Nuvern International

Multi-Sensor Data Fusion and Management Strategies for Robust Perception in Autonomous Vehicles

Meiling Sun[1]

[1]**Hunan Institute of Engineering, Department of Software Engineering, Xiangtan, Hunan, China**

ABSTRACT

Multi-sensor data fusion has emerged as a critical enabler for robust perception in autonomous vehicles, where the reliability and accuracy of environmental understanding directly impact operational safety and efficiency. This paper presents a comprehensive investigation of cutting-edge data fusion approaches and management strategies that address the challenges of sensor heterogeneity, dynamic driving conditions, and computational constraints. We examine various sensor modalities, including LiDAR, radar, and camera systems, and discuss the advantages of combining their complementary strengths to enhance perception and situational awareness. By exploring state-of-the-art algorithms that integrate machine learning models with probabilistic filtering techniques, we illustrate how high-fidelity maps and real-time sensing can be synchronized to form a unified representation of the environment. Detailed mathematical formulations highlight the role of complex transformations and linear algebraic structures in the data alignment and calibration process. Furthermore, we analyze methods for mitigating sensor uncertainties and propose strategies to handle data overload and synchronization issues under real-time constraints. We present approaches for robust machine learning model design, where domain adaptation and multi-task learning methods enable flexible perception pipelines that generalize to diverse traffic and weather conditions. Ultimately, we identify open research directions and highlight the significance of scalable, secure, and adaptive data management in propelling autonomous vehicle perception forward.

1 INTRODUCTION

Autonomous vehicle perception relies upon a synergy of hardware and software that must process vast amounts of environmental data with stringent real-time and safety requirements [1]. The fusion of multiple sensors has become one of the most fundamental strategies to enhance reliability and to provide the vehicle with a thorough understanding of its surroundings [2]. Specifically, LiDAR, radar, and camera systems form the core elements of the sensor array, each with unique advantages and limitations. LiDAR excels in generating high-resolution 3D point clouds, radar exhibits robustness in adverse weather conditions, and cameras capture nuanced color and texture details indispensable for object classification [3]. However, the challenges of integrating these varied sensor streams remain formidable, demanding both advanced algorithmic solutions and optimal system architecture designs.

An effective fusion mechanism must address spatial and temporal misalignments arising from distinct sensor placement, sampling rates, and operational fields of view [4]. These sources of misalignment lead to non-negligible inaccuracies that can cascade through perception pipelines, thereby affecting fundamental tasks such as object detection, tracking, and trajectory estimation. Sensor noise and dynamic factors, such as changes in illumination or rapidly shifting environmental clutter, introduce additional uncertainty in measurements [5]. To manage and reduce these uncertainties, robust Bayesian filtering and state estimation algorithms are commonly employed [6]. It is crucial to formulate them in a mathematically rigorous framework that ensures stable integration of data and enhances reliability.

Efficient data management is pivotal given the high-bandwidth streams that drive contemporary perception systems [7]. The design of real-time data buffering, caching, and processing pipelines forms the backbone of modern autonomous platforms. While hardware acceleration and parallel computing strategies have advanced, they require software-level coordination that is sensitive to the unique demands of multi-sensor fusion [8]. Large-scale deployment of such technology in fleets necessitates scalable data architectures that can handle petabyte-level data volumes from daily operations, as well as robust encryption and anonymization protocols to protect sensitive driver information.

A key objective of multi-sensor fusion is to create a holistic understanding of the environment in the presence of occlusions and unpredictable events [9]. When a single sensor suffers from limitations, additional sensors help fill in the gaps [10]. For instance, cameras might fail in low light, but radar maintains reliable detection; LiDAR point clouds can assist in creating precise geometry, while cameras differentiate intricate texture patterns. Incorporating advanced machine learning techniques, such as deep convolutional networks for computer vision tasks, provides powerful means for feature extraction and classification [11]. Nevertheless, these methods demand substantial labeled data and computational resources, complicating integration with resource-constrained embedded systems. Consequently, there is an ongoing need to balance the benefits of deep learning with the realities of latency and energy constraints in autonomous vehicles. [12]

This paper delves into the intricacies of multi-sensor data fusion and data management techniques, beginning with a discussion of the major sensor modalities and their individual attributes. We investigate the mathematical models underpinning data fusion, including transformations, state estimators, and optimization procedures [13]. We then shift our focus to advanced machine learning frameworks and the role of domain adaptation, in addition to emphasizing sensor calibration and error mitigation methods [14]. We also explore techniques for data organization and retrieval that enable effective real-time perception. The final section offers a conclusion, articulating emerging directions for research that can further fortify robust perception in autonomous vehicles. [15]

2 SENSOR MODALITIES FOR AUTONOMOUS VEHICLE PERCEPTION

Sensor modalities used in autonomous vehicle perception serve as the primary conduit for environmental data capture, each providing unique but complementary streams of information. Cameras furnish high-resolution visual detail, which, when processed through advanced image recognition methods, facilitates the identification of traffic signs, lane markings, and the nuanced shapes of objects [16]. However, camera-based perception is subject to sensitivity in challenging illumination or weather conditions, leading to potential degradation of performance. Meanwhile, radar systems are often favored for their capacity to function efficiently in rain, fog, or other visibility-compromising scenarios [17]. Radar emits radio waves that reflect off objects, enabling the calculation of object velocity and distance [18]. Nonetheless, radar data typically exhibit lower spatial resolution than that of cameras, making it less effective for precise shape delineation.

LiDAR has surged in popularity due to its ability to generate highly detailed 3D point clouds [19]. By measuring the time-of-flight of laser pulses, LiDAR offers accurate distance measurements that yield an explicit geometric representation of the surroundings. Despite their high-resolution range profiles, LiDAR sensors can be sensitive to adverse weather conditions and typically have higher cost and power requirements compared to other sensors [20]. Integrating LiDAR data with camera imagery allows for rich scene interpretation that benefits from both geometric and appearance-based features. For instance, an autonomous system could exploit camera inputs for semantic recognition (distinguishing bicycles from pedestrians) while using LiDAR for precise distance estimation. [21]

The synergy of these sensor modalities, however, cannot be fully harnessed unless data alignment issues are addressed [22]. This alignment, often referred to as sensor calibration, entails verifying that the intrinsic and extrinsic parameters of each sensor are well-determined. Intrinsic parameters, such as focal length or lens distortion for cameras, and range resolution or beamwidth for radar and LiDAR, dictate how raw measurements are generated [23]. Extrinsic parameters characterize the geometric relationship between the sensor and the vehicle frame. These relationships are typically expressed using transformation matrices in homogeneous coordinates [24]. For example, if the transformation from the LiDAR frame to the camera frame is denoted by a matrix $\mathbf{T}_{LC} \in \mathbb{R}^{4\times4}$, then a LiDAR point in homogeneous coordinates $\mathbf{p}_L$ is mapped to the camera coordinate system by $\mathbf{p}_C = \mathbf{T}_{LC}\mathbf{p}_L$. Such transformations must be determined accurately to merge point cloud data with 2D images, aligning geometric and visual information for subsequent perception tasks.

Practical considerations in sensor modality selection extend beyond mere performance metrics and include cost, physical size, power consumption, and manufacturing constraints [25]. High-end LiDAR devices may provide superior detail but can be prohibitively expensive or bulky. Forward-facing radars, in contrast, are more common in consumer vehicles, capitalizing on a better cost-to-performance ratio [26]. Automotive-grade cameras have become increasingly sophisticated, with advanced features such as wide dynamic range imaging and high frame rates, enabling them to address many perception tasks effectively if the software stack is sufficiently robust [27]. Ultimately, an intricate understanding of sensor modality strengths and limitations forms the bedrock for designing fusion architectures that yield robust, real-time perception for autonomous vehicles.

3 DATA FUSION TECHNIQUES: A MULTI-LAYER APPROACH

Effective data fusion in autonomous vehicles requires systematically reconciling and blending diverse sensor streams into a cohesive and actionable representation of the environment [28]. A widely employed approach is hierarchical or multi-layer fusion, where sensor measurements are integrated at multiple stages, ranging from early low-level feature extraction to late high-level decision aggregation. Early fusion strategies often operate on raw sensor data [29]. This can include aggregating LiDAR point clouds and camera pixel data before feature extraction, enabling cross-sensor correlation at an elementary level. While early fusion can yield a richly detailed representation, it can be computationally burdensome due to the high dimensionality of raw sensor data. [30]

Mid-level fusion is often performed on feature representations extracted from each sensor individually [31]. For instance, if a camera-based convolutional neural network extracts a set of high-level features for semantic classification, and a LiDAR-based 3D object detection pipeline extracts geometric descriptors, a mid-level fusion module can integrate these features in a single latent space. This process can be implemented via linear concatenation of the feature vectors or through more sophisticated approaches, such as attention mechanisms that weight features differently based on their reliability [32]. By fusing features rather than raw data, mid-level approaches can reduce computational overhead while preserving important semantic and geometric cues.

Late fusion, in contrast, aggregates decisions or probabilistic estimates from sensor-specific modules [33]. For example, individual object detection algorithms might produce separate lists of bounding boxes with associated confidence levels, which are then reconciled in a late fusion stage. This method can be advantageous in modular system designs, as it allows each sensor-specific module to operate independently with specialized algorithms [34]. However, late fusion may overlook fine-grained correlations among sensor data that earlier integration might exploit more effectively [35]. Another consideration is that late fusion can reduce failure independence, because errors in one sensor-specific module may propagate to the final decision layer without the possibility of corrective cross-sensor data interactions at an earlier stage.

Statistical and probabilistic frameworks have proven instrumental for data fusion [36]. Kalman filters and their variants, such as the Extended Kalman Filter (EKF) and the Unscented Kalman Filter (UKF), are classical approaches for recursively estimating the states of dynamic objects using a combination of motion models and sensor measurements. If $\mathbf{x}_k$ represents the state vector at time step k and $\mathbf{z}_k$ represents the aggregated measurements, then a general fusion strategy can be summarized as:

$$\mathbf{x}_k = \arg\max_{\mathbf{x}} p(\mathbf{x} \,|\, \mathbf{z}_{1:k}) \quad = \arg\max_{\mathbf{x}} p(\mathbf{z}_k \,|\, \mathbf{x})\, p(\mathbf{x} \,|\, \mathbf{z}_{1:k-1}).$$

In the linear Gaussian case, the standard Kalman filter equations apply, with the fused estimate updating through the innovation term $\mathbf{z}_k - \mathbf{H}\mathbf{x}_k^-$, where $\mathbf{H}$ is a measurement model matrix, and $\mathbf{x}_k^-$ is the predicted state prior to assimilation of the new measurement. In practice, sensor modalities may yield non-linear or non-Gaussian measurement likelihoods, necessitating advanced variants such as the UKF or particle filters. [37]

Beyond Bayesian filtering, contemporary machine learning-based fusion approaches continue to gain traction. For example, a neural network architecture might learn sensor-specific embeddings and fuse them by combining latent representations through gated or attention-based fusion layers [38]. These learned fusion methods can adapt to diverse data distributions and complex noise models [39]. Nonetheless, they require massive labeled datasets and can be sensitive to distributional shifts experienced when the autonomous vehicle moves between distinct geographical regions or weather conditions. Techniques such as domain adaptation, adversarial training, or self-supervised learning can help mitigate these challenges, enabling models to maintain reliable performance in a variety of real-world settings. [40]

4 MACHINE LEARNING MODELS FOR DATA FUSION

Machine learning models that handle heterogeneous sensor data play a pivotal role in modern data fusion pipelines. Convolutional neural networks (CNNs) are primarily utilized for image data, extracting hierarchical features from pixel-level inputs in a manner that preserves spatial locality [41]. Recurrent neural networks (RNNs), including gated recurrent units or long short-term memory networks, can track temporal dependencies in time-series data, such as radar or LiDAR scans over time. More recently, attention-based transformers have demonstrated remarkable success in modeling sequence data and multimodal signals [42]. Transformers rely on self-attention mechanisms to assign context-dependent weights to different parts of an input sequence, thereby capturing long-range correlations with fewer assumptions about sequential ordering. [43]

In the context of multi-sensor fusion, neural architectures can be designed to accept fused or separate streams of sensor inputs. For instance, LiDAR point clouds can be encoded using specialized layers such as PointNet or voxel-based 3D CNNs, while images pass through a 2D CNN backbone [44]. The resultant feature embeddings from LiDAR and camera can be concatenated or merged in an attention layer that learns the optimal weighting. Suppose we denote $\mathbf{f}_{\text{LiDAR}}$ as the feature vector extracted from LiDAR data and $\mathbf{f}_{\text{camera}}$ as that from camera data. A fusion mechanism could be described by [45]

$$\mathbf{f}_{\text{fusion}} = \sigma\big(W_1\,\mathbf{f}_{\text{LiDAR}} + W_2\,\mathbf{f}_{\text{camera}}\big),$$

where W_1 and W_2 are trainable parameters, and σ is a non-linear activation function. More sophisticated layers might weigh each feature dimension differently or incorporate learned cross-attention factors to dynamically adjust the fusion process based on the confidence of each sensor modality.

Once a unified representation is obtained, the subsequent task can vary from semantic segmentation of the surrounding environment to predicting the trajectories of nearby objects or planning the vehicle's motion [46]. End-to-end learning frameworks have emerged that integrate all of these tasks into a single deep learning pipeline. However, this holistic approach often demands an enormous quantity of ground-truth data for supervised training and remains challenging to validate in safety-critical scenarios [47]. Domain adaptation techniques aim to address the problem of limited labeled data by transferring knowledge from domains where ample annotated data exist, such as simulation environments or regions with abundant data, to new operational domains [48]. For instance, a domain adaptation procedure might calibrate the feature extractor to align distributions of simulated LiDAR scans with real-world LiDAR scans, thus allowing the network to exploit a large synthetic dataset.

Reinforcement learning has also been explored for sensor fusion tasks, particularly when the objective is not solely to reconstruct or classify the environment but to make control decisions [49]. In these approaches, a policy network might take sensor measurements or their fused feature representations as input, and output the next steering or acceleration command. The reward function in a reinforcement learning framework can incorporate measures of perception quality, collision avoidance, and passenger comfort [50]. Although such methods show promise, their training requires simulation or restricted environments, and ensuring reliability in open-world driving remains an area of active research. As sensor technology evolves, incorporating advanced models that adapt and scale effectively in real-world conditions will be instrumental for robust autonomous vehicle perception. [51]

5 SENSOR CALIBRATION AND ERROR MITIGATION

Calibration and error mitigation are imperative for achieving high-fidelity sensor fusion that translates into accurate perception and control [52]. Even slight misalignments in extrinsic or intrinsic parameters can degrade performance, leading to systematic offsets in object localization or an increased rate of false detections. Calibration processes typically involve a set of known reference patterns or simultaneously visible features in the environment [53]. For cameras, internal parameters such as focal length and principal point can be determined by imaging a chessboard pattern at different orientations, while LiDAR can be calibrated using reflectors at known distances and angles. Multi-sensor calibration often leverages structures visible to multiple sensors [54]. For instance, corner reflectors that produce high-intensity returns in both LiDAR and radar data can facilitate alignment of those frames.

Mathematically, calibration can be formulated as an optimization problem [55]. If $\mathbf{T}_{AB}$ denotes the transformation matrix mapping points from sensor A to sensor B, and $\{(\mathbf{p}_A^{(i)}, \mathbf{p}_B^{(i)})\}$ is a set of corresponding points detected in both sensors, then calibration can be approached by minimizing a cost function of the form

$$\min_{\mathbf{T}_{AB}} \sum_i d\big(\mathbf{T}_{AB}\mathbf{p}_A^{(i)}, \mathbf{p}_B^{(i)}\big),$$

where d is a distance metric appropriate for the sensor data type [56]. For point cloud-based modalities, the cost could be Euclidean distance. For images, the distance might involve reprojected pixel coordinates [57]. Techniques such as Iterative Closest Point (ICP) for point clouds or bundle adjustment for camera images can refine the sensor transformation by iteratively reducing alignment errors. Furthermore, continuous online calibration methods employ adaptive filters or stochastic gradient updates to cope with mechanical vibrations or temperature fluctuations that might alter sensor alignments over time. [58]

Noise and error mitigation is also crucial. Each sensor exhibits a distinct noise profile, which can be characterized statistically or through empirical testing [59]. For LiDAR, random measurement noise might be modeled with Gaussian distributions in radial distance, whereas radar signals could be subject to speckle noise [60]. Camera sensors often have pixel intensity noise sensitive to lighting conditions. When fusing data, the sensor noise parameters are integrated into the state estimation algorithm [61]. In an extended Kalman filter framework, for example, the measurement covariance matrix $\mathbf{R}$ encapsulates the sensor's uncertainty. By tuning $\mathbf{R}$ to reflect empirical noise characteristics, the filter can more accurately weigh incoming measurements during updates.

A further layer of error mitigation involves outlier detection and handling. In cluttered or dynamic environments, spurious reflections or occlusions may corrupt measurements [62]. Statistical methods such as Random Sample Consensus (RANSAC) can identify and remove outliers in geometric fitting tasks, while robust cost functions dampen the effect of large measurement residuals in filter updates. Techniques based on machine learning, such as anomaly detection or confidence estimation networks, can provide sensor-specific or scene-specific measures of data fidelity, allowing the fusion engine to disregard or discount unreliable measurements [63]. Taken together, these calibration and error mitigation strategies form the foundation for maintaining a coherent and accurate sensor fusion pipeline over the long operational life of autonomous vehicles. [64]

6 DATA MANAGEMENT STRATEGIES FOR REAL-TIME PROCESSING

Data management strategies underpin the seamless operation of multi-sensor fusion pipelines in autonomous vehicles, where latency, bandwidth, and storage are tightly constrained. Real-time processing necessitates the efficient orchestration of incoming data streams, from sensor-level readouts to the final decision modules [65]. At the hardware layer, this coordination often involves parallel processing architectures, leveraging multi-core CPUs, GPUs, or specialized accelerators like FPGAs or TPUs. However, hardware acceleration alone is insufficient [66]. Robust middleware and communication frameworks are required to guarantee timely and reliable data transfers among the sensors, perception modules, and control units. Publish-subscribe protocols, message queues, and shared memory approaches can be orchestrated by centralized or decentralized controllers to streamline data flows. [67]

One challenge is the sheer data volume produced by modern high-resolution sensors. Advanced cameras can generate gigabytes of data per second, and LiDAR point clouds can quickly reach tens or hundreds of megabytes per second [68]. Consequently, data buffering strategies are employed to accommodate fluctuations in sensor output and to ensure consistent processing throughput [69]. Ring buffers or sliding windows are commonly used, especially in short-term data retention for tasks like temporal filtering or multi-frame object tracking. These buffers must be carefully sized to avoid underflow or overflow events that could compromise data completeness or introduce latency. [70]

Compression and subsampling techniques are often introduced to reduce network load, but these must be balanced against the risk of losing critical spatial or temporal details required for reliable perception. For instance, LiDAR point cloud compression might employ quantization or voxelization that merges adjacent points, reducing resolution yet retaining the broad structure of the scene [71]. Similarly, image compression must preserve essential features, especially in regions containing potential obstacles, lane markings, or pedestrians. A sophisticated approach is content-aware compression that allocates more bits to regions of interest, guided by neural networks or heuristics. [72]

Data synchronization is another critical aspect, ensuring that sensor measurements used in fusion correspond to the same or closely matching time stamps [73]. Minor temporal misalignments, even on the order of milliseconds, can create significant discrepancies when fusing data from moving vehicles or objects. Timestamping mechanisms that rely on GPS or high-precision clocks can aid synchronization [74]. Alternatively, software-level synchronization aligns data streams through interpolation or extrapolation based on motion models. If $\mathbf{z}_{\text{LiDAR}}(t_L)$ and $\mathbf{z}_{\text{camera}}(t_C)$ are sensor measurements at different timestamps, it may be feasible to estimate $\mathbf{z}_{\text{LiDAR}}(t_C)$ by applying a predictive model. Mathematically, one might use a motion update function $\mathbf{f}$ to bridge the time gap:

$$\mathbf{z}_{\text{LiDAR}}(t_C) \approx \mathbf{f}\big(\mathbf{z}_{\text{LiDAR}}(t_L), t_C - t_L\big).$$

While this approach can mitigate minor misalignments, larger time gaps or abrupt movements can degrade performance. [75]

Long-term data management addresses the accumulation of data for training, validation, and offline analysis. Autonomous vehicle fleets may collectively produce petabytes of data daily, raising questions about storage architectures and cloud-based infrastructures [76]. In many cases, raw data are stored for a limited time, with only selected subsets (e.g., crash scenarios or edge cases) archived for subsequent investigation [77]. Policies for data retention and retrieval must meet privacy and security standards, particularly in regions with stringent data protection regulations. Anonymization of collected data is often required, including obscuring faces and license plates in captured images, or de-linking trajectory data from unique identifiers [78]. Data access control and encryption protocols ensure that only authorized parties can retrieve sensitive information, protecting both end-user privacy and proprietary algorithms.

Ultimately, the efficacy of a data management strategy in autonomous vehicles is judged by its ability to deliver the right data at the right time with sufficient accuracy for safety-critical decisions [79]. Future trends such as distributed edge computing, vehicular cloud offloading, and incremental learning will further complicate data management paradigms. Nevertheless, building a robust real-time pipeline that efficiently manages sensor data remains central to enabling the continued evolution and scaling of autonomous vehicle fleets worldwide. [80]

7 CONCLUSION

Multi-sensor data fusion and robust data management strategies are pivotal for enabling autonomous vehicles to perceive and navigate complex traffic environments with reliability, safety, and efficiency [81]. By carefully selecting and integrating complementary sensor modalities, autonomous systems can compensate for individual sensor weaknesses, achieving a more accurate and comprehensive understanding of their surroundings. Mathematical modeling and optimization, spanning linear algebraic transformations to complex Bayesian filtering, form the theoretical underpinnings that guide the design of fusion algorithms [82]. Advances in machine learning architectures, including attention-based transformers and domain adaptation methods, have introduced unprecedented adaptability to heterogeneous sensor data, though challenges in labeling, distribution shift, and real-time performance persist.

Accurate sensor calibration stands as a cornerstone of successful fusion pipelines, ensuring the alignment of multimodal measurements and reducing systematic biases that might jeopardize downstream perception tasks [83]. Likewise, noise characterization and outlier detection methods bolster data integrity by refining the reliability of sensor inputs. Within this framework, real-time data management strategies address the staggering bandwidth and synchronization demands typical of advanced sensor configurations [84]. The interplay between specialized hardware accelerators, efficient buffering and compression, and robust communication protocols fosters a data flow that must consistently meet the stringent latency requirements of on-road driving scenarios. [85]

Notably, the integration of these technical solutions is governed by overarching considerations in cost, scalability, and compliance with privacy and security regulations. Sophisticated anonymization and encryption methods ensure that the vast quantities of driving data collected do not compromise individual or fleet-wide confidentiality [86]. Furthermore, as sensor technologies evolve and new modalities such as advanced thermal cameras or high-frequency scanning LiDARs gain traction, existing fusion frameworks must adapt to changing data distributions and new functional requirements. This ongoing evolution underscores the importance of flexible architectures and thorough validation strategies. [87]

In the future, deeper levels of cooperation among vehicle platforms and infrastructure, leveraging vehicle-to-vehicle and vehicle-to-infrastructure communication channels, may further enrich the data available for fusion. As global research in artificial intelligence continues to progress, novel techniques in self-supervised learning and incremental model updates show significant promise in resolving persistent issues of training data scarcity and rapidly changing operational environments [88]. Ultimately, forging more resilient, adaptive, and scalable data fusion pipelines stands as a primary objective to guarantee that autonomous vehicles can confidently and reliably operate in diverse real-world conditions. The strategies reviewed and elaborated here represent a robust foundation upon which subsequent innovations in perception can build, advancing the realization of fully self-driving systems that are both safe and efficient. [89]

REFERENCES

[1] L. Daming, L. Deng, and Z. Cai, "Design of traffic object recognition system based on machine learning," *Neural Computing and Applications*, vol. 33, no. 14, pp. 8143–8156, Apr. 30, 2020. DOI: 10.1007/s00521-020-04912-9.

[2] Q. Zheng, X. Tian, M. Yang, and H. Su, "Clmip: Cross-layer manifold invariance based pruning method of deep convolutional neural network for real-time road type recognition," *Multidimensional Systems and Signal Processing*, vol. 32, no. 1, pp. 239–262, Jul. 17, 2020. DOI: 10.1007/s11045-020-00736-x.

[3] M. H. Javed, K. Z. Ibrahim, and X. Lu, "Performance analysis of deep learning workloads using roofline trajectories," *CCF Transactions on High Performance Computing*, vol. 1, no. 3, pp. 224–239, Nov. 29, 2019. DOI: 10.1007/s42514-019-00018-4.

[4] M. Muniswamaiah, T. Agerwala, and C. Tappert, "Data virtualization for analytics and business intelligence in big data," in *CS & IT Conference Proceedings*, CS & IT Conference Proceedings, vol. 9, 2019.

[5] C. Sindermann, P. Sha, M. Zhou, *et al.*, "Assessing the attitude towards artificial intelligence: Introduction of a short measure in german, chinese, and english language," *KI - Künstliche Intelligenz*, vol. 35, no. 1, pp. 109–118, Sep. 23, 2020. DOI: 10.1007/s13218-020-00689-0.

[6] M.-J. Yang, D. Ketcher, K. Witkiewitz, *et al.*, "What happens when you smoke a cigarette mindfully? a deductive qualitative study.," *Mindfulness*, vol. 13, no. 10, pp. 2628–2640, Sep. 19, 2022. DOI: 10.1007/s12671-022-01984-7.

[7] C. Pinard and A. Manzanera, "Does it work outside this benchmark? introducing the rigid depth constructor tool," *Multimedia Tools and Applications*, vol. 82, no. 27, pp. 41 641–41 667, Apr. 4, 2023. DOI: 10.1007/s11042-023-14743-0.

[8] A. N. Muhammad, A. M. Aseere, H. Chiroma, H. Shah, A. Y. Gital, and I. A. T. Hashem, "Deep learning application in smart cities: Recent development, taxonomy, challenges and research prospects," *Neural Computing and Applications*, vol. 33, no. 7, pp. 2973–3009, Jul. 15, 2020. DOI: 10.1007/s00521-020-05151-8.

[9] H. Ratnayake, L. Chen, and X. Ding, "A review of federated learning: Taxonomy, privacy and future directions," *Journal of Intelligent Information Systems*, vol. 61, no. 3, pp. 923–949, Jul. 4, 2023. DOI: 10.1007/s10844-023-00797-x.

[10] A. A. J. Cizotto, R. C. T. de Souza, V. C. Mariani, and L. dos Santos Coelho, "Web pages from mockup design based on convolutional neural network and class activation mapping," *Multimedia Tools and Applications*, vol. 82, no. 25, pp. 38 771–38 797, Mar. 24, 2023. DOI: 10.1007/s11042-023-15108-3.

[11] P. Lyu, M. Wei, and Y. Wu, "High-precision and real-time visual tracking algorithm based on the siamese network for autonomous driving," *Signal, Image and Video Processing*, vol. 17, no. 4, pp. 1235–1243, Aug. 10, 2022. DOI: 10.1007/s11760-022-02331-y.

[12] M. A. Kraus and M. Drass, "Artificial intelligence for structural glass engineering applications — overview, case studies and future potentials," *Glass Structures & Engineering*, vol. 5, no. 3, pp. 247–285, Oct. 7, 2020. DOI: 10.1007/s40940-020-00132-8.

[13] A. Golutvin, A. Iniukhin, A. Mauri, P. Owen, N. Serra, and A. Ustyuzhanin, "The dl advocate: Playing the devil's advocate with hidden systematic uncertainties," *The European Physical Journal C*, vol. 83, no. 9, Sep. 5, 2023. DOI: 10.1140/epjc/s10052-023-11925-w.

[14] A. Ravikumar, H. Sriraman, B. Saleena, and B. Prakash, "Selecting the optimal transfer learning model for precise breast cancer diagnosis utilizing pre-trained deep learning models and histopathology images," *Health and Technology*, vol. 13, no. 5, pp. 721–745, Aug. 19, 2023. DOI: 10.1007/s12553-023-00772-0.

[15] Y. Koo, S. Kim, and Y.-G. Ha, "Opencl-darknet: Implementation and optimization of opencl-based deep learning object detection framework," *World Wide Web*, vol. 24, no. 4, pp. 1299–1319, Feb. 7, 2020. DOI: 10.1007/s11280-020-00778-y.

[16] J. J. Lang, L. F. Yang, C. Cheng, X. Y. Cheng, and F. Y. Chen, "Are algorithmically controlled gig workers deeply burned out? an empirical study on employee work engagement.," *BMC psychology*, vol. 11, no. 1, pp. 354–, Oct. 24, 2023. DOI: 10.1186/s40359-023-01402-0.

[17] C.-H. Lin, Y.-C. Lin, Y. J. Wu, W.-H. Chung, and T.-S. Lee, "A survey on deep learning-based vehicular communication applications," *Journal of Signal Processing Systems*, vol. 93, no. 4, pp. 369–388, Aug. 23, 2020. DOI: 10.1007/s11265-020-01587-2.

[18] J. Kaur and W. Singh, "Tools, techniques, datasets and application areas for object detection in an image: A review.," *Multimedia tools and applications*, vol. 81, no. 27, pp. 38 297–38 351, Apr. 23, 2022. DOI: 10.1007/s11042-022-13153-y.

[19] F. A. Batarseh and L. J. Freeman, "A survey on ai assurance," *Journal of Big Data*, vol. 8, no. 1, pp. 1–30, Apr. 26, 2021. DOI: 10.1186/s40537-021-00445-7.

[20] J. Schmidt, M. R. G. Marques, S. Botti, and M. A. L. Marques, "Recent advances and applications of machine learning in solid-state materials science," *npj Computational Materials*, vol. 5, no. 1, pp. 1–36, Aug. 8, 2019. DOI: 10.1038/s41524-019-0221-0.

[21] T. Herrmann and S. Pfeiffer, "Keeping the organization in the loop: A socio-technical extension of human-centered artificial intelligence," *AI & SOCIETY*, vol. 38, no. 4, pp. 1523–1542, Feb. 18, 2022. DOI: 10.1007/s00146-022-01391-5.

[22] S. Azizi, L. Culp, J. Freyberg, *et al.*, "Robust and data-efficient generalization of self-supervised machine learning for diagnostic imaging.," *Nature biomedical engineering*, vol. 7, no. 6, pp. 756–779, Jun. 8, 2023. DOI: 10.1038/s41551-023-01049-7.

[23] S. A. Sayed, Y. Abdel-Hamid, and H. A. Hefny, "Artificial intelligence-based traffic flow prediction: A comprehensive review," *Journal of Electrical Systems and Information Technology*, vol. 10, no. 1, Mar. 9, 2023. DOI: 10.1186/s43067-023-00081-6.

[24] S. Xia, "What makes chinese rural migrants self-employed: A qualitative perspective," *The Journal of Chinese Sociology*, vol. 9, no. 1, Nov. 2, 2022. DOI: 10.1186/s40711-022-00175-0.

[25] M. Muniswamaiah, T. Agerwala, and C. C. Tappert, "Automatic visual recommendation for data science and analytics," in *Advances in Information and Communication: Proceedings of the 2020 Future of Information and Communication Conference (FICC), Volume 2*, Springer, 2020, pp. 125–132.

[26] E. G. Carayannis and J. Morawska-Jancelewicz, "The futures of europe: Society 5.0 and industry 5.0 as driving forces of future universities," *Journal of the Knowledge Economy*, vol. 13, no. 4, pp. 3445–3471, Jan. 5, 2022. DOI: 10.1007/s13132-021-00854-2.

[27] S. Dhakal, D. Qu, D. Carrillo, Q. Yang, and S. Fu, "Oasd: An open approach to self-driving vehicle," in *2021 Fourth International Conference on Connected and Autonomous Driving (MetroCAD)*, IEEE, 2021, pp. 54–61.

[28] K. Nader, P. Toprac, S. Scott, and S. Baker, "Public understanding of artificial intelligence through entertainment media.," *AI & society*, vol. 39, no. 2, pp. 1–726, Apr. 2, 2022. DOI: 10.1007/s00146-022-01427-w.

[29] W. P. N. dos Reis, G. E. Couto, and O. M. Junior, "Automated guided vehicles position control: A systematic literature review," *Journal of Intelligent Manufacturing*, vol. 34, no. 4, pp. 1483–1545, Jan. 5, 2022. DOI: 10.1007/s10845-021-01893-x.

[30] J. S. Huang, J. X. Liew, A. Ademiloye, and K. Liew, "Artificial intelligence in materials modeling and design," *Archives of Computational Methods in Engineering*, vol. 28, no. 5, pp. 3399–3413, Oct. 11, 2020. DOI: 10.1007/s11831-020-09506-1.

[31] P. Sinha, D. Roshini, V. Daoo, B. M. Abraham, and J. K. Singh, "Integrating machine learning and molecular simulation for material design and discovery," *Transactions of the Indian National Academy of Engineering*, vol. 8, no. 3, pp. 325–340, Jun. 29, 2023. DOI: 10.1007/s41403-023-00412-z.

[32] E. Veitch, H. Dybvik, M. Steinert, and O. A. Alsos, "Collaborative work with highly automated marine navigation systems.," *Computer supported cooperative work : CSCW : an international journal*, vol. 33,

no. 1, pp. 1–38, Oct. 8, 2022. DOI: 10.1007/s10606-022-09450-7.

[33] A. Safari, H. Kharrati, and A. Rahimi, "Multi-term electrical load forecasting of smart cities using a new hybrid highly accurate neural network-based predictive model," *Smart Grids and Sustainable Energy*, vol. 9, no. 1, Dec. 28, 2023. DOI: 10.1007/s40866-023-00188-9.

[34] A. Saxena, "Advancing location privacy in urban networks: A hybrid approach leveraging federated learning and geospatial semantics," *Int. J. Inf. Cybersecur*, vol. 7, pp. 58–72, 2023.

[35] Q. Wang and Y. Ding, "A novel fine-grained method for vehicle type recognition based on the locally enhanced pcanet neural network," *Journal of Computer Science and Technology*, vol. 33, no. 2, pp. 335–350, Mar. 23, 2018. DOI: 10.1007/s11390-018-1822-7.

[36] M. Pasquinelli and V. Joler, "The nooscope manifested: Ai as instrument of knowledge extractivism.," *AI & society*, vol. 36, no. 4, pp. 1–18, Nov. 21, 2020. DOI: 10.1007/s00146-020-01097-6.

[37] Y. Ahmadov and P. Helo, "Deep learning-based approach for forecasting intermittent online sales," *Discover Artificial Intelligence*, vol. 3, no. 1, Dec. 19, 2023. DOI: 10.1007/s44163-023-00085-1.

[38] J.-M. Guo and H. Markoni, "Driver drowsiness detection using hybrid convolutional neural network and long short-term memory," *Multimedia Tools and Applications*, vol. 78, no. 20, pp. 29 059–29 087, Jul. 20, 2018. DOI: 10.1007/s11042-018-6378-6.

[39] P. Wei, R. Liang, A. Michelmore, and Z. Kong, "Vision-based 2d navigation of unmanned aerial vehicles in riverine environments with imitation learning," *Journal of Intelligent & Robotic Systems*, vol. 104, no. 3, Mar. 4, 2022. DOI: 10.1007/s10846-022-01593-5.

[40] L. Cao, "Trans-ai/ds: Transformative, transdisciplinary and translational artificial intelligence and data science," *International Journal of Data Science and Analytics*, vol. 15, no. 2, pp. 119–132, Feb. 22, 2023. DOI: 10.1007/s41060-023-00383-y.

[41] D. S. Asudani, N. K. Nagwani, and P. Singh, "A comparative evaluation of machine learning and deep learning algorithms for question categorization of vqa datasets," *Multimedia Tools and Applications*, vol. 83, no. 19, pp. 57 829–57 859, Dec. 13, 2023. DOI: 10.1007/s11042-023-17797-2.

[42] J.-W. Bian, H. Zhan, N. Wang, *et al.*, "Unsupervised scale-consistent depth learning from video," *International Journal of Computer Vision*, vol. 129, no. 9, pp. 2548–2564, Jun. 18, 2021. DOI: 10.1007/s11263-021-01484-6.

[43] R. L. Abduljabbar, H. Dia, and P.-W. Tsai, "Development and evaluation of bidirectional lstm freeway traffic forecasting models using simulation data.," *Scientific reports*, vol. 11, no. 1, pp. 23 899–, Dec. 13, 2021. DOI: 10.1038/s41598-021-03282-z.

[44] S. Chai and L. Guo, "Edge computing with fog-cloud for heart data processing using particle swarm optimized deep learning technique," *Journal of Grid Computing*, vol. 22, no. 1, Dec. 23, 2023. DOI: 10.1007/s10723-023-09706-6.

[45] X. Zhang, Y. Li, H. Pei, and M. Ding, "Research on chaos of product color image system driven by brand image," *Multimedia Tools and Applications*, vol. 82, no. 16, pp. 24 425–24 444, Feb. 25, 2023. DOI: 10.1007/s11042-023-14549-0.

[46] Y. Wang, Z. Kunxian, D. Tian, X. Duan, and J. Zhou, "Pre-training with asynchronous supervised learning for reinforcement learning based autonomous driving," *Frontiers of Information Technology & Electronic Engineering*, vol. 22, no. 5, pp. 673–686, May 28, 2021. DOI: 10.1631/fitee.1900637.

[47] P. Chiril, E. W. Pamungkas, F. Benamara, V. Moriceau, and V. Patti, "Emotionally informed hate speech detection: A multi-target perspective," *Cognitive computation*, vol. 14, no. 1, pp. 1–31, Jun. 28, 2021. DOI: 10.1007/s12559-021-09862-5.

[48] J. G. Song and J. W. Lee, "Cnn-based object detection and distance prediction for autonomous driving using stereo images," *International Journal of Automotive Technology*, vol. 24, no. 3, pp. 773–786, May 12, 2023. DOI: 10.1007/s12239-023-0064-z.

[49] T. Mortier, M. Wydmuch, K. Dembczyński, E. Hüllermeier, and W. Waegeman, "Efficient set-valued prediction in multi-class classification," *Data Mining and Knowledge Discovery*, vol. 35, no. 4, pp. 1435–1469, May 6, 2021. DOI: 10.1007/s10618-021-00751-x.

[50] J. M. Spector and S. Ma, "Inquiry and critical thinking skills for the next generation: From artificial intelligence back to human intelligence," *Smart Learning Environments*, vol. 6, no. 1, pp. 1–11, Sep. 11, 2019. DOI: 10.1186/s40561-019-0088-z.

[51] A. Saxena and P. Mishra, "Optimizing electric vehicle energy management systems with a hybrid lstm-cnn architecture," *Tensorgate Journal of Sustainable*, 2022.

[52] V. Kocur and M. Ftáčnik, "Detection of 3d bounding boxes of vehicles using perspective transformation for accurate speed measurement," *Machine Vision and Applications*, vol. 31, no. 7, pp. 1–15, Sep. 4, 2020. DOI: 10.1007/s00138-020-01117-x.

[53] K. S. Carter, A. Gondek, W. Underwood, T. Randby, and R. Marciano, "Using ai and ml to optimize information discovery in under-utilized, holocaust-related records," *AI & SOCIETY*, vol. 37, no. 3, pp. 837–858, May 23, 2022. DOI: 10.1007/s00146-021-01368-w.

[54] A. Hashemi, G. Orzechowski, A. Mikkola, and J. McPhee, "Multibody dynamics and control using machine learning," *Multibody System Dynamics*, vol. 58, no. 3-4, pp. 397–431, Feb. 27, 2023. DOI: 10.1007/s11044-023-09884-x.

[55] I. Tiddi, E. Bastianelli, E. Daga, M. d'Aquin, and E. Motta, "Robot–city interaction: Mapping the research landscape—a survey of the interactions between robots and modern cities," *International Journal of Social Robotics*, vol. 12, no. 2, pp. 299–324, Mar. 14, 2019. DOI: 10.1007/s12369-019-00534-x.

[56] K. A. Sonwane, "Simulating self driving car using deep learning," *International Journal for Research in Applied Science and Engineering Technology*, vol. 9, no. VI, pp. 738–743, Jul. 15, 2021. DOI: 10.22214/ijraset.2021.36454.

[57] A. Luczak, B. L. McNaughton, and Y. Kubo, "Neurons learn by predicting future activity.," *Nature machine intelligence*, vol. 4, no. 1, pp. 62–72, Jan. 25, 2022. DOI: 10.1038/s42256-021-00430-y.

[58] R. M. Martins, C. G. von Wangenheim, M. F. Rauber, and J. C. Hauck, "Machine learning for all!—introducing machine learning in middle and high school," *International Journal of Artificial Intelligence in Education*, vol. 34, no. 2, pp. 185–223, Jan. 25, 2023. DOI: 10.1007/s40593-022-00325-y.

[59] J.-L. Shieh, M. A. Haq, Q. M. ul Haq, S.-J. Ruan, and P. Chondro, "Utilizing incremental branches on a one-stage object detection framework to avoid catastrophic forgetting," *Machine Vision and Applications*, vol. 33, no. 2, Feb. 5, 2022. DOI: 10.1007/s00138-022-01284-z.

[60] C. Liu, H. Liu, L. Han, and C. Xiang, "New integrated multi-algorithm fusion localization and trajectory tracking framework of autonomous vehicles under extreme conditions with non-gaussian noises," *International Journal of Automotive Technology*, vol. 24, no. 1, pp. 259–272, Feb. 15, 2023. DOI: 10.1007/s12239-023-0023-8.

[61] L. Meng, B. McWilliams, W. Jarosinski, *et al.*, "Machine learning in additive manufacturing: A review," *JOM*, vol. 72, no. 6, pp. 2363–2377, Apr. 17, 2020. DOI: 10.1007/s11837-020-04155-y.

[62] M. Muniswamaiah, T. Agerwala, and C. Tappert, "Big data in cloud computing review and opportunities," *arXiv preprint arXiv:1912.10821*, 2019.

[63] J. Gao, B. Han, and K. Yan, "Static object imaging features recognition algorithm in dynamic scene mapping," *Multimedia Tools and Applications*, vol. 78, no. 23, pp. 33 885–33 898, Sep. 12, 2019. DOI: 10.1007/s11042-019-08148-1.

[64] A. Agrawal and A. Choudhary, "Deep materials informatics: Applications of deep learning in materials science," *MRS Communications*, vol. 9, no. 3, pp. 779–792, Sep. 20, 2019. DOI: 10.1557/mrc.2019.73.

[65] B. Gezici and A. K. Tarhan, "Systematic literature review on software quality for ai-based software," *Empirical Software Engineering*, vol. 27, no. 3, Mar. 17, 2022. DOI: 10.1007/s10664-021-10105-2.

[66] M. Latzer, "The digital trinity—controllable human evolution—implicit everyday religion," *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, vol. 74, no. S1, pp. 331–354, May 5, 2022. DOI: 10.1007/s11577-022-00841-8.

[67] A. Darbari, K. Kumar, S. Darbari, and P. L. Patil, "Requirement of artificial intelligence technology awareness for thoracic surgeons," *The cardiothoracic surgeon*, vol. 29, no. 1, pp. 1–10, Jul. 3, 2021. DOI: 10.1186/s43057-021-00053-4.

[68] D. Harel, R. Yerushalmi, A. Marron, and A. Elyasaf, "Categorizing methods for integrating machine learning with executable specifications," *Science China Information Sciences*, vol. 67, no. 1, Oct. 17, 2023. DOI: 10.1007/s11432-022-3826-6.

[69] "The company profile: Delphi technologies," *The Hansen Report on Automotive Electronics*, vol. 32, no. 9, pp. 12–20, Nov. 1, 2019. DOI: 10.1007/bf03546059.

[70] S. Dhakal, Q. Chen, D. Qu, D. Carillo, Q. Yang, and S. Fu, "Sniffer faster r-cnn: A joint camera-lidar object detection framework with proposal refinement," in *2023 IEEE International Conference on Mobility, Operations, Services and Technologies (MOST)*, IEEE, 2023, pp. 1–10.

[71] V. Hayyolalam, S. Otoum, and Ö. Özkasap, "Dynamic qos/qoe-aware reliable service composition framework for edge intelligence.," *Cluster computing*, vol. 25, no. 3, pp. 1695–1713, Mar. 26, 2022. DOI: 10.1007/s10586-022-03572-9.

[72] J. Zamboni, D. M. Viana, L. Rodrigues, and M. Consalvo, "Expressive elements of lifelike machines," *International Journal of Social Robotics*, vol. 15, no. 6, pp. 879–891, Apr. 7, 2023. DOI: 10.1007/s12369-023-00994-2.

[73] A. S. Rao and K. Mahantesh, "Learning semantic features for classifying very large image datasets using convolution neural network," *SN Computer Science*, vol. 2, no. 3, pp. 1–9, Mar. 31, 2021. DOI: 10.1007/s42979-021-00589-6.

[74] M. Muniswamaiah, T. Agerwala, and C. C. Tappert, "Integrating polystore rdbms with common in-memory data," in *2020 IEEE International Conference on Big Data (Big Data)*, IEEE, 2020, pp. 5762–5764.

[75] S. Shah and J. Tembhurne, "Object detection using convolutional neural networks and transformer-based models: A review," *Journal of Electrical Systems and Information Technology*, vol. 10, no. 1, Nov. 20, 2023. DOI: 10.1186/s43067-023-00123-z.

[76] A. Leitherer, B. C. Yeo, C. H. Liebscher, and L. M. Ghiringhelli, "Automatic identification of crystal structures and interfaces via artificial-intelligence-based electron microscopy," *npj Computational Materials*, vol. 9, no. 1, Oct. 2, 2023. DOI: 10.1038/s41524-023-01133-1.

[77] O. Kashefi, T. Afrin, M. Dale, *et al.*, "Argrewrite v.2: An annotated argumentative revisions corpus," *Language Resources and Evaluation*, vol. 56, no. 3, pp. 881–915, Jan. 13, 2022. DOI: 10.1007/s10579-021-09567-z.

[78] S. A. F. Harahap, M. F. Rafliansyah, A. Septyanto, and I. Pratama, "Studi perbandingan algoritma statistic, ensemble, dan neural network pada kasus image classification," *Jurnal Rekayasa Elektro Sriwijaya*, vol. 4, no. 2, pp. 46–51, May 15, 2023. DOI: 10.36706/jres.v4i2.64.

[79] R. Lu, B. Chen, G. Liu, Z. Cheng, M. Qiao, and X. Yuan, "Dual-view snapshot compressive imaging via optical flow aided recurrent neural network," *International Journal of Computer Vision*, vol. 129, no. 12, pp. 3279–3298, Oct. 15, 2021. DOI: 10.1007/s11263-021-01532-1.

[80] S. Spanogiannopoulos, Y. Zweiri, and L. Seneviratne, "Sampling-based non-holonomic path generation for self-driving cars," *Journal of Intelligent & Robotic Systems*, vol. 104, no. 1, Dec. 28, 2021. DOI: 10.1007/s10846-021-01440-z.

[81] M. Simmler and N. Markwalder, "Guilty robots? - rethinking the nature of culpability and legal personhood in an age of artificial intelligence," *Criminal Law Forum*, vol. 30, no. 1, pp. 1–31, Dec. 4, 2018. DOI: 10.1007/s10609-018-9360-0.

[82] A. Hsu, Y. H. Lo, S. C. Ke, L. Lin, and P. Tseng, "Variation of picture angles and its effect on the concealed information test," *Cognitive research: principles and implications*, vol. 5, no. 1, pp. 33–, Jul. 31, 2020. DOI: 10.1186/s41235-020-00233-6.

[83] Z. Nadeem, Z. Khan, U. Mir, *et al.*, "Pakistani traffic-sign recognition using transfer learning," *Multimedia Tools and Applications*, vol. 81, no. 6, pp. 8429–8449, Feb. 2, 2022. DOI: 10.1007/s11042-022-12177-8.

[84] M. J. Iqbal, Z. Javed, H. Sadia, *et al.*, "Clinical applications of artificial intelligence and machine learning in cancer diagnosis: Looking into the future.," *Cancer cell international*, vol. 21, no. 1, pp. 270–270, May 21, 2021. DOI: 10.1186/s12935-021-01981-1.

[85] I. Cardoso-Pereira, J. B. Borges, P. H. Barros, A. F. Loureiro, O. A. Rosso, and H. S. Ramos, "Leveraging the self-transition probability of ordinal patterns transition network for transportation mode identification based on gps data," *Nonlinear Dynamics*, vol. 107, no. 1, pp. 889–908, Nov. 30, 2021. DOI: 10.1007/s11071-021-07059-x.

[86] F. H. Shavaki and A. E. Ghahnavieh, "Applications of deep learning into supply chain management: A systematic literature review and a framework for future research.," *Artificial intelligence review*, vol. 56, no. 5, pp. 4447–4489, Sep. 30, 2022. DOI: 10.1007/s10462-022-10289-z.

[87] S. Madakam, T. Uchiya, S. Mark, and Y. Lurie, "Artificial intelligence, machine learning and deep learning (literature: Review and metrics)," *Asia-Pacific Journal of Management Research and Innovation*, vol. 18, no. 1-2, pp. 7–23, Dec. 19, 2022. DOI: 10.1177/2319510x221136682.

[88] M. Abouelyazid, "Comparative evaluation of vgg-16 and u-net architectures for road segmentation," *Eigenpub Review of Science and Technology*, vol. 6, no. 1, pp. 75–91, 2022.

[89] O. P. Swain, H. Hemanth, P. Saran, *et al.*, "Robust and efficient keyword spotting using a bidirectional attention lstm," *International Journal of Speech Technology*, vol. 26, no. 4, pp. 919–931, Nov. 11, 2023. DOI: 10.1007/s10772-023-10067-4.