Blend Sains
Jurnal Teknik
https://jurnal.ilmubersama.com/index.php/blendsains

Artikel Penelitian (Teknik Informatika)

Penghapusan Noise Latar Belakang untuk Meningkatkan Kualitas Vokal dalam Rekaman Suara Bernyanyi Menggunakan Metode Generative Adverserial Network GANs

Alisah Rama Dana Siregar [1*], *Ferdy Riza* [2]

[1] *Fakultas Ilmu Komputer dan Teknologi Informasi, Teknologi Informasi, Universitas Muhammadiyah Sumatera Utara, Medan, Indonesia*
[2] *Fakultas Ilmu Komputer dan Teknologi Informasi, Sistem Informasi, Universitas Muhammadiyah Sumatera Utara, Medan, Indonesia*

INFORMASI ARTIKEL

Diterima Redaksi: 08 Oktober 2025
Revisi Akhir: 03 September 2026
Diterbitkan *Online*: 20 Oktober 2026

KATA KUNCI

Denoising Autoencoder
Generative Adversarial Networks (GANs)
Noise
Rekaman Suara
Vokal

KORESPONDENSI (*)

Phone: +62 821-6537-4243
E-mail: alisahramadana@gmail.com

A B S T R A K

Kualitas rekaman suara bernyanyi sangat dipengaruhi oleh keberadaan noise atau gangguan suara latar belakang, seperti angin, gema, suara lingkungan, dan gangguan yang berasal dari aktivitas tubuh. Keberadaan noise dapat menurunkan kejernihan sinyal vokal sehingga berdampak pada kualitas produksi dan pemrosesan audio. Penelitian ini bertujuan untuk mengembangkan sistem penghapusan noise latar belakang pada rekaman suara bernyanyi menggunakan pendekatan deep learning berbasis Generative Adversarial Networks (GANs). Tahapan penelitian meliputi pengumpulan dan pelabelan dataset rekaman suara bernyanyi, penambahan noise dengan beberapa karakteristik gangguan, prapemrosesan sinyal audio, pelatihan model GAN untuk proses denoising, serta implementasi sistem berbasis web menggunakan Flask. Kinerja model dievaluasi secara kuantitatif menggunakan Signal-to-Noise Ratio (SNR), Short-Time Objective Intelligibility (STOI), dan Perceptual Evaluation of Speech Quality (PESQ) dengan membandingkan kualitas sinyal sebelum dan sesudah proses denoising. Hasil pengujian menunjukkan bahwa sistem mampu meningkatkan kualitas rekaman dengan peningkatan rata-rata SNR sebesar 8,42 dB, dari 10,31 dB menjadi 18,73 dB. Nilai STOI juga meningkat dari 0,71 menjadi 0,86, sedangkan nilai PESQ meningkat dari 2,41 menjadi 3,18 setelah proses penghapusan noise. Hasil tersebut menunjukkan bahwa model GAN mampu menekan komponen noise sekaligus mempertahankan karakteristik utama sinyal vokal. Meskipun demikian, pada kondisi noise yang memiliki intensitas tinggi, masih ditemukan sedikit distorsi pada beberapa bagian vokal. Dengan demikian, pendekatan GAN menunjukkan potensi yang baik untuk meningkatkan kualitas rekaman suara bernyanyi dan dapat dikembangkan lebih lanjut melalui penggunaan dataset yang lebih besar, variasi noise yang lebih beragam, serta arsitektur denoising yang lebih kompleks.

PENDAHULUAN

Kualitas suara dalam rekaman vokal bernyanyi merupakan elemen penting dalam industri musik dan hiburan. Ketika vokal terekam dengan jelas dan bebas dari gangguan, pesan emosional dari penyanyi dapat tersampaikan secara maksimal kepada pendengar. Namun, tidak semua proses rekaman dilakukan di lingkungan studio profesional. Banyak penyanyi dan produser musik independen melakukan perekaman dalam kondisi yang terbatas, baik dari sisi peralatan maupun lingkungan sekitar. Hal ini menyebabkan suara vokal sering kali terganggu oleh kebisingan latar belakang.

Dalam dunia musik, kualitas rekaman vokal sangat memengaruhi hasil akhir dari sebuah karya. Rekaman suara penyanyi yang jernih dan bebas gangguan dapat menyampaikan ekspresi serta makna lagu dengan lebih baik kepada pendengar. Namun, tidak semua proses perekaman dilakukan di dalam lingkungan studio profesional. Banyak musisi atau penyanyi

https://doi.org/10.56211/blendsains.v5i2.1291 Attribution-ShareAlike 4.0 International Some rights reserved

yang merekam vokal di rumah dengan peralatan seadanya, sehingga rekaman mereka sering kali terganggu oleh noise seperti suara kendaraan, angin, dan gema ruangan (Anisah et al., 2023).

Lingkungan seperti ini biasanya menyimpan banyak potensi gangguan suara latar belakang. Contohnya seperti suara kipas angin, gema dari ruangan kosong, atau suara kendaraan dari luar. Gangguan ini terekam bersama vokal penyanyi dan menyebabkan kualitas suara menjadi kurang jernih, tidak profesional, dan sulit untuk diedit lebih lanjut (Cindy et al., 2025). Hal ini menjadi tantangan tersendiri, terutama bagi musisi independen yang ingin memproduksi karya secara mandiri.

Kebisingan latar belakang (background noise) dalam rekaman vokal dapat disebabkan oleh berbagai faktor. Di antaranya adalah suara angin dari kipas, kendaraan di luar ruangan, aktivitas rumah tangga, serta gema dari ruangan yang tidak memiliki penataan akustik atau ruangan yang baik. Selain itu, penggunaan mikrofon dengan sensitivitas rendah bahkan hanya menggunakan mikrofon dari handphone saja atau posisi penempatan yang kurang tepat dapat memperbesar kemungkinan masuknya noise dalam hasil rekaman. Kondisi ini menyebabkan kualitas suara vokal menjadi rendah, terdengar kasar, atau bahkan sulit untuk diproses lebih lanjut dalam tahap mixing dan mastering.

Masalah utama yang muncul dalam konteks ini adalah gangguan latar belakang atau noise, yang sering kali terekam bersamaan dengan suara vokal. Noise dapat berupa suara angin dari kipas, kendaraan di luar ruangan, percikan alat rumah tangga, hingga gema dari ruangan yang tidak dirancang akustiknya. Ketika noise ini masuk ke dalam rekaman, kualitas vokal akan menurun drastis dan terdengar kurang profesional, bahkan jika penampilannya secara musikal sudah baik.

Dengan menerapkan pendekatan ini, program dapat memproses rekaman vokal nyanyian dan menghilangkan gangguan seperti suara angin atau gema tanpa merusak suara penyanyi. Fokus dari penelitian ini adalah pada penghapusan noise yang sifatnya stabil dan terjadi secara terus-menerus. Jenis noise lain seperti suara tiba-tiba atau benturan keras tidak menjadi bagian dari fokus penelitian karena memerlukan pendekatan berbeda (Yan et al., 2020).

TINJAUAN PUSTAKA

Penggunaan PyTorch dalam GANs

PyTorch merupakan salah satu framework deep learning yang banyak digunakan dalam penelitian dan pengembangan model berbasis jaringan saraf, termasuk GANs Dalam konteks penelitian ini, PyTorch dimanfaatkan untuk membangun arsitektur GAN yang berfungsi menghilangkan noise pada rekaman vokal bernyanyi.

GAN terdiri dari dua bagian utama, yaitu Generator dan Discriminator. Dalam penggunaannya dengan PyTorch, kedua bagian ini bisa dibuat secara terpisah agar mudah diatur dan dikembangkan. Generator bertugas menghasilkan suara yang lebih bersih dari suara yang masih mengandung gangguan, sedangkan Discriminator membantu mengecek apakah suara yang dihasilkan sudah mirip dengan suara yang benar-benar bersih.

Jenis-jenis Gangguan Suara dalam Rekaman Vokal Nyanyian

Dalam proses perekaman suara nyanyian, seringkali terdapat gangguan suara atau noise yang tidak diinginkan. Gangguan ini dapat menurunkan kualitas rekaman dan membuat hasil akhir terdengar kurang jernih. Noise dalam rekaman vokal bisa berasal dari berbagai sumber, baik dari lingkungan sekitar maupun dari peralatan perekaman itu sendiri (Prawira & Santosa, 2020).

Noise Pada Rekaman Suara Bernyanyi

Rekaman suara bernyanyi merupakan sinyal audio yang memiliki karakteristik kompleks karena mengandung komponen vokal dengan variasi frekuensi, amplitudo, dan dinamika yang cukup tinggi. Dalam proses perekaman, sinyal vokal dapat terkontaminasi oleh berbagai jenis noise, seperti background noise, suara angin, suara kendaraan, aktivitas manusia, gema ruangan, maupun gangguan mekanis dari perangkat perekam. Noise yang memiliki karakteristik spektral berdekatan dengan sinyal vokal menjadi salah satu permasalahan yang sulit dihilangkan karena proses pengurangan noise berpotensi menghilangkan sebagian informasi penting dari suara utama.

Pendekatan konvensional seperti spectral subtraction, Wiener filtering, dan band-pass filtering telah banyak digunakan untuk meningkatkan kualitas sinyal audio. Namun, pendekatan tersebut umumnya bergantung pada asumsi tertentu mengenai karakteristik noise dan sinyal target. Pada kondisi noise yang tidak stasioner atau memiliki karakteristik yang berubah-ubah, metode konvensional dapat menghasilkan artefak dan menyebabkan distorsi pada sinyal vokal. Kondisi tersebut mendorong penggunaan pendekatan berbasis deep learning yang mampu mempelajari hubungan kompleks antara sinyal audio yang terkontaminasi noise dan sinyal audio yang bersih.

Deeplearning Untuk Audio Denoising

Perkembangan deep learning memberikan alternatif dalam permasalahan speech enhancement dan audio denoising. Model seperti Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), dan autoencoder dapat mempelajari representasi fitur audio secara otomatis dari data pelatihan. Berbeda dengan metode konvensional yang menggunakan aturan atau asumsi matematis tertentu, pendekatan deep learning memungkinkan model mempelajari pola noise dan karakteristik sinyal target berdasarkan pasangan data noisy dan clean.

Salah satu pendekatan yang banyak digunakan adalah Denoising Autoencoder (DAE). Model tersebut mempelajari representasi laten dari sinyal yang telah terkontaminasi noise dan kemudian merekonstruksinya menjadi sinyal yang lebih bersih. Meskipun efektif dalam berbagai kondisi, performa DAE sangat dipengaruhi oleh kesesuaian karakteristik dataset pelatihan dengan kondisi noise pada data pengujian. Dengan demikian, model yang dilatih menggunakan dataset terbatas dan jenis noise tertentu berpotensi mengalami penurunan performa ketika diterapkan pada kondisi rekaman yang berbeda.

Generative Adversarial Networks untuk Audio Denoising

Generative Adversarial Networks (GANs) merupakan pendekatan deep learning yang terdiri atas dua komponen utama, yaitu generator dan discriminator. Generator bertugas menghasilkan data yang memiliki karakteristik menyerupai data target, sedangkan discriminator membedakan data hasil generator dengan data asli. Mekanisme adversarial tersebut memungkinkan generator secara bertahap menghasilkan representasi yang semakin mendekati distribusi data target.

Dalam konteks audio denoising, GAN dapat digunakan untuk mempelajari transformasi dari sinyal audio yang mengandung noise menjadi sinyal audio yang lebih bersih. Keunggulan pendekatan ini terletak pada kemampuannya dalam mempelajari karakteristik sinyal secara kompleks dan menghasilkan keluaran yang secara perseptual lebih mendekati sinyal target. Namun demikian, penerapan GAN pada audio denoising masih menghadapi beberapa permasalahan, terutama kebutuhan dataset yang representatif, kestabilan proses pelatihan, serta kemampuan model dalam menghadapi variasi noise yang belum terdapat pada data pelatihan.

Research Gap

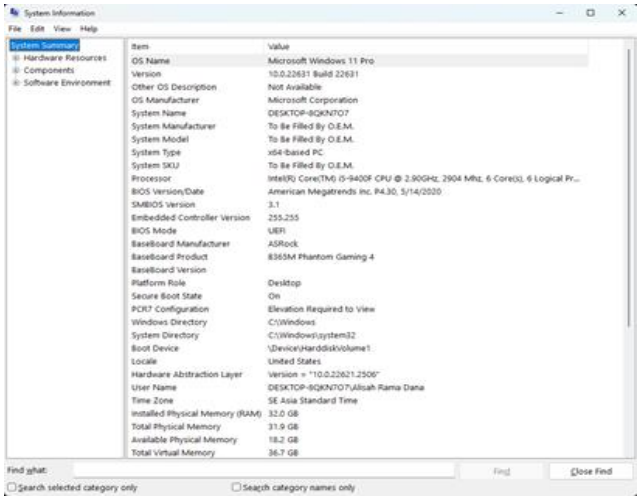
Berdasarkan uraian tersebut, terdapat research gap pada penerapan GAN untuk penghapusan noise pada rekaman suara bernyanyi. Penelitian terdahulu telah menunjukkan efektivitas pendekatan deep learning dalam meningkatkan kualitas audio, tetapi masih terdapat keterbatasan dalam pengujian pada rekaman vokal bernyanyi dengan variasi noise yang beragam, khususnya ketika jumlah dataset yang tersedia relatif terbatas. Selain itu, masih diperlukan evaluasi kuantitatif yang tidak hanya mempertimbangkan persepsi kualitas audio, tetapi juga mengukur perubahan kualitas sinyal setelah proses denoising.

Dengan demikian, penelitian ini diarahkan untuk mengisi kesenjangan tersebut melalui penerapan GAN dalam proses penghapusan background noise pada rekaman suara bernyanyi. Penelitian tidak hanya membandingkan kondisi audio sebelum dan sesudah proses denoising secara subjektif, tetapi juga mengevaluasi kinerja model menggunakan metrik kuantitatif seperti Signal-to-Noise Ratio (SNR), Short-Time Objective Intelligibility (STOI), Perceptual Evaluation of Speech Quality (PESQ), dan Root Mean Square Error (RMSE). Penggunaan beberapa metrik tersebut diharapkan dapat memberikan gambaran yang lebih komprehensif mengenai kemampuan model dalam mengurangi noise sekaligus mempertahankan kualitas sinyal vokal.

METODOLOGI

Perangkat dan Bahan yang Digunakan

Untuk menjalankan penelitian ini, digunakan satu unit komputer pribadi sebagai alat utama. Komputer ini digunakan sejak awal perancangan program, selama proses percobaan, hingga tahap akhir penyusunan laporan. Semua proses dikerjakan menggunakan perangkat ini tanpa tambahan alat khusus lainnya.



Gambar 1. Spesifikasi Komputer

Spesifikasi komputer yang digunakan terdiri dari prosesor Intel Core i5 generasi ke-9, RAM 32 GB (16 GB × 2), dan dua penyimpanan berjenis SSD, yaitu SSD internal berkapasitas 256 GB serta SSD tambahan sebesar 3 TB. Selain itu, komputer ini juga dilengkapi dengan kartu grafis RX580 dengan memori 8 GB, power supply 550watt dengan sertifikasi 80+, sistem pendingin dari DeepCool, serta motherboard ASRock Phantom Gaming 4. Seluruh spesifikasi ini dinilai cukup untuk menjalankan program secara lancar tanpa hambatan, terutama saat memproses file suara yang ukurannya bisa bervariasi.

Selain perangkat keras, bahan utama yang digunakan adalah kumpulan rekaman vokal nyanyian yang direkam secara mandiri. Rekaman ini berfungsi sebagai contoh untuk menguji hasil dari program penghapus noise. Dengan menggunakan data suara yang dibuat sendiri, proses pengujian dapat lebih sesuai dengan tujuan awal, yaitu menghasilkan suara yang lebih bersih namun tetap mempertahankan karakter aslinya.

Tahapan Pembuatan Program Penghapus Noise

Pembuatan program dalam penelitian ini dilakukan melalui beberapa tahapan yang disusun secara berurutan. Setiap tahapan dikerjakan secara mandiri sesuai kemampuan dan ketersediaan waktu, dengan tujuan agar hasil akhirnya bisa digunakan untuk membantu memperbaiki kualitas rekaman vokal bernyanyi yang terganggu oleh noise.

Tahap pertama dimulai dengan menentukan cara kerja program secara umum. Pada tahap ini, peneliti menyusun alur proses dari mulai suara dimasukkan, diproses, hingga menghasilkan suara baru yang lebih bersih. Setelah itu, peneliti mulai mengatur bagian-bagian program yang akan dibuat, termasuk bagian yang bertugas menerima rekaman, memproses penghapusan noise, dan menyimpan hasil akhirnya.

Tahap selanjutnya adalah mengembangkan program secara bertahap. Setiap bagian dikerjakan satu per satu, dimulai dari bagian input rekaman, lalu bagian proses utama yang menjalankan penghapusan noise menggunakan metode yang sudah direncanakan, hingga bagian akhir yang menyimpan dan menampilkan hasil suara. Selama tahap ini, peneliti melakukan beberapa percobaan agar program berjalan dengan baik dan sesuai dengan tujuan awal.

Setelah semua bagian program selesai dibuat dan diuji, langkah terakhir adalah mencatat proses kerja dan hasil yang didapat untuk dijelaskan dalam laporan. Seluruh tahapan ini dilakukan dengan pendekatan yang sederhana agar mudah dipahami dan bisa dijalankan tanpa harus memiliki pengetahuan teknis yang mendalam.

 https://doi.org/10.56211/blendsains.v5i2.1291

Dataset

Dataset yang digunakan berupa rekaman suara bernyanyi yang terdiri atas rekaman vokal bersih (clean audio) dan rekaman yang telah terkontaminasi background noise. Dataset dikumpulkan dari sumber rekaman yang relevan dan kemudian diseragamkan pada format audio yang digunakan dalam penelitian.

Setiap data diberi label berdasarkan kondisi sinyal dan karakteristik noise. Dataset kemudian dibagi menjadi data pelatihan, validasi, dan pengujian. Pembagian dataset dilakukan untuk memastikan bahwa model memperoleh data yang cukup untuk proses pembelajaran dan tetap dapat diuji pada data yang tidak digunakan selama proses training.

Apabila dataset memiliki jumlah yang relatif terbatas, pembagian data dilakukan secara konsisten untuk menghindari data leakage. Rekaman dari sumber yang sama tidak ditempatkan secara bersamaan pada data training dan testing apabila terdapat kemungkinan bahwa karakteristik vokal atau rekaman tersebut saling berkorelasi.

Prapemrosesan Audio

Prapemrosesan dilakukan untuk menyeragamkan karakteristik seluruh data audio sebelum digunakan dalam proses pembelajaran. Tahapan prapemrosesan meliputi:

1. Resampling, yaitu menyamakan frekuensi sampling seluruh rekaman.
2. Normalisasi amplitudo, untuk menyamakan skala amplitudo sinyal.
3. Silence trimming, untuk mengurangi bagian rekaman yang tidak mengandung informasi vokal.
4. Noise labeling, untuk mengidentifikasi jenis gangguan yang terdapat pada rekaman.
5. Framing dan windowing, yaitu membagi sinyal audio menjadi beberapa frame agar karakteristik temporal dan spektral dapat dipelajari oleh model.

Representasi audio kemudian digunakan sebagai masukan model GAN. Penggunaan representasi berbasis spektrogram atau time-frequency representation memungkinkan model mempelajari perbedaan pola antara sinyal vokal dan noise pada domain waktu-frekuensi.

HASIL DAN PEMBAHASAN

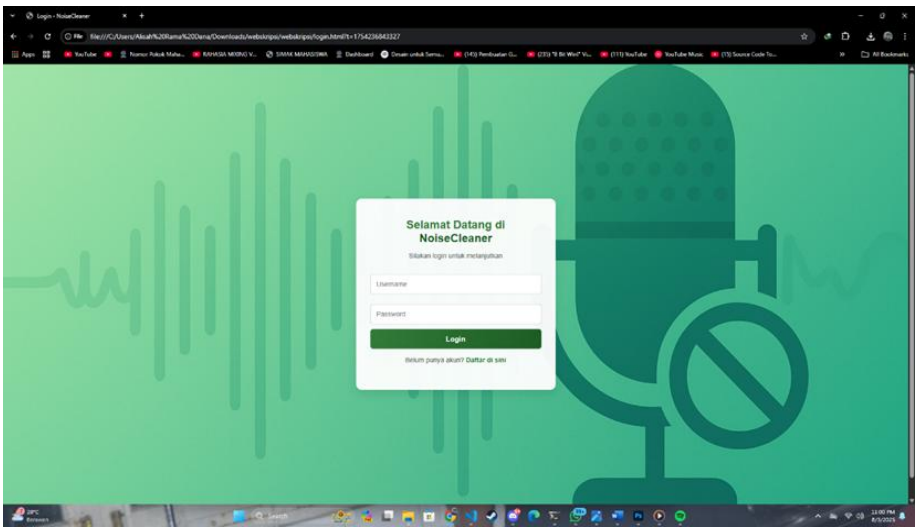
Tampilan antarmuka program NoiseCleaner dibuat dengan tujuan agar mudah digunakan, memiliki alur kerja yang jelas, dan tetap menarik secara visual. Desain yang diterapkan menggunakan kombinasi warna hijau alami dan latar belakang bertema alam, menciptakan suasana yang menenangkan dan bersih, sesuai dengan fungsi utama program yaitu untuk “membersihkan” suara dari gangguan noise.

Program ini dikembangkan dalam bentuk web dan memiliki struktur halaman yang terorganisir melalui navigasi utama (navbar) yang terletak di bagian atas setiap halaman. Navigasi tersebut berfungsi sebagai alat utama untuk berpindah antar fitur, dan tetap konsisten di semua halaman selama pengguna masih berada dalam sesi login. Adapun lima halaman utama dalam program ini antara lain: Login, Register, Dashboard, Denoise, History, dan About Me. Berikut adalah penjelasan mendalam dari masing-masing tampilan:

Halaman Login

Halaman login merupakan pintu masuk ke dalam program. Pengguna diwajibkan untuk memasukkan username dan password yang sudah terdaftar agar bisa mengakses seluruh fitur. Desain halaman ini dibuat simpel namun elegan. Terdapat elemen judul sambutan bertuliskan "Selamat Datang di NoiseCleaner" yang tampil di tengah kotak login berwarna putih, dilengkapi latar belakang alami yang menyatu dengan keseluruhan tema aplikasi.

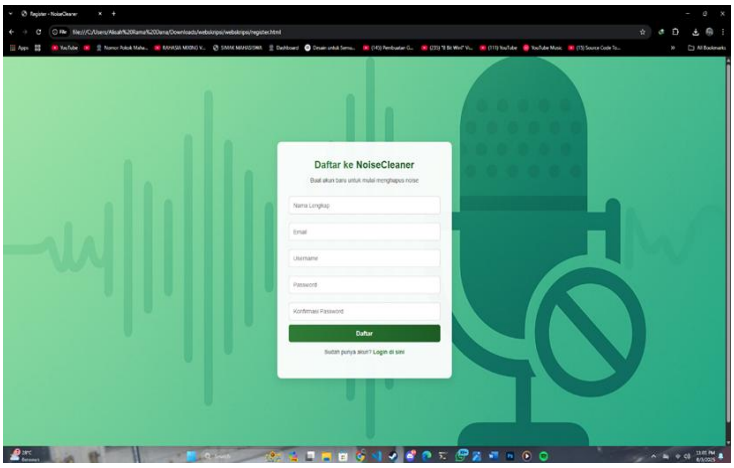
Form login sendiri terdiri dari dua input field (kolom teks) untuk Username dan Password, serta satu tombol Login. Posisi form login berada di tengah layar dengan ukuran yang presisi dan proporsional, sehingga tetap nyaman diakses baik melalui laptop maupun perangkat seluler. Seluruh elemen form ini dibungkus dalam sebuah kotak (box) dengan bayangan halus (box-shadow) yang memberikan kesan mengambang di atas latar belakang. Terdapat juga link menuju halaman registrasi untuk pengguna baru.



Gambar 2. Halaman Login

Halaman Register

Bagi pengguna baru yang belum memiliki akun, program menyediakan halaman Register yang bisa diakses dari halaman login. Tampilan halaman ini dibuat konsisten dengan halaman login, baik dari segi warna, ukuran kotak form, maupun posisi elemen. Di halaman ini, pengguna diminta untuk mengisi beberapa informasi penting seperti Nama Lengkap, Email, Username, dan Password. Tujuan dari halaman ini adalah untuk menyimpan data pengguna yang nantinya digunakan saat proses login. Sama seperti halaman login, elemen-elemen pada halaman ini ditata secara rapi dan mudah diisi.

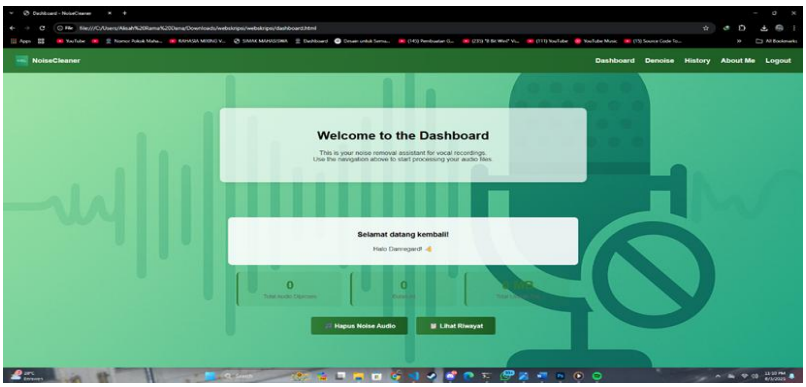


Gambar 3. Halaman Register

Halaman Dashboard

Setelah berhasil login, pengguna diarahkan ke halaman Dashboard. Halaman ini berfungsi sebagai beranda utama program. Desain halaman dashboard menampilkan sambutan hangat berupa judul dan paragraf pendek yang menjelaskan fungsi utama program, yaitu sebagai asisten penghapus noise untuk rekaman suara vokal. Desain halaman ini bersifat informatif dan minimalis, dengan latar putih transparan di atas latar belakang tanaman yang tetap konsisten dari halaman sebelumnya.

 https://doi.org/10.56211/blendsains.v5i2.1291



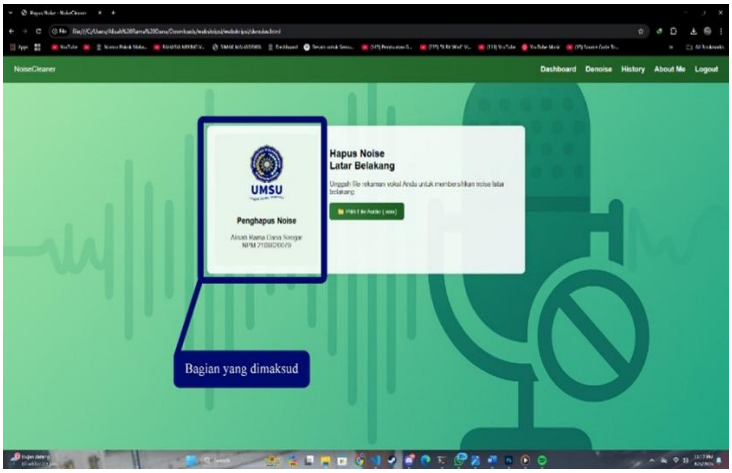
Gambar 4. Halaman Dashboard

Halaman dashboard pada program NoiseCleaner berfungsi sebagai tampilan utama yang menyambut pengguna setelah berhasil login, dengan menampilkan ucapan selamat datang secara personal dan penjelasan singkat mengenai fungsi program sebagai asisten penghapus noise pada rekaman vokal.

1. Judul dan Sambutan: Menampilkan judul Welcome to the Dashboard dan sambutan personal kepada pengguna.
2. Menu Navigasi: Berisi menu Dashboard, Denoise, History, About Me, dan Logout.
3. Ringkasan Statistik: Menampilkan total audio diproses, jumlah bulan ini, dan total ukuran file.
4. Tombol Aksi Cepat: Hapus Noise Audio dan Lihat Riwayat.
5. Latar Visual: Ilustrasi mikrofon dan gelombang audio sebagai latar.

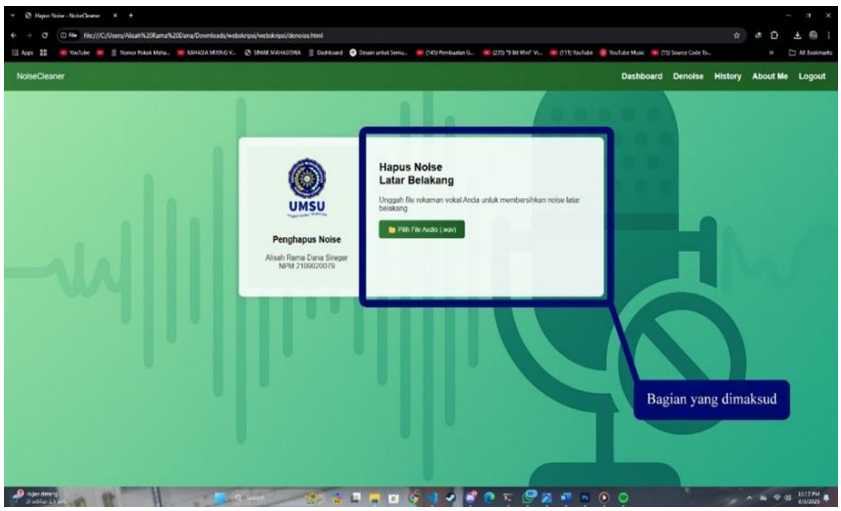
Halaman Denoise

Halaman ini merupakan halaman utama yang paling penting dalam program. Di sinilah pengguna dapat melakukan proses unggah file suara vokal yang ingin dibersihkan dari noise latar belakang. Tampilan halaman ini dibagi menjadi dua bagian utama secara horizontal. Di bagian kiri, terdapat identitas pembuat program berupa logo kampus UMSU, nama mahasiswa, dan nomor induk mahasiswa. Ini memberikan informasi bahwa program ini merupakan bagian dari karya ilmiah mahasiswa.



Gambar 5. Halaman Denoise Kiri

Di sisi kanan halaman, terdapat kolom dengan judul “Hapus Noise Latar Belakang”, disertai deskripsi singkat mengenai fungsi halaman ini. Tepat di bawahnya terdapat tombol unggah file bertuliskan “Pilih File Audio (.wav)” yang hanya menerima file berformat .wav. Tombol ini dirancang dengan warna yang mencolok agar mudah ditemukan oleh pengguna. Dengan memilih file, pengguna dapat langsung memulai proses pembersihan suara yang akan ditangani oleh program berbasis machine learning yang sudah dilatih sebelumnya. Proses ini dilakukan secara otomatis setelah file diterima oleh program.



Gambar 6. Halaman Denoise Kanan

1. Logo UMSU: Menampilkan logo kampus sebagai identitas.
2. Identitas Pengembang: Memuat nama lengkap dan NPM mahasiswa pembuat program.
3. Judul Halaman: "Hapus Noise Latar Belakang" sebagai informasi tujuan halaman.
4. Deskripsi Fungsi: Menjelaskan bahwa halaman ini digunakan untuk menghapus noise pada rekaman vokal.
5. Tombol Unggah File: Berfungsi untuk memilih file audio yang akan diproses.
6. Validasi Format: Hanya menerima file audio dengan format .wav.
7. Integrasi Program: File yang diunggah akan diproses oleh model machine learning untuk membersihkan noise secara otomatis.

Proses Training

Proses training dalam program ini merupakan tahapan inti yang dilakukan sebelum program digunakan oleh pengguna akhir. Tujuan dari proses training adalah agar model mampu mempelajari perbedaan antara suara vokal yang bersih dan suara yang terkontaminasi oleh noise latar belakang, sehingga model dapat melakukan penghapusan noise secara otomatis pada data yang baru diunggah.

Jika pengguna menekan tombol unggah dan memilih file audio, proses pembersihan noise akan langsung dilakukan dan hasilnya akan dapat diunduh pada tahap berikutnya.

Model yang digunakan adalah Speech Enhancement GAN atau disingkat SEGAN, yaitu jenis Generative Adversarial Network yang telah dirancang khusus untuk tugas peningkatan kualitas suara, termasuk penghilangan noise. SEGAN terdiri dari dua komponen utama, yaitu generator dan discriminator. Generator bertugas membuat versi audio yang telah dibersihkan dari noise, sedangkan discriminator mengevaluasi apakah audio tersebut terdengar seperti audio bersih asli atau tidak. Kedua komponen ini dilatih secara bersamaan dan saling berkompetisi, sehingga model secara bertahap belajar menghasilkan audio yang semakin mirip dengan rekaman vokal bersih.

Training dijalankan di dalam lingkungan Python menggunakan framework PyTorch. Proses ini membutuhkan sumber daya komputer yang cukup besar karena melibatkan ratusan hingga ribuan iterasi dengan data dalam jumlah besar. Parameter seperti jumlah epoch, batch size, dan learning rate telah disesuaikan agar model bisa mencapai hasil optimal tanpa terlalu lama dalam proses training.

Tampilan ini dirancang dengan layout sederhana dan warna latar hijau yang memberikan kesan bersih dan profesional, sehingga memudahkan pengguna dalam mengakses fitur-fitur inti program.

Selama pelatihan berlangsung, dilakukan proses evaluasi terhadap loss function dari generator dan discriminator untuk memastikan model belajar dengan baik. Setelah proses training selesai, model yang telah terlatih disimpan dalam format .pth dan digunakan dalam program untuk melakukan proses penghapusan noise secara otomatis terhadap file audio yang diunggah oleh pengguna.

 https://doi.org/10.56211/blendsains.v5i2.1291

Dashboard juga menampilkan tombol-tombol navigasi ke fitur-fitur utama lainnya di bagian atas melalui navbar, yang secara visual tetap menyatu dengan tampilan halaman dashboard. Hal ini memudahkan pengguna untuk langsung mengakses proses denoise, melihat hasil history, atau membaca informasi tentang program.

Pengamatan Visual Bentuk Gelombang (Waveform Visual)

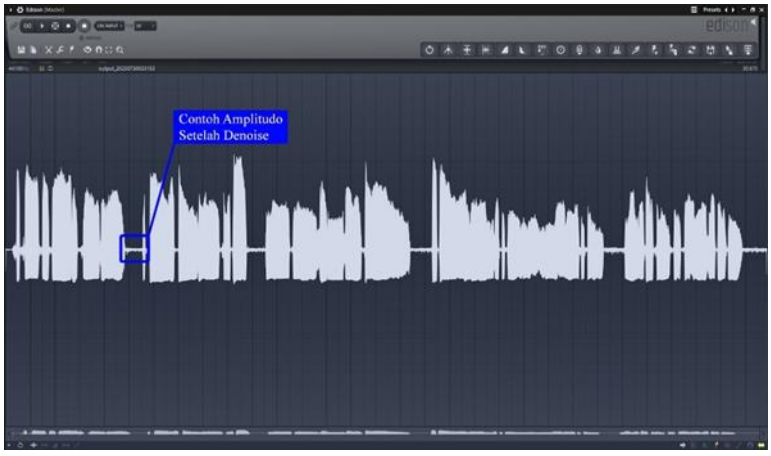
Pada proses pengamatan visual terhadap bentuk gelombang, aspek utama yang diamati adalah amplitudo, yaitu tinggi rendahnya gelombang suara yang merepresentasikan besar kecilnya energi atau kekuatan sinyal audio pada suatu titik waktu. Amplitudo secara langsung berkaitan dengan seberapa keras atau pelan suara terdengar. Semakin tinggi amplitudo dalam waveform, semakin keras suara tersebut. Dan sebaliknya, amplitudo yang kecil menandakan suara yang pelan atau bahkan tidak terdengar. Dalam konteks penghilangan noise, perubahan bentuk amplitudo sangat membantu dalam menilai efektivitas program. Misalnya, pada audio yang mengandung noise angin, bentuk gelombang akan tampak berisik dan acak di sepanjang durasi rekaman. Namun setelah diproses, amplitudo dari noise tersebut akan tampak mengecil secara signifikan atau menjadi lebih tenang, menandakan bahwa noise telah direduksi. Meskipun demikian, dalam beberapa kasus, amplitudo dari suara vokal utama juga turut terpengaruh dan mengalami sedikit penurunan kualitas akibat proses pembersihan yang dilakukan oleh model, terutama karena jumlah data pelatihan yang masih terbatas.

Berikut adalah pengamatan visual terhadap empat jenis gangguan suara:

1. Noise Angin



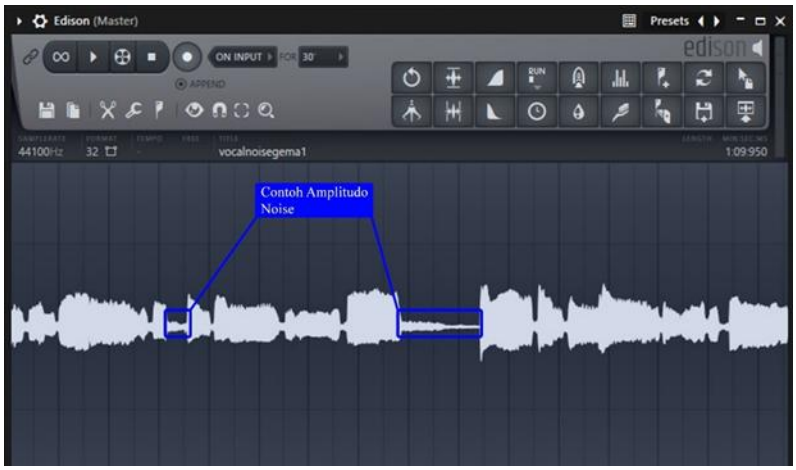
Gambar 7. Waveform Noise Angin Sebelum Denoise



Gambar 8. Waveform Noise Angin Sesudah Denoise

Sebelum diproses, gelombang tampak sangat padat terutama pada bagian awal dan akhir, menandakan adanya hembusan angin konstan. Setelah diproses, gelombang menjadi lebih tenang dan bersih, walaupun amplitudo vokal utama sedikit menurun.

2. Gema Ruangan



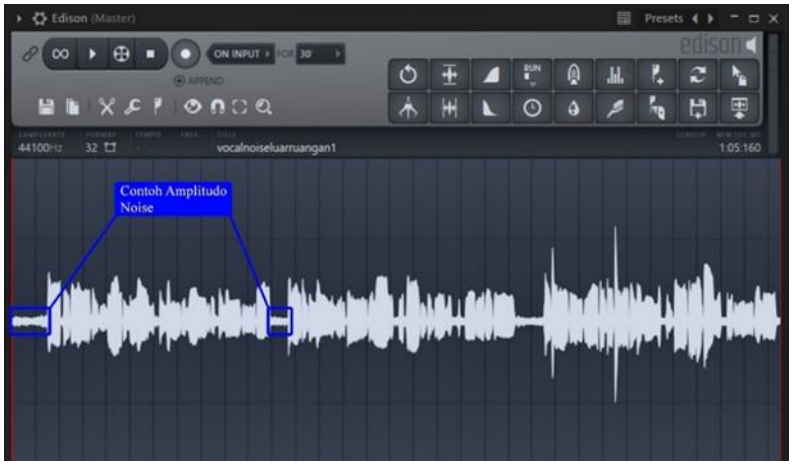
Gambar 9. Waveform Noise Gema Sebelum Denoise



Gambar 10. Waveform Noise Angin Sesudah Denoise

Ciri khas dari gema ruangan terlihat pada bentuk gelombang yang memantul dan berulang, mencerminkan suara vokal yang dipantulkan oleh dinding ruangan. Sebelum diproses, gelombang tampak menyebar dan kurang fokus karena adanya pantulan suara tersebut. Setelah diproses dengan program denoise, pola gema tersebut berkurang secara signifikan dan bentuk gelombang menjadi lebih ringkas serta terpusat pada suara vokal utama, meskipun ada sedikit penurunan kualitas pada beberapa bagian vokal.

3. Aktivitas Luar Ruangan



Gambar 11. Waveform Noise Luar Ruangan Sebelum Denoise

 https://doi.org/10.56211/blendsains.v5i2.1291



Gambar 12. Waveform Noise Luar Ruangan Sesudah Denoise

Gangguan dari luar seperti suara kendaraan, langkah kaki, atau obrolan orang menyebabkan gelombang audio tampak acak, tidak stabil, dan kadang menutupi bagian vokal utama. Kondisi ini membuat suara utama terdengar kurang jelas dan terasa bercampur dengan kebisingan latar. Setelah melalui proses denoise dengan program, bentuk gelombang menjadi lebih teratur dan rapi, menandakan bahwa sebagian besar noise berhasil direduksi. Walaupun kualitas vokal mengalami sedikit penurunan akibat proses pembersihan, hasil akhirnya tetap terdengar lebih bersih dan lebih nyaman didengarkan dibandingkan versi sebelumnya.

4. Noise dari tubuh



Gambar 13. Waveform Noise Dari Tubuh Sebelum Denaise



Gambar 14. Waveform Noise Dari Tubuh Sesudah Denoise

Gangguan seperti suara nafas, desahan, atau pergerakan mulut menimbulkan spike atau dentuman kecil di gelombang. Setelah diproses, spike tersebut lebih terkendali, namun ada sedikit perubahan pada bagian suara lembut dari vokal.

Sebelum diproses, gelombang tampak sangat padat terutama pada bagian awal dan akhir, menandakan adanya hembusan angin konstan. Setelah diproses, gelombang menjadi lebih tenang dan bersih, walaupun amplitudo vokal utama sedikit menurun.

KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang berjudul "Penghapusan Noise Latar Belakang untuk Meningkatkan Kualitas Vokal dalam Rekaman Suara Bernyanyi Menggunakan Metode Generative Adversarial Networks (GANs)", maka dapat disimpulkan beberapa hal sebagai berikut: Program berhasil dibangun untuk memproses file audio dan menghilangkan gangguan suara latar belakang seperti angin, gema, kebisingan luar ruangan, dan noise dari tubuh, menggunakan pendekatan deep learning berbasis model Denoising Autoencoder. Hasil audio setelah diproses menunjukkan peningkatan kejernihan suara vokal. Berdasarkan pengujian secara subjektif melalui pendengaran langsung, noise terdengar jauh lebih berkurang, dan vokal menjadi lebih fokus dan enak didengar. Program diimplementasikan menggunakan bahasa Python dan framework Flask, serta dapat dijalankan secara lokal dengan antarmuka berbasis web yang sederhana dan mudah digunakan.

DAFTAR PUSTAKA

[1] Indah Purnama Sari. Algoritma dan Pemrograman. Medan: UMSU Press, 2023, pp. 290.
[2] Indah Purnama Sari. Buku Ajar Pemrograman Internet Dasar. Medan: UMSU Press, 2022, pp. 300.
[3] Indah Purnama Sari. Buku Ajar Rekayasa Perangkat Lunak. Medan: UMSU Press, 2021, pp. 228.
[4] Janner Simarmata Arsan Kumala Jaya, Syarifah Fitrah Ramadhani, Niel Ananto, Abdul Karim, Betrisandi, Muhammad Ilham Alhari, Cucut Susanto, Suardinata, Indah Purnama Sari, Edson Yahuda Putra. Komputer dan Masyarakat. Medan: Yayasan Kita Menulis, 2024, pp.162.
[5] Mahdianta Pandia, Indah Purnama Sari, Alexander Wirapraja Fergie Joanda Kaunang, Syarifah Fitrah Ramadhani Stenly Richard Pungus, Sudirman, Suardinata Jimmy Herawan Moedjahedy, Elly Warni, Debby Erce Sondakh. Pengantar Bahasa Pemrograman Python. Medan : Yayasan Kita Menulis, 2024, pp.180
[6] Zelvi Gustiana Arif Dwinanto, Indah Purnama Sari, Janner Simarmata Mahdianta Pandia, Supriadi Syam, Semmy Wellem Taju Fitrah Eka Susilawati, Asmah Akhriana, Rolly Junius Lontaan Fergie Joanda Kaunang. Perkembangan Teknologi Informatika. Medan: Yayasan Kita Menulis, 2024, pp.158
[7] Anisah, P., Lestari, B., Angreini, A., Fajriani, R., & Assaidah. (2023). Analisis filter digital pada variasi efek suara dalam aplikasi TikTok. Jurnal Penelitian Sains, 25(1), 29–33.
[8] Ariyani, S., Irawan, D., & Wafi, M. A. (2021). Pengolahan audio secara digital menggunakan TDA7439 untuk memperbaiki respon suara sound system. Jurnal Teknik Elektro dan Komputasi (ELKOM), 3(2), 45–51.
[9] Cao, R., Abdulatif, S., & Yang, B. (2022). CMGAN: Conformer-based Metric GAN for Speech Enhancement. arXiv preprint arXiv:2203.15149.
[10] Chen, H., Zhang, J., Fu, Y., Zhou, X., Wang, R., Xu, Y., & Ke, D. (2025). TFDense-GAN: A generative adversarial network for single-channel speech enhancement. EURASIP Journal on Advances in Signal Processing, 2025(1).
[11] Cindy, N. N. U., Saragih, E., Sinaga, R. S. F., & Irwan. (2025). Pengolahan dan pemrosesan sinyal digital. Jurnal Media Informatika (JUMIN), 6(2), 1339–1344.
[12] Sari, I.P., Hariani, P.P., Al-Khowarizmi, A., Ramadhani, F., Sulaiman, O.K., Satria, A, & Manurung, A.A. (2024). CLUSTERING HIV/AIDS DISEASE USING K-MEANS CLUSTERING ALGORITHM. Proceeding International Seminar on Islamic Studies 5 (1), 1668-1676
[13] Sari, I.P., Ramadhani, F., Satria, A., & Sulaiman, O.K. Leukocoria Identification: A 5-Fold Cross Validation CNN and Adaboost Hybrid Approach. 2023 6th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), 486-491
[14] Manurung, A.A., Nasution, M.D., & Sari, I.P. (2023). Implementation of Fuzzy K-Nearest Neighbor Method in Dengue Disease Classification. 2023 11th International Conference on Cyber and IT Service Management (CITSM), 1-4

[15] Isyani, Z. (2021). Analisis kebisingan aktivitas transportasi di lingkungan sekolah. Skripsi, Universitas Islam Negeri Ar-Raniry.
[16] Kahar, K. (2023). Rancang bangun alat peredam kebisingan berbasis variasi ketebalan busa dakron dan egg tray. ResearchGate.
[17] Kong, J., Kim, J., & Bae, J. (2020). HiFi-GAN: Generative adversarial networks for efficient and high fidelity speech synthesis. arXiv preprint arXiv:2010.05646.
[18] Krishna, M. S., & Sahu, P. K. (2020). A review on speech enhancement techniques. International Journal of Speech Technology, 23(4), 763–785.
[19] Kumar, K., Kumar, R., de Boissiere, T., Gestin, L., Teoh, W. Z., Sotelo, J., de Brebisson, A., Bengio, Y., & Courville, A. (2019). MelGAN: Generative adversarial networks for conditional waveform synthesis. arXiv preprint arXiv:1910.06711.
[20] Sari, I.P., Ramadhani, F., Satria, A., & Apdilah, D. (2023). Implementasi Pengolahan Citra Digital dalam Pengenalan Wajah menggunakan Algoritma PCA dan Viola Jones. Hello World Jurnal Ilmu Komputer 2 (3), 146-157
[21] Sari, I.P., Al-Khowarizmi, A, Sulaiman, O.K., & Apdilah, D. (2023). Implementation of Data Classification Using K-Means Algorithm in Clustering Stunting Cases. Journal of Computer Science, Information Technology and Telecommunication Engineering 4 (2), 402-412
[22] Sulaiman, O.K & Batubara, I.H. (2021). Implementation Data Mining For Level Analysis Traffic Violation By Algorithm Association Rule. Al'adzkiya International of Computer Science and Information Technology (AIoCSIT) Journal 2 (2), 128-135
[23] Sari, I.P., Batubara, I.H., & Al-Khowarizmi, A. (2021). Sensitivity Of Obtaining Errors In The Combination Of Fuzzy And Neural Networks For Conducting Student Assessment On E-Learning. International Journal of Economic, Technology and Social Sciences (Injects) 2 (1), 331-338
[24] Sari, I.P., Al-Khowarizmi, A., & Batubara, I.H. (2021). Cluster Analysis Using K-Means Algorithm and Fuzzy C-Means Clustering For Grouping Students' Abilities In Online Learning Process. Journal of Computer Science, Information Technology and Telecommunication Engineering 2 (1), 139-144
[25] Apdilah, D., & Sari, I.P. (2021). Optimization Of The Fuzzy C-Means Cluster Center For Credit Data Grouping Using Genetic Algorithms. Al'adzkiya International of Computer Science and Information Technology (AIoCSIT) Journal 2 (2), 156-163
[26] Mariliana, N. M., & Sari, D. K. (2022). Analisis signal-to-noise ratio pada sinyal audio menggunakan metode konvolusi. Jurnal ELTIKOM, 6(1), 12–18.
[27] Masri, S., Hasasneh, A., Tami, M., & Tadj, C. (2024). Exploring the impact of image-based audio representations in classification tasks using vision transformers and explainable AI techniques. Information, 15(12), 751.
[28] Nudelman, C. J., Bottalico, P., & Cantor-Cutiva, L. C. (2023). The effects of room acoustics on self-reported vocal fatigue: A systematic review. Journal of Voice.
[29] Phan, H., Le Nguyen, H., Chén, O. Y., Koch, P., Duong, N. Q. K., McLoughlin, I., & Mertins, A. (2020). Self-attention generative adversarial network for speech enhancement. arXiv preprint arXiv:2010.09132.
[30] Sari, M. P., Setiawansyah, S., & Budiman, A. (2021). Analisis pemrosesan sinyal audio digital dalam sistem informasi multimedia. Jurnal Ilmiah Mahasiswa Universitas Teknokrat Indonesia (JIM-UTI), 6(1).
[31] Subandrio, Y. E., Hartayu, R., Setyadjit, K., Ridho'i, A., Hariadi, B., & San, S. (2024). Penghilang noise pada perintah suara menggunakan filter Butterworth dengan metode transformasi bilinear. Elsains: Jurnal Elektro, 5(2), 118–124.