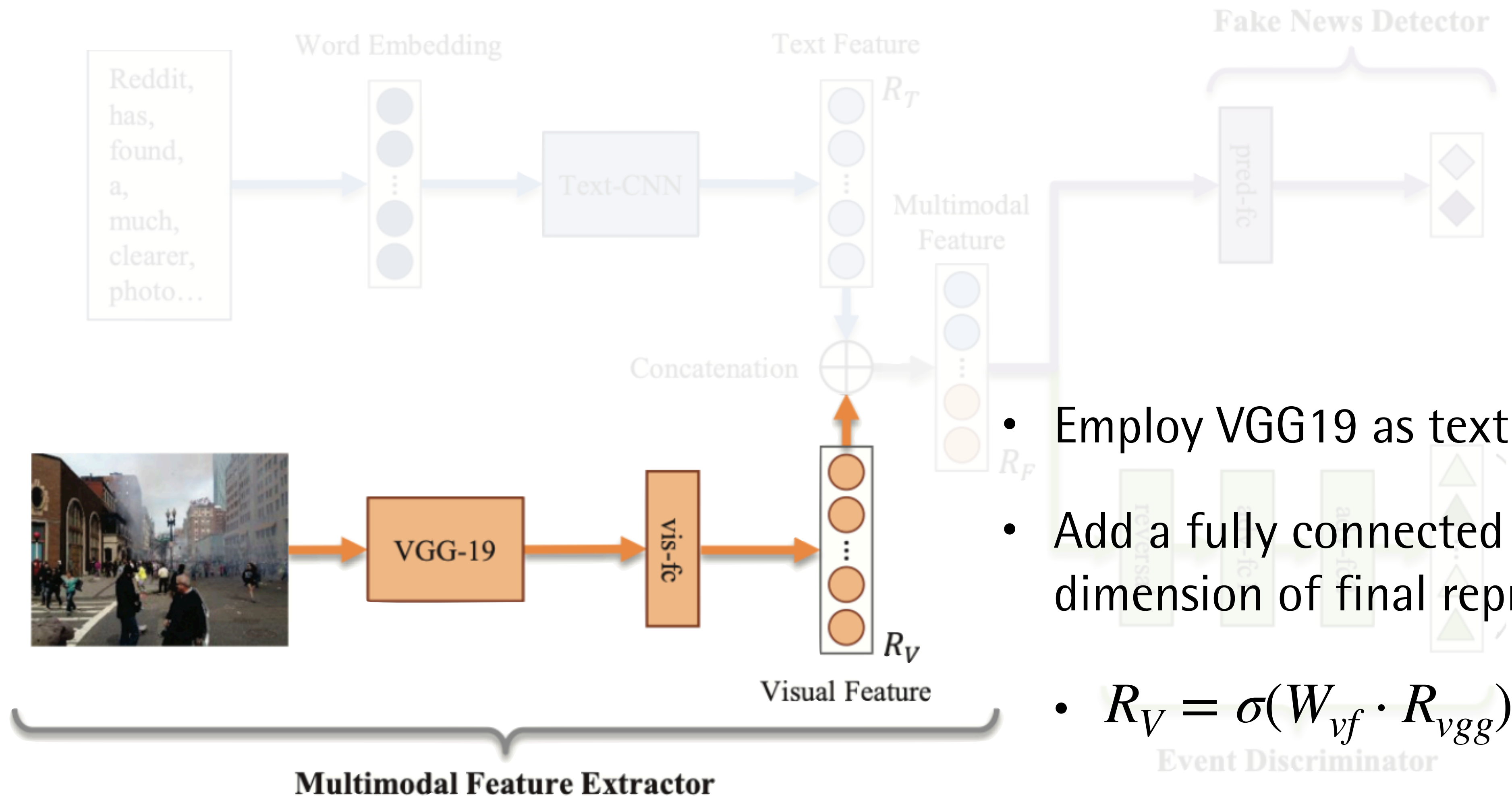# Methodology...
Visual Feature Extractor
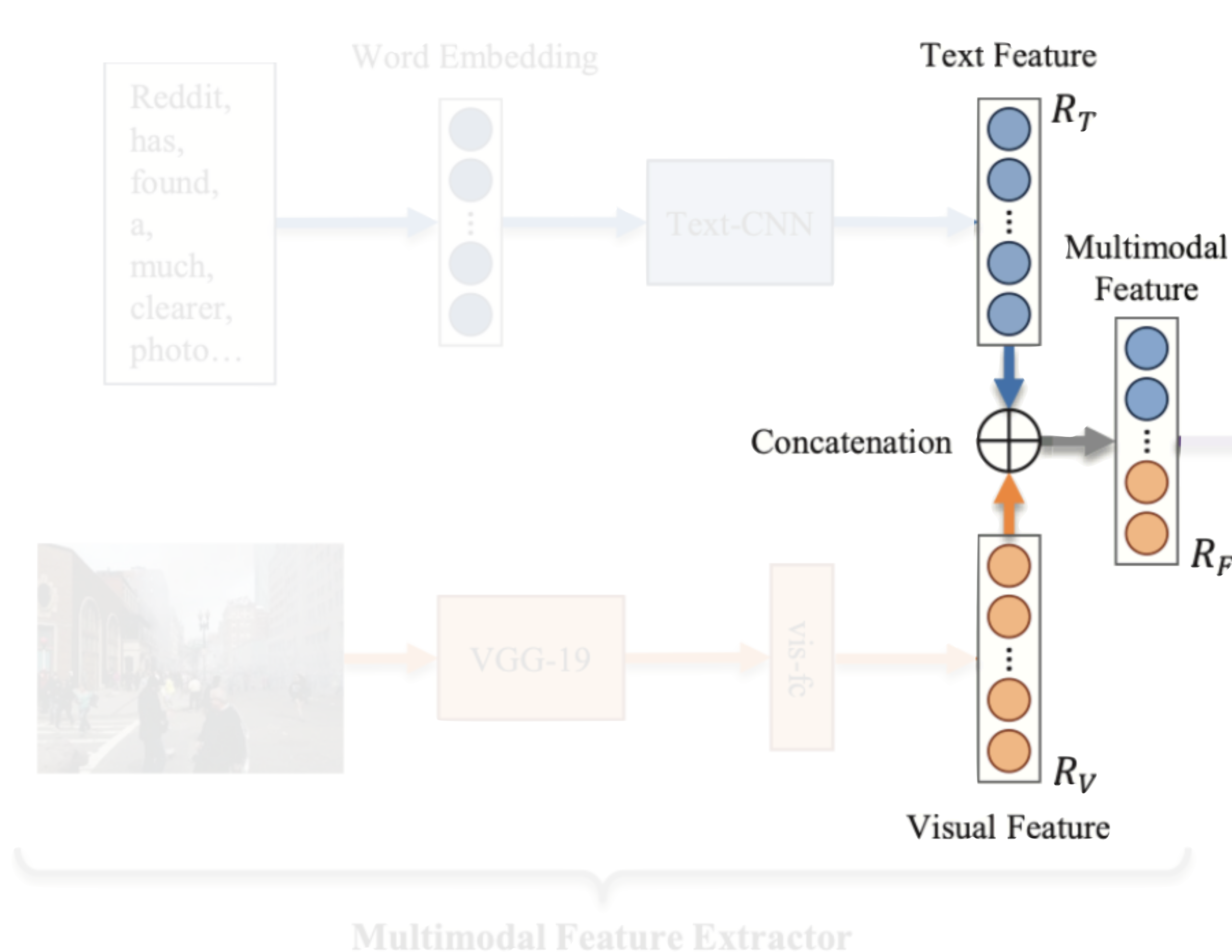

- Employ VGG19 as text feature extractor

- Add a fully connected layer to adjust the dimension of final representation to p.

- $R_V = \sigma(W_{vf} \cdot R_{vgg})$

Methodology....
Multi-model Feature Extractor

- Concatenated to form the multi-model feature representation denoted as

  - $R_F = R_T \oplus R_R \in \mathbb{R}^{2p}$

- Denote the multi-model feature extractor

  - $G_f(M; \theta_f)$

  - M: a set of textual and visual posts

  - θ_f: parameter to be learned