|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Paper | Summary | Novelty | Technologies/  Architecture  Used | Results/  Evaluation/  Benchmark | Limitations | Future Work | Research Gap |
| 1. [**Automatic Player Face Detection and Recognition for Players in Cricket Games**](https://www.researchgate.net/publication/379071780_Automatic_Player_Face_Detection_and_Recognition_for_Players_in_Cricket_Games) | This paper presents an augmented reality system using the AdaBoost algorithm for player face detection and recognition, enhancing live broadcasts with real-time player information | Introduction of an augmented reality overlay in live broadcasts for instant player identification. | AdaBoost algorithm for face detection and recognition | Demonstrated high accuracy under controlled conditions with frontal face orientations | Decreased accuracy with non-frontal face angles and dependency on high-quality video input | Improving algorithm robustness to handle varied face orientations and video qualities | Developing methods to enhance detection accuracy in low-resolution and dynamically changing video feeds |
| 1. [**Automatic Player Detection and Identification for Sports Entertainment Applications**](https://link.springer.com/article/10.1007/s10044-014-0416-4) | Develops a sports broadcasting application for automatic detection and recognition of players, displaying their information in real-time | Real-time overlay of player information in sports broadcasts | Machine learning techniques for face detection | Achieved high accuracy in player detection with frontal and near-frontal facial captures | Requires faces to be frontal or near-frontal for high accuracy. | Expanding the detection capability to include non-frontal facial orientations | Enhancing the system to work effectively under varied lighting and motion conditions |
| 1. [**Past, Present, and Future of Face Recognition: A Review**](https://www.mdpi.com/2079-9292/9/8/1188) | This review paper explores the evolution of face recognition technology, detailing its historical development, current applications, and future potential. It discusses technological advancements, application areas, and emerging trends. | The paper provides a comprehensive historical context, allowing readers to understand how face recognition technology has evolved and what has influenced these changes over time. | Reviews various algorithms and technologies used in face recognition over the years, including but not limited to, neural networks, deep learning techniques, and 3D face recognition methods. | The paper compiles various benchmarks and evaluations from past research to compare the effectiveness and accuracy of different face recognition technologies. | Discusses the challenges related to privacy concerns, ethical issues, and the potential for bias in training datasets that can affect the fairness of face recognition systems. | Suggests areas for further research, particularly in improving the robustness and ethical dimensions of face recognition technology. | Highlights the need for developing more secure and privacy-preserving methods in face recognition, which could be an avenue for future research and development. |
| 1. [**A Comprehensive Review of Computer Vision in Sports: Open Issues, Future Trends and Research Directions**](https://www.mdpi.com/2076-3417/12/9/4429) | Covers the application of computer vision in sports, highlighting current research and future trends | Extensive review of various computer vision applications in sports, outlining future research directions. | Survey of existing computer vision techniques used across multiple sports. | Compilation of findings from numerous studies demonstrating the effectiveness of current technologies and methodologies | Need for more scalable and cost-effective solutions for real-time applications. | Focus on developing lightweight models that can be deployed in real-time sports environments. | Creating cost-effective, scalable computer vision systems that require minimal computational resources |
| 1. [**Multi-camera Multi-player Tracking with Deep Player Identification**](https://www.sciencedirect.com/science/article/abs/pii/S0031320320300650#:~:text=Highlights&text=We%20propose%20a%20robust%20tracking,%2C%20IPOM%20and%20KSP%2DID.) | Proposes a framework for multi-camera tracking and identification of players, improving tracking accuracy using deep learning. | Integration of multi-camera inputs to enhance player tracking and identification accuracy. | Deep learning techniques, multi-camera integration. | Significantly improved tracking accuracy in complex multi-player environments | High computational requirements and complex data integration challenge real-time processing. | Streamlining data integration and processing to enhance real-time performance. | Reducing computational overhead and simplifying system architecture for better real-time applicability. |
|  |  |  |  |  |  |  |  |
| Paper | Summary | Novelty | Technologies/  Architecture  Used | Results/  Evaluation/  Benchmark | Limitations | Future Work | Research Gap |
| 1. [Efficient deep feature selection for remote sensing image recognition with fused deep learning architectures](https://link.springer.com/article/10.1007/s11227-019-03106-y) | This paper presents a method for efficient feature selection in remote sensing image recognition using fused deep learning architectures. The approach combines features extracted from multiple CNN models (e.g., VGG16, GoogleNet, ResNet) and uses the ReliefF algorithm to select the most relevant features. These features are then fed into a support vector machine (SVM) classifier. | The combination of multiple CNN architectures for feature extraction and the use of the ReliefF algorithm for feature selection offer a novel and efficient solution for remote sensing image recognition. | CNN architectures like VGG16, ResNet, GoogleNet; ReliefF algorithm; SVM. | The proposed method significantly reduces the feature set, improving classification performance and efficiency. It achieves a high accuracy of 89% for remote sensing image classification. | The method relies on high-quality, pre-labeled data and struggles when applied to more diverse, noisy datasets. |  Future research could focus on optimizing the method for handling noisy and unlabeled data, improving its robustness in real-world applications.   | The current approach lacks robustness when dealing with noise or data variations in real-world scenarios. There is a need for methods that can handle unstructured or less curated datasets |
| 1. [A Novel and Effective Brain Tumor Classification Model Using Deep Feature Fusion and Famous Machine Learning Classifiers](https://www.researchgate.net/publication/359510680_A_Novel_and_Effective_Brain_Tumor_Classification_Model_Using_Deep_Feature_Fusion_and_Famous_Machine_Learning_Classifiers) | This paper proposes a new approach for brain tumor classification using a combination of deep feature fusion techniques with machine learning classifiers such as SVM and KNN. It utilizes MRI images and applies data augmentation to improve performance | The fusion of deep features extracted from multiple CNN models is a novel approach in this context, providing higher accuracy in tumor classification. | CNN models like AlexNet and VGG16, and machine learning classifiers SVM and KNN. | The proposed method achieves high accuracy in brain tumor classification, reaching 96.8% under certain conditions. | The approach is computationally expensive and requires further testing for real-time applications. | Improvements to make the model computationally efficient and suitable for clinical deployment​ | Despite achieving high accuracy, the model's effectiveness in real-world clinical settings remains untested, especially under varying data qualities. Further research is required to reduce the computational overhead and make the model suitable for real-time clinical applications​ |
| 1. [Multiclass Brain Tumor Classification Using Convolutional Neural Network and Support Vector Machine](https://ieeexplore.ieee.org/document/9526262) | This paper proposes a hybrid model combining convolutional neural networks (CNN) and support vector machines (SVM) for multiclass brain tumor classification. The CNN is used for feature extraction from MRI images, and SVM is applied for final classification. | The hybrid approach leverages the strengths of CNN for automatic feature extraction and SVM for precise classification, offering improved performance in multiclass classification of brain tumors. | CNN for feature extraction, SVM for classification. | The method achieves promising results with an overall classification accuracy surpassing other techniques. It shows improved performance in identifying multiple types of brain tumors. | The model’s performance depends on a well-labeled and curated dataset, and its computational cost is relatively high due to the hybrid nature of CNN and SVM. | Future research could focus on reducing computational complexity and improving the model's performance on less curated and more diverse datasets. | While the hybrid model shows promise, it lacks optimization for real-time clinical use, and its performance on noisy or less structured data remains untested​ |
| 1. **[Hybrid Deep-Learning Framework Based on Gaussian Fusion of Multiple Spatiotemporal Networks for Walking Gait Phase Recognition](https://onlinelibrary.wiley.com/doi/full/10.1155/2020/8672431)** | This study introduces a hybrid deep-learning framework combining multiple spatiotemporal networks using Gaussian fusion for recognizing different gait phases based on IMU sensor data. | The novel use of Gaussian probability fusion optimizes spatiotemporal feature extraction from IMU data to improve gait phase recognition. | CNN, RNN, LSTM, GRU, and a Gaussian fusion module optimized by the Expectation-Maximum algorithm. | The framework achieved over 96.7% accuracy in gait phase recognition, showing significant performance improvements in real-time scenarios. | The model faces challenges with handling noise and complex sensor data, especially when multiple signals are fused. | Future efforts will focus on reducing computational complexity and improving robustness under noisy conditions for better real-time applicatio | While the framework improves gait phase detection, it struggles with the complexity of combining signals from multiple sensors and handling noise effectively. More work is needed to develop methods that can enhance recognition performance in noisy and unconstrained environments​ |
| 1. **[Enhancing Cricket Performance Analysis with Human Pose Estimation and Machine Learning](https://www.mdpi.com/1424-8220/23/15/6839)** | Uses human pose estimation to predict and analyze cricket strokes. | Application of human pose estimation for detailed analysis of cricket strokes | MediaPipe library, machine learning algorithms | Effective prediction of cricket strokes with actionable insights for coaching | Challenges in real-time responsiveness and occlusion handling | Improving real-time processing speeds and accuracy during fast player movements. | Developing enhanced models to handle partial occlusions and multiple player interactions. |
| 1. **[HybridGait: A Benchmark for Spatial-Temporal Cloth-Changing Gait Recognition with Hybrid Explorations](https://ojs.aaai.org/index.php/AAAI/article/view/27926)** | This paper introduces a new benchmark, CCGait, focused on recognizing gait patterns even when individuals change their clothes over time. The benchmark is designed for both indoor and outdoor environments and includes various data formats such as 2D silhouettes, 3D human meshes, and keypoints | The dataset captures diverse temporal and spatial changes in real-world conditions, making it one of the most challenging benchmarks for gait recognition. The proposed framework, HybridGait, employs a combination of 3D human models and temporal dynamics for improved accuracy. | Canonical Alignment Spatial-Temporal Transformer (CA-STT) module, CNN for spatial feature extraction, 3D-to-2D projection models. | The method outperforms previous benchmarks (such as Gait3D) on challenging datasets, showing improved performance in recognizing individuals under varying conditions, including clothing changes. | Handling occlusions and complex lighting conditions remains a challenge in the proposed model. | Further optimization is planned for the model to handle more diverse environments and occlusion issues for better real-time application | Current challenges include handling occlusions and environmental variations more effectively. Further research is necessary to improve the model's adaptability to more dynamic, real-world conditions, and further optimize it for different environments​ |
| 1. [**Gait-DenseNet: A Hybrid Convolutional Neural Network for Gait Recognition.**](https://openurl.ebsco.com/EPDB%3Agcd%3A15%3A21377999/detailv2?sid=ebsco%3Aplink%3Ascholar&id=ebsco%3Agcd%3A157247544&crl=c) | This paper introduces Gait-DenseNet, a hybrid model that combines a pre-trained DenseNet-201 architecture with a multilayer perceptron for gait recognition. The model first extracts features from gait energy images (GEI), which capture both static and dynamic information from a gait sequence. Transfer learning is applied to fine-tune the DenseNet-201 model, and a multilayer perceptron further processes the extracted features. | The use of DenseNet-201 in conjunction with a multilayer perceptron allows for improved feature extraction and classification in the context of gait recognition, particularly when dealing with variations in viewing angle, clothing, and carrying conditions. | DenseNet-201 architecture, multilayer perceptron, Gait Energy Image (GEI), transfer learning. | The proposed model shows significant improvements over existing methods when tested on datasets like CASIA-B and OU-ISIR. It achieves higher accuracy in recognizing gait patterns under varying conditions. | Handling extreme variations in clothing and other factors remains a challenge, although the model performs well on controlled datasets. | Future improvements could focus on enhancing the model's robustness to occlusions and external environmental changes for better real-world application | The model demonstrates strong performance on controlled datasets, but there is a need to extend its robustness to more diverse conditions, such as extreme variations in clothing and environmental factors. The current approach still lacks the ability to handle highly dynamic environments and real-world scenarios |
| 1. [**An Acceleration Based Fusion of Multiple Spatiotemporal Networks for Gait Phase Detection**](https://www.mdpi.com/1660-4601/17/16/5633) | Proposes a hybrid deep-learning framework (FMS-Net) combining CNN and LSTM for gait phase detection from IMU data. It uses skip connections and batch normalization to improve model performance. | Novel fusion of CNN for spatial features and LSTM for temporal features to improve gait phase detection accuracy | CNN, LSTM, IMU sensors, skip connections | Achieves 96.7% macro-F1 score for gait phase recognition, significantly outperforming other methods in controlled environments. | Limited generalization to highly complex environments and walking conditions. | Optimizing the framework for diverse real-world environments, such as rough terrains or various motion patterns. | Lack of robustness in real-world conditions, especially under diverse walking patterns and challenging environments. |
| 1. [**Robust gait recognition using hybrid descriptors based on Skeleton Gait Energy Image**](https://www.sciencedirect.com/science/article/abs/pii/S0167865519301618) | Proposes hybrid descriptors combining skeleton gait energy images with motion features for robust gait recognition. Focuses on skeletal data to improve recognition under occlusion and other challenges. | The fusion of skeleton data and motion features enhances recognition under challenging conditions like occlusion and poor lighting. | Skeleton Gait Energy Image (SGEI), hybrid descriptors, CNN | Outperforms conventional GEI-based methods in gait recognition, especially under occlusion and difficult lighting conditions. | High computational cost due to the combination of multiple feature extraction methods. | Future work could focus on reducing computational complexity and improving recognition performance under extreme occlusion. | Needs better handling of computational cost while maintaining accuracy in extreme occlusion and difficult lighting. |
| 1. [**Model-based person identification in multi-gait scenario using hybrid classifier**](https://link.springer.com/article/10.1007/s00530-022-01041-2) | Introduces a hybrid classifier that integrates model-based gait features with machine learning classifiers for person identification in multi-gait scenarios. | Combination of gait features and machine learning classifiers allows better performance in multi-gait scenarios. | Gait model features, hybrid classifier | Achieves higher accuracy in person identification when dealing with multiple gait variations. | Still limited by the availability of accurate labeled data for training in multi-gait scenarios. | Expand the dataset and apply the classifier to more varied real-world datasets for improved robustness and generalization. | The method requires more diverse, labeled data for effective application in various real-world multi-gait scenarios. |
| 1. [**Hybrid Deep Neural Network Framework Combining Skeleton and Gait Features for Pathological Gait Recognition**](https://www.mdpi.com/2306-5354/10/10/1133) | Proposes a hybrid deep neural network combining skeleton and gait features to improve the recognition of pathological gait patterns. This approach focuses on identifying abnormal walking patterns in medical contexts. | Integration of both skeleton and gait features for more accurate identification of pathological gait patterns in clinical settings. | CNN, RNN, Skeleton data | Demonstrates higher recognition rates of pathological gait patterns compared to other methods, particularly for medical applications. | Still requires large, well-labeled datasets to train the network effectively in various medical contexts. | Further development in gathering more comprehensive medical gait datasets and improving real-time performance in clinical settings. | Requires a larger, diverse dataset specific to pathological gait recognition, as well as real-time testing in clinical settings. |
| 1. [**Gaitcotr: Improved Spatial-Temporal Representation for Gait Recognition with a Hybrid Convolution-Transformer Framework**](https://ieeexplore.ieee.org/abstract/document/10096602) | Introduces a hybrid framework combining convolutional layers and transformers to capture spatial-temporal patterns in gait recognition. | Integration of transformers for enhanced spatial-temporal representation in gait recognition. | Convolutional layers, Transformer networks | Achieved improved accuracy in recognizing subtle gait changes under varying conditions. | Complexity in model design and computation-heavy architecture. | Improve model efficiency to reduce computational requirements without sacrificing accuracy. | Need for optimization of the hybrid model for lower resource environments. |
| 1. [**Multi-view gait recognition system using spatio-temporal features and deep learning**](https://www.sciencedirect.com/science/article/abs/pii/S095741742100498X) | Proposes a multi-view gait recognition system that uses deep learning to capture spatio-temporal features for robust recognition across different camera angles. | Multi-view system enhances recognition by combining spatial and temporal features from various angles. | Multi-view deep learning, Spatio-temporal features, CNN | Demonstrated higher recognition accuracy across multiple angles, making it ideal for real-world security applications. | Limited generalizability in environments with occlusions or poor camera angles. | Adaptation to environments with varying occlusion and extreme lighting conditions to improve robustness. | Handling occlusion and extreme camera angles remain a challenge for accurate identification. |
| 1. [**Gait Recognition Analysis for Human Identification Analysis-A Hybrid Deep Learning Process**](https://link.springer.com/article/10.1007/s11277-022-09758-z) | Combines CNN and RNN to identify human gait for biometric authentication, emphasizing spatio-temporal gait features. | Hybrid approach utilizing both CNN for spatial features and RNN for temporal sequences in human identification. | CNN, RNN, Spatio-temporal features | High accuracy in human identification using gait patterns, showing robustness under controlled conditions. | High dependency on curated datasets; performance drops with noisy or unstructured data. | Explore real-time applications and robustness testing with unstructured, noisy datasets. | The approach needs to handle noisy, real-world data for robust identification in unconstrained settings. |
| 1. [**STAR: Spatio-Temporal Augmented Relation Network for Gait Recognition**](https://ieeexplore.ieee.org/abstract/document/9913216) | Introduces a STAR network that models relationships between different body parts over time, providing improved gait recognition through spatio-temporal relational data. | Focuses on the augmented relation network to improve gait recognition by modeling body part relations over time. | Spatio-Temporal Augmented Relation (STAR) network, Deep learning | Achieved superior recognition accuracy in real-world scenarios, outperforming traditional methods by leveraging body part relational data. | Model complexity increases as more relations are modeled, leading to potential overfitting in smaller datasets. | Focus on optimizing model complexity and expanding relational data for better generalization across different population groups. | Requires optimization to reduce complexity and improve generalization across larger, more diverse datasets. |
| 1. [**Automatic multi-gait recognition using pedestrian’s spatiotemporal features**](https://link.springer.com/article/10.1007/s11227-023-05391-0) | Proposes a system that combines spatiotemporal features from different pedestrian gait patterns for more robust recognition under varied conditions. | Focuses on combining spatiotemporal features from different gait patterns to enhance recognition accuracy. | Spatiotemporal features, Machine Learning | Achieves higher recognition accuracy in dynamic environments with varied pedestrian movements. | Struggles with data generalization in highly noisy environments or occlusion. | Improve robustness of the model by incorporating noise-resilient techniques for better real-world application. | Needs better handling of complex real-world data, especially under occlusion and environmental changes. |
| 1. [**Human Gait Recognition: A Single Stream Optimal Deep Learning Features Fusion**](https://www.mdpi.com/1424-8220/21/22/7584) | Introduces a deep learning framework for human gait recognition, using two pre-trained models and feature fusion for optimal feature selection and classification. | Utilizes modified whale optimization for optimal feature selection and a new fusion approach to enhance accuracy. | Inception-ResNet-V2, NASNet Mobile, Whale Optimization Algorithm, Mean Absolute Deviation Extended Serial Fusion (MDeSF) | Achieved 89% accuracy on the CASIA-B dataset, showing improved precision and recall compared to existing methods. | High computational complexity due to feature extraction from multiple angles and deep learning models. | Future efforts could focus on reducing computational complexity and testing with larger, real-world datasets. | Requires optimization to reduce the high computational cost associated with multi-view feature extraction and fusion processes. |
| 1. [**Skeleton-based abnormal gait recognition with spatio-temporal attention enhanced gait-structural graph convolutional networks**](https://www.sciencedirect.com/science/article/abs/pii/S0925231221018385) | Proposes a framework based on graph convolutional networks (GCNs) that uses spatio-temporal attention mechanisms to detect abnormalities in gait patterns. | Introduction of spatio-temporal attention mechanisms to improve the recognition of abnormal gait patterns in skeleton-based models. | Graph Convolutional Networks (GCNs), Spatio-Temporal Attention Mechanisms | Shows improved accuracy in recognizing abnormal gait patterns, particularly in clinical scenarios for gait disorder diagnosis. | Limited by the quality and resolution of skeletal data, particularly in noisy or low-resolution environments. | Develop methods to handle more complex skeleton structures and environmental noise to improve the robustness of gait recognition. | More work is needed to enhance robustness to noisy skeletal data and handle variability in real-world clinical applications. |
| 1. [**Hybrid LSTM and GAN model for action recognition and prediction of lawn tennis sport activities**](https://link.springer.com/article/10.1007/s00500-023-09215-4) | Combines LSTM networks with generative adversarial networks (GANs) to predict and recognize actions in tennis based on players' movement sequences. | Use of GANs with LSTM for real-time action recognition and prediction in sports, specifically focusing on tennis activities. | LSTM Networks, Generative Adversarial Networks (GANs) | Achieved real-time performance in recognizing tennis activities with a high degree of accuracy, predicting future movements based on current sequences. | Limited applicability outside of tennis due to specific movement patterns, and the complexity of training GANs for real-time application. | Extend the approach to other sports or more generalized action recognition tasks, improving model generalizability and performance across sports. | The model needs better generalization for action recognition in different sports or domains, particularly those with more complex or varied movement patterns. |
| 1. [**INDIVIDUAL ACTION AND GROUP ACTIVITY RECOGNITION IN SOCCER VIDEOS**](https://essay.utwente.nl/84038/1/Gerats_MA_EEMCS.pdf) | Proposes a method that simultaneously recognizes individual actions and group activities in soccer videos using a static panoramic camera. | Combines individual player actions with group activities, which is rare in sports video analytics. | I3D CNN, Optical flow, Graph Attention Networks (GATs) | Achieved 98.7% accuracy for individual action recognition and 75.2% for group activity recognition on a soccer dataset. | Struggles with lower performance in group activities recognition compared to individual actions. | Improve group activity recognition by enhancing temporal dependencies and multi-view analysis. | Needs more robust models for simultaneous detection of both individual and group activities in other sports and video contexts. |
| 1. [**Identification of humans using gait**](https://ieeexplore.ieee.org/abstract/document/1323098) | Focuses on gait as a biometric for human identification, analyzing how the unique patterns of walking can be leveraged to recognize individuals. | Demonstrates the use of gait patterns in human identification as an effective biometric method. | Gait Energy Image (GEI), Principal Component Analysis (PCA), Support Vector Machine (SVM) | Shows high accuracy in controlled environments but struggles when faced with changes in clothing or carrying objects. | Sensitive to changes in external conditions such as clothing or walking surfaces. | Develop methods to handle more varied conditions, including occlusion and environmental changes, for real-world applications. | More robust models needed to handle environmental and personal factors that affect gait recognition in unconstrained environments. |
| 1. [**Concurrent validity of human pose tracking in video for measuring gait parameters in older adults: a preliminary analysis with multiple trackers, viewing angles, and walking directions**](https://jneuroengrehab.biomedcentral.com/articles/10.1186/s12984-021-00933-0) | Explores the validity of using computer vision-based human pose tracking for measuring gait parameters in older adults, comparing various trackers and viewing angles for accuracy. | Evaluates different computer vision trackers for measuring gait in older adults, which has not been extensively studied before. | OpenPose, PoseNet, DeepLabCut, Multi-angle video tracking | Demonstrated that certain trackers are effective for gait analysis in older adults, with limitations based on viewing angles and walking directions. | Some trackers struggled with accuracy depending on the camera angle and walking direction. | Further studies to evaluate the performance of trackers under varied conditions, including real-world clinical applications for older adults. | Research is needed to evaluate the performance of human pose tracking in more diverse clinical environments, such as different patient groups or settings. |
| 1. [**Accuracy of Computer Vision-Based Pose Estimation Algorithms in Predicting Joint Kinematics During Gait**](https://www.researchsquare.com/article/rs-3239200/v1) | Assesses the accuracy of pose estimation algorithms in predicting joint kinematics during gait, comparing algorithms against marker-based motion capture systems. | Focuses on comparing computer vision-based algorithms with traditional motion capture methods for joint kinematics prediction. | OpenPose, AlphaPose, Marker-based motion capture systems | Found that computer vision-based algorithms show promise but are less accurate than marker-based systems, especially in high-motion activities. | Lower accuracy in high-motion scenarios, where joint kinematics predictions are more complex and error-prone. | Improve the robustness of pose estimation algorithms, especially in high-motion and complex gait activities, for real-world applications. | Needs improved accuracy in high-motion or complex scenarios, especially in sports or medical applications where detailed kinematics are crucial. |
| 1. [**Two-dimensional video-based analysis of human gait using pose estimation**](https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1008935#sec001) | This study analyzes human gait using a two-dimensional video-based pose estimation tool (OpenPose) and compares it with 3D motion capture. | Demonstrates the effectiveness of using 2D video-based pose estimation (OpenPose) for clinical gait analysis. | OpenPose, Motion Capture, Google Colaboratory | Achieved comparable results with 3D motion capture in detecting gait events like heel-strikes and toe-offs with high temporal accuracy. | Struggles with false-positive detections and challenges in extreme motion conditions. | Refine pose estimation methods and workflows to better handle real-world clinical applications. | More robust solutions needed for better tracking accuracy in different lighting, viewing angles, and during dynamic movements. |
| 1. [**Gait recognition using spatio-temporal silhouette-based features**](https://www.spiedigitallibrary.org/conference-proceedings-of-spie/8755/87550R/Gait-recognition-using-spatio-temporal-silhouette-based-features/10.1117/12.2017950.full) | Focuses on a silhouette-based method for extracting spatio-temporal features for human gait recognition. | Novel integration of silhouette-based feature extraction with spatio-temporal methods for improved recognition accuracy. | Silhouette extraction, Spatio-temporal feature analysis | Achieves significant improvement in gait recognition accuracy, particularly in controlled environments with clear silhouette data. | Limited performance in occlusion, low-resolution scenarios, and environmental changes that affect silhouette clarity. | Enhance silhouette extraction algorithms to work better in noisy or occluded environments and improve real-time processing capabilities. | Needs to address challenges in real-time applications where silhouette clarity is compromised. |
| 1. [**Ubiquitous Gait Analysis through Footstep-Induced Floor Vibrations**](https://www.mdpi.com/1424-8220/24/8/2496) | This paper introduces a novel method to recognize gait patterns by analyzing footstep-induced floor vibrations. | Innovative use of footstep-induced floor vibrations for gait analysis, allowing contact-free, non-intrusive measurements. | Vibration sensors, Frequency analysis | Successfully identifies gait patterns based on footstep vibrations, offering an alternative to video-based or wearable sensor systems. | Requires precise sensor calibration and struggles in environments with external vibrations or noise. | Further refinement needed to reduce interference from environmental vibrations and improve sensor accuracy for more diverse applications. | Current limitations in sensor technology hinder application in highly dynamic or noisy environments. |
| 1. [**Quantitative and Qualitative Running Gait Analysis through an Innovative Video-Based Approach**](https://www.mdpi.com/1424-8220/21/9/2977) | Proposes a video-based system for analyzing running gait by measuring spatio-temporal parameters and joint angles. | Introduces a combination of quantitative and qualitative measures for running gait analysis using easily accessible video tools. | Video analysis, Joint angle estimation, Spatio-temporal gait metrics | Provides accurate joint kinematic measurements and offers insights into running efficiency and injury prevention. | Difficulties arise in high-speed movements and motion blur, impacting accuracy in fast-paced running scenarios. | Improve algorithms for high-speed movements and increase robustness against motion blur to enhance real-world sports and clinical applications. | Needs further development for real-time and high-speed applications where precise joint kinematics are critical. |
| 1. [**Gait symmetry methods: Comparison of waveform-based Methods and recommendation for use**](https://www.sciencedirect.com/science/article/abs/pii/S1746809419302241) | This paper presents a comparison of waveform-based symmetry methods for assessing gait symmetry and recommends the most suitable methods for different experimental protocols. | Comprehensive comparison of various gait symmetry methods, analyzing factors like phase shift and signal length. | Trend method, Cyclogram, Symbolic method, Cross-Correlation method | Recommends the trend method as the most consistent across different conditions; other methods are context-dependent. | Some methods struggled with signal phase shifts, environment noise, and differences in experimental protocols. | Further development of gait symmetry assessment tools that can better handle real-world noisy environments and signal inconsistencies. | Lack of a universal standard for assessing gait symmetry across different protocols and noise levels in practical settings. |
| 1. [**Sensor-Based Human Activity Recognition with Spatio-Temporal Deep Learning**](https://www.mdpi.com/1424-8220/21/6/2141) | Introduces a sensor-based system for human activity recognition using spatio-temporal deep learning networks. | Combines spatio-temporal deep learning with sensor-based data, providing a robust solution for human activity recognition in real-time. | Spatio-Temporal Convolutional Networks, Wearable sensors, Deep learning | Achieved high accuracy in real-time activity recognition, outperforming previous sensor-based methods. | Performance drops when dealing with noisy sensor data or complex multi-activity scenarios. | Explore more complex activities and environmental settings to further improve real-world applicability and robustness of the recognition model. | Requires enhanced noise handling and multi-activity recognition in more dynamic real-world environments with complex movements. |
| 1. [**Multimodal Low Resolution Face and Frontal Gait Recognition from Surveillance Video**](https://www.mdpi.com/2079-9292/10/9/1013) | Proposes a multimodal recognition approach combining low-resolution facial features and frontal gait patterns for better recognition in surveillance scenarios. | Novel integration of low-resolution facial and gait features for enhanced person identification in challenging surveillance environments. | Facial recognition, Frontal Gait analysis, Surveillance video | Demonstrates improved recognition accuracy in low-resolution surveillance footage, combining two biometric traits for robust identification. | Limited by the resolution of surveillance footage and sensitivity to occlusion and environmental changes. | Improve algorithms to handle occlusions, extreme lighting, and further test in diverse surveillance conditions for improved real-time identification. | Requires further exploration of low-resolution recognition in more diverse surveillance environments, including better handling of occlusions and low-light. |
| 1. [**Real-Time Human Recognition at Night via Integrated Face and Gait Recognition Technologies**](https://www.mdpi.com/1424-8220/21/13/4323) | Focuses on integrating face and gait recognition technologies for human identification in low-light or nighttime environments. | Introduces a hybrid model that combines face and gait recognition for reliable identification in nighttime or low-light conditions. | Infrared cameras, Gait recognition, Facial recognition | Achieved high recognition accuracy in low-light conditions, providing a reliable solution for nighttime surveillance scenarios. | Limited by the range and accuracy of infrared cameras, especially in extremely low-light or complex environments. | Explore new ways to enhance face and gait recognition in more extreme low-light conditions, including enhancing infrared camera sensitivity. | More research needed on improving recognition accuracy and range in very low-light and high-motion conditions, especially for long-range identification. |
| 1. [**Multimodal Adaptive Fusion of Face and Gait Features using Keyless attention based Deep Neural Networks for Human Identification**](https://ar5iv.labs.arxiv.org/html/2303.13814) | Proposes a keyless attention-based method for dynamically fusing face and gait features for human identification. | The keyless attention mechanism dynamically assigns attention weights to facial and gait features for improved accuracy. | Keyless attention, Face and Gait fusion, Convolutional LSTM, Softmax | Achieves superior performance on the CASIA Gait Dataset-A, with significant improvements in multimodal fusion-based human identification tasks. | Limited to specific datasets and dependent on clean video inputs; performance may decrease in noisy or real-world surveillance conditions. | Future research could focus on improving the model's robustness in varied, real-world conditions, including occlusion and changing lighting environments. | Further testing is required to validate the keyless attention method in complex, real-world surveillance settings with noisy or incomplete data. |
| 1. [**Gait Recognition by Jointing Transformer and CNN**](https://link.springer.com/chapter/10.1007/978-981-99-8565-4_30) | This paper combines the strength of Transformer networks for capturing global dependencies with CNN for local feature extraction to improve gait recognition accuracy in various scenarios. | Joint use of Transformer networks with CNN offers a more effective method for recognizing gait patterns by capturing both global and local features. | CNN, Transformer networks | Demonstrates significant improvement in gait recognition across different walking angles and environments when compared to traditional methods. | Computationally expensive and slower due to the combined complexity of CNN and Transformer layers, which may hinder real-time applications. | Optimizing the hybrid model to make it more efficient for real-time use without sacrificing recognition accuracy in practical applications. | The combination of CNN and Transformer introduces computational challenges that need to be addressed for practical deployment in real-time systems. |
| 1. [**Exploring Deep Models for Practical Gait Recognition**](https://ar5iv.labs.arxiv.org/html/2303.03301) | Investigates deep learning models such as ResNet and EfficientNet for practical gait recognition, focusing on real-world application scenarios like varying camera angles, clothing, and occlusions. | Provides insights into using state-of-the-art deep models for recognizing gaits in more practical and less controlled environments. | ResNet, EfficientNet, Deep learning | Achieves strong recognition results even under challenging conditions, such as changes in clothing or camera angles, showing promise for practical use cases. | Performance drops when dealing with occlusion or complex movement patterns, and more testing is required to ensure consistent accuracy across all scenarios. | Further investigation into handling complex movements, occlusion, and multi-view recognition in dynamic, real-world environments for better robustness. | Current limitations in handling occlusion and complex scenarios restrict the real-world applicability of these deep learning models for gait recognition. |
| 1. [**GaitPT: Skeletons Are All You Need For Gait Recognition**](https://ar5iv.labs.arxiv.org/html/2308.10623) | Proposes a novel approach using only skeletal data for gait recognition, emphasizing the importance of skeleton-based models for accurate identification in both controlled and real-world settings. | Emphasizes skeleton-only data for efficient and accurate gait recognition, removing the need for additional biometric cues like face or clothing. | Skeleton-based models, Pose estimation, Gait recognition | Achieves high recognition accuracy using only skeletal data, making it efficient for use in both constrained and unconstrained environments. | Limited performance in cases where the skeleton data is noisy or incomplete, and further refinement is needed to handle more complex environments. | Explore ways to improve the model's robustness in handling noisy or incomplete skeleton data for wider real-world applicability, especially in surveillance. | Needs further testing in real-world surveillance conditions to validate its effectiveness using only skeleton-based gait recognition in varied environments. |
| 1. [**A model-based gait recognition method with body pose and human prior knowledge**](https://www.sciencedirect.com/science/article/abs/pii/S003132031930370X) | Introduces a model-based method that leverages body pose and prior human knowledge for accurate gait recognition. | Combines model-based gait recognition with human body pose data to enhance recognition accuracy in various conditions. | Human body pose estimation, Model-based gait recognition | Demonstrates significant improvements in identifying individuals based on their gait in both controlled and real-world settings. | Struggles with performance when dealing with occluded body parts or incomplete pose data. | Future work could explore better handling of occlusion and missing pose data to improve robustness in complex scenarios. | More research is required to address limitations in pose data completeness, especially in real-world surveillance conditions with occlusion. |
| 1. [**Human gait recognition: A systematic review**](https://link.springer.com/article/10.1007/S11042-023-15079-5) | This paper provides a comprehensive review of human gait recognition methods, exploring various approaches, including model-based and appearance-based techniques. | Systematically analyzes and compares different gait recognition methods, highlighting recent trends and technologies used in gait research. | Gait recognition, Model-based and appearance-based methods | Offers insights into the strengths and weaknesses of current gait recognition techniques, helping guide future research directions. | Limited by the scope of the review; focuses primarily on commonly studied datasets and methods, potentially missing novel approaches or emerging challenges. | Further systematic reviews could focus on new emerging challenges such as real-time recognition in dynamic environments. | Requires more exploration of lesser-known or emerging techniques, particularly those addressing complex environments and real-world constraints. |
| 1. [**Gait analysis for recognition and classification**](https://ieeexplore.ieee.org/abstract/document/1004148) | Focuses on gait analysis techniques for both recognizing individuals and classifying their walking patterns, examining spatio-temporal gait features for improved recognition. | Emphasizes spatio-temporal features for gait recognition and classification, combining biometric traits for more accurate identification. | Spatio-temporal gait features, Gait classification | Achieves improved classification of gait patterns based on different biometric traits, providing more accurate recognition outcomes. | Struggles when dealing with dynamic environments, especially when multiple individuals are walking in the same scene. | Exploring multi-person gait recognition and classification in real-world surveillance environments could enhance practical applications. | Needs to address challenges with handling multiple individuals in dynamic environments, such as crowded or public spaces. |
| 1. [**Human Recognition by Appearance and Gait**](https://link.springer.com/article/10.1134/S0361768818040035) | Proposes a hybrid model that combines appearance-based features with gait recognition to improve human identification in video surveillance systems. | Hybrid model that fuses appearance-based features with gait recognition for better human identification in video-based systems. | Gait recognition, Appearance-based recognition, Hybrid models | Shows strong results in combining appearance and gait features for accurate identification, especially in controlled surveillance environments. | Performance decreases significantly in low-light conditions and with occlusions, highlighting the need for better feature extraction methods in challenging settings. | Future work could focus on improving robustness in low-light and occluded environments, making the model more applicable in real-world surveillance systems. | Further development is needed to handle challenging environments such as low-light conditions or partial occlusions in real-time video surveillance systems. |
| 1. [A robust and consistent stack generalized ensemble-learning framework for image segmentation](https://www.researchgate.net/publication/372195773_A_robust_and_consistent_stack_generalized_ensemble-learning_framework_for_image_segmentation) | Proposes an ensemble-learning framework that stacks different models for image segmentation tasks. | Introduces stack generalization in ensemble-learning to improve segmentation accuracy by combining outputs from different models. | SVM, LightGBM, XGBoost | Achieves improved segmentation accuracy with reduced overfitting compared to traditional individual model approaches. | Requires extensive computational resources for training and testing in large datasets, making it less efficient for real-time applications. | Further optimization of ensemble strategies to handle large-scale data more efficiently and improve real-time segmentation accuracy. | Needs improvement in computational efficiency to make the approach suitable for real-time applications like medical imaging and large-scale image segmentation. |
| 1. [CTransCNN: Combining transformer and CNN in multilabel medical image classification](https://www.sciencedirect.com/science/article/pii/S0950705123007803) | Combines the strengths of Transformer networks for capturing global dependencies with CNN's local feature extraction for multilabel classification in medical imaging. | Novel combination of Transformer and CNN for accurate and robust classification in medical image datasets. | Transformer networks, CNN, Multilabel classification | Demonstrates high performance in multilabel classification tasks, especially in complex medical images like MRI and CT scans. | Computational complexity due to the combination of Transformer and CNN layers, which may hinder practical, real-time application. | Focus on reducing computational cost while maintaining accuracy, enabling the model for real-time medical image analysis. | The combination introduces computational challenges that need to be addressed to ensure its feasibility in real-world clinical applications. |
| 1. [Multi-Class Weed Recognition Using Hybrid CNN-SVM Classifier](https://www.mdpi.com/1424-8220/23/16/7153) | Proposes a hybrid approach using CNN for feature extraction and SVM for classification to recognize different classes of weeds in agricultural fields. | Combines CNN's ability to learn features with SVM's robust classification performance for improved weed recognition accuracy. | CNN, SVM, Image classification | Achieves high accuracy in multi-class weed recognition, offering an efficient solution for precision agriculture. | Struggles with diverse environmental conditions, especially when weeds are occluded or partially visible in images. | Developing more robust recognition models to handle diverse field conditions and different occlusion scenarios in real-time agricultural applications. | Needs further refinement in handling environmental diversity, such as weather variations and occlusions in field images. |
| 1. [Improved Prediction of Ovarian Cancer Using Ensemble Classifier and Shaply Explainable AI](https://www.mdpi.com/2072-6694/15/24/5793) | This paper presents an ensemble classifier that integrates Shaply values for better interpretability and prediction of ovarian cancer outcomes. | Integrates explainability through Shaply values in an ensemble classifier, making the AI’s predictions more interpretable and reliable for clinicians. | Ensemble learning, SHAP (Shaply values), Explainable AI | Provides higher accuracy in predicting ovarian cancer outcomes and improves trust in AI predictions through explainability for medical professionals. | The explainability module increases computational overhead, which may slow down prediction time, especially in real-time clinical scenarios. | Optimization of explainability methods to reduce computational cost without compromising the reliability and transparency of the model. | The model's explainability methods need to be improved to handle large datasets and maintain real-time performance without sacrificing interpretability. |
| 1. [Explainable AI for interpretation of ovarian tumor classification using enhanced ResNet50](https://www.mdpi.com/2075-4418/14/14/1567#:~:text=The%20proposed%20enhanced%20ResNet50%20model,belonging%20to%20the%20malignant%20classes.) | Proposes an ensemble-learning framework that stacks different models for image segmentation tasks. | Introduces stack generalization in ensemble-learning to improve segmentation accuracy by combining outputs from different models. | SVM, LightGBM, XGBoost | Achieves improved segmentation accuracy with reduced overfitting compared to traditional individual model approaches. | Requires extensive computational resources for training and testing in large datasets, making it less efficient for real-time applications. | Further optimization of ensemble strategies to handle large-scale data more efficiently and improve real-time segmentation accuracy. | Needs improvement in computational efficiency to make the approach suitable for real-time applications like medical imaging and large-scale image segmentation. |
| 1. [Understanding Robustness of Transformers for Image Classification](https://arxiv.org/abs/2103.14586) | Investigates the robustness of Vision Transformers (ViT) for image classification tasks, comparing them with ResNet models under different perturbations and real-world conditions. | Analyzes the robustness of ViTs to input and model perturbations, showing that ViTs outperform ResNets when pre-trained on large datasets. | Vision Transformers (ViT), ResNet, ImageNet-C, ImageNet-R, FGSM attacks, PGD attacks | Demonstrates that ViTs are more robust to corruptions when trained on large datasets but less robust in small data regimes compared to ResNets. | High computational cost for training large models like ViT, and performance may decrease in resource-constrained environments. | Optimizing ViTs for smaller datasets and improving robustness in low-resource scenarios could enhance their real-world applicability. | Needs further research on improving the efficiency of ViTs in low-resource settings without sacrificing robustness to adversarial attacks and corruptions. |
|  |  |  |  |  |  |  |  |
| Paper | Summary | Novelty | Technologies/  Architecture  Used | Results/  Evaluation/  Benchmark | Limitations | Future Work | Research Gap |
| 1. [**Enhancing Natural Language Query to SQL Query Generation Through Classification-Based Table Selection**](https://link.springer.com/chapter/10.1007/978-3-031-62495-7_12) | Proposes a classification-based method for selecting relevant tables during the SQL query generation process, improving accuracy in mapping complex natural language queries to SQL. | Introduces classification-based table selection to address the challenge of multi-table queries, improving query accuracy. | Classification models, SQL query generation, Natural language processing | Significantly improves the performance of natural language to SQL query generation systems, particularly in multi-table query scenarios. | Limited generalizability across diverse query structures and databases, especially when dealing with rare or complex user queries. | Focus on extending the model to handle more complex SQL structures and diverse database environments for broader applicability. | More research is needed to handle uncommon SQL structures and rare queries, especially in dynamic environments with rapidly changing datasets. |
| 1. [**Structure Guided Large Language Model for SQL Generation**](https://arxiv.org/abs/2402.13284) | Presents a structure-guided approach for enhancing SQL generation by leveraging large language models (LLMs), improving accuracy in generating complex SQL queries. | Uses a structure-guided mechanism to align SQL generation more closely with database schema, improving the reliability of SQL query results. | Large Language Models (LLMs), SQL generation, Database schema guidance | Achieves higher accuracy in generating complex SQL queries, especially when dealing with schema-specific constraints in databases. | High computational costs and potential overfitting when applied to specific database structures, limiting generalizability. | Develop methods to generalize the model across different database types and schema structures while maintaining performance. | Needs further work on reducing overfitting and improving the generalizability of structure-guided SQL generation across different types of databases. |
| 1. [**MCS-SQL: Leveraging Multiple Prompts and Multiple-Choice Selection For Text-to-SQL Generation**](https://arxiv.org/abs/2405.07467) | Proposes a multi-prompt approach for text-to-SQL generation, using multiple-choice selection to improve the accuracy of SQL queries generated from natural language inputs. | Multi-choice selection improves SQL generation accuracy by allowing the model to choose between multiple candidate SQL queries. | Multiple-choice selection, Text-to-SQL, Natural language processing | Demonstrates improved performance in text-to-SQL generation, particularly for ambiguous or complex natural language queries. | Requires significant computational resources to evaluate multiple candidate queries, which may limit its efficiency in real-time applications. | Future research could focus on improving the efficiency of multi-choice selection mechanisms to make the approach feasible for real-time query generation. | More research is needed to reduce computational complexity and improve the efficiency of evaluating multiple candidates for real-time SQL generation. |
| 1. [**Knowledge-to-SQL: Enhancing SQL Generation with Data Expert LLM**](https://arxiv.org/abs/2402.11517) | Proposes an approach to improve SQL generation by using a Data Expert Large Language Model (DELLM), which incorporates knowledge generation and reinforcement learning to enhance SQL query accuracy. | Introduces a knowledge generation model (DELLM) and uses feedback from database execution and ground-truth SQL to refine generated SQL. | Data Expert LLM, Reinforcement Learning, Supervised Fine-Tuning | Achieves higher accuracy in SQL generation by providing LLMs with domain-specific knowledge, especially in large databases. | High computational cost, particularly when handling large database schemas with multiple tables and columns. | Optimize the DELLM model to reduce computational overhead and generalize across different database structures. | Further research is required to handle complex schemas and improve computational efficiency in real-world scenarios. |
| 1. [**Prompting GPT-3.5 for Text-to-SQL with De-semanticization and Skeleton Retrieval**](https://link.springer.com/chapter/10.1007/978-981-99-7022-3_23) | This paper explores the use of de-semanticization and skeleton retrieval techniques to improve the accuracy of GPT-3.5 for text-to-SQL generation tasks, particularly for complex queries. | Uses de-semanticization to simplify natural language inputs, and skeleton retrieval to provide a structured framework for SQL generation. | GPT-3.5, De-semanticization, Skeleton retrieval | Demonstrates improved performance in generating complex SQL queries from ambiguous or vague natural language inputs. | Complexity increases with query length and database size, leading to longer response times in real-time applications. | Improve the efficiency of skeleton retrieval techniques and extend testing to larger and more complex database environments. | Further research is needed to optimize query generation in real-time applications, especially for very large databases and complex queries. |
| 1. [**RH-SQL: Refined Schema and Hardness Prompt for Text-to-SQL**](https://arxiv.org/abs/2406.09133) | RH-SQL introduces schema refinement and hardness prompting techniques to help large language models better understand complex database schemas and generate accurate SQL queries. | Schema refinement and hardness prompts allow the model to consider schema complexity and difficulty levels when generating SQL queries. | Large Language Models, Hardness Prompting, Schema refinement | Achieves significant improvements in generating accurate SQL for complex schemas, with particular benefits in multi-table queries. | The model requires fine-tuning for each database schema, which reduces its scalability across different databases. | Further work could focus on automating schema refinement for various databases to improve scalability and ease of deployment. | The model needs further testing across a wider variety of database structures and schemas to ensure generalization and scalability. |
| 1. [**Enhancing Text-to-SQL Capabilities of Large Language Models: A Study on Prompt Design Strategies**](https://aclanthology.org/2023.findings-emnlp.996.pdf) | Proposes a multiple-choice selection mechanism for text-to-SQL generation, where multiple query candidates are evaluated to select the most accurate SQL output. | Uses multiple prompts and query selection to refine the SQL output, improving accuracy in ambiguous or complex text-to-SQL generation tasks. | Multiple-choice selection, Large Language Models, Prompting | Demonstrates superior performance in handling complex text-to-SQL tasks, particularly in cases with multiple query interpretations. | Computational costs are high due to the multiple-choice evaluation process, which may limit its practical use in real-time applications. | Focus on improving the efficiency of the multiple-choice selection mechanism to make it more feasible for real-time SQL generation in larger-scale systems. | Needs further research on reducing the computational overhead of evaluating multiple candidate queries in real-time scenarios. |
| 1. [**Next-Generation Database Interfaces: A Survey of LLM-based Text-to-SQL**](https://arxiv.org/abs/2406.08426) | Provides a comprehensive review of LLM-based text-to-SQL systems, analyzing the challenges and advances in question understanding, schema comprehension, and SQL generation. | Highlights the shift from rule-based to LLM-based methods and how these models have improved accuracy in text-to-SQL tasks. | LLMs, In-context learning, Fine-tuning | Identifies that LLMs are more robust in question comprehension and schema handling but still face issues with large, complex datasets. | Scalability and efficiency remain challenges, particularly in large-scale databases where LLMs struggle to generalize across diverse schemas. | Future research should focus on improving scalability and generalizability across various database schemas while maintaining query accuracy. | More work is needed on handling complex schemas and scaling LLMs for real-time and large-scale database queries. |
| 1. [**Bridging Language & Data: Optimizing Text-to-SQL Generation in Large Language Models**](https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1833681&dswid=-1064) | Explores how text-to-SQL generation can be optimized by integrating knowledge graphs and schema-based understanding into LLM training. | Focuses on bridging the gap between natural language queries and database-specific schema understanding through optimization strategies. | Knowledge graphs, Schema-based learning, LLM optimization | Demonstrates that integrating knowledge graphs significantly improves the accuracy of text-to-SQL generation in complex queries. | The model’s performance drops when handling new and unseen schemas, indicating the need for more robust learning techniques. | Further research could explore enhancing LLMs to generalize better to unseen schemas and improve contextual accuracy. | Needs to address generalization issues in unseen database schemas and ensure that LLMs can maintain performance in diverse contexts. |
| 1. [**CodeS: Towards Building Open-source Language Models for Text-to-SQL**](https://dl.acm.org/doi/abs/10.1145/3654930) | Discusses the challenges and strategies for developing open-source language models for text-to-SQL tasks, aiming for wider accessibility and customizability of these models. | Emphasizes open-source models to allow customization for specific datasets and queries, making LLM-based systems more accessible. | Open-source LLMs, Customization for databases | Highlights the potential for open-source LLMs to outperform proprietary models in specific cases due to customization flexibility. | Open-source models may require significant domain-specific knowledge to fine-tune, limiting their utility in broad applications without expert intervention. | Focus on reducing the domain-specific knowledge required to fine-tune open-source models for text-to-SQL tasks to expand usability. | Developing open-source LLMs that require less fine-tuning and can generalize across different database contexts will be crucial for broader adoption. |
| 1. [**[2405.16755] CHESS: Contextual Harnessing for Efficient SQL Synthesis (arxiv.org)**](https://ar5iv.labs.arxiv.org/html/2405.16755) | Proposes CHESS, a system that enhances SQL synthesis by using contextual cues from both the natural language query and the database schema, leading to more accurate SQL generation. | Leverages contextual information from both user queries and database schemas to improve the efficiency of SQL synthesis. | Contextual cues, SQL synthesis, Schema understanding | Demonstrates higher execution accuracy and faster query generation by using contextual information to refine the generated SQL queries. | Limited performance when handling highly ambiguous queries where contextual cues are not sufficient to disambiguate the user’s intent. | Future work could explore integrating additional contextual layers, such as user intent and domain-specific constraints, to further improve query accuracy. | More research is needed on handling ambiguous queries with minimal context and refining SQL synthesis models for such challenging scenarios. |
| 1. [**Retrieval-Augmented GPT-3.5-Based Text-to-SQL Framework** **with Sample-Aware Prompting and Dynamic Revision Chain**](https://link.springer.com/chapter/10.1007/978-981-99-8076-5_25) | This paper proposes a framework that enhances GPT-3.5-based text-to-SQL generation by integrating sample-aware prompting and a dynamic revision chain. These techniques improve the accuracy and efficiency of SQL generation. | Utilizes a dynamic revision chain to iteratively refine SQL queries with minimal human intervention, leveraging retrieval-augmented techniques. | GPT-3.5, Sample-aware prompting, Dynamic revision chain | Demonstrates improved SQL generation accuracy and execution success by iteratively refining queries through feedback loops and sample-aware retrieval techniques. | High computational overhead due to the iterative refinement process, making it less efficient for real-time applications. | Future work could focus on optimizing the retrieval process and reducing computational complexity to make the system more feasible for real-time use. | More research is needed to improve the efficiency of the dynamic revision chain for large-scale real-time applications and reduce computational overhead. |
| 1. [**Evaluating the Text-to-SQL Capabilities of Large Language Models**](https://www.researchgate.net/publication/359709736_Evaluating_the_Text-to-SQL_Capabilities_of_Large_Language_Models) | Evaluates the performance of various large language models (LLMs) in generating SQL queries from natural language questions, comparing accuracy, execution success, and error rates. | Provides a comparative study of different LLMs for text-to-SQL generation, focusing on how well they perform across various database schemas. | Large Language Models (LLMs), SQL evaluation metrics | Shows that while LLMs perform well on simple queries, they often struggle with complex or multi-table queries, with significant variance across models. | LLMs tend to generate erroneous SQL for complex queries, and their performance varies widely depending on the database schema and the complexity of the query. | Future research should focus on improving schema understanding and query complexity handling to enhance the accuracy of SQL generation in complex databases. | Requires more exploration of how to improve LLMs' performance on complex, multi-table queries and better schema comprehension in real-world scenarios. |
|  |  |  |  |  |  |  |  |