Interpretability and fairness are critical in computer vision and machine learning applications, in particular when dealing with human outcomes, e.g. inviting or no