Methodology

Cross-modal Similarity Extraction

- Most method are considered two different modal features (t, v) separately
  - Just concatenating them with no relation between them explored
- However, besides that, the falsity of a news article can be also detected by assessing how (ir)relevant the textual information is compared to its visual information
- Fake news creators sometimes actively use irrelevant image for false statements to attract readers' attention, or passively use them due to the difficulty in finding a supportive non-manipulated image.

Methodology

Cross-modal Similarity Extraction

- Compared to news articles delivering relevant textual and visual information, those with disparate statements and images are more likely to be fake.
- Define the relevance between news textual and visual information as follows by slightly modifying cosine similarity:
  - $M_s(t, v) = \frac{t \cdot v + \|t\|\|v\|}{2\|t\|\|v\|}$ v.s. $\cos(t, v) = \frac{t \cdot v}{\|t\|\|v\|}$
- In such a way, it's guaranteed that $M_s(t, v)$ is positive and $\in [0,1]$
  - $M_s(t, v) \to 0$: t, v are far from being similar, $\to 1$: t, v are exactly the same