It is still challenging to build an AI system that can pe	erform tasks that involve vision	on and language at human lev	el. So far, researchers have	singled out individua
5 5 7 7 mm 2 mm 2 your man out per		5 <u>6 </u>	, 1111111111111111111111111111111111111	J : ::::::::::::::::::::::::::::::::::