



Figure 13 Recognition confusion for providing matching feedback across all exercises and intervention study patients. Numbers in cells show class-relative accuracy (first line) and number of instances (second line).

**Accuracy of system feedback.** Feedback caused by insertion errors represented repetitions that were actually not performed by the user (no ground truth reference available). Insertions were not further considered for the feedback analysis. Deletion errors were handled and taken into account when computing feedback accuracy, however deletions are not visible in the confusion matrix. The confusion matrix illustrating the system accuracy for providing matching feedback is shown in Figure 13. Overall, an accuracy of 87.5% was achieved. Correctly performed repetitions (performance class 1) were recognized at 90%. For Too large (class 4) and Too fast and too small (class 9) a lower accuracy (63% and 64%, respectively) with regard to the healthy participants validation, was obtained. Furthermore, the distribution of the error across exercise repetitions was investigated and is shown in Figure 14. It can be observed that errors occurred mostly during first and last repetition of each exercise session (17% and 21% of the total errors, respectively). We assume that these repetitions may have been influenced by interactions of trainee and smartphone to start and stop recordings. Figure 15 further analyses the system's feedback with regard to individual exercises and patient. As the result shows, for Patient 3, lowest accuracy (84%) was obtained. Leg lifts (Ex. 5) were the most challenging to recognise, resulting in an accuracy of 85%.