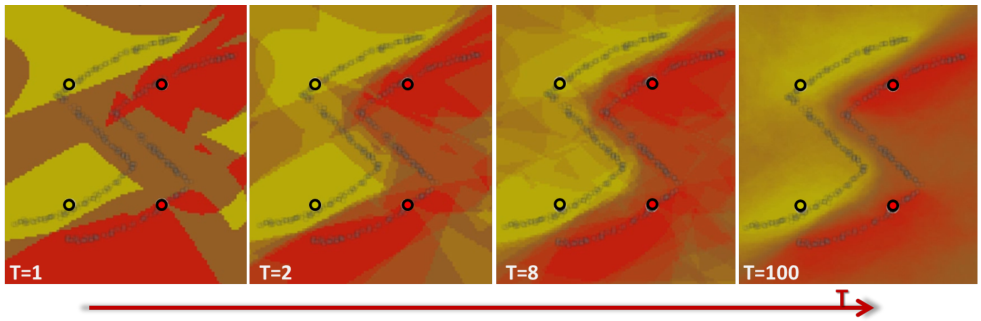

Fig. 7.3: **Learning a generic, inductive classification rule.** Output classification posteriors, tested on all points in a rectangular section of the feature space. Labelled training points are indicated by coloured circles (only four of those per image). Available unlabelled data are shown by small grey squares. Note that a purely inductive classification function would separate the left and right sides of the feature space with a vertical line. In contrast here the separating surface is "S"-shaped because affected by the density of the unlabelled points, thus demonstrating the validity of the use of unlabelled data densities. From left to right the number of trees in the forest increases from $T = 1$ to $T = 100$. See text for details.

within a rectangular section of the feature space. As expected a larger T produces smoother posteriors. Note also how the inferred separating surface is "S"-shaped because it takes into account the unlabelled points (small grey squares). Finally we observe that classification uncertainty is greater in the middle due to its increased distance from the four ground-truth labelled points (yellow and red circles).

Discussion. In summary, by using our mixed information gain and some geodesic manipulation the generic decision forest model can be readily adapted for use in semi-supervised tasks. Semi-supervised forests can be used both for transduction and (refined) induction without the need for a two-stage training procedure. Further efficiency is due to the parallel nature of forests. Both for transduction and induction the output is fully probabilistic. We should also highlight that