

Figure 5. **Example segmentation results on Oxford-IIIT Pet dataset.** The segmentation of the pet from the background was obtained automatically as described in Sect. 3.3.

fication, in which a 37-class SVM is learned directly, using the same method discussed in Sect. 4.2. The relative performance of the different models is similar to that observed in Sect. 4.1 and 4.2. Flat classification is better than hierarchical, but the latter requires less work at test time, due to the fact that fewer SVM classifiers need to be evaluated. For example, using the appearance model with the *image*, *head*, *image-head* layouts for 37 class classification yields an accuracy of 51.23%, adding the shape information hierarchically improves this accuracy to 52.78%, and using shape and appearance together in a flat classification approach achieves an accuracy 54.03%. The confusion matrix for the best result for breed classification, corresponding to the last entry of the eight row of Table 4 is shown in Fig. 4.

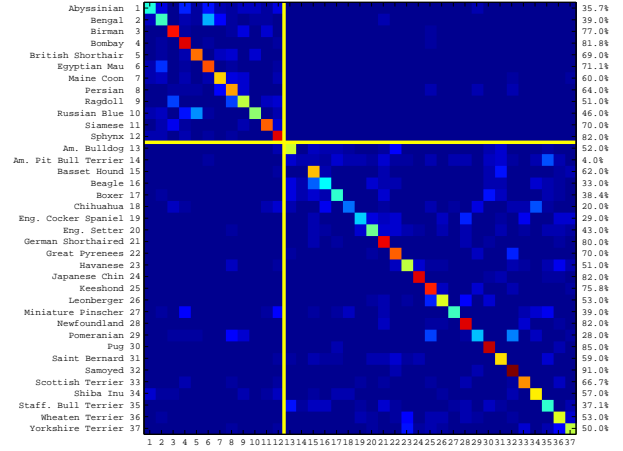


Figure 6. **Confusion matrix for breed discrimination.** The vertical axis reports the ground truth labels, and the horizontal axis to the predicted ones (the upper-left block are the cats). The matrix is normalized by row and the values along the diagonal are reported on the right. The matrix corresponds to the breed classifier using shape features, appearance features with the *image*, *head*, *body*, *body-head* layouts with automatic segmentations, and a 37-class SVM. This is the best result for breed classification, and corresponds to the last entry of row number 8 in Tab. 4.

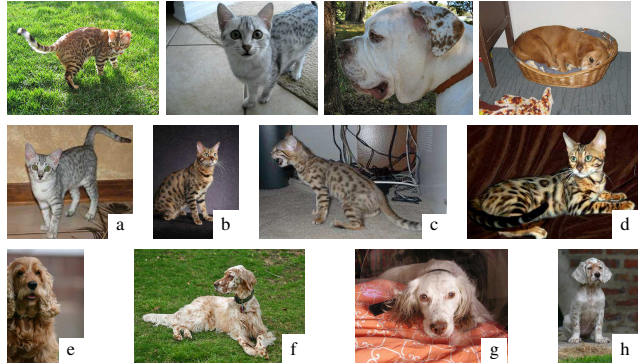


Figure 7. **Failure cases** for the model using appearance only (*image layout*) in Sect. 4.2. *First row:* Cat images that were incorrectly classified as dogs and vice versa. *Second row:* Bengal cats (b-d) classified as Egyptian Mau (a). *Third row:* English Setter (f-h) classified as English Cocker Spaniel (e).

5. Summary

This paper has introduced the Oxford-IIIT Pet dataset for the fine-grained categorisation problem of identifying the family and breed of pets (cats and dogs). Three different tasks and corresponding baseline algorithms have been proposed and investigated obtaining very encouraging classification results on the dataset. Furthermore, the baseline models were shown to achieve state-of-the-art performance on the ASIRRA challenge data, breaking the test with 42%