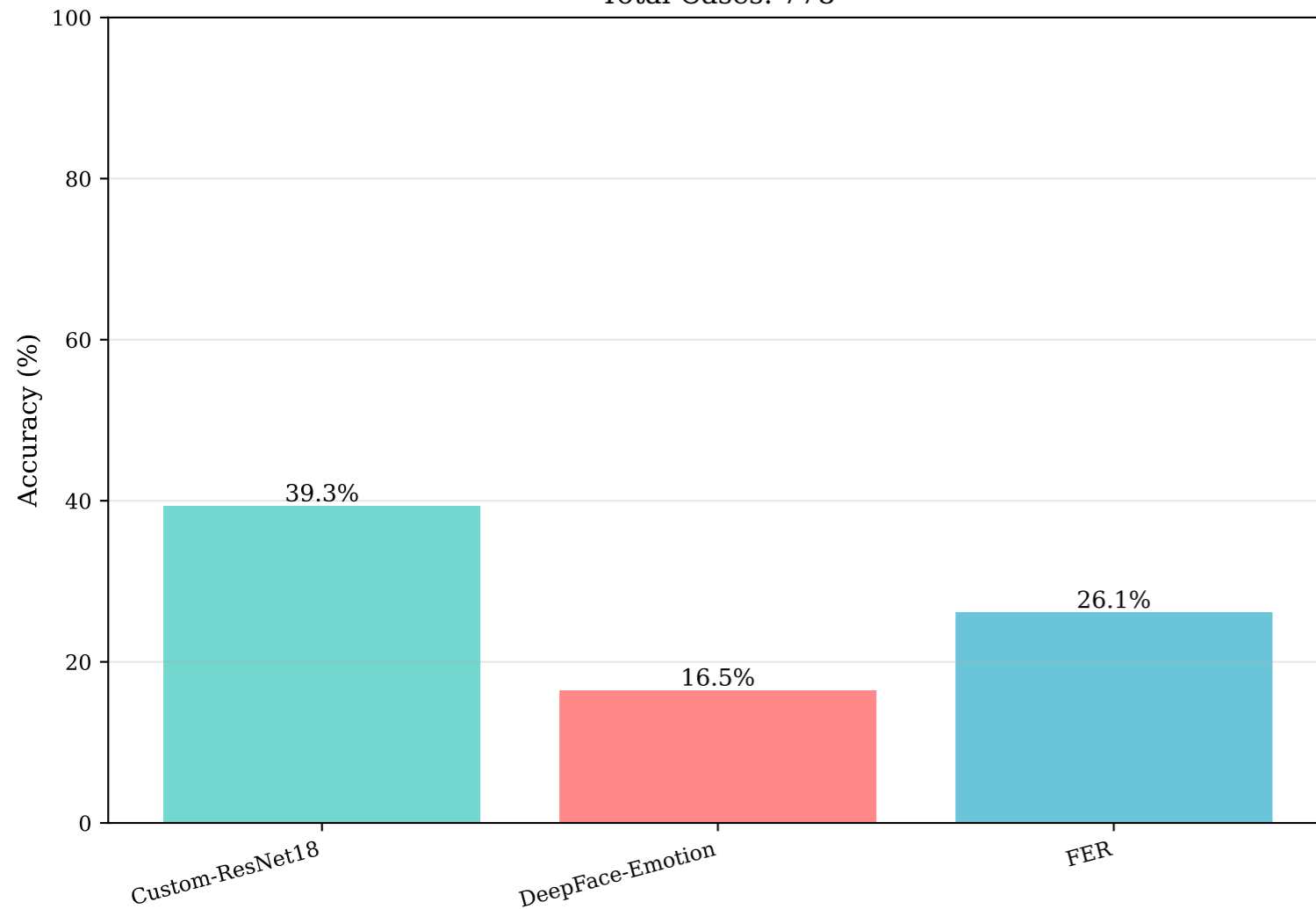


Performance on Difficult Cases (All Models Disagree)

Total Cases: 778



Overall Accuracy vs. Accuracy on Difficult Cases

