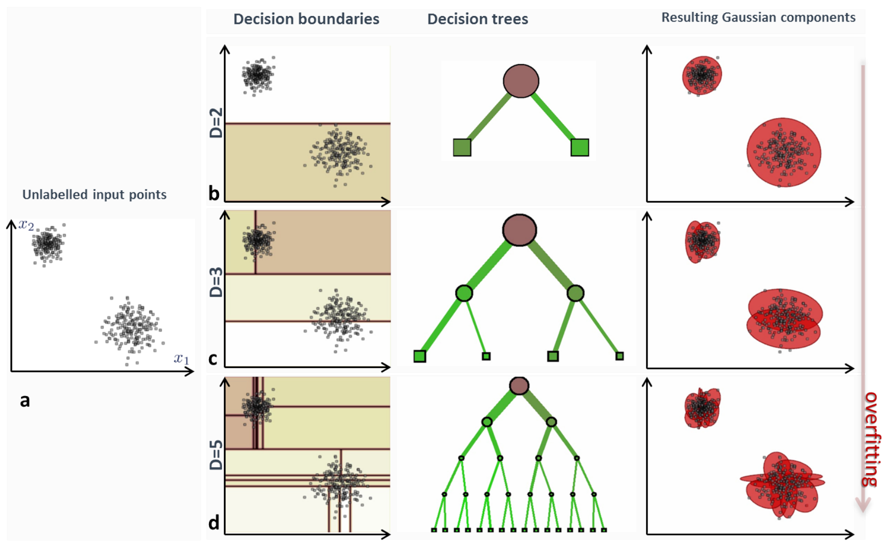

Fig. 5.4: **The effect of tree depth on density. (a)** Input unlabelled data points in a 2D feature space. **(b,c,d)** Individual trees out of three density forests trained on the same dataset, for different tree depths D. A forest with unnecessarily deep trees tends to fit to the training noise, thus producing very small, high-frequency bumps in the density.

In this simple example $D = 2$ (top row) produces the best results.

5.3.2 The effect of forest size

Figure 5.5 shows the output of six density forests trained on the input data in fig. 5.4a for two different values of T and three values of D. The images visualize the output density $p(\mathbf{v})$ computed for all points in a square subset of the feature space. Dark pixels indicate low values and bright pixels high values of density.

We observe that even if individual trees heavily over-fit (*e.g.* for $D = 6$), the addition of further trees tends to produce smoother densities. This is thanks to the randomness of each tree density estimation and reinforces once more the benefits of a forest ensemble model. The tendency of larger forests to produce better generalization has been