





David C. Wyld  
Dhinaharan Nagamalai (Eds)

## **Computer Science & Information Technology**

International Conference on Soft Computing, Artificial Intelligence and Machine Learning (SAIM 2020)  
June 27~28, 2020, Copenhagen, Denmark



**AIRCC Publishing Corporation**

## **Volume Editors**

David C. Wyld,  
Southeastern Louisiana University, USA  
E-mail: David.Wyld@selu.edu

Dhinaharan Nagamalai,  
Wireilla Net Solutions, Australia  
E-mail: dhinthia@yahoo.com

ISSN: 2231 - 5403  
ISBN: 978-1-925953-22-0  
DOI: 10.5121/csit.2020.100801- 10.5121/csit.2020.100807

This work is subject to copyright. All rights are reserved, whether whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the International Copyright Law and permission for use must always be obtained from Academy & Industry Research Collaboration Center. Violations are liable to prosecution under the International Copyright Law.

Typesetting: Camera-ready by author, data conversion by NnN Net Solutions Private Ltd., Chennai, India

## Preface

The International Conference on Soft Computing, Artificial Intelligence and Machine Learning (SAIM 2020) June 27~28, 2020, Copenhagen, Denmark, 8<sup>th</sup> International Conference on Signal Image Processing and Multimedia (SIPM 2020), 8<sup>th</sup> International Conference of Advanced Computer Science & Information Technology (ACSIT 2020), 8<sup>th</sup> International Conference on Foundations of Computer Science & Technology (FCST 2020), 8<sup>th</sup> International Conference on Information Technology in Education (ICITE 2020) was collocated with International Conference on Soft Computing, Artificial Intelligence and Machine Learning (SAIM 2020). The conferences attracted many local and international delegates, presenting a balanced mixture of intellect from the East and from the West.

The goal of this conference series is to bring together researchers and practitioners from academia and industry to focus on understanding computer science and information technology and to establish new collaborations in these areas. Authors are invited to contribute to the conference by submitting articles that illustrate research results, projects, survey work and industrial experiences describing significant advances in all areas of computer science and information technology.

The SAIM 2020, SIPM 2020, ACSIT 2020, FCST 2020 and ICITE 2020 Committees rigorously invited submissions for many months from researchers, scientists, engineers, students and practitioners related to the relevant themes and tracks of the workshop. This effort guaranteed submissions from an unparalleled number of internationally recognized top-level researchers. All the submissions underwent a strenuous peer review process which comprised expert reviewers. These reviewers were selected from a talented pool of Technical Committee members and external reviewers on the basis of their expertise. The papers were then reviewed based on their contributions, technical content, originality and clarity. The entire process, which includes the submission, review and acceptance processes, was done electronically.

In closing, SAIM 2020, SIPM 2020, ACSIT 2020, FCST 2020 and ICITE 2020 brought together researchers, scientists, engineers, students and practitioners to exchange and share their experiences, new ideas and research results in all aspects of the main workshop themes and tracks, and to discuss the practical challenges encountered and the solutions adopted. The book is organized as a collection of papers from the SAIM 2020, SIPM 2020, ACSIT 2020, FCST 2020 and ICITE 2020.

We would like to thank the General and Program Chairs, organization staff, the members of the Technical Program Committees and external reviewers for their excellent and tireless work. We sincerely wish that all attendees benefited scientifically from the conference and wish them every success in their research. It is the humble wish of the conference organizers that the professional dialogue among the researchers, scientists, engineers, students and educators continues beyond the event and that the friendships and collaborations forged will linger and prosper for many years to come.

David C. Wyld  
Dhinaharan Nagamalai (Eds)

## General Chair

David C. Wyld,  
Dhinaharan Nagamalai,

## Organization

Southeastern Louisiana University, USA  
Wireilla Net Solutions, Australia

## Program Committee Members

Abdel-Badeeh M. Salem,  
Abdurrahman Celebi,  
Ahmed Abdelgawad,  
Ahmed Farouk AbdelGawad,  
Ajune Wanis Ismail,  
Alessio Ishizaka,  
Alessio Ishizaka,  
Alexandru Strunga,  
Ali El-Zaart,  
Alicia Trivino,  
Almir Pereira Guimaries,  
Anand Nayyar,  
Anandi Giridharan,  
Anas M.R. AlSobeh,  
Apai,  
ARIDJ Mohamed,  
Ashkan Tashk,  
Asimi Ahmed,  
Athanasios V. Vasilakos,  
Azah Kamilah Muda,  
Azian Azamimi Abdullah,  
Azizollah Babakhani,  
Barbara Pekala,  
Barbaros Preveze,  
Basma BOUKENZE,  
Belobo Belobo Didier,  
Bichitra Kalita,  
Bilal Abu Salih,  
Bilal H. Abed-alguni,  
Bin Cao,  
Biruta SvagZdiene,  
Boukari nassim ,  
BRAHAMI Menaouer,  
Chandan Kumar Karmakar,  
Chen-Chia Chuang,  
Chin-chen Chang,  
Ching-Nung Yang,  
Christian Mancas,  
Christophe Gueyex,  
Chuanzong Zhang,  
Ain Shams University, Egypt  
Bedër University, Albania  
Central Michigan University, USA  
Zagazig University, Egypt  
Universiti Teknologi Malaysia, Malaysia  
Neoma Business School, France  
University of Portsmouth, England  
University of Craiova, Romania  
Beirut Arab University, Lebanon  
University of Malaga, Spain  
Federal University of Alagoas, Brazil  
Duy Tan University, Vietnam  
Indian Institute of Science, India  
Yarmouk University, Jordan  
Universiti Malaysia Perlis, Malaysia  
Hassiba benbouli University, Algeria  
University of Southern Denmark, Denmark  
Ibn Zohr University, Morocco  
Lulea University of Technology, Sweden  
Universiti Teknikal Malaysia Melaka, Malaysia  
Universiti Malaysia Perlis, Malaysia  
Babo Noshirvani University of Tecnology, Iran  
University of Rzeszow, Poland  
Cankaya University, Turkey  
Hassan first University, Morocco  
African Centre for Advanced Studies, Cameroon  
Assam Don Bosco University, India  
University of Jordan, Jordan  
Yarmouk University, Jordan  
Hebei University of Technology, P.R. China  
Lithuanian Sports University, Lithuania  
skikda university, algeria  
National Polytechnic School of Oran, Algeria  
University of Melbourne, Australia  
National Ilan University, Taiwan  
Feng Chia University, Taiwan  
National Dong Hwa University, Taiwan  
Ovidius University, Romania  
University of Bourgogne Franche-Comté, France  
Aalborg University, Denmark

Dallel Sarnou,	Abdelhamid Ibn Badis University,Algeria
Daniela Momete,	University Politehnica Of Bucharest, Romania
Deepak Garg,	Bennett University, India
Desmond Bala,	Cranfield University, United Kingdom
Dilana Hazer-Rau,	University of Ulm, Germany
E.M.Badr,	Benha University, Egypt
Eng Islam Atef,	Alexandria universty, Egypt
Farouq Saber Al-Shibli,	Philadelphia University, Jordan
Fatih Korkmaz,	Cankiri Karatekin University, Turkey
Federico Tramarin,	University of Padova, Italy
Felix Yang Lou,	City University of Hong Kong, China
Franco Frattolillo,	University of Sannio,italy
Gitesh K. Raikundalia,	Victoria University , Australia
Grigorios N. Beligiannis,	University of Patras, Greece
Guezouli Larbi,	University Of Batna 2, Algeria
Guilong Liu,	Beijing Language and Culture University, China
Hamid Ali Abed AL-Asadi,	Basra University, Iraq
Hassan Ugail,	University of Bradford, UK
Henrique Joao Lopes Domingos,	New University of Lisbon, Portugal
HlaingHtakeKhaung Tin,	University of Computer Studies, Myanmar
Ing. habil,	University of Sibiu, Romania
Isaac Agudo,	University of Malaga, Spain
Jafar A. Alzubi,	Al-Balqa Applied University, Jordan
Jafar Mansouri,	Ferdowsi University of Mashhad, Iran
Jamal El Abbadi,	Mohammadia V University Rabat, Morocco
Jasmin Cosic,	DB Netze AG, Germany
Jun Zhang,	South China University of Technology, China
Kamel Hussein Rahouma,	Minia University, Egypt
Karthikeyan,	Mangayarkarasi College Of Engineering, India
Ke-Lin Du,	Concordia University, Canada
Klenilmar Lopes Dias,	Federal Institute of Education, Brazil
Klimis Ntalianis,	Athens University of Applied Sciences, Greece
LABRAOUI Nabila,	University of Tlemcen, Algeria
Liliana Mața ,	Vasile Alecsandri University of Bacau,Romania
Lim Chen Kim,	University Pendidikan Sultan Idris (UPSI), Malaysia
Longzhi Yang,	Northumbria University, UK
Maissa Hamouda,	University of Sousse, Tunisia
Malka N. Halgamuge,	The University of Melbourne, Australia
Maoh-Chin Jiang,	National Ilan University, Taiwan
Maria Hallo,	Escuela Politecnica Nacional, Ecuador
Mario Henrique Souza Pardo,	University of Sao Paulo, Brazil
Marisnel C. Olivares A,	Naval Academy Research Institute, France
Marius CIOCA,	Lucian Blaga University of Sibiu, Romania
Maryam hajakbari,	Islamic Azad University, Iran
Maumita Bhattacharya,	Charles Sturt University, Australia
mohamed anis mastouri,	Syscom lab. enit, Tunisia
Mohamed Elhoseny,	Mansoura University, Egypt
Mohamed HAMLICH,	Hassan II University, Morocco
Mohammed Elbes,	Al-Zaytoonah University, Jordan
Mueen Uddin,	Effat University Jeddah,Saudi Arabia
Muhammad Arif,	Guangzhou University, China
Mussab Alaa,	University of Malaya, Malaysia

Nadia Abd-alsabour ,	Cairo University, Egypt
Noura Taleb,	Badji Mokhtar University, Algeria
O. S. Albahri,	Universiti Pendidikan Sultan Idris, Malaysia
Okba KAZAR,	University of Biskra, Algeria
Olfa Belkahla Driss,	University of Manouba, Tunisia
Ouafa Mah,	Ouargla university, Algeria
Pavel Loskot,	Swansea University, United Kingdom
Pedro Donadio,	Federal University of Amazonas, Brazil
Roberto De Virgilio,	Roma Tre University, Italy
Ruksar Fatima,	KBN College of Engineering, India
Saad M. Darwish,	Alexandria University, Egypt
Saadat Pourmozafari,	Tehran Poly Technique, Iran
Sabina Rossi,	University Ca' Foscari, Venice, Italy
Sabu Mes,	M.E.S College, India
Santhi Balaji,	Bangalore University, India
Savita Wali,	Basaveshwar Engineering College, Bagalkot
Shahid Siddiqui,	Integral University, India
Shaiful Islam Babu,	City University, Bangladesh
Shashank Sharma,	Manipal University, India
Shirish Patil,	Independent/Industry, USA
Shoeib Faraj,	Institute of Higher Education of Miaad, Iran
Smain FEMMAM ,	UHA University, France
Souad Taleb,	Oran 2 university, Algeria
Soubhik Chakraborty,	Birla Institute of Technology, India
Stefania Tomasiello,	University of Salerno, Italy
Subarna Shakya,	Tribhuvan University , Nepal
Suman Deb,	NIT Agartala, India
Svetoslava Saeva,	Neofit Rilski South-West University, Bulgaria
Tapas Si,	Bankura Unnayani Institute of Engineering, India
Tran Cong Manh,	Le Quy Don Technical University, Vietnam
Vinod Pangracious,	Sorbonne University Paris, France
Vishal Shrivastava,	Arya Group of Colleges, India
Walid Barhoumi,	University of Carthage, Tunisia
Wee kuok kwee,	Multimedia University, Malaysia
Wenwu Wang,	University of Surrey, UK
Willie K Ofosu,	Penn State Wilkes-Barre, USA
Yan Li,	University of Southern Queensland, Australia
Yannick Le Moullec,	Aalborg University, Denmark
Yousef FARHAOUI,	Moulay Ismail University, Morocco
Yu-Chen Hu,	Providence University, Taiwan
Yung Gi Wu,	Chang Jung Christian University, Taiwan
Yu-Sheng Lu,	National Taiwan Normal University, Taiwan
Zayar Aung,	National Research University , Russia
Zeshui Xu,	Sichuan University, China
Zhe Chen,	Dalian University of Technology, China
Zoran Bojkovic,	University of Belgrade, Serbia

## Technically Sponsored by

Computer Science & Information Technology Community (CSITC)



Artificial Intelligence Community (AIC)



Soft Computing Community (SCC)



Digital Signal & Image Processing Community (DSIPC)



## Organized By



Academy & Industry Research Collaboration Center (AIRCC)

## TABLE OF CONTENTS

### **International Conference on Soft Computing, Artificial Intelligence and Machine Learning (SAIM 2020)**

**Multi-label Classifier Performance Evaluation with Confusion Matrix.....01 - 14**  
*Damir Krstinić, Maja Braović, Ljiljana Šerić and Dunja Božić-Štulić*

**Developing E-learning System to Support Math in Thai Education Curriculum (Primary Level) .....15 - 21**  
*Piya Techateerawat*

### **8<sup>th</sup> International Conference on Signal Image Processing and Multimedia (SIPM 2020)**

**Simple Face Thermal Features for Gender Discrimination.....23 - 31**  
*Georgia Koukiou and Vassilis Anastassopoulos*

### **8<sup>th</sup> International Conference of Advanced Computer Science & Information Technology (ACSIT 2020)**

**Biomedical Study of Demographics and Clinical Features of Lichen Planopilaris among the Iranian Population ..... 33 - 39**  
*Maryamsadat Nejadghaderi, Ashkan Tashk, Parvin Mansouri and Zahra Safaei Naraghi*

**Tibetan and Chinese Text Image Classification based on Convolutional Neural Network ..... 59 - 67**  
*Jincheng Li, Penghai Zhao, Yusheng Hao, Qiang Lin, Weilan Wang*

### **8<sup>th</sup> International Conference on Foundations of Computer Science & Technology (FCST 2020)**

**Two Evolutionary Hybrid Stages for the Rectangular Bin Packing Problem with Constraints ..... 41 - 49**  
*Slimane Abou-Msabah, Ahmed Riadh Baba-Ali and Basma Sagr*

### **8<sup>th</sup> International Conference on Information Technology in Education (ICITE 2020)**

**Effective Teaching Strategies in an Online Environment..... 51 - 57**  
*Keith Buckley*

# MULTI-LABEL CLASSIFIER PERFORMANCE EVALUATION WITH CONFUSION MATRIX

Damir Krstinić, Maja Braović, Ljiljana Šerić and Dunja Božić-Štulić

Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture,  
University of Split, R. Boškovića 32, Split 21000, Croatia

## **ABSTRACT**

*Confusion matrix is a useful and comprehensive presentation of the classifier performance. It is commonly used in the evaluation of multi-class, single-label classification models, where each data instance can belong to just one class at any given point in time. However, the real world is rarely unambiguous and hard classification of data instance to a single class, i.e. defining its properties with single distinctive feature, is not always possible. For example, an image can contain multiple objects and regions which makes multi-class classification inappropriate to describe its content. Proposed solutions to this set of problems are based on multi-label classification model where each data instance is assigned one or more labels describing its features. While most of the evaluation measures used to evaluate single-label classifier can be adapted to a multi-label classification model, presentation and evaluation of the obtained results using standard confusion matrices cannot be expanded to this case.*

*In this paper we propose a novel method for the computation of a confusion matrix for multi-label classification. The proposed algorithm overcomes the limitations of the existing approaches in modeling relations between the classifier output and the Ground Truth (i.e. hand-labeled) classification, and due to its versatility can be used in many different research fields.*

## **KEYWORDS**

*Classification, multi label classifier, performance evaluation, confusion matrix*

## **1. INTRODUCTION**

Multi-class classification (MCC), where each data instance or object is assigned to a class from the set of a priori known classes, is widely encountered in scientific literature and engineering applications. Regardless of the number of possible classes, they are mutually exclusive and each object can be assigned to only one class. This approach in machine learning, also known as Single-label classification [18], relies on fundamental assumption that each data object belongs to only one concept and has a unique semantic meaning [21]. While this approach is well-known and widely used in supervised learning, there are data sets that are too complex to impose the restriction of only one label for each data instance [5], [12]. These problems include text categorization, sentiment and emotion recognition, semantic scene classification and many other problems. In fact, most of real life data is often difficult to describe with a single distinctive feature. Objects, phenomena, relations, interactions and other manifestations of natural and artificial processes are seldom simple enough to be easily defined and unambiguously classified. Machine learning approaches dealing with data that

cannot be simply classified to a single distinctive class, i.e. described with a unique semantic context, are referred to as Multi-label classification (MLC) [18], [17] or Multi-label learning [5], [19] gain the attention and becomes a relevant research area [1], [21].

Beside assigning more than one semantic concept to each data instance, MLC algorithms have to deal with other difficulties, e.g. correlation between different labels and an uneven number of label occurrences on the data. Representative example is image classification. Aside from rare exceptions, most of the real life images contain multiple objects and regions with their interactions and attributes and can be annotated with multiple labels [19]. Some labels can be very common in the data set and appear in almost all images (e.g. Sky), while others could be rare and appear only on a few images in the whole data set. Moreover, it is not uncommon for different labels to look very similar in the image context (e.g. Clouds, Smoke or Fog), making hard even for a human to create unambiguous Ground Truth (GT) labeling [2]. For a researcher working on a MLC algorithm, it is essential to have a tool that does not only evaluate algorithm effectiveness but can also reveal relations between labels and clearly indicate the weaknesses of the classifier.

Confusion matrices have been present in the evaluation of scientific models and engineering applications for a long time and are commonly used in many different areas such as computer vision [14], natural language processing [10], acoustics [16], etc. In its simplest form a confusion matrix shows a binary classifier performance in table with two rows and two columns [3], [13], [4] and represents the percentages of four possible classification outcomes: True Positive (TP), False Positive (FP), True Negative (TN) and False negative (FN). This principle is easily extended to visualization of results obtained by Multi-class classification model [11], where each object from the data set can belong to just one of many distinctive classes at any given point in time. If, for example, an object of type A is often misclassified as type B, the confusion matrix will clearly reveal this and suggest to a researcher that she or he could improve the classification model by looking for additional features that can help better distinguish classes A and B.

Even though confusion matrices are adequate for the visualization of results obtained by MCC models, they fail when it comes to Multi-label classification where an object from a data set can simultaneously belong to multiple classes. Since the analysis of the confusion matrix could provide insight into the relations between different data features and objects and also reveal inherent structure of the data itself, it is important to find a way in which a confusion matrix can be applied for evaluation of MLC models. In this paper we propose a method for representation of Multi-label classification results with confusion matrices.

The rest of the paper is structured as follows. Multi-label classification paradigm and existing evaluation measures for Multi-label classification model are presented in section II. In section III a novel approach for the computation of confusion matrices for Multi-label classification problems is proposed and elaborated in details. In section IV we give a conclusion and discuss future work.

## **2. MULTI-LABEL CLASSIFIER PERFORMANCE EVALUATION**

Multi-label classification is a supervised learning paradigm where each data instance can be assigned more than one label from the set of predefined labels. This approach gained a lot of attention in recent years as it is applicable whenever the data set is too complex to assign each data instance to a single distinctive class, i.e. characterize each data instance with a single distinctive feature or concept. Reported applications of multi-label classification include [17], [5] text categorization,

image classification, graph classification, bioinformatics, gene function analysis, emotion and sentiment recognition, multimedia annotation, social network analysis and many more. Surveys of multi-label classification techniques and state of the art approaches are given in [18], [17], [5], [21].

We define Multi-label classification problem according to [21], [5]:

Let  $\mathcal{X} = R^d$  be an input space of  $d$ -dimensional data instances and  $\mathcal{Y} = \{\lambda_1, \dots, \lambda_q\}$  the output label space with  $|\mathcal{Y}| = q$  possible labels that can be assigned to each data instance. Multi-label pattern is a pair  $(x, Y)$  where  $x \in \mathcal{X}$  is a data instance and  $Y \subseteq \mathcal{Y}$  is a set of associated true labels. Label set  $Y$  is represented as a  $q$  dimensional binary vector  $Y \in \{0,1\}^q$  where labels relevant to  $x$  are represented by 1 and labels irrelevant to  $x$  are represented by 0.

The task of multi-label training is to learn function  $\mathcal{H}_{\mathcal{ML}}: \mathcal{X} \rightarrow 2^{|\mathcal{Y}|}$  which predicts a set  $Z \subseteq \mathcal{Y}$  of relevant labels for an unseen data instance. Note that Multi-class classification could be defined as a special case of Multi-label classification where  $\mathcal{H}_{\mathcal{MC}}: \mathcal{X} \rightarrow \mathcal{Y}$  predicts a single class associated to data instance [5].

## 2.1. Evaluation Measures

Evaluation measures used to evaluate performance of Multi-class classifier are usually based on hit and miss ratio on an unseen test data with associated Ground Truth classes. Prediction of the classifier  $\mathcal{H}_{\mathcal{MC}}: \mathcal{X} \rightarrow \mathcal{Y}$  is accurate only if the predicted class is the same as the GT class. In Multi-label classification, prediction could be completely accurate if predicted labels  $Z$  are exactly the same as GT labels  $Y$ , partially accurate if  $Y \cap Z \neq \emptyset$  or completely inaccurate if  $Y \cap Z = \emptyset$ . Thorough survey of the evaluation techniques for Multilabel classification with correlation analysis of different performance measures is given in [1].

Let  $D_t = \{(x_i, Y_i) | i = 1, \dots, N\}$  be a set of multi-label patterns where  $Y_i$  is the Ground Truth set of labels for data instance  $x_i$  unseen by the classifier  $\mathcal{H}_{\mathcal{ML}}: \mathcal{X} \rightarrow 2^{|\mathcal{Y}|}$  and  $Z_i = \mathcal{H}_{\mathcal{ML}}(x_i)$  is the prediction of the classifier. Evaluation measures for evaluating performance of the MLC classifier are divided into example-based metrics and label-based metrics. Example-based metrics calculate performance for every data instance and average over the entire data set, while label-based metrics evaluate performance for each label individually and then average across all labels.

### 2.1.1. Example-based evaluation measures

Example-based evaluation measures are calculated by taking into account each instance hit and miss ratio regardless of label and averaging over the entire test set. Example based Accuracy, Precision and Recall are define with [1], [6]:

$$Accuracy_{EB}(\mathcal{H}_{\mathcal{ML}}, \mathcal{D}_t) = \frac{1}{N} \sum_{i=1}^N \frac{|Y_i \cap Z_i|}{|Y_i \cup Z_i|} \quad (1)$$

$$Precision_{EB}(\mathcal{H}_{\mathcal{ML}}, \mathcal{D}_t) = \frac{1}{N} \sum_{i=1}^N \frac{|Y_i \cap Z_i|}{|Z_i|} \quad (2)$$

$$Recall_{EB}(\mathcal{H}_{ML}, \mathcal{D}_t) = \frac{1}{N} \sum_{i=1}^N \frac{|Y_i \cap Z_i|}{|Y_i|} \quad (3)$$

Using example-based Precision and Recall, F<sub>1</sub>score [7] can be computed, representing the weighted average between Precision and Recall [1]:

$$F_{1EB} = \frac{1}{N} \sum_{i=1}^N \frac{2|Y_i \cap Z_i|}{|Y_i| + |Z_i|} \quad (4)$$

Hamming loss:

$$HammingLoss(\mathcal{H}_{ML}, \mathcal{D}_t) = \frac{1}{N} \sum_{i=1}^N \frac{|Y_i \Delta Z_i|}{|L|} \quad (5)$$

where  $\Delta$  stands for a symmetric difference of two sets, which is equivalent to the XOR operation in Boolean logic [1]. Hamming Loss (5) is a widely used evaluation measure for MLC, penalizing difference between predicted and GT labels. Both labels that are predicted and do not exist in GT and labels that are not predicted but exist in GT are taken into the account.

Subset Accuracy:

$$SubsetAccuracy(\mathcal{H}_{ML}, \mathcal{D}_t) = \frac{1}{N} \sum_{i=1}^N \mathbb{I}[Y_i = Z_i] \quad (6)$$

where  $\mathbb{I}[\Delta]$  is Iverson bracket [20], mapping true logic condition to 1 and false to 0. Subset Accuracy or Exact Match is a rigid measure considering prediction accurate only if it is exactly the same as GT.

### 2.1.2. Label-based evaluation measures

Label-based evaluation considers every label separately, reducing Multi-label classifier to a binary classifier for a particular label, with four possible prediction outcomes: True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN). Accuracy, Precision and Recall are defined by:

$$Accuracy = \frac{TP + TN}{N}$$

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

$$Recall = \frac{TP}{TP + FN}$$

$$F_1 = 2 \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

Label-based  $F_1$  score is defined by:

$$F_1 = 2 \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2TP}{2TP + FP + FN} \quad (8)$$

Label-based classification metrics for the classifier  $\mathcal{H}_{\mathcal{ML}}$  and dataset  $\mathcal{D}_t$  could be obtained by using macro or micro averaging techniques. Let  $B$  be any of the measures defined by equation (7).  $B_{macro}(\mathcal{H}_{\mathcal{ML}}, \mathcal{D}_t)$  and  $B_{micro}(\mathcal{H}_{\mathcal{ML}}, \mathcal{D}_t)$  are calculated by [1]:

$$B_{macro}(\mathcal{H}_{\mathcal{ML}}, \mathcal{D}_t) = \frac{1}{q} \sum_{j=1}^q B(TP_j, FP_j, TN_j, FN_j) \quad (9)$$

$$B_{micro}(\mathcal{H}_{\mathcal{ML}}, \mathcal{D}_t) = B \left( \sum_{j=1}^q TP_j, \sum_{j=1}^q FP_j, \sum_{j=1}^q TN_j, \sum_{j=1}^q FN_j \right) \quad (10)$$

### 3. MULTI-LABEL CLASSIFICATION CONFUSION MATRIX

Multi-label classification measures presented in the previous section are widely used in scientific papers. Detailed literature reviews on evaluation techniques for MLC are given in [1], [15]. While wide variety of proposed measures can be used to evaluate performance of a MLC algorithm, there is little information on what is happening with data instances which are labeled inaccurately. If, for example, label  $\lambda \in \mathcal{Y}$  has poor Recall, i.e. if  $\lambda$  is usually not assigned to data instances for which it is relevant, information on what labels are often assigned instead of  $\lambda$  could be very useful. In order to optimize classifier performance it is crucial to gain deeper insight into the internal classifier operations and the relations amongst the different labels. This information could also be used to select new discriminative features on the data set.

#### 3.1. Multi-Class Confusion Matrix

In the case of Multi-class classification (MCC), where each data instance is assigned to a single distinctive class, Confusion matrix is a useful and comprehensive presentation of the classifier performance on a data set with known true labels. Moreover, most of the evaluation metrics can be represented as a function of the Confusion matrix entries [12].

For the Multi-class classifier where  $\mathcal{H}_{\mathcal{MC}}: \mathcal{X} \rightarrow \mathcal{Y}$  predicts a single class, Confusion matrix is constructed by comparing the predicted class with the known GT class [8]. Each row of the matrix represents true label (GT) and each column of the matrix represents the prediction of the classifier  $\mathcal{H}_{\mathcal{MC}}$ . For each data instance  $x$  with GT class  $Y$  and predicted class  $Z$ , matrix cell corresponding to the  $Y$ -th row and  $Z$ -th column is incremented, counting the number of times that the object of class  $Y$  is assigned to a class  $Z$ . This way raw Confusion matrix is constructed where diagonal elements of the matrix represent the number of accurate classifications for each class, while off-diagonal elements represent missclassifications. Precision and Recall for each class separately

can be directly computed from the raw Confusion matrix. Precision for each class is computed by dividing a diagonal element of the Confusion matrix with the sum of all elements in the corresponding column. Recall for a class is computed by dividing a diagonal element of the matrix by the sum of all elements in the row.

An example of the Confusion matrix for MCC data set with  $N = 45$  data instances and  $q = 4$  classes is shown in Table I. Precision for each class is given in the last column, while Recall value for each class is given in the last row.

Table 1. Example of the confusion matrix for four class multi-class classifier. Precision for each class is shown in the last row. Recall for each class is shown in last column.

		Predicted class				Recall
		$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	
GT class	$\lambda_1$	8	0	0	0	1.00
	$\lambda_2$	4	9	1	1	0.60
	$\lambda_3$	3	0	7	0	0.80
	$\lambda_4$	1	0	2	9	0.75
Precision		0.50	1.00	0.70	0.90	

Table 2. Precision matrix computed from the confusion matrix given in Table 1.

	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$
$\lambda_1$	0.5	0	0	0
$\lambda_2$	0.25	1	0.1	0.1
$\lambda_3$	0.19	0	0.7	0
$\lambda_4$	0.06	0	0.2	0.9

Table 3. Recall matrix computed from the confusion matrix given in Table 1.

	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$
$\lambda_1$	1	0	0	0
$\lambda_2$	0.27	0.6	0.07	0.07
$\lambda_3$	0.3	0	0.7	0
$\lambda_4$	0.08	0	0.17	0.75

Confusion matrix normalization provides further information on relations between classes and types of classification errors [9]. Recall matrix is computed by dividing each cell of the raw Confusion matrix by the sum of all elements in the corresponding row. Diagonal elements of the Recall matrix are Recall values computed for each class, given in the last row of the raw Confusion matrix (Table I). Off-diagonal elements for the row representing true class  $Y$  represent the probability that the object of the class  $Y$  will be misclassified as class  $Z$ , where  $Y \neq Z$ . Precision matrix is computed by dividing each cell of the raw Confusion matrix with the sum of the corresponding column. Diagonal elements of the Precision matrix represent precision for the corresponding class given in the last

column of the raw Confusion matrix. Off-diagonal elements of a column representing class  $Z$  are probabilities that the object assigned to class  $Z$  really belongs to class  $Y$ ,  $Y \neq Z$ . Precision matrix computed from the raw Confusion matrix is given in Table 2 and Recall matrix is shown in Table 3.

### 3.2. Confusion-Matrix in MLC Paradigm

As illustrated in the previous example, a Confusion matrix provides comprehensive insight into the classifier performance for the MCC problems. However, there are several obstacles for the extension of this simple yet effective principle to the Multi-label classification paradigm. The contribution of the data instance  $x$  to the raw Confusion matrix is straightforward only in the trivial case where  $|Y| = |Z| = 1$ . If either  $Y$  or  $Z$  has more than one label the situation is not so clear. In the following paragraphs we will propose the algorithm for the computation of Confusion-matrices for Multi-label classification. In this work we assume that MLC predicts at least one label for each data instance and cardinality of both  $Y$  and  $Z$  is greater than 0.

Let us consider four possible scenarios for sets of true labels  $Y$  and predicted labels  $Z$ :

- (i) GT and predicted labels for data instance  $x$  are exactly the same,  $Y = Z$ , i.e. the classifier accurately predicts relevant labels. Contribution  $C$  for data instance  $x$  to the Confusion matrix is accounted by incrementing diagonal elements corresponding to label set  $Y$ :

$$C = \text{diag}(Y) \quad (11)$$

- (ii) Classifier prediction and GT differ. Prediction  $Z$  contains labels that are not relevant to data instance  $x$ . True label set  $Y$  does not contain labels that do not exist in prediction  $Z$ , i.e. all relevant labels are predicted by the classifier:

$$|Y \setminus Z| = 0, \quad |Z \setminus Y| > 0, \quad (12)$$

where  $Y \setminus Z$  represents a set of labels that exist in GT and do not exist in predicted set  $Z$ , and  $Z \setminus Y$  represents a set of labels that are predicted by the classifier and are not relevant to  $x$ . Although all relevant labels from  $Y$  are accurately predicted, the contribution of data instance  $x$  can not be accounted for by simply incrementing the appropriate diagonal elements. It is reasonable to assume that some features of  $x \in \mathcal{X}$  corresponding to concepts and semantic meaning connected with true labels  $Y$  mislead the classifier to predict non-existing labels. Proportional share from each true label contribution should be accounted for to the labels that exist in prediction and do not exist in GT. Contribution  $C$  of the data instance  $x$  to the Confusion matrix is:

$$C = [Y \otimes (Z \setminus Y) + |Y| \cdot \text{diag}(Y)] / |Z| \quad (13)$$

First element in square brackets is the outer product that redistributes the contribution of true labels to inaccurately predicted labels, i.e. accounts part of contribution of each label from  $Y$  equally to all labels in  $Z$  which are not relevant to  $x$ . Second summand in square brackets increments diagonal elements corresponding to labels in  $Y$ . Overall contribution is normalized by  $|Z|$  to accurately model the distribution of  $|Y|$  true labels to contribution  $C$ . The sum of each row representing one true label is equal to one, while the sum of all

elements of  $C$  is equal to number of true labels  $|Y|$ . Equation (13) takes equal parts of the contribution from each true label and transfers it to the inaccurately predicted labels.

- (iii) Prediction and GT differ. GT contains labels that do not exist in prediction. Prediction does not contain labels that do not exist in GT:

$$|Y \setminus Z| > 0, \quad |Z \setminus Y| = 0 \quad (14)$$

It is reasonable to assume that some features of  $x$  corresponding to true labels  $Y$  are not recognized accurately by the classifier and are attributed to other true labels from  $Y$  that are predicted in  $Z$ . Contribution of the labels that are not accurately predicted should be equally distributed to all labels that exist in both GT and prediction. Contribution to the Confusion matrix is:

$$C = [(Y \setminus Z) \otimes Z]/|Z| + \text{diag}(Z) \quad (15)$$

First summand in the equation (15) redistributes the contribution of relevant labels that are not predicted equally to all predicted labels. Second summand accounts the contribution of the predicted labels to the diagonal elements of the Confusion matrix. The sum of rows corresponding to true labels in  $Y$  is one, while sum of all elements in  $C$  is equal to  $|Y|$ .

- (iv) Prediction and GT differ. GT contains labels that do not exist in prediction. Prediction also contains labels that do not exist in GT:

$$|Y \setminus Z| > 0, \quad |Z \setminus Y| > 0 \quad (16)$$

Diagonal elements corresponding to labels that exist in both GT and prediction (if any) should be incremented. We consider these labels accurately classified. Contribution of other labels that exist in GT and do not exist in prediction should be equally distributed among labels that are predicted but do not exist in GT. Contribution to the Confusion matrix is:

$$C = [(Y \setminus Z) \otimes (Z \setminus Y)]/|Z \setminus Y| + \text{diag}(Y \cap Z) \quad (17)$$

First summand equally redistributes contribution of GT labels that are not accurately predicted to predicted labels not relevant to data instance  $x$ , normalized by cardinality of  $Z \setminus Y$  set. Second summand accounts for the contribution of accurately predicted true labels, if any. The sum of each row corresponding to labels in  $Y$  is equal to one. The sum of all elements of  $C$  is equal to cardinality of  $Y$ .

Based on the previous analysis, the algorithm that computes Multi-label Confusion matrix  $\mathcal{M}$  for dataset  $D = \{(x_i, Y_i) \mid i = 1, \dots, N\}$ , where  $x_i$  is data instance and  $Y_i \in \{0,1\}^q$  its true labels represented as  $q$ -dimensional binary vector, and classifier  $\mathcal{H}_{\mathcal{M}\mathcal{L}}$  capable of predicting labels  $Z_i = \mathcal{H}_{\mathcal{M}\mathcal{L}}(x_i)$ ,  $Z_i \in \{0,1\}^q$  for unseen data instances  $x_i$  is given in Algorithm 1.

**Algorithm 1** Multi-label confusion matrix**Input:**  $D = \{x_i, Y_i\} | i = 1, \dots, N\}$ **Output:**  $\mathcal{M}$  $\mathcal{M} \leftarrow \text{zeros}(q \times q)$ **for**  $i \leftarrow 1:N$  **do** $Z_i \leftarrow \mathcal{H}_{ML}(x_i)$ **if**  $Y_i = Z_i$  **then** $C \leftarrow \text{diag}(Y_i)$ **else****if**  $|Y_i \setminus Z_i| = 0$  **then** $C \leftarrow [(Y_i \cap Z_i) \otimes (Z_i \setminus Y_i) + |Y_i| \cdot \text{diag}(Y_i)] / |Z_i|$ **else if**  $|Z_i \setminus Y_i| = 0$  **then** $C \leftarrow [(Y_i \setminus Z_i) \otimes Z_i] / |Z_i| + \text{diag}(Z_i)$ **else** $C \leftarrow [(Y_i \setminus Z_i) \otimes (Z_i \setminus Y_i)] / |Z_i \setminus Y_i| + \text{diag}(Y_i \cap Z_i)$ **end if****end if** $\mathcal{M} \leftarrow \mathcal{M} + C$ **end for**

Algorithm 1 requires single pass through the data set to compute raw Confusion matrix for a Multi-label classifier. Once the confusion matrix is constructed, Precision and Recall matrices can easily be computed by normalizing raw confusion matrix with sum of column elements or sum of row elements, respectively [9].

**3.3. Example**

Let us illustrate Algorithm 1 by using a simple example. Data set  $\mathcal{D}$  with  $|\mathcal{D}| = 7$  samples and  $q = 4$  possible labels is shown in Table 4. Scenario for computing contribution  $C$  for MLC confusion matrix  $\mathcal{M}$ , according to the analysis given in subsection 3.2 is shown below each data instance  $x_i$ . Computed contribution  $C$  for each  $x_i$  is given in the last row (only diagonal and non-zero elements are shown).

Table 4. Example of multi-label Confusion matrix calculation. Top to bottom: GT and prediction vectors for 7 data instances with 4 possible labels; scenario for computing contribution (see subsection 3.2) contribution of each data instance to Confusion matrix.

	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$
$X_i$	1 1 0 0	0 1 1 0	0 0 0 1	1 1 1 1	0 1 1 0	0 1 1 0	0 1 0 1
$Y_i$	1 1 0 0	1 1 1 0	1 0 0 1	0 1 1 1	0 1 0 0	1 1 0 0	1 0 1 0
	(i)	(ii)	(ii)	(iii)	(iii)	(iv)	(iv)
$C$	1	0	0	0	0	0	0
		$\frac{1}{3}$	$\frac{2}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	0
	1	$\frac{1}{3}$	$\frac{2}{3}$	1	1	1	$\frac{1}{2}$
	0	$\frac{1}{3}$	$\frac{2}{3}$	1	1	0	0
	0	0	$\frac{1}{2}$	1	0	0	$\frac{1}{2}$

Contribution of  $x_1$  with  $Y_1 = Z_1$  is straightforward and does not require deeper elaboration. For  $x_2$  prediction contains two labels that exist in GT, and one additional label that does not exist in GT (scenario (ii):  $|Y_2 \setminus Z_2| = 0, |Z_2 \setminus Y_2| > 0$ ). Contribution is computed according to eq. (13):

$$\begin{aligned} Y_2 \cap Z_2 &= [0 \quad 1 \quad 1 \quad 0] \\ Y_2 \setminus Z_2 &= [0 \quad 0 \quad 0 \quad 0] \\ Z_2 \setminus Y_2 &= [1 \quad 0 \quad 0 \quad 0] \end{aligned} \quad (18)$$

$$\begin{aligned} C &= [(Y_2 \cap Z_2) \otimes (Z_2 \setminus Y_2) + |Y_2| \cdot \text{diag}(Y_2)] / |Z_2| \\ &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ \frac{1}{3} & \frac{2}{3} & 0 & 0 \\ 1 & 0 & 2 & 0 \\ \frac{1}{3} & 0 & \frac{2}{3} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{aligned} \quad (19)$$

Resulting contribution computed in eq. (19) takes into the account that true labels are partially misclassified as non-existent labels and redistributes proportional part of the contribution while keeping sum of all elements of contribution matrix equal to  $|Y_2|$ . Contribution of the  $x_3$  is computed according to the same scenario.

For  $x_4$  GT contains label that is not accurately predicted, i.e.  $|Y_4 \setminus Z_4| > 0, |Z_4 \setminus Y_4| = 0$ . Contribution is computed according to eq. (15):

$$\begin{aligned} Y_4 \cap Z_4 &= [0 \quad 1 \quad 1 \quad 1] \\ Y_4 \setminus Z_4 &= [1 \quad 0 \quad 0 \quad 0] \\ Z_4 \setminus Y_4 &= [0 \quad 0 \quad 0 \quad 0] \\ C &= [(Y_4 \setminus Z_4) \otimes Z_4] / |Z_4| + \text{diag}(Z_4) \end{aligned} \quad (20)$$

$$= \begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (21)$$

Equation (21) increments diagonal matrix elements for accurately predicted labels and distributes evenly the contribution of true label that is inaccurately predicted to predicted labels. Contribution of  $x_5$  is computed in the same manner.

Samples  $x_6$  and  $x_7$  represent scenario (iv) where both GT and prediction contain labels that do not exist in the opposite vector, with contribution computed according to the equation (17). For  $x_6$  label that exists in both GT and prediction is considered correctly assigned and corresponding diagonal element is incremented. The contribution of other true label is assigned to inaccurately predicted

label. For  $x_7$  both true labels are inaccurately classified as predicted labels that do not exist in GT. As it is not possible to conclude which features were misinterpreted, contribution of the labels is evenly distributed across the predicted labels.

Table 5. Raw confusion matrix  $\mathcal{M}$  for the example given in Table 4. Precision for each class is shown in the last row, Recall for each class is shown in last column.

		Predicted class				Recall
		$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	
GT class	$\lambda_1$	<b>1.00</b>	0.33	0.33	0.33	0.50
	$\lambda_2$	0.83	<b>4.67</b>	0.50	0.00	0.78
	$\lambda_3$	1.33	1.00	<b>1.67</b>	0.00	0.42
	$\lambda_4$	1.00	0.00	0.50	<b>1.50</b>	0.50
Precision		0.24	0.78	0.56	0.82	

Resulting raw MLC confusion matrix is show in Table 5. Unlike raw multi-class confusion matrix, entries in the raw multi-label confusion matrix are floating point numbers, as contribution of each label in GT could be split amongst labels in prediction. Precision matrix obtained by normalization of raw confusion matrix is shown in Table 6 and Recall matrix is shown in Table 7.

Table 6. Precision matrix computed from the raw confusion matrix  $\mathcal{M}$ .

	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$
$\lambda_1$	0.24	0.06	0.11	0.18
$\lambda_2$	0.20	0.78	0.17	0
$\lambda_3$	0.32	0.17	0.56	0
$\lambda_4$	0.24	0	0.17	0.82

Table 7. Recall matrix computed from the raw confusion matrix  $\mathcal{M}$ .

	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$
$\lambda_1$	0.50	0.17	0.17	0.17
$\lambda_2$	0.14	0.78	0.08	0
$\lambda_3$	0.33	0.25	0.42	0
$\lambda_4$	0.33	0	0.17	0.50

## 4. CONCLUSIONS

Confusion matrix is a useful and comprehensive presentation of classifier performance. It is not just another way of computing Precision, Recall or some other evaluation measure, it is rather a magnifier that provides us with deeper insight into the classifier internal operation. Inspection of the confusion matrix and its derivatives provide us with strong clues on relations between classes and labels representing semantic meanings and concepts assigned to data instances. Confusion matrix reveals classifier weaknesses and suggests guidelines for further research and improvement in model

performance. Analysis of the confusion matrix could also provide insight into relations between different data features and objects and reveal inherent structure of the data itself.

In this paper a method for the computation of a confusion matrix for the Multi-label classification model is proposed. The proposed algorithm overcomes the limitations of the existing approaches in modeling relations between the classifier output and the Ground Truth classification. Multi-class classification can be considered as a special case of multi-label classification with imposed limitation  $|Y| = |Z| = 1$ . Accordingly, evaluation metrics used in Multi-label classification should be generalization of techniques and measures used in Multi-class classification. It is easy to see that by imposing the same limitation, proposed algorithm for the confusion matrix computation reduces to a simpler form which is used to compute confusion matrix for Multi-class classification problem.

The proposed technique emerged from our work on a specific multi-label classification problem. In our problem, classes are loosely defined and often share common features. Our intention was not to propose a new evaluation metrics, but to develop a tool to gain deeper understanding of data set and classifier operation. Therefore, the proposed technique is indeed used in decision making in a specific case study. Specific details from our work are omitted and would exceed the scope of the paper. Intention of this paper is to provide an insight to this technique to the classifier developers community and expect feedback on specific issues of the technique.

The proposed technique emerged from our work on multilabel classifier for scene understanding which will be used as a case study for this method. We expect to propose even more general solution without some minor limitations mentioned in this method description. In the future we also expect to deliver tools for automatic extraction of confusion matrix semantics as a result of comprehensive testing on several case study multi-label classifiers. In that stage we will also consider computational complexity aspect of the proposed solution. According to ours' best knowledge, at this moment there are no other proposed methods for constructing multi-label confusion matrix, and we consider this a good basis for developing a general solution to this problem, or a specialized method for evaluation of a specific problems.

## ACKNOWLEDGEMENTS

This work is partly supported by the Ministry of Science, Education and Sport of the Republic of Croatia under VIF Grant through project ViO - Vision Based Intelligent Observers.

## REFERENCES

- [1] B. Pereira, R., Plastino, A., Zadrozny, B., Merschmann, L.: Correlation analysis of performance measures for multi-label classification. *Information Processing and Management* 54, 359–369 (05 2018). <https://doi.org/10.1016/j.ipm.2018.01.002>
- [2] Braović, M., Stipaničev, D., Krstinić, D.: Cogent confabulation based expert system for segmentation and classification of natural landscape images. *Advances in Electrical and Computer Engineering* 17, 85–94 (2017). <https://doi.org/10.4316/AECE.2017.02012>, <http://dx.doi.org/10.4316/AECE.2017.02012>
- [3] Canbek, G., Sagirolu, S., Taskaya Temizel, T., Baykal, N.: Binary classification performance measures/metrics: A comprehensive visualized roadmap to gain new insights. In: 2017 International Conference on Computer Science and Engineering (UBMK). pp. 821–826 (10 2017). <https://doi.org/10.1109/UBMK.2017.8093539>

- [4] Fawcett, T.: An introduction to roc analysis. *Pattern Recogn. Lett.* 27(8), 861–874 (Jun 2006). <https://doi.org/10.1016/j.patrec.2005.10.010>, <http://dx.doi.org/10.1016/j.patrec.2005.10.010>
- [5] Galindo, E.L.G., Ventura, S.: Multi-label learning: a review of the state of the art and ongoing research. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* 4, 411–444 (2014)
- [6] GiraldoForero, A.F., Jaramillo-Garzón, J., Castellanos-Dominguez, G.: Evaluation of examplebased measures for multi-label classification performance. In: *Lecture Notes in Computer Science*. vol. 9043, pp. 557–564 (04 2015). <https://doi.org/10.1007/978-3-319-16483-054>
- [7] Goutte, C., Gaussier, E.: A probabilistic interpretation of precision, recall and f-score, with implication for evaluation. In: *Proceedings of the 27th European Conference on Advances in Information Retrieval Research*. pp. 345–359. ECIR'05, Springer-Verlag, Santiago de Compostela, Spain (2005)
- [8] Haj Mohamad, T., Chen, Y., Chaudhry, Z., Nataraj, C.: Gear fault detection using recurrence quantification analysis and support vector machine. *Journal of Software Engineering and Applications* 11 (05 2018). <https://doi.org/10.4236/jsea.2018.115012>
- [9] Hardin, P.J., Shumway, M.J.: Statistical significance and normalized confusion matrices. *Photogrammetric Engineering and Remote Sensing* 63, 735–740 (1997)
- [10] Kavitha, A.S., Shivakumara, P., Kumar, G.H., Lu, T.: Text segmentation in degraded historical document images. *Egyptian Informatics Journal* 17(2), 189–197 (2016). <https://doi.org/https://doi.org/10.1016/j.eij.2015.11.003>
- [11] Koço, S., Capponi, C.: On multi-class learning through the minimization of the confusion matrix norm. *Journal of Machine Learning Research* 29 (03 2013)
- [12] Koyejo, O., Ravikumar, P., Natarajan, N., Dhillon, I.S.: Consistent multilabel classification. In: *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 2*. pp. 3321–3329. NIPS'15, MIT Press, Montreal, Canada (2015)
- [13] Labatut, V., Cherifi, H.: Evaluation of performance measures for classifiers comparison. *Ubiquitous Computing and Communication Journal* 6, 21–34 (11 2011)
- [14] Li, S., Deng, W.: Blended emotion in-the-wild: Multi-label facial expression recognition using crowdsourced annotations and deep locality feature learning. *International Journal of Computer Vision* (Nov 2018). <https://doi.org/10.1007/s11263-018-1131-1>, <https://doi.org/10.1007/s11263-018-1131-1>
- [15] Madjarov, G., Kocev, D., Gjorgjevikj, D., Deroski, S.: An extensive experimental comparison of methods for multi-label learning. *Pattern Recogn.* 45(9), 3084–3104 (Sep 2012). <https://doi.org/10.1016/j.patcog.2012.03.004>, <http://dx.doi.org/10.1016/j.patcog.2012.03.004>
- [16] Miller, G.A., Nicely, P.E.: An analysis of perceptual confusions among some english consonants. *The Journal of the Acoustical Society of America* 27, 338–352 (03 1955). <https://doi.org/10.1121/1.1907526>
- [17] Tidake, V.S., Sane, S.S.: Multi-label classification: a survey. *International Journal of Engineering and Technology* 7(4.19), 1045–1054 (2018). <https://doi.org/10.14419/ijet.v7i4.19.28284>, <https://www.sciencepubco.com/index.php/ijet/article/>
- [18] Tsoumakas, G., Katakis, I.: Multi-label classification: An overview. *International Journal of Data Warehousing and Mining (IJDWM)* 3(3), 1–13 (2007), <https://EconPapers.repec.org/RePEc:igg:jdwm00:v:>
- [19] Wang, J., Yang, Y., Mao, J., Huang, Z., Huang, C., Xu, W.: Cnn-rnn: A unified framework for multi-label image classification. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 2285–2294. Las Vegas, Nevada, USA (2016)
- [20] Weisstein, E.W.: Iverson Bracket From Math- World – A Wolfram Web Resource (2019), <http://mathworld.wolfram.com/IversonBracket.html>, [Online; accessed 9-March-2019]
- [21] Zhang, M.L., Cheng Zhou, Z.: A review on multi-label learning algorithms. *IEEE Transactions on Knowledge and Data Engineering* 26(8):1819-1837

**AUTHORS**

**Damir Krstinić** received the Ph.D degree from the University of Split, Split, Croatia in 2008. He is Associate Professor on Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture of the University of Split, Department for Modelling and Intelligent Systems. His main area of research interest are computer vision, image understanding and machine learning. He has been involved in several research and technological projects related to forest fire research and is project leader of Wildfire Early Detection and Monitoring System for Croatian Forests.



**Maja Braović** received her Bachelor's and Master's degrees in Computer Science from the University of Split, Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, in 2008 and 2010, respectively. She received her PhD degree in Artificial Intelligence from the same Faculty in 2015, where she is also currently a postdoctoral researcher. Her main areas of research interest include artificial intelligence, computer vision, image understanding and natural language processing.



**Ljiljana Šerić** is associate professor at University of Split, Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture. She received her PhD in electrical engineering and computer science in 2010. She is a member of Department for Modelling and Intelligent Systems and Centre for wildfire research. She has co-authored more than 50 scientific papers. Her research interests are focused on artificial intelligence and web and distributed information systems.



**Dunja Božić - Štulić** received her Master of Engineering degree in Electronics and Computer engineering in 2014, from Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, University of Split, Croatia. She is currently working as research assistant at Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, University of Split, Croatia and is PhD candidate. Dunja is co-author of 5 conference paper's and 2 scientific paper's. (<http://bib.irb.hr/lista-radova?autor=355870>). Her research interests include machine learning, deep learning and artificial intelligence.



# DEVELOPING E-LEARNING SYSTEM TO SUPPORT MATH IN THAI EDUCATION CURRICULUM (PRIMARY LEVEL)

Piya Techateerawat

Engineering Faculty, Thammasat University, Thailand

## ABSTRACT

*E-learning is a common tool to support the education in variety of scenarios. As the education content can be prepared by the group of specialists, but the skilled teachers are limited in remote area. Also, the contents in most curriculum are planned to distribute to limited skilled people. The gap of education can be full-filled with E-learning system. However, the conventional E-learning is high cost system and not appropriated for rural area. Also, open-source system is complicated to implement and configure in dedicated curriculum. This research is proposed the customized design of E-learning system for primary Thai education curriculum. As Thai education curriculum was updated in 2017, school and teacher need to update the plan and methodology accordingly. Our research is based on the actual Thai school in Ratchaburi province by using the Android framework. The system is designed and implemented from actual requirements from teachers and students in grade 3. As a result, the result show student involvement and continue using of system in both school and extra hours. The feedback from actual usage also is evaluated.*

## KEYWORDS

*E-learning, Android, Thai Curriculum.*

## 1. INTRODUCTION

Electronic learning (E-learning) is a general system to support the learning process and assist the content delivering to the specified learner. The key features of E-learning are to let learner access contents promptly with less constraints e.g. remote area, limited time, lack of teacher and large number of learners. Since network performance and equipment are easily access currently, a number of learners' request more of E-learning system. Also, learners expect to use E-learning for knowledge sharing and on-demand contents.

This paper shows the customized solution of E-learning system that can be implemented in adaptive content with specified need (2017 Thai education curriculum) The solution also can be applied to other scenarios for other organizations.

## 2. E-LEARNING SYSTEM

E-learning is shown as critical system for content sharing in business, research and academic. Since the market share of E-learning system reached \$2 trillions in 2001, many organizations implement the system for accessible and extra class support [1, 2].

David C. Wyld et al. (Eds): SAIM, SIPM, ACSIT, FCST, ICITE - 2020  
pp. 15-21, 2020. CS & IT - CSCP 2020

DOI: 10.5121/csit.2020.100802

The main benefit of E-learning system is cost effective, flexibility, accessibility and content distribution. As the system are content centralized and using the technology platform for access anywhere-anytime. Then overall, it reduced the cost of staff, operation task and open the opportunity to more people [3]. It also adds more features that can provide online chat, AI enquiry, reminder, online quiz and content suggestions.

In general, large system is approaching with framework [4-5]. As developing time can be reduced significantly by using the built-in library. The security is conformed by the framework library as well as updating regularly with framework update. The important key is maintenance as framework using the standard model following with guidance.

The one of challenge of implementing E-learning is cost that increased 83 percent from 1998-2003 [6]. This is a struggle for many organizations especially in education system. However, the need of E-learning also pushes to one of the largest share markets [7]. For, specific small organization and limited budget company are difficult to access to the commercialized platform as well as the open source system also required the skilled person to configure and understand that need follow the scheme of open source package.

### **3. FRAMEWORK**

Android framework is a structure that prepare for programmer to develop application. Adobe Illustrator and Figma are the of tools to design on UI and character based. SQLite is use for database of quiz and content storage. This assist for both user and backend side to communicate and interact and share contents over the network. Most of Android framework is based on JDK model, this scheme let programmer customize the framework to meet the requirement by using the universal programming language and layout organization that is compatible to designated devices. Also, most of the framework is based on open source, so it can minimize the cost of license and maintenance.

The advantage of using this framework is let the structure more robust and more secure from the prepared package compared to developing from bare bone. In addition, the framework has prepared the libraries and themes for developer to choose and apply. So, implementation can be rapid and more convenience to meet the requirement.

SQLite provides the up-to-date content structure that can add, delete and modify the content of application on the fly. These support the upcoming content update from the education curriculum.

The special request is also to add the gaming features. The project designs the main character model of the game in Adobe Illustrator and lay it on Figma to handle the flexible of mobile device. This game contains many scenes such as Main Scene, Menu Scene, Game Scene etc., and the application also includes the question scene.

The gaming section in layout based on the purpose of core education curriculum in math subject. Also, all the main features can be operated in Android system.

### **4. PLANNING & DEVELOPMENT**

This E-learning system is based on the actual case from Thai School in Ratchaburi province. The group of primary students and teacher in 2018 is based on the research. The development and evaluation are also based on the actual activity that interacts with students.

The design must be compatible to the new mobile devices (e.g. smart phone and tablets). The contents in each screen should be simplified with one main idea.

The purpose of this E-learning system is to support the actual class and student has sufficient knowledge in using computer technology. The requirement and design of this system is survey from in-class students.

#### **4.1. Requirement**

- The system must be convenient for in-class student to apply and access.
- Students which do not enrol may apply and register for the interest class.
- Every user can browse the contents in subject.
- The interface should be kids friendly.
- The objective must follow the core curriculum.
- The quiz must the task of activities in the class.
- The flow of application should let the student feel enjoy.

#### **4.2. Application Design**

This application is focused to the math subject from Thai core curriculum. In primary, 5 topics are the main education topics: 1.) Calculation 2.) Measurement 3.) Weight 4.) Time and 5.) Shape. Our design is to support the main class activities that provide the summary video and quiz. For each topic in curriculum the amount of content is weigh by the teacher plan as follows:

The main page of application is focused on 5 main tasks: 1.) Calculation 2.) Measurement 3.) Weight 4.) Time and 5.) Shape. This page let student to navigate and see the overview of task where the character and symbolic to illustrate grade 3 students easily follow the assigned task from teacher as shown in figure 1. The character in the application is from the brainstorm of student and teacher in the school to get the sketch and forming up by the research team. As a result, the collaboration leads the students the feel involves and willing to join the tasks and activities as shown in figure 2.

In calculation section, the tasks are addition, subtraction, multiplication and division. Once, the user selects the topics the detail of given topics will be shown. The drill-down UI is provided and given in each section and provided the given content and video which is approved by teacher as shown in figure 3.

In quiz section, the quiz is based on the learning knowledge by pulling up the question from SQLite and arrange by the selected topics. When the user finishes the quiz, simple mark output is given. The quiz allows the user to re-do in the case the gain the understanding as shown in figure 4.



Fig. 1. Main Application Design



Fig. 2. Sub-Menu: Calculation Section



Fig. 3. Video Content

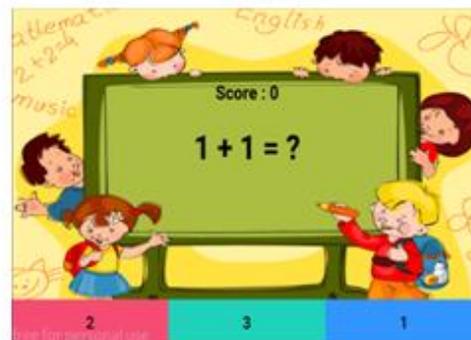


Fig. 4. Quiz Section

To support the flexible in design, background is laid with browser in Android studio. The character is drawn by Adobe Illustrator or equivalent and lay it with Figma.

In quiz section, the quiz is based on the learning knowledge by pulling content from SQLite and using Genymotion to adjust the final result.

## 5. RESULT

The end result of system is based on the actual testing in grade 3 students from primary school in Ratchaburi province. The group of 20 students are using this application along with class study and giving the test result and feedback as follows:

Administrator mode:

- Log-in to the system.
- Registering to the system.
- Access to E-learning contents & downloadable contents.
- Some E-learning subject provide optional quiz.
- Search function.
- Forum for learner to share/discuss knowledge.
- File organization Add, Modify, Delete.
- Profile management.
- History organization (Review Content).

User mode:

- Topic selective.
- Content Summary
- Quiz
- Flexible in browsing
- Simple UI
- Design from user requirements.

The test result is based on the given time allocation of teacher to the students where the two weeks allocation on everyday task. The average time per slot is 78 minutes. Also, the parents of each student are given task to share their child where the average time is spending 52 minutes in 5 assignment tasks across in these two weeks.

Table 1. Summation of Usage Time

	Male	Female
Average usage per week	7	7.7
Average Time per usage at school (minutes)	72	84
Willing to continue to use.	85.71%	87.5%
Average time per usage at home (minutes)	48	54

Table 2. Feedback Result

Feedback (out of 5)	Average	SD
Quality of content	2.95	0.1
To support the class knowledge	2.95	0.1
Quality of application	2.93	0.15
Effective to learning process	2.97	0.07

## 6. DISCUSSION

As this system is created based on the actual user (grade 3 school students) requirement and reviewed by the same group student, our purpose is to gain more contribution of student. The objective is gaining more attention and support the in-class education.

As shown in Table 1, our objective is to evaluate the contribution of learner. As the learners does not intention to use the E-learning before this research. To motivate learners to contribute the new learning tools are shown in both male and female. The result also shows the continuous of the involvement for all study period in both school and home. This usage time includes both quiz and learning content based on 2017 Thai education curriculum.

A large number of involvements also require network traffic and computer resource. With this need, in our research has a large scale of computer resource in research lab. However, in real case, the system requires implementation resource in the organization. Therefore, with limited resource and recent technology, cloud server is suggested for the proposed system.

The quality from feedback is shown as above average due to the design and requirement of their choices, however they suggest that the improvement on the next version of application would be expected.

Overall, the customized system shows the significant improved in main objective of E-learning system. This also provides the solution that new trend of users needs the information in their own need format and their own need. The alternative solution is served their task achieved in this area.

## 7. CONCLUSION

E-learning is challenged in every organization to customize for appropriate for each organization. The difference in objective, members' background and knowledge leads to different expectation. The customization system is a solution with a high cost of out-of-the-box software.

This paper suggest the customized E-learning system based on Android framework. The specific learning group are based on grade 3 primary school in Ratchaburi province. The features of this system are content sharing, online quiz and feedback system to the responding teacher in school. The implementation of deployed system is based on user's requirement that tend to meet with expectation.

The trend of user involvement is improving the participation in most of area especially support Thai core curriculum in primary school.

## ACKNOWLEDGEMENTS

We would like to thank Faculty of Engineering, Thammasat University, Thammasat University, Optical and Quantum Communication Research Lab and National Research Council of Thailand.

## REFERENCES

- [1] Cisco Systems, Cisco Systems IQ Atlas, Cisco, 2001.
- [2] K. Fry, "E-learning markets and providers: some issues and prospects", *Education + Training*, Vol. 43 No. 4/5 pp. 233-9, 2001.
- [3] S.M. Furnell, P.D. Onions, U. Bleimann, U. Gojny, M. Knahl, H. F. Roder and P.W. Sanders, "A security framework for online distance learning and training", *Internet Research: Electronic Networking Applications and Policy*, Vol. 8 No. 3, pp. 236-42, 1998.
- [4] S. Alexander, "E-learning developments and experiences", *Education + Training*, Vol. 43 No. 4/5 pp. 240-8, 2001.
- [5] S.H. Garrison and D.J. Borgia, "Using an Internet based distance learning to teach introductory finance", *Campus-Wide Information Systems*, Vol. 16 No. 4, pp. 136-9, 1999.
- [6] D. Lance, " Venture captial veiwpoints and E-learning futures", *The Business of E-learning: Bringing Your Organizationin the Knowledge Economy*, Univeristy of Technology Sydney, 2000.
- [7] P. Henry, "E-learning Technology, Content and Services", *Education + Training*, Vol. 43 No. 4, pp. 249-55, 2001.

**AUTHOR**

Piya Techateerawat is a lecturer of Computer Engineering at Thammasat University. He received his B.Eng. from University of New South Wales, Australia with Honors in 2004. He continued his PhD study at Royal Melbourne Institute of Technology University, Australia, where he obtained his PhD in Wireless Sensor Network Security. His current interests involve applications of Sensor Network, Security and Quantum Cryptography.



© 2020 By AIRCC Publishing Corporation. This article is published under the Creative Commons Attribution (CC BY) license.



# SIMPLE FACE THERMAL FEATURES FOR GENDER DISCRIMINATION

Georgia Koukiou and Vassilis Anastassopoulos

Electronics Laboratory Physics Department, University of Patras, Greece

## **ABSTRACT**

*A very simple approach is proposed for gender discrimination using thermal infrared images of the persons' face. The selected features are actually based on the mean value of the pixels of specific locations on the face. It is proved that the discrimination is feasible either by simple visualization in the feature space or by using a relatively simple neural network for this purpose. All cases of persons from the used database, males and females, are correctly distinguished based on the mean value of the employed locations of the face. Classification results are verified using two conventional approaches, namely: a. the simplest possible neural network so that generalization is achieved along with successful discrimination between all persons and b. the leave-one-out approach to demonstrate the classification performance on unknown persons using the simplest classifiers possible.*

## **KEYWORDS**

*Thermal Infrared, Face Recognition, Gender Discrimination, Neural Networks.*

## **1. INTRODUCTION**

The approach used in this correspondence for feature extraction was based on a thorough comparison between the available images in the thermal infrared region. The thermal infrared region of the electromagnetic spectrum which lies between 7 and 13 microns, is the region in which radiate all bodies with typical temperature that of the environment. Infrared signatures of the face describe face temperature distribution which depends on physiological and psychological condition of the person.

Infrared images acquired from the face of males and females differ significantly in specific locations. Accordingly, the used features should incorporate this type of information. Thus, the mean value of the pixels in four different locations of the face is used to form the feature vector. This feature vector is fed into Neural Network structures to perform gender identification. In this work, we seek for the simplest neural structure to perform successful gender identification and simultaneously we examine the suitability of the derived features to support correct identification even in the case of unknown persons present. For this purpose the leave-one-out method is employed. The classification performance of the neural structures is very efficient giving very good gender separability in the feature space, which is proved by the distinct locations of the clusters of males and females in the feature space.

The work is organized as follows. Work related to the presented method is extensively exposed in section 2. In section 3 the infrared data employed for gender identification are described. In the

next section the feature extraction approach is outlined and the feature space is demonstrated. Gender identification performance is carried out in section 5, while the conclusions are drawn in section 6.

## 2. RELATED WORK

Various techniques have been presented in the past regarding gender discrimination. In [1] features are combined using fusion techniques from visible and infrared images for gender recognition. The presented technique is very complicated and the features used from specific locations on the face in case of infrared images are not conceptual prominent, while a similar approach for feature extraction has been proposed in the past [2] for drunk person discrimination. Furthermore, thermal features using eye signatures were used with successful results in intoxication identification [3]. Gender recognition using body-based images from visible and thermal cameras was carried out in [4]. The approach is basically employing movement detection and classification algorithms. The possibility of deducing gender from face images obtained in the near-infrared (NIR) and thermal (THM) spectra is elaborated in [5] by means of local binary pattern histogram (LBPH) features. In [6-7] methods for gender and ethnicity identification are described with small emphasis on thermal signatures. In [8] a combination of visible and infrared cameras is employed based on body images and neural networks for gender identification.

Other approaches have also been developed for gender discrimination based on handwriting, voice, walking movements and fingerprints. Specifically, in [9] a system is presented to predict gender of individuals from offline handwriting samples. The technique relies on extracting a set of textural features from handwriting samples of male and female writers and training multiple classifiers to learn to discriminate between the two gender classes. The features include local binary patterns, histogram of oriented gradients, statistics computed from gray-level co-occurrence matrices and features extracted through segmentation-based fractal texture analysis. The work in [10] presents a study to predict gender of individuals from scanned images of their handwritings. The proposed methodology is based on extracting a set of features from writing samples of male and female writers and training classifiers to learn to discriminate between the two. Writing attributes like slant, curvature, texture and legibility are estimated by computing local and global features. The study in [11] investigates the relative importance of temporal and spectral cues in voice gender discrimination and vowel recognition by normal-hearing subjects listening to an acoustic simulation of cochlear implant speech processing and by cochlear implant users. The results suggest that both spectral and temporal cues contribute to voice gender discrimination and that temporal cues are especially important for cochlear implant users to identify the voice gender when there is reduced spectral resolution.

An adaptive three-mode PCA framework was presented in [12] for recognizing gender from walking movements. Prototype female and male walkers are initially decomposed into a sub-space of their three-mode components (posture, time, gender). Then is assigned an importance weight to each motion trajectory in the sub-space and have the model automatically learn the weight values (key features) from labeled training data. In [13] it is revealed that an information-rich latent fingerprint has not been used to its full potential. The content present in the sweat left behind—namely the amino acids—can be used to determine physical such as gender of the originator. As a result, it is possible to focus on the biochemical content in the fingerprint using a biocatalytic assay, coupled with a specially designed extraction protocol, for determining gender rather than focusing solely on the physical image.

### 3. INFRARED DATABASE USED

The thermal infrared images required to assess the proposed gender identification method have been obtained by our research team during a systematic and well organized experimental procedure. Researchers from our university were asked to participate in the experiment. The Thermal Vision Micron/A10 infrared camera was employed for image acquisition. The resolution of the camera is 160x128 pixels and is sensitive in the thermal infrared region of the electromagnetic spectrum i.e. from 7.5 to 13.0  $\mu\text{m}$ . At the middle of this region of wavelengths i.e. at 10  $\mu\text{m}$ , a perfect black body with mean temperature around 300 degrees Kelvin, radiates its maximum according to the Wien Law [14]. The human body as being in the same temperature with the environment i.e. 300 degrees Kelvin, radiates in this exact region of wavelengths.

The participants in the experiment were healthy, calm and in normal physical and psychological condition. No illness or fever, no psychological stress, other pathological reason or any kind of body exercises were recorded for any one of the participants. Thus, the thermal images of the face were suitable for comparisons between males and females without having been affected by other factors.

The thermal infrared images were recorded in a well organized database. The database contains information regarding the age, the weight and the sex of the participant. In each acquisition, a sequence of 50 frames was acquired from each person with a sampling period of 100 msec between the frames.

In order to secure smooth experimental procedure that would lead to reliable data acquisition the participants were asked to be present in the room of the experiment half an hour earlier and to keep calm during the experimental procedure. Simultaneously, the temperature of the room where measurements took place was kept between 23 and 25 degrees Celsius (almost 300 Kelvin). No physical light was used in the room of the experiment. The distance of each face from the camera was kept almost 30 cm for all persons so that the acquired images could be compared. The images of eight females and 8 males from the database were employed in this present gender discrimination procedure. In Figure 1 are shown a thermal image from the face of a male (left) along with one from a female (right).

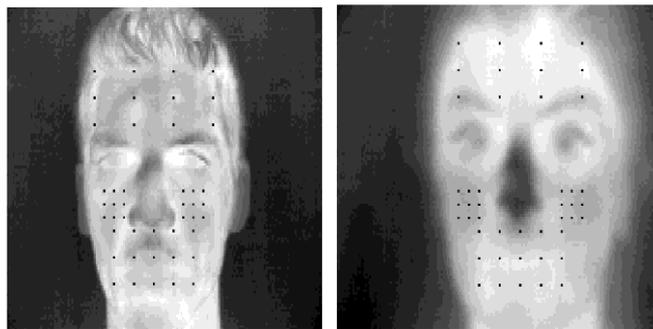


Figure 1: Infrared images of the face of a man (left) and a woman (right). According to the gray level of the regions depicted with the black dots the two genders differentiate significantly in their thermal representation.

#### 4. FEATURE EXTRACTION AND GENDER DISCRIMINATION

In the thermal infrared region of the electromagnetic spectrum (7-13 microns) the face transmits electromagnetic radiation as an almost perfect black body and the amount of transmitted energy depends on the physiological condition of the person and is irrelevant of the external artificial lighting. The features to be selected for gender discrimination have to describe the temperature distribution on the face of the persons and obviously this temperature distribution map has to be different between males and females. In order to check this temperature distribution on the face of the persons the values of the pixels in four specific regions of the face are recorded.

In Figure 1 are shown the faces of a man and a woman obtained from our database. The four inspected regions on each face are: the forehead, the left cheek, the right cheek and the region of the mouth. By simple observation it is obvious that the region of the forehead has different brightness for the females compared to the corresponding of the males. The same happens for the region of mouth as well. Consequently, each person's face is represented in a 4-dimensional space by means of these four values. In practice, if in each face the intensity (temperature) of the forehead is considered as the value of reference for the other locations, then the feature space is of 3 dimensions and the persons can be represented in this space with three values each. This fact is graphically depicted in Figure 2 for eight men (cross) and eight women (diamond) of our database. In this representation can be easily realized that the two populations form distinct clusters in the feature space that can be separated by a relatively simple decision boundary (discriminant function) not necessarily linear.

The data available for the eight males and eight females are given in Tables 1 and 2. For each person a four-point column is provided. The data are given in order the reader to be able to test and verify their capability in gender classification.

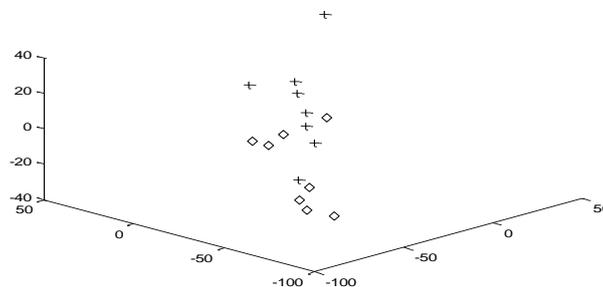


Figure 2. The 3-D feature space. Eight men (crosses) and eight women (diamonds) from our database are depicted. Males and females can be easily separated except the female (diamond) on the upper right.

Table 1. In each column are given the values for each male in the feature space.

1	2	3	4	5	6	7	8
208	227	195	222	204	198	171	221
168	197	196	205	204	173	200	224
168	205	206	208	212	208	193	227
208	198	205	217	207	212	208	214

Table 2. In each column are given the values for each female in the feature space.

1	2	3	4	5	6	7	8
234	236	227	230	225	225	227	215
168	170	183	154	208	202	182	191
176	174	185	142	232	218	216	183
214	211	203	212	205	216	221	223

## 5. GENDER CLASSIFICATION RESULTS

Gender classification results were carried out using two different conventional classification approaches. Firstly, a neural network is employed in order to examine the separability of the feature space into two subspaces corresponding to the two gender classes and simultaneously to study the degree of difficulty in drawing a discriminating hyper plane among the two classes. After that the leave-one-out approach is used to demonstrate the classification performance of the employed features on unknown persons.

### 5.1. Classification using neural networks

The classification attempt started by testing feed forward neural networks having a significant number of neurons in their 3 layers. The training algorithm was the conventional back propagation procedure. All MATLAB available routines were used.

Two different goals we tried to achieve during training. The first one was to select a neural structure capable to converge fast to a minimum giving the best gender classification result. For this purpose we started with a network having a relatively large number of perceptrons i.e. 21 neurons in the input layer, 11 neurons in the hidden layer and 1 neuron in the output layer (binary output - male or female). This network was found very effective in classifying the 16 persons.

The second goal was the selected neural structures to meet the property of generalization. This means to be capable after training to correctly classify the new incoming vectors representing persons of unknown gender. For this purpose a low number of epochs was selected in the training procedure.

Lowering gradually the number of neurons in the input and hidden layers we found that even a very simple neural structure of 3-2-1 can satisfactorily classify all the 16 samples in the 4-D feature space. Even if this simple structure was converging to a minimum after several attempts, it was preferred compared to larger ones since it was very effective in classifying unknown samples according to the leave-one-out approach given in the next subsection. For the reader to be able to test our results, the weights and the biases of this simple structure are given in Tables 3 to 5. In Table 3 is given the information regarding the 3 neurons of the input layer. In Table 4 the relative information for the hidden layer is given, while in Table 5 the weights and the bias for the output neuron are provided.

Table 3. Weights and biases values for the 3 neurons of the input layer

	<b>Neuron 1</b>	<b>Neuron 2</b>	<b>Neuron 3</b>
W1	-0.3419	-0.2219	0.0195
W2	2.8386	-0.2261	0.0045
W3	-1.3043	0.0305	0.0221
W4	-1.0688	0.3757	0.0101
bias	0.3403	1.1603	0.6821

Table 4. Weights and biases values for the 2 neurons of the hidden layer

	<b>Neuron 1</b>	<b>Neuron 2</b>
W1	-1.3745	-1.7958
W2	-0.4105	-1.5866
W3	1.0303	0.0669
bias	1.7587	-1.5691

Table 5. Weights and bias values for the output neuron

	<b>W1</b>	<b>W2</b>	<b>bias</b>
Output Neuron	0.3320	-2.9791	0.9287

## 5.2. Leave-one-out approach

The Leave-one-out method is employed to verify that the used neural structure behave appropriately with the available data in the classification procedure. i.e. the classifier is built using in the training phase all the data except one, and is expected in the test phase to be able to classify correctly the excluded information. This is actually the generalization property which requires the classifier not to be over trained on the available data, so that be able to classify correctly the new incoming vectors representing persons of unknown gender.

For this purpose a neural structure of size 3-2-1 is trained using 15 of the persons and then is tested whether the excluded person is correctly classified. This procedure was repeated 16 times, each time excluding one of the persons. In all cases, the devised neural structures classified with success the remaining person. The classification goal (classification error) achieved during the training procedure by each of the 16 neural structures was very small which reveals that this small neural network adapts fast to the needs of the problem. In Table 6 the classification goals for all 16 structures are given.

Table 6. Classification goals for all 16 structures devised for the leave-one-out method

Neural structure	Person excluded	Classification goal
1	1st male	0.001
2	2nd male	0.009
3	3rd male	0.002
4	4th male	0.038
5	5th male	0.009
6	6th male	0.003
7	7th male	0.001
8	8th male	0.001
9	1st female	0.001
10	2nd female	0.008
11	3rd female	0.030
12	4th female	0.003
13	5th female	0.090
14	6th female	0.130
15	7th female	0.009
16	8th female	0.004

We have to mention at this point that only in the 16th case in Table 6, the neural structure of size 3-2-1 neurons achieved after many attempts a convergence so that it was able to correctly classify the excluded 16th person or otherwise the 8th female. However, this behaviour was expected since according to the feature space depicted in Figure 2, this person was eventually deviating from the distribution of the cluster corresponding to females.

## 6. CONCLUSIONS

Gender discrimination is feasible using thermal images from the face of a person. This is possible since the thermal map of the face differs from males to females. This fact is evident if the temperature of the forehead is compared to that of other locations on the face. In this work sixteen persons were employed to carry out the experimental procedure. The feature selected was very simple since it consists of the mean value of the pixel in specific locations of the face. Examination of the feature space, as it was depicted in Figure 2, shows that there is a significant clustering tendency of the two gender populations. Consequently, the used features are potentially good candidates for gender classification.

In the experimental procedure it is proved that the clusters corresponding to the male and female subjects can be easily separated using a simple neural structure. Experimentally, various neural structures were tested starting from complex ones and concluding to the simplest with 3-2-1 neurons in the corresponding layers. This leads to the conclusion that the classification problem is easily manageable.

Furthermore, it was shown using the leave-one-out method that the classification can be successful for data for which the classifier has not been trained with. In all 16 cases the person excluded each time from the training procedure was correctly recognized by the corresponding neural structure. The data provided throughout the paper can help the reader to verify the results and compare with his approaches. The proposed gender classification procedure can be combined with other features derived from the persons' face, such as geometric and relative distance features in order to further improve gender classification procedures.

**REFERENCES**

- [1] Shangfei Wang, Zhen Gao, Shan He, Menghua He and Qiang Ji, (2016), "Gender recognition from visible and thermal infrared facial images", *Multimedia Tools and Applications*, Volume 75, Issue 14, pp. 8419–8442.
- [2] Georgia Koukiou and Vassilis Anastassopoulos, (2012), "Drunk person identification using thermal infrared images", *Int. J. Electronic Security and Digital Forensics*, Vol. 4, No. 4, pp. 229-243..
- [3] Georgia Koukiou, and Vassilis Anastassopoulos, (2016), "Drunk Person Screening using Eye Thermal Signatures", *J. Forensic Sci*, Vol. 61, No. 1, pp. 259-264.
- [4] Dat Tien Nguyen and Kang Ryoung Park, (2016), "Body-Based Gender Recognition Using Images from Visible and Thermal Cameras", *Sensors (Basel)*, 16(2), p.156.
- [5] Cunjian Chen and Arun RossLane, (2011), "Evaluation of Gender Classification Methods on Thermal and Near-infrared Face Images", *Proc. of International Joint Conference on Biometrics (IJCB)*, Washington DC, USA.
- [6] Xiaoguang Lu, Hong Chen and Anil K. Jain, (2006), "Multimodal Facial Gender and Ethnicity Identification", *International Conference on Biometrics, ICB 2006: Advances in Biometrics* pp. 554-561, Hong Kong, China.
- [7] Mohamed, Abouelenien., Verónica, Pérez-Rosas., Rada, Mihalcea., and Mihai, Burzo., (2017), "Multimodal gender detection", *19th International Conference on Multimodal Interaction, CMI'17*, November 13–17, Glasgow, UK.
- [8] Dat Tien. Nguyen., Ki Wan, Kim., Hyung, Gil Hong., Ja Hyung, Koo., Min Cheol, Kim. and Kang Ryoung, Park., (2017), "Gender Recognition from Human-Body Images Using Visible-Light and Thermal Camera Videos Based on a Convolutional Neural Network for Image Feature Extraction", *Sensors*, 17(3), 637; <https://doi.org/10.3390/s17030637>
- [9] Mahreen Ahmed, Asma Ghulam Rasool, Hammad Afzal, Imran Siddiqi, (2017), "Improving handwriting based gender classification using ensemble classifiers", *Expert Systems with Applications*, Volume 85, Pages 158-168.
- [10] Imran Siddiqi, Chawki Djeddi, Ahsen Raza, Labiba Souici-meslati, "Automatic analysis of handwriting for gender classification", *Pattern Analysis and Applications*, volume 18, pages 887–899 (2015).
- [11] Qian-Jie Fu, Sherol Chinchilla, John J. Galvin, (2004), "The Role of Spectral and Temporal Cues in Voice Gender Discrimination by Normal-Hearing Listeners and Cochlear Implant Users", *Journal of the Association for Research in Otolaryngology*, volume 5, pages 253–260.
- [12] J.W. Davis, Hui Gao, (2004), "Gender Recognition from Walking Movements using Adaptive Three-Mode PCA", : 2004 IEEE Conference on Computer Vision and Pattern Recognition Workshop, 27 June-2 July 2004, Washington, DC, USA, USA, DOI: 10.1109/CVPR.2004.354.
- [13] Crystal Huynh, Erica Brunelle, Lenka Halámková, Juliana Agudelo, Jan Halánek, (2015), "Forensic Identification of Gender from Fingerprints", *Anal. Chem.*, 87, 22, 11531–11536.
- [14] N.A. Diakides, J.D. Bronzino, (2008), *Medical Infrared Imaging*, Chapter 6: Physiology of Thermal Signals CRC Press, Taylor & Francis Group, New York.

**AUTHORS**

**Georgia Koukiou (B.Sc., M.Sc., Ph.D)** was born in Patras. She received the B.Sc. degree in Physics in 2004, the M.Sc. in Electronics and Computers in 2006 and her Ph.D. in Digital Image and Signal Processing in 2014 all from University of Patras. She is a member of the research team of Digital Image and Signal Processing of Electronics Laboratory, Physics Department in the University of Patras. She is currently working in Face Identification using Thermal Infrared, Biomedical imaging, Remote Sensing, SAR imaging, Ground Penetrating Radar methods and Information Fusion specializing in Decision Fusion methods. She has 20 publications in International Journals and Conferences.



Dr. Koukiou is a PostDoc. Scholar of IKY. She is currently giving lectures on Electronics, Programming, Signal and Image Processing, Pattern Recognition and Remote Sensing at University of Patras and University of Thessaly.

**Prof. Vassilis Anastassopoulos** was born in Patras, Greece, in 1958. He received the B.Sc. degree in Physics in 1980 and his Ph.D. in Electronics in 1986, both from the University of Patras, Greece. He worked for two years in Canadian Scientific Institutions. He is an academic staff member in Physics Department, University of Patras, Greece from 1987. He is a Professor from 2005. He was the Head of Physics Department, University of Patras, for the biennium 2001-2003 and 2005-2007, and the Vice Rector for Strategic Research Planning and Development in the University of Patras from 2006 to 2010.



His publication record contains over 100 journal and conference papers with over than 1500 citations. His research interests are within the scope of Digital Signal Processing, Image Processing, Radar Signal Processing and Remote Sensing with emphasis in Multi-spectral, SAR and Infrared Imagery, Signal Detection, Pattern Recognition with emphasis in Handwritten Analysis, and Information Fusion including Image Fusion, Decision Fusion and Sensor Fusion Architectures. Lately, he was involved in image processing techniques for Astro-particle Physics.



# BIOMEDICAL STUDY OF DEMOGRAPHICS AND CLINICAL FEATURES OF LICHEN PLANOPILARIS AMONG THE IRANIAN POPULATION

Maryamsadat Nejadghaderi<sup>1</sup>, Ashkan Tashk<sup>2</sup>, Parvin Mansouri<sup>1</sup>  
and Zahra Safaei Naraghi<sup>1</sup>

<sup>1</sup>Tehran University of Medical Sciences (TUMS), Tehran, Iran  
<sup>2</sup>Applied AI and Data Science Group, Mærsk McKinney Møller Institute  
(MMMI), University of Southern Denmark (SDU), Odense, Denmark

## ABSTRACT

**Introduction:** *The demographic of Lichen PlanoPilaris (LPP) among the Iranian population is unknown. The aim of this study is to describe the clinical, demographic, and histopathologic findings of lichen planopilaris in the Iranian population.*

**Method:** *In this cross-sectional study, all the patients with Lichen planopilaris were referred to the dermatology clinic of Imam Khomeini hospital from 2013 to 2015. Their demographic characteristics, drug histories, onset of disease, and family histories were obtained by written questionnaire. Additionally, this study employed SPSS v.20 as the statistical analysis software.*

**Results:** *One hundred patients were enrolled in this study. With an average age of 47.11 years, 78% of the patients were female, and 50 of these were housewives. The patients included were often from Tehran with Fars ethnicity. Among these patients, 7 had alopecia areata skin disease, and 10 of them suffered from thyroid disease. Most of the histopathology samples collected from these biopsies revealed degeneration of the basal layer of the follicular structure, perifollicular fibrosis, inflammatory cells, and atrophy of the pilosebaceous structures.*

**Conclusion:** *Both the age spectrum and the disease distribution of LPP among the Iranian population were very diverse when compared to previous studies.*

## KEYWORDS

*Lichen PlanoPilaris (LPP), Epidemiologic, Demographics, Clinical Features, histology.*

## 1. INTRODUCTION

Lichen planopilaris (LPP) is a follicular form of lichen planus, an autoimmune disease, which frequently presents as primary cicatricial alopecia and a rare inflammatory scalp disorder [1-4]. Lichen planopilaris is clinically characterized as hair loss, hyperkeratosis and perifollicular erythema [2-5]. The age of onset of LPP is between 40 and 60 years, and women are more affected by it than men [1, 6].

The incidence, prevalence, and physiopathology of LPP is unclear and unknown [7- 9]. North American tertiary hair research centers recommend that patients with existing lichen planopilaris be admitted to their centers to evaluate their demographic, epidemiologic, and clinical features

[10]. Demographic and Epidemiologic data on LPP in Iranian people are limited. The aim of this study is to describe the clinical, demographic, and histopathologic findings of lichen planopilaris among the Iranian population.

## 2. MATERIAL AND METHODS

This cross-sectional study was conducted on patients with the diagnosis of lichen planopilaris who were referred to the dermatology clinic of Tehran Imam Khomeini hospital affiliated by Tehran University of Medical Sciences (TUMS) between Sep 2013 and Sep 2015.

Lichen planopilaris was diagnosed by collecting histological evidence, dermatological examination, and clinical diagnosis. The inclusive criteria were lymphocytic cicatricial alopecia with perifollicular infiltrate. Patients without the histological confirmation were excluded.

Information about demographic characteristics, drug history, onset of disease, family history, the presenting complaint (signs and symptoms), hepatitis vaccination, consumption of gold or Dental amalgam, were obtained by written questionnaire.

This study was approved by the research Ethics Committee of the Tehran University of Medical Sciences, and the participants gave their voluntary participation consent at this study.

## 3. RESULTS

The whole statistical analysis related to this study are done by the means of SPSS v.20 (SPSS Inc, Chicago, IL, USA).

A total number of 118 patients were enrolled in this study, and of these 100 had the inclusive criteria and therefore were included in this study. Of these 100 patients, 78 (78%) were women and 22% were men. The average age was  $47.11 \pm 13.10$  years, with the youngest patient being 14 and the oldest 82. Ninety percent of the patients were married. The age of onset of disease was between 7 to 70 years, and the average age was  $41.19 \pm 13.03$ .

Table 1. Demographic characteristics

title	characteristic	N (%)
Smoking	Yes	4(4%)
	No	96(96%)
Ethnicity	Fars	88(88%)
	Turkic	8(8%)
	Kurd	3(3%)
	Lori	1(1%)
Education	Low literacy	12(12%)
	Diploma	48(48%)
	Bachelor	30(30%)
	Master degree	8(8%)
	Doctorate degree	2(2%)
Domicile	Tehran	73(73%)
	Other city	27(27%)
Occupation	Housekeeper	50(50%)
	Employee	19(19%)
	Self-employed	12(12%)
	Student	4(4%)
	Unemployed	7(7%)
	Another job	8(8%)

Fifty-one percent of the present patients had a history of stress, and 15% of the patients had a long-term duration history to sun exposure. Two of the patients had a history of radiation therapy (RT). In 19% of the patients, there was a positive history of LPP. Twenty-one percent of the patients had received a hepatitis vaccination before the onset of LPP, and none of them had a history of infectious disease. The Signs, symptoms, and medical history of the participating patients in this research are presented in table 2.

Table 2. Medical history

<b>Title</b>	<b>Characteristic</b>	<b>Frequency</b>
<b>History of another skin disease</b>	Psoriasis	3
	vitiligo	3
	Alopecia areata	7
	Alopecia androgenic	3
	Eczema	5
	Rosacea	3
<b>History of another disease</b>	Hypothyroidism	10
	Diabetes mellitus	8
	Lupus Erythematosus	2
	Rheumatism arthritis	1
	Behcet's disease	1
	Epilepsy	1
	Minor thalassemia	2
<b>Symptom</b>	Itching	69
	Local pain	6
	Tenderness	18
	Local irritation	14
	Hair loss	89
	Red spots	11
	Welding	5
<b>Sign</b>	Alopecia	91
	scar	25
	Erythema	59
	pigmentation	22
	scaling	25
	Frontal fibrosis	9
<b>LP in other organ</b>	Nail	10
	oral	22
	skin	58
	genital	10
<b>Type of treatment</b>	Topical	77
	Systemic	85
	Intralesional injection	47
<b>Disease features</b>	First LP then LPP	29
	First LPP then LP	10
	Only LPP	57
<b>Drug history</b>	Dental Gold	0
	Amalgam in dental	70
	Hair color	51
	NSAID	8
	propranolol	8
	Levothyroxine	10

Moreover, table 3 shows the frequency of characteristics related to the histopathology studies.

Table 3. Histopathology report

Characteristic	Frequency
Hyperkeratosis, follicular plugs and degeneration of basal layer	25
pigment incontinence in dermis	33
Degeneration of basal layer of follicular structure	98
Decreased density of hair follicles/Increased in catagen-telogen	87
Perifollicular fibrosis and inflammatory cells and atrophy of pilosebaceous structures	97
Fibrous-tract and stellae	67

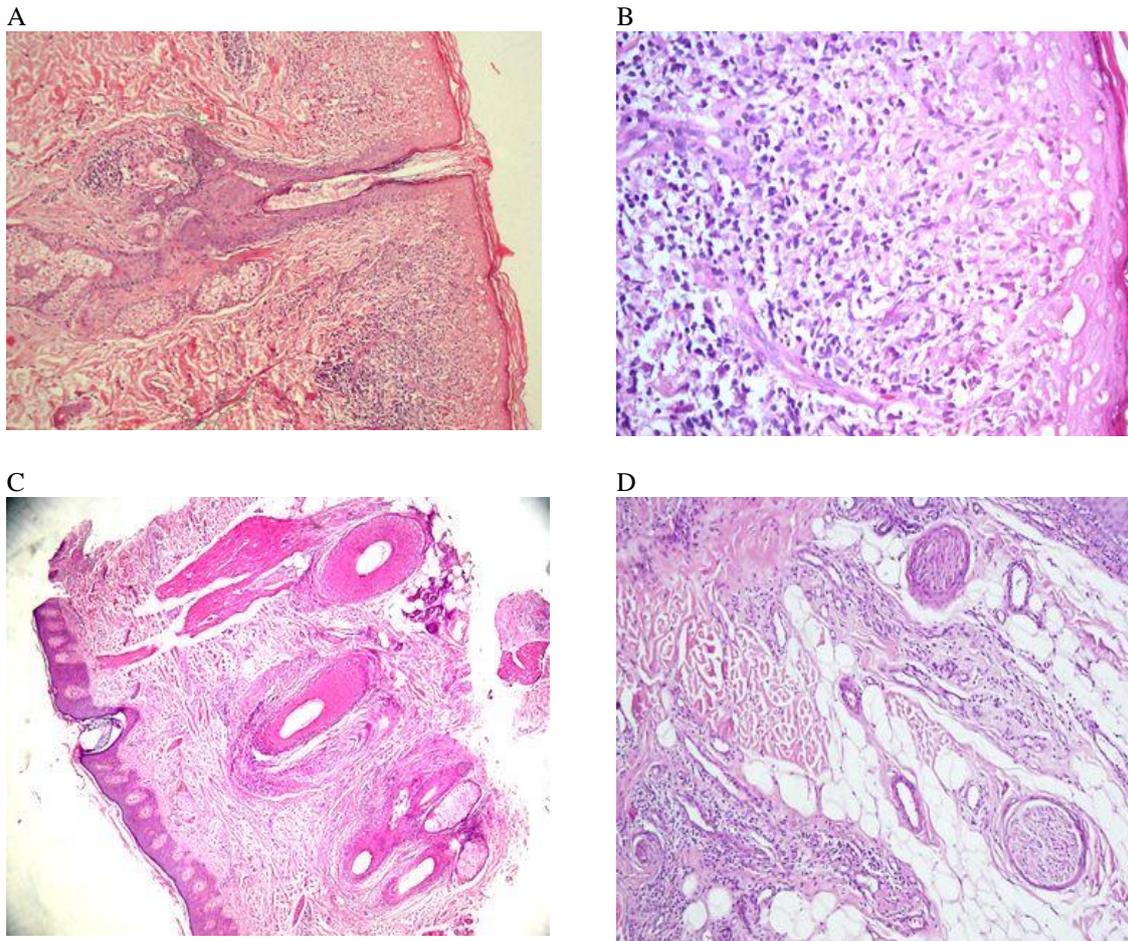


Figure 1. Histologic finding. haematoxylin and eosin (H&E) staining and optical microscope. A, B: interface Lichenoid infiltrate with epidermal atrophy Orthokeratosis, Follicular plugs and vacuolization of basal layer -there is perivascular and perifollicular lymphohistiocytic infiltration. LPP with epidermal involvement. C, D: Normal epidermis -Lichenoid reaction pattern involving the follicular epithelium with an association of perifollicular fibrosis and lymphocytic infiltration. Area loss of hair follicle, replaced by linear tract of fibrosis in deep dermis and fat tissue. LPP without epidermal involvement

#### 4. DISCUSSION

This current, cross-sectional study attempts to explain the demographics and Clinical Features of Lichen planopilaris in the Iranian population.

This study was compared to the results of similar studies, which discovered that usually young women are affected by LPP. These other studies demonstrated that LPP is a rare and uncommon inflammatory disease, which the results of this current study have also demonstrated within the Iranian population. Although the age spectrum for the LPP affected patients varied widely [11]: the ages of the patients participating in this study are between 14 and 82 years old. In the other studies, the range of ages for the LPP affected patients vary between 25 and 70 years old [12]. Both the age spectrum range and distribution in the Iranian population are very wide. In a retrospective study of 355 patients, where the mean age was 56 years old, the patients were between the ages of 21 to 81 [13]. In both this current study and other similar ones, Lichen planopilaris develops in association with lichen planus affecting the nails, skin, and mucosa, demonstrating similar results as the previously conducted studies [14, 15]. In the results of Mehregan et al.'s study, 45 patients, affected by Lichen planopilaris with an average age of onset of 52 years [16], participated. However, the average age of onset in this present study is 41 years.

The 10-year difference in the average age of onset of LPP for the Iranian population compared to the worldwide statistics is a very important factor which needs to be taken into consideration. In Mehregan et al.'s study, the female to male ratio was 4:1, which is similar to the ratio of the present study [16].

In Tan et al.'s study, the gender difference ratio among 25 of the patients with Lichen planopilaris was 1.8:1 [17]. In several studies, thyroid disease was significantly more frequent among the patients with Lichen planopilaris [4]. Hypothyroidism was the most common thyroid abnormality in LPP in this study, which is the same as the other studies [4]. In this study Diabetes mellitus was also common.

## 5. CONCLUSIONS

Due to the low prevalence of LPP among those included in this study, demographic and background information about the patients was collected, and their association with the disease was investigated. It is strongly recommended that further studies on the risk factors for patients newly diagnosed with LPP be done by academic referral centers.

This is also a case control study for patients whose information is in the database of the Skin Research Center, which conducts focused studies on LPP. This study has proven that the demographic criteria among the Iranian population are different, and in some cases early attention is urgent. The differences between the demographic factors can provide further studies and support, but more management models are recommended.

## ACKNOWLEDGEMENTS

The authors would like to thank Dr. Azita Khiltash, M.D. for her really constructive guidance and helps.

## REFERENCES

- [1] Ochoa BE, King LE Jr, Price VH. Lichen planopilaris: Annual incidence in four hair referral centers in the United States. *J Am Acad Dermatol* 2008; 58:352.
- [2] Vañó-Galván S, Molina-Ruiz AM, Serrano-Falcón C, et al. Frontal fibrosing alopecia: a multicenter review of 355 patients. *J Am Acad Dermatol* 2014; 70:670.

- [3] Tan KT, Messenger AG. Frontal fibrosing alopecia: clinical presentations and prognosis. *Br J Dermatol* 2009; 160:75.
- [4] Atanaskova Mesinkovska N, Brankov N, Piliang M, et al. Association of lichen planopilaris with thyroid disease: a retrospective case-control study. *J Am Acad Dermatol* 2014; 70:889.
- [5] Mobini N, Tam S, Kamino H. Possible role of the bulge region in the pathogenesis of inflammatory scarring alopecia: lichen planopilaris as the prototype. *J Cutan Pathol* 2005; 32:675.
- [6] Assouly P, Reygagne P. Lichen planopilaris: update on diagnosis and treatment. *Semin Cutan Med Surg* 2009; 28:3.
- [7] Tosti A, Piraccini BM, Iorizzo M, Misciali C. Frontal fibrosing alopecia in postmenopausal women. *J Am Acad Dermatol* 2005; 52:55.
- [8] Matta M, Kibbi AG, Khattar J, et al. Lichen planopilaris: a clinicopathologic study. *J Am Acad Dermatol* 1990; 22:594.
- [9] Giménez-García R, Lázaro-Cantalejo TE, Sánchez-Ramón S, Velasco Fernandez C. Linear lichen planopilaris of the face. *J Eur Acad Dermatol Venereol* 2005; 19:770.
- [10] Duque-Estrada B, Tamler C, Sodr e CT, et al. Dermoscopy patterns of cicatricial alopecia resulting from discoid lupus erythematosus and lichen planopilaris. *An Bras Dermatol* 2010; 85:179.
- [11] Ioffreda MD. Inflammatory diseases of hair follicles, sweat glands, and cartilage. In: *Lever's Histopathology of the Skin*, 10th ed, Elder DE (Ed), Lippincott Williams & Wilkins, Philadelphia 2009. p.459.
- [12] R acz E, Gho C, Moorman PW, et al. Treatment of frontal fibrosing alopecia and lichen planopilaris: a systematic review. *J Eur Acad Dermatol Venereol* 2013; 27:1461.
- [13] Chiang C, Sah D, Cho BK, et al. Hydroxychloroquine and lichen planopilaris: efficacy and introduction of Lichen Planopilaris Activity Index scoring system. *J Am Acad Dermatol* 2010; 62:387.
- [14] Spencer LA, Hawryluk EB, English JC 3rd. Lichen planopilaris: retrospective study and stepwise therapeutic approach. *Arch Dermatol* 2009; 145:333.
- [15] Baibergenova A, Walsh S. Use of pioglitazone in patients with lichen planopilaris. *J Cutan Med Surg* 2012; 16:97.
- [16] Mehregan DA, Van Hale HM, Muller SA. Lichen planopilaris: clinical and pathologic study of forty-five patients. *J Am Acad Dermatol* 1992; 27:935.
- [17] Tan E, Martinka M, Ball N, Shapiro J. Primary cicatricial alopecias: clinicopathology of 112 cases. *J Am Acad Dermatol* 2004; 50:25.

**AUTHORS**

Maryamsadat Nejadghaderi is a medical doctor (M.D.) from Tehran University of Medical Sciences, Iran.



Ashkan Tashk is a Ph.D. of Electrical engineering and is currently working as a postdoc researcher at SDU in .



Parvin Mansouri (M.D. Professor) is currently working as a Professor of Dermatology in Tehran University of Medical Sciences, Iran.



Zahra Safaei Naraghi is currently working as a Fellow in Dermatopathology in Razi-Skin-Hospital, Tehran University of Medical Sciences, Iran.





# TWO EVOLUTIONARY HYBRID STAGES FOR THE RECTANGULAR BIN PACKING PROBLEM WITH CONSTRAINTS

Slimane Abou-Msabah<sup>1</sup>, Ahmed Riadh Baba-Ali<sup>2</sup> and Basma Sagr<sup>1</sup>

<sup>1</sup>Department of Computer Science, University of Science and Technology Houari Boumedienne, USTHB, Bab Ezzouar, Algiers, Algeria

<sup>2</sup>Department of Electronics, University of Science and Technology Houari Boumedienne, USTHB, Bab Ezzouar, Algiers, Algeria

## ABSTRACT

*The Bin Packing problem is met in several domains of application, especially in the industry of: sheet metal, wood, glass, paper etc. In this article we are interested to the orthogonal cutting problem, with the hold in charge of the constraint of end-to-end cutting, and orientation constraint. The bin packing problem belongs to the class of NP-hard problems; therefore, our work has turned towards heuristic methods of resolution, and more particularly evolutionary methods. The application of genetic algorithms that are part of the evolutionary methods has limitations for solving the bin packing problem with large data. To minimize this disadvantage, we propose an original method which consists in subdividing the initial problem into two sub-problems. The first step tries to apply a hybrid genetic algorithm based on the order of appearance of pieces, to be packed on levels in an infinite band by applying the new placement routine (BLF2G). The second step uses the results of the first, namely the levels, and tries to project them on Bins by applying a second hybrid genetic algorithm. Besides that, we propose a new definition of the problem, it's about seeing the strip not as usual, with a fixed width and infinite height, but with a fixed height and infinite width. And we must apply some improvements, found in the literature, to the classic genetic algorithm to improve results, by introducing greedy heuristics to the population. Results are compared with other heuristic methods on data sets found in the literature.*

## KEYWORDS

*Rectangular Bin Packing, Orthogonal Cutting stock, Combinatorial Optimization, Heuristic, Hybrid Genetic Algorithm.*

## 1. INTRODUCTION

The problem of placement is an optimization problem whose objective is to find a good arrangement of different objects in other wider ones. The main objective is to maximize the exploitation of the raw material, and therefore to minimize losses. It is important for industries of massive production where the optimization of the raw material plays an important role in reducing manufacturing costs. The development of an algorithm for the resolution of an industrial investment problem must take in consideration the complexity of the problem determined by the shape of objects manipulated, and of the bound constraints (imposed by the system of production). In our case we take into consideration an orthogonal cutting problem of which handles raw rectangular bins to generate rectangular items as well. The material used can be sheet metal and the production machines are typically guillotine shears (end-to-end cutting). In order to generalize our approach, we take into consideration a constraint often found in the

literature [1], it is about imposing the orientation of the objects. Therefore, our algorithm will be adapted for a wide range of applications, such as wood, glass, fabric, and in the layout of newspapers, magazines, or web pages, ... where there are motifs and decorations for which the orientation is fixed.

## 2. CONSTRAINTS OF THE PROBLEM

### 2.1. The orthogonal cutting-up

Cuts are made in length or width. Diagonal cuts are not feasible, see figure 1.

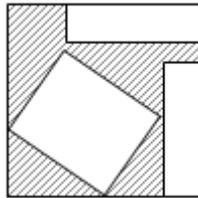


Figure 1. Non feasible shape by the shears

### 2.2. The end-to-end cutting

This constraint is directly related to the guillotine shears, which are the typical machines for handling sheet metal considered by our approach. This is a constraint that ensures the feasibility of cutting formats in the production facility, see figure 2.

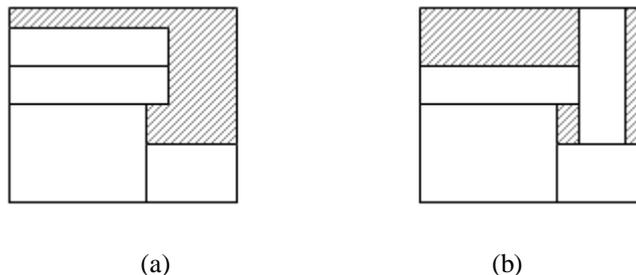


Figure 2. (a) feasible Shape, (b) impracticable Shape

### 2.3. The orientation of pieces

The pieces retain their original orientation, this constraint aims to ensure the feasibility of cutting formats made on textured or decorated surfaces.

## 3. STATE OF THE ART

Few works have used genetic algorithms for the resolution of this problem, recently E. Hopper & B. Turton has been interested, in the application of genetic algorithms for the resolution of this problem. After a state of the art in 1997 that lists the work done in this field [2], they present two hybridized genetic algorithms with placement heuristics in 1999 [3]. This work is crowned by a doctoral thesis in philosophy in 2000 [4], other work has referred to genetic algorithms for solving the problem we can quote [5], [6], [1], [7] and more recently [8] ...

Recently a new tendency has tried to guide the genetic algorithm to improve the results by modifying the evolutionary process, [7] [8].

A common disadvantage confirmed by the authors, ([4], [1], [7], [8] ...), encountered by genetic algorithms, is the problem of the resolution time which becomes increasingly large for big sized problems. It reduces the effectiveness of the genetic algorithms applied to the placement problem for which they give significantly less good results than simple heuristics. Our resolution approach attempts to minimize this disadvantage through the improvements we have proposed to the classical evolutionary process.

#### 4. THE GENETIC ALGORITHM

Generally, evolutionary algorithms developed for the problem of layout, operate a hybridization of the genetic algorithm with a placement heuristic [3]. The genetic algorithm explores the research space to generate individuals. A second non-genetic step is needed to assess the quality of these individuals. The genetic algorithm offers solutions based on the order of appearance of the pieces for the arrangement process, the exact cut-out format is then given by the placement routine.

The quality of an individual depends essentially on the order in which the pieces are presented to the adopted placement routine. The genetic algorithm is the most effective research strategy [3], using order-based representation. This representation requires a new definition of crossing and mutation operators, which are identical to those of sales traveler problem.

##### 4.1. The operator of crossing:

We propose to apply a hybrid genetic algorithm based on the order of appearance of the pieces, using the Partial Mapped Crossover (PMX) [9], given that the PMX crossover is the most effective for order-based representations [10] [3].

However, the conventional PMX crossover operator may give birth to individuals that are not feasible, for example:

<b>parent 1 : 1234 5678</b>		<b>son1 : 1234 4321</b>
and	gives	and
<b>parent 2 : 8765 4321</b>		<b>son2 : 8765 5678</b>

These two new individuals do not represent an attainable order. To do this, we need to modify the two individual sons so that they are feasible. For each son the PMX crossing replaces the double genes with genes missing in the individual according to the order of their appearance in the other parent individual.

We will have:

<b>son 1 :1234 8765</b>
and
<b>son 2 :8756 1234</b>

##### 4.2. The operator of mutation:

The mutation operator can simply be presented as a permutation. Just randomly select two sites in the individual and make the swap between these two sites.

Exemple : 81|67432|5 became s: 81|57432|6

## 5. OUR CONTRIBUTION

Our contribution to solving the problem can be summarized in the following improvements:

### 5.1. The first stage «pre-processing»

The quality of the individuals produced by the genetic algorithm is measured after the arrangement of the pieces according to a placement routine. The pieces are arranged on an infinite band subdivided into levels [1], as shown in figure 3.

The arrangement is done by applying an adequate heuristic, it is about the Bottom Left Fill (BLF) heuristics, which tries to place the pieces in the lowest possible place on the left [3]. This phase allows us to realize the effective arrangement of the pieces on the strip structured in levels, moreover, it gives us a value that is the height of the strip. This value defines the quality of the solution under consideration.

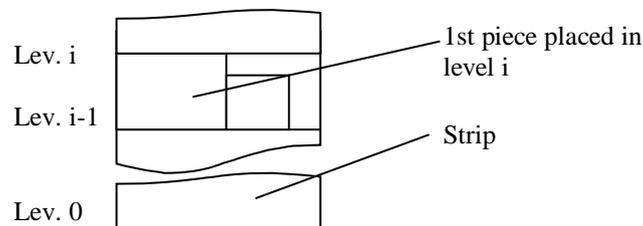


Figure 3. The strip is constituted of several levels.

In order to make better use of the space, the new BLF2G placement routine will be adopted, which carries out vertical and horizontal exploitation of intra-level residues, and which checked the end-to-end cutting constraint [7], see figure 4.

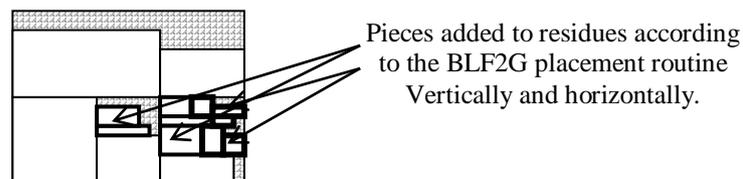


Figure 4. Filling the gaps under BLF2G Policy

### 5.2. The second phase « Packing »

It is now a question of projecting the levels obtained previously into bins of known dimensions. For each individual in the population, the placement routine is applied to assess his quality (fitness). Recall that the levels have an identical width and varied length, the heuristics BL (Bottom Left), which places the current level in the first plate offering a sufficient residue is the most adequate, so it is applied in this phase.

### 5.3. Guide the genetic algorithm

Several studies have found that by applying genetic algorithms to medium and large placement problems, heuristic methods, based on sorting, offer better results. [7] proposes to guide the

genetic algorithm by adding an individual sorted to the population, thus the method takes over from the heuristics, and gives results equal and even better than the conventional heuristic methods regardless of the size of the problem.

Recently, a new technique has been proposed to guide the genetic algorithm by introducing, with control, individuals into the population, it's about csGA. It's a question of controlling the evolution of the genetic phase, and when there is no improvement, a sorted individual is injected into the population, the author finds that the introduction of several sorted individuals can disrupt the genetic process and thus degrade the results [8].

For our case these improvements will be taken into consideration to take advantage of the benefits.

#### 5.4. Optimization in Width OW

In some cases, the gap between the lengths of the pieces is high and the results are therefore unsatisfactory. Our second improvement will therefore be to arrange the pieces on the strip not in length but also in width, that is by fixing the length of the strip, and by arranging the pieces in width.

Figure 5 illustrates the width arrangement, using the following example:

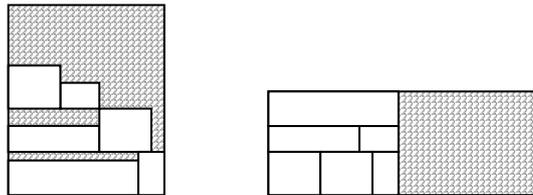


Figure 5. Result of arrangement of a set of pieces in length and width.

We note the contribution of this improvement for this set of pieces. Note that the pieces keep their original orientations, only the band changes direction and the arrangement is made in width.

With this enhancement, the user retains the choice between length and width layout depending on the pieces set used, giving the user more flexibility, as shown in figure 6.

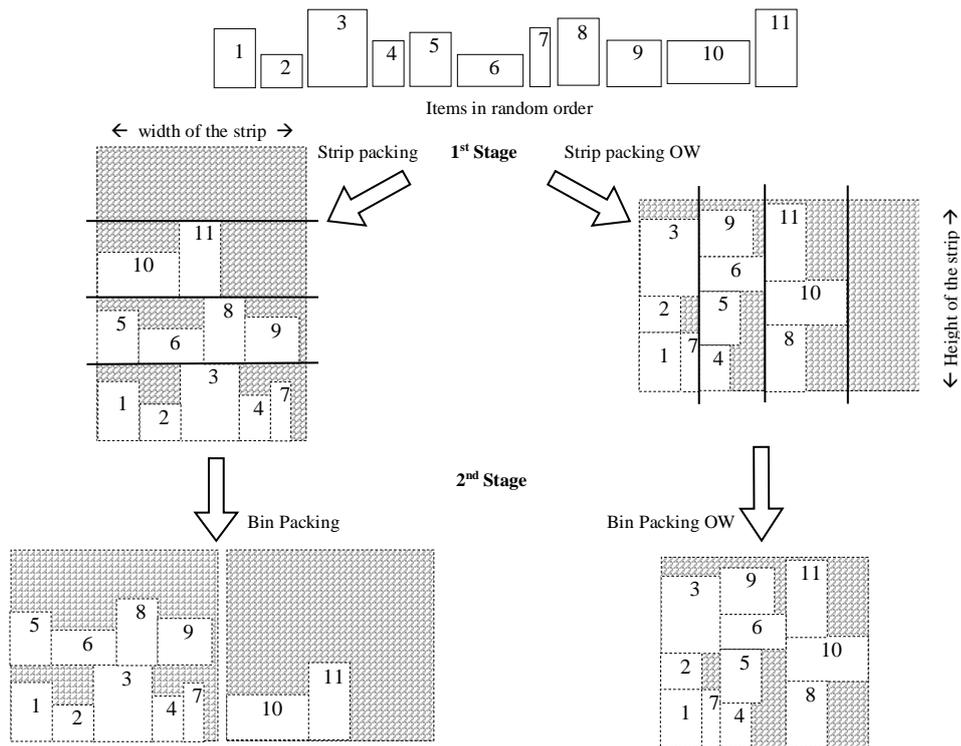


Figure 6. Our Approach 1<sup>st</sup> and 2<sup>nd</sup> stages: Classical Strip packing and Bin Packing vs. Optimization width Strip packing and Bin Packing

## 6. RESULTS

To evaluate our method, we chose a large number of examples found in the literature, for which we know the optimal solution, described in table 1:

Table 1. Data sets.

Reference	Name	Size	Optimal
[11]	Kendall	13	140 x 80
[12]	D2	21	40 x 60
[13]	J1	25	40 x 15
	J2	50	40 x 15
[3]	T1a, T1b, ... T1e	17	200 x 200
	T2a, T2b, ... T2e	25	200 x 200
	T3a, T3b, ... T3e	29	200 x 200
	T4a, T4b, ... T4e	49	200 x 200
	T5a, T5b, ... T5e	73	200 x 200
	T6a, T6b, ... T6e	97	200 x 200
	T7a, T7b, ... T7e	197	200 x 200
[7]	Msa17a, b, c	17	200 x 200
	Msa35a, b, c	35	200 x 200
	Msa75a, b, c	75	200 x 200
	Msa150a, b, c	150	200 x 200

The results obtained with these examples using the csGA = csGA<sub>OH</sub> (Optimization in Height) and csGA<sub>OW</sub> (Optimization in Width) are summarized in table2. Note that the test sets were made in length (i.e. the width arrangement does not guarantee the optimal).

Table 2. Results csGS vs csGA<sub>ow</sub>.

Name	Instance N	W	csGA = csGA <sub>OH</sub>		csGA <sub>ow</sub>	
			Fitness/Percentage	Fitness/Percentage	Fitness/Percentage	Fitness/Percentage
Msa17a	17	200 x 200	<b>200x200</b>	<b>0%</b>	200x222	11%
Msa17b	17	200 x 200	<b>200x200</b>	<b>0%</b>	200x222	11%
Msa17c	17	200 x 200	<b>200x200</b>	<b>0%</b>	200x222	11%
Msa35a	35	200 x 200	<b>200x200</b>	<b>0%</b>	200x221	10.5%
Msa35b	35	200 x 200	210x200	5%	200x213	6.5%
Msa35c	35	200 x 200	213x200	6.5%	200x211	5.5%
Msa75a	75	200 x 200	205x200	2.5%	200x212	6%
Msa75b	75	200 x 200	205x200	2.5%	200x211	5.5%
Msa75c	75	200 x 200	208x200	4%	200x210	5%
Msa150a	150	200 x 200	205x200	2.5%	200x211	5.5%
Msa150b	150	200 x 200	<b>200x200</b>	<b>0%</b>	200x210	4%
Msa150c	150	200 x 200	210x200	5%	200x209	4.5%
10xMsa17	170	2000 x 200	2106x200	5.3%	200x2101	5.05%
Kendall	13	140 x 80	162x80	15.7%	140x102	27.5%
D2	21	40 x 60	<b>40x60</b>	<b>0%</b>	<b>40x60</b>	<b>0%</b>
J1	25	40 x 15	43x15	7.5%	40x18	20%
J2	50	40 x 15	42x15	5%	40x16	6.67%
T1a	17	200 x 200	237x200	18.5%	200x231	15.5%
T3a	29	200 x 200	229x200	14.5%	200x232	16%
T5a	73	200 x 200	217x200	8.5%	200x215	7.5%
T7a	197	200 x 200	212x200	6%	200x210	5%
Optimum reached:			6		1	
Best results:			15		7	

These results show the power of our new method csGA<sub>OW</sub>, we manage to give best results 7 times, and we reach the optimal 1 time, despite the data set are confectioned in length. With our method we give a second alternative to the classical csGA method, and we improve the result by 33.33%.

## 7. CONCLUSION

Our contribution, to the cutting problem, showed its efficiency. Firstly, we use a powerful placement routine (BLF2G) that verify the guillotine constraint. Secondly, we hybridized this routine with an improved genetic algorithm, in a first stage, using existed improvements, that surmounts problems met by the classic genetic algorithms, for problems of large size. After the 1st stage is done, we generate an open-end strip with pieces packed in levels. These levels are used by a 2nd stage hybrid genetic algorithm to be packed on bins.

A new and powerful technic is applied in the 1st stage, by seeing the strip horizontally with a fixed height and infinite width, not vertically as usual, with a fixed width and infinite height. With this improvement we can see the same problem in the same context with different parameters, vertical placement, and horizontal placement. The results show that for some cases

the width placement is more efficient than height placement, and we have an improvement of 33.33%, compared with the standard vertical placement.

## 8. FUTURE WORK

The genetic algorithm is a powerful tool that explore the research space to find the best solution, using genetic operators. We find in the literature that the genetic algorithm knows limits in application to problem of middle and large size. In perspective, and seen the clear improvement brought by the introduction of the sorted individual to the population, the development of a heuristic method that serves to guide the genetic algorithm all along the genetic process since the initial population until the genetic operators, remain to explore.

## REFERENCES

- [1] Valenzuela, C. L., & Wang, P. Y. (2001). Heuristics for large strip packing problems with guillotine patterns: An empirical study. *Proceedings of the 4th Metaheuristics International Conference*, 417–421.
- [2] Hopper, E., & Turton, B. (1998). Application of genetic algorithms to packing problems—a review. In *Soft Computing in Engineering Design and Manufacturing* (pp. 279–288). Springer.
- [3] Hopper, E., & Turton, B. (1999). A genetic algorithm for a 2D industrial packing problem. *Computers & Industrial Engineering*, 37(1–2), 375–378.
- [4] Hopper, E. (2000). Two-dimensional packing utilising evolutionary algorithms and other meta-heuristic methods (Doctoral dissertation, University of Wales. Cardiff).
- [5] Fujita, K., Akagi, S., Hirokawa, N., & others. (1993). Hybrid approach for optimal nesting using a genetic algorithm and a local minimization algorithm. *Proceedings of the 19th Annual ASME Design Automation Conference*, 1, 477–484.
- [6] Raidl, G. R., & Kodydek, G. (1998). Genetic algorithms for the multiple container packing problem. *International Conference on Parallel Problem Solving from Nature*, 875–884.
- [7] Msabah, S. A., & Baba-Ali, A. R. (2011). A new guillotine placement heuristic combined with an improved genetic algorithm for the orthogonal cutting-stock problem. *2011 IEEE International Conference on Industrial Engineering and Engineering Management*, 482–486.
- [8] Abou-Msabah, S., Baba-Ali, A.-R., & Sager, B. (2019). A Controlled Stability Genetic Algorithm With the New BLF2G Guillotine Placement Heuristic for the Orthogonal Cutting-Stock Problem. *International Journal of Cognitive Informatics and Natural Intelligence (IJCINI)*, 13(4), 91–111.
- [9] David, E. G. (1989). *Genetic algorithms in search. Optimization and Machine Learning*, Reading, Massachusetts.
- [10] Michalewicz, Z. (2013). *Genetic algorithms+ data structures= evolution programs*. Springer Science & Business Media.
- [11] Burke, E., & Kendall, G. (1999). Comparison of meta-heuristic algorithms for clustering rectangles. *Computers & Industrial Engineering*, 37(1–2), 383–386.
- [12] Ratanapan, K., & Dagli, C. H. (1997). An object-based evolutionary algorithm for solving rectangular piece nesting problems. *1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation*, 2, 989–994.
- [13] Jakobs, S. (1996). On genetic algorithms for the packing of polygons. *European Journal of Operational Research*, 88(1), 165–181.

**AUTHORS**

**Slimane Abou-Msabab** PHD Student last year, Department of computer science, University of Science and Technology Houari Boumedienne, USTHB. His researches revolve around the application of artificial intelligence methods to solve problems with combinatorial explosion, using greedy heuristics, metaheuristics, and hyperheuristics, to find near optimum solutions. Especially the Bin Packing Problem, using Evolutionary Approaches.



**Ahmed-Riadh Baba-Ali** Professor and director of research within the Faculty of Electronics and computer science in the University of Sciences and Technology of Algiers. His research integrates fast algorithm development applied to several application areas such as CAD, data mining, machine learning cryptography. He is also interested in parallel algorithms development such as evolutionary algorithms.

**Basma Sager** Computer Engineer, Department of computer science, University of Science and Technology Houari Boumedienne, USTHB. Her research turns around applying Evolutionary Approaches to solve Np-Hard problems.

© 2020 By AIRCC Publishing Corporation. This article is published under the Creative Commons Attribution (CC BY) license.



# EFFECTIVE TEACHING STRATEGIES IN AN ONLINE ENVIRONMENT

Keith Buckley

Director of Physical Education Rollins College (USA)  
Winter Park Florida, USA

## **ABSTRACT**

*The online classroom is a powerful teaching and learning arena in which new practices and new relationships can make significant contributions to learning. Instructors must be trained not only to use technology, but also to shift the ways in which they organize and deliver material. Making this shift can increase the potential for learners to take charge of their own learning process and facilitate the development of a sense of community. In constructing an online course, the instructor must take into account: designing the course, implementing content, facilitating learning, relevant assignments and course evaluation. A learner-centered approach acknowledges what students bring to the online classroom, their background, needs, and interests, and what they take away as relevant and meaningful outcomes. This paper will explore how to design implement and instruct a thorough online course that engages the student and gives them a platform to learn and enhance their academic experience.*

## **KEYWORDS:**

*Online Learning, Education, Teaching*

## **1. INTRODUCTION**

In order to create an effective online course, it is very important to understand that it is for the benefit of the student and all design implementation and evaluation should be done that is in the best interests of the student. Online Learning is a mixture of learning methods that incorporate multiple teaching models. It is a natural development to the growing accessibility of eLearning, online resources and the continued need for a human component in the learning experience. An online learning approach ensures that the learner is engaged and driving his or her individual learning experience. This approach also helps cater to the individual needs of the learner; most students have unique learning styles and an online approach is more likely to cater to those needs than a traditional classroom teaching experience. Online learning is important because it breaks down the traditional walls of teaching, ones that don't work for all students and now with access to present day technologies and resources we can tailor the learning experience for each student. Online learning also offers flexible time frames that can be personalized to each person, offering them the ability to learn at their own pace. For universities, online courses can be part of a strategy to compensate for limited classroom space, as well as a way to think differently about encouraging faculty collaboration. For faculty, online courses can be a method to infuse new engagement opportunities into established courses or, for some, provide a transitional opportunity between fully face-to-face and fully online instruction. For students the conveniences of online learning combined plus their social and instructional interactions that may not lend themselves to face to face instruction. All students no matter their age learn differently, and teaching methods

should reflect this, by designing teaching programs in a way that reaches visual, auditory and kinetic learners alike. With the heavy integration of technologies, we'll be able to improve teaching, information retention, engagement, responsibility and enjoyment. Students never outgrow their learning styles, meaning online learning is more important than ever, no matter what the industry is, from schools to corporations, in all walks of life.

When setting the course there are five basic tenants of a complete course: Interactions, Design, Content, Assessment and Evaluation. Each one is extremely important and leads the student to have well defined, organized and pedagogically centered experience.

### Background

When observing the literature on online learning, it is apparent that results have not meet expectations of student learning. Administrators are much more excited than educators in the benefits of Online learning. A 2014 department of education study reported: "The proportion of chief academic leaders reporting online learning is critical to their long-term strategy reached a new high of 70.8 percent. At the same time, only 28 percent of academic leaders say that their faculty accepts the 'value and legitimacy of online education.' Online education is the fastest-growing segment of higher education and its growth is overrepresented in the for-profit sector. Despite this faculty and academic leaders, employers and the general public are skeptical about the quality and value of online education, which they view as inferior to face-to-face education. Students in online education, particularly underprepared and disadvantaged students, have consistently underperformed and on average experience poor outcomes. Very surprisingly Online education has failed to improve affordability, frequently costs more than in-person alternatives and has not produced a positive return on investment. Regular and substantive student-instructor interactivity is a key determinant of quality in online education, leading to improved student satisfaction, learning and outcomes.

## 2. INTERACTIONS

This area clearly explains to the student all the necessary information about times, assignment's, expectations and protocols. In order to learn any academic subject students, need to interact and even struggle with the material being learned (called active learning). Pre-classroom assignments are set, and students are asked to explain concepts and ideas to their peers in class. The online component allows students to work at their own speed, but it should force them to research, reflect and discuss complex issues within the course. The assigned work should tie in with the course goals and objectives.

When developing the course schedule topics and assignment should be distributed evenly. The course schedule also helps students know from Day 1 the exam and due dates, as well as plan their reading to be prepared for online discussions. The official syllabus of most schools is very long with details and policies as required by institutional policy. A way to help the students is to create a mini syllabus to provide quick access to the most important aspects of the course. The syllabus should outline objectives and learning outcomes. It should give students resources for disabilities, technology help and relevant institutional information.

Protocols is a document detailing course expectations and norms. In class preparation, E mail correspondence and online discussions are all discussed. This is where all technical support such as log in procedure and course requirements are stated such as student responsibilities in the classroom and online are stated. Also, the consequences of late/missed assignments. The module interaction allows the alignment of the unit learning objectives with assessment and interaction

activities. The teacher has to carefully consider how to design each week of the course, combining different content resources with faculty-student and student-student interaction activities that foster a sense of community and critical thinking, with formal and informal formative assessments.

There are of course many different technological tools. It is vital that the teacher designs the course around technology that he/she is comfortable with and will aid the student in their learning process and not hinder it. Tools such as Google Docs, Edpuzzle, Padlett and Kahoot. Kahoot is especially effective as you can test students individually or in teams that compete against each other in their comprehension of the class that day. Each of the tools listed has its own functionalities, but all are beneficial when it comes to one crucial aspect of modern schooling: they pave the way for re-examination of the current learning methodologies. These tools open up the space for dialogue on the potential that technological devices can bring to today's learning environments.

In the majority of classes, interaction across all forms is vital. This gives the student a sense of belonging and that the course is important. The more effective interaction the more motivated the student will most likely be. Positive interaction will enhance the learning environment in all kinds of courses. The benefits of an online format allow students to work at their own speed and time in the online format and therefore in the classroom this can maximize in class discussion. This gives students the ability to understand other student's viewpoints and ideas. The teacher's responsibility is to facilitate this discussion and make it relevant to the modules of the course

### **3. DESIGN**

Successful online courses start with effective planning for overall design and outcomes. By identifying your course goals and learning objectives, you achieve a clearer picture of how to best design a structured and engaging course. A strong understanding of the needs of your learners and how your course can meet those needs is important to the success of your course delivery. Therefore, it is essential to consider what you want your student outcomes to be at the conclusion of your online course.

It is obviously important to have an effective design that makes learning intuitive and stimulating. The quote "The focus should be first on the learning, and second on the technologies that will support that learning" is very profound. Different schools have various ways of approaching online learning and the design of the course must fit the mission of the institution. The course should focus on teaching and learning in a very collaborative and informative manner. Discussion and debate are very important parts of the curriculum. If an institution's online learning strategy can be designed to address the needs and dynamics of all three constituencies (institution, faculty, and student) simultaneously, then online learning can become a powerful force for institutional transformation. With the right design and strong learning objectives the students will embrace online learning and it will be a major addition to the curriculum.

Course objectives, instructional strategies, content, and assignments should facilitate learners to attain the course goals. The instructor needs to understand the characteristics of the students in terms of what types of learning are involved. Also, what content and information is needed to facilitate learning, and what tasks do learners need to master to achieve the overall course goals. The instructor needs to have a logical sequence of content and activities, and an awareness of what technologies are available and best suited to present the content to facilitate learning. Functioning as guideposts, learning objectives help students organize their efforts toward accomplishing the desired behaviors. Learning objectives also help the instructor identify whether students have gained the appropriate skills and knowledge. A learning objective is a

statement that: specifies in measurable terms what a learner will be able to do as a result of your instruction; describes the intended outcome of the course rather than a description or summary of the content and details the intended results rather than the means of achieving the results.

Enhancing the course with learning objects and activities will help to communicate course content beyond static readings and lectures. Content presented in a memorable and meaningful way will stimulate the learner making the course more dynamic and successful.

#### **4. CONTENT**

It is vital to stay current with effective online teaching strategies for designing, implementing, and facilitating the course. As the author of an online course, the course content may include text, music, graphics, illustrations, articles, photographs, etc. Some of the content you wish to use may be protected by copyright law. Implementing content needs to be predictable. Write detailed, sensible policies and consistent information to help create predictability from the student's perspective. The content should be organized to provide students guidance about how to use the course management tools effectively (e.g., threaded discussions, assignments, quizzes, etc.). The instructor should be prepared to know the syllabus, resources, and learner support systems to respond to students' questions timely and completely. Resourcefulness and planning are important to identify alternative plans for delivering the course in the event of technology issues or modifying activities if students are not "getting it". Lastly the Instructor needs to be responsive to counteract the negative aspects of the distance separating the instructor and the students. This can be done by monitoring the online courses frequently and responding to student concerns in a timely manner.

#### **5. FACILITATING LEARNING**

In an online classroom, flipping can work in a number of ways. Consider, for example, assigning students to complete a particular reading or lesson. An online discussion can begin where learners are encouraged to delve deeper into the lesson, pose questions, and share their own unique interpretations. This method creates a more diversified learning environment and fosters a deeper understanding of the subject matter being discussed. Mind mapping is probably one of the best online teaching techniques to implement for the visual learners within a class. A mind map is a diagram of related ideas and concepts that can be used as an aid for studying, a way of organizing information or even a springboard for a writing assignment. This helps instructors in structuring classroom discussions, classifying ideas and gradually bringing learners to the central idea. The aim is to uncomplicate complex concepts or issues. The benefits for learners are that it's quick, easy and allows them to "dive right in" the ideas around a central concept and connecting the dots to reach the central idea. In an online course, you can have students create their own mind maps (either on paper or using online mapping software) and share them with the rest of the class, allowing for easier sharing of ideas and interpretations. Mind mapping is ideal for making connections between ideas/concepts, planning out projects or written assignments and a better understanding of the learning material. Yet another concept to implement is that of promoting self-learning in a controlled environment. With self-learning, students are encouraged to explore certain subject matter and decide what aspects are most important or relevant to their own interests. Self-learning is a great way to get learners truly invested in the subject matter in ways that apply to real-life situations. The concept of instructional design is a teaching technique that refers to designing your classroom around your learners' unique backgrounds and your ultimate goals (or what you want students to take away from the course). The technique is not restricted for the use in on-site setting, rather it's equally important in an online learning setting, where

learners have very unique backgrounds and bring a different set of experiences to the table (especially when compared with “traditional” face-to-face classrooms).

As such, instructional design means taking the time to get a better understanding of the learners. What is their current understanding of the subject matter? What are their different learning styles? In simplest terms, instructional design recognizes that there’s no “one-size-fits-all” approach to teaching. Adaptive learning involves utilizing computers and other technologies as viable teaching devices. As an online instructor, adaptive learning will inherently come into play, but it’s up to the teacher to make the most of the technology available to themselves and their students. For some students, learning a particular subject by watching a video may be most effective. Others may learn better through participating in online discussion forums. Ultimately, it’s up to the teacher to transform each student from a passive receptor of information into a willing and active participant in the online classroom. This may mean providing different mediums (video, text, visuals, etc.) for introducing each new lesson or concept.

## **6. ASSESSMENT**

It is crucial to provide clear instructions for all assignments, as well as exam objectives, to help students focus, have a clear understanding of expectations and how the assignments help fulfill course objectives. Written, clear instructions also help faculty minimize students' questions about what to do, when is the assignment due or how is it going to be marked. Part of the process is developing a scoring rubric. This can be a tedious, but it helps the teacher mark assignments consistently and students know how they're going to be evaluated and why they lost points. Formal and informal formative and summative assessment are extremely important in the development of a course. Online assessments are quite personal, and the student gets from it the level of work that they put in. In face to face classroom assessments the student’s personalities are much more evident but sometimes one or two verbally strong students can dictate the class. Carefully assigned projects and discussion boards can create a much fairer and less obtrusive form of class participation that is transparent and equal.

## **7. EVALUATION**

How will the instructor know if the course met the learning goals and was effective? Many evaluations are done right at the end of the class/semester when students are more concerned with finals as opposed to a serious reflection of the class. Like many things as long as we grow in our learning each class should be more productive than the last and by building on a foundation of good teaching practices this should occur. It is advisable to have a colleague evaluate the formatting and design of the course and make comments. It is not easy to share with your peers and the teacher has to be able to take constructive criticism, but it is a very valuable tool. The class needs to be intuitive to a student, many teachers just assume the students understand what the teacher is thinking. To determine the quality of the course and its levels of success is hard to define in a learning environment. Success may be having a class that those students who struggle with a traditional format find more appealing and their motivation to complete the class is enhanced. Quality is having best teaching practices best learning practices and the ability to effectively communicate with the students. The key for having a quality online learning class that is well received by the students doing the research and applying simple design rules. The content, the application of the content and then analyzing the content is the way for any class to be successful. Peer review is a very easy yet significant way of making sure you are on task with your course. Student evaluations are a critical component but sometimes they do offer a short-term reflection.

## 8. MOVING FORWARD

Online courses have expanded rapidly and have the potential to extend further the educational opportunities of many students, particularly those least well-served by traditional educational institutions. By harnessing emerging technologies, universities can reach beyond campus walls to empower diverse learners at global scale. However, in their current design, online courses are difficult, especially for the students who are least prepared. Continued improvement of online curricula and instruction can strengthen the quality of these courses and hence the educational opportunities for the most in-need populations. Currently these students' learning and persistence outcomes are worse when they take online courses than they would have had these same students taken in-person courses. It begins with embracing stackable, online learning, which provides flexibility and affordability that increases access to university curricula and allows students to engage in smaller chunks of learning before committing to larger degree programs.

## 9. CONCLUSION

An online course isn't simply throwing in some PowerPoint presentations, assigned textbook readings, weekly quizzes and exams. The content must be balanced, clear, engaging and diverse. Given the nature of the course, communication is important. Students shouldn't be abandoned, and an effort must be made to both give them the assistance they need and create a sense of community between the students. Assessment activities are more than quizzes and exams. Learning management systems provide a plethora of assessment activities, such as discussion forums, peer reviews, collaborative wikis and other group projects. Assessment needs to go beyond measuring cognitive learning and include other aspects such as the development of affective skills (i.e., working in group projects or integrating use of technology in the assignments). It is imperative that an online or online learning course is organized and has an attractive, accessible design. This means that universal design for learning guidelines should be incorporated into the design. Though it can be tricky, accessibility must be ensured for all students, including those that may have some form of disability. Furthermore, students must be able to easily navigate course content and identify graded assignments to ensure the best learning environment. All legal and ethical aspects must be followed to the law. The only way these can be successfully accomplished is by the institution providing concrete support to the faculty and giving students the right to dispute these standards in a very tolerant atmosphere. When these tools are used effectively the student will be in a very healthy learning environment that is intuitive and encourages growth.

## REFERENCES

- [1] Allen, I. Elaine, Jeff Seaman, Russell Poulin, and Terri Taylor Straut (2016). Online Report Card: Tracking Online Education in the United States. Babson Park, MA: Babson Survey Research Group and Quahog Research Group, LLC.
- [2] Anderson, T. & Elloumi, F. (2004). Theory and practice of online learning. Athabasca: Athabasca University.
- [3] Bates, A. W. & Poole, G. (2003). Effective teaching with technology in higher education: Foundations for success. San Francisco: Jossey-Bass.
- [4] Bawa, Papia (2016). "Retention in Online Courses: Exploring Issues and Solutions—A Literature Review," SAGE Open, 1-11.
- [5] Belsky L (October 04, 2019). Where Online Learning Goes Next. Harvard Business Review. Retrieved from <https://hbr.org/2019/10/where-online-learning-goes-next>
- [6] Bettinger E and Loeb S Friday, June 9, 2017. Promises and pitfalls of online education. Retrived from <https://www.brookings.edu/research/promises-and-pitfalls-of-online-education/>
- [7] Bettinger, E., Fox, L., Loeb, S., & Taylor, E. (2017). Changing Distributions: How Online College Classes Alter Student and Professor Performance. American Economic Review.

- [8] Brown, J.S. (2013). Learning in and for the 21st century. In E. Low (Ed.), CJ Koh Professorial Lecture Series. Singapore: National Institute of Education/Nanyang Technological University. Retrieved from <http://www.johnseelybrown.com/CJKoh.pdf>
- [9] Carman, J.M. (October, 2002). Blended learning design: five key elements. Agilant Learning. Retrieved from <http://www.agilantlearning.com/pdf/Blended%20Learning%20Design.pdf>
- [10] Dietz-Uhler, B. and Hurn, J.(2011). Academic dishonesty in online courses. In Smith, P. (Ed.) Proceedings of the 2011 ASCUE Summer Conference. Myrtle Beach, SC. Retrieved from <http://www.ascue.org/files/proceedings/2011-final.pdf>
- [11] Fain Paul. Jan 16 2019. Online learning fails to deliver. HigherEd. Retrieved from <https://www.insidehighered.com/digital-learning/article/2019/01/16/online-learning-fails-deliver-findsreport- aimed-discouraging>
- [12] Hoffman, B. and Lowe, D. (2011, January). Effective online assessment: Scalable success strategies. In Faculty Seminars in Online Teaching. Seminar series conducted at the University of Central Florida, Orlando, FL. Retrieved from <https://online.ucf.edu/faculty-seminar01/>
- [13] Lieberman, Mark (2018, July 18). Keeping Online Courses Fresh: Valuable but Costly. Inside Higher Ed. Retrieved from <https://www.insidehighered.com/digital-learning/article/2018/07/18/maintaining-onlinecourses- maturity-requires-substantial>
- [14] Littlefield, Jamie. Mar. 15, 2019. What Does Research Say About Online Learning? ThoughtCo. Retrieved from [thoughtco.com/what-research-says-about-online-learning-1098012](http://thoughtco.com/what-research-says-about-online-learning-1098012).
- [15] Littlejohn, A. and Pegler, C. (2007). Preparing for blended e-learning. Routledge, LONDON: UK
- [16] Mitra, S. (2007, June). Technology and higher education — Pedagogy for self organised learning systems. Paper presented at Future of Education Online Conference. Retrieved from <https://sas.illuminate.com/site/external/jwsdetect/playback.jnlp?psid=2007-06-04.0738.M.BB2E854755AAFFF4E1A3E2523C4E54.vcr>
- [17] Riley, J.E., Gardner, C., Cosgrove, S., Olitsky, N., O'Neil, C., and Du, C. (2014). Implementation of blended learning for the improvement of student learning, In A. Picciano, C. Dziuban, and C. Graham (Eds.), Blended learning: Research perspectives, volume 2. NY: Routledge.
- [18] Spiros Protopsaltis and Sandy Baum<sup>1</sup> January 2019 Does Online education Live Up to its Promise?A look at the evidence and implications for federal policy. Retrieved from <http://mason.gmu.edu/~sprotops/OnlineEd.pdf>
- [19] U.S. Department of Education (2016). "Introduction to Competency-Based Education," CBE Experiment Guide.

## AUTHOR

### Dr. Keith Buckley

Dr. Buckley is a Professor who has taught in the Health and Physical Education department for the last 15 years at Rollins College. . He earned his Doctorate in Education from the University of Central Florida in 2003 specializing in Instructional technology. He holds a master's in business administration from the Crummer School of Business at Rollins College (1995) and earned his undergraduate degree from Rollins College. He is the Director of Physical Education at Rollins College and serves on the the Health and Wellness committee for the college. He holds the Greene Chair which was established in 1967 by Raymond W. Greene, who stipulated that the chairholder should encourage reverence for the "Amazing Creation of the Human Body and Its Functions." He also is the soccer coach of the college and has accumulated over 300 wins in his coaching tenure.





# TIBETAN AND CHINESE TEXT IMAGE CLASSIFICATION BASED ON CONVOLUTIONAL NEURAL NETWORK

Jincheng Li<sup>1</sup>, Penghai Zhao<sup>1</sup>, Yusheng Hao<sup>2</sup>, Qiang Lin<sup>2</sup>, Weilan Wang<sup>1\*</sup>

<sup>1</sup>Key Laboratory of China's Ethnic Languages and Information Technology  
(Northwest Minzu University), Ministry of Education,  
Lanzhou, P. R. China

<sup>2</sup>College of Mathematics and Computer Science, Northwest Minzu University,  
Lanzhou, P. R. China

## ABSTRACT

*The first stage of Tibetan-Chinese bilingual scene text detection and recognition is the detection of Tibetan- Chinese bilingual scene text. The detection results are mainly divided into three categories: successfully detected regions of Tibetan text and Chinese text, non-words regions with failed predictions. If the detected text image results are accurately classified, then the non-text images should be filtered in the recognition phase, meanwhile the Tibetan and Chinese text images can be identified by using different classifiers, such procedure can reduce the complexity of classification and recognition of two different characters by one recognition model. An accurate classification of Tibetan and Chinese text images is mattered. Therefore, this paper conducts a research on the classification of Tibetan, Chinese and non-text images by using convolutional neural networks. We perform a series of exploration about the classification accuracy of Tibetan, Chinese text images and non-text images with convolutional neural networks in different depths, and compare the accuracy with the classification results based on the transfer learning then analyze it. The results show that for the classification of Tibetan, Chinese and non-text images in the scene, using 7-layer convolutional neural network has reached saturation, and increasing the network depth does not improve the results, which provides reference values for Tibetan-Chinese text image classification.*

## KEYWORDS

*Convolutional Neural Network, Tibetan-Chinese scene text image, image classification, transfer learning*

## 1. INTRODUCTION

Image classification is one popular direction of the computer vision, which also provides a vital foundation for the application of object detection[1,2], face recognition[3,4], pose estimation[5,6], etc. Therefore, image classification technology has high value in academic research and applied value of science technology[7]. As AlexNet[8] surpassed the traditional methods at the Large Scale Visual Recognition Challenge 2012 and achieved remarkable results in the classification task, the following convolutional neural networks(CNN) model such as VGG[9], GoogLeNet[10], ResNet[11] were proposed. These networks made the CNN-based deep learning technology become the mainstream of the classification task[12-15]. Compared to the general neural networks, the basic structure of the convolutional neural network includes two layers, one is the

feature extraction layer, the other is the feature mapping layer. And from a perspective of model features, convolutional neural network has two particularities which can reduce the complexity of the model, one is the sparse connectivity, the other is the shared weights. With convolutional neural network constantly improving and optimizing by researchers, various excellent convolutional neural network models [16-20] were presented and had achieved acceptable results in the classification task.

In the Tibetan areas of China, almost all the textual information in various scene contains both Tibetan and Chinese characters. The objective of text detection for this kind of scene is to locate the position of Tibetan and Chinese characters, and the follow-up is to put the segment of two detected text regions and non-text regions into the trained classifier, thus scanned images with text is converted into computer-readable data. This paper provides a feasible scheme for Tibetan-Chinese scene text recognition by classifying Tibetan text images, Chinese text images, and non-text images. That means using different classifiers to recognize Tibetan text image and Chinese image text respectively, so as to make Tibetan-Chinese scene text recognition more simple and effective.

Traditional image classification methods are generally divided into two steps: First, calculating artificially designed features from the input image. Second, training a classifier based on the extracted features. The effect of this classification method depends on the artificially designed features, thus it has great uncertainty. For the above situation, this paper conducted an exploratory study on the classification of Tibetan, Chinese, and non-text images by using convolutional neural networks. Regarding the constructed data set as experimental data, the features of text images were extracted by employing convolutional neural networks with different depths. Then applying softmax to classify and comparing the classification results with the pre-trained VGG16 model on which transfer learning method is implemented. The results show that the seven-layer deep convolutional neural network has achieved a 98.28% classification accuracy, and increasing the depth of the network has no significant improvement on the classification accuracy, which indicates that for our classification task, the seven-layer network has reached saturation.

## **2. DATASET AND EVALUATION PROTOCOL**

### **2.1. The Dataset**

For the Chinese text image dataset, 5000 pieces of text were selected from the existing dataset[21] and the images taken in Tibetan area. For the Tibetan text image dataset, the same amount of text was obtained by intercepting from the text images of Tibetan area or synthesizing. For non-text image dataset, we randomly cropped 5000 images from the real background image captured by camera. Then it is divided into three categories: Tibetan text, Chinese text and non-text. Figure 1. shows some samples of Tibetan, Chinese, and non-text images. It can be seen from the figure that the Chinese text image and the Tibetan text image have distinct shape of characters. Comparing with Chinese one, Tibetan text image has more complex backgrounds, various colors, and different scales, image sizes. Non-text images also have various textures, colors, and backgrounds which increase the difficulty of classification. During the training, these data of samples are randomly augmented to improve the generalization ability of the model.



Figure 1. Scene image examples

## 2.2. Evaluation Methods

We randomly divided the Tibetan-Chinese text image dataset into training set and test set at a ratio of 7:3, and then evaluated the experimental results with the following two evaluation methods.

- 1) *Precision*: Number of correctly identified test samples as a percentage of total test samples, they are given by:

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

Where  $TP$  is the number of positive test samples that are correctly classified as positive samples.  $FP$  is the number of negative test samples that are incorrectly classified as positive samples.  $TP + FP$  is all test samples that are classified as positive samples.

- 2) *F1 Score*: It is an index used to measure the accuracy of classification models in statistics. It takes into account both the accuracy and recall of the classification model. They are given by:

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

Where  $FN$  is the number of positive test samples that were misclassified as negative samples;  $TP + FN$  is all the positive test samples; Recall is the recall rate.

## 3. THE ARCHITECTURE

### 3.1. Convnet Configurations

The structure of our convolutional neural network is shown in Table 1. The configurations of four different depth are: 5-layer, 7-layer, 9-layer, and 11-layer (excluding pooling and softmax layers). With image size of input layer being  $180 \times 60$ , the network carries out the convolution and subsampling operation alternately three times, and sends the data to the two followed fully connected layers. All convolution manipulations use a size of  $3 \times 3$  kernel, while subsampling manipulations use a size of  $2 \times 2$  max-pooling kernel of which step size is 2, and the convolution layer depths are 64, 128, and 256 respectively.

Table1. Network Configuration

<b>ConvNet Configuration</b>			
<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>
5 weight layers	7 weight layers	9 weight layers	11 weight layers
Input(180×60 RGB image)			
Conv3-64	Conv3-64	Conv3-64 Conv3-64	Conv3-64 Conv3-64
Maxpool			
Conv3-128	Conv3-128 Conv3-128	Conv3-128 Conv3-128	Conv3-128 Conv3-128 Conv3-128
Maxpool			
Conv3-256	Conv3-256 Conv3-256	Conv3-256 Conv3-256 Conv3-256	Conv3-256 Conv3-256 Conv3-256 Conv3-256
Maxpool			
FC-512			
FC-3			
Softmax			

The network configuration in Table 1 includes the use of convolutional layer, pooling layer, BNlayer and Dropout. The detailed descriptions of each deep network structure are as follows :

1) Input: First, the training images of different sizes are scaled to  $180 \times 60$  by using a bilinear interpolation algorithm. In order to improve the generalization ability of the classification model and avoid overfitting, the training data were augmented by reversal, mirroring, translation and perspective transformation to increase the amount of data. Then send 32 training images in each batch to the network for training.

2) Convolution Layer: The convolutional layer includes convolution, activation functions, batch normalization, and max-pooling. The 5-layer, 7-layer, 9-layer, and 11-layer depth network contains three, five, seven, and nine convolutional layers respectively, and each convolutional layer obtains feature maps with different sizes. Each convolution operation uses a  $3 \times 3$  convolution kernel. The advantage of this method is that it can reduce the number of parameters and implement more nonlinear mapping at the same time. After each convolution, the ReLU function is used to activate the feature map  $F_i$ . The generation process of  $F_i$  is given by:

$$f(x) = \max(0, x) \quad (4)$$

$$F_i = f(F_{i-1} * W_i + b_i) \quad (5)$$

Where  $W_i$  is the weight of convolution in layer  $i$ ,  $b_i$  is the offset in layer  $i$ ,  $*$  is the convolution operation,  $f$  is the nonlinear excitation function of ReLU.

In order to accelerate the convergence of the network and prevent the overfitting of the network, each feature map  $F_i$  is followed by the batch normalization(BN) layer, that is, batch standardization, then proceed to the next layer. As the feature map passed through one to four convolutional layers, max-pooling layer operates on each standardized feature map independently. The purpose of this procedure is reducing the dimension of the feature map and to remaining invariant to changes in scale or rotation.

Obviously, for each depth network, the main features of the image are obtained by convolution operation, nonlinear excitation function, and batch standardization. Figure.2 shows the visual effect of the first convolution operation of each image category. It can be seen that the first convolution layer mainly extracts the edge, corner and color feature information about the image

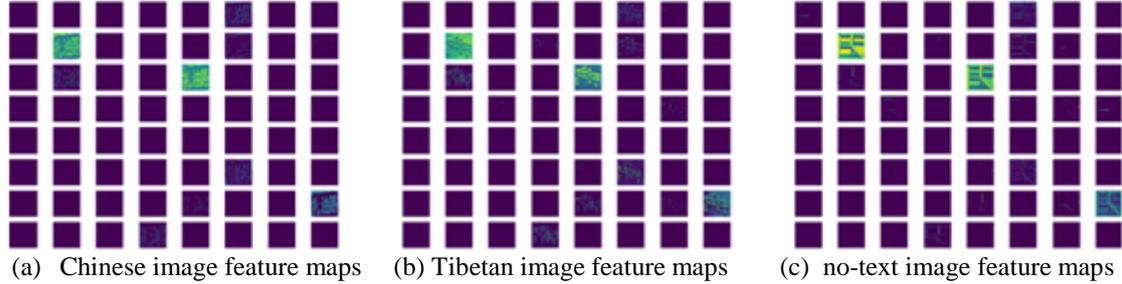


Figure. 2 After the first convolution feature maps

3) **Output Layer:** After multiple convolution and subsampling layers, two full connected layers are connected. The full connection layer is to integrate the local information output from the subsampling layer. The dropout strategy is used in two full connection layers to prevent the overfitting and improve the generalization ability of the model. It will set the output of neurons in the full connection layer to 0 with a probability of 0.6, and these neurons output as 0 will no longer carry out forward propagation and back propagation. At last, the prediction results of the classification output by applying softmax are used to realize the image classification.

### 3.2. Loss Function

During the training, we use the cross entropy loss function to calculate the loss for the classification results. Specifically, the loss function  $L$  of cross entropy is given by:

$$J = - \sum_{i=1}^N y_i \cdot \log(p(y_i)) \quad (6)$$

Where  $N$  is the number of categories,  $y$  represents the label (0 or 1). If the category is the same as the sample, it is 1, otherwise, it is 0.  $p(y_i)$  is the predicted probability that the observation sample belongs to category  $i$ .

## 4. EXPERIMENTAL RESULTS AND ANALYSIS

We adopted Windows10 + Python3.7 + Keras2.2 as the deep learning framework and conducted experiments on this basis. To prepare the training samples, we use the datasets containing 15000 samples from section II.A, 5000 samples for each category. There are 10500 images in the training set and 4500 images in the test set.

### 4.1. Analysis of the Influence of Different Depth Network Structures on Classification Results

We perform experiments on the network of four different depths of 5-layer, 7-layer, 9-layers, and 11-layer to analyze the impact of different depths on the classification results. For each network, the training image is scaled to  $180 \times 60$  by using bilinear interpolation at the first beginning. The training data is augmented by random horizontal flipping, mirroring, and other operations. Among them, the batch size is 32, and each convolution layer is followed by a batch

normalization layer. The initial learning rate is set to 0.0001 and subject to exponential decay every 20 epochs. The Adam optimization algorithm is used and the image is scaled to  $180 \times 60$  during testing. With the same implementation details, the overall average accuracy and F1 evaluation of different depths of the network after 200 epochs are shown in Table 2.

Table 2. Results of networks at different depths

Network layers	5-layer	7-layer	9-layer	11-layer
Precision(%)	97.50	98.28	98.25	98.15
Recall(%)	97.46	98.28	98.24	98.15
F1-score(%)	97.47	98.28	98.25	98.15

In Table 2, from the perspective of precision, while using convolutional neural network with only five layers, the classification accuracy is about 97.5%. By using a 7-layer convolution neural network, the accuracy of correct classification has reached 98.3%. The accuracy of 9-layer and 11-layer networks is similar to that of 7-layer network, which shows that the depth has a great influence on the performance of the convolution neural network, but with the increase of the depth, the network will gradually reach saturation. At the same time, it also reveals that for the Tibetan and Chinese text image classification, the 7-layer network is basically saturated, increasing the depth of the network does not significantly improve the results.

## 4.2. Comparative Analysis of Experimental Results

According to the analysis in section 4.1, we can know that the optimal classification accuracy can be obtained when the depth of the convolutional neural network is seven. In order to verify the effectiveness of the network depth, a comparison experiment is performed with the transfer learning method. Using the pre-trained VGG16[9] model for transfer learning, the parameter values of the convolutional layer (feature layer) are fixed, but the last fully connected layer is retrained from scratch, and let the number of output neurons be consistent with the number of categories of the dataset. The data uses the same set of training samples and test samples, regarding 10500 images as the training set and 4500 images as the test set. The experimental results are shown in Table 3.

Table 3. Results of transfer learning

Network Model	Pretrain VGG16	Depth 7-layer
Precision(%)	97.46	98.28
Recall(%)	97.43	98.28
F1-score(%)	97.43	98.28
speed(s/epoch)	19	23

It can be seen from Table 3. that the classification result of the convolutional neural network with a depth of 7 is compared with the pre-trained VGG16 model for transfer learning. Although transfer learning can turn the learned model parameters to the new model and thus speed up and optimize the learning efficiency of the model, but essentially, VGG16 is a 16-layer convolutional network, the network depth is more than twice that of the 7-layer network, besides, its accuracy is lower than the latter. This further illustrates that deepening the network depth does not improve the accuracy of our classification task, and verifies the effectiveness of the 7-layer network.

### 4.3. Discussion

Through the classification of Tibetan, Chinese and non-text images on convolutional neural networks of different depths, with the increase of the network depth, the classification accuracy is constantly improved. When the network depth is 7, the best accuracy is achieved, and the accuracy will decrease if the depth continues to increase. This could be that the deeper the network, the smaller the size of the feature map, which loses a lot of information, and the phenomenon of gradient disappearance will become more and more obvious, so the classification accuracy will be reduced.

In the process of text image classification, when the text in the image is written horizontally or with a single background, it can be accurately classified, as shown in Figure.3(a). When the text in the image is written in the vertical direction, the non-text background of the image is diverse or other text appears, there will be misclassification occurring, see Figure.3(b). For the problem of vertical text image misclassification, the main reason could be that there are few vertical text images in the training samples, when two types of text appear at the same time in one text image, the misclassification will happen, which is not the result we need obviously.

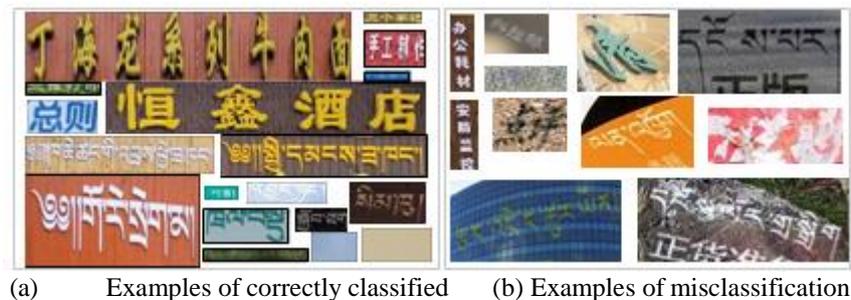


Figure.3 Classification results

## 5. CONCLUSIONS AND FUTURE WORK

Based on the in-depth study of CNN, we use the advantages of convolutional neural networks to conduct an exploratory study of multi-layer deep neural networks to extract and classify text image and non-text image features of scenes in different languages, analyze the impact of different depths on classification results, and simultaneously compare with the classification results of the transfer learning methods. The results show that the network depth has an impact on the classification results. As the network depth increases, the overall classification accuracy increases first and then decreases. Therefore, an appropriate convolutional network depth should be selected for our Tibetan, Chinese, and non-text image classification. For the classification of our task, the shortcomings are that the various non-text background and the paper with low opacity will make the network misclassify. Future studies will focus on exploring the comparison of different deep convolutional neural networks such as ResNet to avoid the gradient disappearing along with network deepening. At the same time, more Tibetan and Chinese text images and non-text image data are collected to train the network, so that it can classify more complex background text images, and final to be able to apply Tibetan, Chinese and non-text image classification to Tibetan- Chinese bilingual scene text detection and recognition.

## ACKNOWLEDGEMENTS

This work is supported by the National Natural Science Foundation of China (No.61772430) and by Program for Innovative Research Team of SEAC ((2018)98). The Gansu Provincial First-Class Discipline Program of Northwest Minzu University (No.11080305). The Northwest Minzu University for Nationalities' Central University Fundamental Research Service Fund Special Funds Support Postgraduate Research and Innovation Projects (Yxm2020108) supports the work also.

## REFERENCES

- [1] Ouyang W, Zeng X, Wang X, et al. DeepID-Net: Object Detection with Deformable Part Based Convolutional Neural Networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(7): 1320-1334.
- [2] Ali Diba, Vivek Sharma, Ali Pazandeh, et al. Weakly super-vised cascaded convolutional networks[C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 5131-5139.
- [3] HU G, YANG Y X, YI D, et al. When face recognition meets with deep learning: an evaluation of convolutional neural net-works for face recognition[C]. *International Conference on Computer Vision*, 2015: 142-150.
- [4] LAWRENCE S, GILES C L, TSOI A C, et al. Face recognition: a convolutional neural-network approach[J]. *IEEE Transactions on Neural Networks*, 1997, 8(1): 98-113.
- [5] Cao Z, SIMON T, WEI S, et al. Realtime multi-person 2D pose estimation using part affinity Fields[C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 1302-1310.
- [6] TOSHEV A, SZEGEDY C. DeepPose: human pose estimation via deep neural networks[C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2014: 1653-1660.
- [7] SuFu, Lv Qin, LuoRenze, Review of image classification based on deep learning[J]. *Telecommunications Science*, 2019,35(11):58-74.
- [8] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Proceedings of the 25th International Conference on Neural Information Processing Systems*. Lake Tahoe, NV, USA. 2012. 1097–1105.
- [9] Simonyan, Karen, Zisserman, Andrew. Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. *Computer Science*, 2014.
- [10] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[J]. *arXiv preprint arXiv:1409.4842*, 2014.
- [11] He KM, Zhang XY, Ren SQ, et al. Deep residual learning for image recognition[C]. *Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. 2015. 770–778.
- [12] Wang JZ, Yang Y, He YH. Pornographic Images Recognition Framework Based on Multi-Classification and ResNet[C]. *Computer Systems and Applications*, 2018, 27(9): 100-106.
- [13] Zhao B, Li P, Dai MR, Ma XN. Research on Optimization Method of Railway Image Scene Classification Based on Deep Learning Method. *Computer Systems and Applications*, 2019, 28(6): 228-234.
- [14] Fang HW, Shi HJ. Satellite Image Recognition and Classification Method Based on Deep Learning. *Computer Systems and Applications*, 2019, 28(10): 27-34.
- [15] Yang Bing, Chen Hao-yue, Wang Xiao-hua, YaoJin-liang. Chinese Painting Image Classification Based on Convolution Neural Network[J]. *Software Guide*, 2019, 18(01):11-14.
- [16] Bai Cong, Huang Ling, Chen jia-nan, Pan Xiang, Chen Shengyong. Optimization of Deep Convolutional Neural Network for Large Scale Image Classification[J]. *Journal of Software*, 2018,29(04):1029-1038.
- [17] Liu Wanjun, Liang Xuejian, Qu Haicheng. Learning performance of convolutional neural networks with different pooling models[J]. *Journal of Image and Graphics*, 2016,21(9):1178-1190.
- [18] Wang Min, Liu Kexin, Liu Li, Yang Runling. Super-Resolution Reconstruction of Image Base on Optimized Convolution Neural Network[J]. *Laser&Optoelectronics Progress*, 2017, 54(11): 111005.
- [19] Li Ming, Zhang Hong. Image classification based on convolution neural network of iterative optimization[J]. *Computer Engineering and Design*, 2017, 38(1): 198-202.
- [20] Guo ST, Luo YX, Song YZ. Random forests and VGG-NET: An algorithm for the ISIC 2017 skin lesion classification challenge. *arXiv preprint arXiv:1703.05148*, 2017.

- [21] Jaderberg M, Simonyan K, Vedaldi A, et al. Synthetic data and artificial neural networks for natural scene text recognition[J]. arXiv preprint arXiv:1406.2227, 2014.

## AUTHORS

**Jincheng Li** received a bachelor's degree in 2018. He started his master's degree in computer technology at Northwestern University for nationalities in 2018. His research interests include image processing, pattern recognition and artificial intelligence. Contact him at ljicher@gmail.com



**Penghai Zhao** received a bachelor's degree in 2019. He started his master's degree in computer technology at Northwestern University for nationalities in 2019. His research interests include image processing, pattern recognition.



**Yusheng Hao** is a lecturer in the School of Mathematics and Computer Science, Northwest Minzu University. He is also a member of CCF and currently pursuing his PhD at the Key Laboratory of China's Ethnic Languages and Information Technology of Ministry of Education, Northwest Minzu University. His research interests include signal processing, document analysis and identification, thangka image processing. Contact him at haoyusheng@xbmu.edu.cn



**Qiang Lin** was born in 1979. He received the Ph.D. degree in computer science and technology from Northwestern Polytechnical University in 2014. Now he is an associate professor and M.S. supervisor at Northwest Minzu University. His research interests include pervasive computing, intelligent information processing, data stream mining, etc.



**Weilan Wang** received a B.S degree in Mathematics from Northwest Normal University, Lanzhou, China, in 1983. She was a visiting scholar with the Sun Yat-sen University, Guangzhou, China, in 1987. From 2001 to 2002, she was a visiting scholar with Tsinghua University, Beijing, China. From 2006 to 2007, she was a visiting scholar with Indiana University, Bloomington, USA. She is currently a Professor with the College of Mathematics and Computer Science, Northwest Minzu University, Lanzhou City, China. Her research interests include image processing, pattern recognition, Tibetan information processing and so on.



## AUTHOR INDEX

<i>Ahmed Riadh Baba-Ali</i>	41
<i>Ashkan Tashk</i>	33
<i>Basma Sagr</i>	41
<i>Damir Krstinić</i>	01
<i>Dunja Božić-Štulić</i>	01
<i>Georgia Koukiou</i>	23
<i>Jincheng Li</i>	59
<i>Keith Buckley</i>	51
<i>Ljiljana Šerić</i>	01
<i>Maja Braović</i>	01
<i>Maryamsadat Nejadghaderi</i>	33
<i>Parvin Mansouri</i>	33
<i>Penghai Zhao</i>	59
<i>Piya Techateerawat</i>	15
<i>Qiang Lin</i>	59
<i>Slimane Abou-Msabab</i>	41
<i>Vassilis Anastassopoulos</i>	23
<i>Weilan Wang</i>	59
<i>Yusheng Hao</i>	59
<i>Zahra Safaei Naraghi</i>	33