Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 7
TPCEE 2024

Leveraging Transfer Learning for Enhanced Internet Security: Methods and Applications

Chang Chen *

Collage of artificial intelligence, Tianjin University of Science and Technology, Tianjin, China

* Corresponding author: 2553728769@mail.tust.edu.cn

Abstract. Internet technology is developing rapidly today. It not only brings new experiences to people but also brings huge security risks. The attack methods of cyber criminals are becoming more and more complex and unpredictable. The rise of artificial intelligence offers promising solutions for detecting and mitigating network attacks, but challenges such as insufficient and imbalanced datasets continue to hinder the effectiveness of these models. Transfer learning has emerged as a valuable technique to address these challenges by transferring knowledge and parameters from a source domain model to a target domain. This approach can reduce training costs, save time, enhance data utilization, and address data imbalance issues. This article provides a detailed examination of transfer learning, including its definition, methodologies, and specific applications within the Internet domain. It explores three key scenarios: network anomaly detection, malicious domain name detection, and rumor detection. By offering a comprehensive overview of transfer learning and its practical implementations, this paper hopes to help readers acquire a clear and comprehensive concept of transfer learning and its relevance to Internet security, serving as a valuable resource for those interested in this field.

Keywords: Transfer Learning; Internet Security; Data Imbalance; Network Attacks.

1. Introduction

With the rapid development of Internet technology, the Internet has brought many conveniences to people. However, some people with ulterior motives have taken advantage of Internet loopholes to attack and steal user data, which has brought people huge security risks. Network intrusion, malicious domain names, abnormal network traffic, and other issues have always bothered people. As long as people are still using the Internet, these problems inevitably affect people's online security. Therefore, people urgently need some efficient and useful methods to solve these problems and detect and prevent cyber-attacks.

In recent years the use of deep learning has provided people with many effective methods to solve these problems. Some scholars have proposed a new deep-learning classification model based on a stacked Neural Autoregressive Density Estimator (NDAE) in order to detect network intrusion [1]. In addition, some scholars have used deep reinforcement learning and semi-supervised dual-depth Q network (SSDDQN) to provide a new detection optimization method for network abnormal traffic monitoring [2]. The method of using convolutional neural networks (CNN) to process data with strong spatial correlation and long short-term memory (LSTM) models to process sequence data has also shown excellent results in the field of malicious domain name detection [3]. However, these methods generally have problems such as insufficient and unbalanced labeled data, high classification ambiguity, and high overhead. In some field training a suitable model can be tough and costly, or the trained model has poor performance. Thus, transfer learning attracts people's attention as an effective way to solve these problems. In recent years, a large number of technologies that use transfer learning technology to solve Internet problems have been proposed, such as the abnormal network traffic detection method based on MobileNet-V2 transfer learning [4], the network traffic classification method based on dynamic balance adaptive transfer learning [5], the migration application of the Efficientnet-B0 model pre-trained based on the ImageNet dataset [6], and the network intrusion detection technology based on Variational Auto Encoder (VAE) oversampling and transfer learning and other technologies [7-11]. The application of transfer learning effectively solves the shortcomings of traditional methods and solves existing problems well.

Transfer learning is suitable for small sample learning. Transfer learning is different from traditional machine learning in that it can directly transfer trained models instead of training a new model from scratch. It transfers the already labeled valid data or trained models in similar fields to the target field, so that effective models for specific fields can be quickly built, greatly saving the cost of training models. It is only necessary to perform simple fine-tuning training on the transferred model under the target data set, and finally connect a prediction layer or classification layer to quickly build a neural network with better results as shown in Fig. 1. Transfer learning has realized the mutual migration and reuse of data in different fields, has well solved the problem of insufficient and unbalanced data, and reduced the possibility of overfitting. Transfer learning can also enhance the module's robustness, strengthen the model's stability, make the model perform better when facing unfamiliar data, and realize the customization of the model. However, transfer learning may not be able to directly complete the transfer when faced with fields with large differences, so researchers need to make sure the selected data sets are accurate and adopt appropriate methods to process them. At the same time, the parameter convergence of transfer learning is not very good, and the selection of weights depends on the personal experience and technology of the trainer. If you choose not to add, negative transfer is likely to occur, which has a negative impact on the learning of the target field. The author of this article introduces the principles of transfer learning to readers in detail and introduces the detailed application of transfer learning in the Internet field. The author hopes this paper can help readers learn more about transfer learning and its application on the Internet.

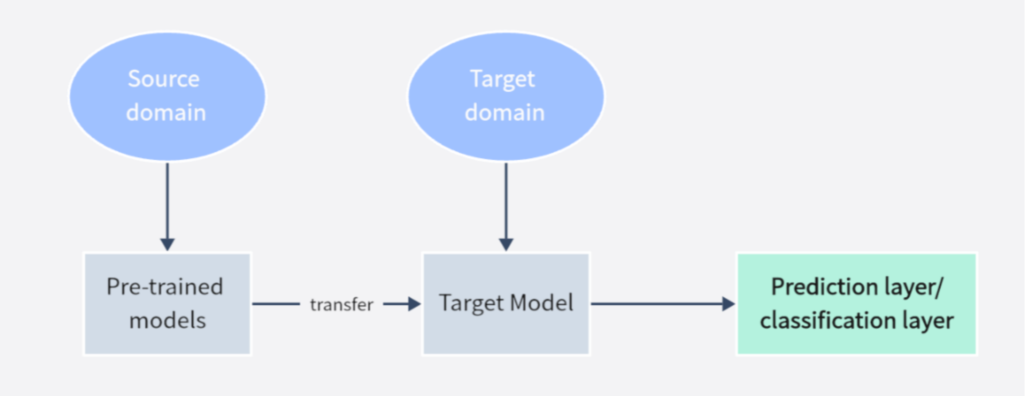

Figure 1. Transfer learning diagram

2. Methodology

2.1. Dataset Description and Preprocessing

In terms of the use of datasets, researchers need to choose appropriate datasets based on their own research directions and fields. Therefore, the datasets used in different research fields are normally different, but in general, they all need to meet the standard that the dataset used is similar to the target field. In the field of network intrusion, datasets such as Knowledge Discovery and Data Mining Cup 1999 (KDD Cup 99）, New Subset of the KDD (NSL-KDD), and Canadian Institute for Cybersecurity Intrusion Detection Dataset 2017 (CICIDS2017) are generally used. In the field of abnormal network traffic detection datasets containing abnormal and normal network traffic. Besides in fields such as rumor detection, special data such as product reviews are used as datasets [8]. In terms of data processing, data cleaning and feature extraction are generally performed on the data, and its format is processed to meet the processing requirements of the model.

2.2. Proposed Approach

This article provides a comprehensive overview of the application methods for transfer learning in various contexts. It outlines the main technical steps involved, which are divided into five key phases: data preprocessing, training source domain models, applying transfer learning, fine-tuning for the target domain, and model verification and evaluation. Each phase is discussed in detail to give readers a thorough understanding of how transfer learning is practically applied. In addition to the general methodology, the article explores three practical scenarios where transfer learning is utilized within the internet domain: anomaly detection in network traffic, malicious domain name detection, and rumor detection. For each scenario, the article describes the specific data processing techniques used, the models employed, the transfer learning processes implemented, and the final verification and evaluation procedures. This detailed examination aims to provide readers with valuable insights and practical references, enhancing their understanding of how transfer learning can be effectively applied in real-world internet applications.

2.2.1. Introduction to basic methods.

This section introduces the specific application steps of transfer learning in detail. The specific steps are shown in Fig. 2. First of all, data preprocessing requires selecting a suitable data set related to the target domain. Generally, a public or previously summarized data set is selected. Then the data needs to be cleaned and preprocessed to remove missing values, outliers, and duplicate values. In the feature extraction operation of the data, researchers can use different models such as Residual Network (ResNet) for processing, and convert the data into a format suitable for model processing. After data processing, the required model can be trained in the source domain. The training models that can be used include CNN, LSTM, etc. Then, the optimization algorithm is used to adjust the model to obtain a model that meets the requirements. Then, transfer learning can be performed to migrate the trained source domain model to the target domain. Direct migration or fine-tuning of parameter models can be adopted to adjust, and the feature distribution is adjusted according to the target domain for domain adaptation. In the fine-tuning part of the target domain, the pre-trained model and other required parts are first combined, and the newly added partial model or the overall model is trained using the target domain data. Parameters such as learning rate and batch size are optimized so that the model can better solve and cope with the problems of the target domain. Finally, the model is verified and evaluated, and reasonable experimental methods and experimental sets are designed, such as cross-validation, holdout method, and other methods for verification, and the experimental results are evaluated. The experimental results are generally judged from several aspects such as accuracy, precision, recall, loss, etc. Through in-depth analysis and processing of experimental results, the model structure hyperparameters, generalization ability as well as robustness of the model are optimized and enhanced.

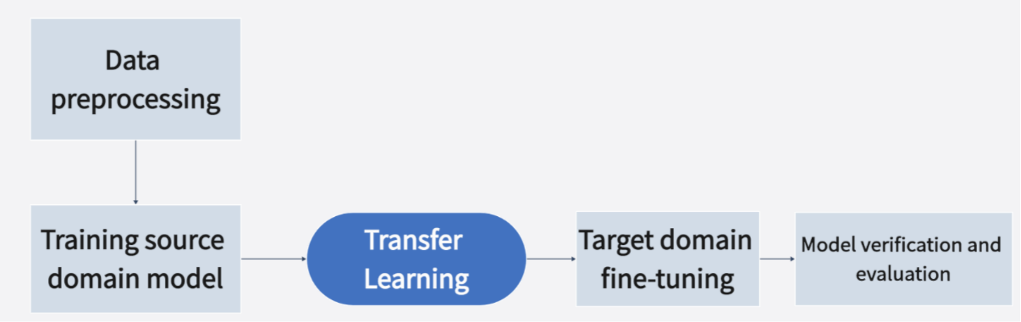
Data preprocessing
Training source domain model
Transfer Learning
Target domain fine-tuning
Model verification and evaluation

Figure 2. Transfer learning application process

2.2.2. Abnormal network traffic detection.

This section introduces how the MobileNet-V2 model combined with transfer learning is used to detect abnormal network traffic [4]. The model of MobileNet-V2 is shown in Fig. 3. The first layer of the model is an expansion layer to expand the feature dimension. The depth wise separable convolution structure is adopted in the second layer, using 3*3 depth wise separable convolution. It not only uses depth wise convolution for feature extraction but also uses point convolution to improve efficiency and reduce computational complexity. Then, the ReLu6 activation function is used for normalization, and the final result is output to a fully connected layer or a global average pooling layer with the activation function of SoftMax [4]. The specific implementation scheme is to first preprocess the data and input the processing results into the feature extraction layer. The first layer of the feature extraction layer is the standard MobileNet-V2. The MobileNet-V2 is migrated to the model required for the experiment. After MobileNet-V2 processing, the output result is connected to a fully connected layer with 32 neurons. After that the activation function ReLU is used in the fully connected layer to distinguish normal network traffic from abnormal network traffic [4]. This implementation method is simple and efficient, can be equipped on any modern device, and has good performance in different situations.

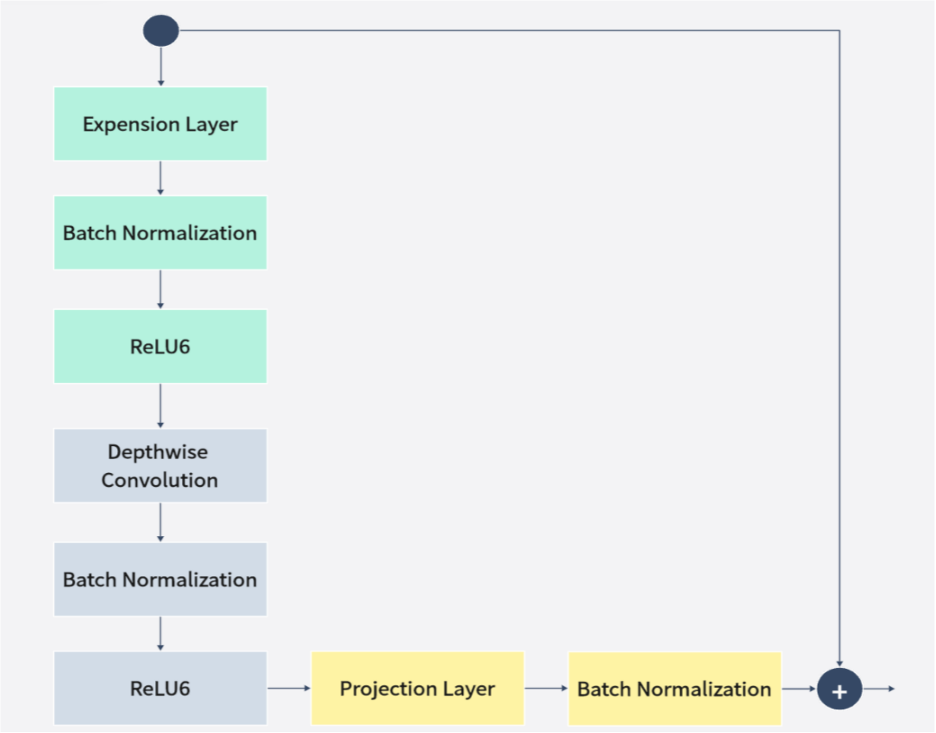

Figure 3. MobileNet-V2 model structure diagram

2.2.3. Introduction to basic methods.

This section introduces the application of the Bidirectional LSTM-CNN (BiLSTM-CNN) model combined with transfer learning technology in malicious domain name detection [12]. The framework of the BiLSTM-CNN detection algorithm based on transfer learning is shown in Fig. 4. First, the second-layer domain name and the remaining sub-layer domain names are extracted, and the data is regularized using the embedding layer. Then, the URL is converted into a numerical vector using distributed encoding DWR, the string length is specified, the part longer than the specified length is cut or truncated, and the part shorter than the specified length is supplemented with a zero vector. The specific formula is shown in Formula 1. In the formula: $F(url_i)$ is the regularized domain name string vector; url_i is the domain name string to be adjusted; V_z is the zero vector [12].

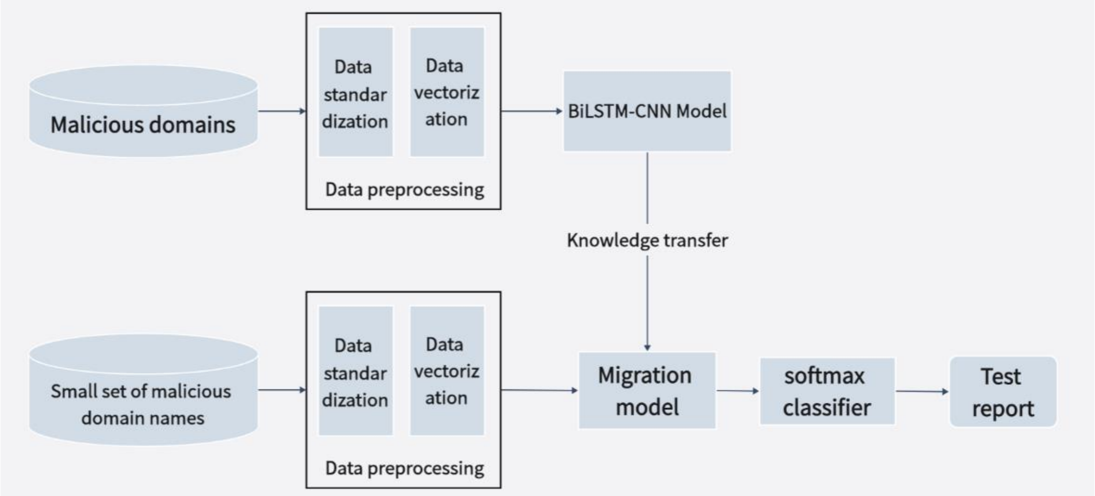

Figure 4. Algorithm framework

$$F(url_i) = \begin{cases} V_z + url_i & length(url_i) < length \\ url_i & length(url_i) = length \\ url_i[0, length - 1] & length(url_i) > length \end{cases} \tag{1}$$

Then, BiLSTM is used to extract the domain name context relationship, and the output result is used as the input content of CNN. The CNN convolution operation extracts the deep features of the semantic space dimension, and then the average pooling method is used to reduce the computational complexity while retaining important features [12].

Finally, Dropout neurons are used to randomly discard data before the fully connected layer which will prevent overfitting of data and the trained model parameters are transferred to the small sample malicious domain name detection task, and the final parameter fine-tuning is performed in combination with the existing data. Finally, it can ensure that the trained BiLSTM-CNN model also has good performance in small sample malicious domain names.

2.2.4. Rumor detection.

This section introduces the application of Weibo rumor detection model (transferring learn-Directional Gated Recurrent Unit and Two-Layer CNN (transferring learn-BiGRU-2-CNN), Two-Layer Bi-Directional Gating Loop Unit (TB2GC)) combined with transfer learning technology in malicious domain name detection [8]. The model that proposed in this experiment is shown in Fig. 5. Here are the specific steps: First, use the word2vec model to vectorize the text, and then input the obtained vector data into the BiGRU2 and CNN joint neural network. Compared with LSTM, the BiGRU model has the characteristics of simple structure and good parameters and can obtain context information in both directions. The CNN layer extracts local features reduces errors through pooling operations, and finally uses the SoftMax function for classification [8].

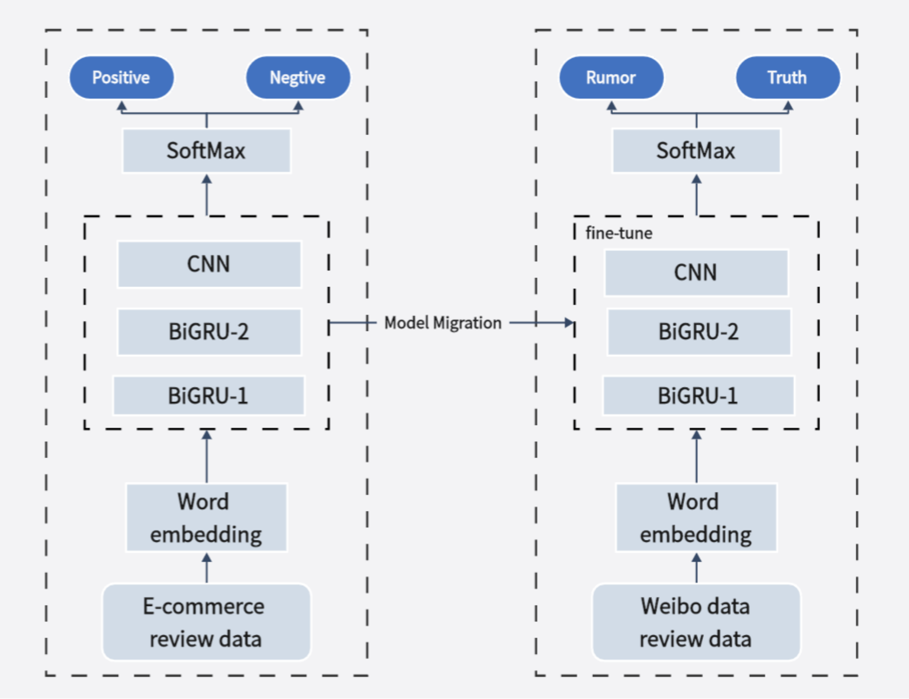

Figure 5. TB2GC model structure

A large amount of comment data is used by the researchers to pre-train the TB2GC neural network. After that, the researchers migrate the model to the target task. For the target detection task, they adopted the oblique triangle learning strategy to quickly determine the appropriate parameter space, and used the discriminative fine-tuning method to set different learning rates for each layer, and finally extracted it to the SoftMax layer, using the ReLU activation function to output whether it is a rumor [8]. The application of transfer learning here has a significant improvement in the accuracy and the model's generalization ability. It also gives the model a chance to perform well when labeled data is scarce.

3. Result and Discussion

The unique advantages of transfer learning have been described in detail in the previous article. In order to allow readers to intuitively understand the advantages of the methods and models used above, this section gives the experimental results and data comparisons of the experimental models and methods introduced in the previous article. At the same time, this section also discusses the development trends and advantages and disadvantages of transfer learning.

3.1. Result Analysis

In the network abnormal traffic detection experiment, the four values of accuracy, precision, recall, as well as F1 score are used as the measurement criteria, and the results of multiple experiments are shown in Table 1 [4]. It can get plenty of information from the table. For example, the model has excellent results in accuracy, precision, recall, as well as F1 score, fully demonstrating the superiority of the model compared with other models in detecting the network abnormal traffic. The model uses transfer learning technology to migrate the pre-trained standard model so that it has good feature extraction capabilities and robustness, which greatly improves the performance of the model.

Table 1. Test set results

Serial number	Accuracy	Recall	Precision	F1 score
1	0.970421	0.877512	0.974283	0.923369
2	0.976284	0.890594	0.978246	0.932364
3	0.980021	0.876249	0.974528	0.922779

In the malicious domain name detection experiment, the experimental standards use Accuracy, Precision, false positive rate (FPR) as well as false negative rate (FNR). The higher the Accuracy and Precision, the lower the FPR and FNR, indicating that the model effect is better. The detection results of the experimental model on the family malicious domain name set with sufficient data and the small sample malicious domain name set are shown in Table 2 and Table 3 [12]. You can see from the table that the model has excellent performance both on the family malicious domain name set with sufficient data and on the small sample malicious domain name set. By migrating the pre-trained BiLSTM-CNN model parameters to the field of small-sample malicious domain names, this experiment greatly improves the model's ability in the field of small-sample malicious domain name detection, demonstrating the model's strong generalization ability when facing different problems.

Table 2. Detection performance of family malicious domain sets with sufficient data

Type	Accuracy	Accuracy	FPR	FNR
Tinba	96. 80	94.04	4. 11	5. 28
Dyre	96. 80	97. 18	3.20	3. 09
Virut	91. 06	90. 42	8 39	8.54
Shifu	97. 45	97. 11	2.22	2.24
Emotet	97. 34	98.53	2.26	2. 08
Simda	90.81	88.06	9. 02	9. 44
Ramdo	93. 56	94.84	7. 22	6. 96
Qakbot	90.37	92.10	9. 49	9. 17
Rovnix	95. 88	94. 06	4. 23	4. 09
Chinad	92.44	93.61	7. 31	7. 28
Necurs	97. 02	96. 83	3.17	3. 11
Locky	97. 64	97.90	3.10	3. 08
Symmi	96 51	95.88	3.50	3.61
Gameover	95.86	96.41	4.66	4.04
Murofet	96.37	97.65	4.23	3.58
Cryptolocker	95.02	96.38	4.36	4.22
Dircrypt	97.19	96.52	3.09	3.16
Suppobox	90.28	90.10	9.64	9.79
Rammit	96.11	94.35	3.14	3.60
Ranbyus	97.01	97.22	3.19	3.15
Kraken	94.54	96.03	5.34	4.92
Pykspa	94.69	93.82	4.75	4.91
Shiotob	97.16	97.02	2.26	2.30
Banjori	94.39	92.14	5.15	5.60
Corebot	94.89	95.13	5.74	5.26
Blacklist Set	97.31	96.59	2.87	3.01

Table 3. Small sample malicious domain detection results

Type	Accuracy	Accuracy	FPR	FNR
Gspy	95.54	94.80	4. 04	4 20
Mirai	95.38	95. 52	4. 33	4. 18
Feodo	94 17	94.01	5. 97	6. 01
Qadars	93.61	93.52	6. 34	6. 34
Tofsee	94.05	93.60	6. 59	6. 98
Matsnu	94.83	94. 91	5.58	5.61
Fobber	96.38	96.23	3. 71	3.77
Bamital	95.07	94.68	4. 05	4 26
Madmax	94.17	94. 11	5.06	5.09
Mydoom	90.44	90.47	9.22	9.13
Bigviktor	93.22	93.26	95.27	7. 92
Blackhole	94.16	94.25	4. 48	4.50
Conficker	95.60	94.38	5.23	5. 94
Tempedreve	95.11	95.64	3.61	3. 52
Tinynuke	93.68	92. 91	7.20	7. 94
Nymaim	93.88	93.12	7. 32	7. 50
Padcrypt	95.01	95.33	4 77	4.68
Proslikefan	92 25	92.18	8.24	8.19
hellghost	94.24	93.66	6.18	6.39
hellghost	93.94	93.22	6.38	6.62
Vawtrak	95.10	95.27	4. 66	4. 64

In the field of rumor detection, the experiment also uses the four indicators of Method, Accuracy, Recall, and F1 as the experimental indicators. The comparison results of the model used in the experiment with the baseline models and the decomposition model are shown in Table 4 and Table 5 [8]. What you can see from the comparison results is that in the field of rumor detection, the model which is used by the researchers has better performance in the four indicators compared with other models. The experiment first uses comment data for model training and then migrates to the field of Weibo rumor detection. The use of transfer learning greatly improves the model's ability to judge and process different rumors, ensuring the robustness and flexibility of the model.

Table 4. Comparison results between TB2GC model and baseline model

Method	Accuracy	Accuracy	Recall	F1
DT Rank	0.732	0.738	0.715	0.72
DTC	0.831	0.847	0.815	0.83
SVM TS	0.857	0.839	0.885	0.86
GRU	0.908	0.871	0. 958	0. 913
GRU-2	0.910	0.876	0. 956	0.914
CNN	0.933	0. 921	0. 945	0.933
TB2GC	0.962	0. 953	0. 963	0.958

Table 5. Comparison results between TB2GC model and decomposition model

Method	Accuracy	Accuracy	Recall	F1
GRU	0.941	0.947	0.926	0.937
CNN	0. 946	0. 945	0. 938	0. 941
Bi-GRU	0. 953	0. 952	0.951	0.948
BiGRU-CNN	0.957	0.951	0.954	0.951
TB2GC	0.962	0.958	0.963	0.961

3.2. Discussion

In complex and dynamic internet environments, transfer learning offers significant advantages by accelerating model development, reducing construction costs, and enhancing data reuse. It also improves model generalization and robustness, enabling effective performance across diverse scenarios. This capability extends its importance beyond internet applications to fields. However, successful implementation of transfer learning hinges on a precise understanding of the relationship between the source and target domains. If the gap between these domains is too wide, negative transfer may occur, which can degrade model performance and hinder construction efficiency. Therefore, careful consideration and judgment are essential when applying transfer learning to ensure that the transfer is beneficial rather than detrimental. Additionally, transfer learning requires meticulous adjustment of various parameters, necessitating coordination across multiple steps and tasks. This process can be challenging and demands substantial expertise and technical skills from researchers. Effectively managing these adjustments is crucial for achieving optimal outcomes and ensuring that the model performs well in its intended application.

4. Conclusion

This article offers a detailed exploration of transfer learning applications within Internet technology, integrating contemporary trends and cutting-edge knowledge. It comprehensively introduces the concept of transfer learning, outlining its advantages, limitations, and specific use cases within Internet technology, including network anomaly detection, malicious domain name detection, and rumor detection. The aim is to provide newcomers to the field with a thorough and accessible understanding of transfer learning and its practical implementations. The article clearly delineates the experimental methods used in each application scenario, highlights the benefits of various models, and details the specific steps involved in applying transfer learning. This analysis underscores the current strengths and versatility of transfer learning in addressing diverse challenges. Additionally, the article acknowledges the limitations of transfer learning, such as the complexity of parameter adjustment and the risk of negative transfer. It is anticipated that future research will address these shortcomings, driving further innovation and expansion in the field. By providing a comprehensive overview and identifying areas for improvement, this article aims to facilitate the broader application and advancement of transfer learning across different domains.

References

[1] Shone N. Ngoc T.N. Phai V.D. et al. A deep learning approach to network intrusion detection. IEEE transactions on emerging topics in computational intelligence, 2018, 2 (1): 41 - 50.

[2] Dong S. Xia Y. Peng T. Network abnormal traffic detection model based on semi-supervised deep reinforcement learning. IEEE Transactions on Network and Service Management, 2021, 18 (4): 4197 - 4212.

[3] Vinayakumar R. Soman K.P. Poornachandran P. Detecting malicious domain names using deep learning approaches at scale. Journal of Intelligent & Fuzzy Systems, 2018, 34 (3): 1355 - 1367.

[4] Chen G. Anomaly traffic detection method based on transfer learning of MobileNet-V2. Ship electronic engineering, 2024, 44 (02): 133 - 137.

[5] Shang F.J. Li S.S. Wang Y. et al. Traffic classification method based on dynamic balance adaptive transfer learning. Journal of Electronics and Information, 2022, 44 (09): 3308 - 3319.

[6] Zhang S.R. Chen B. Bu Y.J. et al. Encryption malicious traffic detection method based on transfer learning. Computer engineering and applications, 2022, 58 (17): 130 - 138.

[7] Huang Z.Y. Yang Y.G. Lei Z.C. Application of VAE oversampling and transfer learning in network intrusion detection. Computer age, 2021, (07): 50 - 54.

[8] Shen R.L. Pan W.M. Zhang H.J. Microblogging rumor detection method based on transfer learning. Computer engineering and design, 2021, 42 (12): 3534 - 3539.

[9] Ma H.Y. Bidirectional authentication of network identity based on transfer learning. Microcomputer application, 2022, 38 (09): 163 - 165.

[10] Wu Y.L. Zhang Z. Li T.Y. et al. Inter-site XSS intrusion detection method based on transfer learning. Science and technology information, 2023, 21 (19): 39 - 42.

[11] Wu P. Guo H. Buckland R. A transfer learning approach for network intrusion detection. IEEE 4th international conference on big data analytics (ICBDA). IEEE, 2019: 281 - 285.

[12] Zhao F. Zhao H. Chang Z.B. Small sample malicious domain name detection based on transfer learning. Computer engineering and design, 2022, 43 (12): 3381 - 3387.