Chapter 7

Conclusions and Future research directions

7.1 General

In this chapter, conclusions are drawn based on the experiments performed using the three datasets described in chapter 5, section 5.2.1. The outcome of this study also suggests a considerable potential for extending the investigations for the development of advanced strategies for improving classification accuracy over urban area. In this direction, topics that are worthy of further studies are also suggested under section ‘scope of future research’.

The principal findings of this work are central to the application of decision tree and textural features in RS classification. The conclusion is made on the performance and employability of the decision tree classification algorithm in respect of training dataset size, class hierarchy levels I and II, and effect of ancillary data on tree complexity (number of rules induced). Further, the effect of GLCM based individual texture measures and their combinations, window size and interpixel distance in texture feature extraction, and evaluation of TD and JM as texture features selection criteria in parametric and non-parametric classification are also dealt under this section.

Another major issue which affects classification in remote sensing is the performance of image fusion algorithm and reliability of employing image quality assessment measures in image fusion. Even though the study carried on image fusion is a pre-requisite in this research, the observations made in respect of the three commonly available image fusion algorithms: the multiplicative technique, Brovey transform and principal component analysis are worth to be noted in the context of classification. Hence, in addition to visual evaluation, the fusion techniques are also analysed quantitatively using TD, histogram-mean, entropy, correlation coefficients to study the spectral distortion and amount of information transferred onto the fused image. Finally, these evaluation parameters are assessed for their reliability in classification by comparing them with the obtained OCA.
7.2 Conclusions

7.2.1 Performance study of image fusion algorithms and evaluation of the reliability of quality assessment measures

The fused images are ranked as PCA (2.69) > BT (2.54) > MT (2.48) according to the diagonal average of the inter-correlation coefficients between all the three bands of fused images and the corresponding bands of MS image. At the same time, it is also interesting to note that the average of the correlation coefficients of three bands of the three fused images with panchromatic band computed for five broad LU/ LC features of the semi-urban area ranks the images in the order of MT (2.86) > BT (2.81) > PCA (2.30). When the histogram was studied to compute the mean, thereby the average band brightness on comparison with MS bands, the fusion techniques are found to follow the order PCA> BT> MT. A closer visual look reveals that the BT fused image is sharper than the PCA and MT; and MT stands low in this quality. The entropy values ranked the fused images as PCA> BT> MT. The best average separability obtained for all the 11 classes indicates that BT (TD: 1905.79) and MT (TD: 1903.92) exhibit better separability in comparison to the PCA (TD: 1886.11). BT also exhibits higher average separability (TD: 1129.19) over MT (TD: 1085.24) and PCA (TD: 1009.45) technique for some of the most important semi-urban LU/ LC classes. The OCA of MLC ranks the images as BT (64.72%) > MT (64.34%) > PCA (62.18%). The prediction accuracy obtained through 10-fold cross validation in DTC ranks the images as MT (83.42%) > BT (83.29%) > PCA (80.35%).

- From the above findings and in the present context of classification of semi-urban land features, it is apparent that the best fusion algorithms are the BT and MT in comparison to PCA. The OCA obtained in both BT and MT fused images are equally good, whereas the OCA of PCA is about 2.25% less than the former two algorithms. Apart from the above, the PCA classified image shows poor class-discrimination between NH, interior road and sheet-roofing - the most important classes among the urban LU/ LC classes.

- The statistical evaluation criteria such as the correlation coefficient, entropy and histogram statistics computed with reference to the original multi-spectral image bands and the subsequent ranking of the fused images have failed to correctly predict the best performing fused image and the corresponding fusion algorithm in RS classification.
The suitability of the fused images in classification quantified in terms of OCA exhibits concurrence only with the distance measure criterion the TD, the correlation coefficient computed with reference to panchromatic band, and the 10-fold cross validation accuracy of the DT learning algorithm. Therefore, TD and 10-fold cross validation accuracy could be used to predict and select the best performing fused MS images in classification, provided the data dimensionality is small.

It can be concluded unequivocally from the results of classification of the three fused images that there is a great need for reliable image quality measures in determining the best fused image while dealing with fine resolution RS data.

7.2.2 Performance study and employability of decision tree classification algorithm on multi-spectral data

The implementation of DT on high resolution satellite imagery for classification of semi-urban scene was carried out and its effectiveness has been investigated in the second phase of the research. In MLC, the largest source of error is the misclassifications between classes: sheet-roofing and NH, interior road and RCC-roofing, and dry grass and open ground. They are spectrally similar and hence show a significant spectral overlap both in TD and quantitative study made on the classified images.

- The DT consistently shows improvement in OCA over MLC (DTC: 80.03% and MLC: 77.88% on MS (1) of 5.8m resolution (Table 6.11), and DTC: 65.04% and MLC: 63.45% on PMS; pan-sharpened MS (2), where DTC was trained with 14.5% of total training instances (Table 6.17)) even when the training and validation sites are changed.

- In general, larger dataset size is favourable in improving the learning and 10-fold cross validation accuracies and the overall classification accuracy of the DT classifier. But, the larger the size of the data set, the larger is the tree size and associated rule sets; however, higher accuracy is ensured.

- The learning accuracy estimated from the training data gets saturated at around 400 training instants and fails to represent the overall behaviour of the DTC and also to predict the conceivable OCA of the classifier. However, the prediction accuracy on validation instances found to be a good indicator of the achievable OCA on entire image data.
The outcome of the second phase of the study also reveals that the ancillary data derived from band transformation and band ratioing technique take decisive role in improving classification accuracy and simplifying rule sets in DTC. The DTC has experienced a two-fold benefit of improved OCA from 73.83% to 80.40% and reduced classification rules from 49 to 40 (approximately 18% less) when IHS and NDVI bands are included. The data combination MS+IHS_DT has made a significant improvement in PA and UA over MS_MLC for classes like dry grass, acacia trees and veg_mix. But NDVI has not shown significant contribution in improving accuracy of vegetation classes, and not even to any particular class in the presence of IHS. Image smoothing improved the OCA of MLC (from 63.45% to 64.81%), on the contrary, smoothing decreased the OCA in DTC (from 65.04% to 61.63). However, data smoothing benefited DTC by reducing its rule sets from 83 to 68 (18% less).

Further, DTC is found to be independent of the thumb rule defined for parametric classifiers and this has turned out to be an advantage in favour of DTC for the reason that collecting training and validation sites in-situ is tedious (inaccessible if the terrain is rough), time-consuming and a costly affair. This inconvenience of collecting training samples encourages the analysts to go for stratified random sampling, instead of stratified cluster sampling the most often demanded by the statistical classifier. Hence, when the classifiers’ performance is viewed in the light of the chosen sampling scheme, the DTC is found performing satisfactorily for stratified random sampling scheme of training instances.

7.2.3 Effectiveness of texture measures, statistical feature selection criteria and border-effect

The study demonstrated the effect of integration of spectral and texture features in the classification of semi-urban LU/ LC features on panchromatic sharpened LISS-IV data of 2.5m spatial resolution at class hierarchy level I and II. Five texture measures, viz. entropy, energy, contrast, mean and homogeneity were investigated with three-band multi-spectral data and various texture feature combinations. The feature selection criteria like TD and JM distance measures (statistical methods) were also evaluated to know their reliability. The following conclusions were drawn from the present study.

The principal finding is that combining texture and spectral features in high resolution satellite data has proven to be highly effective in improving the classification accuracy of spectrally overlapping but spatially homogenous urban LU/ LC features like RCC-roofing, asbestos sheet-roofing, interior road, NH and open grounds (play ground). In contrast to the above, texture combinations have
not made marked difference over spectrally distinct and spatially homogeneous classes like sand and water bodies in comparison to the classification of multi-spectral bands alone. Moreover, integration of texture greatly reduces the salt-and-pepper noise in classified images.

- This study on 2.5m resolution image shows that in general, a window size of 25x25 pixels with interpixel distance of 1 or 2 is found optimal for pan-sharpened LISS-IV data to achieve satisfactory classification results for more than one LU/LC features. However, the optimal window size differs at individual LU/LC class levels and from one texture feature to another. Moreover, literature survey indicates that there is no unique texture window size describing an LU/LC class as it depends on the spatial resolution of the data, the land cover spread which varies from one geographical area to another, and human cultural habitation. Therefore, the optimal window size is solely dependent on the spatial resolution of the data and the classification problems.

- The texture measures entropy and contrast, and their combination ENT+CON provide higher overall classification accuracy among all other feature combinations. The simultaneous integration of more than two texture bands in multi-spectral bands increases data dimensionality, which consequently worsens the performance of the statistical classifier, the MLC. Therefore, the decision of using texture measures and their combinations should be weighed carefully to avoid possible generation of lower accuracy in statistical classifiers.

- In contrast to the statistical classifier, the response of the DTC for combinations of texture bands clearly indicates that addition of texture enhances the class discriminating power of the DT learning algorithm. The number of rules induced in the DT classifier and the complexity of the tree model relatively get reduced as the data dimensionality is increased. This observation is also supported by the discussion made in section 7.2.2 that addition of NDVI and IHS bands into MS bands provides a two-fold benefit of improved OCA and reduced tree complexity in DTC.

- The border-effect in texture-based classification approach is unnoticeable at inter-class (transition region between various classes) boundaries but significant at image border. The approximate width of the border-effect is found to be \((N-1)/2\) pixels, where \(N\) is the size of the moving window chosen to extract texture measures.
At class hierarchy level-I, the OCA obtained for all the datasets is markedly higher than level-II. Incorporation of texture at CHL-I has improved the OCA in the range of 0.57% (with ASM: 81.37%) to 6.74% (with CON: 87.54%) in comparison to MS band at CHL-I (80.8%). Contrary to this observation, at CHL-II, even though the OCA is small for MS band (62.63%), the improvement seen upon adding texture ranges from 6.74% (with MEAN: 69.37%) to as high as 15.54% (with ENT: 78.17%). Hence, incorporation of texture is found to be more effective at higher class hierarchy levels where generally the accuracy decreases for multi-spectral data.

The transformed divergence, Jeffries-Matusita distance measures and the overall classification accuracy obtained for various combinations of texture and multi-spectral bands do not show concurrency in both MLC and DTC. Hence, it is concluded that the feature selection criteria based on the statistical distance measures do not show reliable performance in texture-based classification when large number of texture features are integrated into MS bands. Mostly, it is attributed to the fact that texture distribution is non-Gaussian in nature, whereas the adopted distance measure criteria expect data to follow a Gaussian distribution.

### 7.3 Scope of future research

Finally, this section describes some interesting issues for future work that could not be carried out in this research study due to constraints on access of data and time. This work, although limited to a single study area, illustrates many of the issues related to the feasibility of employing DTC algorithm and incorporation of texture information in classifying satellite data, particularly the fine resolution images. However, to affirm the above research findings and to rule out the possibility of the developed methodology being correct just by chance, at least more than two case studies are recommended under the developed framework over urban areas and images acquired by different sensors at different spatial resolution.

Future study may concentrate on investigating the usefulness of the GLCM based texture measures which were not taken up in this work and also the measures derived from various other methods mentioned by Tuceryan and Jain [123] and Haralick [133].
Since texture features exhibit non-Gaussian distribution, their true impact on the classification accuracy may further be explored using non-parametric classification approach; especially the DTC as it is faster than its counterparts in learning and classification.

More research is required to bring out a reliable generalised relationship between the optimal window size of various texture measures and the classes at various degrees of spatial resolutions of the images.

Incorporation of LIDAR data derived digital elevation model (DEM) or digital surface model (DSM)- as ancillary data to take advantage of height information- would show greater potential in discriminating roads, parking lots (concrete roads, asphalt and tar roads), built-up areas (roof tops with concrete, asphalt sheets, coloured fibre sheets, etc.), lawn and grass, trees, etc [178].

The pixel values (intensity levels) in DEM/ DSM data represent height information. The LIDAR derived elevation data and the MS data are to be re-sampled to bring them to a common spatial and radiometric resolution, and to co-register the data. Once both the data are available in a common format, they can be stacked together to obtain a composite data for classification.

Since image fusion techniques also contribute to the quality of classification, the other conventional approaches like IHS substitution and advanced fusion techniques employing wavelet transform, high-pass filtering (HPF) [31], Laplacian pyramid, smoothing filter-based intensity modulation (SFIM) [30], [40], etc. need to be studied on images acquired from various sensors to test their suitability in classification.

A comparative study needs to be made to understand the performance and behaviour of the DT learning algorithm at varying amount of noise in RS data.

There is also a scope to formulate a reliable feature selection measure, preferably of non-parametric type, to make data analysis time-efficient in non-Gaussian distributed environment.
7.4 Summary

It is well known that image fusion enhances visual interpretation through integrating a high resolution panchromatic image (HRPI) with a low resolution multi-spectral image (LRMI) to produce a high resolution multi-spectral image (HRMI). The fused data helps analysts to identify large number of LU/LC features for detailed study and planning of resources. On the contrary, image fusion or to generalise in a broader way - the fine resolution images, decrease classification accuracy especially in the spectrally overlapping classes. This is due to the fact that within-class spectral variance increases with increase in the spatial resolution of the image, which permits data to follow non-Gaussian distribution. These negative effects can be successfully counteracted in classification by employing texture information, which is inherent and enhanced in the fused image. But, this approach no longer holds good when a large number of texture features or add-on data are integrated into classification since they are basically non-Gaussian in nature, and hence show adverse effect on classification accuracy.

Therefore, the present research work has been taken up to devise an efficient and reliable classification strategy with an attempt to find answers to some of the conflicting issues discussed in literature pertaining to classification of fine resolution RS data. The research investigations were carried on the MS data of IRS LISS-IV sensor of 5.8m spatial resolution and pan-sharpened MS data of 2.5m spatial resolution. The classifications were carried out using the maximum likelihood classifier (representing the statistical classification algorithm) and an advanced classifier of non-parametric type based on the decision tree learning algorithm.

The study has utilised the IHS colour space transformation, NDVI band ratioing and image smoothing techniques, and GLCM texture measures to improve class discrimination power of the classifiers. Literature also indicated that there exists a trade-off between the spatial resolution improvement and the spectral property preservation of the LRMI. Hence, an attempt has been made to throw some light on the effectiveness of employing conventional image fusion algorithms like the multiplicative technique, Brovey transformation and principal component analysis methods in classification. The reliability of employing the statistical evaluation measures in assessing fusion techniques, and as feature selection criteria for texture bands in classification has also been examined.
It is apparent from the study that the decision tree classification algorithm is reasonably effective over MLC and could be a very viable approach if the ancillary data are employed judiciously in classifying high resolution satellite data, especially in urban/semi-urban areas where within-class variability is high. The research findings have also proved that integration of texture information greatly enhances the classification accuracy and makes the classified image more interpretable and free from salt-and-pepper noise.

The work has successfully developed a framework for implementation and evaluation of the decision tree classification algorithm and texture features into classification of semi urban LU/LC features at class hierarchy level I and II (6 and 11 classes respectively). The outcome of this research will help analysts better understand the behaviour of the DTC on RS data and hopefully provide answers to many of the dubious issues related to the employability of statistical quality measures in image fusion, statistical feature selection measures, performance of the statistical classifier on non-Gaussian data, effect of texture and texture window size on classified image, etc. in the course of development of advanced classification strategies.

However, research needs to be continued further in conjunction with the non-parametric classification approach in finding the potential ancillary data and texture features to make classification more accurate and time-efficient on very high resolution satellite data.

As an end note, a remark which can be made from the outcome of the thesis is that the employed image fusion technique, the classification algorithm and the add-on data (ancillary) have a direct impact on classification and are the indispensable issues in improving classification accuracy of multi-spectral data as the spatial resolution becomes finer and finer.