Efficient Transfer Learning using Pre-trained Models on CT/MRI

Nicole Guobadia

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Science in Computer Science & Systems

University of Washington

2023

Committee:

Dr. Juhua Hu

Dr. Vikas Kumar

Program Authorized to Offer Degree:

School of Engineering and Technology

Efficient Transfer Learning using Pre-trained Models on CT/MRI

***Abstract*—The medical imaging field has unique obstacles to face when performing computer vision classification tasks. The retrieval of the data, be it CT scans or MRI, is not only expensive but also limited due to the lack of publicly available labeled data. In spite of this, clinicians often need this medical imaging data to perform diagnosis and recommendations for treatment. This motivates the use of efficient transfer learning techniques to not only condense the complexity of the data as it is often volumetric, but also to achieve better results faster through the use of established machine learning techniques like transfer learning, fine-tuning, and shallow deep learning. In this paper, we introduce a three-step process to perform classification using CT scans and MRI data. The process makes use of fine-tuning to align the pretrained model with the target class, feature extraction to preserve learned information for downstream classification tasks, and shallow deep learning to perform subsequent training. Experiments are done to compare the performance of the proposed methodology as well as the time cost trade offs for using our technique compared to other baseline methods. Through these experiments we find that our proposed method outperforms all other baselines while achieving a substantial speed up in overall training time.**

***Index Terms*—CT Scans, MRI, Transfer learning, Fine-tuning, Feature Extraction**

I. Introduction

Computer vision within the medical field suffers from a variety of issues, namely the lack of labelled medical imaging data, inherent imbalances regarding distribution, and a lack of resources (financial, computational, expertise, etc.) to label what little imaging data already exists. Within the medical imaging domain there is a substantial cost in terms of training time and resources as the data often involves height, width, and an additional dimension: depth [11]. A lack of imaging data is commonly alleviated with the use of transfer learning and data augmentation as demonstrated through [17]. Through this and other works like [13] and [24], common valuable features can be shared across multiple imaging domains, even from unrelated areas like ImageNet [6] to the medical imaging domain. This allows a network to be trained on an abundance of data with the weights being transferred into the relevant domain to be further fine-tuned.

Imaging data is represented using a complex data structure, often represented as H × W where H is the height of the image and W is the width of the image. This gives the data by itself a space complexity of O (H × W). Within the medical imaging domain, it is even higher as there is an additional dimension: depth. This leads to a space complexity of O (H × W × D), where D is the depth. Due to this, the time and space needed to train a model for medical imaging classification tasks is much higher than other similar image classification tasks in different domains [11]. Transfer learning is a technique that has arisen to tackle this exact problem. Transfer learning involves taking the weights of a deep learning model architecture already trained for a specific classification task and applying those weights to another deep learning model upon initialization. The importance of transfer learning comes from the fact that in most real-world problems, we do not have millions of labeled data often needed to train these complex deep learning architectures [3]. Therefore, using transfer learning allows us to train the deep learning model much faster while achieving higher results [3]. Despite this, there is still an issue where we must train the newly initialized deep learning model using 3D data. This means there is still a resource issue involving space and hardware. This is because we still must load the data and train the model while storing and updating weights during the process [19].

A subset of transfer learning is feature extraction; in which image data are passed through a pretrained model, and then fed into a traditional machine learning architecture for classification tasks. This technique helps circumvent the cost of training and resources as it is far less computationally intensive [11]. There is, however, a notable tradeoff between using transfer learning versus feature extraction: performance. Traditional machine learning models are routinely surpassed by their deep learning model counterpart [11]. Despite this, feature extraction is still a viable approach, especially when combined with shallow deep learning, as it still allows for generalization due to the natural limitations of traditional machine learning methods being less complex.

To solve these issues, we propose an efficient transfer learning pipeline that leverages knowledge learned from pretrained models, performs light fine-tuning to allow the model to align with the classification task, extracts the learned features after fine-tuning, and trains a subsequent multi-layer perceptron (MLP) for downstream classification tasks. This process solves the issue of highly complex data structures, as the three-dimensional input becomes one dimension through the process of feature extraction. Likewise, the training of the MLP through shallow deep learning is far less resource intensive in regard to the time taken to complete training, allowing for faster subsequent training for downstream classification tasks while achieving strong performance. We measure this by conducting experiments to evaluate the performance and time cost of the proposed method compared to established techniques like transfer learning and linear probing. Through these experiments, we find that our proposed method outper-

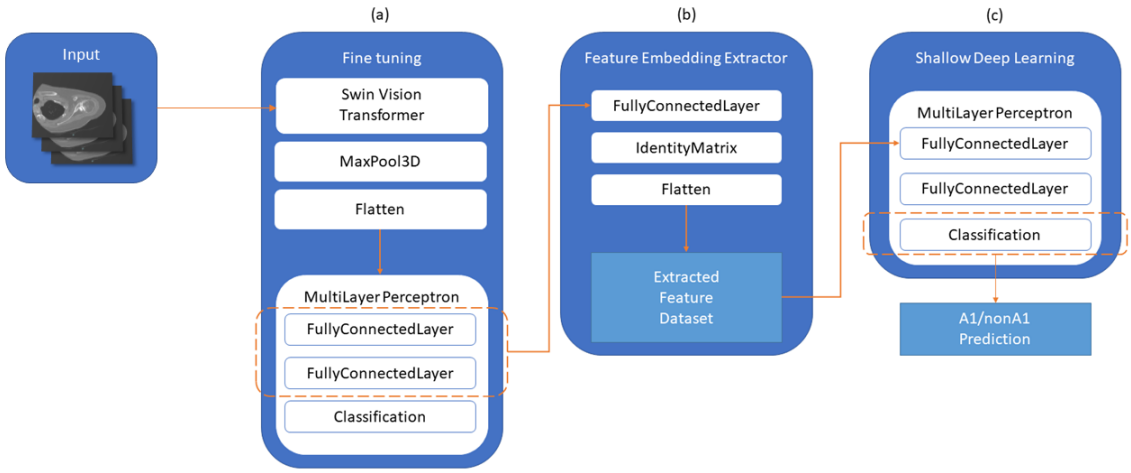

Fig. 1. Overview of methodology using sample Walch CT scan as input. Output is classification of A1 or nonA1.

forms these baselines, while reducing training speeds by 40% when compared to the transfer learning model. Subsequent training using the shallow deep learning component compared to the transfer learning model yields an even higher reduction of 99%.

II. Related work

In this section, we will review works related to our proposed methodology.

A. Transfer learning

In the past, traditional machine learning techniques relied on handcrafted features to derive predictive characteristics from imaging data. This has been surpassed by the use of deep learning as seen in previous works like [12]. Transfer learning has drastically improved the time and space complexity of medical imaging classification tasks as many pretrained networks are available online to the research community [3]. Due to this, handcrafted features have been replaced by these pretrained networks and is the standard approach for many medical imaging classification and segmentation tasks [11].

Traditional fine-tuning using the pretrained model itself is similarly popular and has been quite fruitful. Works like [8], [4], and [18] have used pretrained networks like [21] and [9] for medical imaging classification tasks and demonstrate a clear improvement in performance when compared to its feature extraction counterpart. While using transfer learning and fine-tuning techniques have shown promising results, the training of these models only slightly aids in the time and space constraints that are often faced in these kinds of classification tasks.

B. Feature extraction

An alternative to transfer learning is feature extraction, where a pretrained deep learning model is used to extract features, and those features are then fed into a traditional machine learning model for training. This technique has been used within the medical imaging domain to identify deformities and anomalies. Previous works like [2], [13], and [24], have shown that features extracted from pretrained networks without fine-tuning using both 2D and 3D CNNs achieve high ranks in medical imaging benchmarking competitions like SLIVER07 [20] and PROMISE12 [10].

Transfer learning and feature extraction are key elements to our proposed method, as we leverage the pretrained weights via transfer learning and employ additional fine-tuning to help align the model with our classification task. This additionally aids in the feature extraction process, as those features will now be better fit for the given classification task at hand, resulting in better performance.

III. Methodology

Our proposed methodology consists of three distinct steps: fine-tuning, feature extraction, and shallow deep learning. Figure 1 shows an overview of the proposed methodology. Medical images are fed into the fine-tuning component, which consists of a pretrained SwinViT model and a multi-layer perceptron. The medical imaging data is input to the pretrained model. This output is then pooled and flattened. This flattened output is then the input to the MLP. This model is then fine-tuned to align with the classification task. In the feature extraction component, feature embeddings are extracted from either of the two fully connected layers of the MLP using an identity matrix followed by a flattening layer. This is performed on every patient in the dataset, constructing a new lightweight dataset consisting of extracted features fine-tuned for the classification task at hand. In the shallow deep learning component, the extracted feature dataset is then used to train another smaller MLP for subsequent classification tasks in the shallow deep learning process. In the following sections, we will explore each component in more detail.

A. Fine-tuning

Swin UNETR [5] is a model created for segmentation of brain tumors using CT scans. The encoder is a volumetric Swin transformer [14] pretrained on 5,050 publicly available

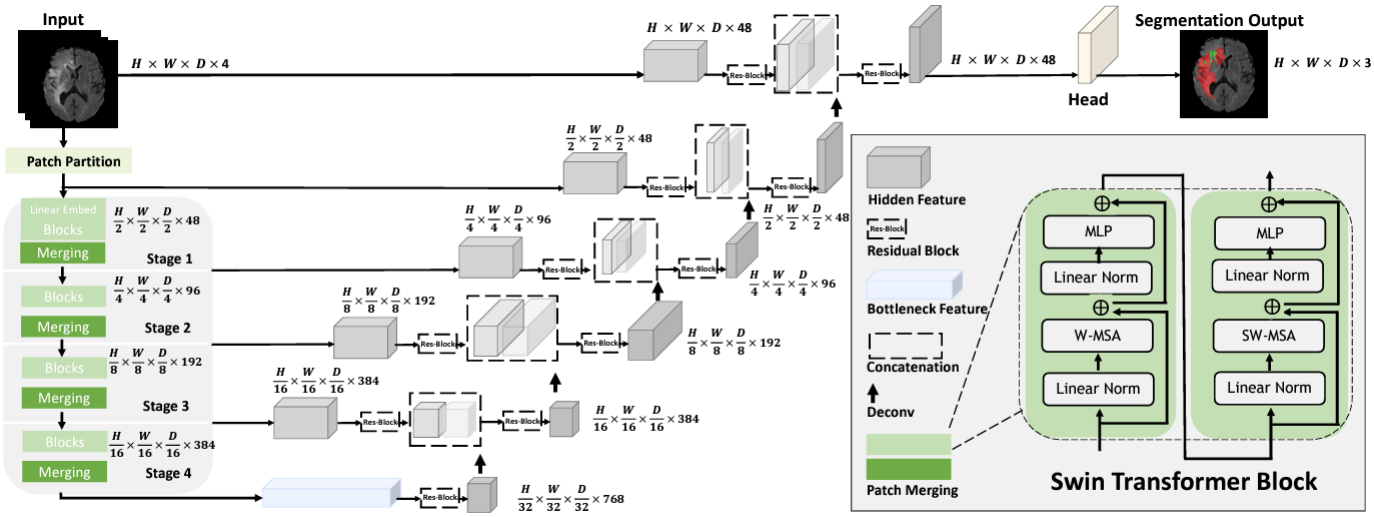

Fig. 2. Pretrained Swin-UNETR model architecture [23].

CT scans from various body organs covering the head, neck, chest, abdomen, and pelvis [23]. We will refer to this pre-trained model as SwinViT. Let x_i be a medical image from dataset $X^{S\times H\times W\times D}$, where H is height, W is width, D is depth, and S is the number of samples within the dataset. Using the output of the Stage 4 of the SwinViT model [23] $((\frac{H}{32} \times \frac{W}{32} \times \frac{D}{32}) \times 768)$, we perform a max pooling and subsequent flattening operation to prepare for input into the multi-layer perceptron for classification. The process is summarized in equation 1.

$$\begin{aligned} SwinViT(x_i) &= \frac{H}{32} \times \frac{W}{32} \times \frac{D}{32} \times 768, \\ z_i &= Flatten(MaxPool3d(SwinViT(x_i))). \end{aligned} \tag{1}$$

The multi-layer perceptron consists of two fully connected layers followed by a classification layer. The process is summarized in equation 2.

$$\hat{y_i} = MLP(z_i) \tag{2}$$

B. Feature extraction

For feature extraction, one of the fully connected layers is taken from the fine-tuned model and an identity matrix is applied so the shape and weights are preserved. For each patient p in the dataset X, their medical imaging data x_p is passed through this feature embedding extraction module. Each feature embedding e is flattened and then saved in a flat file format, along with a primary key, resulting in the extracted feature dataset $L^{S\times F}$, where F is the number of out features of the selected fully connected layer in the MLP. The process is summarized in equation 3.

$$\begin{aligned} e_i &= Identity(MLP(x_i)), \\ l_i &= Flatten(e_i). \end{aligned} \tag{3}$$

C. Shallow deep learning

For shallow deep learning, a patient p_i's extracted features l_i is fed into a multi-layer perceptron. This multi-layer perceptron consists of two fully connected layers followed by a classification layer. This process is summarized in equation 4.

$$\hat{y'} = MLP'(l_i). \tag{4}$$

IV. Experiments

To demonstrate the proposed method in terms of both efficiency and effectiveness, we conduct experiments on two different kinds of downstream tasks, that is, classifications on MRI and CT scans as follows.

A. Datasets

1) Cancer-Net BCa-S: The Breast Cancer Grade Prediction dataset uses Synthetic Correlated Diffusion Imaging MRI sequencing to perform pathological complete response (pCR) prediction [22] [15]. There are a total of 253 patients, 82 have pCR and 171 have no pCR. This dataset highlights a common issue seen within the medical imaging computer vision research community as there is a limited number of samples in this dataset, along with a severe class imbalance of about 1:3.

2) Walch dataset (CT scans): Walch classification is the primary method to describe glenoid morphology and pathology in patients with glenohumeral osteoarthritis, as an attempt to standardize the categorization of shoulder arthroplasty for pathologic glenoid types for diagnostics and treatment purposes [1]. There are a total of 7 classes, but this is broken down into two meta classes for this experiment: A1 (centered humeral head, concentric wear, no fatty infiltration of the humeral head), and nonA1 (some kind of erosion or misalignment of the glenoid and scapula) [1]. For this dataset, there are a total of 1,815 patients. 775 are classified as A1 and 1,040 are classified as nonA1.

B. Evaluation Setup

Each data is split using an 80/10/10 train/validation/test ratio. Oversampling and data augmentation techniques are used on the train split only. For evaluation, common classification metrics are used such as accuracy $\frac{(TP+TN)}{(TP+TN+FP+FN)}$, precision $\frac{TP}{(TP+FP)}$, recall $\frac{TP}{(TP+FN)}$, and F1-score $(2 \times (\frac{precision \times recall}{(precision+recall)})$. Due to some imbalance in the datasets, the area under the curve (AUC) is also reported as accuracy will not be sufficient for model evaluation. Additionally, providing the precision allows us to measure the accuracy of the true positives, while reporting the recall gives insight to the completeness of the true positives. Often, maximizing precision and recall may not be possible as there is a natural trade off that occurs between the two, and thus measuring the F1-score will allow for a more holistic interpretation of the two. For time cost measurement, hours, minutes, and seconds are used where appropriate.

C. Baselines

To verify the effectiveness of each component, a total of three baselines are created: pretrained SwinViT with linear probing, fine-tuning with linear probing, and transfer learning. The pretrained SwinViT with linear probing model simply loads the pretrained weights from [5], performs feature extraction, and then classification using one fully connected layer. This isolates the feature extraction component to see the importance of the fine-tuning process. Fine-tuning with linear probing performs "light" fine-tuning (less than 1,000 epochs), uses the best performing model based on accuracy, performs feature extraction, and then classification using one fully connected layer. This isolates the shallow deep learning component, allowing us to see if simpler models can be used rather than a multi-layer perceptron. Lastly, the transfer learning model simply uses the fine-tuned model as the end-to-end classifier itself. This model is trained using the full 1,000 epochs. This baseline will aid in evaluating the time cost trade-offs, as we expected this model to have higher performance than our proposed method, but also to take the longest amount of time to train. Our proposed method acts similarly to fine-tuning with linear probing, however rather than performing shallow deep learning using the linear probe, we use a multi-layer perceptron.

To show the effectiveness on either modality, extensive tests are run only on the MRI dataset, while the training process for the CT scan dataset is far less expensive. This is done due to computational constraints, however it also aids in demonstrating the effectiveness of our pipeline in a real world scenario given the resource limitations and positive performance. Additional hyper-parameters can be found in Table I.

Additionally, for the Cancer-Net BCa dataset [22], a baseline is provided with accuracy reported. This architecture is a Volumetric ResNet-34 [7] and follows a similar pooling and flattening process as discussed in section III-A.

D. Results

1) Performance comparisons: Table II shows the evaluation metric results. Using this, it is clear that the proposed method outperforms all other baseline methods. The proposed method benefits from additional hidden layers being added to the model as seen from the increase in performance compared to the proposed method with linear probing. Fine-tuning with linear probing performs worse than the baseline from paper. It should be noted that the volumetric ResNet-34 model from [22] uses the same multi-layer perceptron architecture as our proposed method, adding credence to our linear probing observation. Using the AUC value for the pretrained SwinViT results, we see that the linear probe is not able to learn from the extracted features without additional fine-tuning. This further confirms the need of small amount fine-tuning in our proposal, since the pre-trained model can be trained with a very different task.

The transfer learning model was not able to generalize well and quickly overfits to the data as seen in Figure 3. We theorize that this likely occurs due to the small amount of data paired with the complexity of the SwinViT model. Despite employing additional regularization techniques like dropout and data augmentation, the results were never able to compete with the level of the results seen from the proposed methodology's results.

The results for Walch classification are seen in Table III. Here, we see a similar phenomenon as we did with the results for the MRI dataset. It is noteworthy, however, that the performance of the transfer learning model does surpass the pretrained SwinViT model when comparing AUC. We theorize that this is due to there being more data available within this dataset, which emphasizes the importance of the amount of diverse data. Despite this, our proposed method with a multi-layer perception still outperforms all other baselines.

2) Time cost comparisons: Table II additionally displays the execution time for each step of the proposed method process. Note that the costs for fine-tuning and feature extraction are amortized, as this only needs to occur once when preparing for downstream tasks. The costs for the proposed methods are calculated by taking the timestamp of the checkpoint used for feature extraction and subtracting the timestamp at the initial time of training.

While the pretrained SwinViT linear probing has the fastest training time as there is no fine-tuning time cost, it has the second worst performance. The transfer learning fairs worse with not only the worst performance but also the longest training time. When comparing the proposed method variations however, there is only a minor difference between the total time taken to train (3 seconds). This, compared to the AUC performance trade-off of the MLP variant versus the linear probing variant would suggest that the former to be the superior model when evaluated in a holistic manner. This further demonstrates both the efficency and effectiveness of our proposed method.

TABLE I
HYPERPARAMETERS

Model architecture	batch size	epochs	lr	FC1(in,out)	Dropout(p)	FC2(in,out)
MRI Hyperparameters						
Transfer learning	1	1000	$1e^{-6}$	(768,256)	(0.5)	(256,128)
MLP	64	100	0.0001	(256,128)	(0.5)	(128,2)
Linear probe	64	100	0.0001	(256,2)	-	-
CT scan Hyperparameters						
Transfer learning	1	100	$1e^{-5}$	(768,64)	(0.8)	(64,2)
MLP	256	400	0.0001	(256,128)	(0.5)	(128,2)
Linear probe	256	400	0.0001	(256,2)	-	-

TABLE II
PERFORMANCE METRIC COMPARISONS FOR CANCERNET-BCA DATASET

	Evaluation metrics					Execution time			
Model name	***Accuracy***	***AUC***	***Precision***	***Recall***	***F1-score***	***Fine-tuning***	***Feature extraction***	***Shallow deep learning***	***Total***
Volumetric ResNet34	0.8775	-	-	-	-	-	-	-	-
Pretrained SwinViT (w/linear probing)	0.6923	0.5000	0.4793	0.6923	0.5664	**0s**	41s	1m 58s	**2m 39s**
Transfer learning model	0.6154	0.4444	0.4615	0.6154	0.5275	5h 4m 1s	**0s**	**0s**	5h 4m 1s
Proposed method (w/linear probing)	0.8616	0.8028	0.8667	0.8612	0.8552	2h 58m 40s	50s	1m 39s	3h 1m 9s
Proposed method (w/MLP)	**0.9231**	**0.9028**	**0.9269**	**0.9231**	**0.9222**	2h 58m 40s	50s	1m 42s	3h 1m 12s

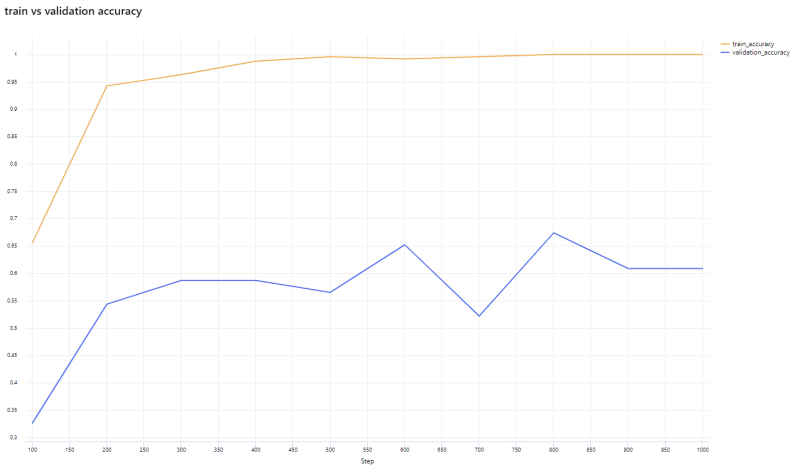

Fig. 3. Transfer learning model training and validation accuracy over 100 epochs.

V. CONCLUSION

Following the design of our methodology and experiments we conclude the following:

- The proposed method is effective on both MRI and CT scan modalities.
- Directly using pretrained models without fine-tuning is ineffective.
- When comparing the transfer learning fine-tuning training speed with the proposed method's fine-tuning training speed, we see a training speed decrease in 40%.
- By using our methodology, subsequent training using extracted features is lightweight, and further decreases training by 99%.

There is future work we would like to focus on. A major contribution of the paper is the reduction of training time for subsequent classification tasks using the extracted features, and thus performing additional classification tasks using those

TABLE III
PERFORMANCE METRIC COMPARISONS FOR WALCH DATASET

Model name	Evaluation metrics				
	Accuracy	***AUC***	***Precision***	***Recall***	***F1-score***
Pretrained SwinViT (w/linear probing)	0.5934	0.5593	0.5817	0.5934	0.5680
Transfer learning model	0.5934	0.5881	0.5962	0.5934	0.5945
Proposed method (w/linear probing)	0.8022	0.7933	0.8016	0.8022	0.8010
Proposed method (w/ MLP)	0.8571	0.8558	0.8600	0.8600	0.8600

extracted features could further demonstrate their usefulness. Additionally, condensing the methodology into an end-to-end pipeline process is imperative, as there are currently three distinct steps in the process. Creating a streamlined process will additionally aid in efficient training, the key goal of our paper.

VI. ACKNOWLEDGMENT

Thank you to Exactech and BlueOrtho for providing access to CT scan data, clinician guidance, and computational resources. Guobadia's research is also partially supported by the Advata Gift Funding. All opinions, findings, conclusions and recommendations in this paper are those of the author and do not necessarily reflect the views of the funding agencies.

REFERENCES

[1] Barnsley L, Knipe H, Weerakkody Y, et al. Walch classification of glenoid morphology. Reference article, Radiopaedia.org (Accessed on 31 May 2023)

[2] Cha, K. H., Hadjiiski, L. M., Samala, R. K., Chan, H. P., Cohan, R. H., Caoili, E. M., Paramagul, C., Alva, A., Weizer, A. Z. (2016). Bladder Cancer Segmentation in CT for Treatment Response Assessment: Application of Deep-Learning Convolution Neural Network-A Pilot Study. Tomography (Ann Arbor, Mich.), 2(4), 421–429. https://doi.org/10.18383/j.tom.2016.00184

[3] Elgendy, M. (2020). Deep Learning for Vision Systems. United States: Manning.

[4] Gensure, R. H., Chiang, M. F., Campbell, J. P. (2020). Artificial intelligence for retinopathy of prematurity. Current opinion in ophthalmology, 31(5), 312–317. https://doi.org/10.1097/ICU.0000000000000680

[5] Hatamizadeh, A., Nath, V., Tang, Y., Yang, D., Roth, H.R., Xu, D. (2022). Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images. In: Crimi, A., Bakas, S. (eds) Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2021. Lecture Notes in Computer Science, vol 12962. Springer, Cham.

[6] J. Deng, W. Dong, R. Socher, L. -J. Li, Kai Li and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 2009, pp. 248-255, doi: 10.1109/CVPR.2009.5206848.

[7] Kataoka, H., Wakamiya, T., Hara, K., Satoh, Y. (2020). Would mega-scale datasets further enhance spatiotemporal 3D CNNs?. arXiv preprint arXiv:2004.04968.

[8] Kim, E., Corte-Real, M., and Baloch, Z., "A deep semantic mobile application for thyroid cytopathology", in Medical Imaging 2016: PACS and Imaging Informatics: Next Generation and Innovations, 2016, vol. 9789. doi:10.1117/12.2216468.

[9] Krizhevsky, A., Sutskever, I., Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In F. Pereira, C. J. Burges, L. Bottou, K. Q. Weinberger (Eds.), Advances in Neural Information Processing Systems (Vol. 25).

[10] Litjens, G., Toth, R., van de Ven, W., Hoeks, C., Kerkstra, S., van Ginneken, B., Vincent, G., Guillard, G., Birbeck, N., Zhang, J., Strand, R., Malmberg, F., Ou, Y., Davatzikos, C., Kirschner, M., Jung, F., Yuan, J., Qiu, W., Gao, Q., Edwards, P. E., … Madabhushi, A. (2014). Evaluation of prostate segmentation algorithms for MRI: the PROMISE12 challenge. Medical image analysis, 18(2), 359–373. https://doi.org/10.1016/j.media.2013.12.002

[11] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. Medical image analysis, 42, 60–88. https://doi.org/10.1016/j.media.2017.07.005

[12] Qureshi, S.A., Hussain, L., Ibrar, U. et al. Radiogenomic classification for MGMT promoter methylation status using multi-omics fused feature space for least invasive diagnosis through mpMRI scans. Sci Rep 13, 5188 (2023).

[13] Liu, F., Yang, L. (2015). A Novel Cell Detection Method Using Deep Convolutional Neural Network and Maximum-Weight Independent Set. In: Navab, N., Hornegger, J., Wells, W., Frangi, A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science(), vol 9351. Springer, Cham. https://doi.org/10.1007/978-3-319-24574-4

[14] Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., ... Guo, B. (2021). Swin transformer: Hierarchical vision transformer using shifted windows. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 10012-10022).

[15] Newitt, D. C., Partridge, S. C., Zhang, Z., Gibbs, J., Chenevert, T., Rosen, M., Bolan, P., Marques, H., Romanoff, J., Cimino, L., Joe, B. N., Umphrey, H., Ojeda-Fournier, H., Dogan, B., Oh, K. Y., Abe, H., Drukteinis, J., Esserman, L. J., Hylton, N. M. (2021). ACRIN 6698/I-SPY2 Breast DWI [Data set]. The Cancer Imaging Archive. https://doi.org/10.7937/TCIA.KK02-6D95

[16] Nicoletti D, Glenoid morphology - Walch B2. Case study, Radiopaedia.org (Accessed on 08 Jun 2023) https://doi.org/10.53347/rID-94407

[17] Pang, S., Yu, Z., Orgun, M. A. (2017). A novel end-to-end classifier using domain transferred deep convolutional neural networks for biomedical images. Computer methods and programs in biomedicine, 140, 283–293. https://doi.org/10.1016/j.cmpb.2016.12.019

[18] Rajkomar, A., Lingam, S., Taylor, A. G., Blum, M., Mongan, J. (2017). High-Throughput Classification of Radiographs Using Deep Convolutional Neural Networks. Journal of digital imaging, 30(1), 95–101. https://doi.org/10.1007/s10278-016-9914-9

[19] Roth, Holger Lee, Christopher Shin, Hoo-Chang Seff, Ari Kim, Lauren Yao, Jianhua Lu, Le Summers, Ronald. (2015). Anatomy-specific classification of medical images using deep convolutional nets. 2015. 10.1109/ISBI.2015.7163826.

[20] Styner, M., Lee, J., Chin, B., Chin, M., Commowick, O., Tran, H.-H., … Warfield, S. (11 2007). 3D Segmentation in the Clinic: A Grand Challenge II: MS lesion segmentation. MIDAS Journal. doi:10.54294/lmkqvm

[21] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... Rabinovich, A. (2015). Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1-9).

[22] Tai, C.-E. A., Hodzic, N., Flanagan, N., Gunraj, H., Wong, A. (2022). Cancer-Net BCa: Breast Cancer Pathologic Complete Response Prediction using Volumetric Deep Radiomic Features from Synthetic Correlated Diffusion Imaging. ArXiv [Cs.CV].

[23] Tang, Y., Yang, D., Li, W., Roth, H. R., Landman, B., Xu, D., ... Hatamizadeh, A. (2022). Self-supervised pre-training of swin transformers for 3d medical image analysis. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 20730-20740).

[24] Zhang, J., Cheng, Z., Ma, Y., He, C., Lu, Y., Zhao, Y., Chang, X., Zhang, Y., Bai, Y., Cheng, N. (2017). Effectiveness of Screening Modalities in Colorectal Cancer: A Network Meta-Analysis. Clinical colorectal cancer, 16(4), 252–263. https://doi.org/10.1016/j.clcc.2017.03.018