

Detecção de Mídias Pornográficas em Dispositivos com Recursos Limitados para Controle Parental

Jhonatan Geremias, Eduardo Kugler Viegas, Altair Olivo Santin, Jackson Mallmann

*Programa de Pós-Graduação em Informática – Pontifícia Universidade Católica do Paraná, Brasil

{jgeremias, eduardo.viegas, santin}@ppgia.pucpr.br

**Instituto Federal Catarinense (IFC), Brusque – Santa Catarina, Brasil

jackson.mallmann@ifc.edu.br

SBSeg 2024

Simpósio Brasileiro em Segurança da Informação (SBSeg)



Agenda

- **Introdução**
- Fundamentação Teórica
- Trabalhos Relacionados
- Proposta
- Avaliação
- Conclusão

Introdução

Contextualização

- Crescente utilização dos dispositivos móveis
 - Inclusão das crianças e adolescentes ao meio digital
 - Tablets e smartphones
 - A exposição das crianças na internet
- Disseminação de conteúdo pornográfico na Internet
 - Fácil acesso, amplamente disponível na rede
 - Grande parte é composto por vídeo sem devida classificação
 - Menores de idade estão suscetíveis a esse tipo de conteúdo
- Monitorar/controlar crianças e adolescentes para que não acessem conteúdos impróprios
 - Classificação do teor do conteúdo do vídeo por idade

Introdução

Contextualização

- A literatura dispõem diversas técnicas para a classificação de vídeos.
 - CNNs se destacam por fornecer alta taxas de acurácia na classificação de vídeos.
 - A utilização de CNNs para classificação de vídeo é complexa e exige recursos computacionais
 - Dificuldade para detecção em tempo real
 - Assumem que o vídeo inteiro está disponível
- Dispositivos com recursos computacionais limitados
 - Inviabilidade de implementação em hardware com recursos limitados

Introdução

Objetivo

- Proposta de uma abordagem baseada em CNN para detecção de pornografia em vídeos em tempo real
 - Abordagem direcionada a dispositivos com recursos computacionais limitados
 - Extração de Características de Movimento
 - Informações de movimento que são obtidas entre dois frames adjacentes
 - Técnicas de fluxo óptico e mapas de similaridade estrutural
 - Análise das Informações de Movimento entre os frames
 - Determinar a direção do movimento, região de deslocamento e similaridade entre os frames

Introdução

Contribuições

- Caracterização do Problema de Contexto em Vídeos
 - Um único frame do vídeo pode não fornecer informações suficientes para detectar conteúdo pornográfico.
 - Sequências de frames oferecem contexto adicional
- Abordagem para extração do contexto dos frames de vídeos
 - Representação baseada em informações de movimento
 - Utiliza fluxo óptico para capturar e analisar o movimento entre frames, melhorando a interpretação do conteúdo
 - Extrair o contexto dos frames de vídeos em tempo real
- Abordagem de Baixo Custo Computacional
 - Método eficiente para a detecção de pornografia em tempo real, utilizando um classificador raso.

Agenda

- Introdução
- **Fundamentação Teórica**
- Trabalhos Relacionados
- Proposta
- Avaliação
- Conclusão

Fundamentação Teórica

Convolutional Neural Network

- CNN são redes neurais voltadas geralmente para tarefas de visão computacional
 - Reconhecimento e classificação de imagens
 - Detecção de objetos
- Capacidade de ser escalável
 - Permite processar uma grande quantidade de informação de entrada advindas dos pixels das imagens/*frames* dos vídeos
- Os neurônios são agrupados em camadas, por sua vez a arquitetura CNN é constituída por essas camadas

Fundamentação teórica

Classificação de Vídeos

- Resolução do vídeo: é uma característica que define as dimensões do vídeo
 - Posição horizontal dos *pixels* são estruturados em colunas
 - Posição vertical são organizados em linhas
- Espaço de cor: modelo matemático utilizado para descrever cada cor por meio de uma fórmula
 - Bits necessários para representar cada uma das cores básicas;
 - Espaço de cor RGB
- Taxa de *frames*: número de *frames* que são exibidos por segundo
 - Em geral, uma taxa de 24 FPS é utilizada para se adequar a percepção humana

Fundamentação teórica

Classificação de Vídeos Utilizando CNN

- Classificação do contexto dos vídeos com CNNs
 - Extração dos frames dos vídeos
 - Na classificação cada frame extraído é classificado individualmente
- Classificação em Tempo Real
 - Os frames devem ser classificados na mesma velocidade de transmissão do vídeo
- Desafios em Dispositivos com Recursos Limitados:
 - A classificação de cada frame deve consumir o mínimo de recursos computacional possível
 - Considerar a execução de outras tarefas em paralelo

Agenda

- Introdução
- Fundamentação Teórica
- **Trabalhos Relacionados**
- Proposta
- Avaliação
- Conclusão

Trabalhos Relacionados

	[Endeshaw et al. 2008]	[Rea et al. 2006]	[Lee et al. 2009]	[Moreira et al. 2009]	[Karpathy et al. 2019]	[Li et al. 2014]	[Yue-Hei Ng et al. 2015]	[Wu et al. 2015]	[Ji et al. 2013]
Informações de movimento	x			x	x		x	x	x
Frêquência	x								
Características visuais		x		x					
Características de áudio		x							
Hierárquico multinível			x	x					
Características cor, texto e forma			x						
SVM			x			x			
Espaço temporal				x	x	x	x	x	x
CNN					x	x	x	x	
LSTM							x	x	
3D CNN									x

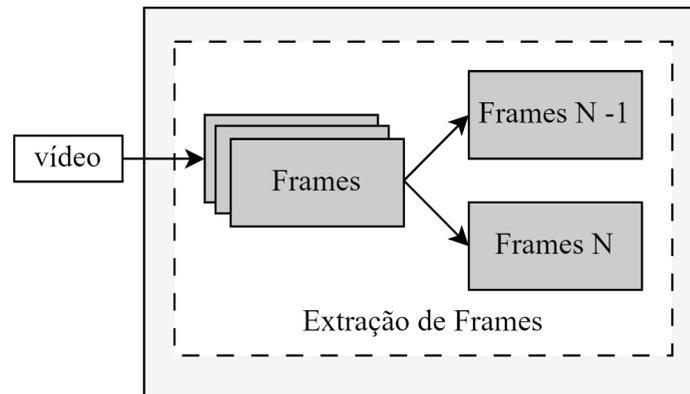
Agenda

- Introdução
- Fundamentação Teórica
- Trabalhos Relacionados
- **Proposta**
- Avaliação
- Conclusão

Proposta

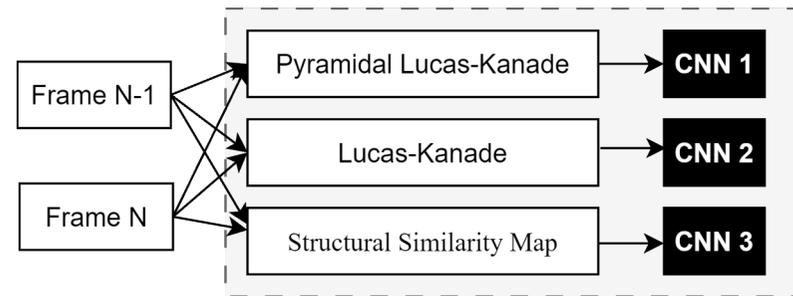
Módulo Extrator de Frames

- Este módulo recebe os frames de vídeos à medida que são capturados
 - Os frames são armazenados em uma janela deslizante
 - A janela deslizante possui um tamanho fixo
 - Comportamento de fila
 - Organizados em sequência conforme são recebidos

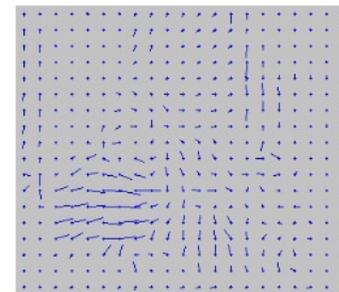


Proposta

Módulo Extrator de Movimento



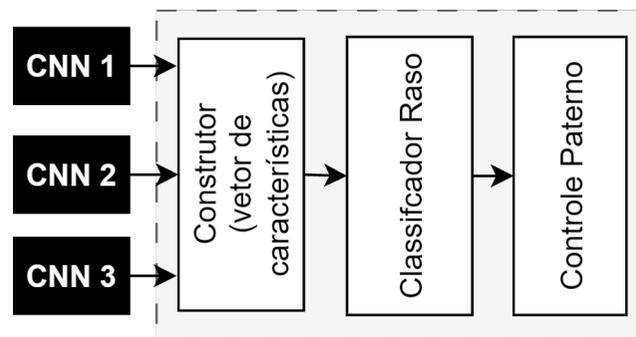
- Este módulo efetua a extração da informação de movimento utilizando três descritores de movimento
 - Os fluxos ópticos obtidos com os algoritmos Lucas-Kanade e PLK
 - Algoritmo de similaridade estrutural entre frames adjacentes, o SSM
- Os descritores de movimento são aplicados sobre dois frames: o atual e o anterior
 - São obtidas novas amostras de frames para cada descritor de movimento



Proposta

Módulo de Classificação de Cena

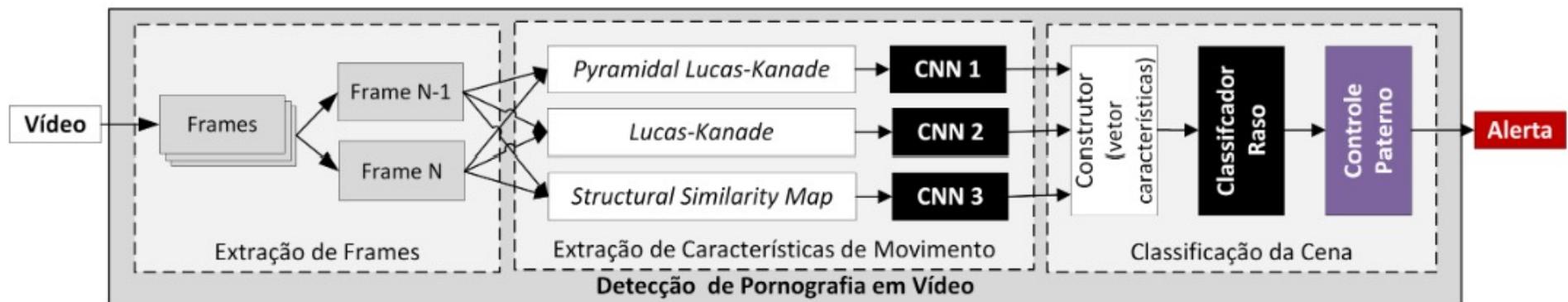
- Responsável em fornecer o resultado final da classificação do frame do vídeo
 - Realizado a partir da classificações individuais da CNN para cada descritor de movimento
 - Obtido pelos rótulos e valores de confiança da distribuição das classes (normal e pornô) atribuídos pela CNN
- Os atributos extraídos das CNNs são utilizados para conceber um único vetor de características para cada frame do vídeo
 - O vetor de características é fornecido para um classificador raso realizar a classificação final



Proposta

Análise de Contexto dos Vídeos

- Abordagem proposta baseada em informação de movimento para detecção de pornografia em vídeos sob a granularidade de frames
 - O modelo proposto avalia múltiplos descritores de movimento
 - Abordagem de detecção sensível ao contexto dos vídeos
 - Classificação precisa e em tempo real



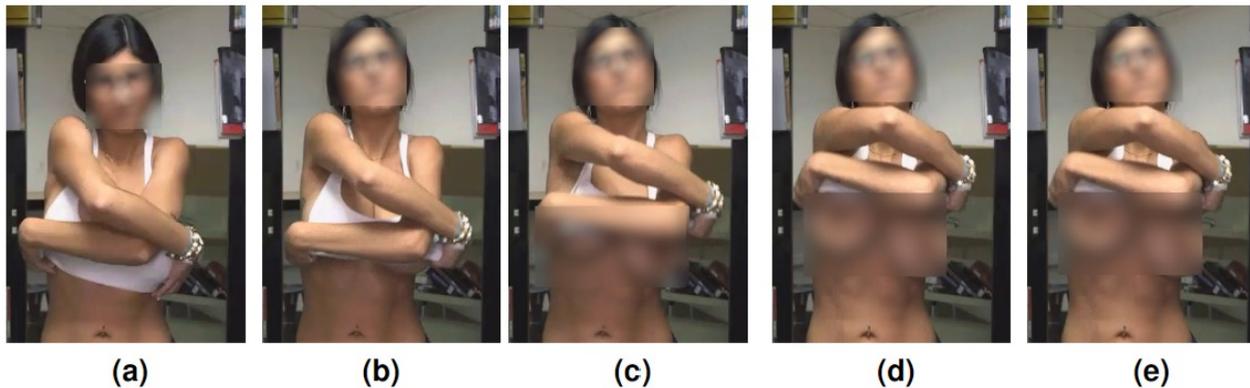
Agenda

- Introdução
- Fundamentação Teórica
- Trabalhos Relacionados
- Proposta
- **Avaliação**
- Conclusão

Avaliação

Dataset FPD

- Limitação dos datasets da literatura para detecção de pornografia
 - Não conseguem fornecer o nível esperado de granularidade
 - **Desprezam o contexto dos vídeos**
 - **Vídeos pornográficos podem conter frames que não tenham teor pornográfico**
- Construído o dataset FPD
 - Composto por 476.482 frames rotulados manualmente
 - Extraído de 14.671 vídeos
 - O dataset foi rotulado em uma granularidade fina (por frame)



Avaliação

Construção do Modelos

- O modelo proposto foi implementado fazendo uso das arquiteturas Mobilinet, Resnet e Squeezenet
 - A arquitetura treinada para cada descritor de movimento
 - Fluxos ópticos obtidos com os algoritmos Lucas-Kanade e PLK
 - Mapa de similaridade estrutural
 - Extração das características a partir da análise do frame atual e o anterior
 - O dataset foi dividido em três partes: treinamento, validação e teste (40%, 30% e 30%, respectivamente)

Avaliação

Questões de Pesquisa

- A classificação baseada em movimento proposta permite a detecção de conteúdo pornográfico em vídeos?

Descritor de Movimento	CNN	Acurácia (%)			
		Normal	Pornô	Dependente de Contexto	Geral
PLK	Mobilenet	99,39	20,06	89,19	69,55
	Resnet	97,87	23,98	80,26	67,37
	Squeezenett	99,90	0,11	99,98	66,67
LK	Mobilenet	99,11	42,52	80,75	74,13
	Resnet	97,62	7,30	94,28	66,40
	Squeezenet	99,41	25,50	81,34	68,75
SSM	Mobilenet	97,99	38,29	78,68	71,65
	Resnet	86,40	25,23	74,67	62,10
	Squeezenet	99,35	12,79	85,31	65,82

- As taxas de detecção utilizando descritores de movimento são mais estáveis do que aquelas obtidas em arquiteturas CNN utilizando apenas imagem
- Ex.: Mobilenet utilizando o descritor de Lucas-Kanade conseguiu classificar uma quantidade maior de frames pornográficos com um *tradeoff* de apenas 0,62%
- A CNN que utiliza descritor de movimento é capaz de classificar pornografia nos vídeos levando em consideração os frames que são dependentes do contexto

Avaliação

Questões de Pesquisa

- O módulo de classificação de cena proposto ajuda na detecção de conteúdo pornográfico em tempo real?

Classificador Raso	Acurácia (%)			
	Normal	Pornô	Dependente de Contexto	Geral
AdaBoost	92,03	93,97	59,39	81,80
Bagging (REPTree)	87,52	93,37	75,01	85,30
BayesNet	88,28	92,81	74,21	85,11
J48	87,87	93,56	66,42	82,62
RandomForest	83,82	92,90	74,80	83,84
REPTree	88,31	93,62	74,09	85,35

- Os experimentos mostram uma melhora na acurácia na classificação obtida para cada um dos classificadores rasos avaliados
- Por exemplo, o classificador raso REPTree teve um aumento 7,61% até 15,26% na avaliação geral da proposta
- Foi evidenciado pelo aumento das taxas de acurácia para a classe de frames pornográficos e dependente de contexto

Avaliação

Questões de Pesquisa

- A abordagem proposta é adequada para utilização em dispositivos com recursos computacionais limitados?

CNN + Descritor de Movimento	Benchmark Jetson TK1 (ms.)	
	CPU	GPU
Pyramidal Lucas-Kanade	1572.02	1170
Lucas-Kanade	1550.95	1168
Structural Similarity Map (SSM)	1548.54	1109
Classificador Raso		
REPTree	0.7631	

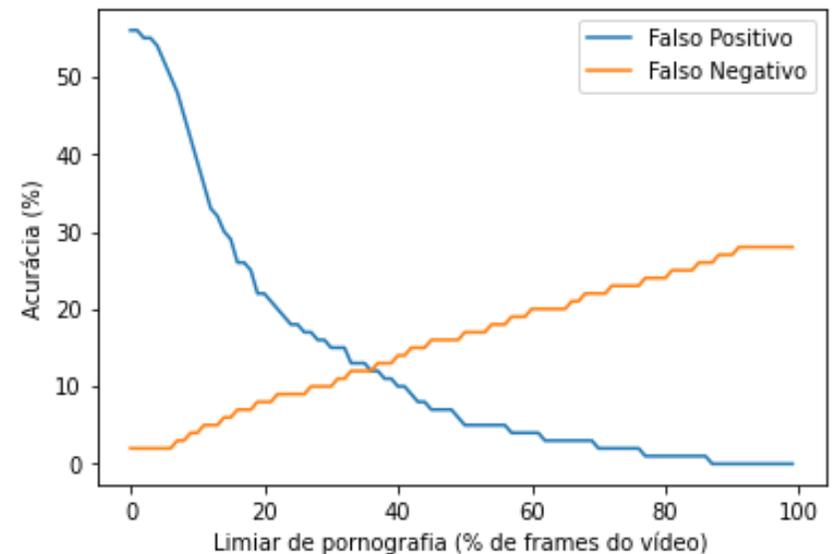
- Os experimentos realizados na placa da NVIDIA Jetson TK1, dispositivo com recursos computacionais limitados
- Selecionado a melhor arquitetura CNN e o melhor classificador raso (REPTree), atingindo somente 0.7631 ms
- Realizado experimentos de benchmark demonstramos que a abordagem proposta é promissora para aplicação em dispositivos com recursos limitados, possuindo um tempo aceitável na detecção em tempo real.

Avaliação

Questões de Pesquisa

- A abordagem proposta pode ser aplicada para detectar conteúdo pornográfico em uma granularidade de vídeo?

- O vídeo é classificado como pornográfico com base na análise de todos os frames e um limiar predeterminado
- O limiar define a proporção de frames pornográficos necessários para classificar o vídeo como pornográfico
- É possível notar que nossa proposta é capaz de atingir uma taxa de FN de 6% e uma taxa de FP de 15% ao usar um limiar pornográfico de 50%.



Agenda

- Introdução
- Fundamentação Teórica
- Trabalhos Relacionados
- Proposta
- Avaliação
- **Conclusão**

Conclusão

- Abordagens atuais na literatura para detecção de pornografia em vídeo são limitadas
 - Não foram projetadas para detecção em tempo real.
 - Abordagens mais promissoras combinam convoluções 3D e arquiteturas CNN e LSTM
 - Tratam o problema de contexto
 - Aplicar essas técnicas em dispositivos com limitações de memória, CPU ou disco ainda é inviável
- Nossa abordagem projetada para detecção em tempo real de conteúdo pornográfico em vídeos
 - Melhora significativamente a acurácia da detecção do pornográfico em vídeos, considerando o seu contexto.
 - Fornece recursos para auxiliar o controle paterno
 - Além da abordagem ser leve, pode ser implementada em dispositivos com restrições de recurso computacional

Detecção de Mídias Pornográficas em Dispositivos com Recursos Limitados para Controle Parental

Perguntas?

Jhonatan Geremias, Eduardo Kugler Viegas, Altair Olivo Santin, Jackson Mallmann

*Programa de Pós-Graduação em Informática – Pontifícia Universidade Católica do Paraná, Brasil

{jgeremias, eduardo.viegas, santin}@ppgia.pucpr.br

**Instituto Federal Catarinense (IFC), Brusque – Santa Catarina, Brasil

jackson.mallmann@ifc.edu.br

SBSeg 2024

Simpósio Brasileiro em Segurança da Informação (SBSeg)