

Transfer Learning in Brain Tumor Detection: from AlexNet to Hyb-DCNN-ResNet

Zhuodiao Kuang *

School of Statistics, Renmin University of China, Beijing, China

* Corresponding Author Email: 2019202274@ruc.edu.cn

Abstract. Detecting abnormalities in the human body with magnetic resonance imaging has long been a challenge in medical computer-aided diagnosis (CAD). This paper presents a comprehensive review of research focusing on transfer learning (TL) in brain tumor detection. Each work starts from collecting MR images and substantial strategies are applied when preprocessing data including data augmentation and image segmentation. Multiple pre-trained models from AlexNet to Hyb-DCNN-ResNet in the latest work are focused. And the results of binary and multiple class classification are compared chronologically. Three pre-trained models which are frequently used to attain a good performance in brain tumor detection are illustrated in detail. And these pre-trained models, GoogLeNet, VGG and ResNet, all are capable to help the proposed systems reach the accuracy of 99%. The challenges even after transferring apposite knowledge to the target domain still exist in pluralistic forms. But the essence of transfer learning can support interdisciplinary research to get better performance.

Keywords: transfer learning; brain tumor detection; convolutional neural network; classification

1. Introduction

A brain tumor is an abnormal mass of tissue growing in cranial cavity. And the symptoms of a brain tumor vary widely depending on the location, size, and rate of growth of the tumor. Headaches with a new onset or a change in pattern are common complaints. And brain tumors are one of the most fatal cancers, despite their rarity. Magnetic Resonance Imaging (MRI) is frequently utilized in modern technology to provide significant information about the shape, size, location, and metabolism of brain tumors, assisting in diagnosis. T1-weighted MRI (T1), T2-weighted MRI (T2), T1-weighted MRI with gadolinium contrast enhancement (T1-Gd), and Fluid Attenuated Inversion Recovery (FAIR) are the four main MRI modalities used for glioma diagnosis (FLAIR). BRATS Challenges' datasets are often used for visual categorization tasks and these datasets always contain four types of MR images mentioned above. Recent work on computer-aided medical diagnosis has improved because of the introduction of deep learning ideas. To track brain metastases, Charron et al. [1] used a deep convolutional neural network (CNN). In the latest research, Kumar, et al. [2] applied a new technique called Hyb-DCNN-ResNet 152 TL on binary classification of tumors. Various datasets are used to test different pre-trained models' ability in both binary and multi-class classification tasks

Transfer learning (TL) is applied to build up a more complicated and accurate convolutional neural networks to cope with problems like data scarcity or computational inefficiency, which is the subfield of deep learning (DL). And in 2009, an early survey [3] discussed about the specific applications of TL and its definition that the training data in another domain of interest, which may be in a different feature space or follow a different data distribution, is used when researchers sometimes have a classification task. More recently, deep transfer learning, has been widely applied on visual categorization, object recognition and image classification problems [4]. And due to data privacy or naturally lack of data, transfer learning has been widely employed in medical field to help improve the performance of learning.

And this paper provides a comprehensive review of the application of TL in brain tumor detection. Several relevant papers focusing on one model or multiple models' comparison with TL are presented. This article referred to some reviews about the application of TL and other tumor detection task, but mainly focuses on evaluate different models classifying MR images. And some pioneering works and

proposed methods which cited the former works are presented in this review. The structure of the present work is as follows: the TL terminology, definition, solution categorization, and analysis are initially given in Section III. Section IV presents data preprocessing in different situations before building up an integrate system. Sections V-VII elaborate three noted pre-trained models and the last one is a combined type which is called ResNeXt. Finally, in Section VIII, the advantages and the limitation of using TL to identify brain tumors are discussed, as well as the domain's future prospects. And tumor detection tasks are not always a binary classification problem, which is also explained in detail in Section III.

2. Methodology

The searching process includes three main parts. Starting from the development history of transferring learning, Scopus database provides relevant research on CAD with transfer learning. To narrow down the list, results with brain tumor detection are considered. Besides, papers which cited pioneering works and generated good performance or new methods are put in the review list. The models with their accuracies are recorded and this part can always be found in the abstract or the introduction.

3. An Overview on Transfer learning

3.1. Transfer Learning Terminology and Definition

Deep learning's success is also strongly linked to vast amounts of data, which means that the paucity of training data can severely limit the performance of deep learning models. As a result, transfer learning is used to solve this problem at a low cost. The learning process of transfer learning is shown in Fig 1. It is necessary to review the basics of transfer learning in this part before moving on to the use of transfer learning in medical image analysis.

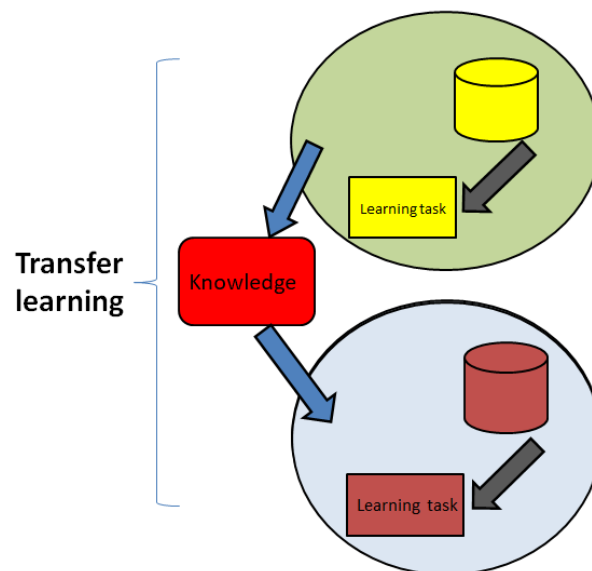


Fig 1. Process of transfer learning.

A domain D , according to Weiss et al. [5], is made up of two parts: a feature space X and a marginal probability distribution $P(X)$, with $X = x_1, x_2, \dots, x_{n-1}, x_n$, and n is the number of feature vectors in X . A task T is made up of two pieces for a given domain D : a label space Y and a prediction function $f(\cdot)$. From pairs of feature vectors and labels x_i, y_i , where $x_i \in X$ and $y_i \in Y$, the prediction function $f(\cdot)$ is trained. As a result, a domain $D = \{X, P(X)\}$ and a task $T = \{Y, f(\cdot)\}$ are required.

As a result, the target domain $D_t = \{X_t, P(X_t)\}$ with a matching source task $T_t(\hat{A}) = \{Y, f_t(\hat{A} \cdot)\}$ can be denoted. Similarly, the source domain $D_s = \{X_s, P(X_s)\}$ and the accompanying source task $T_s(\hat{A}) = \{Y, f_s(\hat{A} \cdot)\}$ can be denoted as above. The process of enhancing the target prediction function $f_t(\cdot)$ based on D_s and T_s can therefore be referred to as transfer learning. $D_s \neq D_t$ or $T_s \neq T_t$ should be noted. Heterogeneous and homogeneous transfer learning are two types of transfer learning. When $X_s \neq X_t$, it's called heterogeneous transfer learning, but when $X_s = X_t$, it's called homogeneous transfer learning.

3.2. Solution Categorization of Transfer Learning

Because the types of transfer learning can be categorized from multiple aspects. And based on the medical problem focused, the solution categorization will be discussed only. Instance-based transfer learning, feature-based transfer learning, parameter-based transfer learning, and relation-based transfer learning are the four general types of information transfer learning. These four transfer learning methods are summarized in Table I.

Table 1. Four categories of transfer learning

Transfer Method	Key Points	Advantages
Relational-based	Transfer relationship or rules learned in the source domain	Preferable for data with dependency and identical distribution.
Feature-based	Asymmetric or Symmetric methods	The performance is based on labeled data, and the reduced disparity between the source domain and the target domain.
Parameter-based	For target domain models, parameters of the source domain models are modified.	Highly improved computational efficiency
Instance-based	Instance weighting strategy	When the source domain feature data cannot be reused, this option is appropriate.

To create instance-based transfer learning systems, an instance weighting approach is used. Feature-based techniques transform the original features to generate new feature representations, which are further divided into asymmetric and symmetric feature-based methods. Asymmetric techniques are used to modify the source features so that the altered source features match the destination features. On the other hand, symmetric techniques look for shared feature spaces into which both source and destination attributes may be mapped.

In parameter-based techniques, knowledge is transferred at the parameter level, where parameters in source domain models have been altered to fit the destination domain. The logical relationship or rules learnt in the source domain are transferred in relation-based transfer learning methods. In a tumor identification job with TL, both feature-based and parameter-based methods are always examined. Similar tasks in the source domain are strongly advised to avoid unfavorable transfer and provide a pleasing result. It is difficult to pursue the dependence and identical distribution in the target domain, especially when various tumors are similar to one other; the efficacy of telling the difference was not satisfying when utilizing conventional approaches.

4. Data Preprocessing

The datasets of MRI used in researchers' work are mainly from 3 different sources. Related work shows that BRATS Challenges' datasets are discussed a lot, not only for classification [6], but also some grading [7] and segmentation work [8]. Besides, the CE-MRI dataset is frequently used from 2019. Vital research of classifying multiple types of brain tumors [9] started the application of

different pre-trained models on multiple-class classification of brain tumors. Other prevailing datasets are either from noted online repositories [10] [11] or private images [12].

Brain MRI images in the target domain always don't match the size of the images designed for the pre-trained network. But before resizing the images, what researchers frequently choose to do is normalizing images in intensity values. A min-max normalization technique is followed to scale the intensity values between 0 and 1. And in some relevant work, model designers usually balance the datasets with unique sampling strategies. New images may be created and this process, which is called data augmentation, will help improve the performance of models, especially when data is insufficient. Image segmentation is not common in brain tumor detection area, because the images are greyscale with MRI and the brain part is conspicuous with a black background. The MRI images in the dataset are pre-processed in the following manner in the most of relevant works (Fig. 2). To cope with imbalanced data problem, some sampling strategies are also applied [13]. And for different classification tasks, datasets containing different manually labeled information will be divided into fixed files or sections. The specific number of sections depends on the type of work.

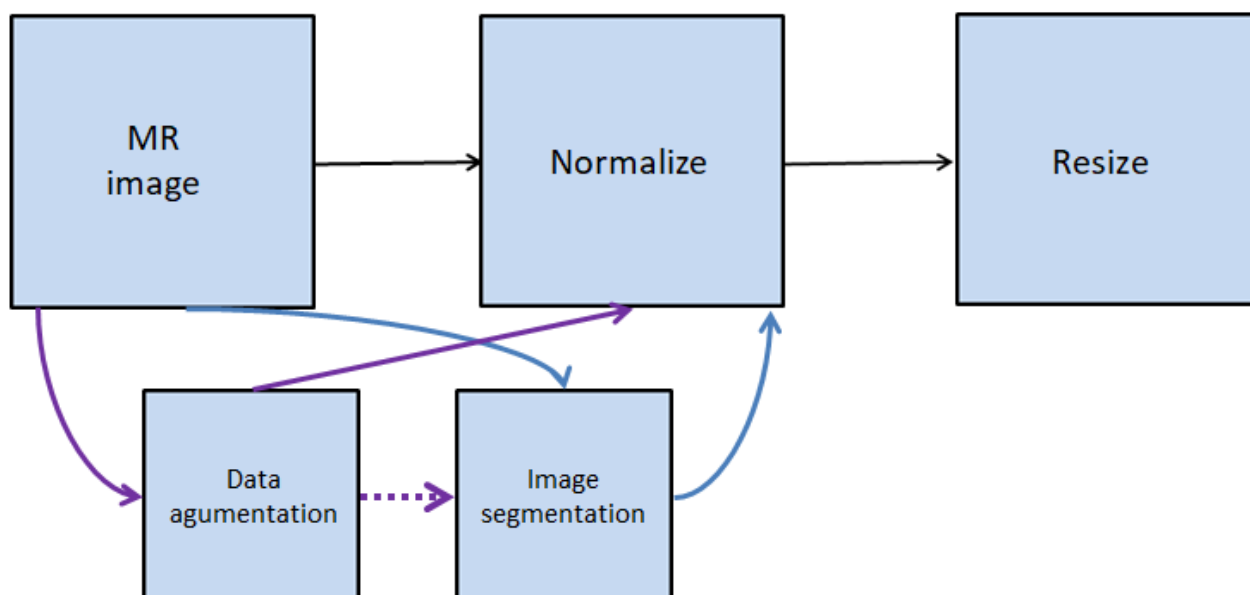


Fig 2. Data preprocessing steps.

After data preprocessing work, a pre-trained model should be considered. And when transferring knowledge from state-of-the-art models to a common CNN system, researchers always redefine train and validation generator since base model expects pixel values in a different range. And in next three parts, pre-trained models will be discussed in detail.

5. GoogLeNet

Inception, which is also called GoogLeNet, is a brand-new deep learning structure proposed by Szegedy et al. [14] in 2014. Scholars call it GoogLeNet rather than GoogleNet because this name is in the memory of the early visual recognition model called LeNet. And before GoogLeNet, some work has been done, which are listed in Table II.

Table 2. Five early models used for transfer learning

Name	Year	Features	Relevant work
AlexNet	2012	First deep CNN for the image classification task	[15, 16]
NIN	2013	Proposition of GAP; Introduction of micro-networks into CNN	[17]
VGGNet	2014	Deeper architecture	[18-25]
GoogLeNet	2014	Expands CNN's width	[9, 14, 12, 22, 23, 26, 27]
Inception-v3	2016	Improved Inception block	[28]

AlexNet, which is the first deep CNN for the image classification task, gains better performance with more layers. And VGGNet is also designed with deeper architecture to get good training performance. But they both face with the problems like overfitting and vanishing or exploding gradients.

5.1. Structure

Inception improves the performance from another perspective by using computational resources wisely, thus, extracts more features with equal computational complexity and improves the results.

Because adding more layers with the same receptive field, deep learning structure can extract more features. (NIN) And Fig. 3 displays how GoogLeNet functions in a unique way. Focusing on yellow 1×1 convolution modules, the author explains how it can extract more features and decrease the dimension, thus, reducing the computational complexity.

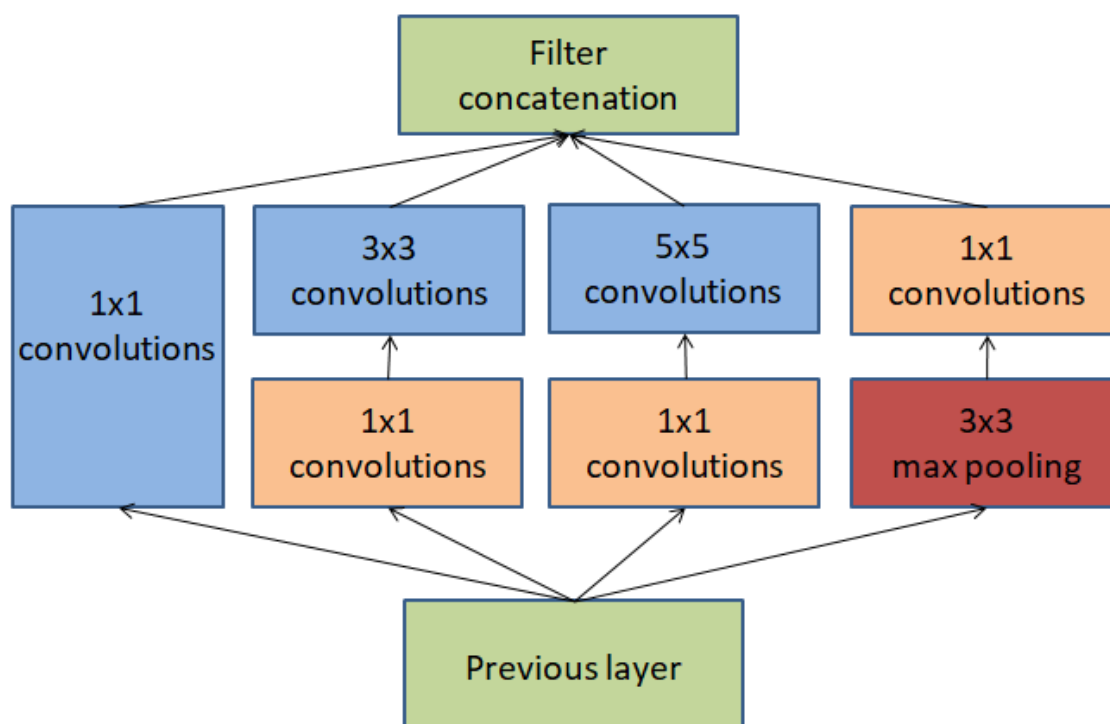


Fig 3. Inception's module.

InceptionV2 was invented with reduced internal covariate shift in 2015. [29] In the next year, Szegedy et al. [28] improved the GoogLeNet's computation efficiency, and for distinguishing the

difference, it is generally called Inception-v3. The Inception is continuously improved and combined with other models to generate better performance in certain tasks.

5.2. Relevant Work and Evaluation

Deepak, et al. [9] introduced the idea of deep transfer learning to the categorization of three types of brain tumors. And, for the first time, GoogLeNet was used as a pre-trained model to classify multiple types of brain tumors. And to adapt to the target domain, the last three layers of GooLeNet was modified and the fully connected layer was removed. Later in 2019, Amin, et al. [26] assess Alex and Google networks with top medical image computing and computer-assisted intervention (MICCAI) challenge datasets and finally get a mean accuracy over 85%. Related work is listed as Table III.

Table 3. Transfer learning with GoogLeNet

Reference	Dataset	Size of Dataset	Performance	Year
Amin, et al. [26]	BRATS Challenge	520,180	Mean accuracy by GoogLeNet: over 85%.	2019
Deepak, et al. [9]	CE-MRI dataset	3064 abnormal brain CE-MRI	GoogLeNet:97.1% for average classification accuracy	2019
Rehman, et al. [22]	CE-MRI dataset	3064 abnormal brain CE-MRI	Fine-tune VGG16: accuracy of 98.69%	2020
Chelghoum, et al. [23]	CE-MRI dataset	3064 abnormal brain CE-MRI	98.71% by using VGG-16 with 90 epochs	2020
KULKARNI, et al. [12]	Private dataset	100 benign and 100 malignant	Fine-tuned AlexNet: 93.7%(precision), 100%(recall), 96.77%(f-measure)	2021
Anjum, et al. [27]	CE-MRI dataset	3064 abnormal brain CE-MRI	Two-class classification: GoogLeNet yielded 99.33% TA; Multiple-class:98.9% TA with ResNet101	2022

CE-MRI dataset from figShare was mentioned again by Rehman, et al. [22] in 2020. And they compared the performance when applying AlexNet, GoogLeNet and VGG16 as pre-trained model. And Chelghoum, et al. [23] presents a fully automatic system for the classification task. And both get a pretty good training result with accuracy of 98.70% by using VGG. But Rehman, et al. [22] focused on the comparison of different transfer learning techniques, i.e., fine-tune and freeze. Based on training duration and epoch numbers, Chelghoum et al. [23] present the overall classification accuracy of the nine pre-trained designs. KULKARNI, et al. [12] used private dataset to valid the performance of state-of-the-art transfer learning methods on the binary task, which showed that all the fine-tuned models performed well under any circumstances.

In more recent work, Anjum, et al. [27] tested the accuracy of binary classification of brain tumor with GoogLeNet, which yielded a fairly good result of 99.33% total accuracy (TA). This work has been tried in 2019 with VGG-19 [21] and visual geometry group will be illustrated in the following part of this article.

6. VGG

6.1. Structure

VGG was invented by Visual Geometry Group [18] in University of Oxford, which was cooperating with Google DeepMind and ranked second in ILSVRC-2014. VGG and GoogLeNet are both powerful in imaging visual recognition work and they get different features for their good performances.

And in some research of brain tumor detection with transfer learning, VGG sometimes got better results than GoogLeNet because of its regular design and stackable convolutional blocks. Due to its 13 convolutional layers and 3 fully connected layers, VGG is widely called as VGG-16 and gets a deeper architecture than AlexNet and more parameters. VGGNet has 6 different structures but VGG16 and VGG19 are frequently used. Related work is listed as Table IV.

6.2. Relevant Work and Evaluation

Table 4. Transfer learning with VGGNet

Reference	Dataset	Size of Dataset	Performance	Year
Cruz-Roa, et al. [19]	Private dataset	10 pathology slide cases	VGG16: 76.60% IBCa-CNN: 89.80%	2015
Chato, et al. [20]	BRATS 2017 Challenge	163 samples in BraTS 2017	91%(Alexnet): Linear Discriminant classifier 86.4%(VGG16):	2017
Swati et al. [21]	CE-MRI dataset	3064 abnormal brain CE-MRI	94.82% for the highest accuracy	2019
Rehman, et al. [22]	CE-MRI dataset	3064 abnormal brain CE-MRI	98.69%: fine-tune VGG16	2020
Chelghoum, et al. [23]	CE-MRI dataset	3064 abnormal brain CE-MRI	98.71% by using VGG-16 with 90 epochs	2020
Ahuja, et al. [24]	(BraTS) 2019 challenge database	155 slices	VGG-19 at epoch 6, 99.82% training accuracy	2020
Kora, et al. [25]	CE-MRI dataset	3064 abnormal brain CE-MRI	VGG-16: 98.16%	2021

Early after VGGNet was invented, Cruz-Roa, et al. [19] applied it with transfer learning on medulloblastoma (a kind of brain tumor) detection task. And, another CNN model, which is called IBCa-CNN was trained previously. But these two models were trained in two different domains and IBCa-CNN, which is a 2-layer CNN, was trained for invasive breast cancer tumor classification. VGG16 failed the competition, and this survey stressed the importance of the similarity between source domain and target one. Besides, this research stimulates academics' interests in how to use transfer learning models wisely. And later relevant work did give out better performance. After that, Chato, et al. [20] launched a further study with VGGNet in transfer learning. 163 samples in BraTS 2017 were considered and due to the speciality of the dataset, it was turned out to be a multi-class

classification task. AlexNet and VGG16 were tried, and they both generated relatively good results, but VGG16 still couldn't entirely show its power due to overfitting and negative transfer problems.

In 2019, CE-MRI dataset from figshare was considered. Different groups tried binary and multi-class classification with the dataset using transfer learning. From 94.82% for the highest accuracy in the work of Swati, et al. [21], Rehman, et al. [22] used freeze and fine-tuned techniques to modify the pre-trained models' layers and got a higher accuracy of 98.69% in even a multi-class classification. And in the same year of 2020, Chelghoum, et al. [23] did a challenging job, which contained 9 pre-trained models for comparison in a multi-class classification task. And they also recorded the epochs for the best performance and drew the conclusion that VGG16 could reach the highest accuracy of 98.71% with 90 epochs.

Almost from the CE-MRI was tried, the performance of VGG in transfer learning got on to a higher stage. In more recent work, Ahuja, et al. [24] used BraTS 2019 dataset and realized brain tumor segmentation and glioma (a kind of) detection with superpixel technique. The proposed system yielded 99.82% training accuracy and illustrated the potential of VGG for tumor detection although it's a binary classification.

In 2021, Kora, et al. [25] retested the CE-MRI dataset, emphasizing the differences when considering VGG-16 as a pre-trained model with a novel approach of looking at convolutional layers established by the University of California. The nonlinearity of a sequence of 3 by 3 convolutions combined into a single layer, which resembles the features of GoogLeNet, provides two advantages: one is that the nonlinearity leads to better discrimination in terms of the receptive field; Another advantage is that researchers can examine the filter kernel size of this layer on the original image. And, when comparing, InceptionV3 and XceptionNet with VGG-16, their work using the VGG-16 architecture produced a relatively higher accuracy of 98.16 percent.

7. ResNeXT

In this section, the module structure and relevant work of ResNeXt, which is a combination of ResNet and Inception, will be explained in detail. And ResNeXt doesn't need any manual designed complicated structure like Inception, instead, each branch of ResNeXt uses the same topology structure. The essence of ResNeXt is Group Convolution which controls the number of groups by adjusting cardinality. If looking back at Fig. 3, it can be found that the inner structure of Inception is quite exquisite and it's also getting hard to adjust those hyper-parameters. Therefore, group convolution is always regarded as a tradeoff between common convolution and depth-wise separable convolution.

Inception has been mentioned in the former part of this review, ResNet will be introduced briefly. The study of ResNet [30] was based on the degradation, which means the accuracy decreases while increasing the depth of neural network. Identity transformation can't be well realized because of nonlinear structure so a one-by-one convolution was added to the module, which is called a shortcut connection. And ResNet won the champion in ILSVRC 2015 and had been widely used and modified in transfer learning tasks. Related work is listed as Table V.

7.1. Relevant Work and Evaluation

Table 5. Transfer learning with ResNet

Reference	Dataset	Size of Dataset	Performance	Year
Chelghoum, et al. [23]	CE-MRI dataset	3064 abnormal brain CE-MRI	98.71% by using VGG-16 with 90 epochs	2020
Kaur, et al. [31]	Multiple resources: Harvard Repository with 3 versions; Clinical; figShare dataset	V1:50 T2, V2:74 T2, V3:160 T2, Clinical:500, figShare:3064	AlexNet: 100%, 94%, and 95.92% for three datasets	2020
Divya, et al. [32]	CE-MRI dataset	3064 abnormal brain CE-MRI	A maximum accuracy as 98.67%	2020
Alnemer, et al. [33]	Kaggle repository	7023 MR images	Overall accuracy:98.9% after data augmentation	2021
Polat, et al. [34]	CE-MRI dataset	3064 abnormal brain CE-MRI	The highest classification performance is 99.02%	2021
KULKARNI, et al. [12]	Private dataset	100benign+100malignant	Fine-tuned AlexNet: 93.7%(precision), 100%(recall), 96.77%(f-measure)	2021
Ananda Kumar, et al. [2]	Brats MRI image dataset 2020HS dataset	Unknown	The proposed method attains highest accuracy of 99.57%	2022

Early in 2020, Chelghoum, et al. [23] and Kaur, et al. [31] have tested multiple transfer learning models on different dataset. Every state-of-the-art model including ResNet shows its powerful ability of classification and prediction. But focusing on ResNet or ResNeXt with multi-class classification, Divya, et al. [32] used CE-MRI dataset and gained a maximum accuracy of 98.67%. And in the next year of 2021, Polat, et al. [34] improved the work by using Adadelta(a kind of optimization algorithm) with the same dataset. Almost at the same time, Alnemer, et al. [33] distinguished four clinical states of brain tumor using 7023 MR images in the Kaggle repository. The proposed system is structured by modified ResNet152V2 network and finally gained overall accuracy of 98.9% after data augmentation.

In more recent work, Kumar, et al. [2] tried a new technique called Hyb-DCNN-ResNet152 TL on binary classification of tumors. The proposed method achieves the best accuracy of 99.57 percent by tuning the weight parameters using the Covid-19 optimization algorithm (Cov-19 OA).

Some classical models have been discussed in detail, but other brilliant models like MobileNet, NASNet, SENet, etc. are not illustrated in this article. Partly because some of the models have their

limitations in TL and fewer academics choose to apply them on brain tumor detection. But they still have unique features to attain a good performance. For example, MobileNet is one of the lightweight neural networks and it immensely improves the computing speed. And just like its name, everyone can use it at any time to cope with a classification work or other ML problems. Instead, NASNet needs much time to attain a finest result. So, this article just analyses the balanced and classical TL pre-trained models. And other potential models for TL are listed in.

8. Discussion

As deep learning and transfer learning improve, more transfer learning studies in the field of medical image analysis will include meta-learning and GNN for the building of higher-performance CAD systems. And still, there are no in-depth research results on the measurement of influence of similarity and commonality in different domains when transferring knowledge, so more accurate measurement methods need to be determined in the future. Secondly, the application of transfer learning is very wide in algorithm research. Transfer learning is mainly used in classification algorithm, but some ideas presented in this article can be transferred to other fields like natural language processing and automatic speech recognition. Thirdly, as for transferable learning conditions, the essential attributes of positive transfer have not been fully understood. How to avoid negative transfer and the effectiveness of transfer are also one of the directions. And negative transfer is an important technical difficulty. Negative transfer refers to the phenomenon that the use of source domain data in target domain training does not improve the model ability, instead, reducing the recognition rate. There have been several explorations of tasks and clustering between tasks to provide guidance on how to automatically avoid negative transfer.

To only address the detection problems, the latest research has reached the very stage. And all computer-aided diagnosis' models get a good performance of accuracy. The true problem is to cope with visual categorization when lack of data to train the proposed system. And this article only summaries about the work on inductive transfer learning. This kind of image classification task is always much more direct and easier to understand than unsupervised transfer learning. The latter one means there is no labeled data in both source and target domain. Unsupervised learning of representations has also been found to be beneficial in some studies when the input distribution $P(x)$ is structurally related to a goal of interest, such as predicting $P(y | x)$. And this problem can be further studied even in medical area. Also, this review helps conclude the base work of TL in brain tumor detection and widely used pre-trained models with satisfying results are summaries as above. And attention can be focused on the preparation before beginning to find a proper model and modify the parameters. Multiple sampling strategies can greatly improve the accuracy.

9. Conclusion

Using magnetic resonance imaging to detect anomalies in the human body has long been a focused area in medical computer-aided diagnosis (CAD). Starting from collecting data prudentially, researchers then preprocess data using manifold approaches to increase the total size of datasets or highlight the features. And the pivotal part, the pre-trained model, which is chosen in a scrutinizing way, are analyzed chronologically. From AlexNet to Hyb-DCNN-ResNet152, three frequently used models of them can all generate a quite good performance. Researchers can choose the suggested models in other research to get work improved. Nonetheless, there are shortcomings in this review, such as a lack of coverage of major works in the subject. This survey, on the other hand, is believed to be suited for readers of various backgrounds and may be used as a primer on interdisciplinary research.

References

- [1] O. Charron, A. Lallement, D. Jarnet, V. Noblet, J.B. Clavier, P. Meyer, "Automatic detection and segmentation of brain metastases on multimodal MR images with a deep convolutional neural network". *Computers in biology and medicine*, 2018, 95: 43-54.
- [2] K.S.A. Kumar, A.Y. Prasad, J. Metan, "A hybrid deep CNN-Cov-19-Res-Net Transfer learning architype for an enhanced Brain tumor Detection and Classification scheme in medical image processing". *Biomedical Signal Processing and Control*, 2022, 76: 103631.
- [3] S.J. Pan, Q. Yang, "A survey on transfer learning". *IEEE Transactions on knowledge and data engineering*, 2009, 22(10): 1345-1359.
- [4] L. Shao, F. Zhu, X. Li, "Transfer learning for visual categorization: A survey". *IEEE transactions on neural networks and learning systems*, 2014, 26(5): 1019-1034.
- [5] K. Weiss, T.M. Khoshgoftaar, D.D. Wang, "A survey of transfer learning". *Journal of Big data*, 2016, 3(1): 1-40.
- [6] S. Ahuja, B.K. Panigrahi, T. Gandhi, "Transfer learning based brain tumor detection and segmentation using superpixel technique" [C]//2020 International Conference on Contemporary Computing and Applications (IC3A). IEEE, 2020: 244-249.
- [7] M.G. Ertosun, D.L. Rubin, "Automated grading of gliomas using deep learning in digital pathology images: A modular approach with ensemble of convolutional neural networks AMIA Annual Symposium Proceedings, American Medical Informatics Association", San Francisco, CA (2015), p. 1899.
- [8] L. Shen, T. Anderson, "Multimodal brain MRI tumor segmentation via convolutional neural networks", 2017.
- [9] S. Deepak, P.M. Ameer, "Brain tumor classification using deep CNN features via transfer learning". *Computers in biology and medicine*, 2019, 111: 103345.
- [10] A. Alnemer, J. Rasheed, "An Efficient Transfer Learning-based Model for Classification of Brain Tumor" [C]//2021 5th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT). IEEE, 2021: 478-482.
- [11] T. Kaur, T.K. Gandhi, "Deep convolutional neural networks with transfer learning for automated brain image classification". *Machine Vision and Applications*, 2020, 31(3): 1-16.
- [12] S.M. Kulkarni, G. Sundari, "Comparative analysis of performance of deep cnn based framework for brain mri classification using transfer learning". *Journal of Engineering Science and Technology*, 2021, 16(4): 2901-2917.
- [13] T.Y. Liu, "Easyensemble and feature selection for imbalance data sets" [C]//2009 international joint conference on bioinformatics, systems biology and intelligent computing. IEEE, 2009: 517-520.
- [14] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, "Going deeper with convolutions" [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 1-9.
- [15] A. Krizhevsky, I. Sutskever, G.E. Hinton, "Imagenet classification with deep convolutional neural networks". *Advances in neural information processing systems*, 2012, 25.
- [16] M.Z. Alom, T.M. Taha, C. Yakopcic, S. Westberg, P. Sidike, M.S. Nasrin, B.C.V. Esesn, A.AS. Awwal, V.K. Asari, "The history began from alexnet: A comprehensive survey on deep learning approaches". *arXiv preprint arXiv:1803.01164*, 2018.
- [17] M. Lin, Q. Chen, S. Yan, "Network in network". *arXiv preprint arXiv:1312.4400*, 2013.
- [18] K. Simonyan, A. Zisserman, "Very deep convolutional networks for large-scale image recognition". *arXiv preprint arXiv:1409.1556*, 2014.
- [19] A. Cruz-Roa, J. Arévalo, A. Judkins, A. Madabhushi, F. González, "A method for medulloblastoma tumor differentiation based on convolutional neural networks and transfer learning"//11th International Symposium on Medical Information Processing and Analysis. SPIE, 2015, 9681: 8-15.
- [20] L. Chato, S. Latifi, "Machine learning and deep learning techniques to predict overall survival of brain tumor patients using MRI images". 2017 IEEE 17th International Conference on Bioinformatics and Bioengineering (BIBE), IEEE, Washington, DC, USA (2017), pp. 9-14.

- [21] Z.N.K. Swati, Q. Zhao, M. Kabir, F. Ali, Z. Ali, S. Ahmed, J.F. Lu, "Brain tumor classification for MR images using transfer learning and fine-tuning". *Computerized Medical Imaging and Graphics*, 2019, 75: 34-46.
- [22] A. Rehman, S. Naz, M.I. Razzak, F. Akram, M. Lmran, "A deep learning-based framework for automatic brain tumors classification using transfer learning". *Circuits, Systems, and Signal Processing*, 2020, 39(2): 757-775.
- [23] R. Chelghoum, A. Ikhlef, A. Hameurlaine, S. Jacquir, "Transfer learning using convolutional neural network architectures for brain tumor classification from MRI images"//*IFIP International Conference on Artificial Intelligence Applications and Innovations*. Springer, Cham, 2020: 189-200.
- [24] S. Ahuja, B.K. Panigrahi, T. Gandhi, "Transfer learning based brain tumor detection and segmentation using superpixel technique" [C]//*2020 International Conference on Contemporary Computing and Applications (IC3A)*. IEEE, 2020: 244-249.
- [25] P. Kora, S. Mohammed, MJS Teja, C.U. Kumari, K. Swaraja, K. Meenakshi, "Brain Tumor Detection with Transfer Learning" [C]//*2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*. IEEE, 2021: 443-446.
- [26] J. Amin, M. Sharif, M. Yasmin, T. Saba, M. A. Anjum, S.L. Fernandes, "A new approach for brain tumor segmentation and classification based on score level fusion using transfer learning". *Med. Syst.*, 43 (11) (2019), p. 326.
- [27] S. Anjum, L. Hussain, M. Ali, M.H. Alkinani, W. Aziz, S. Gheller, A. A. Abbasi, A.R. Marchal, H. Suresh, T.Q. Duong, "Detecting brain tumors using deep learning convolutional neural network with transfer learning approach". *International Journal of Imaging Systems and Technology*, 2022, 32(1): 307-323.
- [28] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, "Rethinking the inception architecture for computer vision" [C]//*Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016: 2818-2826.
- [29] S. Ioffe, C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift" [C]//*International conference on machine learning*. PMLR, 2015: 448-456.
- [30] K. He, X. Zhang, S. Ren, J. Sun, "Deep residual learning for image recognition" [C]//*Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016: 770-778.
- [31] T. Kaur, T.K. Gandhi, "Deep convolutional neural networks with transfer learning for automated brain image classification". *Machine Vision and Applications*, 2020, 31(3): 1-16.
- [32] S. Divya, L.P. Suresh, A. John, "A Deep Transfer Learning framework for Multi Class Brain Tumor Classification using MRI" [C]//*2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*. IEEE, 2020: 283-290.
- [33] A. Alnemer, J. Rasheed, "An Efficient Transfer Learning-based Model for Classification of Brain Tumor" [C]//*2021 5th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*. IEEE, 2021: 478-482.
- [34] Ö. Polat, C. Güngen, "Classification of brain tumors from MR images using deep transfer learning". *The Journal of Supercomputing*, 2021, 77(7): 7236-7252.