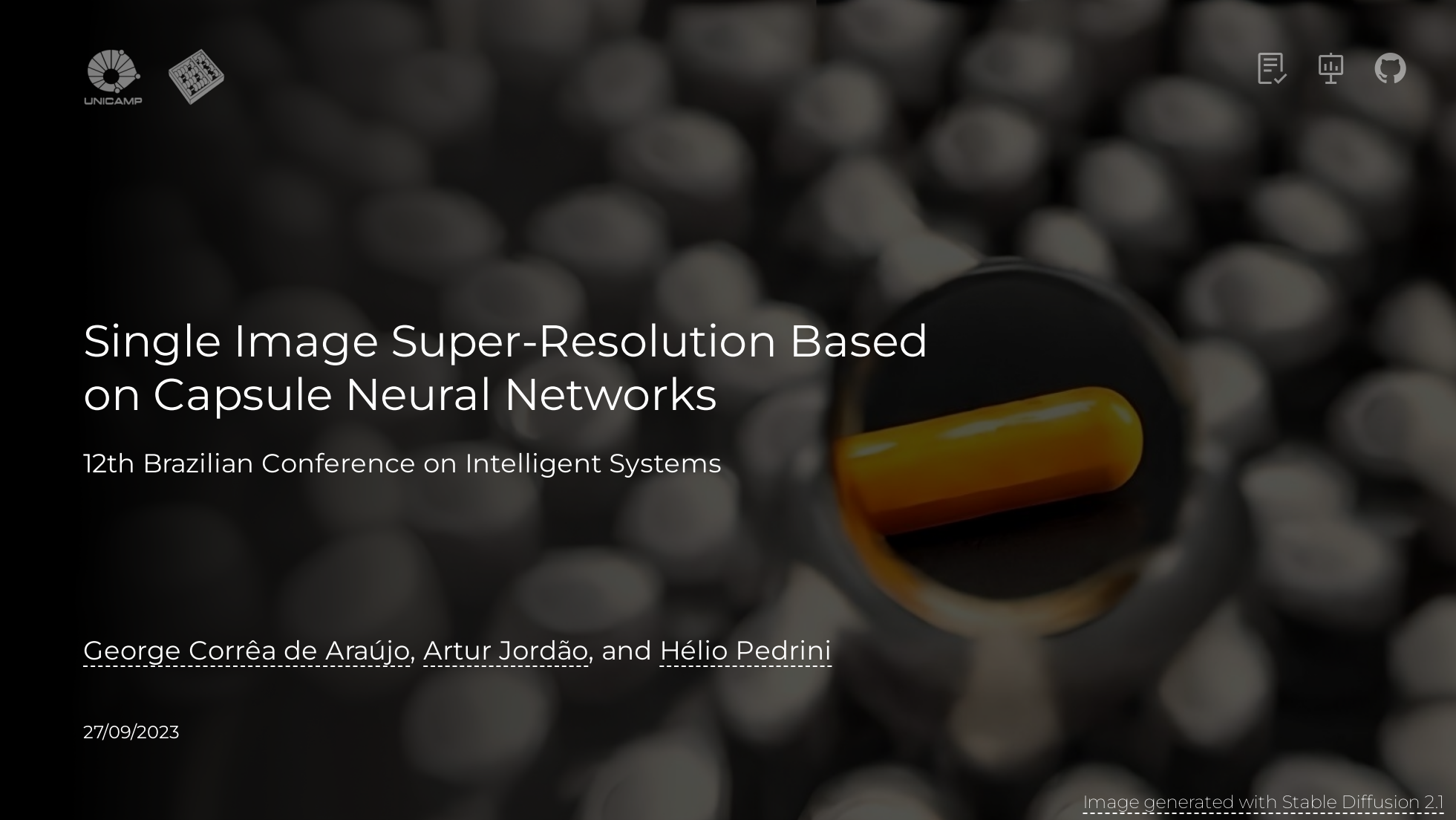


Single Image Super-Resolution Based on Capsule Neural Networks

12th Brazilian Conference on Intelligent Systems

George Corrêa de Araújo, Artur Jordão, and Hélio Pedrini

27/09/2023

Image generated with Stable Diffusion 2.1

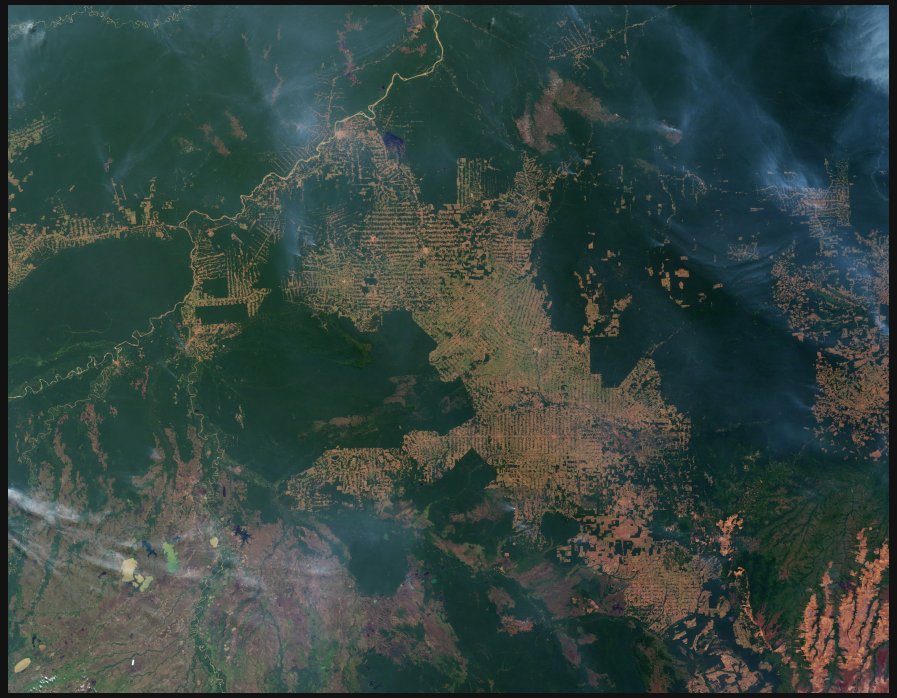
Table of Contents

Computer Vision Tasks

Computer Vision Tasks

MRI of ankle[1]

License plate[2]



Amazon deforestation[3]

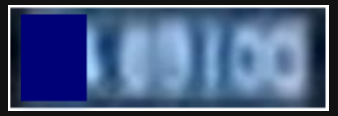
[1]Super Resolution Techniques for Medical Image Processing

[2]Beyond Human-level License Plate Super-resolution with Progressive Vehicle Search and Domain Priori GAN [3]The Earth Observatory

Computer Vision Tasks

Computer Vision Tasks

LR MRI of ankle[1]



LR license plate[2]

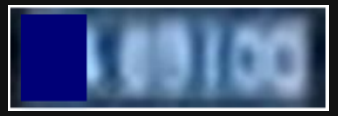
[1]Super Resolution Techniques for Medical Image Processing

[2]Beyond Human-level License Plate Super-resolution with Progressive Vehicle Search and Domain Priori GAN

Computer Vision Tasks

LR MRI of ankle[1]

HR MRI of ankle[1]



LR license plate[2]

HR license plate[2]

[1]Super Resolution Techniques for Medical Image Processing

[2]Beyond Human-level License Plate Super-resolution with Progressive Vehicle Search and Domain Priori GAN

"Enhance"

"Enhance"

Increasing resolution as seen in fiction[1]

[1]Adapted from CSI Zoom Enhance on YouTube

Upscaling

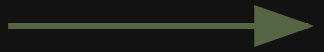
Upscaling

LR image[1]

[1] A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics

Upscaling

LR image[1]



[1] A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics

Upscaling

LR image[1]

Upscaled image

[1] A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics

Super-Resolution

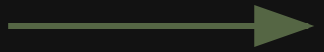
Super-Resolution

LR image[1]

[1] A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics

Super-Resolution

LR image[1]



[1] A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics

Super-Resolution

LR image[1]

Super-resolution image

[1] A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics

Super-Resolution

Super-Resolution

- Most solutions use CNNs
  - SRCNN, EDSR, RDN, RCAN, WDSR, SRGAN, ESRGAN, ...
  - Good visual quality

Super-Resolution

- Most solutions use CNNs
  - SRCNN, EDSR, RDN, RCAN, WDSR, SRGAN, ESRGAN, ...
  - Good visual quality
- CNNs have known drawbacks

Super-Resolution

- Most solutions use CNNs
  - SRCNN, EDSR, RDN, RCAN, WDSR, SRGAN, ESRGAN, ...
  - Good visual quality
- CNNs have known drawbacks

Super-Resolution

- Most solutions use CNNs
  - SRCNN, EDSR, RDN, RCAN, WDSR, SRGAN, ESRGAN, ...
  - Good visual quality
- CNNs have known drawbacks

Super-Resolution

- Most solutions use CNNs
  - SRCNN, EDSR, RDN, RCAN, WDSR, SRGAN, ESRGAN, ...
  - Good visual quality
- CNNs have known drawbacks

Super-Resolution

- Most solutions use CNNs
  - SRCNN, EDSR, RDN, RCAN, WDSR, SRGAN, ESRGAN, ...
  - Good visual quality
- CNNs have known drawbacks

Super-Resolution

- Most solutions use CNNs
  - SRCNN, EDSR, RDN, RCAN, WDSR, SRGAN, ESRGAN, ...
  - Good visual quality
- CNNs have known drawbacks

Super-Resolution

- Most solutions use CNNs
  - SRCNN, EDSR, RDN, RCAN, WDSR, SRGAN, ESRGAN, ...
  - Good visual quality
- CNNs have known drawbacks

Capsules

Capsules

- Introduced by Hinton et al. in 2011
  - First successful implementation by Sabour et al. in 2017

Capsules

- Introduced by Hinton et al. in 2011
  - First successful implementation by Sabour et al. in 2017
- Proposes to solve some of the main flaws found in CNNs
  - Inability to identify spatial hierarchy between elements
  - Lack of rotation invariance

Capsules

- Introduced by Hinton et al. in 2011
  - First successful implementation by Sabour et al. in 2017
- Proposes to solve some of the main flaws found in CNNs
  - Inability to identify spatial hierarchy between elements
  - Lack of rotation invariance
- Inspired by the human visual system

Capsules

- Introduced by Hinton et al. in 2011
  - First successful implementation by Sabour et al. in 2017
- Proposes to solve some of the main flaws found in CNNs
  - Inability to identify spatial hierarchy between elements
  - Lack of rotation invariance
- Inspired by the human visual system
- Achieved good results in classification and segmentation tasks

Why "Capsule"?

Why "Capsule"?

“Instead of aiming for viewpoint invariance in the activities of "neurons" that use a single scalar output to summarize the activities of a local pool of replicated feature detectors, artificial neural networks should use local "capsules" that perform some quite complicated internal computations on their inputs and then encapsulate the results of these computations into a small vector of highly informative outputs.”[1]

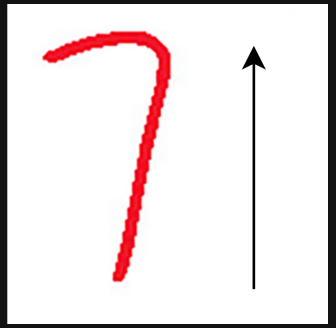
[1] Transforming Auto-Encoders

Convolution VS Capsule

Convolution VS Capsule

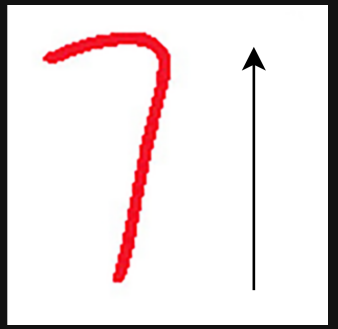
Adapted from Understanding Dynamic Routing between Capsules (Capsule Networks)

Convolution VS Capsule

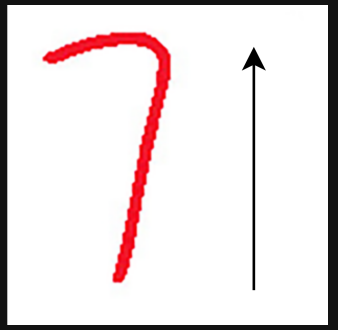


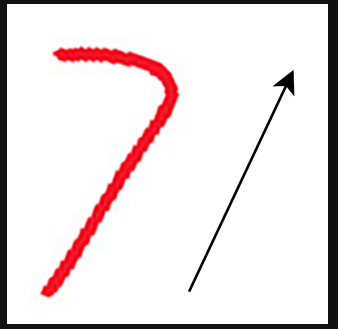
Convolution VS Capsule

Convolution VS Capsule



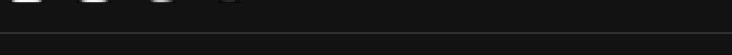
Convolution VS Capsule





Adapted from Understanding Dynamic Routing between Capsules (Capsule Networks)

Capsules' Reconstructions

Properties	Reconstruction from capsules
Scale and thickness	
Localized part	
Stroke thickness	
Localized skew	
Width and translation	
Localized part	

Adapted from Dynamic Routing Between Capsules

Capsules

Capsules

- Originally were used in classification tasks
  - Achieved state-of-the-art results

Capsules

- Originally were used in classification tasks
  - Achieved state-of-the-art results
- Capsules have been explored in other tasks
  - Object detection
  - Image segmentation
  - Visual question answering

Capsules

Capsules

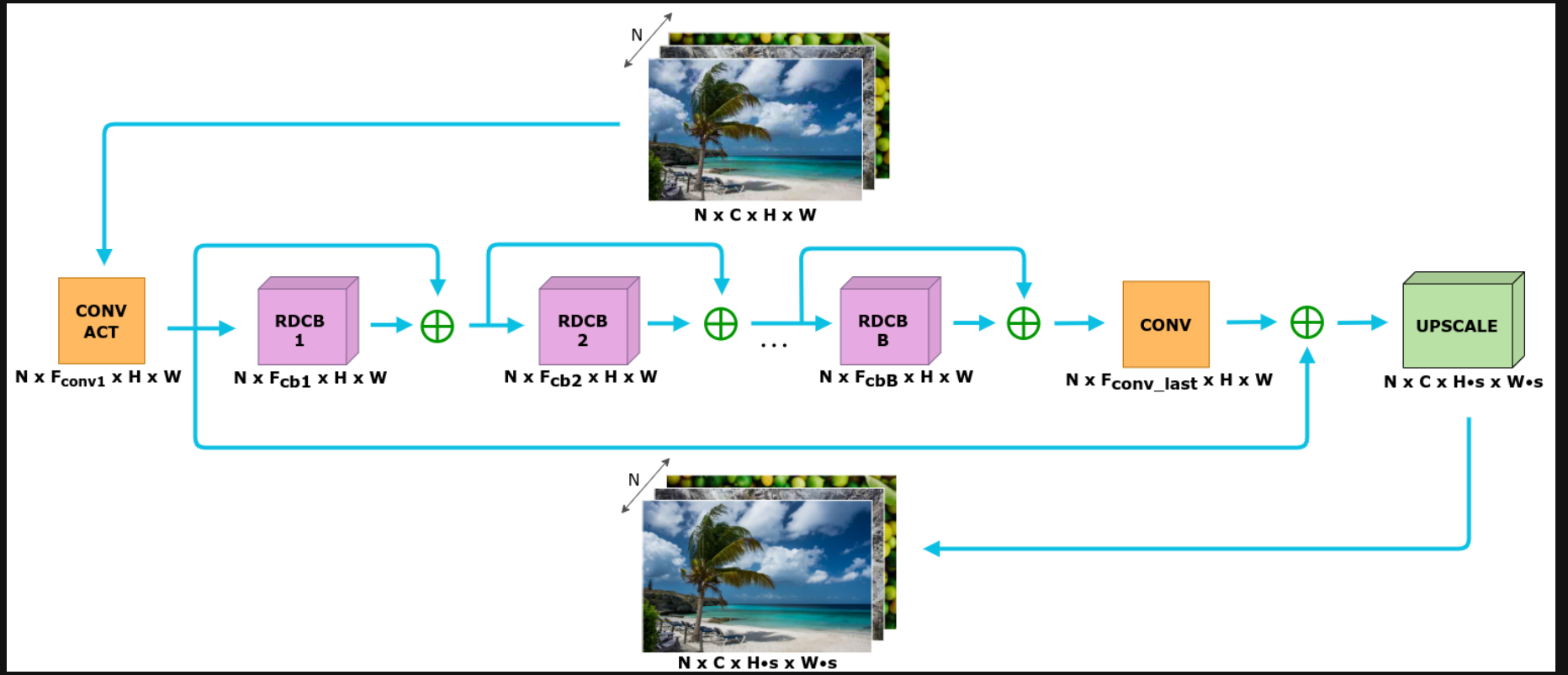
- Few explorations in SISR tasks
  - Little modifications to the original CapsNet

Capsules

- Few explorations in SISR tasks
  - Little modifications to the original CapsNet
- Novel concepts have been applied to CapsNets
  - Different capsules types
  - New routing algorithms

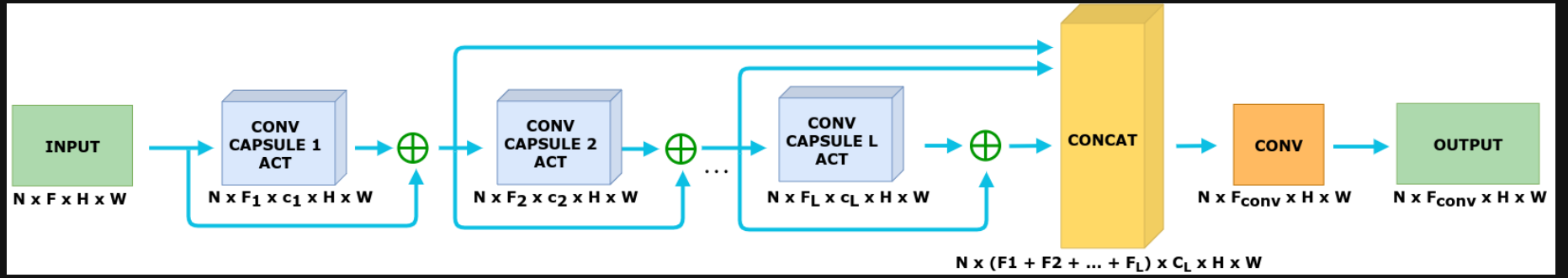
SRCaps

SRCaps



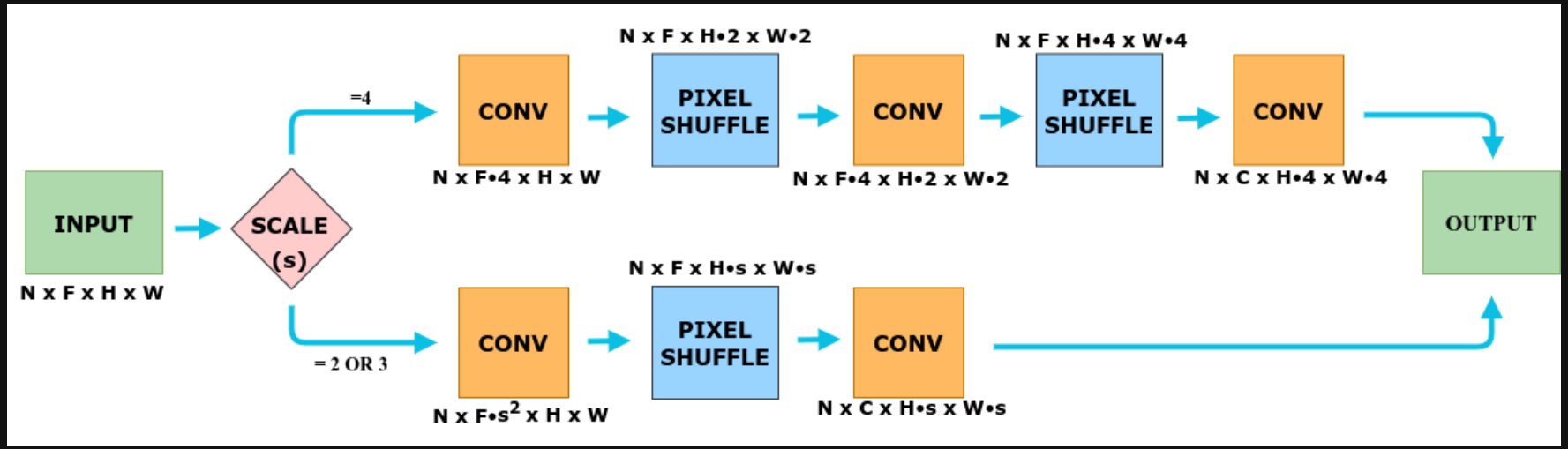
Model diagram

SRCaps



Capsblock diagram

SRCaps



UPNet diagram

Experimental Setup

Experimental Setup

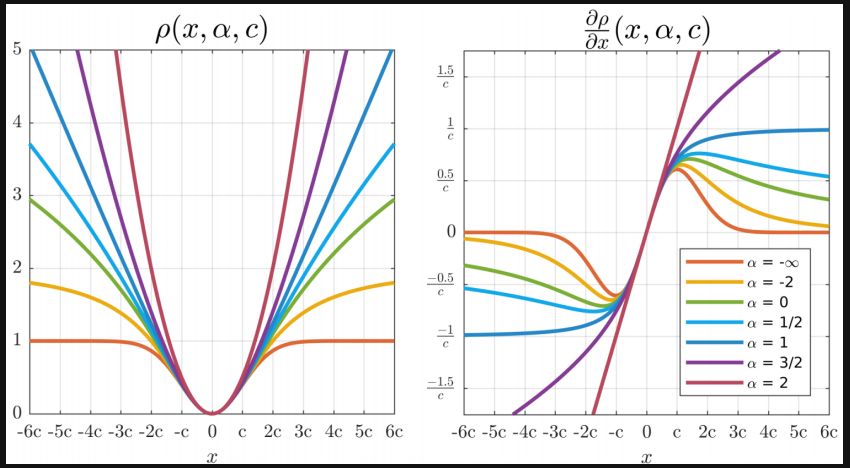
- Training
  - DIV2K training set
  - Losses: L1, SSIM, MS-SSIM, L1 after a few RDCBs, L1 + edge map, 3-PSNR, 3-SSIM, adaptive loss
  - Model configuration: refer to the paper

Experimental Setup

- Training
  - DIV2K training set
  - Losses: L1, SSIM, MS-SSIM, L1 after a few RDCBs, L1 + edge map, 3-PSNR, 3-SSIM, adaptive loss
  - Model configuration: refer to the paper
- Validation
  - DIV2K validation set, Set5, Set14, BSD100 (B100), Urban100
  - Metrics: PSNR, SSIM, MS-SSIM, FLIP

Adaptive Loss

α	Loss function
2	L2
1	Charbonnier / pseudo-Huber / L1-L2
0	Cauchy/Lorentzian
-2	Geman-McClure
$-\infty$	Welsch/Leclerc

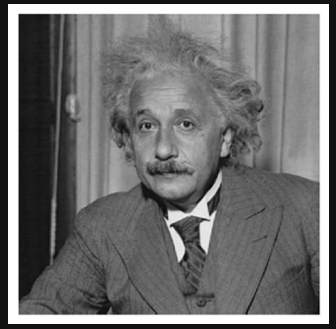


The general loss function (left) and its gradient (right) for different values of its shape parameter α[1]

[1] A General and Adaptive Robust Loss Function

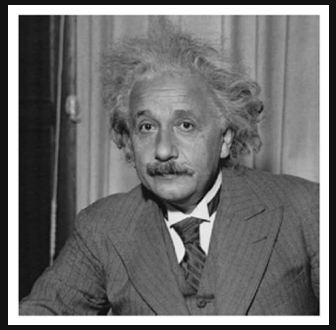
Why not only use PSNR and SSIM?

Why not only use PSNR and SSIM?

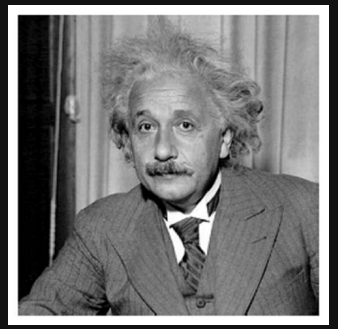


Reference

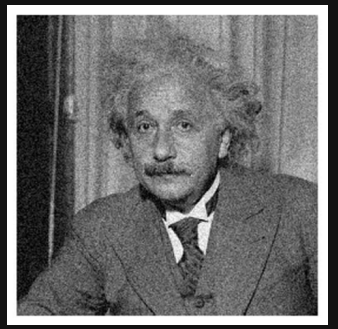
Why not only use PSNR and SSIM?



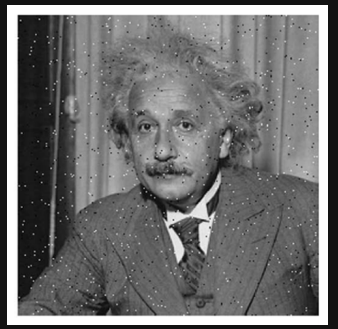
Reference



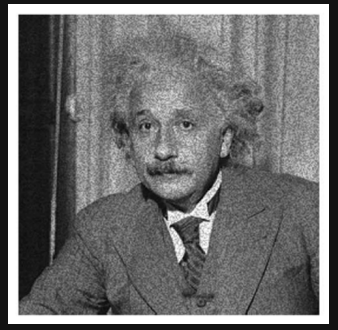
Contrast enhanced



Gaussian noise

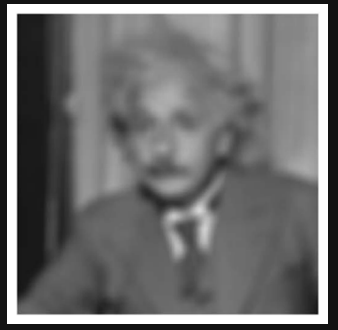


Salt-pepper noise



Speckle noise

JPEG compressed



Blurred

Adapted from Content-weighted video quality assessment using a three-component image model

Model Comparison

	SRCaps	EDSR	RCAN	WDSR	RDN	SRCNN
Number of Parameters	15M	1.5M	12.6M	4.8M	22.3M	20.1K
Number of Blocks	7	16	10 × 16	16	16	1 (not residual)
Number of Layers per Block	4	2	3	3	8	3
Dense Connections	✓	✗	✗	✗	✓	✗
Uses mean RGB	✗	✓	✗	✓	✗	✗
Sub-pixel Convolution	✓	✓	✓	✓	✓	✗
Loss Function	Adaptive	L1	L1	L1	L1	L1

Results

Results

Model results for “barbara” image from Set14 dataset

Results

Model results for “0891” image from DIV2K dataset

Conclusions

Conclusions

- The purpose of this work was to
  - evaluate the use of the capsules in SISR
  - verify new forms of training and validating the results of NNs

Conclusions

- The purpose of this work was to
  - evaluate the use of the capsules in SISR
  - verify new forms of training and validating the results of NNs
- We proposed SRCaps, a new model for SISR based on the CapsNet architecture
  - despite the inferior result, a smaller number of layers obtained a relevant result
  - nonlinearity function applied may be a limiting factor

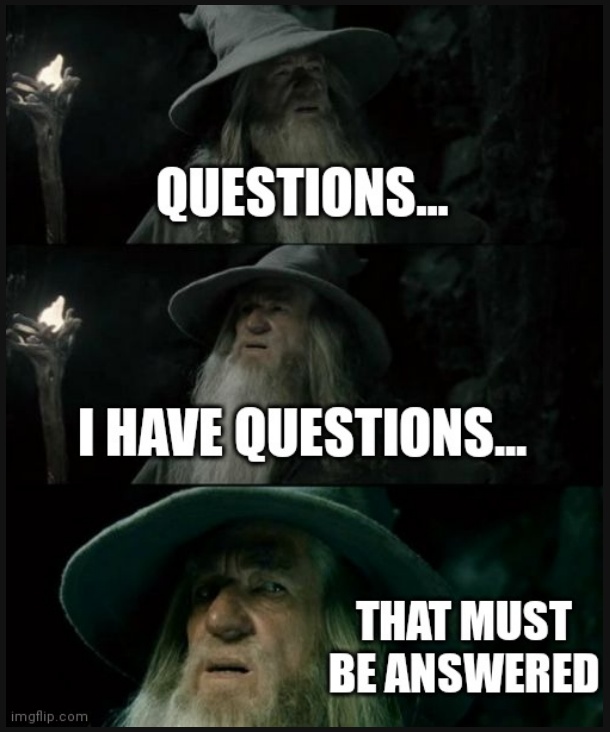
Conclusions

- The purpose of this work was to
  - evaluate the use of the capsules in SISR
  - verify new forms of training and validating the results of NNs
- We proposed SRCaps, a new model for SISR based on the CapsNet architecture
  - despite the inferior result, a smaller number of layers obtained a relevant result
  - nonlinearity function applied may be a limiting factor
- Highlights
  - RCAN ability to create smooth edges
  - adaptive loss function
  - metrics not exactly reflect visual quality

Conclusions

- The purpose of this work was to
  - evaluate the use of the capsules in SISR
  - verify new forms of training and validating the results of NNs
- We proposed SRCaps, a new model for SISR based on the CapsNet architecture
  - despite the inferior result, a smaller number of layers obtained a relevant result
  - nonlinearity function applied may be a limiting factor
- Highlights
  - RCAN ability to create smooth edges
  - adaptive loss function
  - metrics not exactly reflect visual quality
- Future research
  - replace the composition of the UPNet
  - new non-linearity and routing functions for the capsules
  - novel capsule models

Questions?



Gandalf has questions[1]

[1] imgflip