Robustness Study of Edge Segmentation and Completion by CNN Based Method for Tessellation Images

Valentin Penaud-Polge, Yann Gavet

HAL Id: emse-03284617
https://hal-emse.ccsd.cnrs.fr/emse-03284617

Submitted on 3 Feb 2022

Robustness Study of Edge Segmentation and Completion by CNN Based Method for Tessellation Images

Valentin Penaud--Polge
Ecole Nationale Superieure des Mines de Saint-Etienne
Saint-Etienne, France
valentin.penaudpolge@etu.emse.fr

Yann Gavet
Ecole Nationale Superieure des Mines de Saint-Etienne
Saint-Etienne, France
gavet@emse.fr

Abstract—In this paper, the robustness of tessellation images segmentation using Convolutional Neural Network (CNN) is presented. Particularly, this paper aims to study the effect of the quality and number of the images used for training on the robustness of the edge segmentation and completion. Three kinds of image deterioration are considered : the absence of a part of the tessellation, the discontinuity of the edges constituting the tessellation and the presence of a background noise. Five CNNs are trained with several kinds of deteriorated images and the trained models are then compared according to the results obtained with four classes of images.

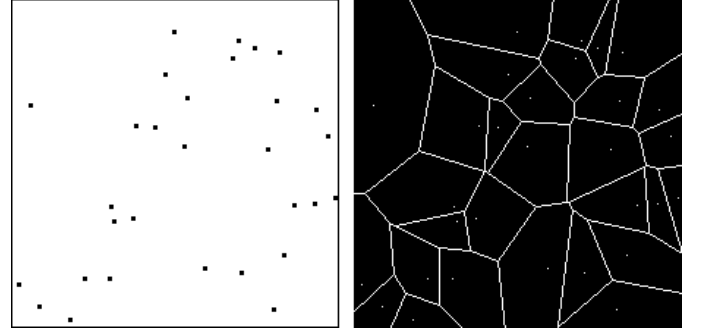*Index Terms*—Tessellation, Cell Segmentation, Edge Completion

I. INTRODUCTION

Segmentation of tessellation or mosaic images can often arise in biological or industrial problems, such as the segmentation of corneal endothelium images [1]. Many articles have been proposed to deal with this issue: using watershed [2], active contours [3] or directional filters [4]. With the recent breakthrough of deep learning methods, several solutions have been studied using neural networks and CNN [5] [6] [7], even though, a common problem that occurs using CNN is the lack of a sufficient number of images and their relatively poor quality. The present paper constitutes a study of the robustness of the segmentation, based on CNN, of tessellation images depending on the quality and the number of images used to train a deep learning model. In order to control the quantity and the diversity (number of cells, deterioration) of images used, the tessellations are synthetic and computed with a Poisson-Voronoi model (Poisson Point Process (PPP) coupled with a Voronoi Diagram). By being spatially completely random, the PPP offers the most diversified Voronoi Diagrams possible.

The present paper is organised as follows. Sect. II presents the tessellation model used to generate the images. Then, Sect. III introduces three different ways of deteriorating the images. Sect. IV offers details about the experimentation procedures. Sect. V presents the results obtained. And finally Sect. VI concludes this study and offers new perspectives.

Fig. 1: Left : Result of a PPP. Right : Voronoi Diagram associated

II. TESSELLATION MODEL

A tessellation is a partition of the plane, defined by a set of points (the seeds) and a distance (generally the Euclidean distance). The set of points is randomly generated by the realization of a PPP, and the partition is defined as a Voronoi diagram on these seeds.

A. Poisson Point Process

The PPP is a fast and simple way to generate seeds for a Voronoi Diagram. The PPP uniformly generates $n_{\lambda S}$ random points $\{p_i\}$ in the spatial domain of the subset $D \subset \mathbb{R}^2$. $n_{\lambda S}$ comes from a Poisson law $\mathcal{P}$ with parameter λS, where S represents the Lebesgue measure of D in $\mathbb{R}^2$ (area) and λ is a density parameter.

$$n_{\lambda S} \sim \mathcal{P}\left(\lambda S\right) \quad (1)$$

For $D = [a,b] \times [c,d]$ with $a, b, c, b \in \mathbb{R}$:

$$\forall i \in [\![1; n_{\lambda S}]\!] \ : \ p_{ix} \sim \mathcal{U}\left([a,b]\right) \ and \ p_{iy} \sim \mathcal{U}\left([c,d]\right) \quad (2)$$

Fig. 1 shows on the left a realization of the PPP with $\lambda = 0.0005$ and $S = 224 \ pixel^2$.

B. Voronoi Diagram

A Voronoi Diagram of a set of points $\{p_i\}$ in a subset $D \subset \mathbb{R}^2$ with regard to a distance d corresponds to the partitioning of D into cells $\{C_i\}$, defined as followed :

$$C_i = \{x \in D \mid d(x, p_i) \le d(x, p_j) \ \forall j \ne i\} \qquad (3)$$

The Voronoi Diagram, as the one in Fig 1, is the intersection of all the cells.

$$Vor(\{p_i\}) = \bigcup_{i,j \ ; \ i \ne j} C_i \cap C_j \qquad (4)$$

For more details about Voronoi Diagrams, the reader can refer to [8].

III. DETERIORATION MODELS

The following section describes the three alterations applied on the images used in order to train the models. We group these three alterations into two kinds of deterioration: a deterioration concerning the entire image, which is presented in Sect. III-A, to study the robustness with regard to an additionnal noise and deteriorations concerning only the edges of the cells, described in Sect. III-B and III-C, in order to study the robustness concerning the lack of edges information and the capacity of the trained models to close edges [1].

A. Gaussian Random Field Background

A Gaussian Random Field (GRF) is a specific Random Field [9] defined by the following covariance:

$$\forall x, y \in \mathbb{R}^2 \ C(x, y) = C(\tau) = e^{-\frac{\tau^2}{2\sigma^2}} \ ; \ \tau = ||x - y||_{\mathbb{R}^2} \quad (5)$$

This defines a continuous and stationary Random Field. A noisy background based on GRF can be applied to a given tessellation image $T_{Vor(\{p_i\})}$ built from a Voronoi Digram $Vor(\{p_i\})$:

$$T_{Vor(\{p_i\})} : D \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$$
$$x \longmapsto \begin{cases} 1 \ if \ x \in Vor(\{p_i\}) \\ 0 \ otherwise \end{cases} \qquad (6)$$

The resulting image $T_{Vor(\{p_i\})}^{GRF}$ is defined as:

$$T_{Vor(\{p_i\})}^{GRF} = \mathcal{F}^{-1}\left(\sqrt{\mathcal{F}(C_0)\mathcal{F}(S)}\right)$$
$$S = W + snr \times \left(1 - T_{Vor(\{p_i\})}\right) \qquad (7)$$

With $C_0 : x \in \mathbb{R}^2 \mapsto C(0_2, x)$. $\mathcal{F}$ corresponds to the Fourier Transform, σ represents the standard deviation, 0_2 is the neutral element of $\mathbb{R}^2$ for the usual addition, W is a white noise and snr is the signal-to-noise ratio.

For all images used with a GRF background deterioration σ follows a Poisson law of parameter 4 and snr follows a Poisson law with parameter 10: $\sigma \sim \mathcal{P}(4)$ and $snr \sim \mathcal{P}(10)$.

The aim of adding a GRF background is to simulate a potential texture of the cells. Fig. 2 represents the result (bottom left) of a tessellation (top left) immersed into a GRF (top right). An example of a real corneal endothelium image is given as a comparison (bottom right).
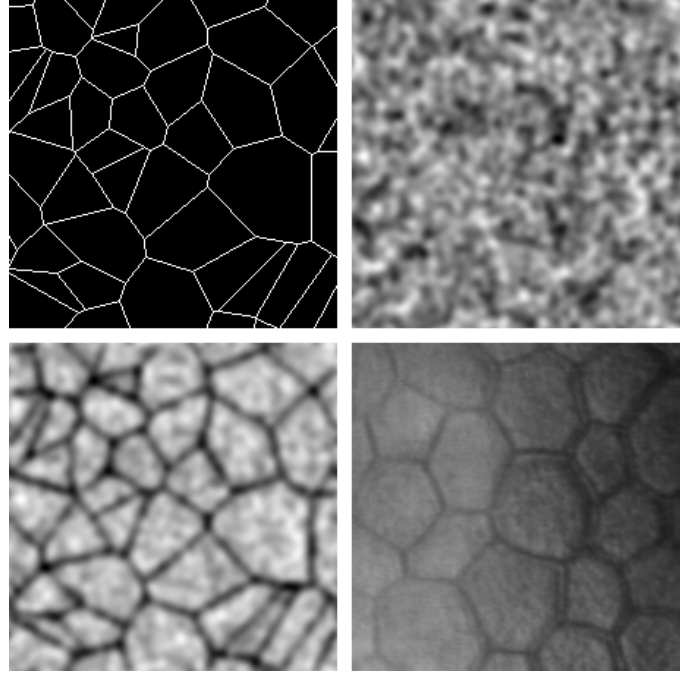

Fig. 2: Top Left: Tessellation. Top Right: Gaussian Random Field. Bottom Left: Tessellation with the Gaussian Random Field background. Bottom Right: Endothelium Corneal Image

B. Boolean Model Holes

This degradation consists in removing the cell edges which intersect a Boolean Model [10] realization whose germs are computed by a PPP with parameter $\lambda_{BM}S$ and whose grains are disks. The radius of each disk is uniformly chosen between two extremum values R_{min} and R_{max}. Given a Voronoi Diagram $Vor(\{p_i\})$, and a boolean model BM, the result $Vor_{BM}(\{p_i\})$ is defined as follows:

$$Vor_{BM}(\{p_i\}) = Vor(\{p_i\}) \cap BM^c \qquad (8)$$

$.^c$ represents the complementary.

An example of a resulting image is showed in Fig. 3 (bottom images), before and after being embedded in a GRF respectively.

C. Density Reduction

Contrary to the holes created by using the Boolean Model Holes (BMH) alteration which concentrates the loss of information in few areas, the Density Reduction (DR) aims to remove uniformly the edges information. Each point of the edges is removed or kept depending on the realization of a Bernoulli law $\mathcal{B}$ of parameter q. We define as $Vor_q(\{p_i\})$ the resulting set.

$$Vor_q(\{p_i\}) = \{x \in Vor(\{p_i\}) \mid u_x = 1 \ ; \ u_x \sim \mathcal{B}(q)\}$$
$$u \sim \mathcal{B}(q) \Leftrightarrow P(u = k) = q^k (1 - q)^{1-k} \ ; \ k \in \{0, 1\} \qquad (9)$$
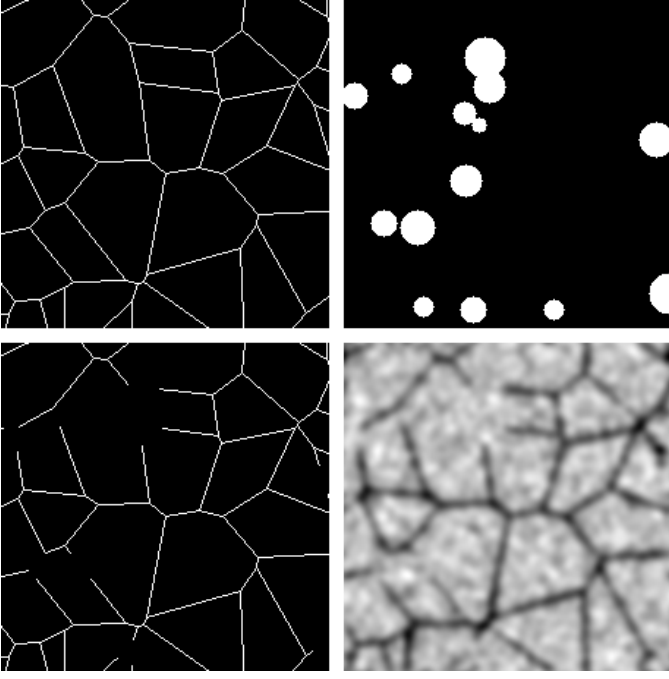
Fig. 3: Top Left: Tessellation. Top Right: Boolean Model. Bottom Left: AND(Tessellation, NO(Boolean Model)). Bottom Right: Resulting Image
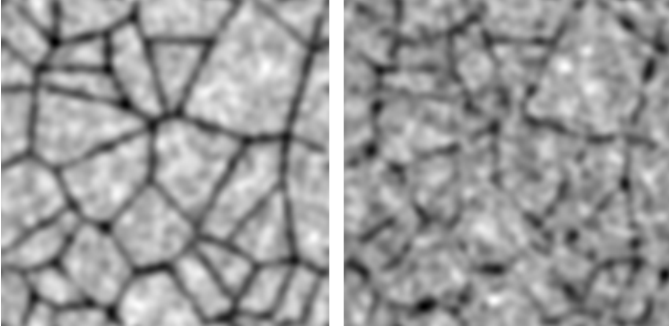

Fig. 4: Left: $Vor\left(\{p_i\}\right)$ with a GRF background. Right: $Vor_{q=0.5}\left(\{p_i\}\right)$ with a GRF background

An example is given in Fig. 4. The image on the left (resp. on the right) corresponds to $Vor\left(\{p_i\}\right)$ (resp. $Vor_{q=0.5}\left(\{p_i\}\right)$) immersed in a GRF.

IV. EXPERIMENTATION

A. Experimental Design

The notations used are the same as in the previous sections. All the images have a size $(M, N) = (224, 224)$ with $S = MN$.

Five CNNs have been trained with different datasets descibed below. The architecture is the same for each CNN and is the U-net [11] based architecture inspired from [6], see Fig. 5.

All datasets are generated with 3 different values of λ : 0.0002, 0.0008, 0.0016 equally split during the image generation.
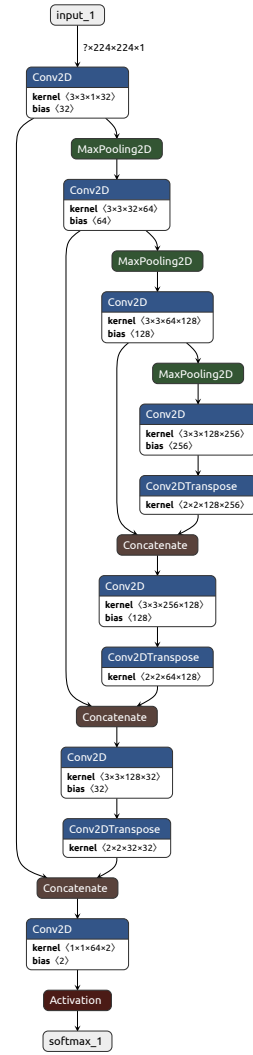

Fig. 5: U-net architecture.

A morphological dilation was applied on each ground truth as better results were observed when using thicker ground truth edges.

- BIN-BMH: The BIN-BMH dataset is composed of 3000 images. The only deterioration used is the BMH process with two similarly divided sets of parameters: $\lambda_{BM}S = 2, R_{min} = 20, R_{max} = 30$ and $\lambda_{BM}S = 15, R_{min} = 10, R_{max} = 15$. As there is no background alteration, the images of this dataset are binary images.
- GRF: The GRF dataset is composed of 3000 images only deteriorated with a GRF background.
- GRF-BMH-3k: This dataset contains 3000 images deteriorated with a GRF background and the BMH using the same two sets described in the BIN-BMH dataset.
- GRF-BMH-9k: This dataset contains the same kind of images used in the GRF-BMH-3k dataset, the difference is that it is composed of 9000 images.
- GRF-BMH-DR: All kinds of deterioration have been used to compute this last dataset. All the images have

a GRF background. 3000 images are only altered by a background, 3000 others are images corresponding to the same kind of images as in the GRF-BMH-3K dataset and the last 3000 other images are deteriorated with a GRF background and a DR process : the following values of q have been evenly used : 0,25 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7.

B. Evaluation

1) Images: In order to evaluate the robustness of the five trained CNN, four classes of images are used. Two of them are computed in the same way as the images used to train the CNN models, but were not used for the training process. The two others are used to determine the ability of the models to close edges.

- EV-GRF-DR: This class of images represents a tessellation immersed in a GRF background, iteratively deteriorated by a DR process. The corresponding values of q go from 0.1 to 0.9 with a step of 0.01 between each image. The background is the same for the 81 images. Fig. 6 shows an example of tessellation for $q = 0.3$, $q = 0.7$ and the associated ground truth. Results presented in Sect. V represent the mean values calculated over 100 iterations of this class.
- EV-GRF-BMH: The images of this class correspond to a tessellation iteratively altered by an BMH deterioration with disks of radius $r = 10$. Each image corresponds to a remaining proportion of edges information. The proportions go from 0 to 0.95 with a step of 0.05 between each image. The background is the same for the 20 images. Fig. 7 represents the image with 90% of the initial edges information, the image with 15% of the initial edges information and the corresponding ground truth. Results presented in Sect. V represent the mean values calculated over 100 iterations of this class.
- BIN-GI: The binary grid images class represents a grid with iteratively growing holes at the intersection points. The holes are disks with radius going from 0 to 15 pixels, with a step of 1 between each image. The length of a side of one square of the grid in the initial situation (before any holes) is 32 pixels. Fig. 8 shows the grid with disk holes having respectively 3 and 13 as radius and the initial grid used as the ground truth.
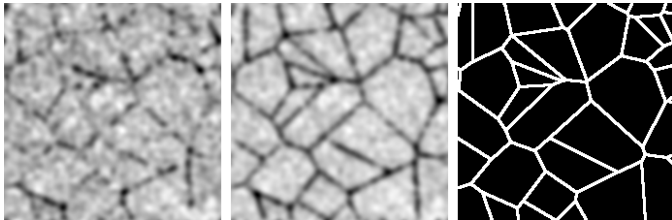- BIN-KAN: This last class of images represents binary Kanizsa triangle patterns [12] as it is a reference in
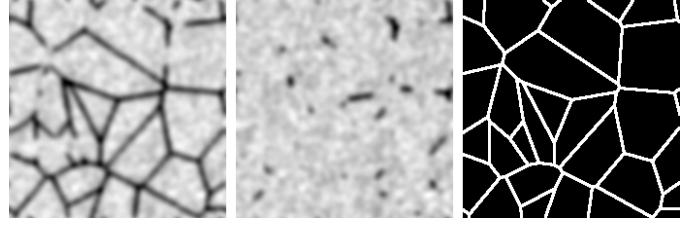

Fig. 7: Left: Tessellation with 90% of edges information. Center: Tessellation with 15% of edges information. Right: Tessellation ground truth.

continuation pattern. The images correspond to the three vertices of an equilateral triangle iteratively moved away by steps of 2 pixels. The initial edge size of the triangle is 80 pixels. Each vertex is moved with its two 40 pixels edges. Fig. 8 displays the patterns having a gap of 40 pixels and 80 pixels with their correponding ground truths. This last class is composed of 71 images : from the equilateral triangle to a pattern with a 140 pixels gap.

2) The ϵ Dissimilarity Criterion: The criterion used to measure the segmentation accuracy is the criterion presented in [13] [14], its definition is the following. A simple definition of this criterion would be that it corresponds to a symetric difference with a tolerance introduced be the Minkowski addition of a ball of radius ρ.

$$\epsilon_G^\rho (X) = \frac{\#\{(X \backslash G \oplus \mathcal{B}(0, \rho)) \cup (G \backslash X \oplus \mathcal{B}(0, \rho))\}}{\#\{G \oplus \mathcal{B}(0, \rho)\}} \quad (10)$$

With #, $\oplus$, $\mathcal{B}(0, \rho)$, G and X representing respectively the cardinal sign, the Minkowski addition, the disk of center 0 and radius ρ (the tolerance) as a structural element for the Minkowski addition, the ground truth and the predicted segmentation.

As all the ground truths in the training datasets were dilated, the ground truths used in Sect. V were morphologically dilated before computing the values of ϵ. The value of the tolerance ρ is fixed to 1.5 for the rest of the paper.

V. RESULTS

Fig. 10 and Tab. I summarize the results obtained.

In the rest of the paper, each trained model holds the name of its training dataset. The BIN-BMH (respectively GRF, GRF-BMH-3k, GRF-BMH-9k and GRF-BMH-DR) model
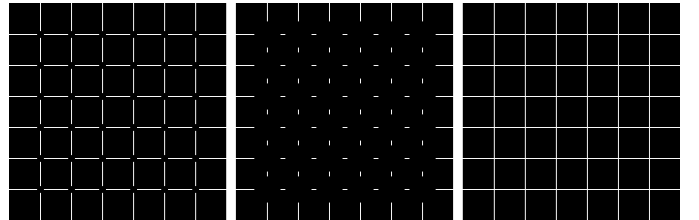

Fig. 6: Left: Tessellation for $q = 0.3$. Center: Tessellation for $q = 0.7$. Right: Tessellation ground truth.

Fig. 8: Left: Grid with disk holes of radius 3. Center: Grid with disk holes of radius 13 Right: Grid ground truth.
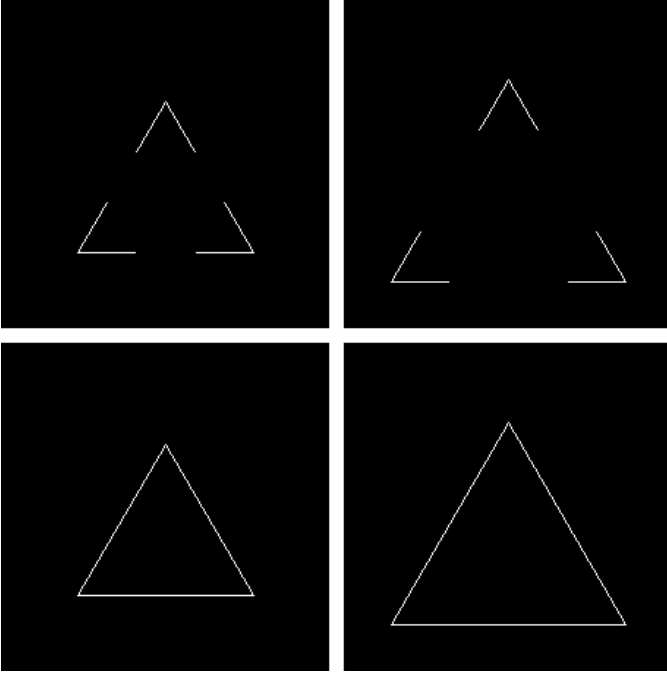
Fig. 9: Top Left: Kanizsa motif with a gap of 40 pixels. Top Right: Kanizsa motif with a gap of 80 pixels. Bottom Left: Ground truth of the Top Left images. Bottom Right: Ground truth of the Top Right image.

Fig. 10: (a): Results of the EV-GRF-DR images. y-axis: ϵ, x-axis: q.
(b): Results of the EV-GRF-BMH images. y-axis: ϵ, x-axis: Proportion of remaining edges information.
(c): Results of the BIN-GI images. y-axis: ϵ, x-axis: Radius of the disk holes.
(d): Results of the BIN-KAN images. y-label: ϵ, x-label: Size of the gap in number of pixels

corresponds to the model trained on the BIN-BMH (respectively GRF, GRF-BMH-3k, GRF-BMH-9k and GRF-BMH-DR) dataset .

- EV-GRF-DR (Fig. 10-a): The results highlight an important difference between the BIN-BMH model and the four others, as it has been trained without images containing a GRF background. The best results are obtained by the GRF-BMH-DR model which is the only model trained with images deteriorated by a DR process. Nevertheless, the gain obtained by training the model with a DR deterioration is smaller than the gain obtained by using a GRF background.
- EV-GRF-BMH (Fig. 10-b): The same comment done for EV-GRF-DR images is observable: the BIN-BMH model has poor results compared to the other models. As the the GRF-BMH-9k model has been trained with the most important number of images deteriorated by a BMH process, it presents the best results. Moreover, the fact that the GRF-BMH-DR model has better results than the GRF-BMH-3k model highlights the importance of the quantity of images.
- BIN-GI (Fig. 10-c): As it has been trained with binary images with BMH deterioration, the results given by the BIN-BMH model are better than the other models, as it can be seen in Fig. 11, only the two most difficult images are not properly completed by the BIN-BMH model. The GRF model which was not trained with images deteriorated by a BMH deterioration gives the
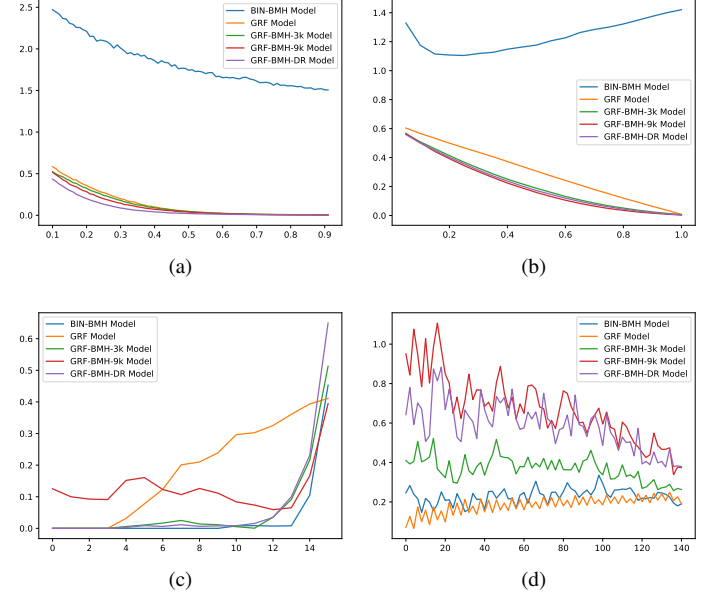
worst results (Fig 11). Finally, as it can be observed in Fig. 11, even though the GRF-BMH-9k properly extends the grid, it obtains worst results than GRF-BMH-3k as it adds false positive pixels.
- BIN-KAN (Fig. 10-d): For this class of images too, the BIN-BMH model gives good results (see Fig. 11), for the same reasons as the ones described for the BIN-GI images. But by adding barbules, the BIN-BMH model obtains scores close to the GRF model which only segments the existing edges (Fig. 11).
  Here again, by the important number of false positive pixels, the GRF-BMH-3k, the GRF-BMH-9k and the GRF-BMH-DR models obtain worst results than the two other models.

Images \ Models	BIN-BMH	GRF	GRF-BMH-3k	GRF-BMH-9k	GRF-BMH-DR
EV-GRF-DR	✗	✓	✓	✓	✓
EV-GRF-BMH	✗	∼	✓	✓	✓
BIN-GI	✓	✗	✓	∼	✓
BIN-KAN	✓	✗	∼	∼	∼

TABLE I: Summarized Results

VI. CONCLUSION

Results showed that a CNN can either learn accurately a segmentation or a completion task if the training dataset is
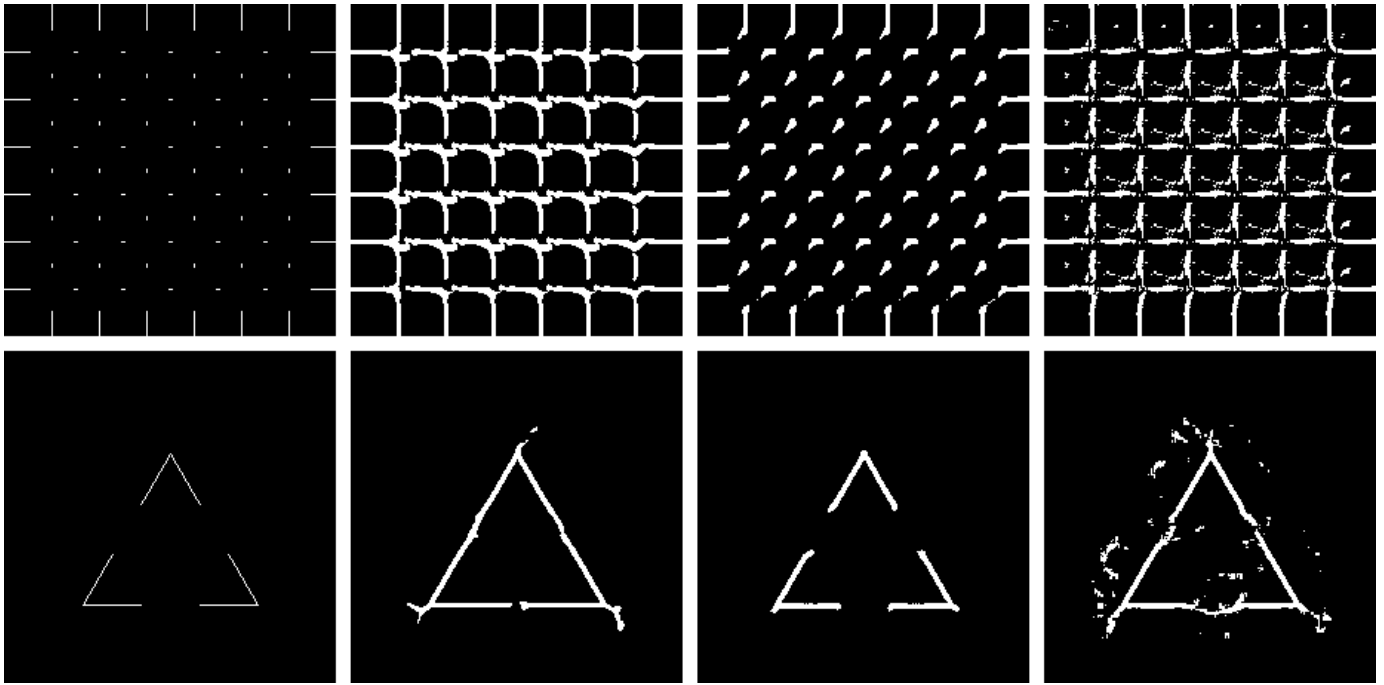
Fig. 11: From top to bottom and from left to rigth: Image 1: Input image from the grid images. Image 2: Output of image 1 by the BIN-BMH model. Image 3: Output of image 1 by the GRF model. Image 4: Output of image 1 by the GRF-BMH-9k model. Image 5: Input image from the Kanizsa images. Image 6: Output of image 5 by the BIN-BMH model. Image 7: Output of image 5 by the GRF model. Image 8: Output of image 5 by the GRF-BMH-DR model.

well build, as it is the case for the BIN-BMH model for the BIN-GI class of images, but without being able to generalize in unknown situations (in the presence of a GRF background). The overall better results of the GRF-BMH-DR model over the GRF-BMH-9k model (except for EV-GRF-BMH) show that the diversity of images is an important factor. Moreover, the completion process tends to penalize the models by adding unwanted edges or isolated pixels. Finally, as the main aspects underlined are the presence of a GRF background and the lack of constraints for the completion process, a new perspective to this work would be to use numerical twin images adapted to a specific application domain to study the capacities of a CNN on real life tessellation images. A first approach could be to use specific shapes of mosaics for a particular application to constrain the completion process to a certain range of angles or situations. Indeed, the Voronoi Diagrams resulting from the realisations of the PPP used in Sect. II, present a high cells shape diversity, due to the lack of spatial interaction in this process.

REFERENCES

[1] Y. Gavet and J.-C. Pinoli, "Visual perception based automatic recognition of cell mosaics in human corneal endothelium microscopy images," *Image Anal Stereol*, vol. 27, pp. 53–61, March 2008.

[2] L. M. Vincent and B. R. Masters, "Morphological image processing and network analysis of cornea endothelial cell images," in *Image Algebra and Morphological Image Processing III*, vol. 1769. International Society for Optics and Photonics, 1992, pp. 212–226.

[3] K. Charłampowicz, D. Reska, and C. Bołdak, "Automatic segmentation of corneal endothelial cells using active contours," *Advances In Computer Science Research*, 2014.

[4] K. Habrat, M. Habrat, J. Gronkowska-Serafin, and A. Piórkowski, "Cell detection in corneal endothelial images using directional filters," in *Image Processing and Communications Challenges 7*. Springer, 2016, pp. 113–123.

[5] A. Fabijańska, "Automatic segmentation of corneal endothelial cells from microscopy images," *Biomedical Signal Processing and Control*, vol. 47, pp. 145–158, 2019.

[6] J. P. Vigueras-Guillén, B. Sari, S. F. Goes, H. G. Lemij, J. van Rooij, K. A. Vermeer, and L. J. van Vliet, "Fully convolutional architecture vs sliding-window cnn for corneal endothelium cell segmentation," *BMC Biomedical Engineering*, vol. 1, no. 1, p. 4, 2019.

[7] A. Fabijańska, "Segmentation of corneal endothelium images using a u-net-based convolutional neural network," *Artificial intelligence in medicine*, vol. 88, pp. 1–13, 2018.

[8] F. Aurenhammer and R. Klein, "Voronoi diagrams," *Handbook of computational geometry*, vol. 5, no. 10, pp. 201–290, 2000.

[9] R. J. Adler and J. E. Taylor, *Random fields and geometry*. Springer Science & Business Media, 2009.

[10] J. Serra, "The boolean model and random sets," in *Image Modeling*. Elsevier, 1981, pp. 343–370.

[11] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[12] G. Kanizsa, "Subjective contours," *Scientific American*, vol. 234, no. 4, pp. 48–53, 1976.

[13] Y. Gavet and J.-C. Pinoli, "A geometric dissimilarity criterion between jordan spatial mosaics. theoretical aspects and application to segmentation evaluation." *JMIV*, 2011, submitted.

[14] Y. Gavet, M. Fernandes, J. Debayle, and J.-C. Pinoli, "Dissimilarity criteria and their comparison for quantitative evaluation of image segmentation: application to human retina vessels," *Machine Vision and Applications*, vol. 25, no. 8, pp. 1953–1966, 2014. [Online]. Available: http://dx.doi.org/10.1007/s00138-014-0625-2