Fast process for classifying structural image pairs using Weibull parameters extracted from undersampled Local Dissimilarity Maps

Moustapha DIAW
CReSTIC EA 3804, 51097 Reims
Univ. de Reims Champagne-Ardenne
IUT Troyes, France
moustapha.diaw@univ-reims.fr

Agnès DELAHAIES
CReSTIC EA 3804, 51097 Reims
Univ. de Reims Champagne-Ardenne
IUT Troyes, France
agnes.delahaies@univ-reims.fr

Jérôme LANDRÉ
CReSTIC EA 3804, 51097 Reims
Univ. de Reims Champagne-Ardenne
IUT Troyes, France
jerome.landre@univ-reims.fr

Frédéric MORAIN-NICOLIER
CReSTIC EA 3804, 51097 Reims
Univ. de Reims Champagne-Ardenne
IUT Troyes, France
frederic.nicolier@univ-reims.fr

Florent RETRAINT
ICD-LM2S
Univ. de Technologie de Troyes
Troyes, France
florent.retraint@utt.fr

***Abstract*—In previous works, the Local Dissimilarity Map (LDM) was proposed to compare two binary and grayscale images. This measure is based on a Hausdorff distance, which allows to quantify locally the dissimilarities between images. In this paper, we proposed the two-parameter Weibull distribution to model the LDM and the undersampled LDMs for two structural images. To classify structural image pairs, we used the two parameters of Weibull distribution for each LDM as descriptors. They are relevant to discriminate image pairs into similar and dissimilar classes. Experiments were made on the MNIST image dataset and in our own old print image dataset. The results shown our approach is more accurate than the other measures in the literature.**

***Index Terms*—Local Dissimilarity Map, undersampled Local Dissimilarity Map, Two-parameter Weibull distribution, Binary classification.**

I. Introduction

In today's times, image comparison becomes essential in view of the number of images circulating on the internet. In this paper, we are interested in images containing only structures. Many techniques are used to compare structural images and to classify them. Among these, we can mention the structural similarity index SSIM [1], [2] and the feature similarity index FSIM [3] which are very well known and widely used in the field of image processing. These measures have given good results in recent years in image comparison. For the SSIM, the similarities are calculated through a local sliding window and the global image quality is evaluated by calculating the average SSIM index: the MSSIM [2]. Neither of the maps of these two measures of similarity was modeled by a statistical distribution. Indeed, the computation of these measures does not depend on the distance between the pixel values of the images to be compared, but on these parameters: the luminance and the contrast.

In this paper, we used a local measure of dissimilarity which depends on the distance of the pixels of images to be compared. This measure is the Local Dissimilarity Map (LDM). As the SSIM, the dissimilarities are calculated through a window: the parts of each image seen through this window are compared by a well-defined local measure of dissimilarity. The Local Dissimilarity Map is constructed from the euclidean distance. We can model it by a two-parameter Weibull distribution. In addition, the undersampled Local Dissimilarity Maps follow the same distribution. These parameters are very relevant for classifying the LDMs of similar and dissimilar images. The performances of the undersampled LDMs are compared to the MSSIM and FSIM similarity measures using the accuracy and the F1-score. They showed good performances compared to other classical methods.

The remainder of this paper is organized as follows. Section II presents the Local Dissimilarity Map for binary and grayscale images. Section III introduces the statistical model. In this section, we validated the model on the undersampled LDMs of two old print images. Section IV reports the experimental results. Finally, section V concludes the paper.

II. Local Dissimilarity Map

A. LDM for binary images

For binary images, A and B take their values in the set $\{0, 1\}$ and $\mathrm{LDM}(A, B)$ takes its values in $\mathbb{R}^+$. It corresponds to a map of distances between A and B.

Let A and B be two binary images and $p = (x, y)$ be the pixel at coordinates (x, y). The LDM is defined from $\mathbb{R}^2 \times$

$\mathbb{R}^2$ to $\mathbb{R}^2$ by [5] :

$$\mathrm{LDM}(A,B)(p) = 1_{A\triangle B}(p)\ \max(\mathrm{dt}_A(p), \mathrm{dt}_B(p)) \quad (1)$$

$$= |A(p) - B(p)|\ \max(\mathrm{dt}_A(p), \mathrm{dt}_B(p)), \quad (2)$$

where

$$1_{A\triangle B}(p) = \begin{cases} 1 & \text{if } p \in A \triangle B \\ 0 & \text{otherwise,} \end{cases}$$

with $A \triangle B$= $(A|B) \cup (B|A)$ and $\mathrm{dt}_A(p)$ is the euclidean distance transform of A, namely the distance between the p-pixel and the nearest non-zero pixel of A. The eq. (2) can be simplified for binary images [7] :

$$\mathrm{LDM}(A,B)(p) = B(p)\ dt_A(p) + A(p)\ dt_B(p). \quad (3)$$

The eq. (3) removes the $\max$ operator and the absolute value. Hence, it represents a great interest in the modeling of the undersampled Local Dissimilarity Maps.

B. LDM for grayscale images

In order to extend the LDM to grayscale images, in eq. (3), we could use the real-valued distance transform RVDT from Molchanov and Teran [6] instead of the classical distance transform in eq. (2) or (3). This distance computes real values using several thresholds on the image :

$$\mathrm{RVDT}(p,A) = \frac{1}{a-b}\sum_{i=a}^{b} d(p,A_i)\omega_i, \quad (4)$$

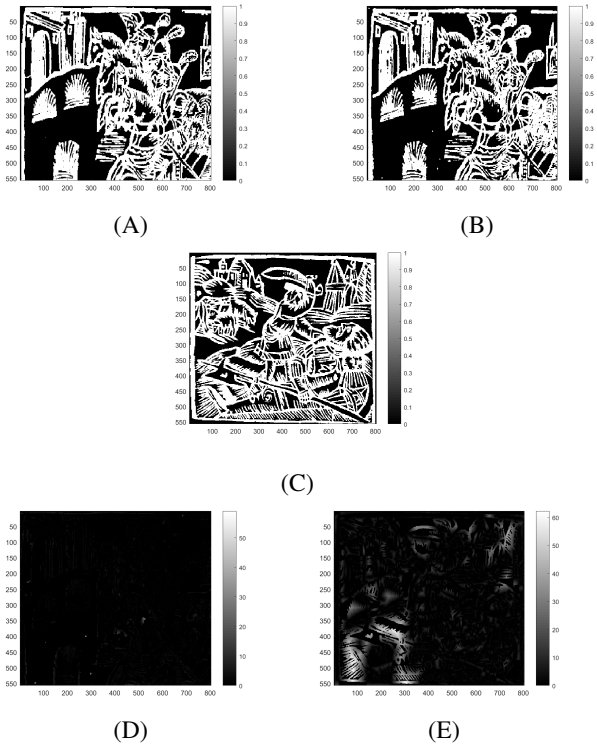
with $A_i = \{p : A(p) \geq i\}, 0 \leq a < b \leq 255$ and ω_i is a weight equal to 1 in the simplest case. A simple solution could be to use this RVDT directly in eq.(2). But in this case, distances between positions (x, y) and pixel values (luminance) are mixed together leading to interpretation and dynamic problems. Indeed, by mixing pure distances $(\max(\mathrm{dt}_A(p), \mathrm{dt}_B(p)))$ and luminance differences $(|A(p) - B(p)|)$, the distances obtained are no longer pure positional deviations between two pixels (p, q). We clearly want in this paper to keep true distances for the LDM, which will be important in the section III. One way to avoid this problem is to use the thresholding techniques of Molchanov by slicing grayscale images into several binary images and then compute simple multiple LDMs between these binary images.

So the Grayscale Local Dissimilarity Map (GLDM) is defined here as a sum of thresholded images:

$$\mathrm{GLDM}(A,B)(p) = \frac{1}{N}\sum_{i=1}^{N} \mathrm{LDM}(A_i,B_i)(p), \quad (5)$$

where N is the number of thresholds used in the sum, t_i is a threshold, $A_i = \{p \in A : A(p) \geq t_i\}$. Each threshold t_i is chosen according to a regular spacing between $m = \max(\min(A), \min(B))$ and $M = \min(\max(A), \max(B))$. So this regular spacing has a step of $s = (M - m)/N$.

The fig. 1 and fig. 2 show examples of LDMs for binary and grayscale images that come from the old print dataset. A, B and C are initially grayscale images and are binarised using the Sauvola's method [8]. The white regions of the LDMs represent large differences between the images while the black regions mean the pixels values are locally close.



(A) (B) (C) (D) (E)

Fig. 1: Binary images: (A), (B) and (C). Image (D): $\mathrm{LDM}(A,B)$, Image (E): $\mathrm{LDM}(A,C)$.

III. Statistical Model

The model used in this paper is the two-parameter Weibull distribution. The probability density and the cumulative distribution function of Weibull are :

$$f(t) = \frac{\alpha}{\beta}\left(\frac{t}{\beta}\right)^{\alpha-1} e^{-\left(\frac{t}{\beta}\right)^{\alpha}}, \quad (6)$$

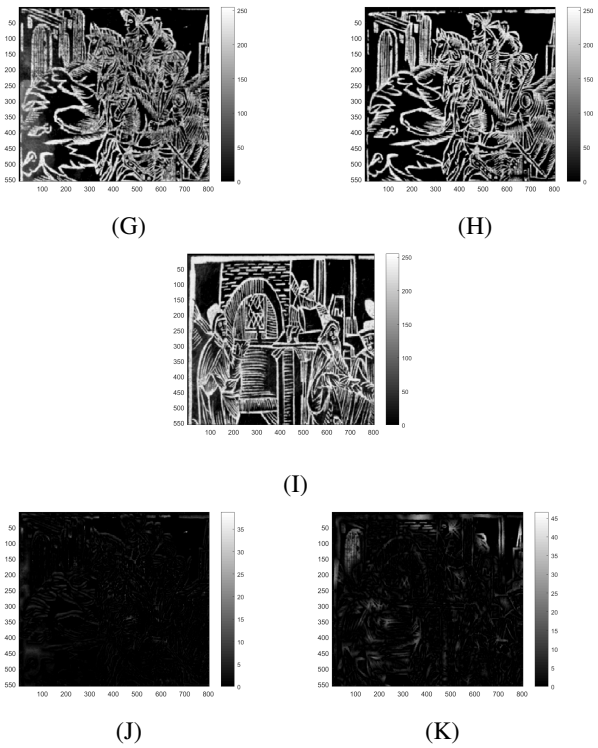
and

$$F(t) = 1 - e^{-\left(\frac{t}{\beta}\right)^{\alpha}}, \quad (7)$$

respectively, for $t > 0$, $\alpha > 0$ and $\beta > 0$. The shape parameter is α and the scale parameter is β.

According to eq. (3), modeling the Local Dissimilarity Map depends on the modeling of the euclidean distance transform of the two compared images. The theorem 1 [4] gives the distribution of the L_p norm and the L_2 norm for $p = 2$.

Theorem 1: For non-identical, correlated and upper-bounded random variables $X_i = |p_i - q_i|^p$, the random variable $\Delta = \sum_{i=1}^{N} X_i$, with finite N, is Weibull-distributed. The density of Δ is given by the eq. (6).



(G) (H)

(I)

(J) (K)

Fig. 2: Grayscale images: (G), (H) and (I). Image (J): $\mathrm{LDM}(G,H)$, Image (K): $\mathrm{LDM}(G,I)$.

p_i and q_i are respectively the vectors of the pixel values of the first image and the second image.
If Δ is Weibull-distributed with parameters α and β, then $\Delta^{1/p}$ is also Weibull-distributed with parameters $p\alpha$ and $\beta^{1/p}$ [11].

The pixel of a distance transform are non-identical and upper-bounded by the distance between the two vertices of the diagonal and the neighboring pixel values of the distance transform are correlated [11].

For $p = 2$ the euclidean distance transform and the Local Dissimilarity Map of binary images are Weibull-distributed (see eq. (3) and [11]). The eq. (3) and eq. (5) show that the LDM of grayscale images is a sum of euclidean distances, so it remains a distance. According the above, LDM of grayscale images follows a Weibull distribution.

IV. Experimental results

In this section, we modeled the undersampled LDMs of the two grayscale images ((G) and (H)) quoted in Fig. 2 by the Weibull distribution and show the relevance of the parameters $(\alpha;\beta)$ to discriminate images pairs to be compared into similar and dissimilar classes. A classification is made using K Nearest Neighbour. The KNN is one of the easiest supervised learning algorithms and gives good scores in the case of binary classification. The performances of the classifier is calculated by taking the accuracy and the F1-scores for both classes.

A. User datasets

In the applications, we used the well-known MNIST dataset [9] and the old print dataset [10]. The MNIST dataset contains $70,000$ grayscale images of size 28*28 (60,000 images for training and $10,000$ for testing). This dataset is composed of 10 classes numbered from 0 to 9. We will deal in our case with the problem of 2-class classification (class of similar image pairs and class of dissimilar image pairs). Unlike the MNIST dataset, the old print dataset contains only 64 images. When comparing images, each image must be compared with the rest of the images in the dataset to provide the LDM of each image pairs. Two images are similar if they belong to the same class. So we have a considerable number of LDMs (see table I). The storage contraint limits the number of usable images in the MNIST dataset. In the rest of the paper, we are random randomly selected $1,300$ images in the MNIST dataset (130 from each class) to avoid bias in the results and the storage contraints.

TABLE I: Number of images, classes and LDMs in each dataset.

Dataset	# of images used	# of classes	# of LDMs
MNIST	1 300	10	844 350
Old print	64	10	2 016

B. Model validation

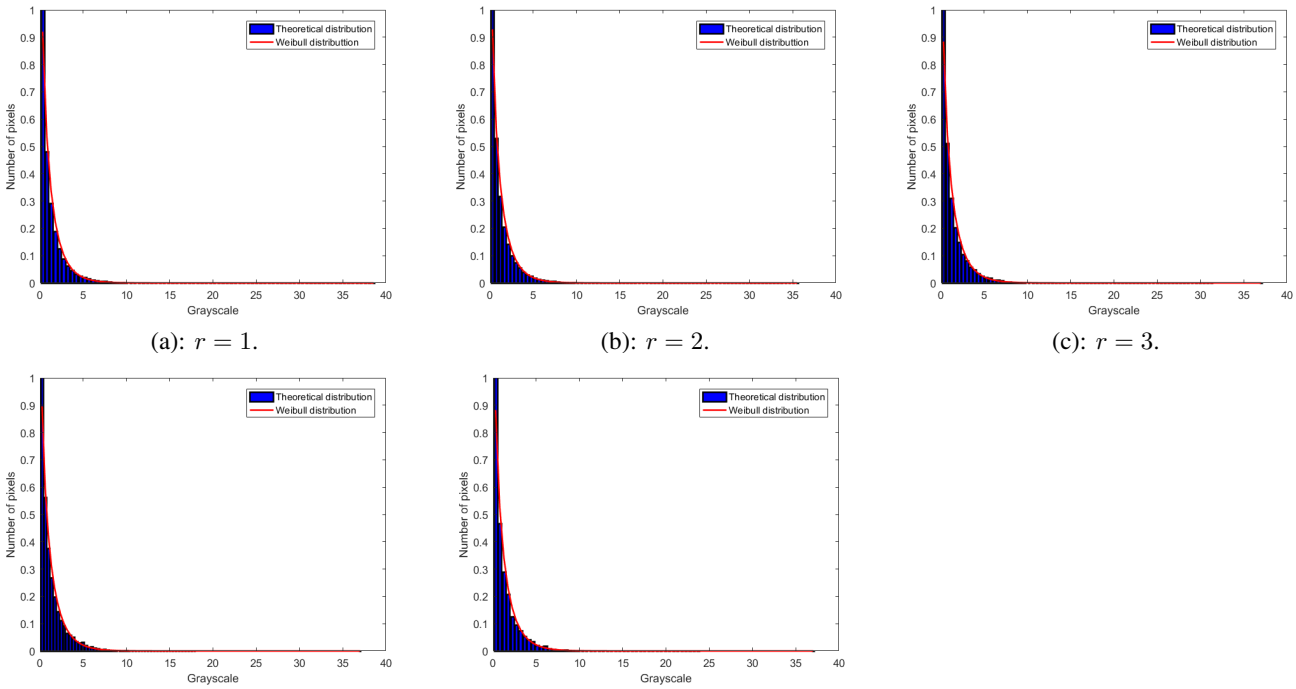
The Local Dissimilarity Map (Fig. 2: $\mathrm{LDM(G,H)}$) is undersampled with r steps, $i.e$ the sampling is done by taking 1 pixel on r in the $\mathrm{LDM(G,H)}$.
The Fig. 3 shows the histograms, fitted by a two-parameter Weibull distribution, of the undersampled LDMs. We have a very good fitting between the pixel values and the Weibull's law of parameters $(\alpha;\ \beta)$ which depend on r.

The table II shows the two parameters of Weibull for each undersampled LDM

TABLE II: The α and β parameters of the undersampled LDMs.

r	α parameter	β parameter
1	1.0870	0.8762
2	1.0886	0.8789
3	1.0831	0.8772
4	1.0835	0.8797
5	1.0878	0.8821

We can see that $\mathrm{LDM}(G,H)$ does not lose much information when we undersampled the LDM with steps of $r = 1, 2, ...$ (see table II), because the parameters of the distributions are approximately equals. Note that if $r = 1$, we have the original LDM (not undersampled).
The next section introduces the classification of image pairs, from the MNIST dataset and old print dataset, into similar and dissimilar classes using the two parameters of the distribution for each undersampled LDM.



(a): $r = 1$. (b): $r = 2$. (c): $r = 3$.

(d): $r = 4$. (e): $r = 5$.

Fig. 3: The histograms, fitted by a two-parameter Weibull distribution, of the undersampled LDMs. The fig. (a), (b), (c),(d) and (e) are respectively for the undersampled LDMs of step $r = 1, 2, 3, 4$ and 5 . For $r = 1$, it corresponds to the original LDM (LDM(G, H)). Red is the theorical distribution and blue the two-parameter Weibull distribution.

C. Classification

In this section the performances of the undersampled LDMs, the MSSIM and the FSIM are measured using the K Nearest Neighbour classifier.
The fig. 4 gives the accuracy of the undersampled LDMs as a function of the number of nearest neighbours K for the MNIST dataset. Performances are better when $K = 9$. Then in the rest of this paper, we take $K = 9$ to compare the classsification scores of LDMs with those of MSSIM and FSIM. The computation of $1,300$ images in the MNIST dataset gives $844,350$ LDMs, $83,850$ are for similar images and $760,500$ for dissimilar images. There is a very large class imbalance. The number of LDMs for dissimilar images is 9 times greater than for similar images. To balance the classes (similar class C_{sim} and dissimilar class C_{dissim}), we randomly undersampled the majority class C_{dissim}. Finally, we have $83,850$ LDMs in C_{sim} and C_{dissim}. The 10-cross validation is done on all data to give the classification scores. The same process was done for the undersampled LDMs. Since the size of images in the MNIST dataset is $28 * 28$, we limited the computation to $r = 3$. The two Weibull parameters extracted in each undersampled LDMs are used as descriptors and are relevant to discriminate the two classes. The fig.5 shows the whole process for classifying an image pair based on undersampled LDMs, MSSIM and FSIM.

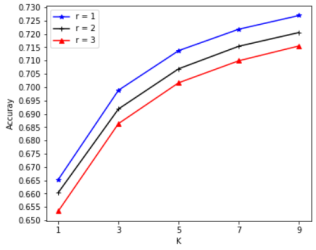


Fig. 4: Accuracy of the undersampled LDMs as a function of the number of nearest neighbours K of the MNIST dataset. Undersampling was done with r steps ($r = 1, 2, 3$).

The table III and table IV show the performances of the similarity measures in the MNIST and old print datasets.
On these two datasets, the undersampled LDMs achieved the highest scores (accuracy and F1-score). Furthermore, when the original LDMs are undersampled, the performances do not degrade significantly (see table III and table IV), $i.e$ they remain stable. Thus, to classify image pairs as similar and dissimilar, it is sufficient to undersample the LDMs by choosing a reasonable step (which depends on the images size) to save computing time. In table III and IV, the inference time is the time to extract the Weibull parameters (α; β) on

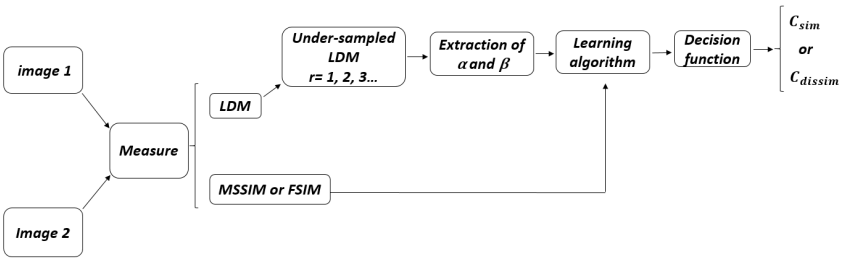


Fig. 5: Process of classifying an image pair image 1 and image 2 based on the undersampled LDMs, the MSSIM and FSIM. For $r = 1$, it corresponds to the original LDM.

TABLE III: Accuracy, F1-score and inference time of the undersampled LDMs, MSSIM and FSIM for the MNIST dataset.

Measures & Metrics		Accuracy	F1-score C_{sim}	F1-score C_{dissim}	Inference time (s)
Undersampled LDM	$r = 1$	0.727	0.714	0.739	564.7
	$r = 2$	0.722	0.709	0.733	400.43
	$r = 3$	0.715	0.702	0.726	376.5
MSSIM		0.613	0.588	0.636	177.6
FSIM		0.674	0.656	0.691	3568.8

TABLE IV: Accuracy, F1-score and inference time of the undersampled LDMs, MSSIM and FSIM for the old print dataset.

Measures & Metrics		Accuracy	F1-score C_{sim}	F1-score C_{dissim}	Inference time (s)
Undersampled LDM	$r = 1$	0.791	0.784	0.797	103.4
	$r = 2$	0.791	0.780	0.801	30.82
	$r = 3$	0.744	0.720	0.765	16.31
	$r = 4$	0.802	0.791	0.811	11.13
	$r = 5$	0.770	0.765	0.786	9.4
MSSIM		0.614	0.593	0.633	26.02
FSIM		0.640	0.631	0.649	720.9

each undersampled LDM and classification with KNN. For instance, if we choose $r = 5$ for the old print dataset, the inference time is at least 10 times faster than for $r = 1$ with comparable results. In addition, in the MNIST dataset, if we choose $r = 3$ the inference time is at least 1.5 times faster than $r = 1$. The undersampled LDMs give a good quality of decision and there is not a wide gap between the performances of the original LDMs ($r = 1$) and the undersampled LDMs ($r = 2, 3, ..$).

V. Conclusion

The objective of this work was the implementation of a method for classifying image pairs into similar and dissimilar classes based on the shape and scale parameters of the Weibull distribution extracted from the undersampled LDMs. The two parameters are used as descriptors for the KNN classifier. The undersampling of the LDMs allowed to save computing time. The performances of our approach was compared with the average structural similarity index (MSSIM) and the feature similarity index (FSIM). The results show that the classification method based on the undersampled LDMs is faster and performs well.

References

[1] B. Kumar, S. B. Kumar, C. Kumar, Development of improved SSIM quality index for compressed medical images, IEEE Second International Conference on Image Information processing (ICIIP-2013), Shimla, pp.251-255, 2013.

[2] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncell, Image quality assessment: from error visibility to structural similarity, Journal in IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600-612, Apr. 2004.

[3] Lin Zhang, Lei Zhang, X. Mou and David Zhang, FSIM: A feature similarity index for image quality assessment, IEEE transactions on image processing, vol.20, no.8, pp. 2378-2386, August 2011.

[4] G. J. Burghouts, A. W. M. Smeulders, J. M. Geusebroek, The distribution family of similarity distances, 20th International Conference on Neural Information Processing, pp. 201-208, 2007.

[5] Etienne Baudrier, Frédéric Nicolier, Gilles Millon, Su Ruan. Binary-image comparison with local-dissimilarity quantification. Pattern Recognition, Elsevier, vol.41 (5), pp. 1461-1478, 2008.

[6] I. S. Molchanov, P. Teran, Distance transforms for real-valued functions, Journal of Mathematical Analysis and Applications, vol. 278, pp. 472–484, 2003.

[7] A. Merdani, A. Kharbach, M. Rahmoun, B. Bellach, M. Elayachi, M. Elhitmy, Local and global measure of dissimilarity between two segmentations, International Journal of Innovative Technology and Exploring Engineering (IJITEE), Volume-4 Issue-6, pp. 2278-3075, 2014.

[8] Z. Hadjadj, A. Meziane, Y. Cherfa, M. Cheriet and I. Setitra, ISauvola: Improved Sauvola's algorithm for document image binarisation, International Conference on Image Analysis and Recognition, vol 9730, pp. 737-745, July 2016..

[9] G. Cohen, S. Afshar, J. Tapson and A. van Schaik, EMNIST: an extension of MNIST to handwritten letters, International Joint Conference on Neural Networks (IJCNN), vol.2, no.1, pp. 2921-2926, 2017.

[10] R. Seulin, C. Stolz, D. Fofi, G. Millon and F. Morain-Nicolier, Three dimensional tools for analysis and conservation of ancient wooden stamps, Imaging Science Journal, vol. 54(2), pp. 111-121, juin 2006..

[11] Moustapha Diaw, Agnès Delahaies, Jérôme Landré, Florent Retraint, Frédéric Morain-Nicolier, Modeling the local dissimilarity map with a two-parameter Weibull distribution – Application to images classification, submitted to IEEE Transactions on Image Processing.