IEEE WORLD CONGRESS ON COMPUTATIONAL INTELLIGENCE
8-13 July 2018, Rio de Janeiro, Brazil
IEEE WCCI 2018 Rio de Janeiro, Brazil

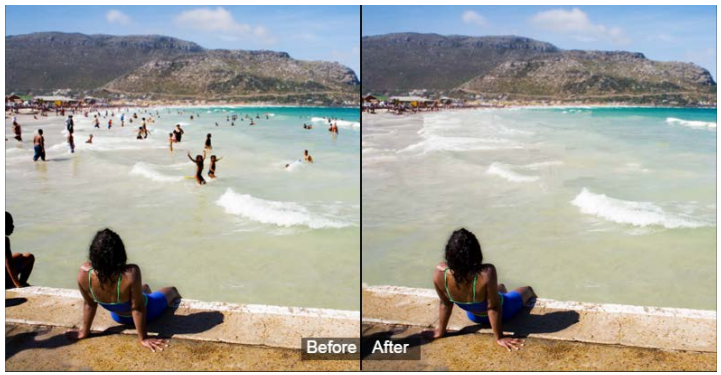
ChaLearn Looking at People
Inpainting Challenge
@WCCI18 @ECCV18

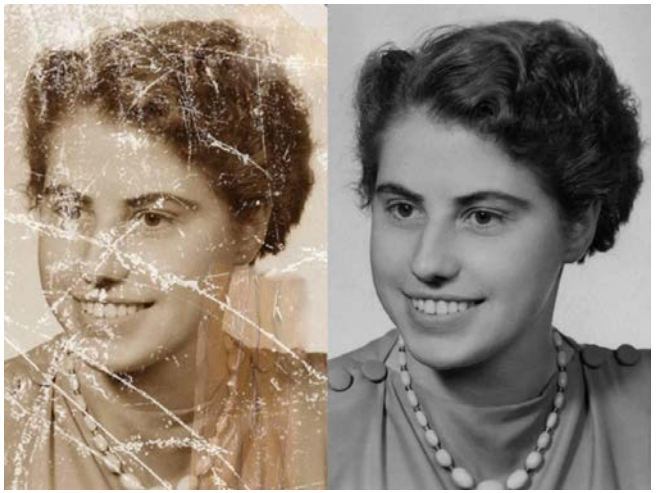
Stephane Ayache, Sergio Escalera, Florin Popescu, Isabelle Guyon, Umut Guclu, Yagmur Gucluturk, Marti Soler, Meysam Madadi, Xavier Baro, and Hugo Jair Escalante

http://chalearnlap.cvc.uab.es/

Image and video inpainting

- We target the visual inpainting task
  - Recovering/reconstructing lost or deteriorated parts of images and videos (also known as image/video interpolation)
- Related tasks and applications
  - Denoising, enhancement, restoration, super-resolution, etc.





ChaLearn LaP inpainting challenge

- We organized a challenge around 3 applications of image/video inpainting:
  1. Pose estimation
  2. Video de-captioning
  3. Fingerprint verification

- We provided datasets, evaluation protocol, baselines, prizes and dissemination opportunities

- The three tracks were launched in the CodaLab platform

- Milestones in WCCI2018 and ECCV2018

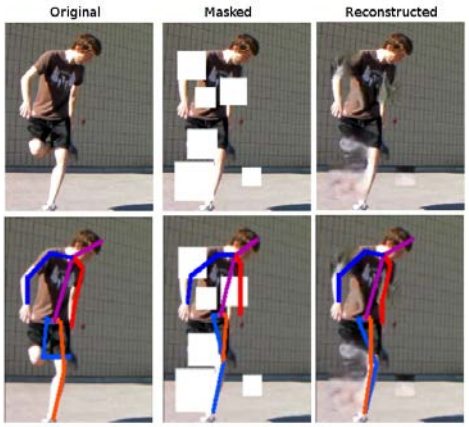
http://chalearnlap.cvc.uab.es/

http://www.chalearn.org/

http://codalab.org/

Pose estimation from occluded images

- Restore occluded parts of images for pose estimation
- A novel data set with 41,706 images taken from other benchmarks was released
  - Artificially introduced occlusions
  - Block based masks near joint positions
  - Large image diversity
- Evaluation: image quality metrics (MSE, PSNR, DSSIM) and performance of a pose estimation method on reconstructed images (WNJD)



Name	#Images Used	Cropped
MPII Human Pose Dataset [10]	26571	Yes
Leeds Sports Pose Dataset [11]	2000	No
Synchronic Activities Stickmen V [12]	1128	Yes
Short BBC Pose [13]	996	No
Frames Labelled In Cinema [14]	10381	Yes

Pose estimation from occluded images

- Baselines:

Name	DSSIM	MSE	WNJD
Context-Encoders (ImageNet Model)	0.3224	0.0524	0.1487
Context-Encoders	0.3425	0.0911	0.1489
Semantic Image Inpainting	0.4533	0.0942	0.2215
Multi-Scale Neural Patch	0.2947	0.0567	0.1491
anubhap93	0.2089	0.0176	0.1488

- Participation:
  - 42 participants registered
  - Top ranked participant: **anubhap93**

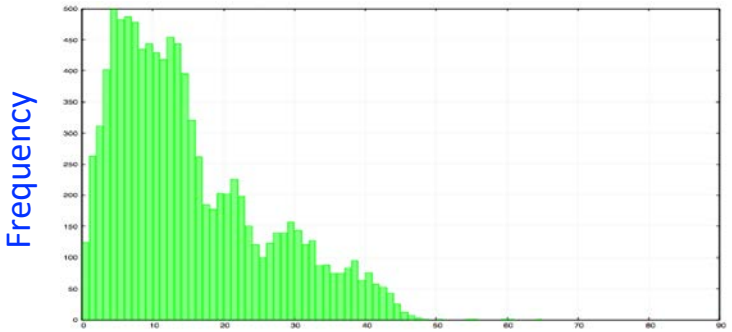
CNN model with regular+dilated convolutions, skip connections from the encoder and deconvolutions+convolutions to generate the full image.

Video de-captioning

- Remove caption from videos!
- Brand new data set comprising
  - 125 hours of video with subtitles
  - 70000 video clips, 125 frames (5 secs at 25 fps), 128x128 pixels

- Evaluation: image quality metrics (MSE, PSNR, DSSIM)

- Baselines:
  - Global and local versions of encoder-decoder with reconstruction loss



Video de-captioning

- About 35 different participants registered
- 7 active participants, 50 submissions, 6 beat our baselines

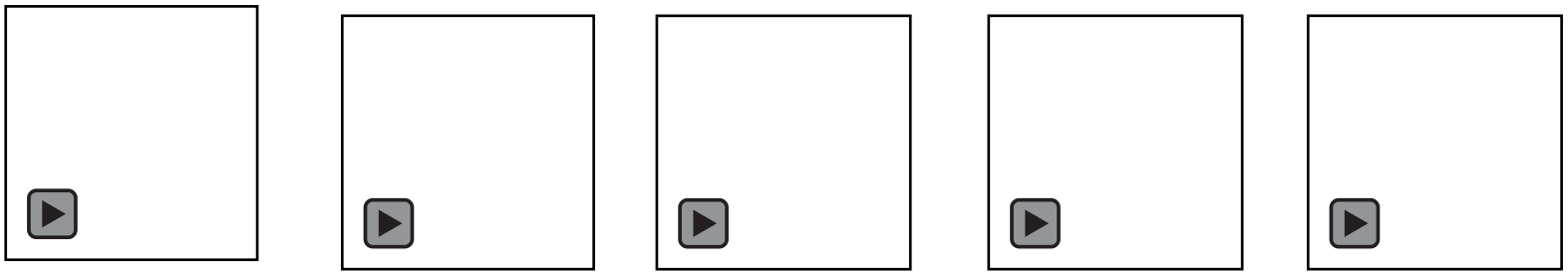
Results							
#	User	Entries	Date of Last Entry	<Rank> ▲	MSE ▲	PSNR ▲	DSSIM ▲
1	arnavkj95	12	05/14/18	1.3333	0.0014 (1)	31.9629 (1)	0.0512 (2)
2	hcilab	2	07/02/18	1.6667	0.0015 (2)	30.9972 (2)	0.0493 (1)
3	vismay	5	06/27/18	3.3333	0.0016 (3)	30.8919 (3)	0.0625 (4)
4	mcahny01	2	07/02/18	4.6667	0.0018 (4)	29.9306 (5)	0.0751 (5)
5	dhkim	7	07/05/18	5.6667	0.0021 (5)	28.5301 (6)	0.0867 (6)
6	ucs	7	07/05/18	6.6667	0.0021 (6)	28.5006 (7)	0.0868 (7)
7	Stephane	1	03/08/18	4.6667	0.0023 (7)	30.0993 (4)	0.0621 (3)
8	SanghyunWoo	9	06/27/18	8.0000	0.0027 (8)	27.5200 (8)	0.0994 (8)
9	mmadadi	4	03/20/18	9.0000	0.0037 (9)	25.8083 (9)	0.1005 (9)

baseline

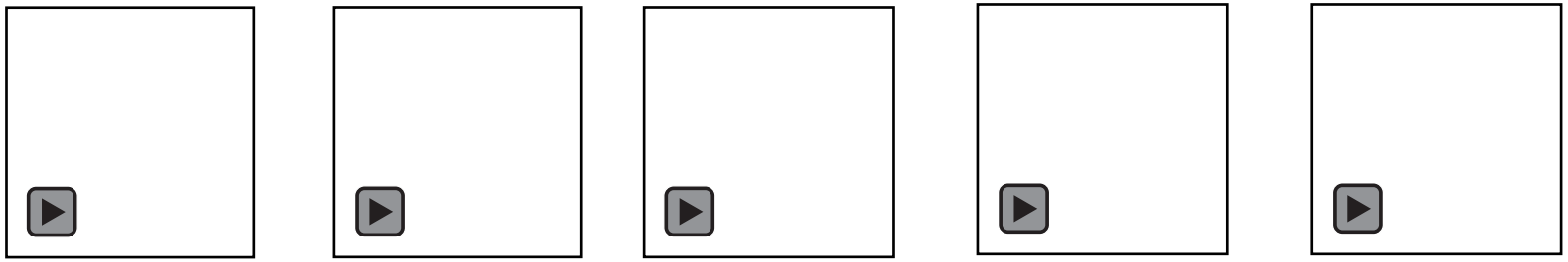
- Top ranked participants relied on deep architectures (e.g., U-Net and other CNNs)

Video de-captioning

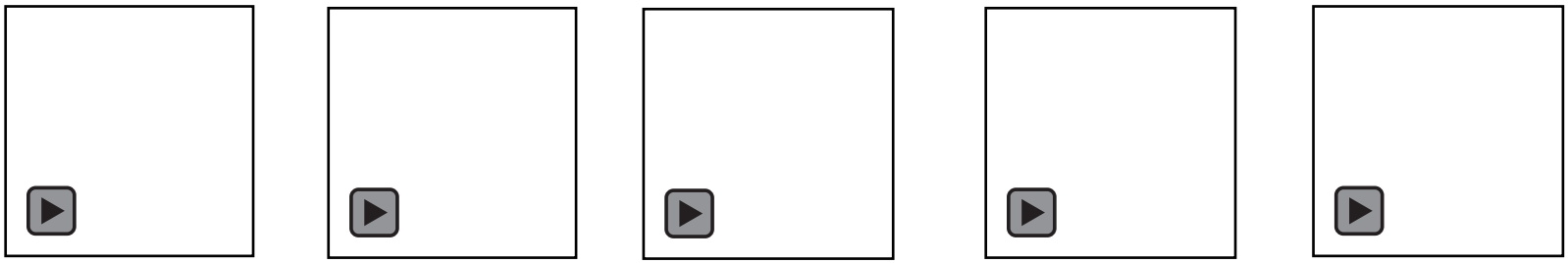
Input | Ground truth | 1st | 2nd | 3rd



« easy sample » : small font, transparent background, few text



« medium sample » : bigger font, transparent background, many text



« hard sample » : medium font, opaque background, more text

Inpainting and denoising for fingerprint verification

- To inpaint and denoise fingerprint images that contain artifacts like noise, scratches, etc.
- A new dataset with more than 150K images was generated:
  - Translations, rotations, blurring, modifying brightness, contrast, elastic transformation, occlusion, scratch, resolution, rotation
- Evaluation: image quality metrics (MSE, PSNR, DSSIM)
- Baseline: straightforward DNN

Inpainting and denoising for fingerprint verification

- More than 50 participants registered for the track

Results							
#	User	Entries	Date of Last Entry	<Rank> ▲	MSE ▲	PSNR ▲	SSIM ▲
1	rgsl888	27	07/07/18	1.6667	0.0239 (2)	16.8363 (1)	0.8069 (2)
2	hcilab	11	07/02/18	2.6667	0.0241 (3)	16.5974 (2)	0.8034 (3)
3	CVxTz	3	05/19/18	2.6667	0.0237 (1)	16.5770 (3)	0.7964 (4)
4	sukeshadigav	2	07/09/18	4.3333	0.0278 (6)	16.3872 (6)	0.8220 (1)
5	umuguc	1	03/13/18	4.6667	0.0252 (5)	16.4098 (4)	0.7954 (5)
6	finlouarn	1	03/24/18	5.0000	0.0251 (4)	16.3992 (5)	0.7904 (6)
7	Xiaojing	2	05/27/18	7.0000	0.0381 (7)	14.6347 (7)	0.6990 (7)
8	BriceRauby	3	05/01/18	8.0000	0.0398 (8)	14.1740 (8)	0.6954 (8)
9	yashkotadia	1	06/23/18	9.0000	0.0564 (9)	12.7785 (9)	0.6417 (9)
10	yg	1	03/14/18	10.0000	0.7282 (10)	1.3781 (10)	0.0001 (10)

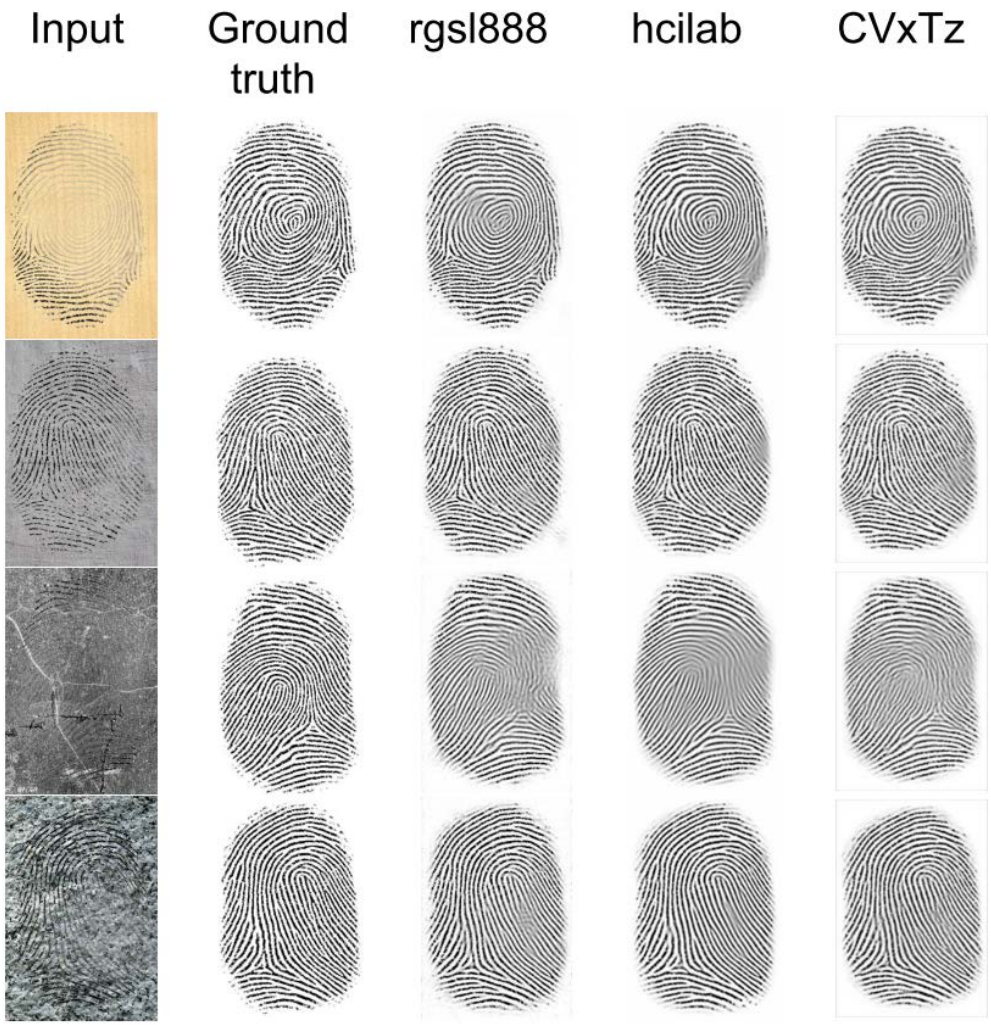
baseline

- Top ranked team: **rgsl888**

A deep encoding-decoding architecture with skip connections (encoding to decoding), optimized on L2 loss. This architecture has been enhanced with dilated convolution network and fine modifications in kernel sizes.

Inpainting and denoising for fingerprint verification

- Qualitative analysis

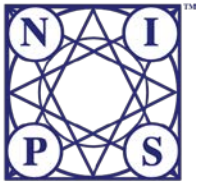


Summary

- So far the three tracks have attracted more than 100 participants with more than a dozen of active teams
- Solutions based on deep learning predominate among top ranked submissions
- Success: the three baselines have been outperformed by participants, we expect further improvements in the next few days

Forthcoming events/activities

- Multimedia Information Processing for Personality & Social Networks Analysis Workshop at ICPR – Beijing, August 2018 (**submission deadline mid July 2018**)
  - http://chalearnlap.cvc.uab.es/workshop/28/description/
- Chalearn Looking at People Satellite Workshop ECCV, Munich, September 2018 (**submission deadline August 5 2018**)
  - http://chalearnlap.cvc.uab.es/workshop/29/description/
- AutoML for Lifelong Machine Learning – NIPS 2018 (**starting July 23, 2018**)
  - https://www.4paradigm.com/competition/nips2018
- IEEE TPAMI SI on Image and Video Inpainting and Denoising (**Submission deadline: December 15, 2018**)
  - http://chalearnlap.cvc.uab.es/special-issue/30/description/
- IJCV SI on Analyzing Human Behavior from Social Media Data (**Submission deadline: March 1, 2019**)
  - http://chalearnlap.cvc.uab.es/special-issue/31/description/



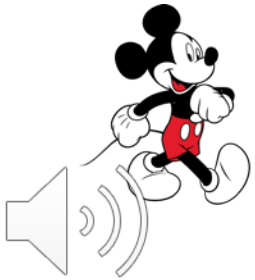
Sponsors!

Microsoft Research

amazon

Google

CHALEARN



Disney Research

IEEE WCCI 2018
Rio de Janeiro, Brazil

IEEE WORLD CONGRESS ON COMPUTATIONAL INTELLIGENCE
8-13 July 2018, Rio de Janeiro, Brazil

@ChaLearn_org

Stephane.Ayache@lif.univ-mrs.fr
http://chalearnlap.cvc.uab.es/