International Journal of Innovative Research in Computer Science and Technology (IJIRCST)
ISSN (Online): 2347-5552, Volume-12, Issue-5 September 2024
https://doi.org/10.55524/ijircst.2024.12.5.13
Article ID IJIRD-1354, Pages 94-99
www.ijircst.org

Dual-Branch Dynamic Graph Convolutional Network for Robust Multi-Label Image Classification

Bingxing Wang[1], Hongye Zheng[2], Yingbin Liang[3] , Guanming Huang[4], and * Junliang Du[5]

[1] Illinois Institute of Technology, Chicago, USA
[2] Chinese University of Hong Kong, Hong Kong, China
[3] Northeastern University, Seattle, USA
[4] University of Chicago, Chicago, USA
[5] Shanghai Jiao Tong University, Shanghai, China

Correspondence should be addressed to Junliang Du: jldu@acm.org

Received 4 September 2024; Revised 18 September 2024; Accepted 29 September 2024

ABSTRACT- For the intricate task of multi-label image classification, this paper introduces an innovative approach: an attention-guided dual-branch dynamic graph convolutional network. This methodology is designed to address the difficulties faced by current models when handling multiple labels within images. By integrating multi-scale features, it enhances the retention of original category information and boosts the robustness of feature learning. Utilizing a semantic attention module, the study dynamically reweights feature categories in the training dataset, enhancing the network's capability to identify smaller objects and generate context-sensitive category representations. The effectiveness of the proposed model was evaluated using the MS-COCO2014 imagery dataset, demonstrating superior performance in critical metrics such as classification precision (CP), recall (CR), and F1 score (CF1), outperforming other state-of-the-art models. Furthermore, a cascaded classification structure was implemented to leverage the prior information from static images to inform the processing of dynamic ones, and to utilize original image category data to augment label correlations, thereby enhancing overall classification accuracy.

KEYWORDS- Multi-label Image Classification, Graph Convolutional Networks, Attention Mechanisms, Dual-Branch Dynamic Networks

I. INTRODUCTION

The multi-label image classification task aims to identify objects in a single image to determine whether there are one or more different object categories. It has a wide range of applications in various computer vision fields, such as image annotation, social network data mining, medical diagnosis, and human attribute recognition.

In the current multi-label image classification research field [1], traditional strategies often rely on manually designed feature extraction techniques, such as scale-invariant feature transform (SIFT) [2] and bag-of-words (BOW) [3]. However, these methods capture superficial image information and are highly dependent on feature selection, making it difficult to adapt to the complex needs of the big data era. In the flood of big data, it has become almost impossible to manually construct a comprehensive and in-depth image feature description. Fortunately, with the rise of neural network models [4], they have demonstrated excellent performance and unlimited potential in visual recognition tasks with their ability to learn automatically, greatly promoting the innovation process in the field of multi-label image classification. However, due to the complex spatial layout, cluttered background and occlusion between different labeled objects in multi-label images, the multi-label image classification task is limited to being treated as a special single-label classification and the traditional binary classification algorithm [5] is used to deal with related problems.

In recent years, graph convolutional networks (GCNs) are often used to model label dependencies in the study of multi-label image classification tasks, and some progress has been made. However, traditional graph network models often calculate the label co-occurrence probability of training data through preset graphs, which causes the loss of underlying information and reduces the generalization of the model [6,7]. At the same time, there is a problem of excessive computational space requirements during model training, which makes the model classification accuracy low [8]. To this end, this paper starts from the use of underlying image information by attention-driven graph convolutional networks and the construction of graph neural network graph structures and proposes a dynamic graph convolutional network. The relevant research content and innovative work are as follows:

- A dual-branch dynamic graph convolutional network based on attention-driven is proposed. By introducing multi-scale features of images [9], the original category information lost in previous dynamic neural networks is protected, providing support for improving the robustness of learning features.
- The model uses the semantic attention module to calculate the training data and redistribute the feature category weights to improve the network's recognition ability for small target labels, thereby generating content-aware category representations. A cascade classification structure is designed. The main path uses the preset static image as prior information to provide guidance for the

dynamic image; the auxiliary branch directly inputs the content-aware category representation into the dynamic image and uses the original image category information to supplement the label correlation.

II. RELATED WORK

Recent The field of multi-label image classification has made significant progress through the use of deep learning models, which are adept at learning complex feature representations from large-scale datasets. Traditional methods, such as scale-invariant feature transform (SIFT) and bag-of-words (BOW), were once prominent but have proven limited in their ability to handle the intricate nature of multi-label datasets where multiple objects with overlapping features may appear in a single image.

Graph neural networks (GNNs) have become an effective solution for modeling relationships between entities and labels in multi-label classification tasks. GNN-based approaches have demonstrated strong performance in domains where capturing the dependencies between labels is crucial, such as sentiment analysis and dynamic data modeling [10-11]. These methods offer a more sophisticated approach to image classification by learning the co-occurrence of labels and relationships between object categories, which is key in multi-label scenarios. Another critical advancement in deep learning for multi-label classification is the introduction of attention mechanisms. Attention mechanisms enhance the model's ability to selectively focus on important regions of the input data. This is particularly useful in tasks where multiple labels are present, and the model must distinguish between overlapping objects within an image. Attention-based models have shown notable improvements in text and image classification tasks, where they help prioritize relevant features for better classification performance [12]. This concept has been applied in various fields, including breast cancer detection using attention mechanisms [13] and image registration tasks utilizing dense U-Net architectures with channel attention [14]. Contrastive learning, another emerging technique, enables models to learn distinct features that differentiate between closely related categories. This is highly relevant for multi-label classification, where objects in the same image might belong to closely related but distinct categories. Contrastive learning allows the model to refine its feature space, leading to more accurate label predictions [15].

The integration of spatiotemporal features is another essential area in multi-label classification, particularly when the data involves dynamic elements or requires an understanding of both local and global contexts. Spatiotemporal feature extraction techniques have proven effective in time-series data and applications like fraud detection, providing insights into how models can handle multi-scale features in image classification tasks [16]. Additionally, hybrid frameworks that integrate LSTM with models like GARCH have been explored for tasks such as financial risk prediction, which is an example of how dynamic data dependencies are captured in various domains [17].

Moreover, deep learning models have benefited from advancements in optimization techniques. Optimized gradient descent methods have enhanced the training efficiency of neural networks, particularly in handling large datasets like those used in multi-label image classification. This helps models converge faster and perform better on complex datasets by improving training processes [18].

Finally, pre-trained models have been increasingly explored in fields such as named entity recognition, showcasing the potential of transfer learning for optimizing performance across diverse tasks [19]. This approach can similarly be applied to image classification, where pre-trained models may help transfer learning features across domains to improve classification accuracy.

Incorporating these advancements, this paper introduces a dual-branch dynamic graph convolutional network (DB-DGCN) for multi-label image classification. The DB-DGCN leverages both GNNs and attention mechanisms to dynamically capture label dependencies and multi-scale image features, resulting in improved classification accuracy and robustness. The ability to focus on relevant parts of the image and model label correlations makes this approach particularly effective in handling the complex nature of multi-label classification tasks.

III. ALGORITHM PRINCIPLE

Multi-label classification tasks are difficult to handle because one image often contains multiple labels [20]. In recent years, relevant theories and methods based on deep learning have received widespread attention. Convolutional neural network (CNN) is currently the most efficient image processing technology. The network model consists of three levels: convolution, activation function, and aggregation, and can represent the feature space of each image. In terms of image classification, the output of CNN is used as input, and the fully connected layer is used to classify the image. Figure 1 shows the entire architecture. Among them, the most important task is to realize the continuous iterative correction of network weights based on the model training of the samples, that is, the backpropagation algorithm.

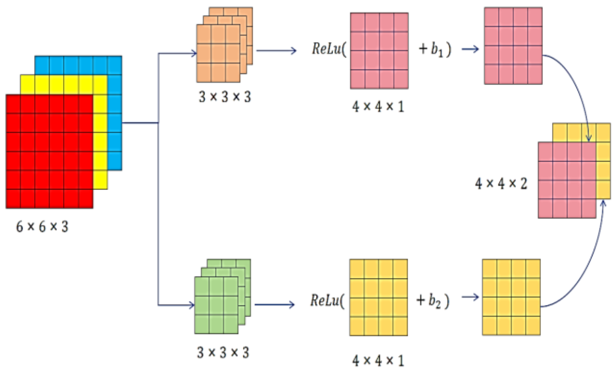

Figure 1: CNN network

In practical applications, it has been proven that deep learning networks can effectively capture more complex information and features. However, experiments have observed a phenomenon: as the network layer deepens, its optimization efficiency weakens, and the accuracy of the validation set and training set also declines simultaneously. This is mainly because when the network structure deepens, it will encounter challenges such as gradient explosion and gradient disappearance [21]. In order to cope with these challenges, ResNet, also known as a residual network, is proposed to enable the creation of a depth-adjustable architecture through the stacking of multiple layers. This paper introduces a novel approach that builds upon the ResNet framework. The network is characterized by a series of residual blocks, where the fundamental equation representing these blocks can be stated as:

$$x_l + 1 = x_l + F(x_l, w_l)$$

In the formula, the residual block is comprised of two key components: the direct mapping portion and the residual portion. $B(x_l)$ represents the direct mapping part, which is represented by the curve on the left in Figure 2; $F'(x_l, w_l)$ represents the residual term, and the residual part generally includes two to three convolution operations, which is the so-called block, as shown in Figure 2.

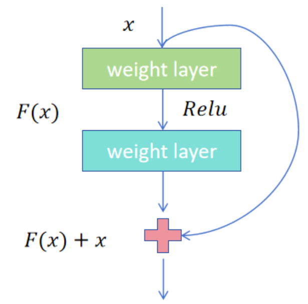

Figure 2: Residual Network Structure

Recent exploration has found that applying graph convolutional networks (GCN) to depict the correlation between labels shows significant effects in improving the recognition accuracy of multi-label image classification. However, relying on label co-occurrence statistics to construct maps may affect the generalizability of the model, thereby losing a large amount of basic category information, especially in scenarios where uncommon joint objects appear in the test images. Therefore, in this chapter, an attention-driven dual-branch dynamic graph convolutional network is proposed to obtain more effective image features through main/auxiliary branch content-aware category representation learning. The model proposed in this article uses a dynamic graph convolutional network. Its main path uses preset static images as prior information to provide guidance for dynamic graphs, while the auxiliary branch directly inputs content-aware category representations into dynamic graphs, using the original image categories. Information supplement tag relevance.

For each image, the convolutional neural network (CNN) is used to obtain the convolutional feature representation of the multi-label image. In this process, the feature representation of the image is transformed into a category representation V that reflects the image content by applying the semantic attention module:

$$V = SAM(CNN(x))$$

Compared with the traditional multi-label image classification algorithm research, the D-GCN module in the dual-branch dynamic graph convolutional neural network proposed in this chapter can perform adjacency matrix transformation according to the input image features, and transform the predefined correlation matrix A into the correlation matrix A_d, which is specially adapted to the corresponding multi-label image samples. The correlation matrix A_d is expressed as follows:

$$Ad = \delta(W_A H')$$

Therefore, this paper designs a multi-label relationship model based on dynamic graph network topology and analyzes it.

As indicated above, the final category representation $Z = [Z_1,\ Z_2, ..., Z_c]$ is used for the final classification. Since the vector feature Z_i accurately corresponds to the category to which it belongs and carries rich and complex inter-class associations, only these category vectors need to be integrated into the binary classifier to effectively improve the accuracy of the classification prediction. In the process of data integration, this paper adopts a unique method to connect the scores of each category in series to form the final score vector $s_r = [s_r^1, s_r^2, ..., s_r^C]$. In addition, an adaptive score is designed, which is derived from the SAM estimate of the category-specific activation map M, and is aggregated in the global space to further enhance the robustness of the prediction. This adaptive score is combined with the original score vector, and a more credible final score $s_m = [s_m^1, s_m^2, ..., s_m^C]$ is extracted through a simple weighted average. The final score s is an important supervisory score during the classification experiment, and the entire network is trained using the traditional multi-label classification loss, as shown below:

$$L(y,s) = \sum_{c=1}^{C} y^c \log(\sigma(s^c)) + (1-y^c)\log(1-\sigma(s^c))$$

In the formula, $\sigma(\cdot)$ is the Sigmoid function, y is the true category label of the sample, and s represents the predicted score for category c.

IV. EXPERIMENTAL DESIGN

A) Experimental setup

The objective of this study is to employ a set of well-established evaluation metrics to ensure a fair benchmarking against existing research outcomes. These metrics include overall and category-specific precision (OP and CP), recall (OR and CR), and F1 scores (OF1 and CF1), alongside the mean average precision (mAP). In the computation of these measures, a prediction is classified as a true positive if its confidence threshold surpasses 0.5. Among these, the OF1, CF1, and mAP are deemed particularly significant for assessing performance.

When constructing the overall structure of the model, this paper uses ResNet-101 as the core architecture. In particular, in the SAM module and D-GCN unit, the category feature expression is carefully designed, the dimension of the V channel is increased to 1024, and the LeakyReLU nonlinear activation function is used with a slope set to -0.2 to enhance the expressiveness of the model. In order to ensure the generalization ability of the model, strategic data enhancement techniques are adopted in the training phase, including randomly cropping the image size to 448x448 pixels and introducing random horizontal flipping operations to reduce the risk of model overfitting. In order to speed up the convergence of the model and improve the classification accuracy, this paper chooses the MS-COCO dataset, which is commonly used and more challenging in multi-label classification tasks, for pre-training the model. For the optimization problem of the network, the stochastic gradient descent (SGD) with an initial momentum of 0.9 and a weight decay value of 0.001 as the default value is used as the optimizer. For each GPU, the batch size is 18. The initial learning rate of the SAM/D-GCN network is set to 0.5, and the learning rate of the initial CNN is set to 0.05. At 30 and 40 epochs, the learning rate of this method decreases by 0.1 percentage points respectively. All the work in this paper is completed based on the PyTorch development environment.

B) Datasets

To further validate the proposed method, this paper verifies it on the MS-COCO2014 image dataset. This dataset contains 80 unique categories, covering a wide range of complex daily environments, including 82,738 images for training and 40,504 images for testing. Each sample has an average of 2.9 labels, making it extremely challenging in terms of diversity. The uneven distribution of training samples between categories has an advantage in improving the accuracy of class imbalance object classification for the test network.

C) Experimental Results

This paper selects several algorithms for comparison with the attention-driven dual-branch dynamic graph convolutional network (TB-DGCN) proposed herein. The selected methods include the repeat discovery of attention regions (RDAR) algorithm, the weakly supervised learning-based multi-label classification (Multi-Evidence) algorithm, the deep residual learning for image recognition (ResNet-101) algorithm, the class independence and autocorrelation decoupling for multi-label classification (DecoupleNet) algorithm, and the global and local label relationship-based multi-label classification (ML-GLLR) method, along with the attention-driven graph convolutional network for multi-label recognition (ADD-GCN). Comparative experiments were conducted on the MS-COCO dataset, with evaluation metrics detailed in Table 1.

Table 1: Model experimental results in the dataset

Model	**mAP↑**	**CP↑**	**CR↑**	**CF1↑**	**OP↑**	**OR↑**	**OF1 ↑**
ME	80.3	80.5	68.1	72.2	82.1	71.1	77.9
Resnet 101	79.8	81.2	68.7	74.3	83.5	72.3	77.9
DecoupleN	80.5	83.6	69.4	75.9	83.6	72.8	78.7
ML-GLLR	81.1	84.1	69.5	76.1	84.1	73.9	79.8
ADD-GCN	81.3	84.3	70.3	76.5	84.3	74.1	80.0
Ours	82.9	85.2	71.5	77.3	85.6	74.3	80.1

According to the experimental results provided, we can observe the performance of different models on multiple evaluation indicators, including average precision (mAP), classification precision (CP), classification recall (CR), and classification F1 score (CF1) , over-prediction (OP), over-recall (OR) and over-F1 score (OF1). As can be seen from the table, all models have good performance in mAP, indicating that the overall target detection accuracy is high. Specifically, our model (Ours) achieved 82.9 mAP, compared to other models such as ME (80.3), Resnet101 (79.8), DecoupleN (80.5), ML-GLLR (81.1) and ADD-GCN (81.3), has significant advantages. In addition, on indicators such as classification precision (CP), classification recall (CR) and classification F1 score (CF1), our model also achieved results of 85.2, 71.5 and 77.3 respectively, which shows that the model is not only good at identifying targets The performance is excellent while also having good control in reducing false positives and false negatives. Although there is a slight increasing trend in the two negative indicators of overprediction (OP) and overrecall (OR), considering that their values still remain at a relatively low level, it can be considered that this impact is limited. To sum up, the experimental results show that our model is better than or at least not inferior to other existing models on a variety of performance indicators, especially on the key mAP indicator, showing that the model is effective in practical applications. potential.
In addition, in order to more intuitively illustrate the advantages of the model proposed in this chapter, the top-3 results of precision/recall/F1 score are also reproduced, as shown in Table 2 below.

Table 2: Results of precision/recall/F1 score

Top-3						
Model	**CP↑**	**CR↑**	**CF1↑**	**OP↑**	**OR↑**	**OF1 ↑**
RDAR	85.1	64.2	74.2	87.1	66.8	74.1
ME	85.2	64.7	74.9	87.5	67.1	74.2
Resnet101	86.6	65.4	75.0	88.4	67.6	75.7
ML-GLLR	87.3	65.8	75.1	89.1	67.9	75.8
ADD-GCN	88.7	66.0	75.9	89.3	68.1	76.0
Ours	89.6	66.3	76.3	90.1	68.8	77.1

According to the provided Top-3 experimental results, we can see the performance of different models in six indicators: classification precision (CP), classification recall (CR), classification F1 score (CF1), overprediction (OP), overrecall (OR), and overF1 score (OF1). The proposed model performs best in all these indicators, especially in CP (89.6), CR (66.3), and CF1 (76.3), which shows that the proposed model not only leads in correctly identifying positive samples, but also does the best in balancing precision and recall, thus obtaining the highest comprehensive score.
Compared with other models, the proposed model also has the highest scores in overprediction (OP) and overrecall (OR), which are 90.1 and 68.8 respectively, which actually means that the model may have more false positives or missed positives in some cases. However, considering that it also achieved the highest score in OF1 (77.1), it shows that despite some overprediction, the model can still effectively balance precision and recall as a whole, providing a good overall performance.
Other models such as Resnet101, ML-GLLR, and ADD-GCN also perform well in some indicators, but are slightly

inferior to our proposed model. For example, ADD-GCN performs the closest among all models, but is still slightly inferior in all key indicators. In contrast, our proposed model shows better overall performance, especially in tasks that require high accuracy, it can provide more reliable results. In summary, our proposed model shows excellent performance in Top-3 tasks, proving its effectiveness and robustness in dealing with complex classification problems.

V. CONCLUSION

This paper introduces a novel multi-label image classification approach, specifically a dual-branch dynamic graph convolutional network (DB-DGCN) enhanced by an attention mechanism. By integrating intrinsic image data with a sophisticated graph neural network (GNN) architecture, this method addresses the limitations found in traditional approaches. The DB-DGCN is designed to capture both local and global contextual information, thereby enriching the feature representation and enhancing the model's capacity to handle complex, multi-label scenarios. The integration of an attention mechanism further refines the model's focus, allowing it to selectively emphasize relevant features while suppressing noise, thus improving its overall robustness and adaptability.

Validation of the proposed model was conducted using the MS-COCO2014 imagery dataset, a collection renowned for its diversity and complexity. The experimental results demonstrate that the DB-DGCN either outperforms or matches existing techniques across a variety of evaluation metrics. Specifically, the model shows significant improvements in classification precision (CP), recall (CR), and the F1 score (CF1). Additionally, it performs well in terms of over-prediction (OP), over-recall (OR), and the over-F1 score (OF1). Though there might be cases where the model exhibits higher over-prediction or over-recall, it generally achieves a well-balanced equilibrium between precision and recall, leading to a commendable overall performance. Consequently, the DB-DGCN not only excels in tackling multifaceted classification challenges but also lays a solid foundation for future advancements in the field of multi-label image recognition.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

[1] R. Abdelfattah, Q. Guo, X. Li, *et al.*, "Cdul: Clip-driven unsupervised learning for multi-label image classification," *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, pp. 1348-1357, 2023. Available From: https://doi.org/10.48550/arXiv.2307.16634

[2] M. Li, D. Wang, X. Liu, *et al.*, "Patchct: Aligning patch set and label set with conditional transport for multi-label image classification," *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, pp. 15348-15358, 2023. Available From: https://doi.org/10.48550/arXiv.2307.09066

[3] J. Zhang, J. Ren, Q. Zhang, *et al.*, "Spatial context-aware object-attentional network for multi-label image classification," IEEE Trans. Image Process., vol. 32, pp. 3000-3012, 2023. Available From: https://ieeexplore.ieee.org/abstract/document/10122681

[4] B. Dang, D. Ma, S. Li, Z. Qi, and E. Zhu, "Deep learning-based snore sound analysis for the detection of night-time breathing disorders," *Appl. Comput. Eng.*, vol. 76, pp. 109-114, Jul. 2024. doi: 10.54254/2755-2721/76/20240574. Available From: https://www.researchgate.net/publication/382297062

[5] J. Yao, T. Wu, and X. Zhang, "Improving depth gradient continuity in transformers: A comparative study on monocular depth estimation with CNN," *arXiv preprint arXiv:2308.08333*, 2023. Available From: https://arxiv.org/abs/2308.08333

[6] J. Ouyang, Q. Lv, S. Zhang, *et al.*, "Energy transfer contrast network for unsupervised domain adaption," *Proc. Int. Conf. Multimedia Model.*, Cham: Springer Nature Switzerland, pp. 115-126, 2023. Available From: https://link.springer.com/chapter/10.1007/978-3-031-27818-1_10

[7] B. Dang, W. Zhao, Y. Li, D. Ma, Q. Yu, and E. Y. Zhu, "Real-time pill identification for the visually impaired using deep learning," *Proc. 2024 6th Int. Conf. Commun., Inf. Syst. Comput. Eng. (CISCE)*, pp. 552-555, 2024. doi: 10.1109/CISCE62493.2024.10653353. Available From: https://doi.org/10.48550/arXiv.2405.05983

[8] D. Lin, "Probability guided loss for long-tailed multi-label image classification," *Proc. AAAI Conf. Artif. Intell.*, vol. 37, no. 2, pp. 1577-1585, 2023. Available From: https://doi.org/10.1609/aaai.v37i2.25244

[9] X. Qu, H. Che, J. Huang, *et al.*, "Multi-layered semantic representation network for multi-label image classification," *Int. J. Mach. Learn. Cybern.*, vol. 14, no. 10, pp. 3427-3435, 2023. Available From: https://link.springer.com/article/10.1007/s13042-023-01841-6

[10] L. Wu, Y. Luo, B. Zhu, G. Liu, R. Wang, and Q. Yu, "Graph neural network framework for sentiment analysis using syntactic feature," *arXiv preprint arXiv:2409.14000*, 2024. Available From: https://doi.org/10.48550/arXiv.2409.14000

[11] W. Gu, M. Sun, B. Liu, K. Xu, and M. Sui, "Adaptive spatio-temporal aggregation for temporal dynamic graph-based fraud risk detection," *J. Comput. Technol. Softw.*, vol. 3, no. 5, 2024. Available From: https://ashpress.org/index.php/jcts/article/view/75

[12] B. Liu, J. Chen, R. Wang, J. Huang, Y. Luo, and J. Wei, "Optimizing news text classification with Bi-LSTM and attention mechanism for efficient data processing," *Proc. 2024 IEEE Int. Conf. Big Data Smart Comput.*, 2024. Available From: https://doi.org/10.48550/arXiv.2409.15576

[13] W. He, R. Bao, Y. Cang, J. Wei, Y. Zhang, and J. Hu, "Axial attention transformer networks: A new frontier in breast cancer detection," *arXiv preprint arXiv:2409.12347*, 2024. Available From: https://doi.org/10.48550/arXiv.2409.12347

[14] Y. Liang, Y. Zhang, Z. Ye, and Z. Chen, "Enhanced unsupervised image registration via dense U-Net and channel attention," *J. Comput. Sci. Softw. Appl.*, vol. 4, no. 5, pp. 8-15, 2024. https://doi.org/10.5281/zenodo.13643091

[15] Z. Zhang, J. Chen, W. Shi, L. Yi, C. Wang, and Q. Yu, "Contrastive learning for knowledge-based question generation in large language models," *arXiv preprint arXiv:2409.13994*, 2024. Available From: https://doi.org/10.48550/arXiv.2409.13994

[16] X. Yan, Y. Jiang, W. Liu, D. Yi, H. Sang, and J. Wei, "Data-driven spatiotemporal feature representation and mining in multidimensional time series," *arXiv preprint arXiv:2409.14327*, 2024. Available From: https://doi.org/10.48550/arXiv.2409.14327

[17] Y. Liang, Y. Zhang, Z. Ye, and Z. Chen, "Enhanced unsupervised image registration via dense U-Net and channel attention," *J. Comput. Sci. Softw. Appl.*, vol. 4, no. 5, pp. 8-15, 2024. Available From:

https://doi.org/10.5281/zenodo.13643091

[18] X. Yan, W. Wang, M. Xiao, Y. Li, and M. Gao, "Survival prediction across diverse cancer types using neural networks," *Proc. 2024 7th Int. Conf. Mach. Vis. Appl.*, pp. 134-138, 2024. Available From: https://doi.org/10.48550/arXiv.2404.08713

[19] Z. Zheng, Y. Cang, W. Yang, Q. Tian, and D. Sun, "Named entity recognition: A comparative study of advanced pre-trained model," *J. Comput. Technol. Softw.*, vol. 3, no. 5, 2024. Available From: https://doi.org/10.5281/zenodo.13624035

[20] I. P. Singh, E. Ghorbel, O. Oyedotun, *et al.*, "Multi-label image classification using adaptive graph convolutional networks: From a single domain to multiple domains," *Comput. Vis. Image Underst.*, vol. 247, p. 104062, 2024. Available From: https://doi.org/10.1016/j.cviu.2024.104062

[21] Y. Ma, D. Sun, E. Gao, N. Sang, I. Li, and G. Huang, "Enhancing deep learning with optimized gradient descent: Bridging numerical methods and neural network training," *arXiv preprint arXiv:2409.04707*, 2024. Available From: https://doi.org/10.48550/arXiv.2409.04707