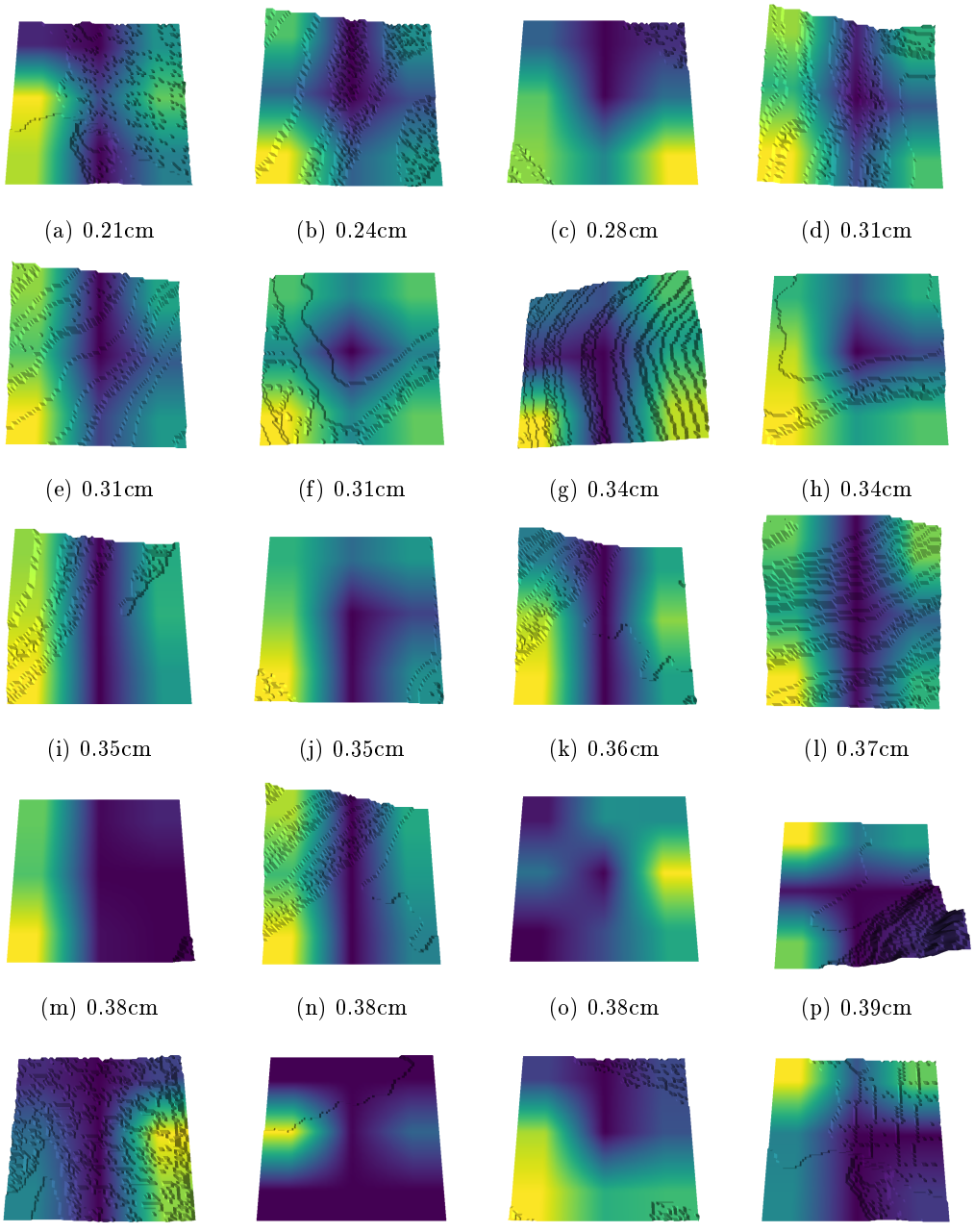
0.1 Quarry dataset

After showing the model's capability of correctly separate classes' features we utilized Grad-CAM to visualize some of the samples in the test set. We divided those inputs in four classes based on the model's performance: worst, best, false positive and folse negative. We expect the worst and the best output to be on the left and the right branch of figure **??** respectively, while the other two categories to be in the mixed points. We randomly sampled twenty inputs from those set and appplied Grad-CAM as texture on the 3D render to better visualize which region of the inputs caused the prediction.

0.1.1 Best



(a) 0.21cm (b) 0.24cm (c) 0.28cm (d) 0.31cm

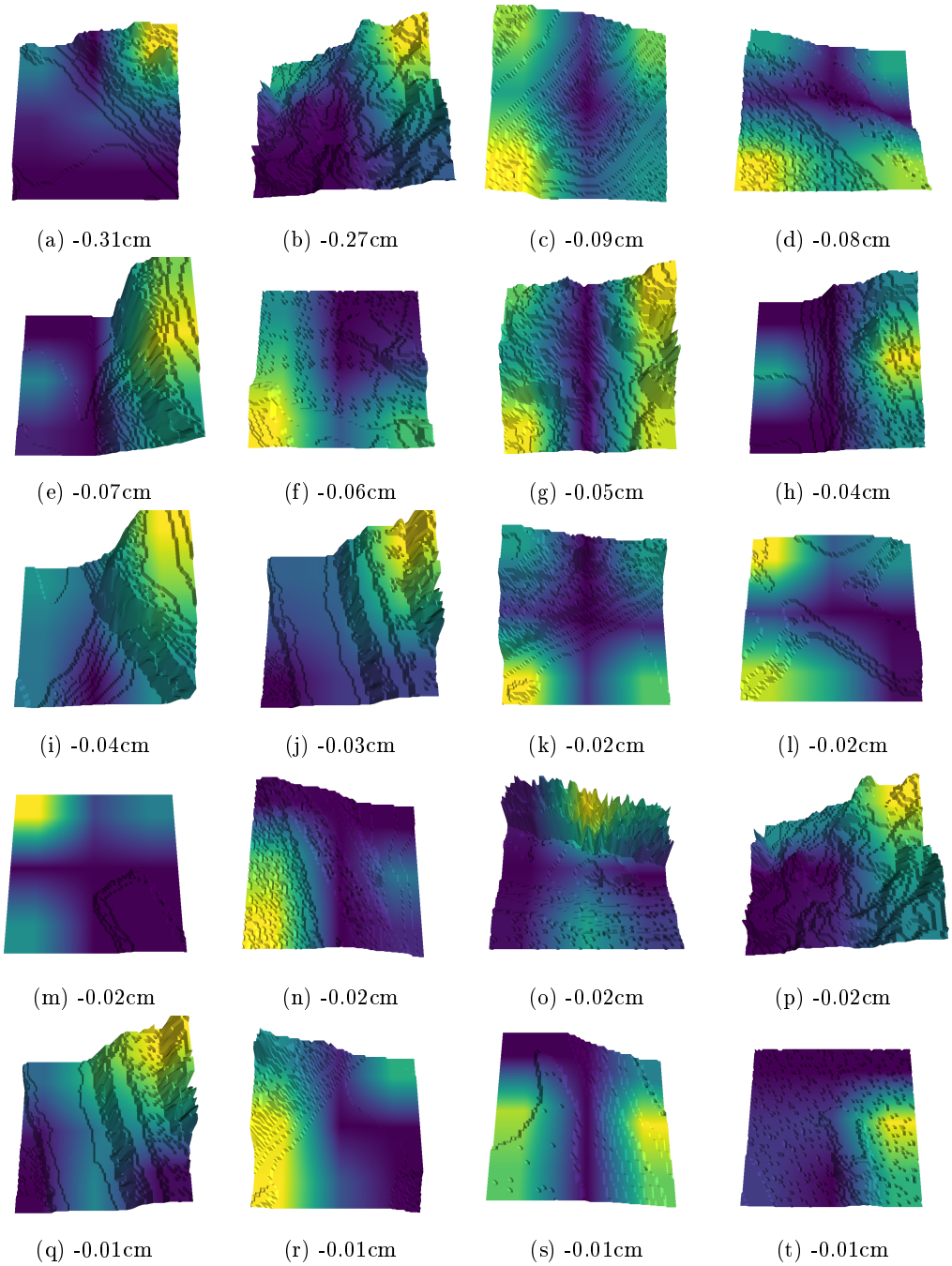
(e) 0.31cm (f) 0.31cm (g) 0.34cm (h) 0.34cm

(i) 0.35cm (j) 0.35cm (k) 0.36cm (l) 0.37cm

(m) 0.38cm (n) 0.38cm (o) 0.38cm (p) 0.39cm

(q) 0.39cm (r) 0.40cm (s) 0.40cm (t) 0.40cm

0.1.2 Worst



(a) -0.31cm (b) -0.27cm (c) -0.09cm (d) -0.08cm

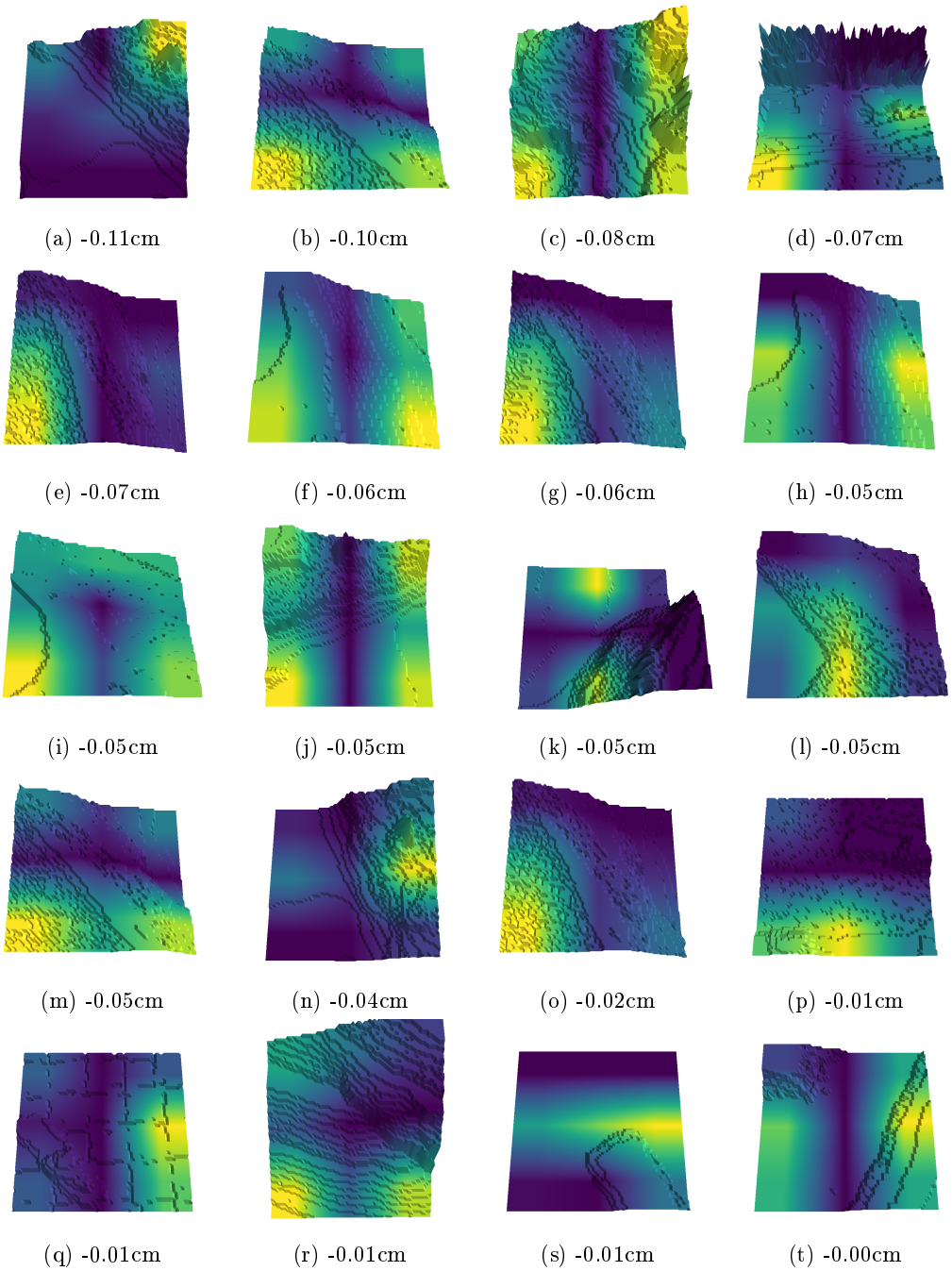
(e) -0.07cm (f) -0.06cm (g) -0.05cm (h) -0.04cm

(i) -0.04cm (j) -0.03cm (k) -0.02cm (l) -0.02cm

(m) -0.02cm (n) -0.02cm (o) -0.02cm (p) -0.02cm

(q) -0.01cm (r) -0.01cm (s) -0.01cm (t) -0.01cm

0.1.3 False Negative



(a) -0.11cm (b) -0.10cm (c) -0.08cm (d) -0.07cm

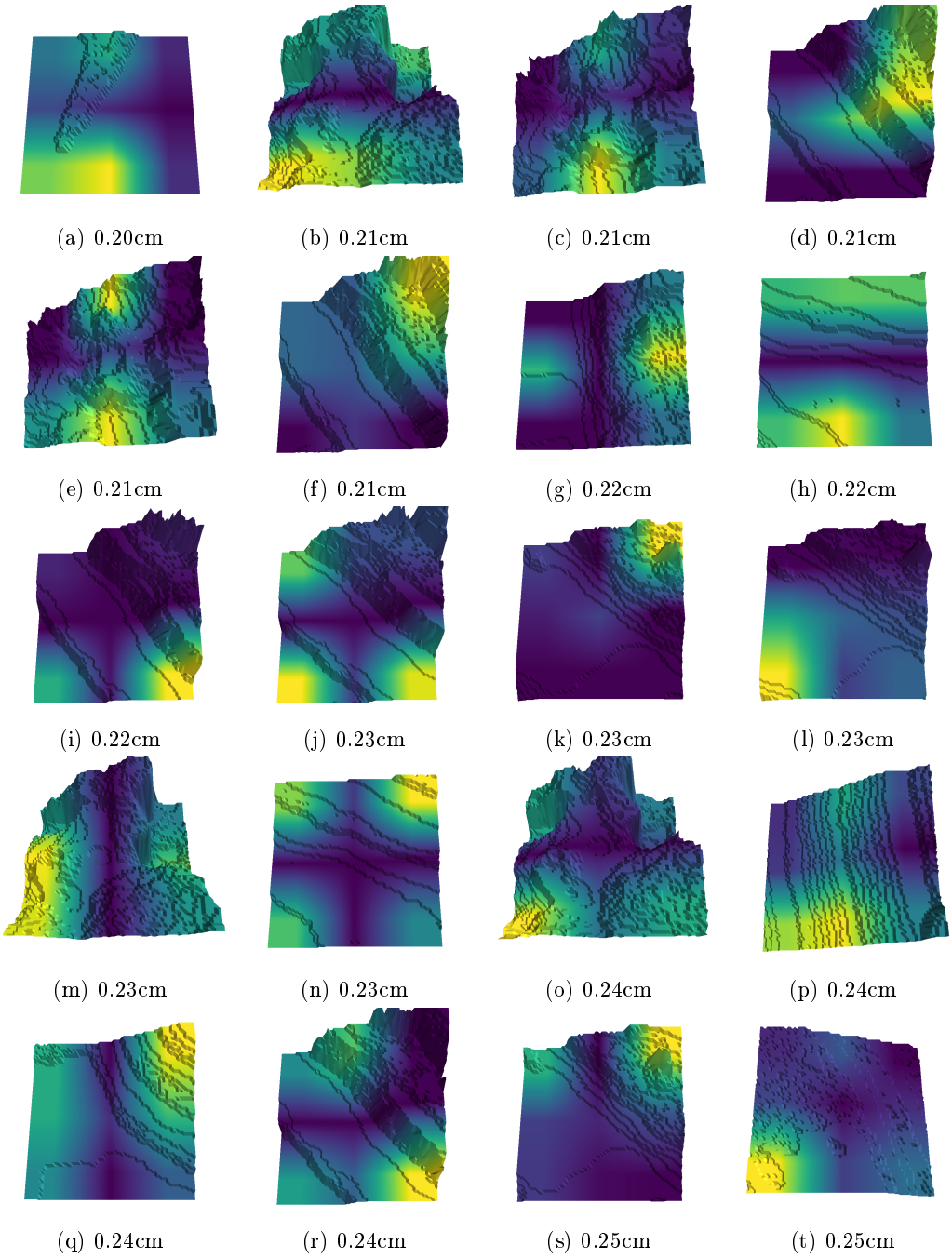
(e) -0.07cm (f) -0.06cm (g) -0.06cm (h) -0.05cm

(i) -0.05cm (j) -0.05cm (k) -0.05cm (l) -0.05cm

(m) -0.05cm (n) -0.04cm (o) -0.02cm (p) -0.01cm

(q) -0.01cm (r) -0.01cm (s) -0.01cm (t) -0.00cm

0.1.4 False Positive



(a) 0.20cm (b) 0.21cm (c) 0.21cm (d) 0.21cm

(e) 0.21cm (f) 0.21cm (g) 0.22cm (h) 0.22cm

(i) 0.22cm (j) 0.23cm (k) 0.23cm (l) 0.23cm

(m) 0.23cm (n) 0.23cm (o) 0.24cm (p) 0.24cm

(q) 0.24cm (r) 0.24cm (s) 0.25cm (t) 0.25cm