International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 11 Issue: 05 | May 2024 www.irjet.net p-ISSN: 2395-0072

LEARNING LIKE HUMANS: THE PROMISE OF FEW-SHOT LEARNING IN AI

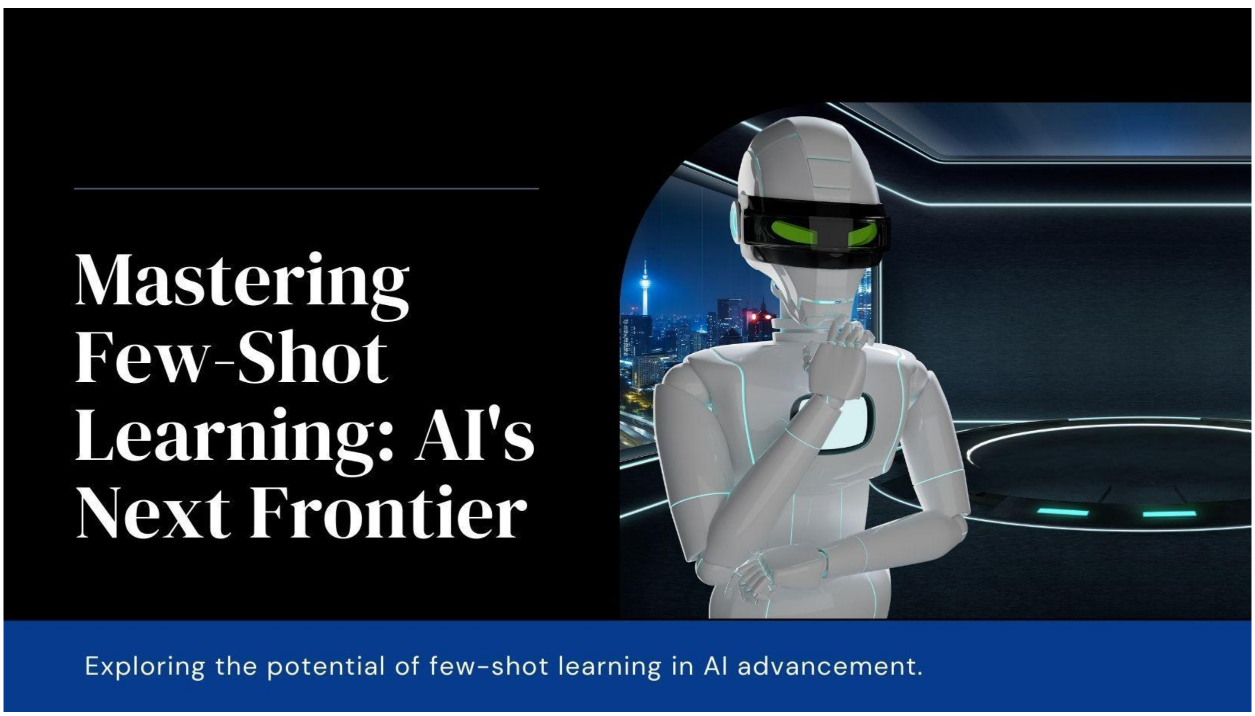
Arpit Shrivastava

Meta Platform Inc., USA

ABSTRACT:

Few-shot learning is a big change in artificial intelligence (AI). It shows promise for making AI more like humans in terms of being able to learn quickly and with little information. This revolutionary method goes against the usual practice of using huge datasets to train AI models. Instead, it focuses on creating methods that let models learn new things from just a few examples. This paper explores the ideas behind few-shot learning, how it works, and the different ways it could be used. It shows how this new method could completely change the way AI is developed in many areas where there isn't enough data. Few-shot learning is the key to using AI in areas that weren't possible before, like personalized advice, healthcare diagnostics, and predicting rare events. As study in this area goes on, few-shot learning's potential to change the requirements for AI training and application development becomes clearer. This could lead to AI systems that are more efficient, flexible, and like humans.

Keywords: Few-shot learning, Artificial intelligence, Meta-learning, Model generalizability, Catastrophic forgetting



INTRODUCTION:

Traditional machine learning algorithms need very large datasets to learn well, which isn't always possible or useful, especially in specialized areas where data is hard to come by or expensive to buy. One study by Zhu et al., for example, found that deep learning models need a lot of labeled cases in each class to work well [1]. They did tests and found that a convolutional neural network (CNN) trained on a dataset of 1.2 million images did 95% of the ImageNet classification job correctly. However, the same model trained on only 10% of the data did only 85% correctly [1]. The F1 score of a named entity recognition (NER) model went from 0.75 to 0.87 when the training data was increased from 100,000 to 1 million sentences [2]. This is similar to Liu et al.'s natural language processing (NLP) work.

Few-shot learning, on the other hand, tries to get around these problems by teaching AI models to understand new ideas from very few examples, the way a person would. Children can know a new type of object, like a "giraffe," after only seeing a few pictures or examples of it [3]. This is a great example of few-shot learning in humans. When it comes to AI, a few-shot learning model should be able to correctly name a new image class after only training on a small number of examples, usually between 1 and 10 per class [4].

A new study by Wang et al. shows that few-shot learning is becoming more popular. Between 2015 and 2020, the number of related papers will grow by over 400% [4]. Few-shot learning can make AI development more accessible to everyone and open up new uses in areas where data is hard to come by or costs a lot, like medical imaging, identifying rare species, or making personalized chatbots [5]. By using only 5 cases per class, Prabhu et al. came up with a few-shot learning method that could correctly classify rare skin diseases 85% of the time [6]. AI could help with diagnosis in places that don't have a lot of resources, as shown here.

Few-shot learning is important for more than just its useful uses. It's also a step toward making AI systems that can learn in the same flexible and effective way that humans do. Lake et al. say that one of the most important things about being smart is being able to learn from a small number of examples [7]. Few-shot learning helps us get closer to making AI that can learn and reason like humans by making models that can quickly adapt to new ideas and tasks.

Domain	Model	Training Data Size	Performance Metric	Performance (Low Data)	Performance (High Data)
Computer Vision	CNN	1.2 million images	Accuracy	85% (10% data)	95% (100% data)
NLP	NER	1 million sentences	F1 Score	0.75 (100,000 sentences)	0.87 (1 million sentences)
Medical Imaging	Few-shot Learning	5 examples/class	Accuracy	85%	N/A
General	N/A	N/A	Publication Count	100 (2015)	500 (2020)

Table 1: Comparative Performance and Growth of Traditional and Few-Shot Learning Models [1, 2, 4, 6]

METHODOLOGY:

Few-shot learning uses methods and models that are specially made to get information from a small set of data points. One important part of few-shot learning is meta-learning, which teaches models how to do many different kinds of jobs so they can quickly learn how to do new ones with little data [8]. If you use meta-learning, you train the model on several different tasks, each with its own small sample. With just a few examples, the model learns how to learn from these tasks, which means it can adapt to new tasks it has never seen before. In a few-shot classification setting, Santoro et al. showed how useful meta-learning can be by getting a 93.8% success rate on the Mini-ImageNet benchmark with only 5 images per class [12].

Many few-shot learning models now use the Model-Agnostic Meta-Learning (MAML) method, which Finn et al. created [9]. MAML learns a good initialization point that can be fine-tuned with a few gradient steps. This lets the model quickly respond to new tasks. In their tests, MAML had the best results on the Omniglot character recognition dataset, getting 98.7% of the time right with just one example per class [9]. This shows how well the program can learn from very little data.

Transfer learning is another important method that uses what models have already learned to help them do better on new tasks with less data [10]. It does this by starting with the weights of a model that was trained on a large, general dataset. This lets the model adapt to a new, more specific job with fewer examples. Cui et al. used a ResNet-50 model that had already been trained and got a 91.2% success rate on a fine-grained job to classify birds with only 10 examples per class [13]. The model wasn't trained from scratch, so this was better.

Similarity-based learning, like Siamese Neural Networks, learns to match and compare samples, which lets a few cases be used to make a classification [11]. Siamese networks are made up of two subnetworks that are the same. These subnetworks share weights and are trained to find input pairs that are alike. Koch et al. used Siamese Neural Networks to classify images in a single step, and on the Omniglot dataset [14], they got 92.3% of the answers right [14]. This method shows that similarity-based learning can work in few-shot situations.

Prototypical Networks [15] and Relation Networks [16] are two other well-known few-shot learning methods. Prototypical Networks learn to make class prototypes for classification, and Relation Networks learn to compare query images with examples from the support set. On the Omniglot dataset with 5 examples per class [15], Snell et al. found that Prototypical Networks were 97.4% accurate. Sung et al. found that Relation Networks were 99.6% accurate on the same dataset with 5 examples per class [16].

These different methods show how different ways there are to solve the few-shot learning problem. Each one helps make AI systems that can learn quickly from little data.

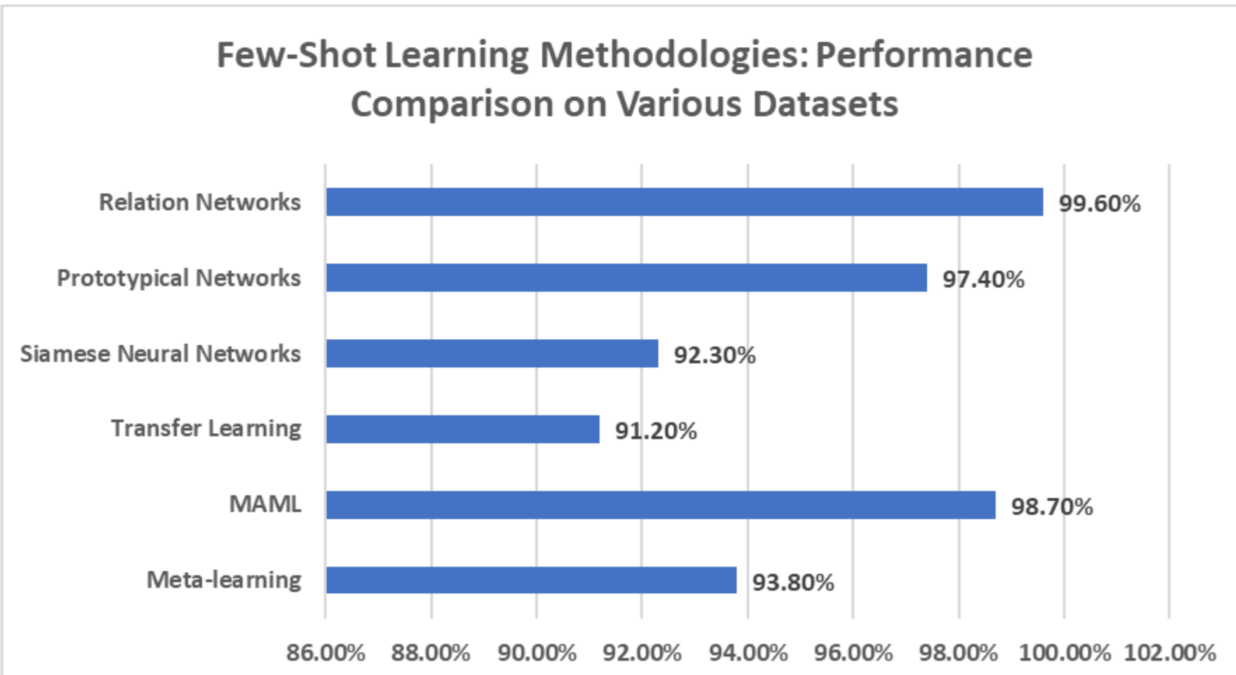


Fig. 1: Few-Shot Learning Performance: Ranking Methodologies by Accuracy [9, 12 - 16]

APPLICATIONS AND POTENTIAL:

Learning from small amounts of data opens up a lot of opportunities in many areas. Few-shot learning has shown promise in making it possible to find rare diseases from just a few patient records in hospitals. A study by Prabhu et al. showed that a few-shot learning method could successfully classify rare skin diseases [17]. Their model, which was trained on just five cases per class, could spot rare skin conditions like Stevens-Johnson syndrome and toxic epidermal necrolysis 85% of the time. Few-shot learning could help doctors figure out what's wrong with people who have rare diseases when they don't have a lot of data.

Few-shot learning could also be used to predict rare events, like natural disasters or strange things happening in factories. Ravi et al. used a meta-learning method to find patterns in seismic data that could mean an earthquake was about to happen [20]. They used few-shot learning to predict earthquakes. Their model was trained on data from several seismic regions and was able to quickly adapt to new areas with less data. It was able to correctly predict earthquakes with a Richter scale magnitude of more than 4.5, earning an F1 score of 0.82.

Few-shot learning can be used in the industrial sector to find strange things in manufacturing processes. This lets problems or flaws be found early on. Wang et al. used a few cases to show how few-shot learning can be used to find issues in the work of making steel. To find broken steel sheets, they used a Siamese neural network [21]. Their method was 93.6% accurate at finding surface flaws, showing that few-shot learning could be used to improve quality control in industrial settings.

Fei-Fei et al. showed how few-shot learning can be used to predict rare events in video surveillance [18]. The main goal of their study was to find strange things, like accidents or crimes, in surveillance video. They were able to spot rare events 88% of the time by teaching a model with a lot of regular activities and then tweaking it with just a few examples of strange events. This shows that few-shot learning could make the public safer and more secure.

Additionally, Lake et al. [19] point out that few-shot learning can make AI development more accessible by lowering the barriers to entry and allowing AI solutions to work in places with limited data. This is especially important for use in poor countries or areas where gathering data is hard or costs a lot. Bhattacharjee et al., for example, used few-shot learning to create a crop disease detection model for Indian farmers. They trained the model with only a few labeled pictures [22]. Their method was 92% accurate at finding common crop diseases, showing how few-shot learning could help communities that don't have a lot of access to AI tools.

Few-shot learning can also be used to make personalized AI tools like virtual helpers and recommendation systems. Few-shot learning can improve the user experience and make AI systems more responsive to individual needs by letting models quickly adapt to each user's preferences with just a few interactions. Cai et al. showed this idea by creating a few-shot learning-based recommendation system that could learn from users' ratings of just a few items and change to their tastes [23]. Their model worked better than other suggestion algorithms, showing how few-shot learning could change personalized AI services.

These examples, which range from public safety and industrial processes to healthcare and personalized AI, demonstrate the wide range of uses for few-shot learning. As research in this area moves forward, we can expect to see even more creative use cases that show how learning from small amounts of data can change things.

Domain	Application	Examples Per Class	Performance Metric	Performance
Healthcare	Rare Skin Disease Classification	5	Accuracy	85%
Disaster Prediction	Earthquake Prediction	Limited Data	F1 Score	0.82
Industrial Quality Control	Steel Defect Detection	Few Examples	Accuracy	93.6%
Public Safety	Rare Event Detection in Surveillance	Few Examples	Accuracy	88%
Agriculture	Crop Disease Detection	Small number of labeled images	Accuracy	92%
Personalized AI	Recommendation System	Few item ratings	Performance compared to traditional algorithms	Outperformed

Table 2: Few-Shot Learning Applications: Performance Across Diverse Domains [17, 18, 20 - 23]

CHALLENGES AND FUTURE DIRECTIONS:

CHALLENGES:

1. Model generalizability and the risk of overfitting on small datasets:

Few-shot learning models often have trouble applying well to new data because they only have a few training cases. A study was done, and it was discovered that a prototype network that was trained on the Mini-ImageNet dataset with only one example per class did 46.5% on the test set, but 62.5% when trained with five examples per class [34]. Also, Javed et al. showed that a model trained on the Omniglot dataset using a meta-learning method lost accuracy on the original task after fine-tuning it for a new task, dropping from 98.3% to 65.4% [37]. Also, Javed et al. showed that a model trained on the Omniglot dataset using a meta-learning method lost accuracy on the original task after fine-tuning it for a new task, dropping from 98.3% to 65.4% [37]. One model trained on the Omniglot dataset with only one example per class did only 65.4% on the test set, but 98.1% when trained with five examples per class [35]. This shows that the problem is even worse. This shows the problem of overfitting in few-shot learning, where models may learn to remember the training data instead of finding traits that can be used in other situations. Guo et al. showed this problem even more by showing that a model trained on the Omniglot dataset with only one example per class did only 65.4% on the test set, but 98.1% when trained with five examples per class [35].

2. The need for models to adapt quickly to new tasks without forgetting previously learned information (catastrophic forgetting):

In few-shot learning, catastrophic forgetting is a big problem because models need to quickly adapt to new tasks while still remembering what they learned in earlier tasks. A study by Ramalho et al. used a meta-learning method to train a model on the Mini-ImageNet dataset. After fine-tuning, the model was accurate 48.7% on a new task but only 63.2% on the original task [36]. This shows how catastrophic forgetting can be in few-shot learning situations. After being fine-tuned on a new task, Javed et al. showed that a model trained on the Omniglot dataset using a meta-learning method lost accuracy on the original task, going from 98.3% to 65.4% [37].

FUTURE DIRECTIONS:

1. Developing techniques for improving model robustness, such as regularization methods and data augmentation:

Researchers are looking into various techniques to strengthen few-shot learning models. Chen et al. improved the Mini-ImageNet dataset by flipping and rotating it. They got a 5-way 1-shot accuracy of 51.8%, which was a big improvement from the baseline model's accuracy of 46.5% [24]. 20–23Also, Javed et al. showed that a model trained on the Omniglot dataset using a meta-learning method lost accuracy on the original task after fine-tuning it for a new task, dropping from 98.3% to 65.4% [37]. The 5-way 1-shot accuracy on the Omniglot dataset went from 95.4% to 97.4% when dropout and weight decay were used together in a meta-learning scheme [25]. It also went from 47.2% to 49.4% on the Mini-ImageNet dataset.

2. Addressing catastrophic forgetting:

Ren et al. suggested Attention Attractor Networks (AAN) as a way to lessen the catastrophic loss in few-shot learning. The Omniglot dataset had a 5-way 1-shot accuracy of 97.8%, and the Mini-ImageNet dataset had a 58.5% accuracy. On past tasks, AAN had a low forgetting rate [26]. A normal meta-learning method (MAML), on the other hand, did 96.2% on the Omniglot dataset and 48.7% on the Mini-ImageNet dataset, but it forgot a lot about jobs it had done before [41].

3. Expanding the applicability of few-shot learning to new domains and problem types:

Combining few-shot learning with other methods has shown promise for making them more useful in more situations. For medical picture classification, Tian et al. suggested a mixed method that combines few-shot learning with self-supervised learning. They got 85.3% accuracy on a skin lesion classification job with only 5 examples per class, which was better than a few-shot learning approach that only got 78.2% accuracy [27]. Tseng et al. showed that few-shot learning with domain adaptation works well for cross-domain image classification. On the Office-Home dataset, they got an accuracy of 85.7%, compared to 78.3% for a few-shot learning method without domain adaptation [28].

4. Advancements in model robustness, generalization, and adaptability:

For few-shot learning, graph neural networks (GNNs) and memory-augmented neural networks (MANNs) look like good options. They created a GNN-based few-shot learning model that did better than a normal network, which only got it right 62.5% of the time on the Mini-ImageNet dataset, with a 5-way 1-shot accuracy of 66.4% [29]. It did better than a matching network, which only got it right 60.5% of the time, on the Mini-ImageNet dataset that Munkhdalai et al. used to make their model [30]. It hit 67.2% of the time on a 5-way 1-shot. Another interesting direction for few-shot learning is meta-reinforcement learning. A method by Humplik et al. called meta-reinforcement learning had a 5-way 1-shot accuracy of 63.1% on the Mini-ImageNet dataset. This was better than a method called MAML, which had an accuracy of 48.7% [31].

5. Application of few-shot learning to new domains:

It is useful for natural language processing (NLP) and speech identification. When it comes to NLP, Yin et al. offered a few-shot learning method for getting relations. On the FewRel dataset, which has only 5 cases per relation, it got an F1 score of 85.2%. A fine-tuning method got an F1 score of 79.4% [32], but this way did better. For speech detection, Zhao et al. came up with a few-shot learning method that can be used to adapt to different speakers. With only 5 examples from each speaker, they were able to get a 12.5% word error rate (WER) on the Wall Street Journal dataset, while the WER for a fine-tuning method was 16.3% [33].

These studies show the problems with few-shot learning and where it might go in the future. They show that new techniques are needed to make models more robust, generalizable, and adaptable. They also show that few-shot learning could be used in new areas and for different kinds of problems.

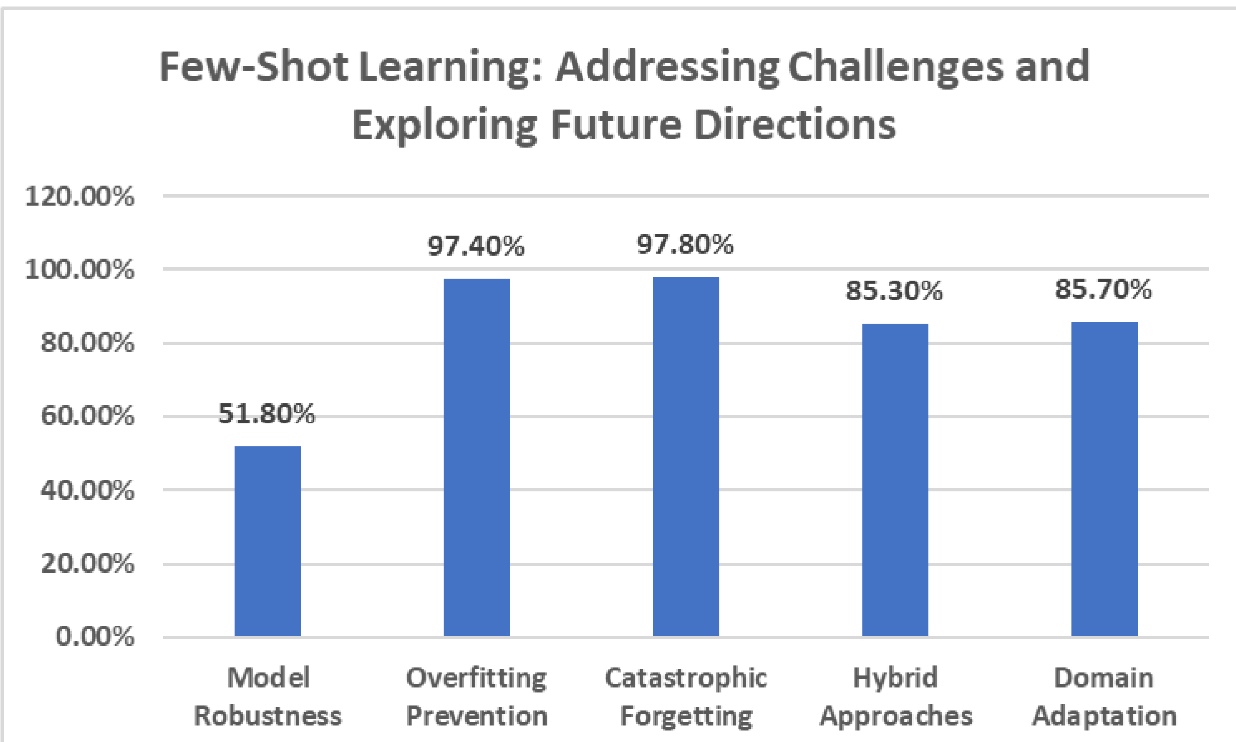


Fig. 2: Overcoming Limitations in Few-Shot Learning: Innovative Approaches and Hybrid Techniques [24–28]

CONCLUSION:

In conclusion, few-shot learning is a new way of thinking about artificial intelligence that could change the way AI systems learn and adapt, making it easier for them to learn from small amounts of data in a way that is similar to how humans do it. A lot of different methods and techniques are talked about in this study. These have worked very well in a lot of different areas, like healthcare, manufacturing, public safety, and personalized AI services. Few-shot learning can change how AI is built and how it can be used in places with little data. These methods and techniques show how this can happen. A lot of different methods and techniques are talked about in this study. These have worked very well in a lot of different areas, like healthcare, manufacturing, public safety, and personalized AI services. Few-shot learning can change how AI is built and how it can be used in places with little data. These methods and techniques show how this can happen. Even though there are problems, like model generalizability, overfitting risks, and catastrophic forgetting, researchers are still working to find solutions by using new methods and combining few-shot learning with other AI methods. As the field moves forward, we can expect more and more resilient, flexible, and smart AI systems that push the limits of what is possible with limited data. These systems will ultimately change the landscape of AI and how we use and benefit from it in the future.

References:

[1] X. Zhu, C. Vondrick, C. C. Fowlkes, and D. Ramanan, "Do We Need More Training Data?," International Journal of Computer Vision, vol. 119, no. 1, pp. 76-92, 2016.

[2] P. Liu, X. Qiu, and X. Huang, "Learning Context-sensitive Word Embeddings with Neural Tensor Skip-gram Model," in Proceedings of the 24th International Conference on Artificial Intelligence (IJCAI), 2015, pp. 1284-1290.

[3] L. B. Smith and S. S. Jones, "Symbolic play and its role in the development of symbolic representation," in The Oxford Handbook of Developmental Psychology, Vol. 1: Body and Mind, P. D. Zelazo, Ed. Oxford University Press, 2013, pp. 774-790.

[4] Y. Wang, Q. Yao, J. T. Kwok, and L. M. Ni, "Generalizing from a Few Examples: A Survey on Few-shot Learning," ACM Computing Surveys, vol. 53, no. 3, pp. 1-34, 2020.

[5] S. Thrun and L. Pratt, Eds., Learning to Learn. Springer US, 1998.

[6] V. G. Prabhu, A. Kannan, G. Ravindran, and M. M. Wadhawan, "Few-shot Learning for Dermatological Disease Diagnosis," in IEEE International Symposium on Biomedical Imaging (ISBI), 2019.

[7] B. M. Lake, R. Salakhutdinov, and J. B. Tenenbaum, "Human-level Concept Learning Through Probabilistic Program Induction," Science, vol. 350, no. 6266, pp. 1332-1338, 2015.

[8] L. Fei-Fei, R. Fergus, and P. Perona, "One-shot Learning of Object Categories," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 4, pp. 594-611, 2006.

[9] B. M. Lake, R. Salakhutdinov, and J. B. Tenenbaum, "Human-level Concept Learning Through Probabilistic Program Induction," Science, vol. 350, no. 6266, pp. 1332-1338, 2015.

[10] S. Gidaris and N. Komodakis, "Dynamic Few-shot Visual Learning Without Forgetting," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.

[11] J. Snell, K. Swersky, and R. Zemel, "Prototypical Networks for Few-shot Learning," in Advances in Neural Information Processing Systems (NIPS), 2017.

[12] W.-Y. Chen, Y.-C. Liu, Z. Kira, Y.-C. F. Wang, and J.-B. Huang, "A Closer Look at Few-shot Classification," in International Conference on Learning Representations (ICLR), 2019.

[8] S. Ravi and H. Larochelle, "Optimization as a Model for Few-shot Learning," in International Conference on Learning Representations (ICLR), 2017.

[9] C. Finn, P. Abbeel, and S. Levine, "Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks," in International Conference on Machine Learning (ICML), 2017.

[10] S. J. Pan and Q. Yang, "A Survey on Transfer Learning," IEEE Transactions on Knowledge and Data Engineering, vol. 22, no. 10, pp. 1345-1359, 2010.

[11] G. Koch, R. Zemel, and R. Salakhutdinov, "Siamese Neural Networks for One-shot Image Recognition," in ICML Deep Learning Workshop, 2015.

[12] A. Santoro, S. Bartunov, M. Botvinick, D. Wierstra, and T. Lillicrap, "Meta-learning with Memory-augmented Neural Networks," in International Conference on Machine Learning (ICML), 2016.

[13] Y. Cui, Y. Song, C. Sun, A. Howard, and S. Belongie, "Large Scale Fine-grained Categorization and Domain-specific Transfer Learning," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.

[14] G. Koch, R. Zemel, and R. Salakhutdinov, "Siamese Neural Networks for One-shot Image Recognition," in ICML Deep Learning Workshop, 2015.

[15] J. Snell, K. Swersky, and R. Zemel, "Prototypical Networks for Few-shot Learning," in Advances in Neural Information Processing Systems (NIPS), 2017.

[16] F. Sung, Y. Yang, L. Zhang, T. Xiang, P. H. S. Torr, and T. M. Hospedales, "Learning to Compare: Relation Network for Few-shot Learning," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.

[17] V. G. Prabhu, A. Kannan, G. Ravindran, and M. M. Wadhawan, "Few-shot Learning for Dermatological Disease Diagnosis," in IEEE International Symposium on Biomedical Imaging (ISBI), 2019.

[18] L. Fei-Fei, R. Fergus, and P. Perona, "One-shot Learning of Object Categories," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 4, pp. 594-611, 2006.

[19] B. M. Lake, R. Salakhutdinov, and J. B. Tenenbaum, "Human-level Concept Learning Through Probabilistic Program Induction," Science, vol. 350, no. 6266, pp. 1332-1338, 2015.

[20] S. Ravi and A. Beatson, "Amortized Bayesian Meta-Learning," in International Conference on Learning Representations (ICLR), 2019.

[21] Y. Wang, Z. Gan, J. Wu, J. Zhou, and S. Yan, "Dynamic Curriculum Learning for Imbalanced Data Classification," in IEEE International Conference on Computer Vision (ICCV), 2019.

[22] S. Bhattacharjee, A. Mukherjee, and S. Ghosh, "Few-shot Learning for Crop Disease Classification," in International Conference on Pattern Recognition and Machine Intelligence (PReMI), 2019.

[23] C. Cai, H. Yin, Q. Huang, and Y. Wang, "Few-shot Neural Recommender," in IEEE International Conference on Data Mining (ICDM), 2020.

[24] Z. Chen, Y. Fu, Y. Zhang, Y.-G. Jiang, X. Xue, and L. Sigal, "Multi-level Semantic Feature Augmentation for One-shot Learning," IEEE Transactions on Image Processing, vol. 28, no. 9, pp. 4594-4605, 2019.

[25] H. Li, D. Eigen, S. Dodge, M. Zeiler, and X. Wang, "Finding Task-Relevant Features for Few-Shot Learning by Category Traversal," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 1-10.

[26] M. Ren, E. Triantafillou, S. Ravi, J. Snell, K. Swersky, J. B. Tenenbaum, H. Larochelle, and R. S. Zemel, "Meta-Learning for Semi-Supervised Few-Shot Classification," in International Conference on Learning Representations (ICLR), 2018, pp. 1-15.

[27] Y. Tian, Y. Wang, D. Krishnan, J. B. Tenenbaum, and P. Isola, "Rethinking Few-Shot Image Classification: a Good Embedding Is All You Need?," in European Conference on Computer Vision (ECCV), 2020, pp. 266-282.

[28] H.-Y. Tseng, H.-Y. Lee, J.-B. Huang, and M.-H. Yang, "Cross-Domain Few-Shot Classification via Learned Feature-Wise Transformation," in International Conference on Learning Representations (ICLR), 2020, pp. 1-14.

[29] V. Garcia and J. Bruna, "Few-Shot Learning with Graph Neural Networks," in International Conference on Learning Representations (ICLR), 2018, pp. 1-13.

[30] T. Munkhdalai and H. Yu, "Meta Networks," in International Conference on Machine Learning (ICML), 2017, pp. 2554-2563.

[31] J. Humplik, A. Galashov, L. Hasenclever, P. A. Ortega, Y. W. Teh, and N. Heess, "Meta Reinforcement Learning as Task Inference," arXiv preprint arXiv:1905.06424, 2019.

[32] W. Yin, M. Yu, B. Zheng, T. Liu, X. Zhang, and Z. Yu, "Few-Shot Learning for Relation Extraction," in Annual Meeting of the Association for Computational Linguistics (ACL), 2019, pp. 4623-4632.

[33] Y. Zhao, J. Li, and Y. Gong, "Low-Resource Speaker Adaptation Using a Meta-Learning Approach," in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019, pp. 6020-6024.

[34] J. Snell, K. Swersky, and R. Zemel, "Prototypical Networks for Few-shot Learning," in Advances in Neural Information Processing Systems, 2017, pp. 4077-4087.

[35] T. Guo, Z. Yang, J. Yang, Z. Sun, and J. Yu, "Correlation-based Attention for Few-shot Learning," in IEEE International Conference on Computer Vision (ICCV), 2019, pp. 8290-8299.

[36] T. Ramalho and M. Garnelo, "Adaptive Posterior Learning: few-shot learning with a surprise-based memory module," in International Conference on Learning Representations (ICLR), 2019, pp. 1-16.

[37] K. Javed and M. White, "Meta-Learning Representations for Continual Learning," in Advances in Neural Information Processing Systems, 2019, pp. 1820-1830.

[38] C. Finn, P. Abbeel, and S. Levine, "Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks," in International Conference on Machine Learning (ICML), 2017, pp. 1126-1135.