

APLICAÇÃO DE BAYESIAN SEARCH NA BUSCA POR HIPERPARÂMETROS EM REDE PRUNADA DE APRIMORAMENTO DE VÍDEOS DECODIFICADOS

LUIS CARLOS RODRIGUES LINARES¹; GILBERTO KREISLER¹; BRUNO ZATT¹; DANIEL PALOMINO¹; GUILHERME CORRÊA¹

¹Universidade Federal de Pelotas – {lcrlinares, gkfneto, zatt, dpalomino, gcorrea}@inf.ufpel.edu.br

1. INTRODUÇÃO

A crescente demanda por conteúdo de vídeo em alta definição, impulsionada por plataformas de *streaming*, videoconferências e redes sociais, consolidou o vídeo como o principal componente do tráfego de dados na internet. Para viabilizar a transmissão desse volume massivo de informações sob larguras de banda limitadas, a aplicação de algoritmos de compressão, como o H.265/HEVC, é necessário. No entanto, processos de compressão introduzem artefatos visuais indesejados como efeito de blocos, borramento e bordas fantasma, que degradam a qualidade da experiência do espectador e podem prejudicar o desempenho de tarefas subsequentes de visão computacional.

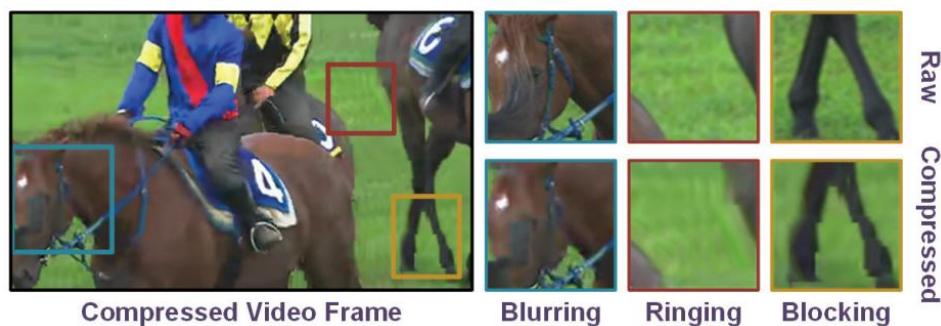


Figura 1: Exemplos de artefatos de compressão (DENG, 2020)

Para mitigar esses efeitos, as técnicas de Aprimoramento da Qualidade de Vídeo (VQE) têm ganhado destaque. Metodologias baseadas em Redes Neurais Convolucionais (CNNs) demonstram sucesso, aprendendo a restaurar a qualidade visual a partir de grandes volumes de dados. Entre elas, a rede STDF (*Spatio-Temporal Deformable Fusion*) (DENG, 2020) se destaca por sua eficácia em explorar a redundância temporal entre quadros de um vídeo para aprimorar a qualidade de quadros degradados.

Contudo, a complexidade computacional e o elevado número de parâmetros desses modelos representam um obstáculo para sua implementação em dispositivos com recursos de hardware limitados, como celulares e sistemas embarcados.

Uma solução para este problema é a aplicação de técnicas de poda de redes neurais (*pruning*), que visam reduzir a complexidade do modelo sem comprometer sua performance de maneira relevante. A técnica *Filter in Filter*, ou Poda por Faixas (*Stripe-Wise Pruning*) (MENG, 2020), oferece uma granularidade mais fina que os métodos tradicionais de poda de filtros. Em vez de remover filtros convolucionais inteiros, ela permite a remoção de "faixas" (componentes 1x1) dentro de cada filtro, preservando a estrutura computacional da rede e alcançando taxas de compressão mais elevadas.

Após a aplicação da poda *Filter in Filter* na rede STDF, possuímos a necessidade de otimizar os hiperparâmetros que comandam o processo de poda, a taxa de esparsidade (*Sparsity Ratio - SR*) e o limiar de poda (*Threshold*). A seleção manual desses valores é um processo trabalhoso e que consome tempo, sendo improvável alcançar uma configuração ótima que equilibre perfeitamente a redução de parâmetros e a manutenção da qualidade de vídeo.

Diante disso, este trabalho propõe a aplicação da técnica de Busca Bayesiana (*Bayesian Search*) para a otimização automática dos hiperparâmetros de poda da rede STDF. Diferente de metodologias como *Grid Search* ou *Random Search*, que exploram o espaço de hiperparâmetros de forma "cega", a Busca Bayesiana é um método de otimização sequencial que constrói um modelo probabilístico do mapeamento entre os valores dos hiperparâmetros e a métrica de objetivo (SNOEK, 2012).

A cada iteração, o algoritmo utiliza os resultados das combinações já testadas para refinar sua "confiança" sobre quais áreas do espaço de busca são mais promissoras. Isso permite que a busca se concentre em regiões com maior potencial de encontrar a solução ótima, equilibrando a exploração de novas combinações de hiperparâmetros com a exploração de combinações que já demonstraram bons resultados. Isso torna o processo mais eficiente, reduzindo o número de avaliações necessárias para convergir para uma solução de alta qualidade.

O objetivo é encontrar a combinação ideal que maximize a relação entre a qualidade do vídeo aprimorado, medida pela métrica PSNR, e a compressão do modelo. Ao final, espera-se obter uma arquitetura de rede leve, rápida e otimizada de forma robusta e reproduzível para aplicações práticas, focado em ambientes com recursos computacionais limitados.

2. METODOLOGIA

A partir dos procedimentos técnicos e as ferramentas empregadas para realizar a otimização dos hiperparâmetros de poda da rede STDF, espera-se garantir a reproduzibilidade dos experimentos e dos resultados.

Todos os experimentos de treinamento e avaliação foram conduzidos em um servidor equipado com uma GPU NVIDIA RTX A6000 com 48 GB de memória. O modelo de rede neural base é o MFVQE com *Spatio-Temporal Deformable Fusion* (STDF), sobre o qual foi aplicada a técnica de poda *Filter in Filter*.

A poda é controlada dinamicamente por dois hiperparâmetros: a taxa de esparsidade (SR) e o limiar de magnitude (THRESHOLD). O desafio do trabalho consiste em encontrar a combinação ótima desses dois valores. Para automatizar essa busca, foi empregada a Busca Bayesiana por meio da ferramenta W&B Sweeps (WEIGHTS & BIASES, 2025).

Como a simples maximização da qualidade de vídeo (PSNR) não considera o objetivo de compressão, foi desenvolvida uma métrica de objetivo composta. Esta métrica estabelece um limiar mínimo de qualidade (um ganho de 0.540 dB no Delta PSNR) e, caso o modelo atinja ou supere esse limiar, sua pontuação é definida como o negativo do número total de parâmetros. Dessa forma, o otimizador bayesiano busca maximizar essa pontuação, o que na prática significa minimizar o tamanho do modelo, desde que a restrição de qualidade seja atendida.

O fluxo de trabalho para cada execução do *sweep* consistiu em: 1) O agente do W&B sugerir uma nova combinação de *SR* e *THRESHOLD*; 2) O modelo ser treinado do início ao fim aplicando a poda; 3) Ao final, um script de avaliação

calcular o número de parâmetros e o Delta PSNR do modelo final; 4) A métrica objetivo ser calculada e registrada na plataforma, retroalimentando o otimizador bayesiano para a próxima iteração.

Este ciclo se repete, explorando o espaço de busca até convergir para uma solução que representa o melhor compromisso entre compressão e qualidade.

3. RESULTADOS E DISCUSSÃO

A aplicação da metodologia de otimização de hiperparâmetros por meio da Busca Bayesiana demonstrou ser altamente eficaz, gerando resultados que validam a abordagem proposta. Até o momento, o processo de busca automatizada convergiu para uma configuração de hiperparâmetros (SR e THRESHOLD) que produziu um modelo de rede STDF mais compacto e eficiente.

O melhor modelo obtido até o momento possui aproximadamente 160 mil parâmetros, uma grande redução quando comparado ao modelo de referência, que continha 367 mil parâmetros. Isso representa uma compressão adicional de 56,4% no número de parâmetros, evidenciando a existência de uma redundância na rede do STDF.

O aspecto mais relevante deste resultado é que essa redução de complexidade foi alcançada sem qualquer sacrifício na performance de aprimoramento de vídeo. O modelo otimizado manteve um Delta PSNR de 0,545 dB, um valor idêntico ao apresentado pelo modelo de referência de 367 mil parâmetros. Este achado comprova que é possível mais do que dobrar a eficiência do modelo (em termos de tamanho) sem incorrer em perdas de qualidade.

Tal resultado seria extremamente difícil de ser alcançado por meio de uma sintonia manual de hiperparâmetros, o que reforça o poder da Busca Bayesiana guiada por uma métrica de objetivo bem definida.

A capacidade de reduzir a complexidade da rede em mais de 50%, mantendo a qualidade intacta, tem implicações diretas para a viabilidade prática da tecnologia, abrindo caminho para sua implementação em dispositivos com recursos computacionais limitados, onde o tamanho do modelo e a velocidade de inferência são fatores influentes.

| Modelo | Tipo | SR | Threshold | Parâmetros | ΔPSNR |
|----------|------------|-------------|-------------|------------|-------|
| Modelo 1 | Referência | - | - | 367.000 | 0,545 |
| Modelo 1 | Podado | 0,000028662 | 0,020684347 | 161.450 | 0,545 |
| Modelo 2 | Referência | - | - | 367.000 | 0,549 |
| Modelo 2 | Podado | 0,000345425 | 0,058167058 | 161.450 | 0,549 |
| Modelo 3 | Referência | - | - | 367.000 | 0,548 |
| Modelo 3 | Podado | 0,000320981 | 0,014769255 | 161.450 | 0,548 |
| Modelo 4 | Referência | - | - | 367.000 | 0,529 |
| Modelo 4 | Podado | 0,000286786 | 0,056582029 | 161.450 | 0,529 |

Tabela 1: Resultados dos melhores modelos até o presente momento

4. CONCLUSÕES

Este trabalho demonstrou com sucesso a viabilidade e a eficácia da aplicação de uma metodologia de otimização automatizada, a Busca Bayesiana, para a sintonia fina de hiperparâmetros em redes neurais de aprimoramento de vídeo submetidas à poda. A principal inovação reside na sistematização de um processo que, tradicionalmente, depende de ajustes manuais e heurísticos, estabelecendo um método robusto para encontrar o equilíbrio ótimo entre a compressão do modelo e a manutenção de sua performance.

Ao definir uma métrica de objetivo que prioriza a minimização de parâmetros sob uma restrição de qualidade, foi possível extrair um potencial de compressão maior da técnica *Filter in Filter* aplicada à rede STDF, superando os resultados de uma poda com hiperparâmetros fixos.

Para trabalhos futuros, recomenda-se a expansão da análise de performance para além das métricas de qualidade e complexidade. A avaliação do consumo de energia e do tempo de processamento, medido em quadros por segundo, é importante para validar a viabilidade do modelo otimizado em um cenário de aplicação real, principalmente em dispositivos com restrições de hardware e bateria.

É importante salientar que os experimentos realizados se concentraram exclusivamente na poda do módulo STDF. Como o módulo de Aprimoramento de Qualidade (QE) não foi submetido ao mesmo processo, ainda existe um espaço para obter resultados ainda mais expressivos, aplicando a mesma metodologia de poda e otimização a essa segunda parte da arquitetura, visando uma compressão ainda maior do modelo como um todo.

5. REFERÊNCIAS BIBLIOGRÁFICAS

DENG, J. et al. Spatio-Temporal Deformable Convolution for Compressed Video Quality Enhancement. In: AAAI CONFERENCE ON ARTIFICIAL INTELLIGENCE, 34., 2020. p. 10696-10703.

MENG, F. et al. Pruning Filter in Filter. In: CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS, 34., 2020, Vancouver. **Proceedings...** Vancouver: NeurIPS, 2020.

SNOEK, J.; LAROCHELLE, H.; ADAMS, R. P. Practical Bayesian Optimization of Machine Learning Algorithms. In: CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS, 25., 2012, Lake Tahoe. **Proceedings...** Lake Tahoe: NeurIPS, 2012. p. 2951-2959.

WEIGHTS & BIASES. **W&B Sweeps.** 2025. Disponível em: <https://docs.wandb.ai/guides/sweeps>. Acesso em: 23 ago. 2025