



Article Title: Hybrid Deep Transfer Learning Framework for Stroke Risk Prediction

Hybrid Deep Transfer Learning Framework for Stroke Risk Prediction

Reshma S V, Gini R

¹Student, Department of MCA, Narayanaguru College of Engineering, Anna University

²Assistant professor, Department of MCA, Narayanaguru College of Engineering.

ABSTRACT

Stroke has become a leading cause of death and long-term disability in the world with no effective treatment. Deep learning-based approaches have the potential to outperform existing stroke risk prediction models. Due to the strict privacy protection policy in health-care systems, stroke data is usually distributed among different hospitals in small pieces. Transfer learning can solve small data issue by exploiting the knowledge of a correlated domain, especially when multiple source of data are available. In this work, we propose a novel Hybrid Deep Transfer Learning-based Stroke Risk Prediction scheme.

Keywords: Stroke risk prediction, transfer learning, generative adversarial networks, active learning, Bayesian optimization.

1 Introduction

Stroke is one of the most prevalent diseases which could lead to death or long-term disability among elderly people all over the world. In a recent report, around 795 000 people experience a new or recurrent stroke each year in the US; one stroke incident occurs in approximately every 40 seconds. Among the patients who suffered strokes, one in five would die within one year. For the survivals, the cost of treatment and rehabilitation becomes an extremely high burden to their families and the health-care system. From 2014 to 2015, the direct and indirect cost due to stroke incidents was about 45.5 billion US dollars. Thus, accurate stroke prediction is highly desirable so that the cost can be reduced with early interventions to delay the onset of and to reduce the risks of stroke. There exist several works which exploit medical data (e.g., electronic health record and retinal image) to develop Stroke Risk Prediction (SRP) Models. These methods can be broadly categorized into classical machine learning approaches (e.g., Support Vector Machine (SVM), Decision Tree, Logistic Regression) and deep learning-based approaches. It is reported that deep neural network (DNN) can achieve best performance in stroke prediction.

However, a well-known drawback is that such model relies on the availability of large well-labeled data. In real-world scenario, the quantity of reliable data that is required may not be readily available. Due to strict privacy protection policy in health-care system, sharing stroke data between hospitals is usually difficult. Thus, the full set of stroke data tends to be distributed among multiple hospitals in small subsets. In addition, stroke data contains extremely



Article Title: Hybrid Deep Transfer Learning Framework for Stroke Risk Prediction

imbalanced positive and negative instances. Thus, the DNN-based SRP models could work poorly in real-world deployment. Though the stroke data is small, some common chronic diseases (e.g., hypertension and diabetes) have sufficiently larger data and are highly correlated with stroke development in clinical trials. When multiple correlated sources are available, Transfer Learning (TL) approaches offer a suitable framework to address small data issue. Most of existing TL works are single transfer approaches including feature transfer instance transfer, network transfer. A recent work proposed a hybrid adapted-embedding method and empirically showed that hybrid transfer outperforms single transfer approaches. Transfer learning is also used in Meta-learning framework for low-resource predictive modeling with patient EHRs. However, existing approaches do not consider the issue of imbalanced labels in the target domain. In contrast, this work proposes a hybrid transfer approach that incorporates generative instance transfer coupled with active selection which can exploit external stroke data to address the label imbalance issue. The generative instance transfer can share high-quality synthetic stroke data for training SRP model while preserving patients' privacy and active instance selection allows the most informative generated instances to be transferred to the target domain. Furthermore, the training and inference of framework are designed in a distributed fashion such that it can take advantage of high data transmission and stringent latency in 5 G/B5G cellular network.

2 Recent Works

2.1 [1] Using Machine Learning to Improve the Prediction of Functional Outcome in Ischemic Stroke Patients, Miguel Monteiro, Ana Catarina Fonseca, 2018.

Ischemic stroke occurs when there is an obstruction of the blood vessels that supply blood to the brain. This medical condition is one of the leading causes of disability and mortality amongst adults worldwide. Despite advances in treatment, around one-third of patients who survive go on to live with long-term disability. Ischemic stroke is a leading cause of disability 12 and death worldwide among adults. The individual prognosis after stroke is extremely dependent on treatment decisions physicians take during the acute phase. In the last five years, several scores such as the ASTRAL, DRAGON and THRIVE have been proposed as tools to help physicians predict the patient functional outcome after a stroke. These scores are rule-based classifiers that use features available when the patient is admitted to the emergency room. In this paper, we apply machine learning techniques to the problem of predicting the functional outcome of ischemic stroke patients, three months after admission. We show that a pure machine learning approach achieves only a marginally superior AUC (0.808 ± 0.085) than that of the best score (0.771 ± 0.056) when using the features available at admission. However, we observed that by progressively adding features available at further points in time, we can significantly increase the AUC to a value above 0.90. We conclude that the results obtained validate the use of the scores at the time of admission, but also point to the importance of using



Article Title: Hybrid Deep Transfer Learning Framework for Stroke Risk Prediction

more features, which require more advanced methods, when possible. The goal of this paper is to use machine learning techniques to predict the functional outcome of a patient three months after the initial stroke. We start by using only the information available at admission to train the classifiers. We then compare the results of the machine learning methods with the results given by the scores, which were designed by domain experts for this specific application. Afterwards, we analyze how the prediction improves as we add more features collected at different points in time after admission. Furthermore, we aim to show how machine learning techniques can be successfully applied to clinical data without losing interpretability of the models, a characteristic which is extremely valuable for medical professionals.

2.2 [2] Revisiting Unreasonable Effectiveness of Data in Deep Learning Era, Chen Sun, Abhinav Shrivastava, 2020.

The success of deep learning in vision can be attributed to: (a) models with high capacity; (b) increased computational power; and (c) availability of large-scale labeled data. Since 2012, there have been significant advances in representation capabilities of the models and computational capabilities of GPUs. But the size of the biggest dataset has surprisingly 13 remained constant. What will happen if we increase the dataset size by $10\times$ or $100\times$? This paper takes a step towards clearing the clouds of mystery surrounding the relationship between ‘enormous data’ and visual deep learning. By exploiting the JFT-300M dataset which has more than 375M noisy labels for 300M images, we investigate how the performance of current vision tasks would change if this data was used for representation learning. Our paper delivers some surprising (and some expected) findings. First, we find that the performance on vision tasks increases logarithmically based on volume of training data size. Second, we show that representation learning (or pretraining) still holds a lot of promise. One can improve performance on many vision tasks by just training a better base model. Finally, as expected, we present new state-of-the-art results for different vision tasks including image classification, object detection, semantic segmentation and human pose estimation. Our sincere hope is that this inspires vision community to not undervalue the data and develop collective efforts in building larger datasets first observation is that large-scale data helps in representation learning as evidenced by improvement in performance on each and every vision task we study. This suggests that collection of a larger-scale dataset to study visual pretraining may greatly benefit the field. Our findings also suggest a bright future for unsupervised or self-supervised representation learning approaches. It seems the scale of data can overpower noise in the label space. Our data has quite a long tail and yet the representation learning seems to work. This long-tail does not seem to adversely affect the stochastic training of ConvNets (training still converges). Finally, our paper presents new state-of-the-art results on several benchmarks using the models learned from JFT-300M. For example, a single model (without any bells and whistles) can now achieve 37.4 AP as compared to 34.3 AP on the COCO detection benchmark.



Article Title: Hybrid Deep Transfer Learning Framework for Stroke Risk Prediction

2.3 [3] Standard Plane Localization in Fetal Ultrasound via Domain Transferred Deep Neural Networks, Hao Chen, Dong Ni, 2016.

Automatic localization of the standard plane containing complicated anatomical structures in ultrasound (US) videos remains a challenging problem. In this paper, we present a learning based approach to locate the fetal abdominal standard plane (FASP) in US videos by constructing a domain transferred deep convolutional neural network (CNN). Compared with 14 previous works based on low-level features, our approach is able to represent the complicated appearance of the FASP and hence achieve better classification performance. More importantly, in order to reduce the overfitting problem caused by the small amount of training samples, we propose a transfer learning strategy, which transfers the knowledge in the low layers of a base CNN trained from a large database of natural images to our task-specific CNN. Extensive experiments demonstrate that our approach outperforms the state-of-the-art method for the FASP localization as well as the CNN only trained on the limited US training samples. The proposed approach can be easily extended to other similar medical image computing problems, which often suffer from the insufficient training samples when exploiting the deep CNN to represent high-level features. Based on the acquired FASP, the clinician can measure abdominal circumference (AC), which is the most important measurement for estimating fetal weight. The accuracy of AC measurement is heavily dependent on both the quality of the FASP and the manual measurement on the FASP by clinicians. Recently, commercial tools have been developed for the automatic AC measurement on several US scanners including Siemens Acuson S2000, GE LOGIQ S8, Mindray DC8, etc. However, little attention has been paid to the prerequisite step, that is, FASP acquisition.

3 Proposed Work

This work proposes a novel Hybrid Deep Transfer Learning based Stroke Risk Prediction framework. The proposed framework can achieve a better ability in establishing SRP model. However, the parameters such as the number of transferred layer and the transferred sequence of different source domains are vital factors for model performance. Common methods such as grid and random search for parameter tuning are often inefficient due to the search space being too large. Bayesian Optimization (BO) is an approach for model-based global optimization of black-box function and the most universally used model for BO is a Gaussian process due to its simplicity and flexibility in constructing a probabilistic model of objective function. Therefore, BO is used to find the best parameter in SRP model.



Article Title: Hybrid Deep Transfer Learning Framework for Stroke Risk Prediction

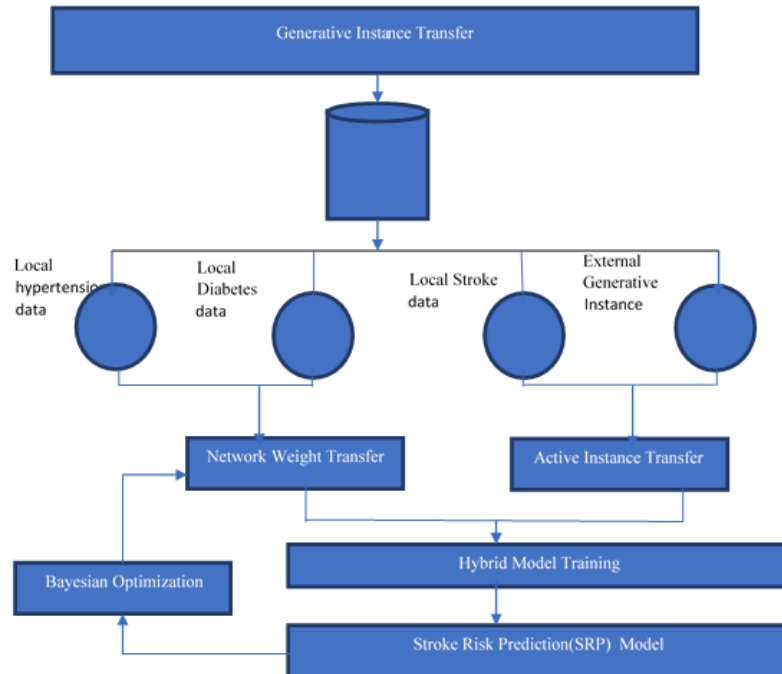


Figure 1: system architecture

4 Conclusion

In this paper has addressed the issues of SRP with small and imbalanced stroke data. We have proposed a novel Hybrid Deep Transfer Learning-based Stroke Risk Prediction (HDTL-SRP) framework which consists of three key components: (1) Generative Instance Transfer (GIT) for making use of the external stroke data distribution among multiple hospitals while preserving the privacy, (2) Network Weight Transfer (NWT) for making use of data from highly correlated diseases (i.e., hypertension or diabetes), (3) Active Instance Transfer (AIT) for balancing the stroke data with the most informative generated instances. It is found that the proposed HDTL-SRP framework outperforms the state-of-the-art SRP models in both synthetic and real-world scenarios.

References

1. Benjamin Emelia J.; Michael J. Blaha; Stephanie E. Chiuve, Year: 2017, "Heart disease and stroke statistics—2017 update a report from the American Heart Association," *Circulation*, Vol: 135, no: 10, pp. e146—e603.
2. Koton Silvia et.al, Year: 2014, "Stroke incidence and mortality trends in US communities, 1987 to 2011," *JAMA*, Vol: 312, no: 3, pp. 259–268.



Article Title: Hybrid Deep Transfer Learning Framework for Stroke Risk Prediction

3. Benjamin Emelia J; Paul Muntner; Alvaro Alonso; Marcio S. Bittencourt; Clifton W. Callaway; April P. Carson; Alanna M. Chamberlain et al, Year: 2019, "Heart disease and stroke statistics—2019 update: A report from the American Heart Association," *Circulation*, Vol: 139, no: 10, pp. e56–e528.
4. M. Monteiro, Year: 2018, "Using machine learning to improve the prediction of functional outcome in ischemic stroke patients," *IEEE/ACM Transaction Computational Biology Bioinformatics*, Vol: 15, no: 6, pp. 1953–1959.
5. S. F. Sung; C. Y. Lin; Y. H. Hu, Year: 2020, "EMR-based phenotyping of ischemic stroke using supervised machine learning and text mining techniques," *IEEE Journal of Biomedical and Health Informatics*, Vol: 24, no: 10, pp. 2922–2931.
6. G. Lim, Year: 2019, "Feature isolation for hypothesis testing in retinal imaging: An ischemic stroke prediction case study," in *Proceeding of AAAI Conference on Artificial Intelligence*, Vol: 33, pp. 9510–9515.
7. S. Cheon; J. Kim; J. Lim, Year: 2019, "The use of deep learning to predict stroke patient mortality," *International Journal of Environmental Research Public Health*, Vol: 16, no: 11.
8. D. Teoh, Year: 2018, "Towards stroke prediction using electronic health records," *BMC Medical Informatics Decision Making*, Vol: 18, no: 1, pp. 1–11.
9. T. Liu; W. Fan; C. Wu, Year: 2019, "A hybrid machine learning approach to cerebral stroke prediction based on imbalanced medical dataset," *Artificial Intelligence in Medicine journal*, Vol: 101, no: 101723.
10. F. Wang; L. P. Casalino; D. Khullar, Year: 2019, "Deep learning in medicine— promise, progress, and challenges," *JAMA International Medicine*, Vol: 179, no: 3, pp. 293–294.
11. C. Sun; A. Shrivastava; S. Singh; A. Gupta, Year: 2017, "Revisiting unreasonable effectiveness of data in deep learning era," in *Proceeding of IEEE International Conference on Computer Vision*, pp. 843–852.
12. O'Brien; C. Rajkumar; C. J. Bulpitt; Year: 1999, "Blood pressure lowering for the primary and secondary prevention of stroke: Treatment of hypertension reduces the risk of stroke," *Journal of Cardiovascular Risk*, Vol: 6, no: 4, pp. 203–205.
13. J. E. Manson, Year: 1991, "A prospective study of maturity-onset diabetes mellitus and risk of coronary heart disease and stroke in women," *Archives Internal Medicine*, Vol: 151, no: 6, pp. 1141–1147.
14. H. Chen, Year: 2015, "Standard plane localization in fetal ultrasound via domain transferred deep neural networks," *IEEE Journal of Biomedical Health Informatics*, Vol: 19, no: 5, pp. 1627–1636.



Article Title: Hybrid Deep Transfer Learning Framework for Stroke Risk Prediction

15. V. Cheplygina; I. P. Peña, J. H. Pedersen; D. A. Lynch, L. Sørensen; M. de Bruijne, Year: 2018
“Transfer learning for multicenter classification of chronic obstructive pulmonary disease,” IEEE
Journal Biomedical Health Informatics, Vol: 22, no: 5, pp. 1486–1496.