

# Stetoskop Digital Paru dengan CNN dan Knowledge Distillation untuk Deteksi Penyakit Pernapasan

Almira Raisa Izzatina<sup>1</sup>, Ziyen Nadia Putri<sup>2</sup>, Giovrey Ernesto Putra<sup>3</sup>, Putri Alief Siswanto<sup>4</sup>

<sup>1</sup>Fakultas Teknologi Elektro dan Informatika Cerdas, Institut Teknologi Sepuluh Nopember, Surabaya, 60298, email: 5025221250@student.its.ac.id

<sup>2</sup>Fakultas Kedokteran dan Kesehatan, Institut Teknologi Sepuluh Nopember, Surabaya, 61482, email: 5049221006@student.its.ac.id

<sup>3</sup>Fakultas Kedokteran dan Kesehatan, Institut Teknologi Sepuluh Nopember, Surabaya, 60177, email: 5049221026@student.its.ac.id

<sup>4</sup>Fakultas Kedokteran dan Kesehatan, Institut Teknologi Sepuluh Nopember, Surabaya, 60296, email: putri.alief@its.ac.id

Corresponding Author: Almira Raisa Izzatina

**INTISARI** Beragamnya pola penyakit pernapasan dan keterbatasan akses layanan kesehatan menuntut hadirnya alat skrining yang cepat, ekonomis, dan mudah digunakan di lapangan. Penelitian ini memperkenalkan prototipe stetoskop digital berbasis edge computing yang dikembangkan untuk mendeteksi kelainan pernapasan secara lokal tanpa bergantung pada koneksi internet, sehingga dapat dioperasikan pada fasilitas primer atau wilayah dengan infrastruktur terbatas. Tujuan penelitian ini adalah merancang dan mengimplementasikan prototipe stetoskop digital berbasis Convolutional Neural Network (CNN) untuk klasifikasi kelainan pernapasan, serta menerapkan teknik knowledge distillation untuk menjadi solusi dalam mengecilkan kapasitas model besar yang masih menjadi permasalahan pada implementasi di perangkat *edge* saat ini. Metodologi melibatkan pemanfaatan dataset ICBHI 2017 Respiratory Sound Database dan Mendeley Pulmonary (Lungs) Sound Database, langkah pra-pemrosesan sinyal, serta konversi sinyal menjadi mel-spectrogram sebagai masukan ke model konvolusional. Model teacher dilatih penuh untuk mencapai performa optimal kemudian pengetahuannya ditransfer ke model student yang lebih ringan melalui distillation dengan kombinasi loss terawasi dan soft-target. Evaluasi eksperimen dilakukan menggunakan metrik akurasi, sensitivitas, spesifisitas, dan F1-score untuk menilai kestabilan hasil. Hasil menunjukkan bahwa pendekatan distillation berhasil menghasilkan model student yang kinerjanya mendekati teacher namun memiliki ukuran dan parameter yang jauh lebih kecil setelah dilakukan optimasi. Model terkompresi ini mampu dijalankan pada perangkat microcontroller dengan waktu respons rendah, sehingga praktis untuk keperluan skrining di lapangan. Secara keseluruhan, integrasi desain perangkat keras dan lunak serta pemanfaatan knowledge distillation menawarkan solusi efisien dan terjangkau untuk memperluas akses deteksi penyakit pernapasan di lingkungan dengan sumber daya terbatas.

**KATA KUNCI** Stetoskop Digital, *Edge Computing*, *Convolutional Neural Network*, *Knowledge Distillation*, Klasifikasi Suara Paru, Optimasi Model

## I. PENDAHULUAN

Penyakit pernapasan seperti pneumonia, bronkitis, dan penyakit paru obstruktif kronis (PPOK) masih menjadi penyebab utama morbiditas dan mortalitas secara global. Pada tahun 2021, lebih dari 468 juta orang di seluruh dunia hidup dengan penyakit pernapasan kronis, yang menyebabkan lebih dari 4,4 juta kematian setiap tahunnya [1]. Kondisi ini semakin menantang di Indonesia, terutama di wilayah 3T (Terdepan, Terpencil, dan Tertinggal) yang mengalami kekurangan tenaga medis dan keterbatasan fasilitas kesehatan [2]. Deteksi dini penyakit ini terbukti menurunkan mortalitas dan memperbaiki prognosis pasien, namun metode auskultasi paru konvensional yang mengandalkan stetoskop analog rentan terhadap variabilitas interpretasi karena bergantung pada pengalaman dokter yang mendengar [3]. Tanpa adanya fitur perekaman, suara paru tidak dapat dianalisis ulang sehingga mengurangi peluang untuk mendapatkan diagnosis yang lebih akurat [4].

Sejak penemuan stetoskop oleh Laennec, evolusi stetoskop digital berkembang pesat. Solusi *smartphone-based* pertama memungkinkan perekaman dan pengiriman suara ke cloud untuk inferensi AI [5]. Prototipe STETHAID bahkan mengintegrasikan stetoskop digital dengan aplikasi iOS untuk auskultasi berbantuan AI [6]. Namun, ketergantungan pada konektivitas internet membatasi penggunaan di daerah dengan sumber daya terbatas.

Dalam ranah Machine Learning (ML), beberapa pendekatan telah dicoba untuk mengklasifikasikan suara paru. Metode klasik seperti Support Vector Machine (SVM) menunjukkan akurasi sebesar 65,5% untuk deteksi *wheeze* dan *crackle*, namun performanya sensitif terhadap noise dan sangat bergantung pada

rekayasa fitur manual [7]. Sementara Random Forest dan ensemble lainnya juga telah digunakan, tetapi menunjukkan performa yang bervariasi tergantung kualitas data dan tuning parameter [8][9]. Sebaliknya, Convolutional Neural Network (CNN) menunjukkan performa yang jauh lebih tinggi. Misalnya, CNN berbasis mel-spectrogram mencapai akurasi 95,56% dalam klasifikasi tujuh jenis suara paru (normal, *coarse crackle*, *fine crackle*, *polyphonic wheeze*, *monophonic wheeze*, *squawk*, dan *stridor*), bahkan dalam kondisi akustik yang beragam [3]. Studi oleh Khalifa et al. (2023) juga melaporkan akurasi 98,6% dan F1-score 98,4% untuk klasifikasi tiga kelas utama: normal, *wheeze*, dan *crackle* [10]. Recurrent Neural Network (RNN) dan Long Short-Term Memory (LSTM) juga dipakai untuk menangkap konteks temporal, dengan sensitivitas tinggi dalam deteksi PPOK. Model hybrid CNN–Random Forest pun berhasil dijalankan secara on-device pada Raspberry Pi 4 Model B. Penelitian oleh Shuvo et al. (2021) menunjukkan bahwa kombinasi CNN dan Best Discrepancy Forest (BDF) dapat mencapai akurasi 99,62% dan F1-score 98,56%, dengan waktu inferensi sekitar 9 detik pada sistem tertanam tanpa koneksi internet [11].

Meskipun akurasinya unggul, parameter pada CNN dengan akurasi tinggi memerlukan daya komputasi yang besar, sehingga sulit dijalankan pada komputer berdaya rendah. Selain itu, model seringkali mengalami penurunan akurasi ketika diterapkan di luar kondisi laboratorium akibat fenomena domain shift atau perbedaan antara data training dan kondisi nyata di lapangan [12][13]. Untuk menjawab tantangan tersebut, pendekatan Knowledge Distillation (KD) dikembangkan

sebagai solusi. Metode ini memungkinkan model besar (teacher) mentransfer "pengetahuan" ke model kecil (student), dengan cara membuat student meniru prediksi teacher secara lebih halus dan terarah seperti ditunjukkan di Gambar 3 [14]. K. Choi *et al.* berhasil memangkas ukuran model CNN lebih dari 80%, menghasilkan model student berukuran  $\leq 5$  MB dengan penurunan akurasi kurang dari 1% serta latensi inferensi di bawah 200 ms [15]. Penelitian Zhong, Tao *et al.* juga menunjukkan bahwa KD meningkatkan generalisasi model di berbagai kondisi akuisisi dan mengurangi dampak domain shift [16].



Gambar 1. Prototipe alat

Berdasarkan tinjauan tersebut, penelitian ini bertujuan merancang Stetoskop Digital Auskultasi Paru Berbasis CNN dan Knowledge Distillation yang sepenuhnya on-edge dan low-cost, khususnya untuk wilayah 3T, sebagaimana ditunjukkan pada Gambar 1. Prototipe pada Gambar 1 terdiri dari stetoskop digital dan case kotak yang berisikan sistem pengolahan dan dilengkapi LCD OLED untuk menampilkan hasil pengukuran. Perangkat merekam suara napas selama 5–10 detik pada titik auskultasi anterior dan posterior paru menggunakan stetoskop [5], kemudian melakukan pra-pemrosesan dengan mengubah audio menjadi representasi visual, mel-spectrogram. Data yang dihasilkan akan diklasifikasikan menggunakan model student CNN hasil knowledge distillation berukuran  $\leq 5$  MB, tanpa mengorbankan akurasi diagnostik [15]. Seluruh pipeline akan dijalankan secara lokal pada Raspberry Pi, mengeliminasi kebutuhan koneksi Internet dan mendukung penggunaan di area minim infrastruktur [11].

Keterbaruan penelitian ini terletak pada penerapan KD secara eksklusif pada sinyal auskultasi paru untuk menghasilkan model CNN yang ringkas namun tetap memiliki kemampuan diagnostik tinggi [12]. Selain itu, penelitian ini mengembangkan hardware-software co-design pada platform low-cost yang mengimplementasikan pipeline machine learning on-device, sehingga memberikan solusi praktis yang memperluas akses diagnosis penyakit pernapasan di lapangan tanpa ketergantungan pada komputasi berskala besar maupun jaringan Internet. Pendekatan ini dapat memberikan diagnosis yang lebih cepat dan objektif, yang berguna khususnya di wilayah 3T yang memiliki keterbatasan tenaga spesialis

## II. METODE

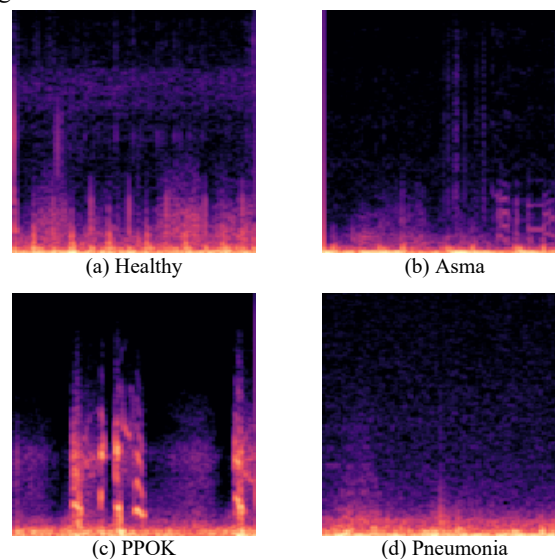
Secara keseluruhan, sistem yang dikembangkan meliputi diagnosis suara paru berbasis Artificial Intelligence (AI), pra-pemrosesan, ekstraksi fitur, representasi mel-spectrogram, pelatihan model, knowledge distillation (KD), evaluasi, serta akuisisi data dan perancangan perangkat keras.

### A. DIAGNOSIS SUARA PARU

Suara pernapasan merupakan sinyal akustik penting dalam penilaian klinis berbagai kondisi paru. Jenis suara ini dapat dikategorikan menjadi suara normal (vesikular) dan abnormal seperti *wheeze*, *crackles*, dan *rhonchi*. *Wheeze* menandakan adanya penyempitan saluran pernapasan, *crackles* mengindikasikan adanya cairan di alveoli atau kolaps jaringan

paru, sedangkan *rhonchi* biasanya terdengar pada obstruksi di saluran napas besar [15].

Klasifikasi ini sangat penting untuk diagnosis penyakit seperti asma, Penyakit Paru Obstruktif Kronis (PPOK), dan pneumonia [4]. Namun, keterbatasan utama dari auskultasi konvensional adalah sifatnya yang sangat subjektif, sehingga akurasi diagnosis dapat bervariasi antar individu, bahkan di antara dokter yang terlatih sekalipun [18]. Representasi visual penyakit yang diklasifikasikan pada penelitian ini ditunjukkan pada gambar 2



Gambar 2. Mel-spectrogram dataset

### B. DIAGNOSIS MENGGUNAKAN ARTIFICIAL INTELLIGENCE (AI)

Arsitektur diagnosis dimulai dari pengumpulan data suara paru melalui rekaman digital. Kemudian, sinyal diubah menjadi fitur visual yang menjadi input model CNN dan dilakukan normalisasi serta transformasi tensor. Model menghasilkan keluaran *Respiratory Disease Recognition* (RDR) yang mengaitkan jenis suara dengan empat kategori, yaitu Healthy (sehat), Asma, Penyakit Paru Obstruktif Kronik (PPOK), dan Pneumonia. Hasil klasifikasi ini dapat menjadi dasar diagnosis awal dan tindak lanjut klinis.

### C. TAHAPAN PEMROSESAN MACHINE LEARNING

Penelitian ini menggunakan dataset ICBHI 2017 Respiratory Sound Database dan Mendeley Pulmonary (Lungs) Sound Database sebagai sumber data utama [19]. Dataset ICBHI 2017 memiliki delapan kategori diagnosis, yaitu kondisi sehat, URTI (*Upper Respiratory Tract Infection*), asma, PPOK, LRTI (*Lower Respiratory Tract Infection*), bronkiektasis, pneumonia, dan bronkiolitis, sementara dataset Mendeley memiliki dua kategori yaitu, kondisi normal dan abnormal.

Meskipun dataset mencakup delapan kategori diagnosis, dalam penelitian ini hanya empat kelas utama (Healthy, Asma, PPOK, dan Pneumonia) yang digunakan untuk klasifikasi, karena keempatnya memiliki **jumlah sampel yang memadai**.

Setiap rekaman suara disimpan dalam format WAV dan kemudian dilakukan proses pengolahan data melalui serangkaian tahapan sebagai berikut:

#### 1) PRA-PEMROSESAN

Rekaman yang dihasilkan dari mikrofon diproses sebelum dapat dianalisis menggunakan model. Tahapan pertama adalah segmentasi sinyal menjadi siklus napas individu menggunakan

metode energy thresholding [20]. Selanjutnya dilakukan band-pass filtering pada frekuensi 100–2000 Hz, sesuai rentang karakteristik suara napas manusia [21].

## 2) REPRESENTASI VISUAL DAN SPEKTOGRAM

Selanjutnya, fitur audio dikonversi menjadi representasi visual dalam bentuk mel-spectrogram. Representasi ini memungkinkan model CNN mengenali pola visual yang berkaitan dengan *wheeze* (pita sempit kontinyu) dan *crackles* (pulsa pendek tak beraturan) [22].

## 3) PEMILIHAN DAN PELATIHAN MODEL

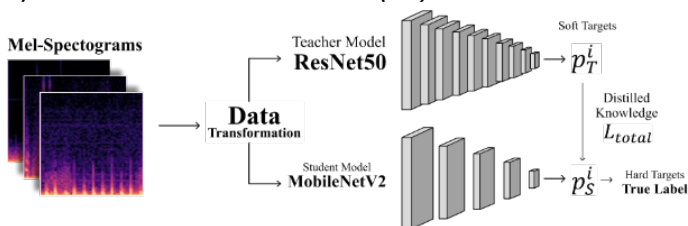
Beberapa model machine learning telah digunakan dalam penelitian suara paru, seperti Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Random Forest (RF), Multi-layer Perceptron (MLP), Long Short-Term Memory (LSTM), dan CNN [23]. Namun, CNN menunjukkan performa paling unggul dalam mengenali pola visual dari spektrogram, dengan akurasi melebihi 90% pada dataset publik seperti ICBHI [24]. Selain itu, CNN merupakan salah satu arsitektur yang paling stabil untuk proses distilasi tanpa penurunan performa.

Model CNN yang digunakan di penelitian ini yaitu ResNet50, pernah digunakan secara luas dalam deteksi pneumonia dari citra X-ray dada untuk diagnosis COVID-19. Arsitektur ini menunjukkan keseimbangan optimal antara kedalaman jaringan dan kompleksitas komputasi, serta kemampuan representasi fitur yang kuat berkat mekanisme residual learning yang memungkinkan pelatihan jaringan yang lebih dalam tanpa mengalami degradasi gradien [25]. Oleh karena itu, ResNet50 dipilih dalam penelitian ini karena kemampuannya menghasilkan representasi fitur yang kaya dan stabil, sekaligus tetap mempertahankan efisiensi dalam pelatihan dan inferensi.

Selain itu, untuk student model menggunakan MobileNetV2 karena karakteristiknya yang ringan dan efisien untuk implementasi di perangkat dengan sumber daya terbatas. Berdasarkan penelitian oleh Velu, MobileNetV2 menunjukkan training accuracy sebesar 92–93% dan precision sebesar 88% pada deteksi COVID-19 yang menggunakan citra X-ray dada [26]. Hasil ini menunjukkan bahwa MobileNetV2 mampu mencapai kinerja tinggi dengan ukuran model yang jauh lebih kecil dibandingkan arsitektur CNN konvensional. Sehingga MobileNetV2 menjadi pilihan ideal sebagai student model dalam skema knowledge distillation yang menuntut keseimbangan antara akurasi dan efisiensi komputasi.

Meskipun CNN sangat efektif dengan akurasi yang tinggi, ukuran model yang besar menjadi tantangan bagi implementasi on-device. Model CNN seperti ResNet-50 atau VGG16 kerap mendapatkan akurasi tinggi, namun memerlukan parameter dan daya komputasi tinggi [27]. Untuk itu, penelitian ini menggunakan pendekatan Knowledge Distillation (KD) untuk menghasilkan model student ringan tanpa penurunan kinerja signifikan.

## 4) KNOWLEDGE DISTILLATION (KD)



Gambar 3. Arsitektur Knowledge Distillation

Dalam KD, model teacher yang kompleks dilatih terlebih dahulu menggunakan dataset lengkap. Kemudian, model student dengan arsitektur lebih ringan dilatih untuk meniru output teacher berupa softmax logits [28]. Gambar arsitektur penelitian ditunjukkan pada Gambar 3.

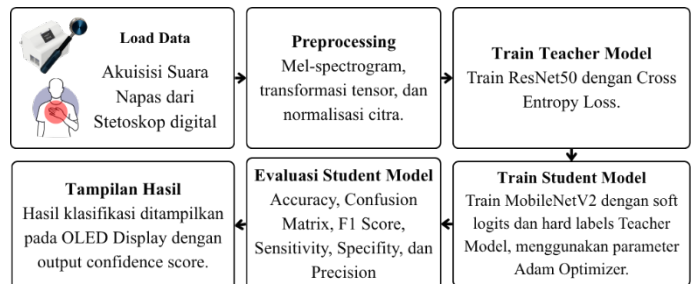
Loss function yang digunakan adalah gabungan cross-entropy loss dengan distillation. Selain itu, digunakan juga intermediate hint loss untuk mempercepat konvergensi student [29]. Target akhir dari model student adalah ukuran maksimal 5 MB agar dapat dijalankan on-device pada perangkat ber-resource terbatas.

## 5) EVALUASI DAN VALIDASI MODEL

Model dievaluasi menggunakan metrik akurasi, sensitivitas, spesifisitas, dan F1-score pada dataset publik ICBHI 2017 Respiratory Sound Database dan Mendeley Pulmonary (Lungs) Sound Database, serta subset data lokal yang dikumpulkan mandiri. Proses evaluasi mencakup k-fold cross-validation untuk menghindari overfitting [30]. Selanjutnya evaluasi kinerja model dilakukan menggunakan beberapa metrik evaluasi, yaitu akurasi, sensitivitas (recall), spesifisitas, precision, dan F1-score. Evaluasi kinerja model dilakukan menggunakan beberapa metrik evaluasi, yaitu akurasi, sensitivitas (recall), spesifisitas, precision, dan F1-score.

## D. AKUISISI DATA DAN PERANGKAT KERAS

Sistem ini dirancang menggunakan perangkat keras berbasis edge computing untuk portabilitas, efisiensi, dan independensi dari koneksi jaringan. Raspberry Pi dipilih karena konsumsi daya rendah dan kemampuan komputasi yang cukup untuk menjalankan inferensi model ringan.



Gambar 4. Diagram Alur Sistem.

Gambar 4 menunjukkan akuisisi suara yang akan diambil menggunakan stetoskop digital. Akuisisi dilakukan dengan menempatkan stetoskop digital pada area anterior dan posterior toraks selama 5–10 detik di tiap titik auskultasi [12]. Data disimpan dalam format WAV melalui modul microSD atau penyimpanan internal, kemudian data akan dikirimkan kepada Raspberry Pi untuk dilakukan pengolahan, hasil pengolahan akan dimunculkan pada layar LCD OLED. Sistem ditenagai oleh baterai li-ion 3.7V rechargeable, yang memungkinkan penggunaan mandiri di lapangan tanpa daya eksternal. Rancangan sistem mencakup integrasi antara stetoskop digital, microcontroller, modul penyimpanan microSD serta LCD OLED untuk menampilkan hasil.

Dengan pendekatan hardware-software co-design ini, prototipe mampu menjalankan seluruh pipeline analisis suara paru secara real-time [5].

Penggunaan TensorFlow Lite dan optimasi kuantisasi parameter memungkinkan inferensi real-time di perangkat dengan RAM terbatas (<512 MB). Dengan demikian, sistem dapat beroperasi secara mandiri tanpa koneksi internet dan cocok untuk layanan kesehatan di daerah terpencil.

### III. HASIL DAN DISKUSI

Bagian ini memaparkan ringkasan hasil eksperimen dan analisis yang diperoleh. Uraian detail disajikan pada subbab berikutnya.

#### A. IMPLEMENTASI DAN PENGUJIAN SISTEM

Prototipe stetoskop digital direalisasikan dengan perangkat portabel meliputi Raspberry Pi 4 Model B, layar OLED, modul penyimpanan microSD, serta baterai Li-ion 3.7V rechargeable.

Pengujian awal di lingkungan terkendali memastikan seluruh pipeline berjalan penuh di perangkat. Sistem mampu melakukan inferensi model CNN hasil knowledge distillation dengan rata-rata 183 ms per input dan konsumsi daya <2,5 W. Seluruh proses dari perekaman suara napas hingga klasifikasi akhir berjalan lokal menggunakan TensorFlow Lite tanpa koneksi internet.

#### B. EVALUASI KINERJA MODEL

Kinerja model dievaluasi berdasarkan kapabilitasnya dalam mengklasifikasikan empat kategori diagnosis dari gabungan dua dataset, ICBHI 2017 Respiratory Sound Database dan Mendeley Pulmonary (Lungs) Sound Database. Proses evaluasi difokuskan pada metrik akurasi yang diukur pada data validasi, yang mencakup 20% dari total dataset. Model teacher (ResNet-50) dan model student (MobileNetV2) dilatih selama 100 epoch untuk memastikan konvergensi yang memadai. Model teacher dilatih terlebih dahulu untuk membangun basis pengetahuan dan berhasil mencapai akurasi validasi yang solid sebesar 90.65%, yang kemudian ditetapkan sebagai baseline untuk proses transfer pengetahuan. Selanjutnya, model student yang memiliki arsitektur lebih ringan dilatih menggunakan metode Knowledge Distillation (KD) dengan parameter suhu ( $T$ ) 3.0 dan faktor penyeimbang ( $\alpha$ ) 0.8.

Hasil akhir pelatihan pada tabel I menunjukkan keberhasilan implementasi Knowledge Distillation, di mana model student berhasil menyamai secara signifikan performa model teacher.

**TABEL I**  
HASIL MODEL

Mo del	Arsitek tur	Par ams	Aku rasi	F1- Sco re	Sensiti fitas	Spesi fitas	Preci sion
Teac her	ResNet 50	23,5 M	0.8167	0.8178	0.8167	0.9389	0.8274
Stud ent	Mobile NetV2	2,2 M	0.8417	0.8436	0.8417	0.9472	0.8495

Fenomena ini menyoroti keunggulan utama dari mekanisme

Knowledge Distillation. Proses ini bekerja dengan melatih model student menggunakan fungsi loss gabungan yang terdiri dari Cross-Entropy Loss terhadap label data asli (hard targets) dan Distillation Loss terhadap distribusi probabilitas keluaran dari teacher (soft targets). Soft targets ini menyajikan sinyal pelatihan yang lebih kaya informasi karena mengandung "pengetahuan gelap" (dark knowledge) mengenai hubungan implisit antar kelas. Sinyal yang lebih bernuansa ini berfungsi sebagai bentuk regularisasi yang efektif, mencegah model student menjadi terlalu percaya diri (overconfident) dan membantunya mempelajari representasi fitur yang lebih tergeneralisasi.

#### C. PERBANDINGAN DENGAN PENELITIAN SEBELUMNYA

Penelitian-penelitian sebelumnya dalam bidang auskultasi digital telah mengeksplorasi berbagai pendekatan machine learning dan deep learning untuk klasifikasi suara jantung dan paru. Namun dalam beberapa tahun terakhir, model CNN telah terbukti menjadi pendekatan paling unggul dalam klasifikasi sinyal auskultasi, baik untuk jantung maupun paru-paru.

Zhang et al. [31] mengembangkan stetoskop digital berbasis AI yang terintegrasi dengan model hibrida CNN dan Best Discrepancy Forest (BDF) untuk klasifikasi simultan penyakit jantung dan paru. Mereka melaporkan akurasi hingga 99,94% dalam klasifikasi 11 kelas sinyal auskultasi. Model CNN digunakan untuk ekstraksi fitur spasial dari representasi spektrum suara, yang kemudian dikombinasikan dengan BDF sebagai classifier akhir. Pendekatan ini menegaskan keunggulan CNN dalam menangkap pola-pola kompleks dalam sinyal audio medis.

Tabel II menjelaskan lebih rinci perbandingan antara beberapa model. Pada tabel tersebut dapat dilihat bahwa CNN Hybrid memiliki skor akurasi paling tinggi dibandingkan dengan model yang lain.

Selain unggul dalam akurasi, CNN juga menjadi model yang paling ideal untuk diintegrasikan dengan pendekatan Knowledge Distillation (KD). CNN memiliki arsitektur hierarkis yang memungkinkan transfer representasi pengetahuan secara efisien dari model teacher ke model student. Dalam konteks implementasi edge computing seperti perangkat stetoskop digital low-cost, Knowledge Distillation menjadi sangat relevan. CNN sebagai teacher model dapat dilatih secara mendalam dan kemudian didistilasi menjadi student model yang lebih ringan dan efisien, seperti yang dilakukan dalam beberapa penelitian terkini [31].

**TABEL II**  
PERBANDINGAN DENGAN PENELITIAN SEBELUMNYA

Penelitian	Model	Dataset	Akurasi (%)	Kelebihan	Kekurangan
Kim et al. [14]	Random Forest + Feature Extraction	ICBHI 2017	84.2	Cepat dilatih dan mudah diinterpretasi.	Akurasi rendah, sensitif terhadap noise.
Perna et al. [32]	CNN sederhana	ICBHI 2017	89.5	Struktur sederhana, training stabil.	Kurang robust terhadap variasi domain.
Shuvo et al. [33]	Hybrid CNN + Transformer	Private (pediatric)	90.3	Kuat dalam long-term dependency.	Kompleks, mahal secara komputasi.
Abeyratne et al. [34]	SVM + Feature Engineering	Private (asma)	86.7	Efektif pada dataset kecil	Sulit generalisasi, butuh pra-pemrosesan banyak
Hasil Penelitian	CNN + Knowledge Distillation	ICBHI 2017 + Mendeley Respiratory Sound Dataset	Teacher: 81.67 Student : 84.17	Menghasilkan model ringan dengan akurasi tinggi	Membutuhkan waktu pelatihan lebih lama dan sensitif terhadap noise rekaman.

## IV. KESIMPULAN

Penelitian ini berhasil merancang stetoskop digital cerdas berbasis komputasi on-edge untuk diagnosis penyakit pernapasan di wilayah dengan keterbatasan sumber daya. Sistem memanfaatkan CNN yang dioptimalkan melalui Knowledge Distillation (KD), memungkinkan kinerja tinggi pada perangkat berdaya rendah. Model student MobileNetV2 mencapai akurasi 84,17%, mendekati model teacher ResNet-50 (81,67%), menunjukkan efektivitas KD sebagai teknik regularisasi yang meningkatkan kemampuan generalisasi. Prototipe berbasis Raspberry Pi 4 Model B mampu menjalankan diagnosis secara lokal tanpa koneksi internet. Hasil ini membuktikan bahwa co-design hardware–software berbasis KD merupakan solusi potensial untuk alat skrining penyakit pernapasan yang akurat, portabel, dan terjangkau. Pada tahap pengembangan selanjutnya, penelitian diarahkan pada pelatihan model dengan data penyakit pernapasan yang lebih beragam guna meningkatkan kemampuan sistem dalam klasifikasi multi-kelas.

## REFERENSI

- [1] GBD 2021 Chronic Respiratory Diseases Collaborators, "Global, regional, and national burden of chronic respiratory diseases, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021," *The Lancet Respiratory Medicine*, vol. 11, no. 10, pp. 879–900, 2023, doi: 10.1016/S2213-2600(23)00248-7.
- [2] M. Faizal, M. R. Islam, and S. M. Rahman, "Deep learning-based multi-class classification of respiratory diseases from lung sound signals," *PLoS ONE*, vol. 19, no. 8, e0308225, 2024, doi: 10.1371/journal.pone.0308225.
- [3] J. L. Arts, E. H. T. Lim, P. M. van de Ven, L. Heunks, and P. R. Tuinman, "The diagnostic accuracy of lung auscultation in adult patients with acute pulmonary pathologies: a meta-analysis," *Sci. Rep.*, vol. 10, no. 1, pp. 1–11, 2020, doi: 10.1038/s41598-020-64405-6.
- [4] H. Lee, G. Kim, and J. S. Bae, "Digital auscultation in clear and present threat of novel respiratory infectious disease: a narrative review," *J. Yeungnam Med. Sci.*, vol. 42, pp. 1–8, 2025, doi: 10.12701/jyms.2025.42.19.
- [5] A. Auscultation, D. K. Pillai, S. J. Teach, S. J. Patel, and R. W. Doroshov, "DMD2023-8881," pp. 1–4, 2023.
- [6] A. Youness, N. T. N. D. R. W. and S. Raj, "Technical characterisation of digital stethoscopes : towards scalable artificial intelligence-based auscultation," *J. Med. Eng. Technol.*, vol. 0, no. 0, pp. 1–14, 2024, doi: 10.1080/03091902.2023.2174198.
- [7] Y. Kim et al., "Respiratory sound classification for crackles, wheezes, and rhonchi in the clinical field using deep learning," *Sci. Rep.*, vol. 11, no. 1, pp. 1–11, 2021, doi: 10.1038/s41598-021-96724-7.
- [8] A. Mondal, P. Banerjee, and H. Tang, "A novel feature extraction technique for pulmonary sound analysis based on EMD," *Comput. Methods Programs Biomed.*, vol. 159, pp. 199–209, 2018, doi: 10.1016/j.cmpb.2018.03.016.
- [9] R. Nadkarni, E. Nikolakis, and R. V. Marinescu, "AFEN: Respiratory disease classification using ensemble learning," *arXiv preprint, arXiv:2405.05467*, 2024. [Online]. Available: <https://arxiv.org/abs/2405.05467>.
- [10] M. Aykanat, Ö. Kılıç, B. Kurt, and S. Saryal, "Classification of lung sounds using convolutional neural networks," *Eurasip J. Image Video Process.*, vol. 2017, no. 1, 2017, doi: 10.1186/s13640-017-0213-2.
- [11] S. B. Shuvo, S. N. Ali, S. I. Swapnil, T. Hasan, and M. I. H. Bhuiyan, "A lightweight CNN model for detecting respiratory diseases from lung auscultation sounds using EMD-CWT-based hybrid scalogram," *IEEE J. Biomed. Heal. Informatics*, vol. 25, no. 7, pp. 2595–2603, 2021, doi: 10.1109/JBHI.2020.3048006.
- [12] J. W. Kim, S. Bae, W. Y. Cho, B. Lee, and H. Y. Jung, "Stethoscope-guided supervised contrastive learning for cross-domain adaptation on respiratory sound classification," in *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, pp. 1431–1435, 2024, doi: 10.1109/ICASSP48485.2024.10447734.
- [13] B. H. M. van der Velden, H. J. Kuijff, K. G. A. Gilhuijs, and M. A. Viergever, "Explainable artificial intelligence (XAI) in deep learning-based medical image analysis," *Med. Image Anal.*, vol. 79, p. 102470, 2022, doi: 10.1016/j.media.2022.102470.
- [14] G. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," 2015. [Online]. Available: <http://arxiv.org/abs/1503.02531>
- [15] K. Choi, M. Kersner, J. Morton, and B. Chang, "Temporal knowledge distillation for on-device audio classification," in *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, vol. 2022-May, pp. 486–490, 2022, doi: 10.1109/ICASSP43922.2022.9747908.
- [16] T. Zhong, Z. Chi, L. Gu, Y. Wang, Y. Yu, and J. Tang, "Meta-DMoE: Adapting to domain shift by meta-distillation from mixture-of-experts," *Adv. Neural Inf. Process. Syst.*, vol. 35, no. NeurIPS, pp. 1–21, 2022.
- [17] T. Grzywalski et al., "Practical implementation of artificial intelligence algorithms in pulmonary auscultation examination," *Eur. J. Pediatr.*, vol. 178, no. 6, pp. 883–890, 2019, doi: 10.1007/s00431-019-03363-2.
- [18] A. Bohadana, G. Izicki, and S. S. Kraman, "Fundamentals of lung auscultation," *N. Engl. J. Med.*, vol. 370, no. 8, pp. 744–751, 2014, doi: 10.1056/nejmra1302901.
- [19] N. Baghel, V. Nangia, and M. K. Dutta, "ALSD-Net: Automatic lung sounds diagnosis network from pulmonary signals," *Neural Computing & Applications*, vol. 33, pp. 17103–17118, 2021, doi: 10.1007/s00521-021-06302-1.
- [20] H. Chen, X. Yuan, Z. Pei, M. Li, and J. Li, "Triple-classification of respiratory sounds using optimized S-transform and deep residual networks," *IEEE Access*, vol. 7, no. April, pp. 32845–32852, 2019, doi: 10.1109/ACCESS.2019.2903859.
- [21] T. Hirose et al., "The utility of real-time remote auscultation using a bluetooth-connected electronic stethoscope: Open-label randomized controlled pilot trial," *JMIR mHealth uHealth*, vol. 9, no. 7, pp. 1–11, 2021, doi: 10.2196/23109.
- [22] R. Palaniappan, K. Sundaraj, and S. Sundaraj, "Artificial intelligence techniques used in respiratory sound analysis - A systematic review," *Biomed. Tech.*, vol. 59, no. 1, pp. 7–18, 2014, doi: 10.1515/bmt-2013-0074.
- [23] X. Xu, I. Ra, and R. Sankar, "Enhancing lung disease diagnosis via semi-supervised machine learning," no. Ifsa, pp. 171–176, 2023.
- [24] Z. Zhu et al., "Development and application of a deep learning-based comprehensive early diagnostic model for chronic obstructive pulmonary disease," *Respir. Res.*, vol. 25, no. 1, pp. 1–12, 2024, doi: 10.1186/s12931-024-02793-3.
- [25] S. Showkat and S. Qureshi, "Efficacy of Transfer Learning-based ResNet models in Chest X-ray image classification for detecting COVID-19 Pneumonia," *Chemometrics and Intelligent Laboratory Systems*, vol. 224, p. 104534, Mar. 2022, doi: 10.1016/j.chemolab.2022.104534.
- [26] M. K. Dutta, V. Nangia, and N. Baghel, "ALSD-Net: Automatic lung sounds diagnosis network from pulmonary signals," *Neural Computing and Applications*, vol. 33, no. 15, pp. 17103–17118, 2021, doi: 10.1007/s00521-021-06302-1.
- [27] S. Mei et al., "Deep learning for detecting and early predicting chronic obstructive pulmonary disease from spirogram time series," *npj Syst. Biol. Appl.*, vol. 11, no. 1, 2025, doi: 10.1038/s41540-025-00489-y.
- [28] A. Moslemi, A. Briskina, Z. Dang, and J. Li, "A survey on knowledge distillation: Recent advancements," *Mach. Learn. Appl.*, vol. 18, p. 100605, 2024, doi: 10.1016/j.mlwa.2024.100605.
- [29] L. Fang et al., "Knowledge distillation and dataset distillation of large language models: Emerging trends, challenges, and future directions," 2025. [Online]. Available: <http://arxiv.org/abs/2504.14772>
- [30] F. S. Hsu et al., "Benchmarking of eight recurrent neural network variants for breath phase and adventitious sound detection on a self-developed open-access lung sound database-HF\_Lung\_V1," *PLoS One*, vol. 16, no. 7, pp. 1–26, 2021, doi: 10.1371/journal.pone.0254134.
- [31] H. A. Sounds, "A Low-Cost AI-Empowered Stethoscope and a Lightweight," 2023.
- [32] D. Perna and A. Tagarelli, "Deep auscultation: Predicting respiratory anomalies and diseases via recurrent neural networks," in *Proc. - IEEE Symp. Comput. Med. Syst.*, vol. 2019-June, pp. 50–55, 2019, doi: 10.1109/CBMS.2019.00020.
- [33] S. B. Shuvo and T. Hasan, "A multi-stage hybrid CNN-transformer network for automated pediatric lung sound classification," 2025. [Online]. Available: <http://arxiv.org/abs/2507.20408>
- [34] U. R. Abeyaratne, A. Swarnkar, and V. Setyati, "Automatic wheeze detection from asthmatic lung sounds," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 3, pp. 597–606, Mar. 2011, doi: 10.1109/TBME.2010.2096509.