3.6 Results

This section describes the results of our system validation with healthy participants and the evaluation with COPD patients. Performance of the system to detect execution errors is detailed for both investigations. For the patient study, we furthermore analysed feedback efficacy. Finally, we compared the therapist's performance in detecting erroneous exercise repetitions with the ones of our system.

3.6.1 Validation with healthy participants

We formally evaluated our system with a dataset acquired from healthy participants, using a scripted exercise and performance protocol. According to the protocol, participants were asked to perform all exercise movement errors, besides the correct execution of each exercise. Thus, the script and supervising trainer provided reference information for the ground truth used. Here we detail the recognition accuracy for all nine performance classes and six exercise classes.

Recognition confusion among performance classes. The average performance classification accuracy of the system was 96.2%, confirming that applicable feedback would have been provided for almost every exercise repetition. The confusion matrix for providing matching feedback (i.e. classify too slow when the recorded motion was slower than in the Teach-mode) for all performance classes is shown in Figure 9. The matrix illustrates very good accuracies for all performance classes, above 90%. Lowest accuracy (92%) was observed for performance class 9 (Too slow & too small), where thirteen repetitions where misclassified. A further analysis attributed these classification errors to challenges in performing Leg lifts (Ex. 5).

Recognition accuracy for exercise classes. To further analyse correctness of the feedback, Figure 10 shows the accuracy distribution across exercises and performance classes. It can be observed that most of the erroneous feedback in performance class 9 was given during exercise Leg lifts (Ex. 5), where the system showed lowest accuracy (77.5%).