High Efficiency Video Coding and Convolutional Neural Networks: A Survey

www.surveyx.cn

Abstract

High Efficiency Video Coding (HEVC), a cornerstone of modern video compression standards, effectively manages the increasing demand for high-quality video content transmission. Supporting resolutions up to 8K, HEVC surpasses predecessors like H.264 in compression efficiency, crucial for bandwidth-limited scenarios. Its integration with Convolutional Neural Networks (CNNs) further enhances compression capabilities, automating complex processes and improving video quality through techniques like content-aware in-loop filtering and adaptive block up-sampling. CNNs also advance optical flow prediction and transform network architectures, optimizing energy compaction and coding efficiency. Neural network optimization is vital for refining architectures to manage the high computational demands of CNNs in video compression. Strategies such as coreset-based compression and fixed-point CNNs reduce computational overhead, while quantization and performance-aware pruning optimize trade-offs between compression and quality. Despite challenges in handling HDR content and optimizing perceptual characteristics, HEVC's contributions to video analytics, such as traffic surveillance, remain indispensable. Future directions include enhancing feedback mechanisms, exploring hybrid approaches, and integrating emerging technologies like reconfigurable computing and machine learning. These advancements promise to refine video compression methodologies, ensuring HEVC's continued relevance and superiority in the evolving landscape of video technologies.

1 Introduction

1.1 Significance of HEVC in Video Compression

High Efficiency Video Coding (HEVC) is essential for modern video compression, effectively addressing the rising demand for high-quality video transmission. As video content increasingly dominates internet traffic, HEVC's enhanced bit rate efficiency is crucial, particularly as it supports resolutions up to 8K, making it suitable for Ultra High-Definition (UHD) videos in bandwidth-constrained environments [1, 2].

HEVC outperforms its predecessor, H.264, in compressing video data, thereby meeting the needs of contemporary applications [3]. Its design incorporates advanced techniques such as the discrete cosine transform (DCT), which is vital for achieving high compression efficiency [4]. This efficiency is critical not only for reducing storage and transmission costs but also for supporting emerging applications that involve complex data types.

Moreover, HEVC enhances secure video distribution by addressing the limitations of existing encryption methods that restrict access to viewable information [5]. The integration of deep learning techniques further strengthens HEVC's capabilities, marking a significant evolution in video compression standards [6].

Despite its advantages, HEVC encounters challenges in processing High Dynamic Range (HDR) content and optimizing perceptual characteristics for human visual systems. Nonetheless, its contribu-

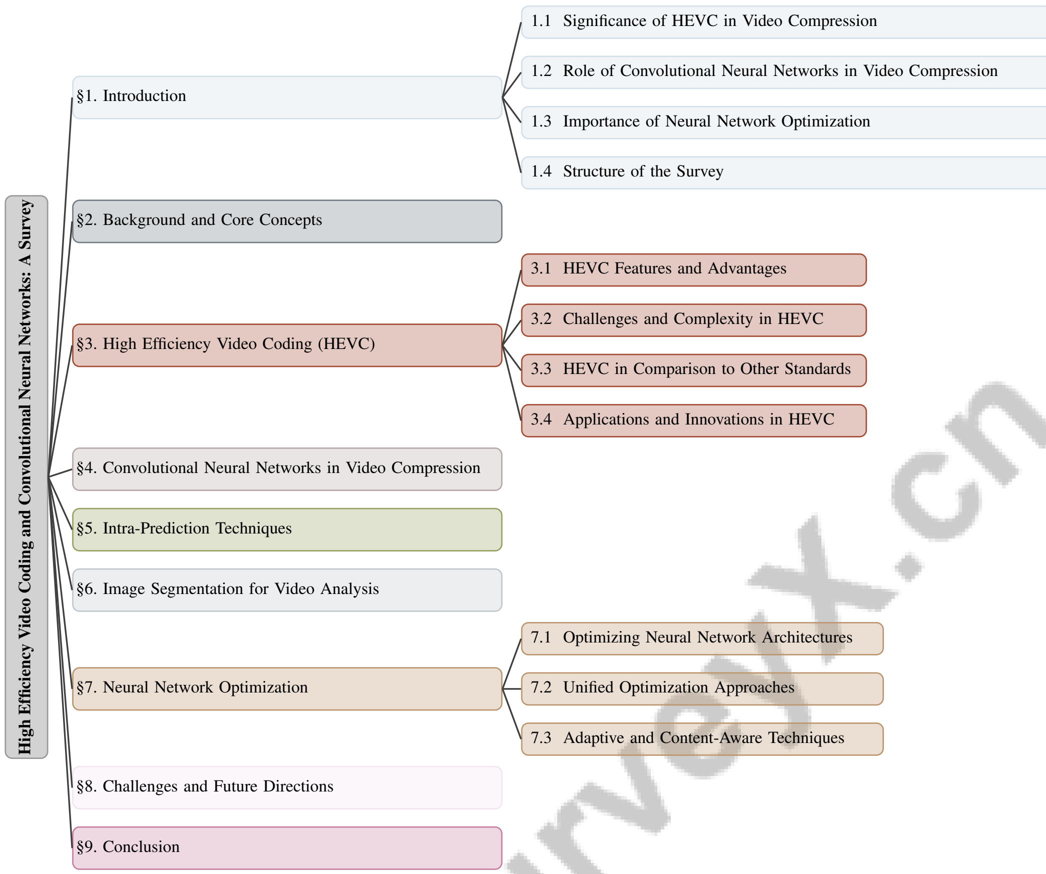


Figure 1: chapter structure

tions to video compression efficiency and its relevance in video analytics, such as traffic surveillance, underscore its importance as the demand for video content continues to grow [7].

1.2 Role of Convolutional Neural Networks in Video Compression

Convolutional Neural Networks (CNNs) significantly enhance HEVC by automating complex processes, thereby improving efficiency and image quality. Notably, content-aware CNN-based in-loop filtering employs multiple CNN models to adaptively filter regions of each coding tree unit (CTU) in HEVC, resulting in superior video quality [8]. This adaptability is further exemplified by CNN-based block up-sampling schemes, which allow for varying sampling rates across different CTUs based on local features, thereby enhancing HEVC's flexibility [2].

CNNs also improve optical flow prediction, enabling motion prediction from static images without prior scene assumptions, which enhances accuracy in video compression [9]. Architectures like MFRNet contribute to post-processing and in-loop filtering, showcasing the critical role of CNNs in video compression tasks [10]. Adjustments in CNN channel configurations are pivotal for refining compression efficiency, as indicated by the heterogeneity hypothesis [11].

Additionally, CNNs facilitate the development of new transform network architectures that directly process YUV 4:2:0 data, achieving better coding efficiency than traditional RGB-focused methods [12]. This advancement highlights CNNs' potential to enhance energy compaction and coding efficiency, crucial for contemporary video compression standards. Deep learning techniques, particularly CNN-based tools, have shown notable improvements in compression efficiency and image quality, especially in intra and inter prediction scenarios [6].

Through these innovative applications, CNNs not only enhance HEVC's technical capabilities but also broaden its applicability across various video compression tasks. By integrating CNNs into the HEVC framework, significant advancements are made in compression efficiency, image quality, and computational complexity. For instance, partition-masked CNNs and Variable-filter-size Residue-learning CNNs have demonstrated reductions in bit-rate by over 9.76

1.3 Importance of Neural Network Optimization

Optimization is crucial for enhancing the efficiency and performance of neural networks in video compression, particularly given the complexity and resource demands of Convolutional Neural Networks (CNNs). In resource-constrained environments, the high computational and memory demands of CNNs necessitate innovative optimization strategies to alleviate these challenges [13]. The need for optimization is further underscored by the management of compression artifacts in advanced super-resolution models for omnidirectional videos [14].

Inefficiencies in CNN architectures often arise from redundant convolutional layers that do not enhance predictive performance, making the streamlining of these architectures essential for improving video compression performance [15]. Effective optimization can yield substantial coding gains while minimizing computational burdens, as demonstrated by the LCCVC method, which highlights the critical role of optimization in achieving efficient video compression [16].

Furthermore, optimization is integral to methods that combine prediction information with model selection strategies to enhance video quality [17]. These strategies are vital for balancing compression efficiency with the computational complexity of encoding and decoding processes [18]. Comparative analyses of memory accesses in Versatile Video Coding (VVC) relative to HEVC provide valuable insights into memory requirements and optimization for energy-efficient applications, emphasizing the necessity for optimization in video compression standards [19].

The development of multiparametric classes of low-complexity 8-point DCT approximations, implementable via addition and bit-shifting operations, further illustrates optimization's potential to enhance computational efficiency [4]. Key challenges in applying deep learning to video coding include model complexity, the need for extensive datasets, and the balance between compression efficiency and encoding complexity [6]. By incorporating advanced optimization techniques, neural networks can achieve improved performance and efficiency in video compression tasks, underscoring their essential role in advancing video compression technologies.

1.4 Structure of the Survey

This survey is meticulously structured to provide a comprehensive exploration of High Efficiency Video Coding (HEVC) and its integration with Convolutional Neural Networks (CNNs) for enhanced video compression. The paper begins with an **Introduction** section, discussing the significance of HEVC in modern video compression, followed by an examination of CNNs' role in augmenting HEVC capabilities and the importance of neural network optimization. A detailed **Background and Core Concepts** section follows, providing an overview of essential concepts such as video compression standards, intra-prediction, image segmentation, and the interrelation between video compression and computer vision.

Subsequently, the survey delves into the specifics of **High Efficiency Video Coding (HEVC)**, discussing its features, advantages, challenges, and comparisons with other standards. The section on **Convolutional Neural Networks in Video Compression** explores the application of CNNs in tasks like intra-prediction and image segmentation, highlighting their impact on compression efficiency. Detailed analysis of **Intra-Prediction Techniques** and **Image Segmentation for Video Analysis** further elucidates the technical advancements in these areas.

The survey then addresses **Neural Network Optimization**, discussing strategies for optimizing architectures and parameters to enhance performance. The final sections, **Challenges and Future Directions**, and **Conclusion**, identify current challenges, potential research areas, and summarize the key insights of the survey, emphasizing the importance of HEVC and CNNs in advancing video compression technologies.The following sections are organized as shown in Figure 1.

AI-generated, for reference only.

2 Background and Core Concepts

2.1 Overview of Video Compression Standards

Video compression standards are pivotal in meeting the increasing demand for high-quality video across various platforms. High Efficiency Video Coding (HEVC), or H.265, is distinguished by its ability to support 4K and 8K resolutions with approximately 50

Emerging standards such as Versatile Video Coding (VVC) and AOMedia Video 1 (AV1) offer new features for specialized applications. VVC, as HEVC's successor, includes innovative coding tools that enhance compression efficiency and support new video formats [20]. AV1, developed by the Alliance for Open Media, aims to deliver high-quality video with improved compression, especially for web applications [21]. The challenges of streaming omnidirectional video demand efficient compression to manage bandwidth and latency [14]. Additionally, effective signal processing on non-Euclidean domains, such as the 2-sphere, indicates the evolving nature of video compression technologies [22].

2.2 Interrelation Between Video Compression and Computer Vision

The convergence of video compression and computer vision is rapidly advancing, significantly driven by Convolutional Neural Networks (CNNs). These networks enhance both domains by providing models that improve video processing efficiency and quality. In video compression, CNNs enable advanced techniques like intra-prediction and image segmentation, optimizing compression and video quality [23]. Concurrently, CNNs serve as mechanistic models in computer vision, enhancing visual recognition accuracy [23].

Hybrid models that integrate CNNs with Recurrent Neural Networks (RNNs) exemplify this synergy, utilizing keyframes to classify video content by merging temporal and spatial features, thereby improving classification accuracy [24]. The relationship between video compression and object tracking is explored through benchmarks assessing tracking performance on uncompressed sequences, highlighting the impact of compression artifacts on accuracy and the necessity for robust models [25].

Shot boundary detection (SBD) also benefits from this integration. Traditional low-level feature methods have struggled with accurate SBD, but CNNs have enhanced detection accuracy and efficiency [26]. Deep learning models facilitate more effective SBD, supporting tasks such as video summarization and indexing.

The incorporation of CNNs in video compression and computer vision significantly enhances these fields' capabilities. For instance, CNNs combined with RNNs efficiently categorize video types, while studies on lossy compression methods like JPEG and H.264 show that retraining on pre-compressed data can recover up to 78.4

In recent years, High Efficiency Video Coding (HEVC) has emerged as a pivotal standard in the realm of video compression, showcasing significant advancements over its predecessors. To better understand the complexities and advantages of HEVC, it is essential to examine its hierarchical structure, which is succinctly illustrated in Figure **??**. This figure elucidates the core features of HEVC, such as its performance enhancements, while also addressing the technical challenges and resource demands associated with its implementation. Furthermore, it outlines recent innovations and techniques aimed at reducing computational complexity, thereby demonstrating HEVC's adaptability and superiority in the field of video compression. By integrating these insights, we can appreciate the multifaceted nature of HEVC and its implications for future developments in video technology.

Figure 2: This figure illustrates the hierarchical structure of High Efficiency Video Coding (HEVC), highlighting its features, challenges, and innovations. It categorizes HEVC's core features and performance enhancements, addresses technical challenges and resource demands, and outlines recent innovations and techniques for reducing complexity, demonstrating HEVC's adaptability and superiority in video compression.

3 High Efficiency Video Coding (HEVC)

3.1 HEVC Features and Advantages

High Efficiency Video Coding (HEVC), also known as H.265, represents a significant advancement in video compression, enabling high-quality video delivery at reduced bit rates. A core feature of HEVC is its flexible coding unit structure, which utilizes variable block sizes and sophisticated motion compensation techniques to maximize compression efficiency across various content types [27]. It employs advanced prediction methods, including hierarchical coding structures that optimize encoding order and reference frame selection, enhancing its compression capabilities [27].

The incorporation of deep learning has further improved HEVC's performance. Techniques like the Partition-Masked Convolutional Neural Network (PMCNN) use partition-derived masks to enhance HEVC-compressed videos, improving video quality [28]. Low-complexity CNNs exploit non-local redundancies, surpassing traditional codecs and boosting compression performance [16].

HEVC outperforms its predecessor, H.264/AVC, with an average quality improvement of 10.18

Innovative approaches like frequency-dependent perceptual quantization (FDPQ) achieve bitrate reductions of up to 41

HEVC's superiority is confirmed through datasets featuring both objective and subjective assessments, establishing its edge over codecs like AV1 and VVC [29]. Its flexible coding structure, neural network integration, and adaptability to high-resolution formats solidify its status as a leading video compression standard. Figure 3 illustrates the key features and advantages of HEVC, highlighting its core features, performance improvements, and innovative approaches that contribute to its status as a leading video compression standard.

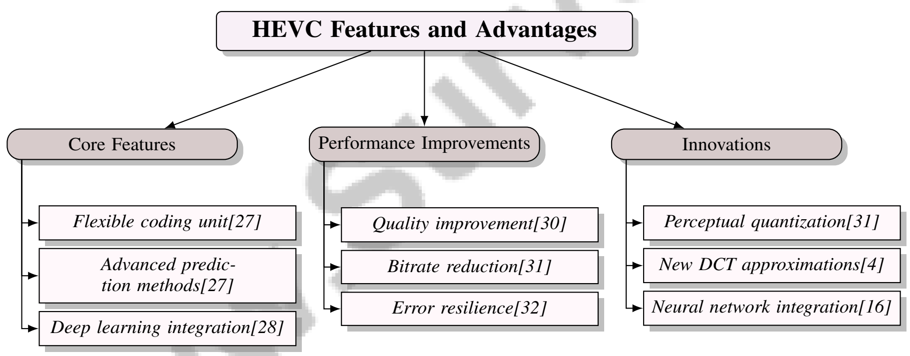


Figure 3: This figure illustrates the key features and advantages of High Efficiency Video Coding (HEVC), highlighting its core features, performance improvements, and innovative approaches that contribute to its status as a leading video compression standard.

3.2 Challenges and Complexity in HEVC

Implementing High Efficiency Video Coding (HEVC) involves challenges due to its complex coding techniques and significant computational demands. A primary challenge is the computational intensity of inter-frame coding, crucial for reducing temporal redundancy [20]. The exhaustive search required for rate-distortion optimization (RDO) in coding unit partitioning complicates real-time applications [33].

HEVC faces issues with compression artifacts, particularly at low bit rates, where intra coding can degrade image quality [34]. Inefficient signaling of intra prediction modes increases both bitrate and complexity due to the growing number of modes [35].

The complexity of HEVC's coding tools results in higher energy demands, challenging existing benchmarks to accurately represent these needs [1]. The intricate partitioning processes, while enhancing compression efficiency, also increase resource demands, posing challenges for devices with limited processing power [36].

Addressing these challenges requires innovative algorithms to optimize HEVC's capabilities, ensuring its effectiveness in modern video compression. HEVC maintains its position as a premier standard, accommodating the increasing demands of High Dynamic Range (HDR) content, with performance improvements of 10.18

3.3 HEVC in Comparison to Other Standards

High Efficiency Video Coding (HEVC) is often compared to other video compression standards to evaluate its performance. AV1, developed by the Alliance for Open Media, consistently outperforms both VP9 and HEVC, achieving lower bitrates while maintaining quality, making it effective for web applications and streaming services [37]. This is crucial given the demand for efficient video compression in digital media.

Despite advancements in neural network-based compression, traditional standards like JPEG and JPEG2000 remain less efficient than HEVC. While neural network-based methods outperform JPEG and JPEG2000, they struggle to surpass HEVC's efficiency, highlighting an area for further research [21]. This underscores HEVC's robust standing in video compression, particularly for high-resolution content.

The SC-SKV codec, designed for light field compression, demonstrates significant improvements over HEVC, achieving a 47.87

3.4 Applications and Innovations in HEVC

High Efficiency Video Coding (HEVC) continues to lead video compression innovations, enhancing its applicability across various domains. A significant innovation is the integration of Convolutional Neural Networks (CNNs) to improve the quality of Versatile Video Coding (VVC)-coded frames, highlighting the potential of CNNs in refining video encoding and elevating compressed video quality [17].

Recent developments focus on optimizing HEVC's decoding for real-time applications. The OpenVVC and VVdeC software decoders use parallel processing for real-time decoding on embedded platforms, enhancing speed and efficiency [38]. This is vital for high-definition video playback on resource-limited devices.

Innovations like the LCCVC method, which trains a lightweight CNN during encoding, yield substantial coding gains [16]. The MFRNet architecture achieves notable coding gains in in-loop filtering and post-processing, with enhancements of up to 16.0

Advanced rate control algorithms, such as those proposed by Tang et al., optimize compression efficiency while maintaining high video quality, approximating fixed QP configurations for random access and low-delay scenarios [39]. Balancing compression efficiency and video quality is critical.

HEVC's applications extend to reducing decoding complexity through methods like Saliency-Guided Complexity Control (SGCC), which focuses on salient regions to minimize perceptual quality loss [40]. This benefits applications with limited computational resources.

Innovations like ETH-CNN and ETH-LSTM significantly reduce encoding complexity for intra- and inter-modes, outperforming existing approaches while maintaining video quality [41]. These efforts highlight ongoing attempts to streamline HEVC's encoding processes, ensuring its relevance in video compression.

While the SC-SKV codec shows improvements over HEVC for light field compression, achieving a 47.87

4 Convolutional Neural Networks in Video Compression

4.1 Impact of CNNs on Video Compression Efficiency

Convolutional Neural Networks (CNNs) significantly enhance video compression efficiency, especially within High Efficiency Video Coding (HEVC). By utilizing deep learning, CNNs improve coding unit (CU) partitioning accuracy, reducing exhaustive searches and boosting compression efficiency [42, 41]. A CNN-based block up-sampling scheme achieves average BD-rate reductions of 5.5

Optimized network architectures like the Layer-Wise Differentiated Network Architecture (LW-DNA) refine channel configurations, enhancing performance while reducing complexity, crucial for efficient compression in resource-limited environments [11]. The Variable-filter-size Residue-learning CNN (VRCNN) mitigates artifacts in HEVC intra coding and enhances video quality through variable filter sizes and residue learning [34].

CNNs also streamline resource management by automating feature extraction in CTU depth decision algorithms, significantly boosting compression efficiency despite requiring larger datasets [33]. Feature-based models for Versatile Video Coding (VVC) exemplify innovations over HEVC models, reflecting continuous evolution in video compression [1].

By exploiting non-local redundancies, minimizing computational complexity, and enhancing video quality through artifact mitigation, CNNs contribute to bitrate savings of up to 10

4.2 Integration of Segmentation and Prediction

The integration of image segmentation and prediction techniques through CNNs enhances video compression efficiency within the HEVC framework. This approach optimizes coding decisions and improves compressed video quality, achieving bit-rate reductions up to 10

A two-stage CU partition decision process using CNNs exemplifies efficient feature extraction and adaptive decision-making in HEVC intra coding [8]. Segmentation techniques like EdgeSegNet optimize network architecture for image segmentation, crucial for maintaining quality during compression [43].

CNNs also enhance prediction techniques in compression. The Residual-Guided In-Loop Filter (RRNet) uses reconstructed frames and prediction residuals for improved frame reconstruction, effectively integrating segmentation and prediction [44]. Neural intra, inter, and residual coding within CNN frameworks enable efficient compression by creating compact intermediate representations [45]. ConvLSTM cells further enhance temporal feature propagation, ensuring consistent semantic segmentation across frames [46].

The LW-DNA approach adjusts channel configurations, integrating segmentation and prediction to boost compression efficiency [11]. This method infers multiple sub-pixel samples from a shared feature map, optimizing prediction [47].

By optimizing coding decisions and enhancing image quality, CNNs significantly advance video compression technologies, facilitating simultaneous extraction of semantic and visual information crucial for both human perception and machine vision applications. Innovative methods like keyframe extraction reduce processing time while maintaining performance, addressing challenges posed by increasing video data volumes [21, 24, 48, 49].

4.3 Feature Propagation Across Video Frames

Feature propagation across video frames is crucial for maintaining consistency in video compression, particularly enhancing HEVC efficiency. This process leverages spatial and temporal redundancies in video data to optimize quality and performance. The CAFBP method exemplifies effective utilization of these redundancies, enhancing performance in standards like H.265/SHVC [50].

A key advancement is the use of convolutional long short-term memory (ConvLSTM) cells for feature propagation through frames, ensuring semantic consistency by combining temporal propagation with a loss function that penalizes prediction inconsistencies [46]. Integrating ConvLSTM cells into CNN architectures enhances feature propagation over time, improving compression accuracy and quality.

These techniques highlight feature propagation's role in addressing temporal redundancy and maintaining visual consistency. By optimizing feature propagation, modern compression technologies significantly improve efficiency and quality, underscoring its essential role in advancing video compression standards. Recent developments, including AI-driven models and collaborative compression frameworks, reflect the evolving landscape of video compression aligned with diverse applications in smart cities and intelligent analytics [48, 51].

AI-generated, for reference only.

5 Intra-Prediction Techniques

5.1 Optimization Techniques for Intra-Prediction

Method Name	Optimization Techniques	Computational Efficiency	Neural Network Integration
ACIP[52]	Attention Mechanisms	Keyframe Extraction	Neural Network
TN[53]	Data Clustering-driven	Fast Termination Strategy	Tree-structured Neural
MFRNet[10]	Deep Learning Techniques	Computational Complexity	Convolutional Neural Network
EPC-IPM[35]	Genetic Algorithm	Keyframe Extraction	-
VRCNN[34]	Residue Learning Techniques	Faster Training Times	Convolutional Neural Networks
NCP[54]	Comprehensive Optimization Techniques	Global Search	Cnn Model Inference

Table 1: Comparison of Various Optimization Techniques for Intra-Prediction in Video Compression. This table outlines different methods, their optimization techniques, computational efficiency, and integration with neural networks, highlighting diverse approaches to enhance prediction accuracy and reduce computational complexity in HEVC and VVC standards.

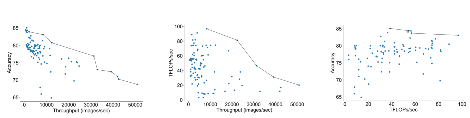
Enhancing video compression efficiency in HEVC and VVC necessitates optimizing intra-prediction techniques. Recent deep learning advancements have introduced methodologies employing neural networks, notably CNNs and RNNs, alongside attention mechanisms and hierarchical data structures. These innovations improve prediction accuracy in applications such as video content classification and biomedical image segmentation while utilizing techniques like keyframe extraction and weight quantization to reduce computational complexity [55, 24, 23, 56].

Attention-based Cross-component Intra-prediction (ACIP) exemplifies leveraging attention mechanisms to model spatial relationships, enhancing intra-prediction accuracy and overall video compression quality by dynamically adjusting to spatial characteristics [52]. Similarly, TreeNet utilizes a tree-structured, data clustering-driven neural network, capturing complex data relationships and optimizing prediction accuracy while reducing computational burdens [53].

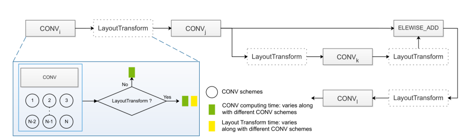
The MFRNet architecture integrates neural networks for enhanced feature reuse and information flow, minimizing artifacts and improving perceptual quality [10]. Genetic algorithms further enhance coding efficiency by selecting optimal intra prediction modes based on contextual information, minimizing rate costs [35].

The Variable-filter-size Residue-learning CNN (VRCNN) effectively reduces compression artifacts through variable filter sizes and residue learning, enhancing HEVC intra coding quality [34]. Moreover, deep learning techniques, including CNNs and LSTM networks, efficiently predict coding unit (CU) partitions, optimizing intra-prediction accuracy and reducing prediction process complexity [41].

These optimization techniques employ multiple neural networks for improved reference block mapping, attention mechanisms for chroma prediction, and hierarchical data structures that optimize CNNs by analyzing receptive fields and eliminating redundant layers, thereby improving computational efficiency and predictive performance [52, 15, 24, 57]. These strategies contribute to more efficient and effective video compression standards. Table 1 provides a comprehensive comparison of recent optimization techniques employed in intra-prediction for video compression, detailing the methods, optimization strategies, computational efficiencies, and neural network integrations.



(a) Performance Trade-offs in Deep Learning Models[58]



(b) LayoutTransform and Convolution Processing Flow[54]

Figure 4: Examples of Optimization Techniques for Intra-Prediction

As illustrated in Figure 4, optimization strategies are crucial for enhancing the efficiency and effectiveness of neural network models in intra-prediction. The first example, "Performance Trade-offs in Deep Learning Models," reveals the complex relationship between key performance metrics such as accuracy, throughput, and TFLOPs/s through scatter plots. This visualization highlights the trade-offs encountered when optimizing for different metrics, notably the negative correlation between accuracy and throughput. The second example, "LayoutTransform and Convolution Processing Flow," presents a flowchart detailing the sequence of convolution operations within a neural network, emphasizing the importance of layout transformations in optimizing convolution processing and improving computational efficiency. Together, these examples illustrate the multifaceted nature of optimization techniques in intra-prediction, balancing performance metrics with the strategic arrangement of computational processes [58, 54].

5.2 Handling Temporal and Spatial Variability

Method Name	Spatial Variability	Temporal Variability	Neural Network Integration
ACIP[52]	Spatial Relationships	-	Attention Module
DOFP[9]	Sparse Predictions	Dense Optical Flow	Cnn Architecture
TN[53]	Complex Patterns	Temporal Prediction Accuracy	Tree-structured Networks
SC-SKV[59]	Spatial Relations	-	-

Table 2: Overview of methods addressing spatial and temporal variability in video compression, highlighting their techniques and neural network integration. The table includes methods such as ACIP, DOFP, TN, and SC-SKV, each employing different strategies to enhance prediction accuracy and efficiency through spatial and temporal analysis.

Managing temporal and spatial variability during intra-prediction is crucial for improving video compression accuracy and efficiency, particularly in complex and dynamic scenes. The ACIP method addresses spatial variability by employing neural networks to extract features from reconstructed samples and predict chroma components, dynamically weighing reference samples based on spatial relevance [52].

Temporal variability poses additional challenges, as existing methods often yield sparse predictions constrained to specific domains. Techniques like dense optical flow prediction capture motion information across frames, improving temporal prediction accuracy and reducing artifacts [9].

The TreeNet approach provides a novel solution for addressing spatial variability by clustering training data hierarchically, enabling precise modeling of spatial correlations and enhancing intra-prediction accuracy [53]. Additionally, the SC-SKV codec employs disparity-guided sparse coding to enhance compression efficiency by addressing spatial and angular correlations in light field data [59].

Integrating sophisticated neural network architectures, such as attention-based convolutional networks, enhances chroma intra-prediction by effectively modeling spatial relationships. Implementing hierarchical data structures combined with RNNs facilitates efficient processing and categorization of video data, further improving predictive accuracy and compression performance in advanced coding frameworks like VVC [52, 24]. These methods significantly enhance video compression quality by effectively modeling and predicting complex temporal and spatial relationships. Table 2 provides a comprehensive comparison of various methods for handling spatial and temporal variability in video compression, illustrating their distinct approaches and integration of neural networks to improve prediction accuracy.

6 Image Segmentation for Video Analysis

6.1 Semantic Segmentation Techniques

Semantic segmentation is pivotal in video analysis, facilitating the division of images into meaningful segments for enhanced comprehension and processing. SegNet, a prominent deep convolutional encoder-decoder architecture, excels in achieving high accuracy and memory efficiency, making it suitable for resource-constrained environments [60]. EdgeSegNet further advances this field by achieving 89.7% segmentation accuracy with a model over 20 times smaller than traditional networks, proving advantageous for edge and mobile applications with limited computational resources [43].

Rebol et al. propose a method enhancing temporal consistency in segmentation, ensuring frame-to-frame coherence and surpassing traditional single-frame approaches, which is crucial for applications like video editing and augmented reality [46]. Segmentation-aware convolutional networks by Harley et al. improve spatial precision and minimize feature blurring, essential for pixel-wise prediction tasks in complex scenes [61].

These techniques significantly advance video analysis by improving segmentation accuracy, efficiency, and temporal consistency. SegNet's encoder-decoder structure efficiently maps low-resolution feature maps to high-resolution pixel-wise classifications, optimizing memory and computational speed for scene understanding [60]. EdgeSegNet, developed through a human-machine collaborative approach, achieves competitive accuracy while being compact, ideal for low-power edge computing [43]. Incorporating temporal information through methods like convolutional long short-term memory (ConvLSTM) cells enhances frame-to-frame consistency and overall segmentation accuracy, highlighting the critical role of these advancements in video analysis [60, 24, 46].

6.2 SegNet and Boundary Information Retention

SegNet, a deep convolutional encoder-decoder architecture, is distinguished by its capability to retain crucial boundary information during decoding, enhancing pixel-wise classification accuracy [60]. This retention is achieved through pooling indices stored during encoding and utilized in decoding to accurately upsample feature maps, ensuring boundary integrity and improving object edge delineation.

The retention of boundary information is vital in applications requiring high precision, such as autonomous driving, medical imaging, and video surveillance. SegNet's architecture efficiently realizes pixel-wise classification, ensuring precise upsampling of features while preserving boundary integrity, making it suitable for real-time applications demanding quick and accurate object recognition [60, 61, 23, 26]. By maintaining high spatial resolution and boundary integrity, SegNet fosters accurate segmentation outcomes critical for subsequent analysis and decision-making processes.

SegNet's sophisticated encoder-decoder design mirrors the 13 convolutional layers of the VGG16 network to generate low-resolution feature maps, while the decoder employs a novel non-linear upsampling technique using pooling indices to produce high-resolution pixel-wise classifications. This design enhances segmentation accuracy and performance across various image processing tasks while optimizing memory usage and computational efficiency, making SegNet a compelling choice for scene understanding and real-time processing environments. Comparative benchmarks confirm its competitive performance against established architectures, underscoring its practicality in both road scene and indoor segmentation tasks [60, 61, 62, 43].

7 Neural Network Optimization

A comprehensive understanding of neural network optimization strategies is vital for enhancing video compression performance. This section delves into methodologies that optimize neural network architectures, emphasizing techniques that improve efficiency while maintaining data integrity. The subsection *Optimizing Neural Network Architectures* highlights pivotal strategies that advance video compression technologies.

7.1 Optimizing Neural Network Architectures

Optimizing neural network architectures is fundamental to enhancing video compression, particularly within High Efficiency Video Coding (HEVC). Coreset-based neural network compression offers high compression rates without retraining, reducing computational overhead across various CNN architectures [13]. Fixed-point CNNs showcase optimization by minimizing weights and using fixed-point calculations, enhancing performance and reducing memory bandwidth requirements, ideal for resource-constrained settings [63]. Quantization plays a crucial role in architectural optimization, with methods that train networks to learn optimal quantization tables from image blocks in an unsupervised manner, balancing compression and quality [64].

Approximate discrete cosine transforms (DCTs), like the 16-point orthogonal DCT, eliminate multiplications, relying on additions to lower computational complexity [65]. NeoCPU exemplifies optimization for specific CPU architectures, enhancing performance through data management and operation optimization [54]. Frequency-dependent perceptual quantization (FDPQ) adjusts transform coefficient quantization based on perceptual relevance, ensuring visually lossless compression [31]. Performance-aware channel pruning balances accuracy and computational efficiency by selectively pruning channels based on performance metrics [66]. The MSE-CNN model predicts coding unit (CU) partitions at multiple stages, optimizing architectures for efficient video compression by reducing redundant computations [42].

These optimization strategies involve compression techniques, hardware alignment, perceptual coding, and efficient computation strategies, ensuring superior performance and efficiency in video compression tasks. Proposed models effectively estimate decoding energy demand for VVC, achieving a mean error below 4

7.2 Unified Optimization Approaches

Unified optimization approaches in neural networks enhance performance across multiple tasks, particularly in video compression. These approaches integrate diverse techniques, such as coreset-based compression, quantization, and receptive field analysis, forming a comprehensive framework that improves neural network operations' efficiency and effectiveness. This framework reduces memory footprint and inference time while enhancing model accuracy across applications, addressing computational challenges in complex neural network architectures [55, 24, 13, 56, 15].

Coreset-based neural network compression achieves high compression rates and reduces computational overhead without retraining [13]. Quantization optimizes the trade-off between compression and quality by training networks to learn optimal quantization tables from image blocks [64]. FDPQ optimizes neural networks for visually lossless compression by adjusting transform coefficient quantization based on perceptual relevance [31]. Performance-aware channel pruning enhances optimization by balancing accuracy and computational efficiency [66]. The MSE-CNN model illustrates unified optimization by predicting CU partitions at multiple stages, allowing early termination of redundant computations [42].

Unified optimization approaches integrate coreset-based CNN compression, quantization strategies, perceptual coding frameworks, and efficient computation strategies, addressing image and video data management challenges. These methodologies leverage deep learning to enhance coding performance and facilitate next-generation compression standards [21, 56, 13, 64], ensuring superior performance and efficiency across tasks.

7.3 Adaptive and Content-Aware Techniques

Adaptive and content-aware techniques enhance neural network performance in video compression by dynamically adjusting to content variability. These techniques leverage unique video content characteristics to improve encoding and decoding processes, enhancing compression efficiency and video quality. Machine learning algorithms optimize motion estimation, maintain learned information for better predictive accuracy, and implement spatial rate control, reducing file sizes and improving visual fidelity compared to traditional codecs. These methods also mitigate blocking artifacts and pixelation, ensuring an aesthetically pleasing viewing experience [49, 48, 51].

Content-aware CNNs adaptively filter different regions of each coding tree unit (CTU) in HEVC, employing multiple CNN models for in-loop filtering, dynamically adjusting based on local features, improving video quality [8]. Content-adaptive block up-sampling schemes utilize CNNs for varying sampling rates across CTUs, enhancing HEVC's adaptability, leading to better compression efficiency and video quality [2]. Adaptive rate control algorithms adjust quantization parameters based on content characteristics, balancing compression efficiency with video quality [39]. Adaptive prediction techniques, like those in the Residual-Guided In-Loop Filter (RRNet), integrate prediction information with model selection strategies, adapting to content variability, improving video quality and compression efficiency [44].

Adaptive and content-aware techniques significantly advance video compression technologies by optimizing neural network operations to accommodate content variability. Dynamic adjustments and content-specific optimizations improve efficiency and effectiveness in video delivery. Machine learning-based approaches outperform traditional codecs, achieving file sizes up to 60

8 Challenges and Future Directions

8.1 Future Directions and Optimization

The advancement of video compression technologies, particularly those incorporating neural networks like High Efficiency Video Coding (HEVC), presents substantial opportunities for optimization. Current research emphasizes refining feedback mechanisms and exploring complexity measures to enhance the robustness of Lookahead Frame-based Adaptive Mode (LFAM) across diverse coding scenarios. This could lead to adaptive algorithms that dynamically respond to varying content characteristics [67, 20]. Optimizing convolutional neural networks (CNNs) for video compression is crucial, with a focus on extending the Variable-filter-size Residue-learning CNN (VRCNN) for HEVC inter coding and simplifying network architectures while maintaining efficiency [34]. Developing robust machine learning models that adapt to diverse video characteristics, alongside hybrid approaches merging statistical and machine learning techniques, are promising areas for further exploration [33].

Enhancements in genetic algorithms and the incorporation of additional contextual information can significantly improve predictive accuracy in video compression [35]. Moreover, refining models to include additional coding tools and validating them on consumer-level decoders is essential to ensure practical applicability [1]. Hybrid approaches that integrate deep learning with traditional methods also show promise, potentially yielding substantial improvements in video compression efficiency and quality [6]. Future research may focus on enhancing encryption security while maintaining flexible visual information access, marking an important area for HEVC advancements.

The literature highlights the transformative potential of advanced video compression technologies through the integration of deep learning and machine learning. Innovations such as learned video compression algorithms and collaborative compression frameworks for intelligent analytics indicate a shift towards more adaptive approaches that enhance compression performance and visual fidelity [21, 24, 48, 51, 68]. Addressing current challenges and exploring new methodologies can lead to significant improvements in efficiency, quality, and applicability across various video compression scenarios.

8.2 Integration with Emerging Technologies

Integrating emerging technologies with existing video compression standards, such as HEVC, offers substantial opportunities to enhance efficiency and quality. Reconfigurable approximate computing frameworks enable dynamic adaptation to varying video content characteristics, thereby improving compression efficiency and quality [69]. Emerging machine learning techniques have notably advanced video compression, with a two-stage method demonstrating an average complexity reduction of 47.4

Exploiting parallelism in the decoding process is critical for reducing real-time video decoding time [38]. Advanced parallel processing techniques ensure efficient handling of high-resolution content, meeting modern application demands. Self-supervised learning techniques mitigate the reliance on manual annotations, enabling advanced video compression systems to autonomously extract meaningful representations from video data, enhancing adaptability and performance across various scenarios, including low-latency applications. This approach outperforms traditional codecs by producing smaller file sizes—up to 60

The development of lightweight CNN-based prediction methods ensures high performance across diverse encoding scenarios. These methods are particularly advantageous for resource-efficient scenarios, such as real-time video analytics and streaming applications, while maintaining high video quality through techniques like collaborative compression and intelligent analytics [70, 24, 29, 48]. Integrating data-independent low-complexity KLT approximations, such as the SKLT, significantly reduces arithmetic complexity, making these methods suitable for real-time applications.

Integrating emerging technologies with current video compression practices offers substantial opportunities for enhancing efficiency, reducing complexity, and improving video quality. By harnessing advancements in reconfigurable computing, machine learning, self-supervised learning, and parallel processing, the field of video compression is poised to meet the increasing demands for high-quality video delivery across diverse applications. Recent developments in learned video compression algorithms demonstrate significant efficiency improvements over traditional codecs, reducing file

AI-generated, for reference only.

sizes while enhancing visual quality and minimizing artifacts. Collaborative compression strategies, such as those seen in Video Coding for Machines (VCM), bridge the gap between machine vision and human perception, facilitating intelligent analytics and feature extraction. This evolution in video compression technology is essential for addressing the growing complexity of video content and its applications in smart cities, entertainment, and beyond [69, 24, 48, 51].

9 Conclusion

The exploration of High Efficiency Video Coding (HEVC) and Convolutional Neural Networks (CNNs) underscores their pivotal roles in advancing video compression technologies. HEVC's superior compression capabilities and support for high-resolution formats address the increasing demand for high-quality video content. The integration of CNNs into HEVC enhances decision-making precision, reduces computational demands, and improves video quality, solidifying its position as a leading standard in video compression.

Innovative techniques like the Block Modulating Video Compression (BMVC) demonstrate the potential of alternative codecs in environments with limited resources, enhancing image quality and reducing quantization artifacts. Additionally, advancements in multi-encoding strategies highlight the need for optimizing encoding processes to facilitate faster video delivery in adaptive streaming contexts.

The development of methods such as the Graph-based Light Field Transform (GLT) illustrates significant improvements in coding efficiency and redundancy reduction, essential for the effective management of high-quality light field images. These advancements reflect the ongoing evolution in video compression technologies and signal promising directions for future research focused on enhancing both efficiency and quality.

The synergy between HEVC and CNNs represents a significant advancement in the field of video compression. Leveraging the strengths of both technologies offers substantial improvements in video quality and compression efficiency, while also enhancing adaptability to new technological challenges. Continued research into innovative methodologies and the integration of emerging technologies promises to drive further innovations in video compression.

References

[1] Matthias Kränzler, Christian Herglotz, and André Kaup. Decoding energy modeling for versatile video coding, 2022.

[2] Yue Li, Dong Liu, Houqiang Li, Li Li, Feng Wu, Hong Zhang, and Haitao Yang. Convolutional neural network-based block up-sampling for intra frame coding, 2017.

[3] Saeed ur Rehman and Gulistan Raja. Performance evaluation of hevc over broadband networks, 2014.

[4] D. R. Canterle, T. L. T. da Silveira, F. M. Bayer, and R. J. Cintra. A multiparametric class of low-complexity transforms for image and video coding, 2020.

[5] Wenying Wen, Rongxin Tu, Yushu Zhang, Yuming Fang, and Yong Yang. A multi-level approach with visual information for encrypted h.265/hevc videos, 2020.

[6] Dong Liu, Yue Li, Jianping Lin, Houqiang Li, and Feng Wu. Deep learning-based video coding: A review and a case study. *ACM Computing Surveys (CSUR)*, 53(1):1–35, 2020.

[7] Muhammet Sebul Beratoğlu and Behçet Uğur Töreyin. Vehicle detection and classification without residual calculation: Accelerating hevc image decoding with random perturbation injection, 2023.

[8] Chuanmin Jia, Shiqi Wang, Xinfeng Zhang, Shanshe Wang, Jiaying Liu, Shiliang Pu, and Siwei Ma. Content-aware convolutional neural network for in-loop filtering in high efficiency video coding. *IEEE Transactions on Image Processing*, 28(7):3343–3356, 2019.

[9] Jacob Walker, Abhinav Gupta, and Martial Hebert. Dense optical flow prediction from a static image, 2015.

[10] Di Ma, Fan Zhang, and David R. Bull. Mfrnet: A new cnn architecture for post-processing and in-loop filtering, 2020.

[11] Yawei Li, Wen Li, Martin Danelljan, Kai Zhang, Shuhang Gu, Luc Van Gool, and Radu Timofte. The heterogeneity hypothesis: Finding layer-wise differentiated network architectures, 2021.

[12] Hilmi E. Egilmez, Ankitesh K. Singh, Muhammed Coban, Marta Karczewicz, Yinhao Zhu, Yang Yang, Amir Said, and Taco S. Cohen. Transform network architectures for deep learning based end-to-end image/video coding in subsampled color spaces, 2021.

[13] Abhimanyu Dubey, Moitreya Chatterjee, and Narendra Ahuja. Coreset-based neural network compression, 2018.

[14] Ahmed Telili, Ibrahim Farhat, Wassim Hamidouche, and Hadi Amirpour. Odvista: An omnidirectional video dataset for super-resolution and quality enhancement tasks, 2024.

[15] Mats L. Richter, Julius Schöning, Anna Wiedenroth, and Ulf Krumnack. Should you go deeper? optimizing convolutional neural network architectures without training by receptive field analysis, 2021.

[16] Jan P. Klopp, Liang-Gee Chen, and Shao-Yi Chien. Utilising low complexity cnns to lift non-local redundancies in video coding, 2020.

[17] Fatemeh Nasiri, Wassim Hamidouche, Luce Morin, Nicolas Dhollande, and Gildas Cocherel. Model selection cnn-based vvc qualityenhancement, 2021.

[18] Emre Can Kaya and Ioan Tabus. Lossless compression of point cloud sequences using sequence optimized cnn models, 2022.

[19] Arthur Cerveira, Luciano Agostini, Bruno Zatt, and Felipe Sampaio. Memory assessment of versatile video coding, 2020.

[20] Yongfei Zhang, Chao Zhang, Rui Fan, Siwei Ma, Zhibo Chen, and C. C. Jay Kuo. Recent advances on hevc inter-frame coding: From optimization to implementation and beyond, 2019.

AI-generated, for reference only.

[21] Siwei Ma, Xinfeng Zhang, Chuanmin Jia, Zhenghui Zhao, Shiqi Wang, and Shanshe Wang. Image and video compression with neural networks: A review. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(6):1683–1698, 2019.

[22] Jianfei Li, Han Feng, and Xiaosheng Zhuang. Convolutional neural networks for spherical signal processing via spherical haar tight framelets, 2022.

[23] Grace W Lindsay. Convolutional neural networks as a model of the visual system: Past, present, and future. *Journal of cognitive neuroscience*, 33(10):2017–2031, 2021.

[24] Pradyumn Patil, Vishwajeet Pawar, Yashraj Pawar, and Shruti Pisal. Video content classification using deep learning, 2021.

[25] Takehiro Tanaka, Hyomin Choi, and Ivan V. Bajić. Sfu-hw-tracks-v1: Object tracking dataset on raw video sequences, 2021.

[26] Michael Gygli. Ridiculously fast shot boundary detection with fully convolutional neural networks, 2017.

[27] Li Li, Zhu Li, Bin Li, Dong Liu, and Houqiang Li. Pseudo sequence based 2-d hierarchical reference structure for light-field image compression, 2016.

[28] Xiaoyi He, Qiang Hu, Xintong Han, Xiaoyun Zhang, Chongyang Zhang, and Weiyao Lin. Enhancing hevc compressed videos with a partition-masked convolutional neural network, 2018.

[29] Angeliki V. Katsenou, Fan Zhang, Mariana Afonso, Goce Dimitrov, and David R. Bull. Bvi-cc: A dataset for research on video compression and quality assessment, 2022.

[30] Amin Banitalebi-Dehkordi, Maryam Azimi, Mahsa T. Pourazad, and Panos Nasiopoulos. Compression of high dynamic range video using the hevc and h.264/avc standards, 2018.

[31] Lee Prangnell. Frequency-dependent perceptual quantisation for visually lossless compression applications, 2019.

[32] Shaoshi Yang, Cheng Zhou, Tiejun Lv, and Lajos Hanzo. Large-scale mimo is capable of eliminating power-thirsty channel coding for wireless transmission of hevc/h.265 video, 2016.

[33] Ekrem Cetinkaya, Hadi Amirpour, Mohammad Ghanbari, and Christian Timmerer. Ctu depth decision algorithms for hevc: A survey, 2021.

[34] Yuanying Dai, Dong Liu, and Feng Wu. A convolutional neural network approach for post-processing in hevc intra coding, 2016.

[35] Kevin Reuzé, Wassim Hamidouche, Pierrick Philippe, and Olivier Déforges. Efficient predictive coding of intra prediction modes, 2023.

[36] Fangyu Shen and Wei Gao. A rate control algorithm for video-based point cloud compression, 2022.

[37] Yue Chen, Debargha Murherjee, Jingning Han, Adrian Grange, Yaowu Xu, Zoe Liu, Sarah Parker, Cheng Chen, Hui Su, Urvang Joshi, et al. An overview of core coding tools in the av1 video codec. In *2018 picture coding symposium (PCS)*, pages 41–45. IEEE, 2018.

[38] Anup Saha, Wassim Hamidouche, Miguel Chavarrías, Guillaume Gautier, Fernando Pescador, and Ibrahim Farhat. Performance analysis of optimized versatile video coding software decoders on embedded platforms, 2022.

[39] Minhao Tang, Jiangtao Wen, and Yuxing Han. A generalized rate-distortion-λ model based hevc rate control algorithm, 2019.

[40] Ren Yang, Mai Xu, Zulin Wang, Yiping Duan, and Xiaoming Tao. Saliency-guided complexity control for hevc decoding, 2018.

[41] Mai Xu, Tianyi Li, Zulin Wang, Xin Deng, Ren Yang, and Zhenyu Guan. Reducing complexity of hevc: A deep learning approach, 2018.

[42] Tianyi Li, Mai Xu, Runzhi Tang, Ying Chen, and Qunliang Xing. Deepqtmt: A deep learning approach for fast qtmt-based cu partition of intra-mode vvc, 2021.

[43] Zhong Qiu Lin, Brendan Chwyl, and Alexander Wong. Edgesegnet: A compact network for semantic segmentation, 2019.

[44] Wei Jia, Li Li, Zhu Li, xiang zhang, and Shan Liu. Residual-guided in-loop filter using convolution neural network, 2021.

[45] Haojie Liu, Tong Chen, Ming Lu, Qiu Shen, and Zhan Ma. Neural video compression using spatio-temporal priors, 2019.

[46] Manuel Rebol and Patrick Knöbelreiter. Frame-to-frame consistent semantic segmentation, 2020.

[47] Sifeng Xia, Wenhan Yang, Yueyu Hu, Siwei Ma, and Jiaying Liu. A group variational transformation neural network for fractional interpolation of video coding, 2018.

[48] Lingyu Duan, Jiaying Liu, Wenhan Yang, Tiejun Huang, and Wen Gao. Video coding for machines: A paradigm of collaborative compression and intelligent analytics. *IEEE Transactions on Image Processing*, 29:8680–8695, 2020.

[49] Matt Poyser, Amir Atapour-Abarghouei, and Toby P. Breckon. On the impact of lossy image and video compression on the performance of deep convolutional neural network architectures, 2020.

[50] L. Balaji and K. K. Thyagharajan. An enhanced performance for h.265/shvc based on combined aegbm3d filter and back-propagation neural network, 2020.

[51] Oren Rippel, Sanjay Nair, Carissa Lew, Steve Branson, Alexander G Anderson, and Lubomir Bourdev. Learned video compression. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 3454–3463, 2019.

[52] Marc Górriz, Saverio Blasi, Alan F. Smeaton, Noel E. O'Connor, and Marta Mrak. Chroma intra prediction with attention-based cnn architectures, 2020.

[53] Hengyu Man, Xiaopeng Fan, Ruiqin Xiong, and Debin Zhao. Tree-structured data clustering-driven neural network for intra prediction in video coding, 2022.

[54] Yizhi Liu, Yao Wang, Ruofei Yu, Mu Li, Vin Sharma, and Yida Wang. Optimizing cnn model inference on cpus, 2019.

[55] Qianru Zhang, Meng Zhang, Tinghuan Chen, Zhifei Sun, Yuzhe Ma, and Bei Yu. Recent advances in convolutional neural network acceleration, 2018.

[56] Wentao Chen, Hailong Qiu, Jian Zhuang, Chutong Zhang, Yu Hu, Qing Lu, Tianchen Wang, Yiyu Shi, Meiping Huang, and Xiaowe Xu. Quantization of deep neural networks for accurate edge computing, 2021.

[57] Heming Sun, Zhengxue Cheng, Masaru Takeuchi, and Jiro Katto. Enhanced intra prediction for video coding by using multiple neural networks, 2020.

[58] Jack Kosaian and Amar Phanishayee. A study on the intersection of gpu utilization and cnn inference, 2022.

[59] Jie Chen, Junhui Hou, and Lap-Pui Chau. Light field compression with disparity guided sparse coding based on structural key views, 2017.

[60] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. Segnet: A deep convolutional encoder-decoder architecture for image segmentation, 2016.

[61] Adam W. Harley, Konstantinos G. Derpanis, and Iasonas Kokkinos. Segmentation-aware convolutional networks using local attention masks, 2017.

[62] Philip G Brodrick, Andrew B Davies, and Gregory P Asner. Uncovering ecological patterns with convolutional neural networks. *Trends in ecology & evolution*, 34(8):734–745, 2019.

[63] Roman Solovyev, Alexander Kustov, Dmitry Telpukhov, Vladimir Rukhlov, and Alexandr Kalinin. Fixed-point convolutional neural network for real-time video processing in fpga, 2020.

[64] Johnathan Chiu. Modeling image quantization tradeoffs for optimal compression, 2021.

[65] T. L. T. da Silveira, F. M. Bayer, R. J. Cintra, S. Kulasekera, A. Madanayake, and A. J. Kozakevicius. An orthogonal 16-point approximate dct for image and video compression, 2016.

[66] Valentin Radu, Kuba Kaszyk, Yuan Wen, Jack Turner, Jose Cano, Elliot J. Crowley, Bjorn Franke, Amos Storkey, and Michael O'Boyle. Performance aware convolutional neural network channel pruning for embedded gpus, 2020.

[67] Hongfei Fan, Lin Ding, Xiaodong Xie, Huizhu Jia, and Wen Gao. Joint rate allocation with both look-ahead and feedback model for high efficiency video coding, 2018.

[68] Hossein Talebi, Damien Kelly, Xiyang Luo, Ignacio Garcia Dorado, Feng Yang, Peyman Milanfar, and Michael Elad. Better compression with deep pre-editing, 2021.

[69] Francesca Palumbo and Carlo Sau. Reconfigurable and approximate computing for video coding, 2021.

[70] Hadi Amirpour, Vignesh V Menon, Ekrem Çetinkaya, Adithyan Ilangovan, Christian Feldmann, Martin Smole, and Christian Timmerer. Fast multi-encoding to reduce the cost of video streaming, 2022.

Disclaimer:

SurveyX is an AI-powered system designed to automate the generation of surveys. While it aims to produce high-quality, coherent, and comprehensive surveys with accurate citations, the final output is derived from the AI's synthesis of pre-processed materials, which may contain limitations or inaccuracies. As such, the generated content should not be used for academic publication or formal submissions and must be independently reviewed and verified. The developers of SurveyX do not assume responsibility for any errors or consequences arising from the use of the generated surveys.

AI-generated, for reference only.