

KLASIFIKASI SINYAL *PHONOCARDIOGRAM* MENGGUNAKAN *SHORT TIME FOURIER TRANSFORM* DAN *CONVOLUTIONAL NEURAL NETWORK*

Muhammad Alwi Adnan Amal¹, Dodi Zulherman^{*2}, Rahmat Widadi³

^{1,2,3}Institut Teknologi Telkom Purwokerto, Kabupaten Banyumas
Email: ¹17101066@itttelkom-pwt.ac.id, ²zulherman.dodi@itttelkom-pwt.ac.id, ³rahmat@itttelkom-pwt.ac.id
^{*}Penulis Korespondensi

(Naskah masuk: 10 Agustus 2021, diterima untuk diterbitkan: 10 April 2023)

Abstrak

Berdasarkan laporan *American Heart Association*, penyakit kardiovaskular menjadi penyebab kematian global tertinggi. *Phonocardiogram* (PCG) dan *electrocardiogram* (ECG) biasanya digunakan untuk mendeteksi penyakit jantung. Penggunaan sinyal PCG memberikan hasil prediksi yang lebih baik pada deteksi penyakit jantung bila dibandingkan dengan ECG. Tetapi, penggunaan PCG secara elektronik membutuhkan analisis sinyal yang kompleks untuk mengklasifikasikan kondisi jantung. Penelitian ini bertujuan merancang suatu sistem klasifikasi sinyal PCG berdasarkan metode ekstraksi fitur menggunakan *Short Time Fourier Transform* (STFT) dan metode klasifikasi menggunakan *Convolutional Neural Network* (CNN). Pengujian rancangan sistem menggunakan dataset sekunder dengan 2.575 rekaman PCG normal dan 665 rekaman PCG abnormal dalam format wav. Pengujian kinerja menggunakan variasi *Hamming*, *Hann* dan *Blackman-Harris Window* pada bagian ekstraksi fitur dan variasi jumlah *layer* konvolusi pada bagian klasifikasi. Berdasarkan hasil pengujian, penggunaan *hamming window* pada proses ekstraksi fitur dan 4 *layer* konvolusi pada proses klasifikasi memberikan hasil terbaik dengan tingkat akurasi 88,11%. Penelitian ini membuktikan bahwa penggunaan *hamming window* pada bagian ekstraksi fitur dan 4 *layer* konvolusi pada bagian klasifikasi sebagai bentuk model terbaik sistem klasifikasi PCG berdasarkan STFT dan CNN.

Kata kunci: *Phonocardiogram*, *Short Time Fourier Transform* (STFT), *Hamming Window*, *Convolutional Neural Network* (CNN)

PHONOCARDIOGRAM SIGNAL CLASSIFICATION USING SHORT TIME FOURIER TRANSFORM AND CONVOLUTIONAL NEURAL NETWORK

Abstract

According to a report by the *American Heart Association*, cardiovascular disease is the leading global cause of death. *Phonocardiogram* (PCG) and *electrocardiogram* (ECG) are commonly used to detect heart disease. The use of PCG signals provides better predictive results in the detection of heart disease when compared to ECG. However, the use of PCG electronically requires complex signal analysis to classify heart conditions. This study aims to design a PCG signal classification system based on the extraction method using the *Short Time Fourier Transform* (STFT) and the classification method using the *Convolutional Neural Network* (CNN). The system design test used a secondary dataset with 2,575 normal PCG records and 665 abnormal PCG records in wav format. Performance testing uses variations of *Hamming*, *Hann* and *Blackman-Harris Window* in the feature extraction section and variations in the number of convolution layers in the classification section. Based on the test results, the use of a *hamming window* in the feature extraction process and 4 convolution layers in the classification process gives the best results with an accuracy rate of 88.11%. This study proves that the use of a *hamming window* in the feature extraction section and 4 convolution layers in the classification section is the best form of the PCG classification system based on STFT and CNN.

Keywords: *Phonocardiogram*, *Short Time Fourier Transform* (STFT), *Hamming Window*, *Convolutional Neural Network* (CNN)

1. PENDAHULUAN

Berdasarkan laporan *American Heart Association*, penyakit kardiovaskular menjadi penyebab kematian global tertinggi dan diperkirakan lebih dari 22,2 juta kematian pada tahun 2030 (Virani et al., 2020). Seseorang yang mengidap penyakit kardiovaskular memerlukan pendeteksian dini untuk mencegah kematian mendadak. *Electrocardiogram* (ECG) dan *Phonocardiogram* (PCG) biasanya digunakan untuk mendeteksi penyakit jantung (Deperlioglu, 2018). Penggunaan PCG pada penyakit jantung memberikan prediksi penyakit yang lebih baik dibandingkan ECG (Bashar, Dandapat and Kumazawa, 2018). Penggunaan PCG secara elektronik membutuhkan analisis sinyal yang kompleks untuk mengekstraksi dan menganalisis fitur-fitur yang terdapat pada sinyal PCG karena *noise* yang ikut dalam perekaman dapat mengurangi tingkat keakuratan dalam diagnosis secara klinis (Chowdhury, Poudel and Hu, 2020). Oleh karena itu, penggunaan sinyal PCG membutuhkan metode ekstraksi dan klasifikasi untuk menginterpretasikan isi informasi yang terkandung di dalam potongan sinyal.

Klasifikasi sinyal PCG secara otomatis dapat dilakukan untuk membedakan kondisi jantung normal dan abnormal dari sinyal yang telah dilakukan ekstraksi fitur. Metode CNN termasuk algoritma yang dapat digunakan dalam melakukan klasifikasi sinyal PCG, khususnya untuk jumlah data yang banyak (Li et al., 2020). CNN merupakan salah satu algoritma *deep learning* yang dikembangkan dari *Multi Layer Perceptron* (MLP) (Putra and Suartika, 2016). Berdasarkan hasil studi literatur, ditemukan beberapa penelitian yang menggunakan CNN untuk klasifikasi suara jantung. Pada penelitian (Ryu, Park and Shin, 2016) sinyal PCG yang telah difilter menggunakan *Windowed-Sinc Hamming Filter* digunakan sebagai masukan *Convolutional Neural Network* (CNN). Hasil yang diperoleh menggunakan model tersebut yaitu tingkat akurasi 79,5%, dengan sensitivitas 70,8%, dan spesifisitas 88,2%. Penelitian (Meintjes, Lowe and Legget, 2018) membandingkan berbagai metode klasifikasi *Support Vector Machine* (SVM), *K-Nearest Neighbor* (KNN), dan CNN untuk mengetahui tingkat akurasi tertinggi. Hasilnya tingkat akurasi CNN lebih tinggi dari metode klasifikasi lain dengan nilai 86 %. Penelitian (Li et al., 2020) juga menggunakan CNN pada tahap klasifikasi. Fitur diperoleh dari berbagai domain yaitu domain waktu, frekuensi, statistik orde tinggi, *entropy*, *Cyclostationarity*, *Cepstrum*, dan energi. Pada penelitian ini diperoleh tingkat akurasi 86,8%, sensitivitas 87%, dan spesifisitas 86,6%.

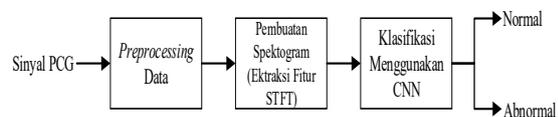
Salah satu metode analisis sinyal yang dapat digunakan dalam proses ekstraksi fitur adalah *Short Time Fourier Transform* (STFT) (Sumarna, Purwanto and Agustika, 2017). STFT mencuplik sinyal dengan durasi waktu tertentu yang selanjutnya dikonversi ke dalam domain frekuensi yang kemudian sinyal

dicuplik selama t detik sehingga posisi sinyal dapat dipetakan dalam domain waktu dan frekuensi secara bersamaan (Sumarna, Purwanto and Agustika, 2017). Dalam hal ini STFT mengubah sinyal satu dimensi menjadi sinyal dua dimensi (citra). Sinyal dua dimensi lebih cocok untuk digunakan pada CNN (Qi et al., 2021). Penggunaan STFT dan CNN pada naskah ini menjadi metode yang diusulkan dan akan diteliti tingkat akurasi, sensitivitas dan spesifisitas model.

Langkah awal dalam penelitian ini, sampel suara jantung yang diperoleh dari *PhysioNet/Computing in Cardiology Challenge 2016* dilakukan *preprocessing*. *Preprocessing* dibagi menjadi beberapa bagian yaitu pelabelan data, pemotongan rekaman suara jantung dan *zero padding*. Setelah *preprocessing*, sinyal PCG dilakukan pemotongan sinyal dengan *Hamming window*, *Hann window* dan *Blackman-Harris window* dengan setiap potongan sinyal ditransformasikan ke dalam transformasi *fourier* sesuai metode STFT, dan hasilnya berupa fitur yang diklasifikasikan menggunakan metode CNN. Pengukuran kinerja rancangan menggunakan variasi *layer* konvolusi pada bagian CNN dengan parameter analisis akurasi, sensitivitas dan spesifisitas.

2. METODE PENELITIAN

Metode penelitian digunakan mempermudah dan sebagai panduan dalam pelaksanaan penelitian. Gambar 1 menunjukkan langkah-langkah yang dilakukan pada penelitian ini. Langkah-langkah tersebut meliputi *input* sinyal PCG, sinyal PCG dilakukan *preprocessing*, hasil dari *preprocessing* akan diteruskan pada proses ekstraksi fitur menggunakan STFT dan tahapan klasifikasi menggunakan CNN. Hasil dari klasifikasi nantinya akan dibandingkan dengan data asli untuk mendapatkan nilai *accuracy*, *sensitivity* dan *specificity*.



Gambar 1. Diagram Blok Perancangan Sistem

2.1 Input sinyal PCG

Sinyal PCG yang akan digunakan adalah sinyal PCG normal dan abnormal (mengandung murmur) yang diperoleh dari *PhysioNet / Computing in Cardiology Challenge 2016* (Clifford et al., 2016). Frekuensi pengambilan sampelnya adalah 2000 Hz dengan panjang rekaman bervariasi dari 5 detik hingga lebih dari 120 detik, semuanya dalam format wav. Dataset terdiri dari 2.575 sinyal PCG normal dan 665 sinyal PCG abnormal, diberi label masing-masing secara manual dengan -1 dan 1.

2.2 Preprocessing Data

Preprocessing dibagi menjadi beberapa bagian yaitu pelabelan data, pemotongan rekaman suara jantung dan *zero padding*. Pelabelan data digunakan untuk mengkategorikan sinyal suara jantung normal dan abnormal berdasarkan *file reference-SQI.CSV* yang ada pada dataset. Setelah pelabelan, suara jantung normal akan memiliki label -1 dan suara jantung abnormal akan memiliki label 1. *Zero padding* digunakan untuk menyamakan ukuran *input* sinyal suara jantung. Karena *sampling rate* ditentukan sebesar 22050 Hz sedangkan *raw* data dari *files* rekamannya hanya memiliki 2000 Hz maka kita memerlukan *zero padding*. Konsep dari *zero padding* adalah menambahkan nilai 0 ke akhir sinyal domain waktu untuk menambah panjangnya. Panjang rekaman suara jantung menjadi 2 detik dengan rentang frekuensi sepanjang 0–44.100 Hz.

2.3 Ekstraksi Fitur STFT

STFT merupakan pengembangan Transformasi Fourier untuk mengatasi sinyal non-stasioner dengan menerapkan *window* pada analisis sinyal tersegmentasi (Krishnan, 2021). Tahapan pertama dari ekstraksi fitur pada STFT adalah menentukan nilai parameter sebagai *input* fungsi STFT, yaitu *hop size*, *window length*, dan *fft points* (*n_fft*). *Window length* ditentukan sebesar 2048. Nilai tersebut dipilih agar diperoleh resolusi waktu yang bagus. Resolusi waktu yang bagus adalah resolusi dengan jarak antar *window* yang lebar. Namun dengan lebarnya *window*, maka resolusi frekuensi pun menjadi tidak bagus. Oleh karena itu diperlukan nilai dari *hop size* yang lebih kecil untuk dapat mengoptimalkan *range* dari resolusi frekuensi. Dimana pada penelitian ini ditentukan *hop size* sebesar 512, dengan *range* frekuensi yang semakin kecil akan memberikan hasil yang semakin akurat. *N_fft* pada penelitian ini ditentukan nilainya sama dengan nilai dari *window length*, hal ini dipilih sesuai dengan nilai default yang disediakan oleh *syntax* untuk kesesuaian dengan pemrosesan sinyal suara yang berasal dari suara fisik.

Tahapan selanjutnya adalah menentukan jenis *window* yang akan digunakan. Pada penelitian ini menggunakan *hamming*, *hann* dan *blackman-harris window*. Alasan mengapa menggunakan tiga *window function* tersebut karena *window function* tersebut paling banyak digunakan, dan pilihan ukuran fungsi jendela didasarkan pada hubungan antara periode sinyal pemindaian dan frekuensi *sampling* (Zhong et al., 2019).

Setelah menentukan jenis *window* yang akan digunakan, kemudian diambil nilai absolut STFT dari masing-masing *window*. Nilai absolut STFT tersebut akan digunakan sebagai masukan pada CNN. Tipe data dari nilai STFT adalah bilangan kompleks dengan sumbu *real* dan *imaginer*. Nilai STFT ini merepresentasikan nilai dalam domain waktu dan frekuensi. Sumbu x (kolom) dari nilai STFT merepresentasikan waktu, sedangkan sumbu y (baris)

merepresentasikan frekuensi. Nilai STFT dapat ditampilkan sebagai grafik yang menunjukkan informasi mengenai perubahan gelombang dalam rentang waktu, frekuensi, dan intensitas amplitudo. Magnitudo pada sebuah titik frekuensi dan waktu (waktu, frekuensi) pada spektrogram dinyatakan dengan nilai warna tertentu (*grayscale* atau RGB).

2.4 Pembuatan Data Latih dan Data Uji

Dataset yang diperoleh dari *PhysioNet/Computing in Cardiology Challenge 2016* dibagi menjadi 2 yaitu data yang digunakan untuk pelatihan dan data yang digunakan untuk pengujian. Pembagian data latih dan data uji dilakukan secara acak dengan jumlah data latih dan data uji memiliki perbandingan 80%:20%. Data latih menggunakan 2.592 *files* rekaman suara jantung dan data uji sebanyak 648 *files* rekaman suara jantung. Dalam penelitian ini, label kelas normal dan abnormal yang semula -1 dan 1 diganti dengan 0 dan 1. Penggantian label dilakukan karena menggunakan fungsi *keras.utils.to_categorical*, dimana fungsi tersebut memahami semua nilai sebagai positif.

2.5 Klasifikasi CNN

Penelitian ini mengusulkan metode untuk klasifikasi sinyal suara jantung normal dan abnormal menggunakan arsitektur CNN. CNN memiliki kemampuan menggeneralisasi lebih baik bila dibandingkan fully connected layers (Nebauer, 1998). CNN dideskripsikan sebagai pendekteksi fitur yang dapat mempelajari fitur abstrak secara efisien (Fieres, Schemmel and Meier, 2006). Dataset akan dilakukan *training* dan *validation* berulang-ulang untuk mendapatkan model CNN yang akan digunakan seperti jumlah layer, jumlah filter, ukuran filter dan *pool size*. *Dropout layer* ditambahkan agar fitur yang tidak dibutuhkan dibuang dan mengurangi terjadinya *overfitting*.

Pada penelitian ini digunakan 3 model CNN dengan jumlah *layer* konvolusi yang berbeda untuk melatih model dan melihat bagaimana kinerja model tersebut. Arsitektur model 1 CNN, model 2 CNN dan model 3 CNN terdapat pada Tabel 1, Tabel 2 dan Tabel 3. Pemilihan model CNN mempertimbangkan proses komputasi, jumlah parameter yang dilatih oleh model serta keakuratan model dalam melakukan proses klasifikasi suara jantung normal dan abnormal.

Tabel 1. Komposisi Arsitektur Model 1

Layer	Size Feature Map	Kernel Size / Pool Size
<i>Input</i>	(1025, 87, 1)	-
<i>Convolution 1</i>	(1023, 85, 32)	3 x 3
<i>Max Pooling 1</i>	(255, 21, 32)	4 x 4
<i>Convolution 2</i>	(253, 19, 64)	3 x 3
<i>Max Pooling 2</i>	(63, 4, 64)	4 x 4
<i>Flatten</i>	16128	-
<i>Dense 1</i>	1024	-
<i>Dense 2</i>	512	-
<i>Dense 3</i>	2	-
<i>(Output)</i>		

Tabel 2. Komposisi Arsitektur Model 2

Layer	Size Feature Map	Kernel Size / Pool Size
Input	(1025, 87, 1)	-
Convolution 1	(1023, 85, 32)	3 x 3
Max Pooling 1	(255, 21, 32)	4 x 4
Convolution 2	(253, 19, 64)	3 x 3
Convolution 3	(251, 17, 64)	3 x 3
Max Pooling 2	(62, 4, 64)	4 x 4
Flatten	15872	-
Dense 1	1024	-
Dense 2	512	-
Dense 3 (Output)	2	-

Tabel 3. Komposisi Arsitektur Model 3

Layer	Size Feature Map	Kernel Size / Pool Size
Input	(1025, 87, 1)	-
Convolution 1	(1023, 85, 32)	3 x 3
Max Pooling 1	(255, 21, 32)	4 x 4
Convolution 2	(253, 19, 64)	3 x 3
Convolution 3	(251, 17, 64)	3 x 3
Convolution 4	(249, 15, 64)	3 x 3
Max Pooling 2	(62, 4, 64)	4 x 4
Flatten	11904	-
Dense 1	1024	-
Dense 2	512	-
Dense 3 (Output)	2	-

2.6 Confusion Matrix

Confusion matrix digunakan untuk membandingkan nilai yang diprediksi dengan nilai aktual dan mendapatkan ukuran kesalahan klasifikasi (Nisbet, Miner and Yale, 2018). Dalam confusion matrix, semakin rendah jumlah kesalahan klasifikasi maka semakin baik kinerjanya. Pada Tabel 4 terdapat confusion matrix untuk model 2 kelas, dalam tabel tersebut terdapat nilai benar dan salah pada setiap kolom dan barisnya. Tabel disusun berdasarkan kelas aktual dan kelas yang diprediksi, sementara di dalam tabel terdapat komponen yang dapat diukur nilainya.

Tabel 4. Confusion Matrix 2 Kelas

	Predicted Negative	Predicted Positive
Actual Negative	True Negatives (TN)	False Positives (FP)
Actual Positive	False Negatives (FN)	True Positives (TP)

Parameter uji yang digunakan sebagai acuan dalam menilai model apakah model tersebut dikatakan baik atau tidak dibagi menjadi 3 yaitu accuracy, sensitivity dan specificity. Perhitungan nilai accuracy, sensitivity dan specificity ditunjukkan oleh persamaan (1), persamaan (2) dan persamaan (3).

$$Accuracy = \frac{TP+TN}{TP+TN+FN+FP} \tag{1}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{2}$$

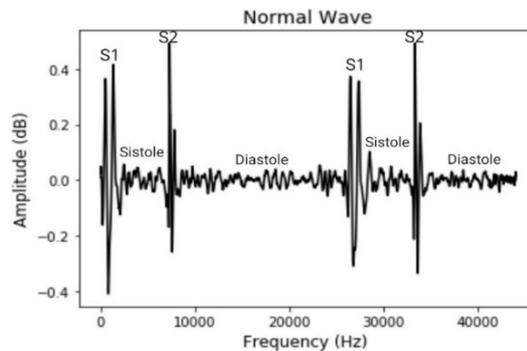
$$Specificity = \frac{TN}{TN + FP} \tag{3}$$

1. Accuracy merupakan kemampuan model dalam memprediksi kelas dengan benar.
2. Sensitivity merupakan perhitungan ketepatan klasifikasi pada jumlah data positif yang teridentifikasi benar sebagai data positif.
3. Specificity merupakan perhitungan ketepatan klasifikasi pada jumlah data negatif yang teridentifikasi benar sebagai data negatif.

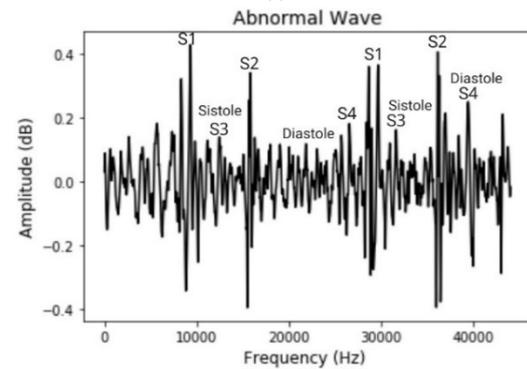
3. HASIL DAN PEMBAHASAN

3.1 Analisis Sinyal PCG

Pada penelitian ini menggunakan dataset training-e hingga training-f yang disediakan oleh PhysioNet Challenge 2016, sebanyak 3.240 data rekaman suara jantung, dengan 2.575 data rekaman suara jantung normal dan 665 data rekaman suara jantung abnormal. Seluruh data suara jantung pada dataset ini menggunakan frekuensi sampling 2000 Hz.



(a)



(b)

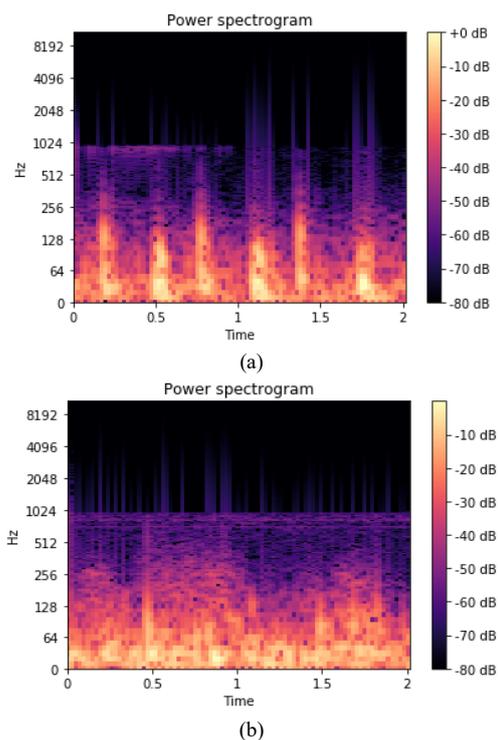
Gambar 2. Sinyal PCG (a) Normal, (b) Abnormal

Pada Gambar 2 (a) dan (b) terdapat tampilan dari sinyal PCG normal dan abnormal yang telah dipotong dengan durasi waktu 2 detik. Gambar tersebut merupakan representasi dari suara jantung normal dan abnormal yang terekam. Data perekaman ini masih direpresentasikan terhadap domain frekuensi. Data awal menunjukkan nilai power frequency-domain dari rentang 0.4 sampai dengan -0.4 dengan rentang frekuensi sepanjang 0 sampai 44.100 Hz. Diantara suara jantung pertama (S1) dan suara jantung kedua (S2) terdapat interval yaitu sistole,

sedangkan diantara S2 dan S1 selanjutnya terdapat interval yaitu diastole. Jika diperhatikan suara jantung normal hanya mempunyai nilai S1 dan S2 saja sedangkan pada suara jantung abnormal mempunyai nilai S1, S2, S3 dan S4.

3.2 Analisis Hasil Ekstraksi Fitur

Sinyal masukan yang diperoleh diolah menggunakan STFT dari suatu sinyal $x(t)$. STFT dihitung berdasarkan nilai sampel pada sinyal suara sebanyak panjang *window* yang digunakan dan kemudian dikalikan dengan fungsi *window*, dan fungsi *window* yang digunakan adalah *Hann*, *Hamming*, dan *Blackman-Harris*. Panjang *window* yang digunakan adalah 2048. Nilai tersebut disesuaikan dengan *range* frekuensi dan waktu yang digunakan pada penelitian ini. Jika menggunakan ukuran *window* yang lebih kecil, maka akan mempengaruhi *range* dari frekuensi dan waktu. Dan banyaknya sample suara yang digunakan adalah 512. Karena jangkauan *range* frekuensi pada penelitian ini sangat besar yaitu dari 0 Hz sampai 22050 Hz, namun jangkauan *range* waktu yang kecil yaitu antara 0 s sampai 2 s, maka, digunakanlah panjang *window* 2048. Hasil dari kalkulasi ini dimunculkan ke dalam gambar yang dikenal dengan sebutan spektrogram. Kemudian program akan menampilkan grafik dan spektrogram pada setiap *file* yang diekstraksi. Banyaknya sinyal suara responden yang akan diekstraksi adalah 3.153 suara, yang terdiri dari 2.575 suara jantung normal dan 665 suara jantung abnormal. Gambar 4.2 adalah spektrogram sinyal suara jantung normal yang diambil dengan menggunakan fungsi *hamming window* dengan panjang *window* 2048.



Gambar 3. Spektrogram (a) Sinyal PCG Normal Menggunakan *Hamming Window*, (b) Sinyal PCG Abnormal Menggunakan *Hamming Window*

Gambar 3 (a) menunjukkan representasi sinyal PCG normal dalam domain waktu pada sumbu x dan domain frekuensi pada sumbu y. Kekuatan sinyal atau *magnitude* direpresentasikan dalam domain warna, dimana warna kuning adalah amplitudo tertinggi dan warna ungu adalah amplitudo terendah. Pada Gambar 3 (b) menunjukkan amplitudo tertinggi pada sinyal PCG abnormal tersebar di beberapa daerah. Perbedaan mendasar antara spektrogram suara jantung normal dan abnormal terlihat pada bias warna kuning yang membentuk pola S1 dan S2 secara jelas. Apabila Gambar 3 (a) dan (b) dibandingkan, gambar (a) terlihat jelas S1 dan S2 pada rentang waktu 0 sampai 1 detik, sedangkan pada gambar (b) tidak terlihat jelas S1 dan S2 yang bias warnanya dikaburkan oleh S3 dan S4.

3.3 Analisis Pengaruh Variasi Window Function Terhadap Hasil Klasifikasi

Hasil pengujian sistem dari skema ini adalah membandingkan nilai *accuracy*, *sensitivity*, dan *specificity* yang dipengaruhi oleh variasi penggunaan *window function* pada proses STFT. Variasi *window function* yang digunakan adalah *hamming*, *hann* dan *Blackman-Harris*.

Tabel 5 Hasil Variasi Penggunaan *Window* dan Model CNN

Model CNN	Window Function	Accuracy (%)	Sensitivity (%)	Spesificity (%)
Model 1	<i>Hamming</i>	87,96	74,45	91,58
	<i>Hann</i>	86,72	82,48	87,86
	<i>Blackman</i>	87,19	75,91	90,21
Model 2	<i>Hamming</i>	87,65	78,10	90,21
	<i>Hann</i>	87,65	62,77	94,32
	<i>Blackman</i>	87,19	64,96	93,15
Model 3	<i>Hamming</i>	88,11	76,64	91,19
	<i>Hann</i>	87,96	67,15	93,54
	<i>Blackman</i>	87,34	72,99	91,29

Pada Tabel 5 menjelaskan hasil dari pengujian sistem pengaruh penggunaan *window function* terhadap nilai *accuracy*, *sensitivity*, dan *specificity*. Bisa dilihat untuk hasil dari nilai dari tiga parameter bervariasi disetiap *window function* dan model CNN yang digunakan, tentunya penggunaan *window function* dan model CNN sangat berpengaruh pada hasil yang didapat. Nilai *accuracy* terendah terdapat pada saat menggunakan *Hann window* dengan model 1 CNN yaitu sebesar 86,72 % dan *accuracy* tertinggi diperoleh pada saat menggunakan *Hamming window* dengan model 3 CNN yaitu sebesar 88,11 %. Nilai *sensitivity* terendah terdapat pada saat menggunakan *Hann window* dengan model 2 CNN yaitu sebesar 62,77 % dan *sensitivity* tertinggi diperoleh pada saat menggunakan *Hann window* dengan model 1 CNN yaitu sebesar 82,48 %. Nilai *specificity* terendah terdapat pada saat menggunakan *Hann window* dengan model 1 CNN yaitu sebesar 87,86 % dan *specificity* tertinggi diperoleh pada saat menggunakan

Hann window dengan model 2 CNN yaitu sebesar 94,32 %.

Berdasarkan hasil pengujian variasi *window function* terhadap tiga arsitektur model CNN, dapat disimpulkan bahwa dengan menggunakan *Hamming window* dalam proses ekstraksi fitur dengan STFT memiliki kinerja yang baik (*accuracy* terbaik) dalam melakukan klasifikasi terhadap sinyal PCG. Secara keseluruhan nilai *accuracy* ketika menggunakan *Hamming window* lebih baik daripada menggunakan *Hann window* dan *Blackman-Harris window*, adapun nilai sensitivitas dan spesifisitas yang diperoleh menggunakan *Hamming window* cukup baik. Setelah menentukan pemilihan *window function* yang digunakan, dilakukan perbandingan terhadap nilai parameter yang dihasilkan menggunakan *Hamming window* dengan tiga model CNN yang diusulkan.

3.4 Analisis Pengaruh Variasi Model CNN Terhadap Hasil Klasifikasi



Gambar 4 Perbandingan Performa Model CNN–*Hamming Window*

Analisis pengaruh variasi model CNN terhadap hasil klasifikasi dilakukan untuk membuat perbandingan hasil uji coba parameter akurasi, sensitivitas dan spesifisitas dari masing-masing model CNN. Berdasarkan hasil pada sub bab sebelumnya dapat disimpulkan bahwa dengan menggunakan *hamming window* pada proses STFT dapat meningkatkan hasil klasifikasi, baik dari segi akurasi, sensitivitas maupun spesifisitas. Penulis mencoba mengimplementasikan *hamming window* dengan menggunakan beberapa aksitektur CNN yang diusulkan. Pada Gambar 4 dapat dilihat bahwa nilai akurasi terbaik diperoleh ketika menggunakan model 3 CNN, sedangkan nilai sensitivitas lebih tinggi pada model 2. Nilai spesifisitas yang diperoleh model 3 diatas 90%. Ukuran filter masing–masing *layer* konvolusi pada model 1, model 2 dan model 3 menggunakan 3x3, dimana semakin kecil ukuran filter, semakin detail hasil yang dihasilkan. Akan tetapi, hal tersebut mengakibatkan waktu komputasi semakin lama. Jumlah parameter yang dilatih pada model 3 lebih sedikit daripada model 1 dan model 2, meskipun demikian nilai akurasi yang diperoleh lebih besar. Hal tersebut dikarenakan jumlah *layer* konvolusi yang digunakan lebih banyak dari model 1 dan model 2, sehingga fitur yang dihasilkan juga semakin banyak. Akan tetapi, hal tersebut berakibat

pada waktu komputasi yang semakin lama. Apabila dibandingkan dengan penelitian sejenis, akurasi yang diperoleh dari metode yang diusulkan memiliki akurasi yang lebih tinggi yaitu 88 %.

4. KESIMPULAN

Berdasarkan hasil percobaan yang telah diujikan, peneliti mengambil kesimpulan bahwa rancangan sistem klasifikasi terbaik diperoleh dengan variasi menggunakan *hamming window* pada proses ekstraksi fitur dan model 3 CNN pada proses klasifikasi. Rancangan model 3 CNN yang memberikan hasil terbaik terdiri dari 4 *layer* konvolusi 3x3, 2 *layer maxpooling* 4x4, 1 *flatten layer*, 2 *dropout layer* dan 3 *dense layer*. Berdasarkan hasil pengujian, dengan variasi menggunakan *hamming window* dan model 3 CNN didapatkan kinerja sistem terbaik dengan nilai akurasi 88,11%, sensitivitas 76,64% dan spesifisitas 91,19%. Hasil tersebut lebih baik dibandingkan dengan variasi menggunakan *hamming window* dan model 1 CNN dengan nilai akurasi 87,96%, sensitivitas 74,45% dan spesifisitas 91,58% serta variasi menggunakan *hamming window* dan model 2 CNN dengan nilai akurasi 87,65%, sensitivitas 78,1% dan spesifisitas 90,21%.

DAFTAR PUSTAKA

- BASHAR, M.K., DANDAPAT, S. AND KUMAZAWA, I., 2018. Heart Abnormality Classification Using Phonocardiogram (PCG) Signals. In: *2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)*. pp.336–340.
- CHOWDHURY, T.H., POUDEL, K.N. and HU, Y., 2020. Time-Frequency Analysis, Denoising, Compression, Segmentation, and Classification of PCG Signals. *IEEE Access*, 8, pp.160882–160890.
- CLIFFORD, G.D., LIU, C., MOODY, B., SPRINGER, D., SILVA, I., LI, Q. AND MARK, R.G., 2016. Classification of normal/abnormal heart sound recordings: The PhysioNet/Computing in Cardiology Challenge 2016. *Computing in Cardiology*, 43, pp.609–612.
- DEPERLIOĞLU, Ö., 2018. Classification Phonocardiograms with Convolutional Neural Networks. *BRAIN. Broad Research in Artificial Intelligence and Neuroscience*, 9(1), pp.5–13.
- FIERES, J., SCHEMMEL, J. AND MEIER, K., 2006. Training convolutional networks of threshold neurons suited for low-power hardware implementation. In: *The 2006 IEEE International Joint Conference on Neural Network Proceedings*. pp.21–28.
- KRISHNAN, S., 2021. 5 - Advanced analysis of biomedical signals. In: S. Krishnan, ed.

- Biomedical Signal Analysis for Connected Healthcare*. [online] Academic Press, pp.157–222. Available at: <<https://www.sciencedirect.com/science/article/pii/B9780128130865000037>>.
- LI, F., TANG, H., SHANG, S., MATHIAK, K. AND CONG, F., 2020. Classification of heart sounds using convolutional neural network. *Applied Sciences (Switzerland)*, 10(11).
- MEINTJES, A., LOWE, A. AND LEGGET, M., 2018. Fundamental Heart Sound Classification using the Continuous Wavelet Transform and Convolutional Neural Networks. In: *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. pp.409–412.
- NEBAUER, C., 1998. Evaluation of convolutional neural networks for visual recognition. *IEEE Transactions on Neural Networks*, 9(4), pp.685–696.
- NISBET, R., MINER, G. AND YALE, K., 2018. Chapter 11 - Model Evaluation and Enhancement. In: R. Nisbet, G. Miner and K. Yale, eds. *Handbook of Statistical Analysis and Data Mining Applications (Second Edition)*, Second Edition. [online] Boston: Academic Press, pp.215–233. Available at: <<https://www.sciencedirect.com/science/article/pii/B9780124166325000116>>.
- PUTRA, E. AND SUARTIKA, W., 2016. Klasifikasi Citra Menggunakan Convolutional Neural Network (CNN) pada Caltech 101. *Jurnal Teknik ITS*, 5(1).
- QI, Y., YANG, L., LIU, B., LIU, L., LIU, Y., ZHENG, Q., LIU, D. AND LUO, J., 2021. Accurate diagnosis of lung tissues for 2D Raman spectrogram by deep learning based on short-time Fourier transform. *Analytica Chimica Acta*, [online] 1179, p.338821. Available at: <<https://www.sciencedirect.com/science/article/pii/S0003267021006474>>.
- RYU, H., PARK, J. AND SHIN, H., 2016. Classification of heart sound recordings using convolution neural network. In: *2016 Computing in Cardiology Conference (CinC)*. pp.1153–1156.
- SUMARNA, S., PURWANTO, A. AND AGUSTIKA, D.K., 2017. Frequency Component Extraction of Heartbeat Cues With Short Time Fourier Transform (Stft). *Jurnal Sains Dasar*, 5(1), p.1.
- VIRANI, S.S., ALONSO, A., BENJAMIN, E.J., BITTENCOURT, M.S., CALLAWAY, C.W., CARSON, A.P., CHAMBERLAIN, A.M., CHANG, A.R., CHENG, S., DELLING, F.N., DJOUSSE, L., ELKIND, M.S.V., FERGUSON, J.F., FORNAGE, M., KHAN, S.S., KISSELA, B.M., KNUTSON, K.L., KWAN, T.W., LACKLAND, D.T., LEWIS, T.T., LICHTMAN, J.H., LONGENECKER, C.T., LOOP, M.S., LUTSEY, P.L., MARTIN, S.S., MATSUSHITA, K., MORAN, A.E., MUSSOLINO, M.E., PERAK, A.M., ROSAMOND, W.D., ROTH, G.A., SAMPSON, U.K.A., SATOU, G.M., SCHROEDER, E.B., SHAH, S.H., SHAY, C.M., SPARTANO, N.L., STOKES, A., TIRSCHWELL, D.L., VANWAGNER, L.B., TSAO, C.W., WONG, S.S. AND HEARD, D.G., 2020. *Heart disease and stroke statistics—2020 update: A report from the American Heart Association. Circulation*.
- ZHONG, Y. HONG, ZHANG, S., HE, R., ZHANG, J., ZHOU, Z., CHENG, X., HUANG, G. AND ZHANG, J., 2019. A convolutional neural network based auto features extraction method for tea classification with electronic tongue. *Applied Sciences (Switzerland)*, 9(12).

Halaman ini sengaja dikosongkan