

Klasifikasi Citra Digitalretina Penderita Diabetes Retinopati Menggunakan Metode Euclidean

Yusuf Fadlila Rachman¹, Kusrini¹, Hanif Al Fatta¹

¹Magister Teknik Informatika, Universitas Amikom, Yogyakarta

Email: ¹yusuf.rachman@students.amikom.ac.id, ¹kusrini@amikom.ac.id,

¹hanif.a@amikom.ac.id

Abstrak. Retinopati diabetes merupakan gangguan yang terjadi pada mata, yang disebabkan oleh penyakit diabetes. Diagnosa Retinopati Diabetes dapat dilakukan menggunakan citra digital retina. Pada penelitian ini dilakukan klasifikasi dan analisa tingkat keparahan Retinopati Diabetes berdasarkan ciri atau fitur gambar retina. Terdapat 2 dataset yang digunakan. Dataset pertama merupakan hasil dari preprocessing gambar yaitu konversi gambar ke *grayscale*. Dataset kedua didapat dengan menghapus optikal disk pada gambar dataset pertama. Penelitian ini menggunakan ekstraksi ciri tekstur metode statistik orde pertama. Ekstraksi ciri orde pertama mengambil ciri citra berdasarkan karakteristik histogram citra. Ciri gambar yang didapatkan adalah *mean*, *variance*, *skewness*, *curtosis*, dan *entropy*. Klasifikasi dilakukan sebanyak 2 kali pada dataset pertama dan kedua menggunakan metode klasifikasi jarak euclidean. Pengujian dilakukan menggunakan confusion matrix dengan menghitung nilai akurasi, presisi, dan recall. Berdasarkan hasil pengujian didapatkan bahwa dataset kedua memiliki nilai akurasi yang lebih tinggi yaitu 64,81%, nilai precision 69,3%, dan recall sebesar 64,8%.

Kata kunci: Retinopati Diabetes, Ekstraksi Ciri, Klasifikasi, *Euclidean*

Abstract. *Diabetic retinopathy is a disorder that occurs in the eye, which is caused by diabetes. Diagnosis of Retinopathy Diabetes can be done using a retinal digital image. In this study the classification and analysis of the severity of Diabetes Retinopathy were carried out based on the features or features of the retinal image. There are 2 datasets used. The first dataset is the result of image preprocessing, namely the conversion of images to grayscale. The second dataset is obtained by removing the optical disk in the first dataset image. This study uses the first-order statistical feature texture extraction. First-order feature extraction takes image characteristics based on image histogram characteristics. Image features obtained are mean, variance, skewness, curtosis, and entropy. Classification is done 2 times in the first dataset and both of them use the euclidean distance classification method. Testing is done using a confusion matrix by calculating the value of accuracy, precision, and recall. Based on the results of testing, the second dataset has a higher accuracy value of 64.81%, a precision value of 69.3%, and a recall of 64.8%.*

Keywords: Diabetes retinopathy, Feature Extraction, Classification, Euclidean

Pendahuluan

Salah satu penyakit akibat dari komplikasi penyakit diabetes adalah Diabetes Retinopati (DR). Diabetes Retinopati ini merusak bagian retina mata dan mengakibatkan terganggunya penglihatan penderita. Dampak paling parah dari komplikasi ini akan menyebabkan

kebutaan permanen jika terlambat ditangani[1]. Penderita yang mengalami diabetes retinopati akan mengalami gejala antara lain: *microneurysm*, *hemorrhages*, *soft exudates*, *hard exudates*, dan *neovascularis*. Berdasarkan gejala-gejala yang muncul, diabetes retinopati dapat diklasifikasikan ke dalam 4 fase (tingkat

keparahan), yaitu normal, *non-proliferative diabetic retinopathy (NPDR)*, *proliferative diabetic retinopathy (PDR)*, dan *macular edema (ME)*[2].

Pengenalan pola merupakan metode yang berguna untuk mendapatkan dan menganalisa pola yang dimiliki data dan menunjukkan suatu informasi tertentu. Metode pengenalan pola bekerja dengan membandingkan kemiripan suatu data berdasarkan informasi yang telah didapat. Salah satu tujuan penggunaan metode pengenalan pola adalah untuk mengklasifikasi data. Tujuan dari penelitian ini adalah untuk menganalisa dan mengklasifikasi fase penyakit retinopati diabetes berdasarkan fitur citra digital retina menggunakan metode euclidean. Adapun fitur-fitur citra yang digunakan adalah *mean*, *variance*, *skewness*, *kurtosis*, dan *entropy*.

Penelitian terdahulu oleh Sapriani Gustina, Abdul Fadlil, Rusydi Umar tahun 2017 menggunakan ekstraksi ciri statistik orde 1 untuk menentukan ciri dari gambar jamur. Tujuan penelitian ini untuk mengklasifikasi jenis jamur berdasarkan ciri yang didapat. Hasil yang diperoleh menunjukkan bahwa sistem memiliki akurasi sebesar 82%[3].

Penelitian selanjutnya oleh Darshankumar Gorasiya, Sanket Dayama tahun 2019. Penelitian bertujuan untuk







mengidentifikasi gejala awal pada pasien retinopati diabetes. Metode klasifikasi yang digunakan adalah CNN. Preprocessing dilakukan menggunakan kontras dan histogram. Akurasi sistem yang dihasilkan sebesar 89%[4].

Berdasarkan pemaparan diatas, maka penelitian ini akan membahas tentang implementasi metode ekstraksi ciri untuk klasifikasi citra digital retina pasien retinopati diabetes.

Landasan Teori Retinopati Diabetes

Retinopati diabetes merupakan salah satu gangguan pada mata yang disebabkan oleh diabetes. Penyumbatan pada pembuluh darah pada bagian retina mata dapat dipicu oleh Retinopati Diabetes[5]. Gejala yang ditimbulkan oleh komplikasi Retinopati Diabetes biasanya berhubungan dengan pembuluh darah, penurunan penglihatan secara bertahap, terlihat bercak hitam pada penglihatan, noda melayang pada penglihatan, penglihatan berbayang, sulit membedakan warna, dan nyeri pada mata. Retinopati Diabetes terbagi menjadi beberapa fase, yaitu fase normal, *fase non-proliferative diabetic retinopathy (NPDR)*, *fase proliferative diabetic retinopathy (PDR)*, dan *macular edema (ME)*[2]. Contoh citra digital retina dapat dilihat pada tabel 1.

Tabel 1. Contoh Citra Digital Retina

Fase Normal	Fase NPDR (Mild)	Moderate NPDR
		
Severe NPDR	PDR	ME
		

Citra Digital

Citra digital dapat dikatakan sebagai sebuah fungsi $f(x,y)$ yang memiliki ukuran M piksel dan N piksel. Memiliki x (baris) dan y (kolom) sebagai koordinat spasial. Amplitudo f pada koordinat (x,y) merupakan tingkat keabuan atau intensitas citra pada titik tersebut. Apabila semua

nilai x,y dan nilai amplitudo f berhingga dan bernilai diskrit, maka citra tersebut dapat disebut sebagai citra digital[6]. Fungsi citra digital $M \times N$ dapat digambarkan dalam bentuk matriks berukuran M baris dan N kolom seperti pada gambar 1.

$$f = [f(x,y)] = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0,M-1) \\ f(1,0) & f(1,1) & \dots & f(1,M-1) \\ \dots & \dots & \dots & \dots \\ f(N-1,0) & f(N-1,1) & \dots & f(N-1,M-1) \end{bmatrix}$$

Gambar 1. Matriks Citra Digital

Ekstraksi Ciri

Ciri (fitur) merupakan karakteristik unik yang membedakan antara satu objek dengan yang lain[7]. Tujuan dari ekstraksi fitur adalah mendapatkan informasi yang terkandung pada suatu citra untuk dijadikan pembeda antara satu citra dengan citra yang lain. Terdapat beberapa metode ekstraksi ciri, salah satunya adalah ekstraksi ciri statistik. Ekstraksi ciri statistik mengukur tingkat kekasaran, granularitas, dan kontras suatu daerah berdasarkan hubungan antar piksel dalam

citra. Terdapat 2 jenis ekstraksi ciri statistik, yaitu ekstraksi ciri orde pertama dan orde kedua. Ekstraksi ciri statistik orde pertama melakukan pengambilan ciri (fitur) citra berdasarkan karakteristik histogram citra. Berdasarkan histogram citra yang didapatkan, maka dapat dicari parameter ciri orde pertama. Adapun parameter ciri orde pertama yaitu, *mean*, *variance*, *skewness*, *kurtosis*, dan *entropy*. Untuk menghitung nilai masing-masing parameter ciri orde pertama dapat digunakan persamaan pada tabel 2.

Tabel 2. Ciri Statistik Orde Pertama

No	Ciri	Rumus	Keterangan
1	Mean	$\mu = \sum_n^n fn P(fn)$	Menunjukkan ukuran dispersi citra
2	Variance	$\sigma^2 = \sum_n^n (fn - \mu)^2 P(fn)$	Menunjukkan variasi elemen pada histogram citra.
3	Skewness	$\alpha_3 = \frac{1}{\sigma^3} \sum_n^n (fn - \mu)^2 P(fn)$	Menunjukkan tingkat kebengkokan relatif kurva histogram citra.
4	Curtosis	$\alpha^4 = \frac{1}{\sigma^4} \sum_n^n (fn - \mu)^4 P(fn) - 3$	Menunjukkan tingkat keruncingan relatif kurva histogram citra
5	Entropy	$H = - \sum_n^n P(fn)^2 \log P(fn)$	Menunjukkan ketidakaturan bentuk suatu citra

Keterangan: F_n merupakan nilai intensitas keabuan, $p(f_n)$ menunjukkan nilai histogram.

Metode Euclidean

Algoritma *K-Nearest Neighbors* (KNN) merupakan metode non prametrik yang digunakan untuk klasifikasi dan regresi[8]. Proses klasifikasi pada KNN dilakukan berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Kelebihan algoritma KNN adalah menghasilkan data yang jelas, kuat dan efektif jika digunakan pada data yang besar. Pada KNN proses penentuan nilai *k* bergantung pada data yang digunakan. Nilai *k* yang lebih besar berguna untuk mengurangi efek *noise* pada klasifikasi, tetapi menyebabkan batasan antar kelas sedikit berbeda.

Metode Penelitian

Penelitian yang akan diusulkan berjenis eksperimen. Penelitian dilakukan menggunakan data citra digital retina. Peneliti ini menggunakan pendekatan kuantitatif berupa citra digital retina yang telah melalui tahap pre-processing dan menghasilkan matriks numerik.

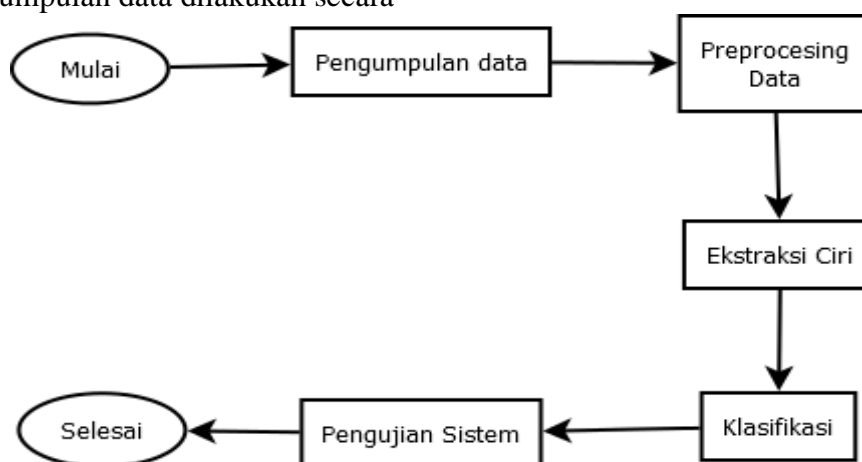
Rancangan Penelitian

Penelitian ini terbagi menjadi 3 bagian, yaitu Pengumpulan data, Preprocessing, Ekstraksi Ciri, Klasifikasi, dan Pengujian sistem. Pengumpulan data dilakukan secara

Untuk menghitung jarak antar data pada algoritma KNN dapat digunakan rumus jarak *Euclidean Distance* atau *Manhattan Distance*. *Euclidean Distance* bekerja dengan merepresentasikan cara berpikir manusia tentang jarak pada kehidupan nyata[9]. *Euclidean Distance* digunakan karena sesuai dengan karakteristik data yang *continue* dan memiliki dimensi yang jelas dan relevan.

$$d_{\text{Euclidean}}(x, y) = \sqrt{\sum_{k=1}^n (x_k - y_k)^2}$$

manual dengan menggunakan dataset *Messidor*. Preprocessing data dilakukan untuk memperbaiki kualitas citra, sedangkan ekstraksi ciri berguna untuk mendapatkan fitur-fitur gambar yang digunakan untuk klasifikasi. Proses klasifikasi dilakukan menggunakan metode *K-Nearest Neighbors*. Pengujian dilakukan dengan confusion matrix untuk menghitung nilai akurasi, presisi, dan recall yang akan dihasilkan. Alur penelitian dapat dilihat pada gambar 2:



Gambar 2. Diagram Alur Penelitian

Sumber dan Teknik Pengumpulan data

Penelitian ini menggunakan dataset yang berasal dari database *Messidor*. *Messidor* merupakan sebuah program penelitian yang dibiayai oleh Departemen Penelitian

dan Pertahanan Prancis pada tahun 2004[10]

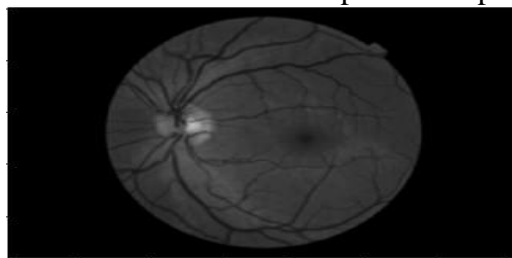
Pengolahan dan analisis data

Metode analisis data yang digunakan pada penelitian ini adalah analisa kuantitatif.

Proses klasifikasi citra yang dilakukan menghasilkan 4 kelas yaitu: kelas normal, kelas *Non-Proriferative Diabetic Retinopathy* (NPDR), kelas *Proliferative Diabetic Retinopathy* (PDR), dan kelas Macular Edema (ME). Data citra retina yang digunakan untuk klasifikasi akan melalui tahap *pre-processing*. Data hasil preprocessing kemudian akan memasuki tahap ekstraksi ciri untuk mendapatkan fitur-fitur citra yang digunakan dalam klasifikasi.

Klasifikasi data akan dilakukan sebanyak dua kali menggunakan 2 dataset yang berbeda. Dataset pertama adalah data gambar setelah dilakukan konversi citra ke *grayscale*. Dataset kedua adalah data gambar setelah dilakukan penghapusan optikal disk. Hasil dari kedua klasifikasi akan dibandingkan dan dianalisis mana yang terbaik.

Hasil dan Pembahasan Pengumpulan Data



Gambar 3. Konversi Citra *Grayscale*

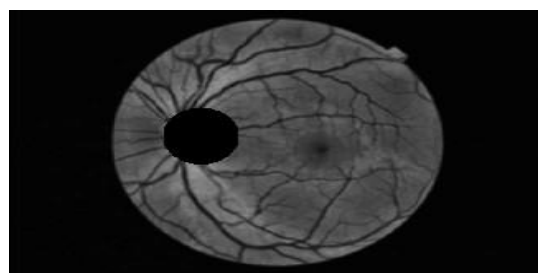
Langkah preprocessing selanjutnya adalah menghilangkan optikal disk pada gambar. Proses penghapusan optikal disk menggunakan dataset hasil preprocessing pertama. Penghapusan optik disk bertujuan untuk menghilangkan objek yang tidak terpakai dalam ekstraksi ciri. Metode yang digunakan dalam penghapusan optikal disk

Messidor menyediakan dataset yang dapat di gunakan untuk tujuan penelitian khususnya penelitian dengan topik Retinopati Diabetes. Data yang digunakan berupa citra digital retina yang diambil menggunakan kamera fundus 3CCD color video Topcon TRC NW6 Non-mydratic Retinograph dengan sudut pandang 45 derajat. Total data yang digunakan sebanyak 160 data. Data yang berhasil dikumpulkan kemudian akan melalui tahap preprocessing dan ekstraksi ciri sebelum kemudian dilakukan klasifikasi.

Preprocessing data

Terdapat beberapa langkah preprocessing data yang dilakukan pada penelitian ini. Langkah preprocessing yang pertama adalah merubah citra warna retina kedalam bentuk *grayscale* (keabuan). Data gambar hasil preprocessing pertama ini akan dijadikan sebagai dataset pertama. Hasil preprocessing konversi citra *grayscale* dapat dilihat pada gambar 3.

adalah *masking*. Metode ini membuat objek (kontur) yang tidak diinginkan tertutupi oleh *mask*. Data hasil preprocessing kedua ini digunakan sebagai dataset kedua. Hasil akhir gambar setelah menghapus optikal disk dapat dilihat pada gambar 4.



Gambar 4. Menghapus Optikal Disk

Ekstraksi Ciri

Pada tahap ini digunakan ekstraksi ciri tekstur metode statistik orde pertama. Adapun ciri (fitur) yang dicari adalah

mean, variance, kurtosis, skewness, dan entropy. Hasil ekstraksi ciri pada dataset pertama dapat dilihat pada tabel 3.

Tabel 3. Ekstraksi Ciri Dataset Pertama (Konversi Grayscale)

No	citra_retina	mean	variance	skewness	curtosis	entropy	target
1	20051019_38557_0100_PP.tif	41,069	7,839	3324,666	68,059	3,175	me
2	20051020_43906_0100_PP.tif	52,359	8,032	5950,999	70,634	3,257	me
3	20051020_44349_0100_PP.tif	33,553	9,291	2456,283	93,912	2,586	me
4	20051020_44598_0100_PP.tif	32,709	9,846	1771,064	104,82	3,058	me
5	20051020_44636_0100_PP.tif	28,576	9,164	1563,003	92,911	2,716	me
...
157	20060410_46893_0200_PP.tif	44,943	12,4	3983,335	173,336	2,941	npdr
158	20060410_47016_0200_PP.tif	59,96	13,94	6786,659	207,57	2,605	npdr
159	20060410_47166_0200_PP.tif	44,02	13,901	3962,656	206,377	2,73	npdr
160	20060410_47331_0200_PP.tif	51,781	13,483	4353,421	195,727	3,092	npdr

Proses ekstraksi ciri selanjutnya dilakukan menggunakan gambar hasil preprocessing.

Hasil ekstraksi ciri pada dataset kedua dapat dilihat pada tabel 4.

Tabel 4. Ekstraksi Ciri Dataset Kedua (Menghapus Optik Disk)

No	citra_retina	mean	variance	skewness	curtosis	entropy	target
1	20051019_38557_0100_PP.tif	45.676	2651.255	8.553	93.127	3.925	me
2	20051020_43906_0100_PP.tif	54.264	3735.06	9.015	99.929	4.076	me
3	20051020_44349_0100_PP.tif	37.116	1653.681	5.967	43.289	3.368	me
4	20051020_44598_0100_PP.tif	36.252	1610.379	6.562	55.226	3.664	me
5	20051020_44636_0100_PP.tif	31.652	1246.735	7.009	61.937	3.745	me
...
157	20060410_46893_0200_PP.tif	50.066	2889.699	6.797	60.404	3.706	npdr
158	20060410_47016_0200_PP.tif	65.205	4879.39	9.022	106.101	3.637	npdr
159	20060410_47166_0200_PP.tif	49.993	2894.901	9.961	124.17	3.723	npdr
160	20060410_47331_0200_PP.tif	57.73	3807.349	7.802	82.673	3.644	npdr

Klasifikasi

Klasifikasi dilakukan menggunakan algoritma k-nearest neighbors. Klasifikasi dilakukan sebanyak dua kali menggunakan nilai k=10. Klasifikasi dilakukan menggunakan metode split. Metode split berguna untuk membagi data menjadi data training dan data testing. Adapun presentasinya yaitu 66% sebagai data training dan 34% sebagai data testing.

Proses pertama akan mengklasifikasi menggunakan dataset pertama, yaitu dataset hasil preprocessing konversi grayscale. Sedangkan proses kedua akan dilakukan klasifikasi menggunakan dataset kedua yaitu dataset hasil preprocessing penghapusan optikal disk. Hasil klasifikasi dataset pertama dan kedua dapat dilihat pada tabel 5 dan tabel 6.

Tabel 5. Hasil Klasifikasi Dataset Pertama

Klasifikasi	Normal	NPDR	PDR	ME
Normal	8	0	2	4
NPDR	3	8	1	1
PDR	5	0	3	4
ME	6	1	0	8

Tabel 6. Hasil Klasifikasi Dataset Kedua

Klasifikasi	Normal	NPDR	PDR	ME
Normal	9	0	2	3
NPDR	3	7	1	2
PDR	2	0	8	2
ME	2	0	2	11

Pengujian

Pengujian dilakukan menggunakan confusion matrix dengan menghitung nilai *accuracy*, *precision*, dan *recall*. Terdapat 2 model knn yang akan di uji coba. Model

pertama adalah klasifikasi gambar tanpa preprocessing dan kedua klasifikasi gambar setelah preprocessing. Hasil pengujian menggunakan confusion matrix dapat dilihat pada tabel 7.

Tabel 7. Confusion Matrix

Data	Accuracy	Precision	Recall
Dataset 1	50%	55%	50%
Dataset 2	64,81%	69,3%	64,8%

Berdasarkan tabel 7, dapat dilihat bahwa hasil klasifikasi pada dataset kedua memiliki nilai yang lebih tinggi daripada dataset pertama. Nilai akurasi sebesar 64,81%, nilai precision sebesar 69,3%, dan recall sebesar 64,8%. Nilai akurasi menunjukkan bahwa kemampuan sistem dalam mengklasifikasi data dengan benar sebesar 64,81%. Nilai Precision berarti sistem memiliki nilai ketepatan sebesar 69,3% terhadap informasi yang telah diketahui terhadap jawaban yang dihasilkan oleh sistem. Nilai recall sebesar 64,8% adalah tingkat keberhasilan sistem dalam menemukan informasi.

Kesimpulan

Berdasarkan analisa pada hasil klasifikasi yang telah dilakukan, maka nilai akurasi, precision, dan recall terbesar didapatkan pada hasil klasifikasi menggunakan dataset kedua, yaitu 64,81%, 69,3% dan 64,8%.

Dataset kedua merupakan gambar hasil dari preprocessing konversi citra ke grayscale dan penghapusan optikal disk.

Nilai akurasi, precision, dan recall pada klasifikasi dataset kedua lebih tinggi disebabkan karena pada dataset pertama masih terdapat objek-objek yang dapat mengganggu proses ekstraksi ciri. Pada penelitian ini objek yang teridentifikasi berupa optik disk pada gambar retina. Maka untuk, meningkatkan nilai akurasi dilakukan penghapusan optik disk tersebut

Penghapusan optikal disk bertujuan untuk menghilangkan objek ataupun kontur yang tidak diperlukan dalam ekstraksi ciri. Hasil klasifikasi dataset pertama lebih rendah dibanding klasifikasi dataset kedua dikarenakan, masih adanya kontur ataupun objek yang tidak diperlukan dalam proses ekstraksi ciri.

Saran

Berdasarkan pemaparan diatas, dapat diketahui bahwa akurasi dapat dihasilkan

sebesar 64,81%. Nilai akurasi ini mungkin dapat ditingkatkan dengan menambahkan langkah-langkah pada tahap preprocessing data. Salah satu metode yang dapat digunakan salah satunya adalah peningkatan kontras (*contrast enhancement*) ataupun *gaussian blur*.

Daftar Pustaka

- [1] Dillak, R. H. (2011). Klasifikasi Fase Retinopati Diabetes Menggunakan Backpropagation Neural Network. *IJEIS*, 89-100.
- [2] Fadzil, A. I. (2011). Analysis of Retinal Fundus Images for Grading of Diabetic Retinopathy Severity. *Med. Biol. Eng. Comput.*
- [3] Gustiani, S. F. (2017). Sistem Identifikasi Jamur Menggunakan Metode Ekstraksi Ciri Statistik Orde 1 dan Klasifikasi Jarak. *Tekno COM*, Vol 16. No 4. 378-386
- [4] Gorasiya, D. D. (2019). Early Detection of Curable Diabetic Retinopathy from Retinal Images
- [5] Rahman, K. A. (2013). *Pedoman Penanganan Retinopati Diabetika*. PERDAMI Seminat Vitreoretina
- [6] Gonzalez, R. &. (2010). *Digital Image Processing, Third edition*. New Jersey: Prentice Hall.
- [7] Coltelli, Primo and Gualtieri, Paulo. (1990). A procedure for the extraction of object features in microscope images. *International Journal of Bio-Medical Computing*, Volume 25, Issues 2–3, Pages 169-176.
- [8] Altman, N. S. (1992). An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression. *The American Statistician*.
- [9] Larose D, T. (2005). *Discovering knowledge in data : an introduction to data mining*. Jhon Wiley & Sons Inc.
- [10] Anonim. (2004). *adcis.net*. Retrieved from Messidor Consortium: adcis/en/third-party/messidor/