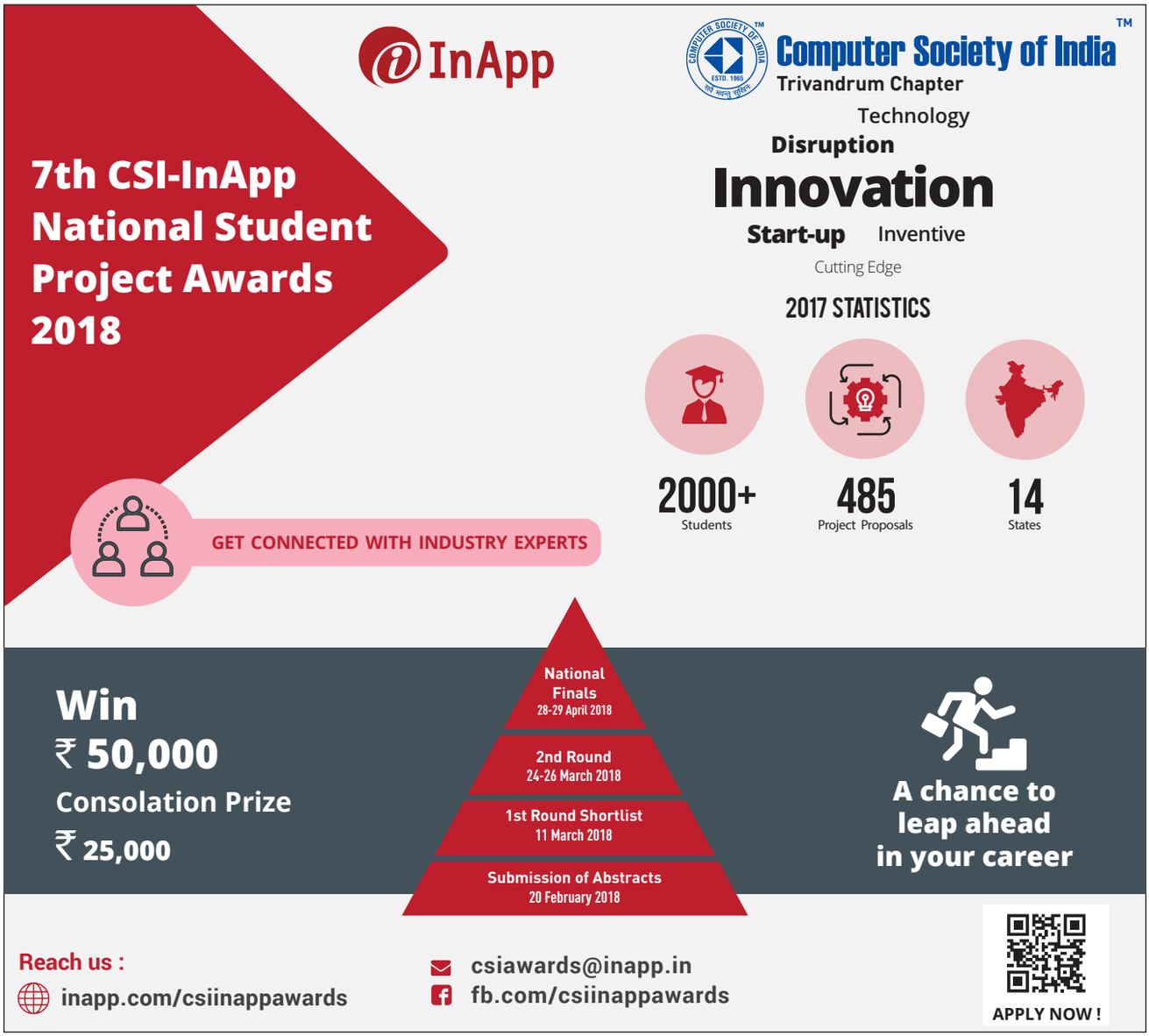
[5] http://sliet.ac.in/health-center/

■

About the Author

Pawan Kumar (CSI Membership No: 1183942) Assistant Professor, School of Computer Applications, Lovely Professional University, Punjab

Research Interests: Machine learning, Image processing, ICT Applications in health care. He have more than 12 years of experience in academics



February 2018.indd 31 2/5/2018 5:55:44 PM

Pattern recognition using Neural Network: A basic introduction

S K Gupta
Dept. of Computer Science and Engineering, UIT, Burdwan University, Burdwan, West Bengal, India, E-mail: skgupta.81@gmail.com

A Mukherjee
Associate, Cognizant, Kolkata, West Bengal, India
E-mail:mukherjeeayan16@gmail.com

Pattern Recognition is a fast-paced developing field of machine learning that needs to learn, adapt and develop new ways to identify patterns. Neural Network provides the same required features. Statistical Method needs to be combined with various other methods for pattern sensing and it is time consuming for counting all the samples. Structural Methods are not effective for Color Patterns, Intensity Patterns and the structural methods also does not evolve with the previous outputs i.e. its static and does not covers the previous outputs to purify the pattern recognition system for next input. Neural network provides solution for all the above disadvantages. All the Major sectors like Space Science, Metrological Science and various other fields of Science and Technology depends on Pattern Recognition like Fingerprints or Robotics. The faster the approach the better the valuation for the technology. Neural network also provides an option to integrate other approaches with neural network to gain finer results. The intent of this chapter is to recapitulate, equate some of the renowned approaches used in numerous phases of a pattern recognition structure and to streamline the users with the details of neural network its methodology and examples.

Keywords : Pattern Recognition, patterns, pattern sensing, Neural network.

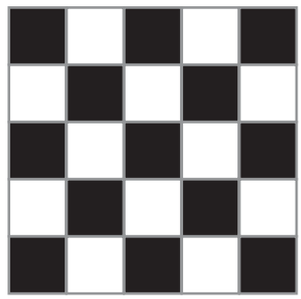
Introduction

As famously quoted by Margaret Heffernan stating that the prodigious covenant of creativeness is about pattern recognition, and what is compulsion to determine patterns is heaps of data. A mind accumulates that data by captivating memo of unsystematic minutiae and anomalies effortlessly seen every day: flukes and changes that, collectively, provides the insights.

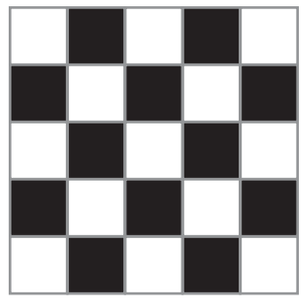
Pattern Recognition is a field of Computer Application's Machine Learning, which focuses on identifying the Patterns i.e. the Similarities and Regularities that is reputable within a collection of Objects. For E.g.: - Image Patterns, Speech or Sound Pattern

Now we can define Pattern as is opposite of a Chaos, i.e. any Entity or Object or Sequence which is vaguely defined and that could be given name and used as an identifying character [1].

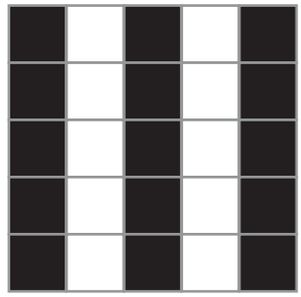
Duda & Hart states as - The Consignment of a corporeal object is proportional to occurrence to one of



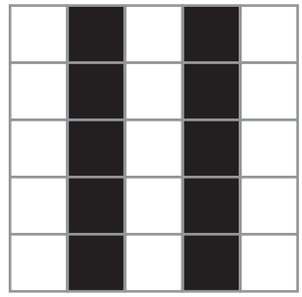
Pattern type 1, version 1



Pattern type 1, version 2



Pattern type 2, version 1



Pattern type 2, version 2

Fig. 1 : Example of Pattern

February 2018.indd 32 2/5/2018 5:55:45 PM

the numerous pre-specified categories. Pattern recognition is used for Machine Learning. There are various phases for Pattern Recognition [2]: -

A. Data Acquisition and Sensing

In this phase, various operations are used to extract the data that will be used as baseline parameters for classification of the required elements. This phase is used for the measurement of physical variables like Size, Resolution etc.

B. Pre-Processing or Segmentation

This phase is used to remove the unwanted elements from the required elements like noise etc. This phase is also used to remove the elements like background and other physical variables that are not the element of interest.

C. Feature Extraction

In this phase, the element of interest is inspected and the features are extracted from the element for further processing and classification in next phases.

D. Classification

This phase uses the output of 1st phase and 3rd phase as input i.e. uses the 1st phases output as base element for the Classifiers and the output of 3rd phase is used as input of the element to be classified.

In this phase, the element is classified and placed into a given set of pattern.

E. Post-Processing

This phase in a numeric phase that calculates and evaluates the correctness of the classification of element into a particular set of pattern

Patter Recognition is used in many practical applications. Like:

A. Optical Character Recognition: - Handwritten: Sorting Letters by postal codes, Input devices for PDAs
B. Diagnostic Systems: - Medical Diagnosis like ECG etc.
C. Biometrics: - Face Recognition, Fingerprint Recognition, Retina Scan etc.
D. Military Application: - Automated Target recognition like Laser Guided Bombs and Image Segmentation and analysis from Aerial Snapshots

Classification phase of Pattern Recognition

There are broadly four types of Approach for Classification phase of Pattern Recognition: -

i. Template Matching
ii. Statistical Approach
iii. Syntactic and Structural Approach
iv. Neural Network Approach

1. Template Matching Approach

This is oldest and the easiest method of classification of pattern according to the classifiers. In this approach, a template or prototype is available and the pattern recognition occurs by determining the similarity between two entities like point, shapes or curves. This similarity check is done at various angles and intensity like after rotation or scale change [2].

Pros of Template Matching

Fast method and does not requires much of computation.

Cons of Template Matching

i. Works on a very confined domains only
ii. Fails if the image is distorted due to image processing, viewpoint change etc.

2. Statistical Approach

It is an approach to pattern recognition grounded on statistical molding of data. It also depends on probabilistic model and decision model to get an algorithm. The efficiency of Statistical approach solely depends on the feature being selected as classifiers.

In this methodology, pattern is characterized in terms of a character. This pattern is regarded as a point in a-dimensional planetary. Primary aim is to select those characters that can differentiate between patterns with higher uniqueness and less similarity I.e. more compactness and disjoint sections in the a-dimensional feature space.

In Statistical methodological tactic, the decision frontiers are resolute by the probabilistic scatterings of the pattern of each class, which necessarily be either itemized or learned.

Another approach in statistical methodology is discriminate exploration based approach in which a parametric form of the decision boundary needs to be specified as first step; then the required decision limit of the specified form is recognized based on the cataloging of pre-processing or training patterns. There are various ways to construct such boundaries like a Mean Squared Error Criterion.

As per Vapnik's Philosophy, If someone retain a constrained quantity of information for explaining some problem, try to resolve the delinquent directly and never decipher a more broad-spectrum problem as an transitional step. It might be possible that the existing information is ample for a direct resolution but is deficient for cracking a more wide-ranging intermediate problem.

Pros of Statistical Method

It has high accuracy because it generally combines with other methods

Cons of Statistical Method

i. Time Consuming as it checks all the combinations i.e. counting samples
ii. Does not evolve with the previous input.

Let us see the comparison between Statistical Approach and Template Matching [5]

Approach	Representation	Recognition Function	Typical Condition
Template Matching	Samples, curves, pixels	Distance Measure, Correlation	Classification Error
Statistical	Features	Discriminant Function	Classification Error

Table 1: Comparison between Template Matching and Statistical Approach

February 2018.indd 33 2/5/2018 5:55:45 PM

Let us see the comparison between Syntactical Approach and Template Matching

Approach	Representation	Recognition Function	Typical Condition
Template Matching	Samples, curves, pixels	Distance Measure, Correlation	Classification Error
Syntactic or Structural	Primitives	Grammar, rules	Acceptance Error

Table 2: Comparison between Template Matching and Syntactical Approach

3. Syntactic and Structural Approach

In this methodology, the object can be characterized by a variable-cardinality set of representative, nominal feature. This provides an option to represent any pattern structure with more complex inter-relationships between attributes than fixed dimensionality using Statistical Approach. In this approach, patterns are considered as composition of sub patterns, which in turn comprises of further simpler sub patterns. The elementary or the base unit i.e. the simplest pattern is called as Primitives where in the complex patterns are considered as inter-relationship of these primitives. A simple and formal resemblance is haggard between the structures of the pattern and the grammar of the language where in patterns characterizes the sentences in the language with primitives as alphabets. The Sentences are created according to a Grammar. Grammar is mostly generated from the available training samples that are provided in Data Acquisition and Sensing phase [5].

In Syntactical approach, the classification does not only refer as which class it refers to i.e. the resultant output is not only that "Yes, it belongs to this class" or "No, it doesn't belong to this class", but also this methodology provides an explanation on derivation of given pattern from the given primitives. This archetype is being used with the patterns with definite structures that can be described with set of rules, textured images etc.

Syntactical method can also be implemented using Graphs. In this, the nodes are connected when the corresponding sub patterns are related. An item belongs to the particular class when the graph is isomorphic with prototype graphs. As Marshall McLuhan says: "When information overload occurs, pattern recognition is how we determine the truth". For e.g.: - Fields like Medical Industry i.e. ECG uses Syntactical Approach. ECG waveforms are represented with mesh of vertical and horizontal lines where in the normal or unhealthy waveforms can be identified with formal grammars. The identification of heart condition is by first describing it in term of the basic line segments and then trying to parse the descriptions according to the grammars.

Pros of Statistical Method

i. Can be used for complex structures
ii. Less time consuming and appropriate for solutions like Character sets.

Cons of Statistical Method

i. Not able to classify patterns related to Colors, intensity etc.
ii. Segmentation of patterns and derivation of grammar is a hefty and time taking process in case of Noisy patterns.

4. Neural Network Approach

Neural network Approach can be termed as hefty cluster of parallel computational system with bulky number of simpler processors with n-number of inter-connections. In this approach, the classifiers can be best described as network of cells modeling neurons of human brain. This implemented using feed-forward neural system has been skilled accordingly. Neural Network model uses principles like Learning, simplification, adaptively, distributed representation, fault tolerance and reckoning in a mesh of weighted graphs. In the graphs, the nodes are the simulated neurons and focused weighted edged are networks between the neuron productivities and neuron responses [5].

The key characteristic of neural networks is they consume the capability to learn multifarious non-linear associations, use chronological teaching techniques and adapt to the data.

Mostly in neural network methodology, one approach as the feed-forward system is implemented for pattern cataloguing tasks. It embraces multi-layer perceptron and Radial-Basis Function (RBF) networks. These grids form a hierarchy of layers having unidirectional networks amongst the layers [5].

Another prevalent grid methodology is the Kohonen-Network also known as Self Organizing Map (SOM). It is primarily used for data bundling and feature plotting. The knowledge procedure includes modernizing of grid architecture and linking weights for the network efficiency in performing a specific classification/clustering chore.

It has squat dependency on realm precise information (comparative to model-based and rule-based methodologies) and due to the accessibility of proficient learning algorithms for practitioner's usage; it

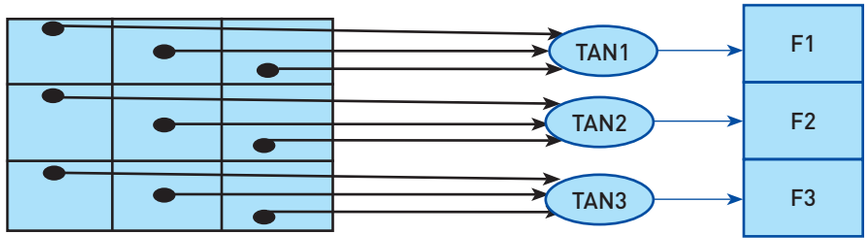


Fig. 2 : Feed-Forward Neural Network

February 2018.indd 34 2/5/2018 5:55:45 PM

Approach	Representation	Recognition Function	Typical Condition
Template Matching	Samples, curves, pixels	Distance Measure, Correlation	Classification Error
Neural Networks	Samples, curves, pixels	Network functions	Mean Square Error

Table 3: Comparison between Template Matching and Neural Network Approach

is widely popular pattern recognition approach.

This approach can implement non-linear algorithms for feature abstraction and cataloging like multilayer perceptron. Prevailing feature abstraction and cataloguing algorithms can also be charted on neural system architectures for higher efficiency.

Neural networks do bid numerous benefits such as, integrated tactics for feature abstraction, classification and malleable methods for finding decent, discreetly nonlinear resolutions.

4.1 Advantages of Neural Network

i. Adaptive Erudition: A knack to acquire knowledge to perform chores grounded on the data given for teaching or preliminary experience.
ii. Auto-Organization: An Advanced Neural Network can craft its own association or depiction of the statistics it acquires throughout learning period.
iii. Real Time Maneuver: Advanced Neural Network reckonings may be performed in analogous, and distinct hardware are being considered and contrived to take advantage of this competency.
iv. Error Lenience via Repetitive Information Coding: Fractional annihilation of a system leads to the equivalent deprivation of Functionality. Though, few system competences may be retained even with utmost network damage.

4.2 Details on Neural Network Approach

There are diverse forms of Neural networks [5]: -

i. Feed-Forward Methodology
ii. Feedback Network Methodology
iii. Network Layers
iv. Perceptron

I. **Feed-forward networks:** Feed forward Artificial Neural Networks permits the unidirectional signals travel only; from feedback to response. There is no feedback (loops) i.e. the resultant response of any layer does not affect that identical layer. Feed Forward Artificial Neural Networks incline to be forthright networks that subordinate inputs with responses. They are comprehensively utilized in pattern acknowledgement. This type of association is also denoted to as bottom-up or top-down.

II. **Feed-back networks:** Feed-back systems can devour Bi-Directional signals i.e. itinerant in both directions by induction of loops in the grid. Feed-back systems are dynamic type of networks i.e. their 'state' is altering incessantly till they attain an equipoise point. They endure at the equipoise point until there is alteration in input and a new stability point needs to be discovered.

III. **Network layers:** One of the most castoff simulated neural system type comprises of three clusters, or layers, of units: a level of "**input**" units is coupled to a level of "**hidden**" units, which is linked to a level of **"output"** units. The bustle of the input units embodies the crude statistics that is served into the system.

- The action of each concealed unit is resolute by the actions of the input layer or entities and the weights on the networks amongst the input layer and the hidden layer.
- The comportment of the output entities relies on the commotion of the hidden layer and the weights amongst the hidden and output elements.
- The hidden layer or entities are permitted to ideate their own depiction of input.

There are two type of cataloging in network layer i.e. Solitary layer architectures and multi-layer architectures. In sole layer, all entities are coupled to one another; it has better prospective computational power compared to hierarchically organized multi-layer establishments. In multi-level systems, entities are numbered by stratum, instead of global numbering.

IV. **Perceptron:** The perceptron is a variation of MCP model with some supplementary, fixed, pre-processing. In the given below figure, Entities characterized A1, A2, Aj, Ap are baptized association units and their functionality is to excerpt precise, confined characters or pattern from the input images. Perceptron impersonates the elementary ideology in the mammalian optical system. They were principally used in pattern acknowledgment even though their competences protracted a lot more.

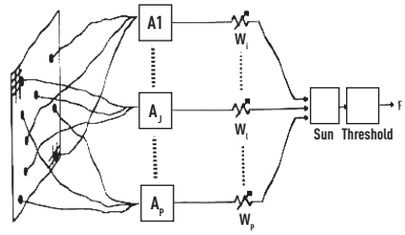


Fig 3: Perceptron Neural Network

4.3 Details on Neural Network Engineering Approach

a) **Simple Neuron Methodology:** In this methodology, an simulated neuron type device is functionalized. A simulated neuron is a expedient with several inputs and one output. The neuron has two approaches of maneuver; the teaching or Learning mode and the usability mode. In the learning approach, the neuron can be made competent to fire (or not), for certain input

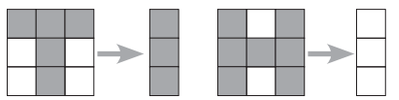
February 2018.indd 35 2/5/2018 5:55:45 PM

patterns. In the usability approach, when a previously trained input outliner is perceived at the feedback, its concomitant output develops as the contemporary output. If the input outliner configuration does not fit in the previously trained group of input arrangements, the firing instruction are used to regulate the trigger points [6].

b) **Firing instructions:** The Triggering instruction are the vital notion in neural systems and responsible for their high elasticity. A trigger instruction governs the tactic of analysis on when a neuron should get triggered for any given input configuration. It relays to all the input configurations, irrespective of which the node was made competent [6].

Pattern Recognition in NN- an example

The grid of figure below for neural system is skilled to distinguish the configurations T and H. The related outliner are all black and all white correspondingly as displayed underneath [6].



If we signify black blocks with 0 and white blocks with 1 then the truth tables for the 3 neurons after generalization are

X11:		0	0	0	0	1	1	1	1
X12:		0	0	1	1	0	0	1	1
X13:		0	1	0	1	0	1	0	1
OUT:		0	0	1	1	0	0	1	1

Top neuron

X21:		0	0	0	0	1	1	1	1
X22:		0	0	1	1	0	0	1	1
X23:		0	1	0	1	0	1	0	1
OUT:		1	0/1	1	0/1	0/1	0	0/1	0

Middle neuron

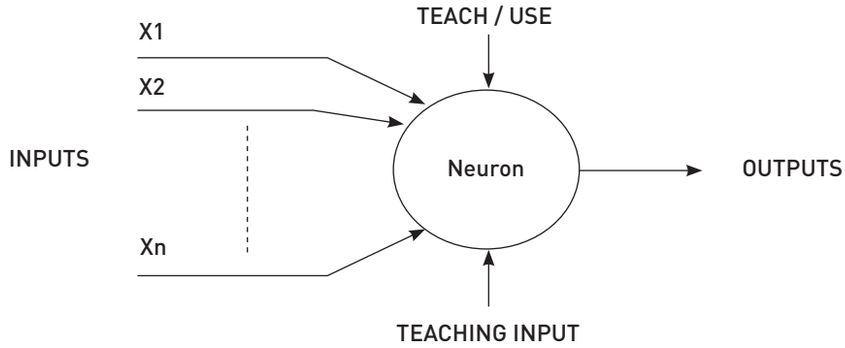


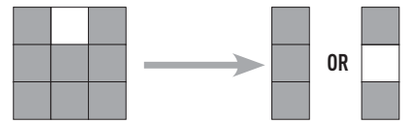
Fig. 4 : Simple Artificial Neuron

X21:		0	0	0	0	1	1	1	1
X22:		0	0	1	1	0	0	1	1
X23:		0	1	0	1	0	1	0	1
OUT:		1	0	1	1	0	0	1	0

As per the boards it is to be perceived the following connotations can be mined:

In this instance, it is apparent that the resultant should be all blacks as the input configuration is utmost similar to the 'T' pattern.

Here also, it is noticeable that the yield should be all whites as the input configuration is similar to the 'H' pattern.



Here, the topmost tuple is 2 faults distant from a T and 3 from an H. Hence, the top response is black. The intermediate tuple is 1 fault distant from both T and H so the response is arbitrary. The end tuple is 1 fault distant from T and 2 distant from H. Therefore, the response is black. The aggregate response of the system still favors the T shape [6].

a. **Complicated Neuron:** A further erudite neuron model is the McCulloch and Pitts model (MCP). The variance from the preceding prototype is that the receiving ends are 'weighted'; the consequence of decision-making for a particular input is reliant on the weight of

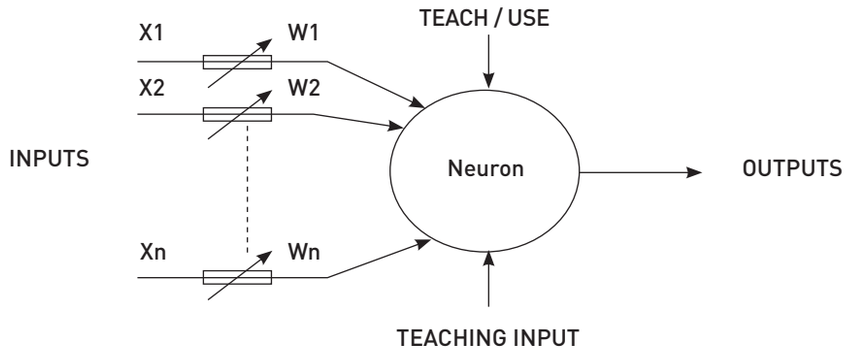


Fig. 5 : An MCP neuron

February 2018.indd 36 2/5/2018 5:55:46 PM

the specific input. The weight of an input is a value which when crossed with the input provides the weighted input. These weighted inputs are further summed up and if they outstrip a defined threshold value, the neuron triggers. In any other circumstance, the neuron does not trigger [6].

In scientific expressions, the neuron triggers if and only if;

$$X1W1 + X2W2 + X3W3 + \ldots > T$$

The totaling of input values and of the edge makes this neuron a very elastic and dominant one. The MCP neuron has the aptitude to acclimatize to a certain circumstance by altering its value or weights and/or the edge limit. Numerous algorithms pertains that roots the neuron to 'adapt'; the utmost castoff ones are the Delta rule and the back error propagation. The prior is utilized in feed forward grids and the latter in feedback systems [7].

Conclusion

Neural Networks is fascinating for user-specific systems in the field of education, entertainment, information processing, genetic engineering, neurology and psychology. Programs are being developed that requires feedback from the user for its efficiency. Unpretentious and "passive" sensors like fingertip sensors, or wristbands to sense pulse, blood pressure etc. can provide effective feedback into a neural control system. Neural Network's ability to learn by example makes them very flexible and powerful. Recent advances and future applications of Neural Networks include:

i. Integration of fuzzy logic into neural networks
ii. Pulsed neural networks
iii. Hardware specialized for neural networks
iv. Improvement of existing technologies

In the future, Neural Networks might allow:

- Advanced Autonomous Robots that can visualize, sense, and predict the world around them
- Upgraded stock forecast
- Conjoint practice of self-driving cars
- Advanced Composition of music
- Handwritten artifacts to be spontaneously renovated into formatted word processing documents
- Developments in the human genome to benefit the learning of the data compiled by the Humanoid Genome Project
- Diagnosis of medical complications using neural networks and much more!

In conclusion, it should be articulated that irrespective of neural systems have an enormous prospective, the best of them can be extracted only when they are unified with computing, fuzzy logic, AI and related subjects.

So this can be deduced that, pattern recognition using Neural network can bring a new era in the field of automatic intelligent Computing.

Reference

[1] A.K.Jain,R.P.W.Duin,J.Mao, "Statistical pattern Recognition: A Review",IEEE Transactions on Pattern analysis and machine Intelligence,Vol22,No:1,2000

[2] W. H. Highleyman,"The design and analysis of pattern recognition experiments", The Bell System Technical Journal (Volume: 41, Issue: 2,IEEE, March 1962.

[3] "Benjamin Franklin Medal in Computer and Cognitive Science". ***Franklin Institute. 2012***. Retrieved April 6, 2013.

[4] Scholkopf, Bernhard et al (eds)). "Preface". Empirical Inference: Festschrift in Honor of Vladimir N. Vapnik. Springer. ISBN 978-3-642-41136-6.2013

[5] Igor Aleksander, Helen Morton, "An introduction to Neural Computing", International Thomson Computer Press, 01-Jan-1995, 2nd edition

[6] Md. Adam Baba, Mohd Gouse Pasha, Shaik Althaf Ahammed, S. Nasira Tabassum, "Introduction to Neural Networks Design", International Journal of Scientific & Engineering Research Volume 4, Issue 2, February-2013

[7] Oludele Awodele, Olawale Jegede, "Neural Networks and Its Application in Engineering", Proceedings of Informing Science & IT Education Conference (InSITE) 2009

Web-Link

[1] https://en.wikipedia.org/wiki/Artificial_neural_network

[2] https://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html

■

4.4 Comparison between different approaches

	Stat PR	Synt PR	Neur PR
1. Pattern Generation (Storing) Basis	Probabilistic Models	Formal Grammars	Stable State or Weight Array
2. Pattern Classification Basis	Estimation/Decision Theory	Parsing	Based on Properties of Neural Network
3. Feature Organization	Feature Vector	Primitives and Observed Relations	Neural Input or Stored States
4. Typical Learning Approaches Supervised:	Density/Distribution Estimation	Forming Grammars	Determining Neural Network System parameters
5. Typical Learning Approaches Unsupervised	Clustering	Clustering	Clustering
6. Limitations	Difficulty in Expressing Structural Information	Difficulty in Learning Structural Rules	Often little semantic information from network

Table 4: Comparison between different Approach

February 2018.indd 37 2/5/2018 5:55:46 PM

About the Authors

Dr. S. K. Gupta is an assistant Professor of CS and Engg. Dept., University Inst. of Tech., Burdwan University,West Bengal,India.He obtained his B. E. degree in CS and Engg. from Burdwan University, West Bengal in 2005 and M.Tech Degree in the same subject in 2007 from KIIT University, Orissa. He received Ph.D. in the field of Multi Agent System, CS. His research interests include Multi Agent System, Intelligent Healthcare System, à Robot systems and swarm robotics. In 2009 He got the best paper award for the paper "A Multi-Agent Based Scheme for Health Care and Clinical Diagnosis System" in the International Conference on Intelligent Agent and Multi-Agent, IEEE. Prof. Gupta is the Recipient of "The Cambridge Certificate" for Outstanding Engineering Achievement for services to Computer Science and Engineering by IBC, Cambridge, England (12th July 2013). His biography was selected in the 30th Pearl Anniversary Edition of Marquis Who's Who in the world, 2012.He is a board member of International Journal of Emerging Technology & Advanced Engineering, International Journal of Recent Advancement in Engineering and Technology. He is an author/co-author of thirty published articles in Multi Agent System field. He is the member of IEEE, ACM, CSTA. He is the supervisor of six PG and three Ph.D. (Cont.) research work.

Ayan Mukherjee is M. Tech Degree holder from BITS Pilani and B.E from University Institute of Technology, Burdwan. He is an Associate and a Product Specialist in Cognizant with experience in Quality Engineering and Analysis, Business Analysis and development and System Architect.

He is worked on topics like Computer-User Interaction Automation, Multi Agent System and UML Based PEGA on Industrial Level in Cognizant. He also has experience in analysis of various Digital Interaction based industrial model. He has currently four years of Industrial Experience.

CSI Transactions on ICT Editorial Board Conference

CSI Transactions on ICT Editorial Board Conference was conducted on 3 February 2018 at IIT Madras Research Park. The Conference was presided by past CSI Presidents & Fellows, Mr. S. Mahalingam and Dr. S.V. Raghavan and attended by distinguished professionals from academia, industry and government such as IIT Madras, IIT Hyderabad, IIT Guwahati, Ministry of Electronics and IT, Cognizant, Tata Consultancy Services etc. Prof. A.K. Nayak, Hon. Secretary, CSI and Prof. Prashant R.Nair, Editor, CSI Communications represented CSI for the conference. Prof. A.K. Nayak also addressed the august gathering and briefed about various publications of CSI such as CSI Journal of Computing, Adhyayan and Communications

February 2018.indd 38 2/5/2018 5:55:49 PM

Pattern Recognition: Overview and Applications

Priya Sen
RCC Institute of Information Technology, Kolkata, India

Siddhartha Bhattacharyya
RCC Institute of Information Technology, Kolkata, India

Pattern recognition is the technique by which computing devices can examine the environment, figure out distinguish patterns of interest from their background and make more sensible and reasonable outcomes about the classes of the patterns. The objective of this widely used technique is to resolve troublesome mechanisms of decision making methods by using of computers. This paper introduced the architecture and classification of pattern recognition, its related fields and real world applications in details.

Keywords : Pattern recognition, features, classification, Template, statistical, syntactic, supervised, and unsupervised learning.

1. Introduction

In the last few years, the applications involving pattern recognition (PR) are quickly increasing in numbers from wider area and has become more popular and important research topic. This standout keystone of Computer Science includes finding the similarities or patterns among small, deteriorated issues that can help us to find out more complex problems in adequately. It is a combination of mathematical, statistical, metaheuristic and inductive methods of fundamental role for the processing of tasks on computing systems, processed like human beings [1].

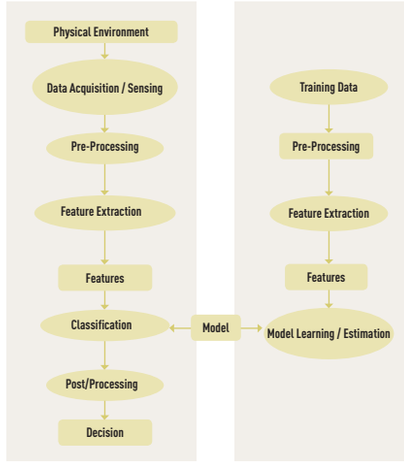
Research in PR has exponentially expanded in the previous decades because of the progress in both nature and determination of imaging sensors and the sensational build over computational control. This increase has also been accompanied by smoothing the boundaries between different applications of PR, making it really interdisciplinary. The objective of this special issue is to invite original research contributions that address the broad challenges faced in PR.

PR is a branch of machine learning that concentrates on the recognition of patterns and regularities on data, in spite of the fact that it is almost synonymous with machine learning. PR algorithms generally give an acceptable respond for all possible inputs and to accomplish "most likely" similar inputs, taking into account their statistical variations. A typical example of a pattern-matching algorithm is finding out the similarity of regular expressions, which searches for patterns of a given sort in textual data and is comprised in the search competencies of various text editors and word processing units. In compare to PR, pattern matching is normally not treated as a kind of machine learning, although pattern-matching algorithms can be fruitful in some cases by implementing similar-quality result of the sort given by pattern-recognition algorithms.

The paper is outlined as follows: Section 2 explains the term PR, Section 3 surveys various works related to this field, Section 4 states the signature of the work. Various research challenges associated with PR has been described in Section 5 followed by conclusion in Section 6.

2. Definition

PR is a methodology that taking in raw data and working based on the class of the pattern. By using our past learning experiences we can effectively distinguish things or objects depending upon its characteristics. Then how about computers? It can be solved by using of two basic factors: feature and classifier. A pattern is an entity, method or fact whereas a type (or category) is a set of designs that share common attribute (features) commonly beside the same statistic sources. During recognition (or classification) training are allocated to the entities. A classifier is a machine that executes such assignments. The scope of a PR system may be examined a view for this present reality and to arrive at a depiction of the scene which may be helpful for the completion of some works. The real world observations are gathered through sensors and PR system classifies or describes these observations. A general PR system is shown in the Figure 1. In the first step, data is acquired and preprocessed which is followed by feature extraction, feature reduction and grouping of features and finally the features are classified. In the classification step, depending upon the measured features input patterns are assigned to the pattern classes by



February 2018.indd 39 2/5/2018 5:55:50 PM

those trained classifiers. The training set, utilized during development of the classifier is different from the test set which is used for evaluation. This ensures different performance environments.

3. Related Works

There are numerous domains where different PRs have been developed like sequence analysis of Bio-informatics, searching for required patterns in Data Mining, Internet searching for Document Classification, Perusing machine for blind, Printed circuit board analysis, Personal identification, speaker identification, disease identification and weather forecasting.

Three aspects have been considered to design of a PR system which are data representation, classification and prototyping. In the problem domain, it describes choosing of the appropriate sensors, pre-processing techniques, representational scheme, and more decision making model.

3.1 Violence Detection From ECG Signals

Harmful action, identified as violence has been detected from not more than 6-second strip of electrocardiogram signals by calculating character-based two ultra-short standard methods: the Bivariate Empirical Mode Decomposition (BEMD) and the Recurrence Quantification Analysis (RQA) [3]. Performance has been evaluated by using above two methods in 10-fold cross validation and Leave-one-subject-out (LOSO) validation, applied to 12 students of two different standard (2nd Std. and 5th Std.) of a primary school depended on k-nearest neighbour (kNN) classifier and verified in two different training models as shown in Fig. 2.

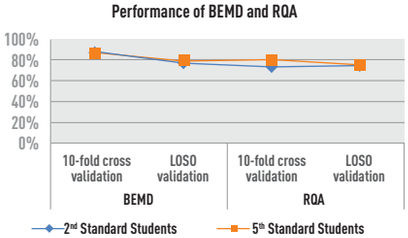


Fig. 2 : Performance of BEMD and RQA

Though a bigger deviation has been noticed in the experimental result, average execution in subject specific field in 10-fold cross validation had been done, displayed in Fig. 3.

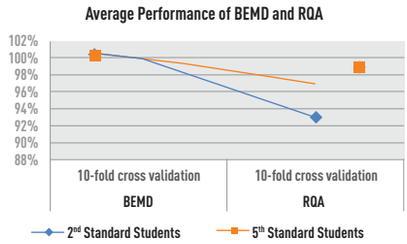


Fig. 3 : Average Performance of BEMD and RQA

To our knowledge, this is the first study to investigate methods to detect violence events based on ultra-short ECG signals analysis from the victims' point of view. The finding should make an important contribution in PR research, especially related to violence detection, which can be used in violence intervention programs in schools. Tests showed that the subject-specific classifier achieved excellent results in the data but it raised practicality issue as the system needed to find the optimum combination of window size and overlap in spectrogram analysis and selected appropriate features for each subject. Both subject-dependent and subject-independent classifiers produced only satisfactory results as the false negative rate was relatively high, meaning that the classifiers misclassified many violent events as non-violent ones. Above all, these facts indicate that the violence events can be detected using ultra-short ECG signal with appropriate signal analysis.

3.2 Deep learning for visual understanding

Deep learning algorithm has been developed by different researchers and generally embraced distinct aspects of computer vision such as image classification, item detection, image recovery and semantic division and human pose evaluation. Yanming Guo et al. [5] have divided the deep learning into four sections: Convolutional Neural Networks, Restricted Boltzmann Machines, Autiencoder and Soarse coding.

3.3 An Efficient painting based Skew Estimation method for scanned Documents by using Piece-Wise Painting Algorithm

An efficient method based on the concept of geometry and computer graphics has been developed by Alireza Alaei et al. [6] which is divided into three phases: 1)Piecewise Painting Algorithm (PPA) 2) selection of candidate bands and 3) determination of the best-fit line and skew estimation. The stated method has been experimented in a large scale on three identical datasets consisting various types of scanned images. Utilizing the PPA on the original report in both even and vertical plane, two separate painted pictures have been considered [6]. To get the best fit line and to assess the skew edge of the input picture, beginning, center and end purposes of the candidate bands are recorded in three isolated groups. As a result, effective outcome were acquired. The skew edge of the document picture has been finally assessed base on the incline of the best-fit line and the document skew is rectified. This method has demonstrated better outcome considering both the precision and calculation time with no extra operations and has been repeatedly tested until the evaluated skew is under 1 degree.

3.4 Biometric Authentication for Smart Phone Security

Tempestt J. Neal et al. [7] have shown the separate modules of the Biometric Authentication for smart phone operating in different modes. Different types of physical biometrics authentication systems like face recognition, fingerprint recognition, palm print recognition, iris recognition, gait recognition, touch gesture and voice/speaker recognition have been introduced to develop the algorithm for this type of research.

3.5 Rotation invariants of vector fields from orthogonal moments

Mr. Bo Yang et al. [8] have proposed a technique for the depiction and coordinating of vector field pattern, which are elements of orthogonal moments and furthermore demonstrated their handiness in a real

February 2018.indd 40 2/5/2018 5:55:51 PM

world layout coordinating applications.

Vector fields carry on uniquely in contrast to gray level and color images under spatial changes though conventional scalar invariants cannot be productively utilized for recognition. They clarified the utilization of orthogonal moments gives critical higher numerical soundness than the steadiness of geometric/complex moment invariants. Two well known sorts of orthogonal minutes named Gaussian-Hermite moments - orthogonal on a square and Zernike moment, orthogonal on a disk have been tested for this purpose. As a result of that experiment, the steadiness of the earlier was slightly somewhat better yet the distinction was not indicative and every kind had its own upsides and downsides, inferred by their different territories of orthogonally. Contrasting with vector field invariants from non-orthogonal moments and to scalar image invariants, the proposed procedure accomplishes essentially better outcomes.

4. State of the Art

To describe PR in a simple way, three questions arrive in our mind: what is it? Where is it? And how is it produced? The answers can be given by three tasks: Classification i.e. analyzing manuscript, Segmentation applied for identifying test alternately face areas in images and Parsing, Syntactic PR for deciding how a group of images are similar and how they design an explanation. Patterns generated from the raw data depending on the nature of the data. It can be in the type of any picture, textual content, video or another layout such that a image of fingerprint, a bar code, a handwritten cursive letter, image of a voice signal or a face of human, given in Fig. 4 [9] .

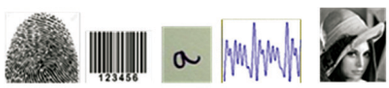


Fig. 4 : Different types of Patterns

Principle of PR commonly gives a proper outcome for all viable inputs by thinking about statistical variations [4]. There are four well-known approaches for the PR which is shown in Fig. 5.

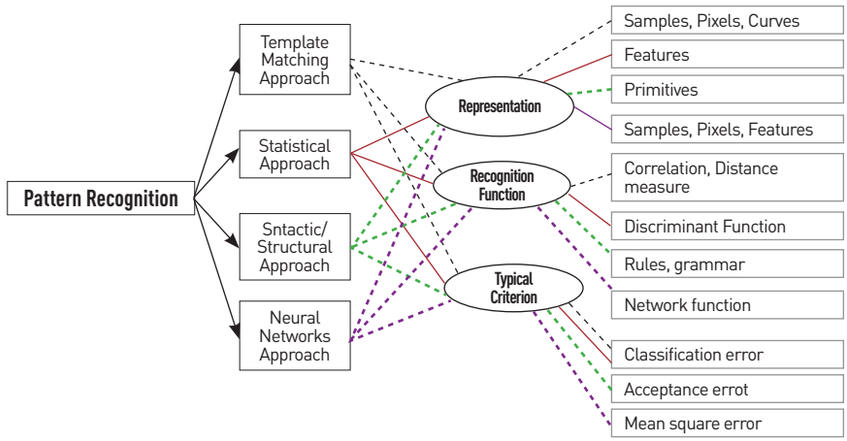


Fig. 5 : Different approaches of PR

4.1 Template Matching

A simple technique is used for finding sample i.e. small parts of input images to be compared against the template image. It has been implemented at the pixel level to higher level also.

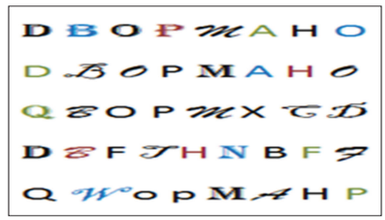


Fig. 6 : Example of Template matching

4.2 Statistical Classification

The objective of Statistical PR (described in Figure 7), known as "machine learning" is to find, learn and analyze the training pattern in terms of feature sets having no relation. Classification, implemented by supervised learning method and clustering, developed by unsupervised learning method are the two paradigm of PR [9].

Fig. 7 : Statistical approach of PR

4.2 Syntactic Matching

In this approach, a pattern is characterized by the structures depends on the inter-relation between features where a matching is linked between the structure of patterns and the logical rules of a language [9].

4.3 Neural Networks

The focal characteristics of the neural network (NN), suitable for developing nonlinear algorithms for feature extraction and classification are to learn complex nonlinear input-output connections, utilization of sequential training methods, and familiarized themselves to the data. Figure 8 depicts a common diagram of NN: three layers with full interconnection. It consists of one input layer with four neurons, one output layer and a hidden layer with five neurons. The action of the NN is decided by weights applied in the hidden and output nodes.

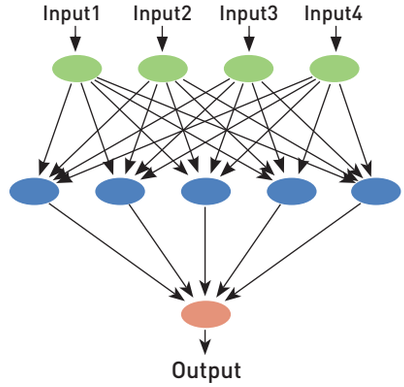


Fig. 8 : Example of Artificial Neural Network

February 2018.indd 41 2/5/2018 5:55:51 PM

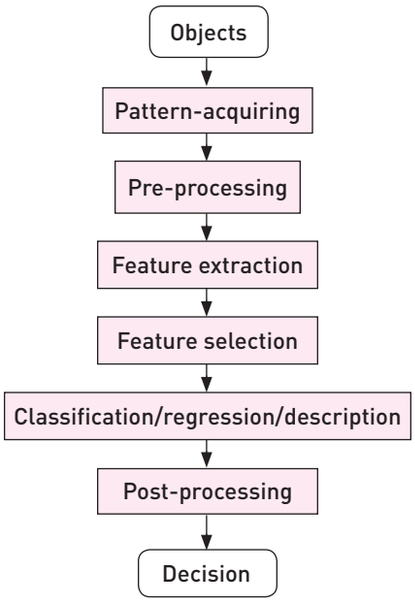


Figure 9: The Construction of PR System [1] [2]

Jie Liu et al. [1] and Vinita Dutt et al. [2] describe PR system, which can be developed through five processes and the classification of this system described in Fig. 10.

A pattern recognition system based on any PR method mainly consists of three mutual-associate and differentiated processes: data building, pattern analysis and pattern classification. Data building convert original information into vector which can be dealt with by computer. Pattern analysis' work is to process the vector, such as feature selection, feature extraction. Pattern classification is used to utilize the information, learnt from pattern analysis to discipline the computer. It is a PR problem of appointing a entity to a class, where the output of the PR framework is an integer label, for characterizing an item as "1" or "0" in a quality control test. A very common description of the pattern recognition system that includes five steps to accomplish which are shown in the Fig. 9.

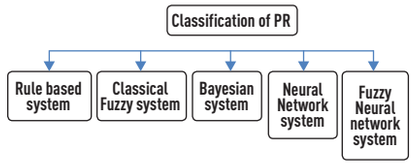


Fig. 10 : The Classification of PR System

Applications

Pattern Recognition is nowadays used in almost all fields starting from natural language processing to stock exchange forecasting. Following table (Table 1) shows the application of Pattern Recognition in different fields.

5. Conclusion

In this paper, an overview of pattern recognition has been presented by describing definition of PR, different methods of PR, the design model of PR system, related fields of PR and the application of pattern recognition.

5. References

[1] Jie Liu, Jigui Sun, and Shengsheng Wang, "Pattern Recognition: An overview", IJCSNS International Journal of Computer Science and Network Security, VOL.6, No.6, June 2006.

[2] Vinita Dutt, Vikas Chaudhry, and Imran Khan "Pattern Recognition: an Overview", American Journal of Intelligent System, 2(1): Pages 23-27, DOI: 10.5923/j.ajis,20120201.04, 2012.

[3] Hany Ferdinando, Guobing Sun, Tuija Huuki, Tapio Seppanen, and Esko Alasaarela, "Violence detection from ECG Signals: a preliminary study", Journal of Pattern Recognition Research 1 Vol 12, No 1; pages 7-18, doi:10.13176/11.790, September 2017.

[4] M. P. Raj, P. R. Swaminarayan, and J. R. Saini and D. K. Parmar, "Application of Pattern Recognition Algorithms in

Table 1: Application of PR System

Fields	Applications	Example of Pattern
Engineering	Safety performance of automobile	Recognition of vehicle type
Science	Quantum computing	Quantum interference pattern
Civil Administration	Traffic Analysis and Control	Classification of signals
Economy	Stock exchange forecast	Predictable market value of financial time period
Geography	Classification of rocks	Earthquake analysis Rocks classification
Multimedia database retrieval	Internet search	Video clip, Video games
Data mining	Searching for meaning full pattern	Points in multi-dimensional space
Bioinformatics	Sequence analysis	DNA/Protein sequence
Remote Sensing	Forecasting crop yield	Multi spectral image
Biometric recognition	Personal identification	Face, iris, fingerprints
Document Classification	Internet search	Text document
Document Image Analysis	Optical character recognition	Document image
Industrial Automation	Printed circuit board inspection	Intensity or range image
Medical	Computer aided diagnosis	Microscopic image
Military	Automatic target recognition	Optical or infrared image
Natural language processing	Information extraction	Sentences
Machine perceptron	Speech Recognition	Speech waveform
Agriculture	Nature of crops	Extraction mineral characterization in coffee and sugar
Astronomy	Classifying galaxies	Image analysis by astronomical telescope and automatic spectroscopy

February 2018.indd 42 2/5/2018 5:55:51 PM

Agriculture: A Review", Int. J. Advanced Networking and Applications, Volume: 6, No 6, Issue: 5, Pages: 2495-2502 , ISSN: 0975-0290, 2015.

[5] Yanming Guo, Yu Liu, Ard Oerlemans, Songyang Lao, Song Wu, and Michael Lew, "Deep learning for visual understanding: a review", Neurocomputing 187, pages 27 - 48, Elsevier, 2016.

[6] Alireza Alaei, P. Nagabhushan, Umapada Pal, and Fumitaka Kimura, "An efficient Skew estimation technique for scanned documents: an application of Piece-Wise Painting Algorithm", JPRR, Vol 11, No 1; doi:10.13176/11.635,2016.

[7] Tempestt Neal and Damon Woodard, "Surveying Biometric Authentication for Mobile Device Security", Journal of Pattern Recognition Research, Vol 11, No 1, doi:10.13176/11.764, November 2016.

[8] Bo Yang, Jitka Kostkova, Jan Flusser, Tomas Suk and Roxana Bujack, Rotation invariants of vector fields from orthogonal moments, Pattern Recognition, Pages 110-121, Elsevier 2017.

■

About the Authors

Mrs. Priya Sen did her Bachelor of Technology in Information Technology in 2006 and Master of Technology in Computer Science and Engineering in 2015 from Maulana Abul Kalam Azad University of Technology, West Bengal (Formerly known as West Bengal University of Technology) in 2006. She has worked as a Junior Research Fellow (GATE) in Central Mechanical Engineering Research Institute (Council of Scientific & Industrial Research), Durgapur from 2008 to 2009. She has done M. Tech thesis, carried out from Variable Energy Cyclotron Centre (VECC), Department of Atomic Energy, Government of India on Event Selection for MUCH of CBM Experiment @FAIR using Heterogeneous Computing from 2014-2015 and this work was published in 12th IEEE India International Conference at Jamia Millia Islamia, New Delhi, India. She has worked as a faculty in department of Electronics and Communication Engineering, RCC Institute of Information Technology, Kolkata, India during the academic session of July to December 2015. She is currently working as a Technical Assistant in department of Computer Science and Engineering, RCC Institute of Information Technology, Kolkata, India from 2016. Her research interests include GPU computing, pattern recognition, hybrid intelligence and quantum computing.

Dr. Siddhartha Bhattacharyya did his Bachelors in Physics, Bachelors in Optics and Optoelectronics and Masters in Optics and Optoelectronics from University of Calcutta, India in 1995, 1998 and 2000 respectively. He completed PhD in Computer Science and Engineering from Jadavpur University, India in 2008. Awards won include University Gold Medal from the University of Calcutta, Adarsh Vidya Saraswati Rashtriya Puraskar, Distinguished HoD Award and Distinguished Professor Award conferred by Computer Society of India, Mumbai Chapter, Bhartiya Shiksha Ratan Award, NACF-SCRA, India award for Best Faculty for Research, Honorary Doctorate Award (D. Litt.) from The University of South America and the South East Asian Regional Computing Confederation (SEARCC) International Digital Award ICT Educator of the Year. He is currently the Principal of RCC Institute of Information Technology, Kolkata, India. In addition, he is also serving as the Professor of Computer Application and Dean (Research and Development and Academic Affairs) of the institute. He is a co-author of 4 books and the co-editor of 8 books and has more than 185 research publications in international journals and conference proceedings to his credit. He has got a patent on intelligent colorimeter technology. He has served as the Editor-In-Chief/Editor/Associate Editor of International Journal of Ambient Computing and Intelligence (IJACI) published by IGI Global, Hershey, PA, USA; International Journal of Pattern Recognition Research; International Journal of Engineering, Science and Technology and ACCENTS Transactions on Information Security (ATIS).

CSI CALENDAR 2017-18

Gautam Mahapatra, Vice President, CSI, Email: vp@csi-india.org

Date	Event Details & Contact Information
FEBRUARY 17-18, 2018	**International Conference on Economics, Business Management and Business Analytics, ICEB'18**, Christ Institute of Management, Lavasa, Pune in association with CSI Division IV, Communication & CSI Pune Chapter. Contact: **Ms. Unnati G Hunjan**, 020 66753814, conference@cimlavasa.in, samiksha.shukla@christuniversity.in
MARCH 08-09, 2018	**National Conference on Challenges and Opportunity in Computer Engineering,** NCCOCE'18, Christ (Deemed to be University), Bengaluru in association with CSI Division IV, Communication & CSI Bangalore Chapter. Contact: **Dr. Samiksha Shukla**, 9880462311, **Contact - Mr. Naveen J**, 9994289908, nccoce@christuniversity.in
APRIL 07-08, 2018	**International Conference on Computational Intelligence and Data Science (ICCIDS2018)**, The NorthCap University, Gurugram, India In association with Computer Society of India Division IV Communications, Website - http://iccids2018.ncuindia.edu/ Contact - **Dr. Vijendra Singh**, email - vsingh.fet@gmail.com

February 2018.indd 43 2/5/2018 5:55:52 PM

Optimizing Deep Convolutional Neural Network for Facial Expression Recognition

Umesh Chavan
Walchand College Of Engineering, Sangli MH, INDIA
umesh.chavan@walchandsangli.ac.in

Dinesh Kulkarni
Walchand College Of Engineering, Sangli MH, INDIA

Facial expression recognition (FER) system detects human's emotion using their expression in the input image. It classifies expression into one of categories -sad, happy, disgust, fear, surprise, angry, neutral. Traditional systems are based on machine learning techniques. The training time is critical issue in these approaches. The advanced development is applying deep learning which is better approach which will accelerate the performance of the model. Convolutional Neural Network (CNN) is often used in object recognition and detection. We describe the basic structure of CNN and exhibit the speedup in training that can be achieved by shifting the computation intensive task of CNN to the GPU. In this paper we studied the method of learning and classification of CNNs on the GPU (Graphical Processing Unit) which accelerate the performance and scalability. We build a deep learning Convolutional Neural Network (CNN) model for facial expression recognition using Theano and Caffe libraries. The result demonstrated that GPU version works better than CPU version of training the model.

Keywords : Deep Learning, Convolutional neural network, Graphical Processing Unit, Facial expression Recognition.

1. Introduction

1.1 Background

Facial expression conveys important non-verbal communication cues in interpersonal relationship. Facial expression recognition can be implemented in all computer interfaces, automated psychological research and treatment, robots or even polygraphs. This task is challenging to machines [1]. Machine learning (ML) is the capability enabled to a computer to learn without being explicitly programmed. It is functionality to learn and make predictions from data. Deep learning is ML technique that uses multiple units of layers of nonlinear processing units to conduct supervised or unsupervised learning from data. Deep Learning (DL) is a new advancement in area of machine learning research whose motivation is moving closer to the objective of Artificial Intelligence (AI).Convolutional neural networks (CNNs) are a special kind of DL method. CNNs are useful in the area of computer vision. The milestones in the development of deep learning from neural network are presented in Fig. 1. Professor Hinton at Toronto University in Canada, and others published a paper on "A fast learning algorithm for deep belief nets" [2]. In this paper, a method called deep belief nets (DBN) was presented. DBN is an expansion of neural networks, a method of machine learning. DBN was tested on MNIST images [4], the standard database for comparing the precision and accuracy of each image recognition method. This dataset includes 70,000; 28x28 pixel hand-written character image data of numbers from 0 to 9. Then they constructed a prediction model based

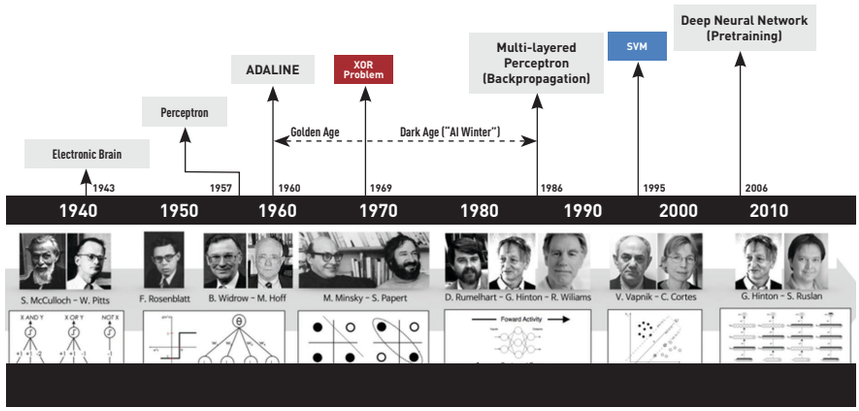


Fig. 1: AI and DL History

(Source: https://beamandrew.github.io/index.html)

February 2018.indd 44 2/5/2018 5:55:53 PM

on the training samples and measured its performance accuracy based on whether a machine could correctly answer which number from 0 to 9 was written in the test case.

The neural network algorithm substitutes the linkage of the network by weighting with numbers. If a network is simple, weights are properly allocated from the learning data and the network can recognize and classify patterns well. However, once a network gets complicated, the linkage becomes too dense and it is difficult to make a difference in the weights. The network can make a proper model by adopting a mechanism that feeds back errors that occurred during training algorithm to the whole network. Again, if the network is simple the feedback can be reflected properly, but if the network has many layers a problem occurs in which the error disappears before it's reflected to the whole network. While deep learning succeeded in making network multi-layered, that is making a network "deep". The key to success is to make each layer learn in stages. Hence deep learning took the approach of making each layer learn in advance. This is known as pertaining. In pre-training, learning starts from the lower-dimension to layer in sequence. Deep CNN is built on hierarchy of layers. Every layer transforms the input data into more hypothetical representations (e.g. edge->eyes->face). The output layer combines those features to make predictions. A pictorial description of typical CNN architecture can be found in Fig. 2.

All the machine learning/deep learning algorithms based on that the type of input data is one-dimensional. For a real-world applications, mostly input data is multidimensional (images, video etc.). It would be preferable to build a model that can handle multi-dimensional data as it is. Otherwise, some information content in the data might be lost when flattened to one dimension. To solve this problem, a model called Convolutional Neural Network (CNN) was suggested. In CNN features are extracted from multi-dimensional input data through convolutional layers and

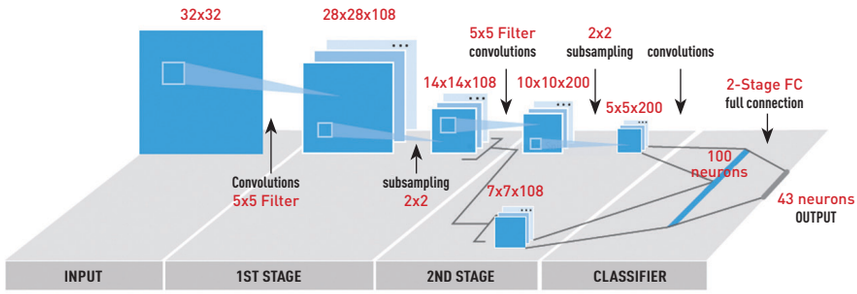


Fig. 2: Typical DL model

pooling layers and then these features are put into general multi-layer perceptions. From vast amounts of data, CNN learn a complex representation of visual data. The human visual system is the motivation. Facial expression recognition (FER) techniques detect people's emotions using their facial expressions. We build a model for FER using deep CNN. CNNs have much fewer parameters as compared to neural networks. So they are easier to train. Training CNN on large data sets is computationally intensive tasks on CPUs. We have adopted inherent multi-level parallelism of CNNs for NVidia's GPU architecture to optimize and accelerate the training. We implemented model in Theano framework (5) and exploited Graphics Processing unit (GPU) computation. The result shows that it is achieving up to 5 times speedup with training model on GPU. The accuracy obtained on validation set is 77%. Our intention is to exhibit the performance and scalability improvement for FER using deep CNN.

2. Convolution Neural Network (CNN)

A CNN is an advance in neural network evolution. It consists of sequences of one or more convolutional layers (CLs). CL's are mostly with pooling layers (PLs). PLs are succeeding by one or more fully connected layers (FCs), FCs are just as standard neural network. Accurate and correct feature extraction is necessary. This is the base for CNN. Input to a neural network is fed from output of feature extractor. It is challenging work to select a "suitable" feature extractor. it cannot adapt to network configuration. It is not part of learning procedure. These layers arranged in feed-forward structure as shown in Fig. 3.The typical neural network has parameters - weights and biases. The model learns these values during the training process and it's continuously updates with each new training examples. However, in the case of CNN , weights and bias values are the same for all hidden neurons in a given layer. This means that all hidden neurons detect the same feature such as an *edge* or blobs in different regions of the input image. This makes network tolerant to translation of object in an image. Activation step applies the transformation to the output of each neuron by using activation functions Rectified Linear Unit commonly known as a ReLUs. Most DL network use ReLU for hidden layers. The power of ReLU is it trains much faster, is more expressive than other alternatives -logistic function. Also ReLU prevents the gradient vanishing problem. It takes output of neurons and maps it the highest positive value or if the output is negative the function maps it to zero. We can further transform the activation step by applying a pooling step. Pooling reduced the dimensionality of the feature map by condensing the output of small regions of neurons into a single output.

2.1 Convolutional Layer

The Convolution is used to find the same feature in different places of an image. Convolution is conducted using learnable filters or kernels that are passed through the input data, the input image. The convolutional layer uses multiple filters where each filter moves sequentially across the input data or image to make a 2-dimensional

February 2018.indd 45 2/5/2018 5:55:53 PM

activation map based on each filter. Feature maps are made from the activation maps of the filters. And the number of learnable filters, or kernels, in the convolution process, determines how many feature maps are generated after the convolutional process. Convolution operation is shown in Fig. 3.

For a two dimensional case, the convolution can be described by the equation:

$$y[m,n] = \emptyset(p) = \emptyset(b + \sum_{k=0}^{k-1} \sum_{l=0}^{k-1} v[k,l]x[l])$$

Where

$$\emptyset(z) = \frac{1}{(1+\exp(-z))}$$

Where k is number of kernel and l is number of layer.

2.2 Pooling Layers (PL)

Pooling Layer compute the max or average value of a particular feature over a region of the input data. It also supports to detect objects in some unusual place and reduce memory size and downsizing of input image. The task performed by this layer is also known as sub-sampling. A typical structure of Pooling Layer is represented in Fig.4

2.3 ReLU (Rectified linear Unit) Layers

After each CLs it get applied a nonlinear layer. The goal of ReLU layer is to introduce nonlinearity to a system. This is due to linear operations in CLs. tank and sigmoid are some traditional activation functions were applied. ReLU layers work better because network is able to train faster with preserving the accuracy [9].

3. GPU (Graphics Processor Units) Programming

GPUs are massively parallel numerical processors. It's programmed in C with extensions for GPU programmers. It has application programming interfaces for programmers. It takes advantages of heterogeneous computing systems that contain both CPUs and massively parallel CPU's. For a GPU developer, the computing environment consists of a host that is traditional CPU and one or more devices that are processors with

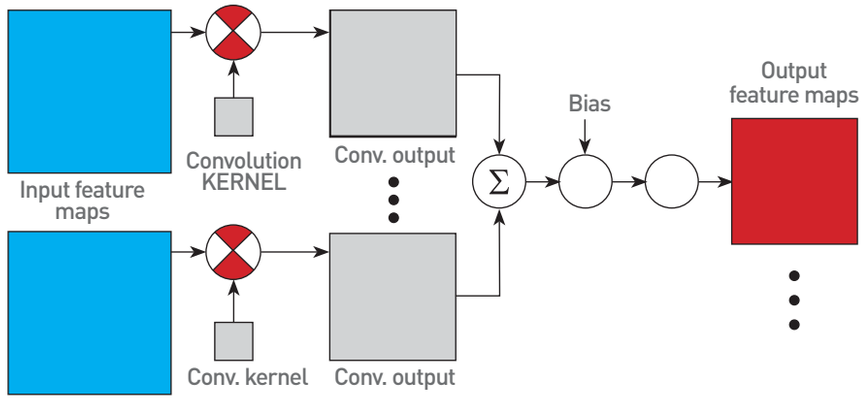


Fig. 3 : Convolution Layer (CN)

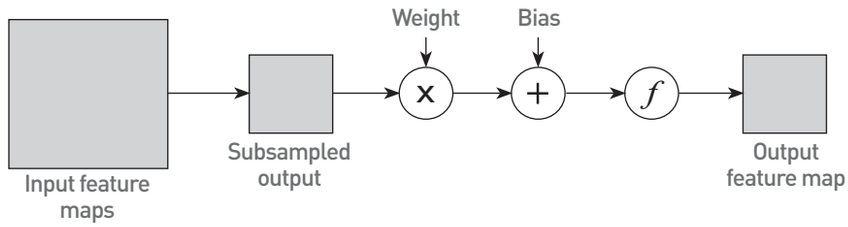


Fig. 4 : Pooling Layer

massive number of arithmetic units. GPU is a typically known as device in CUDA. Use of CPUs together with a CPU is GPU-accelerated computing. It accelerates deep learning applications. This work presents results of accelerated implementation of the deep CNN in graphic processing units (CPUs) for FER.

3.1 CUDA programming Architecture

NVIDIA proposed CUDA [14] (Compute Unified Device Architecture) in 2007. This created a new programming paradigm of GPU computing and make easy and effective in using computational power of GPUs. Extending existing C program for GPU is easy in CUDA programming model. A CUDA device is a coprocessor to the CPU and runs large number of threads in parallel. CUDA execution model involves grids, blocks and threads. Figure shows the CUDA programming execution unit. A kernel in CUDA is data-parallel portion of an execution code. Kernel executes by an array of threads. All these threads run same code but different data. In the model monolithic threads are grouped into blocks. A grid is group of blocks. Thread within a block cooperates via shared memory. Each thread had an ID which used to calculate memory

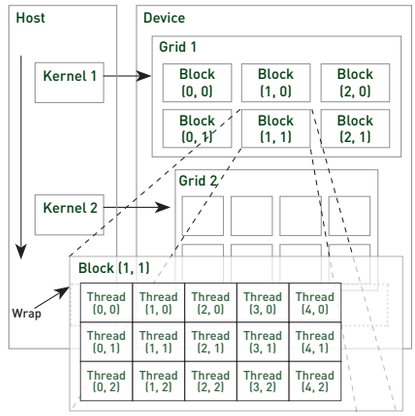


Fig. 5: CUDA Programming Model

address and make control decisions. The ID is 1D, 2D or 3D. Similarly each block has 2D or 1D ID. Group of threads physically runs on SMs (streaming processors) of GPUs. SMs share instructions. A group of threads running on SM is known as wrap. Typically wrap of 32 threads is used. The role of SMs is to schedule and allocation of threads. SM is array of streaming processors. (SPs).We used NVIDIA GeForce GT 520 of Fermi architecture has 48 cores. The memory bandwidth us 14.4 GB/s. GT

February 2018.indd 46 2/5/2018 5:55:54 PM

520 has compute capability 2.1. CUDA architecture is abstracted as a multi-core processor allowing the concurrent execution of the multiple threads. This architecture is structured as a grid of thread blocks where each thread performs a single instruction set called kernel.

4. Methodology

4.1 System Design

We built a CNN that has following structure. The architecture is shown in Fig. 6.

a) Input → Convolution Layer 1 → Max Pooling Layer 1: Input image of size of 48 × 48 pixel is convolved with a filter of 5 × 5 pixels of depth 20followed by rectified linear units and then a max pooling that scales the results by 0.5.

b) Max Pooling Layer 1 → Convolution Layer 2: Input which is now 22 × 22 × 20, is convolved with a filter of size 5, and depth 20 followed by ReLUs (rectified linear units).

c) Convolution Layer 2 → Convolution Layer 3: Input which is now 9 × 9 × 20, is convolved with a filter of size 5, and depth 20 followed ReLUs

d) Convolution Layer 3 → Convolution Layer 4: Input which is now 9 × 9 × 20, is convolved with a filter of size 5, and depth 20 followed by ReLus

e) Fully Connected Layer 1 → Input propagates through 500 input units fully connected to 300 units

f) Fully Connected Layer 2 → Input propagates through 300 input units fully connected to 7 units.

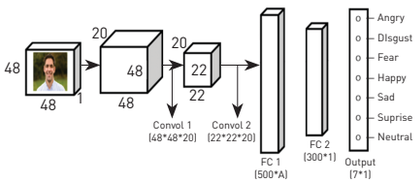


Fig. 6: Network architecture of the proposed basic convolutional neural network.

4.2 CNN Model

For an image with a size of and kernels with a size of, the convolution is represented as

$$z_{ij}^{(k)} = \sum_{s=0}^{m-1}\sum_{t=0}^{n-1} w_{st}^{(k)} x(i+s)(j+t) \quad (1)$$

Here w is the weight of the kernel, which is the model parameter. Above equation (1) is for one kernel. For multi-convolution layers, the equation is

$$z_{ij}^{(k)} = \sum_{c}\sum_{s=0}^{m-1}\sum_{t=0}^{n-1} w_{st}^{(k,c)} x_{(i+s)(j+t)}^{(c)}(i+s)(j+t) \quad (2)$$

Here c denotes the channel of the image. If the number of kernels is k and the number of channels is C, we have. Then we see from the equation (2) that the size of convolved image is (M-m+1) x(N-n+1).

After the convolution, all the convolved values will be activated by the activation function. We will implement CNN with the ReLU (rectified Linear Unit) function. With the activation we have

$$a_{ij}^{(k)} = h(z_{ij}^{(k)} + b^{(k)} = \max(o, z_{ij}^{(k)} + b^{(k)}) \quad (1)$$

Where, a one dimensional array. Next is the max-pooling layer. The propagation can simply be expressed as.

$$y_{ij}^{(k)} = \max(z_{(l_1 i+s)(l_2 j+t)}^{(k)}) \quad (2)$$

Here and l_1 are l_2 the size of pooling layers and s∈[0,l_1], t∈[0,l_2]. Usually l_1 are l_2 are set to same sizes (2 or 4).

The simple Multilayer Perceptron Network (MLP) follows after sequences of convolutional layers and pooling layers to classify data. MLP can accept one-dimensional data. Output of CLs and PLs are two-dimensional, we need to flatten the down sampled /pooled data as preprocessing to adapt it to input to input layer of MLP. The error from input layer of MLP is back propagated to the max-pooling layer, and this time it is un-flattened to two dimensions to be adapted properly to the model. Max –pooling layer simply back-propagates error to its previous layer as max-pooling layer doesn't have parameters. The equation can be expressed as

$$\frac{\partial E}{\partial_{(l_i i+s)(l_j j+t)}^{(k)}} = \begin{cases} \dfrac{\partial E}{\partial y_{ij}^{(k)}} if, y_{ij}^{(k)} = a_{(l_i i+s)(l_j j+t)}^{(k)} \\ 0 \end{cases} \quad (3)$$

Here denotes the evaluation function, the error is then back-propagated to the CL, and with it can calculate the gradient of the weight and bias. Gradient of bias is represented as

$$\frac{\partial E}{\partial b^{(k)}} = \sum_{i=0}^{M-m}\sum_{j=0}^{N-m} \frac{\partial E}{\partial a_{ij}^{(k)}} \frac{\partial a_{ij}^{(k)}}{\partial b^{(k)}} \quad (4)$$

$$\partial_{ij}^{(k)} = \frac{\partial E}{\partial a_{ij}^{(k)}} \quad (5)$$

$$c_{ij}^{(k)} = z_{ij}^{(k)} + b^{(k)} \quad (8)$$

Then we get,

$$\frac{\partial E}{\partial b^{(k)}} = \sum_{i=0}^{M-m}\sum_{j=0}^{N-m} d_{ij}^{(k)} \partial_{ij}^{(k)} \frac{\partial a_{ij}^{(k)}}{\partial c_{ij}^{(k)}} \frac{\partial c_{ij}^{(k)}}{\partial b_{ij}^{(k)}} = \sum_{i=0}^{M-m}\sum_{j=0}^{N-m} d_{ij}^{(k)} h'(c_y^{(k)}) \quad (6)$$

In the same way the gradient for weight (kernel) is

$$\frac{\partial E}{\partial w_{st}^{(k,c)}} = \sum_{i=0}^{M-m}\sum_{j=0}^{N-m} d_{ij}^{(k)} h'(c_y^{(k)}) x_{(i+s)(j+t)}^{(c)} \quad (7)$$

When we think for multi-convolutional layers it is necessary to calculate the error of convolutional layers.

$$\frac{\partial E}{\partial w_{st}^{(c)}} = \sum_{k}\sum_{s=0}^{M-m}\sum_{t=0}^{N-m} \frac{\partial E}{\partial z_{i(i-s)(j-t)}^{(k)}} \frac{\partial z_{i(i-s)(j-t)}^{(k)}}{\partial x_{ij}^{(c)}} = \frac{\partial E}{\partial z_{i(i-s)(j-t)}^{(k)}} w_{st}^{(k,c)} \quad (8)$$

So, the error can be expressed as

$$\frac{\partial E}{\partial x_{ijst}^{(c)}} = \sum_{k}\sum_{s=0}^{M-m}\sum_{t=0}^{N-m} \partial_{(i-s)(j-t)}^{(k)} h'(c_{(i-s,j-t)}^{(k,c)}) \quad (9)$$

5. Experiments And Results

The speed and accuracy implementation was benchmarked on a system on *AMD Processor having 4 Core(s) CPU and NVIDIA with GTX520 GeForce, GPU having compute capability 2.1,48 cores running on Ubuntu 14.04*. All training and testing are in single precisions

5.1 Data Set

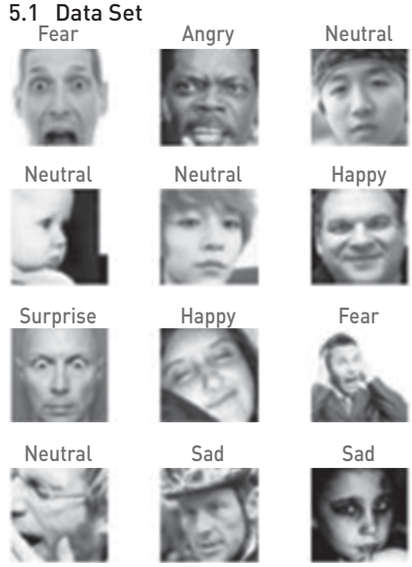


Fig. 7: Example Images from Kaggle dataset

The data we use is comprised of 48 x48 pixel grayscale images of faces from

February 2018.indd 47 2/5/2018 5:55:56 PM

the Kaggle competition Challenges in Representation Learning: Facial Expression Recognition Challenge [5]. The training set consists of 28,709 examples, while both the test and validation sets are composed of 3,589 examples. This data set contains photos and labels of six categories of facial expressions emotions: Anger, Fear, Happy, Sad, Surprise, Neutral. These images have already been somewhat 'preprocessed': they are mostly centered and adjusted so that the face occupies about the same amount of space in each image. Some sample images in the data set is shown in Fig. 7.

5.2 Experiment

We tested CNN for varying number of training epochs with fix batch size of 30 samples.. But to achieve speedup we exploited GPU accelerated deep learning facilities on Theano library in using Python. We fine-tuned the hyper-parameters for the models to optimize performance. The parameters we ended up using are the following:

- Batch size:30
- Learning Rate:0.00001
- Regularization:0.0000001
- Learning Rate decay:0.9999
- Size of filter:5x5

5.3 Results & Evaluation

We implemented the model which was explained in Sec.5. We tried to get the results through validation test. The final validation accuracy we obtained is 77%., which exceeds the oracle method (46%) and mlp-6 (42%) and conv-d (71%) model which is reported in [3]. Basic Comparison (Recall or Accuracy) for three models:mlp-6 model, conv-d and our model is shown in Table. We also benchmarked our results with other methods (10) and (11). The comparative classification accuracies with other methods are shown in Fig 10 .

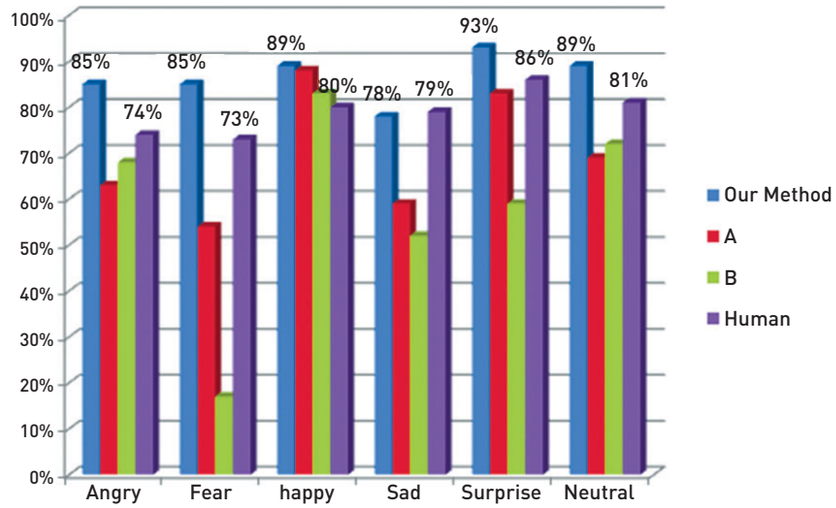


Fig. 8: Classification accuracies of different methods on the FER data set.

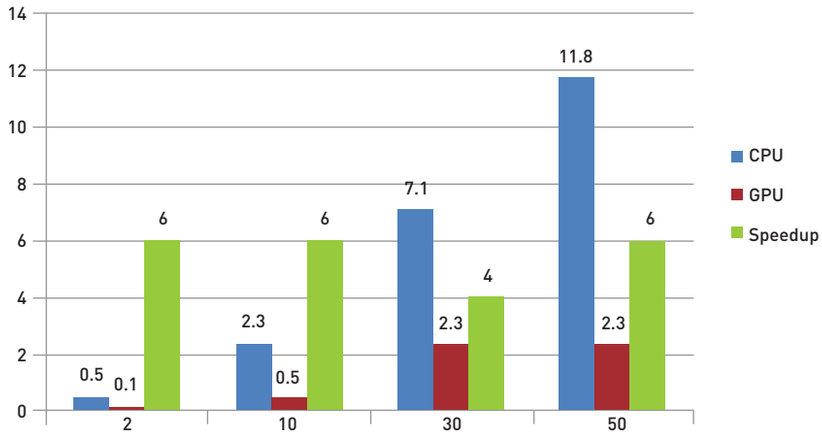


Fig. 9 : Execution Time (in hours) and corresponding Speedup with GPU.

Table : Basic Comparison (Recall or Accuracy) for three models:mlp-6 is the 6 layer mlp model, conv-d is the deep version CNN with our model

	MlP-6v	Conv-d	Our Model
Angry	7.55	68.97	74
Fear	5.31	55.98	71
Happy	79.70	86.20	87
Sad	34.04	52.90	76
Surprise	56.97	81.49	78
Neutral	45.76	68.96	71
Average	38	69	77

The performance speedup gain with GPU for different runs of training epochs are shown in Fig 9.The confusion matrix on the validation set is shown Fig. 10.

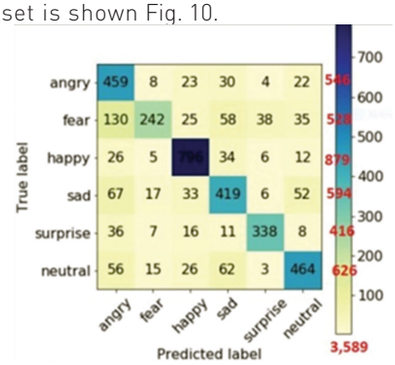


Fig. 10 : Confusion Matrix

February 2018.indd 48 2/5/2018 5:55:56 PM

Conclusions

Our result shows that GPU frameworks with their hierarchical structure of threads are well suited for a parallel implementation of CNN for FER. In comparison with CPU implementation, we were able to get significantly speed up the training algorithm for CNN. The result demonstrated that Deep CNN's are capable of learning facial characteristics and with satisfactory accuracy for facial emotion detection. Some of the difficulties with improving this are that images are very small and some cases it is difficult to distinguish which emotion is on each image. The aim of our future work is to extend this approach to systems with multiple GPUs and to visualize the feature learned by our proposed CNN-deep model by embedding all the testing facial expressions into a two-dimensional space using t-SNE techniques [16]

References

[1] C. C. Claude and F. Bourel, "Facial Expression Recognition: A Brief Tutorial Overview", Stafford: Staffordshire University, 2002.

[2] Geoffrey E. Hinton, Simon Osindero and Yee-Whye Teh, "A Fast Learning Algorithm for Deep Belief Nets", Neural Computation, vol. 18, (7), pp. 1527-1554, 2006

[3] Bergstra, James, Olivier Breuleux, Frédéric Bastien, Pascal Lamblin, Razvan Pascanu, Guillaume Desjardins, Joseph Turian, David Warde-Farley, and Yoshua Bengio,"Theano: A CPU and GPU math compiler in Python", In Proc. 9th Python in Science Conf,2010

[4] Yann LeCun, Corinna Cortes, and Christopher JC Burges,"The mnist database of handwritten digits", 1998.

[5] https://www.kaggle.com/c/challenges-in-representation-learningfacial-expression-recognition-challenge/leaderboard

[6] James Bergstra, Olivier Breuleux, Frédéric Bastien, Pascal Lamblin, Razvan Pascanu, Guillaume Desjardins "Theano: A CPU and GPU Math Compiler in Python". Proc. of the 9th Python in Science Conf. (Scipy 2010). p. 7, 2010.

[7] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey. "Imagenet classification with deep convolutional." Advances in neural information processing systems, 2012.

[8] Rothkrantz, M. Pantic and J.M. "Facial action recognition for facial expression analysis from static face images". IEEE Transactions on Systems, Vol. 34(3), 2004.

[9] Vinod Nair, Geoffrey E. Hinton."Rectified Linear Units Improve Restricted Boltzmann Machines". Toronto: Department of Computer science, University of Toronto.

[10] Zhiding Yu, Zhiding Yu. "Image based Static Facial Expression Recognition with". Seattle, WA, USA : ICMI 2015, November 9–13, 2015,2015. ISBN 978-1-4503-3912-4/15/11.

[11] "Deep Neural Networks with Relativity Learning for facial expression recognition". (ICMEW), IEEE International Conference on Multimedia & Expo Workshops. Seattle : IEEE International Conference on Multimedia & Expo Workshops (ICMEW), 2016. doi: 10.1109/ICMEW.2016.7574736.

[12] Andra Adams, Peter Robinson. "Automated recognition of complex categorical", International Conference on Affective Computing and Intelligent Interaction (ACII), 2015.

[13] H. Kim, H. Nam, W. Jung and J. Lee, "Performance analysis of CNN frameworks for GPUs", IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS), Santa Rosa, CA, 2017.

[14] www.nvidia.com/object/cuda_home.html, 2009

[15] Mo, Kaichun, Yang, Yuxin, Zhang, YunCS 221 Project Final Report:: Facial Expression Automatic Recognition.

[16] L. van der Maaten, "Learning a Parametric Embedding by Preserving Local Structure", Proc. Conf. Artificial Intelligence and Statistics, 2009.

■

About the Authors

Mr. Umesh Chavan [Membership No.N1285682] is currently working as Assistant Professor at Walchand College Of Engineering, Sangli . He has teaching and resrach experience of more than 20 years. He did his Bachelors in Engineering in Computer Science & Engineering from Walchand College. He completed Masters of Engineering in Computer Science and Engineering from Shivaji University, India in 2004. He is currently Professor of Walchand College Of Engineering, Sangli, India. He served as the Head of the Department from January, 2004 to April, 2009 at Bharati Vidyapeeth's College Of Engineering, Kolhapur. Prior to this, he was an Associate Professor of Computer Science & Engineering Department of Bharati Vidyapeeth's College Of Engineering, Kolhapur, India from 2003-2009. Before that, he served as an Lecturer in Computer Science of College Of Engineering, Malegaon (Bk.) affiliated to the University of Pune, India during 1997-2003. He has visited USA and has deceived invited talks at conference. He has more than 10 research publications in international journals and conference proceedings to his credit. His research interests include soft computing, pattern recognition, deep learning, high performance computing.

Dr. Dinesh Kulkarni is currently working as Professor at Walchand College Of Engineering, Sangli . He obtained his Bachelor's Degree in Electronics Engineering from Marathwada Univrsity, Aurangabad. Then He obtained the Master's Degree in Electronic Engineering from Shivaji University, Kolhapur. Also He obtained M. Tech in Computer Science & Engineering from IIT, Bombay. He received Ph.D. degree from Shivaji University, Kolhapur. He has more than 25 years of teaching and research experience and has contributed to several research and consultancy projects. He has received research fellowships from DST's SERC and received post-doctoral research award University of New Brunswick Canada and UGC's. He worked as visiting professor at Technological University, Vienna, Austria.He has visited many countries and has delivered invited talks at conferences. He has published more than 25 papers in National and International Journals. His specialization includes High Performance Computing, Computer Algorithm.

February 2018.indd 49 2/5/2018 5:55:56 PM

Pattern Recognition based approach for Real time recognition of Human Activities in Videos

S N Kakarwal
Professor in Dept. of CS and Engg., P.E.S. College of Engineering, Aurangabad, MS-India. Email: s_kakarwal@yahoo.com

R R Deshmukh
Professor in Dept. of CSIT, Dr. B.A.M. University, Aurangabad, (MS), India. Email: ratnadeep_deshmukh@yahoo.co.in

A. Introduction:

A five year child has no difficulty in identifying different person, animals, plants etc. from their appearances. Also they can analyse the things that are happening around them. To provide similar ability to our computer machine, pattern recognition field came into existence. Biometric verification is one of the applications of pattern recognition. From last few decades lots of work has been done on various biometric verification such as fingerprint recognition, face recognition, eye retina recognition etc. and many of this biometric techniques are now going to use successfully for person identification particularly for attendance monitoring, secure authentication system, criminal verification etc.

Automatic identification Crops, crop disease, nutation deficiency is another growing field in pattern recognition. One of the most emerging topic in pattern recognition domain is "Human activity recognition /prediction from live streaming of pre-recorded videos.

Now a day's surveillance camera is being used everywhere for the purpose of security and for monitoring some abnormal and criminal activities. When some harmful activity is happened the operator need to check all the footage of CCTV camera. This process needs lots of time and concentration to watch all footage and then to investigate the case based on it. Also such system always needs human intervention to monitor the activity recorded in camera after some illegal activity is happened. So researchers have drawn much attention towards developing intelligent video surveillance system which can automatically recognize or predict the ongoing human activity in video. The goal of activity recognition is to automatically understand the ongoing human activity in video. The basic difference between activity recognition and activity prediction is that, activity recognition is after-the- fact classification. It mean that the activity is recognized after it is fully happened while activity prediction means guessing (or predicting) the type of ongoing activity before it is fully observed. In short, activity prediction means inferring the ongoing activity given temporally incomplete video. Such as: in intelligent prediction system can be useful in real world scenario such as healthcare monitoring system for automatically analysing and understanding patient's activities, in intelligent driver assistance system for monitoring driver activities, in ATM surveillance for detecting criminal activities etc.

B. Challenges in Activity Recognition

Activity recognition from videos is very challenging task as unlike image video is dynamic in nature which consists of spatial as well as temporal information. Various challenges in activity recognition are as following:

1. **Viewpoint:** Changes in viewpoint of camera can affect the performance of prediction system. The person's activity may look different from different viewpoint and camera locations.
2. **Background:** If the background is clutter then at first it needs to extract human from its background. Object segmentation from video is again highly depends on camera position i.e. static or moving camera.
3. **Motion Style:** The same activity can be performed in different style by different person. For example some person walks very fast so it may seems similar to running activity of another person.
4. **Speed:** The speed at which human performs the activity may differ from person to person.
5. **Duration:** Performance duration can vary every time, so size of extracted features can vary.
6. **Occlusions:** Occluded part in video is often difficult to monitor. If the video is occluded then it becomes difficult to track targeted human activity which may affect the accuracy of recognition system.
7. **Scale:** The scale of the persons in video can affect the performance of the system.

C. Human Activity Recognition system:

There are lots of variations in human action recognition. For example an activity can be simple short duration activity such as walking, running, sitting, jogging etc or it could be long duration complex activity such as cooking, playing, dancing etc. Long duration complex activity consists of series of short duration activity. Various scenarios for activity recognition system are as follows:

1. Single Person Activity Recognition

This system works for the video which consists of single person. When a video consists of single person then it is easier to track and monitor that person. Such recognition system can be useful to ATM Surveillance, Patient's Activity

February 2018.indd 50 2/5/2018 5:55:57 PM

monitoring in ICU room, Driver action recognition System etc.

2. Multiple person Interaction or crowd activity recognition

This system works for environment when videos which consist of Multiple people interaction and the crowd behaviour. Such Recognition task is very difficult and challenging and it have drawn much attention recently due to the needs of environment security. The main three important issues in such system include people counting, tracking each individual separately and then analyse activity of each individual separately. Such systems can be useful for public places such as hospitals, office etc where limited number of peoples are present in video.

3. Abnormal Activity Recognition:

This system is helpful for recognizing abnormal or criminal activities at public places such as malls, bus stands, railway stations etc. To explicitly define abnormal activity is very challenging, since their definition depends on the contexts and surrounding environments.

D. Application

1. Surveillance System

In surveillance systems human activity recognition can be used for identifying the criminal and detecting suspicious activities. This needs automatically tracking of every individual in crowd and understands activities.

Now days most security surveillance systems are equipped with several cameras and it require human intervention for monitoring of video and for video content understanding when some abnormal or criminal activity is reported. By applying automatic human activity recognition techniques to video-based surveillance systems, we can effectively reduce the workload of security staff as well as systematically creating an alert immediately when security events are detected in order to prevent potentially dangerous situations.

2. Healthcare System

In healthcare environment this system can be used to continuously track patients in ICU and immediately raise an alarm if some critical situation arises.

3. Entertainment Environments

In Entertain environment such recognition system can be used to analyse and predict success rate of sport, dance and gaming etc. in order to enrich lifestyles

4. Daily Life Activity Monitoring

This system can be effectively used to monitor the activity of seniors at home. This would be useful for the safety and immediate help to senior persons.

E. Conclusion:

To develop a system for automatic recognition human action is crucial task because of diversity in different actions performed by different people. Today's word is very fastest growing world and everyone wants enriched life style with emerging technologies along with safety. Real time human action recognition system to investigate the doubtful situations at their early stages and can also reduce human effort and can also help to provide safety.

F. References

[1] ManojRamanathan, Wei-Yun Yau, and Eam Khwang Teoh, "Human Action Recognition With Video Data: Research and Evaluation Challenges", IEEE Transactions On Human-Machine Systems, Vol. 44, No. 5, pp650 - 663 ,October 2014.

[2] Yang Wang and Greg Mori, "Hidden Part Models for Human Action Recognition: Probabilistic versus Max Margin", IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 33, No. 7, pp. 1310-1323, July 2011

[3] Qiuxia Wu, Zhiyong Wang, , Feiqi Deng, , Zheru Chi, , and David Dagan Feng," Realistic Human Action Recognition With Multimodal Feature Selection and Fusion", IEEE Transactions On Systems, Man, And Cybernetics: Systems, Vol. 43, No. 4, 875 - 885 ,July 2013.

[4] An-An Liu, , Yu-Ting Su, Ping-Ping Jia, ZanGao, Tong Hao, and Zhao-Xuan Yang, "Multipe/Single-View Human Action Recognition via Part-Induced Multitask Structural Learning", IEEE Transactions On Cybernetics, Vol. 45, No. 6, 1194 – 1208, June 2015.

[5] H. Wang et al., "Action Recognition by Dense Trajectories," Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 3169-3176,2011.

[6] Y. Cao, "Recognizing Human Activities from Partially Observed Videos," Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp 2658 - 2665,2013.

[7] M.S. Ryoo, "Human Activity Prediction: Early Recognition Ongoing Activities from Streaming Videos," Proc. IEEE Int'l Conf. Computer Vision, pp. 1036-1043, 2011.

[8] M. Pei, Y. Jia, and S.-C. Zhu, "Parsing Video Events with Goal Inference and Intent Prediction," Proc. IEEE Int'l Conf. Computer Vision, pp. 487-494, 2011.

[9] Kang Li and Yun Fu, "Prediction of Human Activity by Discovering Temporal Sequence Patterns", IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 36, No. 8,pp: 1644 – 1657, August 2014

■

About the Authors

Dr. S. N. Kakarwal, received Ph.D., M.E. and B.E. degree in Computer Science and Engineering. She is presently working as Professor in Department of Computer Science and Engineering, P.E.S. College of Engineering, Aurangabad, MS-India. She is having 19 years of experience at graduate and postgraduate levels. Her research interests include Image Processing, Pattern Recognition and Artificial Neural Network. In these areas, she has published 50 research papers in leading Journals, National and International conferences proceedings. She is member of IEEE, Life member of ISTE and also member of CSI (F8000602). She can be reached at s_kakarwal@yahoo.com.

Dr. R. R. Deshmukh, [CSI Membership No. 00100518], is currently working as Professor in Dept. of CSIT, Dr. B.A.M. University, Aurangabad, (MS), India. He has been elected as sectional member of ICT section of Indian Science congress Association. His areas of specialization are Human Computer Interaction, Digital Speech Signal processing, Computational Auditory Scene Analysis (CASA), Neural Networks etc. He can be reached at ratnadeep_deshmukh@yahoo.co.in.

February 2018.indd 51 2/5/2018 5:55:57 PM

Registered with Registrar of News Papers for India - RNI 31668/1978
Regd. No. MCN/222/2018-2020
Posting Date: 10 & 11 every month. Posted at Patrika Channel Mumbai-I
Date of Publication: 10th of every month

If undelivered return to :
Samruddhi Venture Park, Unit No.3,
4th floor, MIDC, Andheri (E). Mumbai-400 093

CSI Nihilent eGovernance Awards 2017 - A Brief Report

20th January 2018, Science City, Kolkata

The Special Interest Group on eGovernance of Computer Society of India (CSI-SIGeGov), successfully conducted the 15th edition of CSI Nihilent eGovernance Awards (CNeGA) ceremony on 20th January 18, during the 52nd CSI Annual Convention in Kolkata. The awards ceremony is a culmination of an elaborate awards process spread over 9 months starting with nominations and ending with awards presentation to the successful entries. Sponsored by Nihilent Technologies Ltd., since the year 2002, the awards ceremony has become integral part of the CSI Annual Convention. This year's event took place at Science City, Kolkata.

To make the event rewarding, an eGovernance track sessions were organized as a parallel session of the CSI 2017 program. This session coordinated by Shri Apoorva Agha, Chairman Division I, CSI was inaugurated by Sri Gautam Mohapatra, Vice President, CSI. The first panel discussion was on 'Digital Empowerment in Agriculture, Education & Health Sectors - Implementing challenges' chaired by Dr G P Sahu (MNNIT, Allahabad). The second session was on 'Digital Empowerment - Learnings and Way forward (Policy Perspectives)' chaired by Dr Ashok Agarwal, Founder CSI SIGeGov and co chaired by Shri D K Dwivedi, a Senior member of the CSI SIG eGov. The third and final session was on 'Research, Innovation & Use of Technology in eGovernance [Partner's perspective]' moderated by Shri Satish Babu, Former President CSI and co chaired by Dr Prateep K Guha, MD Nihilent Data Analytics. Panelists included senior officials of the award winning teams. All the sessions were well attended by the award winners and accompanying teams.

The much awaited, Awards ceremony took place in the Mini Auditorium. Attended with over 280 participants consisting of 45 awardees with their teams and other invitees, the entire program has been well received and appreciated for the ambience and the professional way it got conducted.

Shri GSN Prabhu, Convener, CSI-SIGeGov welcomed the Chief Guest, Gusts of Honor, all the awardees and the invitees. He also thanked all the nominees for continued patronage and participation in the CSI Nihilent eGovernance Awards process and complemented the award winners.

The highlight of this year's ceremony is the presence and participation of Shr Rao Rajendra Singh, Deputy Speaker of Rajasthan Legislature Assembly, as the Chief Guest, who also distributed the awards.to deserving award winners under different award categories from across the States and Central Departments. The distinguished presence of Guests of Honor Dr Amit Mitra, Finance Minister of Government of West Bengal and Prof Mike Hinchey, President, IFIP on the dais, added glitter to the flagship event of the CSI SIG eGov and the annual CSI convention.

For his distinguished contribution and continued patronage of CSI Nihilent eGovernance Awards for 15 long years, Shri L C Singh, Vice Chairman and CEO of Nihilent Technologies was presented with an Award of Recognition by Shri Sanjay Mohapatra, President, CSI. Shri Prateep K Guha, MD of Nihilent Analytics representing Shri L C Singh accepted the same and thanked the Organizers for the recognition on behalf of Shri L C Singh.

Dr K S Vijaya Sekhar, Awards convener presented the Convener's report and stated that the Selection Committee had received over 170 nominations and that identifying the best from the list has been very tough task but for the able support and participation of the Reviewers in three stages of evaluation process.

Shri Sanjay Mohapatra, President-CSI, who presided over the event opined that the foot print of participation from across all the States and Central Government departments is very encouraging and assured CSI will continue to play a pivotal role in such recognition journey that compliments national efforts in bringing 'Digital Empowerment' in the country.

Dr Ashok Agarwal, Founding Member CSI SIG eGov and Dr Prateek Kumar Guha, MD Nihilent Analytics spoke on the occasion.

A compendium in the form of a book titled Digital Empowerment – A Cornerstone for eGovernance was released along with a Souvenir by the Chief Guest on the occasion. The book edited by Dr K S Vijaya Sekhar, Dr G P Sahu and Shri Prabhu Gollamudi consists of over 75 articles that talk of successful eGovernance initiatives from across the country.

The awards program concluded with a Group Photo of awardees along with the Chief Guest and Guest of Honor on the dais. Dr K S Vijaya Sekhar, profusely thanked the award winners for their participation and also thanked the CSI-2017 Local Organizing Committee and Program Committee lead by Dr Phalguni Mukherjee, Prof J K Mandal, Prof D P Sinha and others for their immense support to make the event memorable. He also thanked Nihilent Technologies for the continued support and thanked Mr Shohel Noor, Mr Anoop Bharadwaj for excellent coordination and support during and prior to the event.

This year's event also included presentation of CSI's eRatna award to two distinguished personalities heading two large and important States of Rajasthan and West Bengal for their significant contribution and initiatives towards driving the eGovernance journey in their respective States. Smt Vasundhara Raje, Hon'ble Chief Minister, Rajasthan and Smt Mamatha Banerjee, Hon'ble Chief Minister of West Bengal were honored with the eRatna Awards by President CSI and the awards were received by Shri Rao Rajendra Singhji, Dy Speaker, Rajasthan Legislature Assembly and Dr Amit Mitra, Minister Finance, Govt of West Bengal on behalf of the respected dignitaries.

CSI Nihilent eGovernance Awards 2017 Winners List

(i) State Category (4)
Award of Excellence: Rajasthan (Overall Growth)
Award of Appreciation: Gujarat (Overall Growth)
Award of Recognition: Odisha (Food and Security)
Award of Recognition: Telangana (Health and Wellbeing)

(ii) Theme Category (2)
Award of Excellence: SHE Teams, Hyderabad
Award of Recognition: Raj Mahila Suraksha, Rajasthan

(iii) Project Category (39)
(a) Award of Excellence 3
State Level Projects (3)

(b) Award of Appreciation 17
Central Govt. Projects (3)
State Level Projects (12)
District Level Projects, ULBs (2)

(c) Award of Recognition 16
Central Govt. Projects (2)
State Level Projects (14)

(d) Award of Sustenance 3
Award of Excellence Projects for the year 2012
Total (i+ii+iii): 45 Awardees

More Details are available at www.csinihilent-egovernanceawards.org

- Dr. K S Vijaya Sekhar & GSN Prabhu, Convener's Team, CNeGA 2017