A Comparative Study of Three Image Representations for Population Estimation Mining Using Remote Sensing Imagery

Kwankamon Dittakan[1], Frans Coenen[1], Rob Christley[2], and Maya Wardeh[2]

[1] Department of Computer Science,
University of Liverpool, Liverpool, L69 3BX, United Kingdom
[2] Institute of Infection and Global Health,
University of Liverpool, Leahurst Campus,
Chester High Road, CH64 7TE Neston, Cheshire, United Kingdom
{dittakan,coenen,robc,maya.wardeh}@liverpool.ac.uk

Abstract. Census information regarding populations is important with respect to many governmental activities such as commercial, health, communication, social and infrastructure planning and development. However, the traditional census collection methods (ground surveys) are resource intensive in terms of both cost and time. The resources required for census collection activities are particularly high in rural areas which feature poor communication and transport networks. In this paper the interpretation of high-resolution satellite imagery is proposed as a low cost (but less accurate) alternative to obtaining census data. The fundamental idea is to build a classifier that can label households according to "Family size" which can then be used to generate a census estimation. The challenge is how best to translate the raw satellite data into a form that captures key information in such a way that an appropriate classifier can still be built. Three different representations are considered: (i) Colour Histogram, (ii) Local Binary Pattern and (iii) Graph-based. The representations were evaluated by generating census information using test data collected from a rural area to the northwest of Addis Ababa in Ethiopia.

Keywords: Satellite Image Analysis and Mining, Data Mining Applications, Population Estimation Mining

1 Introduction

Census data collection is an important governmental activity with respect to a wide range of applications such as the provision of public services (education, health care, transportation and water and electricity supply). The traditional approaches to census collection are typically undertaken using either questionnaires or door-to-door surveys. In the case of questionnaires each householder within a given region receives a questionnaire either by post or electronically, members of the household complete the questionnaire and return it (again either by post or electronically). In the case of door-to-door surveys field staff visits every household so as to "interview" the household. Once the data has been collected it needs to be processed, another resource intensive activity.

Whatever the case, traditional census collection methods are expensive in terms of both time and cost. The resource required is exacerbated with respect to rural area where communication and transport networks tend to not be as sophisticated as in urban areas.

In this paper an alternative approach to census collection is proposed whereby population counts are obtained remotely using a classifier applied to satellite imagery. The approach has particular application in the context of sparsely populated rural areas. The fundamental idea is to train a classifier, using labelled satellite imagery of households, which can then be applied to obtain large area population estimates (census data). The proposed approach offers a number of advantages: (i) the resource required, in terms of time and cost, is significantly less than in the case of the traditional methods; (ii) it is efficient (traditionally census data is typically collected following a five or ten year cycle, using the proposed approach census data can be collected as required); (iii) it is non-intrusive in the sense that populations do not know it is happening (populations tend to be suspicious of census collection activities); and (iv) the collected census data is immediately in a format that allows for further processing. Although our proposed approach offers many advantages the drawback is that it is not as accurate as the traditional ("on ground") census collection methods. However, as will be demonstrated later in this paper the error margin is not significant.

The main challenge for satellite image classification (and it can be argued for image mining in general) is how best to represent satellite image data so that machine learning techniques can be applied. The amount of data available in a satellite image collection is typically too large for it to be used in its entirety thus a representation is required that serves to capture the salient information but in a reduced form. At the same time the representation needs to be compatible with the classification process which is usually founded on a feature vector representation. Three satellite image representations are therefore considered in this paper: (i) Colour Histogram, (ii) Local Binary Pattern (LBP) and (iii) Graph-based.

The rest of this paper is organised as follows. In Section 2 some related work is briefly presented. Section 3 then provides a description of the proposed census mining framework. A brief overview of the proposed image segmentation process is presented in Section 4. In Section 5 details of each of the proposed image representations are presented. The performance of the proposed census mining framework, using test data collected from a rural area to the northwest of Addis Ababa in Ethiopia, is then considered in Section 6. Finally Section 7 provides a summary and some conclusions.

2 Previous Work

Population estimation has been the subject of significant research work. Population estimation using satellite images offers the benefit over alternative population estimation methods in that the information is: (i) up-to-date, (ii) (at least to an extent) reliable and (iii) obtained at low cost [1]. Satellite imagery has been used with respect to population estimation in many ways. For example the relationship between light frequency (from night-time satellite imagery) and the distribution of populations has been investigated in [18]. Correlation analyse was used in [2] to analyse the relationship between light intensity from night-light satellite images and population density. Correlation analy-

sis was also applied in [14] to study the connection between remote sensing variables (such as spectral signatures and vegetation indices) and population density with respect to block regions in the city of Indianapolis in The USA. In [10] satellite mage texture features were extracted to provide for a population density estimation model. In [19] Multi-Spectral satellite images data was used to obtain population estimates with respect to the city of Belo Horizonte in Brazil. However, the concept of applying machine learning to satellite imagery for population estimation, to the best knowledge of the authors, has not been previously considered.

A necessary precursor for satellite image classification, and image classification in general, is that the key properties or characteristics of the image set are first extracted. This is typically achieved using some feature extraction method. The idea is to transform the input image into a set of features to produce a feature vector representation. Image features can be categorise into two groups: (i) general feature and (ii) domain-specific feature. General features are application independent and included features such as colour, texture and structure. Domain-specific features are application dependent, for example elements of the human face as used in face recognition. The work presented in this paper is founded on three basic general features: (i) Colour, (ii) Texture and (iii) Structure. Colour features are typically used in the context of colour image representation methods [8], for example using colour histogram. Colour histograms are obtained by discretising the image colours and counting the number of pixels in each of the *quantisation bins* [22]. Colour histograms offer advantages of: (i) rotation and translation invariance, (ii) computational simplicity and (iii) low storage requirements [11]. Texture features give information about the spatial arrangement of colours and/or intensities in a given image. There are many texture extraction methods that may be adopted such as the use of Gray Level Co-occurence Metrics (GLCMs), texture spectrums and wavelet transforms [9, 24]. The texture extraction method of interest with respect to the work described in this paper is the use of Local Binary Patterns (LBPs), as first proposed in [16], where each pixel is define using the relative grayscale of its neighbourhood pixels [15]. Texture features tend to be robust with respect to image rotation, illumination change and occlusion [21]. Structure features are commonly used to describe the "geometry" of the relative position of the elements containing within a given image. Image structure is often represented using a graph representation [20, 7]. Examples of graph-based structural image feature representations include Attributed Graphs (AGs), Function Describe Graphs (FDGs) and Quadtrees. The advantages of graph-based representations are their general applicability [3] and their invariance to rotation and translation [12].

3 Census Mining Framework

An overview of the proposed process for population estimation mining is presented in this section. A schematic of the framework is shown in Figure 1. The framework comprised three phases (as represented by the rectangular boxes): (i) image segmentation, (ii) image representation and (ii) classification. The input to the first phase is a large scale satellite image of a given area covering (say) a number of villages. This image is first coarsely segmented and then finely segmented in order to identify individ-

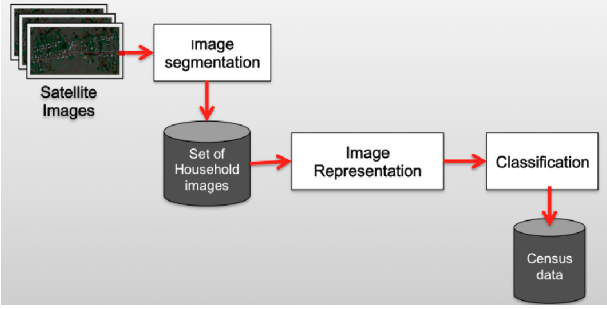

Fig. 1. Proposed population estimation mining from satellite imagery framework

ual households. This segmentation process was described in detail in [5]; however, for completeness, a brief overview of the process is presented in Section 4 below. During image representation the identified household pixel data is translated into a representation that allows for the application of a classifier. Three alternative representations are considered and compared in this paper, these are described in Section 5. After the households have been segmented and appropriately represented the classification phase may be commenced. This is relatively straight forward given that established techniques can be adopted, a back propagation neural network learning method was uitiised with respect to the evaluation presented in Section 6.

4 Image Segmentation

An overview of the adopted satellite image segmentation process is presented in this section (further detail can be found in [5]). The aim of the segmentation is to identify individual households from within satellite imagery. Figure 2 illustrates the image segmentation process. This process consists of three individual steps: (i) coarse segmentation, (ii) image enhancement and (iii) fine segmentation. During the first step (the left rectangle in Figure 2) the input satellite images of a given area are coarsely divided into a series of sub-images each covering a small group of households (between 1 and 4). Each sub-image is then enhanced (the middle rectangle in Figure 2) so as to facilitate the fine segmentation. During the fine segmentation process (the right rectangle in Figure 2) the enhanced coarse segmented sub-images are further segmented so as to identify individual households. The end result is a set of individual household images. (Note that the image segmentation process was implement using the MATLAB, matrix laboratory, workbench[3].)

5 Image Representation

This section considers the three different household image representations which form the central theme of this paper: (i) Colour Histogram, (ii) Local Binary Pattern (LBP) and (iii) Graph-Based. Each is discussed in the flowing three sub-sections, Sub-Sections 5.1 to 5.3 respectively.

[3] http://www.mathworks.com

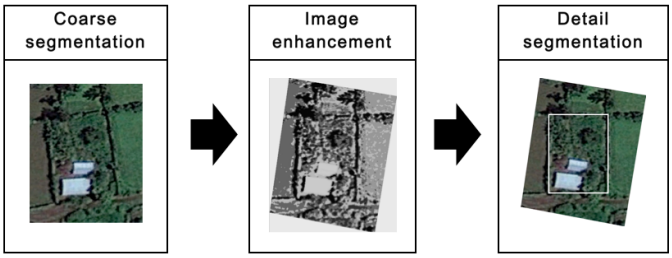

Fig. 2. Schematic illustration the segmentation process (Phase 1 of the overall process)

5.1 Colour Histogram

As noted in Section 2 colour histograms offer advantages of computational simplicity and tolerance against small changes such as rotation, translation and scaling. Colour histograms are wildly used in the context of image analysis and understanding, and hence were adopted for consideration with respect to the work described in this paper. More specifically the colour histogram concept was used to represent individual households in terms of colour distribution. The X-axis of the histogram comprised a number of "bins" such that each bin represented a "colour range". The Y-axis of the histogram then represented the number of pixels falling into each bin. Seven different type of histogram were considered: (i) red channel histogram, (ii) green channel, (iii) blue channel, (iv) hue channel, (v) saturation channel, (vi) value channel and (vii) grayscale. Each histogram was comprised 32 bins, giving a total of 224 ($7x32$) features. In addition 13 colour statistical distribution metrics were extracted, these are listed in Table 1. Thus a feature vector of length 237 ($224 + 13 = 237$) could be generated for each household image.

Table 1. Additional colour based statistical features.

#	RGB color channel descripton	#	HSV color channel description	#	Grayscale channel description
1	Average red	6	Average hue	11	Mean of grayscale
2	Average green	7	Average saturation	12	Standard deviation of grayscale
3	Average blue	8	Average value		
4	Mean of RGB	9	Mean of HSV	13	Average of grayscale histogram
5	Standard deviation of RGB	10	Standard deviation of HSV		

5.2 Local Binary Pattern

As already noted in Section 2, LBPs have also been wildly adopted since they are easy to calculate, robust to illumination changes and are an effective texture description mechanism. Consequently the use of LBPs was also considered with respect to the work described in this paper. In the context of the individual household images each image was first transform into grayscale. A 3×3 pixel window was then considered. For each window the grayscale value of the centre pixel, the threshold value, was compared with surrounding eight neighbour values. A 1 was reccorded if the grayscale value was greater than the threshold and a 0 otherwise. The eight binary digits recorded in this way then

formed an eight bit number, the LBP for the window. Using LBPs in this way $2^8 = 256$ different patterns can be formed. If the number of sampling points (neighbours) is P and the pixel distance from the centre point is R. The notation LBP(P,R) is used to describe variations of the LBP approach. Three different variations were considered: (i) 8 sampling points with a radius of 1 (LBP$(8,1)$), (ii) 8 sampling points with a radius of 2 (LBP$(8,2)$) and (iii) 8 sampling points with a radius of 3 (LBP$(8,3)$). Again additional statistical features were used to augment the LBP representation. Three categories of texture statistic were identified: (i) entropy features (E), (ii) grey-level occurrence matrix features (M) and (iii) wavelet transform features (W). Table 2 lists the statistical features generated (the letter in parenthesis in each case indicates the category of the feature). Thus the LBP based representation could result in feature vectors of length 266 ($2^R + 10 = 2^8 + 10 = 256 + 10 = 266$).

Table 2. Additional texture based statistical features.

#	Description	#	Description	#	Description
1	Entropy (E)	7	Average approximation	10	Average diagonal
2	Average Local Entropy (E)		coefficient matrix, *cA* (W)		coefficient matrix, *cD* (W)
3	Contrast (M)	8	Average horizontal		
4	Correlation (M)		coefficient matrix, *cH* (W)		
5	Energy (M)	9	Average vertical		
6	Homogeneity (M))		coefficient matrix, *cV* (W)		

5.3 Graph-Based Structure

Graph based techniques allow for the structural representation of image objects. In the context of the census estimation application a quad-tree representation was adopted on the grounds that this has been frequently used with respect to many other image mining applications. The desired graph-based representation, for each household image, was generated using a four step process: (i) quadtree decomposition, (ii) tree construction, (iii) frequent subgraph mining and (iv) feature vector transformation. Each household image was first "cropped" to a size of 128x128 pixel (best appropriate size) surrounding the main building of the household by detecting the roof of the building (the largest contiguous homogenous region within the image). Then the cropped household images were recursively subdividing into four quadrate until a minimum 8×8 pixel decomposition was arrived at. During the second step the decomposition was cast into a quad tree data structure. Each node was labelled with one of 32 ranged potential mean grayscale values for the associated quadrant. A set of identifiers, $\{1,2,3,4\}$, representing the north west (NW), north east (NE), south west (SW) and south east (SE) quadrants of a decomposition respectively were used as edge labels.

Although graph-based representations provide for a powerful and flexible structural image model they do not readily lend themselves to input with resect to classification algorithms [17]. However, we can use frequent sub-graph mining techniques to identify frequently occurring sub-graphs across the collection of household image quad trees and consider these frequent sub-graphs to be features within a standard feature space representation. There were a number of different frequent subgraph mining approaches than could have been adopted. From the literature the most frequently referenced frequent sub-graph mining algorithm is the gSpan algorithm [23]. This was

therefore adopted with respect to the work described in this paper. Gspan uses the concept of a support threshold σ to differentiate a frequently occurring sub-graph from an infrequently occurring sub-graph. If the occurrence count (the *support*) for a subgraph is greater than σ then the subgraph is identified as a frequently occurring subgraph.

6 Evaluation

In this section, the evaluation and comparison of the proposed population estimation mining approaches is presented. Extensive evaluation has been conducted so as to compare the operation of the three different representations and their variations, however this paper only reports the most significant results obtained (there is insufficient space to allow for the presentation of all the results obtained). For evaluation purposes, so as to compare the operation of the three different representations and their variations, a back propagation neural network learning method (multilayer perceptron) was used for the classification coupled with Ten-fold Cross Validation (TCV). In each case the effectiveness of the classification was recorded in terms of: (i) accuracy (AC), (ii) area under the ROC curve (AUC), (iii) precision (PR), (iv) sensitivity (SN) and (v) specificity (SP). Note also that χ-squared feature selection was applied to reduce the overall size of the feature space in each case so that the top k features could be selected (the most appropriate value for k was found to be 35). The remainder of this section is organised as follows. Sub-section 6.1 gives an overview of the test data used. Sub-section 6.2 then itemises the results obtained.

6.1 Test Data

To act as a focus for the research a case study was considered directed at a rural area within the Ethiopian hinterland, more specifically two data sets were collected with respect to two villages (Site A and Site B) located within the Horro district in the Oramia Region of Ethiopia (approximately 300 km north-west of Addis Abba) as shown in Figure 3[4]. Site A was bounded by the parallels of latitude 9.312650N and 9.36313N, and the meridians of longitude 37.123850E and 37.63914E and Site B was by the parallels of latitude 9.405530N and 9.450000N, and the meridians of longitude 36.590480E and 37.113550E. Using the know bounding latitudes and longitudes of our two test sites appropriate satellite imagery was extracted from Google Earth. The images were originally obtained using the GeoEye satellite with a 50 centimetre ground resolution. The satellite images for Site A were obtained during the rainy season (lot of green colouring) and were released by Google Earth on 22 August 2009 and those for Site B were obtained during the dry season (lot of brown colouring) and released on 11 February 2012. On-ground household data (including family size and household latitude and longitude) was collected by University of Liverpool field staff in May 2011 and July 2012. The minimum and maximum family size were 2 and 12 respectively, the mean was 6.31, the medium were 6 and standard deviation was 2.56. These two data sets then provided the training and test data required to evaluate our proposed census collection system.

[4] https://maps.google.co.uk/

The benchmark household family size was separated into three classes: (i) Small family size (less than or equal to 5 people), (ii) Medium family size (between 6 and 8 people in the family) and (iii) Large family size (more than 8 people). Some statistics concerning the class distributions for the Site A and B data sets are presented in Table 3.

Fig. 3. The test site location: Horro district in Ethiopia

For the colour histogram representation technique, three data sets were produced: colour histogram (CH), colour statistics (CS) and combine (CH+CS). For the LBP representation seven data sets were produced: LBP(8,1), LBP(8,2), LBP(8,3), texture statistics (TS), combine (LBP(8,1) + TS), combine (LBP(8,2) + TS) and combine (LBP(8,3) + TS). For the graph-based representation a range of σ values, $\{10, 20, 30, 40, 50\}$, were used. The number of sub-graphs generated in each case is presented in Table 4. From the table it can be seen that, as would be expected, the number of identified subgraphs decreases as the value for σ increases (and vice-versa).

Table 3. Class label distribution for data set Site A and data set Site B

Sites	Small family	Medium family	Large family	Total
Site A	28	32	10	70
Site B	19	21	10	50
Total	47	53	20	120

Table 4. Number of identified sub-graphs produced using a range of σ values with respect to the Site A and B data

Support threshold	Site A	Site B
$\sigma = 10$	757	420
$\sigma = 20$	149	119
$\sigma = 30$	49	60
$\sigma = 40$	24	39
$\sigma = 50$	12	19

6.2 Results

The obtained results are presentation in Table 5. Based on these results the following observations can be made:

- With respect to the colour histogram based representation best results were obtained using CH (AUC = 0.754) for Site A and CH + CS (AUC = 0.765) for Site B.
- With respect to the LBP representation best results tended to be produced using LBP(8,1) for both Site A (AUC = 0.88) and Site B (AUC = 0.792).

- With respect to the graph-based structure representation, $\sigma = 10$ produced the best results with respect to both Sites A and B (AUC = 0.730 and 0.744 respectively).

Table 5. Comparison of different variations of the proposed histogram, LPB and Graph-based representations in terms of classification performance

Types	Techniques	Site A					Site B				
		AC	AUC	PR	SN	SP	AC	AUC	PR	SN	SP
Histogram	CH	0.614	0.754	0.615	0.614	0.761	0.560	0.753	0.568	0.560	0.723
	CS	0.400	0.550	0.404	0.400	0.629	0.480	0.645	0.466	0.480	0.693
	CH+CS	0.557	0.741	0.551	0.557	0.713	0.580	0.755	0.602	0.580	0.728
LBP	LBP(8,1)	0.718	0.880	0.758	0.771	0.856	0.600	0.792	0.599	0.600	0.764
	LBP(8,2)	0.614	0.765	0.618	0.614	0.725	0.600	0.705	0.600	0.600	0.762
	LBP(8,3)	0.586	0.718	0.583	0.586	0.710	0.660	0.672	0.679	0.660	0.774
	TS	0.414	0.502	0.379	0.414	0.695	0.500	0.650	0.602	0.500	0.754
	LBP(8,1)+TS	0.757	0.848	0.754	0.757	0.847	0.600	0.768	0.601	0.600	0.757
	LBP(8,2)+TS	0.643	0.770	0.645	0.643	0.746	0.580	0.736	0.580	0.580	0.742
	LBP(8,3)+TS	0.557	0.706	0.553	0.557	0.686	0.620	0.638	0.639	0.620	0.748
Graph-based	$\sigma = 10$	0.586	0.730	0.582	0.586	0.739	0.580	0.744	0.584	0.580	0.756
	$\sigma = 20$	0.500	0.588	0.499	0.500	0.669	0.480	0.606	0.483	0.480	0.698
	$\sigma = 30$	0.400	0.495	0.391	0.400	0.607	0.480	0.648	0.479	0.480	0.726
	$\sigma = 40$	0.357	0.397	0.344	0.357	0.542	0.320	0.521	0.344	0.320	0.656
	$\sigma = 50$	0.314	0.506	0.330	0.314	0.588	0.280	0.459	0.309	0.280	0.643

To conduct a further performance comparison, with respect to the recorded AUC values, the Friedman's test was applied [6]. The AUCs for all 15 classification techniques (Colour Histogram (3), LBP (7) and Graph-based (5)) for both the Site A and the Site B data are listed in Table 6 (columns two and three); the number in parentheses in each case indicates the overall "ranking" of each individual result with respect to the two data sets. The average rank (AR) is given in the fourth column, this is the mean value of the rankings for each classification technique. The Friedman test statistic is based on the AR values, and is calculated using equation 1, where N is the number of data sets (2), K is the total number of classification technique considered (15) and r_i^j is the rank of classification technique j on data set i.

$$\chi_F^2 = \frac{12N}{K(K+1)} \left[\sum_{j=1}^{K} AR_j^2 - \frac{K(K+1)^2}{4} \right] \tag{1}$$

The Friedman test statistic and corresponding p value are presented in the first row of Table 6 . The Friedman test statistic (25.70) and the significance threshold ($p < 0.005$) indicate that the null hypothesis, that there is no difference between the techniques, can be rejected. The highest recorded AUC value for each data set is indicated in bold font in Table 6 indicating that the LBP(8,1) representation produced the overall best performance (AR = 1.0), while the graph-based representation with $\sigma = 40$ produced the worst overall result (AR = 14.5).

A post hoc Nemanyi test [4, 13] was applied to each class distribution to evaluate the performance the approaches. The performance between individual approaches are

Table 6. Area Under the Curve (AUC) results

Friedman test statistic = 25.70 ($p < 0.005$)			
Techniques	Site A	Site B	AR
CH	0.754 (5)	0.753 (4)	4.5
CS	0.550 (11)	0.645 (11)	11
CH+CS	0.741(6)	0.755 (3)	4.5
LBP(8,1)	**0.880** (1)	**0.792** (1)	1
LBP(8,2)	0.765 (4)	0.705 (7)	5.5
LBP(8,3)	0.718 (8)	0.672 (8)	8
TS	0.502 (13)	0.650 (9)	11
LBP(8,1)+TS	0.848 (2)	0.768 (2)	2
LBP(8,2)+TS	0.770 (3)	0.736 (6)	4.5
LBP(8,3)+TS	0.706 (9)	0.638 (12)	10.5
$\sigma = 10$	0.730 (7)	0.744 (5)	6
$\sigma = 20$	0.588 (10)	0.606 (13)	11.5
$\sigma = 30$	0.495 (14)	0.648 (10)	12
$\sigma = 40$	0.397 (15)	0.521 (14)	14.5
$\sigma = 50$	0.506 (12)	0.459 (15)	13.5

significantly different if their ARs are different by more than a Critical Difference (CD) value that can be calculated using equation 2 where the critical difference level α ($\alpha = 0.05$) and the value $q_{\alpha,\infty,K}$ is based on the Studentised range statistic.

$$CD = q_{\alpha,\infty,K}\sqrt{\frac{K(K+1)}{12N}} \tag{2}$$

The significant diagram in Figure 4 illustrates the AUC performance rank of the proposed image classification technique with Nemenyi's critical difference (CD) tail (the calculated CD for the diagram is 10.72). The diagram shows the classification techniques listed in ascending order of ranked performance on the y-axis, and the image classification techniques's average rank across all two data sets displayed on the x-axis. The LBP(8,1) approach produced the best results with $AR = 1$, which was significantly better than graph-based representation ($\sigma = 30, 50, 40$) with $AR = 12$, 13.5 and 14.5, respectively.

7 Conclusion

A framework for population estimation mining using satellite imagery has ben proposed. The framework can operate with a number of different representations, three different categories of representation were considered: (i) Colour Histogram, (ii) Local Binary Pattern (LBP) and (iii) Graph-based. These were used to encode individual segmented household images so that a classifier generation process could be applied. Experiments were conducted using the test data collected from two villages located in a rural part of the Ethiopian hinterland. The main findings with respect to the evaluation (using a neural network learning method and a χ-squared top k feature selection algorithm with $k = 35$) were: (i) with respect to the Colour Histogram based representations the best result was obtained using CH for the Site A (rainy season) data and CH + CS

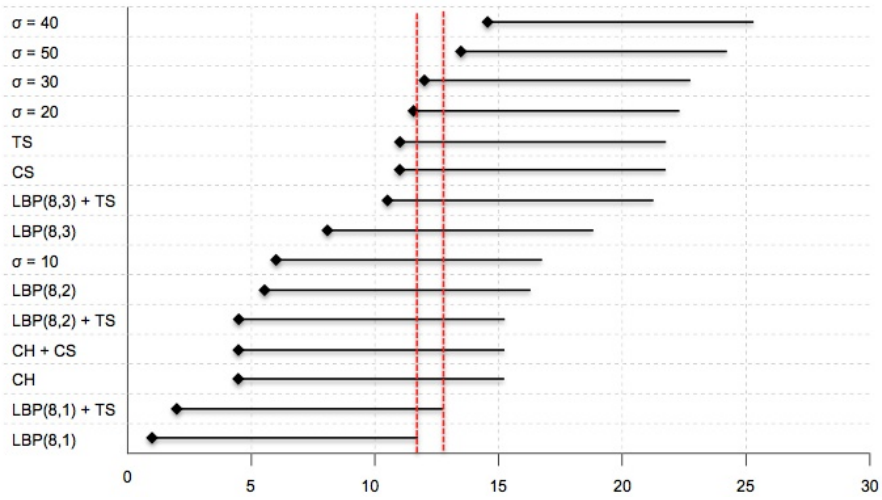

Fig. 4. AR comparison between the proposed image classification approach using Nemenyi's post hoc test with $\alpha = 0.05$

for the Site B (dry season) data, (ii) with respect to the LBP representation the best result was produced using LBP(8,1) for both Site A and Site B, and (iii) with respect to the Graph-based structure representation, $\sigma = 10$ produced the best results with respect to both the Site A and Site B data. However, the LBP(8,1) representation produced the best overall result. For future work the research team intend to conduct a large scale population estimation exercise using the proposed framework.

References

1. A.S. Alsalman and A.E. Ali. Population Estimation From High Resolution Satellite Imagery: A Case Study From Khartoum. *Emirates Journal for Engineering Research*, 16(1):63–69, 2011.
2. L. Cheng, Y. Zhou, L. Wang, S. Wang, and C. Du. An estimate of the city population in china using dmsp night-time satellite imagery. In *Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International*, pages 691–694, 2007.
3. D. J. Cook and L. B. Holder. Substructure discovery using minimum description length and background knowledge. *ournal of Artificial Intelligence Research*, 1(1):231–255, 1994.
4. J. Demšar. Statistical comparisons of classifiers over multiple data sets. *J. Mach. Learn. Res.*, 7:1–30, December 2006.
5. K. Dittakan, F. Coenen, and R. Christley. Towards the collection of census data from satellite imagery using data mining: A study with respect to the ethiopian hinterland. In Max Bramer and Miltos Petridis, editors, *Proc. Research and Development in Intelligent Systems XXIX*, pages 405–418. Springer, London, 2012.
6. M. Friedman. A Comparison of Alternative Tests of Significance for the Problem of m Rankings. *The Annals of Mathematical Statistics*, 11(1):86–92, 1940.
7. E. Ganea and M. Brezovan. Graph object oriented database for semantic image retrieval. In *ADBIS (Local Proceedings)*, pages 151–164, 2010.

8. J. Han and K. Ma. Fuzzy color histogram and its use in color image retrieval. *Image Processing, IEEE Transactions on*, 11(8):944–952, 2002.
9. C.C. Hung, M. Pham, S. Arasteh, B.C. Kuo, and T. Coleman. Image texture classification using texture spectrum and local binary pattern. In *IEEE International Conference on Geoscience and Remote Sensing Symposium, 2006. IGARSS 2006.*, pages 2750–2753, 31 2006-aug. 4 2006.
10. Y. Javed, M.M. Khan, and J. Chanussot. Population density estimation using textons. In *Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International*, pages 2206–2209, 2012.
11. S. Jeong, C.S. Won, and R.M. Gray. Image retrieval using color histograms generated by gauss mixture vector quantization. *Computer Vision and Image Understanding*, 94(3):44–66, 2004.
12. S. Jouili and S. Tabbone. Towards performance evaluation of graph-based representation. In *Graph-Based Representations in Pattern Recognition - 8th IAPR-TC-15 International Workshop, GbRPR 2011, Münster, Germany, May 18-20, 2011. Proceedings*, pages 72–81, 2011.
13. S. Lessmann, B. Baesens, C. Mues, and S. Pietsch. Benchmarking classification models for software defect prediction: A proposed framework and novel findings. *IEEE Trans. Software Engineering.*, 34(4):485–496, July 2008.
14. G. Li and Q. Weng. Using Landsat ETM+ imagery to measure population density in Indianapolis, Indiana, USA. *Photogrammetric Engineering and Remote Sensing*, 71(8):63–69, 2005.
15. P. Liang, S.F. Li, and J.W. Qin. Mulit-resolution local binary patterns for image classification. In *Wevelet Analysis and Pattern Recognition, 2010 20th International Conference on*, pages 164–169, 2010.
16. T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):971–987, 2002.
17. B. Ozdemir and S. Aksoy. Image classification using subgraph histogram representation. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 1112–1115, 2010.
18. F. Pozzi, C. Small, and G. Yetman. Modeling the distribution of hunman population with night-time satellite imagery and gridded population of the world. In *Pecora 15/Land Satellite Information IV/ISPRS Comission I/FIEOS. Conference on*, 2002.
19. I.A. Reis, V.L. Silva, and E.A. Reis. Adjusting population estimates using satellite imagery and regression models. In *Anais XV Simpsio Brasileiro de Sensoriamento Remoto - SBSR*, pages 830–837, 5 April 2011 2011.
20. A. Sanfeliu, R. Alquézarb, J. Andrade, J. Climent, F. Serratosa, and J. Vergésa. Graph-based representations and techniques for image processing and image analysis. *Pattern Recognition*, 35(3):639–650, 2002.
21. C. Song, P. Li, and F. Yang. Multivariate texture measured by local binary pattern for multispectral image classification. In *Proc. IEEE International Conference on Geoscience and Remote Sensing Symposium (IGARSS'06)*, pages 2145–2148, 2006.
22. Michael J. Swain and Dana H. Ballard. Color indexing. *Int. J. Comput. Vision*, 7(1):11–32, November 1991.
23. X. Yan and Jiawei Han. gspan: graph-based substructure pattern mining. In *Data Mining, 2002. ICDM 2003. Proceedings. 2002 IEEE International Conference on*, pages 721–724, 2002.
24. Y. Zhang, R. He, and M. Jian. Comparison of two methods for texture image classification. In *Computer Science and Engineering, 2009. WCSE '09. Second International Workshop on*, pages 65–68, 2009.