

Impact of Deep Learning on Transfer Learning: A Review

Mohammed Jameel Barwary & Adnan Mohsin Abdulazeez

Abstract:

Transfer learning and deep learning approaches have been utilised in several real-world applications and hierarchical systems for pattern recognition and classification tasks. However, in few of the real-world machine learning situations, this presumption does not sustain since there are instances where training data is costly or tough to gather and there is continually a necessity to produce high-performance learners competent with more easily attained data from diverse fields. The objective of this review is to determine more abstract qualities at the greater levels of the representation, by utilising deep learning to detach the variables in the outcomes, formally outline transfer learning, provide information on present solutions, and appraise applications employed in diverse facets of transfer learning and deep learning. This can be attained by rigorous literature exploration and discussion on all presently accessible techniques and prospective research studies on transfer learning solutions of independent as well as big data scale. The conclusions of this study could be an effectual platform directed at prospective directions for devising new deep learning patterns for different applications and dealing with the challenges concerned.

Keywords: *Machine Learning, Transfer Learning, Deep Learning, classifications, Supervised Learning techniques.*



IJSB

Literature Review

Accepted 6 February 2021

Published 24 February 2021

DOI: 10.5281/zenodo.4559668

About Author (s)

Mohammed Jameel Barwary (corresponding author), Duhok Polytechnic University, Duhok, Kurdistan Region, Iraq. Email: mohammed.jameel@uod.ac

Professor Adnan Mohsin Abdulazeez, Duhok Polytechnic University, Duhok, Kurdistan Region, Iraq.

Introduction

There are several applications of Machine Learning (ML), and the most significant one is predictive data mining. Every instance in a dataset utilised by machine learning algorithms is signified through the same suite of attributes. These attributes could be categorical, continuous, or binary. If the correct labels (corresponding output values) are tagged to the instances, learning is supervised; on the other hand, non-labelled instances comprise unsupervised learning (Abdulqader et al., 2020; Bargarai et al., 2020; Jahwar & Abdulazeez, 2020; Salaken et al., 2017; Sulaiman, 2020). Different ML applications comprise tasks which could be established as supervised. The largest advantage of deep learning is its neat depiction of a broader assortment of functions compared to the shallow networks utilised by most conventional learning methodologies. A deeper structure provides better descriptivity compared to a relatively shallow structure comprising identical non-linear units. However, functions having a compact representation using k layers might need the use of the exponential scale if there is a need to express using two layers. Therefore, it may be understood that a network comprising k layers may process functions compactly; however, a system comprising $(k-1)$ layers cannot because the secret unit number is infinitely large (Sudre et al., 2017). Several aspects, like parallel CPU architectures, quicker CPUs, and GPU computing, have facilitated deep network training and turned it computationally viable. It is typical for neural networks to use a matrix of weight vectors that can be multiplied rapidly using GPUs because they are designed for such operations (Lane et al., 2016)

Numerous researchers have proposed different solutions over the previous few decades for automatic screening and classification of cancer using breast cytology images. In view of this, many academics have been exploring nucleus analysis by mining nucleus features to offer considerable details for cell categorisation as malignant and benign (Toprak, 2018; Zebari et al., 2020). Likewise, clustering linked algorithms coupled with circular Hough Transform and several statistical attributes are frequently utilised for cluster segmentation and categorisation (Tiwari, Srivastava, & Pant, 2020). Within medical image studies, algorithms for histopathological images are developing swiftly. Nevertheless, there is a great demand for an automated technique to attain effectual and extremely dependable outcomes (Kong et al., 2020; Zebari et al., 2020). Hence, such methods are required so that qualitative diagnostics are conducted in the precisely to ensure that the results are accurate. The dynamic presence of activities like segmentation, reprocessing, and mining of attributes in orthodox machine learning approaches worsens the device's performance with regards to output and precision. For addressing the complications of orthodox machine learning methodologies, the deep learning norm has been fostered so as to obtain the relevant knowledge from raw images and permit effectual usage for the classification procedure (Affonso et al., 2017; Sulaiman, 2020) Deep learning requires no modifications of the functions; rather, a general learning process is employed for processing the data sets (Kim et al., 2019). In the recent past, deep learning that is focused on multiple algorithms has attained substantial success in the domain of biomedical image processing, like determination of mitosis cells from microscopic images (Janowczyk & Madabhushi, 2016; Othman & Zeebaree, 2020), neural membrane segmentation (Saha & Chakraborty, 2018) and skin ailment categorisation (Sun et al., 2016). Although the deep learning applications function properly on broad data sets, they do not attain considerable gains with small data sets.

Deep learning-based neural network frameworks benefit from the transfer learning principle that may be employed to enhance recognition accuracy and reduce computational requirements by integrating expertise (Park et al., 2016). In this context, generic image information is used for gathering feature information. Subsequently, pre-trained deep-

learning models are integrated with relatively small and highly specific data sets (Abdulazeez et al., 2020; Espejo-Garcia et al., 2020). Context-specific learning provides a different learning transition path where the CNN is trained in two parts. This process has produced acceptable results for overlapping or single patches in identifying and classifying breast tissue (Jouni et al., 2016). Transfer learning performance can be enhanced by integrated several CNN architectures; traditional learning methods can be replaced using the updated technique, as depicted in Figure 1. Along the same lines, cell-specific image classification can be simplified and enhanced by using a combination of Inception V2, Inception V3, and ResNet 50 models trained using ImageNet (Abdulazeez et al., 2020; Khan et al., 2019).

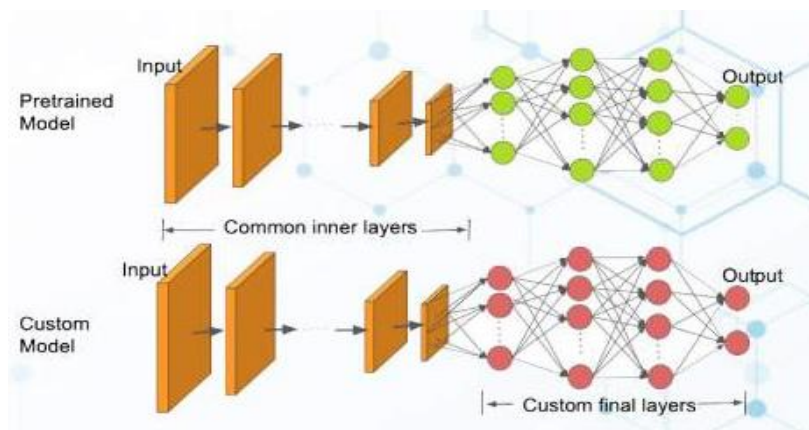


Figure 1: Traditional single model CNN architecture in deep learning (H. Chen et al., 2020)

Transfer learning is a specific kind of machine learning wherein just a portion of the test data is utilised for reproducing the test data. Past studies showed that there are three generic questions about alterations to visual learning issues: 1) when to transfer, 2) where to transfer, and 3) how to transfer. Second, it is vital to ascertain if the information transfer is apt for specific tasks and if the training is pertinent to the new activity. In training circumstances, outstanding progress could be achieved when satisfactory and remarkable input is obtainable. Gathering information using source domain descriptors leads to contrasting information concerning the target (Long et al., 2017). The area where source instances are distributed with the corresponding labels is preferred first. The second preference is given to the case where source instances are passed but the source frameworks that describe information movement concerning instance attributes to the target are used for enhancing target domain performance.

Transfer learning approaches are a subdivision of the extensive range of methodologies and are therefore more vital. Different means of transfer learning have been recommended wherein the process of transfer learning is restricted with regards to dimensionality (Shin et al., 2016; Zebari et al., 2020; Zeebaree et al., 2020). Conventional machine learning frameworks and techniques are used for information exchange. It is feasible to use the correct training data to classify the test sample in the case of conventional machine learning.

Deep learning systems are capable of extracting valuable information from complicated big-data based solutions (Adeen et al., 2018). The deep learning domain has become an active research domain recently; it was first presented in 2006 and has since been of much interest (Shrestha & Solomatine, 2006). The first works on deep learning date as far back as the 1940s. It should be noted that training multi-layer neural networks using conventional training techniques produces locally optimal results or does not converge. Consequently, multi-layer neural networks have not been researched immensely even though they have

better potential at representation learning and performance requirement. Hinton et al. (Ba et al., 2016) suggested a two-stage learning technique in 2016. This technique comprised pre-training and fine-tuning steps so that effective deep learning can be facilitated; this was one major breakthrough in the deep learning domain. In this context, the present study also evaluates the present status of deep transfer learning and suggests directions for additional research in this domain. Furthermore, there is a mention of the background concerning the present state-of-the-art research in this area and a mention about the areas where the literature could be augmented. The discussion also includes the benefits and drawbacks of deep transfer learning so that we can suggest additional areas for future research.

2. RELATED WORK

This analysis aims to address the extensive subject of in-depth learning concerning the movement of scattered information in one paper. This work is new and it builds upon the works of leading academicians who have worked on insightful learning. Several studies emphasise particular domains but omit a comprehensive view of the subject (Weiss et al., 2016). The evaluation comprises an assessment of several deep learning frameworks, methods, optimisation schemes, limitations, recent uses, and application areas.

2.1. Interests on Transfer Learning

The deep learning process is explored by presenting the collective opinion expressed by leading researchers. Evaluating studies (Gao et al., 2014) concerning specific fields allows addressing only a limited part of this wide market. This discussion comprises deep learning network structures, techniques, challenges, and optimisation methods. For instance, playing the keyboard allows an individual to learn from previous experience and apply it to similar areas. Numerous approaches have been formulated to address transfer learning. Widely-used approaches comprise correcting conditional source gap, minimal source variance, or both. Input space alignment pertaining to the source and target domains is used for attempting training of heterogenous-dataset based frameworks (Fang et al., 2015). It is vital to repeat domain adaptation so that optimal matching can be ensured. Information sharing is a critical aspect of the transfer learning scheme, which has four divisions. Group 1 comprises the use of instances for learning. This can be addressed by re-evaluating elements to provide a more refined representation of values in the source domain. The target is then trained using the revised elements (examples in Huang (Huang et al., 2017; Jiang et al., 2017)). The ideal situation is when the two domains have similar conditional distribution. Functionality is the second transformation aspect. Models using features are divided into two classes. The first step uses the mapping function, also called as the transition function, which changes the input features to those of the target domain (e.g., pan) (Yang et al., 2020). The second phase comprises splitting the elements into meaningful groups. The resulting sets are used for predicting future inputs (Zhuang et al., 2020). Figure 2 depicts successful transfer learning. The transformation classifier is trained using labelled source input and target. The transfer learning framework uses unlabelled targets having maximum information; subsequently, the sample labels are produced as output.

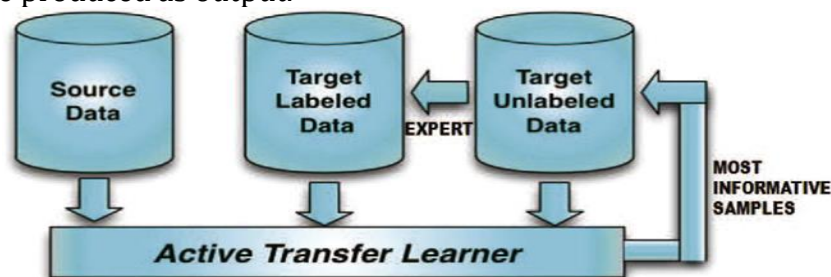


Figure 2: Active transfer learning framework (Zhao et al., 2017)

This is stated as the symmetric transformation function and is depicted in Figure 2. The third transfer variant is the transfer of information by means of common parameters concerning the learning frameworks of the source and target or using the formulation of several source learning frameworks followed by an appropriate combination of reweighted (ensemble) learners to produce a better target learner (examples in (Chen et al., 2020; Segev et al., 2016; Wei et al., 2018)). The least-used transfer approach is information movement considering a specified association between the source and target domains (Chen et al., 2020).

2.2. Negative Transfer Learning

Transfer learning is the high-level concept used for improving the formulated learner using data from an associated source domain. In this scenario, the goal learner would be unfavourably impacted by this frail interaction, which is called as a negative transition. Within a big data setting, there could be a huge data gathering wherein just a subset of the data is pertinent to the intended field of interest. In this instance, it is essential to divide the dataset into diverse sources and employ negative transfer approaches by utilising a transfer learning algorithm. In a circumstance wherein many datasets are reachable, which at first tend to be associated with the target field of interest, it is advisable to select the datasets which present the highest diffusion of knowledge and remove datasets which cause negative transfer. This allows total use of the available huge databases. The domain of negative transition has not been meticulously examined; however, the following papers have started dealing with this concern. A study by Rosenstein deals with the standard of negative transfer in transfer learning and recommends that the source domain should be sufficiently linked to the target domain; else, an attempt to relocate information from the source could exert an undesirable impact on the target learner. Cases of negative transition was shown by Rosenstein (Rosenstein et al., 2005) in studies using the hierarchical classifier Naive Bayes. The researcher also shows that the likelihood of a detrimental transition decreases as the quantity of goal training samples is raised. Eaton's paper (Eaton & Lane, 2008) proposes to create a target learner based on a transferability metric from several linked source domains. First, the technique builds a logistic regression learner for every source domain. A model transition graph is constructed for signifying the transferability of every source learner. In this situation, learner transferability from the first to the second is considered as the second learner's performance using learning information from the first minus its performance without learning from the first learner. Subsequently, the source and target learners are assessed to extract transferability steps that are used to update the Model Transition Map. Spectral graph theory (Chen et al., 2020) is applied to the model transfer graph to obtain a transfer function that retains model transfer graph geometry. The final target learner employs it to assess source-specific transfer level.

Experiments are conducted in the categorisation of records and alphabets. Source domains are outlined, which are either associated or not associated with the target domain. The Eaton technique (Eaton & Lane, 2008) is assessed using a custom method with manually-selected source domains related to the target. A technique uses the average of all present sources and a baseline technique free of transfer learning. Experiments are evaluated using the classification accuracy metric. A function describing a homogenous input space are used to depict the source and destination domains. The results of the tests are pooled together. Generally, the Eaton (Eaton & Lane, 2008) approach performs the best; however, there are instances performed worse than the handpicked, average, and baseline approaches. When deploying the algorithm, the transferability computation between two sources should be the same; furthermore, the transferability from source 1 to source 2 is not essentially equivalent to the transferability from source 2 to source 1. We suggest that future works use directed

graphs to ascertain if bidirectional transferability exists between the two outlets. It is indicated that source data accuracy, target data quality, choice of deep learning technique, and domain divergence are critical factors for addressing negative transition.

2.3. Transfer Learning Applications

The review indicates that the transfer of learning has been applied to several real-world structures. Natural language processing has many applications like document classification, multi-language text classification, feeling classification, and spam email detection. These processes comprise classifying films, photographs, documents, and other artefacts. Applications deliberated earlier comprise muscle fatigue classification, Wi-Fi position classification, human behaviour classification, medication effectiveness classification, machine defect classification, and heart arrhythmia classification (Shao et al., 2014). Most solutions reviewed were general, i.e., the technique could be swiftly applied to a wider assortment of applications. Application-oriented technologies usually focus on image processing and natural language processing. There are a variety of transfer learning solutions relevant to the application of the suggestion systems. Recommendation platforms offer users grades or ratings for a particular field (e.g., books, movies). Nonetheless, with only limited historical instances (epidemiological data) to establish its forecasts, the algorithm lacks dependency. In case there is no adequate domain data for dependable forecasts (for instance, a movie released recently), utilise information from the other domain (for instance, using books). Concerns were presented by utilising transfer learning methodologies and papers (Li et al., 2019).

2.4. Recurring Themes of Deep Learning for Supervised

Dynamic Programming (DP) was introduced in 1957, and it has been consistently associated with deep learning. It can facilitate the enhancement of credit assignment considering specific presuppositions (O'Donoghue, Osband, Munos, & Mnih, 2018). Considering the case of neural networks trained using supervised learning, backpropagation may be employed as a process stemming from deep learning. Traditional RLs emphasising strict Markovian presuppositions benefit from techniques derived using deep learning. These techniques facilitate a noteworthy reduction in deep learning algorithm complexity. Such algorithms are significant for processes or graphics frameworks based on NN fundamentals, such as Secret Markov Models (HMMs) (Ghosh-Dastidar & Adeli, 2009; Hastie et al., 2009; Hinton et al., 2012; Yu et al., 2012). Transfer learning may also be used for classifying aerosol particles suspended in air; this classification facilitates the betterment of global environmental frameworks. The Tr AdaBoost technique is employed and used with SVM classifiers to build upon classification success. Ascertaining low-income areas of emerging nations is vital to food security, humanitarian endeavours, and sustainable progress. Wang et al. formulated a technique (Wang & Saligrama, 2012) matching that of Ribeiro (Ribeiro et al., 2016). The technique proposes the use of convolutional neural networks to estimate poverty. The first prediction framework is trained to forecast light from night-based images. Subsequently, night-time light intensity data is mapped at the source. A deep learning algorithm is utilised for enhancing diagnostics. The rule-dependent learning scheme is prepared to utilise abstract source domain information for modelling distinct types of data pertaining to gene expression.

Video advertisements displayed online are forecast using transfer learning methods; this rapidly growing industry relies on such techniques. A transfer learning technique proposed by (Oquab et al., 2014)) has several source classifiers having weighted outputs to obtain an augmented aim classifier trained to forecast the results of targeted online advertisements. Kan et al. researched facial recognition, where face data belonging to one set is utilised to

prepare a classifier for other classifications (Khan et al., 2019). The paper describes a system formulated for recognising sign language. The technique comprises training to identify several signs captured from different angles. The Widmer study evaluates the use of transfer learning in the genetics domain (Widmer & Rätsch, 2012). Genome splicing sites are modelled using multi-task learning techniques. In another case, data gathered from several hospitals are used to model the infection rate for other hospitals. Romera-Paredes researched a multi-task transfer learning approach that aimed to determine pain level using facial expressions processed using a system trained by labelled facial data of other people (Romera-Paredes, Aung, Bianchi-Berthouze, & Pontil, 2013). Furthermore, Deng et al. used transition learning to identify emotion using speech by employing labelled speech data (Deng et al., 2018). Zhang worked a model that determines wine quality by processing data using a multi-task transfer learning scheme (Ying, Zhang, Huang, & Yang, 2018). The Cook Survey study (Zhang et al., 2020) evaluates the use of transfer learning for behaviour recognition. Shao (Shao et al., 2018) and Patel (Perera & Patel, 2019) researched the use of transfer learning in the image recognition domain. The current study assesses several developments related to transfer learning based on computational deep-learning. We assert that the computational deep learning techniques utilised for transfer learning have real-world applications. Table 1 presents a summary of transfer learning based on computational deep-learning.

Table 1: : Review of previous studies on deep learning in transfer learning

| AUTHOR | AUTHOR YEAR | TITLE | METHOD | OUTCOME |
|---|-------------|--|--|---|
| (CIRILLO, MIRDELL, SJÖBERG, & PHAM, 2019) | 2019 | Time-Independent Prediction of Burn Depth using Deep Convolutional Neural Networks | Uses deep learning models to assess given Colour images of four types of burn depth injured in first few days, including normal skin and background, acquired by a TiVi camera using four pretrained deep CNNs: VGG-16, Google Net, ResNet-50, and ResNet-101. | The finest 10-fold cross-validation outcomes attained from ResNet-101 with an average, minimum, and maximum precision are 81.66%, 72.06%, and 88.06%, respectively; and the average accuracy, sensitivity, and specificity for the four diverse kinds of burn depth are 90.54%, 74.35%, and 94.25%, respectively. |
| (G. ZHOU, XIE, HUANG, & HE, 2016) | 2016 | Bi-Transferring Deep Neural Networks for Domain Adaptation | Uses Bi-Transferring Deep Neural Networks (BTDDNs) to transfer the source domain examples to the target domain, and also transfer the target domain examples to the source domain. | The recommended methodology considerably outclasses the numerous baseline approaches and attains a precision that is competitive with the state-of-the-art technique for domain adaptation. |
| (GLOROT, BORDES, & BENGIO, 2011) | 2011 | Domain Adaptation for Large-Scale Sentiment Classification: A Deep Learning Approach | Recommends a deep learning methodology that learns to mine a meaningful depiction for every review in an unsupervised manner. | This high-level feature delineation used for training sentiment classifiers provides superior performance than the present state-of-the-art. Four different types of products from Amazon were benchmarked for the test. Moreover, this method has superior scalability that allows successful adaptation on an industry-scale dataset comprising twenty-two domains. |
| (KAUR & GANDHI, 2020) | 2020 | Deep convolutional neural networks with transfer learning for automated brain image classification | Recommends a deep learning framework which pools residual thought and dilated convolution to diagnose and identify childhood pneumonia. | Experimental data specific to the test set for pneumonia classification in children indicates a recall of 96.7% and an F1-score of 92.7% |
| (WEN, GAO, DONG, & | 2019 | A negative correlation ensemble transfers learning method for fault | This study used a novel negative correlation ensemble transfer learning technique (NCTE) on the ResNet-50 and to formulate a deep learning model comprising 50 layers. | The KAT Bearing Dataset is used for testing the prediction accuracy using for NCTE; the accuracy was 98.73%. These results show that NCTE has achieved a good result compared with other machine learning and |

| | | | | |
|-----------------------------|------|---|--|---|
| ZHU, 2019) | | diagnosis based on convolutional neural network | Cross-validation is used to ascertain the hyper-parameters of the NCTE. | deep learning method. |
| (Jiao ET AL., 2020) | 2020 | End-to-end prediction of weld penetration: A deep learning and transfer learning based method | Suggests a deep learning technique that has an end-to-end flow for forecasting weld penetration using top-side weld pictures. The top and back-bead are concurrently monitored using two cameras that feed a passive vision sensing mechanism. | Tests indicate classification accuracy as high as 92.70%. A residual neural network (ResNet) based transfer learning method was used to enhance training speed and accuracy. Empirical data suggests that prediction accuracy can be enhanced to 96.35% while effecting a decrease in training time. |
| (DINH, CHU, & NGUYEN, 2015) | 2015 | Transfer learning in genetic programming | Numerous transfer learning techniques were suggested using the foundation of genetic programming (GP). The techniques were executed by moving several good elements and sub-elements from source to target | Genetic programming techniques augmented using transfer learning techniques can have reduced training errors. GP using transfer learning provided better real-world results on new data. Moreover, the effect of transfer learning for assessing GP code bloat indicated that code growth could be regulated by capping the upper limit of transferred individual size. |
| (ABUBAKA ET AL., 2020) | 2020 | Assessment of Human Skin Burns: A Deep Transfer Learning Approach | Employs deep-learning-based transfer learning using the VGG16 and ResNet50 frameworks. Image patterns were extracted from a dataset comprising 2080 RGB uniformly distributed pictures depicting healthy skin, first, second, and third-degree burns. | The recommended methodology yields maximum prediction precision of 95.43% using ResFeat50 and 85.67% using VggFeat16. The average recall, precision and F1-score are 95.50%, 95.50%, 95.50% and 85.75%, 86.25%, 85.75% for both ResFeat50 and VggFeat16, respectively. |
| (ZHOU, ET AL., 2014) | 2014 | Hybrid Heterogeneous Transfer Learning through Deep Learning | Employs the novel Hybrid Heterogeneous Transfer Learning (HHTL) for processing source- or target-biased instances across several domains. Explicitly, the study suggests a deep learning technique for learning to map heterogeneous features from several domains. Additionally, the study aims to have improved feature representation of the mapped instances to lessen the bias arising due to multiple domains. | More layers are associated with higher performance. The results provided by SVM-SC are comparable to single-layer HHTL, where high-level learning does not provide adequate usefulness. Nevertheless, the multiple-language classification performance of HHTL can be enhanced by increasing the number of layers. The resulting framework can reduce bias and provide higher-level features. |
| (LEE ET AL., 2018) | 2018 | Clean Net: Transfer Learning for Scalable Image Classifier Training with Label Noise | The study proposes Clean Net. It is a joint embedding neural network that produces information concerning label noise that can be applied to other classes.. | Clean Net can facilitate 41.5% label noise error reduction concerning held-out classes devoid of human supervision as compared to the presently available weakly-supervised techniques. Image verification performance is enhanced by 47%; approximately 3.2% of the images were considered classification tasks. |

3. Discussion

In terms of the data and model perspectives, the processes and methodologies implemented for the transition of learning have been summarised. Transfer learning applications have since been executed and put to test in several research projects. It was obvious that transfer learning had made noteworthy advances in varied implementations and explorations in an assortment of domains and activities. Although in the face of real research assignments, specific challenges or issues may have possibly been resolved. Some of these concerns might have been mended or alleviated, even though others might not have been addressed. Of late, a

new method known as self-supervised learning has surfaced. Self-supervised learning has the potential to create labels from scratch with non-labelled original data without any human annotation through the design and execution of certain artificial tasks that do not have actual use. For instance, in the case of self-supervised learning, comparative location of patches and prediction of the rotation angle in the picture are the two most common and purposeful tasks (Huang et al., 2020). The automatically generated labels that do not require any human intervention can be acquired using this contrived learning process. Moreover, smart imagery is another methodology that has been used to acquire improved quality image data. It can enable reduction of noise and artefacts, improvement of image resolution, and detection of shadows (Huang et al., 2019). All these influences allow deep learning algorithms to deliver greater precision and swifter procedures.

For many years, a pre-trained coevolutionary neural network has been used in several experiments as a feature extractor for transfer learning, along with the inclusion of other deep learning approaches. Adversarial network and training processes were implemented for adversarial-based unsupervised domain adaptation, which eventually evolved to be a generative adversarial network (GAN). This serves as a characteristic example of the combination of the deep learning models and transfer learning. In reality, many researchers have attempted to execute a transfer learning strategy in order to enhance learning (C. Huang et al., 2018). It is the same that though few papers of medical image on this field have been published, we believe they will be in near future. As highlighted in the earlier relevant works, it is important to note that the efficacy of specific applications and algorithms is not optimal. One of the explanations for this specifies that the default parameter settings present in the original algorithms could possibly not be adequate for the data set that is selected. For example, GFK was originally developed for usage in object recognition, so that it could enable direct integration into the text classification (Jiao et al., 2020), which eventually resulted in an unsatisfactory result (with an average accuracy of 62%). These outcomes indicate that particular algorithms might not be applicable for data sets in such domains. Hence, it is essential to select the appropriate algorithms as the foundation for the research process. Additionally, for functional applications, it is further necessary to identify an effectual algorithm.

4. Conclusion

The inference is that for the diagnosis of machine learning, the impact of deep learning in deep transfer learning has been recorded. Both the academia and the industry have expressed increased interest in intelligent data-driven diagnostic approaches. Several machine learning algorithms have been applied to forecast machine life, track condition, and detect defects. Based on these milestones, deep transfer learning has become the central theme of deep learning diagnostics science. Transfer learning techniques, for example, feature-based, mutual parameter-based, and instance-based have been applied frequently for deep learning diagnostics. Different transfer learning architectures have been created for a range of applications. This paper has focused on the recent advances and various other criteria in deep transfer learning. While the efficacy of these approaches has not been adequately tested, it is still valid to conclude that deep-seated learning has already attained the testing limits of deep-seated learning diagnostics.

The diverse range of data collection makes it more necessary for heterogeneous transfer learning solutions to upgrade. Larger data collection sizes indicate the potential for massive data systems to be applied along with the existing transfer learning solutions. A vital area for potential study includes the variety and scale of data sets that are used to migrate learning

systems. Another field of study involves a scenario where the performance label space differs across realms. As new data sets are being recorded and made accessible, this topic could be an important field of study in the future. To summarise, there are comparatively few transfer learning strategies captured in literature that elaborate the unlabelled source and unmarked goal data situation, and this area holds a lot of potential for extended study.

References

- Abdulazeez, A., Salim, B., Zeebaree, D., & Doghramachi, D. (2020). Comparison of VPN Protocols at Network Layer Focusing on Wire Guard Protocol.
- Abdulazeeza, A. M., Nahmatwllab, L. L., & Qader, D. (2020). Pipelined Parallel Processing Implementation based on Distributed Memory Systems. *International Journal of Innovation*, 13(7), 12.
- Abdulqader, D. M., Abdulazeez, A. M., & Zeebaree, D. Q. (2020). Machine Learning Supervised Algorithms of Gene Selection: A Review. *Machine Learning*, 62(03).
- Abubakar, A., Ugail, H., & Bukar, A. M. (2020). Assessment of human skin burns: A deep transfer learning approach. *Journal of Medical and Biological Engineering*, 40(3), 321-333.
- Adeen, N., Abdulazeez, M., & Zeebaree, D. Systematic Review of Unsupervised Genomic Clustering Algorithms Techniques for High Dimensional Datasets. In: vol.
- Affonso, C., Rossi, A. L. D., Vieira, F. H. A., & de Leon Ferreira, A. C. P. (2017). Deep learning for biological image classification. *Expert Systems with Applications*, 85, 114-122.
- Ba, J. L., Kiros, J. R., & Hinton, G. E. (2016). Layer normalization. *arXiv preprint arXiv:1607.06450*.
- Bargarai, F., Abdulazeez, A., Tiryaki, V., & Zeebaree, D. (2020). Management of Wireless Communication Systems Using Artificial Intelligence-Based Software Defined Radio.
- Chen, C.-L., Hsu, Y.-C., Yang, L.-Y., Tung, Y.-H., Luo, W.-B., Liu, C.-M., . . . Tseng, W.-Y. I. (2020). Generalization of diffusion magnetic resonance imaging-based brain age prediction model through transfer learning. *Neuroimage*, 217, 116831.
- Chen, H., Chen, A., Xu, L., Xie, H., Qiao, H., Lin, Q., & Cai, K. (2020). A deep learning CNN architecture applied in smart near-infrared analysis of water pollution for agricultural irrigation resources. *Agricultural Water Management*, 240, 106303.
- Chen, M., Zhao, S., Liu, H., & Cai, D. (2020). *Adversarial-learned loss for domain adaptation*. Paper presented at the Proceedings of the AAAI Conference on Artificial Intelligence.
- Chen, Y., Qin, X., Wang, J., Yu, C., & Gao, W. (2020). Fedhealth: A federated transfer learning framework for wearable healthcare. *IEEE Intelligent Systems*, 35(4), 83-93.
- Cirillo, M. D., Mirdell, R., Sjöberg, F., & Pham, T. D. (2019). Tensor decomposition for colour image segmentation of burn wounds. *Scientific reports*, 9(1), 1-13.
- Deng, C., Xue, Y., Liu, X., Li, C., & Tao, D. (2018). Active transfer learning network: A unified deep joint spectral-spatial feature learning model for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 57(3), 1741-1754.
- Dinh, T. T. H., Chu, T. H., & Nguyen, Q. U. (2015). *Transfer learning in genetic programming*. Paper presented at the 2015 IEEE Congress on Evolutionary Computation (CEC).
- Eaton, E., & Lane, T. (2008). *Modeling transfer relationships between learning tasks for improved inductive transfer*. Paper presented at the Joint European Conference on Machine Learning and Knowledge Discovery in Databases.
- Espejo-Garcia, B., Mylonas, N., Athanasakos, L., & Fountas, S. (2020). Improving weeds identification with a repository of agricultural pre-trained deep neural networks. *Computers and Electronics in Agriculture*, 175, 105593.
- Fang, M., Guo, Y., Zhang, X., & Li, X. (2015). Multi-source transfer learning based on label shared subspace. *Pattern Recognition Letters*, 51, 101-106.
- Gao, J., Ling, H., Hu, W., & Xing, J. (2014). *Transfer learning based visual tracking with gaussian processes regression*. Paper presented at the European conference on computer vision.
- Ghosh-Dastidar, S., & Adeli, H. (2009). A new supervised learning algorithm for multiple spiking neural networks with application in epilepsy and seizure detection. *Neural networks*, 22(10), 1419-1431.

- Glorot, X., Bordes, A., & Bengio, Y. (2011). *Domain adaptation for large-scale sentiment classification: A deep learning approach*. Paper presented at the ICML.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). Unsupervised learning. In *The elements of statistical learning* (pp. 485-585): Springer.
- Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A.-r., Jaitly, N., . . . Sainath, T. N. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal processing magazine*, 29(6), 82-97.
- Huang, C., Lan, Y., Zhang, G., Xu, G., Jiang, L., Zeng, N., . . . Han, N. (2020). A New Transfer Function for Volume Visualization of Aortic Stent and Its Application to Virtual Endoscopy. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 16(2s), 1-14.
- Huang, C., Tian, G., Lan, Y., Peng, Y., Ng, E., Hao, Y., . . . Che, W. (2019). A new pulse coupled neural network (PCNN) for brain medical image fusion empowered by shuffled frog leaping algorithm. *Frontiers in neuroscience*, 13, 210.
- Huang, C., Xie, Y., Lan, Y., Hao, Y., Chen, F., Cheng, Y., & Peng, Y. (2018). A new framework for the integrative analytics of intravascular ultrasound and optical coherence tomography images. *IEEE Access*, 6, 36408-36419.
- Huang, L., Ji, H., Cho, K., & Voss, C. R. (2017). Zero-shot transfer learning for event extraction. *arXiv preprint arXiv:1707.01066*.
- Jahwar, A. F., & Abdulazeez, A. M. (2020). META-HEURISTIC ALGORITHMS FOR K-MEANS CLUSTERING: A REVIEW. *PalArch's Journal of Archaeology of Egypt/Egyptology*, 17(7), 12002-12020.
- Janowczyk, A., & Madabhushi, A. (2016). Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *Journal of pathology informatics*, 7.
- Jiang, M., Huang, Z., Qiu, L., Huang, W., & Yen, G. G. (2017). Transfer learning-based dynamic multiobjective optimization algorithms. *IEEE Transactions on Evolutionary Computation*, 22(4), 501-514.
- Jiao, W., Wang, Q., Cheng, Y., & Zhang, Y. (2020). End-to-end prediction of weld penetration: A deep learning and transfer learning based method. *Journal of Manufacturing Processes*.
- Jouni, H., Issa, M., Harb, A., Jacquemod, G., & Leduc, Y. (2016). *Neural Network architecture for breast cancer detection and classification*. Paper presented at the 2016 IEEE International Multidisciplinary Conference on Engineering Technology (IMCET).
- Kaur, T., & Gandhi, T. K. (2020). Deep convolutional neural networks with transfer learning for automated brain image classification. *Machine Vision and Applications*, 31(3), 1-16.
- Khan, R. A., Meyer, A., Konik, H., & Bouakaz, S. (2019). Saliency-based framework for facial expression recognition. *Frontiers of Computer Science*, 13(1), 183-198.
- Khan, S., Islam, N., Jan, Z., Din, I. U., & Rodrigues, J. J. C. (2019). A novel deep learning based framework for the detection and classification of breast cancer using transfer learning. *Pattern Recognition Letters*, 125, 1-6.
- Kim, H., Ahn, E., Shin, M., & Sim, S.-H. (2019). Crack and noncrack classification from concrete surface images using machine learning. *Structural Health Monitoring*, 18(3), 725-738.
- Kong, L., Li, C., Ge, J., Zhang, F., Feng, Y., Li, Z., & Luo, B. (2020). Leveraging multiple features for document sentiment classification. *Information Sciences*, 518, 39-55.
- Lane, N. D., Bhattacharya, S., Georgiev, P., Forlivesi, C., Jiao, L., Qendro, L., & Kawsar, F. (2016). *DeepX: A software accelerator for low-power deep learning inference on mobile devices*. Paper presented at the 2016 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN).
- Lee, K.-H., He, X., Zhang, L., & Yang, L. (2018). *Cleannet: Transfer learning for scalable image classifier training with label noise*. Paper presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.
- Li, X., Xiong, H., Wang, H., Rao, Y., Liu, L., Chen, Z., & Huan, J. (2019). Delta: Deep learning transfer using feature map with attention for convolutional networks. *arXiv preprint arXiv:1901.09229*.
- Long, M., Zhu, H., Wang, J., & Jordan, M. I. (2017). *Deep transfer learning with joint adaptation networks*. Paper presented at the International conference on machine learning.

- O'Donoghue, B., Osband, I., Munos, R., & Mnih, V. (2018). *The uncertainty bellman equation and exploration*. Paper presented at the International Conference on Machine Learning.
- Oquab, M., Bottou, L., Laptev, I., & Sivic, J. (2014). *Learning and transferring mid-level image representations using convolutional neural networks*. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.
- Othman, G., & Zeebaree, D. Q. (2020). The Applications of Discrete Wavelet Transform in Image Processing: A Review. *Journal of Soft Computing and Data Mining*, 1(2), 31-43.
- Park, H. J., Lee, S. Y., Park, N. H., Shin, H. G., Chung, E. C., Rho, M. H., . . . Kwon, H. J. (2016). Modified thoracolumbar injury classification and severity score (TLICS) and its clinical usefulness. *Acta Radiologica*, 57(1), 74-81.
- Perera, P., & Patel, V. M. (2019). Learning deep features for one-class classification. *IEEE Transactions on Image Processing*, 28(11), 5450-5463.
- Ribeiro, E., Uhl, A., Wimmer, G., & Häfner, M. (2016). Exploring deep learning and transfer learning for colonic polyp classification. *Computational and mathematical methods in medicine*, 2016.
- Romera-Paredes, B., Aung, H., Bianchi-Berthouze, N., & Pontil, M. (2013). *Multilinear multitask learning*. Paper presented at the International Conference on Machine Learning.
- Rosenstein, M. T., Marx, Z., Kaelbling, L. P., & Dietterich, T. G. (2005). *To transfer or not to transfer*. Paper presented at the NIPS 2005 workshop on transfer learning.
- Saha, M., & Chakraborty, C. (2018). Her2net: A deep framework for semantic segmentation and classification of cell membranes and nuclei in breast cancer evaluation. *IEEE Transactions on Image Processing*, 27(5), 2189-2200.
- Salaken, S. M., Khosravi, A., Nguyen, T., & Nahavandi, S. (2017). Extreme learning machine based transfer learning algorithms: A survey. *Neurocomputing*, 267, 516-524.
- Segev, N., Harel, M., Mannor, S., Crammer, K., & El-Yaniv, R. (2016). Learn on source, refine on target: A model transfer learning framework with random forests. *IEEE transactions on pattern analysis and machine intelligence*, 39(9), 1811-1824.
- Shao, K., Zhu, Y., & Zhao, D. (2018). Starcraft micromanagement with reinforcement learning and curriculum transfer learning. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 3(1), 73-84.
- Shao, L., Zhu, F., & Li, X. (2014). Transfer learning for visual categorization: A survey. *IEEE transactions on neural networks and learning systems*, 26(5), 1019-1034.
- Shin, H.-C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., . . . Summers, R. M. (2016). Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE transactions on medical imaging*, 35(5), 1285-1298.
- Shrestha, D. L., & Solomatine, D. P. (2006). Machine learning approaches for estimation of prediction interval for the model output. *Neural networks*, 19(2), 225-235.
- Sudre, C. H., Li, W., Vercauteren, T., Ourselin, S., & Cardoso, M. J. (2017). Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In *Deep learning in medical image analysis and multimodal learning for clinical decision support* (pp. 240-248): Springer.
- Sulaiman, M. A. (2020). Evaluating Data Mining Classification Methods Performance in Internet of Things Applications. *Journal of Soft Computing and Data Mining*, 1(2), 11-25.
- Sun, X., Yang, J., Sun, M., & Wang, K. (2016). *A benchmark for automatic visual classification of clinical skin disease images*. Paper presented at the European Conference on Computer Vision.
- Tiwari, A., Srivastava, S., & Pant, M. (2020). Brain tumor segmentation and classification from magnetic resonance images: Review of selected methods from 2014 to 2019. *Pattern Recognition Letters*, 131, 244-260.
- Toprak, A. (2018). Extreme learning machine (elm)-based classification of benign and malignant cells in breast cancer. *Medical science monitor: international medical journal of experimental and clinical research*, 24, 6537.
- Wang, J., & Saligrama, V. (2012). Local supervised learning through space partitioning. *Advances in neural information processing systems*, 25, 91-99.
- Wei, Y., Pan, X., Qin, H., Ouyang, W., & Yan, J. (2018). *Quantization mimic: Towards very tiny cnn for object detection*. Paper presented at the Proceedings of the European conference on computer vision (ECCV).

- Weiss, K., Khoshgoftaar, T. M., & Wang, D. (2016). A survey of transfer learning. *Journal of Big data*, 3(1), 1-40.
- Wen, L., Gao, L., Dong, Y., & Zhu, Z. (2019). A negative correlation ensemble transfer learning method for fault diagnosis based on convolutional neural network. *Math. Biosci. Eng*, 16(5), 3311-3330.
- Widmer, C., & Rätsch, G. (2012). *Multitask learning in computational biology*. Paper presented at the Proceedings of ICML Workshop on Unsupervised and Transfer Learning.
- Yang, Q., Zhang, Y., Dai, W., & Pan, S. J. (2020). *Transfer learning*: Cambridge University Press.
- Yu, D., Deng, A. A., Dahl, G., Seide, F., & Li, G. (2012). *More data+ deeper model= better accuracy*. Paper presented at the Keynote at International Workshop on Statistical Machine Learning for Speech Processing 2012 (IWSML).
- Zebari, D. A., Zeebaree, D. Q., Abdulazeez, A. M., Haron, H., & Hamed, H. N. A. (2020). Improved Threshold Based and Trainable Fully Automated Segmentation for Breast Cancer Boundary and Pectoral Muscle in Mammogram Images. *IEEE Access*, 8, 203097-203116.
- Zebari, D. A., Zeebaree, D. Q., Saeed, J. N., Zebari, N. A., & Adel, A. (2020). Image steganography based on swarm intelligence algorithms: A survey. *people*, 7(8), 9.
- Zeebaree, D. Q., Abdulazeez, A. M., Hassan, O. M. S., Zebari, D. A., & Saeed, J. N. (2020). Hiding Image by Using Contourlet Transform. In: press.
- Zhang, W., Li, X., Jia, X.-D., Ma, H., Luo, Z., & Li, X. (2020). Machinery fault diagnosis with imbalanced data using deep generative adversarial networks. *Measurement*, 152, 107377.
- Zhao, L., Pan, S. J., & Yang, Q. (2017). A unified framework of active transfer learning for cross-system recommendation. *Artificial Intelligence*, 245, 38-55.
- Zhou, G., Xie, Z., Huang, X., & He, T. (2016). *Bi-transferring deep neural networks for domain adaptation*. Paper presented at the Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers).
- Zhou, J., Pan, S., Tsang, I., & Yan, Y. (2014). *Hybrid heterogeneous transfer learning through deep learning*. Paper presented at the Proceedings of the AAAI Conference on Artificial Intelligence.
- Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., . . . He, Q. (2020). A comprehensive survey on transfer learning. *Proceedings of the IEEE*, 109(1), 43-76.

Barwary, M. J. & Abdulazeez, A. M. (2021). Impact of Deep Learning on Transfer Learning : A Review. *International Journal of Science and Business*, 5(3), 204-216. doi: <https://doi.org/10.5281/zenodo.4559668>

Retrieved from <http://ijsab.com/wp-content/uploads/698.pdf>

Published by

