Human-Computer Interaction in Multimodal Fake News Detection

Mateusz Walczak
Lodz University of Technology,
Lodz, Poland
mateusz.walczak@dokt.p.lodz.pl

Aneta Poniszewska-Marańda[1]
Lodz University of Technology,
Lodz, Poland
aneta.poniszewska-maranda@p.lodz.pl

ABSTRACT

In the era of digital information, the problem of fake news has become one of the most serious challenges for society. In response to this, researchers from various fields are focusing on developing advanced methods for detecting fake news. Multimodal fake news detection requires advanced algorithms and technologies that can analyze and integrate diverse data. Using a multimodal approach that integrates different data sources, such as text, image, and audio, allows for a more comprehensive analysis and identification of fake news. In the context of Human-Computer Interaction (HCI), it is crucial that these technologies are not only effective, but also intuitive and easy for users to use. Studies show that interactive interfaces that visualize the fake news detection process can significantly improve users' ability to identify disinformation. The paper presents the proposal of the method for evaluating the quality of diagrams visualizing advanced multimodal models for fake news detection that facilitate understanding and verification of information. The proposed method includes nine metrics divided into two categories: qualitative metrics, which relate to the subjective assessment of the model's visualization, and quantitative metrics, which assess whether the diagram includes the key elements of the model architecture.

Keywords: Fake news, Multimodal fake news detection, Human computer interaction, Graphical visualization.

INTRODUCTION

[1] Corresponding author. aneta.poniszewska-marańda@p.lodz.pl + 48 42 631 27 96

One of the biggest problems faced by today's world is the effective verification of information. Therefore, in both social media and traditional media, an increasing number of news articles are fake. Widespread dissemination of fake news carries underlying negative effects on individuals and society. Certain individuals propagate inaccurate or deceptive content across social media platforms with the intent of gaining attention or achieving financial and political benefits.

The recognition of fake news is one of the most popular and fastest growing research areas in natural language processing and beyond. For several years, a growing interest in the issue of recognizing fake news can be observed. Each successive year, more and more effective and advanced methods for recognizing fake news are being developed (Khattar et al. 2019, Wang et al. 2020, Qian et al. 2021, Chen et al. 2022, Xiong et al. 2023). However, there are still many potential directions for the development of algorithms and methods to detect fake news.

State-of-the-art models for detecting fake news use multiple machine learning and deep learning methods, including numerous Natural Language Processing (NLP) techniques. The high complexity of the model architecture is one of the biggest difficulties faced by the reader of an article proposing a new method for fake news detection. To make the structure of the model easier to digest, researchers often include a graphical visualization of the model, taking into account its various modules. To date, no research has focused on assessing the visualization quality of advanced fake news detection models. To address the challenges posed by complex multimodal detection models, integrating Human-Computer Interaction (HCI) principles becomes essential for enhancing user comprehension and interaction.

Human-Computer Interaction (HCI) is a study dedicated to improving user-computer interaction by creating user-friendly computer interfaces. This interdisciplinary field includes computer science, behavioral and cognitive science, psychology, design principles, and ergonomics. One of the research areas covered by the broad spectrum of HCI is the creation and application of methods for the quantitative and qualitative evaluation of human-computer communication interfaces (Keebler et al. 2020, Phan et al. 2016, Setiyawati et al. 2022).

Within the framework of this paper, we discuss an issue at the intersection of the fields of fake news detection and Human-Computer Interaction. The paper presents the proposal of the method for evaluating the quality of diagrams visualizing advanced multimodal models for fake news detection. The structure of the paper is as follows: section 2 presents the background explaining the concept of multimodal models in fake news detection area. Section 3 describes the proposed method in detail, while section 4 deals with the results of conducted experiments for selected papers using the proposed method. Finally, section 5 contains a brief discussion and final conclusions.

MULTIMODAL MODELS FOR FAKE NEWS DETECTION

Depending on the formal representation of news, it is possible to distinguish two general approaches to solving the problem of fake news recognition (Figure 1). The first group of solutions are methods that implement a unimodal approach (Du et al. 2021, Ilie et al. 2021, Palani and Elango 2023, Dahou et al. 2023, Raja et al. 2023, Praseed et al. 2023). In unimodal models, news is represented only by text. In this approach, detecting fake news is in fact a simple classification of text into one of two or more categories that determine its degree of truthfulness.

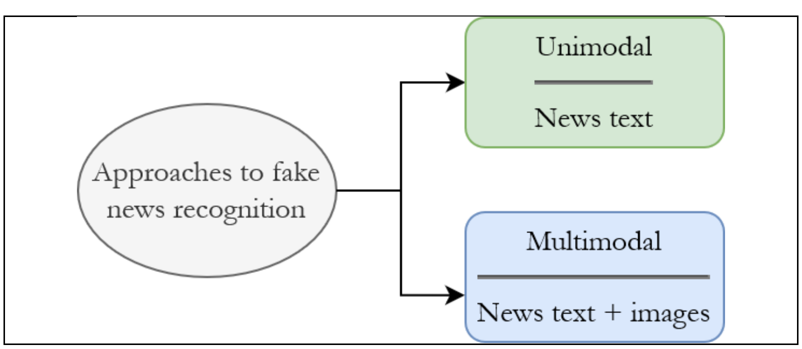

Figure 1. Classification of approaches to solving the problem of fake news recognition

The second approach involves using a more complex news representation, consisting of two modalities. Multimodal models take into account both the text of the news story and the images within it (Khattar et al. 2019, Wang et al. 2020, Qian et al. 2021, Yuan et al. 2021, Qian et al. 2021, Chen et al. 2022, Hua et al. 2023, Obaid et al. 2023, Xiong et al. 2023, Peng et al. 2024, Qu et al. 2024). Solutions of this type are much more complicated than unimodal models and consist of many integrated modules, such as an image feature extraction module, a text feature extraction module, a feature fusion module and a final classification module.

Each of the aforementioned modules can be highly complex and may incorporate a variety of machine learning and deep learning methods. For example, the text feature extraction module frequently utilizes NLP techniques such as Bi-directional Encoder Representation from Transformers (BERT) (Wang et al. 2020, Qian et al. 2021, Yuan et al. 2021) or XLNet (Qian et al. 2021, Peng et al. 2024). In contrast, the image feature extraction module typically relies on image recognition methods like Residual Network (ResNet) (Qian et al. 2021, Hua et al 2023) or VGG-19 (Qu et al. 2024, Xiong et al. 2023), which are both implementations of Convolutional Neural Networks (CNNs). Additionally, the feature fusion module employs a separate set of deep learning methods, combining textual and graphical features through networks such as Long Short-Term Memory (LSTM) or Bidirectional Long Short-Term Memory (BiLSTM) (Yuan et al. 2021, Xiong et al. 2023). The final classification module generally consists of several Fully Connected Layers (FCLs) or a standalone multilayer perceptron (MLP) (Chen et al 2022, Dahou et al 2023, Hua et al 2023).

Visualizing complex multimodal models with diverse modules and an extensive range of machine learning and deep learning techniques, including NLP, poses significant challenges. By contrast, unimodal models that rely solely on textual data are much simpler, making it easier to represent their architecture in a diagram. Therefore, the proposed method focuses exclusively on multimodal solutions.

PROPOSED METHOD

This section provides a detailed description of the proposed method for assessing the quality of Graphical Visualization of multimodal models for Fake News Detection (GVFND). The demonstrated method applies only to diagrams in graphical form. All tables and additional descriptions for visualization of models are not considered. The proposed GVFND method assesses the quality, readability, and informativeness of model visualizations, rather than evaluating the model's accuracy or effectiveness.

The GVFND method consists of 9 metrics divided into 2 categories. Qualitative metrics are related to the reader's subjective assessment of the readability, informativeness and intuitiveness of the model's visualization. Quantitative metrics are used to verify that the various elements of the model architecture are included in the diagram. All metrics allow for the awarding from 0 to 2 points, where 0 points means complete failure to meet the metric's objectives, and 2 points is the highest score indicating that the metric's requirements have been met. Table 1 summarizes and describes all qualitative metrics of the GVFND method. Table 2 provides a similar overview for quantitative metrics.

Table 1. Qualitative metrics of the GVFND method

Id	Name	Description
QL_1	Readability	The overall readability of the visualization: the resolution of the diagram, the color scheme used, the size ratio of captions to graphic elements, etc.
QL_2	Comprehension speed	Time in which the reader gains a sense of understanding of the general way the model works. Under 15 seconds - 2 points, under a minute - 1 point, over a minute - 0 points.
QL_3	Overwhelm level	A high level of overwhelm - 0 points, a low level - 1 point, and no overwhelm at all - 2 points

Table 2. Quantitative metrics of the GVFND method

Id	Name	Description
QN_1	Text features extractor	Visualization of the text feature extraction module. Complete visualization - 2 points, incomplete visualization - 1 point, no visualization - 0 points.
QN_2	Image features extractor	Visualization of the image feature extraction module. Complete visualization - 2 points, incomplete visualization - 1 point, no visualization - 0 points.
QN_3	Feature fusion module	Visualization of the features fusion module. Complete visualization - 2 points, incomplete visualization - 1 point, no visualization - 0 points.
QN_4	Classifier	Visualization of the final classification process. Complete visualization - 2 points, incomplete visualization - 1 point, no visualization - 0 points.
QN_5	Neuron counts	Does the visualization include the number of neurons? Yes, in each module/layer - 2 points, only in some modules/layers - 1 point, not at all - 0 points.
QN_6	Activation functions	Does the visualization include activation functions? Yes, in each module/layer - 2 points, only in some modules/layers - 1 point, not at all - 0 points.

The final model visualization quality score is a weighted average of the points awarded by each metric. Since the main purpose of visualizing a model is to graphically present the most complete possible description of its architecture, it was decided that quantitative metrics have a higher weight (0.6) than qualitative metrics (0.4). Thus, the final score obtained by the GVFND method is calculated according to the following formula:

$$S_{GVFND} = \left(\frac{S_{QL}}{6} * 0.4 + \frac{S_{QN}}{12} * 0.6\right) * 100\%$$

where S_{QL} is the sum of the scores obtained in all three qualitative metrics and S_{QN} is the sum of the scores obtained in all six quantitative metrics. From the above formula, it follows that the final model visualization quality score S_{GVFND} is given on a percentage scale.

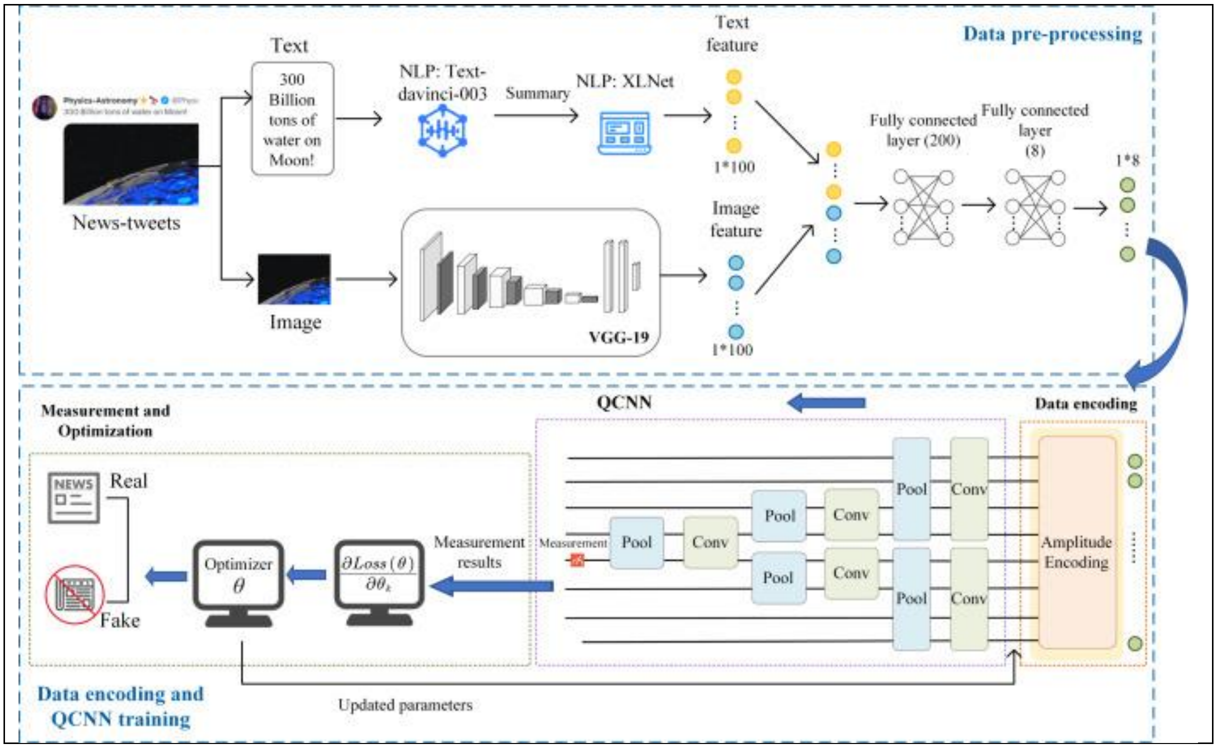

Figure 2. Visualization of quantum multimodal fusion-based fake news detection model (Qu et al. 2024)

Figure 2 provides an example visualization of the multimodal fake news detection model (Qu et al. 2024). The diagram clearly distinguishes all four modules of the model and shows the number of neurons in each fully connected layer. However, it does not include detailed information on the sizes of the convolutional and pooling layers. The visualization also does not include the activation functions in the individual layers of the model. The overall readability of the diagram is very satisfactory, while the comprehension time and level of overwhelm are moderate. Detailed results of the GVFND evaluation of this and other selected visualizations are presented in the next section.

EXPERIMENTAL RESULTS

The GVFND method was used to evaluate the visualization of multimodal models for detecting fake news proposed by 11 selected articles. For each article, a comprehensive evaluation of each quantitative and qualitative metric was conducted, as described in the previous section. Each diagram was evaluated by 3 judges who are experts in multimodal fake news detection. When judges assigned different scores for a single metric, the final score was calculated as the arithmetic mean of the scores, rounded to the nearest integer, in accordance with the scoring guidelines outlined in tables 1 and 2.

Table 3 summarizes the results of the experiments. For each model visualization evaluated, the scores for individual metrics ($QL_1 – QL_3$ and $QN_1 – QL_6$), the sum of scores for qualitative and quantitative metrics (S_{QL}, S_{QN}) and the overall final score (S_{GVFND}) calculated according to formula in page 6 are provided. The final row of the table lists the calculated average values for individual metrics, the sums of S_{QL} and S_{QN}, and the final score S_{GVFND}.

Table 3. GVFND evaluation results for 11 selected articles

Ref.	**QL_1**	**QL_2**	**QL_3**	**S_{QL}**	**QN_1**	**QN_2**	**QN_3**	**QN_4**	**QN_5**	**QN_6**	**S_{QN}**	**S_{GVFND}**
(Chen et al. 2022)	2	1	2	**5**	0	0	2	0	0	1	**3**	**48.3%**
(Hua et al. 2023)	2	2	2	**6**	1	1	2	0	0	0	**4**	**60.0%**
(Khattar et al. 2019)	2	1	1	**4**	2	2	1	2	0	0	**7**	**61.6%**
(Obaid et al. 2023)	2	1	2	**5**	1	0	2	0	0	1	**4**	**53.3%**
(Peng et al. 2024)	0	0	0	**0**	0	0	2	1	0	0	**3**	**15.0%**
(Qian et al. 2021a)	1	1	1	**3**	1	1	2	0	0	0	**4**	**40.0%**
(Qian et al. 2021b)	1	0	0	**1**	2	2	2	0	0	0	**6**	**36.6%**
(Qu et al. 2024)	2	1	1	**4**	2	1	2	1	1	0	**7**	**61.6%**
(Wang et al. 2020)	1	2	2	**5**	0	0	1	0	0	1	**2**	**43.3%**
(Xiong et al. 2023)	2	1	2	**5**	1	1	2	2	0	0	**6**	**63.3%**
(Yuan et al. 2021)	2	2	2	**6**	1	1	2	2	0	0	**6**	**70.0%**
Average	1.54	1.09	1.36	**4.00**	1.00	0.82	1.82	0.73	0.09	0.27	**4.72**	**50.3%**

The final results of model visualization quality for the selected articles range from 15% to 70%. The average final score only slightly exceeds 50%. The sum of scores for qualitative metrics obtained by the analyzed visualizations varies across the possible range: from 0 (Peng et al. 2024) to 6 points (Yuan et al. 2021, Hua et al. 2023). For qualitative metrics, the best score turned out to be 7 points (Khattar et al. 2019, Qu et al. 2024), and the lowest was 2 points (Wang et al. 2020). These results show that the GVFND method effectively differentiates visualization quality while balancing qualitative and quantitative metrics.

It is noteworthy that many papers (Chen et al. 2022, Hua et al. 2023, Obaid et al. 2023, Wang et al. 2020) scored very well on qualitative metrics ($S_{QL} \geq 5$) while scoring low on quantitative metrics ($S_{QN} \leq 4$). These results highlight the considerable challenge of balancing readability and clarity with the informativeness and completeness of the visualization.

Another important observation is the very low average scores for the quantitative metrics QN_5 ($AVG_{QN5} = 0.09$) and QN_6 ($AVG_{QN6} = 0.27$). None of the evaluated diagrams achieved the maximum score ($QN_5 = 2$, $QN_6 = 2$) in these categories. These results suggest that the scale used for QN_5 and QN_6 may be too broad, and that a binary scale (0–1) would likely be more suitable for these metrics.

The best final score of 70% was obtained by the visualization of the model proposed by the authors of (Yuan et al. 2021). The best visualization received the maximum number of points for qualitative metrics and half of the possible points for quantitative metrics. By far the worst model visualization quality at 15% was received for (Peng et al. 2024), which scored no points for qualitative metrics and only 3 points for quantitative metrics. The results presented show that there is still a great potential for progress in terms of the visualization quality of multimodal models for fake news detection.

CONCLUSIONS

The research carried out is a pilot study aimed at combining the fields of fake news detection and Human-Computer Interaction. One of the main goals of the experiments performed was to emphasize the importance of creating clear and understandable visualizations of complex models using deep learning and machine learning methods along with Natural Language Processing techniques. Intuitive, informative and well-organized visualizations of models significantly accelerate the understanding of the principle of the models by the reader of the article.

This paper proposes the GVFND method for assessing the quality of visualization of multimodal models for detecting fake news. The GVFND method can effectively determine both the qualitative and quantitative qualities of multimodal model visualization. The conducted experiments revealed that the final results of the quality of the visualizations obtained using the GVFND method are highly diversified and indicate numerous shortcomings and deficiencies even in the best rated visualizations.

There are many potential directions for the development of research on methods for evaluating the quality of visualizations of complex models. In the future, research can be expanded to test more visualizations of the models proposed in various articles. Secondly, the evaluation of visualizations can be carried out by a large group of readers with different levels of expertise in deep learning and fake news detection. Averaging the results of the metrics determined by a large group of testing readers will avoid the extreme subjectivity of the results obtained, which will translate into an increase in their overall value.

REFERENCES

Chen, Y., Li, D., Zhang, P., Sui, J., Lv, Q., Tun, L., and Shang, L. 2022. "Cross-Modal Ambiguity Learning for Multimodal Fake News Detection," *Proceedings of the ACM Web Conference 2022.WWW'22*, Lyon, France: Association for Computing Machinery, pp. 2897-2905.

Dahou, A., Ewees, A. A., Hashim, F. A., Al-qaness, M. A. A., Orabi, D. A., Soliman, E. A., Tag-eldin, E. M., Aseeri, A. O., and Elaziz, M. A. 2023. "Optimizing fake news detection for Arabic context: A multitask learning approach with transformers and an enhanced Nutcracker Optimization Algorithm," *Knowledge-Based Systems* (280), p. 111023.

Du, J., Dou, Y., Xia, C., Cui, L., Ma, J., and Yu, P. S. 2021. "Cross-lingual COVID-19 Fake News Detection," Proceedings of 2021 International Conference on Data Mining Workshops (ICDMW), pp. 859-862.

Hua, J., Cui, X., Li, X., Tang, K., and Zhu, P. 2023. "Multimodal fake news detection through data augmentation-based contrastive learning," *Applied Soft Computing* (136), p. 110125. ISSN: 1568-4946.

Ilie, V. I., Truica, C. O., Apostol, E., S., and Paschke, A. 2021. "Context-Aware Misinformation Detection: A Benchmark of Deep Learning Architectures Using Word Embeddings," *IEEE Access* (9) pp. 162122-162146. ISSN: 2169-3536.

Keebler, J. R., Shelstad, W. J., Smith, D. C., Chaparro, B. S., and Phan, M. H. (2020). "Validation of the GUESS-18: a short version of the Game User Experience Satisfaction Scale (GUESS)," *Journal of Usability Studies*, (16:1), p. 49.

Khattar, D., Goud, J. S., Gupta, M., and Varma, V. 2019. "MVAE: Multimodal Variational Autoencoder for Fake News Detection," *Proceedings of The World Wide Web Conference. WWW '19*. San Francisco, CA, USA: Association for Computing Machinery, pp. 2915-2921. ISBN: 9781450366748.

Obaid, A. A., Khotanlou, H., Mansoorizadeh, M., and Zabihzadeh, D. 2023. "Robust Semi-Supervised Fake News Recognition by Effective Augmentations and Ensemble of Diverse Deep Learners," *IEEE Access* (11), pp. 54526-54543. ISSN: 2169-3536.

Palani, B., and Elango, S. 2023. "CTrL-FND: content-based transfer learning approach for fake news detection on social media," *International Journal of System Assurance Engineering and Management* (14:3), pp. 903-918. ISSN: 0976-4348.

Peng, L., Jian, S., Kan, Z., Qiao, L., and Li, D. 2024. "Not all fake news is semantically similar: Contextual semantic representation learning for multimodal fake news detection," *Information Processing & Management* (61:1), p. 103564. ISSN: 0306-4573.

Phan, M. H., Keebler, J. R., and Chaparro, B. S. (2016). "The development and validation of the game user experience satisfaction scale (GUESS), " *Human factors*, (58:8), pp. 1217-1247.

Praseed, A., Rodrigues, J., and Thilagam, P. S. 2023. "Hindi fake news detection using transformer ensembles," *Engineering Applications of Artificial Intelligence* (119), p. 105731. ISSN: 0952-1976.

Qian, S., Wang, J., Hu, J., Fang, Q., and Xu, C. 2021. "Hierarchical Multi-Modal Contextual Attention Network for Fake News Detection," P*roceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. SIGIR '21*. Association for Computing Machinery, pp. 153-162. ISBN: 9781450380379.

Qian, S., Hu, J., Fang, Q., and Xu, C. 2021. "Knowledge-Aware Multi-Modal Adaptive Graph Convolutional Networks for Fake News Detection," *ACM Transsaction on Multimedia Computing, Communication and Applications* (17:3). ISSN; 1551-6857.

Qu, Z., Meng, Y., Muhammad, G., and Tiwari, P. 2024. "QMFND: A quantum multimodal fusion-based fake news detection model for social media," *Information Fusion*, p. 102172. ISSN: 1566-2535.

Raja, E., Soni, B., and Borgohain, S. K. 2023. "Fake news detection in Dravidian languages using transfer learning with adaptive finetuning," *Engineering Applications of Artificial Intelligence* (126), p. 106877. ISSN: 0952-1976.

Setiyawati, N., and Bangkalang, D. H. 2022. "The Comparison of Evaluation on User Experience and Usability of Mobile Banking Applications Using User Experience Questionnaire and System Usability Scale," *Proceedings* (82:1), p. 87, MDPI.

Wang, Y., Qian, S., Hu, J., Fang, Q., and Xu, C. 2020. "Fake News Detection via Knowledge-driven Multimodal Graph Convolutional Networks," *Proceedings of 2020 International Conference on Multimedia Retrieval. ICMR '20*. Dublin, Ireland. ssociation for Computing Machinery, pp. 540-547. ISBN: 9781450370875.

Xiong, S., Zhang, G., Batra, V., Xi, L., Shi, L., and Liu, L. 2023. "TRIMOON: Two-Round Inconsistency-based Multi-modal fusion Network for fake news detection," *Information Fusion* (93), pp. 150-158. ISSN: 1566-2535.

Yuan, H., Zheng, J., Ye, Q., Qian, Y., and Zhang, Y. 2021. "Improving fake news detection with domain-adversarial and graph-attention neural network," *Decision Support Systems* (151), p. 113633. ISSN: 0167-9236.