



**HỌC VIỆN
CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG**

Posts and Telecommunications Institute of Technology

BÁO CÁO THỰC TẬP CƠ SỞ

ĐỀ TÀI: LÀM RÕ HÌNH ẢNH SÂU

Sinh viên thực hiện: Nguyễn Thị Lan

Giáo viên hướng dẫn: ThS Đinh Xuân Trường

Mã sinh viên: B21DCCN818 - Lớp: D21CQCN09-B

Khoa CNTT 1 - PTIT

23/04/2024

Nội dung

1. Giới thiệu
2. Cơ sở lý thuyết
3. Làm Rõ Ảnh Không Mùa
4. Làm Rõ Ảnh Mùa
5. Các hàm mất mát
6. Các Bộ Dữ liệu Thử Nghiệm cho Việc Làm Rõ Ảnh
7. Đánh giá hiệu suất
8. Làm Rõ Ảnh Theo Miền Cụ Thể

Giới thiệu

Giới thiệu chung

Làm mờ hình ảnh là một nhiệm vụ kinh điển trong thị giác máy tính cấp thấp, đã thu hút được sự chú ý từ cộng đồng xử lý hình ảnh và thị giác máy tính. Mục tiêu của việc làm mờ hình ảnh là khôi phục hình ảnh sắc nét từ hình ảnh đầu vào bị mờ.



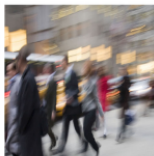
(a) Camera shake blur



(b) Out-of-focus blur



(c) Moving object blur



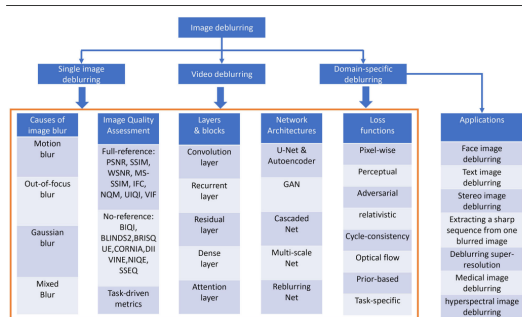
(d) Mixed blur

Hình: Ảnh mờ

Cơ sở lý thuyết

Phát biểu vấn đề

Mờ ảnh có thể phát sinh từ nhiều nguyên nhân khác nhau trong quá trình chụp ảnh, bao gồm rung máy ảnh, chuyển động trong cảnh, và mất nét.



Hình: các kỹ thuật làm mờ ảnh

Đánh Giá Chất Lượng Ảnh

Các phương pháp đánh giá chất lượng ảnh (IQA) có thể được phân loại thành các chỉ số chủ quan và khách quan.

- Chủ quan: dựa trên đánh giá của con người, chỉ số đại diện MOS đánh giá trên thang điểm từ 1-5
- Khách quan: được chia thành hai loại: chỉ số IQA có tham chiếu đầy đủ và không có tham chiếu.

Làm Rõ Ảnh Không Mùa

Làm rõ ảnh không mờ

Trong

trường hợp hạt nhân mờ được cung cấp, gọi là làm rõ ảnh không mờ (gặp khó khăn do nhiễu cảm biến và sự mất thông tin tần số cao) Các phương pháp không mờ:

- Giải Phóng Không Mờ Kèm Làm Sạch Nhiễu
- Học Mô-đun Làm Sạch Nhiễu cho Làm Rõ Ảnh Không Mờ

Method	Category	Blur type	Dataset	Architecture	Key idea
DCNN Xu et al. (2014)		Gaussian, disk			First work to combine traditional optimization-based schemes and neural networks
IRCNN Zhang et al. (2017)		Gaussian, motion			Learn a set of CNN denoisers, use as a modular part of model-based optimization methods to tackle inverse problems
FCNN Zhang et al. (2017)		Motion			Adaptively learn image priors to preserve image details and structure with a robust L1 loss
FDN Kruse et al. (2017)	Uniform	Motion	Convolution	CNN	learn a CNN-based prior with an FFT-based deconvolution scheme
GLRA Ren et al. (2018)		Gaussian, disk, motion			Use generalized low-rank approximations of blur kernels to initialize the CNN parameters
DUBLID Li et al. (2019)		Motion			Recast a generalized TV-regularized algorithm into a deep network for blind image deblurring
RGDN Gong et al. (2020)		Motion			Incorporate deep neural networks into a fully parameterized gradient descent scheme
DWDN Dong et al. (2020)		Motion, Gaussian			Apply explicit deconvolution in feature space by integrating a classical Wiener deconvolution framework
USRNet Zhang et al. (2020)		Motion, Gaussian			End-to-end training of an unfolding network that integrates advantages of model-based and learning-based methods

Hình: Khái quát về các phương pháp làm mờ ảnh đơn hình sâu không mờ đơn

Làm Rõ Ảnh Mù

Các lớp và khối cơ bản được sử dụng trong mạng làm rõ ảnh.

Các lớp và khối này bao gồm:

Lớp tích chập

Lớp đệ quy

Lớp dư thừa

Lớp mật độ

Lớp chú ý

Table 3 Overview of deep video deblurring methods

Method	Category	Blur type	Dataset	Architecture	Key idea
DeBlurNet Su et al. (2017)				DAE	Five neighboring blurry images are stacked and fed into a DAE to recover the center sharp image.
STRCNN Hyun Kim et al. (2017)				RNN-DAE	Recurrent architecture which includes a dynamic temporal blending mechanism for information propagation.
PCNN Attala and Dandekar (2018)				DAE	Permutation invariant CNN which consists of several U-Nets taking a sequence of burst images as input.
DBLRGAN Zhang et al. (2018)				GAN	GAN-based video deblurring method, using a 3D CNN to extract spatio-temporal information.
Reblur2deblur Chen et al. (2018)	Non-uniform	Motion	Averaging	Reblur	Three consecutive blurry images are fed into the reblur2deblur framework to recover sharp images, which are used to compute the optical flow and estimate a blur kernel for reconstructing the input.
IFRNN Nuh et al. (2019)				RNN-DAE	RNN-based video deblurring network, where a hidden state is transferred from past frames to the current frame.
EDVR Wang et al. (2019)				DAE	Pyramidal, Cascading and Deformable (PCD) module for frame alignment and a Temporal and Spatial Attention (TSA) fusion module, followed by a reconstruction module to restore sharp videos.
STFAN Zhou et al. (2019)				DAE	STFAN takes the current blurred frame, the preceding blurred and restored frames as input and recovers a sharp version of the current frame.
CDVD-TSP Pua et al. (2020)				Cascade	Cascaded deep video deblurring which first calculates the sharpness prior, then feeds both the blurry images and the prior into an DAE.

Hình: "Khái quát về phương pháp làm rõ video sâu

Kiến trúc mạng phổ biến nhất được sử dụng cho việc làm rõ ảnh

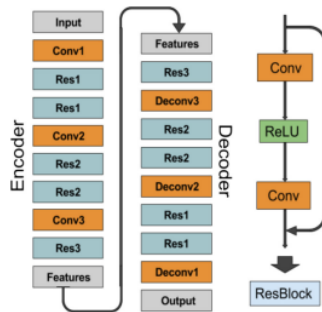
Deep Auto-Encoders (DAE)

Generative Adversarial Networks (GAN)

Cascaded Networks

Multi-Scale Networks

Reblurring Networks



Hình: Mạng làm rõ ảnh đơn sâu dựa trên kiến trúc Mã hóa Tự động Sâu (DAE) của^{8/17}

Các hàm mất mát

Các hàm mất mát

- Hàm mất mát pixel

$$L_{\text{pix1}} = \frac{1}{W \times H} \sum_{x=1}^W \sum_{y=1}^H |I_{\text{source}}(x, y) - I_{\text{db}}(x, y)|$$

$$L_{\text{pix2}} = \frac{1}{W \times H} \sum_{x=1}^W \sum_{y=1}^H (I_{\text{source}}(x, y) - I_{\text{db}}(x, y))^2$$

Ở đây, $I_{\text{source}}(x, y)$ và $I_{\text{db}}(x, y)$ là giá trị của ảnh sắc nét và ảnh đã được làm mờ tại vị trí (x, y) tương ứng.

- Hàm mất mát tri giác

$$L_{\text{per}} = \frac{1}{W \times H \times C} \sum_{x=1}^W \sum_{y=1}^H \sum_{c=1}^C \left(\sqrt{\sum_{i=1}^N \left(\Phi_{(x,y,c)}^I(I_s) - \Phi_{(x,y,c)}^I(I_{\text{db}}) \right)^2} \right)$$

Ở đây, $\Phi_I(x, y, c)(\cdot)$ là các đặc trưng đầu ra của mạng phân loại từ tầng thứ I . C là số kênh màu trong tầng thứ I .

Các hàm mất mát

Hàm mất mát đối kháng

$$L_{\text{adversarial}} = \log(1 - D(G(I_{\text{db}})))$$

Ở đây, $D(G(I_{\text{db}}))$ là xác suất mà ảnh đã được làm mờ là thực tế.

Hàm mất mát tương đối

$$L_{\text{RDBL}} = - [\log (\sigma (C(I_r) - E(C(G(I_{\text{db}}))))) + \log (1 - (\sigma (G(I_{\text{db}})) - E(C(I_s)))))]$$

Ở đây, I_r biểu diễn một ảnh thực tế.

Hàm mất mát dòng chảy quang học

Các Bộ Dữ liệu Thử Nghiệm cho Việc Làm Rõ Ảnh

Bộ dữ liệu làm rõ ảnh

Bộ dữ liệu của Levin et al

Bộ dữ liệu của Sun et al

Bộ dữ liệu của Köhler et al

Bộ dữ liệu của Lai et al

Bộ dữ liệu của GoPro

Bộ dữ liệu HIDE

Bộ dữ liệu RealBlur

Bộ dữ liệu Blur-DVS

International Journal of Computer Vision

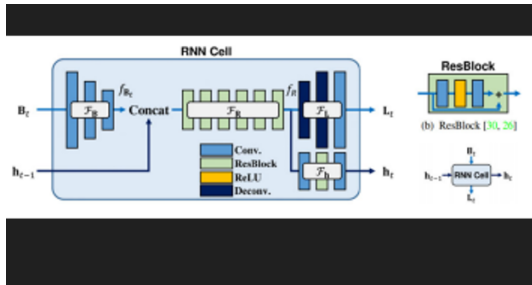
Table 4 Representative benchmark datasets for evaluating single image, video and domain-specific deblurring algorithms

Dataset	Synthetic	Real	Sharp Images	Blurred Images	Blur Model	Type	Train/Test Split
Levin et al. (2009)	×	✓	4	32	Uniform	Single image	Not divided
Sun and Hays (2012)	✓	×	80	640	Uniform	Single image	Not divided
Köhler et al. (2012)	✓	×	4	48	Non-uniform	Single image	Not divided
Lai et al. (2016)	✓	✓	108	300	Both	Single image	Not divided
GoPro Nah et al. (2017)	✓	✓	3,214	3,214	Non-uniform	Single image	2,103/1,111
HIDE Shen et al. (2019)	✓	×	8,422	8,422	Non-uniform	Single image	6,397/2,025
Blur-DVS Jiang et al. (2020)	✓	✓	2,178	2,918	Non-uniform	Single image	1,782/396
Su et al. (2017)	✓	✓	6,708	11,925	Non-uniform	Video	5,708/1,000
REDS Nah et al. (2019)	✓	✓	30,000	30,000	Non-uniform	Video	24,000/3,000
Hradis et al. (2015)	✓	×	3M+35K	3M+35K	Non-uniform	Text	3M/35K
Shen et al. (2018)	✓	×	6,564	130M+16K	Uniform	Face	130M/16K
Zhou et al. (2019)	✓	×	20,637	20,637	Non-uniform	Stereo	17,319/3,318

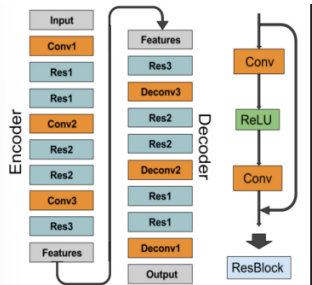
Hình: Bộ dữ liệu thử nghiệm đại diện để đánh giá các thuật toán làm rõ ảnh đơn, video và trong các lĩnh vực cụ thể.

Bộ dữ liệu làm rõ video

Bộ dữ liệu REDS: Nah et al. (2019)

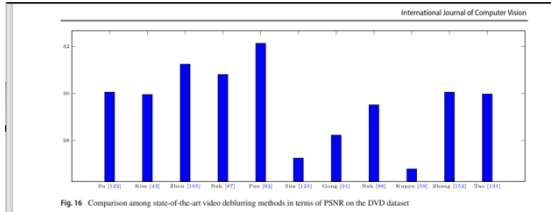


DVD dataset: Su et al. (2017)

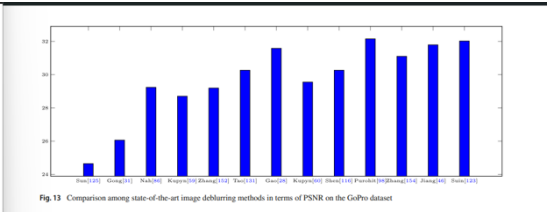


Đánh giá hiệu suất

Deblurring hình ảnh đơn



Deblurring video



Làm Rõ Ảnh Theo Miền Cụ Thể

Deblurring hình ảnh khuôn mặt



Deblurring hình ảnh văn bản

Digital images in this book
els are allowed to be 4-, 8-, 16-
possible. Specifically, a hex
is sometimes used (see the

Digital images in this book
els are allowed to be 4-, 8-, 16-
possible. Specifically, a hex
is sometimes used (see the

Digital images in this book
els are allowed to be 4-, 8-, 16-
possible. Specifically, a hex
is sometimes used (see the

Digital images in this book
els are allowed to be 4-, 8-, 16-
possible. Specifically, a hex
is sometimes used (see the

Digital images in this book
els are allowed to be 4-, 8-, 16-
possible. Specifically, a hex
is sometimes used (see the

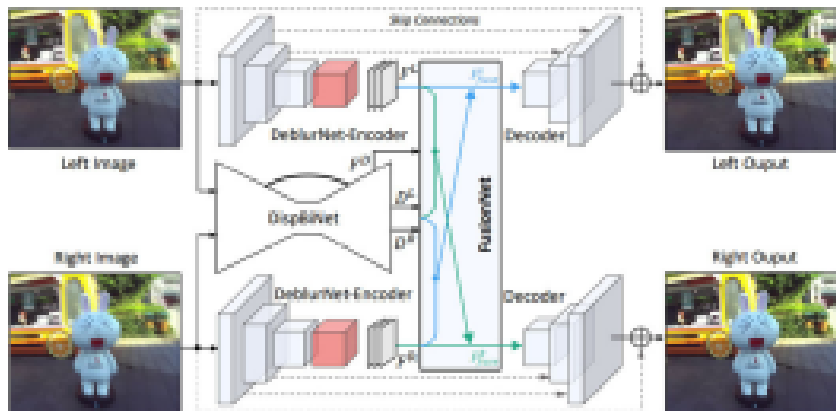
Digital images in this book
els are allowed to be 4-, 8-, 16-
possible. Specifically, a hex
is sometimes used (see the

Digital images in this book
els are allowed to be 4-, 8-, 16-
possible. Specifically, a hex
is sometimes used (see the

Digital images in this book
els are allowed to be 4-, 8-, 16-
possible. Specifically, a hex
is sometimes used (see the

Digital images in this book
els are allowed to be 4-, 8-, 16-
possible. Specifically, a hex
is sometimes used (see the

Deblurring hình ảnh stereo





**HỌC VIỆN
CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG**

Posts and Telecommunications Institute of Technology

Cảm ơn thầy cô đã chú ý lắng nghe!

Sinh viên thực hiện: Nguyễn Thị Lan

Giáo viên hướng dẫn: ThS Đinh Xuân Trường

Mã sinh viên: B21DCCN818 - Lớp: D21CQCN09-B