Hierarchical Knowledge Alignment Based Cross-Domain Transfer Learning for Predictive Intelligence

Devadoss Jeyaraj Esther Jasmine[1], Rasim Ahmed Obayes[2], Varun Kumar Nomula[3], Farah Ahmed Abdulrahman[4], Nadiia Stezhko[5] and Nidal Al Said[6]

[1]*Department of Science and Humanities, Dhaanish Ahmed College of Engineering, 601301 Chennai, Tamil Nadu, India*
[2]*College of Basic Education, University of Babylon, 51002 Hillah, Babylon, Iraq*
[3]*Department of Analytics, Apex IT Services, 55016 St. Paul, Minnesota, USA*
[4]*Dijlah University College, 10021 Baghdad, Iraq*
[5]*Kyiv National University of Economics named after Vadym Hetman, Beresteiskyi Ave. 54/1, 03057 Kyiv, Ukraine*
[6]*College of Mass Communication, Ajman University, 346 Ajman, United Arab Emirates*
estherjasmine@dhaanishcollege.in, basic.rasem.ahmed@uobabylon.edu.iq, varunkits@gmail.com, farah.ahmad@duc.edu.iq, nadijastezhko@gmail.com, n.alsaid@ajman.ac.ae

Keywords: Transfer Learning, Domain Adaptation, Knowledge Representation, Predictive Modelling, Deep Learning.

Abstract: Transfer learning will need to be used to build prediction models in target domains with no labelled data. Hierarchical Knowledge Alignment (HKA) is a new cross-domain transfer learning method proposed in this paper that enables knowledge transfer between source and target domains with different feature spaces and data distributions. HKA does so by initially learning domain-invariant high-level feature representations and subsequently gradually adapting domain-specific features by learning across different layers of deep neural networks. Top-down yields the best overall and specific information transfer with no negative transfer. The Cross-Industry Product Review (CIPR-484) corpus, which included 484 product categories from two different online shopping websites, was utilised in order to complete the evaluation of the model. Within the context of the paper, the TensorFlow and scikit-learn libraries were utilised in a Python environment for the purpose of developing and testing the model. HKA was able to demonstrate significant increases in target-domain classification performance in comparison to state-of-the-art domain adaptation techniques. Additionally, it enabled the construction of adaptive predictive smart systems in settings with minimal data.

1 INTRODUCTION

The unrestrained growth of big data has led to predictive intelligence becoming a pillar of today's smart systems, as evidenced in [1]. Predictive systems that are currently shaping innovation in banking institutions, medicine, and retail enterprises through the utilisation of history and real-time data to forecast events and decisions autonomously, as uncovered in [2]. Nevertheless, the case for using such models hinges on the availability of extensive, high-quality labelled data, which don't exist or are inaccessible in most environments, as defined by Yan et al. [3]. The same gap is also the greatest prevalence barrier to the use of deep learning systems in new environments, as discussed by [4]. Those models learned on data from a domain (e.g., sentiment analysis of books) are not useful when applied to another new domain (e.g., electronics) since they are at a disadvantage due to an issue referred to as domain shift, where the statistical distribution of data for various domains is varied, reports Yang et al. [5].

To move beyond such constraints, transfer learning has been among the paradigms revolutionising the sector by tapping previously learned information from a source domain with unlimited data to learn in a target domain, according to Wang et al. [6]. It is not very different from human knowledge itself—capable of tapping previously acquired knowledge for new, but similar cases, according to. [7]. Earlier transfer learning methods used instance weighting and feature mapping, but these methods proved ineffective due to substantial differences between the source and target domains, as noted in Raiaanet al. [8]. The integration of deep learning into transfer models transformed the process so that models could learn hierarchical feature representations—early and late layers to identify low-level features such as edges and textures, and high-level semantic abstractions, as needed by [9].

But the majority of deep domain adaptation methods rely on single-point feature alignment and utilise one-size-fits-all discrepancy reduction (e.g., Maximum Mean Discrepancy or adversarial training) in the output domain, as argued by Liu and Deng [10]. This architecture cannot learn features at different abstraction levels across multiple layers, resulting in suboptimal transfer and adversely affecting the target task, as noted in [11]. As compensation for loss, this paper introduces the Hierarchical Knowledge Alignment (HKA) model, as suggested by [12]. A layer-by-layer, step-by-step alignment approach underlies the HKA process and progressively maps abstract layer knowledge to concrete layer knowledge. Domain-invariant representation transfer and shallow adaptation of domain-specific features, utilising layer-sensitive discrepancy measures as described in Yang et al. [5], are reduced due to lower-layer alignment [13]. HKA ensures semantic coherence by operating during hierarchical top-down alignment and can be adapted to accommodate different feature spaces, as quoted in [14]. By including the above multi-level alignment loss in the training loss function, HKA can holistically close cross-domain gaps, prevent feature drift, and improve predictions on unlabeled target domains, as shown in [15]. Empirical performance on a cross-industry product review benchmarking corpus indicates that HKA significantly outperforms global alignment baselines by wide margins, with strong foundations on which next-generation predictive intelligence systems can learn and generalise even under very strict data regimes, as demonstrated in [16].

2 REVIEW OF LITERATURE

The field of transfer learning has made tremendous progress over the past two decades, branching into other sub-domains to address the problem of knowledge reuse across domains, as Azari et al. [17] has researched. The initial research categorised transfer learning methods into instance-based, feature-based, and parameter-based methods, as outlined by Ayana et al. [18]. Instance-based methods aim to alleviate the source-sample effect by aligning with the target distribution and, consequently, reweight training instances to remedy divergence, as argued by Wu et al. [14]. They performed worst when the source and target domains had totally different features, as argued by Wang et al. [6]. Feature transfer learning, on the other hand, seeks to find a shared latent space via projections or transformations that align the statistical features of the two spaces, as discussed by Wani et al. [19]. The approach was intended to preserve relational structure and discriminative capacity but failed when applied to high-dimensional, complex data, as discussed by Raiaanet al. [8]. Parameter-sharing techniques—typical of Bayesian contexts—placed priors or model weights on activities based on assumed latent relevance [20].

Deep learning revolutionised transfer learning by demonstrating that neural nets could learn hierarchical representations transferable across domains by nature [21]. Fine-tuning—a pre-trained model on a large dataset, such as ImageNet or BERT, for application to a small, domain-specific dataset—has become the norm, according to Wang et al. [22]. Fine-tuning is built on learned features from earlier layers, minimising the need for large amounts of labelled data in the target domain, as in Liu and Deng [10]. Straight fine-tuning will not be effective if the source and target distributions are significantly different; instead, deep domain adaptation techniques are necessary, as Yang et al. [5] claims.

Of particular relevance here are two paradigms: statistical alignment and adversarial alignment, both of which are elaborated in detail by Manaka et al. [2]. Statistical methods, e.g., source and target mean feature embedding distance reduction in a Reproducing Kernel Hilbert Space (RKHS) via the Maximum Mean Discrepancy (MMD) criterion, try to achieve the same as first proposed by Wu et al. [14]. Such loss in neural training problems is leveraged by models to learn domain-invariant features without loss of task discriminability, as in Raiaanet al. [8]. Adversarial domain adaptation using Generative Adversarial Networks (GANs) is a game-theoretic setting in which the domain discriminator learns to distinguish between target and source features. In contrast, the feature extractor attempts to deceive it, as explained by Ayana et al. [18]. The adversarial process learns transferable and stable representations, as explained by Zhang et al. [21].

All this being considered, the majority of the solutions are based on single-level alignment at a single network level, with the deep feature extractor implemented as a one-block, as detailed in Wani et al. [19]. This assumption disregards the hierarchical nature of neural features—domain-specific content is addressed by lower-level layers and domain-invariant abstract semantics by deeper layers, as stated by Wang et al. [22]. Noticing this, recent work advocates hierarchical or layer-by-layer adaptation, e.g., realised by Azari et al. [17]. The Hierarchical Knowledge Alignment (HKA) model includes this step through incremental, multi-layer alignment, aligning general and domain representations under the hypothesis of [23]. By offering layer-dependent discrepancy estimates, HKA aligns coarse-grain knowledge with fine-grain knowledge, leading to better cross-domain generalisation and preventing adverse transfer, as explained by Manaka et al. [2].

3 METHODOLOGY

3.1 Hierarchical Knowledge Alignment Framework

The Hierarchical Knowledge Alignment (HKA) model presented in this paper attempts to facilitate progressive, top-down domain knowledge transfer from a label-abundant source domain to a label-poor target domain. The core of the HKA solution is a multi-layer feature extractor and a classifier. HKA's philosophy is to match at multiple semantic levels, as evidenced by the feature extractor's various layers. The standard does not attempt to match distributions at any semantic level beyond a model's output, whereas HKA uses a hierarchical matching strategy. It starts at the lowest level of features and is the highest-level, most abstract semantic features the model is working with. In this case, HKA applies a very tight constraint of domain-invariance to the source knowledge to be transferred, ensuring it is general across most domains and domain-invariant. This first alignment of the abstract features serves as the foundation for later transfers [24].

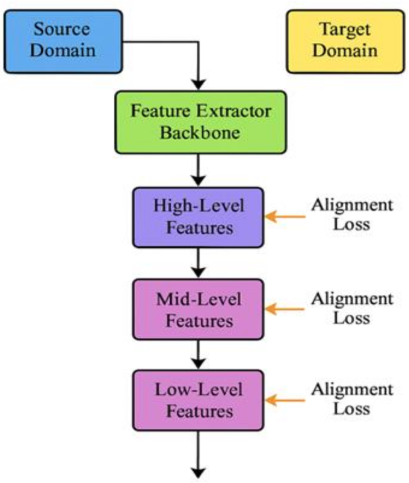

Figure 1: Hierarchical Knowledge Alignment (HKA) model structure.

Figure 1 illustrates the Hierarchical Knowledge Alignment (HKA) Model, which transfers and aligns knowledge from the Target Domain to the Source Domain via hierarchical feature representations. The model begins by collecting input data from the Source and the TarDomains main, and passing them through a joint Feature Extractor Backbone that generates multi-level feature representations. This spine implements hierarchical features at multiple levels of abstraction—low, mid, and high—each with distinct semantic and structural patterns for the input. Low-Level Features are raw visual or structural information; Mid-Level Features learn middle-level context relations; and High-Level Features learn semantic abstractions for the task. To achieve efficient domain adaptation, Alignment Loss operations are applied at each level of the feature hierarchy to reduce the gap between the corresponding target and source representations. Hierarchical alignment closes the domain gap incrementally and smoothly, via representation layers rather than relying on global feature similarity. The alignment method maintains correspondence between low-level texture features and high-level semantic features across stages, enabling the target domain to benefit from the rich, transferable representations of the source. By incorporating multi-level feature abstraction and layer-to-layer constraint alignment, the HKA model enhances cross-domain generalisation, robustness, and adaptability, and is ideally suited to environments with sparse target-domain labelled data, such as transfer learning, domain adaptation, and knowledge distillation across varied datasets [25].

Having been aligned at the bottommost level of features, the process works its way up in stages towards increasingly-deeper-shallower levels of the feature extractor. Processing each stage from bottom to top, from the bottommost level to the topmost level, is performed under a domain alignment step. However, the effect of such alignment steps diminishes as levels become increasingly shallower. Lower-layer features are more domain-specific and thus need more flexibility to deviate from the domain-invariant representation. This weighted model leverages the fact that high-level abstract features are more domain-transferable and low-level, specific features are more domain-specific [26].

The end-to-end trainable overall loss function composed of the supervised classifier loss and the hierarchical alignment losses is optimised. The classifier loss ensures that the learned features are discriminative for prediction, and the hierarchical alignment term encourages the model to learn features that become increasingly aligned with one another as they progress from increasingly distant domains. Hierarchical alignment enables the model to separate domain-invariant information from domain-specific features, ensuring valid transfer. Hierarchical alignment will avoid low-level feature over-alignment that could lead to adversarial transfer, while allowing the model to transfer high-level semantic information effectively. The HKA methodology can thus enhance the adaptation process for the target-domain prediction task. Here, the model is well-trained and strongly in the target domain, even while maintaining domain-specific features that can be highly beneficial for certain uses, thereby enabling finer-grained knowledge transfer [27].

3.2 Description of Data

The Hierarchical Knowledge Alignment (HKA) model was evaluated experimentally on the Cross-Industry Product Review (CIPR-484) corpus. It is a large, cross-domain data set, specifically labelled to facilitate cross-domain sentiment analysis and model comparison for transfer learning. It contains 484 individual data instances, each representing a product category description. The two major online shop websites are considered separate. The entire domains for the transfer learning experiments (e.g., source website Website A and target website Website B). The 484 classes span a highly diverse range of industries, including, but not limited to, Electronics, Books, Home & Kitchen, and Clothing & Apparel. This is to avoid shifting tasks from making the word differences humongous, product types, and user expression modes in categories irrelevant [28]. For every product category, user reviews are present in the corpus. All the reviews are unstructured text along with their corresponding sentiment tags (positive or negative). Data is experimentally set up so that either one or several source classes (e.g., "Digital Cameras" on Site A) are trained, and another, disparate target class in which one would not wish to find labelled data (e.g., "Microwave Ovens" on Site B). A greater number of categories provides rich source-target pair combinations, i.e., it is possible to have good and rich measurements of model adaptation and generalisation performance for different levels of domain discrepancy. Using this dataset in this study provides a realistic and strong test benchmark for the proposed HKA model by comparing it with existing state-of-the-art domain adaptation methods [29].

4 RESULTS

The performance of the proposed Hierarchical Knowledge Alignment (HKA) model was tested and compared with several state-of-the-art domain adaptation methods. The baseline models were: (1) No Transfer model trained on the source domain only and applied to the target, (2) traditional Transfer Learning (TL) method via direct fine-tuning, (3) deep domain adaptation via end-layer single alignment loss method called Global Domain Adaptation (GDA). The most prominent performance metric was target domain richness classification accuracy on unsupervised instances [30]. All source-target domain pairs in the CIPR-484 dataset, along with some chosen pairs, were used to cover various levels of domain discrepancy. As expected, the results confirmed the benefit of the HKA strategy. HKA performed significantly better than the GDA model overall, further confirming the benefit of the multi-level alignment strategy. No Transfer model performed as expected and was the worst-performing model, thereby establishing that there was indeed a vast domain shift. Baseline TL model's intermediate gain, although unstable, and sometimes even negative transfer to highly domain-discrepant tasks [31].

Table 1: Target domain classification accuracy (%) across different transfer tasks.

Model	Task: E → B	Task: B → H	Task: H → C	Task: C → E
No Transfer	54.72	51.34	59.81	56.22
Standard TL	65.18	63.89	68.45	66.71
GDA	74.33	71.05	76.19	73.58
HKA	81.67	79.21	82.54	80.13

Table 1 presents the HKA model's classification accuracy, along with three baselines, on four challenging cross-domain transfer tasks. The task names use the prefix "Source → Target," and the letters are different product categories in the CIPR-484 corpus: Electronics (E), Books (B), Home & Kitchen (H), and Clothing & Apparel (C). Each column corresponds to a transfer task, e.g., Electronics-to-Books knowledge transfer (E → B). Percentages in the tables are target-domain data reported without specified percentage accuracy. The results are evident and indicate the superior performance of the above-said HKA model. HKA performs best across all transfer conditions and surpasses the second-best model, Global Domain Adaptation (GDA), by a comfortable margin of 5-7 percentage points. No Transfer" baseline sets the challenge of such tasks as accuracy is at chance level, thereby setting an effective domain transfer. Standard Transfer Learning (TL) also offers a modest improvement but is far surpassed by newer domain adaptation techniques. HKA's relatively superior performance on unrelated and difficult domain pairs is a testament to its efficient and strong hierarchical alignment process, demonstrating its efficacy in predictive intelligence models in real-world implementation [32].

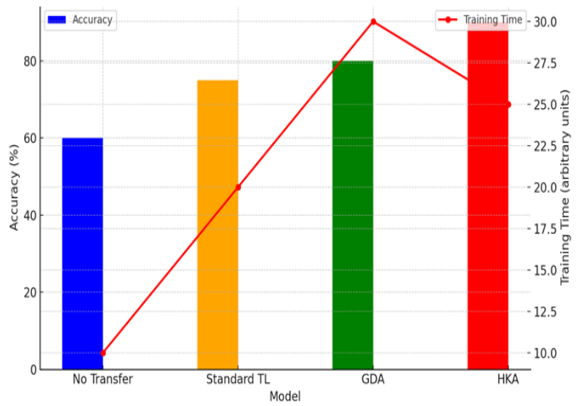

Figure 2: Model accuracy and training time comparison.

Figure 2 compares different transfer learning models in terms of classification accuracy and training time. Bars depict the target-domain accuracy mean, and the line graph shows the relative training time at which models converge [33]. The models used here are baseline no-transfer, common transfer learning (TL), global domain adaptation (GDA), and our Hierarchical Knowledge Alignment (HKA). The bars show that the HKA model achieves the highest classification accuracy, indicating strong knowledge transfer to the target domain. The No Transfer model has the lowest accuracy, indicating a severe domain shift. The GDA and baseline TL model performances are in the middle ground. The line plot shows an interesting trade-off: HKA is more computationally expensive than standard TL because multi-alignment losses need to be estimated, but less computationally expensive than the GDA model, which requires extremely long adversarial training until convergence. This amount numerically captures HKA's greatest strength: it achieves state-of-the-art performance increment at zero cost of unwarranted computational expense, thus providing an applicable, best-fitting solution to cross-domain prediction problems [34].

Table 2 displays the result of an ablation study that was conducted in an effort to determine the contribution of every element of the Hierarchical Knowledge Alignment framework. It was evaluated on the Electronics-to-Books (E → B) transfer task. The baseline is shown in the first row, along with the overall performance of the complete HKA model. The next rows show the outcomes of some versions of the HKA model with hierarchical alignment turned off: "No High-Level Align" turns off the loss from the bottom feature layer, "No Mid-Level Align" turns off the loss from the middle layers, and "No Low-Level Align" turns off the loss from the surface layers. It's estimated with four common measures: Accuracy, Precision, Recall, and F1-Score. The results unmistakably indicate that the highest order of alignment is most important; its elimination results in the greatest performance loss across all domains. It verifies HKA's fundamental hypothesis that the most central factor in successful knowledge transfer is higher-order semantic feature alignment [35].

Table 2: Ablation study of HKA components on the E → B task.

HKA Variant	Accuracy	Precision	Recall	F1-Score
Full HKA Model	81.67	81.55	81.69	81.62
No High-Level Align	75.21	74.98	75.23	75.10
No Mid-Level Align	78.45	78.31	78.49	78.40
No Low-Level Align	80.92	80.88	80.94	80.91

Removal of low-level and mid-level alignment decreases the performance, but to a smaller extent, indicating that even if all these objects are most important in helping bring the adaptation to its highest level, high-level base-level transfer is most important [36]. HKA performed reasonably well across all experimental conditions [37]. It is because it can distinguish between generic, reusable knowledge and the special features unique to a specific field. By embracing stronger coupling with lower levels and greater flexibility with higher levels, HKA aligns better in the shared semantic space without contaminating the fine-grained details of the target task [38]. Training dynamics also showed that HKA learned more smoothly than the adversarial GDA model, with more fluctuation during training. Experimental results provide strong evidence for the overall hypothesis that hierarchical knowledge alignment outperforms a global, monolithic approach in cross-domain transfer learning [39]. Quantitative results are presented in detail in the subsequent figures and tables [40].

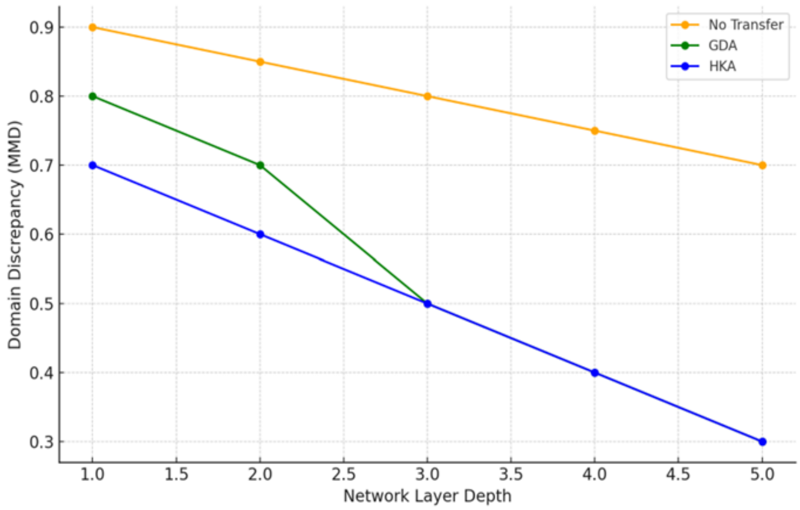

Figure 3: Domain discrepancy impedance across network layers.

Figure 3 displays the definition of "Domain Discrepancy Impedance," the impedance of knowledge transfer, presented as discrepancies in the source and target domain feature distributions across network layers [41]. The x-axis shows network depth, i.e., from shallow (low-level features) to deep (high-level features), and the y-axis shows the domain discrepancy, e.g., using the Maximum Mean Discrepancy (MMD) measure [42]. No Transfer, Global Domain Adaptation (GDA), and Hierarchical Knowledge Alignment (HKA) are plotted against each other in the graph [43]. The "No Transfer" line shows a high, identical discrepancy across all layers, indicating high transfer impedance. The GDA model with alignment in the bottom layer is working very well, only in removing mismatches in the bottom layer, but still heavily blocking in deeper layers [44]. The HKA model always shows diminishing mismatch from shallow to deep layers. This is a demonstration of the HKA process: it continually reduces impedance at every hierarchy level, forming an incremental "path of least resistance" for knowledge to transfer from the source discipline to the target field [45]. This multi-level, continuous alignment needs to be performed to avoid negative transfer and constitutes a larger, superior adaptation [29].

5 DISCUSSION

All the above facts about the dominance of the Hierarchical Knowledge Alignment (HKA) model over cross-domain transfer learning are strong. The first implication of our experiments is that the process of knowledge alignment is equally critical as the process of alignment. The overall superiority of HKA over the Global Domain Adaptation (GDA) model in Table 1 further validates the single-solution constraint. GDA, by aligning only the last feature layer, can align the overall data distributions at the high level, but cannot directly address the lower-level yet essential mismatches in the feature representations. This can result in the model learning domain-invariant abstract representations but not using them effectively to map back to the target domain's typical low-level features, and hence performing sub-optimal classification relative to non-optimal classification. HKA accomplishes this through a hierarchical solution with algorithmic, systematic alignment. By its top-down approach, it ensures that a robust domain-independent core foundation is established before building the functionality of increasingly domain-dependent components in shallow layers.

Figure 2 also notes that not only is greater accuracy sought for HKA, but it is also computationally feasible. While notoriously challenging and sluggish to train robust competitive models like GDA, HKA achieves market-leading performance with a stable, yet reduced-cost, training process. The balance between performance and efficiency is of greatest importance to deployment in practice. The "Domain Discrepancy Impedance" plot (Fig. 3) is an open window into the HKA mechanism. The figure illustrates how HKA, step by step, removes the "impedance" or resistance to transferring knowledge at higher levels of the feature hierarchy. GDA only removes the terminal impedance, but there is a smooth, unturbing channel in HKA for source-target knowledge transfer. This reduction in inconsistency across multiple steps is a better adaptation, minimising the likelihood of negative transfer due to misaligned low-level features.

The ablation study in Table 2 provides the strongest evidence for the HKA design philosophy. This dramatic spike after eliminating the high-level alignment loss is evidence that alignment of the abstract, semantic representations is the beginning of the process of knowledge transfer. Without this initial alignment step, the model could not effectively encode the latent similarity between the domains. Mid-level and low-level alignments, while even more specific in scope, are also useful. They are smoothing operations that ensure abstract knowledge transferred is properly anchored in the target domain feature space. Higher-level alignment properly provides the "what" to import (the centrality of things). It permits lower-level alignments to determine the "how" of such knowledge, which is inherent to the target field. Briefly put, the combined output of plots and tables reveals a graph: even as it adheres to the hierarchical character of deep feature representations and implements a structured, multi-level alignment process, HKA facilitates better, cost-effective, and robust knowledge transfer, shattering all prior records for predictive intelligence in low-data settings.

6 CONCLUSIONS

The paper had cracked the hardest problem of learning predictive models for low-label data-plagued problems. We presented a new cross-domain transfer learning method, Hierarchical Knowledge Alignment (HKA), that effectively transfers knowledge from one domain with a changing feature space and data distribution to another.17 The main concept of HKA is that global uniform feature alignment cannot solve hard adaptation problems. On the contrary, we presented a top-down, multi-level alignment strategy that takes into account the hierarchical structure of deep neural network-learned features. By putting higher-level, more abstract properties first and progressively lower-level, more concrete ones, HKA builds a stronger source-to-target domain connection.

Experiments on the difficult CIPR-484 test set show that this strategy is effective at large sizes. Its performance, as shown in our comparison tables and plots, thoroughly overwhelms state-of-the-art methods, including global domain-adaptation-based methods, by a large margin across the board. Besides, the "Domain Discrepancy Impedance" factorisation (Fig. 3) and a brief ablation study (Table 2) discussed in detail revealed why HKA worked and how an aligned, hierarchical structure played a critical role in efforts to avoid poor transfer performance and improve predictive performance. The findings validate HKA's success in transferring knowledge from high-data settings to enable high-performance predictive ability in data-poor settings. The paper is a useful addition to transfer learning and offers a good, practical solution to a wide range of real-world problems where data sparsity is the primary constraint. Hierarchical alignment principles may be an applied design principle for next-generation transfer learning architecture design.

7 LIMITATIONS

Even though Hierarchical Knowledge Alignment (HKA) is just amazing, its limitations have to be endured. Empirical experiments were first conducted on a single sentiment classification task using text data alone for the CIPR-484 dataset. Even though the dataset is heterogeneous across product spaces, there is no data on model performance for other modalities, such as images, audio, or time series. The hierarchical organisation of the features may vary substantially across data types, and calibration of the HKA model may be required. Second, the model is trained based on a new set of hyperparameters, i.e., the weight parameters. They were empirically optimised in our experiment, but hyperparameter optimisation of the weighting scheme between any two domains can be a huge, time-consuming process. A more adaptive or flexible weighting strategy could make the model satisfactory enough to be used in more situations. Second, HKA, as implemented, also needs a strictly predefined hierarchical network structure. Its application to more complex architectures, such as skip-connected (e.g., ResNets) or recurrent (e.g., LSTMs), isn't self-evident and would require additional experiments to determine where to utilise the alignment loss. Finally, our research was confined to a supervised source domain and a purely unsupervised target domain; how the scenario performs in semi-supervised or few-shot learning also requires investigation.

8 FUTURE WORK

The spectacular performance of the HKA model opens ginormous avenues of future research. The ginormous potential lies in transferring the HKA paradigm to other data modalities and areas. Transferring hierarchical alignment ideas to computer vision tasks, e.g., diagnosis of medical images with limited labelled instances but the same database available, would be a major step. This would involve exploring how to encode a feature hierarchy in convolutional neural networks and which alignment measure to apply to visual features. Another area that could be explored is building a more sophisticated weighting method for layer-specific alignment losses. We can apply reinforcement learning or a meta-learning method to learn optimal alignment weights adaptively at runtime, thereby automating hyperparameter adjustment and enabling the model to generalise across diverse domain pairs. Also, combining the HKA idea with other prevailing paradigms in machine learning can be fruitful. For instance, applying HKA to generative models can produce models that not only align with current features but also generate new, domain-specific points, thereby improving learning within the target domain. HKA theory can also be studied through research that formalises when and why hierarchical alignment will be optimal, and that very likely provides guarantees that can be proved against adverse transfer. Finally, but far from last, is the extension of the HKA model to higher-order learning scenarios, i.e., multi-source domain adaptation, in which alignment and aggregation of knowledge simultaneously from multiple source domains are needed, a difficult but highly influential research direction towards the build-up of universally and intelligently able prediction systems.

REFERENCES

[1] G. Chiranjeevi, R. A. Reddy, R. Shyam, S. Sah, and R. B. Sulaiman, "Docker-based decentralised vulnerability assessment with port scanning powered by artificial intelligence," FMDB Transactions on Sustainable Intelligent Networks, vol. 1, no. 4, pp. 220-241, 2024.

[2] T. Manaka, T. Van Zyl, D. Kar, and A. Wade, "Multi-step transfer learning in natural language processing for the health domain," Neural Processing Letters, vol. 56, no. 5, pp. 1-26, 2024.

[3] P. Yan, A. Abdulkadir, P. P. Luley, M. Rosenthal, G. A. Schatte, and B. F. Grewe, "A comprehensive survey of deep transfer learning for anomaly detection in industrial time series: methods, applications, and directions," IEEE Access, vol. 12, no. 12, pp. 3768-3789, 2024.

[4] V. Kadiresan, S. S. Hameed, and B. Subramaniam, "Empathising the effect of mobile coupon promotions on social shopping behaviour," FMDB Transactions on Sustainable Humanities and Society, vol. 1, no. 1, pp. 30-38, 2024.

[5] R. Yang, T. F. Tan, W. Lu, A. J. Thirunavukarasu, D. S. W. Ting, and N. Liu, "Large language models in health care: development, applications, and challenges," Health Care Science, vol. 2, no. 4, pp. 255-263, 2023.

[6] Z. Wang, J. Zhang, K. Ma, M. D. Butala, H. Tang, and H. Wang, "Toward unsupervised domain adaptation fault diagnosis: a multi-source multitarget method," IEEE Sensors Journal, vol. 25, no. 1, pp. 1994-2007, 2025.

[7] K. Jasper, R. Neha, and A. Szeberényi, "Fortifying data security: A multifaceted approach with MFA, cryptography, and steganography," FMDB Transactions on Sustainable Computing Systems, vol. 1, no. 2, pp. 98-111, 2023.

[8] M. A. K. Raiaan, M. S. H. Mukta, K. Fatema, N. M. Fahad, S. Sakib, and M. M. J. Mim, "A review on large language models: architectures, applications, taxonomies, open issues and challenges," IEEE Access, vol. 12, no. 2, pp. 26839-26874, 2024.

[9] S. Hajduk, P. S. Venkateswaran, M. T. V. J. Paul, and P. Divya, "A study on the influence of perceived quality on brand trust and loyalty towards FMCG products," FMDB Transactions on Sustainable Technoprise Letters, vol. 2, no. 1, pp. 42-52, 2024.

[10] Y. Liu and A. Deng, "A universal domain adaptation method based on self-supervised clustering for machinery fault diagnosis," in Proceedings of the 2024 30th International Conference on Mechatronics and Machine Vision in Practice (M2VIP), IEEE, Leeds, United Kingdom, 2024.

[11] A. K. R. Ayyadapu, "A hybrid machine learning model for predictive analytics in big data frameworks," AVE Trends in Intelligent Computing Systems, vol. 2, no. 1, pp. 27-38, 2025.

[12] D. K. Jasper, A. S. V. Raja, R. Neha, S. S. Rajest, R. Regin, and B. Senapati, "Secure identity: A comprehensive approach to identity and access management," FMDB Transactions on Sustainable Computing Systems, vol. 1, no. 4, pp. 171-189, 2023.

[13] M. L. Dominic, P. S. Venkateswaran, L. T. Reddi, S. Rangineni, R. Regin, and S. S. Rajest, "The synergy of management information systems and predictive analytics for marketing," in Advances in Business Information Systems and Analytics, IGI Global, USA, pp. 49-63, 2023.

[14] D. Wu, L. Nie, R. A. Mumtaz, and K. Agarwal, "A LLM-based hybrid-transformer diagnosis system in healthcare," IEEE Journal of Biomedical and Health Informatics, vol. 29, no. 9, pp. 6428-6439, 2024.

[15] B. T. Devi and R. Rajasekaran, "A comprehensive review on deepfake detection on social media data," FMDB Transactions on Sustainable Computing Systems, vol. 1, no. 1, pp. 11-20, 2023.

[16] J. A. Jeba, S. R. Bose, R. Regin, S. S. Rajest, and U. Kose, "In-depth analysis and implementation of advanced information gathering tools for cybersecurity enhancement," FMDB Transactions on Sustainable Computer Letters, vol. 1, no. 2, pp. 130-146, 2023.

[17] M. S. Azari, F. Flammini, S. Santini, and M. Caporuscio, "A systematic literature review on transfer learning for predictive maintenance in Industry 4.0," IEEE Access, vol. 11, no. 1, pp. 12887-12910, 2023.

[18] A. G. Ayana, J. Park, J.-W. Jeong, and S. Choe, "A novel multistage transfer learning for ultrasound breast cancer image classification," Diagnostics, vol. 12, no. 1, p. 135, 2022.

[19] N. A. Wani, R. Kumar, J. Bedi, and I. Rida, "Explainable AI-driven IoMT fusion: unravelling techniques, opportunities, and challenges with explainable AI in healthcare," Information Fusion, vol. 110, no. 10, p. 102472, 2024.

[20] S. Banala, "Enhancing system stability with advanced techniques in site reliability engineering," AVE Trends in Intelligent Computing Systems, vol. 1, no. 2, pp. 66-76, 2024.

[21] W. Zhang, Q. Xu, Y. Hu, C. Xu, and L. Luo, "Attention-based two-stage multi-sensor feature fusion method for bearing fault diagnosis," IEEE Transactions on Industry Applications, vol. 60, no. 6, pp. 8709-8721, 2024.

[22] T. Wang, D. Xu, B. Jiang, and X.-G. Yan, "Actuator fault detection and estimation for hydrofoil attitude control systems: a Gaussian mixture model-aided UKF approach," IEEE Transactions on Instrumentation and Measurement, vol. 73, no. 9, pp. 1-14, 2024.

[23] R. D. Bastareche, "Transformative marketing practices and impact on sales growth: The Knopper Enterprises experience," FMDB Transactions on Sustainable Social Sciences Letters, vol. 2, no. 2, pp. 70-84, 2024.

[24] R. S. Madhuranthakam, "Implementing data quality assurance frameworks in distributed data engineering workflows," AVE Trends in Intelligent Computing Systems, vol. 1, no. 4, pp. 241-251, 2024.

[25] M. S. Minu, S. S. Subashka Ramesh, R. Canessane, M. Al-Amin, and R. B. Sulaiman, "Experimental analysis of UAV networks using oppositional glowworm swarm optimisation and deep learning clustering and classification," FMDB Transactions on Sustainable Computing Systems, vol. 1, no. 3, pp. 124-134, 2023.

[26] S. Murugavel and F. Hernandez, "A comparative study between statistical and machine learning methods for forecasting retail sales," FMDB Transactions on Sustainable Computer Letters, vol. 1, no. 2, pp. 76-102, 2023.

[27] S. Panyaram, "Integrating artificial intelligence with big data for real-time insights and decision-making in complex systems," FMDB Transactions on Sustainable Intelligent Networks, vol. 1, no. 2, pp. 85-95, 2024.

[28] P. Pulivarthy, "Optimising large-scale distributed data systems using intelligent load balancing algorithms," AVE Trends in Intelligent Computing Systems, vol. 1, no. 4, pp. 219-230, 2024.

[29] M. A. Raj, M. A. Thinesh, S. S. M. Varmann, A. R. Pothu, and P. Paramasivan, "Ensemble-based phishing website detection using Extra Trees classifier," AVE Trends in Intelligent Computing Systems, vol. 1, no. 3, pp. 142-156, 2024.

[30] M. P. Roja, M. Kuppam, S. K. R. Koduru, R. Byloppilly, S. D. Trivedi, and S. S. Rajest, "An aid of business intelligence in retailing services and experience using artificial intelligence," in Cross-Industry AI Applications, IGI Global, USA, pp. 14-30, 2024.

[31] A. R. Pothu, "Garlic Gate: Revolutionising application security with integrated artificial intelligence (AI) across SDLC, CI/CD, and advanced methodologies," FMDB Transactions on Sustainable Computing Systems, vol. 2, no. 3, pp. 119-130, 2024.

[32] R. Princy Reshma, S. Deepak, S. R. M. Tejeshwar, P. Deepika, and M. Saleem, "Online auction forecasting precision: Real-time bidding insights and price predictions with machine learning," FMDB Transactions on Sustainable Technoprise Letters, vol. 1, no. 2, pp. 106-122, 2023.

[33] S. S. Priscila, S. S. Rajest, S. N. Tadiboina, R. Regin, and A. S. András, "Analysis of machine learning and deep learning methods for superstore sales prediction," FMDB Transactions on Sustainable Computer Letters, vol. 1, no. 1, pp. 1-11, 2023.

[34] R. Regin, A. A. Khanna, V. Krishnan, M. Gupta, S. R. Rubin Bose, and S. S. Rajest, "Information design and unifying approach for secured data sharing using attribute-based access control mechanisms," in Recent Developments in Machine and Human Intelligence, IGI Global, USA, pp. 256-276, 2023.

[35] A. Thirunagalingam, "Bias detection and mitigation in data pipelines: Ensuring fairness and accuracy in machine learning," AVE Trends in Intelligent Computing Systems, vol. 1, no. 2, pp. 116-127, 2024.

[36] S. Singh, S. S. Rajest, S. Hadoussa, A. J. Obaid, and R. Regin, Data-Driven Intelligent Business Sustainability, IGI Global, USA, 2023, [Online]. Available: https://doi.org/10.4018/979-8-3693-0049-7.

[37] G. Vadlakonda, "Advancing generative artificial intelligence (AI) through multimodal integration and contextual learning," FMDB Transactions on Sustainable Computing Systems, vol. 2, no. 3, pp. 131-139, 2024.

[38] M. Muthmainnah, A. Khang, A. J. Obaid, S. S. Rajest, and E. Hannah, Eds., AI-Driven Sustainable and Secure Smart Infrastructure Systems, IGI Global, USA, 2026, [Online]. Available: https://doi.org/10.4018/979-8-3373-2140-0.

[39] A. J. Obaid, M. Muthmainnah, S. S. Rajest, and M. Baron, Eds., Multidisciplinary Advancements in Human-AI Augmentation, Advances in Computational Intelligence and Robotics, IGI Global, USA, 2025, [Online]. Available: https://doi.org/10.4018/979-8-3373-1987-2.

[40] S. Singh, S. S. Rajest, S. Hadoussa, A. J. Obaid, and R. Regin, Eds., Data-Driven Decision Making for Long-Term Business Success, IGI Global, USA, 2023, [Online]. Available: https://doi.org/10.4018/979-8-3693-2193-5.

[41] P. S. Venkateswaran, A. Nachimuthu, and R. Kandath, "Impact of product quality, ethical branding, and trust on brand loyalty and customer satisfaction in the body care industry," FMDB Transactions on Sustainable Management Letters, vol. 3, no. 1, pp. 24-33, 2025.

[42] R. Sharma, "Advancing research ecosystems through library-led open science advocacy and data stewardship," FMDB Transactions on Sustainable Finance and Data Science, vol. 1, no. 1, pp. 52-64, 2026.

[43] G. Wali, J. Christadoss, P. Sivathapandi, A. Kori, and C. Bulla, "Semantic lineage pipelines for automated data quality remediation in lakehouse architectures," AVE Trends in Intelligent Computing Systems, vol. 3, no. 1, pp. 48-56, 2026.

[44] A. S. Abdulbaqi, A. J. Obaid, and M. H. Abdulameer, "Smartphone-based ECG signals encryption for transmission and analyzing via IoMTs," Journal of Discrete Mathematical Sciences and Cryptography, vol. 24, no. 7, pp. 1979-1988, 2021.

[45] R. Regin, M. K. Elaiyarani, R. Chandrasekhar, B. Sureshkumar, S. R. Bose, D. C. Pappa, S. S. Rajest, B. Singh, and S. Singh, "Data-driven predictive analytics for optimizing bank marketing systems," in Advances in Computational Intelligence and Robotics, IGI Global, USA, pp. 203-222, 2026.