**SYSTEMATIC LITERATURE REVIEW PERKEMBANGAN METODE COMPUTER  
VISION PADA PENGOLAHAN CITRA TAHUN 2017–2025****Muhammad Arief Akbar<sup>1)</sup>, Fauzan Azim<sup>2\*)</sup>**<sup>1,2</sup>Fakultas Ilmu Komputer, Universitas Muhammadiyah Riau  
email: <sup>1</sup>220401257@student.umri.ac.id, <sup>2</sup>fauzanazim17@umri.ac.id

\*Corresponding Author

**Abstract**

*The development of Computer Vision technology in recent years has shown significant growth along with advances in Deep Learning methods and the availability of large-scale datasets. Numerous studies have produced various approaches, architectures, and evaluation metrics, creating the need for a structured mapping to comprehensively understand the direction of this field. This study aims to analyze methodological trends, research task focuses, and dominant evaluation metrics in Computer Vision research. The method employed is a Systematic Literature Review (SLR) of 20 scientific article published between 2017 and 2025. The analysis process was conducted through data extraction covering method types, task focuses, datasets, and evaluation metrics used in each study. The results indicate that Convolutional Neural Networks and Vision Transformers are the most dominant architectures, with the primary research focuses on object detection, image classification, and video understanding. The most frequently used evaluation metrics are accuracy, mean Average Precision (mAP), and Intersection over Union (IoU). These findings reveal a gradual shift from convolution-based approaches toward transformer-based architectures that are more adaptive to large-scale visual data. This study provides a comprehensive overview of the development direction of Computer Vision and can serve as a reference for future research in selecting relevant methods and research focuses, both in terms of accuracy-oriented performance and computational efficiency.*

**Keywords:** Computer Vision, Systematic Literature Review, Deep Learning, Method Trends, Evaluation Metrics

**Abstrak**

Perkembangan teknologi Computer Vision dalam beberapa tahun terakhir menunjukkan peningkatan yang signifikan seiring dengan kemajuan metode Deep Learning dan ketersediaan dataset berskala besar. Beragam penelitian telah menghasilkan berbagai pendekatan, arsitektur, serta metrik evaluasi yang berbeda sehingga diperlukan suatu pemetaan untuk memahami arah perkembangan bidang ini secara menyeluruh. Penelitian ini bertujuan untuk menganalisis tren metode, fokus tugas, serta metrik evaluasi yang dominan dalam penelitian Computer Vision. Metode yang digunakan adalah Systematic Literature Review (SLR) terhadap 20 artikel ilmiah yang dipublikasikan pada rentang tahun 2017–2025. Proses analisis dilakukan melalui ekstraksi data yang meliputi jenis metode, fokus tugas, dataset, serta metrik evaluasi yang digunakan. Hasil penelitian menunjukkan bahwa arsitektur Convolutional Neural Network dan Vision Transformer menjadi metode yang paling dominan, dengan fokus penelitian terbesar pada object detection, image classification, dan video understanding. Metrik evaluasi yang paling sering digunakan adalah accuracy, mean Average Precision (mAP), dan Intersection over Union (IoU). Temuan ini menunjukkan adanya pergeseran tren dari metode konvolusional menuju arsitektur berbasis transformer yang lebih adaptif terhadap data visual berskala besar. Kesimpulan penelitian ini memberikan gambaran komprehensif mengenai arah perkembangan Computer Vision serta dapat menjadi referensi bagi penelitian selanjutnya dalam menentukan metode dan fokus kajian yang relevan, baik pada pengembangan model yang berorientasi akurasi maupun efisiensi komputasi.

**Kata Kunci:** Computer Vision, Studi Literatur, Deep Learning, Tren Metode, Evaluasi Kinerja

## PENDAHULUAN

Computer Vision merupakan salah satu cabang utama dalam Artificial Intelligence yang berfokus pada pengembangan metode dan algoritma agar komputer mampu memahami serta menafsirkan informasi visual berupa citra dan video secara otomatis. Perkembangan bidang ini mengalami peningkatan signifikan seiring kemajuan metode Deep Learning, khususnya melalui arsitektur Convolutional Neural Networks (CNN) yang terbukti mampu meningkatkan performa pada berbagai tugas visual seperti image classification, object detection, dan semantic segmentation (Zhao et al., 2024).

Seiring waktu, penelitian Computer Vision tidak hanya berfokus pada CNN, tetapi juga berkembang menuju pendekatan yang lebih modern seperti Vision Transformer dan attention mechanism yang menawarkan kemampuan representasi fitur yang lebih adaptif dan efisien. Pergeseran paradigma ini menunjukkan adanya dinamika tren metode yang terus berubah dan memerlukan pemetaan yang terstruktur agar arah penelitian dapat dipahami secara komprehensif (Khan et al., 2022).

Beberapa penelitian sebelumnya telah melakukan kajian literatur maupun implementasi pada domain yang lebih spesifik. Sebagai contoh, Alfian et al. (2024) melakukan kajian literatur mengenai penerapan teknik *machine learning* untuk klasifikasi kualitas udara, sedangkan Tory et al. (2025) mengembangkan sistem pengenalan citra bahasa isyarat menggunakan metode Manhattan. Kedua penelitian tersebut memberikan kontribusi yang signifikan pada bidang masing-masing, namun masih berfokus pada domain atau aplikasi tertentu sehingga belum memberikan gambaran yang komprehensif mengenai perkembangan metode *Computer Vision* secara umum. Akibatnya, belum tersedia pemetaan yang sistematis sehingga peneliti belum memiliki gambaran yang komprehensif mengenai perkembangan metode, perubahan fokus penelitian, serta kecenderungan penggunaan metrik evaluasi sebagai dasar dalam menentukan arah penelitian selanjutnya. Oleh karena itu, penelitian ini menggunakan pendekatan *Systematic Literature Review* (SLR) untuk menyajikan pemetaan yang komprehensif mengenai perkembangan metode *Computer Vision*, fokus tugas penelitian, dan metrik evaluasi yang dominan pada artikel ilmiah yang dipublikasikan selama periode 2017–2025.

Berdasarkan latar belakang tersebut, penelitian ini melakukan *Systematic Literature Review* terhadap 20 artikel ilmiah pada rentang tahun 2017–2025 dengan tujuan untuk: (1) menganalisis tren perkembangan metode Computer Vision, (2) mengidentifikasi metode yang paling dominan digunakan, (3) mengetahui tugas penelitian yang paling sering dibahas, serta (4) mengkaji metrik evaluasi yang umum digunakan dalam penelitian Computer Vision.

## METODOLOGI PENELITIAN

Penelitian ini menggunakan metode *Systematic Literature Review* (SLR) yang bertujuan untuk mengidentifikasi, mengevaluasi, dan mensintesis berbagai penelitian terkait Computer Vision secara sistematis dan terstruktur. Pendekatan SLR dipilih karena mampu memberikan gambaran menyeluruh mengenai perkembangan metode, fokus penelitian, serta metrik evaluasi yang digunakan dalam literatur ilmiah tanpa melakukan eksperimen langsung terhadap model komputasi.

### 1. Jenis Penelitian

Jenis penelitian yang digunakan adalah penelitian kualitatif–kuantitatif berbasis studi literatur. Pendekatan kualitatif digunakan untuk mengelompokkan dan menginterpretasikan fokus penelitian serta metode yang digunakan pada Artikel, sedangkan pendekatan kuantitatif digunakan untuk menghitung frekuensi kemunculan metode, tugas penelitian, serta metrik evaluasi yang dominan.

### 2. Sumber Data

Data penelitian diperoleh dari Artikel ilmiah yang diakses melalui beberapa basis data daring, yaitu Google Scholar, IEEE Xplore, SpringerLink, MDPI, dan arXiv. Pemilihan basis data tersebut didasarkan pada kredibilitas dan cakupan publikasi ilmiah di bidang Computer Vision. Rentang tahun publikasi yang digunakan dalam penelitian ini adalah 2017 hingga 2025 untuk memastikan relevansi dan keterbaruan penelitian yang dianalisis.

### 3. Strategi Pencarian Literatur

Proses pencarian literatur dilakukan menggunakan kombinasi kata kunci yang relevan dengan topik penelitian, antara lain “computer vision survey”, “computer vision review”, “deep learning computer vision”, “video understanding survey”, dan “attention mechanism computer vision”. Kata kunci tersebut digunakan untuk memperoleh artikel yang berfokus pada tinjauan metode dan perkembangan penelitian Computer Vision.

### 4. Kriteria Inklusi dan Eksklusi

Untuk menjaga kualitas dan relevansi literatur yang dianalisis, penelitian ini menerapkan kriteria inklusi dan eksklusi sebagai berikut:

#### Kriteria Inklusi:

- Artikel ilmiah atau prosiding internasional bereputasi
- Topik utama berkaitan dengan Computer Vision
- Jenis artikel berupa survey, review, atau bibliometric analysis
- Rentang tahun publikasi 2017–2025

- e. Artikel tersedia dalam bentuk full paper dan dapat diakses secara daring

#### Kriteria Eksklusi:

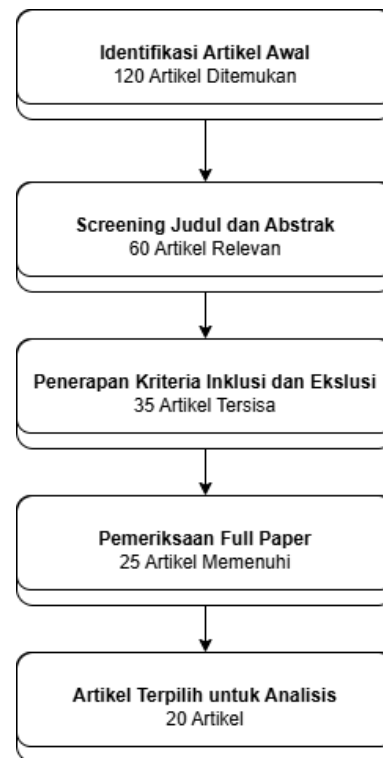
- Artikel dengan fokus utama pada robotics, control system, atau signal processing tanpa pembahasan Computer Vision
- Artikel domain spesifik yang hanya membahas implementasi aplikasi tertentu tanpa membahas tren metode Computer Vision secara umum.
- Buku, laporan teknis, dan artikel non-ilmiah
- Artikel eksperimen tunggal tanpa pembahasan tinjauan metode

## HASIL DAN PEMBAHASAN

### 1. Proses Seleksi Literatur

Tahapan seleksi literatur dilakukan secara bertahap mulai dari identifikasi artikel melalui basis data daring, penyaringan berdasarkan judul dan abstrak, penerapan kriteria inklusi dan eksklusi, hingga pemeriksaan dokumen full paper. Proses pencarian awal menghasilkan 120 artikel ilmiah dari berbagai basis data. Setelah dilakukan penyaringan berdasarkan judul dan abstrak, tersisa 60 artikel. Selanjutnya, penerapan kriteria inklusi dan eksklusi menghasilkan 35 artikel yang memenuhi syarat untuk ditelaah lebih lanjut. Setelah dilakukan peninjauan terhadap naskah lengkap (*full paper*), diperoleh 20 artikel ilmiah yang digunakan sebagai objek kajian dalam penelitian ini. Seluruh artikel yang dipilih kemudian diekstraksi datanya berdasarkan tahun publikasi, fokus tugas Computer

Vision, metode yang digunakan, serta metrik evaluasi yang dilaporkan. Proses seleksi ini dilakukan secara sistematis agar artikel yang digunakan benar-benar relevan dengan topik penelitian. Rangkaian proses seleksi literatur tersebut divisualisasikan pada Gambar 1 sebagai alur seleksi artikel.



Gambar 1. Flowchart Proses Seleksi Literatur (SLR)

### 2. Ekstraksi Data dan Perbandingan Literatur

Setelah diperoleh 20 artikel ilmiah berdasarkan tahapan seleksi literatur, tahap selanjutnya adalah melakukan ekstraksi data untuk mengidentifikasi informasi penting dari masing-masing artikel. Proses ekstraksi dilakukan secara sistematis dengan mencatat beberapa aspek utama, yaitu penulis dan tahun publikasi, judul penelitian, fokus tugas Computer Vision, metode atau model yang digunakan, serta hasil atau kesimpulan utama yang diperoleh. Ekstraksi data ini bertujuan untuk mempermudah proses perbandingan antar penelitian dan menjadi dasar dalam analisis tren metode, fokus task, serta metrik evaluasi yang dominan pada rentang tahun 2017–2025. Hasil ekstraksi data dari seluruh artikel terpilih disajikan pada Tabel 1.

Tabel 1. Hasil Ekstraksi Data dan Perbandingan Literatur Penelitian Computer Vision (2017–2025)

| No | Author (Year)            | Judul Artikel                                           | Fokus/Tugas CV                          | Metode/Model                                      | Hasil/Kesimpulan                                                                                                                                                                |
|----|--------------------------|---------------------------------------------------------|-----------------------------------------|---------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1  | (Zou et al., 2023)       | <i>A Survey of Deep Learning-Based Object Detection</i> | Object Detection                        | CNN-based Deep Learning (YOLO, SSD, Faster R-CNN) | Menunjukkan bahwa metode object detection berbasis deep learning mendominasi penelitian modern dengan peningkatan signifikan pada akurasi deteksi objek melalui arsitektur CNN. |
| 2  | (Ioannidou et al., 2017) | <i>Deep Learning Advances in Computer Vision</i>        | 3D Computer Vision / 3D Data Processing | Deep Learning (CNN untuk 3D data)                 | Penelitian ini membahas perkembangan deep learning untuk pemrosesan data 3D serta menunjukkan bahwa CNN mampu                                                                   |

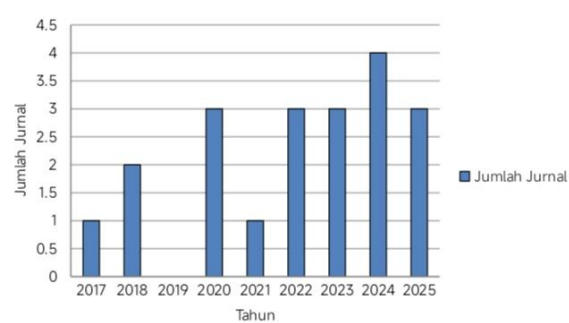
| No | Author<br>(Year)          | Judul Artikel                                                                                | Fokus/Tugas<br>CV                              | Metode/Model                                                 | Hasil/Kesimpulan                                                                                                                                                                                                 |
|----|---------------------------|----------------------------------------------------------------------------------------------|------------------------------------------------|--------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|    |                           | <i>with 3D Data: A Survey</i>                                                                |                                                |                                                              | meningkatkan performa pada tugas berbasis point cloud dan data 3D.                                                                                                                                               |
| 3  | (Ren & Wang, 2022)        | <i>An Overview of Deep Learning-Based Object Detection</i>                                   | Object Detection                               | CNN-based Object Detection Models                            | Menyimpulkan bahwa object detection berbasis deep learning mengalami perkembangan pesat dan metode CNN menjadi pendekatan utama untuk meningkatkan ketepatan deteksi.                                            |
| 4  | (Shao et al., 2022)       | <i>A Review of Deep Learning-Based Object Detection</i>                                      | Object Detection                               | YOLO, Faster R-CNN, SSD                                      | Menunjukkan bahwa model YOLO dan Faster R-CNN menjadi model dominan karena mampu mencapai performa tinggi dalam object detection, khususnya pada dataset skala besar.                                            |
| 5  | (Zhao et al., 2024)       | <i>A Review of Convolutional Neural Networks in Computer Vision</i>                          | Image Classification & General Computer Vision | CNN Architecture                                             | Menjelaskan bahwa CNN masih menjadi metode dominan dalam computer vision karena mampu menghasilkan performa tinggi pada berbagai tugas seperti klasifikasi citra, deteksi objek, dan segmentasi.                 |
| 6  | (Tsirtsakis et al., 2025) | <i>Deep learning for object recognition: A comprehensive review of models and algorithms</i> | Efisiensi Model Computer Vision                | Efficient Deep Learning (compression, pruning, quantization) | Menunjukkan bahwa optimasi model seperti pruning dan quantization menjadi pendekatan penting untuk meningkatkan efisiensi komputasi tanpa mengurangi performa signifikan pada aplikasi Computer Vision.          |
| 7  | (Chahal & Dey, 2018)      | <i>A Survey of Computer Vision Techniques in Object Detection</i>                            | Object Detection                               | CNN-based Detection Models                                   | Menjelaskan bahwa CNN menjadi metode utama pada object detection modern dan mampu meningkatkan performa deteksi dibanding metode tradisional berbasis fitur manual.                                              |
| 8  | (Voulodimos et al., 2018) | <i>Deep Learning for Computer Vision: A Brief Review</i>                                     | General Survey Computer Vision                 | Deep Learning (CNN, DL architectures)                        | Mengidentifikasi bahwa deep learning khususnya CNN menjadi faktor utama dalam kemajuan Computer Vision modern, terutama pada klasifikasi citra, deteksi objek, dan segmentasi.                                   |
| 9  | (Younesi et al., 2024)    | <i>A Comprehensive Survey of Convolutions in Deep Learning</i>                               | General Survey Computer Vision                 | Deep Learning Methods                                        | Menyimpulkan bahwa perkembangan Computer Vision semakin didominasi oleh deep learning dengan peningkatan tren menuju model yang lebih besar, dataset lebih luas, dan integrasi model modern seperti Transformer. |

| No | Author<br>(Year)         | Judul Artikel                                                                                      | Fokus/Tugas<br>CV                                 | Metode/Model                                        | Hasil/Kesimpulan                                                                                                                                                                                                                                                                                     |
|----|--------------------------|----------------------------------------------------------------------------------------------------|---------------------------------------------------|-----------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 10 | (Trigka & Dritsas, 2025) | <i>A Comprehensive Survey of Machine Learning Techniques and Models for Object Detection</i>       | Object Detection                                  | CNN-based Object Detection                          | Menunjukkan bahwa object detection berbasis deep learning berkembang pesat dan menjadi fokus utama penelitian Computer Vision, terutama pada penerapan dunia nyata seperti keamanan dan industri.                                                                                                    |
| 11 | (Goel et al., 2020)      | <i>A Survey of Methods for Low-Power Deep Learning and Computer Vision</i>                         | Efficient Computer Vision / Low-Power AI          | Efficient Deep Learning (optimization, compression) | Menunjukkan bahwa efisiensi energi dan komputasi menjadi fokus penting pada penerapan Computer Vision, khususnya untuk perangkat edge seperti IoT dan embedded systems.                                                                                                                              |
| 12 | (Chai et al., 2021)      | <i>Deep Learning for Computer Vision: A Survey</i>                                                 | General Survey Computer Vision                    | Deep Learning (CNN, DL architectures)               | Menyimpulkan bahwa deep learning telah menjadi pendekatan dominan dalam Computer Vision dan terus berkembang seiring peningkatan dataset besar serta kebutuhan aplikasi real-time.                                                                                                                   |
| 13 | (Han et al., 2023)       | <i>A Survey on Vision Transformer</i>                                                              | Image Classification & General CV                 | Transformer / Vision Transformer (ViT)              | Menunjukkan bahwa Vision Transformer berkembang sebagai metode modern yang mampu menyaingi bahkan melampaui CNN pada beberapa tugas, terutama pada dataset besar dengan pelatihan skala luas.                                                                                                        |
| 14 | (Chen & Deng, 2020)      | <i>Bibliometric Analysis of the Application of Convolutional Neural Network in Computer Vision</i> | Bibliometric Analysis                             | CNN Trend Analysis / Scientometric Analysis         | Menjelaskan bahwa CNN merupakan metode dominan dalam pengembangan Computer Vision dan mengalami peningkatan signifikan dalam publikasi. Studi ini memetakan tren riset CNN, serta menunjukkan bahwa CNN menjadi fondasi utama sebelum berkembangnya metode modern seperti attention dan transformer. |
| 15 | (Lu & Young, 2020)       | <i>A Survey of Public Datasets for Computer Vision Tasks in Precision Agriculture</i>              | Object Detection, Classification, Segmentation    | CNN, YOLO, Faster R-CNN                             | Menunjukkan bahwa Dataset publik membantu perkembangan CV pertanian, namun masih terbatas (ukuran kecil, variasi rendah, kurang standar). Dibutuhkan dataset yang lebih besar dan representatif terhadap kondisi lapangan.                                                                           |
| 16 | (Mettes et al., 2024)    | <i>Hyperbolic Deep Learning for Computer Vision: A Survey</i>                                      | General Computer Vision / Representation Learning | Hyperbolic Deep Learning                            | Menunjukkan bahwa pendekatan hyperbolic deep learning mampu meningkatkan representasi data visual yang bersifat hierarkis dan kompleks, sehingga berpotensi menjadi metode alternatif dalam                                                                                                          |

| No | Author<br>(Year)     | Judul Artikel                                                             | Fokus/Tugas<br>CV                          | Metode/Model                                                       | Hasil/Kesimpulan                                                                                                                                                                               |
|----|----------------------|---------------------------------------------------------------------------|--------------------------------------------|--------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|    |                      |                                                                           |                                            |                                                                    | pengembangan Computer Vision modern.                                                                                                                                                           |
| 17 | (Tang et al., 2025)  | <i>Video Understanding with Large Language Models: A Survey</i>           | Video Understanding                        | Transformer / Video-Language Model (LLM-based)                     | Menjelaskan bahwa integrasi Large Language Models dengan video understanding menjadi tren terbaru yang memperluas kemampuan interpretasi video melalui multimodal learning.                    |
| 18 | (Wang et al., 2023)  | <i>Computation-efficient Deep Learning for Computer Vision: A Survey.</i> | Efficient Computer Vision / General Survey | Efficient Deep Learning (model compression, pruning, quantization) | Menyimpulkan bahwa efisiensi komputasi menjadi tren paling penting dalam pengembangan deep learning untuk Computer Vision.                                                                     |
| 19 | (Madan et al., 2024) | <i>Foundation Models for Video Understanding: A Survey</i>                | Video Understanding                        | Transformer / Foundation Models                                    | Menunjukkan bahwa foundation models berbasis transformer menjadi pendekatan dominan pada video understanding karena mampu mempelajari representasi skala besar dari data video dan multimodal. |
| 20 | (Guo et al., 2022)   | <i>Attention Mechanisms in Computer Vision: A Survey</i>                  | General Survey Computer Vision             | Attention Mechanism                                                | Menjelaskan bahwa attention mechanism berperan besar dalam meningkatkan performa berbagai tugas Computer Vision serta menjadi komponen penting dalam model modern seperti Vision Transformer.  |

Berdasarkan Tabel 1, terlihat bahwa metode CNN dan Transformer mendominasi penelitian Computer Vision, yang selanjutnya dianalisis lebih lanjut pada subbab berikutnya.

3. Tren Publikasi Penelitian



Gambar 2. Tren Publikasi Penelitian Computer Vision tahun 2017-2025

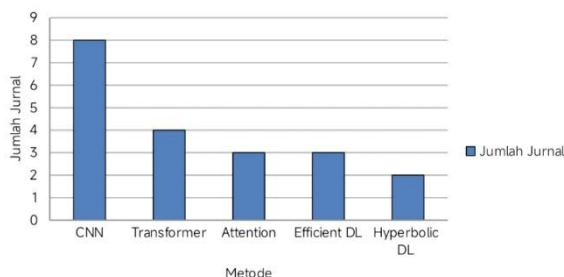
Berdasarkan Gambar 2, tren publikasi penelitian Computer Vision pada rentang tahun 2017 hingga 2025 menunjukkan peningkatan yang cukup signifikan. Pada periode awal, jumlah publikasi masih relatif rendah, yaitu 1 artikel pada tahun 2017 dan meningkat menjadi 2 artikel pada tahun 2018.

Pada tahun 2019 tidak ditemukan artikel yang memenuhi kriteria inklusi penelitian, namun pada tahun 2020 terjadi peningkatan dengan jumlah publikasi mencapai 3 artikel. Setelah itu, jumlah publikasi sempat menurun pada tahun 2021 menjadi 1 artikel, kemudian kembali meningkat secara konsisten pada tahun 2022 dan 2023 dengan masing-masing sebanyak 3 artikel.

Puncak publikasi terjadi pada tahun 2024 dengan total 4 artikel, yang mengindikasikan meningkatnya intensitas penelitian pada bidang Computer Vision. Peningkatan ini sejalan dengan perkembangan metode deep learning yang semakin kompleks, seperti penggunaan Transformer, attention mechanism, serta model berskala besar (foundation models) yang mulai banyak diterapkan dalam analisis citra maupun video. Sementara itu, pada tahun 2025 jumlah publikasi tetap tinggi dengan 3 artikel, yang menunjukkan bahwa penelitian Computer Vision masih menjadi bidang yang aktif dan terus berkembang.

Secara keseluruhan, tren tersebut memperlihatkan bahwa Computer Vision mengalami pertumbuhan yang dinamis, khususnya sejak tahun 2020 hingga 2025. Hal ini menunjukkan bahwa bidang ini memiliki relevansi yang tinggi dalam perkembangan kecerdasan buatan modern dan berperan penting dalam mendukung berbagai aplikasi seperti pengenalan objek, klasifikasi citra, segmentasi, serta pemahaman video.

#### 4. Distribusi Metode



**Gambar 3.** Distribusi Metode Penelitian Computer Vision tahun 2017-2025

Berdasarkan Gambar 3, distribusi metode yang digunakan dalam 20 artikel penelitian Computer Vision pada rentang tahun 2017–2025 menunjukkan bahwa pendekatan berbasis deep learning masih mendominasi pengembangan sistem pengolahan citra dan video. Metode yang paling banyak digunakan adalah Convolutional Neural Network (CNN) dengan total 8 artikel. Dominasi CNN menegaskan bahwa arsitektur ini masih menjadi fondasi utama dalam Computer Vision karena kemampuannya dalam mengekstraksi fitur visual secara otomatis dan efektif. CNN banyak digunakan pada berbagai task seperti image classification, object detection, serta semantic segmentation, dan menjadi dasar pengembangan berbagai model populer seperti YOLO, SSD, dan Faster R-CNN.

Metode Transformer berada pada posisi kedua dengan total 4 artikel, yang menunjukkan adanya tren pergeseran menuju arsitektur modern. Transformer, khususnya Vision Transformer (ViT), mampu memproses informasi visual secara lebih global melalui mekanisme self-attention. Perkembangan Transformer dalam Computer Vision juga didorong oleh meningkatnya kebutuhan pemrosesan dataset berskala besar serta berkembangnya model modern seperti foundation models, terutama pada task yang lebih kompleks seperti video understanding dan multimodal learning. Hal ini mengindikasikan bahwa Transformer mulai menjadi alternatif yang kuat selain CNN dalam penelitian Computer Vision.

Selain CNN dan Transformer, metode Attention Mechanism juga muncul sebanyak 3 artikel. Attention mechanism berperan penting dalam meningkatkan kemampuan model deep learning

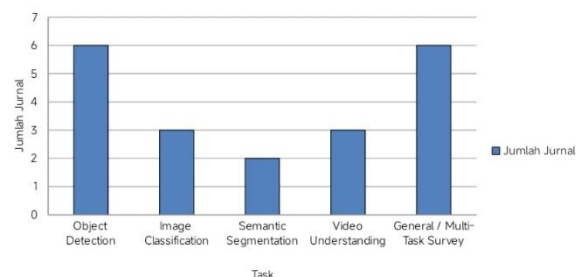
untuk memfokuskan pembelajaran pada bagian citra yang relevan, sehingga dapat meningkatkan performa pada berbagai task visual. Perkembangan attention mechanism juga menjadi dasar penting yang mendukung kemunculan Transformer sebagai salah satu metode dominan dalam Computer Vision modern.

Kategori metode berikutnya adalah Efficient Deep Learning yang juga ditemukan sebanyak 3 artikel. Munculnya kategori ini menunjukkan bahwa penelitian Computer Vision tidak hanya berorientasi pada peningkatan performa akurasi, tetapi juga pada efisiensi komputasi. Teknik seperti pruning, quantization, dan model compression menjadi perhatian utama untuk mendukung implementasi Computer Vision pada perangkat edge, sistem real-time, serta perangkat dengan keterbatasan sumber daya.

Sementara itu, metode Hyperbolic Deep Learning ditemukan sebanyak 2 artikel, yang menunjukkan bahwa pendekatan ini masih tergolong baru namun mulai berkembang. Hyperbolic deep learning memanfaatkan ruang hiperbolik untuk merepresentasikan data hierarkis, sehingga berpotensi meningkatkan kemampuan model dalam memahami struktur visual yang kompleks. Meskipun jumlahnya belum sebanyak metode lain, keberadaan metode ini menunjukkan adanya eksplorasi pendekatan baru dalam penelitian Computer Vision.

Secara keseluruhan, distribusi metode ini menunjukkan bahwa CNN masih menjadi pendekatan dominan, namun penelitian Computer Vision mulai menunjukkan pergeseran menuju metode modern seperti Transformer dan attention mechanism. Di sisi lain, penelitian terkait efficient deep learning dan hyperbolic deep learning menegaskan bahwa bidang ini terus berkembang baik dari aspek inovasi arsitektur maupun efisiensi implementasi.

#### 5. Fokus Task Computer Vision



**Gambar 4.** Distribusi Fokus Task Penelitian Computer Vision tahun 2017-2025

Berdasarkan Gambar 4, fokus penelitian Computer Vision pada 20 artikel yang dianalisis menunjukkan bahwa terdapat beberapa task utama yang dominan pada rentang tahun 2017–2025. Task yang paling

banyak dibahas adalah object detection dengan total 6 artikel. Dominasi object detection menunjukkan bahwa penelitian Computer Vision banyak diarahkan pada pengembangan sistem yang mampu mendeteksi keberadaan objek secara otomatis dalam citra maupun video. Hal ini sejalan dengan kebutuhan aplikasi nyata yang luas, seperti sistem pengawasan keamanan, kendaraan otonom, sistem industri cerdas, hingga pengolahan data visual berbasis real-time.

Selain object detection, kategori yang memiliki jumlah sama besar adalah general survey atau multi-task survey dengan total 6 artikel. Tingginya jumlah artikel survey menunjukkan bahwa perkembangan Computer Vision berlangsung sangat cepat sehingga diperlukan penelitian pemetaan untuk mengidentifikasi tren metode, arsitektur model, dataset yang digunakan, serta tantangan yang masih terbuka. Studi survey juga berperan penting sebagai referensi untuk menentukan arah riset dan metode yang relevan dalam pengembangan Computer Vision modern.

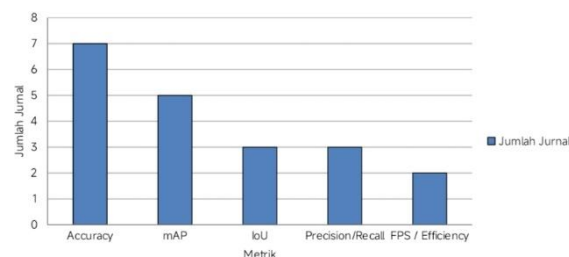
Selanjutnya, task image classification muncul sebanyak 3 artikel. Meskipun jumlahnya tidak sebanyak object detection, image classification tetap menjadi task fundamental dalam Computer Vision karena berfungsi sebagai dasar dalam pembelajaran representasi visual dan banyak digunakan pada berbagai penerapan, termasuk pengenalan objek, analisis citra medis, serta klasifikasi citra berbasis deep learning. Task ini juga sering dijadikan tolok ukur performa model karena bersifat umum dan memiliki dataset benchmark yang luas.

Task video understanding juga ditemukan sebanyak 3 artikel, yang mengindikasikan bahwa penelitian Computer Vision tidak hanya berkembang pada data citra statis, tetapi juga pada analisis data video yang bersifat dinamis. Video understanding menjadi semakin penting karena mampu menangkap informasi temporal dan konteks pergerakan, serta banyak digunakan dalam pengenalan aktivitas manusia, analisis perilaku, serta pengembangan model multimodal berbasis transformer dan foundation models. Hal ini menunjukkan adanya pergeseran tren menuju sistem yang mampu memahami informasi visual secara lebih kompleks.

Sementara itu, task semantic segmentation memiliki jumlah paling sedikit yaitu 2 artikel. Walaupun tidak dominan dalam penelitian ini, semantic segmentation tetap menjadi task yang penting karena berperan dalam segmentasi detail berbasis piksel untuk memisahkan area objek secara lebih akurat. Task ini banyak diterapkan pada bidang yang membutuhkan ketelitian tinggi, seperti analisis citra medis, citra satelit, dan pemetaan lingkungan. Rendahnya jumlah artikel pada kategori ini dapat disebabkan oleh keterbatasan literatur yang sesuai dengan kriteria inklusi pada periode penelitian atau karena fokus publikasi lebih banyak mengarah pada object detection dan survey metode.

Secara keseluruhan, distribusi fokus task tersebut menunjukkan bahwa penelitian Computer Vision pada periode 2017–2025 didominasi oleh penelitian yang berorientasi pada aplikasi utama (object detection) serta kajian literatur yang memetakan tren perkembangan metode (general survey). Hal ini memperkuat bahwa bidang Computer Vision terus berkembang dan memiliki relevansi tinggi dalam pengembangan kecerdasan buatan modern.

## 6. Metrik Evaluasi



**Gambar 5.** Distribusi Metrik Evaluasi Penelitian Computer Vision tahun 2017-2025

Berdasarkan Gambar 5, metrik evaluasi yang paling sering digunakan dalam penelitian Computer Vision adalah Accuracy dengan total 7 artikel. Hal ini menunjukkan bahwa akurasi masih menjadi indikator utama dalam menilai performa model pengolahan citra.

Selanjutnya, metrik Mean Average Precision (mAP) muncul pada 5 artikel, terutama pada penelitian yang berfokus pada object detection. Metrik Intersection over Union (IoU) serta Precision dan Recall masing-masing digunakan pada 3 artikel, sedangkan metrik berbasis efisiensi seperti FPS atau Latency muncul pada 2 artikel.

Distribusi ini menunjukkan bahwa evaluasi performa model Computer Vision tidak hanya berfokus pada ketepatan hasil, tetapi juga mulai mempertimbangkan aspek efisiensi dan kecepatan komputasi.

## KESIMPULAN

Berdasarkan hasil Systematic Literature Review terhadap 20 artikel ilmiah pada rentang tahun 2017–2025, penelitian ini menunjukkan bahwa tren publikasi Computer Vision mengalami peningkatan yang cukup signifikan, khususnya setelah tahun 2022 dengan puncak publikasi pada tahun 2024. Hal ini mengindikasikan bahwa bidang Computer Vision terus berkembang dan menjadi fokus utama dalam penelitian teknologi berbasis kecerdasan buatan.

Dari sisi metode, Convolutional Neural Network (CNN) masih menjadi pendekatan yang paling dominan dengan total delapan artikel, meskipun metode modern seperti Transformer dan Attention Mechanism mulai menunjukkan peningkatan penggunaan. Pada aspek fokus penelitian, tugas Object Detection dan General Survey menjadi topik

yang paling banyak dibahas, sementara Semantic Segmentation dan Video Understanding tetap menunjukkan kontribusi yang konsisten.

Selanjutnya, metrik evaluasi yang paling sering digunakan adalah Accuracy, diikuti oleh Mean Average Precision (mAP) dan Intersection over Union (IoU). Hal ini menunjukkan bahwa evaluasi performa model Computer Vision masih berorientasi pada ketepatan hasil, namun mulai mempertimbangkan aspek efisiensi dan kecepatan komputasi. Secara keseluruhan, penelitian ini memberikan gambaran komprehensif mengenai arah perkembangan metode, fokus tugas, serta metrik evaluasi dalam literatur Computer Vision.

Penelitian ini memiliki keterbatasan karena hanya menganalisis artikel ilmiah yang memenuhi kriteria inklusi pada rentang tahun 2017–2025 dan berasal dari basis data yang telah ditentukan. Oleh karena itu, hasil penelitian ini belum sepenuhnya merepresentasikan keseluruhan perkembangan metode Computer Vision yang dipublikasikan pada berbagai sumber ilmiah. Penelitian selanjutnya disarankan untuk memperluas cakupan basis data, rentang waktu publikasi, serta jenis literatur yang dianalisis agar diperoleh pemetaan perkembangan *Computer Vision* yang lebih komprehensif.

#### DAFTAR PUSTAKA

- Alfian, H., Wahyuni, S., Revalino, A., Mirano, M. F., Rahmayana, E., & Mukhtar, H. (2024). Teknik Machine Learning Untuk Analisa Klasifikasi Kualitas Udara: a Review. *Journal of Software Engineering and Information Systems*, 4(2), 108–118. <https://doi.org/10.37859/seis.v4i2.7617>
- Chahal, K. S., & Dey, K. (2018). *A Survey of Modern Object Detection Literature using Deep Learning*. <https://arxiv.org/abs/1808.07256>
- Chai, J., Zeng, H., Li, A., & Ngai, E. W. T. (2021). Deep learning in computer vision: A critical review of emerging techniques and application scenarios. *Machine Learning with Applications*, 6(March), 100134. <https://doi.org/10.1016/j.mlwa.2021.100134>
- Chen, H., & Deng, Z. (2020). Bibliometric Analysis of the Application of Convolutional Neural Network in Computer Vision. *IEEE Access*, 8, 155417–155428. <https://doi.org/10.1109/ACCESS.2020.3019336>
- Goel, A., Tung, C., Lu, Y. H., & Thiruvathukal, G. K. (2020). A Survey of Methods for Low-Power Deep Learning and Computer Vision. *IEEE World Forum on Internet of Things, WF-IoT 2020 - Symposium Proceedings*. <https://doi.org/10.1109/WF-IoT48130.2020.9221198>
- Guo, M. H., Xu, T. X., Liu, J. J., Liu, Z. N., Jiang, P. T., Mu, T. J., Zhang, S. H., Martin, R. R., Cheng, M. M., & Hu, S. M. (2022). Attention mechanisms in computer vision: A survey. *Computational Visual Media*, 8(3), 331–368. <https://doi.org/10.1007/s41095-022-0271-y>
- Han, K., Wang, Y., Chen, H., Chen, X., Guo, J., Liu, Z., Tang, Y., Xiao, A., Xu, C., Xu, Y., Yang, Z., Zhang, Y., & Tao, D. (2023). A Survey on Vision Transformer. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(1), 87–110. <https://doi.org/10.1109/TPAMI.2022.3152247>
- Ioannidou, A., Chatzilari, E., Nikolopoulos, S., & Kompatsiaris, I. (2017). Deep learning advances in computer vision with 3D data: A survey. *ACM Computing Surveys*, 50(2). <https://doi.org/10.1145/3042064>
- Khan, S., Naseer, M., Hayat, M., Zamir, S. W., Khan, F. S., & Shah, M. (2022). Transformers in Vision: A Survey. *ACM Computing Surveys*, 54(10), 1–30. <https://doi.org/10.1145/3505244>
- Lu, Y., & Young, S. (2020). A survey of public datasets for computer vision tasks in precision agriculture. *Computers and Electronics in Agriculture*, 178(September), 105760. <https://doi.org/10.1016/j.compag.2020.105760>
- Madan, N., Moegelmose, A., Modi, R., Rawat, Y. S., & Moeslund, T. B. (2024). *Foundation Models for Video Understanding: A Survey*. 1–48. <http://arxiv.org/abs/2405.03770>
- Mettes, P., Ghadimi Atigh, M., Keller-Ressel, M., Gu, J., & Yeung, S. (2024). Hyperbolic Deep Learning in Computer Vision: A Survey. *International Journal of Computer Vision*, 132(9), 3484–3508. <https://doi.org/10.1007/s11263-024-02043-5>
- Ren, J., & Wang, Y. (2022). Overview of Object Detection Algorithms Using Convolutional Neural Networks. *Journal of Computer and Communications*, 10, 115–132. <https://doi.org/10.4236/jcc.2022.101006>
- Shao, Y., Zhang, D., Chu, H., Zhang, X., & Rao, Y. (2022). A Review of YOLO Object Detection Based on Deep Learning. *Dianzi Yu Xinxu Xuebao/Journal of Electronics and Information Technology*, 44(10), 3697–3708. <https://doi.org/10.11999/JEIT210790>
- Tang, Y., Bi, J., Xu, S., Song, L., Liang, S., Wang, T., Zhang, D., An, J., Lin, J., Zhu, R., Vosoughi, A., Huang, C., Zhang, Z., Liu, P., Feng, M., Zheng, F., Zhang, J., Luo, P., Luo, J., & Xu, C. (2025). Video Understanding with Large Language Models: A Survey. *IEEE Transactions on Circuits and Systems for Video Technology*, 1–22. <https://doi.org/10.1109/TCSVT.2025.3566695>
- Tory, A. R. A. S., Pradana, A. I., & Maulindar, J. (2025). *Model Klasifikasi Jarak Manhattan Pada Pengenalan Citra Sistem Bahasa Isyarat Bahasa Indonesia*. 5(2), 35–40.

- Trigka, M., & Dritsas, E. (2025). A Comprehensive Survey of Machine Learning Techniques and Models for Object Detection. *Sensors*, 25(1). <https://doi.org/10.3390/s25010214>
- Tsirtsakis, P., Zacharis, G., Maraslidis, G. S., & Fragulis, G. F. (2025). Deep learning for object recognition: A comprehensive review of models and algorithms. *International Journal of Cognitive Computing in Engineering*, 6(December 2024), 298–312. <https://doi.org/10.1016/j.ijcce.2025.01.004>
- Voulodimos, A., Doulamis, N., Doulamis, A., & Protopapadakis, E. (2018). Deep Learning for Computer Vision: A Brief Review. *Computational Intelligence and Neuroscience*, 2018. <https://doi.org/10.1155/2018/7068349>
- Wang, Y., Han, Y., Wang, C., Song, S., Tian, Q., & Huang, G. (2023). *Computation-efficient Deep Learning for Computer Vision: A Survey*. 1(1). <http://arxiv.org/abs/2308.13998>
- Younesi, A., Ansari, M., Fazli, M., Ejlali, A., Shafique, M., & Henkel, J. (2024). A Comprehensive Survey of Convolutions in Deep Learning: Applications, Challenges, and Future Trends. *IEEE Access*, 12, 41180–41218. <https://doi.org/10.1109/ACCESS.2024.3376441>
- Zhao, X., Wang, L., Zhang, Y., Han, X., Deveci, M., & Parmar, M. (2024). A review of convolutional neural networks in computer vision. In *Artificial Intelligence Review* (Vol. 57, Number 4). Springer Netherlands. <https://doi.org/10.1007/s10462-024-10721-6>
- Zou, Z., Chen, K., Shi, Z., Guo, Y., & Ye, J. (2023). Object Detection in 20 Years: A Survey. *Proceedings of the IEEE*, 111(3), 257–276. <https://doi.org/10.1109/JPROC.2023.3238524>