## Hierarchical Multi-modal Contextual Attention Network for Fake News Detection

Shengsheng Qian
National Lab of Pattern Recognition,
Institute of Automation, CAS
University of Chinese Academy of
Sciences
shengsheng.qian@nlpr.ia.ac.cn

Jinguang Wang HeFei University of Technology wangjinguang502@gmail.com

Jun Hu
National Lab of Pattern Recognition,
Institute of Automation, CAS
hujunxianligong@gmail.com

#### Quan Fang

National Lab of Pattern Recognition, Institute of Automation, CAS University of Chinese Academy of Sciences qfang@nlpr.ia.ac.cn Changsheng Xu
National Lab of Pattern Recognition,
Institute of Automation, CAS
University of Chinese Academy of

Sciences
Peng Cheng Laboratory
csxu@nlpr.ia.ac.cn

SIGIR'21
220725 Chia-Chun Ho

# Outline of HMCAN

Introduction

Methodology

Experiments

Conclusion

Comments

#### **Fake News Detection**

- Social media websites are convenient platforms for people to share information, express and exchange opinions in their daily life.
- However, the authenticity of these information is difficult to guarantee since users do not check the reliability of the shared information.
  - Which has led to the widespread dissemination of considerable fake news.
- Therefore, detecting fake news on social media websites to ensure that users obtain true information has become a top priority.

#### **Existing Approaches**

- Traditional learning methods
  - Design plenty of hand-crafted features from the media content of posts and the social content of users.
  - SVM, Decision Tree... etc.
  - However, the content of fake news is highly complicated and hard to be fully captured by hand-crafted features.

#### **Existing Approaches**

- Deep learning methods
  - Many multi-modal representation methods utilize deep schemes to learn the representative features, and obtain superior performance for fake news detection.
  - MVAE, SAME, SpotFake, SpotFake+... etc.
  - Although these approaches show promising performance on fake news detection tasks.
    - They are still insufficient to take advantage of the multi-modal context information and the hierarchical semantics of text content.

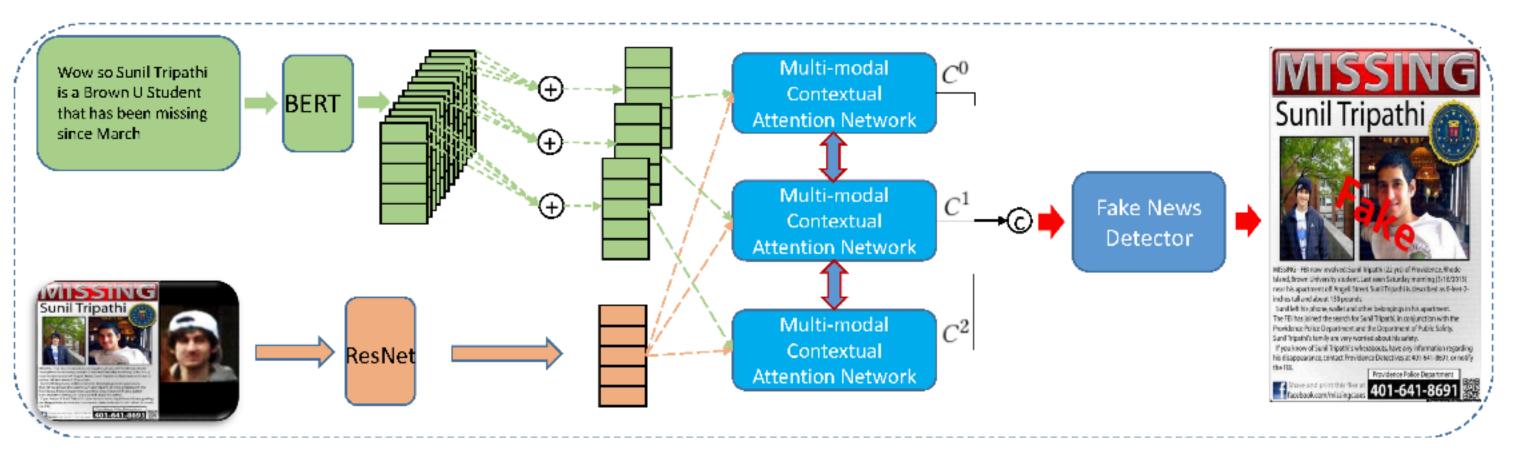
#### Challenges (1/2)

- How to fully utilize the multi-modal context information and extract high-order complementary information from it to enhance the performance of fake news detection?
  - The visual content of news posts usually contain may uncertain elements that are difficult to understand without the help of the text information.
  - The components they employ to capture multi-modal context are too simple to extract high-order complementary information from the multi-modal context.

#### Challenges (2/2)

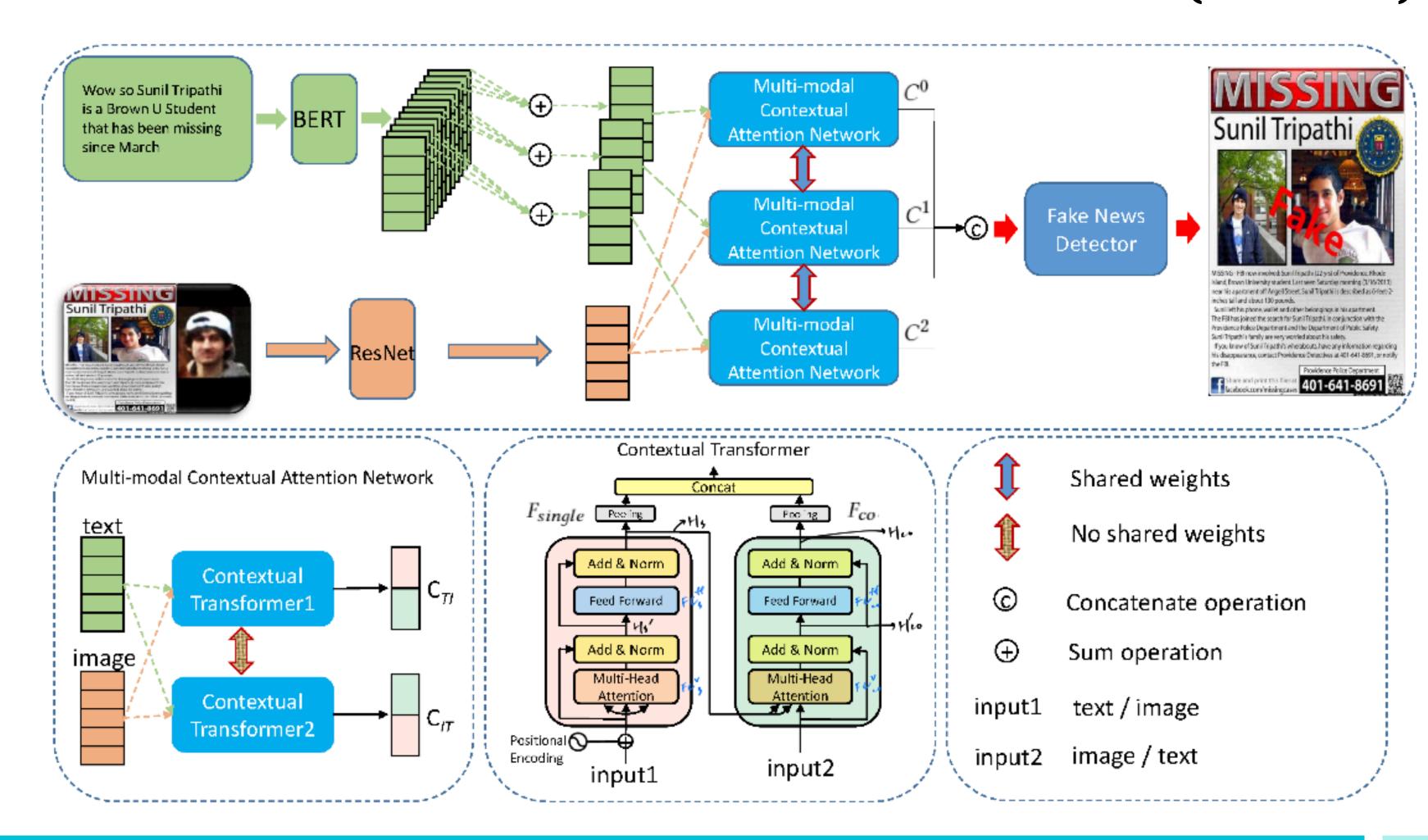
- How to explore and capture the hierarchical semantics of text information to learn a better representation of multi-modal news?
  - Most SOTA models user BERT as text encoders, which can provide hierarchical semantics of text.
    - But most of them only utilize the output of the last layers of these hierarchical models.
    - Some works explore the potential of exploiting the semantic knowledge in the intermediate layers of BERT models, showing that many downstream tasks can benefit from the full hierarchical semantics.

## Introduction HMCAN & Contributions



- Propose a novel Hierarchical Multi-modal Contextual Attention Network for FND.
  - By jointly modeling the multi-modal context information and the hierarchical semantics of text in a unified deep model.
  - Propose a multi-modal contextual attention network.
    - Modeling the multi-modal context for each news posts, where data from different modalities
      can complement each other to provide a better understanding of the multi-modal data.
  - Design a hierarchical encoding network.
    - For capture the rich hierarchical semantics for fake news detection.

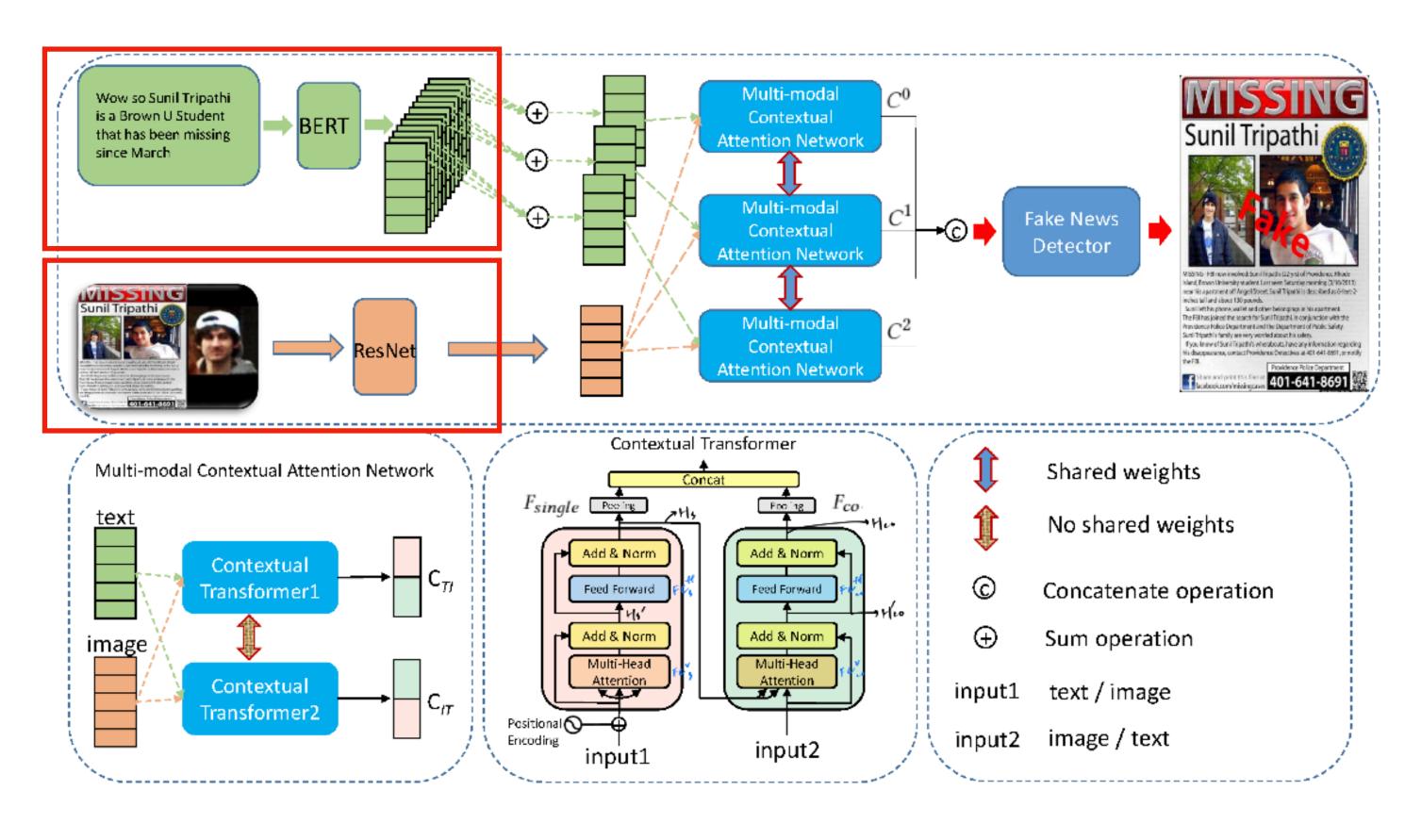
#### Hierarchical Multi-modal Contextual Attention Network (HMCAN)



#### **Problem Definition**

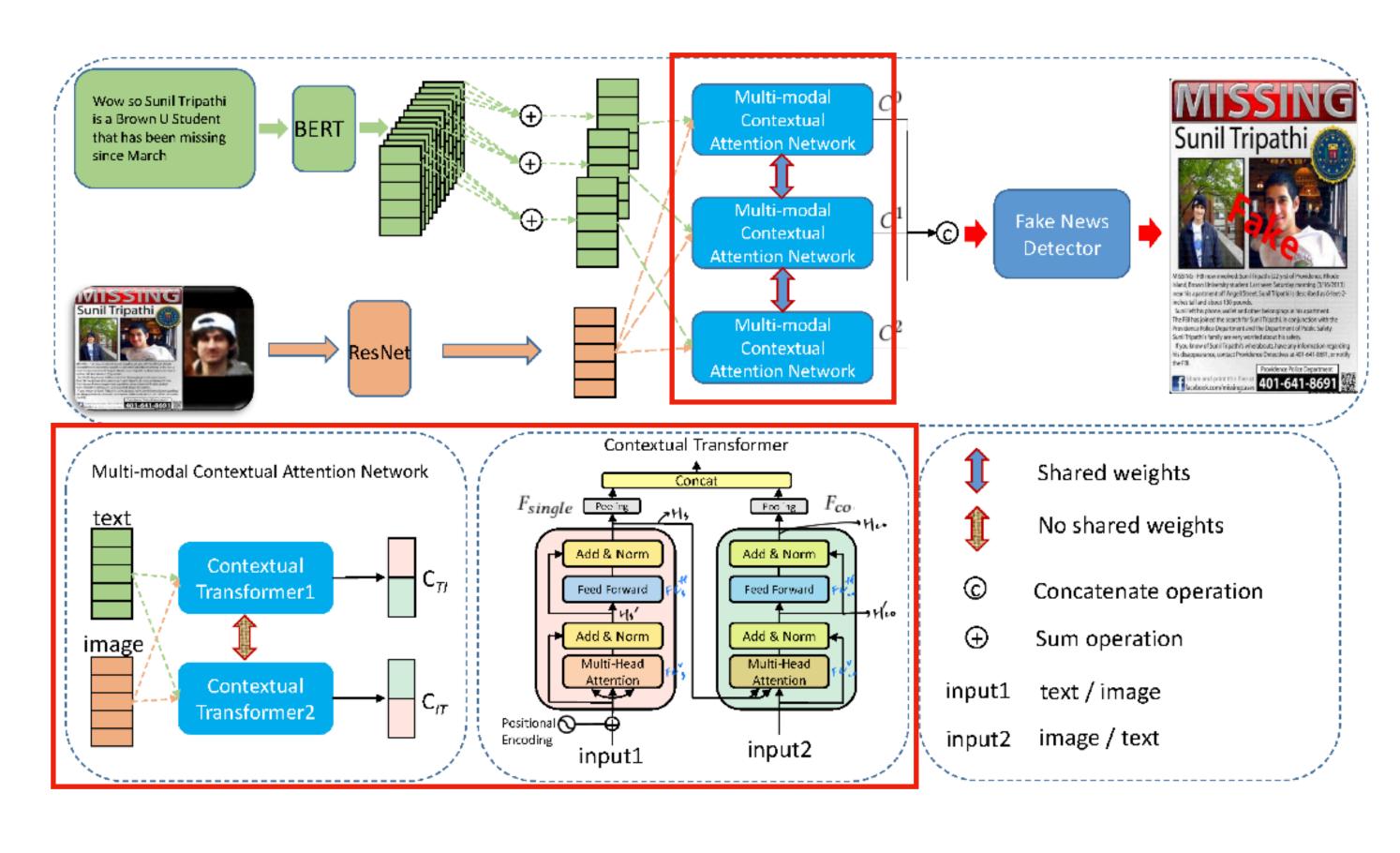
- P: a multi-modal post from social media consisting of text messages and corresponding images.
- The model will output  $Y = \{0,1\}$  to indicate to the label of the post.
  - Y=0: real
  - Y = 1 : fake

#### Text and Image Encoding Network



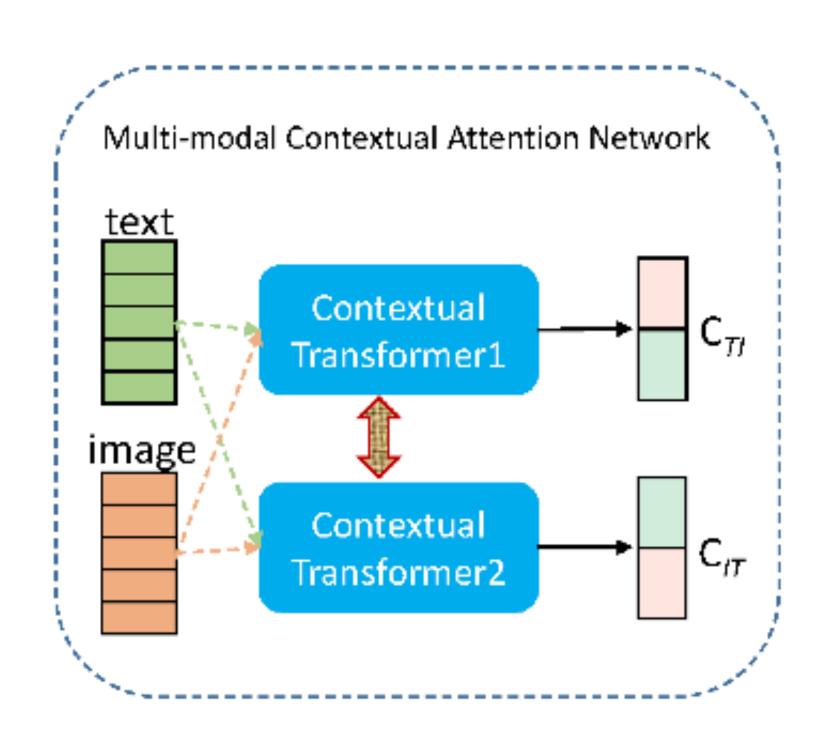
- Text Encoding Network
  - BERT
- Image Encoding Network
  - ResNet-50

#### Multi-modal Contextual Attention Network



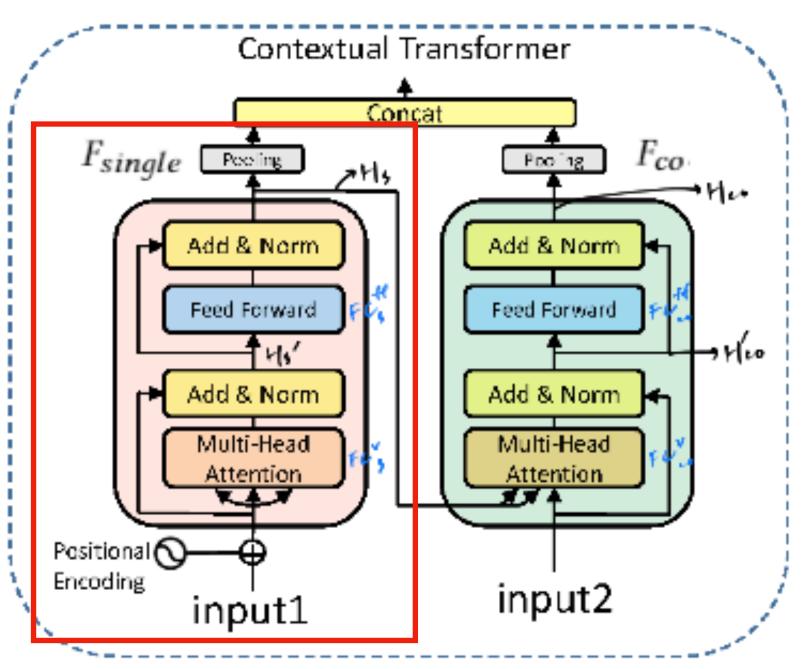
- Build the multi-modal context information.
- Extract high-order complementary information.

#### Multi-modal Contextual Attention Network



- Multi-modal contextual attention network consists of two contextual transformer units (Contextual Transformer1, 2).
- Each context transformer unit focuses on different context information for multi-modal representation learning.

#### Multi-modal Contextual Attention Network - $F_{single}$

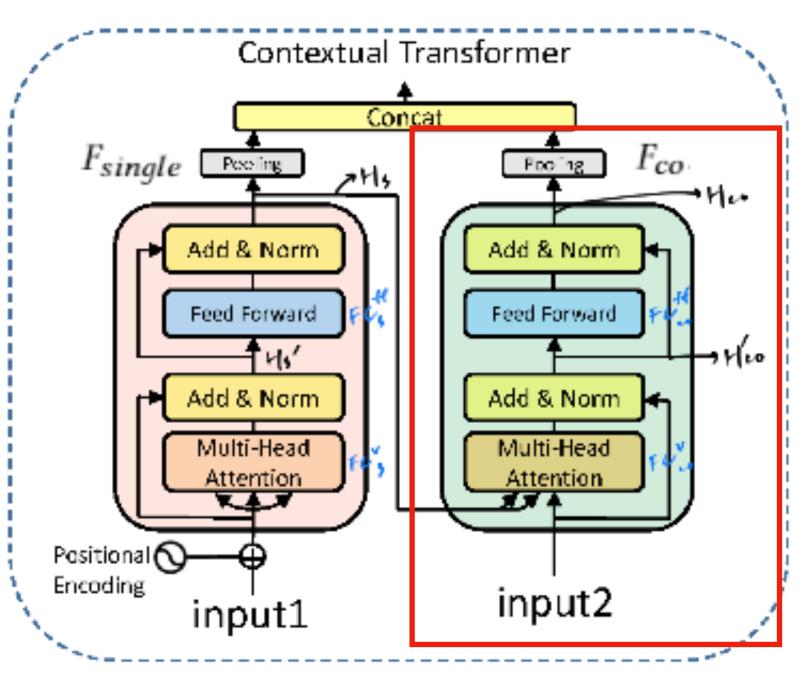


- Self-attention network  $F_{single}$  (the left part) is utilized to learn the representation of text (input1).
- The self-attention network computes a intra-modality affinity matrix  $A_{\mathcal{S}}$ .

$$A_{s} = \operatorname{softmax} \left( \frac{FC_{s}^{Q}(input1) \cdot FC_{s}^{K}(input1)^{\top}}{\sqrt{d}} \right)$$

- Representation of text  $H_s$  can be learned as follows:
  - $H'_s = layer\_norm(input1 + A_s \cdot FC_s^V(input1))$
  - $H_s = layer\_norm(H'_s + FC_s^{ff}(H'_s))$

### Multi-modal Contextual Attention Network – $F_{co}$

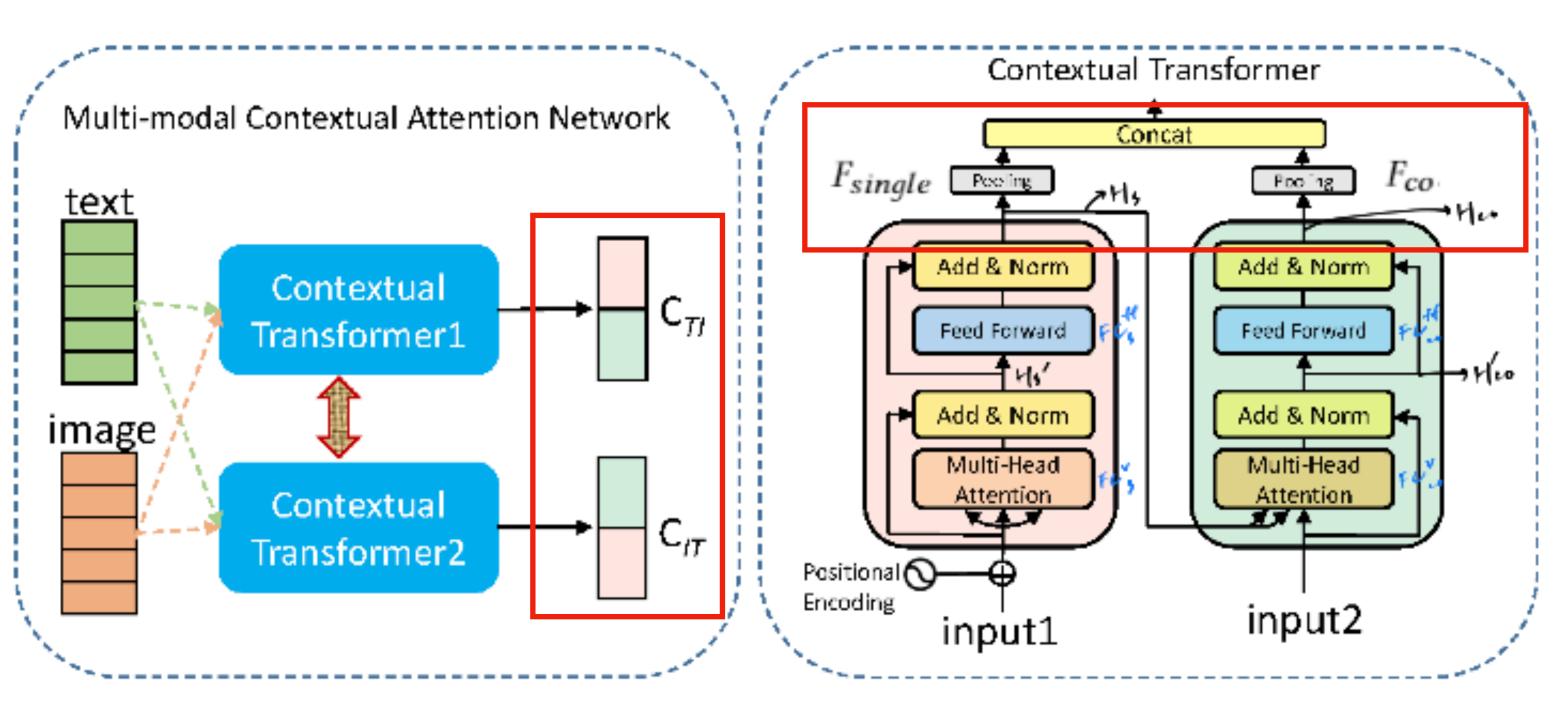


- The core idea is to extract information that is relevant to the image from the learned text representation, which can complement the visual information.
- $F_{co}$  computes a intermodality affinity matrix  $A_{co}$ .

$$A_{co} = \operatorname{softmax} \left( \frac{FC_{co}^{Q}(input2) \cdot FC_{co}^{K}(H_{s})^{\mathsf{T}}}{\sqrt{d}} \right)$$

- $F_{co}$  learns the multi-modal context-aware text representation  $H_{co}$  as follows:
  - $H'_{co} = layer\_norm(input2 + A_{co} \cdot FC^{V}_{co}(H_S))$
  - $H_{co} = layer\_norm(H'_{co} + FC^{ff}_{co}(H'_{co}))$

#### Multi-modal Contextual Attention Network

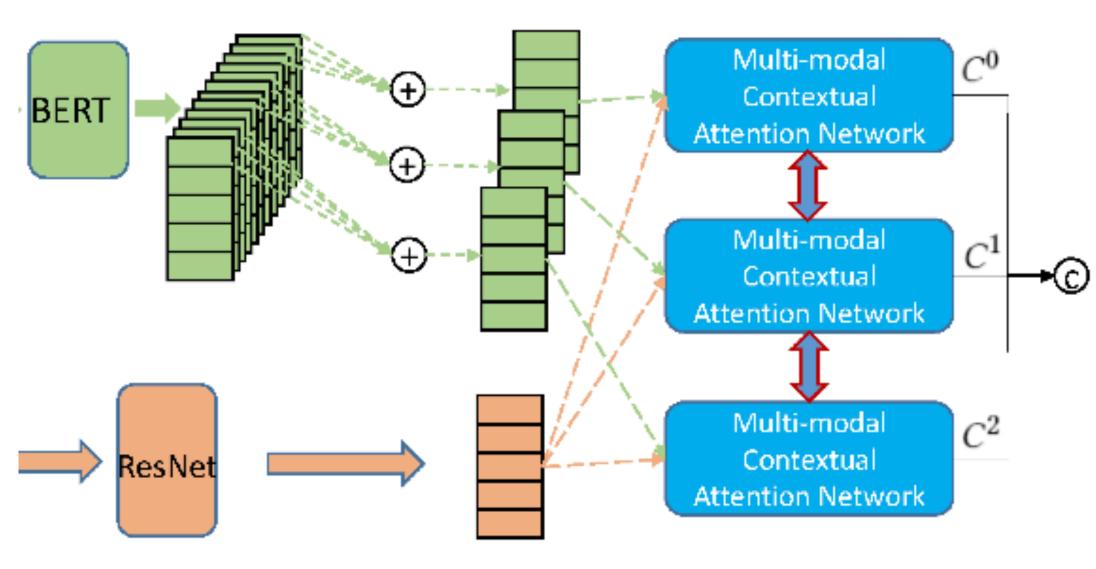


- $H_s$  and  $H_{co}$  are pooled into two feature vectors, which are then concatenated into a feature vector  $(C_{TI}/C_{IT})$  as the multimodal contextual representation of the text.
- Make the output of the multimodal contextual attention network.

• 
$$C = \alpha C_{TI} + \beta C_{IT}$$

• 
$$\alpha + \beta = 1$$

#### Hierarchical Encoding Network

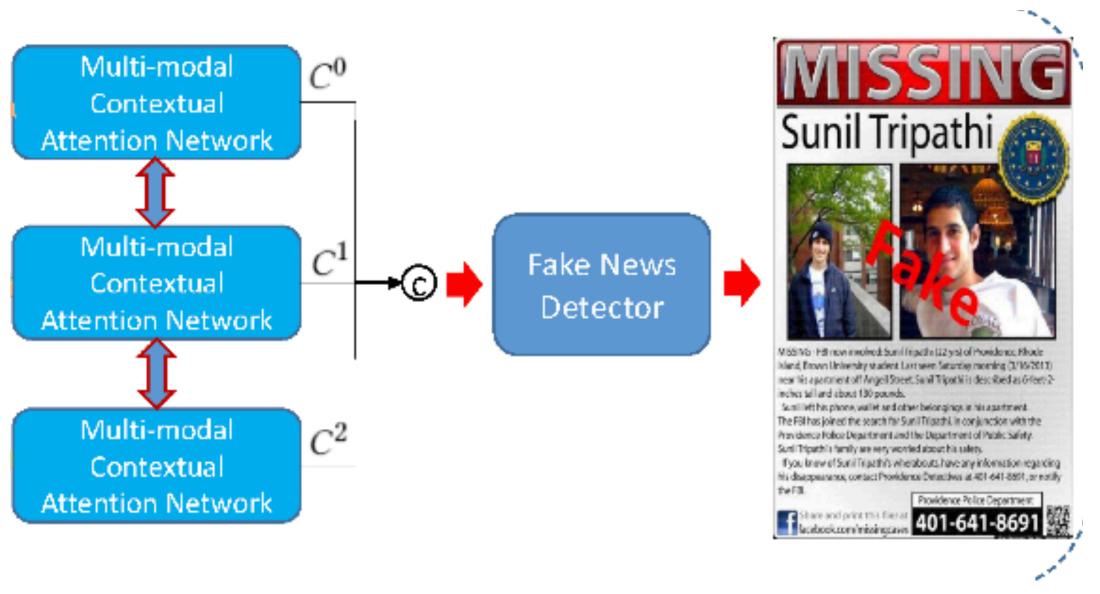


- BERT can provide hierarchical semantics for text, which consists of the outputs of 11 intermediate layers and 1 output layers.
- Reduce the 12 layer outputs into g group outputs by integrating every 12/g adjacent layers of BERT

$$\mathbf{s}_{\mathbf{i}}^{0} = \sum_{j=1}^{4} f_{B}(W)_{j,i}, \mathbf{s}_{\mathbf{i}}^{1} = \sum_{j=5}^{8} f_{B}(W)_{j,i}, \mathbf{s}_{\mathbf{i}}^{2} = \sum_{j=9}^{12} f_{B}(W)_{j,i}$$

• 
$$C = \operatorname{concat}(C^0, C^1, C^2)$$

#### Fake News Detector



 It deploys a fully connected layer with the corresponding activation function.

• 
$$\hat{P}_n = \sigma(W_f C_n + b)$$

Employ cross entropy loss:

$$\mathcal{L}(\Theta) = \sum_{n=1}^{N} - [Y_n \log(\hat{P}_n + (1 - Y_n) \log(1 - \hat{P}_n)]$$

# **Experiments**Datasets

- Weibo
  - Train : Test = 8:2
- Twitter
- PHEME
  - Train: Test = 8:2

News	WEIBO	TWITTER	PHEME
# of Fake News	4749	7898	1972
# of Real News	4779	6026	3830
# of Images	9528	514	3670

# **Experiments**Baselines (1/2)

- Single-modal
  - SVM-TS, CNN, GRU, TextGCN
- Multi-modal
  - EANN
  - att-RNN
  - MVAE
  - SpotFake, SpotFake+
  - SAFE

- SVM-TS performs the worst among all.
  - Indicating that the hand-crafted features are weak and insufficient to identify fake news.
- Deep learning models (CNN, GRU) have better performance than SVM-TS.
  - Indicating the superior advantages over traditional methods.
  - In Twitter dataset, CNN fails to capture long-range semantic relationships between words.
  - TextGCN is better than them show that graph structure can effectively capture word co-occurrence and document-word relationships.

Dataset	Detect	Methods	Λοουνοου	Fake news			Real news		
WEIBO    CRIV   0.702   0.671   0.794   0.727   0.747   0.609   0.671	Dataset		Accuracy	Precision	Recall	F1	Precision	Recall	F1
WEIBO  WEIBO  WEIBO  CNN   0.740   0.736   0.756   0.744   0.747   0.723   0.735   SAFE   0.763   0.833   0.659   0.736   0.717   0.868   0.785   att_RNN   0.772   0.854   0.656   0.742   0.720   0.889   0.795   EANN   0.782   0.827   0.697   0.756   0.752   0.863   0.804   TextGCN   0.787   0.975   0.573   0.727   0.712   0.985   0.827   MVAE   0.824   0.854   0.769   0.809   0.802   0.875   0.837   SpotFake   0.869   0.877   0.859   0.868   0.861   0.879   0.870   SpotFake   0.892   0.902   0.964   0.932   0.847   0.656   0.739   HMCAN   0.885   0.920   0.845   0.881   0.855   0.892   0.873   HMCAN   0.885   0.920   0.845   0.881   0.856   0.926   0.890   SVM-TS   0.529   0.488   0.497   0.496   0.565   0.556   0.561   GRU   0.634   0.581   0.812   0.677   0.758   0.502   0.604   CNN   0.549   0.508   0.597   0.549   0.598   0.509   0.550   SAFE   0.766   0.779   0.795   0.786   0.752   0.731   0.742   SAFE   0.766   0.749   0.615   0.676   0.589   0.728   0.651   EANN   0.648   0.810   0.498   0.617   0.584   0.759   0.660   TextGCN   0.703   0.808   0.365   0.503   0.680   0.939   0.779   MVAE   0.745   0.801   0.719   0.758   0.689   0.777   0.730   SpotFake   0.777   0.751   0.900   0.820   0.832   0.606   0.701   SpotFake   0.779   0.793   0.827   0.810   0.786   0.747   0.766   HMCAN   0.897   0.971   0.801   0.878   0.853   0.979   0.912    PHEME  PHEME  PHEME  ENN   0.681   0.827   0.559   0.667   0.806   0.940   0.866   MVAE   0.852   0.806   0.719   0.770   0.876   0.899   0.888   EANN   0.681   0.685   0.664   0.694   0.701   0.750   0.747   TextGCN   0.828   0.775   0.735   0.737   0.827   0.828   0.868   MVAE   0.852   0.806   0.719   0.760   0.871   0.917   0.893   SpotFake   0.823   0.743   0.745   0.744   0.864   0.863   0.863   SpotFake   0.823   0.743   0.745   0.744   0.864   0.863   0.863   SpotFake   0.800   0.730   0.668   0.697   0.832   0.869   0.850    EANN   0.852   0.806   0.719   0.760   0.807   0.850    MVAE   0.852   0.806   0.719   0.760   0.807   0.808    MVAE   0.852   0.806		SVM-TS	0.640	0.741	0.573	0.646	0.651	0.798	0.711
WEIBO  WEIBO  WEIBO  REANN  WEIBO  WEIBO  WEIBO  WEIBO  WEIBO  REANN  WEIBO  WAIL  WEIBO  WEI		GRU	0.702	0.671	0.794	0.727	0.747	0.609	0.671
WEIBO         att_RNN         0.772         0.854         0.656         0.742         0.720         0.889         0.795           EANN         0.782         0.827         0.697         0.756         0.752         0.863         0.804           TextGCN         0.787         0.975         0.573         0.727         0.712         0.985         0.827           MVAE         0.864         0.869         0.877         0.859         0.808         0.861         0.875         0.837           SpotFake*         0.892         0.902         0.964         0.932         0.847         0.656         0.739           SpotFake+         0.870         0.887         0.849         0.868         0.855         0.892         0.873           HMCAN         0.885         0.920         0.845         0.881         0.856         0.926         0.890           SWM-TS         0.529         0.488         0.497         0.496         0.565         0.556         0.556           GRU         0.634         0.581         0.812         0.677         0.758         0.502         0.604           CNN         0.549         0.508         0.597         0.549         0.598         0.502		CNN	0.740	0.736	0.756	0.744	0.747	0.723	0.735
WEIBO         EANN         0.782         0.827         0.697         0.756         0.752         0.863         0.804           TextGCN         0.787         0.975         0.573         0.727         0.712         0.985         0.827           MVAE         0.824         0.854         0.769         0.809         0.802         0.875         0.837           SpotFake         0.869         0.877         0.859         0.868         0.861         0.879         0.870           SpotFake+         0.870         0.887         0.890         0.868         0.861         0.879         0.870           SpotFake+         0.870         0.887         0.849         0.868         0.855         0.892         0.873           HMCAN         0.885         0.920         0.845         0.881         0.856         0.926         0.890           SVM-TS         0.529         0.488         0.497         0.496         0.565         0.556         0.561           GRU         0.634         0.581         0.812         0.677         0.758         0.502         0.604           CNN         0.549         0.508         0.597         0.549         0.598         0.598         0.55		SAFE	0.763	0.833	0.659	0.736	0.717	0.868	0.785
WEIBO         TextGCN         0.787         0.975         0.573         0.727         0.712         0.985         0.827           MVAE         0.824         0.854         0.769         0.809         0.802         0.875         0.837           SpotFake         0.869         0.877         0.859         0.868         0.861         0.879         0.870           SpotFake+         0.892         0.902         0.964         0.932         0.847         0.656         0.739           SpotFake+         0.870         0.887         0.849         0.868         0.855         0.892         0.873           HMCAN         0.885         0.920         0.845         0.881         0.856         0.926         0.890           SVM-TS         0.529         0.488         0.497         0.496         0.555         0.556         0.556           GRU         0.634         0.581         0.812         0.677         0.758         0.502         0.604           CNN         0.549         0.508         0.597         0.549         0.598         0.509         0.550           SAFE         0.766         0.777         0.795         0.786         0.752         0.731         0.74		att_RNN	0.772	0.854	0.656	0.742	0.720	0.889	0.795
MVAE   0.824   0.854   0.769   0.809   0.802   0.875   0.837	WEIRO	EANN	0.782	0.827	0.697	0.756	0.752	0.863	0.804
SpotFake   0.869   0.877   0.859   0.868   0.861   0.879   0.870	WEIDO	TextGCN	0.787	0.975	0.573	0.727	0.712	0.985	0.827
SpotFake*   0.892   0.902   0.964   0.932   0.847   0.656   0.739     SpotFake+   0.870   0.887   0.849   0.868   0.855   0.892   0.873     HMCAN   0.885   0.920   0.845   0.881   0.856   0.926   0.890     SVM-TS   0.529   0.488   0.497   0.496   0.565   0.556   0.561     GRU   0.634   0.581   0.812   0.677   0.758   0.502   0.604     CNN   0.549   0.508   0.597   0.549   0.598   0.509   0.550     SAFE   0.766   0.777   0.795   0.786   0.752   0.731   0.742     att_RNN   0.664   0.749   0.615   0.676   0.589   0.728   0.651     EANN   0.648   0.810   0.498   0.617   0.584   0.759   0.660     TextGCN   0.703   0.808   0.365   0.503   0.680   0.939   0.779     MVAE   0.745   0.801   0.719   0.758   0.689   0.777   0.730     SpotFake   0.771   0.784   0.744   0.764   0.769   0.807   0.787     SpotFake   0.777   0.751   0.900   0.820   0.832   0.606   0.701     SpotFake   0.770   0.793   0.827   0.810   0.786   0.747   0.766     HMCAN   0.897   0.971   0.801   0.878   0.853   0.979   0.912     SVM-TS   0.639   0.546   0.576   0.560   0.729   0.705   0.717     GRU   0.832   0.782   0.712   0.745   0.855   0.896   0.865     CNN   0.779   0.732   0.606   0.663   0.799   0.875   0.835     SAFE   0.811   0.827   0.559   0.667   0.806   0.940   0.866     att_RNN   0.850   0.771   0.749   0.770   0.876   0.899   0.888     EANN   0.681   0.685   0.664   0.694   0.701   0.750   0.747     TextGCN   0.828   0.775   0.735   0.737   0.827   0.828   0.828     MVAE   0.852   0.806   0.719   0.760   0.871   0.917   0.893     SpotFake   0.823   0.743   0.745   0.744   0.864   0.863   0.863     SpotFake   0.823   0.743   0.745   0.744   0.864   0.863   0.863     SpotFake   0.800   0.730   0.668   0.697   0.832   0.869   0.850     Contact   0.800   0.730   0.668   0.697   0.832   0.869   0.850     Contact   0.800   0.730   0.668		MVAE	0.824	0.854	0.769	0.809	0.802	0.875	0.837
SpotFake+   0.870   0.887   0.849   0.868   0.855   0.892   0.873		SpotFake	0.869	0.877	0.859	0.868	0.861	0.879	0.870
HMCAN   0.885   0.920   0.845   0.881   0.856   0.926   0.890		SpotFake*	0.892	0.902	0.964	0.932	0.847	0.656	0.739
TWITTER    SVM-TS		SpotFake+	0.870	0.887	0.849	0.868	0.855	0.892	0.873
TWITTER  GRU  O.634  CNN  O.549  O.508  O.597  O.549  O.598  O.599  O.550  SAFE  O.766  O.777  O.795  O.786  O.786  O.752  O.731  O.742  att_RNN  O.664  EANN  O.664  O.749  O.615  O.676  O.589  O.589  O.728  O.651  EANN  O.648  O.810  O.498  O.617  O.584  O.759  O.660  TextGCN  O.703  O.808  O.365  O.503  O.680  O.939  O.777  O.795  MVAE  O.745  SpotFake  O.771  O.784  O.774  O.751  O.784  O.774  O.751  O.790  O.820  O.832  O.606  O.747  O.751  O.801  O.786  O.747  O.758  O.606  O.703  O.808  O.365  O.503  O.680  O.939  O.777  O.730  SpotFake  O.771  O.784  O.744  O.764  O.769  O.807  O.787  SpotFake+  O.790  O.793  O.827  O.810  O.882  O.747  O.751  O.801  O.878  O.878  O.879  O.971  O.801  O.878  O.878  O.879  O.971  O.801  O.878  O.879  O.912  SVM-TS  O.639  O.546  O.576  O.560  O.729  O.705  O.717  GRU  O.832  O.782  O.712  O.745  O.855  O.896  O.865  CNN  O.779  O.732  O.606  O.663  O.799  O.875  O.835  SAFE  O.811  O.827  O.559  O.667  O.806  O.806  O.940  O.866  att_RNN  O.681  O.685  O.664  O.694  O.701  O.750  O.747  TextGCN  O.828  O.775  O.735  O.737  O.827  O.828  O.828  MVAE  O.852  O.806  O.719  O.760  O.871  O.817  O.893  SpotFake  O.823  O.743  O.745  O.744  O.864  O.863  O.899  O.885		HMCAN	0.885	0.920	0.845	0.881	0.856	0.926	0.890
TWITTER    CNN   0.549   0.508   0.597   0.549   0.598   0.509   0.550     SAFE   0.766   0.777   0.795   0.786   0.752   0.731   0.742     att_RNN   0.664   0.749   0.615   0.676   0.589   0.728   0.651     EANN   0.648   0.810   0.498   0.617   0.584   0.759   0.660     TextGCN   0.703   0.808   0.365   0.503   0.680   0.939   0.779     MVAE   0.745   0.801   0.719   0.758   0.689   0.777   0.730     SpotFake   0.771   0.784   0.744   0.764   0.769   0.807   0.787     SpotFake*   0.777   0.751   0.900   0.820   0.832   0.606   0.701     SpotFake+   0.790   0.793   0.827   0.810   0.786   0.747   0.766     HMCAN   0.897   0.971   0.801   0.878   0.853   0.979   0.912      SVM-TS   0.639   0.546   0.576   0.560   0.729   0.705   0.717     GRU   0.832   0.782   0.712   0.745   0.855   0.896   0.865     CNN   0.779   0.732   0.606   0.663   0.799   0.875   0.835     SAFE   0.811   0.827   0.559   0.667   0.806   0.940   0.866     att_RNN   0.850   0.791   0.749   0.770   0.876   0.899   0.888     EANN   0.681   0.685   0.664   0.694   0.701   0.750   0.747     TextGCN   0.828   0.775   0.735   0.737   0.827   0.828   0.828     MVAE   0.852   0.806   0.719   0.760   0.871   0.917   0.893     SpotFake   0.823   0.743   0.745   0.744   0.864   0.863   0.863     SpotFake   0.800   0.730   0.668   0.697   0.832   0.869   0.850		SVM-TS	0.529	0.488	0.497	0.496	0.565	0.556	0.561
TWITTER    SAFE		GRU	0.634	0.581	0.812	0.677	0.758	0.502	0.604
TWITTER  att_RNN		CNN	0.549	0.508	0.597	0.549	0.598	0.509	0.550
TWITTER  EANN  0.648  0.810  0.498  0.617  0.584  0.759  0.660  TextGCN  0.703  0.808  0.365  0.503  0.680  0.939  0.779  MVAE  0.745  0.801  0.719  0.758  0.689  0.777  0.730  SpotFake  0.771  0.784  0.744  0.764  0.769  0.807  0.807  0.787  SpotFake*  0.777  0.751  0.900  0.820  0.832  0.606  0.747  0.766  HMCAN  0.897  0.971  0.801  0.878  0.853  0.979  0.912  SVM-TS  0.639  0.546  0.576  0.560  0.729  0.705  0.816  0.865  CNN  0.779  0.732  0.606  0.663  0.799  0.875  0.835  SAFE  0.811  0.827  0.559  0.667  0.806  0.940  0.866  att_RNN  0.850  0.791  0.749  0.770  0.876  0.899  0.888  EANN  0.681  0.685  0.664  0.694  0.701  0.750  0.747  TextGCN  0.828  0.775  0.735  0.737  0.827  0.828  0.828  MVAE  0.852  0.806  0.719  0.760  0.871  0.917  0.893  SpotFake  0.823  0.743  0.745  0.744  0.864  0.863  0.863  0.863  SpotFake  0.823  0.743  0.745  0.744  0.864  0.863  0.869  0.850		SAFE	0.766	0.777	0.795	0.786	0.752	0.731	0.742
TWITTER         TextGCN         0.703         0.808         0.365         0.503         0.680         0.939         0.779           MVAE         0.745         0.801         0.719         0.758         0.689         0.777         0.730           SpotFake         0.771         0.784         0.744         0.764         0.769         0.807         0.787           SpotFake*         0.777         0.751         0.900         0.820         0.832         0.606         0.701           SpotFake+         0.790         0.793         0.827         0.810         0.786         0.747         0.766           HMCAN         0.897         0.971         0.801         0.878         0.853         0.979         0.912           SVM-TS         0.639         0.546         0.576         0.560         0.729         0.705         0.717           GRU         0.832         0.782         0.712         0.745         0.855         0.896         0.865           CNN         0.779         0.732         0.606         0.663         0.799         0.875         0.835           SAFE         0.811         0.827         0.559         0.667         0.806         0.940         0.		att_RNN	0.664	0.749	0.615	0.676	0.589	0.728	0.651
MVAE 0.745 0.801 0.719 0.758 0.689 0.777 0.730 SpotFake 0.771 0.784 0.744 0.764 0.769 0.807 0.787 SpotFake* 0.777 0.751 0.900 0.820 0.832 0.606 0.701 SpotFake+ 0.790 0.793 0.827 0.810 0.786 0.747 0.766 HMCAN 0.897 0.971 0.801 0.878 0.853 0.979 0.912 SVM-TS 0.639 0.546 0.576 0.560 0.729 0.705 0.717 GRU 0.832 0.782 0.712 0.745 0.855 0.896 0.865 CNN 0.779 0.732 0.606 0.663 0.799 0.875 0.835 SAFE 0.811 0.827 0.559 0.667 0.806 0.940 0.866 att_RNN 0.850 0.791 0.749 0.770 0.876 0.899 0.888 EANN 0.681 0.685 0.664 0.694 0.701 0.750 0.747 TextGCN 0.828 0.775 0.735 0.737 0.827 0.828 0.828 MVAE 0.852 0.806 0.719 0.760 0.871 0.917 0.893 SpotFake 0.823 0.743 0.745 0.744 0.864 0.863 0.863 SpotFake+ 0.800 0.730 0.668 0.697 0.832 0.869 0.850	TWITTER	EANN	0.648	0.810	0.498	0.617	0.584	0.759	0.660
SpotFake	IWIIIEK	TextGCN	0.703	0.808	0.365	0.503	0.680	0.939	0.779
SpotFake*   0.777   0.751   0.900   0.820   0.832   0.606   0.701		MVAE	0.745	0.801	0.719	0.758	0.689	0.777	0.730
SpotFake+		SpotFake	0.771	0.784	0.744	0.764	0.769	0.807	0.787
HMCAN         0.897         0.971         0.801         0.878         0.853         0.979         0.912           PHEME         SVM-TS         0.639         0.546         0.576         0.560         0.729         0.705         0.717           GRU         0.832         0.782         0.712         0.745         0.855         0.896         0.865           CNN         0.779         0.732         0.606         0.663         0.799         0.875         0.835           SAFE         0.811         0.827         0.559         0.667         0.806         0.940         0.866           att_RNN         0.850         0.791         0.749         0.770         0.876         0.899         0.888           EANN         0.681         0.685         0.664         0.694         0.701         0.750         0.747           TextGCN         0.828         0.775         0.735         0.737         0.827         0.828         0.828           MVAE         0.852         0.806         0.719         0.760         0.871         0.917         0.893           SpotFake+         0.800         0.730         0.668         0.697         0.832         0.869         <		SpotFake*	0.777	0.751	0.900	0.820	0.832	0.606	0.701
PHEME    SVM-TS		SpotFake+	0.790	0.793	0.827	0.810	0.786	0.747	0.766
PHEME  GRU 0.832 0.782 0.712 0.745 0.855 0.896 0.865  CNN 0.779 0.732 0.606 0.663 0.799 0.875 0.835  SAFE 0.811 0.827 0.559 0.667 0.806 0.940 0.866  att_RNN 0.850 0.791 0.749 0.770 0.876 0.899 0.888  EANN 0.681 0.685 0.664 0.694 0.701 0.750 0.747  TextGCN 0.828 0.775 0.735 0.737 0.827 0.828 0.828  MVAE 0.852 0.806 0.719 0.760 0.871 0.917 0.893  SpotFake 0.823 0.743 0.745 0.744 0.864 0.863 0.863  SpotFake+ 0.800 0.730 0.668 0.697 0.832 0.869 0.850		HMCAN	0.897	0.971	0.801	0.878	0.853	0.979	0.912
PHEME    CNN   0.779   0.732   0.606   0.663   0.799   0.875   0.835     SAFE   0.811   0.827   0.559   0.667   0.806   0.940   0.866     att_RNN   0.850   0.791   0.749   0.770   0.876   0.899   0.888     EANN   0.681   0.685   0.664   0.694   0.701   0.750   0.747     TextGCN   0.828   0.775   0.735   0.737   0.827   0.828   0.828     MVAE   0.852   0.806   0.719   0.760   0.871   0.917   0.893     SpotFake   0.823   0.743   0.745   0.744   0.864   0.863   0.863     SpotFake+   0.800   0.730   0.668   0.697   0.832   0.869   0.850     CNN   0.875   0.835   0.866   0.940   0.876   0.875     O.806   0.791   0.700   0.871   0.917   0.893     SpotFake+   0.800   0.730   0.668   0.697   0.832   0.869   0.850     O.807   0.808   0.809   0.850     O.808   0.809   0.809   0.850     O.809   0.809   0.809   0.809     O.809   0.809     O.809   0.809   0.809     O		SVM-TS	0.639	0.546	0.576	0.560	0.729	0.705	0.717
PHEME  SAFE 0.811 0.827 0.559 0.667 0.806 0.940 0.866  att_RNN 0.850 0.791 0.749 0.770 0.876 0.899 0.888  EANN 0.681 0.685 0.664 0.694 0.701 0.750 0.747  TextGCN 0.828 0.775 0.735 0.737 0.827 0.828 0.828  MVAE 0.852 0.806 0.719 0.760 0.871 0.917 0.893  SpotFake 0.823 0.743 0.745 0.744 0.864 0.863 0.863  SpotFake+ 0.800 0.730 0.668 0.697 0.832 0.869 0.850		GRU	0.832	0.782	0.712	0.745	0.855	0.896	0.865
PHEME  att_RNN		CNN	0.779	0.732	0.606	0.663	0.799	0.875	0.835
PHEME         EANN         0.681         0.685         0.664         0.694         0.701         0.750         0.747           TextGCN         0.828         0.775         0.735         0.737         0.827         0.828         0.828           MVAE         0.852         0.806         0.719         0.760         0.871         0.917         0.893           SpotFake         0.823         0.743         0.745         0.744         0.864         0.863         0.863           SpotFake+         0.800         0.730         0.668         0.697         0.832         0.869         0.850		SAFE	0.811	0.827	0.559	0.667	0.806	0.940	0.866
PHEME         TextGCN         0.828         0.775         0.735         0.737         0.827         0.828         0.828           MVAE         0.852         0.806         0.719         0.760         0.871         0.917         0.893           SpotFake         0.823         0.743         0.745         0.744         0.864         0.863         0.863           SpotFake+         0.800         0.730         0.668         0.697         0.832         0.869         0.850	PHEME	att_RNN	0.850	0.791	0.749	0.770	0.876	0.899	0.888
TextGCN         0.828         0.775         0.735         0.737         0.827         0.828         0.828           MVAE         0.852         0.806         0.719         0.760         0.871         0.917         0.893           SpotFake         0.823         0.743         0.745         0.744         0.864         0.863         0.863           SpotFake+         0.800         0.730         0.668         0.697         0.832         0.869         0.850		EANN	0.681	0.685	0.664	0.694	0.701	0.750	0.747
SpotFake         0.823         0.743         0.745         0.744         0.864         0.863         0.863           SpotFake+         0.800         0.730         0.668         0.697         0.832         0.869         0.850		TextGCN	0.828	0.775	0.735	0.737	0.827	0.828	0.828
SpotFake+   0.800   0.730   0.668   0.697   0.832   0.869   0.850		MVAE	0.852	0.806	0.719	0.760	0.871	0.917	0.893
		SpotFake	0.823	0.743	0.745	0.744	0.864	0.863	0.863
HMCAN   <b>0.881</b>   0.830   0.838   <b>0.834</b>   0.910   0.905   <b>0.907</b>		SpotFake+	0.800	0.730	0.668	0.697	0.832	0.869	0.850
		HMCAN	0.881	0.830	0.838	0.834	0.910	0.905	0.907

- att-RNN has superior performance than GRU.
  - Showing the effectiveness of the attention.
  - It takes into account the text related parts of the image, thus improving the performance of the model.
- MVAE has better performance than single-modal models.
  - Indicates that additional visual information can be used as complementary information.

Dataset	Methods A	Accuracy	Fake news			Real news		
Dataset			Precision	Recall	F1	Precision	Recall	F1
	SVM-TS	0.640	0.741	0.573	0.646	0.651	0.798	0.711
	GRU	0.702	0.671	0.794	0.727	0.747	0.609	0.671
	CNN	0.740	0.736	0.756	0.744	0.747	0.723	0.735
	SAFE	0.763	0.833	0.659	0.736	0.717	0.868	0.785
	att_RNN	0.772	0.854	0.656	0.742	0.720	0.889	0.795
WEIBO	EANN	0.782	0.827	0.697	0.756	0.752	0.863	0.804
WEIBO	TextGCN	0.787	0.975	0.573	0.727	0.712	0.985	0.827
	MVAE	0.824	0.854	0.769	0.809	0.802	0.875	0.837
	SpotFake	0.869	0.877	0.859	0.868	0.861	0.879	0.870
	SpotFake*	0.892	0.902	0.964	0.932	0.847	0.656	0.739
	SpotFake+	0.870	0.887	0.849	0.868	0.855	0.892	0.873
	HMCAN	0.885	0.920	0.845	0.881	0.856	0.926	0.890
	SVM-TS	0.529	0.488	0.497	0.496	0.565	0.556	0.561
	GRU	0.634	0.581	0.812	0.677	0.758	0.502	0.604
	CNN	0.549	0.508	0.597	0.549	0.598	0.509	0.550
	SAFE	0.766	0.777	0.795	0.786	0.752	0.731	0.742
	att_RNN	0.664	0.749	0.615	0.676	0.589	0.728	0.651
TWITTER	EANN	0.648	0.810	0.498	0.617	0.584	0.759	0.660
IWILLER	TextGCN	0.703	0.808	0.365	0.503	0.680	0.939	0.779
	MVAE	0.745	0.801	0.719	0.758	0.689	0.777	0.730
	SpotFake	0.771	0.784	0.744	0.764	0.769	0.807	0.787
	SpotFake*	0.777	0.751	0.900	0.820	0.832	0.606	0.701
	SpotFake+	0.790	0.793	0.827	0.810	0.786	0.747	0.766
	HMCAN	0.897	0.971	0.801	0.878	0.853	0.979	0.912
	SVM-TS	0.639	0.546	0.576	0.560	0.729	0.705	0.717
	GRU	0.832	0.782	0.712	0.745	0.855	0.896	0.865
	CNN	0.779	0.732	0.606	0.663	0.799	0.875	0.835
	SAFE	0.811	0.827	0.559	0.667	0.806	0.940	0.866
PHEME	att_RNN	0.850	0.791	0.749	0.770	0.876	0.899	0.888
	EANN	0.681	0.685	0.664	0.694	0.701	0.750	0.747
	TextGCN	0.828	0.775	0.735	0.737	0.827	0.828	0.828
	MVAE	0.852	0.806	0.719	0.760	0.871	0.917	0.893
	SpotFake	0.823	0.743	0.745	0.744	0.864	0.863	0.863
	SpotFake+	0.800	0.730	0.668	0.697	0.832	0.869	0.850
	HMCAN	0.881	0.830	0.838	0.834	0.910	0.905	0.907

- SAFE outperforms CNN on the three datasets.
  - Because SAFE jointly uses multi-modal (text and visual) and relational information to learn the representation of posts.
- In addition, SpotFake and SpotFake+ achieve better results on all baselines on Twitter and Weibo datasets.
  - Indicating that the pre-trained BERT and XLNet can obtain better textual information to improve model performance.

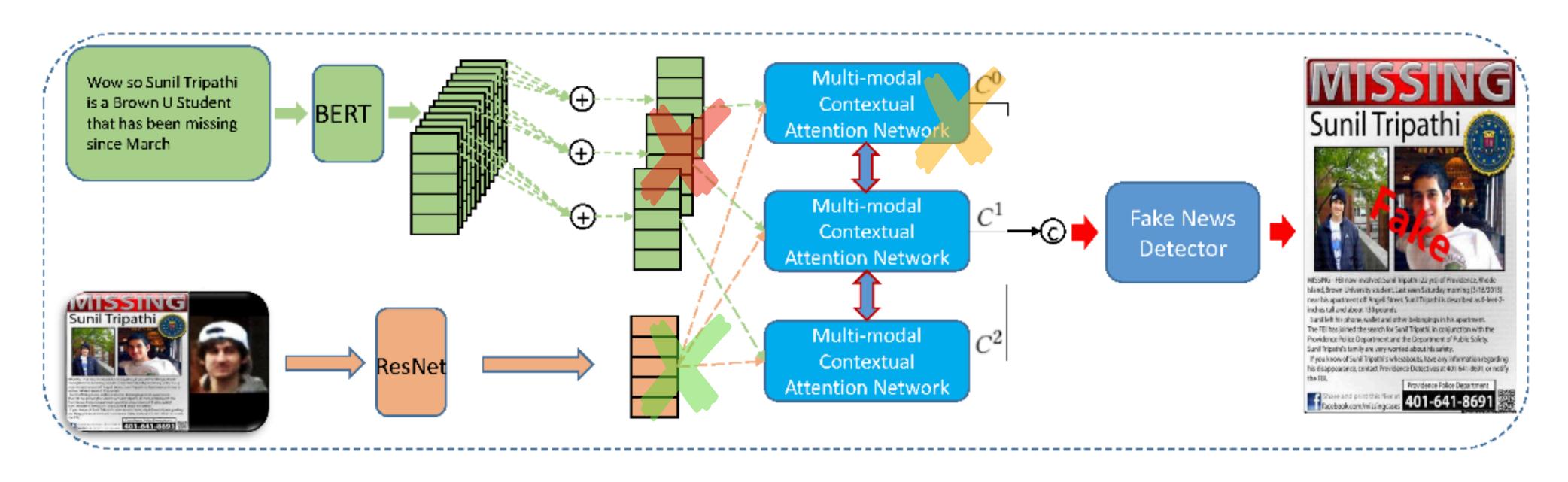
Detect	Methods Ac	A 0011#0017	Fa	Fake news			Real news		
Dataset		Accuracy	Precision	Recall	F1	Precision	Recall	F1	
	SVM-TS	0.640	0.741	0.573	0.646	0.651	0.798	0.711	
	GRU	0.702	0.671	0.794	0.727	0.747	0.609	0.671	
	CNN	0.740	0.736	0.756	0.744	0.747	0.723	0.735	
	SAFE	0.763	0.833	0.659	0.736	0.717	0.868	0.785	
	att_RNN	0.772	0.854	0.656	0.742	0.720	0.889	0.795	
WEIBO	EANN	0.782	0.827	0.697	0.756	0.752	0.863	0.804	
WEIBO	TextGCN	0.787	0.975	0.573	0.727	0.712	0.985	0.827	
	MVAE	0.824	0.854	0.769	0.809	0.802	0.875	0.837	
	SpotFake	0.869	0.877	0.859	0.868	0.861	0.879	0.870	
	SpotFake*	0.892	0.902	0.964	0.932	0.847	0.656	0.739	
	SpotFake+	0.870	0.887	0.849	0.868	0.855	0.892	0.873	
	HMCAN	0.885	0.920	0.845	0.881	0.856	0.926	0.890	
	SVM-TS	0.529	0.488	0.497	0.496	0.565	0.556	0.561	
	GRU	0.634	0.581	0.812	0.677	0.758	0.502	0.604	
	CNN	0.549	0.508	0.597	0.549	0.598	0.509	0.550	
	SAFE	0.766	0.777	0.795	0.786	0.752	0.731	0.742	
	att_RNN	0.664	0.749	0.615	0.676	0.589	0.728	0.651	
TWITTER	EANN	0.648	0.810	0.498	0.617	0.584	0.759	0.660	
IWITER	TextGCN	0.703	0.808	0.365	0.503	0.680	0.939	0.779	
	MVAE	0.745	0.801	0.719	0.758	0.689	0.777	0.730	
	SpotFake	0.771	0.784	0.744	0.764	0.769	0.807	0.787	
	SpotFake*	0.777	0.751	0.900	0.820	0.832	0.606	0.701	
	SpotFake+	0.790	0.793	0.827	0.810	0.786	0.747	0.766	
	HMCAN	0.897	0.971	0.801	0.878	0.853	0.979	0.912	
	SVM-TS	0.639	0.546	0.576	0.560	0.729	0.705	0.717	
	GRU	0.832	0.782	0.712	0.745	0.855	0.896	0.865	
	CNN	0.779	0.732	0.606	0.663	0.799	0.875	0.835	
	SAFE	0.811	0.827	0.559	0.667	0.806	0.940	0.866	
PHEME	att_RNN	0.850	0.791	0.749	0.770	0.876	0.899	0.888	
	EANN	0.681	0.685	0.664	0.694	0.701	0.750	0.747	
	TextGCN	0.828	0.775	0.735	0.737	0.827	0.828	0.828	
	MVAE	0.852	0.806	0.719	0.760	0.871	0.917	0.893	
	SpotFake	0.823	0.743	0.745	0.744	0.864	0.863	0.863	
	SpotFake+	0.800	0.730	0.668	0.697	0.832	0.869	0.850	
	HMCAN	0.881	0.830	0.838	0.834	0.910	0.905	0.907	

- HMCAN outperforms all the baselines on Twitter and PHEME datasets.
- Observe that on Weibo dataset, in the case of fake news, the F1 and accuracy of HMCAN are lower than SpotFake\*, while in the case of real news, the F1 of HMCAN is higher.
- Demonstrate that the proposed model can jointly model multi-modal context information and hierarchical semantics of text in a unified deep model.
  - which can better capture the underlying representation of posts, so as to improve the performance of fake news detection.

Dataset	Methods	Accuracy	Fa	Fake news			Real news		
Dataset			Precision	Recall	F1	Precision	Recall	F1	
	SVM-TS	0.640	0.741	0.573	0.646	0.651	0.798	0.711	
	GRU	0.702	0.671	0.794	0.727	0.747	0.609	0.671	
	CNN	0.740	0.736	0.756	0.744	0.747	0.723	0.735	
	SAFE	0.763	0.833	0.659	0.736	0.717	0.868	0.785	
	att_RNN	0.772	0.854	0.656	0.742	0.720	0.889	0.795	
WEIBO	EANN	0.782	0.827	0.697	0.756	0.752	0.863	0.804	
WEIDO	TextGCN	0.787	0.975	0.573	0.727	0.712	0.985	0.827	
	MVAE	0.824	0.854	0.769	0.809	0.802	0.875	0.837	
	SpotFake	0.869	0.877	0.859	0.868	0.861	0.879	0.870	
	SpotFake*	0.892	0.902	0.964	0.932	0.847	0.656	0.739	
	SpotFake+	0.870	0.887	0.849	0.868	0.855	0.892	0.873	
	HMCAN	0.885	0.920	0.845	0.881	0.856	0.926	0.890	
	SVM-TS	0.529	0.488	0.497	0.496	0.565	0.556	0.561	
	GRU	0.634	0.581	0.812	0.677	0.758	0.502	0.604	
	CNN	0.549	0.508	0.597	0.549	0.598	0.509	0.550	
	SAFE	0.766	0.777	0.795	0.786	0.752	0.731	0.742	
	att_RNN	0.664	0.749	0.615	0.676	0.589	0.728	0.651	
TWITTER	EANN	0.648	0.810	0.498	0.617	0.584	0.759	0.660	
IWITER	TextGCN	0.703	0.808	0.365	0.503	0.680	0.939	0.779	
	MVAE	0.745	0.801	0.719	0.758	0.689	0.777	0.730	
	SpotFake	0.771	0.784	0.744	0.764	0.769	0.807	0.787	
	SpotFake*	0.777	0.751	0.900	0.820	0.832	0.606	0.701	
	SpotFake+	0.790	0.793	0.827	0.810	0.786	0.747	0.766	
	HMCAN	0.897	0.971	0.801	0.878	0.853	0.979	0.912	
	SVM-TS	0.639	0.546	0.576	0.560	0.729	0.705	0.717	
	GRU	0.832	0.782	0.712	0.745	0.855	0.896	0.865	
	CNN	0.779	0.732	0.606	0.663	0.799	0.875	0.835	
	SAFE	0.811	0.827	0.559	0.667	0.806	0.940	0.866	
PHEME	att_RNN	0.850	0.791	0.749	0.770	0.876	0.899	0.888	
	EANN	0.681	0.685	0.664	0.694	0.701	0.750	0.747	
	TextGCN	0.828	0.775	0.735	0.737	0.827	0.828	0.828	
	MVAE	0.852	0.806	0.719	0.760	0.871	0.917	0.893	
	SpotFake	0.823	0.743	0.745	0.744	0.864	0.863	0.863	
	SpotFake+	0.800	0.730	0.668	0.697	0.832	0.869	0.850	
	HMCAN	0.881	0.830	0.838	0.834	0.910	0.905	0.907	

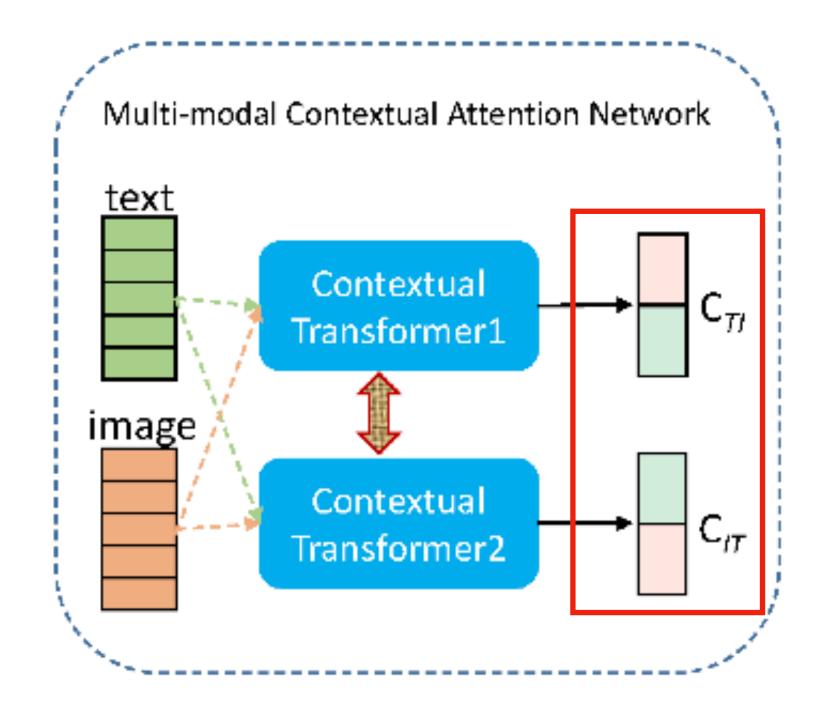
# **Experiments**Ablation Study (1/2)

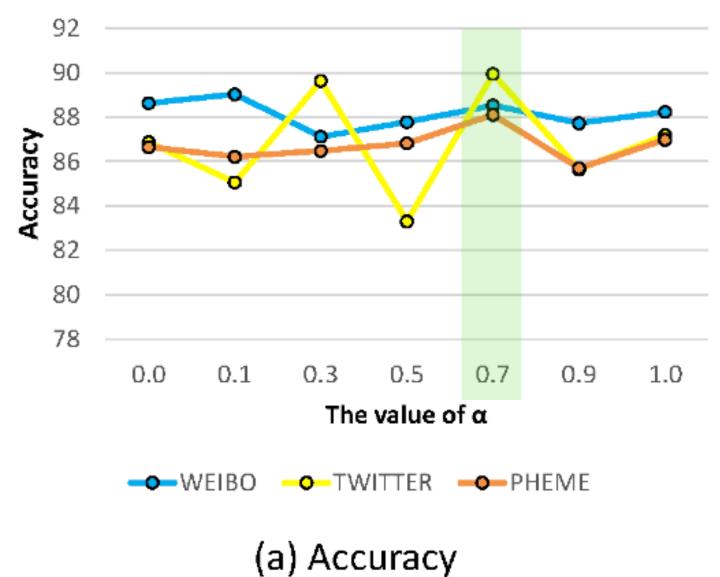
Dataset	Methods	A course our	Fake news			Real news		
Dataset	Wiethous	Accuracy	Precision	Recall	F1	Precision	Recall	F1
	$HMCAN \neg V$	0.809	0.832	0.774	0.802	0.788	0.843	0.815
WEIBO	$HMCAN \neg C$	0.872	0.902	0.836	0.868	0.847	0.909	0.877
WEIBO	$HMCAN \neg H$	0.877	0.871	0.885	0.878	0.883	0.869	0.876
	HMCAN	0.885	0.920	0.845	0.881	0.856	0.926	0.890
	$HMCAN \neg V$	0.755	0.828	0.590	0.689	0.719	0.896	0.798
TWITTER	HMCAN¬C	0.790	0.886	0.622	0.731	0.743	0.932	0.827
IWILLER	$HMCAN \neg H$	0.879	0.884	0.849	0.866	0.875	0.906	0.890
	HMCAN	0.897	0.971	0.801	0.878	0.853	0.979	0.912
	$HMCAN \neg V$	0.854	0.814	0.763	0.788	0.873	0.904	0.888
PHEME	HMCAN¬C	0.858	0.788	0.821	0.804	0.899	0.878	0.888
	$HMCAN \neg H$	0.871	0.808	0.828	0.818	0.906	0.894	0.900
	HMCAN	0.881	0.830	0.838	0.834	0.910	0.905	0.907

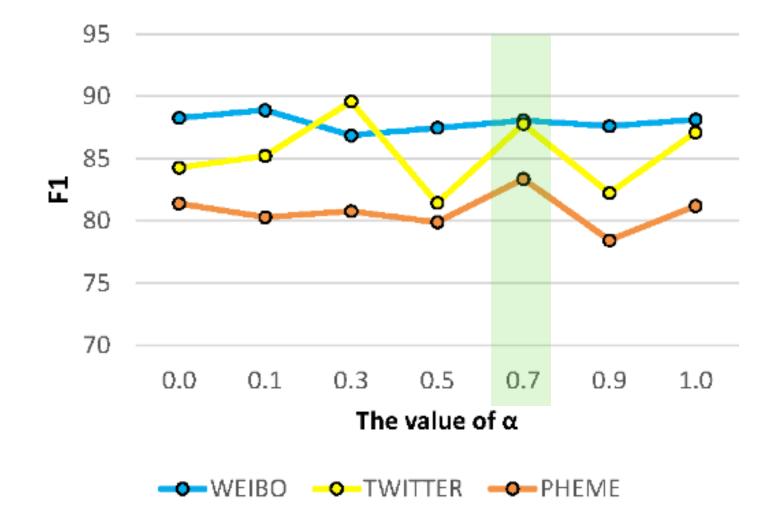


### Experiments

#### Impact of the value $\alpha$





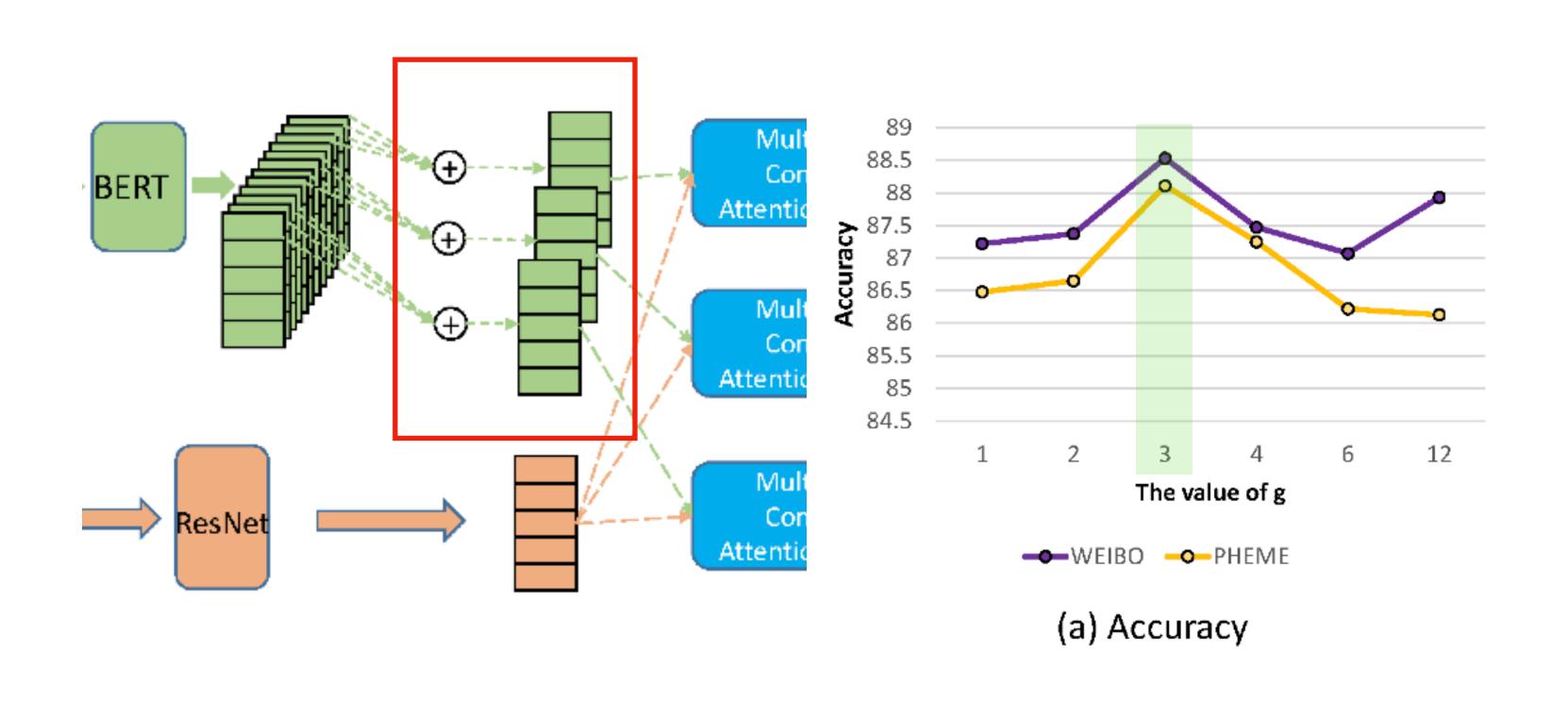


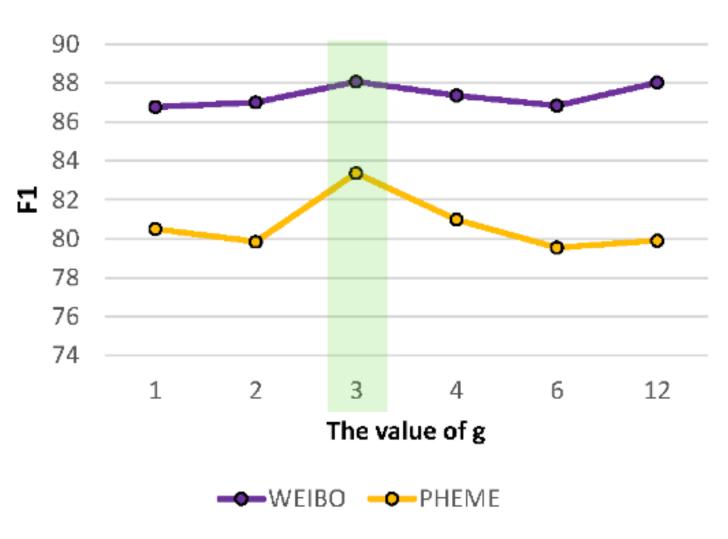
- $C = \alpha C_{TI} + \beta C_{IT}$
- $\alpha + \beta = 1$

(b) F1 score of fake news

## Experiments

### Impact of the value g





(b) F1 score of fake news

# Conclusion of HMCAN

- Propose a novel hierarchical multi-modal contextual attention network (HMCAN) for fake news detection task.
  - To jointly model the multi-modal context information and the hierarchical semantics of text in a unified deep model.
    - Employ ResNet to learn better representations of images and utilize BERT to embed the textual content of news.
    - A multi-modal contextual attention network is proposed to fuse both inter-modality and intramodality relationships.
    - Design a hierarchical encoding network to capture the rich hierarchical semantics for fake news detection.

## **Comments** of HMCAN

- Simple but effective method.
  - Use BERT & ResNet as initial encoder.
  - Design a contextual attention network to get complementary information.
  - Utilize the advantage of BERT intermediate layer to get more information.
- Parameter not enough flexible and a little sensitive.