Original Research Paper

Automatic Identification of Molecules in a Thin-Film Extract Combined with Laser

[1]**Theodore Guié Toa Bi,** [1]**Sié Ouattara and** [2]**Alain Clement**

[1]*Electrical Systems and Signals Laboratory (L2SE),*
Houphouet Boigny's National Polytechnic Institute of Yamoussoukro, Côte d'Ivoire
[2]*University Institute of Technology Angers (IUT), Angers, France*

Abstract: This paper presents the automation of the thin-layer chromatography technique whose separation and identification of molecules present in a mixture are currently done manually and laboriously. We have therefore found an interest in automating this technique. In this part, the method implemented comprises 2 steps. First we proceeded to the segmentation of the spots obtained on the chromatographic plate. We then developed a program to identify families of molecules such as coumarins, terpenes, tannins, flavonoids, polyphenols, etc. by calculating segmentation parameters such as standard deviation, entropy, mean pixel intensity and frontal ratio from an algorithm on the matlab software. Finally our results have been compared to the results obtained by the traditional identification technique in laboratories. Some similarity between the two results obtained shows the reliability and the robustness of our technique.

Keywords: Automation, Chromatography, Thin Film, Identification, Segmentation, Standard Deviation, Entropy, Mean Intensity, Frontal Ratio, Algorithm, Matlab

Introduction

Traditional medicine based on medicinal plants, survives and tends to develop in parallel with modern medicine in developing countries. According to the World Health Organization (WHO) in 2014, nearly 80% of people in developing countries use traditional medicine for primary health care; hence the need for scientific valorization of medicinal plants (Traore *et al.*, 2015). It is important to note that in most African countries, particularly in Côte d'Ivoire, traditional medicine is sometimes the only source of affordable and accessible care, especially for the poorest patients. This medicine uses mostly herbal therapies. However, several anticancer molecules come from medicinal plants (Kabran, 2014a). A majority of the world's population, especially in developing countries, is treated only with traditional herbal remedies (ANSM, 2016). From aspirin to taxol, the modern pharmaceutical industry itself still relies heavily on the diversity of plant secondary metabolites to find new molecules with novel biological properties (Canvalho *et al.*, 2015). Each plant can contain several thousand different compounds (Adjanohoun and Aké-Assi, 2008) whose various uses are aimed at overcoming suffering and improving the health of men (Canvalho *et al.*, 2015).

Moreover, the separation and the identification of the different groups of compounds or molecules in a mixture (extract) have interest in the laboratories of natural substances. In these laboratories all research is based on plants (stem, leaves, roots, seeds) and mostly works in the context of the revaluation of traditional medicine based on plants.

In addition, the identification of families of molecules in the leaves of plants (extracts) are of great interest in laboratories of natural substances. In these laboratories all research is based on plants (stem, leaves, roots, seeds) and mostly works in the context of the revaluation of traditional medicine based on plants (Adjanohoun and Aké-Assi, 2008).

However, the separation and identification of molecules is done manually, always referring to the data in the literature, which is slow and very laborious. We have therefore found an interest in automating the identification technique based on algorithms run on Matlab software version 2016. For this work, we used three (03) plants that are: Alchornea cordifolia, Paullinia pinnata (Sapindaceae), Trema guineensis (Ulmaceae) (Djeddi *et al.*, 2013).

We have first described the material and methods in which we present the segmentation of the digital color images of the extracts acquired on the plate as well as the device of experimentation acquisition of these images. The second step is a low-level image processing of segmentation parameter calculations such as entropy, standard deviation, mean pixel intensity and frontal ratio. Finally, we compare our results with the results obtained by the traditional identification technique.

Materials and Methods

Material

All the chemicals used are of analytical quality. The solvents and reagents employed were purchased commercially from Polychimie (Ivory Coast). For thin-layer chromatography, we used silica gel 60 F254 chromatoplates on an aluminum support. The different devices used for our analyzes are as follows:

- A RETSCH brand electric grinder, type SM 100
- A brand balance Sartorius (Germany)
- Burchi RII Rotary Evaporator

- A magnetic stirrer
- A UV-Visible spectrophotometer
- A camera

  - ✓ Spatial resolution: 2448×3264 (8 M pixels)
  - ✓ Tonal resolution: 24 bits (color)
  - ✓ Color space: RGB
  - ✓ Camera model: Infinix
  - ✓ Focal length: F/2
  - ✓ Exposure time: 1/13 seconds
  - ✓ ISO sensitivity: ISO-864
  - ✓ Focal distance: 4 mm
  - ✓ Picture format: JPG image

- Glassware (round bottom flask, beakers, etc.)
- UV lamp (366 nm, 254 nm)

Plant Material

Plant organs consisting mainly of leaves were all harvested from the forest relic of Nangui Abrogoua University in Côte d'Ivoire (Fig. 1). The plant species have been identified according to the herbaria available at the National Center for Floristics (NCF).

Fig. 1: (a) Leaves of Alchornea cordifolia (b) Leaves of paullinia pinnata (c) Leaves of Trema guineensis
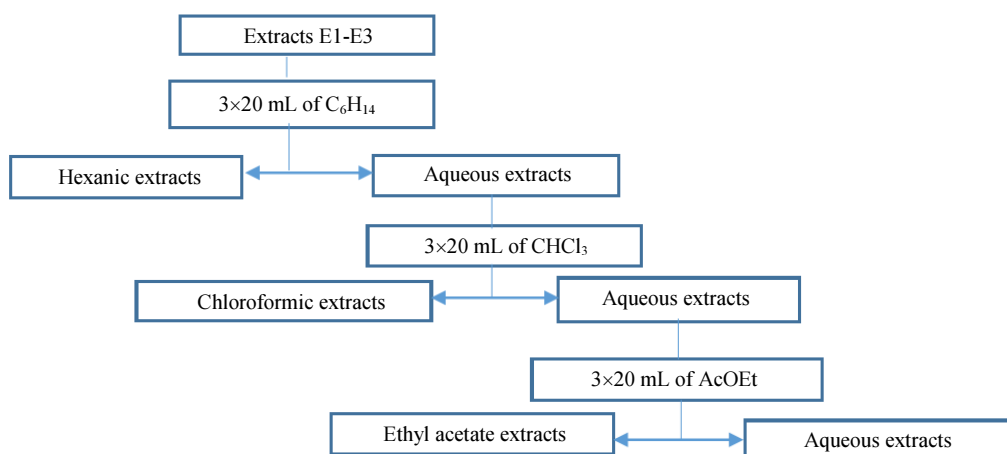

Fig. 2: Synoptic diagram of the preparation of selective extracts

1210

Methods

The hydromethanolic extracts of the various leaves were treated respectively with 3×20 mL of hexane, chloroform and ethyl acetate (Fig. 2). The various organic fractions were concentrated under reduced pressure on a rotary evaporator and then stored in a refrigerator. Thus, for each crude extract, we obtain selective extracts distributed as follows: Hexanic, chloroformic, ethyl acetate which were used for phytochemical screening.

Photochemical Screening and Image Acquisition Device

The photochemical screening was carried out according to the analytical procedures described in the work of (Obeng, 2010; Zimudzi and Cardon, 2013; Benkiki, 2012).

Drops of each selective extract are deposited using a capillary on the points (01 cm from each other) of the baseline drawn 1 cm from the bottom of the chromatoplates (silica gel 60 F254). The deposits are allowed to dry for a moment before introducing the plates into the chromatograph vessel containing the migration solvent. After migration and drying with a hair dryer, the chrotograms are revealed with reagents specific to the phytochemistry groups sought, then visualized first in the visible and then under a UV lamp at 366 nm (Fig. 3). The different developer systems used are:

- Hexane/AcOEt (20: 4, v/v) for hexanic fractions
- CHCl3/AcOEt/hexane (10: 10: 5, v/v/v) for the chloroform fractions
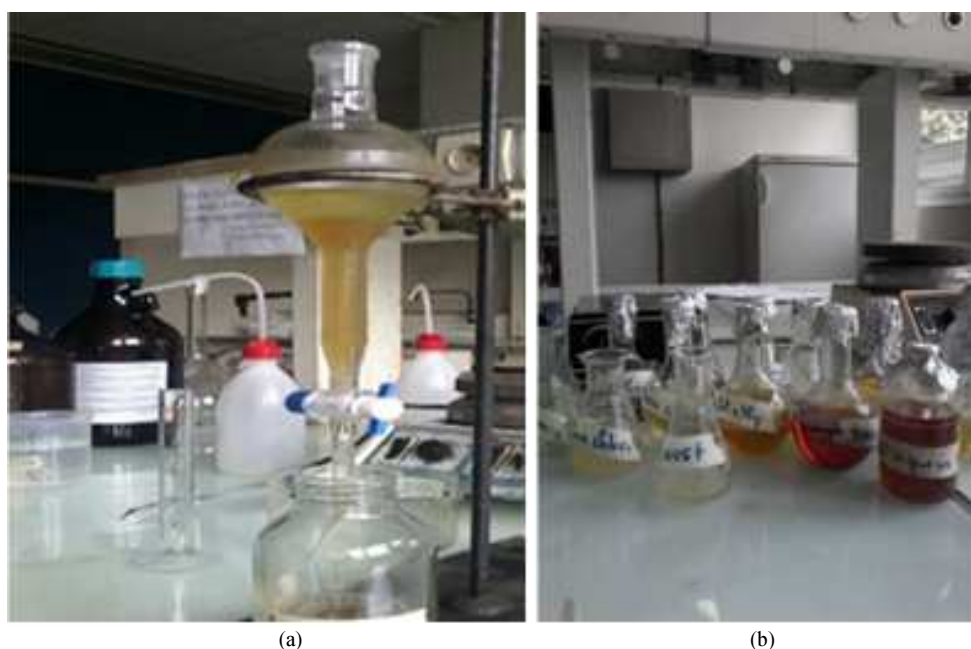- AcOEt/CH3OH/H2O/CHCl3 (18: 2,4: 2,1: 6, v/v/v/v) for the ethyl acetate fractions

Fig. 2: (a) Fractionation process of the hydromethanic extract (b) Hexanic fraction, chloroform, ethyl acetate
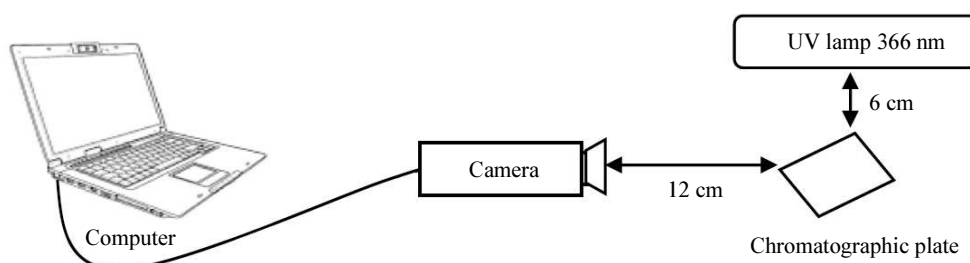

Fig. 3: Image acquisition device on the chromatographic plate

1211

Table 1: Values of the frontal ratios and the colors of the families of molecules

Extract	Without developer Visible Rf	Color	366 nm Rf	Color	Liebermann Buchard Visible Rf	Color	366 nm Rf	Color	KOH Visible Rf	Color	366 nm Rf	Color	Types of compounds
E_1^I					0,33	Blueviolet							Triterpenic Genes
							0,11	Red					Oleanane and ursane
							0,93						Triterpenes type
							0,68	Yellow-					Triterpene of
							0,63	Orangé					Lupan type
					Blue--	0,42	Yellow	Yellow-vert					Sterol steroids
	0,12	Green-Yellow			Yellow--	0,12					0,12	yellow	Cholesterol
									0,52	yellow		Yellow	Coumarin
E_2^I	0,12	Green	0,12	Red					0,12	Green	0,12	Red	Chlorophylle Sterol
			0,00	Jaune									Sterol
	0,65	Green-Yellow			0,65	Brown	0,65	Brown			0,65	Yellow	Sterol
													Polyterpene
			0,60	o-	0,60	Brown	0,60	Yellow					Triterpenic Genes
					0,55	Violet	0,55	Yellow-					
											0,53	Bleu	Coumarin Sterol
					0,63	Violet	0,63	Yellow					Triterpenic Genes
			0,60	B--	0,55	Violet	0,55	Yellow-					
					0,60	Violet	0,65	yellow-					
E_3^I	0,05	j-			0,05	Brown	0,05	Brown	0,05	Yellow	0,05	Orange	Stérol
	0,12	Green	0,12	Red	0,12	Green	0,12	Gray	0,12	Green	0,12	Red	Triterpenic Genes
									0,54	Blue			Coumarin Sterol
					0,27	Brown							
E_1^{II}	0,75		0,75	R					0,75	Yellow	0,75	Brown	Tannin
	0,48		0,48	Blue					0,48	Yellow	0,48	Yellow	Flavonoid
	0,50	Green	0,50	Brown									Polyphenol
E_2^{II}	0,51		0,51	Red					0,51	Gray	0,51	Brown	Tannin
	0,54	J--	0,54	Blue									Coumarin
			0,16	Red					0,16	Green--	0,16	Red--	Hydroflavonol
	0,80	Ve-J	0,80	R									Flavonoid
					0,82	Yellow							Flavonoid
E_3^{II}	0,00	Brown	0,51	Red-					0,00	Orange	0,00	Orange	Tannin
	0,03	Brown-	0,03	Red-					0,03	Brown-			Rutine
	0,54	Green-	0,54	Blue									Coumarine
	0,48		0,48	Red					0,27	Green	0,27	Red-	Flavonoide
													Hydroflavonol
E_1^{III}	0,00	M-	0,00	Brown					0,00	Yellow	0,00	Brown-	Rutin
			0,77	Blue-							0,77	Blue	Tannin
			0,25	Blue-							0,25	Blue	
	0,91	Ve-J--	0,93	Bleu							0,93	Red	Flavonoid
E_2^{III}	0,00	Blue-	0,00	Brown					0,00	Red	0,00	Brown	Rutin, tannin
	0,51	Green	0,41	Red--									Flavonoid
			0,61	Blue--									Coumarin
	0,91	Green	0,89	Red					0,91	Green	0,91	Red	Flavonoid
E_3^{III}	0,00	Brown	0,00	Blue					0,00	Yellow	0,00	Brown	Tannin
	0,40	Yellow--	0,40	Green					0,40	Orange	0,40	Brown	
	Hydroxyflavonol												
	0,49	Yellow--	0,48	Yellow									
											0,47	Brown--	Flavonoid
											0,54	Blue--	Coumarin

- = clear; - - = very clear

Values and Colors of Molecule Obtained

In the various hexanic (E1), chloroformic (E2) and ethyl acetate fractions (E3), we identified the families of molecules under UV 366 nm and with developers.

The sterols revealed with the Liebermann-Bürchard reagent appear in the visible in blue and yellow but fluorescent yellow and yellow-green under UV 366 nm (Newman and Cragg, 2016). These are spots of Rf = 0,42; 0,45 in E1, Rf = 0,42; 0,45 in E2 and Rf = 0,42; 0.45 in E3. Triterpenes of lupan type if the spots have yelloworange fluorescence under UV at 366 nm (Yoshikawa *et al.*, 2014). The Table 1 gives the values of the frontal ratios of the families of molecules and the corresponding colors. These values are in agreement with the work of (Brou *et al.*, 2011).

Results and Discussion

Acquisition of Images on the Chromatographic Plate

For each extract, we obtained images on the chromatographic plates. The Fig. 4 shows some of the images obtained.
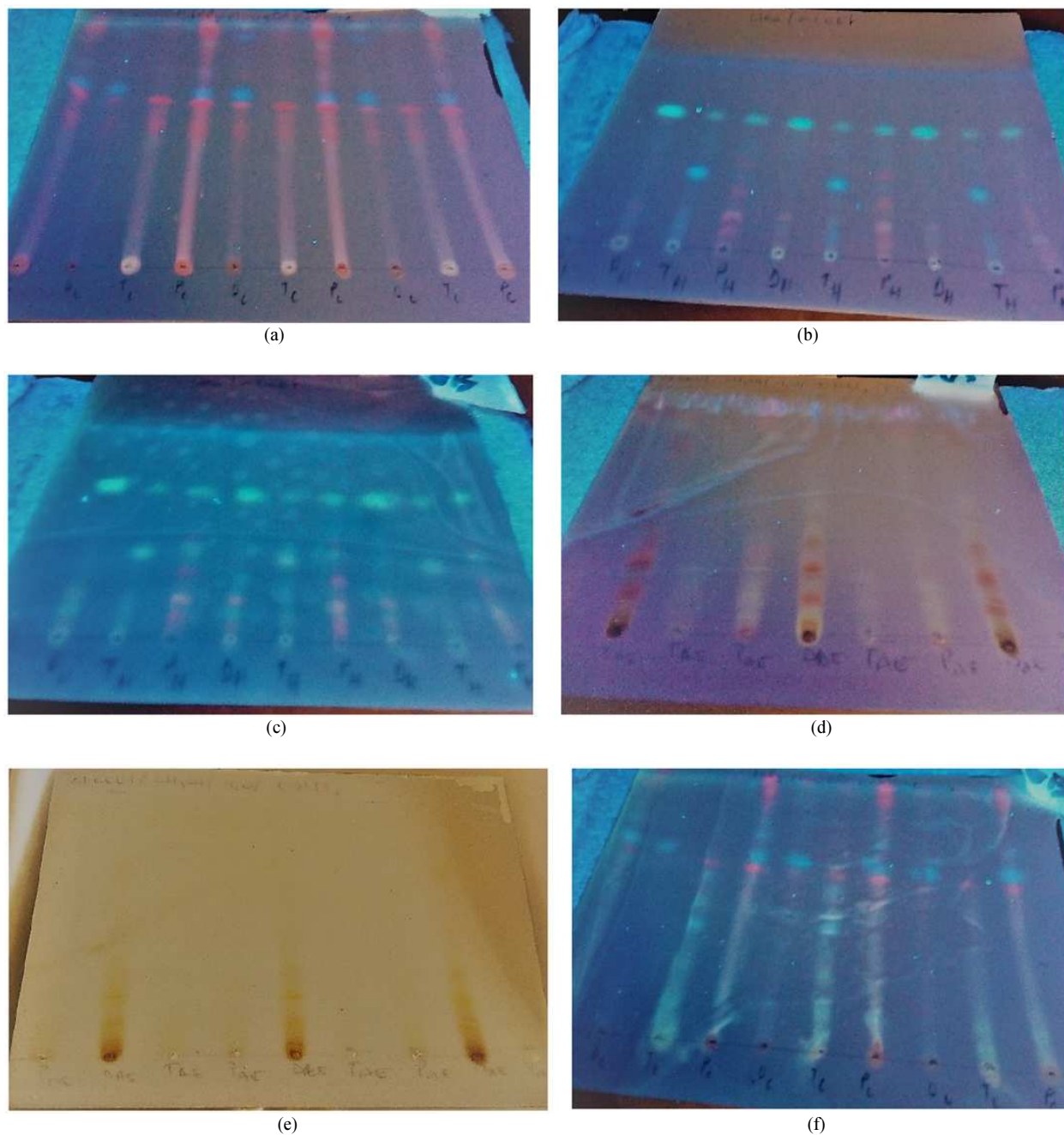
(a)

(b)

(c)

(d)

(e)

(f)

Fig. 4: (a) UV chromatogram at 366 nm chloroformic extracts (b) UV chromatogram at 366 nm of hexane extracts (c) UV chromatogram at 366 nm of hexane extracts after treatment with Libermann-Büchard (d) UV chromatogram at 366 nm Ethyl acetate extracts after treatment with KOH (e) Chromatogram in the visible ethyl acetate extracts after treatment with KOH (f) 366 nm UV chromatogram of the ethyl acetate extracts after treatment with Libermann-Büchard

Parameter Calculations

We segmented each spot of the images in Fig. 4 above, then calculated mean intensity, standard deviation, entropy and frontal report using the algorithms coded in the Matlab 2016 version running under Microsoft Windows 10 Pro on a PC based on

Intel i3. Figures 5-7 illustrate the segmentation of the different spots of the images used. We finally built a database with the values of the parameters of each task. The spots corresponding to families of molecules previously identified by the traditional technique of thin layer chromatography using developers.

1213

Fig. 5: Under image (spot) size 5×5
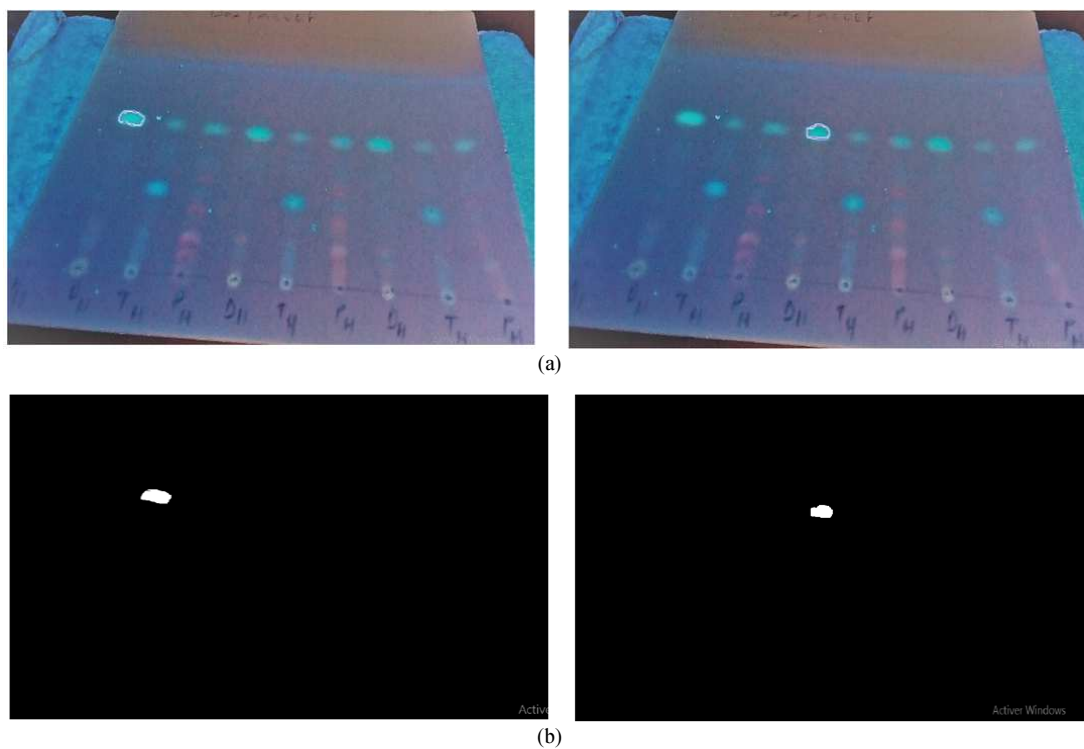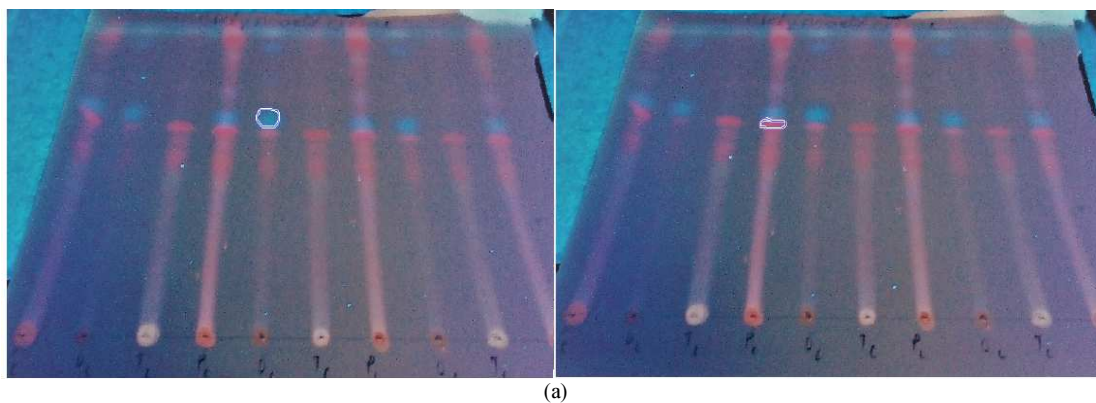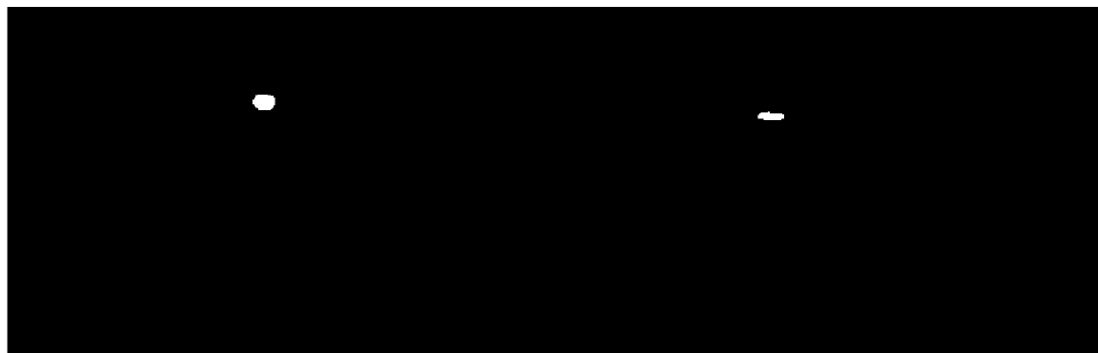

(a)

(b)

Fig. 6: (a) Segmentation of spots under UV 366 nm corresponding to families of molecules; (b) Segmented spots of molecule families

(a)

(b)

Fig. 7: (a) Segmentation of spots under UV 366 nm corresponding to families of molecules (b) Segmented spots of molecule families

Average Pixel Intensity

Each spot on the plate is a grayscale digital image and an array of values. Each box in this table, which stores a value, is called a pixel. Noting n the number of lines and p the number of columns of the image, we manipulate an array of n × p pixels. Figure 5, left, shows a visualization of a square array with n = p = 240, which represents 240×240 = 57600 pixels. For an image, we can calculate the average intensity of pixels characterizing this image (Gabriel, 2011).

The algorithm (a) below calculates the average pixel intensity of the spots (Fig. 6):

```
k = imshow(im1);
im2 = rgb2gray(im1);
h = imfreehand;                                    (a)
bw = createMask(h,k);
figure,imshow(bw)
ht = regionprops(bw,im2,'MeanIntensity');
at = ht.MeanIntensity
```

Standard deviation (std):

$$std = \left(\frac{1}{n-1} \sum\nolimits_{i=1}^{n} \left(x_i - \bar{x} \right) \right)^{1/2} \tag{1}$$

The algorithm (b) below allows to calculate the standard deviation of the spots (Fig. 7):

```
k = imshow(im3);
im2=rgb2gray(im3);
h = imfreehand;
bw = createMask(h,k);                              (b)
figure,imshow(bw)
ht = regionprops(bw,im2,'PixelValues');
ap = ht.PixelValues;
s = std(double(ap))
```

Entropy:

$$H(X) = -\sum\nolimits_{i=0}^{n} P_i \ \log_2 \left(P_i \right) \tag{2}$$

The algorithm (c) below calculates the average pixel intensity of the spots (Fig. 8):

```
k = imshow(im4);
im2 = rgb2gray(im4);
h = imfreehand;
bw = createMask(h,k);
figure,imshow(bw)
ht = regionprops(bw,im2,'PixelValues');           (c)
ap = ht.PixelValues;
E = entropy(ap)
```

Frontal Ratio

After the acquisition of the images (Fig. 9), the algorithm below on the Matlab software made it possible to calculate the frontal ratios of the different families of molecules:

```
[filename,pathname]=uigetfile('*.*','selectionne une image');
        cd(pathname)
        imshow(filename)
        [x,y] = ginput(3);
        h₁ = y(1)-y(2);                            (d)
        h₂ = y(1)-y(3);
        Rf = h₂/h₁;
```

Parameter Values

For a spot on the plate, we identify the corresponding molecule family by the traditional method using the developers and for this task, parameters are calculated constituting a database and corresponding to this family of molecule. The parameter values for the families of molecules are given in Table 2.

For the different fractions, the parameters have the same values, this allows us to develop the identification program (Fig. 10).
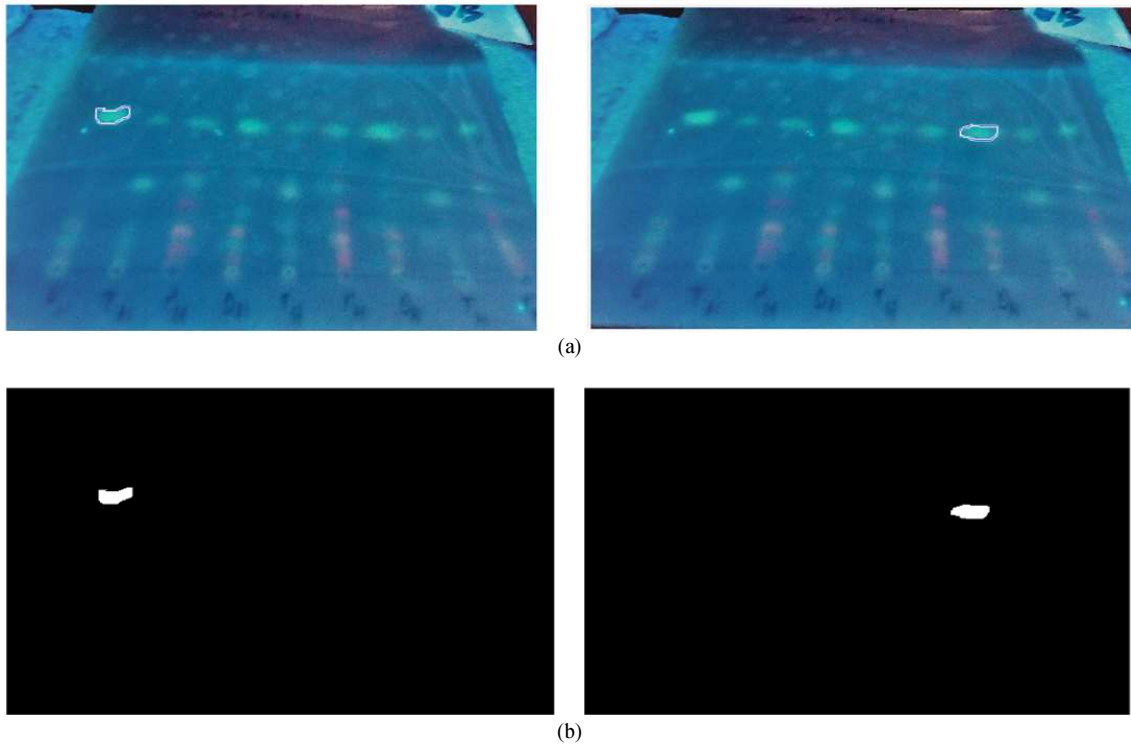
1215

(a)

(b)

Fig. 8: (a) Segmentation of spots under UV 366 nm corresponding to families of molecules; (b) Segmented spots of molecule families
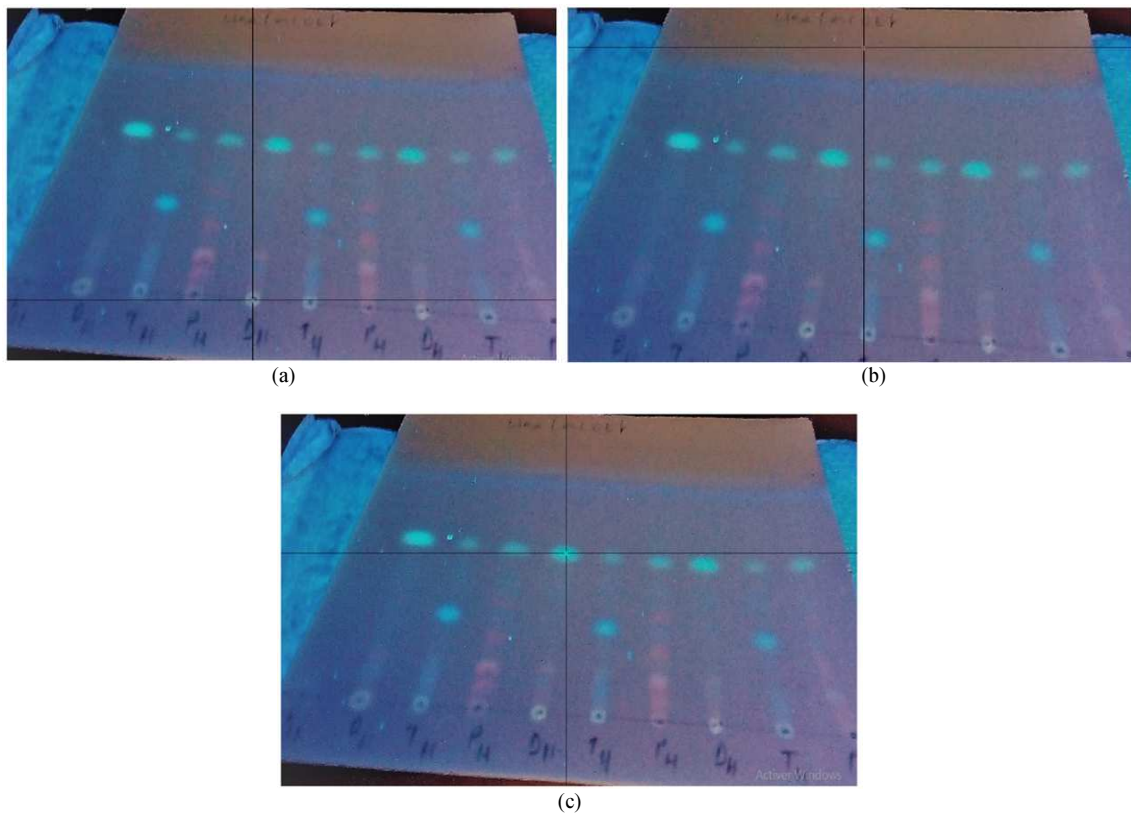

(a)

(b)

(c)

Fig. 9: (a) Deposit line (b) front line (c) height of the spot considered

1216

We tested the presence of coumarins, tannins and terpenes in the hexane fraction of leaves of Blighia unijugata Bak (Sapindaceae).

The protocol of the thin layer chromatography being the same, we made the segmentation of the different spots on the chromatographic plate. We then calculated the different parameters of the image processing namely entropy, front ratio (Rf), standard deviation (s) and mean pixel intensity (at) (Fig. 11 and 12). We obtained the results recorded in Table 3. These values correspond to molecules in the database.

Since the values of the various parameters are defined, we have compared these results with the values in the database.

From Table 3 we can confirm the presence of coumarins, tannins, terpenes in the leaves of Blighia unijugata Bak (Sapindaceae) because the values obtained correspond to the values in the database. These molecules have already been identified by the traditional method with the same plants (Kabran, 2014b).

These results thus show the reliability of this identification technique and its simplicity in its implementation.
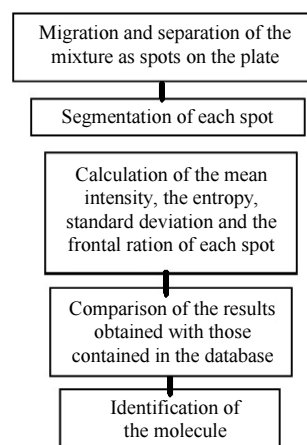

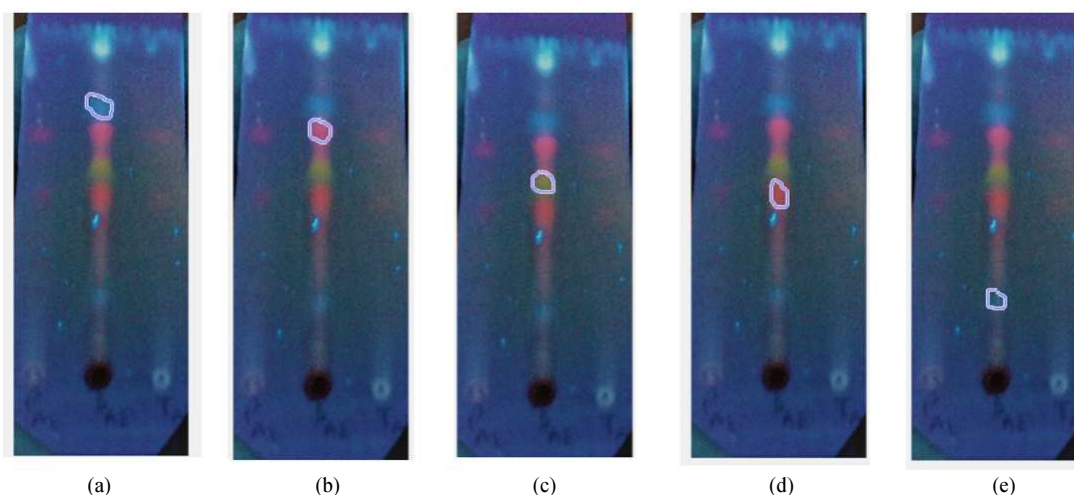**Fig. 10:** Program of identification of the molecule

Fig. 11: (a) Segmentation of the first spot (b) the second spot (c) the third spot (d) the fourth spot (e) fifth spot
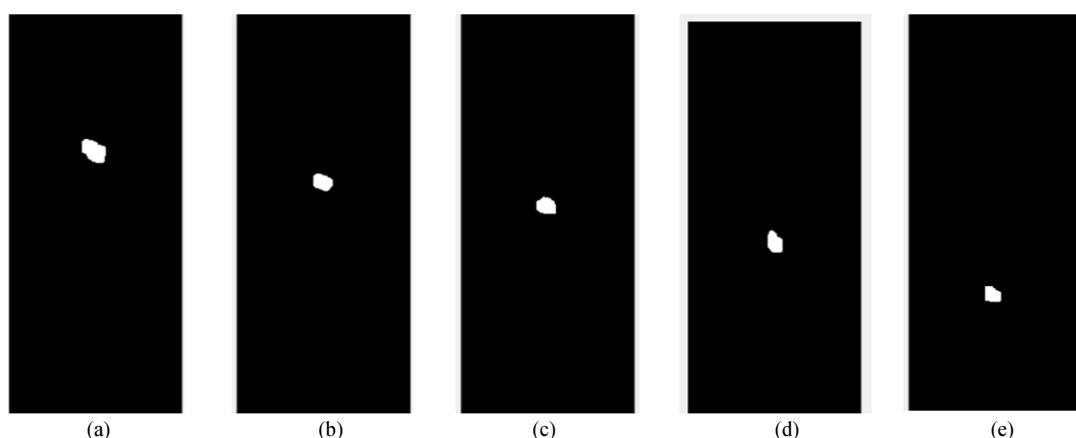

Fig. 12: (a) first segmented spot (b) second segmented spot (c) third segmented spot (d) fourth segmented spot (e) fifth segmented spot

1217

Table 2: Parameter values of the families of molecules

Familles de molecules	MeanIntensity (at)	Standard deviation (s)	Entropy (E)	Frontal ratio (Rf)
Coumarin	[81,700; 87,900]	[4,07; 4,60]	[2,04; 2,11]	[0,52; 0,61]
Tanin	[11,880; 12,710]	[1,57; 1,70]	[4,11; 4,61]	[0,75; 0,77]
Flavonoide	[201,78; 203,92]	[5,87; 5,90]	[3,11; 3,31]	[0,41; 0,51]
Génine triterpénique	[186,04; 196,10]	[4,11; 4,20]	[3,50 3,60]	[0,78; 0,80]
Hydroflavonol	[124,56; 131,60]	[4,50; 4,60]	[5,77; 5,87]	[0,55; 0,58]
Terpene	[134,56; 144,60]	[6,54; 6,60]	[5,90; 5,99]	[0,55; 0,70]
Steroid sterol	[201,88; 203,90]	[2,00; 2,11]	[4,46 4,56]	[0,42; 0,45]
Anthocyanin	[193,11; 194,20]	[4,01; 4,41]	[4,31; 4,41]	[0,12; 0,23]
Polyphenol	[196,04; 198,10]	[4,91; 4,93]	[1,65; 1,71]	[0,31; 0,32]

Table 3: Values for the extract parameters in the hexane fraction of leaves of Blighia unijugata

MeanIntensity (at)	Standard deviation (s)	Entropy (E)	Frontal ratio (Rf)
83,01	4,11	2,08	0,52
11,90	1,61	4,51	0,75
135,04	6,58	5,96	0,61

However we found that when the radius of the segmented spot is small, i.e., when the contour is near the center of the spot, we obtain very satisfactory results. Parameter values for coumarins were not very satirical because the diameters of the segmented areas were not very close to the center of spot.

Conclusion

We exposed a new thin-film family identification technique based on image segmentation and calculation of segmentation parameters such as entropy, standard deviation, mean pixel intensity and frontal ratio. Our method shows quite satisfactory results and have been confronted with the results of the technique traditionally made from the developers and the UV lamp 366 nm and 254 nm. The results of this confrontation showed a similarity between the families of molecules obtained. This shows the reliability and robustness of our technique. In our next work, it will be a question of designing an electronic device coupled to a computer allowing to calculate automatically the entropy, the mean intensity and the standard deviation of each spot in order to identify the family of molecule.

Author's Contributions

Theodore Guié Toa Bi and Sié Ouattara: Performed acquistion of images, simulation work and conducted analysis of the research. All authors contributed to the writing of the manuscript.

Ethics

This article is an original research paper. The authors declare that there are no ethical issues that could arise after the publication of this study.

References

Adjanohoun, E. and L. Aké-Assi, 2008. Contribution to the inventory of medicinal plants of Ivory Coast. Minisère of Scientific Research, National Floristic Center, Abidjan, Ivory Coast.

ANSM, 2016. List a traditionally used medicinal plants - French pharmacopoeia. Agency for the Safety of Medicines and Health Products,

Benkiki, N., 2012. Phytochemical study of Algerian medicinal plants: Ruta Montana, Matricaria pubescens and Hypericum perfoliatum. Ph.D. Thesis, El-Hadj Lakhdar Batna University (Algeria).

Brou, K.G., J.A. Mamyrbekova-Bekro, D.O. Dogbo, S.J. Gogbeu and Y.A. Bekro, 2011. On the qualitative phytochemical composition of crude hydromethanolic extracts of the leaves of 6 varieties of manihot esculenta crantz of ivory coast. Eur. J. Scientific Res., 45: 200-211.

Canvalho, K.R., A.B. Silva, M.C.M. Torres, F.C.L. Pinto and L.A. Guimaraes *et al.*, 2015. Cytotoxic alkaloids from hippeastrum solandriflorum lindl. J. Brazil. Chem. Society, 26: 1976-1980.
DOI: 10.5935/0103-5053.20150176

Djeddi, S., A. Karioti, E. Yannakopoulou, K. Papadopoulos and R. Chatter *et al.*, 2013. Analgesic and antioxidant activities of Algerian retama raetam (Forsk.) Webb and Berthel extracts. Records Natural Products, 7: 169-176.

Gabriel, P., 2011. Digital processing of images. Images of mathematics.

Kabran, 2014a. Chemical and cytotoxic study of ten Ivory Coast plants, used in the traditional treatment of breast cancer.

Kabran, 2014b. Chemical and cytotoxic study of ten plants from Côte d'Ivoire, used in the traditional treatment of breast cancer.

Newman, D.J. and G.M. Cragg, 2016. Natural products as sources of new drugs from 1981 to 2014. J. Natural Products, 79: 629-661. DOI: 10.1021/acs.jnatprod.5b01055

Obeng, E.A., 2010. Blighia unijugata Baker. [Internet] Record from Protabase. Lemmens.

Traore, L., Y. Bekro, J. Pirat and J.A. Mamyrbeva-Bekro, 2015. Study of crude extracts from Cassia sieberiana root bark and Khaya grandifoliola trunk bark: Phytochemical screening, quantitative analysis and radical scavenging activity. Int. J. Curr. Pharmaceutical Res., 7: 22-26.

Yoshikawa, K., M. Okahuji, K. Iseki, T. Ito and Y. Asakawa *et al.*, 2014. Two novel aromatic glucosides, marylaurencinosides D and E, from the fresh flowers of Cymbidium Great Flower 'Marylaurencin'. J. Natural Med., 68: 455-458. DOI: 10.1007/s11418-013-0814-8

Zimudzi, C. and D. Cardon, 2013. Morinda lucida Benth. [Internet] Record from PROTA4U. Jansen, P.C.M. and D. Cardon (Eds.), Plant Resources of Tropical Africa (PROTA), Wageningen, Netherlands.