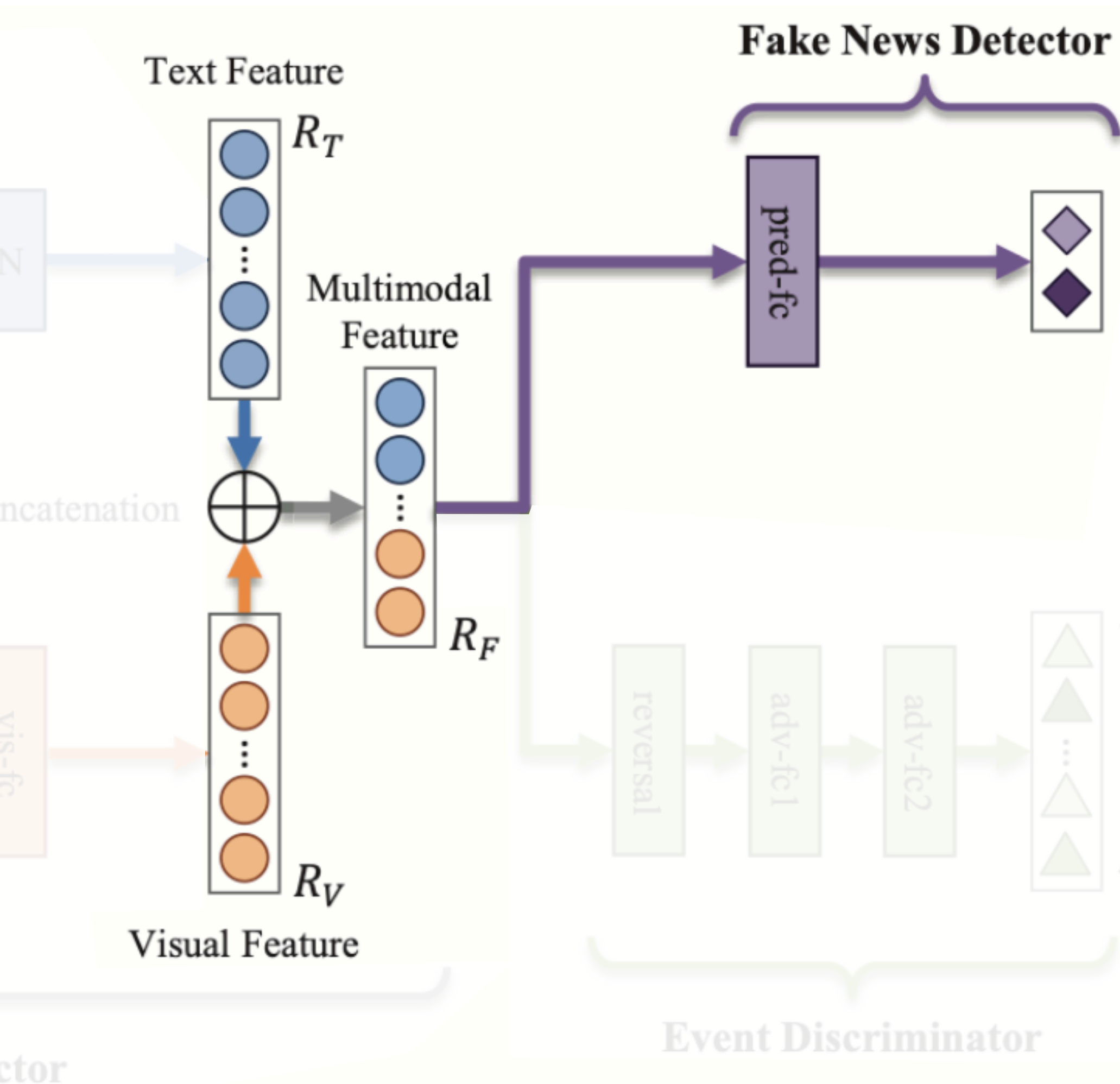


Methodology.....

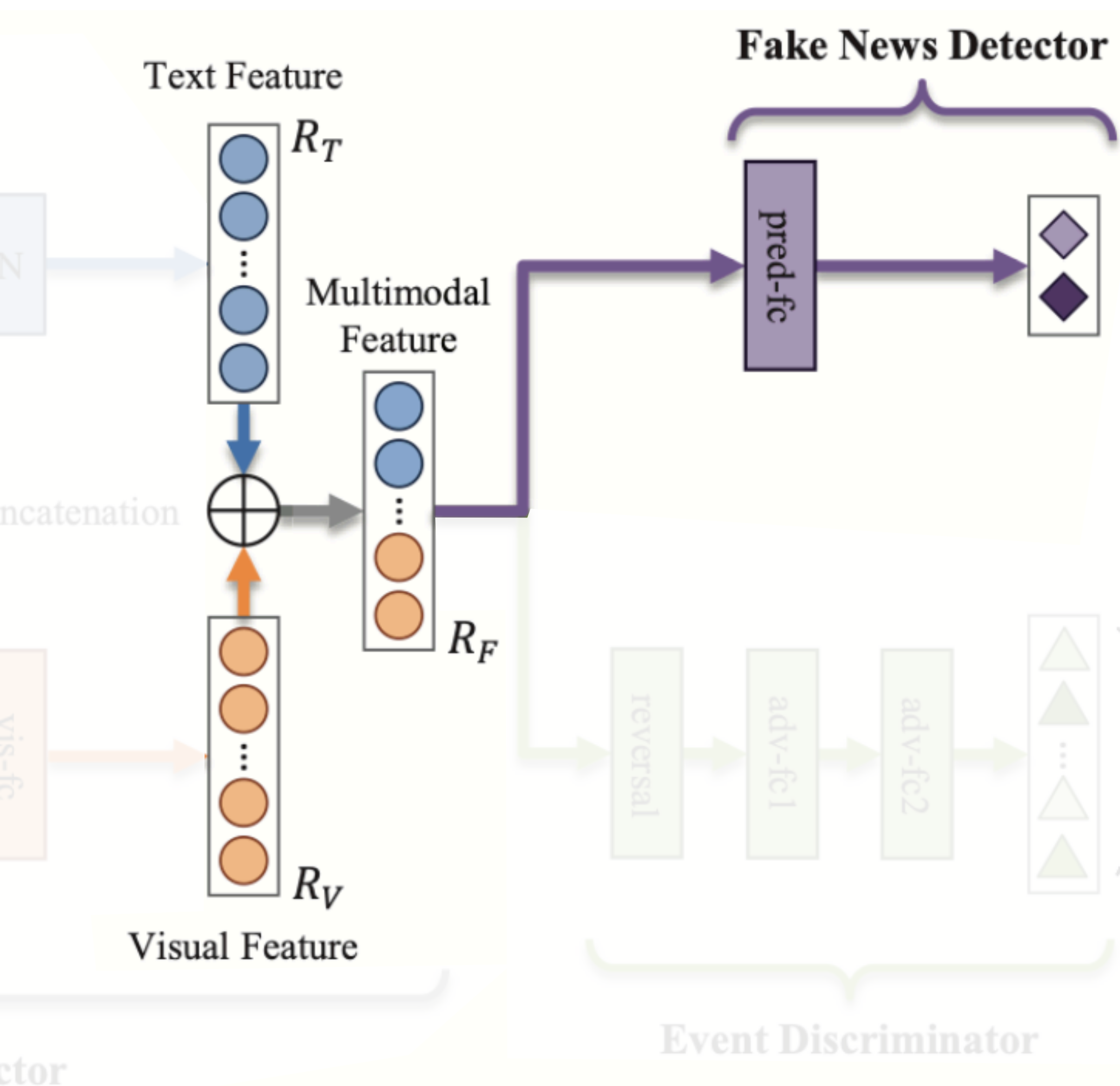
Fake News Detector



- Denote as $G_d(\cdot; \theta_d)$, θ_d : detector parameters
- Deploy a fully connected layer with softmax to predict the post are fake or real.
- Probability of post m_i being a fake one:
 - $P_\theta(m_i) = G_d(G_f(m_i; \theta_f); \theta_d)$
- Employ cross entropy to calculate the detection loss:
 - $L_d(\theta_f, \theta_d) = -\mathbb{E}_{(m,y) \sim (M,Y_d)}[y \log(P_\theta(m)) + (1 - y)(\log(1 - P_\theta(m)))]$
- Minimize loss function by seeking the optimal parameters θ_f, θ_d
 - $(\hat{\theta}_f, \hat{\theta}_d) = \arg \min_{\theta_f, \theta_d} L_d(\theta_f, \theta_d)$

Methodology.....

Fake News Detector



- As main goal of this work is detect the event which not covered by the training dataset.
- Direct minimization of detection loss only helps detect fake news included in the training dataset
 - Capture only event-specific knowledge or patterns
 - Not generalize well
- Need to learn more general feature representations that can capture the common features among all the events.
 - Should be event-invariant and doesn't include any event-specific feature.