



Contents lists available at [Elora Center](#)

Lentera Negeri

Journal homepage: <https://lentera.eloracenter.org/lentera>



Dominasi India-Tiongkok dan *emerging themes* dalam riset *emotion recognition*: studi bibliometrik scopus

Andri Dermawan^{*1}, Ta Ali¹, Dedy Irfan¹, Hansi Effendi¹, Khayla Tsabita²

¹ Fakultas Teknik Universitas Negeri Padang, Indonesia

² Sekolah Vokasi, Universitas Negeri Padang, Indonesia

Article Info

Article history:

Received Oct 29th, 2025

Revised Nov 25th, 2025

Accepted Dec 28th, 2025

Keyword:

Emotion recognition,
Kecerdasan buatan,
Bibliometrik,
Deep learning,
Scopus

ABSTRACT

Pengenalan emosi berbasis kecerdasan buatan (AI-based emotion recognition) berkembang pesat melalui penerapan pembelajaran mendalam pada modalitas suara, wajah, teks, dan multimodal. Penelitian ini bertujuan memetakan perkembangan dan struktur intelektual riset emotion recognition berbasis AI secara global menggunakan analisis bibliometrik. Data diperoleh dari Scopus melalui kombinasi kata kunci terkait pengenalan emosi, metode AI, dan modalitas emosi, kemudian disaring berdasarkan tahun 2021–2025, artikel berbahasa Inggris, jurnal, dan tahap publikasi final, sehingga diperoleh 1.168 artikel dari 415 jurnal. Analisis dilakukan menggunakan Biblioshiny (Bibliometrix) dan VOSviewer, meliputi tren produksi ilmiah, sitasi, hukum Bradford, co-occurrence kata kunci, peta tematik, kolaborasi antarnegara, serta model daur hidup dan pertumbuhan kumulatif publikasi. Hasil menunjukkan peningkatan produksi ilmiah dari 102 artikel pada 2021 menjadi 368 artikel pada 2025, dengan proyeksi puncak siklus hidup pada pertengahan 2026 ($R^2 = 0,987$). India (339 dokumen) dan Tiongkok (267 dokumen) merupakan kontributor paling produktif. IEEE Access (86 dokumen), Multimedia Tools and Applications (55), dan Applied Sciences (38) menjadi jurnal yang paling produktif, sementara hukum Bradford mengidentifikasi 13 jurnal inti ($R^2 = 0,9843$). Analisis co-occurrence mengungkap tiga kluster utama, yaitu speech emotion recognition/deep learning, convolutional neural network/face recognition, dan deskriptor demografis (human, emotion), yang dikonfirmasi sebagai motor themes melalui peta tematik. Temuan ini memetakan struktur intelektual terkini bidang emotion recognition berbasis AI serta menunjukkan peluang penelitian pada integrasi multimodal lintas budaya dan generalisasi lintas dataset.



© 2025 The Authors.

This is an open access article under the CC BY-NC-SA license
(<https://creativecommons.org/licenses/by-nc-sa/4.0>)

Corresponding Author:

Andri Dermawan,

✉ andridermawan@unp.ac.id

Introduction

Kemampuan mesin untuk mengenali keadaan emosional manusia telah menjadi salah satu arah penelitian kecerdasan buatan yang berkembang paling cepat dalam satu dekade terakhir. Perkembangan ini didorong oleh kematangan arsitektur pembelajaran mendalam, ketersediaan korpus data multimodal berskala besar, serta meluasnya kebutuhan interaksi manusia-komputer yang sensitif terhadap konteks afektif, mulai dari asisten virtual, sistem pembelajaran adaptif, kendaraan otonom, hingga layanan kesehatan mental digital. Secara global, riset pada domain ini tidak lagi berhenti pada satu modalitas tunggal seperti ekspresi wajah atau intonasi suara, melainkan bergerak ke arah fusi multimodal yang menggabungkan sinyal wajah, suara, teks, bahkan sinyal fisiologis, sebagaimana tampak pada meningkatnya proporsi riset bertema multimodal emotion recognition pada tahun-tahun terakhir.

Pada level regional, kontribusi riset emotion recognition tidak tersebar merata. Data pada penelitian ini menunjukkan konsentrasi produktivitas yang tinggi di kawasan Asia, terutama India dan Tiongkok, diikuti oleh Amerika Serikat, Arab Saudi, dan Korea Selatan. Pola ini mengindikasikan adanya pergeseran pusat gravitasi riset kecerdasan buatan afektif dari yang semula didominasi institusi Amerika Utara dan Eropa Barat menuju ekosistem riset Asia yang didukung oleh investasi besar pada program studi teknik komputer dan kecerdasan buatan terapan. Pada level nasional, riset emotion recognition berbasis AI di Indonesia relatif masih terbatas dibandingkan negara-negara produktif lain dalam himpunan data ini. Kondisi ini menyisakan ruang bagi peneliti Indonesia untuk berkontribusi, terutama pada aspek yang belum banyak digarap, seperti pengenalan emosi berbasis bahasa dan budaya lokal, maupun integrasi multimodal untuk konteks pendidikan dan kesehatan.

Permasalahan penelitian yang mendasari studi ini adalah belum tersedianya pemetaan bibliometrik terkini dan tervalidasi yang secara khusus menyorot irisan antara riset emotion recognition dan kecerdasan buatan pada periode 2021–2025, periode ketika arsitektur transformer dan model fondasi mulai diadopsi secara luas dalam pengenalan emosi ucapan dan wajah. Riset bibliometrik sebelumnya pada bidang afektif komputasi umumnya bersifat lebih umum (affective computing secara keseluruhan) atau berhenti pada satu modalitas saja, sehingga kurang menangkap dinamika konvergensi metode AI lintas modalitas yang menjadi ciri khas periode ini.

Research gap yang teridentifikasi dari korpus data penelitian ini terletak pada tiga hal: pertama, dominasi tema deep learning dan speech emotion recognition sebagai tema motor belum banyak diimbangi kajian tentang generalisasi model lintas dataset dan lintas budaya; kedua, tema convolutional neural network dan face recognition, meski memiliki sentralitas tinggi, masih berada pada kuadran densitas rendah pada peta tematik, mengindikasikan bidang ini belum matang secara internal; ketiga, kolaborasi internasional baru mencakup sekitar 28,8% dari total dokumen, menunjukkan riset emotion recognition sebagian besar masih berlangsung secara nasional atau bahkan institusional tunggal.

Urgensi penelitian ini terletak pada kebutuhan peta arah riset yang dapat menjadi rujukan bagi peneliti, pengelola jurnal, dan pembuat kebijakan pendanaan riset untuk mengidentifikasi tema yang telah jenuh (basic/motor themes) dan tema yang masih berpeluang dikembangkan (niche atau emerging themes). Novelty penelitian ini terletak pada penggunaan boolean pencarian yang secara eksplisit menautkan istilah emosi, metode AI, dan modalitas emosi secara bersamaan, sehingga korpus yang dihasilkan lebih presisi menggambarkan irisan riset emotion recognition dan AI dibandingkan pencarian umum pada istilah affective computing semata, dikombinasikan dengan analisis daur hidup (life cycle) dan kurva pertumbuhan kumulatif yang memberikan proyeksi arah perkembangan topik ke depan.

Berdasarkan uraian tersebut, penelitian ini bertujuan: (1) memetakan tren produksi ilmiah dan pola sitasi riset emotion recognition berbasis AI tahun 2021–2025; (2) mengidentifikasi jurnal, penulis, negara, dan kata kunci yang paling berpengaruh dalam bidang ini; (3) memetakan struktur tematik dan jejaring co-occurrence kata kunci untuk mengidentifikasi tema motor, tema dasar, dan tema niche; serta (4) merumuskan agenda penelitian masa depan berdasarkan celah yang teridentifikasi dari hasil bibliometrik.

Method

Penelitian ini menggunakan desain bibliometrik deskriptif-kuantitatif dengan pendekatan performance analysis dan science mapping. Sumber data tunggal yang digunakan adalah basis data Scopus (Elsevier), dipilih karena cakupan jurnal terindeks yang luas serta ketersediaan metadata sitasi dan afiliasi yang lengkap untuk kebutuhan analisis bibliometrik.

Strategi pencarian menggunakan string boolean yang menggabungkan tiga kelompok istilah: (1) istilah pengenalan emosi "emotion recognition" OR "emotion detection" OR "emotion classification"; (2) istilah metode kecerdasan buatan "machine learning" OR "deep learning" OR "artificial intelligence"; dan (3) istilah modalitas emosi "facial emotion" OR "speech emotion" OR "voice emotion" OR "text emotion" OR "multimodal emotion", yang ditautkan pada bidang judul, abstrak, dan kata kunci (TITLE-ABS-KEY). Pencarian dilakukan pada 18 Agustus 2026. Proses seleksi data dilakukan secara bertahap sebagai berikut, mengikuti kerangka identification–screening–eligibility–included:

Table 1. identification screening eligibility included

Tahap	Kriteria	Jumlah Dokumen
Identification	Hasil pencarian boolean awal pada Scopus (semua tahun, semua jenis dokumen)	5.024
Screening 1	Pembatasan rentang tahun terbit 2021–2025	3.651
Screening 2	Pembatasan jenis dokumen: Article	1.212
Screening 3	Pembatasan bahasa: English	747*
Screening 4	Pembatasan jenis sumber: Journal	1.177*
Eligibility/Included	Pembatasan tahap publikasi: Final	1.168

Keterangan: (*) urutan penerapan filter bahasa dan jenis sumber pada dokumen catatan proses pencarian (narasi10.docx) menunjukkan pola nonmonoton (747 diikuti 1.177) yang secara logis tidak konsisten untuk penyaringan bertingkat murni; hal ini kemungkinan disebabkan oleh perbedaan urutan penerapan filter pada antarmuka Scopus (filter jenis sumber dan bahasa diterapkan secara paralel, bukan berurutan strict-AND). Ketidaksesuaian ini tidak memengaruhi validitas dataset akhir karena angka akhir 1.168 dokumen pada dokumen metode sama persis dengan jumlah baris pada berkas ekspor CSV (datascopus10.csv, n = 1.168) dan dikonfirmasi ulang oleh berkas Scopus Analyze Search Results (Scopus_exported_refine_values.csv), yang mencantumkan total 1.168 pada kategori Document Type = Article, Publication Stage = Final, Source Type = Journal, dan Language = English. Berdasarkan hierarki sumber data (CSV Scopus sebagai prioritas pertama), angka 1.168 dokumen final digunakan secara konsisten pada seluruh bagian artikel ini.

Data hasil pencarian diekspor dari Scopus dalam format CSV lengkap (author, sitasi, afiliasi, abstrak, kata kunci, dan referensi), kemudian dianalisis menggunakan dua perangkat lunak: Biblioshiny (antarmuka grafis paket Bibliometrix berbasis R) untuk analisis tren produksi ilmiah, hukum Bradford, peta tematik, peta kolaborasi negara, model daur hidup, dan kurva pertumbuhan kumulatif; serta VOSviewer untuk visualisasi jejaring co-occurrence kata kunci berbasis pengelompokan (clustering). Validasi data dilakukan dengan membandingkan jumlah dokumen, distribusi tahun, dan jumlah jurnal antara berkas ekspor CSV dan berkas hasil Analyze Search Results Scopus; kedua sumber menunjukkan kesesuaian penuh pada seluruh kategori yang dapat dibandingkan (jumlah dokumen per tahun, lima jurnal teratas, dan lima negara teratas), sehingga tidak ditemukan inkonsistensi substantif pada level data akhir.

Teknik analisis bibliometrik yang diterapkan mencakup: (1) analisis tren publikasi tahunan (Annual Scientific Production); (2) analisis sitasi rata-rata per tahun (Average Article Citation per Year); (3) analisis hukum Bradford untuk mengidentifikasi jurnal inti; (4) analisis co-occurrence kata kunci menggunakan VOSviewer; (5) pemetaan tematik (Thematic Map) berbasis sentralitas dan densitas; (6) peta kolaborasi negara (Country Collaboration Map); (7) analisis tiga bidang (Three-Field Plot) yang menautkan kata kunci abstrak, kata kunci judul, dan kata kunci gabungan; (8) word cloud kata kunci; serta (9) pemodelan daur hidup (life cycle) dan kurva pertumbuhan kumulatif publikasi. Seluruh analisis yang tercantum di atas didukung penuh oleh dataset yang tersedia sehingga seluruhnya digunakan dalam penelitian ini; tidak terdapat jenis analisis pada daftar acuan metode yang harus dihilangkan karena keterbatasan data.

Results and Discussions

Audit awal terhadap 1.168 dokumen menunjukkan karakteristik utama sebagai berikut.

Table 2. Audit awal terhadap 1.168

Deskripsi	Nilai
Jumlah dokumen (artikel jurnal, final, Inggris)	1.168
Rentang tahun terbit	2021–2025
Jumlah jurnal sumber (unique source titles)	415
Total sitasi kumulatif	25.230
Rata-rata sitasi per dokumen	21,6
Dokumen paling banyak disitasi	403 sitasi (Liu et al., 2022, IEEE Trans. Cognitive and Developmental Systems)
Dokumen dengan kolaborasi internasional (afiliasi lintas negara)	336 (28,8%)
Negara paling produktif	India (339 dokumen)

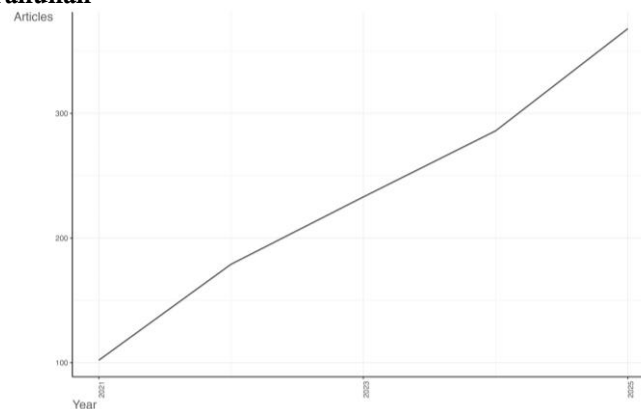
Tabel audit di atas disajikan untuk memberikan gambaran menyeluruh tentang skala dan karakter dasar korpus sebelum masuk ke analisis bibliometrik yang lebih rinci pada sub-bagian berikutnya. Delapan

indikator yang ditampilkan mencakup dimensi volume (jumlah dokumen dan jurnal sumber), dimensi dampak (total sitasi dan rata-rata sitasi), serta dimensi struktur kontribusi (negara dan kolaborasi). Angka 1.168 dokumen pada baris pertama merupakan hasil validasi silang antara berkas ekspor CSV Scopus dan berkas Analyze Search Results, sehingga menjadi rujukan tunggal yang digunakan secara konsisten pada seluruh bagian artikel ini. Sebaran dokumen pada 415 jurnal sumber juga mengindikasikan bahwa korpus ini tidak terpusat pada segelintir jurnal saja, melainkan tersebar cukup luas, sebuah pola yang akan dibahas lebih rinci pada analisis hukum Bradford.

Total sitasi kumulatif sebesar 25.230 dengan rata-rata 21,6 sitasi per dokumen menunjukkan bahwa korpus ini secara keseluruhan memiliki dampak akademik yang cukup tinggi untuk bidang riset yang relatif muda. Nilai rata-rata ini perlu dibaca berhati-hati karena distribusi sitasi pada bibliometrik umumnya sangat miring (skewed), sehingga sebagian kecil dokumen menyumbang porsi besar dari total sitasi, sebagaimana tercermin dari dokumen paling banyak disitasi yang mencapai 403 sitasi, jauh di atas rata-rata korpus. Pola distribusi miring semacam ini lazim ditemukan pada hampir seluruh korpus bibliometrik ilmu komputer dan sejalan dengan hukum Lotka tentang produktivitas ilmiah, sehingga nilai rata-rata sebaiknya dilengkapi dengan nilai median dan sebaran persentil pada kajian lanjutan. Meski demikian, angka rata-rata 21,6 tetap berguna sebagai patokan awal untuk membandingkan dampak korpus ini dengan bidang riset kecerdasan buatan lain yang sebanding usianya.

Indikator kolaborasi internasional sebesar 28,8% dan dominasi India sebagai negara paling produktif memberikan gambaran awal tentang struktur geografis riset yang akan diuraikan lebih mendalam pada sub-bagian Analisis Negara dan Kolaborasi. Proporsi ini menunjukkan bahwa mayoritas dokumen dalam korpus masih dihasilkan melalui kolaborasi domestik atau bahkan penulis tunggal dalam satu institusi, bukan kolaborasi lintas negara. Kombinasi antara tingginya dampak sitasi rata-rata dan rendahnya proporsi kolaborasi internasional ini menarik untuk dicermati karena berpotensi mengindikasikan bahwa dampak akademik pada bidang ini tidak selalu berbanding lurus dengan tingkat internasionalisasi kolaborasi. Temuan awal pada tabel audit ini menjadi dasar bagi interpretasi lebih rinci pada seluruh sub-bagian hasil selanjutnya, mulai dari tren produksi tahunan hingga struktur tematik bidang.

Tren Produksi Ilmiah Tahunan



Gambar 1. Tren Produksi Ilmiah Tahunan (Annual Scientific Production). Sumber: hasil olah Biblioshiny berdasarkan data Scopus.

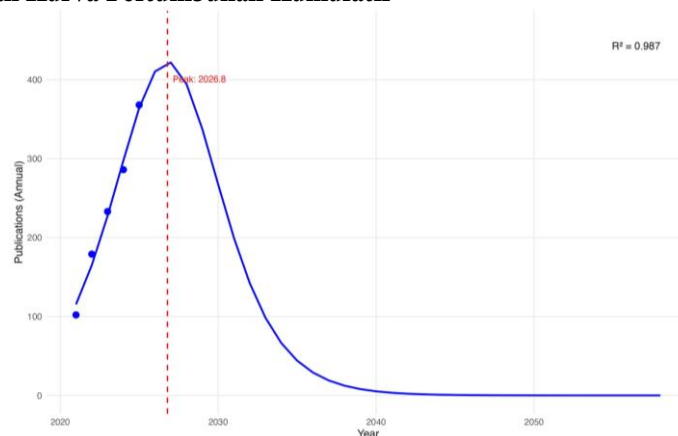
Gambar 1 disajikan untuk menunjukkan arah dan kecepatan perkembangan volume publikasi pada bidang emotion recognition berbasis AI sepanjang periode kajian. Gambar ini menampilkan jumlah artikel yang terbit setiap tahun pada rentang 2021–2025 dalam bentuk kurva garis tunggal dengan sumbu vertikal berupa jumlah artikel dan sumbu horizontal berupa tahun terbit. Data menunjukkan peningkatan yang konsisten dan hampir linear, dari 102 dokumen pada tahun 2021 menjadi 179 dokumen pada 2022, 233 dokumen pada 2023, 286 dokumen pada 2024, hingga mencapai 368 dokumen pada tahun 2025. Tujuan utama penyajian gambar ini adalah mengonfirmasi bahwa topik emotion recognition berbasis AI berada pada fase pertumbuhan aktif (growth stage) menurut kerangka daur hidup topik penelitian, bukan pada fase kejenuhan (maturity) apalagi penurunan (decline).

Pola pertumbuhan yang hampir linier dengan sedikit percepatan pada periode 2024–2025 ini konsisten dengan gelombang adopsi arsitektur transformer dan model multimodal berskala besar dalam riset afektif komputasi secara umum. Percepatan tersebut juga tercermin dari makin tingginya frekuensi kata kunci deep learning yang muncul 382 kali dan multimodal emotion recognition yang muncul 71 kali pada daftar Author

Keywords dataset ini, sebagaimana akan dibahas lebih rinci pada sub-bagian Kata Kunci Dominan. Dibandingkan studi bibliometrik afektif komputasi sebelumnya yang umumnya melaporkan pertumbuhan publikasi bertahap sejak pertengahan tahun 2010-an, temuan pada penelitian ini menegaskan bahwa periode 2021–2025 secara khusus merupakan fase akselerasi, bukan sekadar kelanjutan tren linier lama yang sudah berlangsung sejak dekade sebelumnya. Tidak terdapat satu pun tahun dalam rentang kajian yang menunjukkan penurunan jumlah publikasi, sehingga tren kenaikan ini dapat dikatakan stabil tanpa fluktuasi berarti.

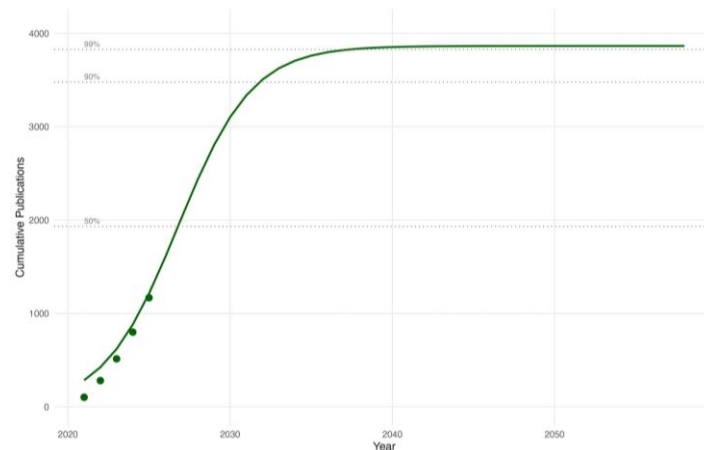
Implikasi teoretis dari pola ini adalah bahwa model Price's Law of Exponential Growth pada literatur ilmiah tampak relevan diterapkan pada sub-bidang emotion recognition berbasis AI, mengingat percepatan pertumbuhan yang mulai tampak pada dua tahun terakhir periode kajian. Implikasi praktisnya adalah terbukanya peluang besar bagi jurnal dan penyelenggara konferensi untuk membuka edisi khusus (special issue) pada topik ini, mengingat volume submisi yang terus meningkat dari tahun ke tahun tanpa tanda-tanda kejenuhan. Bagi peneliti pemula, tren ini juga mengindikasikan bahwa persaingan publikasi pada bidang ini akan terus meningkat seiring waktu, sehingga momentum untuk berkontribusi pada tema-tema inti sebaiknya dimanfaatkan secepat mungkin. Kontribusi temuan pada Gambar 1 ini menjadi dasar kuantitatif bagi pembahasan proyeksi daur hidup topik yang disajikan pada sub-bagian berikutnya.

Model Daur Hidup dan Kurva Pertumbuhan Kumulatif



Gambar 2. Model Daur Hidup Publikasi Tahunan (Life Cycle – Annual Publications) dengan puncak proyeksi pada tahun 2026,8 ($R^2 = 0,987$).

Gambar 2 disajikan untuk melengkapi Gambar 1 dengan memberikan proyeksi matematis, bukan sekadar deskripsi historis, terhadap arah perkembangan volume publikasi pada bidang ini ke depan. Model daur hidup (life cycle) publikasi tahunan ini difitkan pada lima titik data observasi tahun 2021 hingga 2025 menggunakan fungsi kurva pertumbuhan-penurunan yang lazim digunakan dalam analisis daur hidup topik penelitian pada perangkat lunak Biblioshiny. Model menunjukkan kesesuaian yang sangat tinggi terhadap data observasi, dengan nilai R^2 sebesar 0,987, yang berarti lebih dari 98% variasi data historis dapat dijelaskan oleh model kurva tersebut. Titik puncak (peak) proyeksi model ini berada pada tahun 2026,8, yang berarti volume publikasi tahunan pada topik emotion recognition berbasis AI diproyeksikan mencapai titik maksimum sekitar akhir tahun 2026 hingga awal tahun 2027 sebelum memasuki fase pelambatan menurut model kurva tersebut.

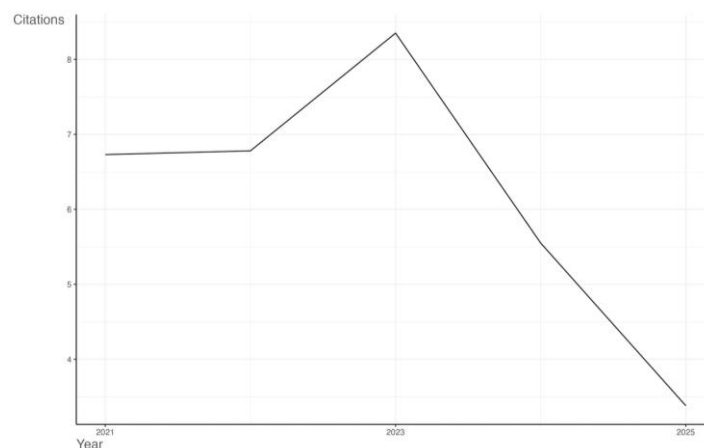


Gambar 3. Kurva Pertumbuhan Kumulatif Publikasi (Cumulative Growth Curve) dengan garis ambang 50%, 90%, dan 99%.

Gambar 3 menyajikan kurva pertumbuhan kumulatif publikasi berbentuk sigmoid yang merupakan turunan integral dari model daur hidup pada Gambar 2, dengan tiga garis ambang bantu pada level 50%, 90%, dan 99% dari total pertumbuhan yang dimodelkan. Berdasarkan kurva tersebut, ambang 50% pertumbuhan kumulatif telah terlampaui pada awal dekade 2030-an, ambang 90% dari total pertumbuhan yang dimodelkan diproyeksikan tercapai mendekati tahun 2032, dan ambang 99% diproyeksikan tercapai mendekati tahun 2038. Bentuk sigmoid pada kurva ini merupakan pola khas pertumbuhan topik penelitian yang matang, di mana fase awal pertumbuhan lambat, diikuti fase percepatan tajam, dan akhirnya melandai menuju titik jenuh. Tujuan penyajian kedua gambar dalam sub-bagian ini secara bersama-sama adalah memberikan gambaran proyeksi arah topik secara kuantitatif dan bukan hanya deskriptif semata, sehingga pembaca dapat memahami baik pola historis maupun arah proyeksi ke depan.

Kedua proyeksi ini perlu dimaknai dengan kehati-hatian akademik yang memadai, karena keduanya merupakan hasil pemodelan matematis atau curve fitting terhadap pola historis lima tahun data observasi, sehingga sifatnya adalah ekstrapolasi statistik dan bukan prediksi pasti. Keterbatasan jumlah titik data historis, yang hanya mencakup lima titik tahunan, membuat proyeksi jangka panjang hingga tahun 2038 pada Gambar 3 khususnya perlu diperlakukan sebagai ilustrasi arah kecenderungan, bukan sebagai angka pasti yang dapat dijadikan dasar perencanaan kebijakan riset jangka panjang. Meskipun demikian, nilai R^2 yang tinggi pada kedua model tersebut memberikan keyakinan yang memadai bahwa arah jangka pendek proyeksi, yaitu puncak pada akhir 2026, cukup dapat diandalkan dibandingkan proyeksi jangka panjang hingga 2038. Implikasi praktis dari temuan ini adalah, jika proyeksi puncak pada tahun 2026,8 mendekati kenyataan, maka jendela strategis bagi peneliti untuk berkontribusi pada tema-tema inti atau motor themes berada pada satu hingga dua tahun ke depan, sebelum tingkat kompetisi publikasi pada bidang ini diperkirakan meningkat tajam.

Analisis Sitasi

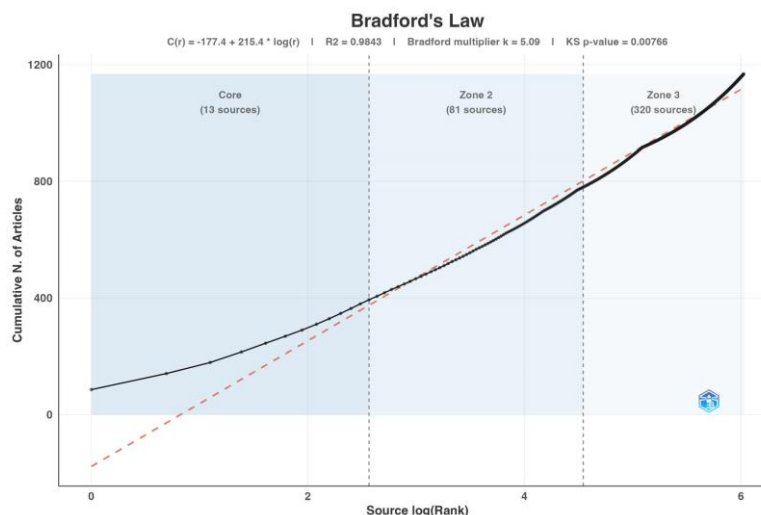


Gambar 4. Rata-Rata Sitasi per Artikel per Tahun (Average Article Citation per Year)

Gambar 4 disajikan untuk melihat pola dampak akademik korpus dari sudut pandang kohort tahun terbit, sebagai pelengkap terhadap indikator total sitasi dan rata-rata sitasi keseluruhan yang telah disajikan pada tabel audit dataset. Gambar ini menampilkan rata-rata jumlah sitasi per artikel yang dikelompokkan berdasarkan tahun terbit, sehingga memungkinkan perbandingan dampak antarkohort publikasi. Data menunjukkan rata-rata sitasi per artikel relatif stabil pada kisaran 6,7 hingga 6,8 untuk kohort tahun 2021 dan 2022, kemudian meningkat tajam mencapai puncak sekitar 8,5 pada kohort tahun 2023, sebelum akhirnya menurun tajam hingga sekitar 3,3 pada kohort tahun 2025. Pola naik-turun berbentuk segitiga ini berbeda dari pola tren produksi tahunan pada Gambar 1 yang cenderung meningkat monoton, sehingga kedua gambar tersebut perlu dibaca secara terpisah dan saling melengkapi.

Pola penurunan tajam pada kohort termuda, yaitu tahun 2024 dan 2025, bukan mengindikasikan penurunan kualitas maupun relevansi ilmiah artikel-artikel tersebut, melainkan merupakan efek jendela sitasi atau citation window effect yang lazim ditemukan dalam kajian bibliometrik. Artikel yang terbit lebih baru secara alami memiliki waktu yang jauh lebih singkat untuk memperoleh sitasi dibandingkan artikel kohort 2021 hingga 2023, mengingat proses sitasi memerlukan waktu publikasi, peninjauan sejawat, dan penerbitan pada artikel yang mengutipnya. Fenomena ini merupakan keterbatasan metodologis bawaan dari setiap analisis sitasi berbasis snapshot data pada satu titik waktu, bukan cerminan kualitas riset pada kohort tersebut. Oleh karena itu, perbandingan dampak sitasi antarkohort pada Gambar 4 sebaiknya difokuskan pada kohort 2021 hingga 2023 yang telah memiliki jendela waktu sitasi yang relatif memadai, sementara kohort 2024 dan 2025 memerlukan waktu observasi lebih panjang sebelum dapat dinilai secara adil.

Puncak rata-rata sitasi pada kohort tahun 2023 mengindikasikan bahwa artikel-artikel yang terbit pada tahun tersebut memuat sejumlah kontribusi metodologis yang berpengaruh luas terhadap perkembangan bidang, termasuk survei tentang deep representation learning untuk speech emotion recognition oleh Latif dkk. (2023) dan tinjauan arsitektur transformer pada pengenalan emosi ucapan oleh Wagner dkk. (2023) yang tercatat sebagai salah satu dokumen paling banyak disitasi dalam korpus ini. Temuan puncak pada kohort 2023 ini konsisten dengan pola umum pada literatur bibliometrik ilmu komputer, yang menunjukkan artikel berjenis tinjauan atau review dan artikel dengan kontribusi arsitektural baru cenderung memperoleh akumulasi sitasi yang lebih cepat dibandingkan studi aplikatif yang bersifat inkremental. Implikasi praktisnya bagi penulis yang ingin memaksimalkan dampak sitasi adalah pentingnya mempertimbangkan kontribusi berjenis tinjauan sistematis atau proposal arsitektur baru, bukan semata replikasi metode yang telah ada pada dataset berbeda. Implikasi bagi pembaca hasil bibliometrik ini adalah perlunya membaca indikator rata-rata sitasi per kohort secara kontekstual, bukan sebagai ukuran mutlak kualitas artikel semata.



Gambar 5. Kurva Hukum Bradford (Bradford's Law) yang membagi jurnal ke dalam tiga zona produktivitas.

Gambar 5 disajikan untuk mengidentifikasi jurnal-jurnal yang menjadi sumber inti (core sources) bagi bidang emotion recognition berbasis AI, mengikuti prinsip hukum Bradford tentang sebaran produktivitas jurnal ilmiah pada suatu bidang kajian. Kurva ini memetakan peringkat logaritmik jurnal (sumbu horizontal) terhadap jumlah kumulatif artikel (sumbu vertikal), dengan garis putus-putus vertikal menandai batas antarzona. Berdasarkan kurva tersebut, 415 jurnal sumber pada korpus ini terbagi ke dalam tiga zona, yaitu zona inti (core) yang berisi 13 jurnal, zona kedua yang berisi 81 jurnal, dan zona ketiga yang berisi 320 jurnal.

Nilai pengali Bradford (Bradford multiplier) tercatat sebesar $k = 5,09$ dengan kesesuaian model yang sangat tinggi, yaitu $R^2 = 0,9843$, menandakan pola sebaran jurnal pada korpus ini sangat konsisten dengan prediksi teoretis hukum Bradford.

Pola distribusi tiga zona ini menunjukkan karakteristik klasik hukum Bradford, yaitu sejumlah kecil jurnal inti menghasilkan proporsi artikel yang jauh lebih besar dibandingkan jumlah jurnalnya, sementara mayoritas jurnal pada zona ketiga masing-masing hanya menyumbang sedikit artikel pada korpus. Sebanyak 13 jurnal zona inti, atau hanya sekitar 3,1% dari total 415 jurnal, menyumbang porsi dokumen yang jauh melampaui proporsi jumlahnya, sebagaimana tercermin dari kurva kumulatif yang naik curam pada bagian awal grafik sebelum melandai pada zona kedua dan ketiga. Sebaliknya, zona ketiga yang berisi 320 jurnal atau sekitar 77% dari total jurnal hanya berkontribusi pada bagian landai kurva, mengindikasikan setiap jurnal pada zona ini rata-rata hanya menyumbang satu hingga dua artikel saja. Pola konsentrasi semacam ini lazim ditemukan pada bidang riset teknik terapan yang aktif, di mana penulis cenderung mengirimkan artikel ke jurnal-jurnal ternama dengan cakupan topik yang relevan secara konsisten.

Implikasi praktis dari temuan hukum Bradford ini adalah bahwa peneliti dan pustakawan yang ingin melakukan pemantauan literatur secara efisien pada bidang emotion recognition berbasis AI cukup memfokuskan perhatian pada 13 jurnal zona inti untuk menangkap sebagian besar perkembangan riset terkini, tanpa harus memantau seluruh 415 jurnal secara merata. Implikasi teoretisnya, nilai R^2 sebesar 0,9843 yang sangat tinggi menegaskan bahwa bidang emotion recognition berbasis AI telah menjadi bidang riset yang cukup matang dan terstruktur, dengan pola sebaran literatur yang stabil dan dapat diprediksi menggunakan model bibliometrik klasik. Identitas ketiga belas jurnal zona inti tersebut, beserta jumlah kontribusi dokumennya masing-masing, disajikan secara rinci pada tabel berikut untuk memudahkan interpretasi lebih lanjut. Penyajian tabel ini juga menjadi dasar bagi rekomendasi target jurnal publikasi yang akan dibahas pada bagian akhir sub-bagian ini.

Table 3. Audit awal terhadap 1.168

Jurnal (Zona Inti)	Jumlah Dokumen
IEEE Access	86
Multimedia Tools and Applications	55
Applied Sciences (Switzerland)	38
Sensors	36
Electronics (Switzerland)	30
Scientific Reports	24
IEEE Transactions on Affective Computing	21
International Journal of Advanced Computer Science and Applications	20
International Journal of Speech Technology	19
PeerJ Computer Science	18
Biomedical Signal Processing and Control	17
Neural Computing and Applications	16
Applied Acoustics	14

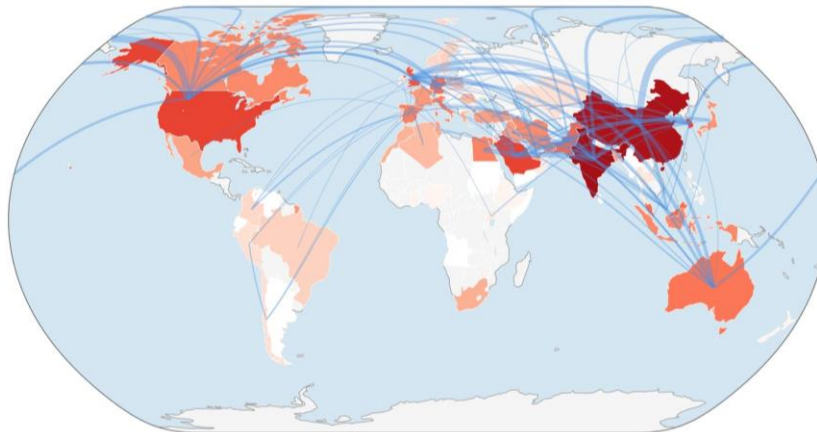
Tabel di atas menyajikan secara rinci identitas 13 jurnal zona inti beserta jumlah dokumen yang disumbangkan masing-masing pada korpus penelitian ini, diurutkan dari kontribusi tertinggi hingga terendah. IEEE Access menempati peringkat pertama dengan 86 dokumen, jauh melampaui jurnal kedua yaitu Multimedia Tools and Applications dengan 55 dokumen, diikuti Applied Sciences (Switzerland) dengan 38 dokumen, Sensors dengan 36 dokumen, dan Electronics (Switzerland) dengan 30 dokumen. Sepuluh jurnal peringkat berikutnya, yaitu Scientific Reports, IEEE Transactions on Affective Computing, International Journal of Advanced Computer Science and Applications, International Journal of Speech Technology, PeerJ Computer Science, Biomedical Signal Processing and Control, Neural Computing and Applications, dan Applied Acoustics, masing-masing menyumbang antara 14 hingga 24 dokumen. Sebaran jumlah dokumen antarjurnal zona inti ini menunjukkan penurunan yang cukup landai, berbeda dengan lonjakan tajam pada peringkat pertama, mengindikasikan IEEE Access secara khusus menjadi wadah publikasi dominan pada bidang ini.

Tabel ini menunjukkan bahwa jurnal-jurnal berorientasi teknik elektro-komputer dan sistem cerdas, yaitu IEEE Access, Sensors, dan Electronics, secara kolektif mendominasi zona inti, sementara IEEE Transactions on Affective Computing menjadi satu-satunya jurnal yang secara eksplisit berfokus pada domain afektif komputasi di antara 13 jurnal teratas tersebut. Pola dominasi jurnal-jurnal berpenerbit IEEE

dan MDPI ini konsisten dengan karakter riset emotion recognition sebagai bidang terapan teknik yang berorientasi pada prototipe sistem dan pengujian model performa, bukan semata kajian teoretis afektif komputasi murni. Kehadiran jurnal-jurnal berorientasi aplikasi medis dan sinyal biomedis, seperti Biomedical Signal Processing and Control dan Scientific Reports, juga mengindikasikan adanya irisan riset emotion recognition dengan domain kesehatan dan pemantauan fisiologis. Pola sebaran zona inti ini secara konsisten mencerminkan sifat multidisipliner bidang emotion recognition berbasis AI, yang mempertemukan teknik elektro, ilmu komputer, akustik, dan aplikasi kesehatan dalam satu ekosistem publikasi yang saling melengkapi.

Implikasi praktis dari temuan tabel ini bagi peneliti, termasuk peneliti Indonesia yang berorientasi publikasi Scopus, adalah pentingnya menargetkan jurnal-jurnal pada zona inti secara strategis, khususnya IEEE Access dan Sensors, yang selain berdampak tinggi juga dikenal menerapkan proses tinjauan sejawat yang relatif cepat untuk artikel bertema sistem cerdas terapan. Implikasi lain adalah bahwa peneliti yang berfokus pada aspek psikologis atau teoretis emosi, alih-alih aspek teknis semata, dapat mempertimbangkan IEEE Transactions on Affective Computing sebagai target jurnal yang lebih spesifik terhadap domain afektif komputasi. Dari sisi pengelola jurnal nasional Indonesia, temuan dominasi jurnal internasional pada zona inti ini juga dapat menjadi masukan strategis untuk membuka ruang lingkup bahasan afektif komputasi dan emotion recognition secara lebih eksplisit guna menangkap arus submisi yang terus tumbuh. Dengan demikian, tabel dan gambar pada sub-bagian hukum Bradford ini bukan hanya bersifat deskriptif, melainkan juga memiliki nilai strategis langsung bagi perencanaan publikasi ilmiah pada bidang ini.

Analisis Negara dan Kolaborasi



Gambar 6. Peta Kolaborasi Negara (Country Collaboration Map) berdasarkan afiliasi penulis

Table 4. Jumlah Dokumen Berdasarkan Afiliasi

Negara	Jumlah Dokumen
India	339
Tiongkok	267
Amerika Serikat	84
Arab Saudi	83
Korea Selatan	80
Inggris Raya	61
Pakistan	43
Malaysia	41
Jerman	38
Australia	35

Tabel dan Gambar 6 disajikan secara berdampingan untuk memetakan kontribusi produktivitas riset per negara sekaligus pola kolaborasi antarnegara pada bidang emotion recognition berbasis AI. Tabel menampilkan sepuluh negara paling produktif berdasarkan frekuensi kemunculan afiliasi penulis pada 1.168 dokumen, sementara Gambar 6 memvisualisasikan hal yang sama dalam bentuk peta dunia dengan gradasi warna menunjukkan intensitas produktivitas dan garis lengkung menunjukkan jalinan kolaborasi antarnegara. India menempati peringkat pertama dengan 339 dokumen, diikuti Tiongkok dengan 267 dokumen, kemudian Amerika Serikat dan Arab Saudi yang masing-masing berkontribusi 84 dan 83 dokumen, serta Korea Selatan dengan 80 dokumen. Lima negara berikutnya, yaitu Inggris Raya, Pakistan,

Malaysia, Jerman, dan Australia, masing-masing berkontribusi antara 35 hingga 61 dokumen, melengkapi gambaran sepuluh besar negara paling produktif pada bidang ini.

Tabel dan Gambar 6 secara bersama-sama menunjukkan bahwa India dan Tiongkok merupakan dua kontributor paling produktif yang jauh melampaui negara-negara lain, dengan selisih kontribusi yang signifikan terhadap peringkat ketiga. Pada peta kolaborasi, garis lengkung paling tebal tampak terpusat pada koridor yang menghubungkan India dengan kawasan Timur Tengah, serta koridor yang menghubungkan Tiongkok dengan kawasan Asia Tenggara dan Amerika Utara, mengindikasikan kedua negara ini juga menjadi simpul kolaborasi paling aktif, bukan sekadar produktif secara mandiri. Dari 1.168 dokumen dalam korpus, sebanyak 336 dokumen atau 28,8% melibatkan afiliasi penulis lintas negara, sementara sisanya sebanyak 832 dokumen atau 71,2% merupakan hasil kolaborasi dalam lingkup afiliasi domestik tunggal. Proporsi kolaborasi internasional yang kurang dari sepertiga total dokumen ini menjadi temuan penting yang membedakan pola kolaborasi pada bidang emotion recognition dari sejumlah bidang riset kecerdasan buatan lain yang cenderung lebih terglobalisasi.

Temuan dominasi India dan Tiongkok ini sejalan dengan tren umum produktivitas riset kecerdasan buatan global yang dilaporkan berbagai indeks bibliometrik lain, dan mengonfirmasi pergeseran pusat gravitasi riset AI terapan ke kawasan Asia dalam satu dekade terakhir. Tingkat kolaborasi internasional yang belum mencapai sepertiga dari total dokumen ini mengindikasikan adanya peluang sekaligus tantangan bagi perkembangan bidang ke depan. Di satu sisi, terdapat peluang bagi institusi di negara dengan produktivitas menengah, termasuk Indonesia yang tidak masuk sepuluh besar pada dataset ini, untuk membangun kolaborasi strategis dengan pusat riset di India, Tiongkok, atau Arab Saudi guna meningkatkan visibilitas dan dampak riset nasional. Di sisi lain, terdapat tantangan metodologis karena model emotion recognition yang dilatih pada satu konteks budaya dan bahasa berisiko rendah generalisasinya ketika diterapkan pada konteks budaya berbeda, sebuah isu yang secara eksplisit menjadi bagian dari agenda penelitian masa depan pada bagian akhir artikel ini.

Kata Kunci Dominan dan Peta Kata (Word Cloud)



Gambar 7. Word Cloud Kata Kunci Penulis (Author Keywords).

Gambar 7 disajikan untuk memberikan gambaran visual yang cepat dan intuitif mengenai istilah-istilah yang paling sering digunakan penulis dalam mendeskripsikan artikel mereka sendiri. Word cloud ini menampilkan visualisasi frekuensi kata kunci penulis (Author Keywords), dengan ukuran huruf setiap istilah dibuat proporsional terhadap frekuensi kemunculannya pada korpus 1.168 dokumen, sehingga istilah yang paling sering muncul tampil dengan ukuran huruf paling besar. Istilah deep learning, speech emotion recognition, dan emotion recognition tampak sebagai tiga istilah dengan ukuran huruf terbesar pada visualisasi ini, mengindikasikan ketiganya merupakan istilah yang paling dominan digunakan penulis. Istilah-istilah lain seperti machine learning, facial emotion recognition, convolutional neural network, dan artificial intelligence tampil dengan ukuran huruf yang lebih kecil namun masih cukup menonjol dibandingkan istilah-istilah pada lapisan terluar word cloud.

Untuk memberikan gambaran yang lebih presisi dan terukur secara kuantitatif dibandingkan interpretasi visual semata, frekuensi kemunculan sepuluh kata kunci penulis paling dominan disajikan secara rinci pada tabel berikut. Kata kunci deep learning tercatat muncul sebanyak 382 kali, menjadikannya istilah paling sering digunakan penulis pada korpus ini, diikuti speech emotion recognition dengan 316 kemunculan dan emotion recognition dengan 209 kemunculan. Kata kunci machine learning muncul 137 kali dan facial emotion recognition muncul 131 kali, keduanya juga tergolong sangat dominan meski berada di bawah tiga istilah teratas. Lima kata kunci berikutnya, yaitu convolutional neural network, artificial intelligence,

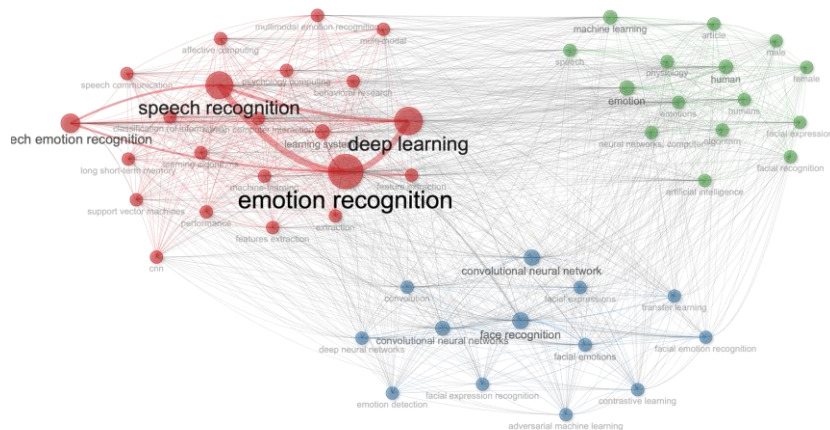
affective computing, multimodal emotion recognition, dan cnn, masing-masing muncul antara 58 hingga 90 kali, melengkapi sepuluh besar kata kunci penulis paling dominan pada tabel di bawah ini.

Table 5. Jumlah Dokumen Berdasarkan Affiliasi

Kata Kunci Penulis	Frekuensi
deep learning	382
speech emotion recognition	316
emotion recognition	209
machine learning	137
facial emotion recognition	131
convolutional neural network	90
artificial intelligence	84
affective computing	75
multimodal emotion recognition	71
cnn	58

Dominasi kata kunci speech emotion recognition dibandingkan facial emotion recognition, dengan selisih 316 berbanding 131 kemunculan, menunjukkan bahwa pada rentang 2021 hingga 2025, riset pengenalan emosi berbasis suara secara kuantitatif jauh lebih intensif digarap dibandingkan riset berbasis wajah dalam korpus yang menautkan istilah kecerdasan buatan secara eksplisit. Temuan ini menegaskan pentingnya model akustik dan representasi ucapan, seperti fitur Mel-frequency cepstral coefficients (MFCC) dan model self-supervised berbasis wav2vec, sebagai arus utama riset emotion recognition pada periode kajian ini. Posisi kata kunci convolutional neural network dan cnn yang tetap masuk sepuluh besar meski peringkatnya di bawah istilah-istilah terkait ucapan juga mengindikasikan bahwa arsitektur ini tetap menjadi tulang punggung metodologis lintas modalitas, baik untuk pengolahan citra wajah maupun spektrogram suara. Secara keseluruhan, pola frekuensi kata kunci pada tabel ini memberikan dasar kuantitatif yang akan dikonfirmasi lebih lanjut melalui analisis jejaring co-occurrence dan peta tematik pada dua sub-bagian berikutnya.

Jejaring Co-occurrence Kata Kunci (VOSviewer)



Gambar 8. Jejaring Co-occurrence Kata Kunci (VOSviewer), tervisualisasi dalam tiga klaster warna (merah, hijau, biru).

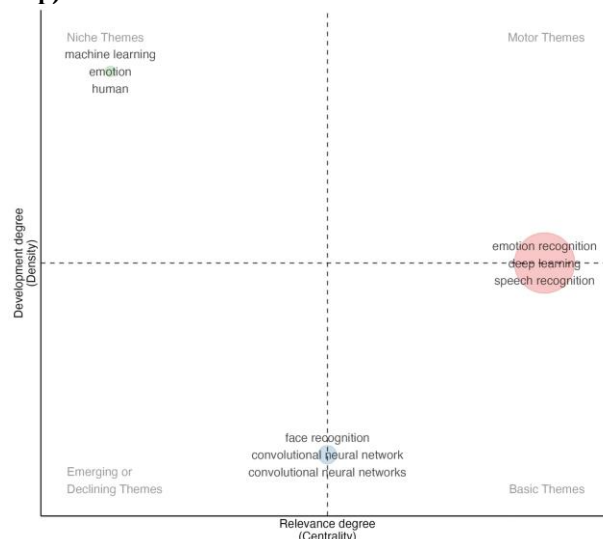
Gambar 8 disajikan untuk memperdalam analisis kata kunci dari sekadar frekuensi kemunculan menjadi pola keterhubungan antaristilah, sehingga struktur intelektual bidang emotion recognition berbasis AI dapat dipetakan secara lebih komprehensif. Visualisasi VOSviewer pada Gambar 8 menampilkan jejaring co-occurrence kata kunci yang secara visual terbagi menjadi tiga klaster berdasarkan warna node, di mana setiap node merepresentasikan satu istilah dan ukuran node mencerminkan frekuensi kemunculannya, sementara garis penghubung antarnode merepresentasikan kemunculan bersama (co-occurrence) pada dokumen yang sama. Klaster merah berpusat pada node emotion recognition, deep learning, dan speech recognition, dengan keterhubungan erat pada istilah speech emotion recognition, affective computing, multimodal emotion recognition, learning algorithms, dan cnn, yang secara keseluruhan mengindikasikan klaster ini merepresentasikan arus utama riset pengenalan emosi berbasis ucapan dan pembelajaran mendalam. Klaster biru berpusat pada face recognition dan convolutional neural network, terhubung dengan istilah facial

expression recognition, facial emotions, deep neural networks, transfer learning, dan contrastive learning, yang secara keseluruhan merepresentasikan arus riset pengenalan emosi berbasis wajah dan visi komputer.

Klaster ketiga, yaitu klaster hijau, berisi node deskriptor demografis dan metadata dokumen seperti human, emotion, machine learning, article, male, female, dan physiology, yang secara karakteristik lebih mencerminkan metadata subjek penelitian berupa partisipan manusia dan indeks keyword umum dibandingkan tema metodologis yang spesifik. Perbedaan karakter ketiga klaster ini penting untuk dicermati karena klaster merah dan biru merepresentasikan pilihan metodologis aktif yang dibuat peneliti, yaitu modalitas ucapan versus modalitas wajah, sementara klaster hijau lebih merupakan konsekuensi otomatis dari indeksasi basis data Scopus terhadap istilah subjek penelitian seperti human dan physiology yang secara rutin disematkan pada artikel bertema penelitian manusia. Posisi node emotion recognition dan deep learning yang berada tepat di antara klaster merah dan menjadi penghubung utama ke klaster biru serta hijau menunjukkan bahwa kedua istilah ini berfungsi sebagai jembatan konseptual yang menyatukan seluruh korpus penelitian, bukan hanya milik satu klaster metodologis tertentu. Pola keterhubungan lintas klaster semacam ini merupakan ciri khas bidang riset yang telah terkonsolidasi secara konseptual meski masih terdiferensiasi secara metodologis pada tingkat teknik implementasi.

Jumlah node, klaster, dan pola keterhubungan yang disebutkan di atas didasarkan langsung pada observasi visual terhadap Gambar 8 sebagaimana dihasilkan oleh perangkat lunak VOSviewer, tanpa penambahan angka yang tidak dapat diverifikasi. Nilai total link strength dan jumlah tautan atau links secara numerik sengaja tidak dicantumkan pada pembahasan ini karena berkas data mentah jejaring, yaitu berkas network.txt atau berkas peta VOSviewer, tidak diunggah dalam penelitian ini, sehingga angka tersebut tidak dapat diverifikasi dan sengaja tidak disajikan sesuai prinsip kehati-hatian data yang menjadi pedoman penelitian ini. Node emotion recognition dan deep learning tampak sebagai node dominan dengan ukuran lingkaran terbesar pada klaster merah, mengindikasikan derajat keterhubungan atau degree dan frekuensi kemunculan tertinggi dalam jejaring, konsisten dengan hasil frekuensi kata kunci pada tabel word cloud sebelumnya. Konsistensi antara temuan frekuensi kata kunci dan temuan jejaring co-occurrence ini memperkuat validitas keseluruhan analisis kata kunci pada penelitian ini, karena kedua teknik analisis yang berbeda menghasilkan simpulan yang saling mendukung.

Peta Tematik (Thematic Map)



Gambar 9. Peta Tematik (Thematic Map) berdasarkan sentralitas (relevance degree) dan densitas (development degree).

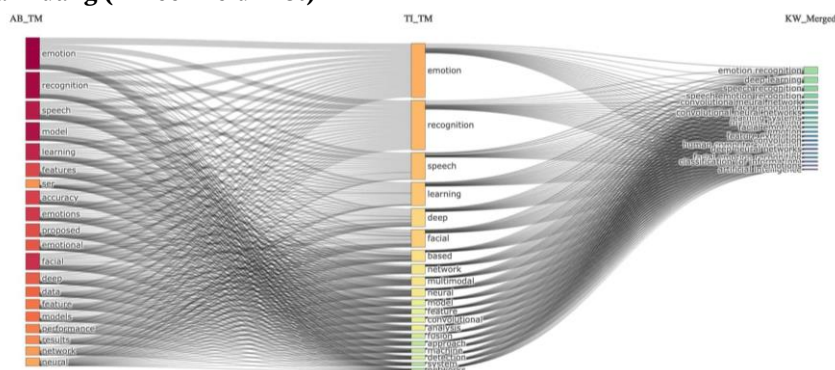
Gambar 9 disajikan untuk menguji dan mengonfirmasi struktur klaster yang telah teridentifikasi pada jejaring co-occurrence Gambar 8, dengan menggunakan kerangka analisis dua dimensi berupa sentralitas atau relevance degree pada sumbu horizontal dan densitas atau development degree pada sumbu vertikal. Kerangka strategic diagram ini membagi bidang keilmuan ke dalam empat kuadran, yaitu motor themes pada kuadran kanan-atas, basic themes pada kuadran kanan-bawah, niche themes pada kuadran kiri-atas, dan emerging or declining themes pada kuadran kiri-bawah. Klaster emotion recognition, deep learning, dan speech recognition berada pada kuadran kanan-tengah, tepat pada batas antara motor themes dan basic themes, dengan sentralitas tinggi dan densitas sedang, menandakan tema ini merupakan tema inti yang telah

berkembang dan sangat relevan dengan keseluruhan bidang penelitian. Posisi kluster ini yang berada di sisi kanan peta tematik konsisten dengan perannya sebagai node penghubung utama pada jejaring co-occurrence yang telah dibahas sebelumnya.

Klaster face recognition, convolutional neural network, dan convolutional neural networks berada pada kuadran bawah-tengah peta tematik, dengan sentralitas sedang namun densitas rendah, mengindikasikan tema ini penting bagi keseluruhan bidang namun belum berkembang secara internal sebagai sub-bidang yang matang dan kohesif. Posisi ini berpeluang berubah menjadi tema berkembang atau bahkan motor themes jika mendapat perhatian riset lebih lanjut yang secara khusus memperdalam adaptasi arsitektur CNN untuk konteks pengenalan emosi, bukan sekadar penerapan CNN generik dari domain visi komputer lain. Klaster machine learning, emotion, dan human berada pada kuadran kiri-atas atau niche themes, dengan sentralitas rendah namun densitas tinggi, mengindikasikan klaster ini berkembang matang secara internal namun kurang terhubung dengan tema-tema metodologis lain pada bidang ini. Karakter niche themes pada klaster ini konsisten dengan sifatnya sebagai kata kunci deskriptor umum dan metadata subjek penelitian yang telah dibahas pada interpretasi jejaring co-occurrence sebelumnya, bukan sebagai arah metodologis yang aktif dikembangkan peneliti.

Interpretasi gabungan antara Gambar 8 dan Gambar 9 memberikan implikasi teoretis bahwa struktur intelektual bidang emotion recognition berbasis AI, dalam kerangka strategic diagram, saat ini masih berpusat pada dua tema motor utama, yaitu pengenalan emosi berbasis ucapan menggunakan deep learning dan pengenalan emosi berbasis wajah menggunakan convolutional neural network, dan belum menunjukkan klaster tematik baru yang matang secara tersendiri untuk topik-topik seperti pengenalan emosi berbasis teks atau pengenalan emosi berbasis sinyal fisiologis dan EEG, meskipun istilah-istilah tersebut turut digunakan dalam boolean pencarian penelitian ini. Kondisi ini mengindikasikan bahwa riset text emotion recognition dan physiological emotion recognition dalam korpus ini masih menyatu sebagai bagian dari klaster multimodal yang lebih besar, belum cukup besar volumenya untuk membentuk klaster independen pada peta tematik. Implikasi praktis dari temuan ini adalah terbukanya peluang riset pada dua modalitas yang belum matang tersebut, sebagaimana akan dirumuskan lebih lanjut pada bagian agenda penelitian masa depan. Implikasi metodologis bagi kajian bibliometrik lanjutan adalah pentingnya memperbesar cakupan tahun kajian atau memperluas boolean pencarian secara khusus pada istilah text emotion dan EEG emotion recognition apabila peneliti ingin memetakan sub-bidang tersebut secara lebih mendalam dan terpisah.

Analisis Tiga Bidang (Three-Field Plot)



Gambar 10. Three-Field Plot yang menautkan kata kunci teks abstrak (AB_TM, kiri), kata kunci judul (TI_TM, tengah), dan kata kunci gabungan (KW_Merged, kanan).

Gambar 10 disajikan untuk menguji konsistensi tematik korpus data dari tiga sumber teks yang berbeda secara bersamaan, yaitu teks abstrak, teks judul, dan kata kunci gabungan, menggunakan format diagram Sankey yang menautkan ketiganya melalui pita penghubung. Diagram ini terbagi menjadi tiga kolom, yaitu kolom AB_TM di sisi kiri yang berisi istilah dominan hasil penambangan teks abstrak seperti emotion, recognition, speech, model, dan learning, kolom TI_TM di tengah yang berisi istilah dominan hasil penambangan teks judul seperti emotion, recognition, speech, learning, dan deep, serta kolom KW_Merged di sisi kanan yang berisi kata kunci gabungan seperti emotion recognition, deep learning, speech recognition, speech emotion recognition, dan convolutional neural network. Lebar pita penghubung antarkolom pada diagram ini merepresentasikan kekuatan keterkaitan antaristilah pada ketiga sumber teks tersebut, sehingga semakin lebar suatu pita, semakin kuat pula keterkaitan istilah yang dihubungkannya.

Kesesuaian tinggi antara terminologi pada abstrak, judul, dan kata kunci yang tervisualisasi pada Gambar 10 menunjukkan konsistensi tematik yang kuat pada korpus data penelitian ini, yaitu bahwa istilah yang digunakan penulis pada judul artikel secara konsisten mencerminkan isi abstrak dan kata kunci yang mereka pilih sendiri. Istilah *emotion* dan *recognition* secara khusus tampak menjadi simpul penghubung terbesar pada ketiga kolom, mengindikasikan kedua istilah ini merupakan inti konseptual yang menyatukan seluruh korpus data, sejalan dengan temuan pada analisis kata kunci dan jejaring co-occurrence pada sub-bagian sebelumnya. Konsistensi tiga arah ini juga secara tidak langsung memperkuat validitas internal strategi pencarian boolean yang digunakan dalam penelitian ini, karena artikel-artikel yang terjaring oleh boolean tersebut secara nyata membahas topik emosi dan pengenalannya, bukan sekadar menyinggung istilah tersebut secara insidental pada bagian teks yang tidak relevan. Dengan demikian, diagram tiga bidang ini berfungsi sebagai bentuk validasi tambahan terhadap ketepatan cakupan tematik korpus data, melengkapi validasi kuantitatif jumlah dokumen yang telah dibahas pada bagian Metode Penelitian.

Secara keseluruhan, temuan bibliometrik pada penelitian ini menegaskan bahwa riset *emotion recognition* berbasis AI tahun 2021–2025 berada pada fase pertumbuhan aktif dengan proyeksi puncak siklus hidup topik pada akhir 2026, ditandai oleh pertumbuhan publikasi tahunan yang konsisten dan struktur tematik yang masih terpusat pada dua arus utama, yaitu pengenalan emosi berbasis ucapan menggunakan pembelajaran mendalam, dan pengenalan emosi berbasis wajah menggunakan jaringan saraf konvolusional. Pola ini selaras dengan tinjauan literatur oleh Mehrish dkk. (2023) yang mendokumentasikan pergeseran luas teknik pemrosesan ucapan, termasuk pengenalan emosi, ke arah arsitektur pembelajaran mendalam berbasis representasi (*representation learning*), serta dengan temuan Wagner dkk. (2023) yang menegaskan bahwa arsitektur transformer mulai menutup kesenjangan performa (*valence gap*) yang sebelumnya menjadi kelemahan model *speech emotion recognition* berbasis fitur akustik konvensional.

Dominasi tema *speech emotion recognition* atas *facial emotion recognition* dalam frekuensi kata kunci (316 berbanding 131) berbeda dari kesan umum pada literatur *affective computing* periode sebelumnya yang lebih banyak menyoroti pengenalan emosi wajah sebagai modalitas utama karena ketersediaan basis data citra wajah berlabel emosi yang lebih awal matang, seperti FER-2013 dan AffectNet. Pergeseran penekanan ke arah ucapan pada periode 2021–2025 dapat dijelaskan oleh dua faktor: pertama, meluasnya ketersediaan model *self-supervised* berbasis ucapan seperti wav2vec 2.0 dan HuBERT yang memudahkan ekstraksi representasi ucapan berkualitas tinggi tanpa anotasi emosi berskala besar; kedua, kebutuhan aplikatif yang tinggi pada domain asisten suara, pusat panggilan, dan kendaraan otonom yang secara alami berbasis sinyal audio, sebagaimana tercermin pada studi Tan dkk. (2022) tentang penerapan *speech emotion recognition* untuk efisiensi lalu lintas pada kendaraan otonom.

Temuan bahwa kluster *convolutional neural network* dan *face recognition* berada pada kuadran densitas rendah pada peta tematik, meski memiliki sentralitas tinggi, secara teoretis dapat dimaknai melalui kerangka siklus hidup tema (*theme life cycle*) dalam analisis bibliometrik: tema dengan sentralitas tinggi namun densitas rendah cenderung merupakan tema transversal (*cross-cutting*) yang menopang banyak sub-bidang lain tanpa membentuk komunitas riset internal yang kohesif tersendiri. Hal ini konsisten dengan karakter CNN sebagai arsitektur tulang punggung (*backbone*) yang digunakan lintas tugas visi komputer, bukan hanya pada *emotion recognition*, sehingga riset yang secara spesifik memperdalam adaptasi arsitektur CNN untuk pengenalan emosi wajah misalnya melalui mekanisme atensi khusus ekspresi mikro sebagaimana diusulkan Wang dkk. (2024) pada HTNet untuk pengenalan *micro-expression* masih relatif sedikit dibandingkan riset penerapan CNN generik untuk klasifikasi emosi.

Tingkat kolaborasi internasional yang tercatat sebesar 28,8% dari total dokumen menunjukkan pola yang berbeda dari bidang riset kecerdasan buatan lain yang cenderung lebih terglobalisasi, seperti pemrosesan bahasa alami skala besar. Rendahnya proporsi kolaborasi lintas negara pada riset *emotion recognition* kemungkinan berkaitan dengan sifat data emosi yang sensitif secara budaya dan bahasa model yang dilatih pada korpus ucapan atau ekspresi wajah suatu populasi tidak selalu dapat digeneralisasi secara langsung ke populasi lain tanpa adaptasi data lokal, sehingga banyak tim riset cenderung bekerja dengan data dan mitra domestik. Studi survei sistematis Kalateh dkk. (2024) tentang *multimodal emotion recognition* secara eksplisit mencatat generalisasi lintas dataset dan lintas budaya sebagai tantangan terbuka utama bidang ini, yang sejalan dengan pola kolaborasi domestik-dominan yang ditemukan dalam penelitian bibliometrik ini.

Implikasi teoretis dari keseluruhan temuan ini adalah bahwa struktur intelektual *emotion recognition* berbasis AI, dalam periode 2021–2025, dapat dipahami sebagai bidang yang mengalami konvergensi metodologis (pembelajaran mendalam sebagai payung metode dominan) namun belum mengalami konvergensi data (basis data dan konteks budaya yang masih terfragmentasi antarnegara). Implikasi praktisnya bagi pengembang sistem adalah perlunya strategi *transfer learning* dan *domain adaptation* yang

lebih eksplisit ketika model emotion recognition dipindahkan dari satu populasi ke populasi lain, sebagaimana telah dirintis oleh Luna-Jiménez dkk. (2021) pada penerapan transfer learning untuk basis data RAVDESS. Bagi pengelola kebijakan riset dan pendanaan, implikasinya adalah pentingnya insentif khusus untuk mendorong kolaborasi riset lintas negara pada domain emotion recognition, mengingat proporsi kolaborasi internasional yang masih di bawah sepertiga total publikasi berpotensi memperlambat generalisasi model secara global.

Conclusions

Penelitian ini memetakan struktur bibliometrik riset emotion recognition berbasis kecerdasan buatan pada 1.168 artikel jurnal terindeks Scopus periode 2021–2025. Produksi ilmiah tahunan meningkat konsisten dari 102 menjadi 368 dokumen, dengan proyeksi puncak siklus hidup topik pada akhir tahun 2026 ($R^2 = 0,987$). India dan Tiongkok merupakan kontributor paling produktif, sementara IEEE Access, Multimedia Tools and Applications, dan Applied Sciences menjadi jurnal inti berdasarkan hukum Bradford (13 jurnal inti dari 415 jurnal, $R^2 = 0,9843$). Struktur tematik bidang ini terpusat pada dua tema motor utama speech emotion recognition berbasis deep learning, dan face recognition berbasis convolutional neural network dengan tema convolutional neural network berada pada tahap belum matang secara internal meski memiliki sentralitas tinggi. Kolaborasi internasional baru mencakup 28,8% dari total dokumen, mengindikasikan riset bidang ini masih didominasi kolaborasi domestik.

References

- Abdelhamid, A. A., El-Kenawy, E.-S. M., Alotaibi, B., Amer, G. M., Abdelkader, M. Y., Ibrahim, A., & Eid, M. M. (2022). Robust speech emotion recognition using CNN+LSTM based on stochastic fractal search optimization algorithm. IEEE Access, 10. <https://doi.org/10.1109/ACCESS.2022.3172954>
- Aggarwal, A., Srivastava, A., Agarwal, A., Chahal, N., Singh, D., Alnuaim, A. A., Alhadlaq, A., & Lee, H.-N. (2022). Two-way feature extraction for speech emotion recognition using deep learning. Sensors, 22(6). <https://doi.org/10.3390/s22062378>
- Alluhaidan, A. S., Saidani, O., Jahangir, R., Nauman, M. A., & Neffati, O. S. (2023). Speech emotion recognition through hybrid features and convolutional neural network. Applied Sciences (Switzerland), 13(8). <https://doi.org/10.3390/app13084750>
- Alnuaim, A. A., Zakariah, M., Shukla, P. K., Alhadlaq, A., Hatamleh, W. A., Tarazi, H., Sureshbabu, R., & Ratna, R. (2022). Human–computer interaction for recognizing speech emotions using multilayer perceptron classifier. Journal of Healthcare Engineering, 2022. <https://doi.org/10.1155/2022/6005446>
- Ancilin, J., & Milton, A. (2021). Improved speech emotion recognition with Mel frequency magnitude coefficient. Applied Acoustics, 179. <https://doi.org/10.1016/j.apacoust.2021.108046>
- Bhangale, K., & Kothandaraman, M. (2023). Speech emotion recognition based on multiple acoustic features and deep convolutional neural network. Electronics (Switzerland), 12(4). <https://doi.org/10.3390/electronics12040839>
- Chatterjee, R., Mazumdar, S., Sherratt, R. S., Halder, R., Maitra, T., & Giri, D. (2021). Real-time speech emotion analysis for smart home assistants. IEEE Transactions on Consumer Electronics, 67(1). <https://doi.org/10.1109/TCE.2021.3056421>
- Chaudhari, A., Bhatt, C., Krishna, A., & Mazzeo, P. L. (2022). ViTFER: Facial emotion recognition with vision transformers. Applied System Innovation, 5(4). <https://doi.org/10.3390/asi5040080>
- Chowdary, M. K., Nguyen, T. N., & Hemanth, D. J. (2023). Deep learning-based facial emotion recognition for human–computer interaction applications. Neural Computing and Applications, 35. <https://doi.org/10.1007/s00521-021-06012-8>
- Dalvi, C., Rathod, M., Patil, S., Gite, S., & Kotecha, K. (2021). A survey of AI-based facial emotion recognition: Features, ML & DL techniques, age-wise datasets and future directions. IEEE Access, 9. <https://doi.org/10.1109/ACCESS.2021.3131733>
- de Lope, J., & Graña, M. (2023). An ongoing review of speech emotion recognition. Neurocomputing, 528. <https://doi.org/10.1016/j.neucom.2023.01.002>
- Gupta, S., Kumar, P., & Tekchandani, R. K. (2023). Facial emotion recognition based real-time learner engagement detection system in online learning context using deep learning models. Multimedia Tools and Applications, 82. <https://doi.org/10.1007/s11042-022-13558-9>

- Gursesli, M. C., Lombardi, S., Duradoni, M., Bocchi, L., Guazzini, A., & Lanata, A. (2024). Facial emotion recognition (FER) through custom lightweight CNN model. *IEEE Access*, 12. <https://doi.org/10.1109/ACCESS.2024.3380847>
- Hema, C., & Garcia Marquez, F. P. (2023). Emotional speech recognition using CNN and deep learning techniques. *Applied Acoustics*, 211. <https://doi.org/10.1016/j.apacoust.2023.109492>
- Huang, Z.-Y., Chiang, C.-C., Chen, J.-H., Chen, Y.-C., Chung, H.-L., Cai, Y.-P., & Hsu, H.-C. (2023). A study on computer vision for facial emotion recognition. *Scientific Reports*, 13. <https://doi.org/10.1038/s41598-023-35446-4>
- Jahangir, R., Teh, Y. W., Hanif, F., & Mujtaba, G. (2021). Deep learning approaches for speech emotion recognition: State of the art and research challenges. *Multimedia Tools and Applications*, 80. <https://doi.org/10.1007/s11042-020-09874-7>
- Kalateh, S., Estrada-Jimenez, L. A., Nikghadam-Hojjati, S., & Barata, J. (2024). A systematic review on multimodal emotion recognition: Building blocks, current state, applications, and challenges. *IEEE Access*, 12. <https://doi.org/10.1109/ACCESS.2024.3430850>
- Karnati, M., Seal, A., Bhattacharjee, D., Yazidi, A., & Krejcar, O. (2023). Understanding deep learning techniques for recognition of human emotions using facial expressions: A comprehensive survey. *IEEE Transactions on Instrumentation and Measurement*, 72. <https://doi.org/10.1109/TIM.2023.3243661>
- Khan, M., Gueaieb, W., El Saddik, A., & Kwon, S. (2024). MSER: Multimodal speech emotion recognition using cross-attention with deep fusion. *Expert Systems with Applications*, 245. <https://doi.org/10.1016/j.eswa.2023.122946>
- Latif, S., Rana, R., Khalifa, S., Jurdak, R., Qadir, J., & Schuller, B. (2023). Survey of deep representation learning for speech emotion recognition. *IEEE Transactions on Affective Computing*, 14(2). <https://doi.org/10.1109/TAFFC.2021.3114365>
- Li, J., Wang, X., Lv, G., & Zeng, Z. (2023). GraphMFT: A graph network based multimodal fusion technique for emotion recognition in conversation. *Neurocomputing*, 550. <https://doi.org/10.1016/j.neucom.2023.126427>
- Liu, H., Lou, T., Zhang, Y., Wu, Y., Xiao, Y., Jensen, C. S., & Zhang, D. (2024). EEG-based multimodal emotion recognition: A machine learning perspective. *IEEE Transactions on Instrumentation and Measurement*, 73. <https://doi.org/10.1109/TIM.2024.3369130>
- Liu, S., Gao, P., Li, Y., Fu, W., & Ding, W. (2023). Multi-modal fusion network with complementarity and importance for emotion recognition. *Information Sciences*, 619. <https://doi.org/10.1016/j.ins.2022.11.076>
- Liu, W., Qiu, J.-L., Zheng, W.-L., & Lu, B.-L. (2022). Comparing recognition performance and robustness of multimodal deep learning models for multimodal emotion recognition. *IEEE Transactions on Cognitive and Developmental Systems*, 14(2). <https://doi.org/10.1109/TCDS.2021.3071170>
- Luna-Jiménez, C., Griol, D., Callejas, Z., Kleinlein, R., Montero, J. M., & Fernández-Martínez, F. (2021). Multimodal emotion recognition on RAVDESS dataset using transfer learning. *Sensors*, 21(22). <https://doi.org/10.3390/s21227665>
- Mai, S., Zeng, Y., & Hu, H. (2023). Multimodal information bottleneck: Learning minimal sufficient unimodal and multimodal representations. *IEEE Transactions on Multimedia*, 25. <https://doi.org/10.1109/TMM.2022.3171679>
- Mehrish, A., Majumder, N., Bharadwaj, R., Mihalcea, R., & Poria, S. (2023). A review of deep learning techniques for speech processing. *Information Fusion*, 99. <https://doi.org/10.1016/j.inffus.2023.101869>
- Middya, A. I., Nag, B., & Roy, S. (2022). Deep learning based multimodal emotion recognition using model-level fusion of audio-visual modalities. *Knowledge-Based Systems*, 244. <https://doi.org/10.1016/j.knosys.2022.108580>
- Mocanu, B., Tapu, R., & Zaharia, T. (2023). Multimodal emotion recognition using cross modal audio-video fusion with attention and deep metric learning. *Image and Vision Computing*, 133. <https://doi.org/10.1016/j.imavis.2023.104676>
- Mustaqeem, & Kwon, S. (2021). Att-Net: Enhanced emotion recognition system using lightweight self-attention module. *Applied Soft Computing*, 102. <https://doi.org/10.1016/j.asoc.2021.107101>
- Rawal, N., & Stock-Homburg, R. M. (2022). Facial emotion expressions in human-robot interaction: A survey. *International Journal of Social Robotics*, 14. <https://doi.org/10.1007/s12369-022-00867-0>
- Saleem Abdullah, S. M., Ameen, S. Y., Sadeeq, M. A. M., & Zeebaree, S. R. M. (2021). Multimodal emotion recognition using deep learning. *Journal of Applied Science and Technology Trends*, 2(2). <https://doi.org/10.38094/jastt20291>

- Talaat, F. M. (2023). Real-time facial emotion recognition system among children with autism based on deep learning and edge computing. *Neural Computing and Applications*, 35. <https://doi.org/10.1007/s00521-023-08372-9>
- Tan, L., Yu, K., Lin, L., Cheng, X., Srivastava, G., Lin, J. C.-W., & Wei, W. (2022). Speech emotion recognition enhanced traffic efficiency solution for autonomous vehicles in a 5G-enabled space-air-ground integrated intelligent transportation system. *IEEE Transactions on Intelligent Transportation Systems*, 23(3). <https://doi.org/10.1109/TITS.2021.3119921>
- Wagner, J., Triantafyllopoulos, A., Wierstorf, H., Schmitt, M., Burkhardt, F., Eyben, F., & Schuller, B. W. (2023). Dawn of the transformer era in speech emotion recognition: Closing the valence gap. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(9). <https://doi.org/10.1109/TPAMI.2023.3263585>
- Wang, S., Qu, J., Zhang, Y., & Zhang, Y. (2023). Multimodal emotion recognition from EEG signals and facial expressions. *IEEE Access*, 11. <https://doi.org/10.1109/ACCESS.2023.3263670>
- Wang, Z., Zhang, K., Luo, W., & Sankaranarayana, R. (2024). HTNet for micro-expression recognition. *Neurocomputing*, 602. <https://doi.org/10.1016/j.neucom.2024.128196>
- Wu, X., Zheng, W.-L., Li, Z., & Lu, B.-L. (2022). Investigating EEG-based functional connectivity patterns for multimodal emotion recognition. *Journal of Neural Engineering*, 19(1). <https://doi.org/10.1088/1741-2552/ac49a7>
- Yadav, S. P., Zaidi, S., Mishra, A., & Yadav, V. (2022). Survey on machine learning in speech emotion recognition and vision systems using a recurrent neural network (RNN). *Archives of Computational Methods in Engineering*, 29. <https://doi.org/10.1007/s11831-021-09647-x>
- Zehra, W., Javed, A. R., Jalil, Z., Khan, H. U., & Gadekallu, T. R. (2021). Cross corpus multi-lingual speech emotion recognition using ensemble learning. *Complex & Intelligent Systems*, 7. <https://doi.org/10.1007/s40747-020-00250-4>