



**HAL**  
open science

# Knowledge Learning Without Forgetting for the Detection of Alzheimer's Disease

Ruotong Liu, Yue Yin, Jing Bai, Xu Wang

► **To cite this version:**

Ruotong Liu, Yue Yin, Jing Bai, Xu Wang. Knowledge Learning Without Forgetting for the Detection of Alzheimer's Disease. 5th International Conference on Intelligence Science (ICIS), Oct 2022, Xi'an, China. pp.441-449, 10.1007/978-3-031-14903-0\_47 . hal-04666412

**HAL Id: hal-04666412**

**<https://hal.science/hal-04666412v1>**

Submitted on 1 Aug 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

# Knowledge Learning without Forgetting for the Detection of Alzheimer’s Disease<sup>\*</sup>

Ruotong Liu<sup>1</sup>, Yue Yin<sup>2</sup>, Jing Bai<sup>1</sup>(✉)<sup>[0000–0001–5412–7793]</sup>, and Xu Wang<sup>1</sup>

<sup>1</sup> Xidian University, Xi’an 710071, China

baijing@mail.xidian.edu.cn(J. Bai), rtlou1004@stu.xidian.edu.cn(R. Liu),  
xuwangxd@stu.xidian.edu.cn(X. Wang)

<sup>2</sup> The First Affiliated Hospital of Air Force Medical University, Xi’an 710071, China  
yiny110@sina.com

**Abstract.** Alzheimer’s disease (AD) is an extremely damaging, slow-progressing neurological disease that causes tremendous inconvenience to patients’ lives. Numerous medical professionals have researched that timely diagnosis and early therapy of AD when it is in its early stages could slow down the progression of AD and even be cured. Therefore, early diagnosis of AD is in urgent need of significant advancement. Nevertheless, there are problems such as brain images with many similar features that are difficult to extract and classify, and insufficient data for training. In this paper, a transfer learning-based knowledge learning without forgetting method we proposed for AD detection, which can preserve the learned knowledge during the transfer process so that it will not be excessively forgotten and this method achieve promising outcomes. The classification accuracy of our method based on resnet50 and resnet18 on the ADNI dataset reached 96.15% and 96.39%, compared to training directly on resnet50 and resnet18, our method increased the classification accuracy by 2.16% and 3.61%, which achieved remarkable results and contributes greatly to the development of AD detection.

**Keywords:** Alzheimer’s disease · Transfer learning · Convolutional neural network (CNN) · MRI · ADNI.

## 1 Introduction

Alzheimer’s disease (AD) is a slowly developing, irreversible neurological disease with unknown causes. It affects mainly the aged population. The symptoms of AD usually include memory impairment, executive dysfunction, and impaired cognitive behavior, etc. This disease dramatically influences the patient’s normal life. A large number of medical researchers have found that there are three primary categories of the early AD: late mild cognitive impairment (LMCI), early

---

<sup>\*</sup> This work was supported in part by the Key Research and Development Program of Shaanxi under Grant 2022GY-062, in part by the National Natural Science Foundation of China under Grant 61772401, and in part by the Science and Technology on Communication Information Security Control Laboratory. (Ruotong Liu and Yue Yin contributed equally to this work.)(Corresponding author: Jing Bai.)

mild cognitive impairment (EMCI) and subject memory concern (SMC) [1]. Therefore, how to efficiently and accurately detect AD disease has become a crucial challenge to be overcome. Currently, numerous scientists are trying to explore new features that may emerge in brain regions during the early stages of AD.

With the continuous development of artificial intelligence, deep learning has made significant achievements in the field of computer vision [2]. Accordingly, in recent years, an increasing number of researchers have attempted to use deep learning with its powerful feature extraction ability on the detection of AD [3–7].

However, the acquisition and collation of medical images is labor-intensive and time-consuming, therefore, inadequate data for training for training has become a problem which should be seriously concerned. On the basis of this problem, in particular, transfer learning [8,9] have been widely utilized for AD detection. In further, [10,11] made enhancements to the fine-tuning department and achieved good results in image classification.

In this paper, we propose a knowledge learning without forgetting method that keeps as much of the knowledge learned in the pre-trained network as possible and ensures that knowledge would not inordinately be dissipated during the transfer learning. We are inspired by [10,11] and chiefly use transfer learning based on its promising performance in AD detection and cleverly integrate multiple losses which includes distillation loss, cross-entropy loss, and improved contrast loss used exclusively for self-supervision. We select Resnet as the pre-training network, use the Alzheimer’s Dataset as the source domain, the classification results are obtained by training on the source domain with the pretrained network and then importing our target domain ADNI to the network with fine-tuned weights. We obtained 96.39% classification accuracy on ADNI for AD, MCI and NC. The improved transfer learning we proposed could perform well in the AD detection and greatly contribute to the early detection of AD.

## 2 Method

### 2.1 Transfer learning for AD detection

Our improved transfer learning method mainly contains two stages: In the first stage we take the Alzheimer’s Dataset as the source domain and use Resnet as our backbone for pre-training. Second stage, we initialize the weights and reset the fc layer, then fine-tune the weights and apply them to the target domain ADNI for training to effectively classify the categories AD, MCI, and NC on MRI images. Two stages of transfer learning is shown in Figure 1.

On the classification mission of the target domain ADNI dataset, we mainly categorize into one primary task and two auxiliary tasks. We input the target domain ADNI data into the network that has been pre-trained in the source domain, and the principal task is to fine-tune the pre-trained network to get the normal three-class output, doing cross-entropy loss and contrast cross-entropy loss with the normal labels. The four-class output obtained by the primary task

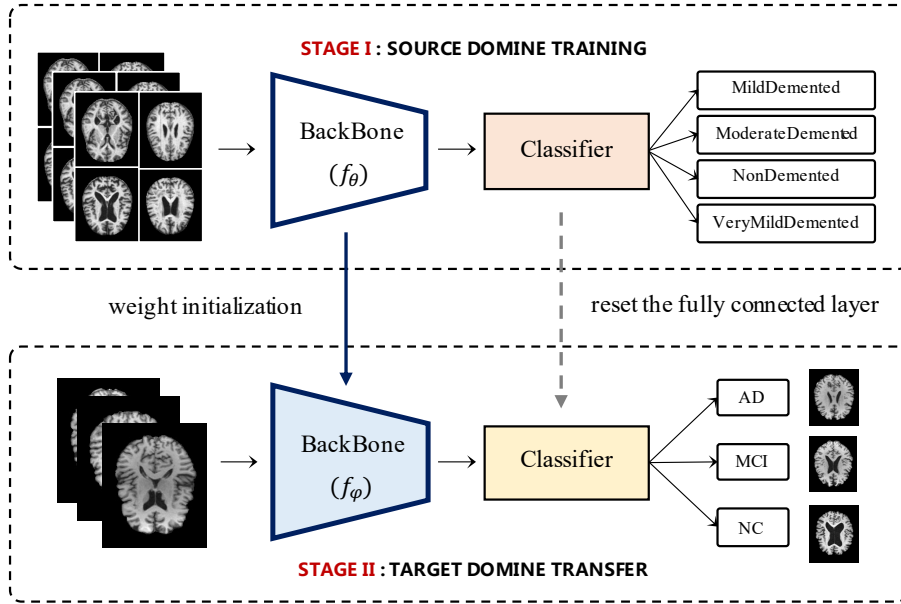


Fig. 1: Two stages of transfer learning for AD detection

then does distillation loss with the four-class source domain output of the auxiliary task 1. Additionally, another branch of the main task takes the classification results into a projection head and processes them for output, doing improved classification contrast loss with the projection head output in auxiliary task 2. The overview of knowledge learning without forgetting is shown in Figure 2.

Based on this transfer learning architecture, we skillfully combine distillation loss, cross-entropy loss, and improved contrast loss to achieve preservation of learned knowledge from over forgetting, realize knowledge learning without forgetting.

## 2.2 Knowledge learning without forgetting

In the auxiliary task 1, the distillation loss is used to ensure that the response of the new data ADNI on the new model retains as much of the response of the old model as possible. First, the target domain ADNI is fed into the pre-trained network, to extract its response at the last layer of the old model, which is the confidence  $y'$  of the new class of data over the old class. Afterwards, the new model is trained with transfer learning, which sets the classifier with the correct number of categories for the new class while introducing a new branch to generate the response with the new model on the classification by the old class. The distance class between  $y'$  and  $y''$  is measured with relative entropy,

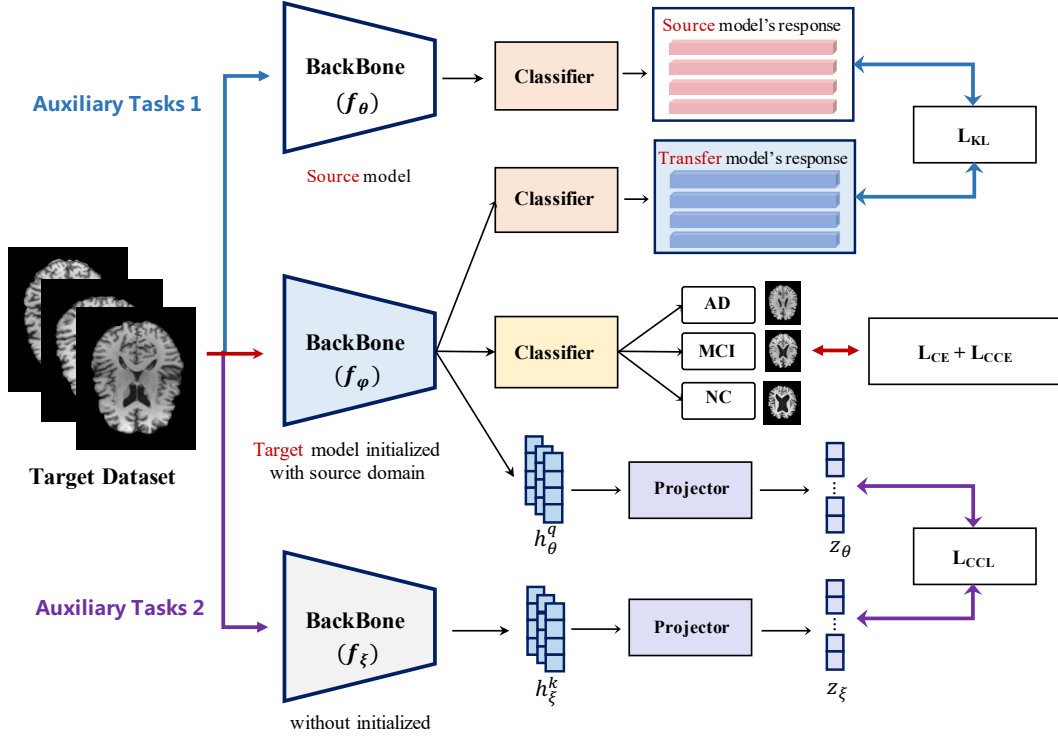


Fig. 2: Knowledge learning without forgetting method overview

also known as KL loss, the formula is as follows:

$$L_{KL} = - \sum_{i=1}^k y^{(i)} \log y''(j) \quad (1)$$

KL loss ensures that the knowledge already learned will not be easily lost after fine-tuning. It keeps the majority of information learned from the old model.

### 2.3 Transfer based on improved contrastive loss

Based on the utilization of the source domain network, we consider the optimal exploitation of the intrinsic structure of the target domain data, ADNI. Inspired by self-supervised learning that uses unlabeled samples for pre-training, we construct the internal structure between the target domain ADNI data by adopting a contrast learning approach. The loss of contrastive learning  $L_{self}$  is expressed as:

$$L_{self} = -\log \frac{\exp(q \cdot s_+ / \tau)}{\sum_{i=0}^K \exp(q \cdot s_i / \tau)} \quad (2)$$

$q$  means the query sample, the positive key  $s_+$  is the data-augmented positive sample, and  $s_i$  is the negative sample, and the contrast loss is to obtain the maximum value of similarity between the query  $q$  and the corresponding positive key  $s_+$ .

Based on this contrast learning approach, the improved contrastive loss are utilized in the main task and auxiliary task 2 separately. It contains contrastive cross-entropy loss  $L_{CCE}$  and improved classification contrast loss  $L_{CCL}$ .

In the main task, the target domain ADNI is entered into the fine-tuned network to obtain the classification results, which are applied to the corresponding labels with normal cross-entropy loss  $L_{CE}$  and contrast cross-entropy loss  $L_{CCE}$ .

We search for the most similar prediction to the ground truth of each instance by using the contrastive cross-entropy loss  $L_{CCE}$ . It permits better utilization of instance comparisons to exploit label information. It allows better use of the formula is as shown below:

$$L_{CCE} = - \sum_{i=1}^N \sum_{k=0}^K \Pi(h_k^k \in S_i) \log \frac{\exp(c_{y_i} \cdot h_i^q / \tau)}{\sum_{j=0}^K \exp(c_{y_i} \cdot h_j^k / \tau)} \quad (3)$$

$K$  is the candidate keys' size,  $h_k^k$  denotes samples generated by the key generating mechanism from the hidden key pool,  $S_i$  denotes the set of positive keys for example  $i$ ,  $h_i^q$  denotes the features extracted by the neural network,  $c_{y_i}$  denotes the prediction result of the classifier.

In the auxiliary task 2, the strategy we used in consonance with the self-supervised approach is to export the target domain ADNI to the old model for extracting features, and input the extracted features to a projector, this projector maps the pre-trained representation  $h_i^q$  to the embedding space  $z_i^q$ . However, the implicit requirement of contrastive loss that each instance belongs to a separate class is not suitable for supervised classification tasks. Therefore, we use an improved classification contrast loss  $L_{CCL}$ , which takes other samples from the identical category to use as positive keys for query matching.

Like the format of  $L_{self}$ , the improved classification contrast loss  $L_{CCL}$  formula is as follows:

$$L_{CCL} = - \sum_{i=1}^N \sum_{k=0}^K \Pi(z_k^k \in S_i) \log \frac{\exp(z_i^q \cdot z_k^k / \tau)}{\sum_{j=0}^K \exp(z_i^q \cdot z_j^k / \tau)} \quad (4)$$

The improved contrast learning loss allows for the full utilization of the inner structure of the learned knowledge in a category-consistent manner.

Consequently, our knowledge learning without forgetting method based on transfer learning can be expressed as:

$$L = L_{CE} + \alpha L_{KL} + \beta (L_{CCE} + L_{CCL}) \quad (5)$$

where  $\alpha, \beta$  are hyperparameters, the value range of  $\alpha, \beta$  are set to 0-1.

### 3 EXPERIMENTS

#### 3.1 Datasets

**ADNI** This dataset is a publicly available dataset used specifically for the AD detection experiments. It mainly consists of three categories: AD, MCI and NC. In our experiments, the ADNI dataset is utilized as the target domain in transfer learning. The three classes of AD, MCI, and NC images in ADNI are shown in Figure 3:

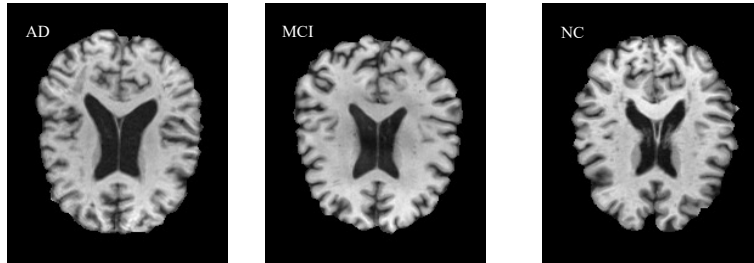


Fig. 3: The three classes of AD, MCI, and NC image in ADNI

**Alzheimer’s Dataset** This dataset is used exclusively in the Kaggle competition. It contains four classes of AD images: mild demented, moderate demented, non demented, very mild demented. In our experiments, this dataset is used as the source domain in transfer learning. The four classes images in Alzheimer’s Dataset are shown in Figure 4:

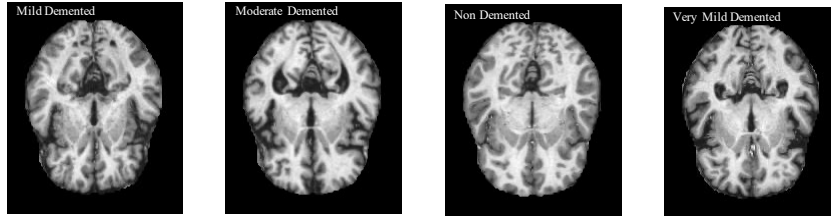


Fig. 4: The four classes images in Alzheimer’s Dataset

In our experiment, we first process 3D MRI images into 2D slices before sending the data to the network for training. This pre-processed dataset procedure to some extent could solve the problem of insufficient medical images for training. Then we pre-train the weights in Resnet by using Alzheimer’s Dataset, and after fine-tuning the weights, the ADNI is fed into the pre-training network for training to obtain the classification results.



### 3.2 Experimental setting

In our experiments, we use the optimizer is SGD, the learning rate is set to 0.001, and we set the batch size to 32. The operating environment for our experiments is mainly as follows: Ubuntu 18.04, GPU 3090, CUDA 11.0, torch 1.7.1+, torchvision 0.8.2+, numpy 1.21.6, and python 3.8.

### 3.3 Experimental results

We apply Alzheimer’s Dataset as source domain for training, the ratio of the divided training set to the test set is 4:1. And we use ADNI as target domain for training, the ratio of the divided training set to the test set is 4:1, either. In the experiment, we found that by setting the learning rate to 0.001, the loss started to decay to 0.8 times of the original per 10 epochs. The classification accuracy on ADNI is shown in Table 1.

Table 1: Accuracy of classification on ADNI

Backbone	Method	25%	50%	75%	100%
Resnet50	train	65.69%	84.89%	91.71%	93.99%
	train_tsfr	67.37%	86.29%	93.29%	94.38%
	train_lwf	71.08%	87.39%	94.12%	95.08%
	train_con	72.53%	87.63%	94.71%	95.45%
	<b>ours</b>	<b>72.61%</b>	<b>87.76%</b>	<b>94.36%</b>	<b>96.15%</b>
Resnet18	train	65.23%	82.12%	90.75%	92.78%
	train_tsfr	66.91%	84.11%	91.47%	94.71%
	train_lwf	64.99%	85.48%	91.84%	94.95%
	train_con	66.30%	86.58%	92.46%	95.08%
	<b>ours</b>	<b>67.72%</b>	<b>87.06%</b>	<b>92.56%</b>	<b>96.39%</b>

In Table 1, 100%, 75%, 50%, 25% indicate the percentage of the data excluding the test data, train means direct training, train\_tsfr means using transfer learning, train\_lwf means using transfer learning with unforgetting loss, train\_con means using transfer learning with improved contrast loss, ours means our method.

Our method enables to achieve 96.39% accuracy of classification on ADNI. In particular, the accuracy can be enhanced by 7% when only 25% of the data is used. The accuracy of classification results indicate that transfer training can lead to a significant improvement over direct training. And both transfer learning optimization strategies can bring increases compared to direct transfer learning.

In addition, the best accuracy results demonstrate the advanced and efficient performance of the better method we proposed to transfer learning, and achieve a great integration of the two mentioned strategies.

## 4 CONCLUSION

In this paper, a transfer learning-based knowledge learning without forgetting method we proposed for the detection of AD, which could ensure that the knowledge acquired during pre-training is well maintained and does not be excessively dissipated during the transfer process. Our substantial experimental results demonstrate that the method we proposed can work effectively in the detection of AD, with an accuracy of 96.15% and 96.39% in the classification of AD, MCI and NC on the ADNI dataset based on resnet50 and resnet18, compared to training directly on resnet50 and resnet18, our method increased the classification accuracy by 2.16% and 3.61%, and it enables to solve the problem of insufficient training data for AD detection.

In forthcoming studies, we intend to pursue our attempts to explore innovative methods on self-supervised learning to achieve greater performance that could be valuable for AD detection.

## References

1. Kruthika, K., Maheshappa, H., Initiative, A.D.N., et al.: Multistage classifier-based approach for alzheimer’s disease prediction and retrieval. *Informatics in Medicine Unlocked* **14**, 34–42 (2019)
2. Yang, K., Mohammed, E.A.: A review of artificial intelligence technologies for early prediction of alzheimer’s disease. *arXiv preprint arXiv:2101.01781* (2020)
3. Sarraf, S., Tofighi, G.: Classification of alzheimer’s disease using fmri data and deep learning convolutional neural networks. *arXiv preprint arXiv:1603.08631* (2016)
4. Amin-Naji, M., Mahdavinataj, H., Aghagolzadeh, A.: Alzheimer’s disease diagnosis from structural mri using siamese convolutional neural network. In: 2019 4th International Conference on Pattern Recognition and Image Analysis (IPRIA). pp. 75–79. *IEEE* (2019)
5. Fulton, L.V., Dolezel, D., Harrop, J., Yan, Y., Fulton, C.P.: Classification of alzheimer’s disease with and without imagery using gradient boosted machines and resnet-50. *Brain sciences* **9**(9), 212 (2019)
6. Liu, S., Yadav, C., Fernandez-Granda, C., Razavian, N.: On the design of convolutional neural networks for automatic detection of alzheimer’s disease. In: *Machine Learning for Health Workshop*. pp. 184–201. *PMLR* (2020)
7. Fedorov, A., Wu, L., Sylvain, T., Luck, M., DeRamus, T.P., Bleklov, D., Plis, S.M., Calhoun, V.D.: On self-supervised multimodal representation learning: An application to alzheimer’s disease. In: 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI). pp. 1548–1552. *IEEE* (2021)
8. Ebrahimi, A., Luo, S., Chiong, R.: Introducing transfer learning to 3d resnet-18 for alzheimer’s disease detection on mri images. In: 2020 35th international conference on image and vision computing New Zealand (IVCNZ). pp. 1–6. *IEEE* (2020)
9. Shanmugam, J.V., Duraisamy, B., Simon, B.C., Bhaskaran, P.: Alzheimer’s disease classification using pre-trained deep networks. *Biomedical Signal Processing and Control* **71**, 103217 (2022)
10. Li, Z., Hoiem, D.: Learning without forgetting. *IEEE transactions on pattern analysis and machine intelligence* **40**(12), 2935–2947 (2017)
11. Zhong, J., Wang, X., Kou, Z., Wang, J., Long, M.: Bi-tuning of pre-trained representations. *arXiv preprint arXiv:2011.06182* (2020)