

IDENTIFIKASI PEMBICARA MENGGUNAKAN MODEL *HYBRID PROBABILISTIC NEURAL NETWORK* DAN *FUZZY C-MEANS*

SPEAKER IDENTIFICATION USING HYBRID MODEL OF PROBABILISTIC NEURAL NETWORK AND FUZZY C-MEANS

Vicky Zilvan*, Agus Buono**, dan Sri Nurdiati**

*UPT Loka Pengembangan Signal & Navigasi, Lembaga Ilmu Pengetahuan Indonesia, Jln. Rangamalela No. 11, Bandung, Jawa Barat, Indonesia
Telp./Faks. +62 22 4205425

**Departemen Ilmu Komputer, FMIPA, Institut Pertanian Bogor, Kampus Darmaga, Jln. Meranti Wing 20 Level 5-6, Bogor, Jawa Barat, Indonesia
Pos-el: vick001@lipi.go.id

ABSTRACT

A hybrid model of Probabilistic Neural Network and Fuzzy C-Means has been developed. The model has been applied using Mel Frequency Cepstrum Coefficients (MFCC) as feature extraction for identification. In addition to the natural voice, the effect of noise has also been taken into account. It has been shown that the model has good accuracy at 96% for voice without noise, 85.5% for voice with noise at the level of signal to noise ratio 30, and 60% for voice with noise at the level of signal to noise ratio 20. It has also been concluded that the clustering procedure using Fuzzy C-Means could improve the accuracy up to 96% for large number of training data.

Keywords: *Speaker Identification, Probabilistic Neural Network, Fuzzy C-Means, Signal to Noise Ratio, Mel Frequency Cepstrum Coefficients (MFCC)*

ABSTRAK

Telah dikembangkan model *Hybrid Probabilistic Neural Network* dan *Fuzzy C-Means*. Model ini menggunakan *Mel-Frequency Cepstrum Coefficients* (MFCC) sebagai pengkstraksi ciri suara untuk identifikasi. Selain percobaan pada data suara asli, penelitian ini juga melakukan percobaan menggunakan suara ber-*noise* untuk melihat kehandalan model yang dikembangkan dalam mengidentifikasi suara ber-*noise*. Dari hasil pengujian didapatkan bahwa model yang telah dikembangkan ini memiliki akurasi cukup tinggi dengan akurasi tertinggi sebesar 96% untuk data suara tanpa *noise*. Sementara untuk suara ber-*noise* dengan tingkat *signal to noise ratio* sebesar 30, akurasi tertinggi mencapai 85,5% dan mencapai 60% untuk suara ber-*noise* dengan tingkat *signal to noise ratio* sebesar 20. Selain itu, dari penelitian ini diketahui bahwa prosedur pengelompokan dengan *Fuzzy C-Means* dapat meningkatkan akurasi hingga 96% untuk ukuran data pelatihan yang besar.

Kata kunci: *Identifikasi Pembicara, Probabilistic Neural Network, Fuzzy C-Means, Signal to Noise Ratio, Mel Frequency Cepstrum Coefficients (MFCC)*

PENDAHULUAN

Pengenalan pola masih menjadi kajian yang menarik bagi para peneliti. Hal ini dilakukan tidak hanya untuk keperluan penelitian ataupun

peningkatan produktivitas kerja, tetapi lebih mengarah pada kebutuhan di era *high technology*, sehingga kemajuan di bidang elektronika dan teknologi informasi diharapkan secara sinergi memenuhi tuntutan kebutuhan manusia. Oleh

karena itu, permasalahan pengenalan pola telah berkembang pesat dan telah banyak digunakan dalam berbagai bidang.

Di perkotaan ataupun industri, baik suara, sidik jari, pola geometri telapak tangan, maupun wajah bias digunakan sebagai mesin kehadiran. Di dunia perbankan, suara digunakan untuk berbagai kepentingan, seperti pelayanan bank melalui telepon. Di dunia kedokteran, iris mata digunakan untuk mengidentifikasi adanya faal pada organ tubuh dan pupil mata untuk mengidentifikasi tingkat kelelahan pada seseorang.

Penelitian yang dilakukan dengan menggunakan data sinyal suara umumnya disebut dengan pemrosesan sinyal suara (*speech processing*). *Speech processing* sendiri memiliki beberapa cabang kajian. Salah satu kajian dalam *speech processing* adalah identifikasi pembicara. Identifikasi pembicara (*speaker identification*) adalah suatu proses mengenali seseorang berdasarkan suaranya.¹

Banyak sekali metode yang dikembangkan para peneliti untuk melakukan identifikasi, antara lain *Hidden Markov Model* (HMM)², *Genetic Algorithm*³, ataupun jaringan saraf tiruan (*Neural Networks*)⁴. Di samping itu, gabungan dari beberapa metode tersebut pun sering digunakan.

Salah satu metode yang digunakan untuk identifikasi pembicara adalah *Probabilistic Neural Network* (PNN). Metode ini banyak digunakan untuk mengenal pola karena model yang fleksibel dibandingkan dengan metode *neural network* lainnya.⁵ Suhartono pada tahun 2007 telah mengembangkan metode *Probabilistic Neural Network* dengan *Mel-Frequency Cepstrum Coefficients* (MFCC) sebagai pengekstraksi ciri suara untuk identifikasi pembicara. Dari penelitian yang telah dilakukan didapat tingkat akurasi yang cukup tinggi, yaitu sebesar 94%.⁶ Salah satu hal yang dapat memengaruhi tingkat akurasi ini adalah gaya berbicara (seperti cara pengucapan, intonasi, dan logat) yang berbeda dari seorang pembicara pada saat pengumpulan data untuk membangun pola.⁷

Untuk meminimalkan pengaruh gaya berbicara yang berbeda dari pembicara, salah satu yang cara yang dapat ditempuh adalah dengan melakukan pengelompokan data dari setiap pembicara berdasarkan gaya berbicara tiap-tiap

pembicara. Salah satu metode pengelompokan terbaik dan sering digunakan dalam banyak aplikasi seperti pengenalan suara dan pengenalan pola adalah *Fuzzy C-Means*.⁸

Dengan melihat kemungkinan pengembangan yang dapat dilakukan untuk meningkatkan akurasi model untuk identifikasi pembicara, pada penelitian ini akan dikembangkan model *hybrid Probabilistic Neural Network* dan *Fuzzy C-Means* (FCM) dengan *Mel-Frequency Cepstrum Coefficients* (MFCC) sebagai pengekstraksi ciri suara untuk identifikasi pembicara. Selain itu, pada penelitian ini juga akan dicobakan pengaruh data *ber-noise* terhadap akurasi model.

TINJAUAN PUSTAKA

Fuzzy C-Means (FCM)

Menurut Jang *et al.*, *Fuzzy C-Means* merupakan algoritma *clustering* data, yaitu setiap titik data masuk ke sebuah *cluster* dengan ditandai oleh derajat keanggotaan. FCM membagi sebuah koleksi dari n data vektor x_j ($j=1, 2, \dots, n$) menjadi c *cluster* dan menemukan sebuah pusat *cluster* (*center*) untuk tiap kelompok dengan meminimalkan ukuran dari fungsi objektif. Pada FCM hasil *clustering* adalah sebuah titik data dapat menjadi anggota untuk beberapa *cluster* yang ditandai oleh derajat keanggotaannya antara nol dan satu.⁹ Berikut ini tahapan *clustering* menggunakan algoritma FCM:

Langkah ke-1: Inisialisasi keanggotaan matriks U yang berisi derajat keanggotaan terhadap *cluster* dengan nilai antara 0 dan 1, sehingga:

$$\sum_{i=1}^c u_{ij} = 1, \quad \forall j = 1, \dots, n \quad (1)$$

Langkah ke-2: Penghitungan c sebagai pusat *cluster*, c_i , $i = 1, \dots, c$ dengan menggunakan

$$c_i = \frac{\sum_{j=1}^n (u_{ij})^m x_j}{\sum_{j=1}^n (u_{ij})^m} \quad (2)$$

Langkah ke-3: Penghitungan fungsi objektif (J_i):

$$J(U, c_1, \dots, c_c) = \sum_{i=1}^c J_i = \sum_{i=1}^c \sum_j^n u_{ij}^m d_j^2 \quad (3)$$

dengan u_{ij} adalah elemen matriks U yang bernilai antara 0 dan 1, $d_{ij} = \|c_i - x_j\|$ adalah jarak antara pusat *cluster* ke- i dan titik data ke- j , c_i adalah pusat *cluster* ke- i , $m \in [1, \infty]$ adalah parameter *fuzzifikasi*. Normalnya, nilai m berada pada selang $[1.25, 2]$.¹⁰ Kemudian kondisi berhenti dicek dengan syarat: Jika $(|J_t - J_{t-1}| < \text{nilai toleransi terkecil yang diharapkan})$ atau $(t > \text{maksimal iterasi})$ maka proses berhenti. Jika tidak memenuhi syarat tersebut, dilakukan operasi penjumlahan $t = t + 1$.

Langkah ke-4: Sebelum perhitungan diulang kembali dari langkah ke-2, matriks U baru dihitung lebih dulu dengan menggunakan formula:

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{d_{jk}}{d_{ij}} \right)^{\frac{2}{m-1}}} \quad (4)$$

Probabilistic Neural Network (PNN)

Probabilistic Neural Network diperkenalkan oleh Donald F. Specht tahun 1990 dalam tulisannya berjudul “*Probabilistic Neural Network*” yang merupakan penyempurnaan ide-ide sebelumnya yang telah dilakukannya sejak 1966. *Probabilistic Neural Network* dirancang berdasarkan ide dari teori probabilitas klasik yaitu Bayesian dan estimator pengklasifikasi Parzen untuk *Probability Density Function*. Dengan menggunakan pengklasifikasi Bayesian dapat ditentukan bagaimana sebuah data masukan diklasifikasi sebagai anggota suatu kelas dari beberapa kelas yang ada, yaitu yang mempunyai nilai maksimum pada kelas tersebut.

Adapun struktur dari PNN ini dapat dilihat pada Gambar 1, yang terdiri atas empat *layer* yaitu *input layer*, *pattern layer*, *summation layer*, dan *decision layer*. Dengan menerima vektor tes x dari *input layer*, keluaran dari *pattern layer* dapat dihitung melalui persamaan:

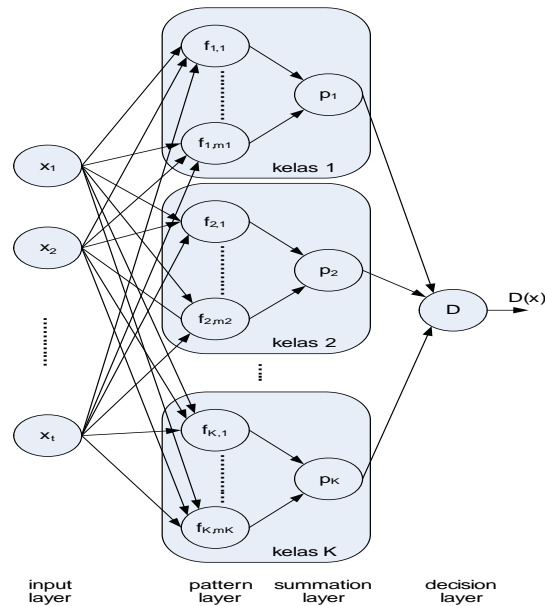
$$f(x) = \prod_{j=1}^d k \left(\frac{x_j - x_h}{h_j} \right) \quad (5)$$

dengan $d = \text{dimensi vektor}$, $k(z) = e^{-0.5z^2}$, $x_j = \text{vektor input kolom ke-}j$, $x_{ij} = \text{vektor bobot baris ke-}i \text{ kolom ke-}j$, $h_j = \text{smoothing parameter}$ ($a \times \text{simpangan baku ke-}j \times n^{1/5}$). Di sisi lain, untuk *summation layer* dihitung dengan persamaan:

$$p(x) = \frac{1}{(2\pi)^{d/2} h_1 h_2 \dots h_d n} \sum_{i=1}^n \left(\prod_{j=1}^d k \left(\frac{x_{ij} - x_j}{h_j} \right) \right), \dots \dots (6)$$

dengan n adalah banyaknya observasi.

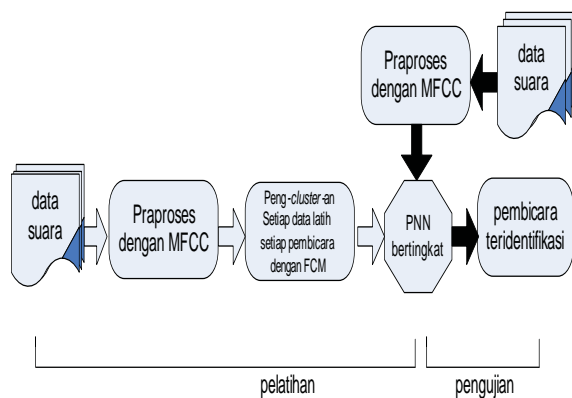
Suatu vektor tes x diklasifikasikan pada *decision layer* sebagai kelas Y jika nilai $D_Y(x)$ paling besar untuk kelas Y .



Gambar 1. Struktur *Probabilistic Neural Network*¹¹

METODOLOGI

Pada model yang akan dikembangkan ini, proses identifikasi terdiri atas dua fase, yaitu fase pelatihan dan fase pengujian. Pada fase pelatihan, data suara dari setiap pembicara dikumpulkan kemudian akan di-*cluster*-kan menggunakan FCM. Hasil peng-*cluster*-an inilah yang akan digunakan oleh PNN sebagai data pelatihan untuk membentuk pola setiap pembicara. Di sisi lain, pada fase pengujian diberikan contoh data suara hasil perekaman untuk diketahui pemilik suara tersebut. Untuk lebih jelasnya, kedua fase ini dapat dilihat pada Gambar 2.



Gambar 2. Blok diagram sistem identifikasi pembicara pada penelitian ini.

Data Suara

Data yang digunakan pada penelitian ini adalah gelombang suara yang telah didigitasi dan direkam dari sepuluh pembicara, yaitu lima pembicara laki-laki dan lima pembicara perempuan dengan rentang usia 0–25 tahun. Setiap pembicara diambil suaranya dalam jangka waktu yang sama dan tanpa pengarahan (*unguided*). Yang dimaksud tanpa pengarahan adalah pembicara dapat menggunakan cara, intonasi, dan logat apa pun pada saat merekam data.

Jenis identifikasi pengucapan pembicara yang dilakukan bersifat bergantung pada teks, maka kata yang diucapkan baik untuk pelatihan maupun pengujian telah ditentukan yaitu “komputer”. Kata tersebut diucapkan sebanyak 60 kali oleh setiap pembicara, sehingga terdapat 600 *file* data. Di samping itu, diperlukan juga data ber-*noise* dengan jumlah yang sama untuk mengetahui pengaruh *noise* terhadap akurasi model yang dikembangkan. Besarnya *noise* yang ditambahkan pada suara asli yang akan dicobakan adalah dengan tingkat *signal to noise ratio* (SNR) sebesar 20 dan 30.

Untuk melihat pengaruh banyaknya proporsi data pelatihan terhadap tingkat akurasi model yang dihasilkan akan dicobakan tiga kombinasi proporsi pembagian data pelatihan. Kombinasi ini juga dipakai untuk data yang diberi tambahan *noise*. Kombinasi proporsi tersebut dapat dilihat pada Tabel 1.

Tabel 1. Kombinasi Proporsi Data Pelatihan dan Data Pengujian

Data Pelatihan	Data Pengujian	Jenis Data
20	20	Data asli
30	20	
40	20	
20	20	Data dengan SNR 30
30	20	
40	20	
20	20	Data dengan SNR 20
30	20	
40	20	

Praproses Sinyal Suara dengan MFCC

Praproses ini merupakan ekstraksi ciri sinyal suara menggunakan MFCC. Pada implementasi MFCC ini, kecuali tahap *frame blocking*, digunakan fungsi dari *Auditory Toolbox* yang dikembangkan oleh Slaney.¹²

Dalam penelitian ini, parameter yang digunakan disesuaikan dengan yang digunakan pada penelitian Suhartono tahun 2007, pertama adalah *input* masukan suara yang berasal dari tiap pembicara, *sampling rate* sebesar 16.000 Hz, *time frame* sebesar 30 ms, *lap* sebesar 0.5, dan *cepstral coefficient* sebanyak 13.

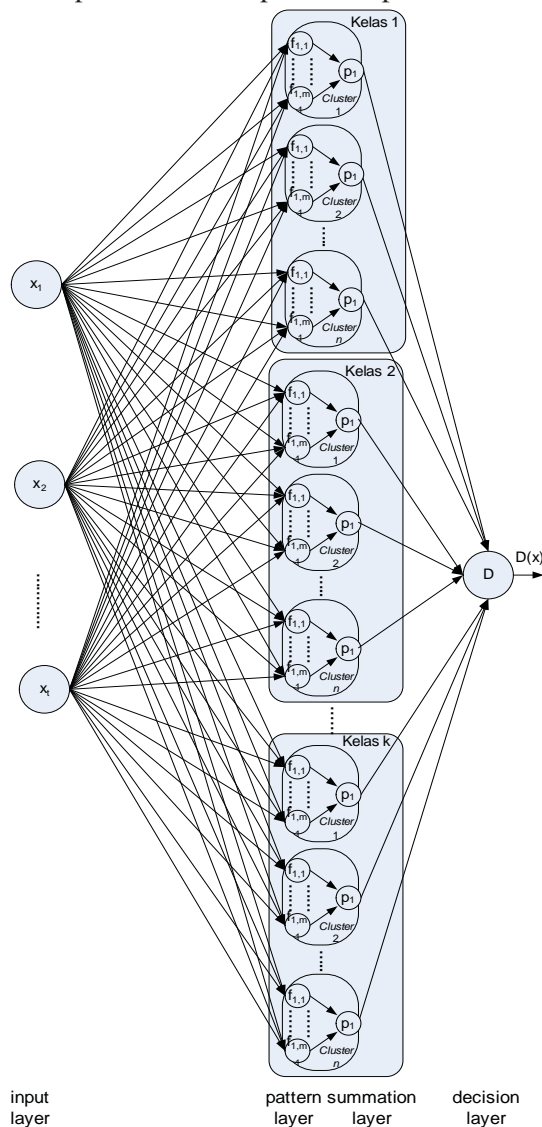
Model Hybrid Probabilistic Neural Network dan Fuzzy C-Means (FCM)

Pada model yang akan dikembangkan ini, data untuk pelatihan dari tiap-tiap pembicara di-*cluster*-kan menggunakan *Fuzzy C-Means* (FCM) dengan parameter standar yang digunakan oleh MATLAB, yaitu *cluster* sebanyak tiga, *error* terkecil yang diharapkan sebesar 1×10^{-5} , pangkat pembobot sebesar dua, dan maksimum iterasi sebanyak seratus. Selanjutnya, data setiap *cluster* yang terbentuk ini akan digunakan oleh *pattern layer* PNN sebagai data latih setiap kelas.

Proses identifikasi model ini dimulai pada *input layer* PNN yang berupa data tes dari suara pembicara yang akan diidentifikasi setelah melalui praproses. Setelah itu, pada *pattern layer* dilakukan penghitungan dengan menggunakan persamaan (5). Selanjutnya, pada *summation*

layer PNN akan dilakukan penjumlahan setiap keluaran dari *pattern layer* yang satu kelas dengan menggunakan persamaan (6).

Dari bagian *decision layer* PNN akan diperoleh nilai terbesar berdasarkan hasil perhitungan *summation layer* PNN. Nilai terbesar yang ini mencerminkan bahwa suara yang diujikan diidentifikasi sebagai pembicara kelas tersebut. Dari penjelasan itu, struktur model *hybrid Probabilistic Neural Network* dan *Fuzzy C-Means* (FCM) untuk identifikasi pembicara dalam penelitian ini dapat dilihat pada Gambar 3.



Gambar 3. Struktur model *hybrid Probabilistic Neural Network* dan *Fuzzy C-Means* untuk identifikasi pembicara pada penelitian ini.

Hasil Identifikasi

Hasil identifikasi merupakan bagian akhir dari identifikasi pembicara yang berupa identifikasi pemilik suara berdasarkan *input* suara yang diujikan. Identifikasi yang dimaksud adalah apakah suara yang diujikan diidentifikasi sebagai pembicara 1, 2, 3, 4, 5, 6, 7, 8, 9, atau 10.

Selain itu, dari hasil identifikasi ini juga dapat dihitung tingkat keberhasilan dari model. Tingkat keberhasilan model ini dapat dilihat dari besarnya tingkat akurasi identifikasi yang dihasilkan oleh model. Tingkat akurasi dihitung sebagai rasio antara jumlah data pembicara yang diidentifikasi secara benar dan jumlah semua data pembicara yang diujikan.

Perhitungan Akurasi

Hasil yang diamati pada penelitian ini adalah tingkat akurasi model dalam mengidentifikasi data pengujian. Tingkat akurasi diperoleh dengan perhitungan:

$$\text{akurasi} = \frac{\sum \text{data uji benar diidentifikasi}}{\sum \text{data uji}} \times 100\% \quad (7)$$

HASIL DAN PEMBAHASAN

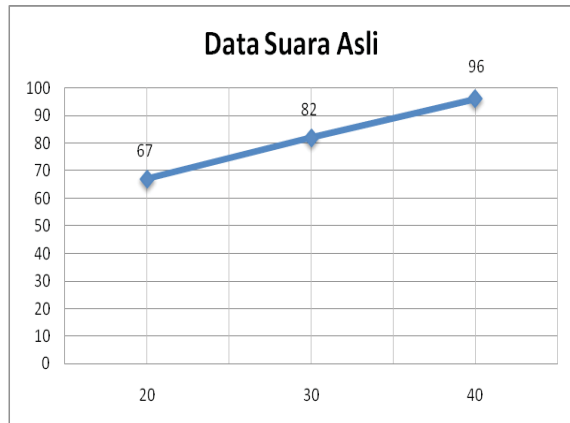
Setelah model selesai dibangun, tahap selanjutnya adalah melakukan tes pada model untuk mengetahui tingkat akurasi model dengan menghitung hasil identifikasi terhadap setiap data tes.

Hasil identifikasi terkait erat dengan *decision layer* pada model. Dari *decision layer* akan diperoleh nilai maksimum untuk suatu kelas. Nilai terbesar ini mencerminkan bahwa suara yang diujikan diidentifikasi sebagai pembicara kelas tersebut.

Setelah data suara tes melalui setiap *layer* model, maka akan dapat diketahui identitas pemilik suara yang diujikan. Identitas yang dimaksud adalah apakah sebagai pembicara 1, 2, 3, 4, 5, 6, 7, 8, 9, atau pembicara 10.

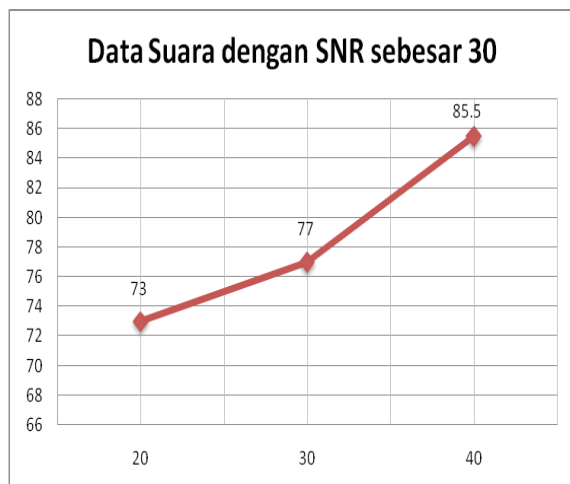
Untuk model dengan menggunakan data asli, tingkat akurasi hasil identifikasi model ditunjukkan pada Gambar 4. Dari hasil identifikasi tersebut, untuk model dengan data latih sebanyak 20 data per pembicara menghasilkan akurasi identifikasi sebesar 67%, dengan data latih sebanyak 30 menghasilkan akurasi identifikasi 82%, dan

dengan data latih sebanyak 40 menghasilkan akurasi identifikasi sebesar 96%. Dari gambar tersebut terlihat peningkatan akurasi yang cukup tinggi seiring dengan bertambahnya jumlah data pelatihan yang digunakan.



Gambar 4. Tingkat akurasi model dengan menggunakan data asli.

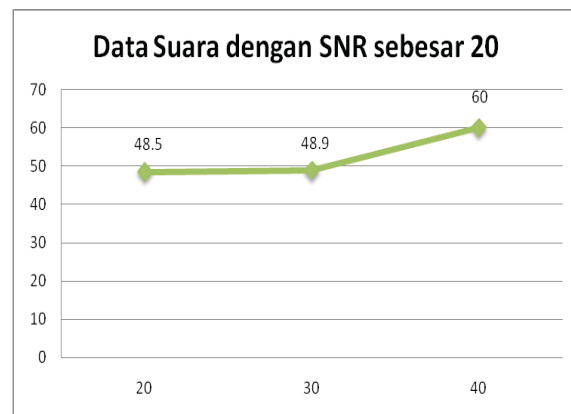
Dari hasil identifikasi untuk model dengan menggunakan data ber-*noise* dengan SNR sebesar 30 sebagaimana terlihat pada Gambar 5, untuk model dengan data latih sebanyak 20 data per pembicara menghasilkan akurasi identifikasi sebesar 73%, dengan data latih sebanyak 30 menghasilkan akurasi identifikasi 77%, dan dengan data latih sebanyak 40 menghasilkan akurasi identifikasi sebesar 85,5%. Dari gambar tersebut terlihat juga peningkatan akurasi yang cukup tinggi seiring dengan bertambahnya jumlah data pelatihan yang digunakan, meskipun data tersebut dipengaruhi oleh *noise* dengan SNR sebesar 30.



Gambar 5. Tingkat akurasi model dengan meng-

gunakan data dengan SNR sebesar 30.

Untuk model dengan menggunakan data ber-*noise* dengan SNR sebesar 20 sebagaimana terlihat pada Gambar 6, tingkat tersebut, untuk model dengan data latih sebanyak 20 data per pembicara menghasilkan akurasi identifikasi sebesar 48,5%, dengan data latih sebanyak 30 menghasilkan akurasi identifikasi 48,9%, dan dengan data latih sebanyak 40 menghasilkan akurasi identifikasi sebesar 60%. Dari gambar tersebut terlihat juga peningkatan akurasi meskipun tidak terlalu tinggi dibandingkan dengan jenis data yang lain seiring dengan bertambahnya jumlah data pelatihan yang digunakan. Hal ini disebabkan *noise* yang diberikan pada jenis data suara yang digunakan ini lebih besar dibandingkan dengan kedua jenis data sebelumnya.

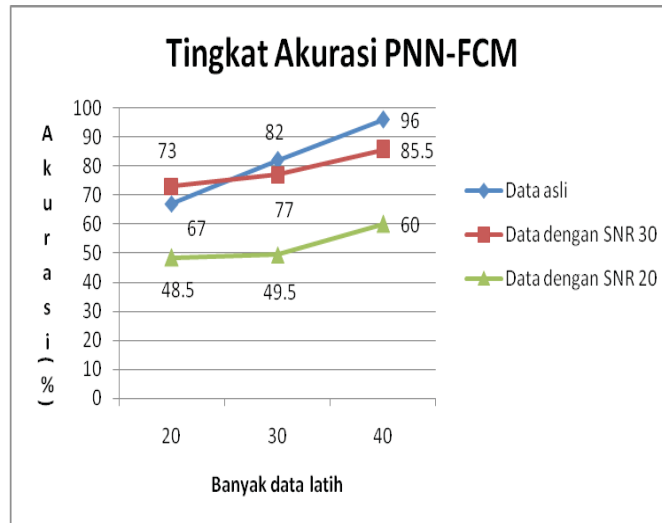


Gambar 6. Tingkat akurasi model dengan menggunakan data dengan SNR sebesar 30.

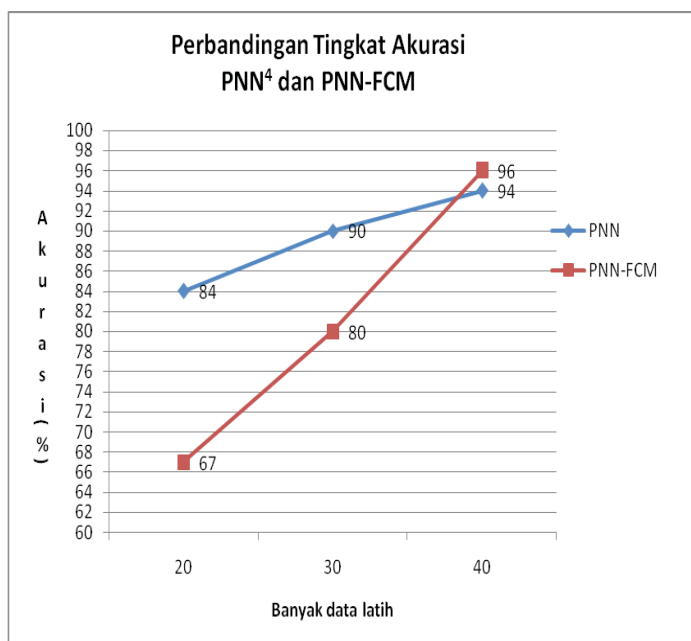
Perbandingan ketiga jenis data yang digunakan model dengan jumlah data pengujian yang sama terlihat pada Gambar 7.

Dengan jumlah data pengujian yang sama, dari Gambar 7 terlihat bahwa untuk setiap jenis data yang digunakan pada model terjadi peningkatan akurasi identifikasi seiring dengan bertambahnya jumlah data pelatihan yang diberikan. Peningkatan akurasi ini disebabkan oleh semakin banyaknya data pelatihan yang digunakan akan membuat semakin banyak pola yang dapat dikenali oleh model.

Selain itu, dari Gambar 7 terlihat bahwa secara umum terjadi penurunan akurasi model apabila suara yang digunakan mengandung *noise*. Semakin tinggi *noise* yang tergabung dalam suara, semakin kecil tingkat akurasi



Gambar 7. Perbandingan tingkat akurasi antara data data asli dan data yang ditambahkan *noise*.



Gambar 5. Perbandingan tingkat akurasi antara model PNN⁶ dan PNN-FCM.

model, terkecuali untuk data dengan SNR 30 pada komposisi data pelatihan sebanyak 20 dan data pengujian sebanyak 30 mengalami peningkatan akurasi dibandingkan dengan data aslinya. Hal ini kemungkinan disebabkan oleh sedikit penambahan *noise* (dengan SNR 30) yang diberikan mengakibatkan ekstraksi ciri suatu data dapat terbaca dengan benar dibandingkan dengan data aslinya, sehingga data tersebut dapat diidentifikasi dengan benar.

Perbandingan dengan Penelitian Sebelumnya

Untuk melihat pengaruh pengelompokan dengan metode *Fuzzy C-Means* (FCM), Gambar 5 disajikan tingkat akurasi model PNN saja tanpa pengelompokan yang dikembangkan oleh Suhartono pada tahun 2007 dan tingkat akurasi model *hybrid* PNN-FCM yang telah dikembangkan pada penelitian ini. Dari Gambar 5 terlihat adanya perbedaan tingkat akurasi dari kedua model pada setiap data pelatihan yang sama. Dengan ukuran

data pelatihan yang lebih sedikit, tingkat akurasi model *hybrid* PNN-FCM lebih rendah dibandingkan dengan model PNN saja. Hal ini disebabkan pada model *hybrid* PNN-FCM data pelatihan setiap pembicara menjadi lebih sedikit akibat adanya pengelompokan dibandingkan dengan model PNN. Akan tetapi, seiring dengan bertambahnya data pelatihan, akurasi model *hybrid* PNN-FCM terus meningkat dan pada akhirnya dengan jumlah data pelatihan yang sama (40 data pelatihan), pengaruh pengelompokan dengan metode *Fuzzy C-Means* terlihat dan berhasil menaikkan akurasi identifikasi menjadi 96%.

KESIMPULAN

Dari penelitian yang telah dilakukan, diperoleh suatu model *hybrid Probabilistic Neural Network* dan *Fuzzy C-Means* (FCM) untuk Identifikasi Pembicara. Tingkat akurasi dengan jumlah data yang sama untuk model yang menggunakan data asli mencapai nilai tertinggi pada proporsi data dengan 40 data pelatihan yaitu sebesar 96%. Di sisi lain, tingkat akurasi untuk model yang menggunakan data dengan tambahan *noise* dengan *signal to noise ratio* (SNR) sebesar 30 mencapai nilai tertinggi pada proporsi data latih sebanyak 40, yaitu sebesar 85,5%. Tingkat akurasi untuk model yang menggunakan data dengan tambahan *noise* dengan *signal to noise ratio* (SNR) sebesar 20 mencapai nilai tertinggi pada proporsi data dengan 40 data pelatihan, yaitu sebesar 60%. Selain itu, dari penelitian ini diketahui bahwa prosedur pengelompokan dengan *Fuzzy C-Means* dapat meningkatkan akurasi hingga 96% untuk ukuran data pelatihan yang besar.

UCAPAN TERIMA KASIH

Saya ucapkan terima kasih kepada Pak Agus Buono dan Ibu Sri Nurdiati selaku pembimbing saya dalam penelitian ini, Departemen Ilmu Komputer IPB yang telah menyediakan sarana dan prasara dalam pelaksanaan penelitian ini, dan Pak Handoko yang telah membimbing dan memberikan banyak masukan dalam penulisan ini.

DAFTAR PUSTAKA

- ¹Campbell, Jr. J.P. 1997. *Speaker Recognition: A Tutorial*. Proceeding IEEE, Vol. 85 No. 9, hLm. 1.437–1.461, September 1997.
- ²Buono, A., Wisnu J., dan Benyamin K. 2009. Perluasan Metode MFCC 1D ke 2D sebagai Ekstraksi Ciri pada Sistem Identifikasi Pembicara Menggunakan Hidden Markov Model (HMM). *Makara, Sains* 13(1): 87–93.
- ³Islam, R. Md. dan Md. Fayzur Rahman. 2010. Noise Robust Speaker Identification using PCA based Genetic Algorithm. *International Journal of Computer Applications* (0975–8887) 4(12).
- ⁴Pawar, R.V. dan S.N. Mali. 2005. Speaker Identification using Neural Networks. *World Academy of Science, Engineering and Technology* 12 2005. (<http://www.waset.org/journals/waset/v12/v12-7.pdf>, diakses 10 Maret 2012).
- ⁵Sarimollaoglu, M., Serhan D., Kamran I., dan Coskun B. 2004. A Text-independent Speaker Identification System Using Probabilistic Neural Network. (http://bayrak.ualr.edu/symsel/mustafa/docs/CCCT_2004.pdf, diakses 19 September 2006).
- ⁶Suhartono N. 2007. *Pengembangan Model Probabilistic Neural Network untuk Identifikasi Pembicara*. Skripsi. Bogor: Departemen Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Pertanian Bogor.
- ⁷Purnamasari W. 2006. *Pengembangan Model Markov Tersembunyi untuk Identifikasi Pembicara*. [Skripsi]. Bogor: Departemen Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Pertanian Bogor.
- ⁸Yeh Yun-Chi, Wen-June Wang, dan Che Wun Chio. A novel fuzzy c-means method for classifying heartbeat cases from ECG signals. *Measurement* 43 (2010): 1.542–1.555.
- ⁹Jang J.S.R., Sun C.T., dan Mizutani Eiji. 1997. *Neuro-Fuzzy and Soft Computing*. London: Prentice-Hall International, Inc.
- ¹⁰Cox, E. 2005. *Fuzzy Modelling and Genetic Algorithms for Data Mining and Exploration*. USA: Academic Press.
- ¹¹Ganchev, T.D. 2005. *Speaker Recognition*. [Tesis]. Greece: Wire Communication Laboratory, Department of Computer and Electrical Engineering, University of Patras.
- ¹²Slaney, Malcom. 1998. *Auditory Toolbox: A MATLAB Toolbox for Auditory Modeling Work Version 2*. Interval Research Corporation.