

Klasifikasi Varietas Tanaman Kelengkeng Berdasarkan Morfologi Daun Menggunakan *Backpropagation Neural Network* dan *Probabilistic Neural Network*

H. Syahputra dan A. Harjoko

Abstract— Pengenalan daun memainkan peran penting dalam klasifikasi tanaman dan isu utamanya terletak pada apakah fitur yang dipilih stabil dan memiliki kemampuan yang baik untuk membedakan berbagai jenis daun. Pengenalan tanaman berbantuan komputer merupakan tugas yang masih sangat menantang dalam visi komputer karena kurangnya model atau skema representasi yang tepat. Fokus komputerasi pengenalan tanaman hidup adalah untuk mengukur bentuk geometris berbasis morfologi daun. Informasi ini memainkan peran penting dalam mengidentifikasi berbagai kelas tanaman. Pada penelitian ini dilakukan pengenalan jenis tanaman berdasarkan fitur yang menonjol dari daun seperti fisiologis panjang (*physiological length*), lebar (*physiological width*), diameter, keliling (*leaf perimeter*), luas (*leaf area*), faktor mulus (*narrow factor*), rasio aspek (*aspect ratio*), faktor bentuk (*form factor*), *rectangularity*, rasio perimeter terhadap diameter, rasio perimeter panjang fisiologi dan lebar fisiologi yang dapat digunakan untuk membedakan satu sama lain. Berdasarkan hasil pengujian, ditunjukkan bahwa hasil pencocokkan daun kelengkeng dengan menggunakan neural network lebih baik dibandingkan dengan hasil pencocokkan daun kelengkeng dengan menggunakan probabilistic neural network. Akan tetapi ekstraksi fitur dengan menggunakan morfologi belum dapat memberikan informasi pembeda yang signifikan bagi pengenalan tanaman varietas kelengkeng berdasarkan daunnya.

Keywords— klasifikasi, morfologi daun, neural network, probabilistic neural network

1. PENDAHULUAN

1.1. Latar Belakang

Banyaknya aneka ragam tanaman yang dapat dijumpai dimana saja membuat sulitnya penentuan jenis tanaman. Salah satu masalah nyata yang ditemukan di bidang biologi atau

pertanian saat ini adalah sulitnya mengidentifikasi jenis tanaman dari beberapa varietas, misalnya varietas tanaman kelengkeng.

Di dalam ilmu botani, identifikasi varietas tanaman diperlukan untuk keperluan persilangan tanaman. Selain itu, identifikasi varietas tanaman dilakukan untuk menjamin kemurnian suatu kelompok varietas terhadap varietas lainnya pada suatu lahan pertanian.

Untuk dapat mengidentifikasi varietas tanaman, dibutuhkan cukup pengetahuan, seperti pengetahuan mengenai informasi atau ciri-ciri unik dari tiap-tiap varietas tersebut. Faktor ciri tanaman dapat diperoleh melalui struktur organ tanamannya, seperti batang, daun, bunga, biji atau struktur lainnya. Fitur daun dalam beberapa varietas memiliki warna, tekstur, dan tulang daun yang cenderung sama. Bentuk daun merupakan salah satu faktor penting yang menyusun bentuk tanaman. Bentuk daun adalah kunci utama untuk identifikasi tanaman (Neto *et al.* 2006).

Untuk mengatasi pengklasifikasian tanaman, beberapa penelitian telah dilakukan dengan menggunakan penggambaran ilmu botani (Groom, Lamont, Markey 1997). Akan tetapi, hal ini tidak mudah dalam mengekstrak dan mentransfer fitur ke dalam komputer secara otomatis. Untuk mengurangi kesalahan dalam mengidentifikasi perlu dicegah intervensi manusia dalam pengekstrasian fitur.

Dalam mengekstrak atau mengukur fitur daun merupakan suatu studi yang cukup rumit. Hal ini membuat aplikasi dari pengenalan pola pada bidang ini adalah suatu tantangan baru. Dalam beberapa tulisan menyatakan bahwa perolehan data dari tanaman secara otomatis dengan komputer belum dapat diimplementasikan. Beberapa pendekatan telah dilakukan dengan menggunakan fitur pradefinisi. Yonekawa (1996) mengidentifikasi tipe daun ideal dengan menggunakan faktor bentuk tak berdimensi sederhana dengan analisis citra. Tsukaya (2006) mengembangkan bagaimana mekanisme penentuan bentuk daun. Plotze *et al.* (2005)

Hermawan Syahputra mahasiswa s3 UGM, yang juga bekerja di Universitas Negeri Medan, Jl Pringgondani no 825 Karang Rejo Stabat Sumatera Utara, 20851 (telp, e-mail: ma1_unimed@yahoo.co.id).

Agus Harjoko staf pengajar Jurusan Ilmu Komputer dan Elektronika, FMIPA UGM, Sekip Utara, Yogyakarta, 55281 (telp, e-mail: aharjoko@ugm.ac.id, aharjoko@yahoo.com).

menganalisis bentuk daun menggunakan skala multi Minkowski dimensi fraktal. Scottferson *et al.* (1985) mengembangkan pengukuran variasi bentuk dari outline dua dimensi.

Pada penelitian ini, morfologi tanaman akan digunakan untuk ekstraksi citra daun untuk mengklasifikasikan varietas tanaman kelengkeng serta menggunakan *neural network* propagasi balik dan *Probabilistic Neural Network* untuk pencocokan basis data citra terhadap input citra yang akan diklasifikasikan.

1.2. Tujuan

Adapun tujuan dari penelitian ini adalah :

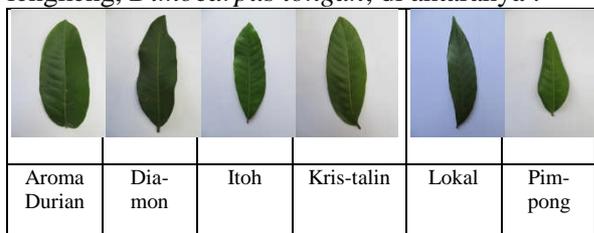
1. Mengekstraksi fitur daun dari varietas tanaman kelengkeng dengan menggunakan morfologi daun yang dapat dijadikan sebagai ciri dalam mengenali varietas tanaman kelengkeng.
2. Mengklasifikasi varitas tanaman kelengkeng dengan menggunakan *neural network* dan *probabilistic neural network*.

1.3. LANDASAN TEORI

1.3.1. Tanaman Kelengkeng

Lengkeng (jugadisebut kelengkeng, matakucin g, atau longan, *Dimocarpus longan*, suku lerak-lerakan atau *Sapindaceae*) adalah tanaman buah-buahan yang berasal dari daratan Asia Tenggara.

Pohon lengkung dapat mencapai tinggi 40 m dan diameter batangnya hingga sekitar 1 m. Berdaun majemuk, dengan 2-4(-6) pasang anak daun, sebagian besar berbulu rapat pada bagian aksialnya. Tangkai daun 1-20 cm, tangkai anak daun 0,5-3,5 cm. Anak daun bulat memanjang, panjang lk. 1-5 kali lebarnya, bervariasi 3-45 × 1,5-20 cm, mengertas sampai menjangat, dengan bulu-bulu kempa terutama di sebelah bawah di dekat pertulangan daun. Ada beberapa jenis lengkung, *Dimocarpus longan*, di antaranya :

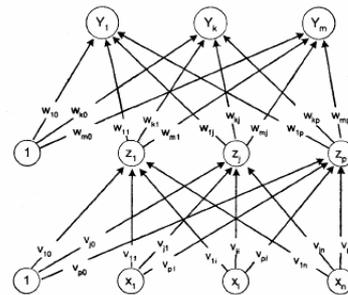


Sumber : Kebun kharisma, jl.bugisan selatan, jogja

1..3.2. Algoritma Pelatihan JST Propagasi balik

Back-Propagation Neural Network (BPNN) atau juga disebut Jaringan Syaraf tiruan (JST) propagasi balik adalah JST dengan topologi

multi-lapis (*multilayer*) dengan satu lapis masukan (lapis X), satu atau lebih lapis *hidden* atau tersembunyi (lapis Z) dan satu lapis keluaran (lapis Y). Setiap lapis memiliki neuron-neuron (unit-unit) yang dimodelkan dengan lingkaran (lihat Gambar 1). Di antara neuron pada satu lapis dengan neuron pada lapis berikutnya dihubungkan dengan model koneksi yang memiliki bobot-bobot (*weights*), *w* dan *v*. Lapis tersembunyi dapat memiliki *bias*, yang memiliki bobot sama dengan satu.



Gambar 1. Arsitektur Backpropagation

Algoritma pelatihan JST Propagasi balik pada dasarnya terbagi menjadi 2 langkah, yaitu: langkah maju (*feedforward*) dan propagasi balik (*back propagation*). Pada langkah maju, perhitungan bobot-bobot neuron hanya didasarkan pada vektor masukan, sedangkan pada propagasi balik, bobot-bobot diperhalus dengan memperhitungkan nilai target atau keluaran. Algoritma pelatihan maju dan propagasi balik ini ditunjukkan pada Algoritma 1 [1]. Nilai *mean square error* (MSE) pada satu siklus pelatihan (langkah 2 – 10, dimana seluruh record dipresentasikan satu kali) adalah nilai kesalahan (*error* = nilai keluaran-nilai masukan) rata-rata dari seluruh record (*tupple*) yang dipresentasikan ke JST dan dirumuskan sebagai:

$$MSE = (\text{rekordjumlaherror} \sum 2) .$$

Semakin kecil MSE, JST semakin kecil kesalahannya dalam memprediksi kelas dari record yang baru. Maka, pelatihan JST ditujukan untuk memperkecil MSE dari satu siklus ke siklus berikutnya sampai selisih nilai MSE pada siklus ini dengan siklus sebelumnya lebih kecil atau sama dengan batas minimal yang diberikan (epsilon).

Algoritma BPNN

Step 0 : Inisial bobot (weight) dengan nilai random kecil

Step 1 : Selama kondisi berhenti belum terpenuhi, lakukan langkah 2 sampai 9.

Step 2 : untuk setiap pasangan training, lakukan langkah 3 sampai 8

(Feedforward)

Step 3 : Untuk tiap unit *input* ($X_i, i=1, \dots, n$) menerima sinyal *input* x_i dan menyebarkan sinyal itu keseluruh unit pada lapis atasnya (lapis tersembunyi)

Step 4 : Untuk tiap unit tersembunyi ($Z_j, j=1, \dots, p$) dihitung nilai input dengan menggunakan nilai bobotnya :

$$z_in_j = v_{0j} + \sum_{i=1}^n x_i w_{ij}$$

Kemudian dihitung nilai *output* dengan menggunakan fungsi akti-vasi yang dipilih : $z_j = f(z_in_j)$. Hasil fungsi tersebut dikirim ke semua unit pada lapis di atasnya

Step 5 : Untuk tiap unit *output* ($Y_k, k=1, \dots, m$) dihitung nilai *input* dengan menggunakan nilai bobot-nya :

$$y_in_k = w_{0k} + \sum_{j=1}^p z_j w_{jk}$$

• Kemudian dihitung nilai *output* dengan menggunakan fungsi aktivasi :

$$y_k = f(y_in_k)$$

(Backpropagation error)

Step 6 : Untuk tiap unit *output* ($Y_k, k=1, \dots, m$) menerima pola target yang bersesuaian dengan pola *input*, dan kemudian dihitung informasi kesalahan :

$$\delta_k = (t_k - y_k) f'(y_in_k)$$

• Kemudian dihitung koreksi nilai bobot yang kemudian akan digunakan untuk memperbaharui nilai bobot w_{jk} :

$$\Delta w_{jk} = \alpha \delta_k z_j$$

• Hitung koreksi nilai bias yang kemudian akan digunakan untuk memperbaharui nilai w_{0k} :

$$\Delta w_{0k} = \alpha \delta_k$$

• dan kemudian nilai dikirim ke unit pada lapis sebelumnya.

Step 7 : Untuk tiap unit tersembunyi ($Z_j, j=1, \dots, p$) dihitung delta *input* yang berasal dari unit pada lapis di atasnya :

$$\delta_in_j = \sum_{k=1}^m \delta_k w_{jk}$$

Kemudian nilai tersebut dikalikan dengan nilai turunan dari fungsi aktivasi untuk menghitung informasi kesalahan :

$$\delta_j = \delta_in_j f'(z_in_j)$$

• Hitung koreksi nilai bobot yang kemudian digunakan untuk memperbaharui nilai

$$\Delta v_{ij} = \alpha \delta_j x_i$$

• dan hitung nilai koreksi bias yang kemudian digunakan untuk memperbaharui :

$$\Delta v_{0j} = \alpha \delta_j$$

(Memperbaharui nilai bobot dan bias)

Step 8 : Tiap nilai bias dan bobot ($j=0, \dots, p$) pada unit *output* ($Y_k, k=1, \dots, m$) diperbaharui :

$$w_{jk}(\text{new}) = w_{jk}(\text{old}) + \Delta w_{jk}$$

$$v_{ij}(\text{new}) = v_{ij}(\text{old}) + \Delta v_{ij}$$

Step 9 : Menguji apakah kondisi berhenti sudah terpenuhi.

Kondisi berhenti ini terpenuhi jika nilai kesalahan yang dihasilkan lebih kecil dari nilai kesalahan referensi atau training telah mencapai epoch yang ditetapkan.

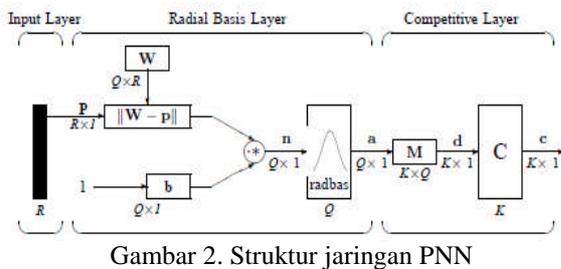
Keterangan notasi untuk Algoritma 1: x = vektor masukan = $(x_1, \dots, x_i, \dots, x_n)$, t = vektor keluaran $(t_1, \dots, t_i, \dots, t_n)$, δ_k = nilai koreksi bobot *error* untuk w_{jk} yang disebabkan oleh *error* pada unit keluaran Y_k . δ_j = nilai koreksi bobot *error* untuk v_{ij} yang disebabkan oleh informasi propagasi balik dari *error* pada lapis keluaran ke unit tersembunyi. Z_j α = konstanta laju pembelajaran (*learning rate*). X_i = unit masukan i. v_{0j} = bias pada unit tersembunyi j. Z_j = unit tersembunyi j. w_{0k} = bias pada unit keluaran k. Y_k = unit keluaran k.

1.3.3. Pengantar Jaringan Syaraf Probabilistik (PNN)

PNN berasal dari Jaringan Radial Basis Function (RBF) yang merupakan JST menggunakan RBF. RBF adalah fungsi *bell shape* yang merupakan variabel skala nonlinear. PNN diadopsi untuk memiliki banyak keuntungan. Kecepatan pelatihannya berlipat kali lebih cepat dari jaringan BPNN. Keuntungan yang paling penting dari PNN adalah pelatihan yang bersifat mudah dan seketika. Bobot tidak "terlatih" tapi ditugaskan. Bobot yang ada tidak

akan pernah berganti-ganti tetapi hanya vektor baru dimasukkan ke dalam matriks bobot saat pelatihan. Sehingga dapat digunakan secara real-time. Karena pelatihan dan berjalan prosedur dapat diimplementasi kan dengan manipulasi matriks, yang kecepatan PNN sangat cepat.

Pengklasifikasi jaringan memasukan vektor ke dalam kelas tertentu berdasarkan kelas yang memiliki probabilitas maksimum untuk menjadi benar. Dalam makalah ini, PNN memiliki tiga lapisan: lapisan input, Dasar Radial Layer dan Layer Kompetitif. radial Basis Layer vektor jarak mengevaluasi antara vektor input dan baris bobot vektor dalam matriks bobot. Ini adalah skala jarak Fungsi Dasar Radial nonlinear. Kemudian Kompetitif Layer menemukan jarak terpendek di antara mereka, dan dengan demikian menemukan pola pelatihan yang paling dekat dengan pola input berdasarkan jarak.

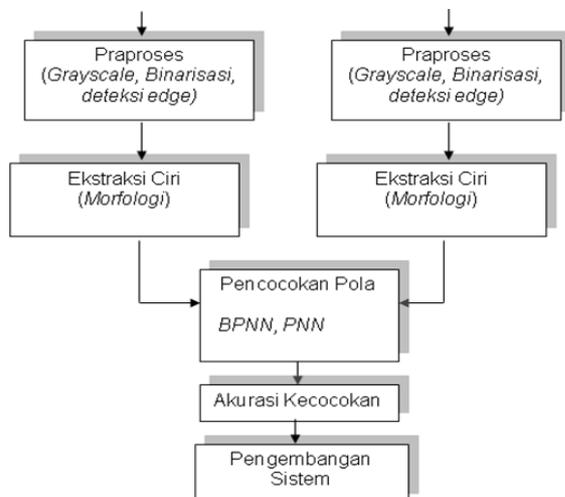


Gambar 2. Struktur jaringan PNN

2. METODE PENELITIAN

2.1. Desain Sistem

Pada bagian ini akan dibahas mengenai sistem perangkat lunak yang dirancang. Hasil keluaran dari sistem ini akan mengklasifikasikan kondisi paru-paru dalam tiga kondisi, yakni kanker, normal, atau efusi. Secara umum sistem penelitian yang dirancang adalah seperti berikut :



Gambar 3. Desain sistem klasifikasi tanaman

2.2 Praproses

2.2.1. Akuisisi Citra dan Praproses Daun

Sumber yang digunakan untuk citra daun adalah citra daun kelengkeng dengan citra warna RGB, sehingga dibutuhkan untuk mengkonversi warna dari RGB menjadi keabuan, dengan gangguan warna dapat diabaikan. Citra-citra ini adalah berbentuk JPEG. Semua citra daun adalah beresolusi 2000 x 3000.

2.2.2. Konversi Citra RGB menjadi Citra Grayscale

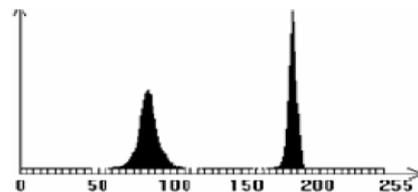
Formula yang digunakan untuk mengkonversi nilai RGB dalam piksel ke dalam nilai grayscale adalah sebagai berikut :

$$\text{Gray} = 0.2989 * R + 0.5870 * G + 0.1140 * B$$

Pada persamaan ini, R, G, B berkaitan dengan warna piksel berturut-turut.

2.2.3. Mengkonversi Citra Grayscale menjadi Biner

Citra biner yang diperoleh dengan mengatur histogram luminasi dari citra grayscale telah dihasilkan. Histogram luminasi dari citra grayscale ditunjukkan seperti gambar berikut :



Gambar 4. Histogram Luminasi citra grayscale

Terdapat dua puncak pada setiap histogram luminasi dari citra grayscale. Menurut luminasi citra grayscale, puncak sebelah kiri merujuk kepada piksel yang memuat citra daun sedangkan puncak sebelah kanan merujuk pada piksel yang mengandung *background*.

Teknik praproses citra dapat diilustrasikan sebagai berikut :



Gambar 5. Prapemrosesan citra daun

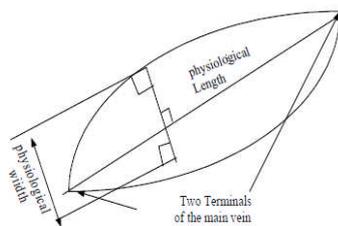
2.3. Ekstraksi Fitur

Beberapa fitur morfologi digital diekstraksi dari citra daun, dimana fitur-fitur tersebut diturunkan dari 5 fitur dasar.

2.3.1. Basic Geometric Features

Pertama, dihasilkan 5 fitur geometri dasar.

- **Diameter (D)**: didefinisikan sebagai jarak terjauh dari dua titik dari margin daun.
- **Physiological Length (L_p)**: Jarak antara dua terminal venasi utama dari daun yang didefinisikan sebagai panjang fisiologi.
- **Physiological Width (W_p)**: Panjang maksimum dari garis, yang tegak lurus dengan venasi utama, yang didefinisikan sebagai lebar fisiologi. Ketika koordinat piksel adalah diskrit, maka dua garis merupakan orthogonal jika sudutnya 90°+0.5°. Hubungan antara panjang fisiologi dan lebar fisiologi ditunjukkan pada gambar.
- **Leaf Area (A)**: Nilai area daun mudah dievaluasi, hanya dengan menghitung banyaknya piksel bernilai biner 1 pada citra daun yang diperhalus.
- **Leaf Perimeter (P)**: perimeter daun dihitung dengan menghitung banyaknya piksel yang memuat margin daun.



Gambar 6. Hubungan antara panjang fisiologi dan lebar fisiologi.

2.3.2. Fitur Morfologi Digital

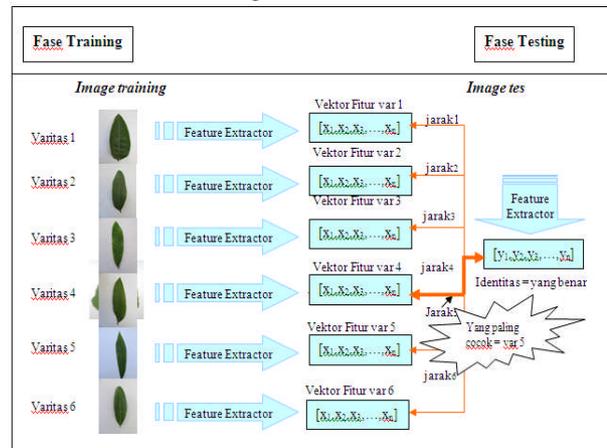
Berdasarkan pada 5 fitur dasar yang diperkenalkan sebelumnya, dapat didefinisikan beberapa fitur morfologi digital yang diperkenalkan untuk deteksi daun.

- **Aspect ratio**: Rasio Aspek didefinisikan sebagai rasio panjang fisiologi L_p dengan lebar fisiologi W_p, maka L_p/W_p
- **Form factor**: Fitur ini digunakan untuk menggambarkan perbedaan antara daun dan lingkaran, yang dinotasikan dengan $4\pi A/P^2$, dimana A adalah area daun dan P adalah perimeter margin daun.
- **Rectangularity**: menggambarkan kemiripan antara daun dan persegi, yang dinotasikan dengan L_pW_p/A, diman L_p adalah panjang fisiologi, W_p adalah lebar fisiologi dan A adalah area fisiologi.
- **Narrow factor**: Narrow factor didefinisikan sebagai rasio antara diameter D dan panjang fisiologi L_p, maka D/L_p.

- **Rasio Perimeter terhadap diameter**: Rasio perimeter terhadap diameter, merepresentasikan rasio perimeter daun P dan diameter daun D, dihitung dengan P/D.
- **Rasio perimeter panjang fisiologi dan lebar fisiologi**: fitur ini didefinisikan sebagai rasio dari perimeter daun P dan jumlah panjang fisiologi L_p dan lebar fisiologi W_p, maka P/(L_p + W_p).

2.4. Rancangan Sistem Klasifikasi Tanaman Berdasarkan Daun

Pada penelitian ini citra daun akan diklasifikasikan ke tiga varitas, yang dibedakan berdasarkan morfologi daun.



Gambar 7. Pencocokan Pola

3. HASIL DAN PEMBAHASAN

Implementasi program dilakukan dengan menggunakan perangkat lunak Matlab 7.04. Berdasarkan data yang diperoleh sebanyak 600 data untuk 6 varitas kelengkeng, yang masing-masing varitas memiliki 100 sampel data. Dari 100 sampel data dari setiap varitas, 80 sampel data digunakan sebagai data training dan 20 sampel data digunakan sebagai data testing. Dari hasil proses training, selanjutnya dilakukan pengujian untuk melihat kemampuan pengenalannya. Untuk pengujian digunakan 1/3 data untuk masing-masing kelompok.

3.1. Hasil Pencocokan dengan Neural Network

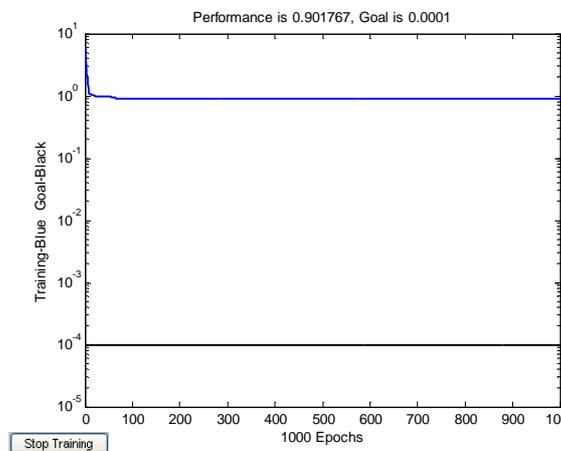
Hasil pengujian dapat dilihat sebagai berikut :

hidden	epoch	lr	MSE	akurasi
10	1000	0.1	0.00202098	91

Pada pengujian dengan bentuk 10-10-1 (menggunakan 10 hidden layer), jumlah epoch 1000, lr 0.1 dengan MSE 0.90 diperoleh tingkat

akurasi sebesar 46.154, ini menunjukkan tingkat akurasi yang cukup baik.

Pada tingkat akurasi yang paling tinggi, diperoleh grafik sebagai berikut :



Gambar 3. grafik training 10-10-1 (10 hidden layer), eph 1000, lr 0.1

3.2 Hasil Pencocokan dengan Probabilistic Neural Network

Hasil pengujian dengan menggunakan probabilistic Neural Network hanya mendapatkan kecocokan 33.3%. Rendahnya hasil pengenalan ini diasumsikan karena kurang diperolehnya fitur yang dapat membedakan daun kelengkeng.

4. KESIMPULAN

Berdasarkan hasil pengujian, ditunjukkan bahwa hasil pencocokkan daun kelengkeng dengan menggunakan *neural network* lebih baik dibandingkan dengan hasil pencocokkan daun kelengkeng dengan menggunakan *probabilistic neural network*. Akan tetapi ekstraksi fitur dengan menggunakan morfologi belum dapat memberikan informasi pembeda yang signifikan bagi pengenalan tanaman varietas kelengkeng berdasarkan daunnya. Sehingga diperlukan ekstraksi fitur lainnya seperti tekstur, bentuk, atau bahkan peningkatan ekstraksi terhadap venasi daun untuk mendapatkan fitur yang lebih detail.

5. SARAN

Perlu dilakukan penelitian lebih lanjut tentang metode atau model yang lebih akurat untuk ekstraksi fitur daun.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Bapak Agus Harjoko, Ph.D yang telah membimbing penulis terhadap penelitian ini.

DAFTAR PUSTAKA

- [1] Wu SG, Bao FS, Xu EY, Wang YX, Chang YF, Xiang QL. 2007. A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network. *arXiv [CS.AI]* 1 : 4289
- [2] Gonzalez RC, Woods RE, Eddins SL. 2004. *Digital Image Processing*. Third Edition. New Jersey : Prentice Hall.
- [3] Duda RO, Hart PE, Stork DG. 2001. *Pattern Classification 2nd*. USA : John Wiley & Sons, Inc.
- [4] J.-X. Du, X.-F. Wang, and G.-J. Zhang, 2007. "Leaf shape based plant species recognition," *Applied Mathematics and Computation*, vol. 185,.
- [5] Y. Ye, C. Chen, C.-T. Li, H. Fu, and Z. Chi, 2004. "A computerized plant species recognition system," in *Proceedings of 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing*, Hong Kong.