

Article

Characterization of Prints Based on Microscale Image Analysis of Dot Patterns

Indrama Das ¹, Swati Bandyopadhyay ¹  and Alain Trémeau ^{2,*}

¹ Printing Engineering Department, Salt Lake Campus, Jadavpur University, Kolkata 700098, India; indrama.das@gmail.com (I.D.); swatib1@yahoo.com (S.B.)

² Hubert Curien Laboratory, University Jean Monnet, 18 Rue Professeur Benoît Lauras, 42100 Saint-Etienne, France

* Correspondence: alain.tremeau@univ-st-etienne.fr

Featured Application: This project was developed in the frame of the project PACKMARK, in collaboration with Sergusa Solutions Pvt. Ltd. which provided all printed samples necessary for this study. The aim of this project is to identify fake pharmaceuticals packaging products printed using a CMYK (Cyan Magenta Yellow Black) gravure printing machine from color information and microscale analysis. Most often fake packaging printed on blister foils using rotogravure process results from the use of counterfeited cylinder and rotogravure press.

Abstract: Identifying a print document (original) from a reprint document (copy or fake) can be a challenge. The analyse at microscopic scale of print documents shows random dot shapes which depend on the printing parameters as well as the printing device used. We can, therefore, draw the assumption that the dot shapes can be used as a fingerprint to differentiate a print from a reprint. In this paper, we explore several shape indexes that were not investigated until now to analyse at microscopic scale documents printed on aluminium foils using rotogravure printing process. This paper presents a statistical analysis which is based on a pattern recognition process defined by three steps. First, a new image processing pipeline is used to segment automatically disconnected dots. Next, new dot pattern features are used to characterize automatically dot patterns. Six types of dot patterns (including four types of doughnut patterns) are introduced. Lastly, a new statistical analysis method is used to characterize a printed sample from the set of dots printed on it. The experiments done demonstrate the relevance of the analytical method proposed. Results shows the potential of this method to identify a reprint from a print.

Keywords: rotogravure printing process; source printer identification; modeling of printed dots; microscale image analysis; pattern recognition



Citation: Das, I.; Bandyopadhyay, S.; Trémeau, A. Characterization of Prints Based on Microscale Image Analysis of Dot Patterns. *Appl. Sci.* **2021**, *11*, 6634. <https://doi.org/10.3390/app11146634>

Academic Editor: Donato Cascio

Received: 6 June 2021

Accepted: 16 July 2021

Published: 20 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The objective of this paper is to explore statistically geometrical shape indexes that were not investigated until now to analyse at microscopic scale documents printed on aluminium foils using rotogravure printing process.

The rotogravure printing process is not a very common printing process except for printing in large quantities, as in the printing packaging industry, especially when it comes to printing on complex surfaces such as thin aluminium foils used in pharmaceutical packaging. One of the main property of foils is that they are non-porous substrates, consequently they do not cause optical dot gain. Optical dot gain happens with porous paper, due to a lateral propagation of light printed dots appear larger than their physical size [1]. On foils, due to the high reflectance of the surface, any surface irregularity (for example due to material handling) may have an impact on surface reflectance even under well-controlled illumination conditions. In our study case this effect was minor as the printed samples measured were kept flat. On the other hand, some surface irregularities

or topographical depressions may be at the origin of randomly occurring small holes in printed surfaces. Similar observation was done by Hamblyn in [2] who reported a very good correlation between flexo-printed uncovered areas in dots and the location of topographical depressions in paper and paperboard surfaces.

The objective of this paper is also to demonstrate that a printed document can be characterized by a set of statistical features, computed from size, shape and distribution of printed dots, which can be considered as a fingerprint of the print process. To reach this objective we propose a pattern recognition based process which consists of: (1) an new image processing pipeline defined to segment automatically, for each primary color, each individual dot (see Section 3.1, Section 3.2, Section 3.3, Section 3.4, Section 3.5); (2) new pattern features to automatically characterize dots at microscopic scale (see Section 3.6) and to classify dots in six types of dot patterns (including four types of doughnut patterns); (3) a new statistical analysis rules-based method to characterize a printed sample from the set of dots printed on it (see Section 3.7).

There is a large variety of segmentation methods in the literature. Each method has some advantages and also few drawbacks. Some perform better than others, depending on the characteristics of the images to process and of the purpose of the segmentation. In our study case, neither edge-based methods nor region-based methods are adapted to our images, as the contrast between printed dots and the background is too low meanwhile the signal-to-noise ratio is low. On the other hand, threshold methods are well adapted to convert grey-scale images corresponding to primary color images to binary images. In this study, we aim to discriminate printed dots from the print substrate, so a binarization method was used. A similar strategy was implemented by Nguyen et al. in [3]. We used the Otsu global thresholding method [4] for two main raisons: (1) it is based on an unsupervised histogram thresholding approach minimizing intra-class intensity variance; (2) the performance of this method is among the best to differentiate ink coverage from paper regions [4]. However, this method suffers from a drawback, when background values are not homogeneous it performs lower than local threshold methods. In the context of our study case, the Otsu method works well, but a postprocessing based on mathematical morphological tools is nevertheless necessary to increase the robustness of this method. This consists of filtering noise and outliers, such a fuzzy pixels on the dot edges, to locally correct the shape of dots and to perform more accurate measurements (to reduce the impact of physical dot gain on dots shape and dots size). A similar postprocessing approach is reported in [5].

One of the most important phenomena that affects dot segmentation and, consequently, the dot structure analysis is the physical dot gain, meaning that printed dots appear larger than the dots in the digital bitmap [1]. This phenomenon is due to the ink spreading on the surface, which results in an enhancement of the physical dot size.

There is a huge variety of shape indexes in the literature, but few are used to characterize dot patterns, such as dot area (in μm^2), convex area (in μm^2), perimeter (in μm), diameter (in μm), angle (in degree), circularity, roundness, aspect ratio, convexity, etc. [3,5,6]. Several are well designed to characterize simple geometric shapes, such a circles or ellipses. On the other hand, they are not well adapted to characterize complex geometric shapes, such as doughnut shapes. In this study, we investigate several other pattern features to discriminate doughnut shapes from simple geometric shapes and to classify printed dots in six types of dot patterns. A few other features can also be used to characterize dot patterns, such as probabilistic modeling [7]; however, probabilistic features are more adapted to describe non-spherical dots with some degree of surface roughness than dot patterns with doughnuts. Likewise, the probabilistic model investigated by Vallat in [5] is not adapted for doughnut dots or non-spherical patterns. The paper [2] was one of the first papers which investigated dot patterns with concave geometry and local deformation. Doughnuts correspond to ring-shape printed halftone dots with a lack of ink in their centre [2]. This effect is caused by the ink being displaced outward and away from the top of the printing feature during substrate contact. This effect is also due to the rotogravure printing process,

which is a pressure-sensitive process. According to Hamblyn et al. [2] doughnuts are more prominent at high impression pressure—that is, the origin of the doughnut printing defect is still uncertain and the influence of the dot top geometry on the defect [2] is still unknown. The problem of doughnuts on the boundary of dots was reported in [8] but, to the best of our knowledge, was never being investigated in the context of dot pattern characterization. In [8], a study performed with two cylinders engraved using chemical etching and two rotogravure presses, we also observed many doughnuts. According to Mathes [9] the most likely cause of doughnuts is cylinder or plate swelling. Aggressive inks and solvents can also be at the origin of swelling and distortion. Ink with too much draw and ink viscosity too low can be also at the origin of doughnuts. Several studies have shown that several printing parameters can have a significant impact on the quality of a print, for example, on printed dot size or on doughnut effect. For example, Sosa reported in [10] that the temperature from the surroundings and/or from the operating equipment has an impact on the viscosity of ink and consequently on the spreading of the ink on the substrate. The viscosity of inks changes about 3 to 4% per degree Fahrenheit change. Sosa showed in [10] that, when the ink temperature increases, dots are deformed, doughnut effect increases and dot perimeter increases. Kader reported in [11] that, when the ink temperature increases, the ink filling the gravure cylinder cells takes a concave shape, leaving an air bubble trapped between the ink and the substrate. Kader also reported in [11] that the lightness of the dots can be explained by the ample quantity of solvent in the ink, reducing the solid ink constituent. The dot structure analysis performed by Kader is limited to a visual analysis; no quantitative evaluation of doughnuts was performed.

Tkachenko et al. [8] and Joshi et al. [12] demonstrated the potential of local deformations of printed patterns at macroscopic scale, such as “e” letter, to identify unique print signature (see example in Figure 1). On the other hand, Oliver et al. [13] demonstrated the potential at the microscopic scale of roundness and the perimeter of printed dots in discriminating one printer model from another. From the experimental results reported in this paper (see Section 4), we can hypothesize that the deformations at the microscopic scale of the printed dots induced by the printing process and modeled by our shape indexes can be used as a fingerprint to authenticate a print document (see Section 4.1).

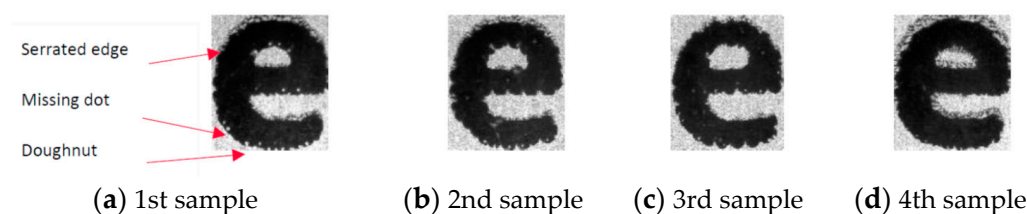


Figure 1. Examples of sample letters “e” printed (or reprinted) with: (a) an engraved cylinder and a rotogravure press, (b) same rotogravure press but using another engraved cylinder, (c) same engraved cylinder as in (a) but using another rotogravure press, (d) same engraved cylinder as in (b) and same rotogravure press as in (c).

The main technical information related to this study is summarized in Section 2. Then, the different methods proposed are described in detail in Section 3. Next, in Section 4 we present experimental results and in Section 5 we discuss the robustness and the efficiency of these methods. Lastly, in Section 6, we discuss the outcomes of the proposed work and draw some conclusions.

2. Technical Information

According to Nguyen et al. [3], at the microscopic scale, a printed dot is a random pattern, the shape of which depends on the technology, the setting of the printer, the ink quality and/or the paper properties. From this statement, Nguyen et al. claimed in [3] that the shapes of printed dots can be considered, at microscopic scale, as the intrinsic feature of the printing process. Cells engraved on the gravure cylinder can have an elongated shape

(cell angle of 60°), a normal shape (cell angle of 45°) or a compressed shape (cell angle of 30°). Cell angle has an impact on ink transfer on the printed substrate. In our case study, the cell angle was of 30° (using a 150 LPI screen ruling and a 130° engraving needle angle) and the ink coverage was of 20%. We will see in this paper that, with a theoretical ink coverage of 20%, due to the physical dot gain, we have already, in some cases, difficulties separating some of the printed dots that are spatially connected. Printed samples were captured by a camera (resolution 8.889×6.667 inches) mounted on a microscope (magnification $4\times$).

Three gravure printing presses with foil inks were used to print the samples used in this study on a blister foil substrate. The gravure cylinders used were engraved by electromechanical engraving process with a dot coverage of 20%. The engraving process was done at 150 LPI screen ruling and a 130° stylus angle. Solvent-based magenta, cyan and yellow gravure ink, specially recommended for blister foils, were used to print the color dots at 60 m/min speed. The thicknesses of the foils were of $25\ \mu\text{m}$. The printing speed, printing pressure and doctor blade pressure were kept constant during the printing of the samples. The drying temperature was set at 50°C to 60°C for the printing. The ambient temperature and humidity were $17 \pm 3^\circ\text{C}$ and $35 \pm 5\%$ inside the press, respectively. Several color print and reprint samples were collected for analysis (either monochromatic samples (e.g., C, M or Y); bi-chromatic (e.g., CM); tri-chromatic (CMY) or quadri-chromatic (CMYK)). These color samples are related to selected samples of the IT8.7/3 color chart, which is commonly used to characterize a CMYK printer (see example in Figure 2).

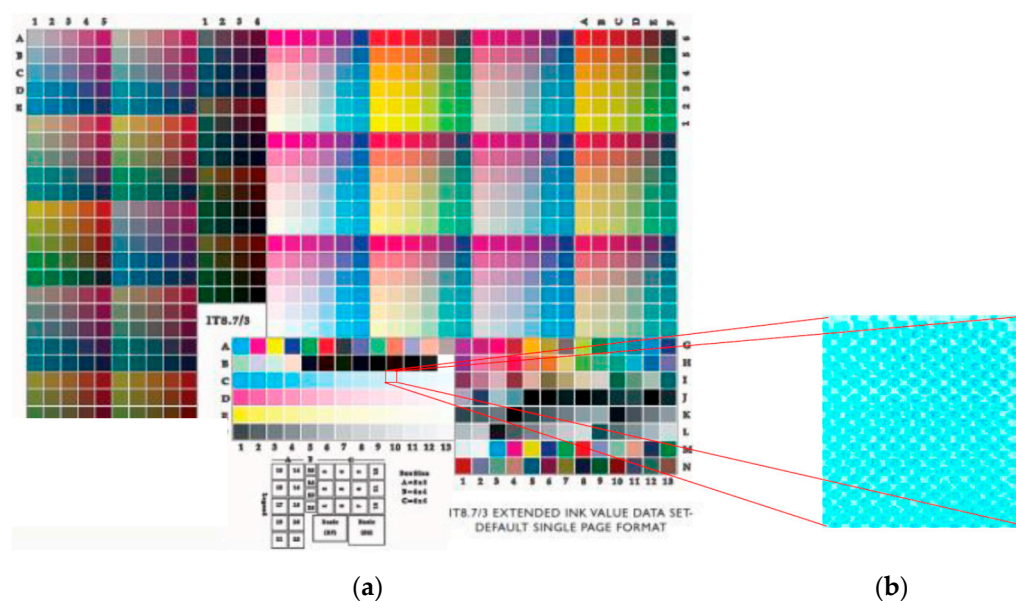


Figure 2. (a) Example of selected samples from the IT8.7/3 Color Chart: sample C10—Cyan with 20% ink coverage; sample D10—Magenta with 20% ink coverage, sample E10—Yellow with 20% ink coverage, (b) cropped image of a rotogravure printed cyan sample magnified $\times 5$.

3. Measurement, Modelling and Analysis at the Microscale of Printed Dots

The image processing pipeline that we propose is illustrated in Figure 3 and detailed in the following sections.

3.1. Separation of Ink Colors (Step 1)

The separation of ink colors is not an easy task as there is often some mis-registration between the printed color channels. Figure 4 shows examples of printed cyan, magenta and yellow dots shifted in position.

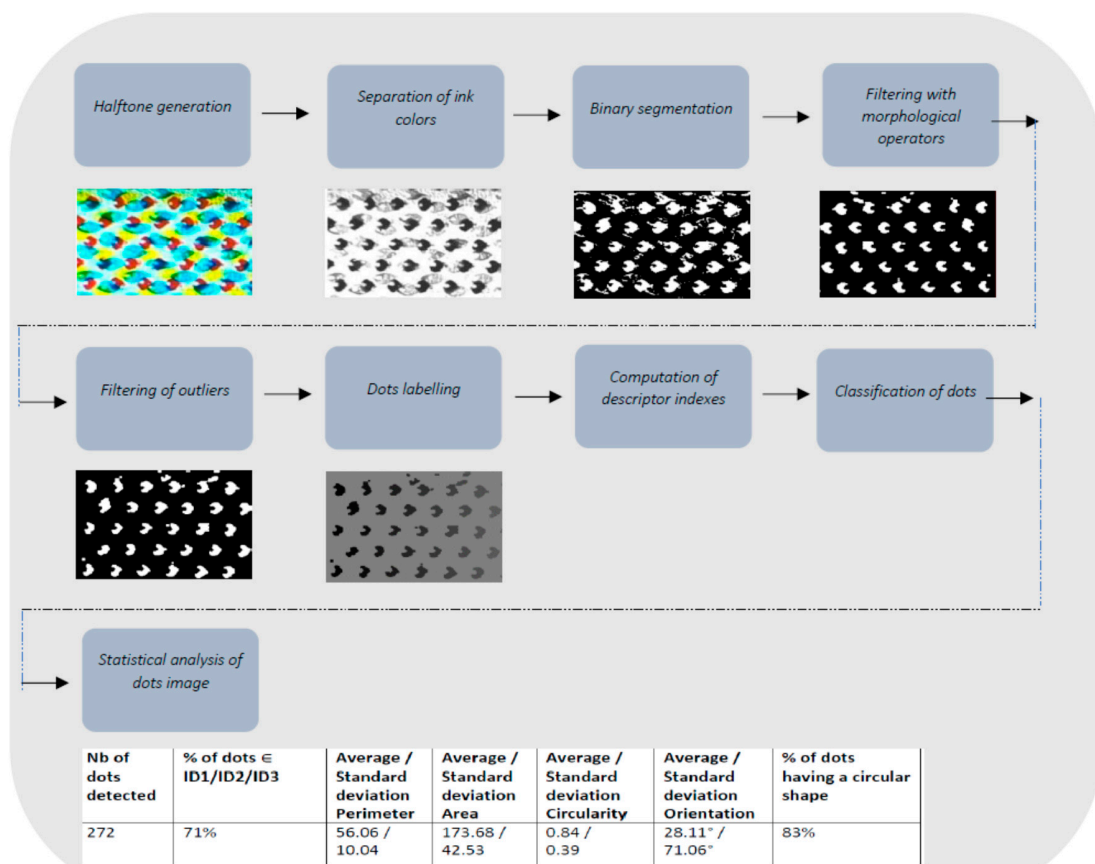


Figure 3. Framework of the proposed image processing pipeline.

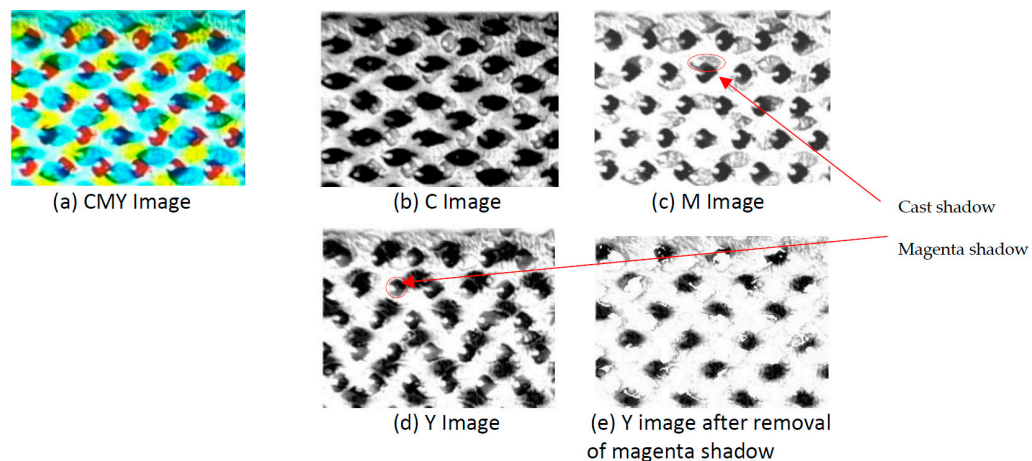


Figure 4. (a) Example of halftone misregistration (of offset along X and Y axis) for a CMY image with 20% of ink coverage for each channel. (b) Even if the magenta dots are perceptible in the cyan image, there is enough contrast between the cyan dots and the background to segment cyan dots. (c) Even if the cyan dots are perceptible in the magenta image, there is enough contrast between the magenta dots and the background to segment magenta dots. (d) Magenta dots cannot be dissociated from the yellow dots in the yellow image. (e) Magenta dots segmented in the magenta image are removed from the yellow image in order to properly segment yellow dots.

The problem of halftone misregistration was studied in [1]. Figure 4a shows an example of a halftoning scheme with different screen angles for cyan, magenta, yellow and black channels. According to Namedanian [1], black, the color with the strongest

contrast, is halftoned and placed at 45° , due to the lower sensitivity of eye at this angle; meanwhile, yellow, the weakest color, is halftoned at 0° , where the human eye is most sensitive. According to Namedanian [1] using a different angle for each channel reduces the effect of misregistration but, on the other hand, introduces a type of pattern, named a rosette pattern, which is quite visible at microscale.

The problem of halftone misregistration was studied in [1]. Figure 5a shows an example of the halftoning scheme with different screen angles for cyan, magenta, yellow and black channels. Black, the color with the strongest contrast, is halftoned and placed at 45° due to the lower sensitivity of eye at this angle, meanwhile yellow, the weakest color, is halftoned at 0° , where the human eye is most sensitive [1]. Using a different angle for each channel reduces the effect of misregistration, on the other hand it introduces rosette patterns, which are quite visible at microscale [1].

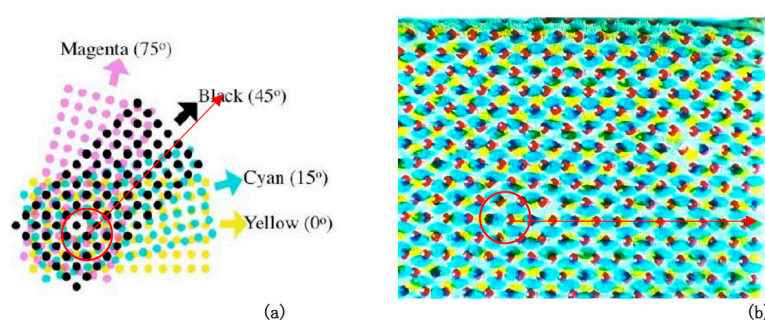


Figure 5. (a) Typical halftoning scheme for yellow (0°), cyan (15°), black (45°) and magenta (75°) (credit [1]). (b) Example of rosette pattern that may occur with another color halftoning scheme (with CMYK at 0° and a shift of few μm along the X and Y axis).

Namedanian proposed, in [1], using different wavelengths to separate colors, but the proposed approach for separating ink colors in color prints was limited to two colors. Moreover, this solution is limited to the available wavelengths, and sometimes it is hard to find good wavelengths to separate more than two color inks. The main issue with the separation of ink colors in color prints when the dots are printed on top of each other (dot-on-dot) is that the effect of one ink cannot be completely removed when sending the light at the absorbing wavelength band of the other ink, as a consequence a shadow effect appears. The term shadow effect of one ink means the contribution of that ink to the reflectance spectra of the other ink [1].

Considering that, in our study case, even if cyan dots were printed on the top of the magenta layer, there is no problem of dot detection in both the Cyan and Magenta channels. Due to the misregistration between these two channels, we process these two channels independently from the others. On the other hand, considering that the yellow dots were printed on the top of the cyan layer and that there is a problem of dot detection on the Yellow channel due to the misregistration between these two channels, we decide to mask the dots detected on the cyan image in the yellow image in order to be able, in the next step, to process the Yellow channel independent from the others. Images corresponding to samples number 7, 8, 9, 10, 14, 15, 16 and 18 in Tables 1–3 show a similar misregistration problem. In our experiments, we only deal with CMY images, and not CMYK images.

Table 1. (a) Statistical values computed for printed magenta dots. The most significant values are highlighted in bold. (b) Statistical values computed for reprint magenta dots. The most significant values are highlighted in bold.

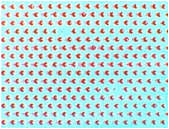
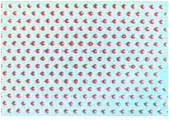
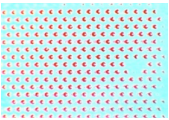
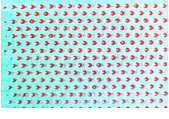
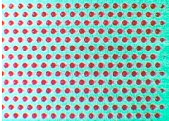
N°	Print Image	Features	Nb of Dots Detected	% of Dots ∈ ID1/ID2/ID3	Avg ± Std Perimeter	Avg ± Std Area	Avg ± Std Circularity	Avg ± Std Orientation	% of Circular Dots
(a)									
1			258	96%	49.52 ± 5.92	121.79 ± 18.68	0.79 ± 0.26	−23.04° ± 74.94°	99%
2			261	96%	51.6 ± 7.69	120.96 ± 23.91	0.76 ± 0.30	12.40° ± 79.12°	99%
3			237	84% (10% ∈ ID6)	42.05 ± 8.75	90.34 ± 25.78	0.80 ± 0.37	−24.17° ± 74.18°	90%
4			242	95%	54.55 ± 7.76	135.68 ± 25.68	0.76 ± 0.31	−6.32° ± 80.11	96%
5			241	94%	54.98 ± 5.19	143.34 ± 19.19	0.77 ± 0.24	40.56° ± 57.43°	94%
6			251	3% (40% ∈ ID4)	62.48 ± 6.51	247.12 ± 40.79	0.89 ± 0.39	22.13° ± 46.09°	94% (49% ∈ ID5)

Table 1. Cont.

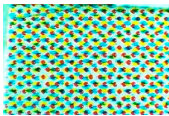
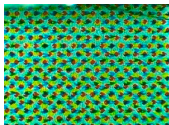
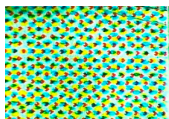
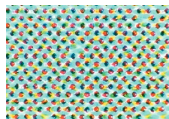
N°	Features		Nb of Dots Detected	% of Dots ∈ ID1/ID2/ID3	Avg ± Std Perimeter	Avg ± Std Area	Avg ± Std Circularity	Avg ± Std Orientation	% of Circular Dots
	Reprint Image								
(b)									
7			272	71%	56.06 ± 10.04	173.68 ± 42.53	0.84 ± 0.39	28.11° ± 71.06°	83%
8			278	76%	55.20 ± 9.35	161.01± 36.29	0.82 ± 0.37	35.49° ± 61.78	82%
9			287	66%	49.84 ± 10.79	137.19 ± 37.66	0.84 ± 0.42	31.8° ± 63.0°	72%
10			258	32% (42 % ∈ ID4)	63.95 ± 7.68	199.48 ± 33.24	0.79 ± 0.31	10.7° ± 75.43°	75%

Table 2. (a) Statistical values computed for print cyan dots. The most significant values are highlighted in bold. (b) Statistical values computed for reprint cyan dots. The most significant values are in bold.

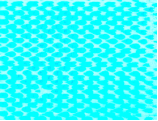
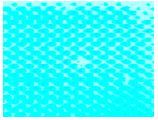
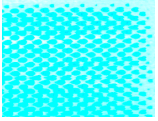
N°	Print Image	Features	Nb of Dots Detected	% of Dots ∈ ID1/ID2/ID3	Avg ± Std Perimeter	Avg ± Std Area	Avg ± Std Circularity	Avg ± Std Orientation	% of Elliptic Dots
(a)									
11			267	34%	86.51 ± 20.40	373.45 ± 105.36	0.81 ± 0.42	− 1.20 ° ± 7.82°	18% (37% have a circular shape) (55% ∈ ID5)
12			265	34% (4% ∈ ID6) (6% of connected dots not analyzed)	94.78 ± 27.20	487.83 ± 170.23	0.84 ± 0.48	− 2.0 ° ± 18.80°	35% (10% have a circular shape) (45% ∈ ID5)
13			249	8% (6% of connected dots not analyzed)	93.50 ± 22.67	481.32 ± 153.78	0.83 ± 0.39	1.81 ° ± 5.56°	24% (31% have a circular shape) (86% ∈ ID5, among them 31% are nor circular nor elliptic)

Table 2. Cont.

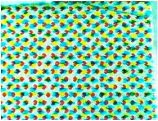
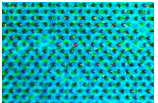
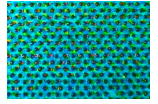
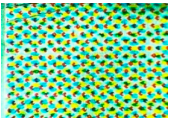
N°	Reprint Image	Features	Nb of Dots Detected	% of Dots ∈ID1/ID2/ID3	Avg ± Std Perimeter	Avg ± Std Area	Avg ± Std Circularity	Avg ± Std Orientation	% of Elliptic Dots
(b)									
14			246	40% (35% ∈ ID6)	63.80 ± 9.81	203.20 ± 38.51	0.80 ± 0.31	5.44° ± 13.90°	21% (25 % ∈ ID5)
15			250	4% (15% ∈ ID6) (28% of connected dots not analyzed)	86.79 ± 67.14	483.47 ± 205.37	0.92 ± 0.53	3.63° ± 19.2°	31% (17% have a circular shape) (53 % ∈ ID5)
16			250	8% (23% ∈ ID6) (22% of connected dots not analyzed)	87.84 ± 65.70	464.39 ± 232.34	0.93 ± 0.56	4.23° ± 21.2°	28% (15% have a circular shape) (47 % ∈ ID5)

Table 3. Statistical values computed for print yellow dots. The most significant values are in bold.

N°	Print Image	Features	Nb of Dots Detected	% of Dots ∈ID1/ID2/ID3	Avg ± Std Perimeter	Avg ± Std Area	Avg ± Std Circularity	Avg ± Std Orientation	% of Elliptic Dots
17			156	6% (7% ∈ ID6)	91.04 ± 13.58	484.19 ± 91.77	0.86 ± 0.35	−1.01° ± 11.65°	75% (80% ∈ ID5)
18			199	30% (30% ∈ ID6)	56.08 ± 23.90	188.63 ± 121.81	0.87 ± 0.52	5.96° ± 44.76°	32% (38% ∈ ID5)

The knowledge that we have acquired about the printing process used enables us to propose an ad hoc solution to separate color dots from color channels. In future work, we will study if this approach could be extended to CMYK images.

3.2. Binarization of Color Images (Step 2)

Figure 6b illustrates how the quality of the printed images studied in our study case is low, due to a high signal-to-noise ratio. Consequently, neither edge-based methods nor region-based methods are adapted to this type of images. On the other hand, the global thresholding method proposed by Otsu [4] is well adapted. This is for two main reasons: (1) it is based on an unsupervised histogram thresholding approach, minimizing intra-class intensity variance; (2) the performance of this method is among the best to differentiate ink coverage from paper regions [5]. However, this method suffers from a drawback, when the background is not homogeneous it performs lower than local threshold methods. To compensate the local variations related to the background and to improve the efficiency of the Otsu segmentation method, we propose to filter the segmented image a posteriori.

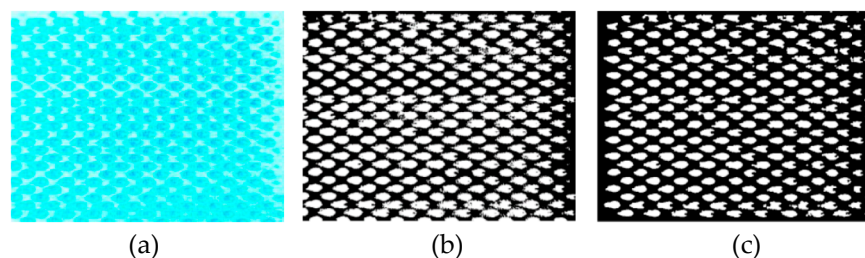


Figure 6. (a) Cyan dots; (b) binary image (with physical dot gain); (c) binary image after removal of dots connected to the borders and noise around dots. Due to the physical dot gain, we have a deformation of the printed dots. The shape and the size of the dots are thus deformed. Physical dot gain led to light cyan elliptic regions; meanwhile, printed dots correspond to dark cyan circular regions.

3.3. Filtering of Color Channels (Step 3)

Several morphological treatments were used in order to reduce the noise contained in the image. These treatments consist of:

1. A removal of dots connected to the image border.
2. An erosion with a disk-shaped structuring element of one-pixel radius. This erosion removes isolated pixels (punctual noise) and reduces the border effects due to the physical dot gain on the size of the dots (see Figure 6).
3. An opening (erosion followed by a dilation) with a disk-shaped structuring element of one-pixel radius. This opening removes noise and removes small objects from the foreground.
4. A removal of small holes in foreground areas with a disk-shaped structuring element of radius 1.
5. A watershed transform based on the Euclidean distance. This transform segments contiguous areas into distinct objects, it enables the better separation of connected dots.

Figures 6c, 7e, 8d and 9e illustrate some results obtained with these operations.

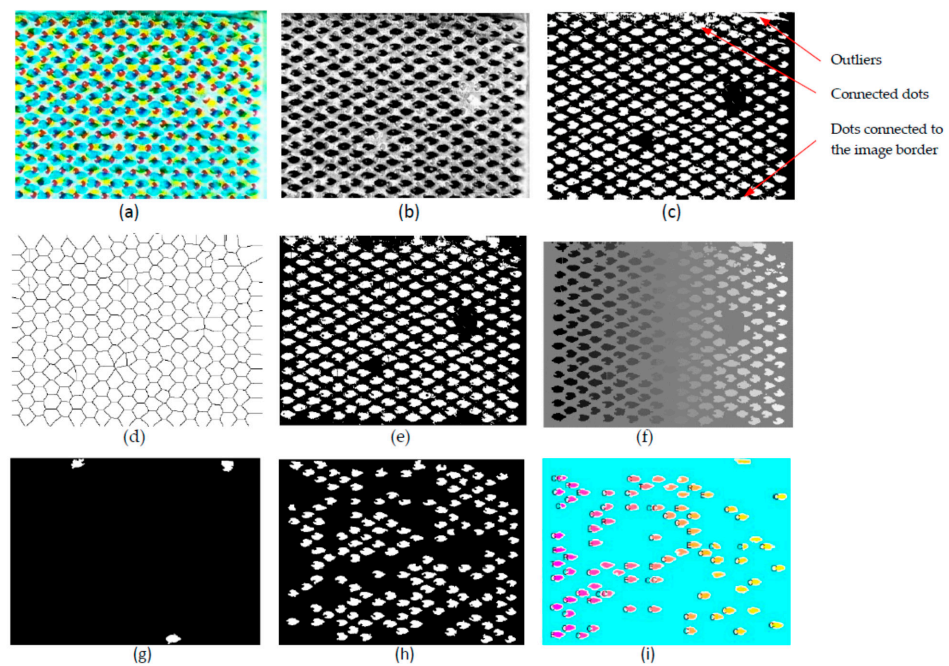


Figure 7. Illustration of the results obtained step by step. In this example, cyan dots are extracted, measured and analyzed. (a) CMY image, (b) Step 1: extraction of C channel, (c) Step 2: binarization, (d) Step 3.5: isosurface of distance transform, (e) Steps 3.1 to 3.4: removal of noise and irrelevant areas, (f) Step 4: labelling of regions of interest, (g) Step 5: removal of outliers, (h) Step 7: extraction of dots with concavity ($\in$ ID1/ID2/ID3) or an irregular shape ($\in$ ID4) defined by a solidity upper than 0.75, (i) extraction of dots with a convex shape ($\in$ ID5) with a circular equivalent shape, or an elliptic equivalent shape, or another pattern shape.

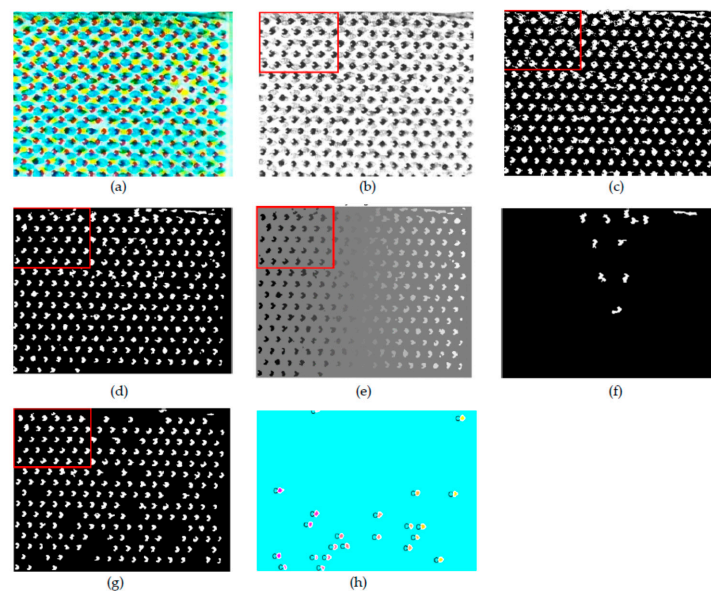


Figure 8. Illustration of the results obtained step by step. In this example magenta dots are extracted, measured and analyzed. The red windows correspond to the image areas shown in Figure 3. (a) CMY image, (b) Step 1: extraction of Magenta channel, (c) Step 2: binarization, (d) Steps 3.1 to 3.4: removal of noise and irrelevant areas, (e) Step 4: labelling of regions of interest, (f) Step 5: removal of outliers, (g) Step 7: extraction of dots with a concavity ($\in$ ID1/ID2/ID3) or an irregular shape ($\in$ ID4) defined by a solidity upper than 0.75, (h) extraction of dots with a convex shape ($\in$ ID5) with a circular equivalent shape, or an elliptic equivalent shape, or another pattern shape.

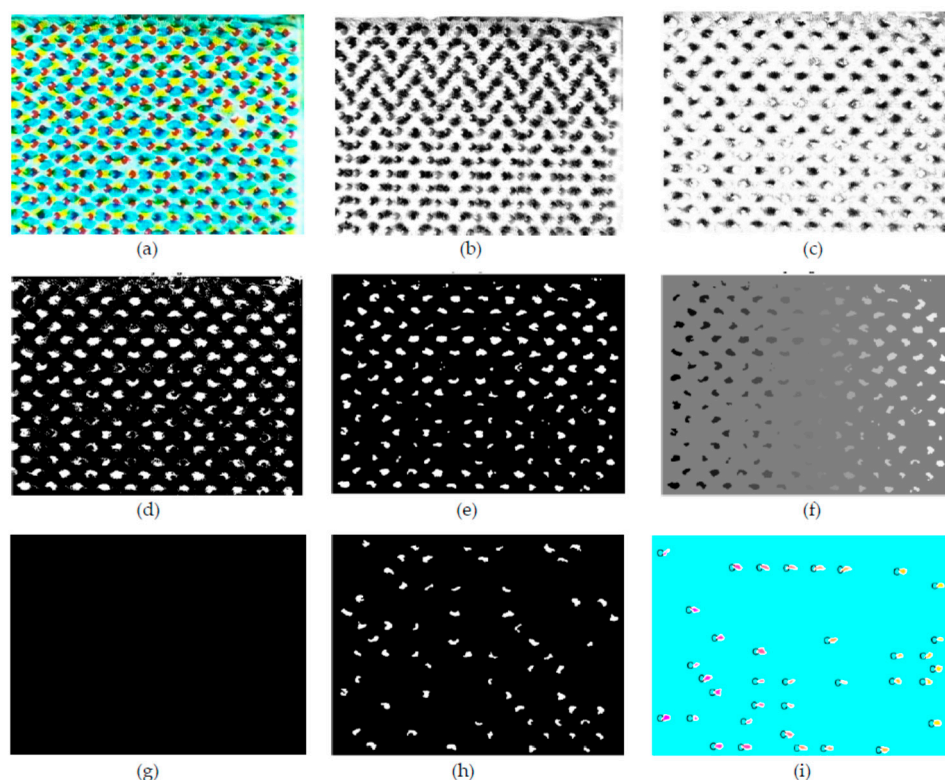


Figure 9. Illustration of the results obtained step by step. In this example yellow dots are extracted, measured and analyzed. (a) CMY image, (b) Step 1: extraction of the Yellow channel, (c) Step 1: Y image after removal of Magenta shadow, (d) Step 2: binarization, (e) Steps 3.1 to 3.4: removal of noise and irrelevant areas, (f) Step 4: labelling of regions of interest, (g) Step 5: removal of outliers, (h) Step 7: extraction of dots with concavity ($\in$ ID1/ID2/ID3) or an irregular shape ($\in$ ID4) defined by a solidity upper than 0.75, (i): extraction of dots with a convex shape ($\in$ ID5) with a circular equivalent shape, or an elliptic equivalent shape, or another pattern shape.

3.4. Labelling of Regions of Interest (Step 4)

Connected pixels belonging to a region of interest of the foreground are labeled as dots. If the area of a region of interest is too large in comparison with the average area of the other dots, then this region is considered an outlier. If two dots overlap, it is very challenging to disconnect them, as this statistically concerns a few numbers of dots. Thus, we propose removing these regions from our analysis process. When two dots are adjacent, it is possible to disconnect them using the watershed transform (see step 3.5). Thus, pixels belonging to the frontier of the two dots (defined by the isosurface of their distance transform) are removed from the foreground in order to separately label each of these two dots.

On the other hand, if the area of a region of interest is too small in comparison with the average area of the other dots, then this region is considered an outlier.

3.5. Filtering of Outliers (Step 5)

If the area of a region of interest is larger than $1.5 \times$ the average area of all dots or smaller than $0.5 \times$ the average area of all dots, then it is removed from the foreground—see Figures 7g, 8f and 9f for examples. This step can be split in two sub-steps: first, the removal of two small dots and updating of the average area of remaining objects in the foreground, followed by the removal of two big dots.

3.6. Computation of Dot Descriptors (Step 6)

We defined six categories of dot patterns (see Figure 10):

- ID1 narrow doughnut (with a hole inside the convex hull of the dot) with a symmetric shape;
- ID2 concave shape (with a large hole) with a symmetrical shape;
- ID3 concave shape (with a hole inside the convex hull of the dot) with an asymmetric shape;
- ID4 elliptic/circular shape with a hole inside the dot;
- ID5 elliptic/circular shape;
- ID6 random shape;

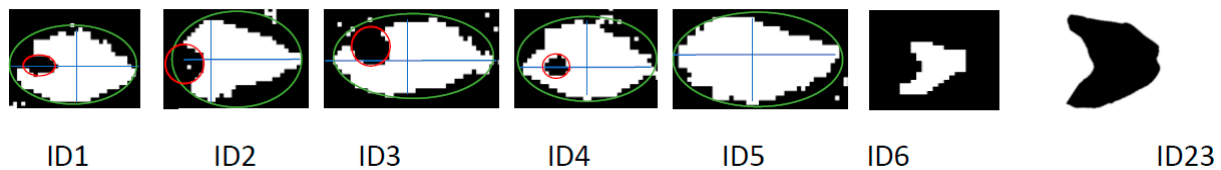


Figure 10. Categories of dot patterns (convex hulls of dots are highlighted in green, major/minor axis are highlighted in blue, convex hulls of holes are highlighted in red).

The ID6 pattern is similar to the ID23 pattern defined in [5].

To analyze each individual dot, we propose computing the following shape descriptors:

- Centroid (X, Y coordinates in μm);
- A Area of the dot surface in μm^2 ;
- A_h Area of the hole in the dot in μm^2 ;
- A_c Area of the convex hull of the dot in μm^2 ;
- A_e Area of the elliptic shape bounding the dot in μm^2 ;
- Solidity = A/A_c ;
- Perimeter of the dot surface in μm ;
- P_c Perimeter of the convex hull of the dot in μm ;
- P_e Perimeter of the elliptic shape bounding the dot in μm ;
- Convexity = P_c/P ;
- Circularity = $\sqrt{4\pi A/P^2}$;
- Roundness = $\sqrt{4\pi A/P_e^2}$;
- $DF_{\min \max}$ Feret diameters in μm ;
- Aspect ratio = $DF_{\min}/DF_{\max}$;
- Orientation = the angle between the x-axis and the major axis of the ellipse that has the same second-moments as the region;
- Major and minor axis;
- A^{x1} Area of the dot upper the major axis;
- A^{x2} Area of the dot lower the major axis;
- A^{y1} Area of the dot at left of the minor axis;
- A^{y2} Area of the dot at right of the minor axis;
- Eccentricity = length of the minor axis / length of the major axis;
- Elongation along X axis = width of the elliptic bounding box / width of the highest inscribed ellipse;
- Symmetry relatively to X axis, to Y axis = A^{x1}/A^{x2} , A^{y1}/A^{y2} , respectively.

Only a few of these descriptors were also used by Valla and Olson [5,6].

To categorize a dot in one of the six categories of dot patterns, we defined the following rules:

- Concave shapes belonging to category ID1 are defined by an elongation upper than 1.75, a symmetry value A^{x1}/A^{x2} upper than 0.85 and lower than 1.15, and a solidity upper than 0.8.

- Concave shapes belonging to category ID2 are defined by a solidity upper than 0.75 and a symmetry value A^{x1}/A^{x2} upper than 0.85 and lower than 1.15.
- Concave shapes belonging to category ID3 are defined by a solidity upper than 0.75 and a ratio A^{ij}/A upper than 0.15 (assuming that the concavity is in the region R^{ij}).
- Dots belonging to category ID4 have a ratio A_h/A upper than 0.15. Below this value, holes are insignificant.
- Circular equivalent dots (see example shown in Figure 11) belonging to category ID5 are defined by a roundness value higher than 0.90 (and an eccentricity upper than 0.85).
- Elliptic equivalent dots (see example shown in Figure 11) belonging to category ID5 are defined by a roundness value between 0.8 and 0.9 (and an eccentricity between 0.6 and 0.85).
- Elongated dots (belonging to category ID6) are defined by an eccentricity lower than 0.6 and a roundness value lower than 0.8 (and an elongation upper than 1.3).
- Any other shape which does not satisfy any of the above rules belongs to category ID6.

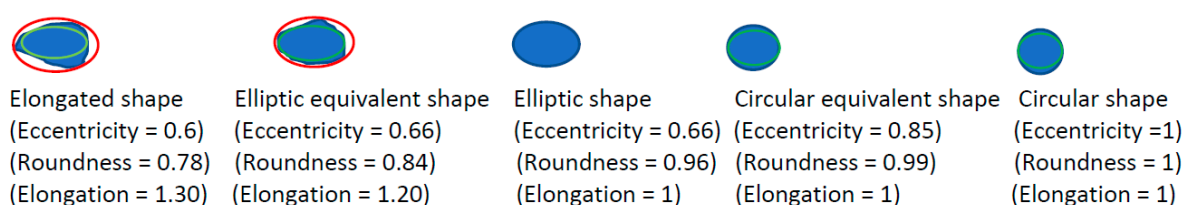


Figure 11. Examples of descriptor values for various patterns. The elliptic bounding box corresponds to the red ellipse. The inscription corresponds to the green ellipse.

These rules are based on threshold values experimentally set from a statistical analysis of 4467 dots from 18 print samples. These thresholds are only valid for the test samples considered in this study case. These threshold values and rules must be adapted to each study case considered. For example, ID1, ID2, ID3 and ID4 patterns may occur only in few study cases, for example when the viscosity of the ink is low, as shown in images corresponding to samples number 1, 2, 3, 4 and 5 (see Table 1 and Figures 12 and 13). These thresholds are dependent of the quality of the print (noise, contrast, physical dot gain, rate of ink covering, overlap of inks, etc.) and, therefore, of the printing parameters, such as the ink viscosity or the ink temperature. The setting of these descriptors results in a trade-off between the accuracy of the descriptors and their robustness against print noise from one dot to another one.

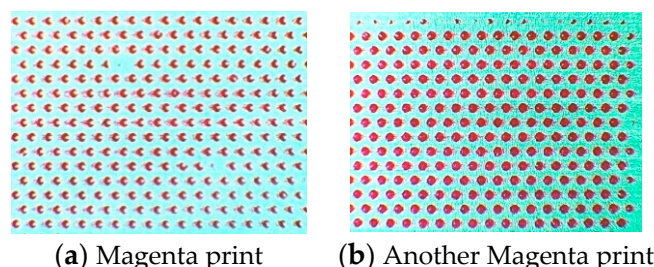


Figure 12. Example of results for two print samples printed with a cell angle of 60°. (a) % of dots belonging to category ID1 (0%), ID2 (92%), ID3(2%), ID4 (less than 1%), ID5 (0%), ID6 (5%, most of them are at right of the figure). There are two missing dots. In total, 92% of dots have a circular shape—see statistical values report in Table 1. (b) % of dots belonging to category ID1 (less than 1%), ID2 (0%), ID3 (less than 1%), ID4 (40%), ID5 (49%, all have a circular shape), ID6 (6%, most of them are the top of the figure). There is no missing dot. In total, 94% of dots have a circular shape.

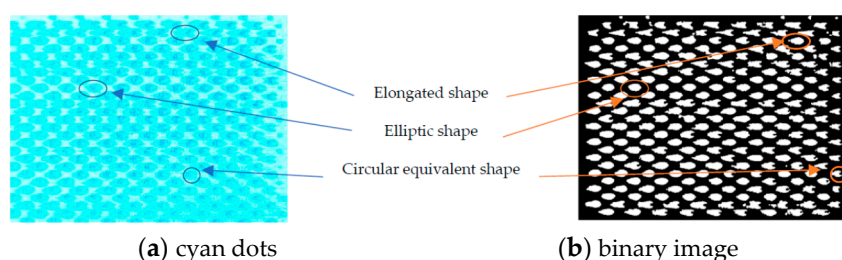


Figure 13. Another example of results for a cyan print with a cell angle of 60° . % of dots belonging to category ID1 (34%), ID2 (0%), ID3 (0%), ID4 (0%), ID5 (55%), ID6 (11%, most of them are at the border of the figure). There is no missing dot. In total, 62% of dots have an elliptic shape, while 25% of dots have an elongated shape (less than 1% has a circular shape).

From this study we claim that: several dot descriptors must be used to properly evaluate the shape of the dots printed; several dot descriptors must be analyzed conjointly to classify properly the dots in predefined dot patterns; the robustness of these thresholds and rules greatly depends on the print parameters.

3.7. Global Analysis vs. Individual Analysis (Step 7)

Considering the heterogeneity of dot patterns and the random distribution of these dot patterns on print samples (as illustrated for example in Figure 13), we hypothesize that, to properly characterize and identify a print, it is necessary to analyze it at the global level—that is to say, to statistically analyze the distribution of the dots printed on these samples. The experimental results reported in Tables 1–3 show some statistical results for 18 print samples.

4. Experimental Results

The efficiency of the dot segmentation method was evaluated by visual evaluation, meanwhile the accuracy of some computed values (size of the major and minor axis, area, radius of the circumscribe circle) was evaluated using the software ScopeImage 9.0. The computed values reported in Tables 1–3 are coherent with measurement values. The problem of threshold methods evaluation and comparison is discussed in [5], meanwhile the problem of shape measurements calibration is discussed in [6]. In our study, the objective was only to make a relative comparison of computed values between print samples. The accuracy of the computations was secondary.

From Table 1 we can draw several observations:

- Samples corresponding to samples number 1, 2, 3, 4, 5 and 6 are very similar in terms of dots shape. As most of the dots are circular, the parameter “orientation” is irrelevant (its standard deviation is very high).
- Samples corresponding to samples number 1, 2, 3, 4, and 5 are very similar (high proportion of doughnuts), meanwhile sample corresponding to sample number 6 is unique (in comparison with these samples) as a significant proportion of its dots have a hole (low proportion of doughnuts).
- The threshold separating “circular equivalent” shapes from “elliptic equivalent” shapes associated with the parameter “circularity” is less robust than the parameter “roundness” which is more relevant to characterize the circularity of non-circle shapes such as the one $\in$ ID1/ID2/ID3). As illustration, compare the “circularity” of the sample corresponding to sample number 3 (0.80) with the “circularity” of the sample corresponding to sample number 11 (0.81, see Table 2), these values are similar but visually dots of the first sample are more circular than the second one.
- The shape and size of dots of samples corresponding to samples number 1, 2, 3, 4, 5 and 6 is statistically quite homogeneous (standard deviation values are quite low, for example less than $10\ \mu\text{m}$ for the perimeter).

From Table 2, we can draw several observations:

- Samples corresponding to samples number 11, 12 and 13 (and also to sample number 17, see Table 3) are very similar in terms of dots shape, most of the dots have a “elliptic equivalent” shape oriented in the direction of the print (see Figure 14), which is coherent with the values computed for the parameter “orientation”.
- The proportion of doughnuts is lower than for samples corresponding to samples number 1, 2, 3, 4, 5 and 6 (only 34% for samples number 11 and 12, and less than 8% for samples number 13 and 17). These doughnuts are less visible to the naked eye but are noticeable at microscopic scale. As there are less doughnuts in these samples the “circularity” parameter is more accurate (in the range between 0.81–0.86), which is coherent with an “elliptic equivalent” shape. On the other hand, the standard deviation of “circularity” values is higher (in the range 0.35–0.48), as the dot patterns are more heterogeneous. This is due to a higher standard deviation of the “perimeter” and of the “area” of the dots.

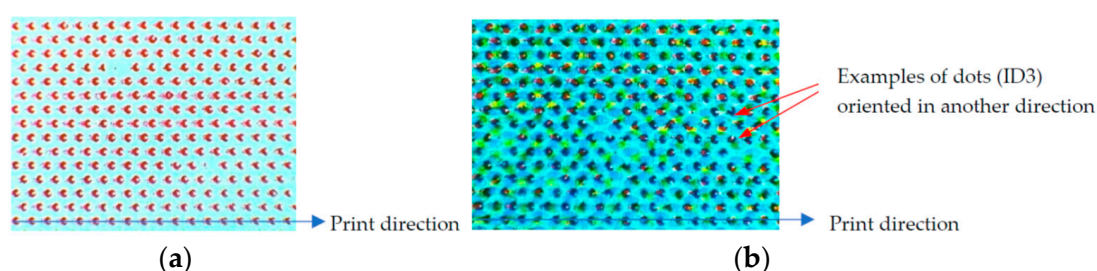


Figure 14. (a) Magenta dots with 95% of doughnuts oriented in the print direction; (b) Magenta dots with 85% of doughnuts oriented in the print direction. Orientation of the doughnuts vs. print direction.

Unfortunately, the amount of data and of rotogravure printers used in this preliminary study (4467 dot samples vs. two printers) are insufficient to precisely evaluate the accuracy and robustness of all dot pattern parameters proposed; nevertheless, the amount of data was sufficient to demonstrate the relevance of these parameters.

In the next section, we compare the robustness of these parameters versus different printing systems.

4.1. Prints vs. Reprints

Values reported in Table 2 samples number 11, 12 and 13 (see also Figure 15a–c) show that the dots distribution in these samples are statistically very similar (elliptic shape), nevertheless we noticed in previous section that the proportion of doughnuts in sample number 13 (printed with printer P2) is lower than for the two other samples (both were printed with printer P1).

Likewise, values reported in Table 1 for samples number 1, 2 and 3 (see also Figure 15d–f) show that the dot distribution in these samples is statistically very similar (circular shape), nevertheless we notice that the proportion of doughnuts in sample number 3 (printed with printer P2) is once again lower than for the two other samples (both were printed with printer P1). This is coherent with our hypothesis that the deformations at microscopic scale of printed dots, induced by the printing process, can be used as a fingerprint to differentiate two prints.

Let us now compare three reprint samples of the same artwork. Images shown in Figure 16 show that the dot distribution in these samples is very similar (Cyan dots have an elliptic shape, Magenta dots have a circular shape and Yellow dots have a random shape), results shown in Figure 16d–f demonstrate the dot patterns detection process is quite robust against noise, non-uniformity of lighting field, dots connectivity, dot-on-dot misregistration, inks, etc. Here, the shape of Yellow dots was badly estimated due to the overlap of Cyan and Magenta dots on the top of Yellow dots. The dot distribution of these reprint samples does not correspond to the dot distribution of the print sample scanned

(sample [11]), due to the high reflectance of the substrate, which complexifies the color calibration process and to an inaccurate color measurement of the dot distribution.

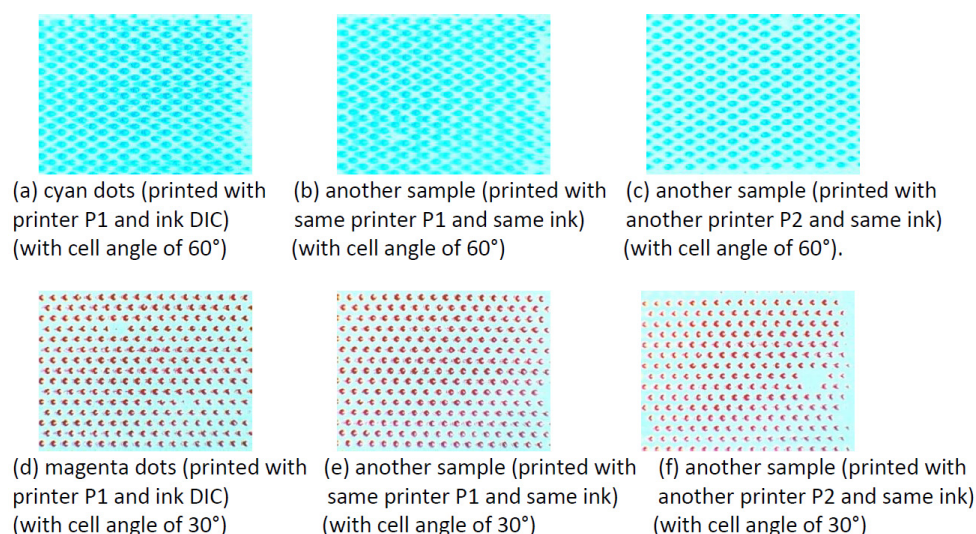


Figure 15. Robustness of the dot patterns detection method relative to the printing device.

If we compare samples corresponding to [11–13] (printed samples) with samples corresponding to [14,15] (reprinted samples), we can notice a higher proportion of dots $\in$ ID6 for [14,15] than for the three other samples (see Table 3). This can be explained by the fact that, for the reprint samples, it is more difficult to estimate properly the shape of the dots due to a higher standard deviation of “area” and “circularity” parameters (and consequently a higher connectivity rate of dots). We hypothesis that this is due to the printing parameters used for the reprint rather and not to the scan process (as the density of Cyan dots was well estimated). We can, therefore, draw the assumption that the deformations at microscopic scale of reprinted dots (induced by the print and scan and reprint process) can be used as a fingerprint to differentiate a print from a reprint.

Figure 17 illustrates two examples of dots’ deformation, one for Magenta print dots (see Figure 17c), another for Magenta reprint dots (see Figure 17f). The overgrowth of dots due to the printing process are smoothed, thanks to morphological operations applied before computing shape indexes, but they are not removed as they are intrinsic of the printing process. The “elongation” parameter enables these local deformations to be quantified (see examples highlighted with an orange arrow). The size of the Magenta dots is larger in Figure 17f than in Figure 17c; meanwhile, elongated dots with an overgrowth in the opposite direction of the print direction are more regularly distributed (see dot line highlighted with a blue arrow).

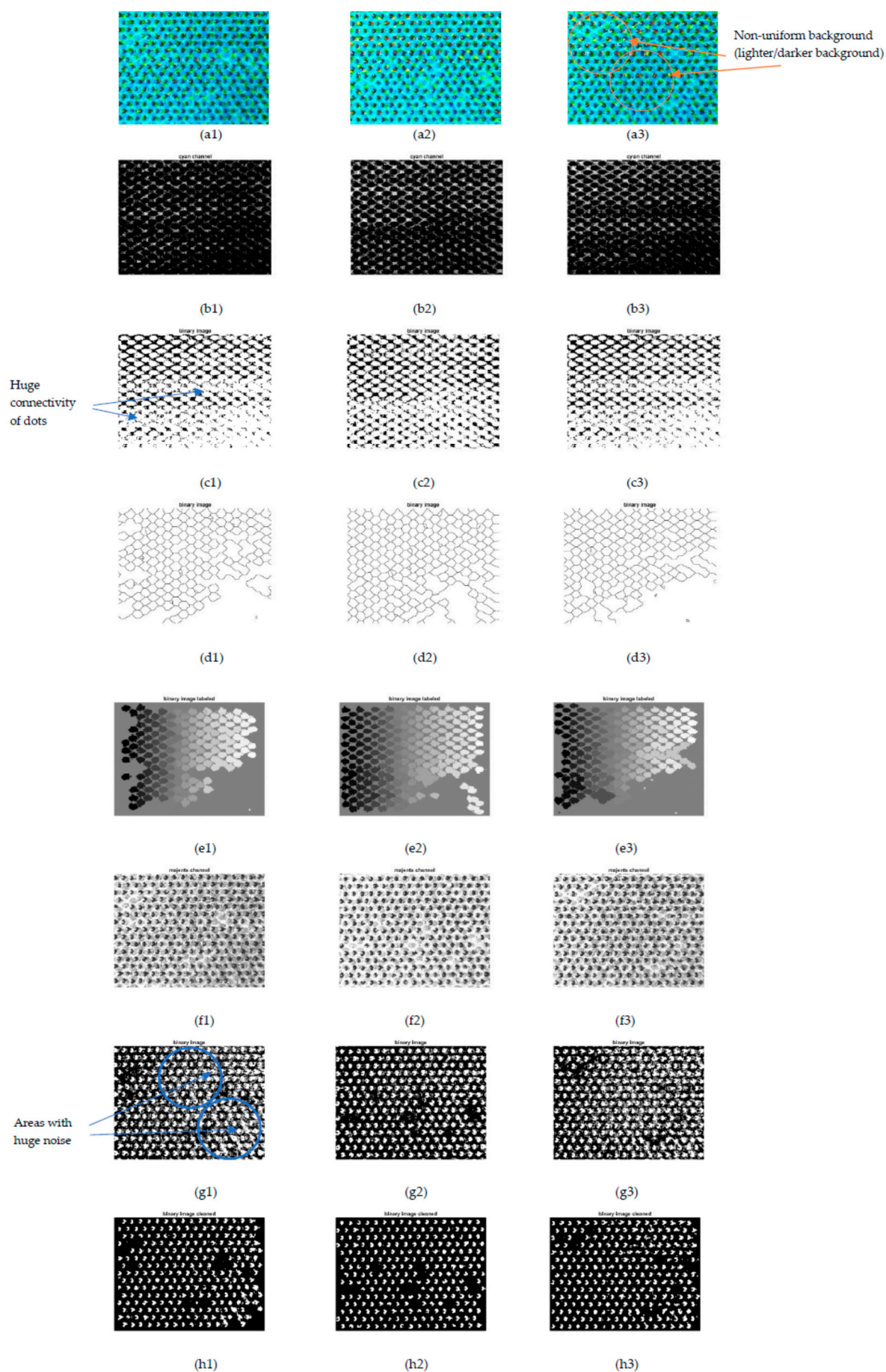


Figure 16. Cont.

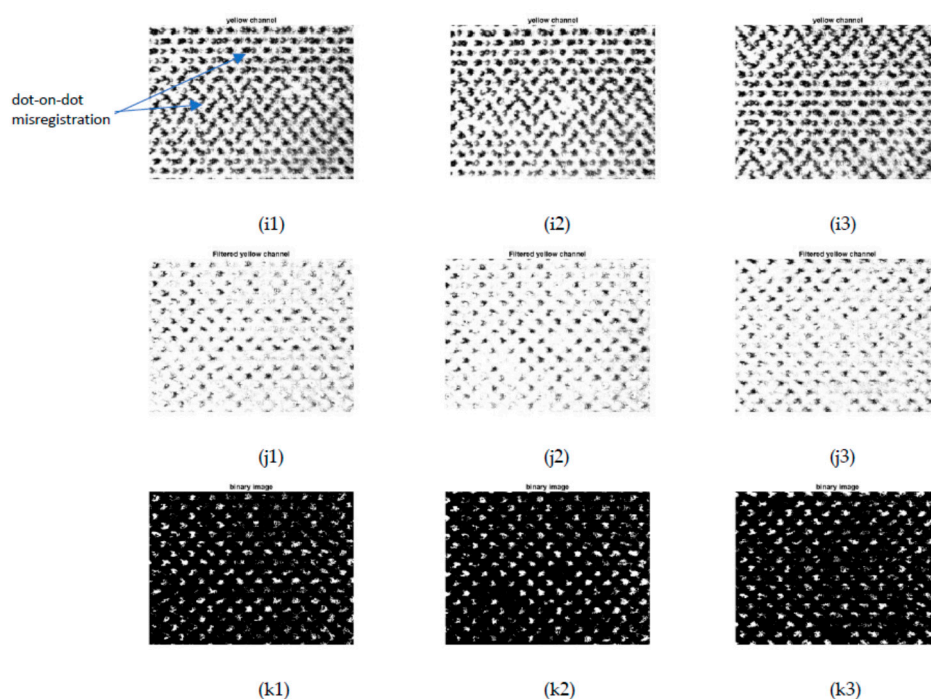


Figure 16. Robustness of the dot patterns detection method relative to the printing device. (a1) CMYK Image (printed with printer P2 and ink DIC) (cyan dots have a cell angle of 30° and magenta dots have a cell angle of 60°), (a2) Another CMYK Image (cyan dots have a cell angle of 30° and magenta dots have cell angle of 60°), (a3) Another CMYK Image (cyan dots have a cell angle of 30° and magenta dots have cell angle of 60°), (b1–e3) results step by step for C color channel, (f1–h3) results step by step for M color channel, (i1–k3) results step by step for Y color channels.

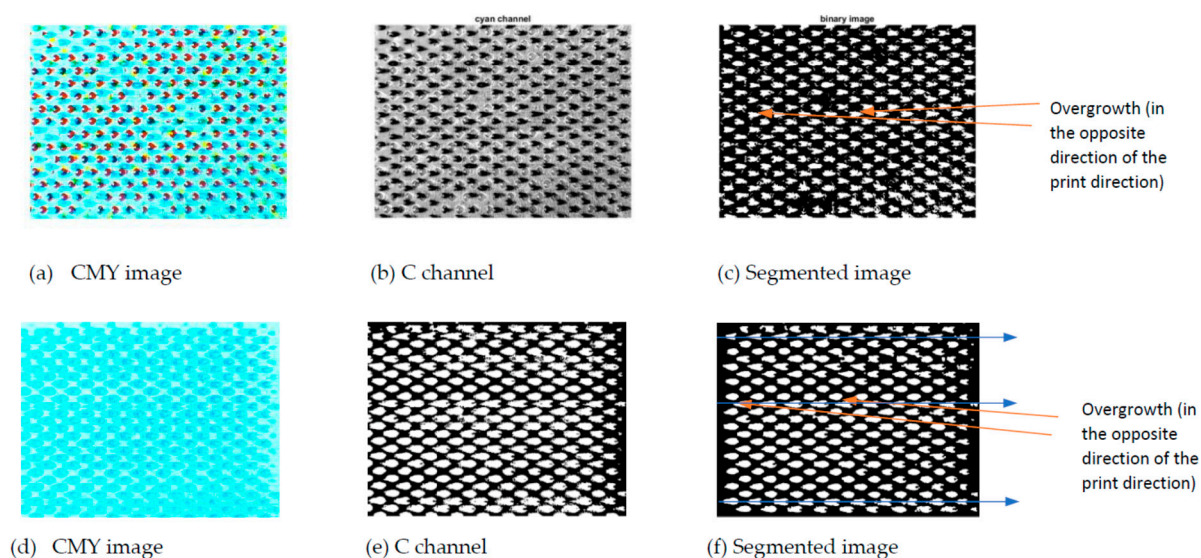


Figure 17. Examples of dot deformation.

5. Discussion

In this study, we did not compare the performance (nor the accuracy) of our thresholding method to other thresholding methods. Likewise, we did not compare the performance (nor the accuracy) of our size and shape indexes with other indexes of the state of the art. Our main objective was to demonstrate that the use of size and shape indexes is not sufficient to properly characterize the shape of print dots (printed using a rotogravure

process), and that the categorization of dots in predefined dot patterns combined with a statistical analysis of relevant shape indexes is more relevant to properly characterize a print from its dots' distribution. In Section 4, we discussed the limitations of the different methods used for steps 1 to 7 of the framework that we propose and of the threshold values that we used. Our objective was not to optimize the efficiency of each of these steps but to propose a relevant framework to our study case that could be extended to other study cases similar to the one studied. Several limitations pointed out in Section 4 would need improvements that could be obtained only with a higher number of print samples.

This preliminary study was done from a "limited" set of prints (with, in total, 4467 dots). To evaluate the accuracy and robustness of the pattern recognition process proposed, more tests should be done with annotated data. More samples should be analyzed and also more printers (and printing parameters) should be compared. This is costly and time consuming. Likewise, to evaluate the efficiency of our hypothesis according with print and reprint samples can be differentiate from dots' deformations at microscopic scale more prints and reprints (and printing parameters) should be compared. That is, the first results shown in this paper are very promising and demonstrate that the study parameters and the proposed methodology are relevant enough to support the hypothesis. In a future work we will compute the confusion matrix (in %) of the proposed classification method and will evaluate its robustness using ROC Curves and Precision–Recall curves. In the case of overlapping problems between classes, we will investigate machine-learning-based methods to optimize soft decision.

The experimental analysis performed in Section 4 shows the importance of considering, in the analytical process, the theoretical/practical knowledge/expertise that we have about the rotogravure printing process and about the shape of the dots printed. This justifies the use of a pattern-recognition-based process rather than a machine-learning-based process. Moreover, machine-learning-based methods, such as deep-learning methods, need lot of data.

6. Conclusions

In this paper, we investigated a promising solution based on a pattern recognition process to identify a print document (original) from a reprint document (copy or fake). We demonstrated that the geometrical shape of printed dots can be analyzed at microscopic scale to differentiate a print from a reprint, as a fingerprint.

To improve the efficiency of the dot segmentation process, we propose a new image processing pipeline. To better characterize complex geometrical dot shapes, we defined six types of dot patterns, including four types of doughnut patterns, and defined a new rules-based characterization method which combines several dot pattern features. To better characterize each printed sample from the set of dots printed on it, we proposed a new statistical analysis method based on a pattern recognition process. The preliminary experiments demonstrated the relevance of the criteria/rules/methods proposed and of the statistical analyze method implemented. In a future work, we will perform complementary experiments to study the robustness of our solution in the function of the printing parameters, as well as the printing device used. We will also study the impact of the scan process involved in the reprint process (based on a scan-and-print process). The implemented solution suffers from some limitations that were discussed in the paper. In future work we will study how to improve some steps of the image processing process, especially the segmentation process when neighboring dots are overlapping or when color dots (primary colors) are overlapping.

If one want to generalize the rules suggested in this paper for new data without overfitting, it could be interesting to investigate an unsupervised machine learning method [14]. Noise learning is one important cause of overfitting. Pruning could be used to eliminating less meaningful or irrelevant data and, finally, to prevent overfitting and to improve the classification accuracy [15]. During the learning process, stopping criteria could be used to determine when to stop adding conditions to a rule, or adding rule to a model description.

Pruning could be also used to test the significance of rules based on significant differences between the distribution of positive and negative examples. It will be interesting to investigate if other combinations of rules could improve the classification accuracy of dots regardless of their shape. Parameter tuning is another important cause of overfitting; it needs sufficient samples for learning. To expand the training set, few (data augmentation) strategies could be used: acquire more training data; add some random noise to existing dataset; reacquire some data from existing data set through some image processing; produce some new data based on the distribution of an existing data set or from simulated data [15]. Well-tuned parameters make a good balance between training accuracy and regularity [15]. As the number of dot patterns resulting from the rotogravure printing process is very limited, the number of useful features will not increase with new data. However, the weight of these features could have an influence on the final classification. A regularization process could be used to minimize the weights of the features which have little influence on the final classification.

Author Contributions: Conceptualization/methodology/formal analysis, A.T.; Software/investigation/resources, A.T. and I.D.; Validation, I.D. and S.B.; Writing—original draft preparation, A.T.; Writing—review and editing, I.D. and S.B.; Supervision/project administration/funding acquisition: A.T. and S.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work developed in the frame of the project NO. IFC/7127-PACKMARK was funded by the Indo-French Centre for the Promotion of Advanced Research (IFCPAR/CEFIPRA).

Data Availability Statement: In this study, all print samples were designed by the authors and images of these samples were captured by them. All samples were printed thanks to the help of Sergusa Solutions Pvt. Ltd. The images of these samples were captured and analyzed by Jadavpur University. These images are available upon request.

Acknowledgments: The authors would like to express their thanks to the Indo-French Centre for the Promotion of Advanced Research (IFCPAR/CEFIPRA) for funding and also thanks to Industry Sergusa Solutions Pvt. Ltd. for providing printing facilities to carry this work.

Conflicts of Interest: The authors have no conflict of interest to declare.

References

1. Namedanian, M. *Characterization of Halftone Prints Based on Microscale Image Analysis*; Linköping University Electronic Press: Norrköping, Sweden, 2013.
2. Hamblyn, A. Effect of Plate Characteristics on Ink Transfer in Flexographic Printing. Ph.D. Thesis, Swansea University, Swansea, UK, 2015. Available online: <http://cronfa.swan.ac.uk/Record/cronfa42827> (accessed on 16 July 2021).
3. Nguyen, Q.; Mai, A.; Chagas, L.; Reverdy-Bruas, N. Microscopic Printing Analysis and Application for Classification of Source Printer. *Comput. Secur.* **2021**, *108*, 102320. [CrossRef]
4. Otsu, N. A threshold selection method from gray-level histograms, *IEEE Trans. Sys. Man. Cyber.* **1979**, *9*, 62–66. [CrossRef]
5. Vallat-Evrard, L. Measurement, Analysis and Modelling at the Microscale of Printed Dots to Improve the Printed Anti-Counterfeiting Solutions. Doctoral Dissertation, Chemical and Process Engineering Université Grenoble-Alpes, Grenoble, France, 2019.
6. Olson, E. Particle Shape Factors and Their Use in Image Analysis—Part 1: Theory. *J. GXP Compliance* **2013**, *15*, 85.
7. Nguyen, Q.T.; Delignon, Y.; Septier, F.; Phan-Ho, A.T. Probabilistic modelling of printed dots at the microscopic scale. *Signal Process. Image Commun.* **2018**, *62*, 129–138. [CrossRef]
8. Tkachenko, I.; Trémeau, A.; Fournel, T. Authentication of Medicine Blister Foils: Characterization of the Rotogravure Printing Process. In Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2019), Prague, Czech Republic, 25–27 February 2019; pp. 577–583. [CrossRef]
9. Mathes, H. Flexo troubleshooting. Most common flexo printing issues, Part 3. *Flexo Grav. Int.* **2011**, *4*, 22–23.
10. Sosa, R. Effects of Temperature Control on Gravure Packaging Inks. Master's Thesis, Western Michigan University, Kalamazoo, Michigan, 1999. Available online: https://scholarworks.wmich.edu/masters_theses/4935 (accessed on 15 July 2021).
11. Kader, M.E.A. The Impact of Ink Viscosity on the Enhancement of Rotogravure Optical Print Quality. *Int. Des. J.* **2017**, *7*, 103–107.
12. Joshi, S.; Saxena, S.; Khanna, N. Source Printer Identification from Document Images Acquired using Smartphone. *arXiv* **2020**, *2003*, 12602.
13. Oliver, J.; Chen, J. Use of Signature Analysis to Discriminate Digital Printing Technologies. In Proceedings of the IS&T's NIP18, International Conference on Digital Printing Technologies, San Diego, CA, USA, 29 September–4 October 2002; pp. 218–222.

-
14. Sun, S.; Cao, Z.; Zhu, H.; Zhao, J. A Survey of Optimization Methods from a Machine Learning Perspective. *IEEE Trans. Cybern.* **2020**, *50*, 3668–3681. [[CrossRef](#)] [[PubMed](#)]
 15. Ying, X. An Overview of Overfitting and its Solutions. *J. Phys. Conf. Ser.* **2019**, *1168*, 022022. [[CrossRef](#)]