

Supplementary

Table S1 Kappa coefficients with linear weighting between the different participants, for stage 1 and for stage 2

Stage 2	KAPPA	Stage 1								
		MR	SC	obs1	obs2	obs3	obs4	obs5	obs6	CNN
	MR	0.80 (0.76–0.85)	0.46 (0.39–0.052)	0.66 (0.60–0.72)	0.77 (0.72–0.83)	0.64 (0.58–0.70)	0.84 (0.80–0.89)	0.83 (0.78–0.87)	0.67 (0.61–0.73)	0.64 (0.58–0.69)
	SC	0.44 (0.38–0.50)		0.57 (0.51–0.63)	0.37 (0.30–0.43)	0.31 (0.25–0.36)	0.43 (0.37–0.49)	0.44 (0.37–0.50)	0.49 (0.43–0.56)	0.54 (0.48–0.60)
	obs1	0.67 (0.62–0.73)	0.51 (0.45–0.57)	0.76 (0.71–0.81)	0.49 (0.42–0.56)	0.37 (0.31–0.43)	0.59 (0.53–0.66)	0.61 (0.55–0.67)	0.73 (0.68–0.79)	0.61 (0.55–0.66)
	obs2	0.83 (0.78–0.88)	0.38 (0.32–0.45)	0.59 (0.52–0.65)	0.70 (0.64–0.76)	0.66 (0.60–0.72)	0.71 (0.65–0.76)	0.68 (0.62–0.74)	0.51 (0.44–0.57)	0.52 (0.46–0.58)
	obs3	0.54 (0.48–0.61)	0.30 (0.24–0.35)	0.35 (0.30–0.41)	0.56 (0.50–0.62)	0.85 (0.80–0.89)	0.60 (0.54–0.66)	0.57 (0.50–0.63)	0.40 (0.33–0.46)	0.43 (0.38–0.49)
	obs4	0.78 (0.73–0.83)	0.41 (0.35–0.47)	0.68 (0.61–0.74)	0.68 (0.62–0.74)	0.43 (0.37–0.50)	0.72 (0.66–0.77)	0.69 (0.63–0.75)	0.62 (0.56–0.68)	0.61 (0.55–0.66)
	obs5	0.82 (0.77–0.87)	0.40 (0.34–0.46)	0.56 (0.49–0.62)	0.72 (0.66–0.78)	0.59 (0.52–0.65)	0.65 (0.58–0.71)	0.68 (0.63–0.74)	0.62 (0.56–0.68)	0.60 (0.54–0.66)
	obs6	0.74 (0.68–0.79)	0.48 (0.41–0.55)	0.57 (0.51–0.64)	0.62 (0.56–0.68)	0.40 (0.34–0.46)	0.61 (0.55–0.68)	0.64 (0.58–0.71)	0.73 (0.68–0.79)	0.60 (0.54–0.66)
	CNN	0.57 (0.52–0.63)	0.54 (0.48–0.60)	0.58 (0.52–0.64)	0.53 (0.47–0.59)	0.39 (0.33–0.44)	0.53 (0.47–0.59)	0.55 (0.50–0.61)	0.57 (0.51–0.63)	

*, the variability of the automated methods is nil. CNN, Convolutional Neural Network; MR, majority report; SC., Commercial software application; Obs, Observer.