International Journal of Computer Engineering and Technology (IJCET)
Volume 15, Issue 4, July-Aug 2024, pp. 184-197, Article ID: IJCET_15_04_016
Available online at https://iaeme.com/Home/issue/IJCET?Volume=15&Issue=3
ISSN Print: 0976-6367 and ISSN Online: 0976-6375
Impact Factor (2024): 18.59 (Based on Google Scholar Citation)
DOI: https://doi.org/10.5281/zenodo.13149144

TRANSFER LEARNING'S TRIUMPH IN COMPUTER VISION: A BOON FOR EFFICIENCY AND PERFORMANCE

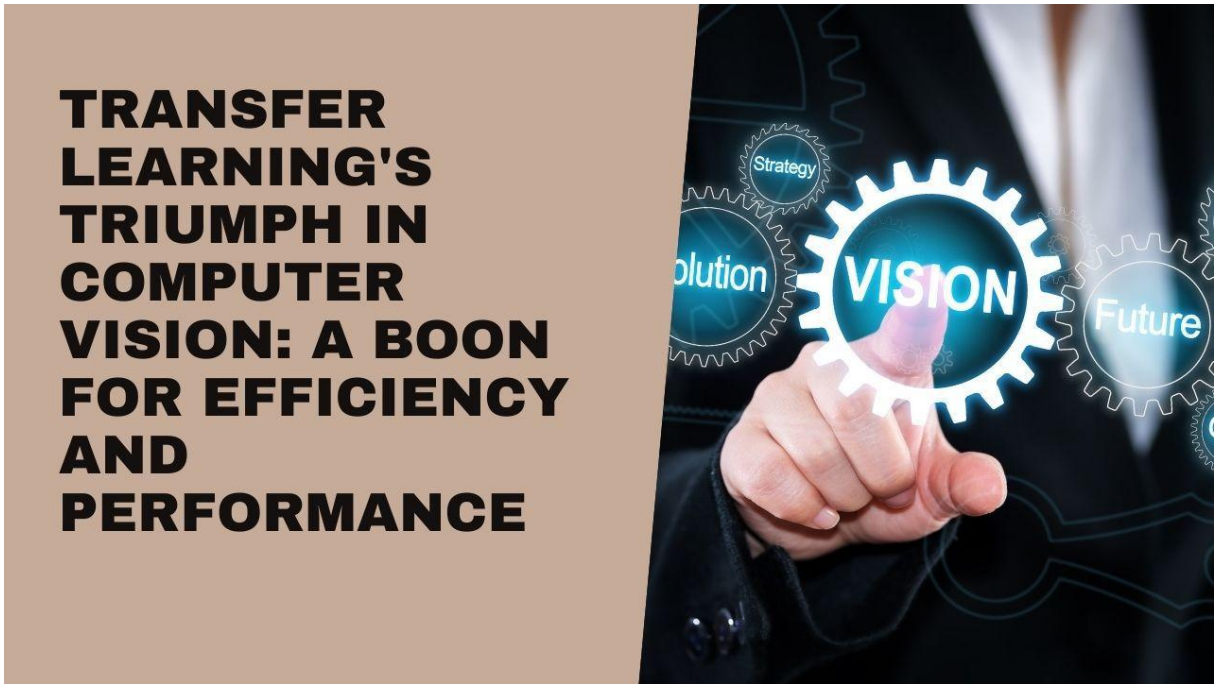
Rathish Mohan
Lore Health, USA

ABSTRACT

This article explores the transformative impact of transfer learning in computer vision, addressing the challenges of extensive data requirements and computational demands in traditional approaches. It discusses the concept and methodology of transfer learning, including model selection, fine-tuning techniques, and layer-freezing strategies. The article highlights the significant advantages of transfer learning, such as reduced data requirements, accelerated training times, and improved model performance. It presents real-world applications across object detection, image classification, and other computer vision tasks. Finally, the article examines future directions in transfer learning research, including advanced techniques and cross-domain transfer, emphasizing the potential for further advancements in the field.

Keywords: Transfer Learning, Computer Vision, Deep Learning, Pre-trained Models, Feature Extraction, Efficiency, Performance

Cite this Article: Rathish Mohan, Transfer Learning's Triumph in Computer Vision: A Boon for Efficiency and Performance, *International Journal of Computer Engineering and Technology (IJCET)*, 15(4), 2024, pp. 184-197.
https://iaeme.com/MasterAdmin/Journal_uploads/IJCET/VOLUME_15_ISSUE_4/IJCET_15_04_016.pdf



1. INTRODUCTION

Computer vision tasks have long been a cornerstone of artificial intelligence research, with applications ranging from facial recognition to autonomous driving [1]. However, the field has been consistently challenged by two primary obstacles: the need for vast labeled datasets and the computational resources required to train deep learning models from scratch.

Traditional computer vision approaches often require millions of labeled images to achieve satisfactory performance. For instance, the ImageNet dataset, which has been instrumental in advancing the field, contains over 14 million hand-annotated images across 20,000 categories [2]. This labeling process is time-consuming and expensive, often requiring domain expertise to ensure accuracy.

Moreover, the computational demands of training complex neural networks on such large datasets are substantial. High-performance GPUs or specialized hardware accelerators are typically necessary, leading to significant energy consumption and environmental concerns. For example, training a single large language model can produce carbon emissions equivalent to five cars over their lifetimes [1].

This article introduces transfer learning as a paradigm shift in addressing these challenges, offering improved efficiency and performance. Transfer learning leverages knowledge gained from solving one problem and applies it to a different but related problem. In computer vision, this often involves using a pre-trained model that has already learned general features from a large dataset like ImageNet, and fine-tuning it for a specific task with a smaller, specialized dataset.

By adopting transfer learning techniques, researchers and practitioners can significantly reduce the data requirements and computational costs of developing high-performance computer vision models. This approach accelerates the development process and democratizes access to advanced AI capabilities, allowing smaller organizations with limited resources to benefit from state-of-the-art computer vision technologies.

The subsequent sections of this article will delve deeper into the mechanics of transfer learning, its advantages over traditional training methods, and its transformative impact on various computer vision applications.

2. CHALLENGES IN TRADITIONAL COMPUTER VISION TRAINING:

2.1. Data Requirements:

One of the most significant challenges in traditional computer vision training is the need for vast amounts of labeled data. Deep learning models, particularly convolutional neural networks (CNNs), require enormous datasets to learn robust and generalizable features [3]. This requirement stems from the complexity of visual data and the need to capture various variations in lighting, angle, occlusion, and other factors that affect image perception.

For instance, the development of AlexNet, a breakthrough in computer vision, relied on the ImageNet dataset containing over 1.2 million labeled images across 1,000 categories [3]. Similarly, more recent models like YOLO (You Only Look Once) for object detection were initially trained on datasets containing hundreds of thousands of annotated images [4].

Creating such large-scale labeled datasets is time-consuming and expensive. Human annotators must manually label each image, often requiring domain expertise for specialized tasks like medical imaging. This labeling process can take months or even years for large datasets, significantly slowing the development of new computer vision models and applications.

Moreover, the need for diverse and representative data to avoid biases and ensure model generalization further complicates the data collection. Researchers must carefully curate datasets to include a wide range of scenarios, demographics, and edge cases, which can be challenging and resource-intensive.

2.2. Computational Resources:

Training complex computer vision models from scratch requires substantial computational resources, presenting another major challenge in traditional approaches. The depth and complexity of modern neural networks and the size of training datasets result in extremely computationally intensive training processes.

For example, training a ResNet-50 model on the ImageNet dataset can take several days on a single high-end GPU [4]. More advanced models like GPT-3, while not specifically for computer vision but illustrative of the scale of modern AI models, required an estimated 355 GPU-years and millions of dollars in computing costs for a single training run [3].

This high computational demand has several implications:

1. Hardware Requirements: Training state-of-the-art models often necessitates access to high-performance GPUs or specialized hardware like Google's Tensor Processing Units (TPUs). This hardware can be prohibitively expensive for many researchers and smaller organizations.

2. Energy Consumption: The energy required to train large models is substantial, raising concerns about the environmental impact of AI research. Some estimates suggest that training a single large AI model can emit as much carbon as five cars over their lifetimes [3].
3. Time Constraints: Even with powerful hardware, training complex models can take days or weeks, slowing down research iterations and model development.
4. Scalability Issues: As models grow more complex and datasets larger, the computational requirements scale non-linearly, potentially hitting physical and economic limits.

These challenges in data requirements and computational resources have led researchers to explore more efficient approaches, such as transfer learning, which we will discuss in subsequent sections.

Model/Dataset	Number of Images	Training Time	GPU Requirements
AlexNet/ImageNet	1,200,000	Several weeks	2 GPUs
YOLO	Hundreds of thousands	Several days	1 high-end GPU
ResNet-50/ImageNet	1,200,000	Several days	1 high-end GPU
GPT-3	Not applicable	355 GPU-years	Multiple GPUs

Table 1: Data and Resource Demands in Modern Computer Vision Training [3, 4]

3. TRANSFER LEARNING: CONCEPT AND METHODOLOGY:

3.1. Definition:

Transfer learning is a machine learning technique that leverages knowledge gained from solving one problem and applies it to a different but related problem. In computer vision, transfer learning typically involves using a pre-trained model on a large, general dataset (such as ImageNet) and adapting it to a new, specific task with a smaller dataset [5].

This approach is based on the observation that neural networks learn hierarchical features, with earlier layers capturing generic, low-level features (like edges and textures) applicable across a wide range of visual tasks. By reusing these learned features, transfer learning allows models to perform highly on new tasks with significantly less training data and computational resources.

The process generally involves two main steps:

1. Taking a pre-trained model (e.g., a CNN trained on ImageNet)
2. Fine-tuning this model on a new, task-specific dataset, often by retraining only the last few layers while keeping the earlier layers frozen

3.2. Pre-trained Models:

Several popular pre-trained models have emerged as go-to choices for transfer learning in computer vision. These models are typically trained on large-scale datasets like ImageNet and have proven effective starting points for various tasks. Some of the most widely used pre-trained models include:

1. ResNet (Residual Networks): Introduced by He, ResNet models use skip connections to enable the training of very deep networks. ResNet-50 and ResNet-101 are commonly used variants [6].
2. VGGNet: Developed by the Visual Geometry Group at Oxford, VGG networks are known for their simplicity and depth. VGG16 and VGG19 are popular choices.
3. Inception: Google's Inception models use inception modules with multiple convolution filters of different sizes to capture features at various scales.
4. EfficientNet: These models, proposed by Tan and Le, use compound scaling to balance network depth, width, and resolution, achieving state-of-the-art performance with fewer parameters.
5. DenseNet: Dense Convolutional Networks connect each layer to every other layer in a feed-forward fashion, strengthening feature propagation and encouraging feature reuse.

3.3. Feature Extraction:

The power of transfer learning in computer vision lies in the hierarchical nature of features learned by deep neural networks. Pre-trained models capture generalizable lower-level features in their earlier layers, which can be effectively repurposed for new tasks [5].

The feature hierarchy in CNNs typically follows this pattern:

- Early layers: Detect basic features like edges, corners, and color blobs
- Middle layers: Combine these to recognize more complex shapes and textures
- Later layers: Identify high-level concepts specific to the training task (e.g., object parts or entire objects)

The assumption is that the lower and middle-level features learned from a large, diverse dataset like ImageNet will be useful for various visual tasks when using transfer learning. Visual elements (like edges, textures, and basic shapes) are common across images and tasks.

For example, a model pre-trained on ImageNet might have learned to detect fur textures, eye shapes, and ear contours. These features would be useful for classifying animals in the original task. They could be leveraged for more specific tasks like detecting particular dog breeds or identifying wildlife in camera trap images.

By reusing these pre-learned features, transfer learning allows models to perform highly on new tasks with much less task-specific training data. This approach is particularly powerful for domains where large labeled datasets are complex or expensive, such as medical imaging or specialized industrial applications [6].

Pre-trained Model	Year Introduced	Typical Layers	Key Features	Common Applications
ResNet	2015	50-152	Skip connections	Image classification, Object detection
VGGNet	2014	16-19	Simplicity, Depth	Image classification, Transfer learning
Inception	2014	22-48	Multi-scale processing	Image classification, Object detection
EfficientNet	2019	Variable	Compound scaling	Mobile/edge applications
DenseNet	2017	121-201	Dense connectivity	Image classification, Segmentation

Table 2: Comparison of Popular Pre-trained Models for Transfer Learning in Computer Vision [5, 6]

4. THE TRANSFER LEARNING PROCESS:

4.1. Model Selection:

Choosing an appropriate pre-trained model is crucial for successful transfer learning. The selection process should consider several factors:

1. Task Similarity: The pre-trained model should ideally be trained on a task similar to the target task. For instance, a model pre-trained on natural images might not be the best choice for medical imaging tasks [7].
2. Model Architecture: Different architectures have varying strengths. For example, ResNet models are known for their ability to train deep networks, while EfficientNet models offer a good balance between accuracy and computational efficiency.
3. Model Size and Computational Requirements: Larger models may offer higher accuracy but require more computational resources. The choice should balance performance with available hardware constraints.
4. Dataset Size: A smaller pre-trained model might be preferable to avoid overfitting if the target dataset is small.
5. Performance on Benchmark Datasets: Models with higher accuracy on standard benchmarks like ImageNet often transfer well to other tasks.
6. Availability of Pre-trained Weights: Models with readily available pre-trained weights in popular deep learning frameworks are more convenient.

Research by Kornblith [7] suggests that models with better ImageNet performance generally transfer better to other tasks, but this is not always the case. They recommend considering both ImageNet performance and the specific characteristics of the target task when selecting a pre-trained model.

4.2. Fine-tuning:

Fine-tuning involves adapting the pre-trained model to the specific task at hand. This process typically involves the following steps:

1. Replace the Final Layer: The last fully connected layer of the pre-trained model, which was specific to the original classification task, is replaced with a new layer appropriate for the target task.
2. Data Preparation: The target dataset is prepared, often involving resizing images to match the input size of the pre-trained model and applying appropriate data augmentation techniques.
3. Learning Rate Selection: A lower learning rate is typically used for fine-tuning than training from scratch to avoid drastically altering the pre-trained weights.
4. Gradual Unfreezing: As suggested by Howard and Ruder [8], gradually unfreezing layers from top to bottom can improve performance. This technique, known as "discriminative fine-tuning," applies different learning rates to different layers.
5. Regularization: Techniques like weight decay and dropout are often employed to prevent overfitting, especially when the target dataset is small.
6. Monitoring and Early Stopping: The model's performance on a validation set is closely monitored, and training is stopped when performance plateaus to prevent overfitting.

4.3. Layer Freezing/UnFreezing:

Layer unfreezing is a key strategy in transfer learning. It involves selectively training specific layers of the model while keeping others fixed. This approach is based on the observation that earlier layers of a neural network capture more general features, while later layers are more task-specific [8].

Common layer freezing strategies include:

1. Freeze All But the Last Few Layers: This approach keeps the pre-trained weights for most of the network and only trains the final layers. It's particularly useful when the target dataset is small or very similar to the pre-training dataset.
2. Gradual Unfreezing: Starting with all layers frozen except the last one, gradually unfreezing more layers as training progresses. This allows the model to slowly adapt to the new task while retaining useful pre-trained features.
3. Fine-tune All Layers: When the target dataset is large and significantly different from the pre-training dataset, fine-tuning all layers may yield better results.
4. Layer-wise Fine-tuning: Apply different learning rates to different layers, typically using lower rates for earlier layers and higher rates for later layers.

The optimal freezing strategy often depends on the specific task and dataset. Experimentation and cross-validation are typically used to determine the best approach for a given problem.

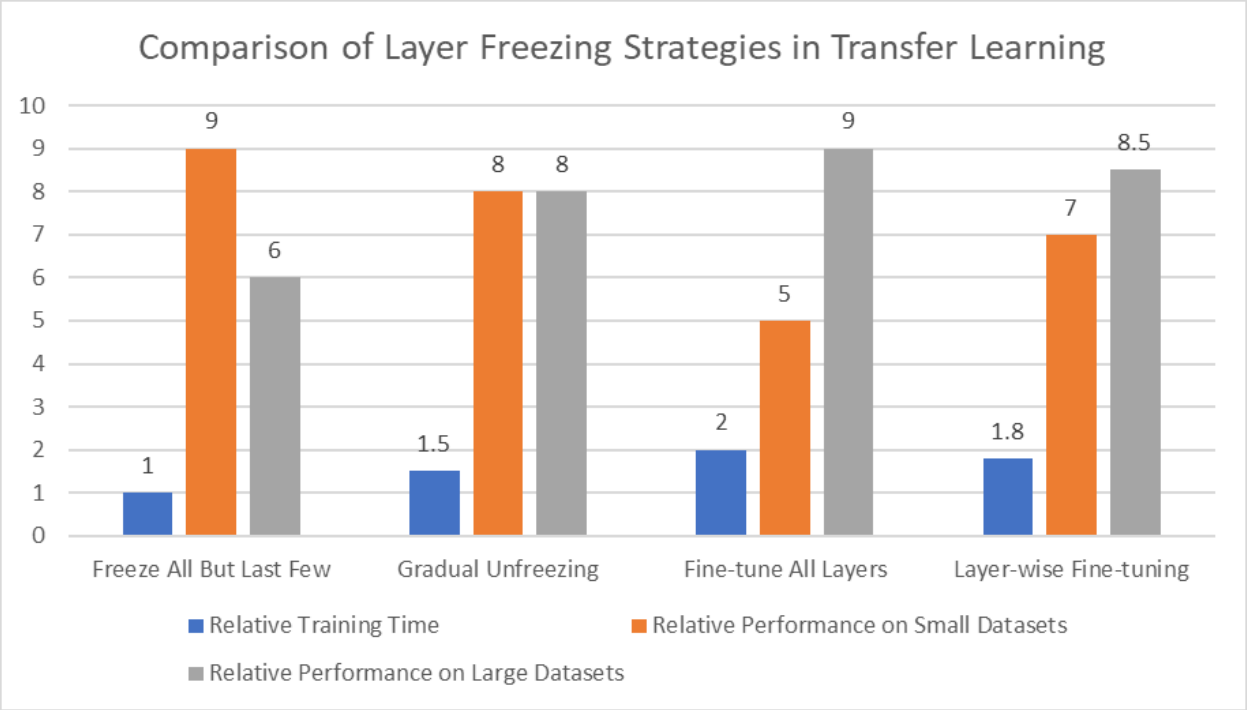


Fig. 1: Impact of Different Layer Freezing Approaches on Training Time and Model Performance [7, 8]

5. ADVANTAGES OF TRANSFER LEARNING:

5.1. Reduced Data Requirements:

One of the most significant advantages of transfer learning is its ability to dramatically reduce the amount of labeled data required for training effective models. This is particularly crucial in domains where large-scale labeled datasets are scarce or expensive.

Transfer learning leverages the knowledge embedded in pre-trained models, which have already learned generalizable features from vast datasets like ImageNet. These pre-learned features can be repurposed for new tasks, even with limited task-specific data.

A study by Kornblith [9] demonstrated that transfer learning could achieve competitive performance on various image classification tasks using only a fraction of the data required for training from scratch. In some cases, models fine-tuned on just 1% of the full dataset achieved performance comparable to models trained from scratch on the entire dataset.

This data efficiency is particularly valuable in specialized domains such as medical imaging, where large annotated datasets are often unavailable due to privacy concerns and the high cost of expert annotation. For instance, Raghu [10] showed that transfer learning could perform state-of-the-art chest X-ray classification tasks using significantly smaller datasets than training from scratch.

5.2. Accelerated Training:

Transfer learning substantially reduces the time required to train models for new tasks. By starting with pre-trained weights, the model begins from a much more informed state, requiring fewer iterations to converge on the new task.

The acceleration in training time can be orders of magnitude. For example, He [9] reported that fine-tuning a ResNet-50 model pre-trained on ImageNet for a new task took only a few hours on a single GPU, compared to days or weeks required for training the same architecture from scratch.

This speedup is particularly pronounced for larger models and datasets. In a study on natural language processing tasks, Howard and Ruder [10] demonstrated that their transfer learning approach could reduce training time by up to 100 times while achieving better performance than models trained from scratch.

The reduced training time has several benefits:

1. Faster iteration in research and development
2. Lower computational costs
3. Reduced energy consumption and environmental impact
4. Increased accessibility of deep learning techniques to researchers with limited computational resources

5.3. Improved Performance:

The most compelling advantage of transfer learning is its ability to improve model performance, especially in scenarios with limited task-specific data.

Numerous studies have shown that models leveraging transfer learning consistently outperform those trained from scratch, particularly when working with smaller datasets. For instance, Kornblith [9] demonstrated that transfer learning from ImageNet pre-trained models improved performance across a wide range of image classification tasks, with the most significant gains observed on smaller datasets.

In the medical imaging domain, Raghu [10] showed that transfer learning matched and often exceeded the performance of models trained from scratch, even when using only a fraction of the available data. Their study on chest X-ray classification achieved state-of-the-art results using transfer learning, outperforming models trained from scratch on the full dataset.

The performance benefits of transfer learning are attributed to several factors:

1. Leveraging pre-learned features that capture general visual patterns
2. Starting from a better initialization point in the parameter space
3. Regularization effect of the pre-trained weights, reducing overfitting on small datasets
4. Ability to capture task-relevant features more efficiently during fine-tuning

It's worth noting that the performance gains from transfer learning can vary depending on the similarity between the source and target tasks. However, even when tasks are quite different, transfer learning often provides a better starting point than random initialization.

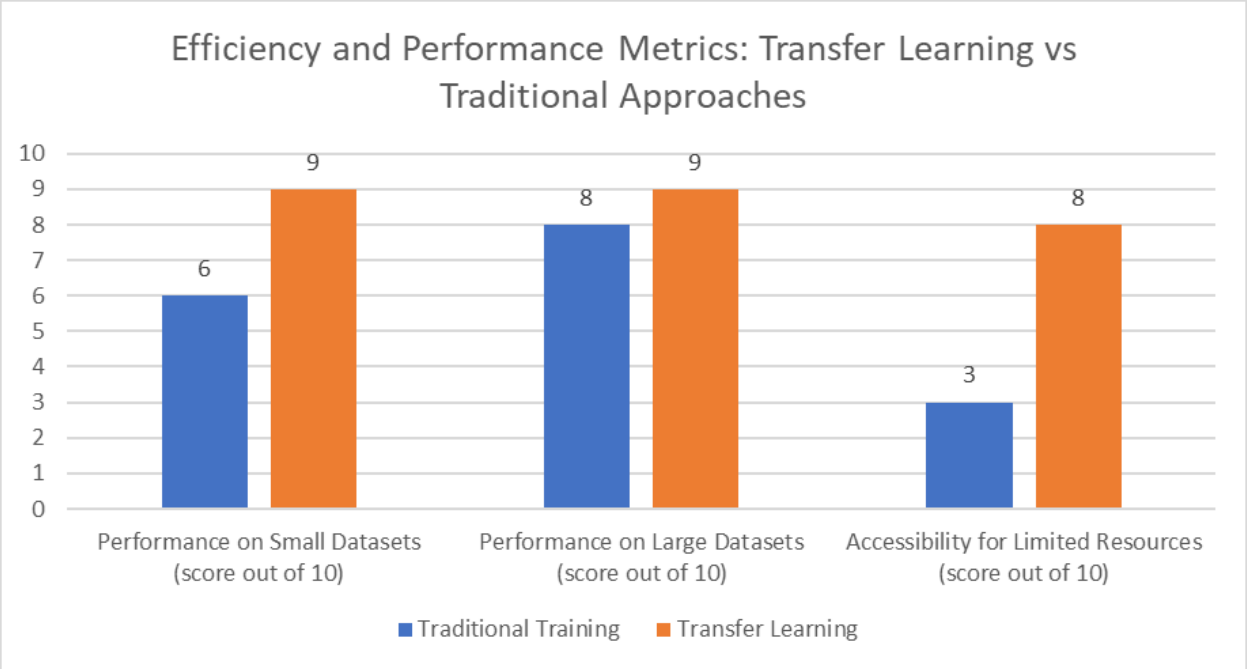


Fig. 2: Comparison of Traditional Training vs Transfer Learning in Deep Learning Models [9, 10]

6. REAL-WORLD APPLICATIONS

6.1 Object Detection:

Transfer learning has significantly advanced the field of object detection, enabling faster development of accurate models for various applications. One notable case study is the development of the YOLOv3 (You Only Look Once version 3) object detection system by Redmon and Farhadi [11]. The authors utilized transfer learning by pre-training their model on the ImageNet dataset before fine-tuning it for object detection tasks. This approach allowed YOLOv3 to achieve state-of-the-art performance while maintaining real-time detection speeds.

In another compelling application, researchers at Stanford University leveraged transfer learning to develop a system for detecting pneumonia in chest X-rays [12]. To identify signs of pneumonia in medical images, they fine-tuned a DenseNet-121 model, pre-trained on ImageNet. The resulting model achieved performance on par with expert radiologists, demonstrating the power of transfer learning in specialized domains with limited labeled data.

6.2. Image Classification:

Image classification is the most common application of transfer learning in computer vision. A prominent example is the use of transfer learning in skin cancer detection. Esteva [12] fine-tuned a Google Inception v3 model, pre-trained on ImageNet, to classify skin lesions. Their model achieved dermatologist-level accuracy in distinguishing between benign and malignant skin lesions, showcasing the potential of transfer learning in medical diagnostics.

In environmental conservation, transfer learning has been applied to classify and count animals in camera trap images. Norouzzadeh [11] used transfer learning with several popular architectures (including ResNet and Inception) to automatically identify species in the Snapshot Serengeti dataset. Their approach significantly reduced the manual labor required for wildlife monitoring, processing years' worth of images in weeks.

6.3. Other Computer Vision Tasks:

Transfer learning has proven valuable across a wide range of computer vision tasks beyond object detection and classification:

1. Semantic Segmentation: In medical imaging, transfer learning has been applied to segment organs and tumors in MRI and CT scans. For instance, researchers have fine-tuned U-Net architectures, pre-trained on large natural image datasets, to accurately segment brain tumors in MRI images [12].
2. Facial Recognition: Transfer learning has enhanced facial recognition systems by allowing models to generalize better across different demographics and imaging conditions. By fine-tuning models pre-trained on large, diverse face datasets, researchers have improved recognition accuracy while reducing bias [11].
3. Style Transfer: Transfer learning has created sophisticated style transfer algorithms in computational art. By leveraging features learned from large art datasets, these models can apply one image's style to another's content with impressive results.
4. Action Recognition in Videos: Researchers have applied transfer learning to recognize human actions in video sequences. By fine-tuning 3D convolutional networks pre-trained on large video datasets, they've achieved state-of-the-art performance in action recognition tasks [12].
5. Autonomous Driving: Transfer learning plays a crucial role in developing perception systems for autonomous vehicles. Models pre-trained on diverse driving scenarios can be fine-tuned for specific environments or weather conditions, improving the robustness and safety of self-driving systems.
6. Satellite Image Analysis: In remote sensing, transfer learning has been used to analyze satellite imagery for tasks such as land use classification, crop yield prediction, and urban planning. Models pre-trained on large-scale satellite image datasets can efficiently adapt to specific geographic regions or detection tasks.

These diverse applications demonstrate the versatility and power of transfer learning in computer vision. Researchers and practitioners can rapidly develop high-performance solutions for various visual tasks by leveraging pre-trained models, even in domains with limited labeled data or specialized requirements.

7. FUTURE DIRECTIONS:

7.1. Advanced Transfer Learning Techniques:

As transfer learning continues to evolve, researchers are exploring advanced techniques to enhance its effectiveness further and broaden its applicability. Some promising directions include:

1. Meta-learning: Also known as "learning to learn," meta-learning aims to develop models that can quickly adapt to new tasks with minimal fine-tuning. Finn [13] proposed Model-Agnostic Meta-Learning (MAML), which learns a set of initial parameters that can be rapidly adapted to new tasks. This approach has shown promising results in few-shot learning scenarios, where only a handful of examples are available for the target task.
2. Multi-task Learning: This approach involves simultaneously training models on multiple related tasks, allowing the model to learn more general and robust features. Zamir [14] demonstrated that multi-task learning can improve performance and faster adaptation to new tasks in various computer vision domains.

3. Unsupervised and Self-supervised Learning: These techniques leverage large amounts of unlabeled data to learn useful representations. Recent advancements, such as contrastive learning methods like SimCLR, have shown that models pre-trained on unlabeled data can achieve performance comparable to or even surpass supervised pre-training in certain scenarios [13].
4. Neural Architecture Search (NAS) for Transfer Learning: NAS techniques are being applied to discover optimal architectures for transfer learning automatically. This approach promises to create more efficient and effective models tailored to specific transfer learning scenarios.
5. Attention Mechanisms: Incorporating attention mechanisms into transfer learning models can help focus on the most relevant features for the target task. This is particularly useful when transferring knowledge between domains with significant differences [14].

7.2. Cross-domain Transfer

Cross-domain transfer learning, which involves transferring knowledge between significantly different domains, remains a challenging and active area of research. Some key areas of exploration include:

1. Domain Adaptation: Techniques are being developed to bridge the gap between source and target domains with different data distributions. Ganin [13] proposed domain adversarial training, which aims to learn domain-invariant features that perform well on source and target domains.
2. Zero-shot and Few-shot Learning: These approaches aim to recognize or classify objects from categories not seen during training or for which only a few examples are available. Recent work by Xian [14] has shown promising results in transferring knowledge from seen to unseen categories in image classification tasks.
3. Cross-modality Transfer: Researchers are exploring ways to transfer knowledge between different modalities, such as from image to text or vice versa. This has applications in areas like image captioning and visual question answering.
4. Transfer Learning in 3D Vision: As 3D computer vision tasks gain importance, there's growing interest in developing transfer learning techniques specifically for 3D data. This includes transferring knowledge from 2D to 3D domains and between different types of 3D representations (e.g., point clouds, voxels, meshes) [13].
5. Federated Learning: This approach allows training models across multiple decentralized devices or servers holding local data samples without exchanging them. It's particularly relevant for scenarios where data privacy is crucial, such as healthcare applications [14].
6. Continual Learning: Also known as lifelong learning, this area focuses on developing models that can continuously learn from data from different tasks or domains without forgetting previously learned knowledge.
7. Transferring Knowledge to Resource-Constrained Devices: With the increasing deployment of AI on edge devices, there's a growing need for techniques that can effectively transfer knowledge from large, powerful models to smaller, more efficient ones suitable for mobile and IoT devices.

These future directions highlight the dynamic nature of transfer learning research in computer vision. As the field progresses, we can expect more powerful and versatile transfer learning techniques that push the boundaries of what's possible in computer vision applications across various domains.

8. CONCLUSION

Transfer learning has emerged as a powerful paradigm in computer vision, offering solutions to longstanding challenges in data scarcity and computational efficiency. Researchers and practitioners can develop high-performance computer vision systems with reduced resources and time by leveraging pre-trained models and adapting them to new tasks. The diverse applications of transfer learning across medical imaging, environmental conservation, autonomous driving, and other domains demonstrate its versatility and effectiveness. As research in advanced transfer learning techniques and cross-domain transfer progresses, we can anticipate more powerful and adaptable computer vision models. These developments will continue democratizing access to state-of-the-art AI capabilities, enabling innovations across various industries and pushing the boundaries of what's possible in computer vision.

REFERENCES

[1] E. Strubell, A. Ganesh, and A. McCallum, "Energy and Policy Considerations for Deep Learning in NLP," in Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 2019, pp. 3645-3650. [Online]. Available: https://aclanthology.org/P19-1355/

[2] J. Deng, "ImageNet: A large-scale hierarchical image database," in 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 248-255. [Online]. Available: https://ieeexplore.ieee.org/document/5206848

[3] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in Advances in Neural Information Processing Systems 25, 2012, pp. 1097-1105. [Online]. Available: https://papers.nips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf

[4] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779-788. [Online]. Available: https://ieeexplore.ieee.org/document/7780460

[5] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?," in Advances in Neural Information Processing Systems 27, 2014, pp. 3320-3328. [Online]. Available: https://papers.nips.cc/paper/2014/file/375c71349b295fbe2dcdca9206f20a06-Paper.pdf

[6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770-778. [Online]. Available: https://ieeexplore.ieee.org/document/7780459

[7] S. Kornblith, J. Shlens, and Q. V. Le, "Do Better ImageNet Models Transfer Better?," in 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 2661-2671. [Online]. Available: https://ieeexplore.ieee.org/document/8953714

[8] J. Howard and S. Ruder, "Universal Language Model Fine-tuning for Text Classification," in Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2018, pp. 328-339. [Online]. Available: https://aclanthology.org/P18-1031/

[9] S. Kornblith, J. Shlens, and Q. V. Le, "Do Better ImageNet Models Transfer Better?," in 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 2661-2671. [Online]. Available: https://ieeexplore.ieee.org/document/8953714

[10] M. Raghu, C. Zhang, J. Kleinberg, and S. Bengio, "Transfusion: Understanding Transfer Learning for Medical Imaging," in Advances in Neural Information Processing Systems 32, 2019, pp. 3347-3357. [Online]. Available: https://papers.nips.cc/paper/2019/file/eb1e78328c46506b46a4ac4a1e378b91-Paper.pdf

[11] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," arXiv preprint arXiv:1804.02767, 2018. [Online]. Available: https://arxiv.org/abs/1804.02767

[12] A. Esteva, "Dermatologist-level classification of skin cancer with deep neural networks," Nature, vol. 542, no. 7639, pp. 115-118, 2017. [Online]. Available: https://www.nature.com/articles/nature21056

[13] C. Finn, P. Abbeel, and S. Levine, "Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks," in Proceedings of the 34th International Conference on Machine Learning, 2017, pp. 1126-1135. [Online]. Available: http://proceedings.mlr.press/v70/finn17a.html

[14] A. R. Zamir, "Taskonomy: Disentangling Task Transfer Learning," in 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 3712-3722. [Online]. Available: https://ieeexplore.ieee.org/document/8578523

Citation: Rathish Mohan, Transfer Learning's Triumph in Computer Vision: A Boon for Efficiency and Performance, International Journal of Computer Engineering and Technology (IJCET), 15(4), 2024, pp. 184-197

Abstract Link: https://iaeme.com/Home/article_id/IJCET_15_04_016

Article Link:
https://iaeme.com/MasterAdmin/Journal_uploads/IJCET/VOLUME_15_ISSUE_4/IJCET_15_04_016.pdf

✉ **editor@iaeme.com**