

Deteksi Kanker Serviks Otomatis Berbasis Jaringan Saraf Tiruan LVQ dan DCT

Dhimas Arief Dharmawan¹

Abstract—Cervical cancer has become the common women disease in the world. Mostly, cervical cancer has been already known lately, because it is very difficult to detect this in early stage. In this work, a computer based software using Learning Vector Quantization (LVQ) has been designed as the early cervical cancer detection aid tool. There are six methods before the detection is performed, namely preprocessing, contrast stretching, median filtering, morphology operation, image segmentation, and Discrete Cosine Transform based feature extraction. In this work, 73 cervical cell images that consist of 50 normal images and 23 cancer images are used. 35 normal images and 14 cancer images are used to train the LVQ. Then, 23 normal images and 9 cancer images are used in the testing process. Our results show 88,89 % cancer image can be detected correctly (sensitivity), 100 % normal image can be detected correctly (specificity), and 95,83 % for overall detection (accuracy).

Intisari—Kanker serviks telah menjadi penyakit yang banyak diderita kaum wanita di dunia. Secara umum, kanker serviks baru terdeteksi setelah memasuki stadium lanjut, sebab kanker ini sulit teramati pada stadium awal. Pada penelitian ini dirancang perangkat lunak dengan jaringan saraf tiruan *Learning Vector Quantization* (LVQ), sebagai alat bantu deteksi kanker serviks. Sebelum dideteksi, dilakukan pengolahan citra terhadap citra sel serviks, yaitu *preprocessing*, peregangan kontras, median filter, operasi morfologi, segmentasi, dan ekstraksi fitur dengan *Discrete Cosine Transform* (DCT). Citra sel serviks yang digunakan berjumlah 73 buah yang terdiri atas lima puluh buah citra sel normal dan 23 buah citra sel kanker. Proses pelatihan LVQ menggunakan 35 buah citra sel normal dan empat belas buah citra sel kanker. Proses pengujian LVQ menggunakan 15 buah citra sel normal dan sembilan buah citra sel kanker. Dari hasil pengujian, didapatkan nilai sensitivitas, spesifisitas, dan akurasi sebesar 88,89 %, 100 %, dan 95,83 %.

Kata Kunci— Citra *Pap Smear*, *Learning Vector Quantization* (LVQ), Kanker Serviks, *Discrete Cosine Transform* (DCT)

I. PENDAHULUAN

Kanker serviks termasuk dalam salah satu jenis kanker ganas yang menyerang kaum wanita. Menurut data WHO, jumlah penderita kanker serviks di Indonesia sangat tinggi. Setiap tahun, tidak kurang dari 15.000 kasus kanker serviks terjadi di Indonesia. Hal ini membuat WHO menempatkan Indonesia sebagai negara dengan penderita kanker serviks terbanyak di dunia [1].

Sebagian besar kasus kanker serviks disebabkan oleh virus, yaitu *Human Papilloma Virus* (HPV). Se jauh ini, terdapat

lebih dari 150 tipe virus HPV yang telah diketahui. Di antara tipe-tipe virus yang telah diketahui tersebut, HPV tipe 16 dan 18 lah yang paling banyak diketahui sebagai penyebab timbulnya kanker serviks. Penyebaran HPV dapat terjadi karena kontak antar kulit. Pada kebanyakan kasus, penyebaran HPV lebih sering terjadi karena hubungan seksual [1].

Pencegahan penyakit ini dapat dilakukan dengan deteksi dini melalui tes *pap smear*. Keluaran dari tes *pap smear* adalah citra sel serviks. Dari citra sel serviks yang dihasilkan, dapat ditentukan sel serviks tersebut masuk dalam kategori normal atau kanker. Deteksi sel serviks sebagai sel normal atau sel kanker dapat dilakukan melalui dua cara, yakni melalui pengamatan visual, maupun dengan bantuan komputer. Pada umumnya, deteksi kanker serviks dilakukan oleh dokter melalui pengamatan visual.

Deteksi melalui pengamatan visual oleh dokter mempunyai beberapa kelemahan, salah satunya adalah. Untuk mengatasi hal tersebut, adanya sebuah perangkat lunak berbasis komputer guna deteksi kanker serviks berdasarkan sel serviks sangat diharapkan kehadirannya. Adanya perangkat lunak deteksi kanker serviks berbasis komputer ini diharapkan mampu mengekstrak dan mengolah ciri kanker serviks secara lebih efektif dibandingkan dengan pengamatan visual, sehingga hasil deteksi yang didapatkan memiliki tingkat akurasi yang tinggi.

Beberapa penelitian mengenai deteksi kanker serviks otomatis diantaranya yang dilakukan oleh [1] dan [2]. Referensi [2] menggunakan fitur area dan keliling inti sel serviks sebagai metodenya. Penelitian ini menghasilkan akurasi sebesar 82 %. Referensi [1] menggunakan fitur area, keliling, bentuk dan kebulatan inti sel serviks serta LVQ sebagai metodenya. Penelitian ini menghasilkan akurasi yang lebih baik sebesar 93.3 %. Pada penelitian tersebut, penulis menyatakan bahwa diperlukan adanya metode ekstraksi ciri sel serviks lain yang mampu meningkatkan akurasi dari perangkat lunak. Untuk meningkatkan tingkat akurasi seperti saran [1] tersebut, pada penelitian ini dirancang suatu perangkat lunak untuk upaya klasifikasi sel serviks dari citra *pap smear* menggunakan jaringan saraf tiruan *Learning Vector Quantization* (LVQ). Metode ini dikenal lebih efektif dari sisi sensitivitas, spesifisitas, akurasi, dan beban komputasi dibandingkan dengan metode *backpropagation* [3].

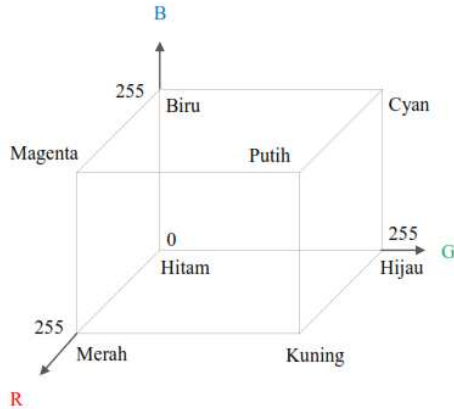
II. METODOLOGI

Proses deteksi kanker serviks terdiri dari delapan buah tahapan, yaitu *preprocessing*, peregangan kontras, median filter, operasi morfologi, segmentasi, ekstraksi fitur dengan *Discrete Cosine Transform* (DCT), pelatihan jaringan LVQ, dan deteksi kanker serviks dengan LVQ.

Tahap pertama dari metode deteksi kanker serviks adalah *preprocessing*. *Preprocessing* terbagi atas dua tahapan, yaitu *cropping* dan *grayscale*. Tahap *cropping* bertujuan untuk menyeragamkan ukuran citra kanker serviks menjadi

¹Mahasiswa Pasca Sarjana, Jurusan Teknik Elektro dan Teknologi Informasi, Fakultas Teknik, Universitas Gadjah Mada, Jalan Grafika 2 Yogyakarta 55281 INDONESIA (telp: 0274-555 225; fax: 0274-4321982; e-mail: dhimas.aries.d@mail.ugm.ac.id)

berukuran 150 x 150 pixel. Tahap kedua *preprocessing* adalah *grayscale*. *Grayscale* bertujuan untuk mengubah citra kanker serviks, dari citra RGB menjadi citra *grayscale*. Citra RGB merupakan jenis citra yang menyajikan warna dalam bentuk komponen R (merah), G (hijau), dan B (biru). Setiap komponen warna menggunakan 8 bit (nilainya berkisar antara 0 sampai dengan 255). Representasi warna dari citra RGB ditunjukkan oleh Gbr 1 [4].



Gbr. 1 Warna RGB dalam ruang berdimensi tiga [4]

Selain citra RGB, terdapat juga citra *grayscale* atau citra berskala keabuan. Citra jenis ini menangani gradasi warna hitam dan putih, yang tentu saja menghasilkan efek warna abu-abu. Pada jenis citra ini, warna dinyatakan dengan intensitas antara 0 sampai dengan 255. Nilai 0 menyatakan hitam dan nilai 255 menyatakan putih. Transformasi citra RGB menjadi citra *grayscale* dilakukan menurut persamaan di bawah ini [4].

$$I = 0,2989R + 0,5870G + 0,1141B \quad (1)$$

Tahap selanjutnya adalah peregangan kontras. Kontras pada suatu citra dapat didefinisikan sebagai distribusi pixel terang dan gelap pada suatu citra. Citra dengan kontras bagus menampilkan rentang nilai pixel yang lebar. Histogram dari citra dengan kontras yang bagus menunjukkan distribusi nilai pixel yang seragam, tidak memiliki puncak utama, dan juga tidak memiliki lembah.

Peregangan kontras adalah salah satu teknik untuk memperbaiki kontras citra. Teknik ini bekerja dengan baik pada citra yang memiliki distribusi *gaussian* atau mendekati *gaussian*. Pada penelitian ini, peregangan kontras dihitung sesuai dengan persamaan berikut [5].

$$o(i, j) = \frac{u(i, j) - c}{d - c} (L - 1) \quad (2)$$

Dengan $o(i, j)$ dan $u(i, j)$ berturut-turut adalah pixel sesudah dan sebelum transformasi pada koordinat (i, j) , c dan d berturut-turut menyatakan nilai maksimum dan minimum pixel citra masukan, sedangkan L menyatakan nilai maksimum pixel citra keluaran [5].

Median filter menjadi tahapan selanjutnya. Median filter adalah salah satu jenis filter non linier. Median filter merupakan salah satu filter yang umum digunakan, khususnya untuk menangani *salt and pepper noise*. Median filter diimplementasikan melalui tiga langkah di bawah ini [6].

1. Tentukan ukuran kernel median filter yang akan digunakan.

2. Untuk setiap kernel, urutkan semua nilai pixel, dan tentukan nilai median dari kumpulan nilai pixel tersebut.
3. Nilai median yang diperoleh, digunakan untuk menggantikan nilai pixel pada pusat kernel.

Untuk upaya lokalisasi inti sel serviks, dibutuhkan operasi morfologi. Dalam penelitian ini, operasi morfologi yang digunakan adalah operasi opening dan closing. Kedua operasi tersebut merupakan operasi tingkat kedua, karena dibentuk oleh operasi morfologi lain, yaitu erosi dan dilasi. Opening didefinisikan sebagai operasi erosi yang diikuti dengan operasi dilasi menggunakan struktur elemen tertentu. Sebaliknya, closing didefinisikan sebagai operasi dilasi yang diikuti dengan operasi erosi menggunakan struktur elemen tertentu. Kedua operasi tersebut dapat dirumuskan sebagai berikut [7].

$$A \circ B = (A \ominus B) \oplus B \quad (3)$$

$$A \cdot B = (A \oplus B) \ominus B \quad (4)$$

Dengan $A \oplus B$ dan $A \ominus B$ masing-masing menunjukkan operasi dilasi dan erosi A atas B .

Tahap segmentasi diinterpretasikan dengan metode pengambangan Otsu. Pada metode Otsu, perhitungan nilai suatu ambang (*threshold*) pada sebuah proses segmentasi didasarkan pada proses pengklusteran. Citra diasumsikan memiliki dua kelompok pixel menurut histogramnya, yaitu *foreground* dan *background*. Nilai ambang optimum yang dihasilkan adalah nilai ambang yang mampu memisahkan kedua kelompok nilai pixel tersebut semaksimal mungkin [8].

Discrete cosine transform (DCT) memiliki aplikasi yang luas dalam bidang pengolahan citra. DCT merupakan dasar pengodean citra JPEG dan video dengan format MPEG. DCT merupakan bentuk sederhana dari *Discrete Fourier Transform* (DFT), dengan menghilangkan unsur imajiner dan mempertahankan sisi riil DFT. Secara umum, DCT dapat dinyatakan melalui persamaan berikut [9].

$$A \circ B = (A \ominus B) \oplus B \quad (5)$$

$$A \cdot B = (A \oplus B) \ominus B \quad (6)$$

Pada penelitian ini, DCT digunakan untuk mengekstrak fitur inti sel serviks.

Pelatihan LVQ menjadi tahapan selanjutnya dari deteksi kanker serviks. LVQ adalah sebuah metode klasifikasi pola yang setiap unit outputnya mewakili kelas tertentu [10].

Tatanama yang digunakan ialah:

X = vector yang dilatih

T = Kelas untuk vector yang dilatih

W_j = Vektor bobot untuk unit output ke- j

C_j = Kelas yang direpresentasikan oleh unit output ke- j .

$\|x - w_j\|$ = Jarak Euclidian antara vector input dan (vektor bobot) unit output ke- j .

Proses pelatihan LVQ dilakukan dengan memasukkan fitur citra pelatihan ke dalam jaringan saraf tiruan LVQ. Proses pelatihan dilakukan untuk mendapatkan nilai bobot optimal yang digunakan untuk mengklasifikasi citra pengujian. Selama pelatihan, unit output diposisikan dengan cara mengatur nilai bobotnya melalui pelatihan terawasi untuk memperkirakan *decision surfaces* dari pengklasifikasi Bayes.

Sebagai langkah awal pelatihan, perlu dilakukan inialisasi sejumlah nilai parameter tertentu. Nilai parameter yang harus

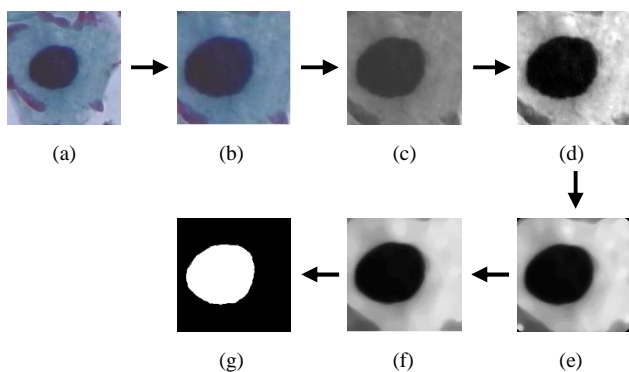
diinisialisasikan ialah bobot awal, maksimal epoch, nilai laju pembelajaran (α) awal, nilai laju pembelajaran (α) minimal, aturan perubahan nilai laju pembelajaran (α), dan syarat penghentian proses pelatihan.

Pada penelitian ini, nilai-nilai parameter yang telah disebutkan di atas diinisialisasi sebagai berikut: Bobot diinisialisasi dengan mengambil dua fitur citra yang terdiri atas satu buah fitur citra sel serviks normal dan satu buah fitur citra sel serviks kanker; Nilai maksimal epoch sebanyak 100 epoch; Nilai laju pembelajaran (α) awal sebesar 0,5; Nilai laju pembelajaran (α) minimal sebesar 10^{-6} ; Nilai laju pembelajaran (α) diperbarui dengan aturan $\alpha(t+1) = 0.5\alpha(t)$; Syarat penghentian ialah apabila jumlah epoch telah melebihi nilai maksimal epoch, atau nilai laju pembelajaran (α) telah kurang dari nilai laju pembelajaran (α) minimal; Proses terakhir adalah deteksi kanker serviks. Jaringan LVQ mend vektor masukan dengan menempatkannya ke kelas yang sama seperti unit output yang mempunyai vektor bobot terdekat dengan vektor masukan.

III. HASIL DAN PEMBAHASAN

Pada paper ini, metode deteksi kanker serviks diterapkan pada 73 buah citra sel serviks. 73 buah tersebut kemudian dibagi menjadi dua kelompok, yakni citra pelatihan dan citra pengujian, masing-masing sebanyak lima puluh dan 23 buah.

Pada 73 buah citra tersebut telah dilakukan rangkaian pengolahan citra sebelum memasuki tahap ekstraksi fitur. Contoh hasil dari pengolahan citra sel serviks, dimulai dari *preprocessing* hingga segmentasi dengan metode otsu dapat dilihat pada Gbr. 2.



Gbr. 2 Citra Hasil dari pengolahan citra sel serviks, dimulai dari *preprocessing* hingga segmentasi dengan metode otsu

Pada Gbr. 2, terlihat bahwa tujuan dari rangkaian pengolahan citra adalah mendapatkan informasi utama dari citra sel serviks, yaitu citra inti sel serviks. *Preprocessing* yang terdiri dari *cropping* dan *grayscale* bertujuan untuk menyeragamkan ukuran serta format citra. Selain itu, *preprocessing* juga berfungsi untuk menurunkan beban komputasi yang karena ukuran citra keluaran yang mengecil dan kanal citra keluaran yang berkurang.

Selanjutnya, peregangan kontras dan median filter digunakan meningkatkan kualitas citra, seperti yang terlihat pada Gbr. 2 (d) dan (e).

Median filter yang diterapkan adalah *multiple* median filter dengan ukuran kernel berturut-turut 3, 5, 7, dan 9. Penerapan

multiple median filter disebabkan oleh banyaknya derau yang terdapat pada citra sel serviks. Dengan dua operasi tersebut, sel serviks dan *background* menjadi mudah untuk dibedakan. Selain itu, kedua operasi tersebut membuat daerah terang dan gelap masing-masing memiliki intensitas pixel yang merata.

Operasi morfologi dilakukan dengan menerapkan operasi *opening* yang diikuti dengan operasi *closing*. Operasi morfologi berfungsi untuk mengisolasi inti sel serviks, seperti yang ditunjukkan oleh Gbr. 2 (f).

Inti sel serviks terisolasi membuat proses segmentasi dengan metode Otsu berjalan lebih maksimal. Adapun penggunaan metode Otsu sebab metode otsu menerapkan pengembangan adaptif dalam proses segmentasi. Hal tersebut membuat setiap citra memiliki nilai ambang masing-masing, yang lebih mewakili karakter citra tersebut, daripada pengembangan global.

Ekstraksi fitur citra sel serviks dilakukan menggunakan *Discrete cosine transform* (DCT). Sebelum DCT diterapkan, citra masukan terlebih dahulu dibagi ke dalam sejumlah jendela berukuran $n \times n$. Tiap pixel pada masing-masing jendela dihitung nilai DCT-nya, untuk kemudian dihitung nilai rata-rata dari kumpulan nilai DCT pada jendela tersebut. Dari proses ini akan dihasilkan r buah nilai rata-rata, untuk kemudian disusun menjadi vektor fitur dari masing-masing citra berukuran $1 \times r$.

Proses pelatihan LVQ dilakukan dengan melatihkan vektor fitur dari 35 buah citra sel normal dan 14 buah citra sel kanker. Karena jaringan LVQ diharapkan memiliki dua buah kelas keluaran, maka bobot awal yang digunakan berukuran $2 \times r$, dimana baris pertama matriks berasal dari salah satu vektor masukan citra sel kanker, dan baris kedua matriks berasal dari salah satu vektor masukan citra sel normal.

Proses pelatihan dilakukan dengan sejumlah ketentuan yang telah disebutkan pada bagian sebelumnya. Proses pelatihan ini menghasilkan matriks bobot berukuran $2 \times r$, dimana baris pertama matriks menunjukkan vektor bobot citra sel kanker, dan baris kedua matriks menunjukkan vektor bobot citra sel normal. Akhirnya, proses deteksi kanker serviks dilakukan dengan memasukkan vektor fitur dari 15 buah citra sel normal dan 9 buah citra sel kanker.

Jaringan LVQ mendeteksi keluaran dari masing-masing vektor masukan dengan menempatkannya pada kelas yang memiliki jarak terdekat dengan vektor bobot. Hasil deteksi kanker serviks oleh jaringan LVQ dapat dilihat pada Tabel 1.

Dari Tabel 1 terlihat bahwa kesalahan deteksi hanya terjadi pada citra Test 8, dimana citra kanker dideteksi sebagai normal. Dengan begitu, berdasarkan hasil deteksi tersebut, didapatkan nilai *True positive* (TP), yaitu banyaknya citra sel kanker yang terdeteksi sebagai kanker sebanyak 8. Selanjutnya, tidak ada nilai *False positive* (FP) yaitu banyaknya citra sel normal yang terdeteksi sebagai kanker. Sedangkan nilai *True Negative* (TN) dan *False Negative* (FN), yaitu banyaknya citra sel norma yang terdeteksi sebagai normal, dan banyaknya citra sel normal yang terdeteksi sebagai kanker masing masing sebesar 15 dan 1. Berdasarkan nilai TP, FP, TN, dan FN di atas, dapat dihitung nilai sensitivitas, spesifisitas, dan akurasi sebagai berikut.

$$\text{Sensitivitas} = \frac{TP}{TP + FN} \times 100\% \quad (7)$$

$$\text{Spesifisitas} = \frac{TN}{TN + FP} \times 100\% \quad (8)$$

$$\text{Akurasi} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (9)$$

Sesuai dengan persamaan 7, 8, dan 9, didapatkan nilai sensitivitas, spesifisitas, dan akurasi masing-masing sebesar 88,89 %, 100 %, dan 95,83 %.

TABELI
HASIL DETEKSI CITRA UJI

Nama Citra	Jarak ke Kanker	Jarak ke Normal	Hasil Deteksi	Deteksi Sebenarnya	Keterangan
Test 1	2,23	4,31	Kanker	Kanker	Cocok
Test 2	2,24	5,78	Kanker	Kanker	Cocok
Test 4	4,66	7,44	Kanker	Kanker	Cocok
Test 5	1,96	5,76	Kanker	Kanker	Cocok
Test 6	1,91	6,87	Kanker	Kanker	Cocok
Test 7	5,29	5,70	Kanker	Kanker	Cocok
Test 8	4,95	3,46	Normal	Kanker	Tidak Cocok
Test 9	3,88	9,58	Kanker	Kanker	Cocok
Test 10	3,78	4,00	Kanker	Kanker	Cocok
Test 11	5,41	0,88	Normal	Normal	Cocok
Test 12	5,67	1,65	Normal	Normal	Cocok
Test 13	6,79	0,97	Normal	Normal	Cocok
Test 14	5,24	1,94	Normal	Normal	Cocok
Test 15	6,06	1,41	Normal	Normal	Cocok
Test 16	5,28	1,63	Normal	Normal	Cocok
Test 17	4,41	1,79	Normal	Normal	Cocok
Test 18	7,00	2,66	Normal	Normal	Cocok
Test 19	5,53	0,88	Normal	Normal	Cocok
Test 20	6,09	1,23	Normal	Normal	Cocok
Test 21	4,71	1,85	Normal	Normal	Cocok
Test 22	6,32	1,67	Normal	Normal	Cocok
Test 23	5,65	1,78	Normal	Normal	Cocok
Test 24	4,37	1,44	Normal	Normal	Cocok

KESIMPULAN

Perangkat lunak guna Deteksi Kanker Serviks Otomatis Berbasis Jaringan Saraf Tiruan *Learning Vector Quantization* (LVQ) dan *Discrete Cosine Transform* (DCT) dari Citra *Pap Smear* berhasil dirancang melalui delapan buah tahapan metode, yaitu *preprocessing*, peregangan kontras, median filter, operasi morfologi, segmentasi, ekstraksi fitur dengan *Discrete Cosine Transform* (DCT), pelatihan jaringan LVQ, dan deteksi kanker serviks menggunakan jaringan LVQ. Dari proses pengujian yang dilakukan, didapatkan nilai sensitivitas, spesifisitas, dan akurasi yang tinggi yaitu masing-masing sebesar 88,89 %, 100 %, dan 95,83 %.

REFERENSI

- [1] E. M. Dewi, "Ekstraksi Fitur dan Klasifikasi Sel Serviks Dengan Metode Learning Vector Quantization (LVQ) Untuk Klasifikasi Dini Kanker Serviks," Universitas Airlangga, 2013.
- [2] J. Norup, "Classification of Pap Smear Data by Transductive NeuroFuzzy Methods," University of Denmark, 2005.
- [3] L. Listyalina, "Implementasi Learning Vector Quantization (LVQ) Untuk Klasifikasi Kanker Paru dari Citra Foto Rontgen," Universitas Airlangga, 2013.
- [4] A. Kadir and A. Susanto, *Pengolahan Citra*. Yogyakarta: Penerbit ANDI, 2012.
- [5] D. Putra, *Pengolahan Citra Digital*. Yogyakarta: Penerbit ANDI, 2009.
- [6] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2nd ed. New Jersey: Prentice Hall, 2002.
- [7] R. C. Gonzales, R. E. Woods, and S. L. Eddin, *Digital Image Processing Using MATLAB, 2nd ed*, 2nd ed. India: McGraw-Hill Education, 2009.
- [8] N. Otsu, "No Title," *IEEE Tras. Syst. Man Cybern*, pp. 62–66, 1979.
- [9] W. K. Pratt, *Digital Image Processing*, 4th ed. California: WILER-INTERSCIENCE, 2007.
- [10] L. V Fausett, *Fundamental of Neural Network: Architectures, Algorithms, and Applications*. Prentice-Hall, 1994.