SLSWCS
Student Workshop on Computer Science

2019

Sri Lanka Student Workshop on Computer Science

07 December 2019

Department of Computer Science,
University of Jaffna

99X Technology WSO2 virtusa Accelerating Business Outcomes

SLSWCS

Student Workshop on Computer Science

Sri Lanka Student Workshop on Computer Science

07 December 2019

Department of Computer Science

University of Jaffna

99X Technology WSO2 virtusa

Contents

Agenda

08:15 – 08:45	**Registration**
08:45 – 08:55	**Welcome Address** Dr. A. Ramanan General Chair/SL-SWCS'19
08:55 – 09:00	**Dean's Address** Prof. J.P. Jeyadevan Dean/Faculty of Science, University of Jaffna
09:00 – 09:50	**Keynote Speech - I** Dr. Sanjiva Weerawarana Founder, Chairman & Chief Architect, Lanka Software Foundation
09:50 – 10:15	**Tea Break**
10:15 – 10:30	**Oral Presentation** Ms. F. Dushmanthi
10:30 – 10:45	**Oral Presentation** Ms. K. Archchitha
10:45 – 11:00	**Oral Presentation** Ms. H.P. Malki Maduka
11:00 – 11:50	**Keynote Speech - II** Prof. Mahesan Niranjan University of Southampton, UK.
11.50- 12:05	**Oral Presentation** Ms. T.C. Kasthuriarachchi
12:05- 12:20	**Oral Presentation** Ms. R. Nirthika
12:20– 12:35	**Oral Presentation** Mrs. S. Majuran
12:35 – 13:15	**Lunch Break**
13:15 – 14:05	**Keynote Speech - III** Dr. G.A. Nalin Asanka La Trobe University, Melbourne, Australia.
14:05 – 14:30	**Spotlight Session**
14:30 – 16:30	**Poster Session**
16:30	**Awarding and Closing Ceremony**

Message from the General Chair

As the founder of the Sri Lanka Student Workshop on Computer Science (SL-SWCS), I have the honour to welcome you all to the fifth national workshop aimed at bringing together research students, leading academics, and industry leaders across the country and foreign institutions to meet and share ideas with stimulating discussions in computing areas. The unique structure of the SL-SWCS allows students to interact with academic and industrial communities that help to identify and explore areas of mutual interests for collaboration. The SL-SWCS is conducted as a biennial workshop since 2011. Moreover, I am glad to inform you that SL-SWCS'19 has obtained recognition from the IEEE Sri Lanka Section this year.

We are delighted and honoured to have Prof. Mahesan Niranjan (Professor of Electronics and Computer Science at the University of Southampton, United Kingdom), Dr. Sanjiva Weerawarana (Founder and Chairman of Lanka Software Foundation, Co-Founder and Former CEO of WSO2), and Dr. Nalin A.G. Arachchilage (Senior Research Fellow at La Trobe University, Melbourne, Australia) with us as our keynote speakers. Heartfelt thanks to our keynote speakers who in spite of their busy schedule manage their times and have kindly agreed to deliver highly stimulating talks. We are also glad to have Dr. Jeevani Jayasinghe the Secretary of the IEEE Sri Lanka Section as an evaluator for the poster presentations.

SL-SWCS'19 has received 48 posters from young research students. All posters were evaluated on the originality, presentation, and empirical results by local and foreign institutional reviewers in the field of computing. Each poster has been reviewed by two to three reviewers and based on the reviews we have accepted 32 posters out of 48 for today's event. I wish to express my sincere gratitude to the reviewers for freely providing their time in reviewing posters and to share their constructive feedback to those students.

We could not have built such an event without the help and guidance from the members of our organising committee of SL-SWCS'19 and the administration of the University of Jaffna. SL-SWCS'19 is also grateful to its sponsors: 99X Technology, WSO2, and Virtusa for their generous support. The cash prize to the student winners of SL-SWCS'19 is sponsored by the Department of Mechanical Engineering, University of Melbourne, Australia, through the generous support of Prof. Saman Halgamuge from the same Department.

I hope this workshop will be joyful and provide all our students with a good opportunity to network, communicate the results of their research, and promote international collaboration. On behalf of the SL-SWCS'19 organising committee, I thank all the reviewers, keynote speakers, sponsors, the IEEE Sri Lanka Section, volunteers, and student authors in making SL-SWCS to be a benchmark for the future events in computing of this University.

Dr. A. Ramanan (General Chair/SL-SWCS'19) 07 Dec 2019

Keynote Speaker – I

Dr. Sanjiva Weerawarana

Founder, Chairman & Chief Architect, Lanka Software Foundation

Dr. Sanjiva Weerawarana completed his Bachelor's degree in 1988, and he was always focused on creating software technology tools that help other technologists build solutions in various domains. He was at the center of several major technology waves in that period including the Web, Web services, cloud computing and data analytics. His recent technology creation contribution is Ballerina, https://ballerina.io, a new programming language optimized for writing network distributed applications.

In 2018, after 30 years of creating technology, he changed his focus to applying software to make the world a better place. His immediate focus is on improving the technology that runs Sri Lanka by voluntarily helping the Government of Sri Lanka to digitally transform itself. However, all the work he is doing is intended to be useful for other parts of the world as well since all countries face similar challenges. This work is being done as part of the Code for Sri Lanka project of the Lanka Software Foundation, a non-profit organization that he co-created in 2003.

Keynote Speaker – II

Prof. Mahesan Niranjan

University of Southampton, UK

Mahesan Niranjan is Professor of Electronics and Computer Science at the University of Southampton, UK. Prior to this appointment in 2008, he has held faculty positions as Lecturer in Information Engineering at the University of Cambridge and as a Professor of Computer Science at the University of Sheffield. At Sheffield, he has also served as Head of the Department of Computer Science and Dean of the Faculty of Engineering. He has a long track record of research in Machine Learning, and has contributed to both the algorithmic and applied aspects of the subject. His current focus of research is in inference problems in the domain of Computational Biology.

Keynote Speaker – III

Dr. Nalin A. G. Arachchilage

La Trobe University, Melbourne, Australia

Dr. Nalin Asanka Gamagedara Arachchilage is a Senior Research Fellow (Research Associate Professor) in Cyber Security within the Department of Computer Science and Information Technology at La Trobe University, Australia, where he currently leads the Usable Security Engineering Research Group (USERGroup) in the Optus La Trobe Cyber Security Research Hub. Previously, he was a Lecturer in Cyber Security in the School of Engineering and Information Technology of the University of New South Wales at the Australian Defence Force Academy (ADFA), where he led the Usable Security research group. He holds a PhD in Computer Science (Usable Security) from Brunel University London, UK. He worked as Research Fellow in Usable Security and Privacy in the Laboratory of Education and Research in Software Security Engineering (LERSSE) at the University of British Columbia (UBC), Canada. Before moving to Vancouver, he was a Postdoctoral Researcher in Systems Security Engineering in the Cyber Security Centre, Department of Computer Science at Oxford University. Nalin has presented his research at Facebook Headquarters, Menlo Park, California, USA and collaborated with HP in a research capacity at the HP Lab, Bristol, UK. His research has been featured in numerous media outlets including ABC News Radio, Sky News Australia and UNSW TV.

List of Reviewers

Dr. (Mrs.) Amalka J. Pinidiyaarachchi	Department of Statistics and Computer Science, University of Peradeniya, Peradeniya, Sri Lanka.
Mr. T. Arudchelvam	Department of Computing & Information Systems, Wayamba University of Sri Lanka, Kuliyapitiya, Sri Lanka.
Dr. Chulantha Kulasekere	Faculty of Engineering, Sri Lanka Institute of Information Technology (SLIIT), Malabe, Sri Lanka.
Prof. Clinton Fookes	Faculty of Science and Engineering, Queensland University of Technology (QUT), Australia.
Dr. H.M.N. Dilum Bandara	Department of Computer Science & Engineering, University of Moratuwa, Moratuwa, Sri Lanka.
Prof. Duc Truong Pham	Department of Mechanical Engineering, University of Birmingham, United Kingdom.
Prof. Gihan Dias	Department of Computer Science and Engineering, University of Moratuwa, Moratuwa, Sri Lanka.
Dr. T. Ketheesan	Faculty of Technology, University of Jaffna, Sri Lanka.
Prof. Kitsuchart Pasupa	King Mongkut's Institute of Technology Ladkrabang (KMITL), Thailand.
Dr. T. Kokul	Department of Physical Science, Vavuniya Campus of the University of Jaffna, Sri Lanka.
Prof. Mahesan Niranjan	School of Electronics and Computer Science, University of Southampton, Southampton, United Kingdom.
Dr. R. Nagulan	Department of Physical Science, Vavuniya Campus of the University of Jaffna, Sri Lanka.
Dr. Nalin A.G. Arachchilage	Department of Computer Science and Information Technology, La Trobe University, Australia.

List of Reviewers ...

Dr. D. Nalin Ranasinghe	University of Colombo School of Computing (UCSC), Colombo, Sri Lanka.
Dr. Y. Pratheepan	School of Computing, Engineering & Intelligent Systems, Ulster University, Northern Ireland, United Kingdom.
Mr. M. Ramanan	Department of Computer Science, Trincomalee Campus of Eastern University, Sri Lanka.
Prof. Roshan G. Ragel	Department of Computer Engineering, University of Peradeniya, Sri Lanka.
Dr. A. Ruvan Weerasinghe	University of Colombo School of Computing (UCSC), Colombo, Sri Lanka.
Prof. Saluka R. Kodituwakku	Department of Statistics and Computer Science, University of Peradeniya, Peradeniya, Sri Lanka.
Prof. Saman Halgamuge	Department of Mechanical Engineering, University of Melbourne, Melbourne, Australia.
Dr. S. Sivasuthan	Cummins Allison, Illinois, USA.
Mr. S. Sotheeswaran	Department of Mathematics, Eastern University, Sri Lanka (EUSL), Sri Lanka.
Dr. S. Suganthan	Senior Image Processing Engineer, De Beers Technologies, Maidenhead, United Kingdom.
Mr. S. Thirukumaran	Department of Physical Science, Vavuniya Campus of the University of Jaffna, Sri Lanka.
Dr. Thrishantha Nanayakara	Faculty of Engineering, Imperial College London, United Kingdom.

Accepted Posters

Poster ID	Title	Authors
1	Opinion Spam Detection in Online Reviews Using Neural Networks	Archchitha, K. and Charles, E.Y.A.
2	Silhouette Image Classification Using Bag of Local Features	Bamini, T. and Mayurathan, B.
3	A Community Based Routing Algorithm for Mobile Opportunistic Networks	Fernando, D. and Thabotharan, K.
4	A Lossy Grayscale Image Compression based on Delaunay Triangulation	Gunasekara, T.M.V.D. and Ramanan, A.
5	Tamil Font Type Identification from Text Images	Janani, P. and Charles, E.Y.A.
6	A Performance Evaluation Study of Selected TCP Protocols on Wireless AD-HOC Networking Environments	Kaluarachchi, K.A.K.M. and Thabotharan, K.
7	Deep Learning Approach to Detect Plagiarism in Sinhala Text	Kasthuriarachchi, T.C. and Charles, E.Y.A.
8	A Novel Approach for Tamil - English Translation and Vice versa Using RNN	Kasthuriraajan, R. and Mahesan, S.
9	Predicting the Outcome of the Cricket Matches Using Machine Learning Techniques	Kausik, M. and Siyamalan, M.
10	Flower Classification Using Multiple Feature Set	Kishotha, S. and Mayurathan, B.
11	Detection of Red Ripe Tomatoes on Plants Using Image Processing Techniques	Kshithija, T.G.A.D. and Mayurathan, B.
12	Automatic Facial Makeup Detection	Ligitha, Y. and Ramanan, A.
13	A Hybrid Data Forwarding Approach for Opportunistic Networks	Malki Maduka, H.P. and Thabotharan, K.
14	Analysis of Methods to Handle Medical Sensor Data Towards Health Disorder Identification	Meruja, S. and Charles, E.Y.A.
15	A Web-based Dengue Monitoring and Warning System	Nirthika, R., Ramanan, A. and Surendran, S.N.
16	Is Soft Pooling better than Max and Average Pooling? A Comparative Study on HEp-2 Cells and Retinal Image Classification Tasks	Nirthika, R., Siyamalan, M. and Ramanan, A.

Accepted Posters ...

Poster ID	Title	Authors
17	Copy-Move Image Forgery Detection using SIFT Descriptors	Parkavi, K. and Ramanan, A.
18	An Attention-based Convolutional Neural Network for Landmark Recognition in Asian Region	Perera, S. and Ramanan, A.
19	Fake News Detection	Prithweeraj, R. and Mayurathan, B.
20	An Improved Approach of Iterative Keypoint Selection with Spatial Pyramid Matching for Visual Object Classification	Ranathunga, R.M.S. and Ramanan, A.
21	Image Reconstruction Using Spatial and Geometrical Information	Ranushka, P. and Mayurathan, B
22	Solar Energy Forecasting with Machine Learning Approaches	Sabbir Hossain, M.D and Siyamalan, M.
23	Unsupervised Sentiment Analysis on Tamil Texts	Sajeetha, T. and Mahesan, S.
24	Sentiment Analysis on Tamil Texts Using K-means and k-Nearest Neighbor	Sajeetha, T. and Mahesan, S.
25	A Robust Parallel Implementation of Active Contours	Saranya, B and Suthakar, S.
26	A Novel Approach of Voice Recognition Using MFCC and GMM, Speech Recognition and Text Recognition to Assist for Email Communication for Visually Impaired People	Senthuja, K. and Mahesan, S.
27	A Multiscale Contextual Technique for Fashion Clothes Landmark Localisation	Shajini, M. and Ramanan, A.
28	HEp-2 Specimen Classification Using Deep CNN	Shawmiya, Y. and Siyamalan, M.
29	Speech Emotion Recognition Using Deep Learning on Audio Recordings	Suganya, S. and Charles, E.Y.A.
30	A Deep Learning Approach for Anomaly Detection in Data Communication Network	Thameera, T. and Thabotharan, K
31	Action Recognition in Videos Using Convolutional and Spatial-Temporal Interest Point Features	Tharmini, T. and Ramanan, A.
32	An Efficient Approach for Patch-based Visual Object Classification	Vinoharan, V. and Ramanan, A.

Opinion Spam Detection in Online Reviews using Neural Networks

K. Archchitha#1, E.Y.A. Charles#2,
#Department of Computer Science, University of Jaffna
Jaffna, Sri Lanka
1archchithak@gmail.com
2charles.ey@univ.jfn.ac.lk

Introduction

- More than 2 billion people use e-commerce websites.
- 91% of consumers read online reviews for local businesses either regularly occasionally.
- Writing positive reviews come with monetary gain.
- Opinion spam activities have increased dramatically on online review websites.

Objective

→ Opinion spam detection mechanism can be stated as identifying whether a review is fake or truthful.

→ This research work:

- Proposes a convolution neural network based model to detect opinion spam
- Analyses the effect of word embedding features and text based features

Existing Solutions

Study	Features	Learning method	Dataset
Ott et al. (2011) Finding Deceptive Opinion Spam by Any Stretch of the Imagination	- genre identification - detection of psycholinguistic deception - text categorization	- SVM - NB	Deceptive opinion spam corpus
Song Feng et al., (2012) Syntactic Stylometry for Deception Detection	- POS features driven from Context Free Grammar (CFG) parse trees	- SVR classifier	Deceptive opinion spam corpus
Vlad Sandulescu and Martin Ester (2015) Detecting Singleton Review Spammers Using Semantic Similarity	- bag-of words and topics extracted from the reviews text with parts-of speech (POS) as patterns	- Latent Dirichlet Allocation (LDA) model	- Deceptive opinion spam corpus

Data Set

→ Publicly available gold standard corpus of deceptive opinion spam.

→ This corpus contains:

- 400 truthful positive reviews from TripAdvisor
- 400 deceptive positive reviews from Mechanical Turk
- 400 truthful negative reviews from Expedia, Hotels.com, Orbitz, Priceline, TripAdvisor and Yelp
- 400 deceptive negative reviews from Mechanical Turk

Novelty of proposed method

	Existing works	Proposed work
Feature Extraction	**Bag of Word vectorization** Words are themselves have meanings associated with them is not adequately represented.	**Word Embedding** Inherently capture the context of word, semantic and syntax similarity and relation with other words.
Learning method	**Classification using traditional method (SVM, NB)** Not scalable	**Classification using Neural Network** Scalable (appropriate for large dataset)

Methodology

Data preprocessing

- Reviews in the data set are cleaned and punctuations removed.
- Then each review text is transformed into a sequence of tokens and the text sequence is zero padded to make all the sequences to have fixed length M.

Feature Representation

- Tokens are mapped to their respective embedding using the GloVE pre-trained word embedding.
- Each token is represented with a vector of N dimension. In order to use text documents to train a machine each of the preprocessed review is converted into numerical features.
- The pre-trained GloVe model provides a dense vector of dimensions 50, 100 and 300 for each English word in a very large corpus.

Model construction & Training

- Build a CNN model with three parallel convolution layers with different filter sizes for text classification.
- The model is trained using the data set of review text where each review text is represented by a vector of M x N dimension.

Performance improvement

- The study also tried to verify the effect of incorporating the traditional text based features in addition to the word embedding features.
- Features tested are number of words, number of characters, number of stop words, number of special characters, number of upper cases and some part of speech tags frequencies.
- The selected features are concatenated with the features obtained from the flatten layer.

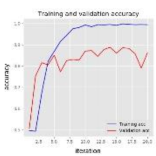
Experimental setup & Testing results

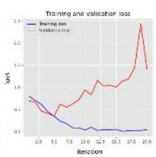
model construction: Keras open source library using Tensorflow backend.

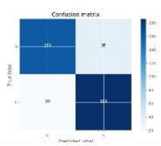
optimization algorithm: Adam

loss function: binary cross-entropy

Training Accuracy: 100%
Testing Accuracy: 86.25%





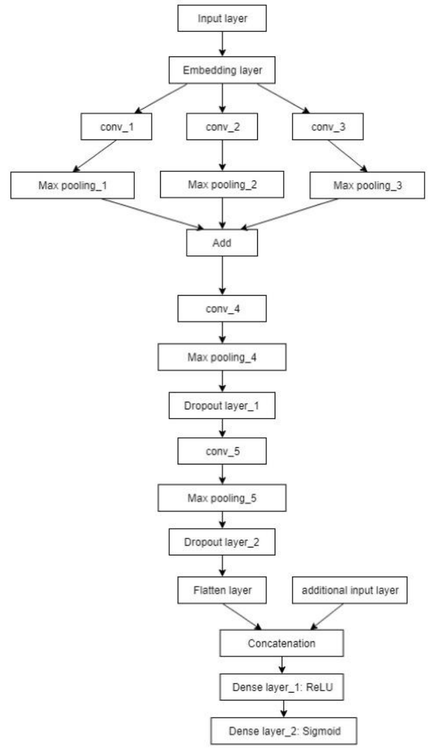


Evaluation result

evaluate the accuracy of the proposed models, test score and confusion matrices are used

Features	Training accuracy	Testing accuracy
Count vectors as features	99.50%	81.25%
TF-IDF as features	99.25%	80.50%
Word embedding as features	100.00%	86.25 %
Word embedding + additional features	99.58%	88.25%
Word embedding + TF-IDF features	99.92%	88.75%
Word embedding + Count vectors features	99.83%	88.50%

CNN model



Discussion & Conclusion

- In this study, a convolution neural network based model is developed to detect opinion spam with the pretrained GloVe word embedding model which is used to effectively represent the review content features.
- Moreover some useful additional features which are applied in the existing research are concatenated with the model.
- By making a feature combination, the accuracy was enhanced to 88.75%.The proposed model outperforms the state-of-the-art method and the additional features improved the performance.

References

1. Myle Ott, Yejin Choi, Claire Cardie and Jeffrey T. Hancock, "Finding Deceptive Opinion Spam by Any Stretch of the Imagination", 2011
2. Yoon Kim, "Convolutional Neural Networks for Sentence Classification", Empirical Methods in Natural Language Processing (EMNLP),2014
3. Jeffrey Pennington, Richard Socher, Christopher D. Manning, "GloVe: Global Vectors for Word Representation", in Empirical Methods in Natural Language Processing (EMNLP), pages:15321543, 2014.

Silhouette Image Classification using Bag of Local Features

T. Bamini and B. Mayurathan
Department of Computer Science, Faculty of Science, University of Jaffna
Baminit@univ.jfn.ac.lk , barathym@univ.jfn.ac.lk

Abstract

Shape is an important feature to identify an object in the image. Classifying objects by using their shape has been an interesting and important area in computer vision. It has improved a lot in the last decades.
In this work, an approach is proposed for shape classification that uses both local and global image representation using Histogram of Oriented Gradient (HOG) Pyramid of Histogram of Orientation Gradients (PHOG) and CENsus Transform hISTogram(CENTRIST).
By evaluating these descriptors it can be concluded that the combined HOG, PHOG gives better performance than CENTRIST in the context of silhouette image classification.

Introduction

- Shape feature is undoubtedly transcending landmark in its ability to produce a complete description of an object where texture or color cannot be used as a cue for recognition.
- Shape representation methods can be classified into two main categories: Contour-based methods and Region-based methods.
  - Contour-based shape techniques use shape boundary information.
  - Region based techniques use all the pixels within a shape region are taken into account to obtain the shape representation.
  - ✓ Contour-based methods are widely used than region-based methods since individuals can easily identify shapes by using their contour features.
- There are three levels of feature extraction such as pixel, global and local.
  - By using pixel values of the image we can identify the features.
  - The global feature can be extracted to describe the whole image.
  - Local features are extracted from small sub region of interest from the original image.
  - ✓ Local features can improve the computational speed and may focus on the object rather than the background.

Methodology

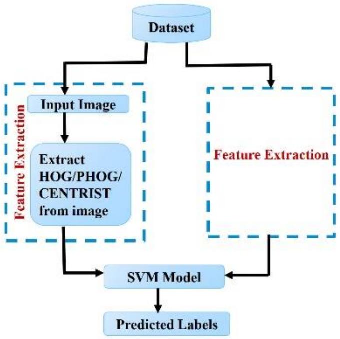


Figure 1: Global representation of image

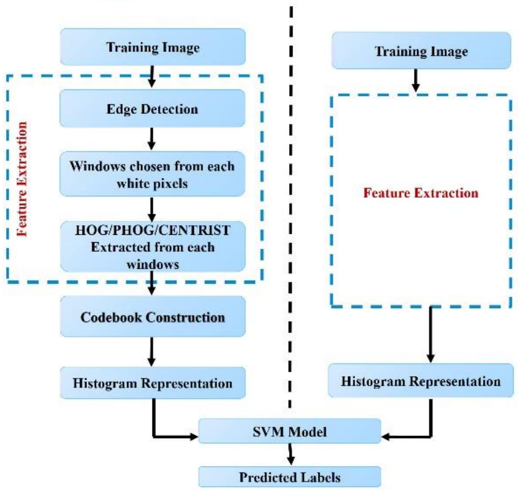


Figure 2: Local representation of image

Testing Results

Descriptor	Performance
HOG	84.14
PHOG	82.29
CENTRIST	55.43
HOG + PHOG	84.14
HOG + CENTRIST	79.86
PHOG + CENTRIST	55.43

Table1: Classification performance for Global representation of image

Descriptor	31,31 sized window	61,61 sized window
HOG	78.14	80.00
PHOG	79.76	79.71
CENTRIST	72.00	74.57
HOG + PHOG	92.14	93.14
HOG + CENTRIST	76.14	80.14
PHOG + CENTRIST	89.14	91.57

Table2: Classification performance for Local representation of image

Methods		Performance (%)
Proposed method	HOG	84.15
	PHOG	82.29
	CENTRIST	74.57
	HOG + PHOG	**93.14**
	PHOG + CENTRIST	91.57
	HOG + CENTRIST	80.14
State-of-the-art method	Sirin et al (2017) [2]	92.70
	Shekar et al (2015) [3]	91.05
	Wang et al (2014) [4]	97.16
	Shu & Wu (2011) [5]	76.56
	Gopalan et al (2010) [6]	93.67

Table3: Performance comparison of the proposed method with state-of-the-art methods

Experimental Setup

- **Dataset**

- The proposed method was tested on MPEG-7 Part-B [1] silhouette dataset .
- The dataset has been split into 50 % for training and 50 % for testing.
- Edges are detected using sobel edge detector with one pixel thinning.
- 31×31, 61×61 sized window is selected from each white pixel.
- HOG, PHOG, CENTRIST descriptors are extracted from each of selected window.
- Codebook Construction: K-means is run K = 50, 100, 150, 200, 300, 500 and found the best K to be at 100.
- Linear OVA-SVM is used to determine the category of test image.

Conclusion

- Proposed method which integrates the use of the HOG, PHOG and CENTRIST in the bag of words
- Silhouette image classification improved by combining different feature of shape descriptors rather than individual feature.
- Future work aims to explore methods for efficient windows and feature selection and it will focus on experiment with other databases

References

1) JS. Jeannin and M. Bober, "Description of core experiments for MPEG-7 motion/shape", Seoul, 1999.
2) Y. Sirin, M.F.Demirci, "2D and 3D shape retrieval using skeleton filling rate", Multimed Tools Appl, Vol. 76,pp. 78237848, 2017.
3) B.H.Shekar, B.Pilar, J.Kittler, "An unification of inner distance shape context and local binary pattern for shape representation and classification" 2nd International Conference on Perception and Machine Intelligence, pp. 4655, 2015.
4) X. Wang, B. Feng, X. Bai, W. Liu, and L. Latecki, "Bag of contour fragments for robust shape classification", Pattern Recognition, vol. 47, no. 6, pp. 2116-2125, 2014.
5) X. Shu and X.-J. Wu, "A novel contor descriptor for 2D shape matching and its application to image retrieval", Image Vis. Comput., vol. 29, no. 4, pp. 286-294, 2011.
6) R.Gopalan, P.Turaga, R.Chellappa, "Articulation-invariant representation of non-planar shapes", European Conference on Computer Vision Springer, pp. 286299, 2010.

A Community Based Routing Algorithm for Mobile Opportunistic Networks

D. Fernando & Dr. K. Thabotharan
Department of Computer Science, Faculty of Science, University Of Jaffna

Introduction

- Opportunistic networking is a kind of delay-tolerant networking in which a number of wireless mobile nodes that communicate with each other, without the support a network infrastructure.
- Opportunistic networking uses locally available wireless technologies such as Bluetooth for pair-wise data forwarding hoping that the data will ultimately reach the destination.
- Intermittent connectivity and long delays in data delivery are inherent properties of this kind of opportunistic networking and they pose us challenges in data delivery.

What is the problem with the existing methods?

- The existence of communities among larger groups of people presents us a use case for opportunistic networking where content of interest could be exchanged among the members of communities.
- In such communities members are not exactly fixed to a single community and are usually connected to several communities based on their interests.
- Forwarding and routing content of interest among these communities should take care of the interests of the members of communities and other inherent properties of opportunistic networking.
- Therefore a more efficient routing algorithm that can overcome the inherent problems of such a set up is needed.

In this work we propose a community-based forwarding approach which we name as SWift routing algorithm (SW algorithm) that can be used to send messages among the community members in an opportunistic manner. Our simulation-based results show that our proposed algorithm outperforms three well-known algorithms in the field under varying network conditions.

Methodology

The proposed SW routing algorithm is a multi-copy algorithm and it ensures that when a node receives a message from its neighbor, the node forwards the message only to half the number of its neighbors compared to the previous node which has just forwarded the message to this node. By doing this it ensures that the message is not forwarded infinite number of times among the nodes in the network. At the same time when a message is received the algorithm also compares whether the current time is greater than the messages total accept time. This Total Accept Time (TAT) is defined as,

TAT = MAST + MCT + MAT

where,

- MAST: (Message Accept Start Time) is the time when a new message is accepted.
- MCT: (Message Checking Time) is the amount of time a node will reject an incoming message it has already received.
- MAT: (Message Accept Time) is the amount of time that is calculated by adding the message generate time with the message time to live and the message check time.

```
Algorithm 1: SW Routing Algorithm
receive a message from the neighbour ;
msg ← message containing the value for count;
c ← message checking time;
a ← message accept time;
if msg > 1 then
    if the node is interested in viewing the message then
        view the message;
        estimate the new value of count;
        send a copy of the message to neighbouring
          nodes;
    else
        estimate the new value of count;
        send a copy of the message to neighbouring
          nodes;
        end;
    end
else
    total_accept_time ←
      message_accept_start_time + c + a;
    if message_get_time > total_accept_time then
        remove it;
    else
        view the message;
        estimate the new value of count;
        send a copy of the message to neighbouring
          nodes;
    end
```

Our SW algorithm uses the following design principle:

1. When a node generates a message, it also determines the number of neighbors the message (which we name as the variable count) needs to be forwarded.
2. The node then passes the generated message to its neighbor along with the variable count.
3. If the receiving node is interested in this message it can view this message and before can forwarding the message to its neighboring.
4. If the node is not interested it can estimate the new value of the variable count by dividing current value of it by two and then can forward the message to its neighbor.
5. At the time of forwarding the node compares the value of the Total Accept Time (TAT) with the current time (message get time plus the duration taken by the message to be sent from the beginning up to this step) and if the current time is greater than the TAT then the message is removed from forwarding. Otherwise, it the message gets forwarded.

Experiments

For our simulation based experiments we use the ONE simulator to implement our proposed SW routing algorithm, Spray and wait, Direct Delivery and the Wave algorithm and we ran our experiments with selected simulation parameters listed in Table. During the experiments we also varied the such as message TTL from 300 to 500 minutes, message size from 400 KB to 1MB to 700 KB to 1MB, and the buffer size from 5MB to 7MB. We have collected the test results in trace files and have analyzed them for their performance.

Parameters	Values
Simulator	The ONE
Parameters Simulation Time	2532 sec
Message TTL	300 min
Buffer size	5 MB
Message size (Event size)	400kB - 1MB
Movement Model	Shortest path map based Movement
Community groups	6
Multi copies	10
MessageCheckingTime	100 min
MessageAcceptTime	200 min
Reports	Message Stats Report Delivered Messages Report

Performance Metrics

For the comparison of our proposed SW algorithm along with the three well known algorithms we use the following performance metrics:

Overhead Ratio: The overhead ratio reflects how many redundant packets are relayed to deliver one packet.
Overhead Ratio = (N - D)=D
(where N is the number of messages forwarded by a node, and D is the number of messages that are delivered to their destinations.)

Delivery probability: The delivery probability is the probability between the total number of messages delivered to their destinations and the total number of messages created at the source node.
Delivery ratio = P=T
(where P is a number of messages delivered to the destination and T is a number of messages made.)

Average latency: Average latency is the time between creating messages and receiving messages of a destination.
Average latency= (where n is the number of messages delivered to their destination, Ri is the moment when a message i reaches to its destination, and Si is the moment the message i is created.)

Discussion

Figures Fig. 1 to Fig. 3 show the results of comparison of message overhead with MessageTTL, Message size and Buffer size of Spray and Wait, Direct delivery, Wave and the SW protocols. When comparing the above four, our results show that the overhead of SW and Direct delivery are better than Wave, Spray and wait. The Direct delivery and the SW algorithm both get zero overhead ratio with MessageTTL, Message size and Buffer size. Compared to other algorithms, the SW algorithm shows the best performance.

Figures Fig. 4 to Fig. 6 show minimum latency with MessageTTL, Message Size and Buffer size in Spray and wait algorithm.

Simulation Results

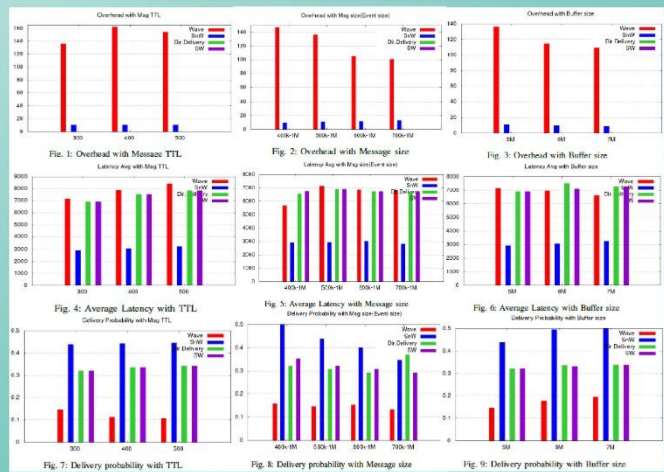


Fig. 1: Overhead with Message TTL
Fig. 2: Overhead with Message size
Fig. 3: Overhead with Buffer size
Fig. 4: Average Latency with TTL
Fig. 5: Average Latency with Message size
Fig. 6: Average Latency with Buffer size
Fig. 7: Delivery probability with TTL
Fig. 8: Delivery probability with Message size
Fig. 9: Delivery probability with Buffer size

SW algorithm performs closer to Direct delivery algorithm. These two perform better when compared to the Wave algorithm.

Figures Fig. 7 to Fig. 9 show a higher delivery probability in Spray and Wait, where as the SW algorithm shows a steady delivery probability in these cases. The SW and the Direct delivery algorithm perform better when compared to the Wave algorithm. In overall, the SW algorithm outperforms the other three algorithms for the overhead with MessageTTL, Message size and Buffer size.

Since SW sends messages based on the node's interest, it was able to achieve this. We were also able to observe that the SW algorithm always shows a better performance than the Wave algorithm in all the test cases.

Conclusion

- Our SW routing algorithm's test results show that the outperforms three existing algorithms when compared with overhead with message TTL, Message size and Buffer size.
- In some of the case the proposed routing algorithm exhibits a steady performance when compared with the three algorithms.
- SW routing algorithm can use to promote business ideas based on customer's interest. We can send promotion messages among community members.
- As a future work we would like to improve the proposed algorithm for larger communities of people.
- Therefore different kinds of communities and their interests vary greatly, and the algorithm accurately takes care of routing in a more efficient manner.

A Lossy Grayscale Image Compression Based on Delaunay Triangulation

T.M.V.D. Gunasekara and A. Ramanan
Department of Computer Science, Faculty of Science, University of Jaffna
Vikum.dheemantha@gmail.com

Introduction

As a result of technical advancements in digital imaging devices, the amount of images that any single person handles is increasing continuously. High quality images not only overuse the space but also becoming the main reason that causes the Internet traffic by transmitting the images through the Internet. Thus, there is a demand for a novel image compression techniques. Basically, image compression techniques can be categorised into two types: Lossy and lossless [1]. In lossless image compression, all information of an original image are retained in the compressed image. On the other hand, lossy image compression do not consider to retain all the information on the image, it mainly focuses on saving useful information while allowing the irrelevant information to be removed from the compressed image. Even for lossy image compression [2] there is a demand because there are plenty of applications of digital images that only consider about providing the visibly equivalent images to the users. Among several image compression techniques, block-based image compression [3] is enabled by the possibility of selecting large block from the image that can be represented as a single unit. Delaunay triangulation provides a better opportunity to create triangular mesh over an image [2, 4, 5]. Delaunay triangulation generates a mesh over a region to provide a set of non-overlapping triangular elements. The generation of non-overlapping elements can be utilised to provide a better image compression technique.

Methodology

Encode:

- Construct an initial lattice by performing Canny algorithm to detect the edges on an image and extract a set of points by performing run length algorithm on the edge map returned from Canny algorithm.
- Perform Delaunay triangulation on the initial lattice.
- Split triangles by adding new vertex to the barycentre of the triangles that are not homogeneous and perform Delaunay triangulation again on the new lattice. The splitting algorithm is continuously executed until convergence.
- Merge algorithm is performed and get rid of any unwanted vertices (vertex is considered as unwanted if all the triangular elements that share the vertex have near intensity value). Then perform Delaunay triangulation on the final lattice.
- Store the triangular elements with their mean intensity value as the result of the image compression.

Decode:

- Retrieve encoded data and re-construct triangulation mesh over an empty image.
- Fill each triangle with respective intensity value.

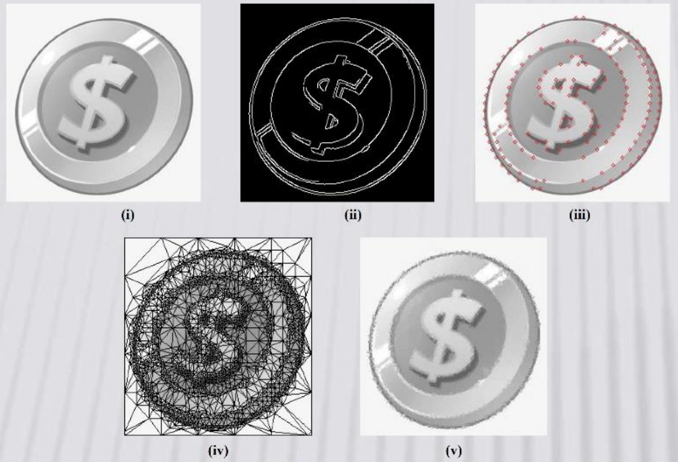


Figure 1: (i) Original image in grayscale, (ii) Canny edge map, (iii) Initial set of points extracted from the edge map in (ii) using run-length algorithm, (iv) Triangulation after performing split and merge algorithms (v) Decoded image after compression.

Results

Proposed method is tested with several grayscale images and some of the experimental results are given below.

Table 1. Experimental results of a compressed image using the proposed method with respect to the storage requirement and number of triangular elements in the generated mesh.

Image	Storage required (in Bytes)		#Triangular elements in mesh
	Original	Compressed	
Coin	65,536	25,809	3,687
French Flag	65,648	9,024	1,052
Artificial Image	326,312	225,408	25,340
Cartoon Girl	50,440	21,048	2623

- Triangles are stored using python data dump method. Hence, extra storage is consumed for store data structure details of python.

Discussion

Based on the experimental results, this study achieved better compression ratio with near quality image. But the quality and/or the compression ratio highly depends on the nature of the image. In our future work, we wish to reduce the storage requirement of 7 Bytes to store a single triangular element of a grayscale image so that a better compression ratio can be achieved.

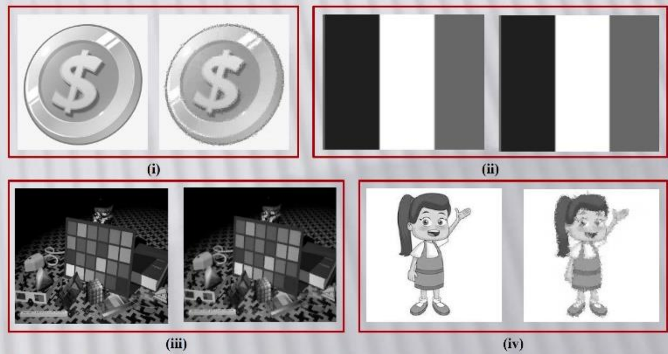


Figure 3: Visual comparisons of original images and decoded images, originals are on left and decoded images on right. (i) Image of a coin (digitally portrayed), (ii) Image of a flag, (iii) Image of some artificial component, (iv) Image of a cartoon girl.

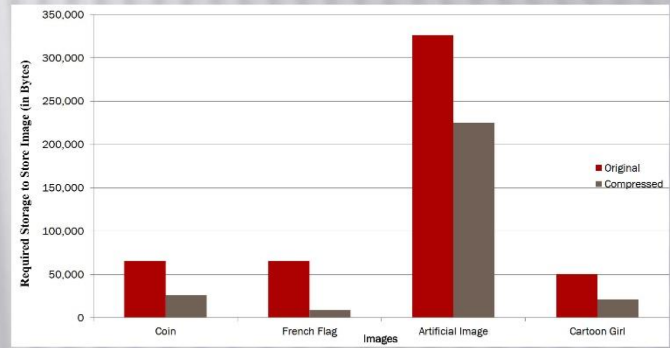


Figure 4: Comparison of the image sizes between original image and compressed image. (Size is calculated in bytes)

References

1. A.J. Hussain, A. Al-Fayadh and N. Radi, “Image compression techniques: A survey in lossless and lossy algorithms”, pp.44-69, 2018.
2. D. Marwood, P. Massimino, M. Covell and S. Baluja, “Representing Images in 200 Bytes: Compression via Triangulation”, 25th IEEE International Conference on Image Processing (ICIP), pp.405-409, 2018.
3. Qiang Du, M. Gunzburger, Lili Ju and X. Wang, “Centroidal Voronoi Tessellation Algorithms for Image Processing”, Journal of Mathematical Imaging and Vision - JMIV, 2005.
4. D. Laurent, N. Dyn and A. Iske, “Image compression by linear splines over adaptive triangulations”, Signal Processing, 86(7), pp.1604-1616, 2006.
5. F. Davoine, M. Antonini, J-M. Chassery, and M. Barlaud, “Fractal image compression based on Delaunay triangulation and vector quantization”, IEEE Transactions on Image Processing, vol. 5, pp.338-346, 1996.

TAMIL FONT TYPE IDENTIFICATION FROM TEXT IMAGES

P.Janani*[1], E.Y.A.Charles*[2]

*Department of Computer Science, Faculty of Science, University of Jaffna

[1]jenipathmanathan3@gmail.com, [2]charles.ey@univ.jfn.ac.lk

Student Workshop on Computer Science

Introduction

- Font type identification is a process of finding the font style of texts in images.
- Font type can be manually identified with experience to some extent.
- It is difficult to differentiate a font type from another due to the vast number of available fonts.
- This work proposes a method to automate the font type identification using machine learning.

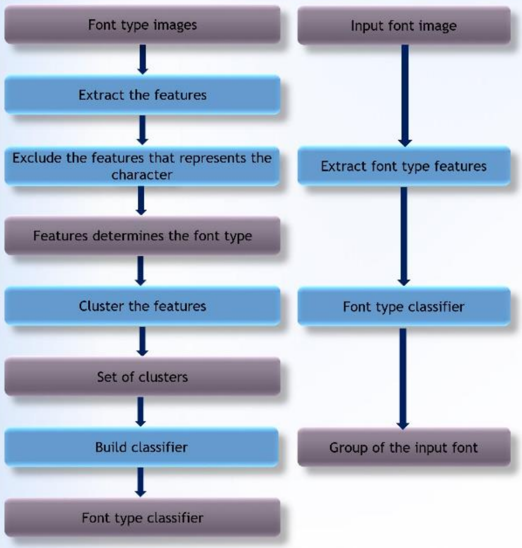
Motivation

- Graphic designers need to identify font types in images they encounter in their day to day life for later use.
- During the designing process of printed material, there is a need to identify similar fonts for a selected font.
- During OCR, the font style information are lost since the characters are identified using selected features. Font type information is needed to recreate a document from an image.

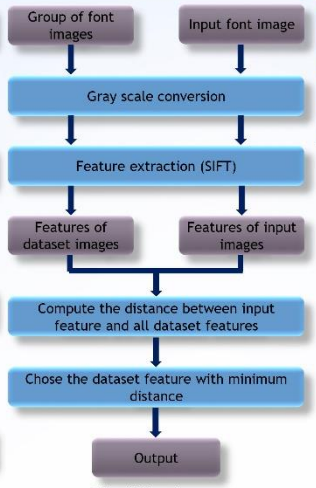
Challenges & Solution

- Font type identification can be done by finding the closest match of features of a given font among the available fonts.
- When the number of fonts increases, comparing a font with other fonts will be a time consuming and computationally intensive task.
- A better approach is to cluster fonts based on their style features and matching a given font image to a potential cluster.
- To cluster font types, style features and character type features should be distinguished from the available features.

Methodology



Identifying group of font type



Identifying font type

Dataset

To assist this research work, a Tamil font type dataset was created. This contains 1540 images covering 10 font styles.

Font Types	Number of Images
Anusha	148
Kamaas	155
Modern Tamil	156
Pravi	149
Rasihapriya	156
Sahaanaa	158
Siva-0002	152
Sivagami	152
Trinco Normal	157
Vairamani	157
Total	1540

Count of dataset images for each font styles

Examples of two character images
i)Kamaas
ii)Modern Tamil
iii)Rasihapriya
iv)TrincoNormal

Sample testing images

Experiments

- Experiments carried out on Tamil font dataset with MATLAB.
- Dataset contains 1540 images and 400 testing images covering 10 font styles.
- Each test images was compared to the dataset images and their K-Nearest neighbors were calculated.
- This test was conducted using various K values, K=1, K=5, K=10, K=15 K=20 and K=25.
- Among the testing images curved texts, rotated texts, texts with different font colors and background colors, texts with different font size and bold italic texts are included.

Confusion Matrix

Font Styles	Anusha	Kamaas	ModernTamil	Pravi	Rasihapriya	Sahaanaa	Siva-0002	Sivagami	TrincoNormal	Vairamani
Anusha	36	0	0	0	0	0	2	1	1	0
Kamaas	0	37	0	0	0	1	2	0	0	0
ModernTamil	0	3	31	1	0	1	1	3	0	0
Pravi	0	1	2	29	2	1	0	3	1	1
Rasihapriya	1	2	1	0	34	1	0	1	0	0
Sahaanaa	0	0	0	1	0	36	2	0	1	0
Siva-0002	0	1	1	1	0	2	33	1	1	0
Sivagami	1	0	2	1	0	1	2	31	2	0
TrincoNormal	0	1	3	0	0	0	1	0	34	1
Vairamani	0	0	2	0	0	0	0	2	0	36

Results

Results given in the following table shows the average recognition rate of each font face.

Fonts	K=1	K=5	K=10	K=15	K=20	K=25
Anusha	85.00%	77.50%	77.50%	77.50%	75.00%	75.00%
Kamaas	87.50%	87.50%	90.00%	90.00%	90.00%	90.00%
ModernTamil	75.00%	72.50%	77.50%	75.00%	77.50%	77.50%
Pravi	77.50%	82.50%	85.00%	85.00%	87.50%	87.50%
Rasihapriya	82.50%	82.50%	80.00%	82.50%	82.50%	87.50%
Sahaanaa	90.00%	87.50%	90.00%	87.50%	87.50%	87.50%
Siva-0002	87.50%	85.00%	87.50%	87.50%	87.50%	87.50%
Sivagami	80.00%	85.00%	87.50%	82.50%	82.50%	82.50%
TrincoNormal	87.50%	87.50%	87.50%	92.50%	92.50%	90.00%
Vairamani	82.50%	80.00%	80.00%	80.00%	80.00%	80.00%
Overall Accuracy	83.50%	82.75%	84.25%	84.00%	84.25%	84.50%

Related Works

- Deepfont : Identify Your Font from An Image
  - Authors : Zhangyang Wang, Zhangyang Wang, Jianchao Yang, Jonathan Brandt
  - Font identification from English Text images
  - Dataset : AdobeVFR (616 font styles with 4385 images)
  - Used CNN model
  - Accuracy : 80%
- Thai Font Type Recognition using SIFT
  - Authors : P. Jamjuntr and N. Dejdumrong
  - Font identification from Thai Document images
  - Dataset : 10 font styles and 10 text images in each font styles (100 images)
  - Features used: SIFT
  - Accuracy:97.37%

Conclusion & Future Work

- K-Nearest neighbor can identify the font styles of Tamil text images with an accuracy rate of 84.5%.
- This was done by comparing the input image features with features of each font type in the set of font types.
- Further research needed to be performed to distinguish the features that represent the style of a font type.
- Using the font style features we can recognize the font faces within a short period of time.

A PERFORMANCE EVALUATION STUDY OF SELECTED TCP PROTOCOLS ON WIRELESS AD-HOC NETWORKING ENVIRONMENTS

Kaluarachchi K.A.K.M. and Thabotharan K.
Department of Computer Science, University of Jaffna.
kasunmkaluarachchi@gmail.com & thabo@univ.jfn.ac.lk

SLSWCS Student Workshop on Computer Science

Introduction

- In the present time many people prefer to use the modern wireless mobile devices for their day today use.
- Wireless networking is very flexible in usage and it can support more than one device in one instance and it also covers larger geographical area.
- With the development of this networking process, people started to face for many problems with respect to their network connectivity with compared to the wired network connectivity.
- The major issue caused by this wireless connectivity is the connectivity speed. Because in wired connectivity most of the time has a constant speed where as in wireless connectivity the connection speed changes frequently.
- Therefore many researchers around the world are interested in finding the network issues with respect to the wireless connectivity in different scenarios. Previous research studies investigating the performance of the transport layer protocols have just looked at the TCP protocol without considering much of its properties.
- In this study we investigate the performance of the TCP and two of its well known properties TCP/Tahoe and TCP/Reno.
- With the two chosen protocols we carryout simulation based experiments to investigate the performance of TCP/Reno and TCP/Tahoe, under different testing conditions.
- We also consider the well known adhoc networking protocols such as AODV,DSR and DSDV for the simulation based experiments. Our simulation based experiments indicate that DSDV performs well rather than AODV and DSR.

Methodology

- TCP-Tahoe and TCP-Reno are used as the testing protocols for this work
- NS2 used for implementation and testing performance under varying conditions
- Ad-hoc network is considered
  - Nodes make intermediate connectivity among themselves
  - Assigned node works as source node for the implementation and also another assigned node be the destination node
- Linux Ubuntu 18.10 used as the operating system
  - core i3 CPU
  - 4GB RAM
- network performance metrics
  - Throughput: Defines the rate of something can be processed; it means in the network, the amount of effective message delivery over a communication channel, perhaps the delivery over a physical or logical link

$$Throughput = \frac{Number\ of\ received\ packets}{Last\ Packet\ sent\ Time\ -\ First\ packet\ sent\ Time}$$

  - Packet loss: For one reason or another, the packets are dropped from node. This causes unreliable delivery in the network. packet loss happens in the wireless network more than the wired network because of sharing media among nodes.

$$PacketLoss = \sum Send\ Packets - \sum Received\ Packets$$

  - Packet Delivery Ratio (PDR): It is referred to the number of packets effectively delivered to an endpoint as compared to the amount of packets that has been sent out by the sender

$$PDR = \frac{\sum Total\ number\ of\ Received\ Packets}{\sum Total\ number\ of\ Send\ Packets}$$

  - Delay: Delay is the time faced by a packet to move or travel across the network from one node to another

$$Delay = Tr - Ts$$

Where Tr - receiving time of that packet , Ts - sending time of a particular packet

$$Mean\ Delay = \frac{Total\ Delay}{N}$$

Where N - total number of packets received during simulation time

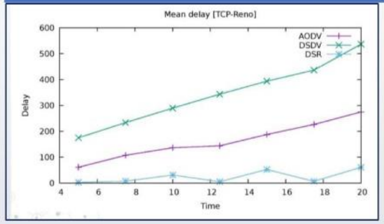
Experimental Setup

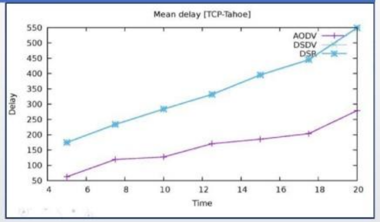
- Creating dataset for the performance evaluation
  1. Building mobile nodes movements source file by varying time
     - 50 mobility nodes , 1000m x 1000m of area size, min speed: 10.00ms-1 , max speed: 50.00ms-1 , avg speed: 23.72ms-1 , Uniform speed
     - Time Duration (Minutes) – 5, 7.50, 10, 12.50, 15, 17.50, 20
- Creating 6 Difference operation files as follows

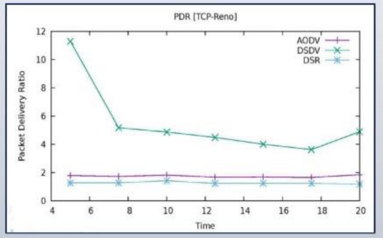
Packet Type	Routing Protocol
TCP-Reno	AODV
	DSDV
	DSR
TCP-Tahoe	AODV
	DSDV
	DSR

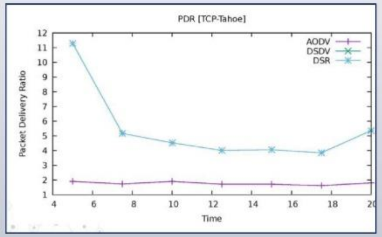
- Extracting required columns from trace files
- Use the extracted column data for calculations on behalf of the network parameters
- Repeat the process on basic of the difference time durations

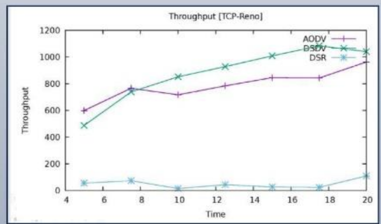
Results

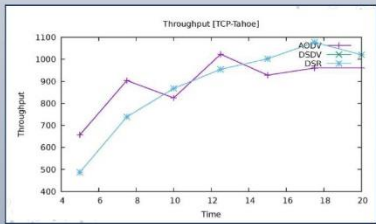


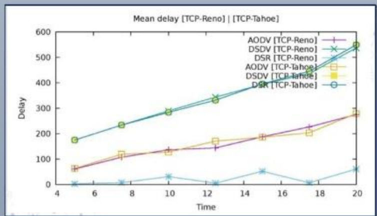


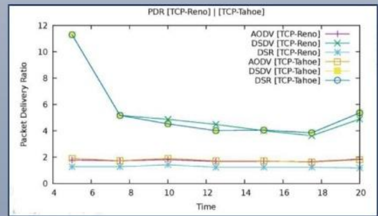


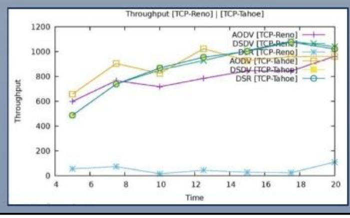












Packet Type	Routing Protocol	Drop Rate(%)	Delivery Ratio (%)
TCP-Reno	AODV	5.969 %	94.031 %
	DSDV	1.649%	98.351%
	DSR	3.156 %	96.844 %
TCP-Tahoe	AODV	6.177 %	93.823 %
	DSDV	1.762%	98.238%
	DSR	1.759 %	98.241 %

Discussion and Conclusion

- The goal of this research was to evaluate network performance of TCP under different traffic behavior and rank various techniques for network performance within that context.
- For this study two protocols of TCP were selected as TCP-Reno and TCP-Tahoe
- The 50 nodes are placed on the area size of 1000m x 1000m with minimum speed of node as 10 ms-1, maximum speed of node as 50 ms-1, average speed of node as 23.72 ms-1 and with uniform speed over the simulation time as 5min, 7.5min, 10min, 12.5min, 15min, 17.5min and 20min.
- Mobile nodes continuing to deliver data from source nodes to their respective destinations.
- PDR, Throughput and Mean delay were used to clarify the results (Network performance metrics).
- According to considered scenario TCP-Reno with DSDV routing protocol performance better than TCP-Tahoe (Routing Protocols - AODV, DSR and DSDV) and TCP-Reno (Routing Protocol - AODV, DSR)
- Consider on packet delivery ratio and throughput, DSDV performs well on them rather than AODV and DSR.
- Finally, conclude that DSDV with TCP-Reno performs better.

References

- T. Gopinath, A. S. R. Kumar, and R. Sharma, “Performance Evaluation of TCP and UDP over Wireless Ad-hoc Networks with Varying Traffic Loads,” 2013 International Conference on Communication Systems and Network Technologies, 2013.
- S. Narayan, “Improving Network Performance: An Evaluation of TCP/UDP on Networks,” 2013 International Conference on Communication Systems and Network Technologies, 2013.
- W. A. Kamil, S. A. Nor, and R. Alubady, “Performance Evaluation of TCP, UDP and DCCP Traffic Over 4G Network,” Research Journal of Applied Sciences, Engineering and Technology, vol. 11, no. 10, pp. 1048–1057, May 2015.
- A. Milanovic, S. Srbljic, and V. Sruk, “Performance of UDP and TCP communication on personal computers,” 2000 10th Mediterranean Electrotechnical Conference. Information Technology and Electrotechnology for the Mediterranean Countries. Proceedings. MeleCon 2000 (Cat. No.00CH37099).
- G. Xylomenos and G. Polyzos, “TCP and UDP performance over a wireless LAN,” IEEE INFOCOM 99. Conference on Computer Communications. Proceedings. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies. The Future is Now (Cat. No.99CH36320), 1999.
- K. Laeeq, N. Islam, and J. Bhayo, “Performance Analysis of TCP and UDP over Mobile Ad hoc Network,” Journal of Independent Studies and Research - Computing, vol. 14, no. 1, Jan. 2016.

DEEP LEARNING APPROACH TO DETECT PLAGIARISM IN SINHALA TEXT

T.C. Kasthuriarachchi , E.Y.A.Charles

Department of Computer Science, Faculty of Science, University of Jaffna
tharuka.ckasthuri@gmail.com,charles.ey@univ.jfn.ac.lk

ABSTRACT

Drawbacks and inefficiency of existing language independent plagiarism detection tools and lack of adequate research for Sinhala plagiarism detection this research work was motivated.

Proposed method focuses on developing a deep learning based approach for plagiarism detection in Sinhala documents

To improve the efficiency of the model Natural Language Processing techniques have been applied.

The proposed model was implemented and tested on an in house data set and found to be capable of detecting plagiarism with an accuracy of 97%.

The model is capable to detect direct and sophisticated copying such as replacing the words with their synonyms as well as changing the order of words in a sentence.

OBJECTIVE

The main objective of this research is to address the actual need of automated plagiarism detection tool for Sinhala language and developing a plagiarism detection corpus for future researches.

Corpus for Developing Sinhala Word2vec model

- Used the Sinhala News Corpus published the Language Technology Research Laboratory of University of Colombo School of Computing.
- The documents of the corpus were in utf-8 and utf-16 Unicode formats, so all the documents were converted to utf-8 format using a Linux script.
- Since the documents in the corpus contained lots of noises such as HTML syntax and punctuations, they were cleaned using NLTK python library.

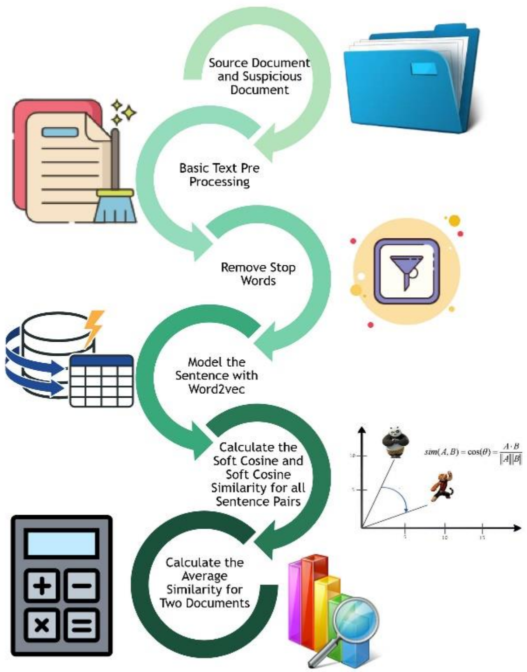
Plagiarism Detection Corpus

- From the UCSC Sinhala News Corpus, 50 documents were randomly selected and the sentences were plagiarised by a selected group of students.
- Plagiarising included changing word order and replacing words with similar words.
- In order to reduce the error probability, the plagiarized text has been rechecked by a Sinhala language experts. In addition the sentences were tagged as whether similar or not manually by the expert.

METHODOLOGY

Phase I

- word2vec model was constructed using a selected Sinhala text corpus.
- Before training the word2vec model, common bigram phrases from the text were extracted. Hence some common phrases like (Sri Lanka) are considered as one word.
- Basic text pre-processing steps such as tokenization, punctuation removal and stop word removal are applied to achieve maximum possible accuracy.
- After experimenting with CBOW and Skip Gram, CBOW method is selected.
- Training loss and perplexity were used as performance measure of the training.



Phase II

- Word vectors for the words in a source sentence (S1) and target sentence (S2) are computed using the word2vec model (see below illustration).

$$WV_1 = [w_{11}\ w_{12}\ \dots\ w_{1j}\ w_{1j+1}\ \dots w_{1d}]$$
$$WV_2 = [w_{21}\ w_{22}\ \dots\ w_{2j}\ w_{2j+1}\ \dots w_{2d}]$$
$$\vdots$$
$$WV_i = [w_{i1}\ w_{i2}\ \dots\ w_{ij}\ \ w_{ij+1}\ \dots w_{id}]$$
$$\vdots$$
$$WV_n = [w_{n1}\ w_{n2}\ \dots w_{nj}\ w_{nj+1}\ \dots w_{nd}]$$
$$SV_k = [sv_{k1}\ sv_{k2}\ \dots sv_{kj}\ sv_{kj+1}\ \dots sv_{kd}]$$

- These vectors are used to generate the word vectors for the sentences using (1).

$$sv_{kj} = \frac{\sum_{i=1}^{n} w_{ij}}{n} \quad (1)$$

- After modelling the sentence with word vectors, all sentences in a target document are compared with all sentences in the source document using the cosine similarity metric (2) and the soft-cosine (3) similarity metric, (generalization of cosine similarity).

$$cosine(S1,S2) = \frac{\sum_{i=1}^{k} S1_i S2_i}{\sqrt{\sum_{i=1}^{k} S1_i^2}\sqrt{\sum_{i=1}^{k} S2_i^2}} \quad (2)$$

$$softcosine(S1,S2) = \frac{\sum_{i=1}^{k}\sum_{j=1}^{k} S1_i M_{ij} S2_j}{\sqrt{\sum_{i=1}^{k}\sum_{j=1}^{k} M_{ij} S1_i S1_j}\sqrt{\sum_{i=1}^{k}\sum_{j=1}^{k} M_{ij} S2_i S2_j}} \quad (3)$$

- If the similarity measure of two sentences is higher than a set threshold value, then the target sentence is considered as plagiarized from the source sentence.
- Finally, the average similarity scores are calculated for the target document.

Sample Results Taken from Word2vec Model

Most Similar words for word දෙවියන්	
දෙවියන්ගේ	0.6743
විශ්ණු	0.6706
යෙහෝවා	0.6517
දෙවිවරුන්	0.6422
දෙවියන්ට	0.6364
දෙවි	0.6158
දෙවියන් වහන්සේ	0. 6025

RESULTS

Suspicious Sentence	Original Sentence	Soft Cosine	Cosine	Expert Judge ment
දෙදියගල දකුණු පළාතේ පිහිටි ස්භාවික සෞන්දර්යයෙන් අනූන සුන්දර ගම්මානයකි.	දකුණු පළාතේ මාතර දිස්ත්‍රික්කයේ අකුරැස්ස ප්‍රාදේශීය ලේකම් කොට්ඨාසය තුළ පිහිටි ස්වාභාවික සෞන්දර්යෙන් අනූන දෙදියගල සුන්දර ගම්මානයකි.	0.6757	0.7741	Similar
දෙදියගල ස්වභාවික සෞන්දර්යයෙන් අනූන දකුණු පළාතේ මාතර දිස්ත්‍රික්කයේ අකුරැස්ස ප්‍රාදේශීය ලේකම් කොට්ඨාසය තුල පිහිටි සුන්දර ගම්මානයකි.	දකුණු පළාතේ මාතර දිස්ත්‍රික්කයේ අකුරැස්ස ප්‍රාදේශීය ලේකම් කොට්ඨාසය තුළ පිහිටි ස්වාභාවික සෞන්දර්යෙන් අනූන දෙදියගල සුන්දර ගම්මානයකි.	0.8999	0.9790	Similar
පවුල් 404ක් පමණ ජීවත්වන මෙම ගම්මානය සිංහරාජ අඩවිය පාමුල උඩුගම කල්එළිය ගමට මායිම්ව පිහිටා තිබේ.	දකුණු පළාතේ මාතර දිස්ත්‍රික්කයේ අකුරැස්ස ප්‍රාදේශීය ලේකම් කොට්ඨාසය තුළ පිහිටි ස්වාභාවික සෞන්දර්යෙන් අනූන දෙදියගල සුන්දර ගම්මානයකි.	0.1896	0.5664	Not Similar

Threshold value	False Positive	False Negative	Accuracy
Using Cosine Similarity			
0.55	2	115	0.9539
0.69	38	41	0.9689
0.90	386	3	0.8466
0.99	507	0	0.8001
Using Soft Cosine Similarity			
0.30	10	71	0.9681
0.32	11	63	0.9708
0.39	44	45	0.9649
0.50	132	18	0.9409

CONCLUSIONS

Performance of the proposed model is based on a relatively small data set which was created by a group of students.

Annotating a sentence whether it is a plagiarised one or not in comparison with original sentences was performed by a single expert.

Work is still needed to construct a large and well evaluated Sinhala plagiarism corpus, such that the performance of the model can be well evaluated.

A Novel Approach for Tamil - English translation and vice versa using RNN

R.Kasthuriraajan and Dr.S.Mahesan
Department of Computer Science, Faculty of Science, University of Jaffna
kasthuriraajan94@gmail.com

Abstract

This study focused on improving a better approach for Tamil-to-English translation and vice versa using RNN. End of the study a novel approach for Tamil-to-English translation and another for English-to-Tamil translation were found to build a Neural Machine Translation system. Here optimizers and bridges had an impact on performance. BLEU scores were used to measure the performance of the system. Finally, the best performing model for Tamil-to-English translation was obtained with a BLEU score of 8.13. The best performing model for English-to-Tamil translation was obtained with a BLEU score of 4.66 which outperforms Google translator that has the score of 4.06. It shows that models with less number of layers can perform better than a high number of layers in terms of computing power while using appropriate optimizers and bridging technologies.

Introduction

Nowadays people unavoidably needs to use machines for translation purposes. In order to meet this need, the studies on machine translation systems emerge in recent years. Big companies like Google, Microsoft also take much effort into building efficient machine translation systems. There are several machine translation techniques such as rule-based translation techniques, and statistical machine translation techniques. Even though Neural Machine Translation (NMT) is getting attention because of its accuracy and behaviour like human translation. So it is an active research topic all over the world. Google Neural Machine Translation (GNMT) system is a well known NMT system that was introduced in 2016 and used in Google translator by Google. More than 100 languages are supported by Google translator including Tamil and English. However, there is a need for many improvements in its performance. So this topic was

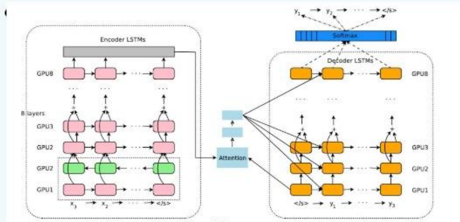


Figure01: GNMT architecture

Objective

The objective of this research project is to build a neural machine translation system for Tamil to English and vice versa using recurrent neural network (RNN).

Methodology

There are three major components in this study.

1)Pre-processing the dataset
2)Training Neural Machine Translation models with training dataset and validation dataset.
3)Testing the trained models with the testing dataset and obtain results.

- ❖ Recurrent Neural Network (RNN) was selected to build neural machine translation models in this research project.
- ❖ A publicly available Tamil-to-English parallel corpus from various domains (EnTam V2) which was compiled by Loganathan Ramasamy was used for this study.
- ❖ Byte Pair Encoding (BPE) was selected to learn encoding and applied to all datasets except the target test data. Vocabulary was created from source and target datasets. All training and validation datasets were changed into torch tensors.
- ❖ In neural machine translation systems, the encoder-decoder mechanism is used to translate language pairs. First source language is encoded by RNN encoders and then RNN decoders decode them into target language.
- ❖ Long Short Term Memory (LSTM) was used in this research experiment to overcome the Long term Dependency Problem. Two layers of bidirectional LSTMs (Bi-LSTM) were selected as encoder with 500 hidden layers and two layers of LSTMs were selected as decoders.
- ❖ Two optimization methods were experimented here. One is ***adam*** with learning rate 0.001 and the other one is ***sgd*** with learning rate 1.0. A bridge is an additional layer between an encoder and decoder that defines how information is passed from encoder to decoder. Here two models were trained with bridge and two models without bridge.
- ❖ Finally, the translated sentences were compared with target test data. Using the BLEU scoring system, the accuracy of each model was measured and compared with each other.

Figure02 :Basic Architecture of the System

Experimental Setup

In this research project, OpenNMT-py was used to do experiments. This is a research-friendly Pytorch port of OpenNMT which is an open source (MIT license) ecosystem for neural machine translation and neural sequence learning.

The following models were created with the help of OpenNMT-py. Word vector size for source and target was defined as 500. Every model was trained with 2 layers Bidirectional RNN, 2 layers of RNN decoders, 500 hidden layers, 100,000 training steps and ***mlp*** attention type with the specific attributes mentioned below. After every 5000 steps, the trained models were saved.

Table 01:Models

	Model 1 (EnTa)	Model 2 (EnTa)	Model 1 (TaEn)	Model 2 (TaEn)
RNN type	LSTM	LSTM	LSTM	LSTM
Optimizer	adam	sgd	adam	sgd
Bridge	False	False	True	True

Table 02:Data set

Data Set	
Training Data	166,871 Sentences
Testing Data	2,000 Sentences
Validation Data	1,000 Sentences

Results

The test data were translated using the trained models mentioned above with the help of OpenNMT-py. The translated corpus was then evaluated using BLEU scoring method. BLEU is measured out of 100 and the better performance will have higher score. According to the BLEU score, the best performing trained models among which were saved after every 5000 steps, were selected in each model. The test data were also translated by Google Translator and evaluated with the BLEU scoring method.

Table 03: English-to-Tamil translation results

Model	BLEU Score
LSTM+mlp+adam(EnTa)	4.59
LSTM+mlp+sgd(EnTa)	4.66
GNMT(Google Translator)	4.06

Table 04:Tamil-to-English translation results

Model	BLEU Score
LSTM+mlp+adam+bridge (TaEn)	8.13
LSTM+mlp+sgd+bridge (TaEn)	7.81
GNMT(GoogleTranslator)	21.16

The OpenNMT-py frame work is available at https://github.com/OpenNMT/OpenNMT-py

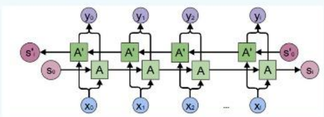


Figure03 : A bidirectional RNN

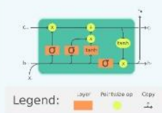


Figure04 :An LSTM Cell

Discussion & Conclusion

Some Neural Machine Translation systems use more layers in their model and a big parallel corpus for training (E.g. GNMT uses 8 layers). But here only two layers were used with a limited number of parallel corpora and a better performance was gained.

Our best performing English-to-Tamil translation model gained a BLEU score of 4.66 and Tamil-to-English gained a BLEU score of 8.13.

We could thus conclude that an NMT system can be implemented using this technique with low resources

References

1. Yonghui Wu, Mike Schuster, Zhifeng Chen, Quoc V Le, Mohammad Norouzi, Wolfgang Macherey, Maxim Krikun, Yuan Cao, Qin Gao, Klaus Macherey, et al. Google's neural machine translation system: Bridging the gap between human and machine translation. arXiv preprint arXiv:1609.08144, 2016.
2. Dzmitry Bahdanau, Kyunghyun Cho, Yoshua Bengio,Neural Machine Translation by Jointly Learning to Align and Translate. arXiv:1409.0473v7 [cs.CL], last revised 19 May 2016.
3. Kyunghyun Cho, Bart van Merrienboer, Dzmitry Bahdanau, Yoshua Bengio.On the Properties of Neural Machine Translation: Encoder-Decoder Approaches.arXiv:1409.1259 [cs.CL], 7 Oct 2014.

PREDICTING THE OUTCOME OF THE CRICKET MATCHES USING MACHINE LEARNING TECHNIQUES

Kausik M. , Siyamalan M.
Department of Computer Science, Faculty of Science , University of Jaffna
kausik219@gmail.com, siyam@univ.jfn.ac.lk

Student Workshop on Computer Science

Abstract

This poster proposes two novel approaches for predicting the outcome of cricket matches by modeling the team performance based on the performances of its players in other matches. Our first approach is based on feature encoding, which assumes that there are different categories of players exist and models each team as a composition of player–category relationships. The second approach is based on a shallow Convolutional Neural Network (CNN) architecture, which contains only four layers to learn an end-to-end mapping between the performance of the players and the outcome of matches. Both of our approaches give considerable improvement over the baseline approaches we consider, and our shallow CNN architecture performs better than our proposed feature encoding based approach. We show that the outcome of a match can be predicted with over 70% of accuracy.

Methodology

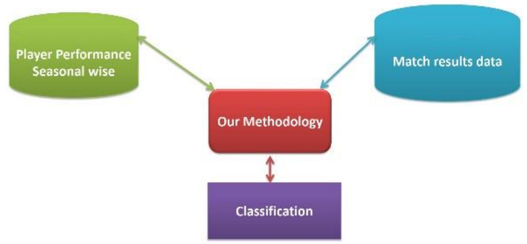
Let $X_{i,t}^{s} \in \mathbb{R}^{d}$ represents the performance of a player P_i from the team T_t at the season s, where $X_{i,t}^{s}$ can be represented by the player P_i's batting, bowling and fielding capability at the season s.

Team T_t^{s} can be represented as { $X_{1,t}^{s-1}, X_{2,t}^{s-1}$, ... , $X_{11,t}^{s-1}$ }; Here the important thing is that the season should be *s-1*, because the match season is *s* therefore we need to get the player performance from the previous season to represent the team composition. Then only we can get the recent performance of the player. The outcome is represented as either 0 or 1; if the outcome is 1 then the Team 1 is won against the Team 2(Opponent); if the outcome is 0 then the Team 2 (Opponent) is won against Team 1. We eliminated the matches if the outcome is draw or no result.

Mainly, The following are the two different approaches carried out:

- Feature encoding approach
- CNN approach

Basic idea of the methodology

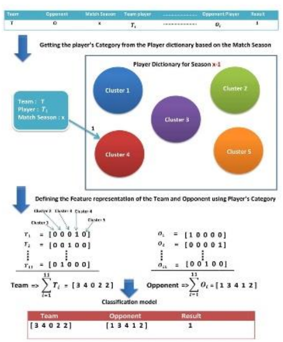


Feature Encoding Approach

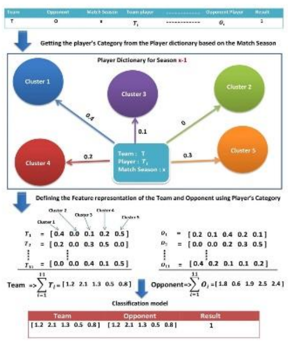
We used two types clustering techniques define the player categories

- Hard Assignment model
- Soft Assignment model

We used linear SVM for the classification process with this trained model to get the output.



Hard Assignment model



Soft Assignment model

Convolution Neural Network approach

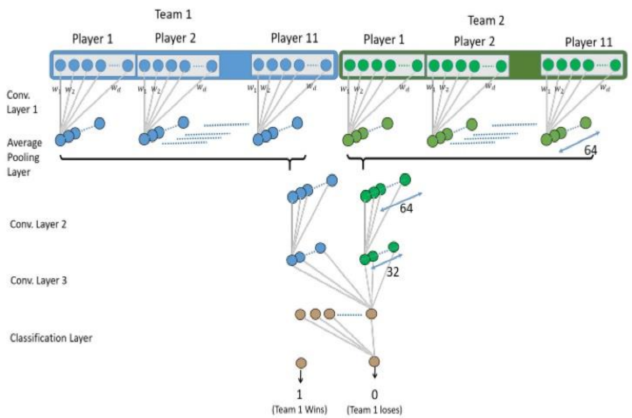


Fig. 1. The proposed CNN architecture. Some connections of the CNN in this figure are not shown to avoid clutter.

TABLE I
PROPOSED CNN ARCHITECTURE. THERE WAS NO PADDING USED IN ALL THE LAYERS.

	Conv. Layer 1	Avg. Pooling	Conv. Layer 2	Conv. Layer 3	Linear classification layer
Input dimension	$d \times 1 \times 1$	$22 \times 1 \times 64$	$2 \times 1 \times 64$	$2 \times 1 \times 32$	$16 \times 1 \times 1$
Output dimension	$22 \times 1 \times 64$	$1 \times 1 \times 64$	$2 \times 1 \times 32$	$1 \times 1 \times 16$	2
Filter size	$d \times 1 \times 1$	$22 \times 1 \times 1$	$1 \times 1 \times 64$	$2 \times 1 \times 32$	$16 \times 1 \times 1$
Stride	d	1	1	1	–
No. of filters	64	–	32	16	–

Dataset

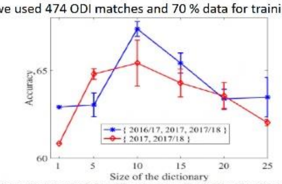
We crawled the website www.espncricinfo.com using our Python wrapper codes to collect the dataset. We were able to collect players information and the details of ODI matches from May 2013 to October 2017. Our dataset contains 2, 581 players and 474 ODI matches in total.

Season	2013	2013/14	2014	2014/15	2015	2015/16	2016	2016/17	2017	2017/18
No. of matches	47	59	35	91	33	90	37	60	51	12

Testing results

We use *Accuracy* as the evaluation measure. For the proposed Feature encoding-based approach, we iterate each experiment 10 times and report the average and standard deviation of accuracy values over these iterations.

In these approaches, we used 474 ODI matches and 70 % data for training and 30 % data for testing.



Effect of dictionary size when tested on different seasons. Vertical bars shows the standard errors.

Testing set (seasons)	N_{tr}	N_{te}	Baseline 1	Baseline 2	FE	CNN	CNN ensemble
{2016/17, 2017, 2017/18/}	634	314	56.68	62.42	67.36 ± 1.02	68.11 ± 1.07	**69.58 ± 3.33**
{2017, 2017/18}	708	240	62.02	62.50	65.42 ± 3.21	71.02 ± 3.33	**75.00**

Comparison of different approaches for cricket outcome prediction: baseline 1 and baseline 2 represent the concatenated feature representation and the averaged representation approaches respectively. FE represents the proposed feature encoding approach with a fixed dictionary size of 10. N_{tr} AND N_{te} represent the number of training and testing matches (augmented) respectively.

Conclusion and Discussion

In this poster, we presented two novel approaches for predicting the outcome of Cricket matches. Both of our approaches perform considerably better than the baseline approaches we considered. To the best of our knowledge, we are the first group to apply CNN for Cricket outcome prediction. We showed that the outcome of a Cricket match can be predicted with an accuracy of over 70%. In this work, we modeled each team based on the performance of the players play for that team. We considered features such as number of matches played in each season, number of runs scored, number wickets captured, and so on. However, we haven't considered some of the crucial factors such as the ranking of the players, player's strike rate, etc. These factors along with a larger dataset will be considered in our future work.

References

- Kaluarachchi, Amal and Varde Aparna .S, "CricAI: A classification based tool to predict the outcome in ODI cricket", *2010 Fifth International Conference on Information and Automation for Sustainability*, IEEE 2010.
- Shah, P. , "Predicting Outcome of Live Cricket Match Using Duckworth- Lewis Par Score", *International Journal of Systems Science and Applied Mathematics*, 2017.
- Madan Gopal Jhanwar and Vikram Pudi, "Predicting the outcome of ODI cricket matches: A Team Composition based Approach", *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD 2016)*, 2016.
- Sankaranarayanan, Vignesh Veppur, Junaed Sattar, and Laks VS Lakshmanan, "Auto-play: A Data Mining Approach to ODI Cricket Simulation and Prediction", SDM. 2014.

Flower Classification Using Multiple Feature Set

Kishotha. S, *and* **B. Mayurathan**
Department of Computer Science, University of Jaffna
kisho1504@gmail.com, and barathym.univ.jfn.ac.lk

Abstract

Flower image classification is still a challenging task because of the wide range of flower species, which have similar shape, appearance or surrounding things such as leaves, and grass.
The goal of this poster is to analyze the effect of multiple local features for flower image classification. Different local features are extracted from the flower images, each describing different aspects such as shape, texture and color. The performance of proposed method is compared with state-of-the-art method and analyzed the performance of the feature descriptors in flower image classification. By evaluating these descriptors it can be concluded that the combined SURF+ CTM gives better performance than other combination of features in the context of flower image classification.

Introduction

- Flower classification is a challenging task due to the large variety of flower classes that share similar features: several flowers from different types share similar color, shape and appearance. Furthermore, images of different flowers usually contain similar surrounding objects such as leaves, grass, etc.
- Hence, many flower classification techniques depend on extracting their features from a segmented flower region to improve accuracy [1], [2].
- Figure 1 illustrates an example of the difficulties of recognizing flower categories. These problems lead to a confusion across classes and make the task of flower classification more challenging.

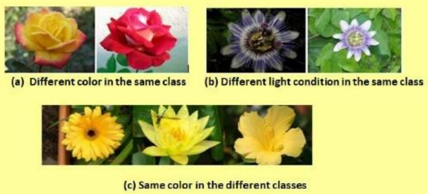


Figure 1: Here (a) and (b) are different color and different light condition in same class, (c) is same color in the different classes.

- An efficient flower classification system is an important task in various applications such as plants monitoring systems, content-based image retrieval for flower representation and indexing [3], floriculture industry, live plant identification and educational resources on flower taxonomy [4].
- Thus, novel convenient method would be of great benefit for flower classification.

Methodology

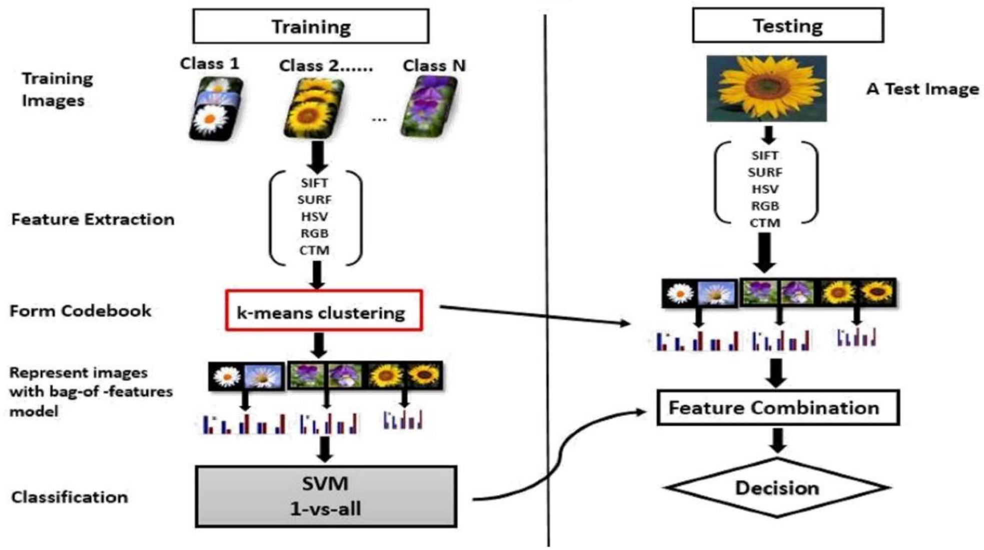


Figure 2: Proposed methodology for flower classification

- The details of the proposed methodology for flower classification is presented in Figure 2. During the whole process, multiple feature descriptors such as SIFT, SURF, RGB, HSV and CTM are used to represent flower images.
- In this experiment, 17 Flower Category Database is used. In 17 Flower Category Database, its consisting of 17 flower categories, where each category is represented by 80 different images.
- For each of these descriptors, K-means clustering algorithm is used on the entire feature database to obtain set of clusters.
- In K-means clustering algorithm, user needs to specify the number of clusters in its initial stage and there is no guarantee that the obtained clusters are visually compact. Due to that reason, K-means is run K = 500, 1000 and 1500 and found the best K to be at 1000. Finally, classifier is constructed based on the histogram of each image class.
- In this experimental setup, in order to identify the appropriate feature descriptor for flower classification, the performances of different combinations of feature descriptors that are used in this experiment are compared.
- In [5], different combinations of feature descriptors are considered to calculate the performance of flower classification and 17 Flower Category Database is also used here. So, we follow the same experiments in [5] in order to compare our proposed method with the performance done in [5].

Testing Results

- Table I shows the performance of the different combinations of feature descriptors that are used in this experiment and the method proposed in [5].
- According to the performance shown in Table I, **it can be seen that SURF + CTM and SIFT + CTM give better performance than other combinations of features**.
- Based on the recognition rate given in Table I, **proposed method gives better performance than [5].**

Features	Recognition rate [5]	Recognition rate (ours)
SIFT internal	55.1	-
SIFT boundary	32.0	-
SIFT	-	68.71
HSV	43.0	47.06
RGB	-	37.88
CTM	-	52.81
SIFT int + HSV	66.4	-
SIFT bdy + HSV	57.0	-
SIFT+HSV	-	68.71
SIFT+RGB	-	70.02
HSV+RGB	-	47.53
SIFT+CTM	-	**73.88**
SURF	-	69.18
SURF+SIFT	-	69.18
SURF+HSV	-	49.88
SURF+RGB	-	68.24
SURF+CTM	-	**74.59**

Table 1 : Proposed methodology for flower classification

Conclusion

- Flower classification method is proposed based on multiple feature descriptors.
- In this work, performance of SIFT, SURF, HSV, RGB and CTM features are analyzed in flower classification.
- According to the experimental results, we observe that multiple features empower the classifier to train a better model and achieve a better classification accurate on test sets.
- In addition, the experimental results have shown that the combined (SURF + CTM) features outperform the individual features.
- The important thing to be noted in this work is that only color features with the combination of SIFT and SURF have given a good classification accuracy when compared to other results in this experiment.

References

[1] Nilsback, M-E. and Zisserman, A., Automated Flower Classification over a Large Number of Classes, Proceedings of the Indian Conference on Computer Vision, Graphics and Image Processing, 2008.
[2] Chai, Y., Lempitsky, V., Zisserman, A., Bicos: A bi-level cosegmentation method for image classification, International Conference on Computer Vision, 2011.
[3] Das, M., Manmatha, R., and Riseman, E., Indexing flower patent images using domain knowledge, In IEEE Intelligent Systems and their Applications, vol. 15, no. 5, pp. 24 33, 1999.
[4] Chi, Z., Data management for live plant identification, Engineering online library, Springer, 2003.
[5] Nilsback, M., F. and Zisserman, A., Automated flower classification over a large number of classes, Proceedings of the Indian Conference on Computer Vision, Graphics and Image Processing, 2008.

Detection of Red Ripe Tomatoes on Plants using Image Processing Techniques

G.A.D. Kshithija Tharaka and B. Mayurathan
dushantharaka2@gmail.com, barathym@univ.jfn.ac.lk
Department of Computer Science
University of Jaffna

Abstract

- Tomatoes are well-known grown fruit in farming and cultivation of tomatoes has greatly increased in past few years. It is necessary to develop an algorithm to differentiate red ripe tomatoes because of tomato fruit does not ripe simultaneously.
- This poster presents a method based on image processing techniques to recognize ripen tomato from others. According to the experimental results, proposed methodology shows encouraging accuracy (87%) for detecting the red ripen tomato.

Introduction

- Tomatoes are world wide popular fruit and cultivate mainly in green houses. Nowadays, high labour cost is needed to increase the size of the greenhouse and it's extremely time consuming process.
- Based on the Figure 1, colour is a most important and effective feature to identify the tomato quality.

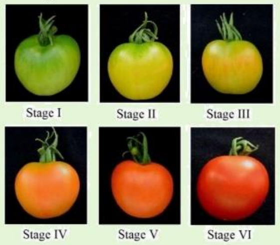


Figure 1: Different stages of tomato ripening

- Several features such as RGB model [1] and colour based algorithm [2] are used to recognize the fruits in the state-of-the-art methods.
- The main goal of this poster is to develop an algorithm to differentiate the red ripen tomatoes from the others.

Methodology

- The diagrammatic representation of the proposed methodology is given in Figure 2. It describes the steps that were involved in this research.

Original image
↓
Background removal using HSV colour model
↓
Identify the noisy tomatoes and remove it
↓
Remove remaining unwanted things using blue and green colour component
↓
Touching tomatoes separation and identify the red ripen tomatoes

Figure 2: Main stages of the proposed methodology

- Ripen tomato can be determined by its colour and the colour of a ripen tomato region is defined as at least 95 percentage red colour value of the whole image.
- First of all, the proposed algorithm blurred the original image using Gaussian blur. So, neighbouring pixels become a little more uniform in colour, so it will ease brighter and darker spots on the image and keep holes out of the mask. Then, HSV colour model is applied to segment similar colour objects using H components of the HSV colour space.
- Then, common morphology operators are applied with a structuring element in order to remove the noises. In order to find out the red region parts only in the image, bitwise operation is carried out between the resultant image and original image.
- Finally, blue and green components are subtracted from the resultant image in order to remove background noises.
- It is a very challenging task to separate touching tomatoes and recognize red tomatoes from non-red ones. Next step of this proposed methodology is separating touching tomatoes and then identify red tomatoes from the non-red ones.
- Watershed algorithm is used to segment and Euclidean Distance Transform is computed. Here, closest foreground pixels are computed and distance map is generated. After identifying the contours, largest contour represent boundary of a given object in the image. Finally, put a circle surrounding the object for a given the contour of the object. To extract the individual objects, bitwise operation is applied.
- Also, Figure 3 describes the main steps of this proposed methodology.

Figure3: Diagrammatic representation of proposed methodology

Dataset

- In this experiment, 100 images are collected from different sources and each images are captured by camera with 16 mega pixel on natural light condition. Single tomatoes, ripe tomatoes, non-ripe tomatoes and touching tomatoes are included in this collected dataset.
- For each image, total number of ripen tomatoes, non-ripen tomatoes and touching tomatoes are recorded in this dataset.
- Figure 4 shows some sample images from this dataset.

Figure 4: Some sample images from the collected dataset

Experimental Design and Testing Results

- Proposed methodology is applied to detect the ripen tomato and then circles are placed on the region of interest. Finally, count the total number of circles in the image and checked with dataset information.
- Based on our testing results, 87 images are correctly classified out of 100 images.
- The Figure5 shows some sample output images from this experimental design.
- **According to the experimental results, proposed methodology shows encouraging accuracy (87%) for red ripen tomato detection.**

Figure 5: Some sample detection outputs for red ripen tomatoes

Conclusion

- In this poster, a method for detecting red ripen tomato is described.
- Our future work is mainly focused on to develop an algorithm to check whether the picked tomato is good or damage.

References

1) Wang, Y., Zhu, X., and Ji, C., "Machine vision based cotton recognition for cotton harvesting robot", Computer and Computing Technologies in Agriculture, pp.1421–1425, 2008.
2) Bulanon, D.M., Kataoka, T., Ota, Y. and Hiroma, T., "A Segmentation Algorithm for the Automatic Recognition of Fuji Apples at Harvest", Biosystem Engineering, vol. 83(4), pp. 405–412, 2002.
3) Khoshroo, A., Arefi, A. and Khodaei, J., "Detection of Red Tomato on Plants using Image Processing Techniques", Agricultural Communications, pp. 9–15, 2014.
4) Monavar, H.M., Alimardani, R. and Omid, M., "Detection of red ripe tomatoes on stem using Image Processing Techniques", Journal of American Science, pp. 376–379, 2011.

Automatic Facial Makeup Detection

Y. Ligitha and A. Ramanan
Department of Computer Science, Faculty of Science, University of Jaffna
ligithay@gmail.com, a.ramanan@univ.jfn.ac.lk

1. Introduction

Recognising face images is a main research area based on many practical applications where human identification is needed.

Makeup can fall under two categories: Light makeup (the makeup cannot easily perceived since the applied colours correspond to natural skin colour), and heavy makeup (the makeup is easily perceptible).

Experiments are conducted on three challenging and unconstrained datasets : YouTube Makeup database (YMU), Makeup In the Wild database (MIW), and Virtual Makeup database (VMU)[1].

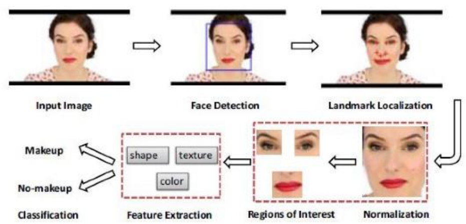
2. Objective

To detect makeup by selecting the best features that lead to the best classification result on images of human faces by image processing and pattern recognition techniques.

3. Methodology

- The Adaboost face detector in OpenCV is used to automatically detect the face[2].
- Given a face image, the proposed method first estimates the feature landmarks within the facial region and Haar-like filters are used for locating and characterising the appearance of each landmark.
- This is followed by cropping region of interest (ROI) by using Viola Jones algorithm (face, the regions around the left eye, the right eye, and the mouth).
- Then a set of shape, colour and texture features are extracted from the face and ROIs by using Hue Saturation Value (HSV) colour space, tessellation, Watershed transform, Canny edge detector and Local Binary Pattern (LBP) histogram[3] [4].
- Feature set is then fed to Support Vector Machine (SVM) classifier to detect the presence or absence of makeup in the input face image.

Methodology…



4. Testing Result

Table 1 : Test results as accuracy for colour descriptor

Technique	YMU	MIW	VMU
HSV	56.04%	51.30%	58.82%
Tessellate	49.47%	51.28%	60.72%
Watershed Transform	56.59%	56.49%	96.08%

Table 2 : Test results as accuracy for shape descriptor

Technique	YMU	MIW	VMU
Canny edge detector	52.20%	53.90%	90.20%

Table 3 : Test results as accuracy for texture descriptor

Technique	YMU	MIW	VMU
LBP histogram	68.13%	69.48%	62.75%

Table 4 : Overall accuracy obtained by using colour, shape, texture descriptors

DATASET	OVERALL ACCURACY
YMU	74.06%
MIW	75.49%
VMU	90.76%

5. Experimental setup

- The proposed method was tested on YMU, MIW, and VMU datasets.
- The proportion of YouTube makeup dataset (YMU) division for the training and testing was 70% and 30%, respectively, Makeup in the wild (MIW) dataset and Virtual makeup dataset (VMU) used for testing only[1].
- YMU dataset contains 604 images.
- For this research from YMU dataset 422 images were used for training and 182 images were used for testing.
- MIW dataset contains 154 images, VMU contains 102 images, and both images were used for testing.
- Classifier: Linear OVA-SVMs

6. Discussion

- If we consider about overall accuracy, higher value was obtained for VMU because YMU and MIW dataset have images with light makeup whereas VMU has heavy makeup images.
- The features obtained through each subsystem were finally given to a classifier in order to categorize them.
- In this study, 860 images (YMU, MIW, VMU) were used in the SVM classifier.
- The overall accuracy obtained is 74.06% for YMU, 75.49% for MIW, and 90.76% for VMU.

7. Conclusion

- Overall accuracy was obtained by combining all methods (HSV, Tessellate, watershed transform, Canny edge detector, LBP histogram) and then fed to Support Vector Machine (SVM) classifier.
- The proposed method can be further improved by considering Eye shadow detection, Lipstick detection, and Liquid foundation detection.

8. References

1. Dantcheva, C. Chen, and A. Ross. Can facial cosmetics affect the matching accuracy of face recognition systems? In BTAS, 2012.
2. C. Chen, A. Dantcheva, A. Ross, "Automatic Facial Makeup Detection with Application in Face Recognition," in Proceeding of 6th IAPR International Conference on Biometrics (ICB), (Madrid, Spain), pages 1-12, June 2013.
3. T. Ahonen, A. Hadid, and M. Pietik¨ainen. Face description with local binary patterns: Application to face recognition. *IEEE Trans.on PAMI*, 28(12):2037–2041, 2006.
4. S. Varshovi. Facial makeup detection using HSV color space and texture analysis. Master's thesis, Concordia University, Canada, 2012.

A Hybrid Data Forwarding Approach For Opportunistic Networks

H. P. Malki Maduka and **K. Thabotharan**
Department of Computer Science, Faculty of Science, University of Jaffna
malki.maduka@gmail.com, thabo@univ.jfn.ac.lk

ABSTRACT

Opportunistic networks is one of the recent paradigm of mobile Ad hoc networks. It is a network of wirelessly connected nodes. Nodes are connected only for a shorter period of time. In this research we developed a hybrid-data forwarding approach Spray with Probability Routing Protocol(SPROP) routing which utilizes the Spray and Wait routing forwarding strategy that is "sprays" a number of copies into the network, and then "waits" till one of these nodes meets the destination, and exploits an important social features of it and apply the PRoPHET routing forwarding strategy, to the spraying phase.

According to the Simulation results, SPROP increases the delivery ratio and decreases the average latency, compared to PRoPHET and Spray and Wait routing forwarding strategies.

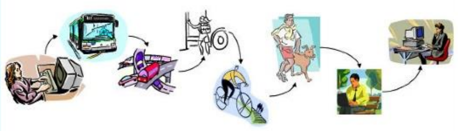


Figure 1: Example of communication in Opportunistic Network

OBJECTIVES

The random way-point mobility model is popular to use in evaluations of mobile ad hoc protocols, real users are not likely to move around randomly, but rather move in a predictable fashion based on repeating behavioural patterns such that if a node has visited a location several times before, it is likely that it will visit that location again. So here we used that concept to achieve following objectives,

- The main objective of this research is to propose a hybrid forwarding algorithm that can maximize the message delivery ratio, and reduced packet duplication and the average latency in to the network.
- And the other objective is understanding the background of PRoPHET routing and Spray and Wait routing to extract related information to the new protocol.

METHODOLOGY

Experimental Setup

- Simulator: The ONE
- Wireless technology : Bluetooth
- Transmission range : 10m
- Transmission speed : 250 kbps
- Map Size
  - ✓ Width : 4500m
  - ✓ Height : 3400m
- Simulation time : 5000 seconds
- Simulation configuration consists of
  - ✓ Varying message size
  - ✓ Varying time-to-live time
  - ✓ Varying buffer size

SPROP Routing

- The source node **a** and the encountered node **b**, and their delivery probabilities $P_{a,b} \epsilon$ [1, 0].
- Node **a** wants to send the message to the destination node **d**.
- If node **a** encounters node **b**, node **a** checks that encountered node b is destination node or not.
- If it is destination node, then node **a** hands over the messages to the node **b**.
- If it is not, then node **a** checks if $P_{b,d} > P_{a,d}$, that is node **b** has more chances to meet destination node **d** than meeting the destination node to node **a** .
- If it is true, then node **a** will hand over half of its copies to node **b** and update both nodes, their delivery predictabilities and the summary vectors.
- Otherwise, node **a** will not give any copy to node **b**. After that nodes **a** and **b** will both act like source nodes and hand over half of their copies to encountered nodes, if they are satisfied with the above condition.
- The process continues to find the destination node until the source node have one copy. It is the final stage of the spray phase. After that these nodes enter the wait phase. That means nodes opt for the direct transmission to find the destination node.

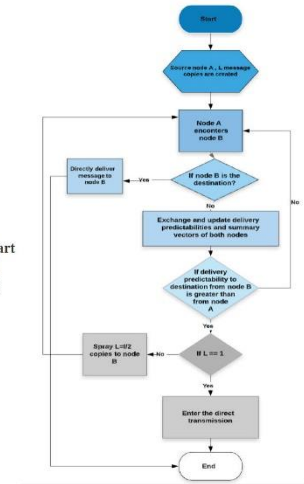


Figure 2 : Flow chart of the proposed SPROP routing

RESULTS

Performance Metrics

$$Delivery\ Probability = \frac{Number\ of\ message\ received}{Number\ of\ message\ sent}$$

$$Overhead\ Ratio = \frac{Number\ of\ message\ forwarded - Number\ of\ message\ received}{Number\ of\ message\ received}$$

$$Latency\ Average = \sum_{i=1}^{n} \frac{Time\ when\ message\ received - Time\ when\ message\ produced}{Number\ of\ message\ received}$$

Results

Varying message sizes

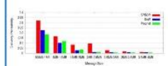


Figure 3: Delivery Probability vs Message sizes

Increasing value of message size gives impact to decrease of delivery probability. It is since where unsuccessful to relay message to other nodes.

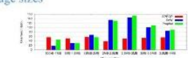


Figure 4: Overhead ratio vs Message sizes

Increasing value of message size gives impact to tend for increasing overhead ratio. It is since where unsuccessful to relay message to other nodes because of buffer already full due to of those message.

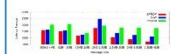


Figure 5: Latency Average vs Message sizes

Result of experiment describes increasing value of message size gives impact to decrease the latency average.

Varying time-to-live time

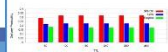


Figure 6: Delivery Probability Vs Time To Live

Increasing value of time to live gives impact to increase the delivery ratio. TTL helps limiting number copies of message in network.

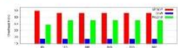


Figure 7: Overhead Ratio Vs Time To Live

Increasing value of time to live gives impact to decrease the overhead ratio.

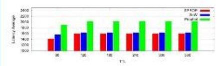


Figure 8: Latency Average Vs Time To Live

Increasing value of time to live gives impact to increase the latency average. It is since lifetime of message increase the message have to wait longer in buffer before it is either delivered toward destination node or dropped message when lifetime expired.

Varying Buffer size

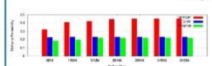


Figure 9: Delivery Probability Vs Buffer Size

Increasing value of buffer size gives impact to increase of delivery probability. It means if the network has high buffer size then many messages can stored on buffer.

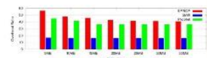


Figure 10: Overhead Ratio Vs Buffer Size

Increasing value of buffer size made value of overhead ratio decreasing. It is because when buffer size of network increases then messages dropped is less then delivery probability will increase.

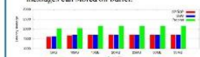


Figure 11: Latency Average Vs Buffer Size

Increasing the value of buffer size make the latency average is higher also.

CONCLUSION

- In this research, we propose the Spray with Probability Routing Protocol (SPROP) for opportunistic networks. In SPROP, a delivery probability function is set up to direct the different number of copies to the destination during the spray phase; and the last one copy is directly delivered to the destination node in the wait phase. We evaluate the proposed SPROP routing under the ONE simulator in different scenarios, simulation experiments show that the proposed SPROP outperforms the other routings (Spray and wait, PRoPHET) in terms of the delivery probability, the overhead ratio and latency average.
- Compare with other algorithms, the proposed routing SPROPs' overhead ratio is relatively high, but its other two performance merits make up the shortcoming. The analysis shows that the proposed SPROP specially adapted for the frequently disconnected opportunistic network.

REFERENCES

[1]. T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Spray and wait: an efficient routing scheme for intermittently connected mobile networks," in 51 Department of Computer Science, University of Jaffna Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking. ACM, 2005, pp. 252–259.

[2]. A. Lindgren, A. Doria, and O. Schelén, "Probabilistic routing in intermittently connected networks," in ACM International Symposium on Mobile Ad Hoc Networking and Computing, MobiHoc 2003: 01/06/2003-03/06/2003, 2003.

[3]. D. Yulianti, S. Mandala, D. Nasien, A. Ngadi, and Y. Coulibaly, "Performance comparison of epidemic, prophet, spray and wait, binary spray and wait, and prophetv2," Faculty of Computing, Universiti Teknologi Malaysia.

ANALYSIS OF METHODS TO HANDLE MEDICAL SENSOR DATA TOWARDS HEALTH DISORDER IDENTIFICATION

S. Meruja and E.Y.A. Charles
Department of Computer Science, Faculty of Science, University of Jaffna
merujajoseph@gmail.com, charles.ey@univ.jfn.ac.lk

SLSWCS Student Workshop on Computer Science

Introduction

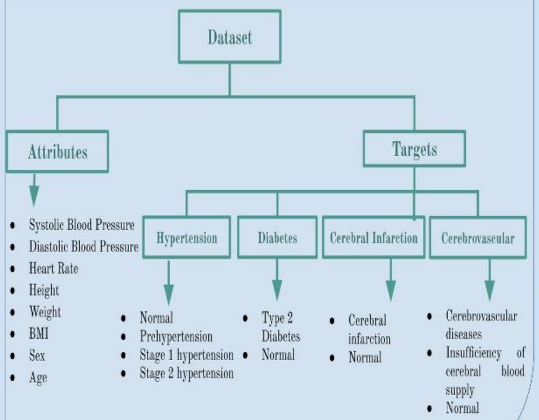
Medical sensors are important for monitoring the health condition of a person. The huge amount of data generated by these sensors, has a great potential for early detection of disorders. This study focuses on analysing methods to handle medical sensor data, particularly the data from Photo-plethysmograph (PPG). PPG quantifies the volumetric change of the heart by measuring the light transmission or reflection on arteries.

Objective

- Analyse various pre-processing approaches to handle issues arising in medical datasets, especially the imbalance of data items.
- Compare various classifiers to identify an optimal classifier for medical sensor data.

Dataset

- This study has used PPG-BP dataset containing features namely, Systolic Blood Pressure, Diastolic Blood Pressure, Heart Rate, Age, Sex and Body Mass Index (BMI) which were collected from 219 persons.
- Systolic blood pressure, diastolic blood pressure and heart rate were calculated from the PPG recordings.
- The dataset contains data for normal persons and persons with four disorders namely; hypertension, diabetes, cerebral infarction, and cerebrovascular disease.



Methodology

➔ Analyse the data and identify appropriate methods to represent data suitable for Machine Learning.

➔ Identify appropriate methods to handle class imbalances in the data set.

The dataset selected for this study is highly imbalanced. Dataset contains a higher number of records from normal persons. To avoid the effect of ***overfitting*** during learning, this study used:

Undersampling technique: Taking a random number of data with target value 'Normal' according to the ratio of other classes.

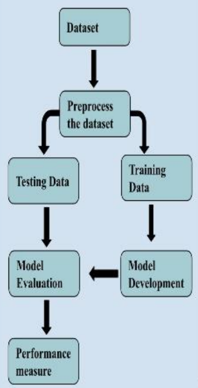
Duplicating the dataset: Make a copy of dataset and scaled it and added to the dataset. Again it leads to overfitting.

Undersampling and duplicating: Taking a random number of data with target value 'Normal' according to the ratio of other classes and a copy of the dataset is scaled and added to the dataset.

➔ Identify appropriate Machine learning methods suitable for the selected dataset. This study selected SVM, K-Nearest Neighbor and Naïve Bayes algorithms.

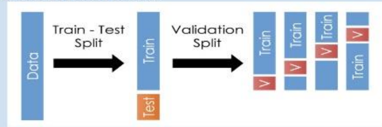
➔ Construction of classifiers - Relationship between attributes and target values. Four different models were constructed, each for one disorder.

➔ Evaluate the constructed models and revise the whole process to identify methods and models for optimal performance.



Experimental setup

Proposed methods and models were implemented and tested using sci-kit learn library.



- The data set is divided into four parts, each for identifying a particular disorder.
- As the standard practice, 70% of the dataset was used to train the models and 30% was used for testing.
- Feature Selection : Recursive Feature Elimination - To reduce the number of parameters used to construct the classifiers.
- Classifiers: Linear SVM-OvA, Linear SVM-OvO, K-Nearest Neighbour, and Naive Bayes

Testing Result

Accuracy, Precision, F1 score and Recall are used as the performance metric to validate the models and confusion matrix to analyse the performance of the proposed approaches.

Accuracy = (TP + TN) /(TP + TN + FP + FN)
Precision = TP /(TP + FP)
Recall = TP /(TP + FN)
F1 Score = 2*[(Precision * Recall)/(Precision + Recall)]

TP – True Positive, TN – True Negative, FP – False Positive, FN – False Negative

Table 1: The classification rate of the proposed method for each disease.

Classifier	*Hypertension*		*Diabetes*		*Cerebrovascular*		*Cerebral Infarction*	
	Overall	*N-Fold*	*Overall*	*N-Fold*	*Overall*	*N-Fold*	*Overall*	*N-Fold*
SVM-OvA	85.45%	86.79%	63.16%	59.89%	66.67%	88.57%	90.00%	87.50%
SVM-OvO	98.18%	96.82%	63.16%	59.89%	71.43%	74.29%	90.00%	80.00%
K-Nearest Neighbour	83.64%	83.22%	84.21%	57.81%	90.48%	62%	100.00%	83%
Naive Bayes	92.73%	90.91%	52.63%	65.89%	71.43%	85.71%	90.00%	92.50%

Discussion

➔ This study concentrated on two aspects:
- ◆ Identifying the morphological information of PPG signals to identify disorders.
- ◆ Implementation of machine learning models to identify the type of disorder.

➔ KNN provides the highest accuracy for identifying Diabetes (84.21%), Cerebral infarction (100%), and Cerebrovascular disease (90.48%) than other models for different K.

➔ Even though SVM is considered as a superior classifier compared to KNN and Naive Bayes it has achieved the highest accuracy only for identifying Hypertension (98.41%)

Conclusion

- This study analyzed the various methods and models on using PPG signals for Cardiovascular disorder identification.
- Learning models are able to perform better on imbalanced data sets in terms of prediction accuracy when the proposed techniques are applied.
- This study shows that the PPG signal data can be used for prediction of certain types of Cardiovascular disorders with adequate accuracy.

References

[1]. N.D.K.G. Dharmasiri and S. Vasanthapriyan, "Approach to Heart Diseases Diagnosis and Monitoring through Machine Learning and iOS Mobile Application," Proceedings of the International Conference on Advances in ICT for Emerging Regions ICTers, vol. 18, pp. 407–412, 2018.
[2]. Liang, Y. et al, "A new, short-recorded photoplethysmogram dataset for blood pressure monitoring in China," Sci. Data 5:180020 doi: 10.1038/sdata.2018..
[3]. Arrozag Ave, Hamdan Fauzan S. Rhandy Adhitya and Hasballah Zakaria, "Earlier detection of cardiovascular diseases with photoplethysmogram (PPG) sensor," Proceedings of the International Conference on Electrical Engineering and Informatics, pp. 676 681, 2015.

Is Soft pooling better than Max and Average pooling?
A comparative study on HEp-2 cells and Retinal image classification tasks

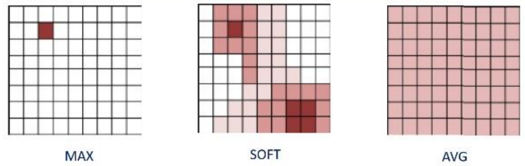
Nirthika R., Siyamalan M., and Ramanan A.
Department of Computer Science, University of Jaffna
{nirthika, siyam, a.ramanan}@univ.jfn.ac.lk

SLSWCS Student Workshop on Computer Science

Abstract

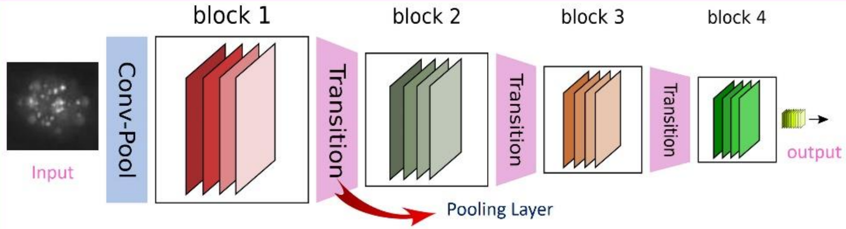
Convolutional Neural Network (CNN) has been widely used for medical image classification [1], where, pooling layers are used for down sampling the feature maps by summarizing the presence of features in local regions of the feature map. Average and Max are the widely used pooling methods. Since average pooling summarizes all the features in the feature map, background regions may dominate in the pooled representation. On the other hand, max pooling can capture noisy features as it focuses on the most activated features. To overcome this, Soft pooling have been proposed [2]. However, soft pooling has not been well explored for medical image analysis. Therefore, this work focuses on investigating its performance on two different medical image classification tasks, i.e., cell image classification and diabetic retinopathy image classification. Our experiments show that soft pooling does not produce significant improvement in performance compare to max and average pooling.

Max, Average and Soft Pooling



- **Max** pooling considers only the maximum element and ignores all the rest of the feature map. Hence, captures noisy features.
- **Average** pooling considers all the elements in the feature map. Hence, captures unwanted background information.
- **Soft** pooling captures information from a set of maximum activated elements of the feature map. Hence, can overcome the above mentioned problems with the max and average pooling.

Network Architecture



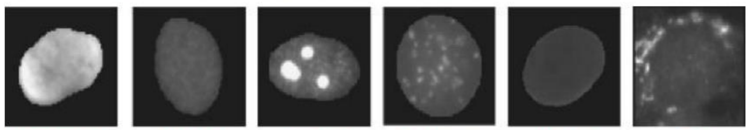
Soft pooling output over region R_i for input x:

$$f_{\text{soft}}(X) = \frac{\ln\left[\frac{1}{|R_i|}\sum_{x_j \in R_i} \exp(\lambda x_j)\right]}{\lambda}$$

where, λ is a hyper-parameter, when $\lambda \rightarrow \infty$ max pooling and $\lambda \rightarrow 0$ average pooling can be obtained.

Dataset & Experimental Setup

1. HEp-2 cell dataset [3] contains Human Epithelial type 2 (HEp-2) cell patterns from six categories



Homogeneous, Speckled, Nucleolar, Centromere, NuMem, Golgi

We used 15,314 images for training and 10,764 images for testing. Each experiment was iterated three times and the average and the standard deviation of mean per-class accuracies (MCA) over iterations are reported.

2. Diabetic Retinopathy (DR) dataset [4] contains images indicating the presence of diabetic retinopathy.

4-Proliferative DR, 3-Severe, 2-Moderate, 1-Mild, 0-No DR

We used 8,837 images for training and 4,771 images for testing. Each experiment was iterated three times and the average and the standard deviation of quadratic weighted kappa over iterations are reported.

Experiments and Results

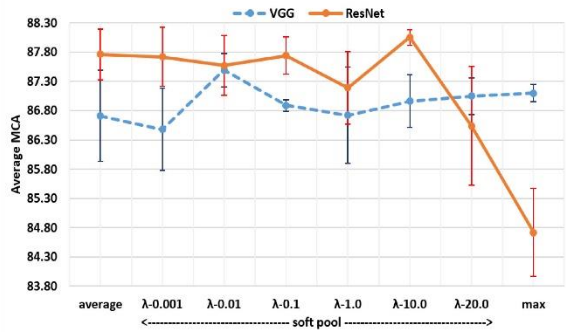


Figure 1: HEp-2 cell image dataset with different pooling strategies.

HEp2 Cell images: For small architecture (VGG), soft pooling performs approximately same as max and average pooling, but for large architecture (ResNet), soft pooling performs approximately same as average pooling and better than max pooling. This result is expected as all the contents in the cell images are necessary to determine its class.

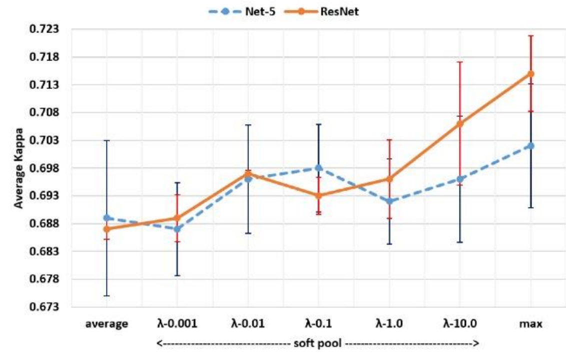


Figure 2: DR image classification with different pooling strategies.

DR dataset: Soft pooling performs better than average pooling and worse than max pooling for both small (Net-5 [1]) and large (ResNet) networks. For both architectures max pooling performs better as the lesions usually occupy small regions in the DR images.

Conclusion

This work compares max, average, and soft pooling for HEp-2 cell and DR image classification tasks with CNN. For each of the task we used two different CNN architectures. For the DR dataset, max pooling performs better than average pooling since the lesions cover only part of the images, and for the cell dataset average pooling performs better as the images cover all the cell regions. We found that soft pooling does not perform better than average and max pooling.

References

[1] Z. Wang and J. Yang, "Diabetic retinopathy detection via deep convolutional networks for discriminative localization and visual explanation," arXiv preprint arXiv:1703.10757, 2017.
[2] Wei, Z., *et al.*, "Building detail-sensitive semantic segmentation networks with polynomial pooling", IEEE CVPR, 2019.
[3] https://mivia.unisa.it/datasets/biomedical-image-datasets/hep2-image-dataset
[4] https://www.kaggle.com/c/diabetic-retinopathy-detection/data

A WEB-BASED DENGUE MONITORING AND WARNING SYSTEM

Nirthika R., Ramanan A., and Surendran S.N.
Department of Computer Science, University of Jaffna
{nirthika, a.ramanan, noble}@univ.jfn.ac.lk

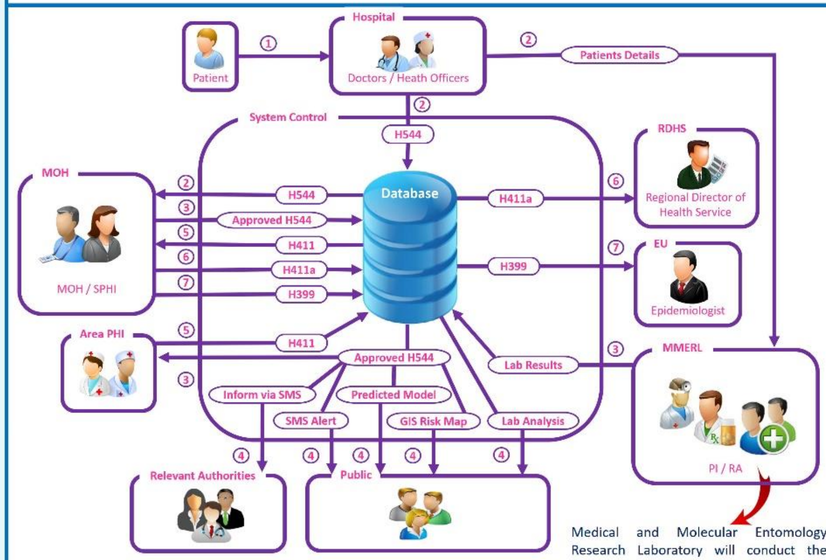
Introduction

Dengue is the most prevalent mosquito-borne viral disease that spreads rapidly among people [1]. In such case, disease prevention and control measures should be improved via early detection and monitoring of outbreaks. Early recognition of an outbreak supports the health officers to plan pre-emptive measures. However, there is no completely computerised dengue surveillance system in Sri Lanka to notify and prevent these kind of communicable diseases [2]. Present manual dengue outbreak prediction is incomplete and time consuming [3]. This study proposes a system that automates the entire manual surveillance system which deals with various level of health related officers by providing them facilities such as monitoring dengue risk areas and reporting dengue outbreaks to the public.

Contributions

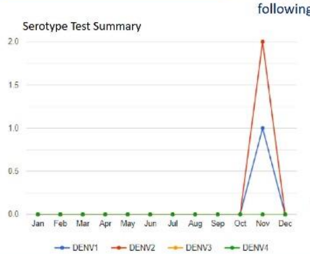
- In order to make an efficient reporting and early warning system, a gap analysis was conducted with the help of northern region health officers and doctors of Jaffna Teaching Hospital to improve the solution.
- The developed dengue surveillance system mainly targets to alert the public.
- In the system, a dynamic GIS dengue risk map is integrated and a heatmap can be visualised to alert the public.
- Medical and Molecular Entomology Research Laboratory do the additional dengue confirmative and serotype test and provide the details through the system.

Methodology

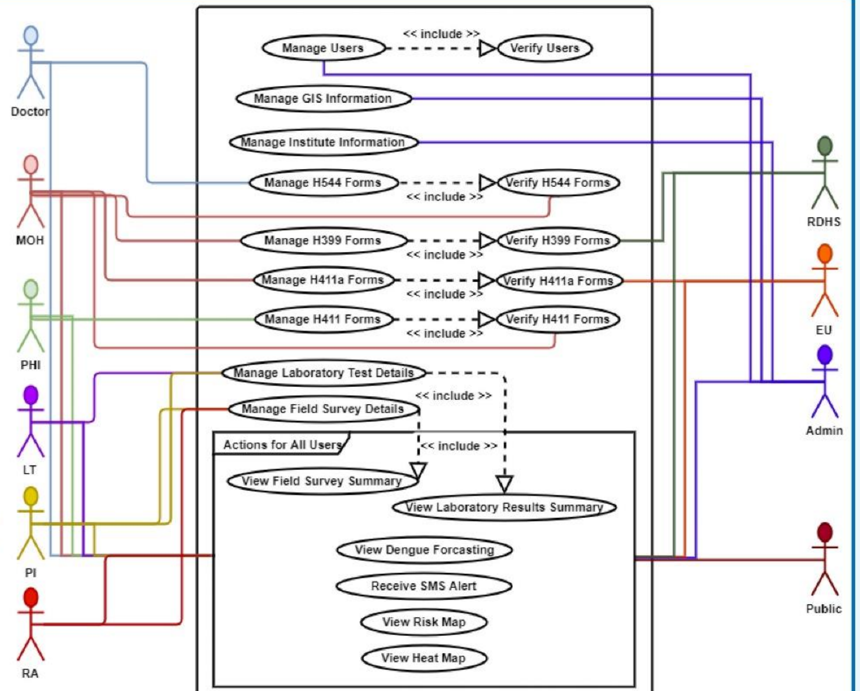


Medical and Molecular Entomology Research Laboratory will conduct the following tests

H399 : Weekly Return of Communicable Diseases
→ MOH will enter the data
H411 : Communicable Disease Report - Part I
→ PHI will enter the data
H411a: Communicable Disease Report - Part II
→ MOH will enter the data
H544 : Notification of Communicable Disease
→ Doctor will enter the data



1. Dengue confirmative test:
   - NS-1
   - IgM
   - IgG
2. Dengue virus serotype test:
   - DENV1
   - DENV2
   - DENV3
   - DENV4



MOH : Medical Officers of Health
SPHI : Supervising Public Health Inspector
PHI : Public Health Inspector
LT : Laboratory Technician
RA : Research Assistant
EU : Epidemiology Unit
MMERL: Medical and Molecular Entomology Research Laboratory
PI : Principal Investigator
RDHS : Regional Directorate of Health Services

Strengths of the System

- Active surveillance
- Inclusion of dengue risk notifications from private hospitals
- Integrating laboratory test results in to the notification process
- Centralised storage of dengue details that ensures data privacy
- Reduced time of notification process enables fast prevention
- Bridging the gap between medical officers and health officers.
- Summarising the data more user friendly
- Dengue risk predictions will be effective

Innovativeness

The proposed system is developed to be complete, timeliness and sensitivity that influence the information flow in the notification process when comparing to the existing manual notifiable disease surveillance system at different levels. The systems speeds-up the detection and prevention of dengue cases via a user friendly web-based intelligent system. In addition, public will be alerted through dengue risk map. Furthermore, the system is expected to predict future dengue outbreak through a mathematical model.

Future Work

As a near future work, the risk factors for dengue is being analysed and will be modelled to produce more valid prediction of outbreaks.

References

[1] Pathirana, S., Kawabata, M., and Goonetilake, R., "Study of Potential Risk of Dengue Disease Out break in Sri Lanka using GIS and Statistical Modelling", Journal of Rural and Tropical Public Health, vol. 8, pp. 8, 2009.
[2] Munasinghe, A., Premaratne, H. L., and Fernando, M. G. N.A.S., "Towards an Early Warning System to Combat Dengue", International Journal of Computer Science and Electronics Engineering, vol. 1(2), pp. 252-256, 2013.
[3] Kalpana, C., Mahesan, S., and Bath, P. A, "Notifiable Disease Surveillance in Sri Lanka and the United Kingdom: A comparative study", Sri Lanka Journal of Bio-Medical Informatics, vol. 4, 2013.

COPY-MOVE IMAGE FORGERY DETECTION USING SIFT DESCRIPTORS

K. Parkavi and A. Ramanan
Department of Computer Science, Faculty of Science, University of Jaffna
parkavi113@gmail.com

Introduction

Image manipulation and editing is very common in this multimedia era. Forgery images are known as manipulated images if the semantics of the original image is changed. This study is developed to detect copy-move image forgery detection under keypoint-based approach using SIFT descriptors since block-based techniques have high computational complexity.

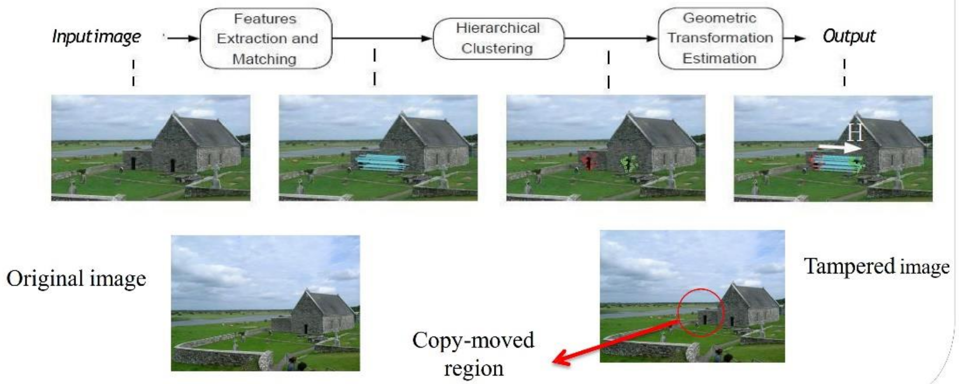
Background

There are two main approaches for image forgery detection: Active and passive approaches [1]. Copy-move falls under passive approach. Image manipulation can be achieved through image enhancing, image retouching, image splicing, image morphing and copy-move. Certain region(s) can be copied and pasted in another region of the same image so as to hide or misinterpret a particular image. This manipulation technique is known as copy-move forgery. To detect copy-move images there are two main approaches: Keypoint feature extraction and block-based features. In keypoint-based detection, keypoints are extracted using either scale invariant feature transform (SIFT) or speeded up robust features (SURF), and in block-based detection a given image is subdivided into blocks using various block-based techniques. Even though block based detection gives better performance, it has high computational complexity.

Methodology

- Given test image is first preprocessed. Preprocessing includes changing the RGB image into grayscale image. This step is optional if the image is already a grayscale image.
- For the preprocessed image, a set of keypoints and corresponding SIFT descriptors are extracted. Matching operation is performed in the SIFT space among the descriptors in order to identify local patches that are similar. The best candidate match for each keypoint is found by identifying its nearest neighbor from all the rest of the keypoints of the image, which is the keypoint with the minimum Euclidean distance in the SIFT space. To identify among multiple copied regions, g2NN-ratios [2] between the adjacent pairs of distances are found. g2NN ratio is the ratio between two adjacent distances.
- The ratios that are greater than a predefined threshold value are chosen for the next step clustering. An agglomerative hierarchical clustering is performed on spatial locations of the matched keypoints to identify possible cloned areas. Based on the adopted linkage method, a specific tree structure is obtained. An appropriate cut-off value is chosen and number of matching keypoints is determined.
- Forgery is detected through previously obtained number of clusters and number of keypoints in each cluster. If there are at least two clusters and in each cluster if there are at least three pairs of matching keypoints, then such image is detected as a forged image. If an image has copy-moved region, there should be at least one similar cluster and to detect forgery there should be at least three pairs of matching keypoints. Therefore this limitation is achieved.

Methodology…



Objective

To improve the overall performance in detecting the copy-move forgery images using SIFT features.

Experimental Setup

The experiment has been carried out in such a way training data is 70% of the dataset and testing is 30% of the dataset.

Dataset:

MICC-F220 dataset [2]

There are 220 images with 110 tampered and 110 original images with the resolution varies from 722 × 480 pixels to 800 × 600 pixels and on average, 1.2% size of the whole image is covered by forged region.

The dataset is divided into two classes namely "Original" and "Tampered" where all the non-tampered images were grouped into the first category and the tampered images were grouped into the other.

Testing Results

Performance in detecting the forged image using, 'Ward linkage' hierarchical clustering and one-versus-one SVMs.

An image is considered as a copy-move attacked image if the method detects two or more clusters with at least three pairs of matching keypoints.

The performance indicates the True Positive Rate (TPR) which correctly classifies the tampered images in the dataset.

The table shows that choosing cut-off threshold of 2.2 in hierarchical clustering gives better performance in detecting copy-moved images. The cut-off threshold here is referred to the point at which the obtained dendrogram is cut to determine the number of clusters in a particular image.

Cut-off threshold	Performance
2.2	38%
2.5	51%
2.8	42%
3.2	35%

Conclusion

- This method shows a good performance in detecting copy-moved forgery images even though the performance can be further improved.
- The performance can be further improved by iterating the method for various linkage methods and estimate the best cut-off threshold.
- This method falsely detect an original image as tampered when there are two identical regions or objects placed in a particular image.

Methodology…

Input greyscale image
↓
Keypoint extraction using SIFT
↓
Keypoint matching
↓
Hierarchical clustering using Ward linkage
↓
Forgery Detection

References

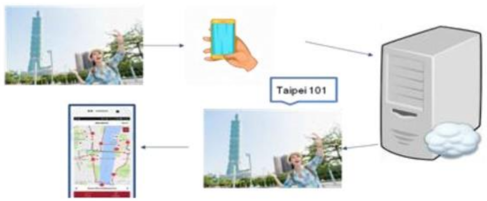
1. V. Christlein and J. Jordan, "An Evaluation of Popular Copy-move Forgery Detection Approaches," IEEE Transactions on information forensics and security, pp. 1-26, 2012.
2. I. Amerini, L. Ballan, R. Caldelli, A. D. Bimbo, and G. Serra, "A SIFT-based Forensic Method for Copy-Move Attack Detection and Transformation Recovery," IEEE Transactions on Information Forensics and Security, vol. 6, no. 3, pp. 1099–1110, Sep.2011.
3. S.Kumar, J.Desai., and S.Mukherjee , "A Fast Keypoint Based Hybrid Method for Copy Move Forgery Detection," International Journal Computing Digital System, 4(2), 2015.
4. B. Yang, X.Sun, H. Guo, Z. Xia and X.Chen, "A copy-move forgery detection method based on cmfd-sift," Multimedia Tools and Applications pp. 1–19, 2017.
5. J.Zhao and W. Zha, "Passive forensics for region duplication image forgery based on Harris feature points and Local Binary Patterns," in Mathematical Problems in Engineering, pp. 1–12, 2013.

AN ATTENTION BASED-CONVOLUTIONAL NEURAL NETWORK FOR LANDMARK RECOGNITION IN ASIAN REGION

S. Perera and A. Ramanan
Department of Computer Science, Faculty of Science, University of Jaffna
shehanperera.office@gmail.com, a.ramanan@univ.jfn.ac.lk

Introduction

Landmark recognition is very helpful in many ways. Tourist can find out new attractive locations from social media and they can plan for a visit. Landmark recognition may also increase the interests of tourists to visit certain places through which it can contribute to the economy of that country. Also landmark recognition greatly helps people to better understand and organise their photo collections.



Objective

To build an intelligent system for recognising landmarks in the Asian region to improve tourism and business near to landmarks.

Dataset

- The Asian landmark dataset chosen from the **Google Landmark Recognition Challenge** dataset [4]. We choose 30 different landmarks from 30 different Asian countries.
- We used 35 images per landmark, for training and 15 images for testing the model.

Methodology

- At the initial stage of this study, we have tested different classifiers on landmark classification and found CNN to outperform SVM, k-NN and Random Forests (See Table 1).
- We fine-tuned the pre-trained VGG-11 model to obtain better results for the Google Landmark Recognition dataset. In fine-tuning process we note that using Optimizer as Adamax, Learning rate as 0.001, Loss function as Cross Entropy Loss give better result. Since we have small number of data we used 100 epochs.
- We also add an attention branch to the framework that combines the predicted *conv5* features and fine grained features to gate or magnify the *conv5* features to improve the precision of landmark classification (See Figure 1).
- We utilize the $14\times14\times512$ predicted feature map of *conv5* to max pool the features in to $1\times1\times512$. After that we used one convolution layer to obtain different number of features. From several trials we found that $1\times1\times1024$ is the best (See Table 2). Thereafter the features are reconstructed through upsampling to yield $14\times14\times1024$, thus producing dense feature map $14\times14\times(512+1024)$. The proposed method shows 94% of classification accuracy by contributing the CNN to yield attentive image features (See Table 2).

Methodology

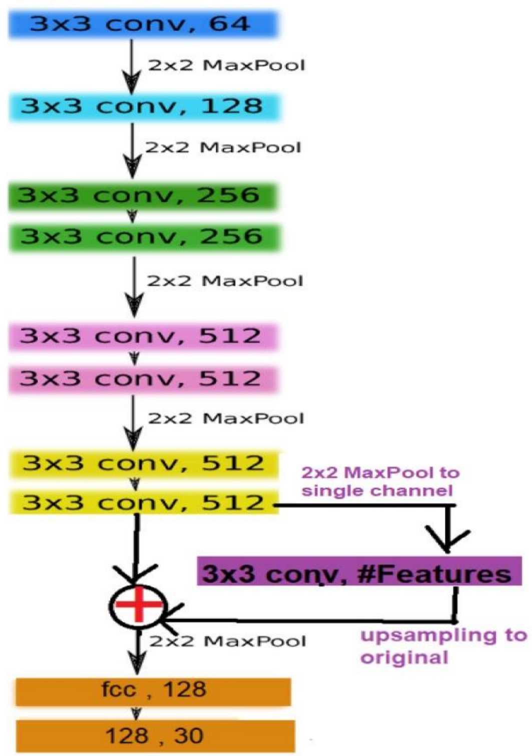


Figure 1: Modified VGG-11 Model

Conclusion

In this study we showed a modification to VGG-11that can recognize Asian landmarks with 94% classification accuracy. Table 1 shows that basically deep learning techniques like CNN works better on image classification rather than shallow learning methods. Even from several CNN models selected in this study VGG-11 model performs better in this landmark dataset. Adding attention branch for original network extracts more important features using max pooling. These important features combined with original features works to improve the knowledge of the network.

Test Results

Table 1: Performance comparison of different classifiers on Landmark classification

Classifier	Classification rate
Nearest Neighbor	35%
Random Forest	42%
SVM	70%
CNN	78%
Fine tune CNN	90%
Proposed Model	94%

Table 2: Performance comparison of different number of features selected at the attention branch on Landmark classification

#Features	Classification rate
16	89%
32	88%
64	90%
128	84%
256	91%
512	87%
1024	94%

Reference

[1]. O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang,A.Karpathy, A. Khosla, M. Bernstein, A.C. Berg and L. Fei-Fei., ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision (IJCV), pp:25-32,2015.

[2]. A. Karpathy, G. Toderici, S. Shetty, T. Leung,R. Sukthankar, and L. Fei-Fei. Large-scale video classification with convolutional neural networks. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1725–1732, 2014.

[3]. H. Noh, A. Araujo, J. Sim, T. Weyand, and B. Han, "Large-Scale Image Retrieval with Attentive Deep Local Features", In Proceedings of IEEE Conference on Computer Vision (ICCV), pp:789-793,2017.

[4]. Google Landmark Recognition Challage Dataset : www.kaggle.com/c/landmark-recognition-challenge

Fake news Detection

R.Prithweeraj and B.Mayurathan

Department of computer Science,University of Jaffna,Sri Lanka.

Prithweeskash93@gmail.com , barathym@univ.jfn.ac.lk

Introduction

- The continuous growth of social media has provided users with more convenient ways to access news than ever before.
- As people continue to benefit from the convenience and easy accessibility of social media, they also expose themselves to certain noisy and inaccurate information spread on social media, especially fake news, which consists of articles intentionally written to convey false information for a variety of purposes such as financial or political manipulation.
- Due to extensive spread of fake news on social and news media it became an emerging research topic now a days that gained attention.

Objective

- In this work, an automatic identification of fake news in online news is proposed.
- In this poster a model is proposed based on some readability matrices such as the TF–IDF, Count-Vectorizer and Hashing-Vectorizer in order to differentiate fake news from real news.

Methodology

- The diagrammatic representation of the proposed methodology is given in Figure 1. It describes the steps that were involved in this research.
- Initially starts with collecting the data from multiple sources, then removing unnecessary characters and words from the data.
- The collected dataset is split into training and testing sets. For instance around 80% of the dataset is used for training and 20% for testing.
- Term Frequency–Inverse Document Frequency (TF–IDF), Count–Vectorizer (CV) and Hashing–Vectorizer (HV) features are extracted from all the terms/words involved in all the documents in the training corpus.
- During the classification process Naive Bayes and Passive Aggressive Classifiers are used to classify the fake news from real news.

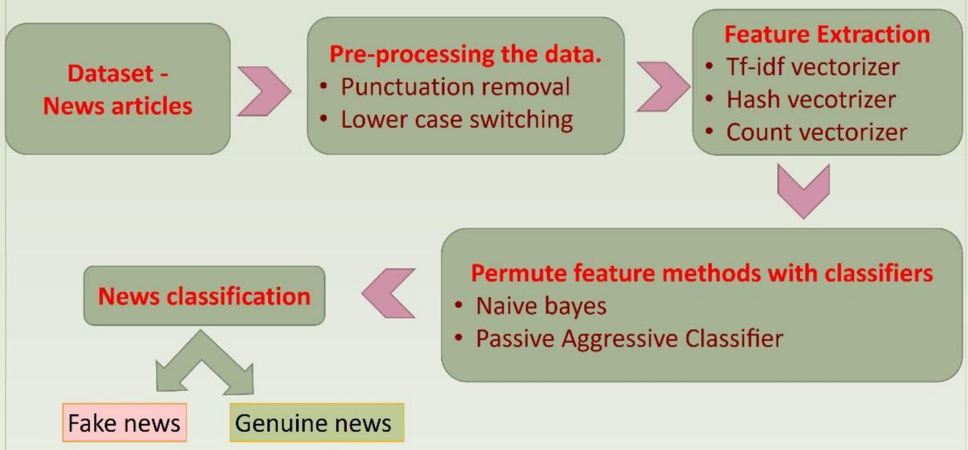


Figure 1: diagrammatic representation of the proposed methodology

Conclusion

- According to this experimental results, tf-idf feature with passive aggressive classification gives a better performance compared to other methods. Also, passive aggressive classifier gives better performance than Naïve Bayes.

Experimental results

- In this experiment, a database is collected from kaggle. It is a collection of data which mainly comprises of single statistical data matrix, database table where every row corresponds to each member in datasets and each column represents variable.
- The dataset list values for each variable such as title, id, author and label with 6335 data.
- Table I shows the performance of the different features that are used in this experiment.

Features	Naive Bayes (%)	Passive Aggressive Classifier (%)
TF–IDF	86.3	**94.1**
Count-Vectorizer	90.2	89.9
Hashing-Vectorizer	90.8	92.5

Table 1: Classification performance of the proposed methodology

References

1. Jiawei, Zhang & Cui, Limeng & Fu, Yanjie & Gouza, Fisher, Fake News Detection with Deep Diffusive Network Model, 2008
2. Rubin, Victoria L., Yimin Chen, and Niall J. Conroy. "Deception detection for news: three types of fakes." In Proceedings of the 78th ASIST Annual Meeting: Information Science with Impact: Research in and for the Community, p. 83. American Society for Information Science, 2015.
3. Pérez-Rosas, Verónica, Bennett Kleinberg, Alexandra Lefevre, and Rada Mihalcea. "Automatic detection of fake news." , Proceedings of the 27th International Conference on Computational Linguistics, 2018.

An Improved Approach of Iterative Keypoint Selection with Spatial Pyramid Matching for Visual Object Classification

R.M.S.Ranathunga and A. Ramanan
Department of Computer Science, University of Jaffna, Sri Lanka
rmshashi5@gmail.com

Introduction

The generic framework of Bag-of-Features (BoF) is depicted in Figure 1. However, one of the problems with this paradigm raise is the number of keypoint that need to be detected from images to generate the Bag-of-Features is usually very large which causes two problems. First, the computational cost during the feature vector generation step is high and Second, some of the detected keypoint are not helpful for recognition. Therefore, this study introduces a framework called Iterative Keypoint Selection (IKS) [4] to select representative keypoints for reducing the computational time to generate the Bag-of-Features. Also this work introduces another technique called Spatial Pyramid Matching (SPM) [3] to retrieve more image details in higher resolutions.

Objectives

To make Bag-of-feature representation to be efficient with stable performance by using Iterative Keypoint Selection and Spatial Pyramid Matching techniques .

Methodology

The overall framework is depicted in Figure 2 and the proposed techniques are depicted in Figure 3 and 4.

1.Iterative Keypoint selection:

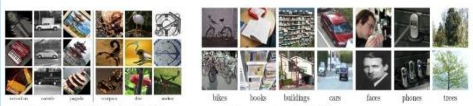
Resulting in fewer but more representative keypoint descriptors in an image.

2.Spatial Pyramid Matching:

Partitioning the image into increasingly fine sub- regions and computing histograms of local features found inside each sub-region. Resulting spatial pyramid is a simple and computationally efficient extension of an orderless BoF image representation.

Experimental Setup

Caltech 101 Xerox7



Caltech101: 9,146 images ; **Xerox7:**1776 images

- Caltech101: 30 images per class training and testing on the rest.
- Xerox7: 70% training , 30% testing
- Features: Dense SIFT Descriptors
- Vocabulary Construction: K-means algorithm
- Classification: Linear OVA-SVMs
- Distance thresholds in IKS: 0.5, 0.6, 0.7
- L=2 in spatial pyramid matching

Traditional Bag-of-Features (BoF) Approach

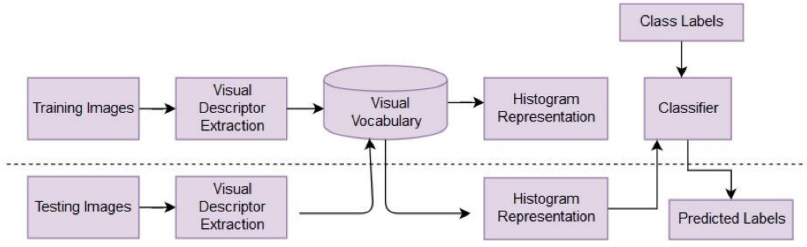


Figure 1. Traditional Bag-of-Features Approach

Iterative Keypoint Selection (IKS)

Input: training dataset I_i (i.e., the i-th image that contains m keypoints)
Output: selected keypoints of I_i (i.e., SK)

Reduced set of keypoints (*Reduced_Keypoints*) ← training dataset
Threshold T ← the distance parameter
While any keypoint can be found from *Reduced_Keypoints*
Get the size of *Reduced_Keypoints*
Get the random number (*random_number*) between 1 and the size of *Reduced_Keypoints*
Randomly find a keypoint as the representative keypoint (RK) from *Reduced_Keypoints* through *random_number*
Put RK in SK
For f from 1 to the size of *Reduced_Keypoints*
If f is not equal to *random_number* then
Find the distance between RK and the f-th keypoint from *Reduced_Keypoints*
If the distance > T then
Put the f-th keypoint from *Reduced_Keypoints* in a temporary matrix
End if
End if
End for
Reduced_Keypoints ← find all keypoints which are put in the temporary matrix
End while
Return SK

Figure 3. Iterative Keypoint Selection Algorithm

Spatial Pyramid Matching (SPM)

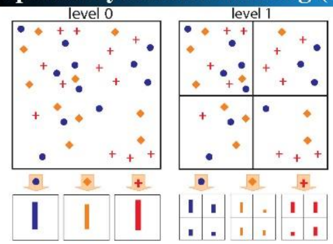


Figure 4. Toy example of constructing a three-level pyramid. The image has three feature types indicated by circles, diamonds and crosses. Subdivide the image at two different levels of resolution. In this study, we choose two levels as not observe any significant in performance beyond two levels.

Construct a sequence of grids at resolutions 0,. . .,L such that the grid at level ℓ has 2^{ℓ} cells along each dimension, where $\ell=0,\ldots,L-1$

The weight associated with each level ℓ is given by the equation,

$$\frac{1}{2^{L}-\ell} \qquad (1)$$

For each level of resolution and each channel, count the features that fall in each spatial bin and weight each spatial histogram according to equation 1.

References

[1] T. Kirishanthy and A. Ramanan, "Creating compact and discriminative visual vocabularies using visual bits," in 2015 International Conference on Digital Image Computing: Techniques and Applications (DICTA). IEEE, 2015, pp. 1–6.

[2]V. Vinoharan and A. Ramanan, "Keypoints and codewords selection for efficient bag-of-features representation," in Future of Information and Communication Conference (FICC). Springer, 2018, pp. 378–390.

[3]S. Lazebnik, C. Schmid, and J. Ponce, "Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories," in 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06),vol. 2. IEEE, 2006, pp. 2169–2178.

[4] W.-C. Lin, C.-F. Tsai, Z.-Y. Chen, and S.-W. Ke, "Keypoint selection for efficient bag-of-words feature generation and effective image classification," Information Sciences, vol. 329, pp. 33–51, 2016.

Proposed methodology

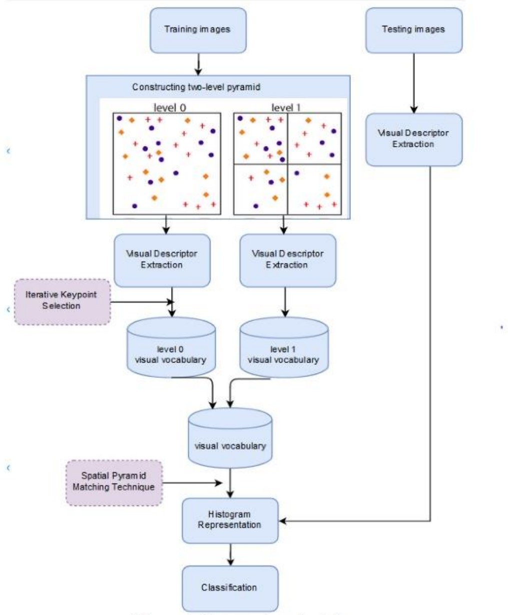


Figure 2. Proposed methodology

Testing Results

Method	Dataset	Classification Rate
Traditional	Caltech101	36.32%
	Xerox7	84.99%
IKS	Caltech101	18.19%
	Xerox7	58.72%
SPM	Caltech101	36.90%
	Xerox7	86.49%
IKS+SPM	Caltech101	23.12%
	Xerox7	81.61%

Table 1: Comparison of classification rates between Standard BoF approach and proposed techniques; IKS and SPM

Discussion and Conclusion

- IKS extracts spatial-based BoF that can provide greater discriminative power and there is a great reduction in the computational time for generating the BOF and spatial-based BoF.
- SPM improves the performance of BoF approach.
- To improve the performance, a supervised learning based keypoint selection approach can be considered for IKS and Convolution Neural Network (CNN) based features can be used for image classification.

Image Reconstruction using spatial and geometrical information

Ranushka Pasindu, Barathy Mayurathan

rdharmaranga@gmail.com, barathym@univ.jfn.ac.lk

Department of Computer Science, University of Jaffna

Abstract

Nowadays, image reconstruction is widely used in many engineering and medical applications. In this work, an approach for reconstructing images is presented and demonstrated.

In this approach, images are reconstructed using its local feature descriptors and its geometric information. For each of the region of interest, visually similar patches are identified from the external image database. Based on the experimental results, an image can be approximately reconstructed using image local feature descriptors like SIFT.

Methodology

A technique based on local feature descriptors and its geometrical information, X,Y coordinates of local features is proposed to reconstruct images. This geometrical information is used to locate the exact point of the local feature descriptor

1. In the initial stage, training images are used to extract local feature descriptors. Extracted local feature descriptors with their corresponding geometrical information is used to generate a database of descriptors. SIFT descriptors are used to describe images. For example, each extracted descriptor is named as

$$D(i) = \{f(i), x(i), y(i), o(i), s(i), index(i)\}$$

where
   - $f(i) \in R^d$ - is the d dimensional of feature descriptor.
   - $x(i), y(i)$ - are the spatial coordinates of the region of interest.
   - $o(i), s(i)$- are the orientation and scale of the extracted feature descriptor.
   - $index(i)$- is the index of the source image from which the feature descriptor was extracted.
2. During the testing time, local feature descriptors are extracted from the query image. Then, extract the suitable image patch from the original image database.
3. Mean Squared Error (MSE) is used to find out the overlapping areas of patches between the new patch that we want to add and patch already existing in the query image. In this experimental design, the upper threshold value of MSE is set as the default threshold (DT) ssin order to reduce overlapping patches.
4. If the threshold value of the new patch is greater than DT, then the new patch is considered as a non-overlapping patch with the existing patches that we already placed it into the query image. So, the new patch is fixed into the query image. Otherwise, the new patch is considered as a overlapping patch and no need to **fixed** into the query imagepairwise matching is used to find out the unknown nearest neighbor descriptor to build up the unknown image nearest neighbor descriptor for $T(j)$ from D.
5. The Nearest neighbor descriptor is used to

Dataset

- The performance of the proposed methodology is evaluated using ZuBuD Image Database. This database includes 201 building classes each have 5 images.
- Figure 1 depicts some example images from ZuBuD Image Database.

Fig. 1. Some sample images from Zurich Building (ZuBuD) Image dataset

Results

Fig. 2. Some sample reconstructed images using the proposed methodology

Experimental Design and testing results

- In these experiments, 20 building classes are randomly selected from the ZuBuD Image Database to calculate the performance of the proposed methodology. Set of parameters such as default threshold (DT) and size of the image patch are tunned with different values to get better reconstruction images.
- A Suitable image patch is selected using the nearest neighbour descriptor from the original image database. To select the suitable size of the image patch which is used to reconstruct images approximately, $3\times3, 6\times6$,and 11×11 sized patches are selected from the original image.
- Also, based on our testing results, 11×11 sized patches gives better-reconstructed images than other sized patches. So, 11×11 sized image patch is used in this experiment.
- Figure 2 gives some testing outputs of the proposed experimental design. Based on our testing outputs, this proposed approach progressively develops an approximation of the unknown image by constructing its region of interest one by one.

Conclusion

- This poster shows that an intensity image can be reconstructed using its spatial and geometrical information.
- The future work for this poster is mainly focusing on algorithm development to reconstruct images that doesn't have enough geometrical information.

References

[1]A., "HOGgles: Visualizi Hiroharu, K. and Tatsuya, H., "Image Reconstruction from Bag-of- Visual-Words", Conference on Computer Vision and Pattern Recogni- tion, pp. 955–962, 2014.

[2] Calonder, M., Lepetit, V., Strecha, C. and Fua, P., "Brief: Binary robust independent elementary features". 11th European conference on Computer vision, page 778–792, 2010.

[3] Alahi, A., Ortiz, R. AND Vandergheynst, "FREAK: Fast Retina Key- point", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 510–517, 2012.

[4] Vondrick, C., Khosla, A., Pirsiavash, H. Malisiewicz, T. and Torralba, ng Object Detection Features", International Conference on Computer Vision ICCV, pp. 1–8, 2013

SOLAR ENERGY FORECASTING WITH MACHINE LEARNING APPROACHES

MD Sabbir Hossain[1] and M. Siyamalan[2]
Department of Computer Science, Faculty of Science, University of Jaffna
[1]falaksabbir3@gmail.com, [2]siyam@univ.jfn.ac.lk

Abstract

Solar power has been used as an energy source by electric utility companies. These companies must forecast accurate production of solar energy to determine fuel needs. Machine Learning plays an important role for this forecasting. In the last decade many research done for forecasting solar energy using various machine learning techniques. In this work, we compare three machine learning approaches such as XGBoost, Random Forest and Linear Regression for solar energy prediction, where, XGBoost and Random Forest are non-linear algorithms and linear regression is a linear one. We show that XGBoost performs better than others.

Dataset and Experimental Settings

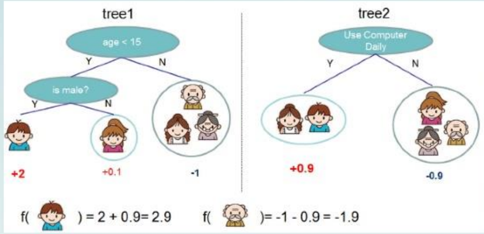
This public dataset is obtained from University of Illinois in Urbana Champaign site and detailed by National Oceanographic and Atmospheric Administration (NOAA). This is an hourly based data contains 165,803 instances. This table shows Weather Features from obtained dataset.

Weather Features	Unit
Cloud Coverage	% range
Visibility	Miles
Temperature	^0C
Dew Point	^0C
Relative Humidity	%
Wind Speed	Mph
Station Pressure	inchHg
Altimeter	inchHg

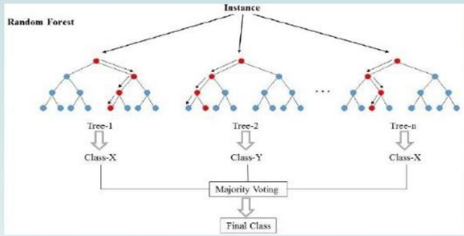
We split the dataset into training and testing set as 77% and 33%. We used Root Mean Squared Error (RMSE) as the evaluation measure.

Methodology

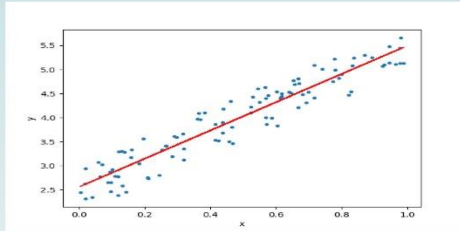
XGBoost and Random Forests are ensemble tree learners. In both cases the final prediction can be obtained by combining the predictions of individual trees. However, they differ from each other by the way the trees are learned. In XGBoost each tree is learnt in an additive manner. At each iteration a tree is found in a greedy way and added to the existing model. On the other hand, in Random Forests each tree is learnt independently from others using a subset of the entire training set. In XGBoost each tree is a weak-learner, but in Random Forests each one is a strong model. XGBoost and Random Forests are non-linear algorithms, on the other hand, Linear Regression is a linear one, Linear Regression models the relationship between a the features and the outcomes.



XGBoost



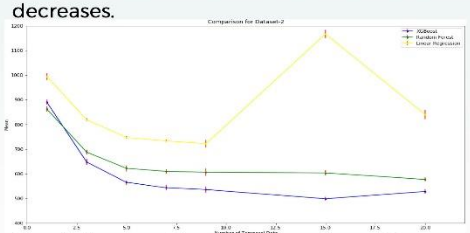
Random Forest



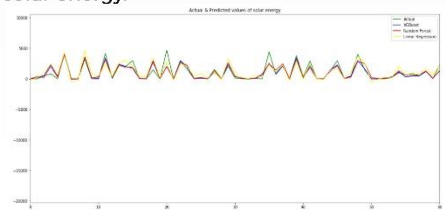
Linear Regression

Results

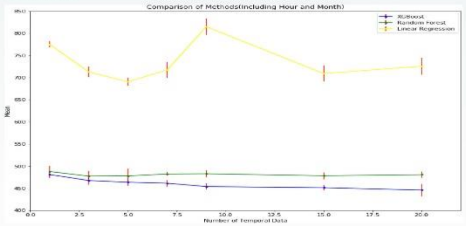
The following figure shows the RMSE values for each method. We compared all the methods with and without temporal data. We found that with temporal data all methods perform better. When the number of temporal data increases, the value of RMSE decreases.



The following figure shows the difference between actual solar energy and predicted solar energy.



There is a big difference between starting and ending index for each month and hour. To improve the performance, we removed this difference by mapping month and hour. The following figure shows the RMSE value for each method with mapped hour and month.



The following table shows the RMSE value for each method with mapped hour and month.

Methods (With Hour and Months)	RMSE with Standard Deviation
XGBoost	480.530 ± 7.650
Random Forest	487.682 ± 13.435
Linear Regression	774.602 ± 6.826

Discussion & Conclusion

After including mapped hour and month, all the methods performance has been improved and gave lower RMSE value with Standard Deviation for all the methods. XGBoost gave the lower RMSE value than other approaches. So, for this research we concluded that XGBoost performs better than other machine learning approaches.

References

[1] Akuzmiakova A.K., Colasg G.C. and McKeehan A. "Short-term Memory Solar Energy Forecasting at University of Illinois", Stanford University. 2017.
[2] Yinghao Chu, Bryan Urquhart, Seyyed M.I., Gohari, Hugo, T.C.Pedro, Jan Kleissl, Carlos and Coimbra F.M. "Short-term reforecasting of power output from a 48 MWe solar PV plant", Solar Energy, vol. 112, 2015, pp. 68-77.
[3]https://github.com/ColasGael/Machine-Learning-for-Solar-Energy-Prediction/tree/master/Datasets

Unsupervised Sentiment Analysis on Tamil Texts

Sajeetha Thavareesan and Sinnathamby Mahesan
sajeethas@esn.ac.lk, mahesans@univ.jfn.ac.lk

SLSWCS Student Workshop on Computer Science

① Introduction

Sentiments are central to almost all human activities and are key influencers of individuals or organizations. It is in the core of understanding sentiments expressed in the text. It groups opinions of written text into *positive*, *negative* or *neutral*.
Several algorithms are there in grouping opinions into positive or negative or neutral. This study uses clustering and lexicon based approaches to determine a suitable approach to performing Sentiment Analysis in Tamil text using a corpus and lexicon.

Clustering is the task of grouping a set of objects in such a way that objects in the same group are similar to each other with respect to certain features. There may be several clusters in a set of objects.

Bag of Words is the representation of the words by their counts appeared in a document.

② Corpus and Lexicon

UJ_Corpus_Opinions corpus consists of reviews and comments with tags of positive/negative [Positive- 1518 and negative- 1173].
For example,

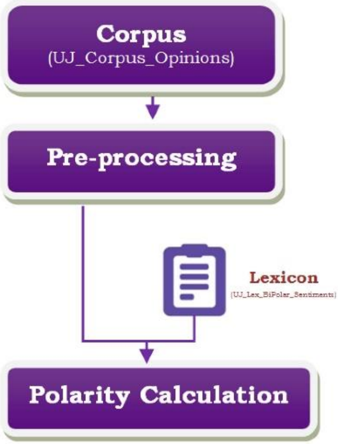
- பயிற்சி ஆட்டத்தில் நியூசிலாந்து அணி வெற்றி 😃 ஆரம்பமே ரொம்ப அமர்களமாயிருக்கு 😂😂😂
- படம் முன் பாதியைக் காட்டிலும் ,பின் பாதி காட்சிகள் வேக வேகமாக ஜம்ப் ஆவது சற்றே பலவீனம்

UJ_Lex_BiPolar_Sentiments lexicon consists of words and emojis expressing sentiments with tags positive/negative [Positive- 820 and negative- 1237].
For example,

- சூப்பர், பெரிய, சிறந்த, கேவலம், குறைவான, எதிரான

③ Methodology

In this work three approaches are experimented: Lexicon based, K-means with BoW and K-modes with BoW. Approaches are trained and tested using *UJ_Corpus_Opinions* corpus taking 70% and the remaining 30%.

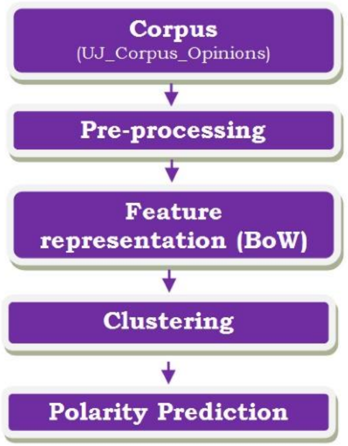


Algorithm 1 Lexicon based approach

Require: UJ_Lex_BiPolar_Sentiments lexicon(UJLexicon)
UJ_Corpus_Opinions (UJCorpus)
Ensure: Accuracy (Acc)

```
Step1: for each comment ∈ UJCorpus do
           words ← tokenised comment
Step2: Initialise variables: pos, neg for positive
       and negative polarity count with zero.
           pos ← 0
           neg ← 0
Step3: for each word ∈ words do
           if word in UJLexicon then
               if Polarity(word) is Positive:
                   pos ← pos+1
               else: neg ← neg+1
Step4: polarity ← positive if pos>neg else negative
```

Step5: Acc $\leftarrow \frac{\text{No. of correctly classified comments}}{\text{Total no. of comments in the corpus}} \times 100$



Algorithm 2 K-means with BoW approach

Require: UJ_Corpus_Opinions (UJCorpus)
Ensure: Accuracy (Acc)

```
Step1: Split UJCorpus into training and testing sets
Step2: for each comment ∈ UJCorpus do
           words ← tokenised comment
Step3: feature vector (BoW)
Step4: centroids ← K-means cluster (K=2) and
       feature vector (train)
Step5: for each vector ∈ feature vector (test) do
           distances ← euclidean(centroids, vector)
           polarity ← Label of the centroid with minimum
           euclidean distance
```

Step6: Acc $\leftarrow \frac{\text{No. of correctly classified comments}}{\text{Total no. of comments in the corpus}} \times 100$

K-modes with BoW approach: This approach is same as K-means with BoW approach but here instead of means mode is used.

④ Results

Tests results of the three approaches:

Approach	Accuracy
Lexicon based approach	57
K-means with BoW approach	61
K-modes with BoW approach	62

⑤ Discussion and Conclusion

Lexicon based Sentiment Analysis approach gives low accuracy (57%) compared with other two models due to the limited size of the lexicon. We are working on this to increase the accuracy of this approach by enhancing lexicon. K-modes with BoW based approach achieved highest accuracy of 62%.

In K-means and K-modes approaches one centroid is used to represent positive class and the other one is used to represent negative class. Each class contains the texts with different patterns thus, these two centroids in both models failed to capture the patterns of the two classes. Increased number of cluster centers K can be used to capture different patterns of text.

⑥ References

[1] E. Nivedhitha, S. P. Sanjay, M. Anandkumar, and K. P. Soman. Unsupervised word embedding based polarity detection for tamil tweets. International Journal of Computer Technology and Applications (IJCTA), 9(10):4631–4638, 2016.

[2] B. G. Patra, D. Das, A. Das, and R. Prasath. Shared task on sentiment analysis in indian languages (sail) tweets- an overview. In International Conference on Mining Intelligence and Knowledge Exploration, pages 650–655. Springer, 2015.

[3] S. Phani, S. Lahiri, and A. Biswas. Sentiment analysis of tweets in three indian languages. In Proceedings of the 6th Workshop on South and Southeast Asian Natural Language Processing (WSSANLP2016), pages 93–102, 2016.

SENTIMENT ANALYSIS ON TAMIL TEXTS USING K-MEANS AND K-NEAREST NEIGHBOR

Sajeetha Thavareesan, Sinnathamby Mahesan

sajeethas@esn.ac.lk, mahesans@univ.jfn.ac.lk

① Introduction

Sentiment Analysis is an application of Natural Language Processing which identifies and categorises the opinions into positive or negative.
In our model,

Bag of Words (BoW) and fastText vectors are used to represent features. These features are clustered using K-means clustering and the cluster centers are used to build the Sentiment Analysis model using K-Nearest Neighbour (K-NN).

BoW is used to represent the number of times a word appears in a document.
fastText treats each word as composed of character ngram. The vector for a word is made of the sum of the character ngram. Each word is represented using a 300 dimension vector.

④ Methodology

Three models are built using two types of feature vectors: *BoW* and *fastText*. *UJ_Corpus_Opinions* corpus is used to train and test these three models.

Model1: In this model K-NN is used as the classifier. K-NN is trained and tested on *UJ_Corpus_Opinions* corpus. Accuracy of this model is evaluated for different number of neighbours Kn in K-NN.

Model2: In this model feature vectors of training set are clustered using K-means with various number if clusters Km and the cluster centers are used to train K-NN.

Model3: Training set is split into groups based on class label, and these feature vectors of these groups are created and clustered these groups separatly using K-means clustering. We have tested this approach with different values of Kn and Km.

The general structure of *Model2* and *Model3* is described in Figure 1.

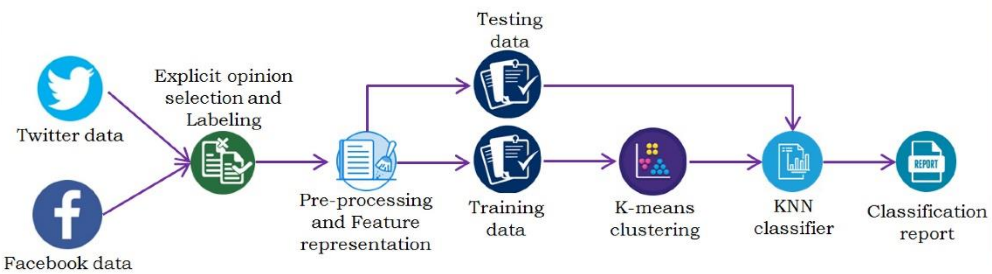


Figure 1: Structure of Model2 and Model3

② Problem Specification

The aim of this research is to build a suitable model with less number of training samples to perform Sentiment Analysis in Tamil text.

③ Contribution

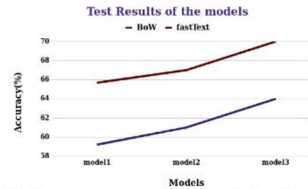
Constructed *UJ_Corpus_Opinions* corpus to tackle the inavailability of the opinion corpus, that contains 1518 positive and 1173 negative reviews and comments.
Proposed three models to perform Sentiment Analysis:

- *Model1: Sentiment classification using K-NN*
- *Model2: Sentiment classification using K-means clustering with K-NN*
- *Model3: Sentiment classification using class-wise K-means clustering with K-NN*

⑤ Results

Table1: Tests results of the models

Model	BoW			fastText		
	Accuracy	Kn	Km	Accuracy	Kn	Km
Model1	59	1	-	66	1	-
Model2	61	1	10	67	1	10
Model3	64	1	6	70	1	16



Test results of three models are listed in Table1. 70 as the highest accuracy is found for *Model3*.

⑥ Discussion and Conclusion

- We considered *Model1* as our base model and obtained 59% and 66% of accuracies for *BoW* and *fastText* feature vectors.
- We tested the models using different values of Km to check their influence in the accuracy and noticed that the accuracy increases with the values of Km.
- In *Model2* and *Model3* we used centroids as training set for K-NN and obtained better results compared with *Model1*. We obtained 61% and 67% of accuracies for *Model2* as we used centroids as training set of K-NN.
- *Model3* outperformed other two models as we used centroids of class-wice K-means clustering to train K-NN. It shows that class-wise clustering performs better than global clustering. Highest accuracy is found for *Model3* for both features *BoW* (64%) and *fastText* (70%) .
- High accuracy occurred with Kn=1 in K-NN for all three models.
- *fastText* as features gives better results than BoW for all Km.
- Furthermore, *Model3* outperformed the other models for all Km. Thus, *fastText* and class-wise clustering with increased number of clusters can be used to classify the sentiments expressed in the Tamil text.

⑦ References

[1] E. Nivedhitha, S. P. Sanjay, M. Anand Kumar, and K. P. Soman. Unsupervised word embedding based polarity detection for tamil tweets. International Journal of Computer Technology and Applications (IJCTA), 9(10):4631–4638, 2016.

[2] Shanta Phani, Shibamouli Lahiri, and Arindam Biswas. Sentiment analysis of tweets in three indian languages. In Proceedings of the 6th Workshop on South and Southeast Asian Natural Language Processing (WSSANLP2016), pages 93–102, 2016.

[3] Braja Gopal Patra, Dipankar Das, Amitava Das, and Rajendra Prasath. Shared task on sentiment analysis in indian languages (sail) tweets- an overview. In International Conference on Mining Intelligence and Knowledge Exploration, pages 650–655. Springer, 2015.

[4] N. Ravishankar and R. Shriram. Corpus based sentiment classification of tamil movie tweets using syntactic patterns. IIOAB Journal: A Journal of Multidisciplinary Science and Technology, 8(2):172–178, 2017.

A Robust Parallel Implementation of Active Contours

B. Saranya and S. Suthakar

Department of Computer Science, Faculty of Science, University of Jaffna

balansaranya99@gmail.com

Abstract

The main intention of this project is to speed up the performance of active contour process by parallelizing it. In traditional methods an algorithm namely 'Snake' has been applied for representing object contours. An energy minimizing technique is being used in this method. To speed up the convergence process I've parallelized the process by dividing the initial contour into sub-contours, converging them and combining them. As a result there will be a big change in the performance within a short period of time than in serial processing.

Keywords: Parallelize, contour, convergence

Introduction

- ❖ Active contours are computer-generated deformable curves which are being used for energy segmentation of objects or to locate object boundaries in the field of computer vision and image processing applications.
- ❖ These contours converge under the influence of energy minimizing technique.
- ❖ The energy is computed by minimizing a function of internal and external forces.
- ❖ The internal forces depend on the curve and the external forces are computed from the image.

Snake Energy = Internal energy + External energy

$$E_{snake} = \int_0^1 \left(E_{snake}(v(s))\right) ds$$

$$E_{snake} = \int_0^1 \left(E_{int}(v(s))\right) + \left(E_{img}(v(s))\right) + \left(E_{con}(v(s))\right) ds$$

- E_{int}: The internal elastic energy term.
  Continuity of the contour + smoothness of the contour.

$$E_{int} = {(\alpha(s)|v_s(s)|^2 + \beta(s)|v_{ss}(s)|^2)}/{2}$$

$\alpha(s)$ and β(s) are user defined weights.

vs(s) : First derivative term controlled by $\alpha(s)$.
$vss(s)$) : Second derivative term controlled by β(s).

- E_{img}: Combination of the forces due to the image itself.

$$E_{img} = W_{line}\, E_{line} + W_{edge}\, E_{edge} + W_{term}\, E_{term}$$

E_{line} : The intensity of the image.
$E_{edge} = -|\nabla I(x,y)|$, It is based on the image gradient.

$$E_{term} = \frac{\partial\theta}{\partial n_\perp} = \frac{{\partial^2 c}/{\partial n_\perp}^2}{{\partial c}/{\partial n}}$$

- E_{con}: Constraint forces introduced by the user.

Methodology

As a parallel processing solution for active contour model, rather than dividing the image into sub-images, we divide the initial contours into sub-contours in radial direction and process them independently. To define the sub-contours, the bounding box of the initial contour is used (see Fig.1.)

Then each sub-contour is given to different processing elements, and the contour points simultaneously move towards the edge of the sub-objects iteratively. Even though the divider lines are used to form closed contours, the points on the divider lines are kept unchanged while the sub-contour points are being shifted.

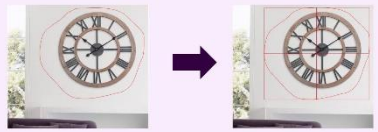


Fig. 1. Dividing initial contour into sub-contours in radial direction

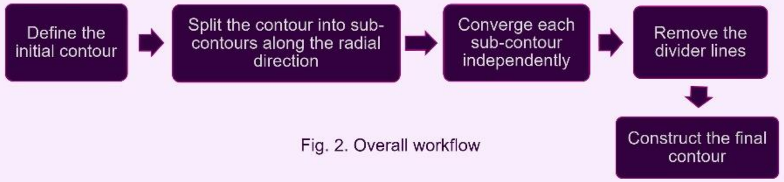


Fig. 2. Overall workflow

Merging Sub-Contours

When the sub-contour points are moving towards the boundary of the object simultaneously, the contour points on the divider lines are kept unchanged. Once the convergence is completed, the points on the opposite sides of each divider lines are connected to construct the final contour Fig. 3 illustrates this process.

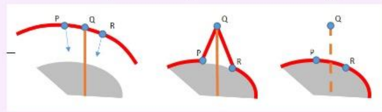


Fig. 3. Merging two sub-contours.

Sample Output

Some intermediate results from the start till the end of the proposed method are shown in Fig. 4.

Fig. 4. Intermediate results for an image.

Results and Discussion

For the evaluation, the proposed method was implemented with the help of Message Passing Interface (MPI) using two computers (Intel(R) Core(TM) i5-7200U, 8 GB RAM). Five different images of same dimension 3968x2976 were tested in serial and parallel environments. The execution times of serial and parallel implementation are given in Table 1. In order to compare the performances, same initial contours were used for both serial and parallel cases.

Table 1. Serial and parallel run times in seconds

	T_S	T_P	$S = T_S/T_P$
Image 1	4.523	3.112	1.453406
Image 2	3.425	2.314	1.480121
Image 3	5.281	3.998	1.32091
Image 4	4.910	3.428	1.432322
Image 5	3.911	2.855	1.369877

T_p - Parallel run time, T_s - Serial run time, S - Speed up

The proposed parallel method can be further improved by implementing using GPU.
In future there are possibilities to perform with concave objects.

Conclusion

- ➢ A parallel method is approached for implementing active contour model for making it more efficient.
- ➢ The given method segments in lesser time comparing to the serial active contour model.
- ➢ Using two PCs, the speed ups were roughly around 1.4. By using GPU, communication overheads can be reduced, and hence, the performance can be further improved

Reference

[1] P. L. Jerry and X. Chenyang, *Gradient Vector Flow Deformable Models, 2000.*
[2] A. Kass, A. Witkin and A. Terzopoulos, "Snakes: Active Contour Models," *International Journal of Computer Vision, 1988.*
[3] D. Cohen and Laurent, "On active contour models and balloons," in *CVGIP: Image Understanding, 1991.*
[4] S. Priya P Sajan, "GVF Snake Algorithm-a parallel approach," *International Journal of Engineering & Technology, 2018.*
[5] D. Oberhelman and Steven, "Active Contour Implementation," 2018.
[6] P. Jiang and Q. D. a. X. Hu, "A Parallel Realization of the Active Contour Model," *Applied Mathematics & Information Sciences, An International Journal, 2013.*

A Novel Approach of Voice Recognition Using MFCC and GMM, Speech Recognition and Text Recognition to Assist for Email Communication for Visually Impaired People

Senthuja Karunanithy

University of Jaffna

senthunithy@gmail.com

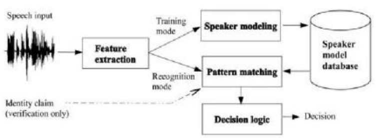
Abstract

Nowadays, Human-computer interaction plays a prominent role in the day to day life. However, it has become a challenging task for visually impaired people to get involved with computers in their day-to-day activities because of limited accessibility to the input mechanism. This work proposes speech-to-text, text-to-speech, and voice recognition techniques giving access to blind people to interact with Email communication. Voice recognition helps to recognize the voice of a specific person from the audio recording as voice is different from each other than the fingerprint where speech recognition helps to disregard the language and meaning to detect by the person behind the speech. The proposed model is based on the classification of MFCC coefficients obtained from speech signals with GMM for voice recognition. The proposed method is evaluated using VoxForge Dataset; containing the 340 voices of 34 speakers and obtained the result with 100% success.

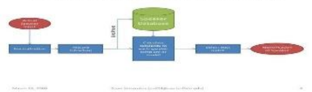
Keywords: Text Recognition, Speech Recognition, Voice Recognition, Mel Frequency Cepstral Coefficient (MFCC), Gaussian Mixture Modelling (GMM)

Introduction

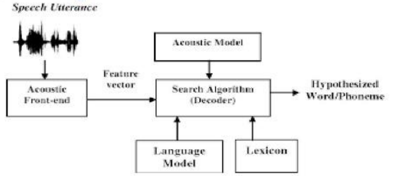
Speaker Recognition



Speaker Identification Diagram



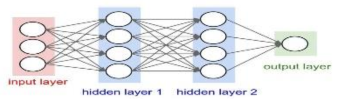
Speech Recognition



Objective

The objective of the research is to develop a voice-based email system that would help blind people to access email. The system will not let the user makes the use of the keyboard instead will work on speech recognition and voice recognition.

Background

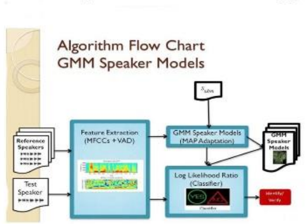


Methodology

The methodology can be summarized in six basic phases:

- Data Acquisition
- Data pre-processing
- Feature Extraction
- Model Training
- Perform Testing (identification)
- Application

Proposed Methodology



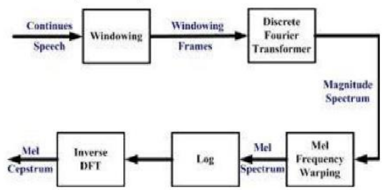
Data Acquisition

Methodology is assessed with the VoxForge DATASET

Data preprocessing

The data must be preprocessed in order to achieve better outputs and prediction results. This is to ensure that the model is trained with minimum errors. Vox-Forge dataset was already clean and noise free.

Feature Extraction: Mel frequency Cepstral coefficient (MFCC) estimation



Functions provided in python_speech_features module

```
pythonspeechfeatures.base.fbank(signal,
samplerate=16000, winlen=0.025, winstep=0.01,
nfilt=26, nfft=512, lowfreq=0, highfreq=None,
preemph=0.97,winfunc=<function <lambda>>)
```

Model Training: Front-end processing

The objective in the front-end processing is to modify the speech signal, so that it will be more suitable for feature extraction analysis. The front-end processing operation based on noise cancelling, framing, windowing and pre-emphasis.

Speaker modelling

The objective of modeling technique is to generate models for each speaker using specific feature vector extracted from each speaker. It performs a reduction of feature data by modeling the distributions of the feature vectors. The speaker recognition is also divided into two parts that means speaker dependent and speaker independent. In the speaker independent mode of the speech recognition the computer should ignore the speaker specific characteristics of the speech signal and extract the intended message .on the other hand in case of speaker dependent mode speech recognition machine should extract speaker characteristics in the acoustic signal.

Speaker database

The speaker models are stored here. These models are obtained for each speaker by using feature vector extracted from each speaker. These models are used for identification of unknown speaker during the testing phase.

Decision logic

It makes the final decision about the identity of the speaker by comparing unknown speaker to all models in the data base and selecting the best matching model.

Perform Testing

The log-likelihood for each GMM of every speaker was calculated in the model training phase. It was stored as a database in a separate folder. This data dictionary is used for matching 1:N speaker's file. The speaker with the highest score is chosen and identified.

Results

Vox-Forge: 34 speakers each accompanied seven voice samples for training data, and three voice samples for testing data totaling 340 cleaned and preprocessed voice sample Data. Each voice sample was around 4 sec. in length, and thus the default value of `nfft=512` in `mfcc()` worked fine.

Training corpus: It has been developed from audios taken from 'on-line Vox-Forge speech database' and consists of seven speech utterances for each speaker, spoken by 34 speakers (20-30 seconds/speaker).

Test corpus: This consists of remaining three unseen utterances of the same 34 speakers taken in train corpus. All audio files are of 10 seconds duration and are sampled at 16000 Hz. Thus, speaker identification was successfully conducted with an outstanding result on the dataset. The accuracy was 100% in case of VoxForge Dataset. MFCC- GMM model gives satisfactory results.

Why it is specific

It is unique, because here we do not consider the language of the speaker. Whatever the language spoken by the user is not the matter. This research mainly focus on the tone, frequency *etc*.

Discussion & Conclusion

MFCC algorithm is used in our system as it has the least false acceptance ratio. In order to improve system performance and also to achieve high accuracy GMM model can be used in the feature matching technique.

Reference

Alif Khan, Shah Khusro, Badam Nasi, Jamil Ahmad, Iftikhar Alam and Inayat Khan, *"Tetramail: a usable email client for blind people"*, Universal Access in the Information Society, September 2018.

Acknowledgement

Foremost, I would like to sincerely thank my supervisor Dr. S. Mahesan and all other lecturers who continually supported me during my research.

A Multiscale Contextual Technique for Fashion Clothes Landmark Localisation

Shajini Majuran and Amirthalingam Ramanan

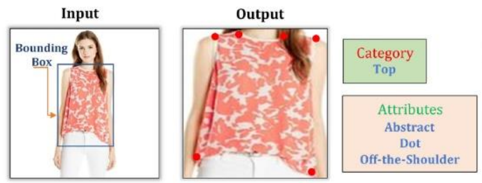
shayu.kiri@gmail.com, a.ramanan@univ.jfn.ac.lk

Introduction

Nowadays, researchers pay attention towards fashion clothes classification using machine learning and deep learning in order to make people's lives better with help of key factors such as image recognition, process of massive data, and enhance the facility of personalisation. In this regard, localising landmarks (collar, sleeves, waistline, and hem) in clothes can be an important attention in clothing classification. *Landmark localisation* involves global integration of information and the ability to retain local pixel-level details. It adjoins more challenges due to the various appearances, deformation, and occlusion of clothes. To conquer these complications, this study utilises multiscale contextual information without losing resolution of feature maps.

Applications: Automated fashion stylists, outfit recommendation, discovering similar fashion pieces, surveillance context, automatic annotation of images with tags or descriptions, context-aided people identification, occupation recognition, and improvement in information retrieval from social media.

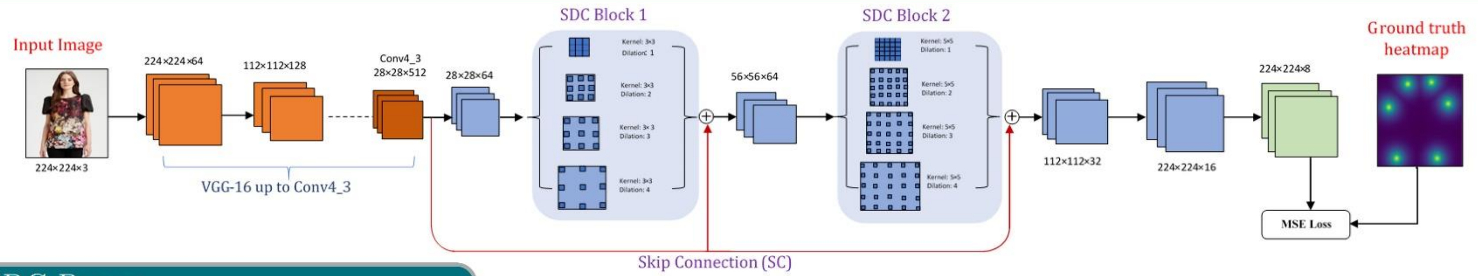
Problem Specification



Contribution

1. A scale-driven structure through deep neural network is proposed using *stacked dilated convolutional (SDC) block* which utilises parallel convolutions with different dilation rates.
2. A much simplified framework for better fashion clothing landmark detection is proposed with the help of *upsampling feature maps* which enhance the performance by transposing to high resolution feature maps.

Methodology



Dataset

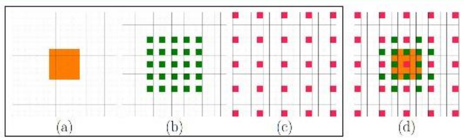
DeepFashion-C Dataset [1]

- # Images: 289,222
- # Categories: 50
- # Landmarks: 8
- # Attributes: 1000

Fashion Landmark Detection (FLD) Dataset [3]

- # Images: 123,016
- # Landmarks: 8

SDC Block



Stacked Dilated Convolution (SDC) Block: Dilated convolution is a convolution, applied to the input feature map with designated distance between the kernel points determined by the dilation coefficient $\mathcal{D}$. (a), (b), and (c) show the example operations using 5×5 kernel with $\mathcal{D}$=1, 2, and 4, respectively. (d) is the overall SDC operation on a feature map with receptive field size of 17×17.

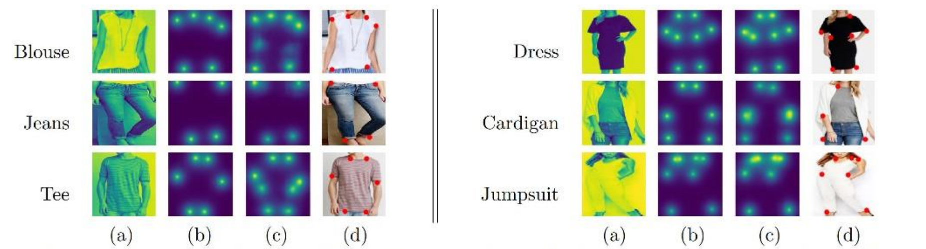


Figure: Examples shows in (a), (b), (c) and (d) are the images of normalised input image, ground truth landmark heatmap, output heatmap, and output landmark locations plotted on original input image for different clothes categories, respectively.

Experimental Setup

- The structure is based on *VGG-16* network.
- Each image is cropped using annotated bounding boxes and resized to 224×224.
- (Training,Validation,Testing)=(209222,40000,40000) images for DeepFashion-C, and (83033,19992,19991) images for FLD.
- *Optimiser*: Adam & *Mini-batch size*: 16
- *Learning rate*: 0.0001 & drop by 0.1 while validation plateaus.
- *Loss*: MSE loss
- *Evaluation*: Normalised distance metric

Results

Table I: Experimental reults on the *DeepFashion-C* dataset for landmark localisation in normalised distance metric

Methods	L.Collar	R.Collar	L.Sleeve	R.Sleeve	L.Waist	R.Waist	L.Hem	R.Hem	Avg.
Liu *et al.*[1]	0.0854	0.0902	0.0973	0.0935	0.0854	0.0845	0.0812	0.0823	0.0872
Liu *et al.*[3]	0.0628	0.0637	0.0658	0.0621	0.726	0.0702	0.0658	0.0663	0.0660
Yan *et al.*[5]	0.0570	0.0611	0.0672	0.0647	0.0703	0.0694	0.0624	0.0627	0.0643
Wang *et al.*[6]	0.0415	0.0404	0.0496	0.0449	0.0502	0.0523	0.0537	**0.0551**	0.0484
Lu *et al.*[7]	0.0332	0.0346	0.0487	0.0519	0.0422	0.0429	0.0620	0.0639	0.0474
Ours	**0.0323**	**0.0334**	**0.0443**	**0.0472**	**0.0368**	**0.0370**	**0.0533**	0.0558	**0.0425**

Table II: Experimental reults on the *FLD* dataset for landmark localisation in normalised distance metric

Methods	L.Collar	R.Collar	L.Sleeve	R.Sleeve	L.Waist	R.Waist	L.Hem	R.Hem	Avg.
Liu *et al.*[1]	0.0784	0.0803	0.0975	0.0923	0.0874	0.0821	0.0802	0.0893	0.0859
Liu *et al.*[3]	0.0480	0.0480	0.0910	0.0890	-	-	0.0710	0.0720	0.0680
Yan *et al.*[5]	0.0531	0.0547	**0.0705**	0.0735	0.0752	0.0748	**0.0693**	**0.0675**	0.0672
Ours	**0.0419**	**0.0424**	0.0733	**0.0718**	**0.0648**	**0.0652**	0.0738	0.0755	**0.0635**

Discussion and Conclusion

The output feature maps from each dilated operation of SDC block is concatenated together to make the subsequent convolution layer to learn features from different scales. By adding the SDC blocks, we can produce model which can lead to attain commendable regression results. Dilated convolutions show significant increase in performance for fashion landmark localisation. We demonstrate our experiments on two benchmark datasets and our model outperforms recently proposed state-of-the-art techniques.

References

[1] Z. Liu, P. Luo, S. Qiu, X. Wang, and X. Tang, "DeepFashion: Powering Robust Clothes Recognition and Retrieval with Rich Annotations", In proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1096-1104, 2016.

[2] E. Shelhamer, J. Long, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation", in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 4, pp. 640-651, 2017.

[3] Z. Liu, S. Yan, P. Luo, X. Wang, and X. Tang, "Fashion Landmark Detection in the Wild." In European Conference on Computer Vision (ECCV), pp. 229-245. 2016.

[4] H. Zhang, Q. Li, Z. Sun, and Y. Liu, "Combining Data-driven and Model-driven Methods for Robust Facial Landmark Detection", In IEEE Transactions on Information Forensics and Security, 13(10), pp.2409-2422, 2018.

[5] S. Yan, Z. Liu, P. Luo, S. Qiu, X. Wang, and X. Tang, "Unconstrained Fashion Landmark Detection via Hierarchical Recurrent Transformer Networks." In Proceedings of the 25th ACM International Conference on Multimedia, pp. 172-180, 2017.

[6] W. Wang, Y. Xu, J. Shen, and S. Zhu. "Attentive Fashion Grammar Network for Fashion Landmark Detection and Clothing Category Classification." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4271-4280, 2018.

[7] J. Liu, and H. Lu, "Deep Fashion Analysis with Feature Map Upsampling and Landmark-driven Attention." In Proceedings of the European Conference on Computer Vision (ECCV), pp. 30-36, 2018.

HEp-2 Specimen Classification Using Deep CNN

Shawmiya Yogaratnam and Siyamalan Manivannan
Department of Computer Science, Faculty of Science, University of Jaffna

INTRODUCTION

Indirect Immunofluorescence (IIF) on Human Epithelial-2 (HEp-2) cells is the most commonly used methodology to diagnose autoimmune diseases. The recognition of HEp-2 cell pattern in IIF images is one of the core challenges for antinuclear antibody (ANA) tests. Traditional approach requires experienced physicians to manually identify the cell patterns, which is extremely laborious and suffers from the inter-observer variability. Consequently, developing an automatic and reliable system for HEp-2 images processing tasks, e.g. cell and specimen image classification, becomes an attractive research topic. In this work a Deep Residual Network is used for classifying specimens into predefined classes. The effects of data augmentation via rotating specimen images was also investigated.

OBJECTIVE

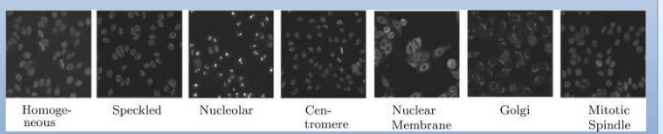
In cell image classification first individual cells must be extracted from the specimen images, and then a system is trained on the extracted cell images to predict the class of the new cell images. But in specimen image classification, such cell extraction is unnecessary, instead, a system can be directly trained on the specimen images to predict any unknown specimen image into one of the predefined classes.
The goal of this work is to investigate a pre-trained CNN architecture for specimen classification and to evaluate the role of data augmentation for network training.

METHODOLOGY

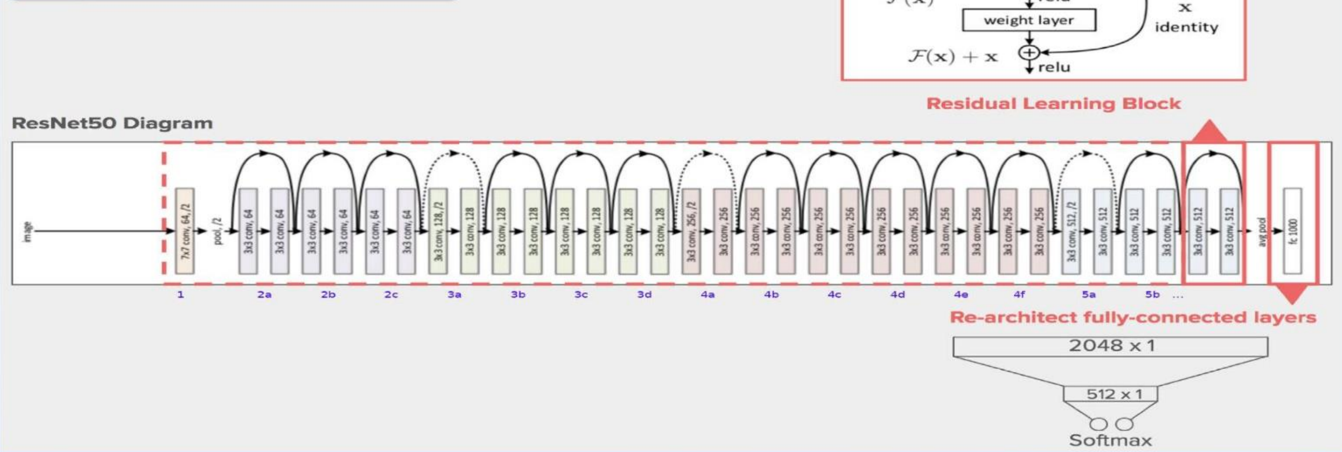
The ImageNet pre-trained Residual Network ResNet-50, was used. The last classification layer with 1000 nodes was replaced by a classification layer with 7 nodes.

DATASET

To evaluate, publicly available I3A-2014 dataset (1008 images) was used to train the model to classify HEp-2 specimen images into seven categories.



ResNet_50 Architecture



EXPERIMENTAL SETUP

80 percentage images of the dataset was used for training and the rest was used for testing. The Mean Class Accuracy (MCA) was used as the evaluation measure.
The model is trained with batch size of 200 and 50 epoches, and the learning rate is set to 0.001.

RESULT AND DISCUSSION

- By using ResNet-50 architecture, the testing accuracy achieved 86.1% for HEp-2 specimen image classification.

- Comparison of data augmentation strategy:
  To evaluate the effect of data augmentation, the proposed model was trained with data augmentation via random rotation.

Augmentation scheme	Result accuracy
Without augmentation	86.1%
Random rotate by 90°	87.4%

CONCLUSION

This study proposes an automatic classification model for HEp-2 specimen images by using ResNet architecture with transfer learning.

Experiments shows that data augmentation improves the classification accuracy.

REFERENCES

1. Yuexiang, L., Linlin, S. and Shiqi, Y., “HEp-2 Specimen Image Segmentation and Classification Using Very Deep Fully Convolutional Network”, in IEEE Transactions on Medical Imaging, pp. 1561 – 1572, 2017.
2. Yuexiang, L., Linlin, S., Xiande, Z. and shiqi, Y., “HEp-2 specimen classification with fully convolutional network”, in 2016 23rd International Conference on Pattern Recognition , 2016.
3. Hongwei, L., Wei-shi, Z. and Jianguo, Z., “Deep CNNs for HEp-2 Cells Classification: A Cross-specimen Analysis”, in research gate, 2018.

Speech Emotion Recognition Using Deep Learning on audio recordings

S.Suganya and **E.Y.A.Charles**
suganyasuven@gmail.com
charles.ey@univ.jfn.ac.lk

Abstract

Speech emotion recognition plays a prominent role in human-centred computing. However, It is still unclear that, which features of a human speech are robust enough to distinguish emotions. This work proposes an end-to-end deep learning approach which applies deep neural network on a raw audio recording directly. Proposed model was assessed on USC-IEMOCAP and EmoDB and obtained accuracy of 68.6% for IEMOCAP and 85.62% for EmoDB.

Objective

To propose a method that applies deep neural network to raw waves directly to perform emotion recognition.

Proposed Approach

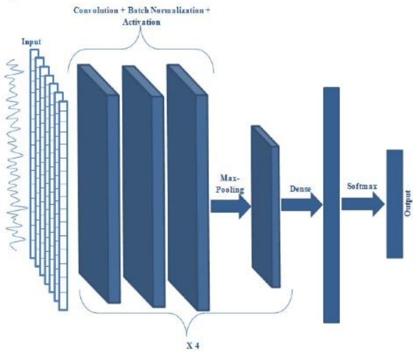


Figure 2 Flow diagram of Proposed method

- In most of the studies, MFCC features and spectrograms are used for the experiments. However it is still challenging to choose an optimal feature set for this task and time consuming procedure.
- Aims to develop a deep neural network which takes raw waveforms that represented as a long vector of values as input, instead of handcrafted features or spectrograms.
- Speech recordings in both IEMOCAP and EmoDB datasets were sampled at 16 kHz for this study.
- Proposed model contains seven convolutional layers, one fully connected layer and a softmax layer.

CNN model

M9
Input : 96000 x 1 time-domain waveform
[80/4, 64]
Max_pooling : 4 x 1 (output : 6000 x 64x n)
[3/1, 128] x 2
Max_pooling : 4 x 1 (output : 1500x 64x n)
[3/1, 256] x2
Max_pooling : 4 x 1 (output : 375x 128 x n)
[3/1, 512] x 2
Global average pooling (output : (1 x 512 x n)
FC(1024)
Softmax

([80/4, 64] denotes a convolutional layer with 64 filters and kernel size 80 with stride 4

Figure 2 Proposed model

Dataset

To evaluate our methodology, the Berlin Database of Emotional Speech (EmoDB) database [1] and the Interactive Emotional Motion Capture dataset (IEMOCAP) [2] published by the University of Southern California, are used to train and evaluate the proposed CNN model. The following figures illustrate the distribution of classes in the the datasets.

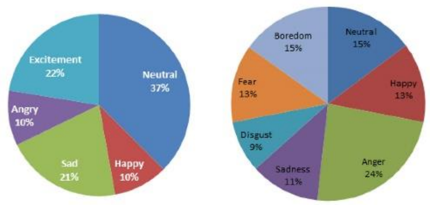


Figure 3 Distributions of datasets

Experiments

- For IEMOCAP dataset, experiments conducted for different sets of emotion classes,
  - [Anger, Happiness, Sadness, Neutral]
  - [Anger, Excitement, Sadness, Neutral]
  - [Anger, Sadness, Neutral]
- For EmoDB database, experiments conducted over 7 emotions.
  - [Anger, Happiness, Sadness, Neutral, Fear, Disgust and Boredom]

Experimental Results

- IEMOCAP
  - Emotion Classes
    - [Anger, Happiness, Sadness, Neutral] – 68.6%
    - [Anger, Excitement, Sadness, Neutral] – 64.3%
    - [Anger, Sadness, Neutral] – 79.3%
- EmoDB
  - [Anger, Happiness, Sadness, Neutral , Fear, boredom, Disgust] - 85.62%

Confusion Matrices

- For EmoDB

Class Labels	Neutral	Happiness	Sadness	Anger	Disgust	Fear	Boredom
Neutral	91.7	0	4.2	4.2	0	0	0
Happiness	4.8	47.6	0	42.9	0	4.8	0
Sadness	0	0	100	0	0	0	0
Anger	2.6	5.3	0	92.1	0	0	0
Disgust	0	7.1	0	7.1	78.6	0	7.1
Fear	5.3	5.3	0	0	0	89.5	0
Boredom	4.0	0	0	0	0	4.0	92.0

- For IEMOCAP
  - [Anger, Happiness, Sadness, Neutral]

Class Labels	Anger	Happiness	Neutral	Sadness
Anger	59.2	3.1	36.0	1.7
Happiness	11.2	14.3	69.2	5.2
Neutral	4.7	4.1	79.9	11.3
Sadness	1.8	1.6	20.2	76.4

- [Anger, Excitement, Sadness, Neutral]

Class Labels	Anger	Excitement	Neutral	Sadness
Anger	40.1	24.9	34.3	0.7
Excitement	11.2	45.8	39.8	3.2
Neutral	2.4	11.0	75.8	10.8
Sadness	0.9	1.6	22.8	74.7

- [Anger, Sadness, Neutral]

Class Labels	Anger	Neutral	Sadness
Anger	62.6	35.3	2.1
Neutral	3.7	82.2	14.1
Sadness	1.8	17.1	81.1

Analysis

- In IEMOCAP,
  - Neutral and sadness classes shows high true positive.
  - Happiness and anger are more classified as Neutral emotions.
  - According to the results, it can be observed that the correlation between anger, happiness and neutral are less compared to Happiness and Excitement emotion classes.
- In EmoDB
  - All emotions except happiness showed high class accuracy.
  - For emotion class sadness, the model achieved 100% accuracy and happiness was heavily confused with anger emotion.

Discussion & Conclusion

- Usage of Mel-scale spectrograms on a deep CNN and combinations of CNN and LSTM achieved a recognition rate in between 62 – 70% .
- In a recent study, phoneme features is combined with the spectrogram features achieved a accuracy of 73.9%.
- Thus, it can be noticed that achieved results are close to the accuracy of CNNs on spectrogram.
- Since the computation of spectrogram is costly and time consuming task, it can be concluded that the proposed approach is highly feasible for the emotion recognition task.

Reference

[1]. F. Burkhardt, A. Paeschke, M. Rolfes, W. Sendlmeier, and B. Weiss, "A database of German emotional speech," *INTERSPEECH*, pp.1517–1520, 2005.
[2]. C. Busso, M. Bulut, C.-C. Lee, A. Kazemzadeh, E. Mower, S. Kim, J. N. Chang, S. Lee, and S. Narayanan, "Iemocap: Interactive emotional dyadic motion capture database," *Language resources and evaluation*, vol. 42, no. 4, pp. 335–359, 2008.
[3]. G. Trigeorgis, F. Ringeval, R. Brueckner, E. Marchi, M. A. Nicolaou, S. Zafeiriou, and B. Schuller, "Adieu features? end-to-end speech emotion recognition using a deep convolutional recurrent network," *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5200–5204, 2015.
[4]. W. Dai, C. Dai, S. Qu, J. Li, and S. Das, "Very deep convolutional neural networks for raw waveforms," *2017 IEEE International Conference on Acoustics, Speech and Signal Processing(ICASSP)*, pp. 421–425, 2017.
[5]. M. B. Mustafa, A. M. Yusoof, Z. M. Don, and M. Malekzadeh, "Speech emotion recognition research: An analysis of research focus," *International Journal of Speech Technology*, vol. 21, pp. 137–156, 2018.

A Deep Learning Approach For Anomaly Detection in Data Communication Network

T.Thameera and K.Thabotharan
Department of Computer Science, University of Jaffna
thameera808@gmail.com,thabo@univ.jfn.ac.lk

Introduction

Cyber attack incidents are rapidly increasing with the use of internet. A Network Intrusion Detection System (NIDS) monitors network traffic searching for suspicious activity and known threats, sending up alerts when it finds such items. It can be categorized into anomaly detection and misuse detection.

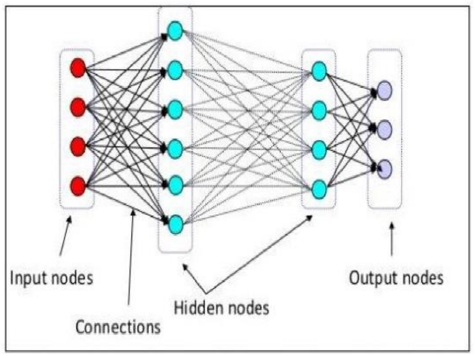
- Misuse detection uses the known attack patterns to identify attacks and shows high accuracy with less false alarm rates. However its performance suffers during the detection of new emerging threats due to the limitation of known attack patterns.
- Anomaly detection (ADNIDS) uses the deviation from normal patterns to identify intrusions. Although ADNIDS produces high false positive alarms it is theoretically potential in the identification of novel attacks.

Objectives

Recurrent Neural Network model is used in a wide range of applications such as Intrusion Detection System. Recurrent Neural Networks has become famous due to the excellence performance and recurrent layers, uses previous inputs to compute the next output. RNNs were developed to work with sequence prediction problems.

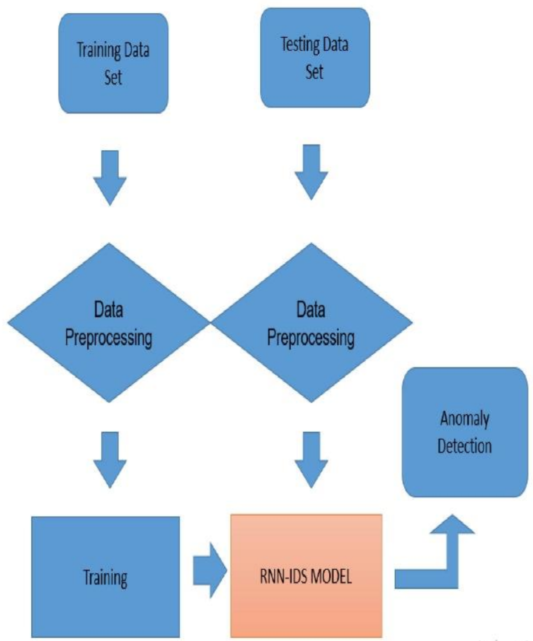
We use RNN for intrusion detection because network traffic is typically a sequential data. Our goal in this report is to improve the performance of Intrusion detection system using Recurrent Neural Network and reduce false positive alerts, also apply RNN with different variations such as LSTM, Simple RNN, and GRU.

Recurrent Neural Networks



Methodology

- Categorical data in the dataset is changed into model-understandable numerical data by label encoder.
- In our dataset protocol type, flags, services are categorical data. The problem here is, if we consider protocol type there are three protocol types(ftp, icmp, udp).so it will be numbered as 0,1 and 2 in any order. Since there are different numbers in the same column, the model will misunderstand that the data has some kind of order like 0<1. Hence hot encoding is performed.
- Preprocessed data is trained by RNN-IDS model.
- Using test data set accuracy is computed.



- Here NSLKDD and KDDcup have separate testing and training sets. Both data are preprocessed and trained by RNN-IDS model.

Testing Results

Experiments are performed on NSL-KDD and KDDCUP99 datasets. The testing results are presented in Table I, for NSLKDD dataset with binary classification. Here LSTM shows more accurate than Simple RNN and GRU shows more accuracy than the LSTM.

	SimpleRNN	LSTM	GRU
Classification rate	86.20%	81.64%	85.59%

The testing results are presented in Table 2, for KDDCUP dataset with binary classification. Here Simple RNN and GRU shows more accuracy than LSTM .

	SimpleRNN	LSTM	GRU
Classification rate	73.79%	75.85%	77.08%

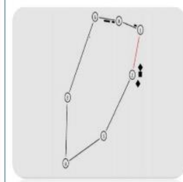
Conclusion and Discussion

In this work, different Deep Recurrent Neural Networks models are proposed to detect intrusions. The models have been implemented and tested on a benchmark dataset NSL-KDD.

- LSTM and GRU provides more accuracy than Simple RNN because they avoid vanishing gradient problem.
- Even Though ANN based model using KDD99 dataset gives high accuracy we can't consider it as a good measure because dataset has some redundancy problem.

In near future, we can configure the generated trace file in the simulation.

For the purpose of real data generation simulation set up for intrusion detection is needed. Here a simulation set up for malicious node in a wired network was carried out. The steps includes the following.



- Model a network.
- Configure the network according to research requirement.
- Add a malicious node in the network model.
- Generate trace file.

References

1. Chuanlong Yin , Yuefei Zhu, Jinlong Fei, and Xinzheng He, A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks,,pp- 2169-3536, November 7, 2017.
2. Ralf C. Staudemeyer, Applying long short-term memory recurrent neural networks to intrusion detection, SACJ No. 56, July 2015
3. P. Garcia-Teodoro, J. Diaz-Verdejo and E. Macia-Fernandez, G. and Vazquez. Anomaly-based network intrusion detection: Techniques, systems and challenges". Computers & security, vol. 28, no. 1-2, pp. 18{28, 2009. DOI http://dx.doi.org/10.1016/j.cose.2008.08.003.
4. Gozde Karatas , Onder Demir and Ozgur Koray Sahingoz,Deep Learning in Intrusion Detection Systems, DOI: 10.1109/IBIGDELFT.2018.8625278,December 2018.

Action Recognition in Videos Using Convolutional and Spatial-Temporal Interest Point Features

T. Tharmini and A. Ramanan
Department of Computer Science, Faculty of Science, University of Jaffna
tharmini7@gmail.com

Introduction

Recently human action recognition is an emerging topic of research in the field of computer. The factors such as occlusion of objects, camera parameters, scene clutter, illumination, body posture size and gender are increasing the complexity of action recognition. Recent trend in computer vision is the usage of convolutional neural network (CNN) which has been successful in image analysis like object recognition. On the other hand, action recognition using handcrafted features showed that space time interest point (STIP) performs better when compared to other local features. The proposed method in this study improves capturing spatial-temporal variation in human actions from videos using the STIP and convolutional features.

Objective

To improve the overall performance of action recognition task by using Convolution and STIP features.

Methodology

- Given a video, *fc7* features are computed for each video frame which is then mapped into a short binary code space using Iterative Quantization (ITQ) [4] , which is a method based on local sensitive hashing method (LSH).
- Key-frames are selected across the time space by picking the frames that their binary codes are different from their previous frames in a given video.
- A subset of key-frames (i.e., a snippet) is constructed using a fixed-sized window which is applied to the initial set of key-frames by striding the window with a constant factor. CNN flow is computed from the difference between the last key-frame and the first key-frame of a snippet.
- A vector representation is computed by stacking the CNN flows of each snippet.
- In the final step, all the videos are represented as Bag-of-Features (BoF) of temporal words.
- On the other hand, space-time interest points are searched in the video frames and feature descriptors are computed. These descriptors were also represented as BoF.
- Action-specific codebooks were constructed using K-means algorithm for BoF representation. The concatenated feature vector was then classified by a linear one-versus-all SVMs.

Methodology...

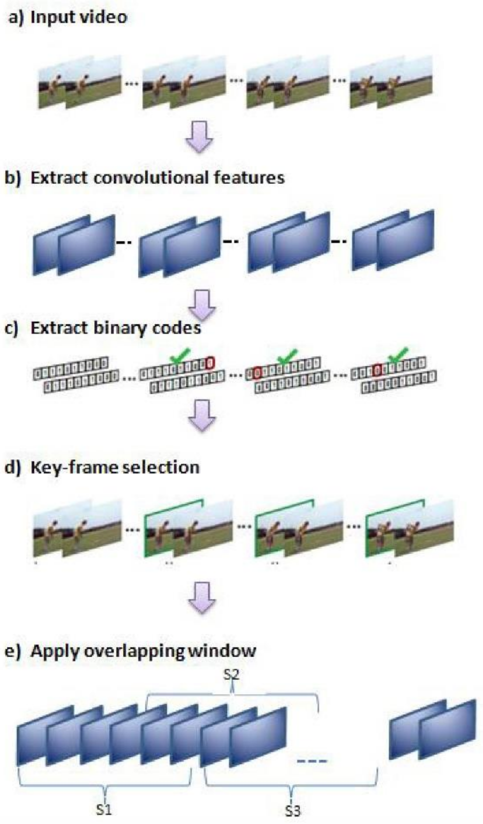


Figure 1: Snippet selection by selecting Key frames using LSH

Experimental setup

The proposed method was tested on KTH dataset [2]

- Testing included 9 subjects (2,3,5,6,7,10,22) and training set included rest of subject (16 subjects)
- Features: For CNN *fc7* features of VGG-F and for STIP HOG and HOF descriptors
- Codebook Construction: K-means algorithm
- Classifiers: Linear OVA-SVMs

Testing Result

Table 1: Comparison between action recognition rates of 24-bits key-frame selection with overlapping window size 8 (stride 4) with different K of K-means using convolutional features in classification

K–100	K–150	K–200
83.79%	**85.69%**	84.72%

Table 2: Comparison of STIP features in classification using action-specific and global codebook with different K of K-means

Action-specific Codebook		Global Codebook	
K–500	K–1000	K–500	K–1000
93.07%	**94.44%**	90.13%	91.67%

Finally, we combined STIP and CNN flow features for action classification by following the best parameter settings and type of codebook obtained in Table 1 and Table 2 as indicated in bold.
The combined feature set of STIP and CNN flow yields a classification rate of **94.91%** which is slightly better than the usage of an independent feature set.

Discussion

- Experiment reveals that how to recognize a video with feature representation. The overall performance of action classification has been improved when STIP and convolutional features are used together.
- This study evaluated the performance in two steps:

(i) Finding key-frames and applying overlapping windows to extract CNN flow features that captures sub-actions in a video sequence
(ii) Detecting local structures in space-time using STIP features.

- Each of the features was represented as bag-of-features.

Conclusion

Based on our experimental results, we conclude that the combined set of convolutional and STIP features perform better in action classification using the KTH dataset.

References

1. B. Banerjee and V. Murino, "Efficient pooling of image based CNN features for action recognition in videos," in Proceedings of IEEE International Conference on Accoust., Speech Signal Process. (ICASSP), pp. 2637–2641, Mar. 2017.
2. C. Schuldt, I. Laptev and B. Caputo, "Recognizing human actions: a local SVM approach in Pattern Recognition," in Proceedings of the 17th International Conference on Pattern Recognition, Aug 2004.
3. I. Laptev and T. Lindeberg, "Space-time interest points," in Proceeding on International Conference on Computer Vision, pp. 432–439, 2003.
4. Y. Gong and S. Lazebnik, "Iterative quantization: A procrustean approach to learning binary codes," in Proceedings of the International conference on CVPR, 2011.

An Efficient Approach for Patch-based Visual Object Classification

Veerapathirapillai Vinoharan and Amirthalingam Ramanan
Department of Computer Science, Faculty of Science, University of Jaffna
{vvinoharan,a.ramanan}@univ.jfn.ac.lk

SLSWCS Student Workshop on Computer Science

Introduction

In this work, we propose a two stage approach to optimize the process of building visual codebooks with discriminative power and compactness in the classification of visual objects.

Contribution

- A one-pass feature selection which is followed by an entropy-based feature selection approach is proposed to filter out ambiguous descriptors from initially extracted large descriptors set.
- Statistical-based measures and Visual-bit representation of codewords is proposed to select informative codewords from an initially constructed large codebook.

Methodology

- Unambiguous descriptors are selected from initially extracted SIFT descriptors using a one-pass feature selection (OPFS) method which is then followed by an entropy-based feature selection (EBFS) method to increase the discriminative power of the codebook.
- A codebook is then constructed by means of K-means approach.
- Indistinctive codewords are eliminated based on statistical measures (inter, intra, and combined category confidence) [1] or visual-bit representation of codeword to obtain a compact codebook [2].
- A histogram representation is created for each descriptor set of images and linear SVM classification algorithm is applied to those fixed-length feature vectors.

The overall framework of the proposed method is illustrated in Figure 1.

Methodology...

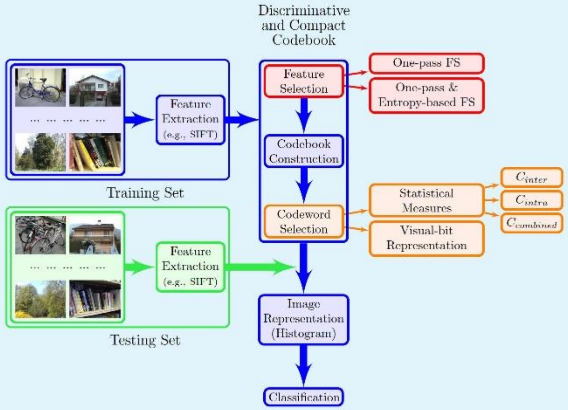


Figure 1: Overview of the proposed framework

⊙ **One Pass Feature Selection (OPFS)**

⋆ **Input:** $trainingFeatures$
⋆ **Output:** $selectedFeatures$
⋆ $r \leftarrow$ radius of the hypersphere,
⋆ $selectedFeatures \leftarrow trainingFeatures\{1\}$,
⋆ **for all** $feature \in trainingFeatures$ **do**
  if $min\| feature - trainingFeatures \|^2 > r^2$ **then**
   Create a new hypersphere of r such that,
   $selectedFeatures \leftarrow \{selectedFeatures \cup feature\}$
  end if
 end for

⊙ **Entropy-based feature selection (EBFS)**

$$E(F) = -\sum_{i=0}^{255} p_i(F) log_2 p_i(F) \quad (1)$$

where, $p_i(F)$ - $\frac{|\{k|f_k=i\}|}{128}$, descriptors are treated as 128 samples of discrete random variable k in {0, 1, 2, $\cdots$, 255}.

⊙ **Statistical Measures**

$$C_{inter,i} = \sum_{j-1}^{N} max\left(\frac{f_{ij}}{n_i} - \frac{1}{m_i}, 0\right) \quad (2)$$

$$C_{intra,i} = \frac{1}{\sum_{j-1}^{N} var(h_{ij})} \quad (3)$$

$$C_{combined,i} = \alpha C_{inter,i} + \beta C_{intra,i};\ 0 \leq \alpha, \beta \leq 1 \quad (4)$$

where,
f_{ij} - number of training features in the i^{th} codeword & j^{th} category.
h_{ij} - i^{th} codeword value of each image belonging to the j^{th} category in the BoF histogram domain, i = 1, 2, ..., K, & j = 1, 2, ..., N.
n_i - is the total number of features in the i^{th} codeword.
m_i - is the number of object categories in the i^{th} codeword.
N - is the number of object categories in classification.
K - is the size of the codebook.

⊙ **Visual-bit representation of codewords**

$$h_i = \begin{cases} 1 : \text{if } C_i \geq t_0 \\ 0 : \text{otherwise} \end{cases} \forall i = 1, \ldots, K \quad (5)$$

$$t_1 = \frac{\lambda p_0 + p_1}{\lambda + 1} \quad (6)$$

$$Compact_{CB} = \begin{cases} \text{eleminate } C_i & : \text{if } C_i \geq t_0 \\ \text{retain } C_i & : \text{Otherwise} \end{cases} \quad (7)$$

where,
p_0 is $min_{1\leq i\leq K}(SB_i)$, p_1 is $max_{1\leq i\leq K}(SB_i)$
SB_i - sum of visual bits associated with the i^{th} codeword.
λ - weighting parameter for a rare informative word.
t_i - level of significant activation of a codeword in a codebook.

Test Results

Comparison of average precision (AP) with number of training features and codebook size: Traditional BoF approach and proposed feature selection method *with* and *without* codeword selection (CS)

Approach	Dataset	#Descriptors	Without CS		Statistical Measures with CS: inter		intra		combined		Visual bit with CS	
			CB	AP	CB	AP	CB	AP	CB	AP	CB	AP
Traditional	Xerox7	4,046,578	987	84.21	803	83.68	740	87.89	902	82.41	286	83.85
OPFS		212,294	500	94.11	400	93.31	375	94.69	409	93.72	191	93.42
OPFS+EBFS		172,006	500	94.04	400	93.40	**375**	**94.79**	406	93.41	201	94.13
Traditional	UIUCTex	4,543,590	1032	82.73	835	81.94	774	86.40	842	81.53	387	90.25
OPFS		314,724	500	93.73	400	94.56	**375**	**95.51**	401	94.27	264	92.45
OPFS+EBFS		157,094	500	94.17	400	92.95	374	94.08	404	92.88	257	93.48
Traditional	PASCAL VOC 2007	1,760,400	1049	71.78	847	72.41	787	73.71	953	71.99	421	71.69
OPFS		245,327	500	72.93	400	73.16	375	73.47	**405**	**73.91**	262	72.88
OPFS+EBFS		181,248	500	72.58	400	72.90	375	73.64	414	72.71	252	72.20
Traditional	Caltech101	5,659,137	925	84.72	742	82.87	694	84.80	850	82.30	336	84.32
OPFS		393,024	500	86.01	400	85.17	375	85.97	408	85.83	289	85.48
OPFS+EBFS		286,925	500	86.02	400	85.36	**375**	**86.34**	407	85.50	249	85.35

Experimental setup

- For the image sets: Xerox7, UIUCTex, and Caltech101 we used 70% for training and 30% for testing from each class.
- For PASCAL VOC 2007, the training was performed on the provided 'trainval' set and evaluated on the testing set.
- We used SIFT descriptors in extracting the features from those image sets.
- The codebook is constructed by using the K-means algorithm with K = 500 for all datasets.
- The OVA-SVMs with RBF kernel was used for classification and the reported classification rates are of average precision (AP) [3].

Discussion and Conclusion

- The proposed ideas in this paper are to generate a compact and discriminative codebook, that can be obtained by selecting representative keypoints and eliminating indistinctive codewords.
- These processes not only reduces the overall computational complexity but also maintains the BoF model to be efficient with stable performance.
- As a near future work we will incorporate another set of detector-descriptors: SURF and ORB.

References

[1] V. Vinoharan and A. Ramanan, "Keypoints and Codewords Selection for Efficient Bag-of-Features Representation", In IEEE International Conference on Future of Information and Communication Conference (FICC), pp. 203-208, 2018.

[2] T. Kirishanthy and A. Ramanan, "Creating Compact and Discriminative Visual Vocabularies Using Visual Bits", In Proceedings of the IEEE Digital Image Computing: Techniques and Applications (DICTA), pp. 258–263, 2015.

[3] K. H. Brodersen, C. S. Ong, K. E. Stephan, and J. M. Buhmann, "The Binormal Assumption on Precision-Recall Curves", In Proceedings of the International Conference on Pattern Recognition (ICPR), pp. 4263–4266, 2010.

Organising Committee

General Chair	Dr. A. Ramanan
Programme Chairs	Dr. S. Mahesan & Dr. E. Y. A. Charles
Finance Chair	Dr. (Mrs.) B. Mayurathan
Session Chairs	Dr. K. Thabotharan, Dr. M. Siyamalan, & Ms. J. Samantha Tharani
Editor	Mr. S. Suthakar
Publicity Chair	Mr. V. Visithan
Web Chair	Mr. N. Thileepan

Achievers of SL-SWCS

SL-SWCS'11	
Best Presentation	Ms. Yawwani Gunawardana University of Colombo School of Computing (UCSC)
Best Poster	Mr. S. Sivasuthan University of Jaffna
SL-SWCS'13	
Best Presentation	Mr. K. Kanarupan University of Moratuwa
Best Poster	Mr. N. Nilashan, Mr. S. Thenuzan, Mr. S. Mayooran and Ms. D. Lukshica, University of Jaffna
SL-SWCS'15	
Best Presentation	Mr. T. Kokul University of Peradeniya
Best Poster	Ms. S. Charini, Ms. K. Anusha and Ms. W.M.B.K. Weerasooriya University of Peradeniya
SL-SWCS'17	
Best Presentation	Mrs. S. Majuran University of Jaffna
Best Poster	Mr. R. Miller University of Jaffna

"Computers are good at following instructions,
but not at reading your mind"

- Donald E. Knuth

SLSWCS
Student Workshop on Computer Science

www.slswcs.jfn.ac.lk

Department of Computer Science
University of Jaffna