Advances in Intelligent Systems and Computing 1393

Aruna Tiwari · Kapil Ahuja ·
Anupam Yadav ·
Jagdish Chand Bansal · Kusum Deep ·
Atulya K. Nagar *Editors*

Soft Computing for Problem Solving

Proceedings of SocProS 2020, Volume 2

Springer

Advances in Intelligent Systems and Computing

Volume 1393

The series "Advances in Intelligent Systems and Computing" contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing such as: computational intelligence, soft computing including neural networks, fuzzy systems, evolutionary computing and the fusion of these paradigms, social intelligence, ambient intelligence, computational neuroscience, artificial life, virtual worlds and society, cognitive science and systems, Perception and Vision, DNA and immune based systems, self-organizing and adaptive systems, e-Learning and teaching, human-centered and human-centric computing, recommender systems, intelligent control, robotics and mechatronics including human-machine teaming, knowledge-based paradigms, learning paradigms, machine ethics, intelligent data analysis, knowledge management, intelligent agents, intelligent decision making and support, intelligent network security, trust management, interactive entertainment, Web intelligence and multimedia.

The publications within "Advances in Intelligent Systems and Computing" are primarily proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

Indexed by DBLP, INSPEC, WTI Frankfurt eG, zbMATH, Japanese Science and Technology Agency (JST).

All books published in the series are submitted for consideration in Web of Science.

More information about this series at http://www.springer.com/series/11156

Aruna Tiwari · Kapil Ahuja · Anupam Yadav ·
Jagdish Chand Bansal · Kusum Deep ·
Atulya K. Nagar
Editors

Soft Computing for Problem Solving

Proceedings of SocProS 2020, Volume 2

Springer

Editors
Aruna Tiwari
Computer Science and Engineering
Indian Institute of Technology Indore
Indore, Madhya Pradesh, India

Kapil Ahuja
Computer Science and Engineering
Indian Institute of Technology Indore
Indore, Madhya Pradesh, India

Anupam Yadav
Dr. B. R. Ambedkar National Institute
of Technology
Jalandhar, India

Jagdish Chand Bansal
Department of Mathematics
South Asian University
New Delhi, India

Kusum Deep
Department of Mathematics
Indian Institute of Technology Roorkee
Roorkee, India

Atulya K. Nagar
School of Mathematics, Computer Science
and Engineering
Liverpool Hope University
Liverpool, UK

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-981-16-2711-8 ISBN 978-981-16-2712-5 (eBook)
https://doi.org/10.1007/978-981-16-2712-5

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

IIT Indore and the Soft Computing Research Society (SCRS), New Delhi, co-hosted the "10th International Conference on Soft Computing for Problem Solving (SocPros 2020)" from 18 December to 20 December 2020 in a Virtual Mode. The seeds for this conference were laid more than a year ago at the 9th conference in this series at the Liverpool Hope University, UK (in September 2019).

The conference opening was done by Prof. Deepak B. Phatak (BoG Chairman, IIT Indore), Prof. Neelesh K. Jain (Director, IIT Indore), and Prof. Ajit K. Chaturvedi (Director, IIT Roorkee and IIT Mandi). These esteemed guests appreciated the efforts and highlighted the need of taking the technology to the common people.

This mega event, which happens to be the first international conference of Computer Science and Engineering at IIT Indore, covered recent developments in the interdisciplinary areas of Artificial Intelligence, Machine Learning, Optimization, and Soft Computing. The conference received 334 papers from participants belonging to 13 different countries, which went through a very stringent blind review process. This was done by the international expert committee and had a very good acceptance rate of 37%. These papers would be published as two books by Springer.

This year the conference had many innovative features. Prof. Chandra Mohan Gold Medal for excellence in Soft Computing was instituted, which was given to Prof. Sankar Pal of ISI Kolkatta. Twelve eminent academicians gave keynote talks. There was a big industry participation with four keynote talks by distinguished industrialists. Fourteen outstanding paper awards and 5 best paper awards (sponsored by Springer) were given. The conference had two special sessions on "Cognitive Science" and "Remote Sensing" as well as a panel discussion on the future of soft computing with ten renowned panelists from academia and industry.

Indore, India	Aruna Tiwari
Indore, India	Kapil Ahuja
Jalandhar, India	Anupam Yadav
New Delhi, India	Jagdish Chand Bansal
Roorkee, India	Kusum Deep
Liverpool, UK	Atulya K. Nagar

Contents

About the Editors

Dr. Aruna Tiwari is Associate Professor in Computer Science & Engineering at IIT Indore, India. Her all degrees B.Tech., M.Tech., Ph.D. are in Computer Science & Engineering from SGSITS Indore. She has more than 20 years of research experience in the areas such as artificial intelligence, machine learning, and soft-computing. Her work is around artificial neural networks, fuzzy clustering, evolutionary computation and their applications to bioinformatics, and medical diagnosis. Her research directions are toward the design of intelligent computing algorithms for classification, clustering, and feature selection which can handle big data. She has more than 30 journal publications of international repute including more than 5 IEEE transactions and Elsevier journals of high impact factor and has published more than fifty research articles in the ranked conferences. She has established a big data handling lab in IIT Indore which is funded by Council of Science & Industrial Research (CSIR) Government of India in 2017. She has enabled two MOUs, collaborating with Indian Institute of Soybean Research, Indore, and CSIR-Central Electronics Engineering Research Institute, Pilani. She has recently initiated a consortium project on Artificial Intelligence which is approved from Ministry of Electronics & Information Technology (MeitY) Government of India. There are seven PhDs and more than 25 masters awarded under her guidance.

Dr. Kapil Ahuja holds master's and Ph.D. degrees in Mathematics and Computer Science (from Virginia Tech, USA) and has a strong interdisciplinary focus. After graduating earlier this decade from VT, he received his postdoctoral training from the Max Planck Institute in Magdeburg (Germany). Since then, he has established his independent research program in Mathematics of Data Science and Computational Science at IIT Indore, where he is currently working as Associate Professor. Dr. Ahuja is solving challenging problems that are at both the ends of the research spectrum, i.e., theoretical as well as applicable. His core research interests are in artificial intelligence, machine learning, numerical methods, and optimization. He believes that it is necessary to collaborate globally to solve challenging research problems. Hence, he has multiple active international research collaborations (USA, Germany, India, France, and UK). In the recent past, he has also held visiting professor positions at TU Braunschweig (Germany), TU Dresden (Germany), and Sandia National Labs

(USA). Dr. Ahuja's overall research output includes thirty publications (including eighteen in reputed journals) as well as external funding worth more than half-a-million USD from twelve projects. While achieving this, he has graduated two Ph.D. students with three more to graduate soon. Since teaching and service are essential for a fruitful and satisfying research career, he is very committed to these aspects as well. Dr. Ahuja has received the Best Teacher Award four times at IIT Indore. In the past, he has held many administrative positions. Since the past three years, he is heading International Affairs at IIT Indore as Founding Dean for the same.

Dr. Anupam Yadav is Assistant Professor in the Department of Mathematics at Dr. BR Ambedkar National Institute of Technology Jalandhar, India. His research area includes numerical optimization, soft computing, and artificial intelligence, he has more than ten years of research experience in the areas of soft computing and optimization. Dr. Yadav has done Ph.D. in soft computing from Indian Institute of Technology Roorkee, and he worked as Research Professor at Korea University. He has published more than twenty-five research articles in journals of international repute and has published more than fifteen research articles in conference proceedings. Dr. Yadav has authored a text book entitled "An introduction to neural network methods for differential equations". He has edited three books which are published by AISC, Springer Series. Dr. Yadav was General Chair, Convener, and Member of steering committee of several international conferences. He is Associate Editor in the journal of the experimental and theoretical artificial intelligence. Dr. Yadav is Member of various research societies.

Dr. Jagdish Chand Bansal is Associate Professor at South Asian University, New Delhi, and Visiting Faculty at Maths and Computer Science, Liverpool Hope University, UK. Dr. Bansal has obtained his Ph.D. in Mathematics from IIT Roorkee. Before joining SAU, New Delhi, he has worked as Assistant Professor at ABV Indian Institute of Information Technology and Management Gwalior and BITS Pilani. His primary area of interest is swarm intelligence and nature-inspired optimization techniques. Recently, he proposed a fission–fusion social structure-based optimization algorithm, spider monkey optimization (SMO), which is being applied to various problems from engineering domain. He has published more than 60 research papers in various international journals/conferences. He is Series Editor of the book series such as Algorithms for Intelligent Systems (AIS) and Studies in Autonomic, Data-driven and Industrial Computing (SADIC) published by Springer. He is Editor-in-Chief of International Journal of Swarm Intelligence (IJSI) published by Inderscience. He is also Associate Editor of IEEE ACESSS published by IEEE and ARRAY published by Elsevier. He is Steering Committee Member and General Chair of the annual conference series SocProS. He is General Secretary of Soft Computing Research Society (SCRS). He has also received Gold Medal at UG and PG level.

Prof. Kusum Deep is Full Professor in the Department of Mathematics, Indian Institute of Technology Roorkee, India, and Visiting Professor, Liverpool Hope University, UK, and University of Technology Sydney, Australia. She has completed her B.Sc. (Hons) and M.Sc. (Hons) from Centre for Advanced Studies, Panjab University, Chandigarh. She is an M.Phil. Gold Medalist. She earned her Ph.D. from UoR (now IIT Roorkee) in 1988. She has been a national scholarship holder and a postdoctoral from Loughborough University, UK assisted by International Bursary funded by Commission of European Communities, Brussels. She has won numerous awards like Khosla Research Award, UGC Career Award Starred Performer of IITR Faculty, many best paper awards by Railway Bulletin of Indian Railways, special facilitation in memory of late Professor M. C. Puri, and AIAP Excellence Award. She has authored two books, supervised 20 Ph.Ds, and published 125 research papers. She is Senior Member of IEEE, ORSI, CSI, IMS, and ISIM. She is Executive Editor of International Journal of Swarm Intelligence, Inderscience. She is Associate Editor of Swarm and Evolutionary Algorithms, Elsevier. She is on the editorial board of many journals. She is Founder President of Soft Computing Research Society, India. She is General Chair of series of International Conference on Soft Computing for Problems Solving (SocProS). Her research interests are evolutionary algorithms, swarm Iitelligence, and nature-inspired optimization techniques and their applications.

Prof. Atulya K. Nagar holds Foundation Chair as Professor of Mathematical Sciences and is Pro-Vice-Chancellor for Research at Liverpool Hope University, United Kingdom. He has been Dean of the Faculty of Science and Head of the School of Mathematics, Computer Science and Engineering which he established at the University. He received a prestigious Commonwealth Fellowship for pursuing his doctorate (D.Phil.) in Applied Nonlinear Mathematics, which he earned from the University of York (UK) in 1996. He holds B.Sc. (Hons), M.Sc. (Mathematics), and M.Phil. (with distinction) in Mathematical Physics from the MDS University of Ajmer, India. Prior to joining Liverpool Hope, he was at the Department of Mathematical Sciences and later at the Department of Systems Engineering, at Brunel University, London. He is internationally respected Scholar working at the cutting edge of theoretical computer science, applied mathematical analysis, and systems engineering with his research expertise spanning both applied mathematics and computational methods for nonlinear, complex, and intractable problems arising in science, engineering, and industry. He has edited volumes on Intelligent System, and Applied Mathematics. He is Editor-in-Chief of the International Journal of Artificial Intelligence and Soft Computing (IJAISC) and serves on editorial boards for a number of prestigious journals. He is well published with over 450 publications in prestigious publishing outlets. Professor Nagar sits on a number of strategic UK wide research bodies including the JISC Research Strategy group, and he is Fellow of the Institute of Mathematics and its Applications (FIMA) and Fellow of the Higher Education Academy (FHEA).

A Deep Semi-supervised Approach for Multi-label Land-Cover Classification Under Scarcity of Labelled Images

Shounak Chakraborty, Nilesh Agarwal, and Moumita Roy

Abstract In this manuscript, a land-cover classification (LCC) mechanism has been investigated for the practical situations where a remotely sensed aerial image can be annotated by more than one land-cover class. A striking factor for the development of the proposed technique is its ability to operate under situations when there is scarcity of labelled images in the training set, thereby alleviating extensive manual collection of multi-label ground truth information for LCC. The solution using very few labelled training images has been envisioned through a semi-supervised deep learning-based methodology where templates have been generated corresponding to each of the classes present in the training images. Thereafter, each of the test images is assigned to multiple classes based on their similarity with each of the class templates. Experimentation conducted on UCM and AID benchmark multi-label aerial image datasets suggests promising results for the proposed approach.

1 Introduction

In the past few decades, the advancement in remote sensing machinery has resulted in the generation of a massive number of digital images captured periodically over the planet's surface, rendering extensive scope for Earth observation. The images captured by satellites as well as by manned or unmanned aerial vehicles have provided scope for the development of software products capable of automatic surveillance of land-covers (i.e. bare soil, forest, agriculture land, water body, etc.) with precise accuracies. These automatic land-cover classification (LCC) techniques are useful in diverse domains like forest and crop monitoring, thermal mapping, land-cover map generation, biogeochemical cycling, biodiversity to name a few [1]. In the current trends, the conventional single-label classification techniques have been in the limelight (where a pixel or an image is annotated with only a single class) since the spatial resolutions of the satellite images have been limited. Accordingly, these

S. Chakraborty (✉) · N. Agarwal · M. Roy
Indian Institute of Information Technology, Design and Manufacturing Kurnool, Kurnool, Andhra Pradesh, India
e-mail: shounak@iiitk.ac.in

A. Tiwari et al. (eds.), *Soft Computing for Problem Solving*, Advances in Intelligent Systems and Computing 1393, https://doi.org/10.1007/978-981-16-2712-5_1

techniques can be broadly categorised [2] as pixel-level classification of satellite images where every pixel is assigned with a class label [3], scene-level classification [4] where each image obtained from an aerial vehicle flying at lower altitude is associated with a label, and object-level classification where each land-objects like cars, buildings, etc., are detected from an image scene [5]. Nevertheless, annotating a pixel or an image with a single class label is far from practical as there should be more than one land-cover class always present in the pixel or the image. For example, each pixel in the Landsat-8 satellite image has a spatial resolution of 15–30 m where more than one land class (like soil and water, vegetation and buildings) can easily co-exist. Regardless, obtaining ground truth information about multiple classes corresponding to each of the pixels of low (spatial) resolution satellite images is extremely daunting due to spectral mixing of signals [6].

A recent expedition in the use of (un)manned aerial vehicles for remote sensing has made a large-scale collection of high-resolution images over the Earth surface possible [7–9]. Though expensive in terms of human effort, these images have multiple land-cover classes clearly distinguishable which makes the collection of ground truth possible. Conventionally, the task of multi-label classification has been handled in a supervised fashion and the contributions therein can be categorised either as problem transformation (i.e. decomposing the multi-label problem as a set of single-label problems) [10, 11] and algorithm adaptation (i.e. adaptation of classification algorithms to work on multi-label data directly) [12–14]. More specifically, the former reduces the multi-label problem into a set of binary classification problem to tackle the problem, which however ignores the dependencies that may be present among the class labels occurring together in the input image [15]. As a solution to this, algorithm adaptation techniques, extend the traditional classification approaches like k-nearest neighbour [12], boosting [13] and neural networks [14] for solving the multi-label classification problem.

In specific context to remote sensing, one of the pioneering works has introduced multi-label LCC as a single-label classification problem where each combination of the co-existing classes have been treated as a separate class [11]. However, this had an obvious drawback in the unavailability of sufficient training examples for each newly defined composite class. To solve this, Moranduzzo et al. [16] have partitioned the images into multiple tiles and assigned coarse classes corresponding to each of them. Similarly, a radial basis function neural network [17] has been used with customised thresholds to identify the possible land classes in an image [18]. Furthermore, the multi-class problem has been solved deliberately by exploring the spatial characteristics and output label structure through a structured SVM [19]. Nevertheless, a semi-supervised graph-theoretic approach has been explored in [20] for extraction of multiple class labels from the aerial images. Recently, deep convolution neural network (CNN) [21] has been investigated for this task using augmented images to supplement the training set [8]. Moreover, recent investigations have suggested exploitation of the strong co-relation among the co-existing labels in case of the remotely sensed images. For example, the occurrence of ship class strongly indicates the presence of water body in the same image; same can be said about the co-occurrence of building-pavements or car–road label pairs [9] within the same image. In this direction, recurrent neural networks [21] have been coupled with CNN to explore spatial dependencies among the image segments using bi-directional long short-term memory neural networks (LSTM) [7, 21].

Most of the multi-label LCC techniques developed so far indicate considerably high prediction accuracies while operating with sufficient availability of labelled training images (assuming standard training–testing image ratio of 80:20). However, the accuracies drop significantly when the amount of training images to be used are reduced. Since the manual generation of such training set containing multiple annotations for every image is expensive, infeasible and time-consuming (even many times more daunting than obtaining single-labelled training set), the development of an approach requiring less training images is of utmost importance. As per the knowledge of the authors, the investigation of a multi-label LCC approach operating with a reduced training set is yet to be explored. In this regard, a deep learning-based solution to the problem of multi-label LCC has been proposed in this manuscript to work under a situation when there is a severe scarcity of labelled images. At the onset, a multi-channel CNN has been investigated for extraction of the 'most confident' training images from each class in the training set. These confident training images from each class have been used to generate a corresponding class template through U-NET CNN network [22]. Here, the training images having the highest of confidences to contain a major class is segmented to extract the major object (class). Thereafter, the collection of the segmented objects is saved as the templates for each class. These class templates are then input to multi-channel CNN for one-versus-all classification training. However, the class-correlation probability matrix is also considered during training to capture the inter-label dependencies. During testing, each of the test images is passed through each of the class-specific U-NET to obtain its map corresponding to each class. These are then placed as input to the series of one-versus-all multi-channel CNN for the identification of multiple objects present in the image.

2 Proposed Methodology

As already mentioned, a multi-labelled LCC technique has been introduced in this manuscript through a deep semi-supervised class template matching to operate in situations where there is a severe scarcity of labelled images. Here, class templates have been generated through the 'most confident' training images in each of the classes; whereas, the label dependencies have been captured using the class-correlation probability matrix. A detailed training and testing methodologies have been presented in Fig. 1.

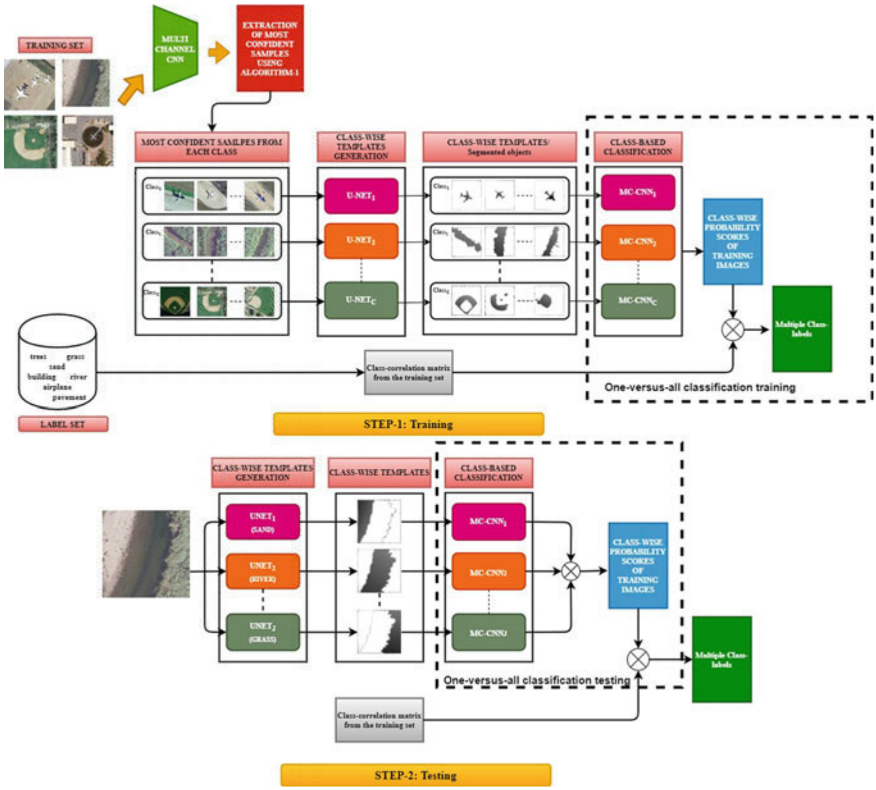

Fig. 1 Diagrammatic representation of the proposed scheme. Step-1: selection of the most confident training images, template generation and training. Step-2: Obtaining feature map of the test images and class label assignment

2.1 *Selection of the Most-Confident Images from the Training Set*

As already mentioned, a multi-channel CNN has been used to select the 'most confident' training images. For this purpose, the multi-labelled training images are used to train a multi-channel CNN, as shown in Fig. 2. In the proposed multi-channel CNN architecture, a global average pooling operation is performed after each block to covert the feature matrix into a one-dimensional vector. These one-dimensional vectors are concatenated and then fed into fully connected layers for classification where available multi-labels have been supplied for the purpose of training. The training images are then applied as input to the trained CNN to obtain the probability scores indicating their belongingness to each of the classes. Thereafter, a semi-supervised threshold selection algorithm [23], as shown in Algorithm 1, has been applied to find out the images having maximum belongingness to each of the classes. More

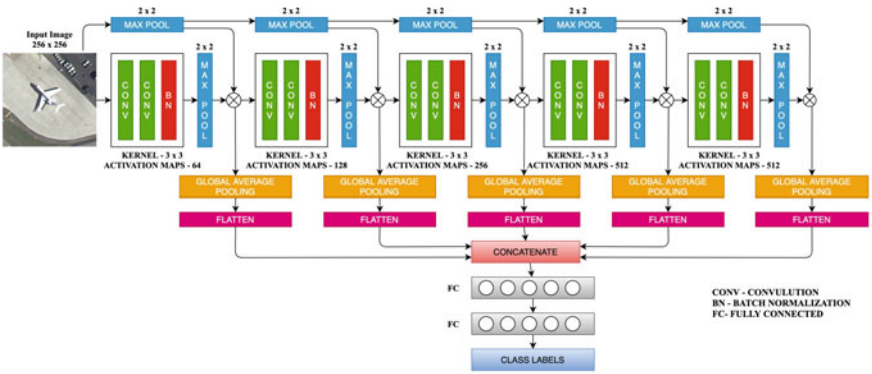

Fig. 2 Architecture of multi-channel CNN

specifically, jth image, denoted as X_j is selected amongst the 'most confident' ones in the mth class if its probability score is greater than the threshold (th_m) for the mth class. Here, the total number of class labels is denoted by C.

---

for all $m = 1, 2, ..., C$ **do**
  for all X_j in $m^{th} class$ **do**
    $min_m = \min_{j=1,2,...,\beta m} membership_j$
    $max_m = \max_{j=1,2,...,\beta m} membership_j$
    $mean_m = \frac{1}{\beta m} \sum_{j=1}^{\beta m} membership_j$
    $diffmin_m = \mid mean_m - min_m \mid$
    $diffmax_m = \mid max_m - mean_m \mid$
    if $diffmin_m > diffmax_m$ **then**
      $th_m = mean_m$
    else
      $th_m = \frac{1}{2}(mean_m + max_m)$
    end if
  end for
end for

In this phase of the proposed methodology, the most confident images from each of the classes have been obtained for further processing in the consequent steps. The set of the 'most confident' images for the mth class is denoted by $Conf_m = \{X_{jm}\}$.

2.2 *Segmentation and Generation of Class Templates*

In this stage of the investigation, the 'most confident' images obtained from the previous stage are segmented using a U-NET CNN, as shown in Fig. 3, to get segmented major object present in the image. A U-net is an unsupervised neural network con-

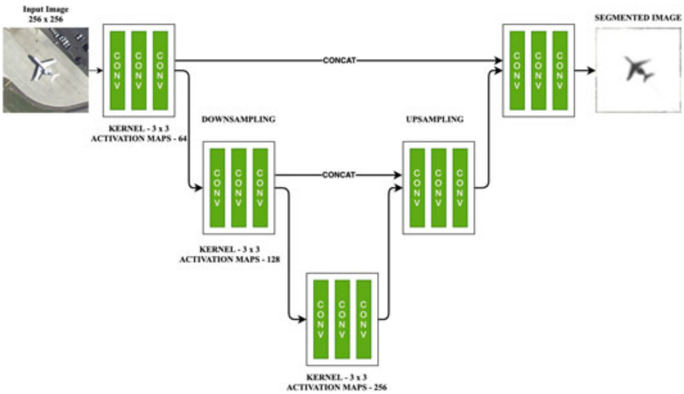

Fig. 3 U-net for image segmentation

sisting of two paths: an encoder and a decoder [22]. The encoder path, similar to a traditional stack of convolution operations, is a contraction path which means that every layer downsample the original image to capture the contents of the major object in the image. On the other hand, the decoder path symmetrically expands the image to store the localisation information from the image. As shown in the figure, the final segmented image is extracted from the final layer of this end-to-end deep network. The image is compressed in the encoder layer to get its symmetric down-scaled activation maps; the same kernel filters and activation maps of the same sizes are used during decompression (expansion) of the image. Each of the confident images from $Conf_m$ is segmented using a U-net (U-NET$_m$) specific to the mth class and hence there are C U-nets corresponding to each class for segmentation and template generation. Finally, the segmented one-class and mono-chromatic maps have been obtained for each of the confident images with respect to their corresponding class, which can now be used for training. Here, a specific U-net has been used to generate templates corresponding to each of the classes. The details of the training process have been presented in the consequent section.

2.3 One-Versus-All Training

At this stage, the segmented images have been used to train the proposed multi-channel CNN (MC-CNN) already described in Sect. 2.1. A one-versus-all training has been adopted for the present investigation, which means that there is a multi-channel CNN assigned to identify images from each of the classes, as shown in Fig. 1. More elaborately, mth MC-CNN (i.e. MCCNN$_m$) is trained with the segmented confident images (or class templates) from all the classes; it is assigned a label “1” during training if the image is a template of the mth class and “0” otherwise. During testing, a trained CNN responds to the test images with a higher probability if they contain the object (class) corresponding to the one it is trained with. However, the catch is that the class-wise probability scores obtained from the CNN have been multiplied with the class-correlation matrix obtained from the labels of the training image. As already mentioned, this is done to capture the inter-dependencies among the class labels for multi-label LCC. It is to be noted that the class-correlation matrix generated from the labels available in the training set only is used for both training and testing processes.

2.4 Assignment of Multiple Class Labels to the Test Image

During testing, an activation map corresponding to each of the classes is generated using the trained U-net. The activation map is the segmented image from the test image which is then fed to the correspondingly trained multi-channel CNN for prediction. The probability score corresponding to the positive class index is recorded and multiplied with the class correlation matrix available from the training labels. Thereafter, a semi-automatic threshold selection technique, as already described in Algorithm 1, has been applied to this modified probability score to predict the classes which are present in the test images. In other words, a label is assigned to a test image only if the modified probability score is greater than a customised automatic threshold from Algorithm 1.

3 Experimental Results

3.1 Description of the Datasets

The effectiveness of the proposed methodology has been validated through experimentation conducted on two very high resolution (VHR) remotely sensed image benchmark datasets. A detailed description of the datasets is given as follows:

3.1.1 UCM Merced Multi-label Dataset

In this dataset, images are available from 21 land-cover classes, with each class containing 100 remotely sensed aerial images [24]. Each of the images has a size of 256 $\times$ 256 $\times$ 3 with a spatial resolution of 0.3 m. Further, the images of this benchmark remote sensing dataset are captured over various regions of the United States. The multi-label ground truth information have been collected manually through photo interpretation [20]. Here, one or more labels have been assigned to each image with a maximum of seven classes. In this regard, the dataset has 17 distinct classes. For experimental purposes, the entire dataset has been divided into an 80–20 training–

test ratio. For evaluation of the proposed scheme under a severe scarcity of labelled images, the evaluation of the proposed approach has been carried out by selecting 80, 40 and 20 percent of the labelled samples for training. The samples not used for training have been used as test samples.

3.1.2 AID Dataset

AID remotely sensed benchmark dataset [25] is a combination of 10,000 images from Google Earth™ with size of each image as $600 \times 600 \times 3$. The spatial resolutions vary from 0.5 to 8 m. The image scenes have been captured over China, United States, England, France, Italy, Japan, and Germany. Out of the 10,000 images, multi-labelled ground truth information has been made available for 3,000 images through photo-interpretation [7]. Similar to the UCM dataset, here also, 80, 40 and 20

3.2 Details of the Implementation

For the implementation of multi-channel CNN, five blocks have been used. Each of the blocks further has two convolution layers followed by a batch normalisation layer to achieve faster convergence. The convolution blocks have a kernel size of 3×3 and a stride of 2×2. The number of activation maps increases across each of the convolution blocks as 64, 128, 256 and 512, respectively. Further, each block is followed by a max-pooling layer of kernel 2×2 and stride of 2×2. Here, skip connections have been introduced in the CNN architecture where there is a max-pooling layer of stride 2×2 and kernel size of 2×2. Afterwards, we fine-tune the entire network in the training phase with the Nestro Adam optimizer, and the initial learning rate is set as 0.0001. The loss is calculated with the binary cross-entropy with an input batch size of 32 for training up to 400 epochs. Moreover, early stopping [17] has been enabled to avoid over-fitting. For U-net, five convolution blocks (i.e. first two for down-sampling and the rest for up-sampling) have been used. Each of the u-net consists of five convolution blocks. Here also, the images are down-sampled using max-pooling layer of kernel size 3×3 with a stride 2×2 and the size of activation maps are 64, 128 and 256, respectively. The fine-tuning of the network has been carried out using the Nestro Adam optimizer [21] keeping the initial learning rate is set as 0.0001. Here also, the linear decay function and early stopping [17] has been utilised to avoid over-fitting of the model.

3.3 Analysis of the Results

As already mentioned, the evaluation of the proposed scheme has been carried out on two benchmark multi-label datasets of aerial (VHR) images. To validate the effectiveness of the proposed scheme in situations when there is severe scarcity of (multi)labelled images for training, the performance analysis has been carried out by varying the training–testing ratio of the datasets, as 80:20, 40:60 and 20:80. Moreover, the performance of the prposed scheme has been compared with that of data augmentation [8]-based and LSTM-based [7] approaches for multi-labelled land-cover classification. In this regard, the comparative analysis has been carried out on performance evaluation metrics [8] like precision, recall, F-score and accuracy as follows:

$$\text{Precision}= \frac{1}{n_t}\sum_{i=1}^{n_t}\frac{|T_i \cap P_i|}{|P_i|} \quad \text{Recall} = \frac{1}{n_t}\sum_{i=1}^{n_t}\frac{|T_i \cap P_i|}{|T_i|}$$

$$\text{Accuracy} = \frac{1}{n_t}\sum_{i=1}^{n_t}\frac{|T_i \cap P_i|}{|T_i \cup P_i|} \quad \text{F-score} = 2 \times \frac{\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}}$$

Here, n_t denotes the number of unlabelled images in the test set. T_i and P_i denotes the true and predicted label set of the ith image in the test set. Moreover, union ($\cup$) and intersection ($\cap$) operators denote the bit-wise OR and AND operations, respectively, on the one-hot-encoded labels.

3.3.1 Performance Evaluation on UCM Multi-labelled Dataset

As shown in Table 1, the performance of the proposed scheme has been compared with that of others in scenarios when there is a scarcity of multi-labelled images in the training set. It can be noted that all the compared schemes show sufficiently high accuracies in captioning the test images when 80% of the labelled images have been used for training. In this regard, the proposed scheme has obtained an improvement of ≈10–12%, ≈7–11%, ≈9–12% and ≈12–13% in precision, recall, F1-score and accuracy, respectively. However, by reducing the number of training images to 40% of the total images, the improvement enhances using the proposed scheme. The same improvement in terms of the four metrics then stand at ≈13–15%, ≈7–9%, ≈ 10–13% and ≈12–15% for the proposed scheme. The proposed scheme is shown to be effective when the training percent is further reduced to 20%. In such a scenario, the proposed scheme outperforms the other compared techniques by ≈14%, ≈ 10–17%, ≈12–15% and ≈15–17% in terms of precision, recall, F1-score and accuracy.

3.3.2 Performance Evaluation on AID Multi-labelled Dataset

To further validate the effectiveness of the proposed scheme, experimentation has been carried out on another benchmark multi-label remotely sensed dataset of AID, as shown in Table 2. On a standard training–test ratio of 80:20, the improvement in case of the proposed scheme stands at ≈11–13%, ≈8–16%, ≈9–15% and ≈10–17% in terms of precision, recall, F1-score and accuracy. However, the same improvement is shown to be ≈12–16%, ≈11–16%, ≈8–13% and ≈11–15% on decreasing the number of training images to 40%. On further reducing the training images to

Table 1 Comparison of results on UCM multi-labelled dataset

Percentage of images used for training (%)	Scheme	Precision (%)	Recall (%)	F1-score (%)	Accuracy (%)
80	Augmentation based [8]	80.78	82.39	81.58	71.57
	LSTM based [7]	78.52	78.68	78.59	70.68
	Proposed scheme	**91.11**	**89.45**	**90.27**	**83.12**
40	Augmentation based [8]	75.95	77.14	76.54	65.18
	LSTM based [7]	73.89	73.45	73.66	62.88
	Proposed scheme	88.04	84.42	86.19	77.30
20	Augmentation based [8]	68.00	71.04	69.49	56.68
	LSTM based [7]	68.62	64.78	66.64	54.57
	Proposed scheme	82.50	81.33	81.91	71.82

around 20%, the proposed scheme still manages an improvement of ≈9–16%, ≈11–20%, ≈10–18% and ≈6–10% in terms of precision, recall, F1-score and accuracy, respectively, over the other compared schemes.

4 Conclusion

The manuscript deals with a deep learning-based technique for multi-label land-cover classification when a few multi-labelled images available for training. This technique has been developed for the identification of multiple land-classes that may be present in an aerial (VHR) image without the necessity of extensively collected manually labelled images for training. Experiments carried out on two benchmark multi-labelled remotely sensed datasets show the effectiveness of the proposed scheme when the number of labelled images in the training set is sufficiently less. A comparative analysis of the proposed approach carried out against the other state-of-the-art techniques suggests a better performance of the former when there is a scarcity of manually labelled training images.

Table 2 Comparison of results on AID multi-labelled dataset

Percentage of images used for training (%)	Scheme	Precision (%)	Recall (%)	F1-score (%)	Accuracy (%)
80	Augmentation based [8]	81.48	84.90	83.16	78.43
	LSTM based [7]	78.32	76.80	77.70	71.32
	Proposed scheme	92.73	92.60	92.90	88.90
40	Augmentation based [8]	75.95	77.81	76.81	71.77
	LSTM based [7]	71.11	72.41	71.70	66.48
	Proposed scheme	**87.95**	**88.22**	**88.09**	**80.12**
20	Augmentation based [8]	73.09	71.48	72.27	66.39
	LSTM based [7]	66.47	62.31	64.55	62.44
	Proposed scheme	82.70	82.79	82.74	72.82

References

1. Meher SK, Kumar DA (2015) Ensemble of adaptive rule-based granular neural network classifiers for multispectral remote sensing images. IEEE J Sel Topics App Earth Obser Remote Sens 8(5):2222–2231
2. Ma L, Liu Y, Zhang X, Ye Y, Yin G, Johnson BA (2019) Deep learning in remote sensing applications: a meta-analysis and review. ISPRS J Photogramm Remote Sens 152:166–177
3. Kim M, Lee J, Han D, Shin M, Im J, Lee J, Quackenbush LJ, Gu Z (2018) Convolutional neural network-based land cover classification using 2-D spectral reflectance curve graphs with multitemporal satellite imagery. IEEE J Sel Topics Appl Earth Obser Remote Sens 11(12):4604–4617
4. Ammour N, Bashmal L, Bazi Y, Al Rahhal MM, Zuair M (2018) Asymmetric adaptation of deep features for cross-domain classification in remote sensing imagery. IEEE Geosci Remote Sens Lett 15(4):597–601
5. Zhang L, Shi Z, Wu J (2015) A hierarchical oil tank detector with deep surrounding features for high-resolution optical satellite imagery. IEEE J Sel Topics Appl Earth Obser Remote Sens 8(10):4895–4909 Oct
6. Karalas K, Tsagkatakis G, Zervakis M, Tsakalides P (2016) Land classification using remotely sensed data: going multilabel. IEEE Trans Geosci nd Remote Sens 54(6):3548–3563
7. Hua Y, Mou L, Zhu XX (2019) Recurrently exploring class-wise attention in a hybrid convolutional and bidirectional lstm network for multi-label aerial image classification. ISPRS J Photogramm Remote Sens 149:188–199
8. Stivaktakis R, Tsagkatakis G, Tsakalides P (2019) Deep learning for multilabel land cover scene categorization using data augmentation. IEEE Geosci Remote Sens Lett 16(7):1031–1035

9. Hua Y, Mou L, Zhu XX (2020) Relation network for multilabel aerial image classification. IEEE Trans Geosci Remote Sens
10. Brinker K, Hüllermeier E (2007) Case-based multilabel ranking, pp 702–707
11. Learning multi-label scene classification. Pattern Recogn, 37
12. Zhang M-L, Zhou Z-H (2007) Ml-knn: a lazy learning approach to multi-label learning. Pattern Recogn 40(7):2038–2048
13. Schapire RE, Singer Y (2000) Boostexter: a boosting-based system for text categorization. Mach Learn 39(2–3):135–168
14. Zhang M-L, Zhou Z-H (2006) Multilabel neural networks with applications to functional genomics and text categorization. IEEE Trans Knowl Data Eng 18(10):1338–1351
15. Nasierding G, Kouzani AZ (2010) Empirical study of multi-label classification methods for image annotation and retrieval. In: 2010 international conference on digital image computing: techniques and applications, pp 617–622
16. Moranduzzo T, Melgani F, Mekhalfi ML, Bazi Y, Alajlan N (2015) Multiclass coarse analysis for UAV imagery. IEEE Trans Geosci Remote Sens 53(12):6394–6406
17. Haykin S (2007) Neural networks: a comprehensive foundation. Prentice-Hall of India, New Delhi
18. Zeggada A, Melgani F, Bazi Y (2017) A deep learning approach to UAV image multilabeling. IEEE Geosci Remote Sens Lett 14(5):694–698
19. Koda S, Zeggada A, Melgani F, Nishii R (2018) Spatial and structured SVM for multilabel image classification. IEEE Trans Geosci Remote Sen 56(10):5948–5960
20. Chaudhuri B, Demir B, Chaudhuri S, Bruzzone L (2018) Multilabel remote sensing image retrieval using a semisupervised graph-theoretic method. IEEE Trans Geosci Remote Sens 56(2):1144–1158
21. Goodfellow I, Bengio Y, Courville A (2016) Deep Learning. MIT Press
22. Ronneberger O, Fischer P, Brox T, U-net: Convolutional networks for biomedical image segmentation. In: Medical image computing and computer-assisted intervention, pp 234–241, Cham, Springer International Publishing
23. Chakraborty S, Roy M, Melgani F (2019) Semisupervised two-level fusion-based autoencoded approach for low-cost domain adaptation of remotely sensed images. IEEE Geosci Remote Sens Lett 16(7):1041–1045 July
24. Yang Y, Newsam S (2010) Bag-of-visual-words and spatial extensions for land-use classification. In: Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems, pp 270–279
25. Xia G, Hu J, Hu F, Shi B, Bai X, Zhong Y, Zhang L, Lu X (2017) Aid: a benchmark data set for performance evaluation of aerial scene classification. IEEE Trans Geosci Remote Sens 55(7):3965–3981

Role of Individual Samples in Modified Possibilistic *c*-Means Classifier for Handling Heterogeneity Within Mustard Crop

Mragank Singhal, Ashish Payal, and Anil Kumar

Abstract In raster remote sensing images within class have variations represented as heterogeneity. Pixel-based classifiers use means/variance-covariance (DVC) statistical parameters, generated from training sample datasets. These parameters do not represent in totality about variations within class. This research paper explains the role of each sample in handling heterogeneity without using statistical parameters from the training samples. Modified possibilistic *c*-means fuzzy algorithm capable of mapping single class to handle heterogeneity has been experimented. Multi-spectral temporal images of Sentinel-2A/B of Banasthali, Rajasthan region acquired from 1 November 2019 to 24 February 2020 have been used for mustard class mapping. It has been observed that while using individual samples in place of statistical parameters in fuzzy-based classifiers, individual class identified has been least affected due to heterogeneity within class.

Keywords Heterogeneity · Soft classification · Modified possibilistic *c*-means · Maximum sample

1 Introduction

Earth's observation by remote sensing methods can be used for a wide range of quantitative measurements. These measurements can be related to vegetation canopy structure or for different LULC applications. The incorporation of human interpretations with various sensors to examine digital images has resulted in lower quantitative accuracy. Digital image classification is one of the prominent application domains to map and extort data about remote areas from satellite imagery. Higher accuracy

M. Singhal (✉) · A. Payal
Guru Gobind Singh Indraprastha University, Delhi, India

A. Payal
e-mail: ashish@ipu.ac.in

A. Kumar
IIRS, ISRO, Dehradun, India
e-mail: anil@iirs.gov.in

A. Tiwari et al. (eds.), *Soft Computing for Problem Solving*, Advances in Intelligent Systems and Computing 1393, https://doi.org/10.1007/978-981-16-2712-5_2

can be achieved with the intervention of computers to process a digital image [16]. Lillesand and Kiefer [14] have mentioned digital image classification as a quantitative technique to classify image data into various categories [14]. Supervised and unsupervised image classifications are two broad categories of classification procedure [6]. When training data are available, supervised classification is widely used and when training data are unavailable, unsupervised classification is incorporated in remotely sensed imagery.

A digital image is a combination of pixels. In the case of hard classification, every pixel is assumed to be pure and is classified into one class. In reality, mixed pixel problems exist and each pixel, through fuzzy classification, can be assigned to multiple class memberships. A fuzzy-based classifier was observed to be generally effective in handling mixed pixels, to produce precise and reasonable outcomes from image classification [7]. Land cover varieties are not consistent, instead are dissimilar because of which the classes have no crisp boundaries. This becomes a prime reason behind the evolution of fuzzy-based classifiers. Another reason is that a pixel may comprise two or more classes because of the coarse or medium spatial resolution of sensors. Bezdek presented a fuzzy *c*-means (FCM) algorithm with the thought of fuzzy sets to solve mixed pixel problem [3]. Later, to overcome the drawbacks of FCM, Krishnapuram proposed an algorithm based on possibilistic concept and improvement in objective function which was labeled as PCM [11]. Hybridization of different methodologies like entropy-based, contextual-based and many more with these classifiers has been eminent in fields of study. The applicability of artificial intelligence in SAR imagery processing has been evaluated using automated SAR image processing (ASIP) system [9].

To classify nonlinear data, a kernel-based fuzzy *c*-means (KFCM) was developed in the year 2007. For KFCM, sample data that appear to be nonlinear in the input space are mapped to a higher dimensional feature space where the sample points are considered to be linearly separable [20]. Because of the ability of KFCM methods to cluster more shapes in the input dataset, their classification accuracies are much higher as compared to FCM [20]. Ben-hur discussed different types of kernels, such as stationary kernels and definite kernels [1]. There is an improvement in the accuracy with the incorporation of eight kernels with fuzzy *c*-mean classification, in order to handle nonlinearity among the classes [5].

Similarly, possibilistic *c*-means (PCM) has been modified with KPCM by replacing the Euclidean norm with the Gaussian kernel, resulting in an increase in robustness to noise [10]. Supervised noise clustering has been chosen as the base classifier, and adding nine different kernel functions as the distance functions with it leads to derive a kernel-based classifier, termed as, KNC [18].

The objective function of PCM was revised by Li et al. [12], and an efficient clustering algorithm named modified possibilistic *c*-means (MPCM) was presented by them [12]. This algorithm saves the amount of running time by eliminating the computation of membership parameters in every iteration. MPCM not only has the properties of PCM, i.e., resisting noisy and avoiding trivial solutions, but also is one of the fast clustering methods.

PCM and MPCM algorithms are capable of mapping specific classes of interest from temporal datasets [15, 19]. Mapping of various stages of a specific class requires temporal data to incorporate changes happening within the class. Also, single date images may have spectral overlap while mapping the second/third level of classification.

This research paper experimented to use individual samples as statistical parameters of fuzzy MPCM classifier, in place of mean. The objective of this experiment was to handle heterogeneity within class. The proposed approach has been compared while using only mean or mean with variance–covariance (DVC) parameters in the MPCM classifier. Secondly, the spectral overlap between classes like wheat, mustard and grass has been handled while using the temporal indices database. This temporal indices database for the mustard class has been generated during the class-based sensor-independent (CBSI)-NDVI approach and compared with NDVI temporal database.

2 Vegetation Indices and Heterogeneity

Many scientists have extracted and modeled various vegetation biophysical variables using remotely sensed data since 1960. A lot of efforts have been put toward the development of vegetation indices and are defined as dimensionless, radiometric measures that function as indicators of relative abundance and activity of green vegetation. Although there are more than 20 different vegetation indices in use, in this research work NDVI was considered. Cohen [8] suggests that the first true vegetation index was the *simple ratio* (SR), which is the near-infrared (NIR) to red reflectance ratio described in Birth and Mc-Vey as mentioned in Eq. (2.1) [4]:

$$\mathrm{SR} = \frac{\mathrm{NIR}}{\mathrm{Red}} \tag{2.1}$$

Rouse developed the generic *normalized difference vegetation index* (NDVI) [17] as mentioned in Eq. (2.2):

$$\mathrm{NDVI} = \frac{\mathrm{NIR} - \mathrm{red}}{\mathrm{NIR} + \mathrm{red}} \tag{2.2}$$

The NDVI was widely used and applied to the original Landsat MSS digital remote sensor data. A novel class-based sensor-independent indices (CBSI) image has been generated to reduce spectral dimensionality of the dataset. The CBSI-NDVI formula is mentioned in Eq. (2.3):

$$\text{CBSI-NDVI} = \frac{\rho_{\max} - \rho_{\min}}{\rho_{\max} + \rho_{\min}} \tag{2.3}$$