

**(a) Global cross-modal relation scheme
(Vempala and Preoțiuc-Pietro, 2019)**

[Oklahoma] **Neutral** legend Bob Stoops reportedly had interest in an NFL job before retiring



Image adds to the tweet meaning & Text is not represented in image.

(b) Aspect-level cross-modal relation scheme regarding explicit alignment (Yu et al., 2022b)

HangUp NBA superstar **[LeBron James]**
Neutral: Racism is part of America



The image is relevant to aspect (Explicit Alignment).

(c) Our aspect-level emotional cross-modal relation scheme (AECR)

Me listening to **[Trey Songz]** **Positive** sing at BETAwards17



The image is relevant to aspect (The image is the source of emotional expression).