

Evaluasi ICA dan NMF pada Pemisahan Sinyal Audio Menggunakan BSS Metrics dan MFCC

Ni Ketut Sukardiasih^{a1}, I Gede Arta Wibawa^{a2}

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana, Bali
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹sukardiasih.2308561036@student.unud.ac.id
²gede.arta@unud.ac.id

Abstract

Blind Source Separation (BSS) is an important problem in audio signal processing, particularly for stereo audio containing multiple sound sources. This study evaluates and compares the performance of Independent Component Analysis (ICA) and Non-negative Matrix Factorization (NMF) in separating two-source mixed audio signals. ICA is applied directly to stereo signals by exploiting spatial information, while NMF is performed on mono signals obtained by averaging stereo channels. Three pairs of audio data with diverse acoustic characteristics are used as test scenarios. Performance evaluation is conducted using Blind Source Separation metrics, namely Signal-to-Distortion Ratio (SDR), Signal-to-Interference Ratio (SIR), and Source-to-Artifacts Ratio (SAR), as well as spectral similarity analysis based on Mel-Frequency Cepstral Coefficients (MFCC). Robustness tests are carried out by adding noise at Signal-to-Noise Ratio (SNR) levels of 10 dB and 5 dB. The results show that ICA consistently outperforms NMF and demonstrates better robustness against noise.

Keywords: blind source separation, ICA, NMF, SDR, MFCC, audio stereo, robustness.

1. Pendahuluan

Blind Source Separation (BSS) merupakan teknik pemrosesan sinyal yang bertujuan untuk memisahkan sinyal-sinyal sumber dari sinyal campuran tanpa memerlukan informasi awal mengenai proses pencampuran. Teknik ini memiliki peran penting dalam berbagai aplikasi pemrosesan audio, seperti pemisahan suara pembicara, pengolahan musik, dan analisis suara lingkungan. Seiring meningkatnya kebutuhan akan sistem audio cerdas yang mampu bekerja pada kondisi nyata, metode pemisahan sinyal yang akurat dan robust menjadi semakin krusial. Salah satu metode yang banyak digunakan dalam BSS adalah *Independent Component Analysis (ICA)*, yang mengekstraksi sinyal sumber berdasarkan asumsi independensi statistik antar sumber. ICA telah terbukti efektif pada pemisahan sinyal audio multikanal karena mampu memanfaatkan informasi spasial dari sinyal stereo [1]. Selain ICA, pendekatan *Non-negative Matrix Factorization (NMF)* juga banyak diterapkan dalam pemrosesan sinyal audio, khususnya pada domain waktu-frekuensi. NMF merepresentasikan sinyal dalam bentuk matriks nonnegatif yang bersifat *sparse* dan mudah diinterpretasikan, sehingga efektif untuk analisis spektral dan sinyal mono [2], [3]. Berbagai penelitian sebelumnya telah mengkaji penggunaan ICA dan metode sejenis dalam pemisahan sinyal audio. Vincent et al. [4] serta Le Roux et al. [5] menunjukkan bahwa ICA mampu memberikan performa yang baik berdasarkan metrik evaluasi standar BSS. Di sisi lain, Stöter et al. [6] mengembangkan pendekatan pemisahan sumber audio berbasis spektral yang banyak digunakan sebagai pembanding dalam penelitian pemisahan sinyal audio. Namun, sebagian besar penelitian tersebut masih berfokus pada evaluasi numerik semata dan jarang mengombinasikannya dengan evaluasi spektral yang berkaitan dengan persepsi pendengaran manusia. Selain itu, perbandingan langsung antara metode berbasis informasi spasial seperti ICA dan metode berbasis spektral seperti NMF pada skenario audio stereo dengan gangguan *noise* masih relatif terbatas. Kondisi ini menunjukkan adanya celah penelitian yang menuntut evaluasi kinerja metode BSS secara lebih komprehensif [6]. Meskipun berbagai metode pemisahan sinyal telah dikembangkan, perbandingan kinerja metode berbasis informasi spasial

seperti ICA dan metode berbasis spektral seperti NMF pada skenario audio stereo dengan gangguan *noise* masih terbatas. Selain itu, evaluasi kinerja sering kali hanya mengandalkan metrik numerik tanpa mempertimbangkan kemiripan spektral sinyal hasil pemisahan. Kondisi ini menunjukkan perlunya evaluasi yang lebih komprehensif untuk menilai efektivitas metode pemisahan sinyal audio secara menyeluruh. Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk mengevaluasi dan membandingkan kinerja ICA dan NMF dalam pemisahan sinyal audio dua sumber menggunakan evaluasi numerik dan spektral, serta menguji ketahanan masing-masing metode terhadap penambahan *noise*.

2. Metode Penelitian

2.1 Data dan Metode Pengumpulan Data

Penelitian ini menggunakan tiga pasangan data audio stereo berdurasi maksimum 30 detik yang terdiri atas kombinasi suara hujan, vokal opera, *beatbox*, tepuk tangan, suara kendaraan, dan dialog. Data diperoleh dari platform Freesound yang menyediakan audio bebas lisensi untuk keperluan penelitian. Seluruh sinyal diproses pada laju sampling 22050 Hz.

Tabel 1. Daftar Pasangan Data Audio

Pasangan	Sumber 1	Sumber 2	Deskripsi
1	rain.wav	opera.wav	Suara hujan dan vokal opera
2	beatbox.wav	applause.wav	Suara beatbox dan tepuk tangan
3	motorbike.wav	dialogue.wav	Suara motor dan suara orang berdialog

Pemilihan tiga pasangan data dilakukan dengan mempertimbangkan variasi karakteristik akustik, seperti sinyal stasioner dan nonstasioner, serta perbedaan pola spektral yang signifikan. Pendekatan ini bertujuan untuk menguji konsistensi kinerja metode pada kondisi suara yang beragam. Oleh karena itu, meskipun jumlah data terbatas, hasil penelitian tetap relevan untuk analisis komparatif antar metode, namun tidak dimaksudkan untuk generalisasi statistik.

2.2 Tahapan Penelitian

Tabel 2. Tahapan Penelitian

No	Tahapan	Deskripsi
1	<i>Pre-processing</i>	Pemotongan durasi, resampling, dan normalisasi RMS terhadap data audio stereo.
2	<i>Mixing</i>	Pencampuran dua sinyal sumber menggunakan proporsi 60:40 dan 40:60 untuk membentuk sinyal stereo
3	Pemisahan Sinyal	Proses pemisahan menggunakan FastICA pada sinyal stereo dan NMF pada sinyal mono hasil rata-rata
4	Evaluasi Kinerja	Pengukuran SDR, SIR, SAR serta kemiripan spektral dengan MFCC Euclidean
5	Visualisasi	Pembuatan grafik SDR, waveform, dan spektrogram untuk analisis visual hasil pemisahan

Tahapan penelitian meliputi *pre-processing* data, *mixing*, pemisahan sinyal menggunakan ICA dan NMF, evaluasi kinerja, serta visualisasi hasil. Ringkasan tahapan penelitian disajikan pada Tabel 2. Pengujian ketahanan metode terhadap gangguan *noise* dilakukan pada tahap evaluasi kinerja.

2.3 Pre-processing Data

Setiap sinyal audio stereo dari pasangan sumber dipotong hingga maksimal 30 detik dan disamakan laju sampling-nya menjadi 22050 Hz. Kemudian dilakukan normalisasi energi dengan metode RMS (Root Mean Square) agar level amplitudo antar sinyal seragam dan tidak mendistorsi saat dicampur.

2.4 Proses Mixing

Dua sinyal sumber dicampur secara linier dengan proporsi berbeda untuk membentuk sinyal stereo:

$$\text{mix}_1 = 0,6 \cdot s_1 + 0,4 \cdot s_2 \quad (1)$$

$$\text{mix}_2 = 0,4 \cdot s_1 + 0,6 \cdot s_2 \quad (2)$$

Sinyal mix_1 dan mix_2 digunakan sebagai input stereo untuk proses pemisahan menggunakan ICA. Sementara itu, untuk metode NMF, kedua sinyal tersebut dirata-ratakan sehingga diperoleh satu sinyal mono sebagai masukan pemrosesan.

2.5 Proses Pemisahan Sinyal

a. Proses ICA

Metode yang digunakan dalam penelitian ini adalah *Fast Independent Component Analysis* (FastICA), yaitu algoritma populer dalam pemisahan sumber buta yang mengasumsikan bahwa sumber-sumber sinyal saling independen secara statistik. FastICA diterapkan langsung pada sinyal stereo mix_1 dan mix_2 dalam format dua dimensi (2-channel input). Hasil dari FastICA adalah dua sinyal terpisah yang diasumsikan sebagai representasi estimasi dari masing-masing sumber asli.

b. Proses NMF

Berbeda dari ICA, pendekatan *Non-negative Matrix Factorization* (NMF) diterapkan pada representasi spektral dari sinyal audio. Oleh karena itu, sinyal stereo hasil pencampuran terlebih dahulu dikonversi menjadi mono dengan merata-ratakan kedua kanal. Kemudian, dilakukan transformasi ke domain frekuensi menggunakan *Short-Time Fourier Transform* (STFT), yaitu teknik yang membagi sinyal menjadi segmen-segmen kecil agar analisis frekuensi dapat dilakukan secara lokal terhadap waktu. Dari transformasi ini diperoleh magnitude spektral, yaitu representasi intensitas tiap frekuensi terhadap waktu

NMF kemudian digunakan untuk mendekomposisi matriks magnitude STFT menjadi dua matriks:

- W (basis spektral): menyatakan pola atau bentuk dasar dari spektrum suara.
- H (aktivasi temporal): menunjukkan kapan dan seberapa kuat setiap pola (basis) tersebut muncul seiring waktu.

Dua hasil estimasi sinyal diperoleh dengan merekonstruksi magnitude hasil dekomposisi dan mengembalikannya ke domain waktu menggunakan *Inverse STFT* (ISTFT).

2.6 Evaluasi Kinerja

a. Evaluasi Kuantitatif (SDR, SIR, dan SAR)

Evaluasi kuantitatif dilakukan menggunakan pustaka `mir_eval` untuk menghitung tiga metrik utama dalam *Blind Source Separation* (BSS), yaitu *Signal-to-Distortion Ratio* (SDR), *Signal-to-Interference Ratio* (SIR), dan *Source-to-Artifacts Ratio* (SAR) [6]. SDR digunakan untuk mengukur tingkat distorsi total antara sinyal hasil pemisahan dan sinyal referensi, SIR mengukur kemampuan metode dalam menekan interferensi dari sumber lain,

sedangkan SAR digunakan untuk menilai tingkat artefak yang dihasilkan selama proses pemisahan sinyal.

Evaluasi kinerja dilakukan pada tiga kondisi pengujian, yaitu kondisi tanpa *noise* (clean), serta dengan penambahan *noise* pada tingkat *Signal-to-Noise Ratio* (SNR) sebesar 10 dB dan 5 dB. Pengujian ini bertujuan untuk mengevaluasi ketahanan (robustness) masing-masing metode pemisahan sinyal terhadap gangguan *noise*.

b. Evaluasi Spektral (MFCC)

Mel-Frequency Cepstral Coefficients (MFCC) diekstraksi dari sinyal hasil pemisahan dan dibandingkan dengan sinyal sumber asli menggunakan jarak Euclidean antar vektor koefisien rata-rata. Semakin kecil jaraknya, semakin mirip karakteristik spektral sinyal hasil terhadap sumber aslinya. Evaluasi ini dilakukan untuk menilai kesesuaian akustik hasil pemisahan yang berkaitan dengan persepsi pendengaran manusia.

c. Visualisasi

Visualisasi digunakan untuk mendukung analisis kualitatif. Tiga bentuk visual yang digunakan dalam penelitian ini adalah:

- Waveform: untuk melihat pola amplitudo terhadap waktu.
- Spektrogram: untuk melihat distribusi energi dalam domain waktu-frekuensi.
- Grafik batang SDR: untuk membandingkan performa antar metode dan kondisi.

d. Analisa Data

Hasil evaluasi numerik (SDR, SIR, SAR) dan spektral (MFCC) dibandingkan secara kuantitatif dan kualitatif untuk menilai keunggulan dan kelemahan masing-masing metode. Analisis difokuskan pada ketahanan terhadap *noise* dan kemiripan spektral hasil pemisahan terhadap sumber asli.

3. Hasil dan Diskusi

3.1. Hasil Evaluasi Kuantitatif (SDR, SIR, SAR)

Tabel 2 menyajikan hasil evaluasi metrik Blind Source Separation (BSS) berupa SDR, SIR, dan SAR terhadap metode ICA dan NMF pada tiga kondisi: tanpa *noise* (clean), dengan *noise* 10 dB, dan 5 dB.

Tabel 3. Hasil Evaluasi SDR, SIR, SAR untuk ICA dan NMF

Data	Kondisi	Metode	SDR(1)	SDR(2)	SIR(1)	SIR(2)	SAR(1)	SAR(2)
rain.wav + opera.wav	Clean	ICA	59.43	45.41	59.47	55.43	80.26	45.86
	Clean	NMF	-25.52	-24.44	-0.65	0.99	-22.16	-21.88
	SNR 10 dB	ICA	3.52	1.51	8.09	29.44	6.00	1.52
	SNR 10 dB	NMF	-25.62	-24.28	-0.97	1.23	-22.09	-21.83
	SNR 5 dB	ICA	0.47	-3.41	3.87	26.96	4.61	-3.40
	SNR 5 dB	NMF	-25.62	-24.11	-1.67	1.70	-21.68	-21.86

beatbox.wav+applause.wav	Clean	ICA	47.26	52.16	50.43	52.69	50.12	61.56
	Clean	NMF	-24.20	-27.24	4.28	2.38	-22.81	-25.26
	SNR 10 dB	ICA	-2.98	1.42	4.59	2.15	-0.84	11.59
	SNR 10 dB	NMF	-24.23	-27.33	4.11	2.43	-22.81	-25.36
	SNR 5 dB	ICA	0.49	-3.06	4.34	24.15	4.16	-3.03
	SNR 5 dB	NMF	-24.28	-27.54	4.31	2.18	-22.90	-25.48
motorbike.wav+dialogue.wav	Clean	ICA	55.06	23.11	55.10	60.03	75.56	23.11
	Clean	NMF	-33.71	4.18	-24.83	40.56	-8.26	4.18
	SNR 10 dB	ICA	2.77	2.56	17.62	11.12	2.30	3.54
	SNR 10 dB	NMF	-33.76	4.39	-25.66	40.58	-7.36	4.39
	SNR 5 dB	ICA	-7.30	-0.99	5.55	0.60	-6.00	6.86
	SNR 5 dB	NMF	-33.92	4.78	-26.74	40.53	-6.24	4.78



Gambar 1. Diagram Batang SDR pada Kondisi Clean, SNR 5 dB dan 10 dB

Gambar 1 menunjukkan bahwa metode ICA secara konsisten menghasilkan nilai SDR yang lebih tinggi dibandingkan NMF, terutama pada kondisi tanpa *noise* (clean). Ketika *noise* ditambahkan pada tingkat SNR 10 dB dan 5 dB, performa kedua metode mengalami penurunan, namun degradasi kinerja pada NMF terlihat lebih signifikan. Hasil ini sejalan dengan karakteristik dasar masing-masing metode. ICA memanfaatkan independensi statistik serta informasi spasial dari sinyal stereo, sehingga kinerjanya relatif lebih stabil terhadap gangguan *noise*. Sebaliknya, NMF bekerja pada sinyal mono dan tidak mengeksplorasi informasi spasial, yang menyebabkan performanya kurang optimal pada skenario pemisahan sinyal stereo [6].

3.2. Hasil Evaluasi Spektral (MFCC)

Evaluasi spektral menggunakan MFCC menunjukkan bahwa ICA menghasilkan sinyal dengan jarak Euclidean yang lebih rendah terhadap sinyal asli dibandingkan NMF pada hampir semua

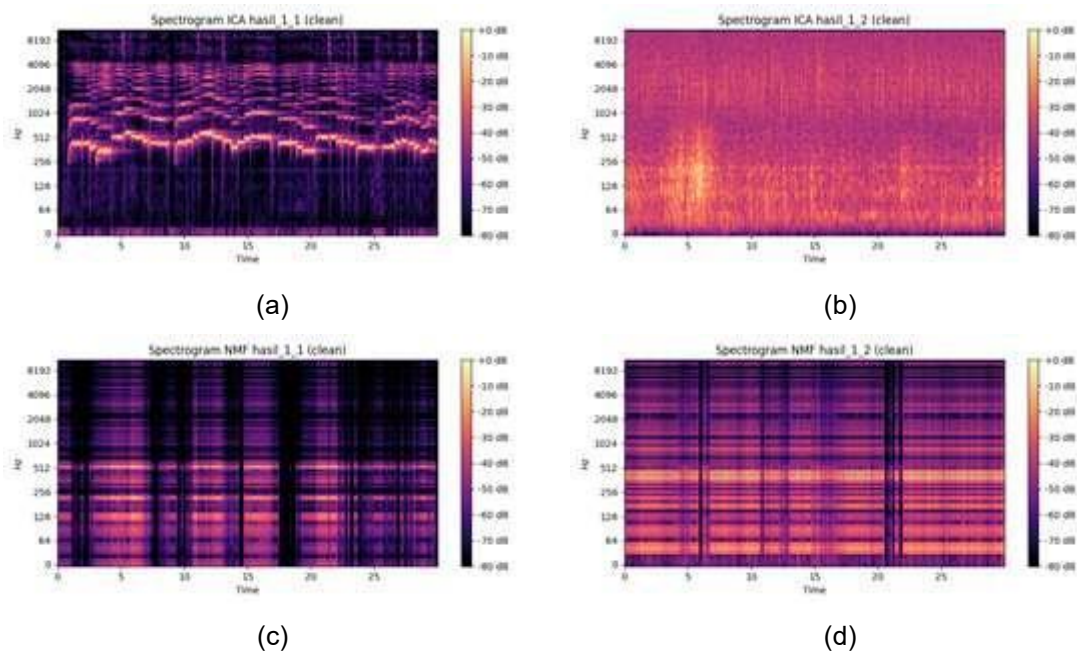
pasangan, terutama pada kondisi clean. Hal ini menunjukkan kemiripan karakteristik spektral yang lebih baik oleh ICA.

Tabel 4. Jarak Euclidean MFCC antara hasil pemisahan dan sumber asli

Data	Kondisi	Metode	Euclidean (hasil 1 ke sumber 1)	Euclidean (hasil 1 ke sumber 2)	Euclidean (hasil 2 ke sumber 1)	Euclidean (hasil 2 ke sumber 2)
rain.wav + opera.wav	Clean	ICA	188.18	63.50	120.91	236.14
	Clean	NMF	-262.36	181.04	161.27	92.59
	SNR 10 dB	ICA	147.33	258.56	126.70	231.81
	SNR 10 dB	NMF	261.06	181.19	155.96	90.54
	SNR 5 dB	ICA	163.85	273.37	164.34	270.47
	SNR 5 dB	NMF	249.91	170.47	155.80	95.12
beatbox.wav+applause.wav	Clean	ICA	52.56	104.24	12.99	61.78
	Clean	NMF	208.02	151.28	154.80	97.88
	SNR 10 dB	ICA	123.63	177.12	166.85	215.70
	SNR 10 dB	NMF	200.03	143.58	148.09	91.38
	SNR 5 dB	ICA	140.33	189.73	162.60	208.54
	SNR 5 dB	NMF	187.95	131.92	141.15	84.35
motorbike.wav+dialogue.wav	Clean	ICA	150.42	150.49	60.63	231.31
	Clean	NMF	351.50	93.67	236.53	124.20
	SNR 10 dB	ICA	142.24	391.94	137.96	397.99
	SNR 10 dB	NMF	289.59	58.00	189.31	138.79
	SNR 5 dB	ICA	127.59	389.44	183.96	440.48
	SNR 5 dB	NMF	158.15	158.07	239.91	64.76

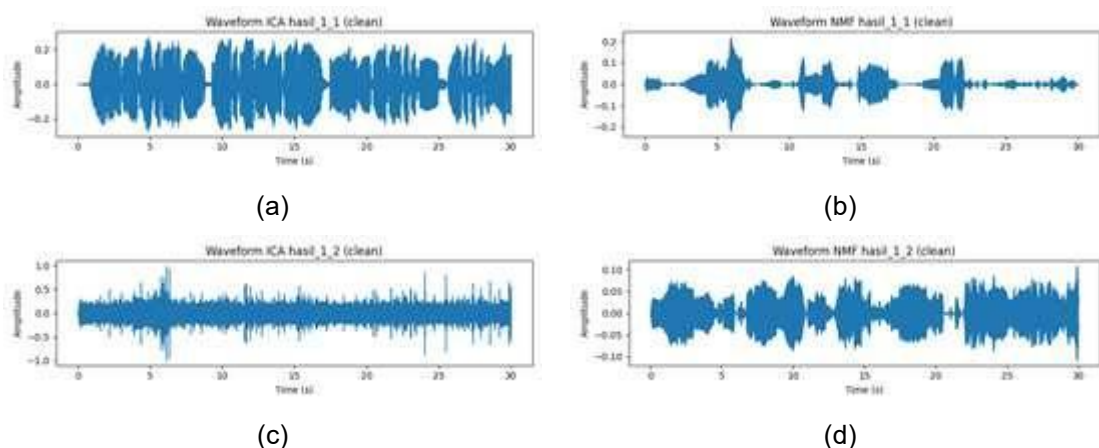
3.3 Hasil Visualisasi Sinyal

Gambar 2 menampilkan contoh spektrogram hasil pemisahan sinyal dari salah satu pasangan data, yaitu rain.wav dan opera.wav. Secara visual, hasil pemisahan menggunakan ICA menunjukkan struktur spektral yang lebih terdefinisi dan menyerupai pola sumber aslinya, sedangkan hasil pemisahan menggunakan NMF cenderung menghasilkan kontur spektral yang lebih menyebar dan kurang terfokus.



Gambar 2. (a) spektrogram hasil pemisahan 1 ICA (b) spektrogram hasil pemisahan 2 ICA (c) spektrogram hasil pemisahan 1 NMF (d) spektrogram hasil pemisahan 2 NMF

Pada spektrogram, sumbu horizontal merepresentasikan waktu, sumbu vertikal menunjukkan frekuensi, sedangkan intensitas warna menggambarkan kekuatan energi pada setiap komponen frekuensi. Warna yang lebih terang mengindikasikan energi yang lebih tinggi. Visualisasi ini digunakan untuk mengamati perbedaan distribusi energi dan pola spektral hasil pemisahan sinyal.



Gambar 3. (a) Waveform Hasil 1 Pemisahan ICA (b) Waveform Hasil 2 Pemisahan ICA (c) Waveform Hasil 1 Pemisahan NMF (d) Waveform Hasil 2 Pemisahan NMF

Gambar 3 memperlihatkan waveform hasil pemisahan sinyal untuk metode ICA dan NMF pada pasangan data yang sama. Waveform hasil ICA menunjukkan pola amplitudo yang lebih terstruktur dan stabil dibandingkan dengan NMF. Pola amplitudo yang lebih jelas mengindikasikan hasil pemisahan yang lebih baik.

4. Kesimpulan

Penelitian ini menyimpulkan bahwa *Independent Component Analysis* (ICA) memiliki kinerja yang lebih unggul dibandingkan *Non-negative Matrix Factorization* (NMF) dalam pemisahan sinyal audio stereo. Berdasarkan hasil evaluasi numerik dan spektral, ICA secara konsisten menunjukkan kemampuan pemisahan yang lebih baik serta ketahanan yang lebih tinggi terhadap gangguan *noise*.

Sebaliknya, keterbatasan NMF dalam memanfaatkan informasi spasial menyebabkan performanya kurang optimal pada skenario pemisahan sinyal stereo. Meskipun demikian, NMF tetap relevan untuk pemrosesan sinyal mono dan analisis berbasis spektral. Kontribusi utama penelitian ini terletak pada penerapan pendekatan evaluasi yang komprehensif, yang mengombinasikan metrik Blind Source Separation, analisis spektral berbasis MFCC, serta pengujian ketahanan terhadap *noise*. Hasil penelitian ini diharapkan dapat menjadi dasar untuk pengembangan metode pemisahan sinyal audio yang lebih lanjut.

Daftar Pustaka

- [1] T. SHIRAISHI and T. DOURA, "Blind source separation by multilayer neural network classifiers for spectrogram analysis," *Mechanical Engineering Journal*, vol. 6, no. 6, pp. 18-00527-18-00527, 2019, doi: 10.1299/mej.18-00527.
- [2] A. Tharwat, "Independent component analysis: An introduction," Aug. 2018, *Emerald Group Holdings Ltd*. doi: 10.1016/j.aci.2018.08.006.
- [3] I. Sobieraj, "Environmental Audio Analysis by Non-negative Matrix Factorization," Apr. 2020.
- [4] A. A. M. Alrustim, A. Elbir, and F. Karabiber, "Blind Audio Source Separation Using Independent Component Analysis and Independent Vector Analysis," 2016. [Online]. Available: <http://ijamec.atscience.org>
- [5] A. P. Putra, N. W. Wiantari, N. P. M. N. Dewi, and I. D. M. B. A. Darmawan, "Independent Component Analysis (ICA) dan Sparse Component Analysis (SCA) dalam Pemisahan Vokal dan Instrumen pada Seni Geguntangan," *Jurnal Elektronik Ilmu Komputer Udayana*, vol. 8, Aug. 2019.
- [6] F.-R. Stöter, S. Uhlich, A. Liutkus, and Y. Mitsufuji, "Open-Unmix - A Reference Implementation for Music Source Separation," *J. Open Source Softw.*, vol. 4, no. 41, p. 1667, Sep. 2019, doi: 10.21105/joss.01667.