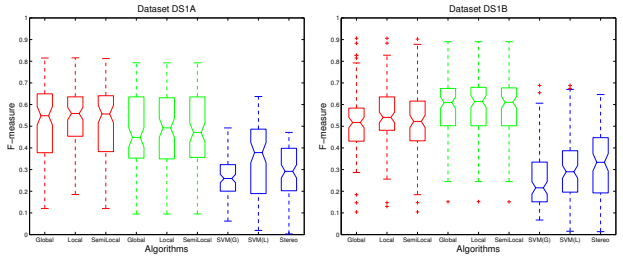
regions (see figure 4(a)). Under such circumstances, both the global and the semi-local models behave as the local model.

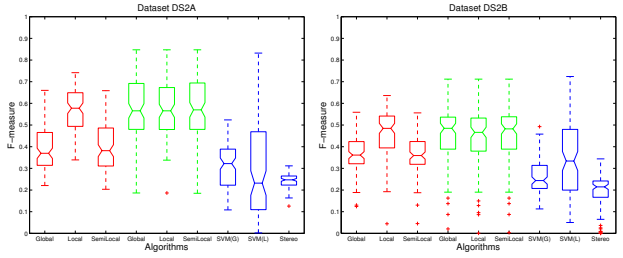
In data sets DS3A and DS3B however, the stereo labels are more fragmented. Both data sets have fairly uniform within-class colour content. However, DS3B exhibits darker lighting conditions . Figure 7(a) shows that the global and the semi-local models outperform the local model in DS3A. This reinforces our intuition that in cases where the intra class variance is low, a global data model aids the classification process. DS3B, which exhibits less vivid colours and poorer class discriminability, causes the global model to perform worse than the semi-local or local models. This effect is particularly visible when the models use graph based segmentation.

In general, it appears that the choice of the proximity model depends on the data set. We expect the local model to perform well for data sets with higher intra class variance, while the global model is expected to work better with comparatively smaller intra class variance. The semi local model should work best for cases with an intermediate amount of intra class variance. The experiments presented in this section though, remain somewhat inconclusive and fail to fully corroborate the above intuitions, primarily due to the lack of fragmented stereo labeled regions in the first four data sets.



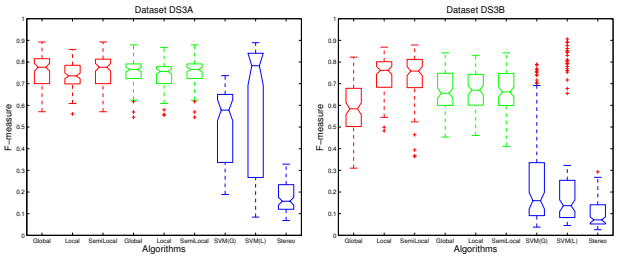
(a) (b)

Fig. 5. F_1 measure of competing algorithms across datasets DS1A and DS1B. The F_1 values for each model have been averaged over all segmentation parameter settings.



(a) (b)

Fig. 6. F_1 measure of competing algorithms across datasets DS2A and DS2B. The F_1 values for each model have been averaged over all segmentation parameter settings.



(a) (b)

Fig. 7. F_1 measure of competing algorithms across datasets DS3A and DS3B. The F_1 values for each model have been averaged over all segmentation parameter settings.

B. Analysis of segmentation results

In this sub-section we compare the two approaches to image segmentation utilized by our proposed algorithm. The primary points of investigation are: the performance of the segmentation algorithms and the sensitivity of the results to the segmentation parameter settings.

To answer these question the segmentation algorithms described in section II-B are compared against one another. To make the comparisons tangible, we make certain simplifying assumptions and fix certain parameters. The graph based segmentation's implementation is controlled by two parameters [17] the minimum number of pixels in a segment min and a parameter m which controls how aggressively pixels are merged into a segment. As suggested in [26], in our experiments we tie these two parameters together $min = m = K$ and vary K over the range $\{50, 100, 150, 200, 250, 300, 350\}$. The mean shift algorithm [18] has three tunable parameters, the spatial bandwidth (h_s), the colour bandwidth (h_r) and the smallest significant feature size (M). In the experiments presented here we fix $h_s = 7$; we found the produced segmentation to be fairly insensitive to spatial bandwidth. Furthermore, we found that fixing $h_r = 7$ and varying $M = K = \{50, 100, 150, 200, 250, 300, 350\}$ produces segmentations which are best comparable to the ones produced by the graph based segmentation. Fixing the colour bandwidth h_r results in more stable segmentations, reducing fluctuations in segmentation granularity. Our inclination toward fairly high values of h_r and M are necessary to discard the effects of small variations present in the kind of complex images that are dealt with in this paper. From here on, we refer to the tunable parameter of the mean shift algorithm, simply as K.

Our objective is to classify image far-field as accurately as possible and not necessarily to produce the best possible segmentation along the way. A segmentation is only as good as the final prediction it produces. Hence, we simply use the far-field prediction performance to evaluate the performance of a segmentation algorithm. The segmentation stability is assessed by evaluating how the performance and the number of segments varies by varying the segmentation parameters.

Figure 8 illustrates the segmentations produced by the two algorithms on representative image frames from each of the 3 environments.