

Score vs Confidence Threshold: test, hellaswac

