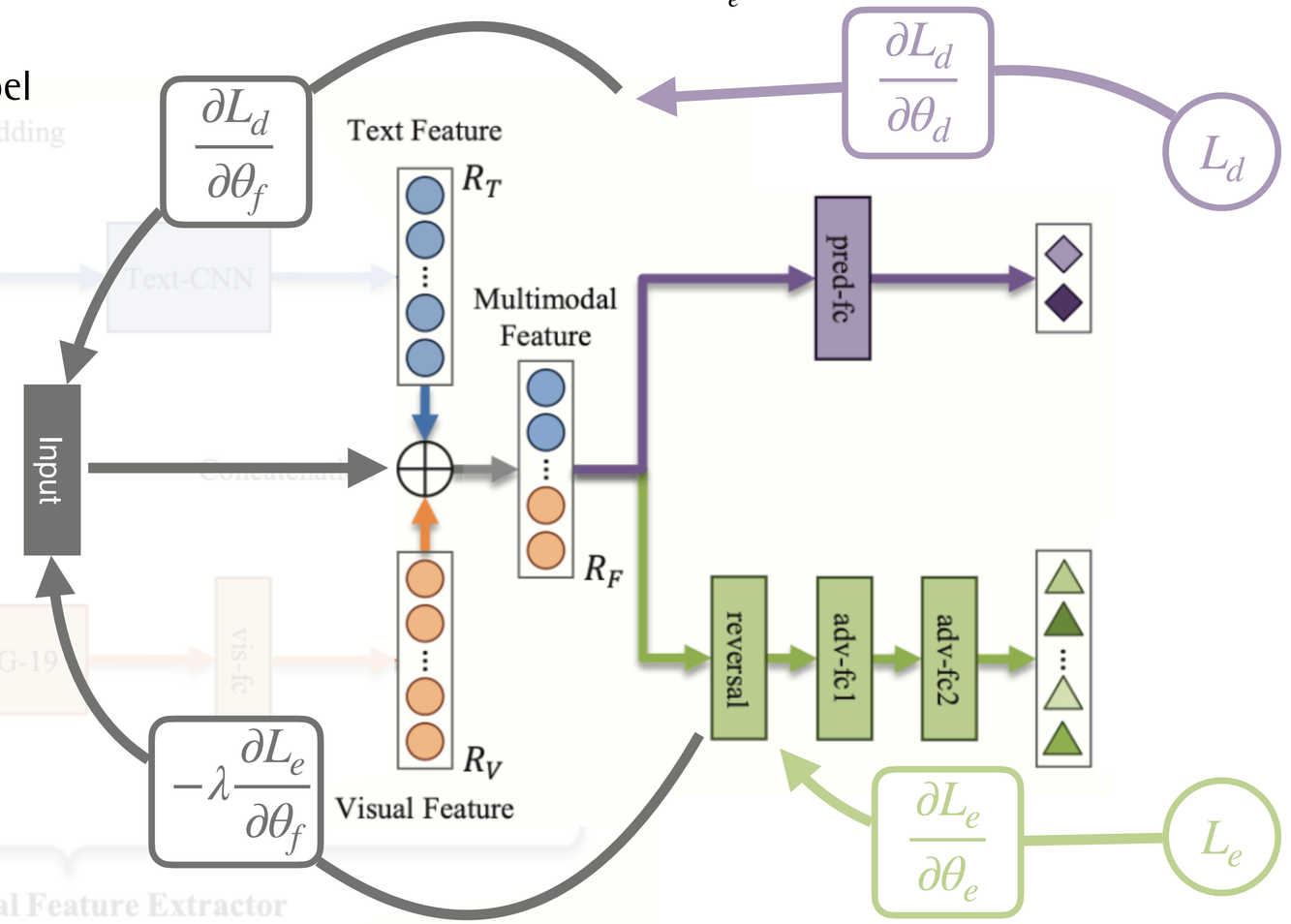
Methodology.....

Gradient Reversal Layer

- $L_{final}(\theta_f, \theta_d, \theta_e) = L_d(\theta_f, \theta_d) - \lambda L_e(\theta_f, \theta_e)$
- $(\hat{\theta}_f, \hat{\theta}_d) = arg \min_{\theta_f, \theta_d} L_{final}(\theta_f, \theta_d, \hat{\theta}_e)$
- $\hat{\theta}_e = arg \max_{\theta_e} L_{final}(\hat{\theta}_f, \hat{\theta}_d, \theta_e)$

Input: The multi-modal input $\{m_i\}_{i=1}^N$, the auxiliary event label $\{e_i\}_{i=1}^N$, the detection label $\{y_i\}_{i=1}^N$ and the learning rate η

1. **for** number of training iterations **do**
2. Decay learning rate: $\eta' = \frac{\eta}{(1 + \alpha \cdot p)^\beta}$
3. Update $\theta_f \leftarrow \theta_f - \eta(\frac{\partial L_d}{\partial \theta_f} - \lambda \frac{\partial L_e}{\partial \theta_f})$
4. Update $\theta_e \leftarrow \theta_e - \eta \frac{\partial L_e}{\partial \theta_e}$
5. Update $\theta_d \leftarrow \theta_d - \eta \frac{\partial L_d}{\partial \theta_d}$
6. **end for**



Experiments.

Dataset

Method	Twitter	Weibo
# of fake News	7898	4749
# of real News	6026	4779
# of image	514	9528

- **Twitter dataset**
- from MediaEval Verifying Multimedia Use benchmark
- Contain text, attach image/video and additional social context information
  - Focus on text and image
  - Remove the tweets without any text or image
- Has two parts: the **development and test set, there is no overlapping events among them.**

- **Weibo dataset**
- Fake news posts: 2012.05 ~ 2016.01 verified by Weibo official rumor debunking system
- Real news posts: 2012.05 ~ 2016.01 from Weibo verified by Xinhua News Agency
  - removed duplicated and very small images
- Use single-pass clustering and split whole dataset into **training, validation, testing sets = 7:1:2 to ensure that they don't not contain any common event.**