

Codecs de Vídeo Tradicionais vs. Neurais: Uma Análise de Eficiência e Qualidade utilizando PSNR e outras métricas

Leandro Tavares¹, Ruhan Conceição¹, Victor Costa¹, Luciano Agostini¹,
Marcelo Porto¹, Guilherme Corrêa¹

¹VideoTechnology Group – Universidade Federal de Pelotas (UFPel)
Pelotas – RS – Brazil

{lwtavares, radconceicao, vcrcosta, agostini, porto, gcorrea}@inf.upel.edu.br

Abstract. *The rise of Neural Video Codecs (NVCs) represents a paradigm shift from traditional codecs, such as VVC. This paper provides a comparative analysis by comparing the traditional codecs HEVC and VVC against the NVCs DCVC-FM and DCVC-RT. We evaluate compression efficiency using BD-Quality with multiple metrics, including PSNR, SSIM, VMAF, and LPIPS. The results demonstrate the superiority of NVCs, with DCVC-FM achieving gains of over 10 VMAF points compared to HEVC. This significant advantage, especially on perceptually-aligned metrics, highlights the limitations of relying solely on PSNR for modern codec evaluation and reinforces the maturity of NVCs.*

Resumo. *A ascensão dos Codecs de Vídeo Neurais (NVCs) representa uma mudança de paradigma em relação aos codecs tradicionais, como o VVC. Este trabalho realiza uma análise comparativa, comparando os codecs tradicionais HEVC e VVC contra os NVCs DCVC-FM e DCVC-RT. Avaliamos a eficiência de compressão com BD-Quality com múltiplas métricas, incluindo PSNR, SSIM, VMAF e LPIPS. Os resultados demonstram a superioridade dos NVCs, com o DCVC-FM alcançando ganhos de mais de 10 pontos em VMAF sobre o HEVC. Essa vantagem expressiva, especialmente em métricas alinhadas à percepção, evidencia as limitações do uso exclusivo do PSNR para a avaliação de codecs modernos e reforça a maturidade dos NVCs.*

1. Introdução

A crescente demanda por conteúdo de vídeo digital impulsiona a evolução contínua dos padrões de compressão. Codecs tradicionais, como o H.266/VVC [VVC 2020], representam o estado-da-arte em arquiteturas híbridas baseadas em blocos. Em paralelo, os avanços em *deep learning* deram origem a uma nova classe de Codificadores de Vídeo Neurais (NVCs), que substituem módulos tradicionais por redes treináveis de ponta a ponta. Abordagens de estado-da-arte como DCVC-FM [Li et al. 2024] e DCVC-RT [Jia et al. 2025] já demonstram um desempenho competitivo, focando, respectivamente, na máxima eficiência de compressão e na decodificação em tempo real.

Apesar dos NVCs frequentemente superarem os codecs tradicionais em métricas objetivas como o PSNR, as duas abordagens produzem artefatos de compressão com características distintas, que são percebidas de forma diferente pelo Sistema Visual Humano (*Human Visual System* – HVS). Isso levanta questões sobre a justiça e a abrangência das comparações baseadas unicamente em métricas tradicionais.

Neste contexto, este trabalho realiza uma análise comparativa entre codecs de vídeo tradicionais (HEVC, VVC) e neurais (DCVC-FM, DCVC-RT). A avaliação de desempenho inclui a métrica padrão PSNR, bem como métricas projetadas para melhor refletir a percepção humana, como SSIM, VMAF e LPIPS.

2. Configuração Experimental

Aqui detalhamos o planejamento experimental usado para comparar as abordagens de compressão de vídeo tradicionais e neurais.

2.1. Codecs

Para avaliar o desempenho dos *frameworks* de compressão de vídeo neural em relação às abordagens tradicionais, selecionamos dois padrões representativos: o *High Efficiency Video Coding* (HEVC) [HEVC 2013] e o *Versatile Video Coding* (VVC) [VVC 2020]. O HEVC é amplamente utilizado em eletrônicos, transmissão e plataformas de streaming, enquanto o VVC representa o atual estado-da-arte em codificação de vídeo tradicional. Para os NVCs, selecionamos duas implementações recentes e avançadas da família DCVC: DCVC-FM [Li et al. 2024], escolhido por sua notável eficiência de compressão, e o DCVC-RT [Jia et al. 2025], selecionado por suas capacidades de tempo real.

2.2. Conjuntos de Dados e Configuração dos Codificadores

Selecionamos um conjunto de dados de teste diversificado, composto por sequências do *dataset Ultra Video Group* (UVG) [Mercat et al. 2020] e das classes de teste padrão HEVC *Common Test Conditions* [Bossen et al. 2013] (Classes B, C, D e E). O conjunto UVG inclui 7 sequências de alta resolução em 1920×1080 e altas taxas de quadros, enquanto os conjuntos HEVC, compostos por um total de 16 vídeos, cobrem uma gama mais ampla de resoluções. Para consistência e viabilidade computacional, usamos apenas os primeiros 96 quadros de cada sequência em todos os experimentos, e todas as sequências foram codificadas usando um período intra de -1 . Para os NVCs, testamos 10 configurações de qualidade diferentes (0, 7, 14, 21, 28, 35, 42, 49, 56 e 63), enquanto para os codecs tradicionais, utilizamos seis parâmetros de quantização (*quantization parameters* – QPs) padrão: 22, 27, 32, 37, 42 e 47. Os codecs tradicionais foram configurados usando o perfil *Low Delay* para corresponder à estrutura de codificação dos modelos neurais. Adicionalmente, todos os vídeos foram codificados no formato YUV.

2.3. Métricas de qualidade

Para avaliar o desempenho dos codecs, empregamos uma combinação de métricas de qualidade tradicionais e alinhadas à percepção. Estas incluem o *Peak Signal-to-Noise Ratio* (PSNR), *Structural Similarity Index* (SSIM), *Learned Perceptual Image Patch Similarity* (LPIPS) e *Video Multi-method Assessment Fusion* (VMAF), cada uma capturando diferentes aspectos da fidelidade visual e da qualidade perceptual.

3. Eficiência de Compressão

Esta seção avalia a eficiência de compressão dos codecs de vídeo selecionados. O desempenho é mensurado utilizando a métrica *Bjontegaard Delta Rate* (BD-Rate) [Bjontegaard 2001]. Essa métrica pode tanto medir a economia de bits de um codec

**Tabela 1. Média de BD-Quality no dataset (dB para PSNR, score para as outras).
(^t) denota tradicionais e (ⁿ) neurais.**

	HEVC ^t (Âncora)	VVC ^t	FM ⁿ	RT ⁿ
BD-PSNR (dB)	-	0,9317	1,8625	<u>1,5013</u>
BD-SSIM	-	0,0130	0,0224	<u>0,0221</u>
BD-LPIPS*	-	0,0159	0,0327	<u>0,0319</u>
BD-VMAF	-	4,8911	10,3056	<u>9,2936</u>

em relação a outro mantendo a mesma qualidade (BD-Rate), quanto o ganho de qualidade para uma determinada taxa de bits (BD-Quality). Neste trabalho, o foco é a segunda abordagem. Portanto, avaliamos o desempenho dos codecs analisando os ganhos de qualidade para uma taxa de bits equivalente, medidos através da BD-Quality (ex: BD-PSNR).

Para realizar as avaliações, escolhemos um codificador como âncora para as comparações e, como o HEVC é o codec mais antigo e mais estabelecido na indústria entre os avaliados, ele foi selecionado como âncora para todas as comparações.

Os resultados da avaliação de eficiência de compressão estão consolidados na Tabela 1, que apresenta a média dos valores de cada métrica em relação ao codec âncora. Esses resultados destacam uma tendência consistente: os NVCs superaram significativamente os codecs tradicionais em termos de ganho de qualidade. Para a mesma taxa de bits que o HEVC, o DCVC-FM proporciona um aumento substancial de qualidade de 10,3 pontos VMAF e 1,86 dB em PSNR. Esses ganhos são quase o dobro dos fornecidos pelo VVC. Esta tendência é consistente em todas as métricas, com os NVCs entregando consistentemente uma qualidade de imagem significativamente maior do que os codecs tradicionais para a mesma taxa de bits. A Figura 1 apresenta uma comparação entre cortes de um quadro decodificado por cada codec, onde é possível perceber a qualidade de reconstrução superior dos NVCs indicada pelas métricas. Também é possível observar artefatos de bloco introduzidos pelos codecs tradicionais, fato que não ocorre nos NVCs.



Figura 1. Comparação entre quadros decodificados, processados por cada codec à uma taxa de bits comparável.

4. Considerações Finais

Neste trabalho, conduzimos uma análise comparativa abrangente entre codecs de vídeo tradicionais (HEVC e VVC) e codecs de vídeo neurais de estado-da-arte (DCVC-FM e DCVC-RT). Nossos resultados quantitativos demonstram a superioridade do paradigma de compressão neural. Do ponto de vista de taxa-distorção, os NVCs oferecem melhoras substanciais de qualidade para a mesma taxa. Essa vantagem foi mais pronunciada quando medida com métricas modernas como VMAF e LPIPS, onde o DCVC-FM alcançou ganhos de qualidade de mais de 10 pontos VMAF e 0,0327 pontos em LPIPS em comparação ao HEVC. Isso indica que os benefícios dos NVCs se estendem além de uma simples fidelidade de sinal e se traduzem em uma melhor experiência perceptual.

Estes resultados não apenas validam os estudos em direção a codecs neurais, mas também expõem as limitações de depender apenas do PSNR para a avaliação de codecs. Para trabalhos futuros, planejamos estender esta análise explorando a eficácia de técnicas de melhoria da qualidade visual como filtros de pós-processamento para mitigar artefatos específicos dos NVCs.

Referências

- (2020). Recommendation H.266: Versatile video coding. International Telecommunication Union (ITU). Available: <https://www.itu.int/rec/T-REC-H.266>.
- Bjontegaard, G. (2001). Calculation of average psnr differences between rd-curves. *ITU SG16 Doc. VCEG-M33*.
- Bossen, F., Li, B., Lim, K., and Norkin, A. (2013). Common test conditions and software reference configurations. Technical Report JCTVC-L1100, Joint Collaborative Team on Video Coding (JCT-VC). Meeting 12, Geneva, Switzerland.
- HEVC (2013). Recommendation H.265: High efficiency video coding. International Telecommunication Union (ITU). Available: <https://www.itu.int/rec/T-REC-H.265>.
- Jia, Z., Li, B., Li, J., Xie, W., Qi, L., Li, H., and Lu, Y. (2025). Towards practical real-time neural video compression. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2025, Nashville, TN, USA, June 11-25, 2024*.
- Li, J., Li, B., and Lu, Y. (2024). Neural video compression with feature modulation. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2024, Seattle, WA, USA, June 17-21, 2024*.
- Mercat, A., Vaittinen, M., and Vanne, J. (2020). Uvlg dataset: 50/120fps 4k sequences for video codec analysis and development. In *Proceedings of the 11th ACM Multimedia Systems Conference (MMSys)*, pages 297–302.