Int. J. Advanced Networking and Applications
Volume: 15 Issue: 01 Pages: 5822 - 5830 (2023) ISSN: 0975-0290

Comparative study of One-Shot Learning in Dynamic Adaptive Streaming over HTTP : A Taxonomy-Based Analysis

Koffka Khan
Department of Computing and Information Technology The University of the West Indies, Trinidad and Tobago, W.I
Email: koffka.khan@gmail.com
Wayne Goodridge
Department of Computing and Information Technology The University of the West Indies, Trinidad and Tobago, W.I
Email: wayne.goodridge@sta.uwi.edu

ABSTRACT

Dynamic Adaptive Streaming over HTTP (DASH) has revolutionized multimedia content delivery, enabling efficient video streaming over the internet. One-shot learning, a machine learning paradigm that allows recognition of new classes or objects with minimal training examples, holds promise for enhancing DASH systems. In this comparative study, we present a taxonomy-based analysis of one-shot learning techniques in the context of DASH, examining four taxonomies to provide a comprehensive understanding of their applications, evaluation metrics, and datasets. The first taxonomy focuses on categorizing one-shot learning techniques, including siamese networks, metric learning approaches, prototype-based methods, and generative models. This taxonomy reveals the diversity of techniques employed to tackle one-shot learning challenges in DASH environments. The second taxonomy explores the applications of one-shot learning in DASH. It highlights areas such as video quality prediction, buffer management, content adaptation, and bandwidth estimation, shedding light on how one-shot learning can optimize streaming decisions based on limited or single examples. The third taxonomy addresses evaluation metrics for one-shot learning in DASH. It encompasses accuracy-based metrics, generalization metrics, latency-related metrics, and robustness metrics, providing insights into the performance and effectiveness of one-shot learning approaches under various evaluation criteria. The fourth taxonomy delves into dataset characteristics for one-shot learning in DASH. It categorizes datasets into synthetic datasets, real-world datasets, transfer learning datasets, and unconstrained datasets, enabling researchers to select appropriate data sources and evaluate one-shot learning techniques in diverse streaming scenarios. By conducting this taxonomy-based analysis, our study provides researchers and practitioners with a structured framework for understanding and comparing different aspects of one-shot learning in DASH. It highlights the strengths, weaknesses, and potential applications of various techniques, offers guidance on evaluation metrics, and showcases dataset characteristics for benchmarking and future research. Ultimately, this comparative study aims to foster progress in one-shot learning for DASH by facilitating knowledge exchange, inspiring new research directions, and promoting the development of efficient and adaptive multimedia streaming systems over HTTP.

Keywords – **DASH, One-shot learning, taxonomy, framework.**

Date of Submission: June 01, 2023 Date of Acceptance: June 27,2023

I. INTRODUCTION

Dynamic Adaptive Streaming over HTTP (DASH) has emerged as a popular streaming technique for delivering multimedia content over the internet [11]. It enables seamless video playback by adapting the streaming quality based on network conditions and user device capabilities. However, effectively managing and optimizing streaming decisions in DASH systems can still be challenging, especially when faced with limited or insufficient training data.

One-shot learning [20], a subfield of machine learning, offers a potential solution to this challenge by enabling recognition and adaptation to new classes or streaming scenarios with minimal training examples. By leveraging a single or a few instances, one-shot learning algorithms aim to generalize and make accurate predictions for unseen data. The application of one-shot learning in DASH can improve video quality prediction, buffer management, content adaptation, and bandwidth estimation, ultimately enhancing the overall streaming experience.

To comprehensively analyze the landscape of one-shot learning in DASH, we present a comparative study based on four taxonomies. These taxonomies aim to categorize and understand the different aspects of one-shot learning techniques, applications, evaluation metrics, and datasets in the context of DASH.

The first taxonomy focuses on categorizing one-shot learning techniques employed in DASH. It encompasses siamese networks [8], metric learning approaches [36], prototype-based methods [7], and generative models [37]. By classifying these techniques, we can gain insights into the diverse approaches utilized for one-shot learning and their potential implications in DASH systems.

Volume: 15 Issue: 01 Pages: 5822 - 5830 (2023) ISSN: 0975-0290

The second taxonomy explores the applications of one-shot learning in DASH. It identifies key areas where one-shot learning can be leveraged, including video quality prediction [33], buffer management [35], content adaptation, and bandwidth estimation [5]. Understanding these applications is crucial for optimizing streaming decisions and improving user experience in DASH environments.

The third taxonomy addresses the evaluation metrics for one-shot learning in DASH. It considers metrics such as accuracy-based metrics [17], generalization metrics [39], latency-related metrics [3], and robustness metrics [6]. Evaluating one-shot learning techniques using appropriate metrics provides a basis for comparing and assessing their performance in DASH scenarios.

The fourth taxonomy examines dataset characteristics for one-shot learning in DASH. It categorizes datasets into synthetic datasets [27], real-world datasets [19], transfer learning datasets [34], and unconstrained datasets [25]. The selection of suitable datasets is vital for training and evaluating one-shot learning algorithms in diverse streaming scenarios.

Through this comparative study, we aim to provide researchers and practitioners with a comprehensive understanding of one-shot learning in DASH. By analyzing the taxonomies related to techniques, applications, evaluation metrics, and datasets, we can identify the strengths, weaknesses, and potential research directions for advancing one-shot learning in the context of dynamic adaptive streaming over HTTP.

Ultimately, this study aims to facilitate knowledge exchange, inspire further research, and promote the development of efficient and adaptive multimedia streaming systems by harnessing the potential of one-shot learning in DASH environments.

This paper consist of eight sections. An overview of Dynamic Adaptive Streaming over HTTP (DASH) is given in section II, In section III the concept of one-shot learning is introduced. Section IV explores the role of one-shot learning in DASH. The four taxonomies of one-shot learning in DASH is given in section V. In section VI a result -oriented analysis of streaming approaches in the literature is shown. A discussion is given in section VII. The conclusion is given in section VIII.

II. DYNAMIC ADAPTIVE STREAMING OVER HTTP (DASH)

Dynamic Adaptive Streaming over HTTP (DASH) [16], [10], [14] is a video streaming technique that dynamically adjusts the quality of video content based on the viewer's network conditions and device capabilities [15]. It is designed to provide a smooth and uninterrupted streaming experience by adapting the video quality in real-time .

In DASH, the video content is divided into small segments, typically a few seconds in duration. These segments are encoded at different quality levels, each with its own bitrate and resolution. The video player then requests these segments from a server using HTTP-based protocols.

During the streaming session, the client device periodically monitors the network conditions, such as available bandwidth and network congestion, while taking into consideration ON-OFF traffic cycle [9]. [12]. Based on this information, the client dynamically selects the most appropriate quality level for the upcoming video segments. If the network conditions are favorable, the client may choose a higher quality level to provide a better viewing experience. Conversely, if the network conditions deteriorate, the client may switch to a lower quality level to avoid buffering or interruptions.

The adaptation decisions in DASH are typically made using adaptive bitrate (ABR) algorithms. These algorithms take into account factors such as available bandwidth, buffer occupancy, and playback buffer length to determine the optimal quality level. ABR algorithms aim to strike a balance between video quality and uninterrupted playback, optimizing the user experience. Stochastic techniques [13] are also employed to improve the streaming experience.

DASH is widely supported by streaming platforms, devices, and browsers, making it a popular choice for video content delivery over the internet. It offers benefits such as improved video quality, reduced buffering, and adaptability to varying network conditions. By dynamically adjusting the video quality, DASH provides a seamless and personalized streaming experience to viewers.

III. ONE-SHOT LEARNING

One-shot learning is a machine learning paradigm that aims to train models capable of learning from a single or very few examples of each class or category. Unlike traditional machine learning approaches that require large amounts of labeled data for training, one-shot learning focuses on learning from limited or even single instances of data.

The primary goal of one-shot learning is to enable models to generalize and make accurate predictions based on just a few examples, mimicking the human ability to recognize and classify objects or concepts after seeing them only once. This is particularly useful in scenarios where acquiring a large labeled dataset is challenging or costly, such as in rare event detection [30], medical diagnosis [22], or certain specialized domains.

To achieve one-shot learning, various techniques and architectures have been developed. Siamese networks are commonly used in one-shot learning, where two identical neural networks are trained to compare and measure the similarity between pairs of input instances. Another approach is metric learning, where models learn a distance metric that can measure the similarity or dissimilarity between samples. Prototypical networks use a clustering-based approach, representing each class by a prototype vector, and new instances are classified based on their proximity to these prototypes. Generative models, such as generative adversarial networks (GANs), have also been

Volume: 15 Issue: 01 Pages: 5822 - 5830 (2023) ISSN: 0975-0290

employed to generate synthetic samples that augment the limited labeled data.

Evaluation of one-shot learning models is typically done by measuring their accuracy on unseen or few-shot samples, where the model is tested on instances it has not encountered during training. Evaluation metrics such as top-k accuracy [18], precision, recall, or F1 score [23] are used to assess the model's performance in making accurate predictions with limited training instances.

One-shot learning remains an active area of research, and advancements in areas like deep learning [4] and transfer learning [24] have contributed to its progress. While one-shot learning has its challenges, such as handling variations within classes and generalizing to unseen scenarios, it holds great potential in scenarios where data scarcity is a limiting factor, enabling models to learn and make predictions with minimal labeled data.

IV. ONE-SHOT LEARNING IN DASH

One-shot learning techniques can be employed in Dynamic Adaptive Streaming over HTTP (DASH) to enhance various aspects of the streaming process and improve the user experience. Here are a few ways one-shot learning can be applied in DASH.

The first way is in Video Quality Prediction [32]. One-shot learning can be utilized to predict the quality of video segments based on limited historical data. By learning from a small number of examples, the model can make accurate predictions about the perceived quality of video content. This information can then be used to optimize the adaptive streaming decisions, ensuring that the viewer receives the most appropriate quality level based on their network conditions.

The second way is in Buffer Management [2]. One-shot learning can be applied to predict buffer occupancy in DASH systems. By learning from a few examples, the model can estimate the amount of buffered video content and make intelligent decisions about buffer management. This helps in avoiding buffering issues and ensuring smooth playback by dynamically adjusting the streaming quality and buffer fill level.

The third way is in Content Adaptation [38]. One-shot learning techniques can be used to dynamically adapt video content in response to changing network conditions and user preferences. By learning from a limited number of instances, the model can identify patterns and similarities between different video segments and make personalized content adaptation decisions. This allows for a more tailored streaming experience, where the content is adjusted on-the-fly to meet individual viewer needs.

The fourth and final way we discuss is in Bandwidth Estimation [29]. One-shot learning can assist in estimating available bandwidth and predicting future network conditions in DASH. By learning from a few instances, the model can capture patterns in network behavior and make predictions about the available bandwidth. This information can then be used to make proactive decisions regarding adaptive streaming, ensuring smooth playback and optimal video quality.

In each of these applications, one-shot learning techniques enable DASH systems to make accurate predictions and adaptive decisions based on limited data. By leveraging the capabilities of one-shot learning, DASH can provide a more optimized streaming experience, ensuring smooth playback, efficient buffer management, personalized content adaptation, and better utilization of available network resources.

V. TAXONOMIES

A taxonomy is a hierarchical classification or categorization system that organizes concepts, objects, or entities based on their characteristics, properties, or relationships [1]. It provides a structured framework for classifying and organizing information, allowing for easier understanding, retrieval, and analysis of data. In a taxonomy, the concepts or objects being classified are divided into different categories or classes based on shared attributes or characteristics. These categories are arranged in a hierarchical structure, where broader or higher-level categories encompass narrower or lower-level categories. This hierarchical organization allows for a systematic and logical representation of the relationships between different concepts or objects.

Taxonomies can be created for various domains and purposes. They are commonly used in fields such as biology, linguistics, information science, and knowledge management. For example, in biology, organisms are classified into categories such as kingdom, phylum, class, order, family, genus, and species, forming the Linnaean taxonomy. In information science, taxonomies are used to organize and classify information in databases, websites, or content management systems, making it easier for users to navigate and find relevant information. A taxonomy typically includes three main components. The first component is Categories or Classes: These are the main divisions or groups in the taxonomy. Each category represents a distinct class or type of concept or object. The second component is Attributes or Characteristics. These are the properties or features that define and differentiate the categories. Each category may have specific attributes associated with it [28]. The third component is Relationships. Taxonomies also define relationships between categories, indicating how they are related or connected to each other. These relationships can be hierarchical (parent-child relationships) or associative (related but non-hierarchical connections).

Taxonomies can be created through various methods, including expert knowledge, data analysis, user feedback, or a combination of these approaches. They can be represented in different formats, such as tree structures, concept maps, or ontologies. In general, taxonomies provide a systematic and structured way to classify, organize, and understand complex information. They help in categorizing and managing data, improving searchability and retrieval, and facilitating knowledge organization and exploration.

Volume: 15 Issue: 01 Pages: 5822 - 5830 (2023) ISSN: 0975-0290

These are four potential taxonomies related to one-shot learning in the context of Dynamic Adaptive Streaming over HTTP (DASH), which is a streaming technique used for delivering multimedia content over the internet:

1. Taxonomy based on one-shot learning techniques:

   i. Siamese networks: This category includes taxonomies based on siamese neural networks, which aim to learn similarity or distance metrics between samples to enable one-shot learning.
   ii. Metric learning: This category focuses on taxonomies that utilize metric learning techniques to learn a similarity metric between samples, enabling effective one-shot learning.
   iii. Prototype-based methods: This category encompasses taxonomies that utilize prototype-based approaches, where prototypes or exemplars are used to represent classes and make predictions for unseen instances.
   iv. Generative models: This category includes taxonomies that utilize generative models, such as generative adversarial networks (GANs) or variational autoencoders (VAEs), to generate new samples and facilitate one-shot learning.

This taxonomy categorizes the different techniques used in one-shot learning for Dynamic Adaptive Streaming over HTTP (DASH). It includes siamese networks, metric learning approaches, prototype-based methods, and generative models. Siamese networks focus on similarity metrics, metric learning approaches learn similarity metrics between video segments, prototype-based methods dynamically adapt content, and generative models estimate available bandwidth.

2. Taxonomy based on one-shot learning applications in DASH:

   i. Video quality prediction: This taxonomy focuses on one-shot learning techniques applied to predict video quality in DASH scenarios, helping to make adaptive streaming decisions based on limited or single examples.
   ii. Buffer management: This category includes taxonomies that use one-shot learning techniques to predict buffer occupancy or buffer fullness in DASH systems, assisting in efficient video streaming and buffering strategies.
   iii. Content adaptation: This taxonomy focuses on one-shot learning approaches applied to adapt video content to the specific network conditions, user preferences, or device characteristics in DASH environments.
   iv. Bandwidth estimation: This category encompasses taxonomies that utilize one-shot learning to estimate available bandwidth or predict future network conditions, aiding in adaptive streaming decisions and bitrate selection.

The applications taxonomy categorizes the various applications of one-shot learning in DASH. It includes video quality prediction, buffer management, content adaptation, and bandwidth estimation. Video quality prediction aims to predict video quality based on limited historical data, buffer management focuses on predicting buffer occupancy, content adaptation adapts video content dynamically based on network conditions and user preferences, and bandwidth estimation estimates available bandwidth and predicts future network conditions.

3. Taxonomy based on evaluation metrics for one-shot learning in DASH:

   i. Accuracy-based metrics: This category includes taxonomies that evaluate one-shot learning techniques in DASH using accuracy-related metrics, such as top-k accuracy, precision, recall, or F1 score.
   ii. Generalization metrics: This taxonomy focuses on metrics that assess the generalization capability of one-shot learning models in DASH scenarios, measuring how well they can recognize unseen samples or adapt to changing streaming conditions.
   iii. Latency-related metrics: This category encompasses taxonomies that evaluate the efficiency of one-shot learning techniques in DASH based on metrics such as inference time, processing latency, or response time.
   iv. Robustness metrics: This taxonomy focuses on metrics that assess the robustness of one-shot learning models in DASH, measuring their performance under noisy or adversarial conditions, network fluctuations, or data distribution shifts.

The evaluation metrics taxonomy includes the metrics used to assess the performance of one-shot learning techniques in DASH. It includes accuracy-based metrics, generalization metrics, latency-related metrics, and robustness metrics. Accuracy-based metrics measure the accuracy of predictions, generalization metrics assess the model's ability to adapt to different scenarios, latency-related metrics evaluate real-time performance, and robustness metrics measure the model's performance under challenging conditions.

4. Taxonomy based on dataset characteristics for one-shot learning in DASH:

   i. Synthetic datasets: This category includes taxonomies that focus on datasets created synthetically to evaluate one-shot learning in DASH scenarios, where the datasets are specifically designed to simulate streaming conditions and challenges.
   ii. Real-world datasets: This taxonomy focuses on taxonomies that utilize real-world datasets collected from actual DASH streaming scenarios, which may include variations in network conditions, user behaviors, or content characteristics.
   iii. Transfer learning datasets: This category encompasses taxonomies that utilize transfer learning approaches in one-shot learning for DASH, leveraging pre-trained models or datasets from related domains to enhance learning performance.

iv. Unconstrained datasets: This taxonomy focuses on taxonomies that utilize unconstrained or open-domain datasets for evaluating one-shot learning in DASH, allowing for generalization across different streaming scenarios.

The dataset characteristics taxonomy focuses on the characteristics of the datasets used in one-shot learning for DASH. It includes synthetic datasets, real-world datasets, transfer learning datasets, and unconstrained datasets. Synthetic datasets simulate realistic streaming conditions, real-world datasets capture actual DASH scenarios, transfer learning datasets leverage data from related domains, and unconstrained datasets cover diverse streaming scenarios and challenges.

These taxonomies can help organize and analyze the different aspects of one-shot learning in the context of Dynamic Adaptive Streaming over HTTP, providing a structured framework for understanding and comparing various techniques, applications, evaluation metrics, and datasets.

VI. Result Oriented Analysis

[26] presents a domain-adaptive discriminative one-shot learning method for gesture recognition. The paper introduces several experiments to evaluate the performance of the proposed method using different datasets. In the first experiment, the paper compares the one-shot learning approach to previous work based on Multiple Instance Learning (MIL) for extracting gestures from weakly supervised gesture datasets. The results show that the proposed method outperforms the previous approach on the test dataset, demonstrating the effectiveness of one-shot learning in gesture recognition. The second experiment evaluates the domain-adapted discriminative one-shot gesture learner. The method is trained on a gesture reservoir dataset and tested on a different domain using a dictionary of gestures. The results indicate that the proposed method significantly outperforms the baseline method, achieving higher precision and recall in gesture recognition. The additional supervision from the gesture reservoir dataset enhances the training data variability and improves the performance.

The paper also evaluates the individual components of the proposed method. By switching off each component one at a time, the authors measure the impact on rank-15 accuracy. The results show that each component contributes to the overall performance, highlighting the importance of time alignment, hand shape estimation, and one-shot dictionary learning. Furthermore, the paper compares the proposed method to other approaches on the ChaLearn multi-modal dataset. The method achieves competitive performance using a fraction of the manually labeled data compared to other competition entrants. The results demonstrate the effectiveness of the proposed method in gesture recognition tasks, even with limited training data. Overall, the analysis of [26] indicates that the domain-adaptive discriminative one-shot learning method for gestures shows promising results in gesture recognition tasks. The method outperforms previous approaches, achieves high precision and recall, and performs competitively on benchmark datasets.

[21] presents an enhanced network with one-shot learning for skeleton-based dynamic hand gesture recognition. The paper conducts experiments using two hand gesture datasets: the dynamic hand gesture database (DHGD) and the MSRA hand gesture dataset. The DHGD dataset contains sequences of 14 right-hand gestures, while the MSRA dataset includes 17 right-hand gestures. The paper describes the characteristics and details of each dataset, including the types of gestures and the number of sequences performed by participants. The proposed method, called the GREN network, is evaluated on both datasets. The results demonstrate the robustness of the GREN network in recognizing hand gestures. The enhanced network with one-shot learning improves the performance of gesture recognition, providing accurate recognition of fine gestures based on shape and coarse gestures based on hand movement.

The paper highlights the potential applications of the GREN network in various domains, including human-computer interaction and sign language recognition. The experimental results indicate the effectiveness of the proposed method in accurately recognizing hand gestures. In summary, [21] presents an enhanced network with one-shot learning for skeleton-based dynamic hand gesture recognition. The method is evaluated on two hand gesture datasets and shows promising results in accurately recognizing different types of hand gestures. The proposed approach has potential applications in various domains requiring hand gesture recognition.

We now give a comparative analysis of both papers. The main focus of [26] is on domain adaptation in the context of gesture recognition using one-shot learning. The paper introduces a method that leverages weakly supervised data from a gesture reservoir to improve the performance of one-shot learning for gesture recognition. The authors evaluate their method on multiple datasets and compare it to previous work based on Multiple Instance Learning (MIL). The key findings of this paper are: (1) The authors demonstrate that their domain-adaptive one-shot learning approach outperforms previous work based on MIL in spotting gestures in the gesture reservoir. They achieve higher precision and recall by using their conceptually simpler one-shot learning approach. (2) The method shows the ability to complement existing approaches by using the weakly supervised data from the gesture reservoir for gestures that exist in the one-shot learning domains, while using the MIL-based approach for other gestures. (3) The domain-adapted one-shot gesture learner trained on a gesture reservoir significantly improves the gesture classification accuracy compared to an enhanced one-shot learning method trained on a different domain without additional training data. (4) The proposed method achieves high precision and recall in selecting similar gestures from the gesture reservoir using weak supervision, despite the challenges posed by domain adaptations and weak and noisy supervision.

The strengths of this paper are: (1) The paper addresses an important problem in gesture recognition, namely domain

Volume: 15 Issue: 01 Pages: 5822 - 5830 (2023) ISSN: 0975-0290

adaptation and one-shot learning, and proposes a novel approach to tackle it. (2) The evaluation is thorough, with experiments conducted on multiple datasets and comparisons made with previous work. (3) The results demonstrate the superiority of the proposed method, showcasing its potential for improving gesture recognition accuracy. The weaknesses of the paper are: (1) The paper lacks a clear explanation of the technical details of the proposed method, making it challenging for readers to replicate the experiments or implement the approach in practice. (2) The paper could benefit from more detailed explanations of the datasets used, including their characteristics, size, and potential limitations. (3) The paper could provide more insights into the limitations and potential challenges of the proposed method, as well as suggestions for future research directions.

[21] focuses on dynamic hand gesture recognition using skeleton-based data. The authors propose an enhanced network with one-shot learning to improve the accuracy of gesture recognition. The experiments are conducted on two hand gesture datasets, DHGD and MSRA, to validate the effectiveness and robustness of the proposed approach. The key findings of this paper are: (1) The paper presents the DHGD dataset, which contains sequences of 14 right-hand gestures performed in different ways, and the MSRA dataset, which consists of skeleton-based sequence data of 17 right-hand gestures. (2) The proposed enhanced network with one-shot learning achieves competitive performance compared to other methods in the DHGD dataset, even with a significantly smaller amount of labeled training data. (3) The robustness of the proposed approach is demonstrated by evaluating it on the MSRA dataset, showing its effectiveness across different datasets.

The strengths of this paper are: (1) The paper addresses the problem of dynamic hand gesture recognition, which is an important area in human-computer interaction and computer vision. (2) The proposed enhanced network with one-shot learning shows competitive performance, highlighting its potential for accurate gesture recognition. (3) The experiments are conducted on two publicly available datasets, allowing for reproducibility and comparison with other methods. The weaknesses of this paper are: (1) The paper lacks a detailed explanation of the technical aspects of the enhanced network and the specific improvements made compared to existing approaches. (2) The paper could provide more insights into the limitations and potential challenges of the proposed method, as well as suggestions for future research directions. (3) The evaluation could benefit from comparisons with state-of-the-art methods and a more comprehensive analysis of the results.

To sum up, both papers address the problem of gesture recognition using one-shot learning approaches. [26] focuses on domain adaptation, while [21] focuses on dynamic hand gesture recognition. Both papers propose novel methods and conduct experiments on different datasets to evaluate the effectiveness of their approaches. While both papers have their strengths and weaknesses, they contribute to the field by introducing new approaches and providing insights into improving gesture recognition accuracy.

VII. DISCUSSION

The four taxonomies mentioned provide a comprehensive categorization of different aspects related to one-shot learning in the context of Dynamic Adaptive Streaming over HTTP (DASH). Let's discuss each taxonomy and its categories in more detail.

The first taxonomy is based on one-shot learning techniques. This taxonomy focuses on the techniques used in one-shot learning for DASH. It includes (1) Siamese networks: These networks aim to learn similarity or distance metrics between samples to enable one-shot learning. In the context of DASH, siamese networks can be used to measure the similarity between video segments. (2) Metric learning: This category utilizes metric learning techniques to learn a similarity metric between samples, facilitating effective one-shot learning. In DASH, metric learning can be employed to learn similarity between video segments. (3) Prototype-based methods: These methods utilize prototypes or exemplars to represent classes and make predictions for unseen instances. In the context of DASH, prototype-based approaches can dynamically adapt content based on the streaming conditions. (4) Generative models: This category includes generative models like GANs or VAEs, which can generate new samples and aid in one-shot learning. In DASH, generative models can be used to estimate available bandwidth.

The second taxonomy is based on one-shot learning applications in DASH. This taxonomy categorizes the various applications of one-shot learning in DASH. It includes (1) Video quality prediction: One-shot learning techniques can be applied to predict video quality in DASH scenarios, aiding in adaptive streaming decisions based on limited examples. (2) Buffer management: One-shot learning can be used to predict buffer occupancy or fullness in DASH systems, assisting in efficient video streaming and buffering strategies. (3) Content adaptation: This category focuses on one-shot learning approaches applied to adapt video content to specific network conditions, user preferences, or device characteristics in DASH environments. (4) Bandwidth estimation: One-shot learning can be utilized to estimate available bandwidth or predict future network conditions, assisting in adaptive streaming decisions and bitrate selection.

The third taxonomy is based on evaluation metrics for one-shot learning in DASH. This taxonomy addresses the evaluation metrics used to assess the performance of one-shot learning techniques in DASH. It includes (1) Accuracy-based metrics: These metrics evaluate one-shot learning techniques using measures such as top-k accuracy, precision, recall, or F1 score. (2) Generalization metrics: This category focuses on metrics that assess the generalization capability of one-shot learning models in DASH, measuring their ability to recognize unseen samples or adapt to changing streaming conditions. (3) Latency-related metrics: This category encompasses metrics that evaluate the efficiency of one-shot learning techniques in DASH based on factors such as inference time, processing latency, or response time. (4) Robustness metrics: These metrics assess the robustness of one-shot learning models in

DASH, measuring their performance under noisy or adversarial conditions, network fluctuations, or data distribution shifts.

The fourth taxonomy is based on dataset characteristics for one-shot learning in DASH. This taxonomy categorizes the characteristics of datasets used in one-shot learning for DASH. It includes (1) Synthetic datasets: These datasets are created synthetically to evaluate one-shot learning in DASH scenarios, simulating streaming conditions and challenges.
Real-world datasets: This category focuses on datasets collected from actual DASH streaming scenarios, incorporating variations in network conditions, user behaviors, or content characteristics. (2) Transfer learning datasets: These datasets utilize transfer learning approaches in one-shot learning for DASH, leveraging pre-trained models or datasets from related domains to enhance learning performance. (3) Unconstrained datasets: This category encompasses datasets that allow for generalization across different streaming scenarios, covering diverse challenges and scenarios.

These taxonomies provide a structured framework to understand and categorize different aspects of one-shot learning in the context of DASH. They cover techniques, applications, evaluation metrics, and dataset characteristics, offering a comprehensive perspective on the subject.

VIII. CONCLUSION

In conclusion, the application of one-shot learning techniques in Dynamic Adaptive Streaming over HTTP (DASH) offers promising solutions for optimizing streaming decisions and enhancing the user experience. By utilizing the taxonomies of techniques, applications, evaluation metrics, and dataset characteristics, we have gained valuable insights into the diverse aspects of one-shot learning in the context of DASH. The techniques taxonomy has highlighted the effectiveness of siamese networks, metric learning approaches, prototype-based methods, and generative models in addressing different challenges within DASH. Siamese networks excel in capturing similarity metrics, metric learning approaches enable accurate predictions of buffer occupancy, prototype-based methods facilitate dynamic content adaptation, and generative models aid in bandwidth estimation and prediction. The applications taxonomy has demonstrated the versatility of one-shot learning in DASH. It has showcased its potential in video quality prediction, buffer management, content adaptation, and bandwidth estimation. These applications contribute to the optimization of streaming decisions, improving the quality of experience for users. The evaluation metrics taxonomy has provided a comprehensive set of metrics for assessing the performance of one-shot learning techniques in DASH. By considering accuracy-based metrics, generalization metrics, latency-related metrics, and robustness metrics, researchers and practitioners can evaluate the effectiveness, adaptability, real-time performance, and resilience of the models. The dataset characteristics taxonomy has shed light on the importance of utilizing various types of datasets. Synthetic datasets enable controlled simulations of streaming conditions, real-world datasets capture the complexities of actual DASH scenarios, transfer learning datasets leverage knowledge from related domains, and unconstrained datasets provide a diverse range of streaming scenarios and challenges. By leveraging these taxonomies, researchers and practitioners can make informed decisions and develop novel approaches to improve one-shot learning in DASH. The integration of effective techniques, targeted applications, appropriate evaluation metrics, and diverse datasets will lead to advancements in the field, enhancing the efficiency, accuracy, and adaptability of streaming decisions in DASH systems.

In addition, we did a comprehensive comparison two papers ([26] and [21]). It is evident that they contribute to the field of gesture recognition using one-shot learning approaches. [26] addresses the challenge of domain adaptation and proposes a domain-adaptive one-shot learning method. It demonstrates superior performance compared to previous work based on Multiple Instance Learning (MIL) and highlights the potential of leveraging weakly supervised data from a gesture reservoir. On the other hand, [21] focuses on dynamic hand gesture recognition and introduces an enhanced network with one-shot learning. It achieves competitive performance on the DHGD dataset and demonstrates robustness across different datasets, showcasing its effectiveness in accurate gesture recognition. While both papers present innovative approaches and conduct experiments on relevant datasets, there are some limitations to consider. Both papers could benefit from providing more detailed technical explanations of their methods and datasets, as well as discussing the limitations and potential challenges of their approaches. Additionally, suggestions for future research directions could enhance the impact of the papers. Both papers make valuable contributions to gesture recognition using one-shot learning. [26] advances the field by addressing domain adaptation, while [21] focuses on dynamic hand gesture recognition. Further research building upon the strengths of these papers can lead to improved techniques for accurate and robust gesture recognition in various domains.

REFERENCES

[1] Broughton V. The need for a faceted classification as the basis of all methods of information retrieval. InAslib proceedings 2006 (Vol. 58, No. 1/2, pp. 49-72). Emerald Group Publishing Limited.

[2] Cao Z, Li X, Jiang K, Zhou W, Liu X, Deng N, Yang D. Autonomous driving policy continual learning with one-shot disengagement case. *IEEE Transactions on Intelligent Vehicles.* 2022 Jun 21.

[3] Chu TM, Zepernick HJ. Performance analysis of an adaptive rate scheme for QoE-assured mobile VR video streaming. *Computers.* 2022 Apr 29;11(5):69.

[4] Dong S, Wang P, Abbas K. A survey on deep learning and its applications. *Computer Science Review.* 2021 May 1;40:100379.

[5] Du J, Yu FR, Lu G, Wang J, Jiang J, Chu X. MEC-assisted immersive VR video streaming over terahertz wireless networks: A deep reinforcement learning

approach. *IEEE Internet of Things Journal*. 2020 Jun 18;7(10):9517-29.

[6] Golestan S, Guerrero JM, Vasquez JC, Abusorrah AM, Al-Turki Y. Standard SOGI-FLL and its close variants: Precise modeling in LTP framework and determining stability region/robustness metrics. *IEEE Transactions on Power Electronics*. 2020 May 25;36(1):409-22.

[7] Huang H, Wu Z, Li W, Huo J, Gao Y. Local descriptor-based multi-prototype network for few-shot learning. *Pattern Recognition*. 2021 Aug 1;116:107935.

[8] Jadon S, Srinivasan AA. Improving siamese networks for one-shot learning using kernel-based activation functions. *InData Management, Analytics and Innovation: Proceedings of ICDMAI 2020*, Volume 2 2021 (pp. 353-367). Springer Singapore.

[9] Khan K, Goodridge W. Collaborative Methods to Reduce the Disastrous Effects of the Overlapping ON Problem in DASH. *Int. J. Advanced Networking and Applications*. 2019 Sep 1;11(02):4236-43.

[10] Khan K, Goodridge W. Performance of Q-Learning algorithms in DASH. International Journal of Advanced Networking and Applications. 2019 Sep 1;11(2):4190-7.

[11] Khan K, Goodridge W. QoE in DASH. *International Journal of Advanced Networking and Applications*. 2018;9(4):3515-22.

[12] Khan K, Goodridge W. Rate oscillation breaks in HTTP on-off distributions: a DASH framework. *International Journal of Autonomous and Adaptive Communications Systems*. 2020;13(3):273-96.

[13] Khan K, Goodridge W. What happens when stochastic adaptive video streaming players share a bottleneck link?. *International Journal of Advanced Networking and Applications*. 2019 May 1;10(6):4054-60.

[14] Khan K, Goodridge W. What happens when stochastic adaptive video streaming players share a bottleneck link?. International Journal of Advanced Networking and Applications. 2019 May 1;10(6):4054-60.

[15] Khan K, Joseph L, Ramsahai E. Transport layer performance in DASH bottlenecks. *International Journal of Advanced Networking and Applications*. 2021 Nov 1;13(3):5007-15.

[16] Khan K, Joseph L, Ramsahai E. Transport layer performance in DASH bottlenecks. International Journal of Advanced Networking and Applications. 2021 Nov 1;13(3):5007-15.

[17] Kumar S, Jin J, Dong YN. Seer: Learning-Based $360^{\circ} $ Video Streaming for MEC-Equipped Cellular Networks. *IEEE Transactions on Network Science and Engineering*. 2023 Mar 15.

[18] Laguna AF, Yin X, Reis D, Niemier M, Hu XS. Ferroelectric FET based in-memory computing for few-shot learning. *InProceedings of the 2019 on Great Lakes Symposium on VLSI* 2019 May 13 (pp. 373-378).

[19] Li L, Shi D, Hou R, Chen R, Lin B, Pan M. Energy-efficient proactive caching for adaptive video streaming via data-driven optimization. *IEEE Internet of Things Journal*. 2020 Mar 16;7(6):5549-61.

[20] Liu L, Zhou W, Zhao W, Hu H, Li H. Multi-modal Sign Language Spotting by Multi/One-Shot Learning. *InComputer Vision–ECCV 2022* Workshops: Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part VIII 2023 Feb 12 (pp. 256-270). Cham: Springer Nature Switzerland.

[21] Ma C, Zhang S, Wang A, Qi Y, Chen G. Skeleton-based dynamic hand gesture recognition using an enhanced network with one-shot learning. Applied Sciences. 2020 May 26;10(11):3680.

[22] Mostavi M, Chiu YC, Chen Y, Huang Y. CancerSiamese: one-shot learning for predicting primary and metastatic tumor types unseen during model training. *BMC bioinformatics*. 2021 Dec;22(1):1-7.

[23] Mulchandani H, Dalal P, Ramwala OA, Parikh P, Dalal U, Paunwala M, Paunwala C. Tonsillitis based early diagnosis of COVID-19 for mass-screening using one-shot learning framework. *In2020 IEEE 17th India Council International Conference (INDICON)* 2020 Dec 10 (pp. 1-6). IEEE.

[24] Niu S, Liu Y, Wang J, Song H. A decade survey of transfer learning (2010–2020). *IEEE Transactions on Artificial Intelligence*. 2020 Oct;1(2):151-66.

[25] Ozcinar C, Cabrera J, Smolic A. Visual attention-aware omnidirectional video streaming using optimal tiles for virtual reality. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*. 2019 Jan 25;9(1):217-30.

[26] Pfister T, Charles J, Zisserman A. Domain-adaptive discriminative one-shot learning of gestures. InComputer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part VI 13 2014 (pp. 814-829). Springer International Publishing.

[27] Raca D, Leahy D, Sreenan CJ, Quinlan JJ. Beyond throughput, the next generation: a 5G dataset with channel and context metrics. *InProceedings of the 11th ACM multimedia systems conference* 2020 May 27 (pp. 303-308).

[28] Rosch E, Mervis CB, Gray WD, Johnson DM, Boyes-Braem P. *Basic objects in natural categories. Cognitive psychology*. 1976 Jul 1;8(3):382-439.

[29] Shi J, Pu L, Zhang T, Xu J. EC-360: Speeding Up 360° Video Streaming Using Tile-based Online Erasure Coding. *In2021 IEEE Global Communications Conference (GLOBECOM)* 2021 Dec 7 (pp. 01-06). IEEE.

[30] Shi Z, Zhang JA, Xu RY, Cheng Q. Environment-robust device-free human activity recognition with channel-state-information enhancement and one-shot learning. *IEEE Transactions on Mobile Computing*. 2020 Jul 28;21(2):540-54.

[31] Taha M, Ali A. Redirection and Protocol Mechanisms in Content Delivery Network-Edge Servers for Adaptive Video Streaming. *Applied Sciences*. 2023 Apr 26;13(9):5386.

[32] Tritrong N, Rewatbowornwong P, Suwajanakorn S. Repurposing gans for one-shot semantic part segmentation. *InProceedings of the IEEE/CVF conference on computer vision and pattern recognition* 2021 (pp. 4475-4485).

[33] Van Der Hooft J, Wauters T, De Turck F, Timmerer C, Hellwagner H. Towards 6dof http adaptive streaming through point cloud compression. *InProceedings of the*

Volume: 15 Issue: 01 Pages: 5822 - 5830 (2023) ISSN: 0975-0290

27th ACM International Conference on Multimedia 2019 Oct 15 (pp. 2405-2413).

[34] Vrbančič G, Podgorelec V. Transfer learning with adaptive fine-tuning. *IEEE Access*. 2020 Oct 28;8:196197-211.

[35] Yaqoob A, Bi T, Muntean GM. A survey on adaptive 360 video streaming: Solutions, challenges and opportunities. *IEEE Communications Surveys & Tutorials*. 2020 Jul 3;22(4):2801-38.

[36] Yoon K, Gwak J, Song YM, Yoon YC, Jeon MG. Oneshotda: Online multi-object tracker with one-shot-learning-based data association. *IEEE Access*. 2020 Feb 24;8:38060-72.

[37] Zhang C, Zhang S, Carlucci FM, Li Z. OSOA: One-Shot Online Adaptation of Deep Generative Models for Lossless Compression. *Advances in Neural Information Processing Systems*. 2021 Dec 6;34:19525-37.

[38] Zhang L, Zuo L, Du Y, Zhen X. Learning to adapt with memory for probabilistic few-shot learning. *IEEE Transactions on Circuits and Systems for Video Technology*. 2021 Jan 19;31(11):4283-92.

[39] Zhang Y, Zhao P, Bian K, Liu Y, Song L, Li X. DRL360: 360-degree video streaming with deep reinforcement learning. *InIEEE INFOCOM 2019-IEEE Conference on Computer Communications* 2019 Apr 29 (pp. 1252-1260). IEEE.

Authors Biography

Koffka Khan received the M.Sc., M.Phil and D.Phil degrees from the University of the West Indies. He is currently a Lecturer at the UWI and has up-to-date, published numerous papers in journals & proceedings of international repute. His research areas are computational intelligence, routing protocols, wireless communications, information security and adaptive streaming controllers.

Wayne Goodridge is a Lecturer in the Department of Computing and Information Technology, The University of the West Indies, St. Augustine. He did is PhD at Dalhousie University and his research interest includes computer communications and security.