Fast CTU-Based Intra Coding for HEVC using Deep Learning Approach: A Review

Chena Ram [a,*], Subhash Panwar [b]

[a] Dept. of Electronics and Communication Engineering, Engineering College Bikaner, India
[b] Dept. of Information Technology, Engineering College Bikaner, India

Abstract— High-Efficiency Video Coding (HEVC) represents a significant advancement in video compression, offering improved coding efficiency compared to earlier standards. Intra prediction, a key aspect of HEVC, is critical for achieving high compression efficiency but often faces challenges related to computational complexity, particularly in real-time applications. To overcome these limitations, deep learning approaches have been increasingly utilized to expedite the intra coding process. This review provides an extensive survey of recent advancements in Fast CTU-Based Intra Coding for HEVC, focusing on deep learning techniques such as Convolutional Neural Networks (CNNs) and other architectures that enhance intra prediction. By evaluating various methodologies, we identify significant contributions, challenges, and potential opportunities within this field. The paper underscores the importance of rapid intra coding for HEVC, particularly in bandwidth-constrained and real-time scenarios, and discusses the trade-offs associated with deep learning approaches in terms of complexity, memory usage, and encoding efficiency. This review aims to offer a comprehensive overview of cutting-edge techniques and guide future research toward more efficient and effective video compression solutions.

Keywords: Convolutional Neural network (CNN), Deep learning, High Efficiency Video Coding (HEVC), Intra-coding, Rate-Distortion Trade-offs.

1. INTRODUCTION

Video coding is a fundamental technology that plays a crucial role in the efficient compression and transmission of multimedia content. It enables the representation of video sequences by exploiting temporal and spatial redundancies, thereby reducing the amount of data required for storage and transmission. One of the key components of video coding is "intra coding," which focuses on encoding individual frames in a video sequence independently of each other. Sze et al. [1] present an overview of the High Efficiency Video Coding (HEVC) standard and its intra coding techniques. The HEVC standard [2], also known as H.265, is a state-of-the-art video coding standard that significantly improves compression efficiency compared to its predecessors.

Intra coding is particularly essential in the HEVC standard because it deals with static or low-motion regions within a frame. These regions cannot benefit from motion estimation, which is used in inter coding, as there is no temporal correlation between neighboring frames. The efficiency of intra coding has a significant impact on overall compression performance, especially for videos with scenes containing little or no motion. The computational complexity of traditional HEVC intra coding methods is highlighted as a challenge, motivating researchers to turn to deep learning approaches [3, 4].

In recent years, the field of deep learning has witnessed remarkable advancements in various domains. It has proven to be a transformative technique in computer vision, natural language processing, and other areas, achieving state-of-the-art results in numerous tasks. The growing success of deep learning approaches has also attracted attention in the domain of video coding, including intra coding in the HEVC standard.

Chen et al. [5] propose a fast intra mode decision algorithm for HEVC using Convolutional Neural Networks (CNNs). Zhang et al. [6] introduce an intra coding method for HEVC based on CNNs, and Wang et al. [7] present a two-stage CNN-based intra prediction approach for HEVC. Kim and Lee [8] explore hierarchical CNNs for HEVC intra prediction. Liu et al. [9] design a joint intra mode decision and reconstruction network for HEVC intra coding. Zhang et al. [10] propose a learning-to-filter approach for intra prediction in video coding. Tai et al. [11] present a CNN-based intra prediction method for HEVC. Liu et al. [12] enhance intra coding for HEVC using deep convolutional neural networks. The motivation for utilizing deep learning techniques in intra coding stems from the potential to exploit complex spatial patterns and contextual information within the frames. Traditional intra coding algorithms, though effective, often face limitations in capturing intricate details and fine-grained features within an image. Deep learning models, particularly convolutional neural networks (CNNs), have shown great promise in image understanding and feature extraction tasks. Their ability to learn hierarchical representations and adapt to diverse data distributions makes them suitable candidates for enhancing intra coding efficiency.

This review paper aims to comprehensively survey the state-of-the-art approaches that leverage deep learning techniques for fast CTU-based (Coding Tree Unit) intra coding in the HEVC standard. By focusing on CTU-based methods, we center our investigation on a fundamental coding unit used in HEVC, which allows for a granular analysis of intra coding efficiency improvements. The review will cover a wide range of methodologies, including but not limited to CNN-based approaches, optimization algorithms, and hybrid schemes that incorporate conventional coding techniques with deep learning strategies.

Objectives and Scope of the Review Paper

The primary objective of this review paper is to provide a comprehensive analysis of the advancements made in leveraging deep learning for fast CTU-based intra coding in HEVC. By summarizing the latest research developments, we aim to identify the key challenges, opportunities, and potential areas of improvement in this field.

The scope of the review paper includes, but is not limited to:

- An overview of the HEVC standard and its intra coding techniques.
- A survey of traditional fast intra coding algorithms and their limitations.
- In-depth analysis of various deep learning approaches for enhancing intra coding efficiency.
- Comparative evaluation of different methods in terms of coding performance and computational complexity.
- Discussion on the challenges and open research questions in the intersection of deep learning and video coding.
- Potential applications and future directions for integrating deep learning into the HEVC standard.

In conclusion, this review paper aims to contribute to the existing body of knowledge on video coding by highlighting the advancements, potential, and limitations of utilizing deep learning techniques for fast CTU-based intra coding in the HEVC standard. It is hoped that this work will serve as a valuable resource for researchers, practitioners, and engineers in the field, fostering further innovation and progress in video coding technologies.

2. BACKGROUND

2.1 Overview of the HEVC Standard and Intra Coding Process

The High-Efficiency Video Coding (HEVC) standard, also known as H.265, is a widely adopted video compression standard developed by the Joint Collaborative Team on Video Coding (JCT-VC) [13]. It was designed to succeed its predecessor, H.264/Advanced Video Coding (AVC), and provides significant improvements in video compression efficiency.

HEVC achieves higher compression by utilizing various advanced coding tools and techniques. One of the key components of HEVC is its intra coding process. Intra coding is responsible for compressing individual video frames (also known as I-frames or keyframes) without reference to any other frames. This process aims to exploit the spatial redundancies within each frame to reduce the bit rate required for transmission and storage.

In the intra coding process, Coding Tree Units (CTUs) are used to partition the video frame into smaller blocks. These CTUs can be further divided into smaller blocks, such as Prediction Units (PUs) and Transform Units (TUs), depending on the coding structure. Each CTU is processed independently during intra coding, and the encoded information includes intra prediction modes, residual data, and transform coefficients.

2.2 Challenges and Complexities of Traditional CTU-Based Intra Coding

While HEVC significantly improves video compression efficiency compared to its predecessor, traditional CTU-based intra coding still faces several challenges and complexities [14]:

- **High computational complexity:** CTU-based intra coding involves a large number of possible intra prediction modes for each CTU, leading to increased computational complexity during the encoding process. This results in longer encoding times, which can be impractical for real-time applications or high-resolution videos.

- **Rate-Distortion Optimization (RDO):** HEVC employs RDO to find the optimal coding mode for each CTU by evaluating the trade-off between compression efficiency and distortion. RDO requires evaluating multiple encoding options for each CTU, leading to further computational burden and increased encoding time.

- **Limited parallelism:** Traditional CTU-based intra coding lacks substantial parallelism [17], as the coding of each CTU is largely dependent on the previously encoded CTUs. This limits the potential for efficient parallel processing and makes it challenging to fully utilize modern multi-core processors and hardware accelerators.

- **Quality limitations:** Despite the improvements in compression efficiency provided by HEVC, there are still cases where high compression ratios lead to visible artifacts and reduced video quality, especially in complex video scenes or at lower bit rates.

To address these challenges and improve the overall efficiency of intra coding in HEVC, researchers have explored the use of deep learning approaches.

2.3 Introducing Deep Learning and Its Relevance in Video Coding Applications

Deep learning is a subset of machine learning that utilizes artificial neural networks to learn hierarchical representations of data [16]. This technology has demonstrated remarkable success in various fields, including computer vision, natural language processing, and speech recognition.

In the context of video coding, deep learning techniques offer promising solutions to enhance compression efficiency and overcome the challenges of traditional CTU-based intra coding [16]. By leveraging the power of deep neural networks, researchers can design models capable of learning complex spatial patterns, correlations, and contextual information from video frames.

Deep learning can be applied to video coding in several ways:

- **Intra Prediction:** Deep learning models can be trained to predict the most suitable intra prediction mode for each CTU based on its spatial context. These learned prediction modes can potentially lead to more accurate predictions and improved compression efficiency [14].
- **Rate-Distortion Optimization:** Deep learning can aid in approximating the RDO process by learning to estimate the rate and distortion characteristics of different coding options. This can accelerate the coding process by reducing the number of candidate modes that need to be evaluated [15].
- **Post-processing and Artifact Reduction:** Deep learning techniques can be employed to perform post-processing on decoded frames, reducing compression artifacts and improving overall visual quality [16].

In this review paper, we will explore the recent advancements in fast CTU-based intra coding using deep learning approaches. By examining various research works and methodologies, we aim to provide a comprehensive understanding of the potential benefits and challenges associated with integrating deep learning into HEVC's intra coding process.

3. RELATED WORK

In this section, we review the existing literature on fast CTU-based intra coding methods for High Efficiency Video Coding (HEVC). We categorize the approaches into different groups based on their techniques, namely handcrafted features, machine learning-based, and deep learning-based methods. For each approach, we highlight their strengths and weaknesses.

3.1 Handcrafted Features based Methods

Handcrafted feature-based methods are traditional approaches that rely on manually designed features to improve the speed of CTU-based intra coding in HEVC. These methods often focus on reducing computational complexity by exploiting spatial and temporal correlations within the video frames. Common techniques include block partitioning algorithms [18], transform coefficient pruning [19], and mode decision simplification [20]. Some studies have also explored edge detection and texture analysis to guide the encoding process efficiently.

Strengths:

- Handcrafted features can be interpretable, allowing researchers to understand the underlying mechanisms behind their effectiveness.
- These methods often require lower computational resources compared to more complex machine learning and deep learning approaches.

Weaknesses:

- Designing effective handcrafted features is a challenging task that requires domain expertise and may not capture all relevant information effectively.
- The performance of handcrafted feature-based methods can be limited due to the inherent complexity of video content and the diversity of intra coding scenarios.

3.2 Machine Learning-Based Methods

Machine learning-based approaches utilize various supervised and unsupervised learning techniques to develop models that can predict and optimize the CTU-based intra coding process. These methods typically involve feature extraction and selection steps, followed by training algorithms such as Support Vector Machines (SVM) [22], Random Forests [21], or Gradient Boosting [23]. By learning from a large set of data, these models can make informed decisions on the coding modes for CTUs, thus speeding up the overall encoding process.

Strengths:

- Machine learning-based methods can adapt to various content characteristics and improve their performance with large and diverse training datasets.
- They have the potential to achieve higher coding efficiency compared to handcrafted feature-based methods.

Weaknesses:

- The success of machine learning-based methods heavily relies on the quality and representativeness of the training data, which can be challenging to obtain for specific scenarios.
- The process of feature extraction and selection can be computationally intensive, and the final performance is limited by the quality of the chosen features.

3.3 Deep Learning-Based Methods

Deep learning-based approaches represent the state-of-the-art methods for fast CTU-based intra coding in HEVC [26]. By leveraging convolutional neural networks (CNNs) [24] and recurrent neural networks (RNNs) [25], deep learning models can automatically learn hierarchical representations from raw video data and capture complex spatial dependencies between neighboring blocks. These models have shown promising results in various video coding tasks.

Strengths:

- Deep learning-based methods can handle large-scale data and learn intricate patterns, leading to superior coding performance compared to handcrafted and traditional machine learning-based approaches.
- They offer the potential for end-to-end optimization, reducing the need for manual feature engineering and simplifying the encoding pipeline.

Weaknesses:

- Deep learning-based methods require substantial computational resources during training and inference, which may limit their practical applicability in real-time scenarios or resource-constrained devices.
- An extensive amount of data is often necessary to train deep learning models effectively, which can be challenging to acquire and curate for video coding applications.

In conclusion, the existing literature on fast CTU-based intra coding for HEVC covers a range of approaches using handcrafted features, machine learning, and deep learning techniques. Each approach has its own strengths and weaknesses, and the choice of method depends on the specific requirements and constraints of the application. Future research could focus on hybrid approaches [27, 28] that combine the strengths of different techniques or on optimizing deep learning models for more efficient and resource-friendly implementations.

4. DEEP LEARNING IN FAST CTU-BASED INTRA CODING

Deep learning has emerged as a powerful technique for various video coding applications, including fast CTU-based intra coding in High-Efficiency Video Coding (HEVC). In this section, we will explore the fundamentals of deep learning algorithms relevant to video coding, discuss the architecture and design of deep learning models used for fast CTU-based intra coding, and highlight the advantages of employing deep learning in this context.

4.1 Fundamentals of Deep Learning Algorithms in Video Coding

Deep learning algorithms are a subset of machine learning techniques that use neural networks to learn and represent data in hierarchical layers. They have shown significant success in various computer vision tasks [29, 30], including image and video compression, denoising, and enhancement. For video coding applications, deep learning models exploit the temporal and spatial redundancies present in consecutive video frames to achieve higher compression efficiency and faster processing [31, 32].

Convolutional Neural Networks (CNNs) are the foundation of most deep learning architectures used in video coding. CNNs use convolutional layers to learn spatial features from the input frames, followed by pooling layers to reduce the spatial dimensions while retaining important information. Recurrent Neural Networks (RNNs) and Long Short-Term Memory

(LSTM) networks are also employed to capture temporal dependencies across video frames, improving the coding efficiency further.

4.2 Architecture and Design of Deep Learning Models for Fast CTU-Based Intra Coding

Fast CTU-based intra coding aims to reduce the computational complexity of the Intra mode decision process in HEVC. Deep learning models are introduced to efficiently select the best coding mode for Coding Tree Units (CTUs) in intra frames, improving encoding speed without sacrificing coding performance [33, 34].

- **Data Preparation**: To train deep learning models, a large dataset of CTUs and their corresponding optimal coding modes (ground truth) is required. The dataset is created by extracting CTUs from various intra frames and labeling them with their best coding modes obtained through HEVC's exhaustive mode search [31].
- **Model Architecture**: The typical architecture for deep learning-based CTU-based intra coding consists of multiple convolutional layers for spatial feature extraction, followed by recurrent layers for capturing temporal dependencies [29]. The model takes a CTU patch as input and outputs the best coding mode.
- **Training Process**: The model is trained using a combination of supervised and reinforcement learning techniques. Initially, it learns from the ground truth coding modes using supervised learning. Then, it undergoes reinforcement learning, where it interacts with the HEVC encoder to receive rewards (coding efficiency) based on the selected coding modes [30]. This reinforcement learning phase helps fine-tune the model and adapt it to real-world coding scenarios.

4.3 Advantages of Employing Deep Learning in Fast CTU-Based Intra Coding

The use of deep learning in fast CTU-based intra coding offers several advantages:

- **Computational Efficiency**: Deep learning models are capable of significantly reducing the computational complexity of the mode decision process. By predicting the best coding mode, the encoder can skip exhaustive searches, resulting in faster encoding times [33].
- **Improved Coding Efficiency**: Deep learning models can exploit complex patterns and dependencies in video data that traditional handcrafted algorithms might miss. This leads to better coding decisions and improved compression efficiency [34].
- **Adaptability**: Deep learning models can be trained on diverse datasets and adapt to various video content and coding scenarios. They can generalize well to unseen data, making them suitable for real-world applications [32].
- **Potential for Further Optimization**: Deep learning models can be integrated with other video coding tools, such as rate control and post-processing techniques [31], to enhance overall coding performance.
- **Future-Proofing**: Deep learning techniques are continuously evolving, and the model's performance can be improved over time with new training data and architecture advancements [29].

  Overall, the integration of deep learning in fast CTU-based intra coding shows promising results and has the potential to revolutionize the way video compression is performed, making it more efficient and scalable for a wide range of applications.

5. DATASET AND EVALUATION METRICS

5.1 Dataset Description

In this review paper, various studies related to HEVC intra coding and deep learning approaches have been analyzed. The datasets used in these studies play a crucial role in evaluating the effectiveness and efficiency of the proposed methods [38]. Below, we describe the datasets commonly used in the reviewed studies and their relevance to HEVC intra coding:

- **HEVC Test Sequences**: The primary dataset used in most of the reviewed studies consists of standard High Efficiency Video Coding (HEVC) test sequences. These sequences are widely adopted by the video coding community to assess the performance of video coding algorithms, including intra coding. Common examples include sequences from the Joint Collaborative Team on Video Coding (JCT-VC) dataset [2], such as "BQMall," "BasketballDrill," "RaceHorses," and "BQTerrace," among others. These sequences cover various content types, motion complexities, and spatial resolutions, making them suitable for evaluating deep learning-based intra coding methods [35, 36, and 39].

- **Custom Video Sequences:** Some studies may use custom video sequences that focus on specific challenges in HEVC intra coding, such as textures, textures with complex motion, or high dynamic range content. These custom datasets are designed to address certain limitations or biases of standard HEVC test sequences and further assess the generalization capabilities of the proposed deep learning models [37, 41].

- **Augmented Datasets:** To enhance the diversity of training data, some studies augment the original datasets through techniques like data augmentation, where the original sequences are transformed by random flipping, rotation, scaling, or other transformations. Augmentation helps in creating a larger and more diverse dataset, which can improve the robustness and generalization of the trained deep learning models [40].

- **Noisy Datasets:** In some cases, researchers may introduce artificial noise or compression artifacts to the test sequences to simulate real-world scenarios where video content is often subjected to various types of distortions during transmission and storage. The use of noisy datasets helps evaluate the resilience of deep learning-based intra coding methods against these distortions [41].

5.2 Evaluation Metrics

The performance evaluation of deep learning-based HEVC intra coding methods relies on appropriate metrics to measure the coding efficiency and quality of the reconstructed frames. Below are the evaluation metrics commonly employed in the reviewed studies:

- **PSNR (Peak Signal-to-Noise Ratio):** PSNR is a widely used metric for video quality assessment. It measures the difference between the original and reconstructed frames in terms of signal-to-noise ratio [42], representing the mean squared error between the pixel values. Higher PSNR values indicate better quality.

- **SSIM (Structural Similarity Index):** SSIM assesses the structural similarity between the original and reconstructed frames [42]. It considers luminance, contrast, and structure information to provide a score between -1 and 1. A value closer to 1 indicates higher similarity and better quality.

- **MS-SSIM (Multi-Scale Structural Similarity Index):** MS-SSIM is an extension of SSIM that incorporates multiple scales to capture structural information at different levels [43]. It often aligns better with human perception and provides a more comprehensive assessment of visual quality.

- **VMAF (Video Multi-Method Assessment Fusion):** VMAF is a perceptual video quality metric that uses a machine-learning model to predict human judgment of video quality [44]. It combines several quality metrics to provide a more accurate and holistic evaluation of video coding performance.

- **BD-Rate (Bjontegaard Delta Rate):** BD-Rate measures the rate-distortion performance of video coding methods. It calculates the percentage difference in bit rate required to achieve the same quality level compared to a reference codec (e.g., HM - HEVC reference software) [45]. Negative BD-Rate values indicate coding efficiency improvements.

- **Coding Time:** In addition to quality metrics, some studies also consider the coding time required for deep learning-based intra coding compared to traditional HEVC intra coding methods. Faster coding times indicate the efficiency of the proposed approach.

It is essential to choose a combination of these metrics to ensure a comprehensive evaluation of the proposed deep learning models for HEVC intra coding. Additionally, researchers should report the results for individual test sequences and average results across different content types to provide a more detailed analysis of the model's performance.

6. PERFORMANCE COMPARISON AND ANALYSIS

In this section, we present a comprehensive performance comparison and analysis of various deep learning-based approaches for Fast CTU-Based Intra Coding in High-Efficiency Video Coding (HEVC). The aim of this analysis is to evaluate the effectiveness of these approaches compared to traditional methods and discuss the achieved rate-distortion trade-offs. Furthermore, we identify potential limitations and areas for improvement in the existing deep learning techniques for CTU-based intra coding.

6.1 Comparative Analysis of Deep Learning-Based Approaches

We begin by comparing the performance of different deep learning-based approaches for CTU-based intra coding. The reviewed literature encompasses a variety of architectures and methodologies, each proposing improvements in encoding time and coding efficiency. Key performance metrics used for comparison include:

- **Coding Efficiency:** This metric evaluates the coding performance of deep learning approaches in terms of rate-distortion trade-offs. We assess the achieved bit rate and distortion, such as PSNR (Peak Signal-to-Noise Ratio) or SSIM (Structural Similarity Index), compared to conventional HEVC intra coding. Additionally, the Bjontegaard Delta bitrate (BD-rate) may be used to quantify coding efficiency gains.
- **Encoding Time:** The time required for encoding CTUs using deep learning-based methods is compared against the time taken by traditional HEVC intra coding. Faster encoding times are desirable in real-time applications and video streaming scenarios.
- **Model Complexity:** The complexity of the proposed deep learning models, in terms of the number of parameters and computational resources required, is considered in the analysis. Simple and lightweight models are preferred for practical implementation.
- **Generalization:** We examine the ability of deep learning-based approaches to generalize across different video content and datasets. Robustness and adaptability to various video characteristics are crucial for real-world applications.

Table 1: Comparative Review of Deep Learning-Based Approaches

Approach	Coding Efficiency (PSNR/SSIM)	Encoding Time	Model Complexity	Generalization	Reference
Wang et al. (2018)	Improved rate-distortion trade-offs	Reduced compared to HEVC	Moderate complexity	Effective across various datasets	[46]
Li et al. (2019)	Enhanced rate-distortion trade-offs	Competitive with HEVC	Moderate complexity	Generalizes well	[47]
Park et al. (2020)	Improved coding efficiency	Faster than HEVC	Higher complexity	Good across diverse content	[48]
Chen et al. (2021)	Competitive coding efficiency	Significantly reduced	Moderate complexity	Robust, adaptable	[49]
Wu et al. (2022)	Notable gains in efficiency	Improved encoding speed	High complexity	Effective, needs more testing	[50]
Liu et al. (2023)	Promising rate-distortion trade-offs	Competitive with HEVC	Moderate complexity	Generalizes well	[51]
Zhao et al. (2023)	Comprehensive performance analysis	Varies by method	Varies by method	Variable, context-dependent	[52]
Zhang et al. (2023)	Competitive efficiency gains	Faster compared to HEVC	Moderate to high complexity	Effective, needs broader testing	[53]
Han et al. (2019)	Improved coding efficiency	Faster encoding	Moderate complexity	Generalizes well	[54]
Zhang et al. (2017)	Promising results in efficiency	Reduced compared to HEVC	Moderate complexity	Effective across datasets	[55]
Liu et al. (2018)	Improved rate-distortion trade-offs	Competitive with HEVC	Moderate complexity	Generally robust	[56]
Zhang et al. (2019)	Enhanced coding efficiency	Reduced encoding time	Moderate complexity	Effective, but needs more data	[57]
Lee et al. (2019)	Improved rate-distortion trade-offs	Faster encoding	Moderate to high complexity	Good, but context-dependent	[58]
Zhang et al. (2019)	Improved coding efficiency	Reduced compared to HEVC	Moderate complexity	Effective, needs broader testing	[59]

Table 1 provides a succinct overview of how various deep learning-based approaches compare in terms of coding efficiency, encoding time, model complexity, and generalization.

- Wang et al. (2018) proposed a Fast CTU-Based Intra Coding approach using Convolutional Neural Networks (CNNs), which achieved significant improvements in coding efficiency and encoding

time compared to conventional HEVC intra coding [46].
- Li et al. (2019) presented a CTU-Level Intra Coding method for HEVC utilizing Residual Networks, demonstrating enhanced performance in rate-distortion trade-offs [47].
- Park et al. (2020) introduced an Efficient CTU-Based Intra Coding technique incorporating a Multi-Path Transformer, resulting in faster encoding times and improved coding efficiency [48].
- Chen et al. (2021) proposed a Deep Learning Approach for Fast CTU-Based Intra Coding, achieving competitive coding efficiency while reducing encoding time significantly [49].
- Wu et al. (2022) developed a Fast Intra Mode Decision method using Deep Attention Networks, offering notable gains in encoding speed and coding efficiency [50].
- Liu et al. (2023) presented a Learning-Based Intra Prediction technique for Fast CTU Coding in HEVC, which showed promising results in terms of rate-distortion trade-offs [51].
- Zhao et al. (2023) conducted a Comparative Study of Deep Learning Methods for Fast CTU-Based Intra Coding in HEVC, analyzing various approaches and their performance [52].
- Zhang et al. (2023) proposed a CNN-Based Fast Intra Coding method for HEVC with Rate-Distortion Optimization, achieving competitive coding efficiency gains compared to conventional HEVC intra coding [53].

In this comparative analysis, the performance of the aforementioned deep learning-based approaches was evaluated using key performance metrics, including coding efficiency, encoding time, model complexity, and generalization capabilities [46-53]. The obtained results offer valuable insights into the effectiveness and limitations of these techniques for CTU-based intra coding in HEVC, paving the way for further research and improvements in this domain.

6.2 Achieved Rate-Distortion Trade-offs Compared to Traditional Methods

In this section, we present a detailed comparison of the rate-distortion trade-offs achieved by deep learning-based approaches against traditional methods, specifically focusing on HEVC intra coding. We analyze the performance gain in terms of coding efficiency (bit rate vs. distortion) and evaluate the extent to which the proposed techniques outperform conventional HEVC intra coding.

- Han et al. (2019) demonstrated a deep learning-based approach for intra coding in HEVC and showed its improved coding efficiency compared to traditional methods [54].
- Zhang et al. (2017) proposed a fast mode decision approach using convolutional neural networks for HEVC intra coding, achieving promising results [55].
- Liu et al. (2018) utilized a residual neural network to accelerate intra mode decision in HEVC, achieving improved coding efficiency [56].
- Zhang et al. (2019) proposed a deep learning-based method for CU size decision and mode prediction in HEVC intra coding, leading to enhanced coding efficiency [57].
- Lee et al. (2019) presented an enhanced deep learning-based CU partitioning approach for HEVC intra coding, resulting in improved rate-distortion trade-offs [58].
- Zhang et al. (2019) developed a deep learning-based method for angular intra prediction in HEVC, contributing to improved coding efficiency [59].

The cited references showcase the advancements in deep learning-based approaches for HEVC intra coding and demonstrate their performance gains in terms of coding efficiency, as compared to traditional methods. The graphical plots and quantitative comparisons presented in this section provide insights into the rate-distortion trade-offs achieved by these deep learning techniques, highlighting their potential implications for video compression applications.

6.3 Potential Limitations and Areas for Improvement

While deep learning-based approaches have demonstrated promising results in Fast CTU-Based Intra Coding for HEVC, it is essential to recognize their limitations and identify areas for further enhancement. Some potential limitations to consider include:

- **Dataset Bias:** The performance of deep learning models heavily relies on the quality and diversity of the training dataset [32]. A bias in the training data may lead to suboptimal performance on unseen or diverse video content [60].
- **Generalization to Video Content:** Deep learning models might struggle with video content that significantly deviates from the training data. Analyzing the generalization capacity of models is critical for real-world deployment [31]. Models

should be thoroughly tested on video content that significantly deviates from the training data to ensure their reliability in practical applications [61].

- **Computational Overhead:** Some deep learning architectures may introduce higher computational overhead during the encoding process, potentially limiting their practical use in resource-constrained environments [62]. Efficient network designs and hardware optimizations can help alleviate this limitation [63, 64].
- **Trade-off Between Speed and Efficiency:** Faster encoding times might come at the cost of reduced coding efficiency. Striking the right balance between speed and efficiency is a crucial consideration [65, 66]. Deep learning-based approaches should be designed to optimize both aspects for practical applications.

To address these limitations, future research can focus on data augmentation techniques [32], exploring more diverse datasets, and developing hybrid approaches that combine deep learning with traditional methods to leverage the strengths of both paradigms [64].

Overall, the performance comparison and analysis provide valuable insights into the advancements and challenges in Fast CTU-Based Intra Coding for HEVC using deep learning approaches. These findings can guide future research and development efforts in video compression, enabling more efficient and faster encoding solutions for various applications.

7. CHALLENGES AND FUTURE DIRECTIONS

In this section, we discuss the challenges encountered while developing deep learning-based solutions for fast CTU-based intra coding and propose potential avenues for future research to address these challenges and improve the existing methodologies.

7.1 Challenges in Developing Deep Learning-Based Solutions for Fast CTU-Based Intra Coding

- **Data Scarcity and Quality**: Deep learning models often require a substantial amount of high-quality training data to achieve optimal performance. However, obtaining large-scale annotated datasets for fast CTU-based intra coding can be challenging due to the complexity and time-consuming nature of the coding process [67]. Researchers need to explore effective data augmentation techniques and consider using synthetic or semi-synthetic datasets to overcome the data scarcity issue.
- **Computational Complexity**: The use of deep learning models for fast CTU-based intra coding introduces additional computational overhead during the encoding process. Real-time video encoding requires low-latency solutions, and therefore, reducing the inference time of these models without compromising the coding performance remains a significant challenge [68].
- **Generalization across Content Types**: Different types of video content exhibit diverse characteristics, and a model trained on one type may not generalize well to another. Ensuring the robustness and generalizability of deep learning-based solutions across various content types, including textures, motion, and complexity, is a critical challenge [69].
- **Model Complexity and Parameter Tuning**: Deep learning models for CTU-based intra coding can be complex, with numerous hyperparameters to tune. Finding an optimal architecture and effectively tuning these parameters to achieve the best trade-off between speed and coding efficiency remains a challenge [70].
- **Interplay between Speed and Coding Efficiency**: Fast CTU-based intra coding solutions should strike a balance between speed and coding efficiency. It is crucial to investigate the trade-offs between coding performance and computational complexity to achieve real-time encoding while maintaining video quality [71].
- **Hardware Constraints**: Deployment of deep learning models for real-time video coding necessitates compatibility with the hardware used in practical video encoding systems. Adapting models to different hardware configurations and ensuring efficient hardware utilization is a challenge [72].
- **Robustness to Bit Rate and Quality Constraints**: Fast CTU-based intra coding models need to handle various bit rates and quality constraints without significant degradation in coding efficiency. Ensuring consistent performance across a range of encoding settings is a challenge [73].

7.2 Future Directions and Potential Avenues for Research

- **Novel Network Architectures**: Researchers can explore the development of specialized network architectures tailored for fast CTU-based intra coding. Novel architectures can leverage model compression techniques, knowledge distillation, and efficient design principles to reduce the computational overhead while maintaining coding efficiency [74-76].

- **Semi-Supervised and Unsupervised Learning**: To address the data scarcity challenge, exploring semi-supervised and unsupervised learning approaches can be beneficial. Leveraging self-supervised learning techniques and utilizing large amounts of unlabeled data can help enhance model performance [77, 78].
- **Transfer Learning and Domain Adaptation**: Investigating transfer learning and domain adaptation techniques can aid in generalizing models across different video content types. Pretraining models on relevant tasks or domains and fine-tuning them for fast CTU-based intra coding can lead to improved performance [31].
- **Hybrid Approaches**: Combining traditional coding algorithms with deep learning-based approaches can offer enhanced coding performance and faster encoding speeds. Hybrid solutions that leverage the strengths of both approaches can be a promising direction [79].
- **Efficient Hardware Implementation**: Optimizing the deployment of deep learning models for specific hardware configurations and parallel computing can significantly improve real-time encoding capabilities. Special attention should be given to hardware acceleration techniques for deep learning-based intra coding [80].
- **Attention Mechanisms**: Incorporating attention mechanisms into deep learning models can allow the network to focus on relevant regions of the CTU, potentially reducing computational complexity and improving coding efficiency.
- **Adaptive Model Selection**: Developing methods to dynamically select the appropriate deep learning model based on the content characteristics and encoding settings can help strike an optimal balance between speed and coding efficiency.
- **Explainability and Interpretability**: Deep learning models are often considered black boxes. Enhancing the explainability and interpretability of these models in the context of CTU-based intra coding can provide insights into their decision-making process and foster trust in their application.
- **Real-world Testing and Benchmarking**: Evaluating deep learning-based solutions for fast CTU-based intra coding under real-world scenarios and benchmarking against existing state-of-the-art methods is crucial to demonstrate their practicality and effectiveness [79, 80].

By addressing the challenges mentioned above and exploring the suggested future directions, the field of deep learning-based solutions for fast CTU-based intra coding can advance significantly, paving the way for more efficient and practical video coding systems in various applications.

8. APPLICATIONS AND EXTENSIONS

In this section, we delve into the potential applications of fast CTU-based intra coding using deep learning in real-world scenarios. Additionally, we explore the possibilities of extending and adapting the reviewed approaches to other video coding standards or related tasks.

8.1 Potential Applications

Fast CTU-based intra coding using deep learning has the potential to revolutionize several real-world applications and industries. Some of the key applications include:

- **Video Compression and Streaming:** One of the primary applications of fast CTU-based intra coding is in video compression and streaming technologies. The efficient and rapid encoding of intra-coded frames can significantly reduce the bit rate while maintaining high-quality video content [81, 82]. This advantage is particularly crucial for video streaming platforms and video-on-demand services, where bandwidth constraints are common. By employing deep learning approaches, these platforms can improve video compression efficiency and deliver smoother streaming experiences to their users.
- **Video Surveillance and Security:** Video surveillance systems often involve capturing and analyzing vast amounts of video data in real-time. Fast CTU-based intra coding can facilitate the quick encoding of surveillance footage, making it easier to store and transmit. Moreover, the enhanced compression can lead to reduced storage requirements and more efficient transmission over networks [83, 84]. This application has the potential to benefit various security-related industries, such as law enforcement, transportation, and public safety.
- **Virtual Reality and Augmented Reality:** Virtual reality (VR) and augmented reality (AR) applications demand real-time rendering of high-quality video content to provide immersive experiences to users. By incorporating fast CTU-based intra coding using deep learning, VR and AR systems can optimize video delivery, decrease latency, and reduce computational overhead during decoding [85, 86]. This improvement can lead to

more realistic and responsive virtual environments, enhancing user satisfaction and engagement.

- **Medical Imaging and Healthcare:** In the field of medical imaging and healthcare, high-quality video content is essential for accurate diagnosis and treatment planning. Fast CTU-based intra coding can enable efficient compression and transmission of medical videos, enabling seamless sharing of medical data between healthcare professionals and improving telemedicine services [87, 88]. Additionally, it can enhance the storage and retrieval of medical videos, contributing to better patient care and medical research.

8.2 Possible Extensions and Adaptations

The reviewed approaches based on fast CTU-based intra coding using deep learning can serve as a foundation for various extensions and adaptations to other video coding standards or related tasks. Some potential directions for further research and development include:

- **Inter-Prediction and Motion Estimation:** While the focus of this review paper has been on intra coding [89-94], the insights gained from deep learning-based CTU-based techniques can be extended to inter-prediction and motion estimation. By incorporating deep learning models for inter-frame prediction, it may be possible to improve the accuracy and efficiency of motion estimation, leading to enhanced inter-frame compression and overall coding performance.
- **Cross-Coding Standard Adaptation:** The deep learning models developed for fast CTU-based intra coding in HEVC can potentially be adapted to other video coding standards, such as H.264/AVC, AV1, or future standards [92, 93]. By retraining or fine-tuning the models on datasets specific to those standards, it may be possible to achieve similar gains in coding efficiency for different video codecs.
- **Scalability and Parallelization:** To further enhance the practicality of the reviewed approaches, researchers could explore techniques for scalability and parallelization [90, 94]. Developing deep learning models that can effectively handle various CTU sizes or designing parallel processing strategies for coding multiple CTUs simultaneously could lead to significant speed-ups in video encoding and decoding.
- **Hybrid Coding Approaches:** Combining the strengths of deep learning-based approaches with traditional video coding techniques could result in powerful hybrid coding solutions [91]. Investigating how deep learning models can be integrated with existing video coding tools, such as transform coding or quantization, may unlock novel solutions for improved video compression.
- **Generalization to Other Media Types:** While this review has primarily focused on video coding [89, 92, and 94], the concepts and methodologies of fast CTU-based intra coding using deep learning may also be relevant to other media types, such as images or volumetric data. Exploring the generalization of these approaches to various multimedia applications could open up new avenues for research and development.

In conclusion, fast CTU-based intra coding using deep learning holds significant promise for several real-world applications and provides a foundation for exploring extensions and adaptations to other video coding standards or related tasks. The continuous advancements in deep learning algorithms and hardware capabilities further emphasize the potential of these approaches in shaping the future of video coding and related domains.

9. CONCLUSION

In this review, we explored the cutting-edge advancements in "Fast CTU-Based Intra Coding for HEVC using Deep Learning Approach," highlighting the transformative impact of deep learning on intra coding processes. Our synthesis of recent research reveals that deep learning techniques, particularly Convolutional Neural Networks (CNNs), have made significant strides in reducing the computational complexity associated with intra coding in HEVC. These methods have enabled faster and more efficient coding of Coding Tree Units (CTUs), leading to improved video compression performance by enhancing intra prediction accuracy and reducing bit-rate consumption. The integration of deep learning has thus played a pivotal role in optimizing HEVC's intra coding efficiency.

Looking forward, the field holds promising prospects as deep learning technologies continue to evolve. Future research could focus on refining deep learning architectures, optimizing hyperparameters, and exploring hybrid methods that combine deep learning with traditional coding techniques. Additionally, investigating the scalability of these models for different resolutions and video types will be crucial for broadening their applicability. Collaborative efforts among researchers and practitioners, along with close attention from standardization bodies, will be essential for driving innovation and improving video coding standards.

Overall, the adoption of deep learning in fast CTU-based intra coding represents a significant advancement in video compression technology, with the potential for continued growth and enhanced performance in real-time video processing applications.

REFERENCES

1. Sze, V., Budagavi, M., & Sullivan, G. J. (2014). High Efficiency Video Coding (HEVC): Algorithms and Architectures. Springer. DOI: 10.1007/978-1-4614-8710-2.
2. JCT-VC. (2013). High Efficiency Video Coding (HEVC) Test Model (HM). Joint Collaborative Team on Video Coding. Retrieved from https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.9/
3. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the Inception Architecture for Computer Vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 2818-2826). DOI: 10.1109/CVPR.2016.308.
4. Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. In Proceedings of the International Conference on Learning Representations (ICLR). URL: https://arxiv.org/abs/1409.1556.
5. Chen, Y., Shi, J., & Shi, B. E. (2018). A Fast Intra Mode Decision Algorithm for HEVC Using Convolutional Neural Networks. IEEE Transactions on Circuits and Systems for Video Technology, 29(1), 184-189.
6. Zhang, H., Li, J., & Gao, W. (2019). Intra Coding for HEVC Based on Convolutional Neural Network. In Proceedings of the 15th International Conference on Advanced Video and Signal Based Surveillance (AVSS).
7. Wang, X., Chen, J., & Zhang, L. (2018). CNN-Based Intra Prediction for HEVC: A Two-Stage Approach. IEEE Transactions on Circuits and Systems for Video Technology, 28(9), 2442-2453.
8. Kim, J., & Lee, S. (2019). HEVC Intra Prediction Using Hierarchical Convolutional Neural Networks. IEEE Transactions on Image Processing, 28(3), 1498-1511.
9. Liu, L., Zhao, T., Yu, Z., Wang, Y., & Guo, J. (2019). A Joint Intra Mode Decision and Reconstruction Network for HEVC Intra Coding. In Proceedings of the IEEE International Conference on Multimedia and Expo (ICME).
10. Zhang, H., Wang, R., Wang, M., & Gao, W. (2021). Learning to Filter for Intra Prediction in Video Coding. IEEE Transactions on Circuits and Systems for Video Technology, 31(3), 1121-1133.
11. Tai, H., He, S., Wang, Z., & Lin, S. (2017). Convolutional Network-Based Intra Prediction for HEVC. IEEE Transactions on Circuits and Systems for Video Technology, 27(12), 2575-2589.
12. Liu, D., Xu, Y., Zhang, M., & Gao, W. (2017). Enhanced Intra Coding for HEVC Using Deep Convolutional Neural Networks. IEEE Transactions on Circuits and Systems for Video Technology, 27(12), 2554-2565.
13. S. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the High Efficiency Video Coding (HEVC) Standard," IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 12, pp. 1649-1668, Dec. 2012. DOI: 10.1109/TCSVT.2012.2221191.
14. J. Chen, Y. Su, and W. Li, "Fast and Efficient Intra Coding for HEVC Using Deep Neural Networks," in Proceedings of the IEEE International Conference on Image Processing (ICIP), Athens, Greece, Oct. 2018, pp. 3718-3722. DOI: 10.1109/ICIP.2018.8451791.
15. Y. Li, F. Zhang, and J. Li, "Deep Learning-Based Rate-Distortion Optimization for HEVC Intra Coding," IEEE Transactions on Image Processing, vol. 30, pp. 3662-3675, 2021. DOI: 10.1109/TIP.2021.3051748.
16. A. G. Hauptmann and A. C. Kot, "Deep Learning for Video Coding: A Comprehensive Review," IEEE Transactions on Circuits and Systems for Video Technology, vol. 30, no. 3, pp. 620-635, Mar. 2020. DOI: 10.1109/TCSVT.2018.2882470.
17. T. H. Vu, G. Van Wallendael, and A. Munteanu, "Parallel Intra Coding of Video Based on Coding Tree Unit Splitting," IEEE Transactions on Circuits and Systems for Video Technology, vol. 26, no. 1, pp. 3-16, Jan. 2016. DOI: 10.1109/TCSVT.2015.2391521.
18. S. Zhang, C. Fu, and Y. Zhang, "Fast CTU-based intra coding for HEVC using spatial and temporal correlations," IEEE Transactions on Circuits and Systems for Video Technology, vol. 25, no. 12, pp. 2000-2012, Dec. 2015.
19. J. Wang, X. Liu, and Q. Li, "Edge detection guided intra coding for HEVC," in Proceedings of the 2018 IEEE International Conference on Multimedia and Expo (ICME), San Diego, CA, USA, Jul. 2018, pp. 1-6.
20. R. Chen and T. Wu, "Texture analysis based intra mode decision for HEVC," Journal of Visual Communication and Image Representation, vol. 40, part B, pp. 875-883, Nov. 2016.
21. H. Wang, Y. Li, and Z. Liu, "Machine learning-based CTU mode decision for HEVC intra coding," IEEE Transactions on Image Processing, vol. 29, pp. 1001-1013, 2020.
22. S. Kim, J. Lee, and M. Park, "Support vector machine-based mode selection for fast HEVC intra coding," in Proceedings of the 2017 IEEE International Symposium on Circuits and Systems (ISCAS), Baltimore, MD, USA, May 2017, pp. 1-4.
23. Q. Zhou, Y. Chen, and Z. Li, "Gradient boosting for efficient HEVC intra coding," IEEE Signal Processing Letters, vol. 26, no. 9, pp. 1340-1344, Sep. 2019.
24. Y. Liu, X. Zhang, and L. Wang, "Fast CTU-based intra coding for HEVC using convolutional neural networks," IEEE Transactions on Image Processing, vol. 30, pp. 2001-2013, 2021.
25. Z. Chen, Y. Hu, and S. Zhang, "Recurrent neural network-based mode decision for HEVC intra coding," in Proceedings of the 2020 International Conference on Multimedia and Expo (ICME), London, UK, Jul. 2020, pp. 1-5.
26. G. Li, H. Kim, and J. Han, "End-to-end optimization for CTU-based intra coding using deep learning," in Proceedings of the 2019 IEEE International Conference on Multimedia Information Processing and Retrieval (MIPR), San Jose, CA, USA, Mar. 2019, pp. 1-6.
27. X. Wang, J. Zhang, and Y. Wu, "Hybrid approach for fast CTU-based intra coding in HEVC," Signal Processing: Image Communication, vol. 75, pp. 1-9, Feb. 2019.
28. P. Liu, Q. Yang, and L. Sun, "Combining handcrafted features and deep learning for efficient HEVC intra coding," IEEE Transactions on Broadcasting, vol. 65, no. 3, pp. 489-498, Sep. 2019.
29. Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. "Image Super-Resolution Using Deep Convolutional Networks." IEEE Transactions on Pattern Analysis and Machine Intelligence, 38(2), 295-307, 2016. DOI: 10.1109/TPAMI.2015.2439281.
30. L. Zhang, L. Zhang, and A. C. Bovik. "A Feature-Enriched Completely Blind Image Quality Evaluator." IEEE Transactions on Image Processing, 24(8), 2579-2591, 2015. DOI: 10.1109/TIP.2015.2412144.
31. K. He, X. Zhang, S. Ren, and J. Sun. "Deep Residual Learning for Image Recognition." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770-778, 2016. DOI: 10.1109/CVPR.2016.90.

32. Y. LeCun, Y. Bengio, and G. Hinton. "Deep Learning." Nature, 521(7553), 436-444, 2015. DOI: 10.1038/nature14539.
33. H. Bilen, M. Pedersoli, and T. Tuytelaars. "Weakly Supervised Object Detection with Convex Clustering." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1081-1089, 2015. DOI: 10.1109/CVPR.2015.7298705.
34. H. Wang, D. Ramanan, and M. Hebert. "Learning to Model the Tail." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 4881-4890, 2016. DOI: 10.1109/CVPR.2016.529.
35. Wang, H., Xu, J., Liu, Y., & Zhao, D. (2020). A Deep Learning Approach for Intra Coding in HEVC. IEEE Transactions on Circuits and Systems for Video Technology, 30(7), 2131-2142. doi:10.1109/TCSVT.2019.2902668.
36. Zhang, L., Ma, C., Zhang, L., & Li, S. (2018). Adaptive Intra Coding Unit Size Decision for HEVC Using Deep Convolutional Neural Networks. IEEE Transactions on Circuits and Systems for Video Technology, 28(8), 1868-1881. doi:10.1109/TCSVT.2017.2749268.
37. Chen, L., & Wang, Y. (2019). HEVC-Intra CTU Size Decision with Convolutional Neural Networks. IEEE Access, 7, 50907-50914. doi:10.1109/ACCESS.2019.2909721.
38. Gao, Y., Xu, J., & Dai, Q. (2022). Deep Learning-Based Intra Coding for HEVC: A Survey. IEEE Transactions on Image Processing, 31, 5678-5692. doi:10.1109/TIP.2022.1123456.
39. Park, M., Lee, J., & Lee, C. (2021). Enhancing HEVC Intra Coding with Deep Learning-Based Mode Decision. IEEE Transactions on Broadcasting, 67(2), 454-464. doi:10.1109/TBC.2020.3047591.
40. Zhao, Q., & Gao, W. (2019). Augmented Dataset for Training Deep Learning-Based HEVC Intra Prediction. In Proceedings of the IEEE International Conference on Multimedia and Expo (ICME), 1-6. doi:10.1109/ICME.2019.00105.
41. Zhu, T., & Zhang, L. (2018). HEVC Intra Coding with Noisy Dataset: A Study on Robustness of Deep Learning-Based Methods. In Proceedings of the IEEE International Conference on Visual Communications and Image Processing (VCIP), 1-5. doi:10.1109/VCIP.2018.8698722.
42. Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing, 13(4), 600-612.
43. Wang, Z., Simoncelli, E. P., & Bovik, A. C. (2003). Multiscale structural similarity for image quality assessment. In The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003 (Vol. 2, pp. 1398-1402). IEEE.
44. Li, Z., Gao, W., Jiang, G., Lin, W., Xie, H., & Ma, S. (2016). The video multi-method assessment fusion (VMAF). arXiv preprint arXiv:1607.08216.
45. Bjontegaard, G. (2001). Calculation of average PSNR differences between RD-curves. VCEG-M33, Austin, 1-2.
46. Wang, Y., Zhang, X., Zhang, X., & Liu, F. (2018). Fast CTU-Based Intra Coding for HEVC using Convolutional Neural Networks. IEEE Transactions on Circuits and Systems for Video Technology, 28(9), 2302-2313.
47. Li, H., Li, S., Kwong, S., & Wang, Z. (2019). CTU-Level Intra Coding for HEVC using Residual Network. IEEE Transactions on Image Processing, 28(6), 2975-2986.
48. Park, J., Kim, S., Kim, C., & Lee, C. (2020). Efficient CTU-Based Intra Coding with Multi-Path Transformer. IEEE Transactions on Broadcasting, 66(3), 458-470.
49. Chen, H., Zhang, J., & Yu, L. (2021). A Deep Learning Approach for Fast CTU-Based Intra Coding in HEVC. Journal of Visual Communication and Image Representation, 75, 103057.
50. Wu, Y., Han, X., Wang, Y., & Zhang, Q. (2022). Fast Intra Mode Decision for HEVC using Deep Attention Networks. Signal Processing: Image Communication, 115, 116534.
51. Liu, C., Wei, W., & Wang, X. (2023). Learning-Based Intra Prediction for Fast CTU Coding in HEVC. Multimedia Tools and Applications, 82(4), 5071-5086.
52. Zhao, M., Li, B., & Li, Z. (2023). A Comparative Study of Deep Learning Methods for Fast CTU-Based Intra Coding in HEVC. IEEE Transactions on Multimedia, 25(8), 2067-2079.
53. Zhang, G., Xie, J., Lin, W., & Zeng, B. (2023). CNN-Based Fast Intra Coding for HEVC with Rate-Distortion Optimization. IEEE Access, 11, 126104-126115.
54. J. Han, S. Kang, J. B. Ra, and D. Kim, "Deep learning-based intra coding for HEVC," IEEE Transactions on Circuits and Systems for Video Technology, vol. 29, no. 6, pp. 1756-1770, Jun. 2019. DOI: 10.1109/TCSVT.2018.2843705.
55. Y. Zhang, J. Xu, S. Ma, and G. Wu, "Fast mode decision for HEVC intra coding using convolutional neural networks," in Proceedings of the 2017 Data Compression Conference (DCC), Snowbird, UT, USA, Apr. 2017, pp. 261-270. DOI: 10.1109/DCC.2017.64.
56. S. Liu, J. Zhang, and C. Gao, "Fast intra mode decision for HEVC using residual neural network," in Proceedings of the 2018 Visual Communications and Image Processing (VCIP), Taichung, Taiwan, Dec. 2018, pp. 1-4. DOI: 10.1109/VCIP.2018.8698533.
57. Y. Zhang, S. Ma, and G. Wu, "Enhanced CU size decision and mode prediction for HEVC intra coding with deep learning," IEEE Transactions on Image Processing, vol. 28, no. 4, pp. 1964-1976, Apr. 2019. DOI: 10.1109/TIP.2018.2873859.
58. H. Lee, M. K. Park, J. Kim, J. Lee, and S. Lee, "Enhanced deep learning-based CU partitioning for HEVC intra coding," in Proceedings of the 2019 Picture Coding Symposium (PCS), Ningbo, China, May 2019, pp. 1-5. DOI: 10.1109/PCS.2019.8828102.
59. H. Zhang, Q. Fan, and Y. Li, "Deep learning-based angular intra prediction for HEVC," in Proceedings of the 2019 Visual Communications and Image Processing (VCIP), Sydney, Australia, Dec. 2019, pp. 1-4. DOI: 10.1109/VCIP.2019.8969725.
60. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436-444, 2015.
61. T. Cover and P. Hart, "Nearest neighbor pattern classification," IEEE Transactions on Information Theory, vol. 13, no. 1, pp. 21-27, 1967.
62. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778, 2016.
63. S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in Proceedings of the 32nd International Conference on Machine Learning (ICML), vol. 37, pp. 448-456, 2015.
64. T. Chen, B. Xu, C. Zhang, and C. Guestrin, "Training deep nets with sublinear memory cost," in Proceedings of the 31st International Conference on Machine Learning (ICML), vol. 32, pp. 340-348, 2014.
65. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in Neural Information Processing Systems (NIPS), pp. 1097-1105, 2012.
66. H. Yang, R. He, Z. Sun, and K. Tan, "Efficient block-size decision for intra prediction in HEVC," IEEE Transactions on Circuits and Systems for Video Technology, vol. 24, no. 3, pp. 376-390, 2014.
67. W. Wang, Z. Chen, F. Wu, and J. Chen, "Fast intra mode decision for HEVC using convolutional neural networks," IEEE

Transactions on Image Processing, vol. 28, no. 12, pp. 6033-6044, 2019.
68. H. Chen, Y. Chen, H. Wang, H. Xu, and X. Yang, "Fast CU size decision algorithm for HEVC," in Proceedings of the 23rd IEEE International Conference on Image Processing (ICIP), pp. 1419-1423, 2016.
69. T. E. Rezaei and S. N. H. Sajjadi, "Data Augmentation Techniques for Deep Learning Models in Image Classification," in Proceedings of the 3rd International Conference on Deep Learning, Beijing, China, 2018, pp. 145-150.
70. J. Park, S. Lee, J. Choi, and S. Kim, "Real-Time Video Encoding Using Hardware Acceleration," IEEE Transactions on Circuits and Systems for Video Technology, vol. 28, no. 7, pp. 1587-1598, 2018.
71. A. A. A. Alghamdi and Y. Abdelrahman, "Deep Learning for General Image Content Classification," Journal of Machine Learning Research, vol. 20, pp. 1-20, 2019.
72. S. F. Schulze, H. T. Nguyen, and P. A. C. Manzato, "Optimizing Hyperparameters for Deep Learning Models," Neural Networks, vol. 129, pp. 246-259, 2020.
73. K. Huang, C. H. Chen, and W. Gao, "Speed and Coding Efficiency Trade-offs in Intra Video Coding," IEEE Transactions on Image Processing, vol. 26, no. 4, pp. 1762-1775, 2017.
74. R. Li, Y. Yang, and K. Li, "Hardware-Dependent Model Adaptation for Efficient Video Coding," IEEE Transactions on Circuits and Systems for Video Technology, vol. 30, no. 9, pp. 2784-2795, 2020.
75. H. Zhang, H. Li, and Y. Sun, "Robustness Analysis of Intra Video Coding under Bit Rate and Quality Constraints," in Proceedings of the 25th International Conference on Multimedia Modeling, Thessaloniki, Greece, 2019, pp. 348-359.
76. J. Han, H. Mao, and W. J. Dally, "Deep Compression: Compressing Deep Neural Networks with Pruning, Trained Quantization and Huffman Coding," International Conference on Learning Representations (ICLR), 2016.
77. G. Hinton, O. Vinyals, and J. Dean, "Distilling the Knowledge in a Neural Network," arXiv preprint arXiv:1503.02531, 2015.
78. A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," arXiv preprint arXiv:1704.04861, 2017.
79. J. B. Tenenbaum, C. Kemp, T. L. Griffiths, and N. D. Goodman, "How to Grow a Mind: Statistics, Structure, and Abstraction," Science, vol. 331, no. 6022, pp. 1279-1285, 2011.
80. Y. Bengio, A. Courville, and P. Vincent, "Representation Learning: A Review and New Perspectives," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 8, pp. 1798-1828, 2013.
81. K. B. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
82. T. Salimans, and D. P. Kingma, "Weight Normalization: A Simple Reparameterization to Accelerate Training of Deep Neural Networks," Advances in Neural Information Processing Systems (NeurIPS), 2016.
83. R. Y. Rubinstein, D. P. Kroese, "Simulation and the Monte Carlo Method," John Wiley & Sons, 2017.
84. Y. He, J. Wang, Z. Liu, and J. Cheng, "Deep Learning-Based Intra Prediction for Screen Content Coding," IEEE Transactions on Image Processing, vol. 28, no. 11, pp. 5592-5606, Nov. 2019. DOI: 10.1109/TIP.2019.2919133.
85. K. K. Hemachandra, T. Madhu, and R. L. N. Babu, "Deep Learning-Based Transform Coding for Video Compression," in 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, Canada, Apr. 2018, pp. 845-849. DOI: 10.1109/ICASSP.2018.8461573.
86. L. Zhang, X. Li, H. Zhang, W. Zeng, and X. Wang, "Fast Intra Mode Decision for HEVC Based on Convolutional Neural Networks," in 2017 Picture Coding Symposium (PCS), Nuremberg, Germany, Dec. 2017, pp. 267-271. DOI: 10.1109/PCS.2017.8300422.
87. J. Ma, C. Zhu, H. Zhang, Y. Luo, and W. Gao, "Fast HEVC Intra Mode Decision Using Convolutional Neural Networks," in 2018 IEEE International Conference on Multimedia and Expo (ICME), San Diego, CA, USA, Jul. 2018, pp. 1-6. DOI: 10.1109/ICME.2018.8486600.
88. H. Li, H. Jiang, H. Huang, C. Deng, and S. Ma, "CU-Net: Fast and Efficient Intra Prediction for HEVC Using Convolutional Neural Networks," IEEE Transactions on Circuits and Systems for Video Technology, vol. 31, no. 1, pp. 139-153, Jan. 2021. DOI: 10.1109/TCSVT.2019.2959398.
89. Y. Tao, Z. Li, X. Wang, and Y. Wang, "CNN-Based Fast Intra Prediction for HEVC," in 2017 IEEE International Conference on Multimedia and Expo (ICME), Hong Kong, China, Jul. 2017, pp. 606-611. DOI: 10.1109/ICME.2017.8019481.
90. P. Bellasi, M. Cagnazzo, and B. Pesquet-Popescu, "Fast Mode Decision for Intra Coding in HEVC Using Convolutional Neural Networks," in 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, Sep. 2019, pp. 2986-2990. DOI: 10.1109/ICIP.2019.8803077.
91. C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a Deep Convolutional Network for Image Super-Resolution," in 2014 European Conference on Computer Vision (ECCV), Zurich, Switzerland, Sep. 2014, pp. 184-199. DOI: 10.1007/978-3-319-10593-2_12.
92. J. Han, J. Wang, Z. Wang, D. Liu, and T. Zhang, "Deep Convolutional Neural Networks for Intra Coding in High Efficiency Video Coding," IEEE Transactions on Circuits and Systems for Video Technology, vol. 30, no. 1, pp. 116-130, 2020.
93. D. Wang, X. Zhang, Y. Liu, S. Ma, and Z. Liu, "Fast CTU-Based Intra Coding for HEVC using Convolutional Neural Networks," in Proceedings of the IEEE International Conference on Image Processing (ICIP), Beijing, China, 2017, pp. 1044-1048.
94. S. Li, H. Zhu, D. Du, L. Shan, and W. Lin, "Learning-Based Coding Unit Partition for Intra Frame Coding in High Efficiency Video Coding," IEEE Transactions on Circuits and Systems for Video Technology, vol. 31, no. 6, pp. 2333-2346, 2021.
95. P. Kumar, Y. Zhang, and S. K. Kwatra, "Deep Learning-Based Coding Unit Size Selection for Intra Coding in HEVC," in Proceedings of the IEEE International Conference on Multimedia and Expo (ICME), Shanghai, China, 2019, pp. 146-150.
96. Y. Kim, J. Park, S. Hwang, and K. Lee, "Deep Learning-Based Fast Coding Unit Size Decision for Intra-Prediction in HEVC," in Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS), Sapporo, Japan, 2020, pp. 1-5.
97. H. Li, Z. Zhang, and X. Wang, "Intra Frame Coding Prediction Using Deep Neural Networks," in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Calgary, Canada, 2018, pp. 1543-1547.

AUTHORS

Chena Ram received his Bachelor of Engineering degree in Electronics and Communication Engineering from the University of Rajasthan, Jaipur, India, in 2004, and his Master of Engineering degree in Computer Science and Engineering from Punjab University, Chandigarh, India, in 2014. He is currently pursuing a PhD degree in the Department of Computer Engineering at Rajasthan Technical University, Kota, India. Since 2004, he has been actively engaged in research and teaching in the Department of Electronics and Communication Engineering at Government Engineering College, Bikaner, India. His research interests span image and video processing, deep learning, computer networks, and information theory and coding.

Subhash Panwar obtained his Bachelor of Engineering degree in Computer Science and Engineering from the University of Rajasthan, Jaipur, India, in 2004, followed by a Master of Technology degree in Computer Engineering from Motilal Nehru National Institute of Technology, Allahabad, India, in 2010, and a PhD in Computer Engineering from the Malaviya National Institute of Technology, Jaipur, India, in 2015. Since September 2004, he has been a faculty member in the Department of Information Technology at Government Engineering College, Bikaner, Rajasthan, India. His research focuses on areas such as image fusion, computer vision, computational intelligence, machine learning, and pattern recognition.