

PERANCANGAN SISTEM DETEKSI KATARAK MENGGUNAKAN METODE *PRINCIPAL COMPONENT ANALYSIS (PCA)* DAN *K-NEAREST NEIGHBOR (K-NN)*

Analisis Performansi Sistem Deteksi Katarak

Riski Wahyu Hutabri ¹⁾, Ir. Rita Magdalena, M.T. ²⁾, R Yunendah Nur Fu'adah, S.T., M.T. ³⁾

^{1),2),3)} Teknik Telekomunikasi, Universitas Telkom Bandung
Jl. Telekomunikasi No.1 Terusan Buah Batu Bandung
Email : riskiwahyuas@gmail.com

Abstrak. Katarak adalah kondisi mengeruhnya bagian lensa mata yang biasanya jernih tertutupi dengan suatu noda putih yang dapat mengganggu penglihatan. Katarak muncul disebabkan oleh berbagai hal, salah satunya adalah penuaan. Katarak adalah salah satu penyebab kebutaan tertinggi di Indonesia. Pemeriksaan katarak dilakukan oleh tenaga ahli medis menggunakan alat slit lamp yang masih sulit ditemukan di setiap rumah sakit atau balai pengobatan sejenis. Oleh karena itu perlu adanya sistem yang dapat mendeteksi katarak yang dapat menolong masyarakat untuk lebih cepat dalam mengetahui kondisi matanya apabila mengalami gejala kabur dalam penglihatan sehingga dapat mengambil tindakan yang tepat, sistem deteksi katarak dirancang dengan menggunakan metode *Principal Component Analysis (PCA)* dan *K-Nearest Neighbour (K-NN)*. Pada penelitian ini, sistem dapat mendeteksi dan mengklasifikasikan ke dalam mata katarak imatur, katarak matur dan mata normal dengan akurasi 70,27% dan waktu komputasi rata-rata 0.331172937 detik.

Kata kunci : Katarak, PCA, K-NN

1. Pendahuluan

Tingginya tingkat kebutaan di Indonesia yang disebabkan oleh katarak sudah tidak diragukan lagi, masyarakat yang tinggal di kota besar biasanya dapat mengikuti bakti sosial sebagai tempat untuk mengikuti operasi katarak secara gratis. Menurut Wakil Wali Kota Bandung Oded M. Danial yang dilansir pada media pikiran-rakyat mengakui, kasus kebutaan di Kota Bandung yang disebabkan oleh katarak cukup tinggi. Karena itu, ia menyambut positif kegiatan bakti sosial untuk membantu pasien katarak yang kurang mampu [1]. Pada pendeteksian katarak, seorang dokter mata menggunakan *slit lamp* untuk mendiagnosis katarak, namun hal ini memerlukan teknik khusus dan harga yang mahal untuk sekali operasi.

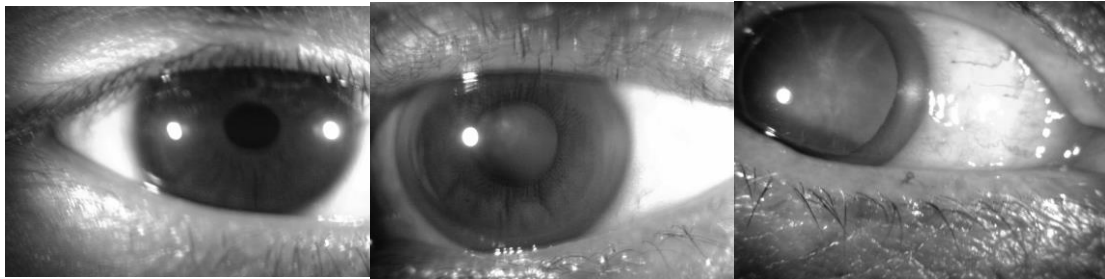
Indonesia sebagai sebuah negara berkembang memiliki jumlah dokter mata dan fasilitas yang terbatas. Beberapa peneliti mengembangkan aplikasi pendeteksi katarak berdasarkan pengolahan citra digital seperti dalam penelitian [2,3,4,5]. Penelitian yang telah dilakukan sebelumnya menunjukkan bahwa dapat dilakukan pendeteksian katarak menggunakan *image processing*. Pada penelitian ini, fokus utama adalah untuk menguji metode *Principal Component Analysis (PCA)* dan *K-Nearest Neighbor (K-NN)* dan mengoptimalkan tingkat akurasi menggunakan *statistical feature*.

Berdasarkan uraian fenomena diatas, penelitian ini bertujuan untuk dapat merancang sebuah sistem deteksi katarak berbasis matlab, menganalisis kombinasi *stastical feature* yang digunakan, menganalisis parameter yang mempengaruhi akurasi sistem, dan memilih parameter terbaik sehingga dapat membuat sistem deteksi dengan performansi yang baik.

2. Kajian Pustaka

A. Katarak

Katarak merupakan suatu keadaan mengeruhnya sebagian atau seluruh lensa mata sehingga cahaya tidak dapat menembus lensa. Kondisi ini membuat pandangan mata penderita katarak terganggu dan melihat benda secara kabur seperti tertutup kabut [2]. Mata normal memiliki kondisi lensa mata yang jernih atau tidak terdapat kekeruhan pada bagian lensa mata. Terdapat beberapa jenis katarak, fokus penelitian ini adalah pada katarak senilis. Katarak senilis merupakan katarak pada orang lanjut usia diakibatkan oleh proses penuaan [2]. Katarak senilis terbagi atas beberapa jenis yaitu, katarak imatur, katarak matur, dan katarak hiper matur. Adapun yang menjadi fokus penelitian ini adalah mengklasifikasikan ke dalam imatur, matur dan mata normal.



Gambar 1. Mata Normal

Gambar 2. Mata Imatur

Gambar 3. Mata Matur

B. Pengumpulan Data

Pengumpulan data menggunakan *digital iriscope* yaitu sebuah alat yang mengambil citra mata secara langsung dengan jarak yang tetap, dan pencahayaan yang cukup sehingga pengambilan citra menjadi lebih efisien.

C. Citra Digital

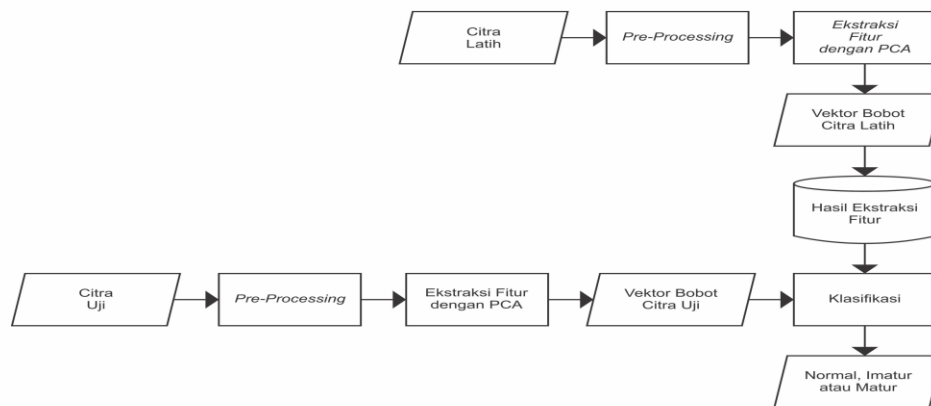
Citra digital adalah sebuah larik (*array*) yang mempunyai nilai-nilai real maupun kompleks yang direpresentasikan dengan deretan bit tertentu [3].

D. Citra RGB dan Citra Berskala Keabuan

Citra RGB adalah suatu ruang warna yang memiliki susunan warna yang luas yang dibentuk oleh komponen warna merah, hijau, biru. *Pixel depth* adalah jumlah bit yang digunakan untuk mewakilkan setiap *pixel* dalam ruang RGB. Setiap warna pokok mempunyai intensitas dengan nilai maksimum 8-bit atau 255. Citra dengan aras keabuan adalah citra yang mempunyai satu nilai kanal pada setiap *pixel*, dengan kata lain citra yang hanya menggunakan warna pada tingkatan warna abu-abu [3].

3. Metode

Diagram alir sistem ditunjukkan pada gambar 4. Menunjukkan proses umum dari sistem. Proses dilakukan pada citra latih dan citra uji dengan perlakuan yang sama, memiliki tahapan yang sama, penamaan, tahapan *pre-processing*, ekstraksi fitur dan klasifikasi. Penelitian menggunakan *statistical feature* untuk memperkecil nilai perbandingan ciri citra untuk digunakan pada pengklasifikasian.



Gambar 4. Diagram Alir Sistem Katarak

Pada perancangan sistem secara umum, sistem akan mengakuisi data dengan menggunakan *digital iriscope*. Data yang digunakan sebanyak 74 data, dibagi atas tiga kategori: matur (20 citra), imatur (34 citra), dan normal (20 citra). Data latih terdiri atas 10 citra matur, 17 citra imatur dan 10 citra normal. Selanjutnya tahapan *pre-processing* berupa *cropping* dengan mengambil pada bagian ROI atau *region of interest*, dilanjutkan dengan *resize* dengan mengubah citra menjadi beberapa ukuran yaitu 64 x 64 piksel, 128 x 128 piksel dan 256 x 256 piksel, setelah citra melewati proses *resize* dilanjutkan dengan *image enhancement* dengan mengubah gambar menjadi format keabuan, setelah citra melewati seluruh tahapan *pre-processing*, selanjutnya adalah ekstraksi fitur dengan menggunakan algoritma *Principal Component Analysis (PCA)* dengan mengambil nilai *eigen value* dan *eigen vektor* yang akan menjadi karakteristik dari citra dan citra akan diklasifikasikan menggunakan algoritma *K-Nearest Neighbor (K-NN)* dengan mencoba beberapa nilai k untuk menguji sistem.

A. Ekstraksi Fitur

Principal Component Analysis (PCA) adalah metode yang baik dalam menganalisis data dan pengenalan pola yang paling sering digunakan dalam pemroses sinyal atau citra. PCA adalah metode statistik dan dianggap sebagai transformasi linear sehingga digunakan dalam analisis data dan kompresi. PCA didasarkan pada representasi statistik dari variabel acak [4]. PCA akan bekerja dengan mencari *mean* dari citra, menentukan matriks normalisasi, menghitung *covariance* matriks, mengambil nilai *eigen value* dan *eigen vector*, dan menyusun nilai *eigen vector* menurut nilai *eigen value* dari yang terbesar menuju yang terkecil. Nilai *eigen vector* tersebut yang disebut dengan *principal component (PC)*. Berikut adalah nilai *eigen value* dari *covariance* matriks data citra katarak matur, terdapat 10 *eigen value*.

-4.23E-09	0	0	0	0	0	0	0	0	0
0	19009.53	0	0	0	0	0	0	0	0
0	0	56718.98	0	0	0	0	0	0	0
0	0	0	227132.1	0	0	0	0	0	0
0	0	0	0	321807.5	0	0	0	0	0
0	0	0	0	0	743948.7	0	0	0	0
0	0	0	0	0	0	1968932	0	0	0
0	0	0	0	0	0	0	3667796	0	0
0	0	0	0	0	0	0	0	3803462	0
0	0	0	0	0	0	0	0	0	68657035

Gambar 5. Nilai *eigen value*

68657035
3803462
3667796
1968932
743948.7
321807.5
227132.1
56718.98
19009.53
-4.23E-09

Gambar 6. Nilai *eigen value* yang disusun menurun dari nilai terbesar

3360.191	-722.403	-394.937	90.50293	345.8526	-29.1023	425.1802	2.388637	93.89825	6.65E+11
3817.543	-670.784	-18.198	-242.557	201.5572	535.1151	234.9701	-200.176	88.97411	7.66E+11
9195.882	-398.571	-395.867	-965.304	207.1726	139.6012	3.974487	-26.5112	95.4451	1.8E+12
1033.812	27.2035	-127.385	-47.2429	-8.59487	87.36205	-50.5226	-83.2993	181.912	2.02E+11
2611.871	-244.96	-9.12016	-314.28	56.83685	566.1577	155.6283	122.8184	112.9802	5.24E+11
2725.774	-2066.99	524.6345	-428.317	-234.726	163.8072	64.67252	-22.1427	94.2427	5.86E+11
8463.349	-205.445	440.4559	599.7823	-352.889	235.8878	151.164	-33.7547	91.39511	1.64E+12
3187.055	136.5793	258.2602	-977.215	-481.451	157.4144	379.3081	-58.8809	105.1717	6.37E+11
1133.284	91.1183	71.74524	-212.034	-7.80947	141.3194	-0.29422	-40.8055	-9.2421	2.25E+11
4268.459	-122.567	1806.126	-421.096	355.9098	163.3315	166.206	-32.3406	106.1003	8.59E+11

Gambar 7. Nilai *principal component*

B. Klasifikasi

K-Nearest Neighbor (K-NN) adalah sebuah metoda klasifikasi terhadap sekumpulan data berdasarkan pembelajaran data yang sudah terklasifikasikan sebelumnya. K-NN termasuk dalam metoda klasifikasi *supervised learning*, dimana hasil *query instance* yang baru diklasifikasikan berdasarkan mayoritas kedekatan jarak dari kategori yang ada dalam K-NN [2]. K-NN dapat digunakan dengan berbagai macam *distance*, pada penelitian ini penulis menguji *euclidean distance*, *chebychev distance*, *cityblock distance*, dan *minkowski distance*.

1. Euclidean distance

Euclidean distance adalah pengukuran untuk menemukan jarak diantara dua titik [6].

$$d_{st}^2 = (x_s - y_t)(x_s - y_t)' \dots\dots\dots(1)$$

2. Chebychev distance

Chebychev distance adalah pengukuran untuk menemukan jarak antar dua vektor atau titik dengan *standard coordinates* [6].

$$d_{st} = \max_j \{|x_{sj} - y_{tj}|\} \dots\dots\dots(2)$$

3. Cityblock distance

Cityblock distance adalah pengukuran jarak berdasarkan jumlah perbedaan dari seluruh dimensi [6].

$$d_{st} = \sum_{j=1}^n |x_{sj} - y_{tj}| \dots\dots\dots(3)$$

4. Minkowski distance

Minkowski distance adalah pengukuran yang menemukan jarak berdasarkan ruang *Euclidean* [6].

$$d_{st} = \sqrt[p]{\sum_{j=1}^n |x_{sj} - y_{tj}|^p} \dots\dots\dots(4)$$

4. Pembahasan

Pengujian sistem dilakukan dengan tujuan untuk mengetahui performansi sistem. Citra yang digunakan dalam format JPG, citra yang akan diuji sebanyak 37 data citra mata, terdiri atas 10 citra mata normal, 17 citra mata katarak imatur, dan 10 citra mata katarak matur. Adapun pengujian sistem menggunakan parameter akurasi dan waktu komputasi dengan skenario sebagai berikut.

1. Pengaruh kombinasi fitur pada performansi sistem.
2. Pengaruh nilai k pada saat pengklasifikasian menggunakan algoritma K-NN dan perubahan *distance metric*.

4.1 Hasil dan Analisis Sistem

Pengujian sistem dilakukan dengan mengubah parameter fitur ekstraksi ciri dan parameter klasifikasi untuk mendapatkan nilai akurasi yang lebih baik. Pada bagian ini akan dilihat pengaruh perubahan parameter fitur ekstraksi ciri dan klasifikasi terhadap akurasi yang diperoleh.

1. Pengujian Skenario 1

Pengujian fitur untuk mendapatkan fitur yang terbaik yang akan digunakan dalam perancangan sistem,

Tabel 1. Perbandingan hasil seluruh kombinasi fitur dari setiap pengujian

Fitur	Persentase Akurasi	Waktu Komputasi
Kurtosis	59,46%	0,280724015
Standard Deviation, Kurtosis	67,57%	0,307011085
Variance, Standard Deviation, Kurtosis	64,86%	0,308305433
Mean, Standard Deviation, Kurtosis	64,86%	0,334169094
Mean, Variance, Standard Deviation, Kurtosis	64,86%	0,298119415

Dapat disimpulkan dari tabel 1. menunjukkan bahwa fitur terbaik dilihat dari akurasi yang dihasilkannya adalah 67,57 %, fitur yang dapat digunakan dalam sistem deteksi katarak ini adalah fitur *standard deviation* dan *kurtosis*.

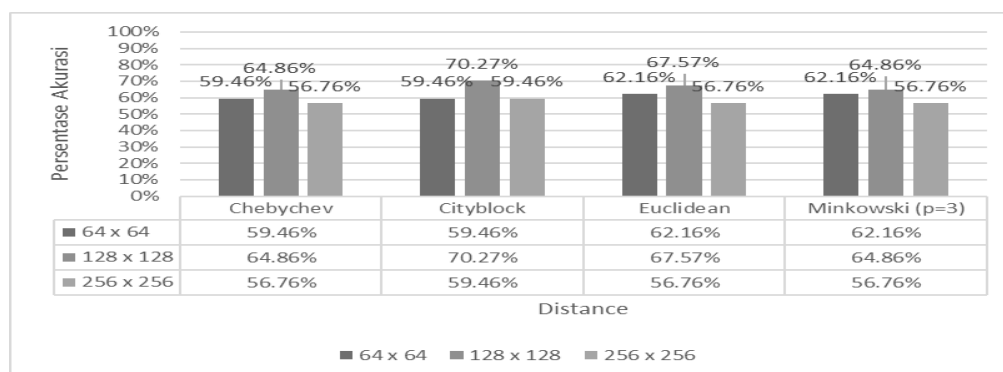
2. Pengujian Skenario 2

Pengujian dilakukan dengan menggunakan fitur *standard deviation* dan *kurtosis* dan ukuran yang tetap. Berikut adalah hasil pengujian nilai k pada K-NN terhadap hasil akurasi sistem.

Tabel 2. Pengaruh nilai k terhadap akurasi sistem

Nilai k	Persentase Akurasi	Waktu komputasi
k = 1	67,57 %	0,290186 s
k = 3	64,86 %	0,287798 s
k = 5	62,16 %	0,32056 s
k = 7	62,16 %	0,292825 s

Berdasarkan hasil pengujian nilai k diperoleh hasil akurasi tertinggi pada saat nilai k = 1. Disaat k = 3 akurasi sistem menurun sebesar 2,71 % dari 67,57 % menjadi 64,86 %, pada nilai k = 5 dan k = 7 akurasi sistem kembali mengalami penurunan. Dapat disimpulkan bahwa, semakin besar nilai k maka akan membuat batasan antar setiap klasifikasi semakin tidak jelas, hal ini sesuai dengan tabel 2. pada saat k = 1 akurasi memiliki hasil yang paling maksimal. Selanjutnya akan dilakukan pengujian pengaruh penggunaan *distance* terhadap akurasi sistem dengan menggunakan k = 1.



Gambar 8. Pengujian pengaruh *distance* terhadap akurasi sistem

Berdasarkan gambar 8. dapat disimpulkan bahwa perubahan pada jenis *distance* yang digunakan menunjukkan bahwa penggunaan jenis *distance* yang tepat pada sistem deteksi katarak ini adalah *cityblock* dengan akurasi tertinggi sebesar 70,27 % pada ukuran 128 x 128 piksel.

5. Simpulan

Pada penelitian ini, kita dapat mengambil kesimpulan kombinasi dari *statistical feature: standard deviation*, dan *kurtosis*. Nilai k optimal adalah 1. Akurasi sistem sebesar 70,27 %. Hasil penelitian ini menunjukkan bahwa metode yang diajukan dapat mengklasifikasikan kondisi mata ke dalam mata katarak imatur, katarak matur dan mata normal. Untuk penelitian selanjutnya diharapkan dapat lebih meningkatkan akurasi sistem dan mengimplementasikan sistem ke dalam *smartphone* berbasis Android.

Ucapan Terima Kasih

Alhamdulillahirrabal'alam, kata yang harus selalu kita ucapkan kepada Allah SWT atas kekuatan, kesabaran, kemudahan, kemampuan bahkan juga atas kelemahan, kegelisahan, kesulitan, dan ketidakmampuan, karena pada akhirnya Allah SWT yang membukakan jalan. Terimakasih kepada orang tua dan keluarga yang selalu mendukung untuk menyelesaikan *paper* ini. Ucapan terimakasih juga kepada pembimbing saya Ibu Rita Magdalena dan Ibu Yunendah Nur Fu'adah. Dan saya ucapkan terimakasih kepada teman-teman yang selalu mendukung saya.

Daftar Pustaka

- [1]. Bandung Raya. 9000 Warga kota Bandung Menderita Katarak, <http://www.pikiran-rakyat.com/bandung-raya/2017/11/03/9000-warga-kota-bandung-menderita-katarak-412921> diakses tgl 31 Desember 2017.
- [2]. Y. N. Fuadah, A. W. Setiawan, & T. L. R. Mengko, *Design and Implementation of Mobile Cataract Detection using Statistical Texture Analysis*, 2014.
- [3]. S. Pramesthi, A. Rizal, & R. D. Atmaja, Deteksi Penyakit Katarak Berbasis Perbandingan Piksel Citra Biner Dengan Menggunakan Android, 2013.
- [4]. A. Amira, S. Sourav, S. Abdel-Meggeed, D. Nilanjan, *Principal Component Analysis in Medical Image Processing : A Study*, 2015.
- [5]. R. Supriyanti, H. Habe, dan M. Kidode, *Extracting Appearance Information inside the Pupil for Cataract Screening*, MVA 2009 IAPR Conference on Machine Vision Application, 2009.
- [6]. C. Kittipong, C. Pasapitch, T. Ponsakorn, K. Kittisak, K. Nittaya, *An Empirical Study of Distance Metrics for k-Nearest Neighbor Algorithm*, Proceeding of the third International Conference in Industrial Application Engineering, 2015.