Virtual Restoration of Paintings Based on Deep Learning

Roman Sizyakin[a], Viacheslav Voronin[b] and Aleksandra Pižurica[a]

GHENT UNIVERSITY

STANKIN MOSCOW STATE UNIVERSITY OF TECHNOLOGY

[a]-Department Telecommunications and Information Processing, TELIN-GAIM, Ghent University, Ghent, Belgium
[b]- Center for Cognitive Technology and Machine Vision, Moscow State University of Technology "STANKIN", Moscow, Russian Federation

Motivation

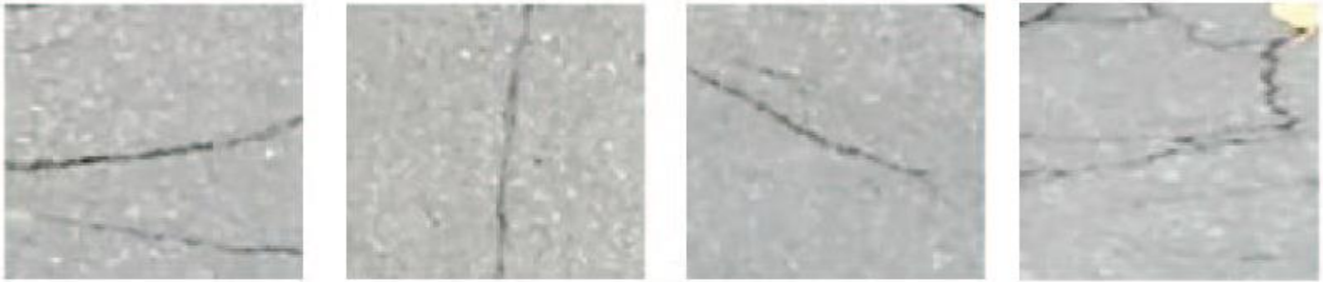
- Altarpiece, publicly available on the Closer to Van Eyck website http://closertovaneyck.kikirpa.be/

Crack detection challenges and why using deep neural networks?

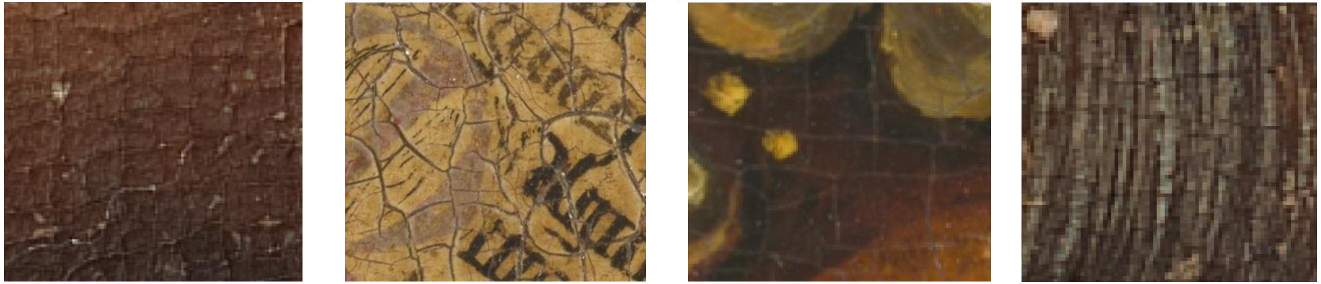
- Dealing with multi-modal data
- Feature selection (leaves many possibilities, ad-hoc choices)
- High-resolution images: computational complexity is a limitation
- Continuous learning is desired to cope with scarce training data

- Deep learning have great performance demonstrated in many computer vision tasks
- Detecting cracks in paintings is much more challenging

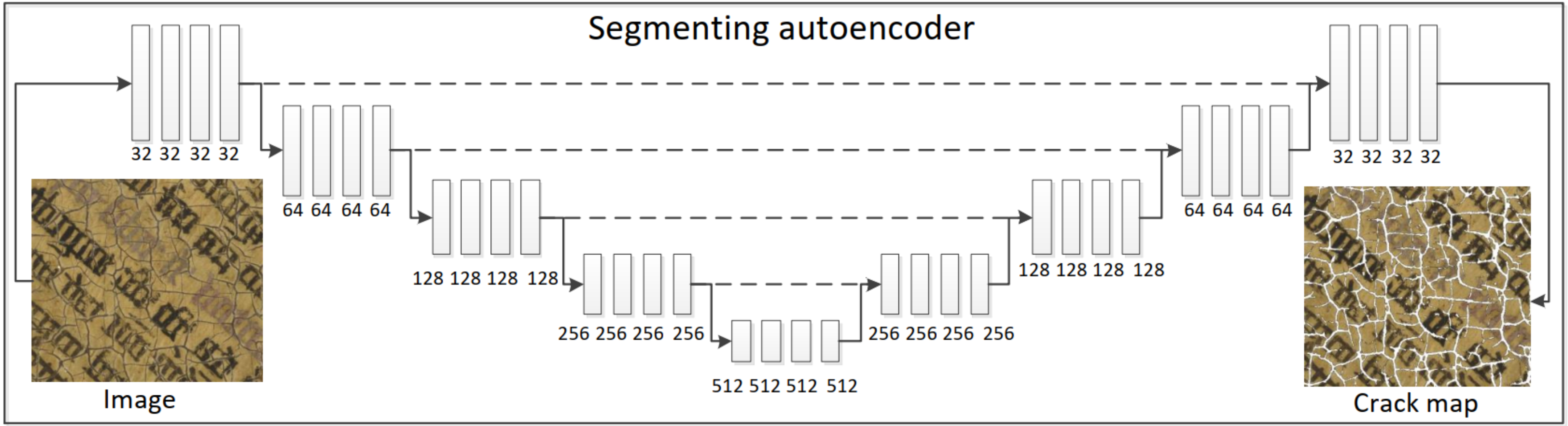
Simple background



Complex background



The proposed architecture of the U-Net based segmenting autoencoder.



Sorensen–Dice coefficient was used for loss estimation

$$Loss = \frac{2|x \cap d|}{x + d}$$

where *x* and *d* - is estimated and ground truth crack maps

All convolution layers use the exponentially linear unit (ELU)

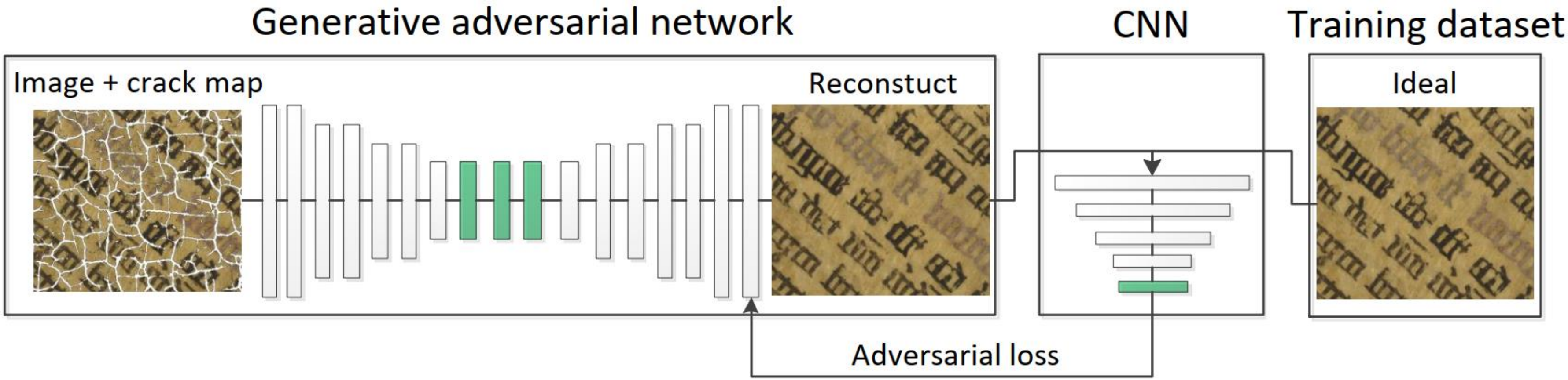
$$f(x) = \begin{cases} x & \text{if } x > 0 \\ a(e^x - 1) & \text{if } x \le 0, \end{cases}$$

where *a* is a hyperparameter

For training we use *Adam* optimization

For final layer was used logistic sigmoid activation function

Proposed GAN-based model for virtual restoration



Combined loss function was used

$$Loss_G = \lambda L_{abs} + L_{adv},$$

$$L_{abs} = |x_{trn} - G(x_{def})|,$$

$$L_{adv} = \mathbb{E}[\log(1 - D(G(x_{def})))]$$

where x_{trn} – ideal image, x_{def} - img. for reconstruction, λ is a hyperparameter

All convolution layers use the exponentially linear unit (ELU)

$$f(x) = \begin{cases} x & \text{if } x > 0 \\ a(e^x - 1) & \text{if } x \leq 0, \end{cases}$$

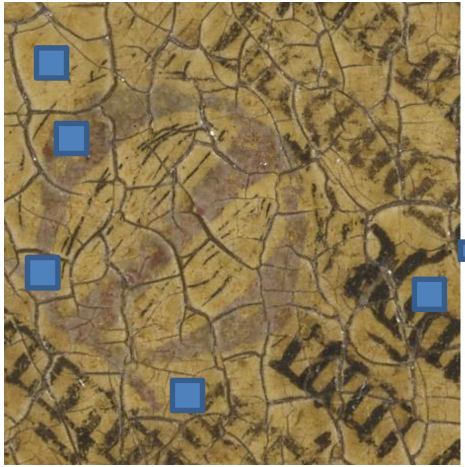
where a is a hyperparameter

For training we use *Adam* optimization

For final layer was used logistic sigmoid activation function

Workflow for virtual restoration

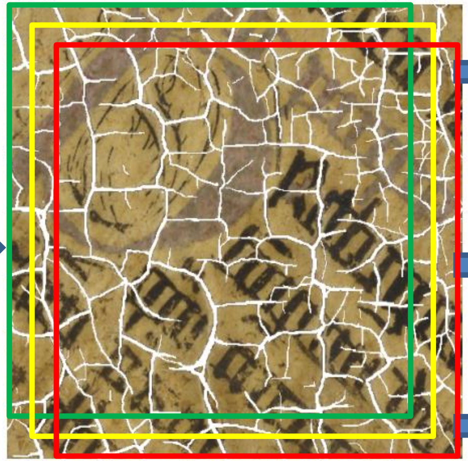
Forming dataset



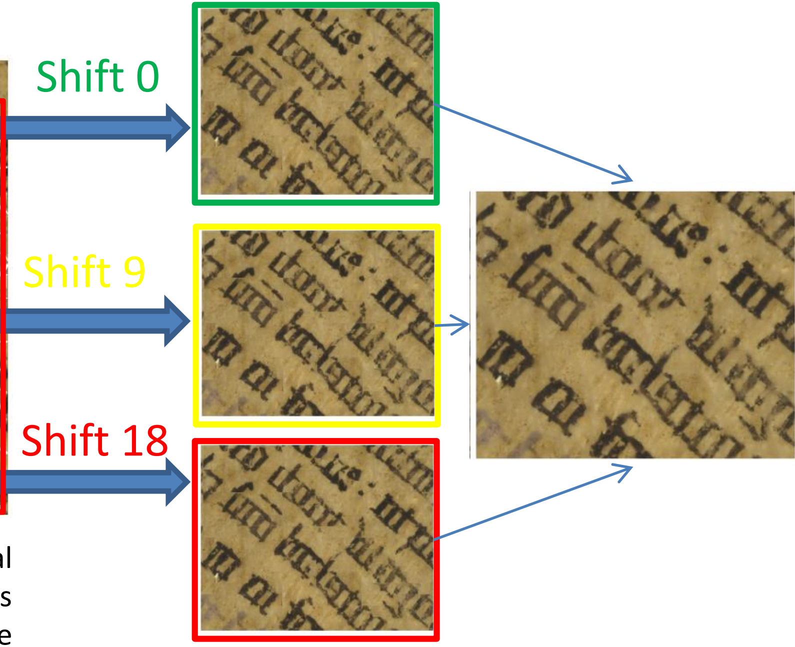
-Example of undamaged patch. This patches was used for training

Apply segmenting autoencoder for detecting crack

Reconstructing using crack map



We process the full image several times using a small shift of 3 pixels for each iteration of the restoration, using GAN-based model for virtual restoration. **This is important because some areas of the patch may be completely covered by the lost area**



After that, the 8 versions of the reconstructed images are combined into one using the **median filter**

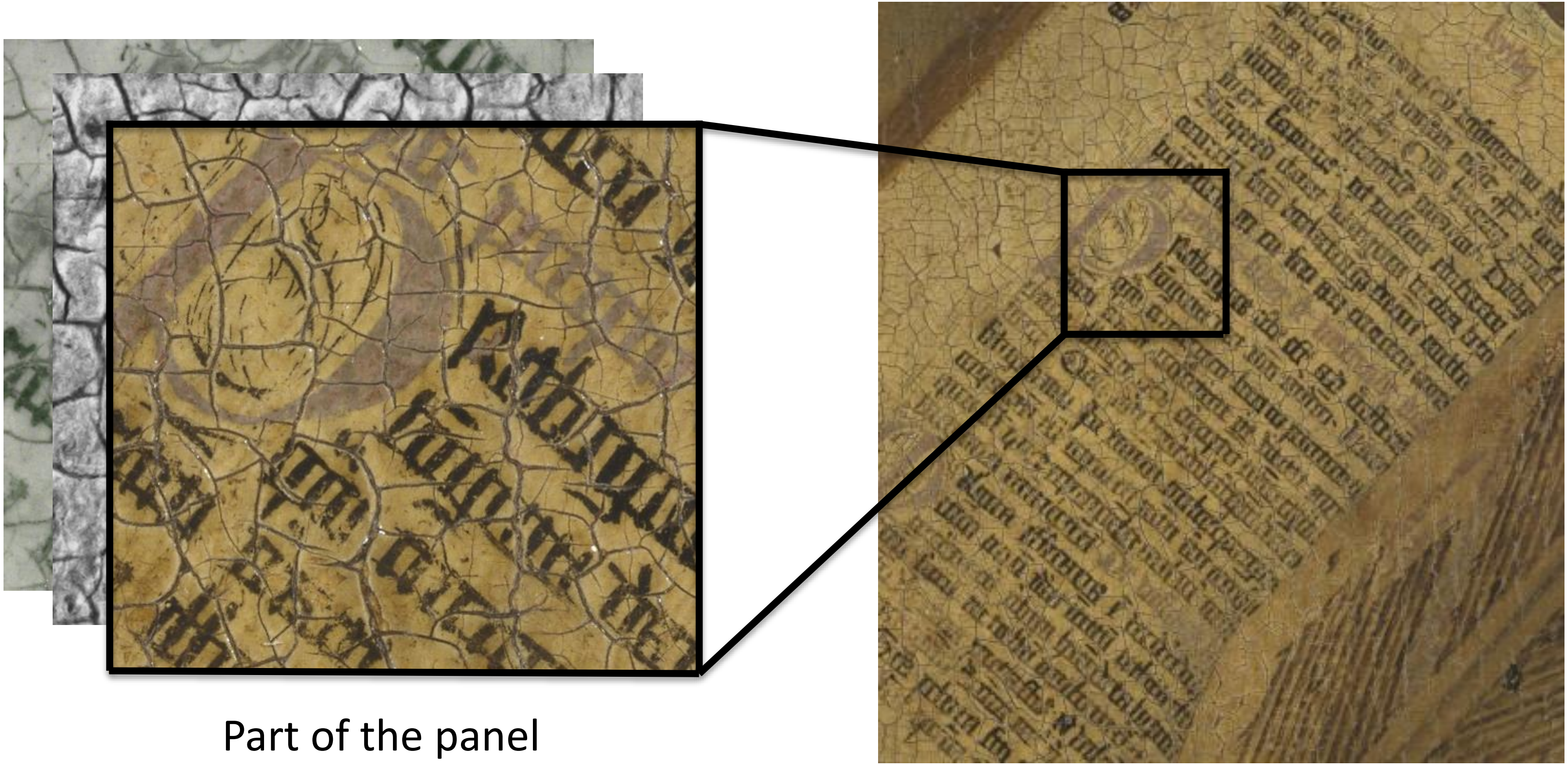
Experimental setup for crack detection

For comparison, we use the following well-known techniques for detecting cracks:

1 - Method with improved crack boundary localization (MCNC)[1],
2 - Bayesian Conditional Tensor Factorization method (BCTF) [2],
3 - CNN-based method that was proposed for crack detection in roads (CNN) [3]
4 - Deep feature fusion network classifer from (DFFN) [4].

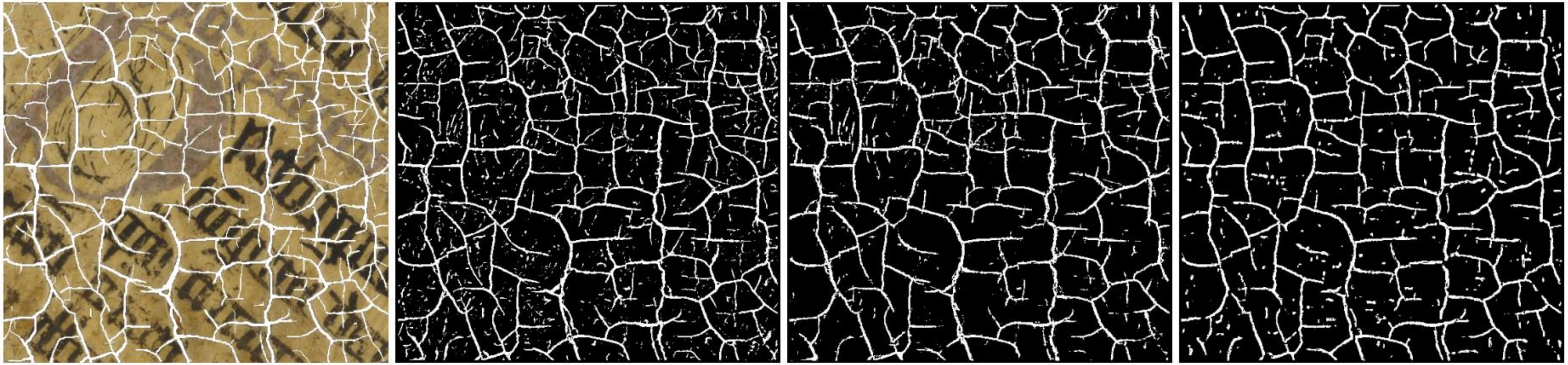
[1] Sizyakin, R., Cornelis, B., Meeus, L., Dubois, H., Martens, M., Voronin, V., and Pizurica, A., "Crack detection in paintings using convolutional neural networks," *IEEE Access* **8**, 74535–74552 (2020).
[2] Cornelis, B., Yang, Y., Vogelstein, J. T., Dooms, A., Daubechies, I., and Dunson, D. B., "Bayesian crack detection in ultra high resolution multimodal images of paintings," *IEEE, 18th International Conference on Digital Signal Processing* (2013)
[3] Lei, Z., Fan, Y., Yimin, D., and Ying, J. Z., "Road crack detection using deep convolutional neural network," *IEEE International Conference on Image Processing (ICIP)* , 3708–3712 (2016)
[4] Song, W., Li, S., Fang, L., and Lu, T., "Hyperspectral image classification with deep feature fusion network," *IEEE Transactions on Geoscience and Remote Sensing* **56**, 3173–3184 (2018).

Example multimodal data



Part of the panel
"Annunciation virgin Mary"

Illustration of the crack detection maps for part of the panel *Annunciation virgin Mary*



RGB + Ground Truth BCTF MCNC mUNET+C

Annunciation virgin Mary panel					
Method	Recall	False alar.	False miss.	Precision	F_1-m.
CNN	**0.8481**	0.0777	**0.1519**	0.5989	0.7020
DFFN	0.7488	0.0422	0.2512	0.7081	0.7279
BCTF	0.7896	0.0535	0.2104	0.6686	0.7241
MCN	0.8161	0.0540	0.1839	0.6741	0.7383
MCNC	0.7673	**0.0375**	0.2327	**0.7365**	**0.7516**
mUNET	0.8034	0.0702	0.1966	0.6101	0.6935
mUNET+C	0.7437	0.0417	0.2563	0.7090	0.7259

Experimental results for the first test image

Example multimodal data

Part of the panel “Singing Angels"

Example labeling data for training

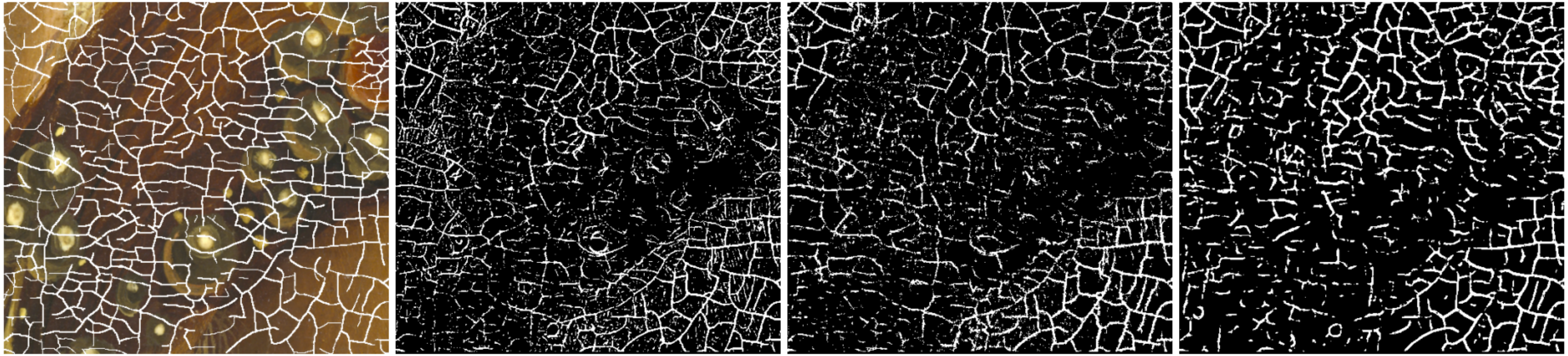
Existing training dataset

Using an incompletely labeling dataset, the autoencoder may not converge

Extending training dataset

For solving this limitation was used extending training dataset

Illustration of the crack detection maps for part of the panel *Singing Angels*



RGB + Ground Truth | BCTF | MCNC | mUNET+C

Existing training dataset

Extending training dataset

Singing angels panel					
Method	Recall	False alar.	False miss.	Precision	F_1-m.
CNN	0.6119	0.0999	0.3881	0.4680	0.5304
DFFN	0.6242	0.0966	0.3758	0.4814	0.5436
BCTF	0.6150	0.0905	0.3850	0.4941	0.5479
MCN	0.6340	0.0894	0.3660	0.5048	0.5621
MCNC	0.6083	**0.0681**	0.3917	**0.5622**	**0.5843**
mUNET	-	-	-	-	-
mUNET*	**0.6441**	0.1104	**0.3559**	0.4559	0.5339
mUNET+C*	0.6140	0.0916	0.3860	0.4905	0.5454

Experimental results for the second test image

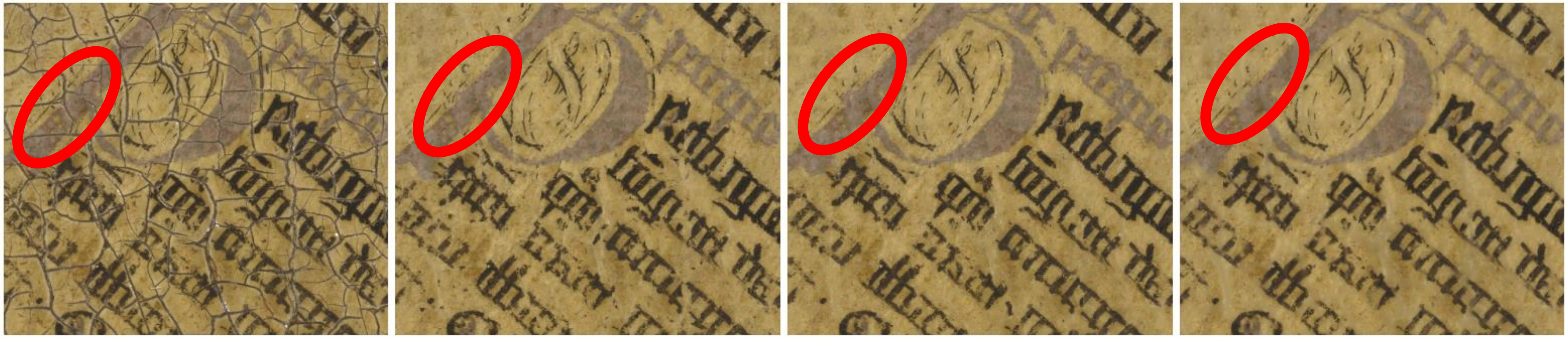
Experimental setup for inpainting

For comparison, we use the following well-known techniques for removing cracks:

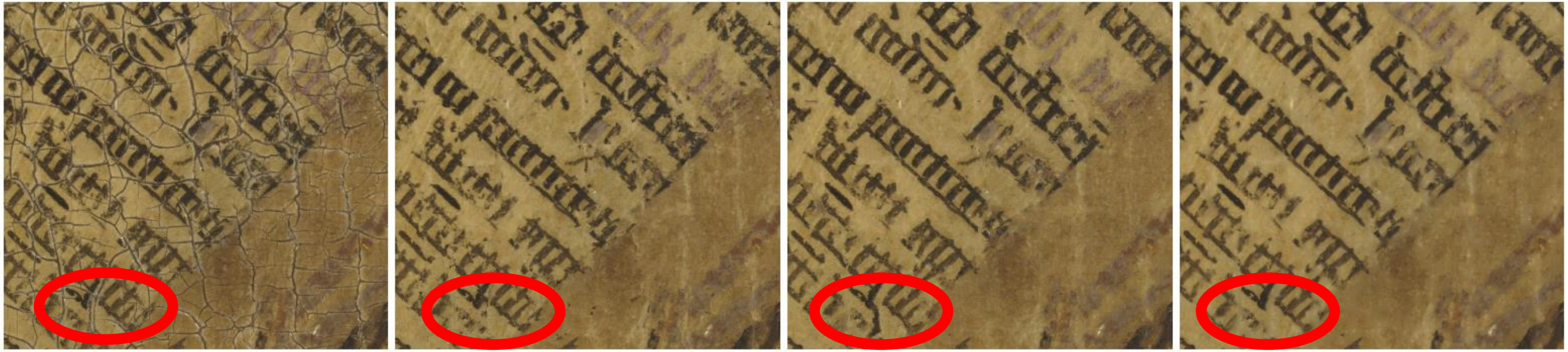
1 - Exemplar-based image inpainting (EBM)[5]
2 - Context-aware patch-based image inpainting using MRF modeling (CA-MRF) [6]

[1] Criminisi, A., Perez, P., and Toyama, K., "Region filling and object removal by exemplar-based image inpainting," *IEEE Transactions on Image Processing* , 1200–1212 (2004)
[2] Ruzic, T. and Pizurica, A., "Context-aware patch-based image inpainting using Markov random field modeling," *IEEE Transactions on Image Processing* , 444–456 (2015)

Illustration of the crack detection maps for part of the panel *Singing Angels*



RGB EBM CA-MRF Prop. GAN



RGB EBM CA-MRF Prop. GAN

Conclusion

- Was designed a two deep neural network architecture for crack detecting and removing
- Designed extended version of the segmenting autoencoder U-Net based has high resistance for noise-like painted objects.
- An additional important advantage with respect to earlier CNN methods based is that there is no excessive, false thickening of the boundaries of the detected cracks, if provided very accurate labeling of training data
- If such a high-quality training data set is not available it is still necessary to use techniques to refine the boundaries of the detected cracks
- The obtained results confirm the high efficiency of the designed architecture of the GAN-based network and the proposed training method. The results of inpainting appear visually consistent and better than the results of patch-based methods that were earlier used to restore digitized paintings

Thank you!