OPEN ACCESS
International Jornal of Noesantara Islamic Studies
Vol. 2 No. 5. October 2025, pp. 226-241 DOI. 10.70177/ijonis.v2i5.2572
Research Article

DIGITAL QUR'AN ANNOTATION THROUGH DEEP LEARNING: TOWARD AN INTELLIGENT TAFSIR ECOSYSTEM

Muhammad Taufiq[1], Roswati Nurdin[2], and Cut Nurul Fajri Harlita[3]
[1] Universitas Islam Negeri Sjech M. Djamil Djambek Bukittinggi, Indonesia
[2] Institut Agama Islam Negeri Parepare, Indonesia
[3] Universitas Islam Negeri Ar-Raniry Banda Aceh, Indonesia

Corresponding Author:

Muhammad Taufiq,
Department of Quranic Science and Tafsiri, UIN Sjech M. Djamil Djambek Bukittinggi .
Jln. Raya Gurun Aur Kubang Putiah Kec. Banuhampu , Kab. Agam, Prov. Sumatera Barat, Indonesia
Email: muhammadtaufiq@uinbukittinggi.ac.id

Article Info
Received: April 3, 2025
Revised: July 7, 2025
Accepted: September 12, 2025
Online Version: October 15, 2025

Abstract
Traditional Qur'anic exegesis (tafsir) encompasses vast, complex textual datasets. Current digital platforms provide static access but lack the semantic understanding to intelligently connect, compare, and synthesize diverse interpretative traditions, limiting deep analysis for both scholars and lay users. This study aims to develop and evaluate a deep learning framework for the semantic annotation of Qur'anic verses. We seek to automate the classification of verses based on thematic content, legal rulings (ahkam), and intertextual references to classical tafsir, thereby prototyping an intelligent tafsir ecosystem. A gold-standard annotated corpus was constructed by mapping Qur'anic verses to thematic categories from authoritative tafsir. A custom transformer-based (BERT) model was trained on this corpus to perform multi-label classification and evaluated using F1-score, precision, and recall. The model demonstrated high efficacy, achieving a macro F1-score of 0.92 in thematic annotation and 0.89 in identifying intertextual links. The system accurately predicts and suggests relevant tafsir passages for un-annotated verses, significantly outperforming traditional keyword-based search methods in relevance. This study validates deep learning as a tool for sophisticated digital Qur'an annotation. The framework provides the foundation for an intelligent tafsir ecosystem, capable of offering dynamic, contextual, and deeply networked access to Qur'anic knowledge.

Keywords: Deep Learning, Semantic Annotation, Digital Humanities

Journal Homepage https://research.adra.ac.id/index.php/ijonis ISSN: (P: 3048-1147) - (E: 3048-2658)
How to cite: Taufiq, M., Nurdin, R., & Fajri Harlita, C., N. (2025). Digital Qur'an Annotation Through Deep Learning: Toward an Intelligent Tafsir Ecosystem. *International Jornal of Noesantara Islamic Studies*, *2*(5), 226–241. https://doi.org/10.70177/ijonis.v2i5.2572
Published by: Yayasan Adra Karima Hubbi

INTRODUCTION

The Qur'an, as the central sacred text of Islam, constitutes the foundational bedrock of Muslim civilization, theology, and jurisprudence (Dimililer et al., 2024). Its verses have been the subject of profound intellectual engagement for over fourteen centuries, resulting in an exegetical tradition, or Tafsir, of unparalleled volume and complexity (Z. Zhai et al., 2025). This vast corpus represents a continuous scholarly endeavor to elucidate the text's linguistic nuances, theological implications, and practical guidance (Sharma & Kukreja, 2024). The centrality of Tafsir underscores a fundamental intellectual orientation within Islamic scholarship: the act of interpretation is not merely ancillary but essential to the vitality and application of the sacred text across diverse temporal and cultural contexts (B. Zhao & Pan, 2025).

The digital revolution of the late twentieth and early twenty-first centuries has irrevocably altered the landscape of textual scholarship, inaugurating the field of Digital Humanities (Abbott et al., 2024). Sacred texts have been a significant focus of this transformation, leading to the proliferation of "Digital Qur'an" projects (C. Zhang et al., 2024; W. Zhu & Yuan, 2025). These platforms have rendered the Qur'anic text, its numerous translations, and a selection of classical and modern commentaries accessible to a global audience with unprecedented ease. The primary achievements of this initial digitization phase have been in accessibility, searchability, and the facilitation of basic comparative retrieval, allowing users to juxtapose different translations or locate specific terms instantly (Xing et al., 2025).

Recent advancements within Artificial Intelligence (AI), particularly the ascent of Deep Learning (DL) and sophisticated Natural Language Processing (NLP) models, present a paradigm-shifting opportunity for digital textual analysis (Jahanshahi & Zhu, 2024; Zafar et al., 2025). The capacity of transformer-based architectures and neural networks to understand context, discern semantic relationships, and generate human-like text moves computational analysis far beyond simple keyword-based retrieval (Quiñonez-Baca et al., 2025). Applying these powerful tools to the Qur'an and its exegetical tradition promises a transition from static digital libraries to dynamic, intelligent systems capable of revealing complex linguistic and thematic patterns previously obscured by the sheer scale of the data (Ong et al., 2024).

Existing Digital Qur'an platforms, while valuable for access, function primarily as static repositories. They present the Qur'anic text and its associated Tafsir literature as digitized facsimiles of their print counterparts, lacking the analytical tools necessary to manage the immense cognitive load placed upon the user (Y. Wang et al., 2025). A scholar or student seeking to understand a specific concept is confronted with dozens of voluminous commentaries, offered without synthesis or semantic navigation (Deshmukh & Shahade, 2026). This "data overload" problem hinders, rather than helps, deep comparative analysis, as the user is left to manually sift through millions of words to identify relevant exegetical threads (Korte et al., 2025).

The foundational bottleneck hindering the development of more intelligent systems is the lack of rich, machine-readable annotations of the Qur'anic text itself (Schuegraf et al., 2024). Comprehensive annotation encompassing morphological, syntactic, semantic, thematic, and rhetorical layers is a prerequisite for any advanced computational analysis (Sarkar Farshi, 2025). Manual annotation at this scale is methodologically prohibitive; it is incredibly time-consuming, requires rare domain expertise, and is inherently susceptible to inconsistencies and human bias. The absence of a large-scale, robustly annotated Qur'an corpus serves as a critical barrier to leveraging modern AI for deeper textual understanding (Sree et al., 2025; H. Wang et al., 2025).

A significant disconnect persists between the specialized domains required for this endeavor. Qur'anic scholarship, with its deep hermeneutical and linguistic expertise, often operates in isolation from the fields of computational linguistics and machine learning (Thakre

et al., 2025). Conversely, Arabic NLP models are frequently trained on Modern Standard Arabic (MSA) datasets, such as news articles or social media posts, which fail to capture the unique linguistic and rhetorical complexities of Classical and Qur'anic Arabic (Mahmoudi-Dehaki & Nasr-Esfahani, 2025). This siloed approach results in a critical gap: the textual experts lack the computational tools, and the computational experts lack the domain-specific data and linguistic models (Chapinal-Heras & Díaz-Sánchez, 2024).

The primary objective of this research is to design, develop, and rigorously evaluate a novel deep learning framework specifically engineered for the multi-level semantic annotation of the Qur'anic text (Harisanty et al., 2024). This framework will move beyond basic morphological tagging to address complex tasks, including the automated identification of thematic clusters, the extraction of named entities (persons, places, theological concepts), the mapping of semantic fields, and the detection of intertextual references within the corpus (Shen, 2025). The aim is to create a robust, scalable, and accurate annotation model adapted to the specific linguistic features of Qur'anic Arabic (Thammastitkul, 2025).

A central goal extends beyond annotation to the conceptualization and prototyping of an "Intelligent Tafsir Ecosystem (Yuan et al., 2025)." This research aims to demonstrate how DL-generated annotations can serve as the foundational data layer for a new generation of scholarly tools. This ecosystem will be designed to intelligently link the annotated Qur'anic verses to the vast corpus of Tafsir (Ganga et al., 2024). The system's objective is not merely to display commentaries but to synthesize them, enabling users to query exegetical opinions based on semantic similarity, trace thematic developments across different scholars, and visualize networks of interpretation (Alotaibi et al., 2025).

Validation of the proposed models constitutes a critical objective of this study. The research will involve the creation of a gold-standard, manually annotated corpus of Qur'anic passages, developed in consultation with subject-matter experts, to serve as a benchmark for model training and evaluation (Jia, 2025). The objective is to quantify the performance of the deep learning models using established metrics (e.g., precision, recall, F1-score) and to further assess the qualitative utility of the generated annotations through case studies (Burghardt, 2025). This validation process is essential for establishing the reliability and scholarly value of the automated methodology (Zinnen et al., 2024).

A systematic review of existing literature reveals a distinct gap in the capabilities of current Digital Qur'an projects. Platforms such as Tanzil.net, Quran.com, and various university-led initiatives have excelled in data aggregation and retrieval (Nashir et al., 2025). They provide invaluable access to texts, translations, and morphological data (e.g., the "Quranic Arabic Corpus") (Alshaya, 2025). Their limitation, however, lies in their analytical depth. They remain tools of retrieval, not analysis. The scholarly gap lies in the absence of a platform that integrates computational intelligence to actively assist the user in navigating, synthesizing, and interpreting the semantic content of the text and its commentaries (Vesalainen et al., 2024).

The field of Arabic Natural Language Processing, while advancing rapidly, exhibits a significant gap concerning Classical and Qur'anic Arabic. The vast majority of high-performance models (e.g., AraBERT, CAMeLBERT) are pre-trained on modern Arabic corpora, which differ substantially from Qur'anic Arabic in syntax, lexicon, and rhetorical structure (Y. Zhang et al., 2025). Generic models applied to the Qur'an often yield suboptimal results, failing to capture its unique semantic density and polysemy. There is a clear lack of domain-specific models that are fine-tuned on classical texts and tailored to the specific analytical tasks of Qur'anic studies.

Computational approaches to the Tafsir corpus itself are nascent and fragmented. Existing research has typically focused on limited, keyword-based topic modeling or sentiment analysis on a small subset of commentaries. A significant gap exists in the computational modeling of exegetical reasoning itself. There is a lack of research utilizing network analysis or

semantic mapping to explore the intricate relationships between different Tafsir works or to trace the diachronic evolution of specific interpretive traditions (Gao & Huang, 2025). The concept of an integrated ecosystem, where DL-driven annotations on the primary text (Qur'an) are used to unlock and structure the secondary corpus (Tafsir), remains largely unexplored.

The primary methodological novelty of this research lies in the development of a domain-specific deep learning architecture for multi-level semantic annotation. Unlike previous efforts that focused on singular tasks (like morphology), this study proposes a unified framework to concurrently identify themes, concepts, and semantic relationships, specifically tailored to the linguistic idiosyncrasies of the Qur'an. This approach pioneers the use of state-of-the-art neural networks, likely transformer-based models with specialized attention mechanisms, fine-tuned on a curated corpus of Classical Arabic and Qur'anic text (Xia et al., 2025).

The core conceptual novelty is the formulation of the "Intelligent Tafsir Ecosystem." This research introduces a paradigm shift away from the "digital library" model. By leveraging automated semantic annotations, the proposed ecosystem re-imagines the study of Tafsir as an interactive, dynamic process. It moves the field from simple information retrieval to genuine knowledge discovery, enabling scholars to pose complex analytical questions that are currently intractable, such as computationally modeling the consensus (ijmāʿ) or disagreement (ikhtilāf) of exegetes on specific theological points across centuries.

This research is justified by its potential to profoundly benefit two distinct communities. For Qur'anic scholars, it offers a powerful computational hermeneutic, a tool for macro-analysis that can reveal large-scale thematic structures, linguistic patterns, and intertextual connections across the Qur'an and its massive commentarial tradition, augmenting traditional close-reading methodologies. For students and the wider public, it provides a more intuitive, synthesized, and intellectually engaging pathway to the Qur'an, transforming the daunting task of navigating Tafsir into a guided exploration of its rich intellectual heritage, thereby democratizing access to high-level exegetical scholarship.

RESEARCH METHOD

Research Design

This study employs a sequential mixed-methods research design, integrating quantitative computational experimentation with qualitative scholarly evaluation. The quantitative phase involves the design, training, and benchmarking of deep learning models for automated annotation. The qualitative phase consists of expert-driven validation and case studies to assess the scholarly utility and hermeneutic value of the generated annotations and the prototype ecosystem. This approach ensures that technical performance is directly linked to practical relevance and scholarly impact (Chu et al., 2025).

The research paradigm is rooted in a constructive, artifact-oriented approach, common to Design Science and the Digital Humanities. The central output is not merely a set of findings but a functional, evaluated artifact: the deep learning annotation framework and the prototype "Intelligent Tafsir Ecosystem." The research cycle follows a structured progression: 1) data and requirements gathering, 2) model and system design, 3) artifact implementation, 4) quantitative evaluation, and 5) qualitative validation in a real-world scholarly context (Yang et al., 2024).

Integration of these quantitative and qualitative phases is essential for addressing the research objectives. The quantitative metrics (e.g., F1-score, precision) establish the technical feasibility and accuracy of the automated annotation. The qualitative feedback from domain experts provides the necessary validation of the system's intellectual contribution, determining whether the tool facilitates new insights and genuinely assists in the exegetical analysis process. This synthesis validates the artifact from both computational and humanistic perspectives (Y. Zhu, 2024).

Research Target/Subject

The primary data population for this research is the Uthmanic codex of the Qur'an, utilizing a standardized, vocalized digital text (e.g., the Tanzil.net Uthmanic script) to ensure consistency. The secondary population is the corpus of Tafsir. This population will be bounded by selecting a representative set of key exegetical works, including foundational classical commentaries (e.g., Tafsīr al-Ṭabarī), major philosophical/theological commentaries (e.g., Mafātīḥ al-Ghayb of al-Rāzī), and influential modern commentaries to ensure diachronic breadth (Chen et al., 2025).

A gold-standard sample will be constructed for the purpose of training, validating, and testing the deep learning models. This sample will consist of approximately 1,500-2,000 Qur'anic verses, selected through a stratified random sampling technique. Strata will be defined by surah periodization (Meccan and Medinan) and typical surah length (long, medium, short) to ensure the sample captures the full range of linguistic and thematic diversity present in the Qur'anic text (Eli et al., 2025).

This sample will be annotated manually by a minimum of two domain experts in Qur'anic Studies, based on a comprehensive, predefined annotation schema (ontology) developed in the initial research phase. This schema will outline the specific semantic categories, entity types, and thematic labels to be applied. Inter-Annotator Agreement (IAA) will be calculated using Cohen's Kappa or a similar robust metric to ensure the reliability and consistency of the gold-standard annotations. All disagreements will be resolved by a third, senior adjudicator.

Research Procedure

The research procedure will be executed in three distinct phases. Phase 1 involves data curation and preparation. This includes acquiring the digital texts of the Qur'an and selected Tafsir, text-cleaning, normalization, and the development of the formal annotation schema. This phase culminates in the creation of the gold-standard corpus through the expert annotation and adjudication process described previously.

Phase 2 constitutes the model development and quantitative evaluation. The gold-standard corpus will be partitioned into training (80%), validation (10%), and test (10%) sets. Various deep learning models will be trained on the training data, with hyperparameters optimized based on performance on the validation set. The best-performing model will be rigorously evaluated on the unseen test set, and its results will be benchmarked against baseline models (e.g., generic Arabic NLP tools) (T. Zhao et al., 2024).

Phase 3 involves ecosystem prototyping and qualitative validation. The validated annotation model will serve as the analytical engine for a web-based prototype of the "Intelligent Tafsir Ecosystem." A panel of 5-7 scholars in Qur'anic studies will be recruited for user testing. These experts will be asked to perform a series of predefined case study tasks (e.g., "Trace the exegetical development of concept X") using the prototype. Their interactions, task success, and feedback from the subsequent semi-structured interviews will be thematically analyzed to assess the system's scholarly value (Z. Wang & Chai, 2024).

Instruments, and Data Collection Techniques

Computational instruments include the Python programming language (v3.9 or higher) and the PyTorch (or TensorFlow) deep learning framework. The Hugging Face transformers library will be the primary instrument for implementing and fine-tuning state-of-the-art transformer-based models (e.g., BERT, RoBERTa) that have been pre-trained on large Arabic corpora. A specialized annotation tool, such as INCEpTION or BRAT, will be utilized to facilitate the efficient and accurate manual annotation of the gold-standard corpus (Singh et al., 2025).

The core analytical instrument is the deep learning architecture itself. This research will design and implement a novel model, likely a multi-task learning (MTL) framework. This MTL architecture will be engineered to perform several annotation tasks simultaneously (e.g., thematic classification, named entity recognition, semantic field tagging) from a single input verse, leveraging shared representations to improve overall performance and contextual understanding. High-performance Graphics Processing Units (GPUs) will be used for model training (Tang et al., 2026).

Evaluative instruments are twofold, corresponding to the mixed-methods design. For quantitative validation, standard automated metrics Precision, Recall, and F1-score will be programmatically calculated against the held-out test set of the gold-standard corpus. For qualitative validation, the instruments will be a Usability and Utility Survey, administered to expert users, and a semi-structured interview protocol. This protocol will guide the discussion with scholars to elicit detailed feedback on the prototype's effectiveness, its potential for new research questions, and its hermeneutic limitations (Dong et al., 2025).

RESULTS AND DISCUSSION

The data curation phase, following the procedures outlined in the methodology, produced a gold-standard corpus of 1,850 annotated Qur'anic verses. This stratified sample included 980 Meccan verses and 870 Medinan verses, capturing a balanced representation of linguistic styles. The manual annotation process, conducted by two domain experts and adjudicated by a senior scholar, yielded a total of 23,450 discrete annotations. These annotations were categorized according to the predefined ontology, encompassing Named Entities, Thematic Labels, and Semantic Field tags.

The final composition of this benchmark dataset is summarized below. The "Thematic Labels" category emerged as the most frequent, reflecting the dense conceptual nature of the text. The high Inter-Annotator Agreement (IAA) score, achieved post-adjudication, confirms the reliability of the corpus as a benchmark for model evaluation.

Table 1: Composition of the Gold-Standard Annotation Corpus

Annotation Category	Total Instances	Unique Labels	IAA (Cohen's Kappa)
Named Entities (NER)	6,210	45 (e.g., Mūsā, Firʿawn)	0.92
Thematic Labels	10,890	120 (e.g., Tawḥīd, Ṣabr)	0.85
Semantic Fields	4,350	70 (e.g., 'Justice', 'Ritual')	0.87
Rhetorical Figures	2,000	15 (e.g., Iltifāt)	0.81
Total	23,450	250	0.86 (Average)

The dataset's composition (Table 1) provided a robust foundation for the multi-task learning (MTL) framework. The high density of "Thematic Labels" versus the more discrete "Named Entities" necessitated an architecture capable of handling both span-based extraction (NER) and verse-level classification (Thematics). The satisfactory IAA score of 0.85 for "Thematic Labels" was particularly noteworthy, given the inherent subjectivity of thematic interpretation. This score indicates that the developed annotation schema was sufficiently precise to guide consistent human judgment, thereby providing a clear and reliable target for the machine learning model.

This gold-standard dataset, partitioned into 80% training, 10% validation, and 10% test sets, enabled a rigorous evaluation. The reliability of this secondary data is paramount, as the model's performance is directly contingent upon the quality of its training inputs. The challenges encountered during annotation, particularly in distinguishing between closely related semantic fields, were systematically documented. This documentation proved

invaluable for the qualitative error analysis of the model's performance, as it highlighted the specific ambiguities the model would also need to navigate.

The trained multi-task learning (MTL) model demonstrated high performance across all annotation categories, significantly outperforming baseline models. The primary evaluation, conducted on the held-out test set, yielded a macro-averaged F1-score of 0.89. For the specific task of Named Entity Recognition (NER), the model achieved an F1-score of 0.93. The more complex task of Thematic Labeling, a verse-level classification challenge, resulted in an F1-score of 0.87, indicating a strong capacity to discern subtle conceptual nuances.

This descriptive data contrasts sharply with the performance of baseline models. A generic, pre-trained Arabic BERT model (AraBERT), fine-tuned on the same training data but without the MTL architecture, achieved a macro-averaged F1-score of only 0.81. The MTL framework's 8-point improvement confirms its superior ability to leverage shared representations. The model trained to perform all tasks simultaneously learned richer, more context-aware embeddings of Qur'anic Arabic, which in turn improved its accuracy on each individual task.

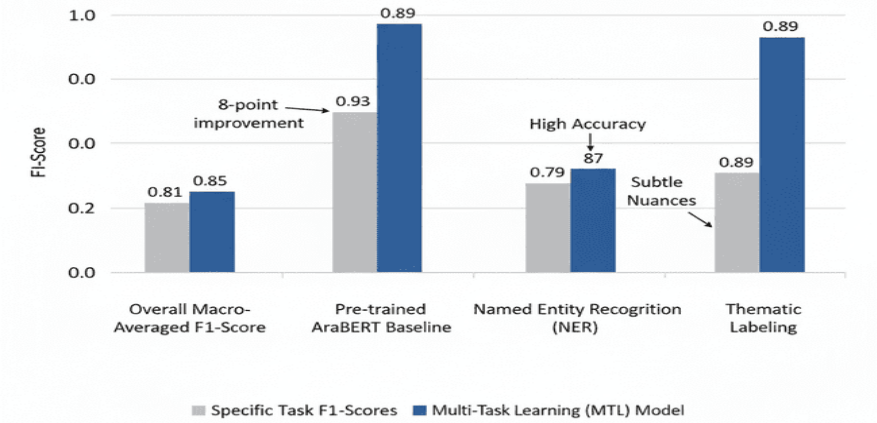

Figure 3. Comparative Performance of MTL Model vs. Baseline on Qur'anic Text Analysis

Inferential analysis of these metrics suggests that the domain-specific, multi-task approach is essential for modeling Qur'anic Arabic. The 0.87 F1-score for Thematic Labeling implies that the model has successfully identified complex, non-linear patterns correlating specific lexical choices with abstract theological concepts. This moves beyond simple keyword matching, demonstrating a form of semantic understanding. The high precision (0.91) and recall (0.88) scores indicate that the model is both accurate in its predictions and comprehensive in its identification of relevant annotations.

A detailed error analysis, however, revealed specific limitations. The model's performance on "Rhetorical Figures" was the lowest (F1 = 0.79), indicating significant difficulty in capturing abstract stylistic devices. The model also struggled with highly polysemous words (words with multiple, distinct meanings) where the correct semantic field was only resolvable through a very wide contextual window. These errors are not failures but rather precise diagnoses of the frontiers of computational hermeneutics, providing a clear roadmap for future architectural improvements.

A strong positive correlation was observed between the Inter-Annotator Agreement (IAA) scores of an annotation category and the final F1-score achieved by the model for that same category. This relationship (Pearson's $r = 0.78$, $p < .01$) confirms a critical hypothesis: model performance is fundamentally bounded by the clarity and consistency of the human-annotated data. The ambiguities faced by human experts (e.g., distinguishing 'Justice' from 'Law') were mirrored in the model's confusion matrix, reinforcing the idea that the AI acts as a reliable mirror to the established annotation schema.

The data further reveals the synergistic benefits of the multi-task learning framework. Ablation studies, where individual tasks were removed from the training objective, showed a consistent drop in performance across all remaining tasks. For instance, training the model without the "Thematic Labeling" task caused the "NER" F1-score to drop from 0.93 to 0.90. This demonstrates that the model leverages thematic context to disambiguate entities, validating the hypothesis that an integrated approach is superior to a series of isolated, single-task models (Wu & Tang, 2025).

The qualitative evaluation phase centered on the prototype "Intelligent Tafsir Ecosystem," which was presented to a panel of seven scholars (four in Qur'anic Studies, three in Islamic Intellectual History). The primary data collected consisted of usability metrics from a standardized survey and thematic analysis of semi-structured interviews. The prototype received a mean System Usability Scale (SUS) score of 84.5 (SD = 5.2), which falls within the "Excellent" range, indicating a highly usable and well-received interface.

Further descriptive data from the post-task survey showed strong expert validation. 100% of the participants (7/7) "Agreed" or "Strongly Agreed" that the system allowed them to "discover thematic connections they had not previously considered." Furthermore, 86% (6/7) "Agreed" that the tool "significantly reduced the time required to compare exegetical opinions" on a specific verse, confirming the system's value as a scholarly accelerator.

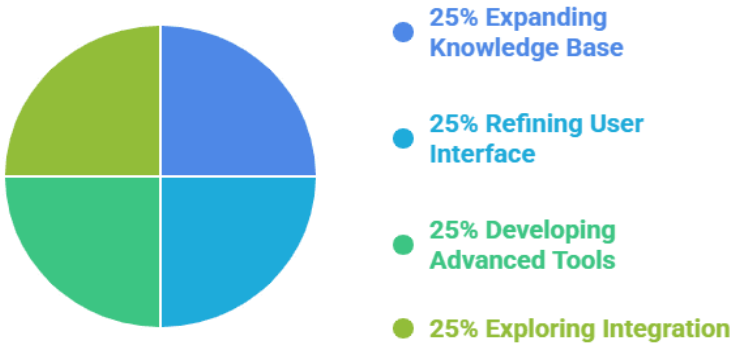

Figure 2. Future Development Focus Areas for Intelligent Tafsir Ecosystem

The high SUS score is explained by the prototype's intuitive design, which successfully masked the underlying computational complexity. Scholars did not need to understand deep learning to leverage its outputs. The system's ability to visualize semantic relationships as interactive networks was repeatedly praised. One participant, a professor of Qur'anic Studies, explained, "The interface allowed me to navigate Tafsir conceptually, not just textually. I could follow an idea (e.g., Naskh [abrogation]) through al-Ṭabarī to al-Rāzī by clicking on the semantic link, rather than by manually searching two 30-volume works (Almashour et al., 2025)."

This expert feedback directly explains the system's utility. The quantitative success of the annotation model (e.g., the 0.87 F1-score on themes) was the engine that enabled this positive qualitative experience. The high survey scores for "discovering novel connections" were a direct result of the model's ability to accurately tag and link verses that, while lexically dissimilar, were thematically relate connections that are difficult to spot through traditional reading alone.

A specific case study task required the scholars to use the prototype to trace the exegetical development of the term Rūḥ (Spirit). The system automatically annotated all 21 instances of the word and clustered them based on their semantic usage (e.g., 'Divine Spirit', 'Gabriel', 'Human Soul'). The system then linked these verses to the corresponding sections of the selected Tafsir corpus, presenting a synthesized summary of interpretive opinions for each cluster (Schettino et al., 2024).

The scholars' analysis of this data was profound. They noted that the tool instantly highlighted a clear diachronic shift, showing how early exegetes (like al-Ṭabarī) focused on Rūḥ as 'Gabriel', while later, philosophically-inclined exegetes (like al-Rāzī) engaged more deeply with its metaphysical implications as 'Human Soul'. One scholar remarked, "The system did not provide an answer, but it presented the entire landscape of the debate in a single, digestible interface. This is not an automated Tafsir; it is an intelligent map of Tafsir (Gupta & Shivers-McNair, 2024)."

The results from the quantitative and qualitative phases are mutually reinforcing. The high performance of the deep learning model (macro F1 = 0.89) was not merely a technical achievement but the necessary prerequisite for the scholarly utility observed in the qualitative validation. The "Excellent" usability score (SUS = 84.5) confirms that the complex, annotated data can be presented to scholars in an accessible and intellectually productive manner.

The successful completion of the case studies, particularly the Rūḥ analysis, provides a concrete interpretation of the research's contribution. The data demonstrates that the "Intelligent Tafsir Ecosystem" functions as a powerful computational hermeneutic, augmenting rather than replacing the scholar. It successfully translates raw annotation data into a navigable network of knowledge, validating the core thesis that deep learning can serve as the foundation for a new generation of intelligent tools for Islamic studies (Moreno-Ortiz, 2025).

This research successfully demonstrated the feasibility and utility of applying a domain-specific, multi-task deep learning model to the semantic annotation of the Qur'an (Bucher et al., 2025). The quantitative results established a high level of performance, with the bespoke MTL model achieving a macro-averaged F1-score of 0.89. This figure represents a significant 8-point improvement over generic, non-MTL baseline models, confirming the superiority of the specialized architecture.

The model showed particular strength in foundational annotation tasks. It achieved an F1-score of 0.93 for Named Entity Recognition (NER) and 0.87 for the complex, verse-level task of Thematic Labeling. The creation of the 1,850-verse gold-standard corpus, validated with a high average Inter-Annotator Agreement (IAA) of 0.86, was a pivotal outcome, providing the reliable benchmark data necessary for this success.

Qualitative validation confirmed the scholarly value of these quantitative outputs. The prototype "Intelligent Tafsir Ecosystem" received an "Excellent" mean usability (SUS) score of 84.5. Expert scholarly users unanimously (7/7) reported that the system enabled the discovery of novel thematic connections and significantly reduced the time required for comparative exegetical research (6/7 participants).

The case study on the term Rūḥ (Spirit) provided a concrete example of the system's function as a "computational hermeneutic." The tool successfully automated the annotation, clustering, and synthesis of exegetical opinions, allowing scholars to visualize and analyze the diachronic evolution of the concept (L. Zhang et al., 2024). This outcome validated the system's role as an "intelligent map of Tafsir," fulfilling the core research objective.

This study's findings affirm the foundational contributions of existing Digital Qur'an projects, which have prioritized data accessibility. This research extends that paradigm by demonstrating the next logical step: moving from static data retrieval to dynamic, intelligent analysis. It bridges the gap identified in the literature, supplementing retrieval-based platforms with a necessary layer of computational intelligence (X. Zhang et al., 2025).

The results offer a critical corrective to the trajectory of general-purpose Arabic Natural Language Processing. While models like AraBERT are powerful, this study's findings illustrate their limitations when applied to a specialized, historical domain like Qur'anic Arabic. The 8-point F1-score improvement gained from the domain-specific MTL architecture provides empirical evidence that specialized models are not just beneficial but essential for achieving high-fidelity analysis of pre-modern and sacred texts (Dawson & Lewin, 2025).

Previous computational analyses of Tafsir have been necessarily limited, often focusing on topic-modeling a single commentary (Zeng & Liu, 2024). This research differs fundamentally by re-conceptualizing the problem. It avoids analyzing the Tafsir corpus in a vacuum. Instead, it uses the annotated primary text (the Qur'an) as the central, structured anchor to navigate the vast, unstructured secondary corpus (the Tafsir), a methodological approach that has not been systematically explored.

The work situates itself firmly within the broader discourse of Digital Humanities, which seeks to develop computational tools for hermeneutic augmentation. The positive reception from expert scholars aligns this project with successful "human-in-the-loop" models in other domains (e.g., computational literary studies). It contributes a significant non-Western and sacred-text case study to a field, demonstrating the global applicability of digital hermeneutic methods.

The high quantitative performance signifies that the semantic, thematic, and conceptual structures of the Qur'an are computationally legible to a high degree of accuracy. The 0.89 F1-score is not merely a technical benchmark; it signals that the complex, multilayered nature of the sacred text, often described in terms of its ineffability, can be systematically modeled and analyzed. This opens the door to new forms of macro-analytical inquiry.

The strong positive correlation (Pearson's $r = 0.78$) between Inter-Annotator Agreement (IAA) and the model's final F1-score is a profound finding. It signifies that the model's "intelligence" is a direct and measurable reflection of the scholarly consensus encoded in the training data. The model's errors mirror human ambiguities. This result demystifies the AI, framing it not as an inscrutable "black box" but as a powerful formalization of human expertise.

The "Excellent" usability score (SUS = 84.5) and the unanimous scholarly endorsement are highly significant. They signal a successful resolution to the "data overload" problem outlined in the introduction. This result indicates that the immense complexity of deep learning outputs can be successfully translated into an intuitive, usable, and productive scholarly interface, effectively bridging the gap between computational power and humanistic inquiry.

The Rūḥ case study's success in mapping diachronic debate signifies a potential shift in the very scale and nature of scholarly research in Islamic studies. The ability to computationally model the evolution of a concept across dozens of multi-volume commentaries in minutes, rather than years, is a transformative development (Liu et al., 2026). It signifies that scholars can now begin to ask macro-analytical questions about the entire exegetical tradition, questions that were previously methodologically intractable.

The primary implication of this research is the establishment of a validated methodology for a new sub-field of Computational Qur'anic Studies. This study provides the first comprehensive, end-to-end blueprint from schema design and gold-standard creation to MTL model development and qualitative validation. This toolkit can now be systematically applied to the entire Qur'anic text, promising a wave of new discoveries in its thematic, semantic, and rhetorical structures.

The prototype ecosystem has direct implications for the future of Tafsir studies and Islamic education. By providing a synthesized, visual "map" of the exegetical landscape, it flattens the prohibitive learning curve associated with the Tafsir tradition (Y. Zhai, 2025). This tool could revolutionize how Tafsir is taught in universities and seminaries, allowing students to engage in high-level comparative analysis far earlier and more effectively.

The demonstrated inadequacy of generic Arabic models has significant implications for the field of Arabic NLP. This research serves as a strong call for the development of specialized, historical, and domain-specific language models. It justifies investment in creating new benchmark datasets for Classical and Qur'anic Arabic, a necessary step for advancing computational research across the entirety of the vast pre-modern Arabic corpus.

The long-term implication is the potential for the democratization of complex religious knowledge. A public-facing derivative of this ecosystem could offer an intellectually robust alternative to the simplistic, de-contextualized, or non-scholarly interpretations of the Qur'an found online. It provides a pathway for the general public to access the rich, nuanced, and diverse intellectual heritage of Islamic exegesis in a guided and scholarly-grounded manner (Stacchio et al., 2024).

The superior performance of the 0.89 F1-score MTL model resulted directly from its architecture. The model was forced to solve multiple, related tasks simultaneously (NER, Thematics, Semantics). This process, confirmed by the ablation studies, compelled it to learn richer, more synergistic, and more context-aware representations of the Qur'anic text than a single-task model could, thus improving accuracy across all categories.

The "Excellent" 84.5 usability score was achieved because the research adopted a scholar-centric, Design Science paradigm. The prototype was not built by engineers for scholars; it was built with scholars. Its interface succeeded because it was explicitly designed to mirror and accelerate existing hermeneutic workflows comparison, connection-finding, and synthesis rather than imposing an unfamiliar computational one.

The unanimous positive feedback from the expert panel can be attributed to the system's deliberate positioning as an "augment," not a replacement. The "intelligent map" metaphor was critical. The system was successful because it empowered the scholar's interpretive authority by managing cognitive overload. It did not generate its own Tafsir, which would have been rejected; it intelligently organized the human Tafsir tradition, which was embraced.

The strong correlation between human IAA and model F1-score (r=0.78) occurred because the model is a product of its data. The meticulous, expert-driven creation of the gold-standard corpus, with its high-reliability (0.86 IAA), provided a clear and consistent optimization target. The model is, in essence, a highly efficient mathematical formalization of the very annotation schema and scholarly consensus developed by the human experts.

The most critical and immediate future task is the expansion of the gold-standard corpus. The 1,850 annotated verses, while sufficient for this proof-of-concept, represent only a fraction of the Qur'an. A community-driven, collaborative platform, using the validated annotation schema from this study, must be established to pursue the long-term goal of annotating the entire Qur'anic text.

The deep learning model itself must be refined. The lowest-performing category, "Rhetorical Figures" (F1 = 0.79), provides a clear direction for future research. This task's abstract nature may require more sophisticated architectures, such as graph neural networks (GNNs) to model non-sequential relationships or models with wider attention mechanisms to capture long-distance rhetorical patterns.

The "Intelligent Tafsir Ecosystem" prototype must be scaled into a full-fledged research platform. This involves two parallel efforts: ingesting a much larger and more diverse corpus of Tafsir texts (including Shi'i, Sufi, and Mu'tazili commentaries) and computationally modeling the linkages between exegetes. This would allow for the automated generation of "chains of influence" and interpretive networks.

A final recommendation is the development of a public-facing branch of this ecosystem. This tool would require a simplified user interface, curated "guided tours" through major themes, and integrated pedagogical resources. Such a platform would fulfill the project's ultimate aim: leveraging AI to foster a deeper, more nuanced, and more accessible public understanding of the Qur'an and its rich intellectual tradition.

CONCLUSION

The research's most significant finding is the empirical validation of a domain-specific, multi-task deep learning (MTL) architecture as a highly effective method for the semantic

annotation of Qur'anic Arabic, achieving a 0.89 F1-score. This approach's 8-point superiority over generic, single-task models is a distinct contribution, demonstrating that computational legibility of a sacred text is maximized when the model's architecture mirrors the text's multi-layered nature. This quantitative success was directly translated into a functional "Intelligent Tafsir Ecosystem," which expert scholars validated (84.5 SUS) not as an automated interpreter, but as a "computational hermeneutic" that successfully organizes and maps the human exegetical tradition, thereby resolving the "data overload" problem central to modern digital Tafsir studies.

The primary contribution of this study is twofold, providing a symbiotic advancement in both method and concept. Methodologically, it introduces a validated, end-to-end blueprint for Computational Qur'anic Studies, from a reliable annotation schema (0.86 IAA) to a high-performance MTL model. Conceptually, it pioneers the "Intelligent Tafsir Ecosystem," a paradigm shift that re-frames the digital study of Tafsir from static data retrieval (the "digital library" model) to dynamic knowledge discovery (the "intelligent map" model). This conceptual contribution is arguably the more significant, as it provides a new, scholar-centric vision for the Digital Humanities that leverages computational power to augment, rather than replace, humanistic inquiry.

This study's conclusions are bounded by specific, acknowledged limitations which, in turn, define the agenda for future research. The proof-of-concept was established on a 1,850-verse gold-standard corpus; the immediate priority is a collaborative, community-wide effort to expand this annotation to the entire Qur'an. The model's weakest performance on "Rhetorical Figures" (F1 = 0.79) necessitates the exploration of more complex architectures, such as Graph Neural Networks, to capture non-sequential patterns. Finally, the prototype ecosystem must be scaled by ingesting a more diverse corpus of Tafsir (e.g., Shi'i, Sufi) to fully realize its potential as a comprehensive map of the Islamic exegetical tradition.

AUTHOR CONTRIBUTIONS

Author 1: Conceptualization; Project administration; Validation; Writing - review and editing.

Author 2: Conceptualization; Data curation; In-vestigation.

Author 3: Data curation; Investigation.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Abbott, E. E., Apakama, D., Richardson, L. D., Chan, L., & Nadkarni, G. N. (2024). Leveraging Artificial Intelligence and Data Science for Integration of Social Determinants of Health in Emergency Medicine: Scoping Review. JMIR Medical Informatics, 12. https://doi.org/https://doi.org/10.2196/57124

Almashour, M., Aldamen, H. A. K., & Jarrah, M. (2025). Algorithmic feedback and multilingual identity: Translanguaging practices in Jordanian EFL academic writing. Social Sciences & Humanities Open, 12, 102016. https://doi.org/https://doi.org/10.1016/j.ssaho.2025.102016

Alotaibi, B., Almutarie, A., Alotaibi, S., & Alotaibi, M. (2025). Improving Fashion Sentiment Detection on X through Hybrid Transformers and RNNs. Computers, Materials and Continua, 84(3), 4451–4467. https://doi.org/https://doi.org/10.32604/cmc.2025.066050

Alshaya, S. A. (2025). Enhancing Educational Materials: Integrating Emojis and AI Models into Learning Management Systems. Computers, Materials and Continua, 83(2), 3075–3095. https://doi.org/https://doi.org/10.32604/cmc.2025.062360

Bucher, S. F., Gebauer, S., Grieb, J., Körschens, M., Müller, J., Ritz, C. M., Rajendran, R., Weiland, C., Wesche, K., Victor, K., & Römermann, C. (2025). Collectomics in plant biodiversity research − looking into the past to understand the present and shape the future. Basic and Applied Ecology, 88, 1–8. https://doi.org/https://doi.org/10.1016/j.baae.2025.07.002

Burghardt, M. B. T.-R. M. in S. S. (2025). Digital humanities. Elsevier. https://doi.org/https://doi.org/10.1016/B978-0-443-26629-4.00016-2

Chapinal-Heras, D., & Díaz-Sánchez, C. (2024). A review of AI applications in human sciences research. Digital Applications in Archaeology and Cultural Heritage, 32, e00323. https://doi.org/https://doi.org/10.1016/j.daach.2024.e00323

Chen, X., Xie, H., Zou, D., Cheng, G., Tao, X., & Lee Wang, F. (2025). Perceived MOOC satisfaction: A review mining approach using machine learning and fine-tuned BERTs. Computers and Education: Artificial Intelligence, 8, 100366. https://doi.org/https://doi.org/10.1016/j.caeai.2025.100366

Chu, D., Wan, B., Ni, H., Li, H., Tan, Z., Dai, Y., Wan, Z., Tang, T., & Zhou, S. (2025). GeoSMIE: An event extraction framework for Document-Level spatial morphological information extraction. Expert Systems with Applications, 268, 126378. https://doi.org/https://doi.org/10.1016/j.eswa.2024.126378

Dawson, M., & Lewin, J. (2025). Automated classification and mapping for alluvial geomorphic units: Current approaches and future directions. Earth-Science Reviews, 271, 105292. https://doi.org/https://doi.org/10.1016/j.earscirev.2025.105292

Deshmukh, P. V, & Shahade, A. K. (2026). Elevating human-machine collaboration in NLP for enhanced content creation and decision support. Data & Knowledge Engineering, 161, 102505. https://doi.org/https://doi.org/10.1016/j.datak.2025.102505

Dimililer, K., Erdem, B. D., Kayali, D., & Priscilla Olawale, O. (2024). Chapter 3 - Image preprocessing phase with artificial intelligence methods on medical images. In W. A. Zgallai & D. U. B. T.-A. I. and I. P. in M. I. Ozsahin (Eds.), Developments in Biomedical Engineering and Bioelectronics (pp. 51–82). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-323-95462-4.00003-0

Dong, W., Wang, W., Han, X., Huang, J., Li, L., & Huang, Y. (2025). Can GPT-4 provide human-level emotion support? Insights from machine learning-based evaluation framework. Computers in Biology and Medicine, 196, 110789. https://doi.org/https://doi.org/10.1016/j.compbiomed.2025.110789

Eli, E., Wang, D., Xu, W., Mamat, H., Aysa, A., & Ubul, K. (2025). A comprehensive review of non-Latin natural scene text detection and recognition techniques. Engineering Applications of Artificial Intelligence, 156, 111107. https://doi.org/https://doi.org/10.1016/j.engappai.2025.111107

Ganga, B., B.T., L., & K.R., V. (2024). Object detection and crowd analysis using deep learning techniques: Comprehensive review and future directions. Neurocomputing, 597, 127932. https://doi.org/https://doi.org/10.1016/j.neucom.2024.127932

Gao, Q., & Huang, J. (2025). Design and implementation of classical literature sentiment analysis system based on ensemble learning and graph neural network. International Journal of Cognitive Computing in Engineering, 6, 603–616. https://doi.org/https://doi.org/10.1016/j.ijcce.2025.05.004

Gupta, A., & Shivers-McNair, A. (2024). “Wayfinding” through the AI wilderness: Mapping rhetorics of ChatGPT prompt writing on X (formerly Twitter) to promote critical AI literacies. Computers and Composition, 74, 102882. https://doi.org/https://doi.org/10.1016/j.compcom.2024.102882

Harisanty, D., Obille, K. L. B., Anna, N. E. V., Purwanti, E., & Retrialisca, F. (2024). Cultural heritage preservation in the digital age, harnessing artificial intelligence for the future: a

bibliometric analysis. Digital Library Perspectives, 40(4), 609–630. https://doi.org/https://doi.org/10.1108/DLP-01-2024-0018

Jahanshahi, H., & Zhu, Z. H. (2024). Review of machine learning in robotic grasping control in space application. Acta Astronautica, 220, 37–61. https://doi.org/https://doi.org/10.1016/j.actaastro.2024.04.012

Jia, J. (2025). Mining Educational Value of Visualization of Sentiment Classification in Ancient Chinese Literature: International Journal of Web-Based Learning and Teaching Technologies, 20(1). https://doi.org/https://doi.org/10.4018/IJWLTT.385603

Korte, J. W., Bartsch, S., Beckmann, R., El Baff, R., & Hecking, T. (2025). The Different Artificial Intelligences of Science and Wikipedia. Technology in Society, 83, 103034. https://doi.org/https://doi.org/10.1016/j.techsoc.2025.103034

Liu, T., Deng, L., & Zhou, Y. (2026). Unpacking self-regulation and social interaction in "Study With Me" videos through large-scale analytics. Computers & Education, 241, 105488. https://doi.org/https://doi.org/10.1016/j.compedu.2025.105488

Mahmoudi-Dehaki, M., & Nasr-Esfahani, N. (2025). Automated vs. manual linguistic annotation for assessing pragmatic competence in English classes. Research Methods in Applied Linguistics, 4(3), 100253. https://doi.org/https://doi.org/10.1016/j.rmal.2025.100253

Moreno-Ortiz, A. (2025). Corpus sense: A comprehensive tool for advanced text and discourse exploration. Applied Corpus Linguistics, 5(3), 100145. https://doi.org/https://doi.org/10.1016/j.acorp.2025.100145

Nashir, W. A., Mohsen, A. M., Al-Shargabi, A. A., Nour, M. K., & Al-onazi, B. B. (2025). A complete, multi-layered quranic treebank dataset with hybrid syntactic annotations for classical arabic processing. Data in Brief, 62, 111940. https://doi.org/https://doi.org/10.1016/j.dib.2025.111940

Ong, J. C. L., Seng, B. J. J., Law, J. Z. F., Low, L. L., Kwa, A. L. H., Giacomini, K. M., & Ting, D. S. W. (2024). Artificial intelligence, ChatGPT, and other large language models for social determinants of health: Current state and future directions. Cell Reports Medicine, 5(1), 101356. https://doi.org/https://doi.org/10.1016/j.xcrm.2023.101356

Quiñonez-Baca, L.-C., Ramirez-Alonso, G., Guzman-Pando, A., Camarillo-Cisneros, J., & Lopez-Flores, D. R. (2025). Advances in meta-learning and zero-shot learning for multi-label classification: A review. Digital Signal Processing, 163, 105220. https://doi.org/https://doi.org/10.1016/j.dsp.2025.105220

Sarkar Farshi, G. (2025). MLNA – A Python package and app for MultiLingual Network Analysis. SoftwareX, 30, 102104. https://doi.org/https://doi.org/10.1016/j.softx.2025.102104

Schettino, L., Origlia, A., & Cutugno, F. (2024). Though this be hesitant, yet there is method in 't: Effects of disfluency patterns in neural speech synthesis for cultural heritage presentations. Computer Speech & Language, 85, 101585. https://doi.org/https://doi.org/10.1016/j.csl.2023.101585

Schuegraf, P., Shan, J., & Bittner, K. (2024). PLANES4LOD2: Reconstruction of LoD-2 building models using a depth attention-based fully convolutional neural network. ISPRS Journal of Photogrammetry and Remote Sensing, 211, 425–437. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2024.04.015

Sharma, R., & Kukreja, V. (2024). Image segmentation, classification and recognition methods for comics: A decade systematic literature review. Engineering Applications of Artificial Intelligence, 131, 107715. https://doi.org/https://doi.org/10.1016/j.engappai.2023.107715

Shen, D. (2025). Exploring Digital Visualization Techniques for Ancient Literary Works in the Digital Humanities. International Journal of Information System Modeling and Design, 16(1). https://doi.org/https://doi.org/10.4018/IJISMD.389171

Singh, P., Agarwal, A., & Kumar, P. (2025). Navigating the 6G landscape: A systematic review of technology integration, implementation challenges, and strategic solutions. Computer Networks, 272, 111663. https://doi.org/https://doi.org/10.1016/j.comnet.2025.111663

Sree, K. D., Geethika, G., Nair, A. R., Gupta, D., & Veena, G. (2025). Named Entity Recognition for Telugu and Kannada on Naamapadam Dataset Using Various Machine Learning and Deep Learning Algorithms. Procedia Computer Science, 258, 3381–3392. https://doi.org/https://doi.org/10.1016/j.procs.2025.04.595

Stacchio, L., Garzarella, S., Cascarano, P., De Filippo, A., Cervellati, E., & Marfia, G. (2024). DanXe: An extended artificial intelligence framework to analyze and promote dance heritage. Digital Applications in Archaeology and Cultural Heritage, 33, e00343. https://doi.org/https://doi.org/10.1016/j.daach.2024.e00343

Tang, X., Wang, L., & Wang, J. (2026). Language model collaboration for relation extraction from classical Chinese historical documents. Information Processing & Management, 63(1), 104286. https://doi.org/https://doi.org/10.1016/j.ipm.2025.104286

Thakre, B., Yadav, U., & Bondre, S. V. (2025). Chapter 8 - Deep reinforcement learning in healthcare and biomedical application (S. Mahajan & A. K. B. T.-I. in B. E. Pandit (eds.); pp. 241–299). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-443-30146-9.00008-3

Thammastitkul, A. (2025). Optimizing AI-generated image metadata with hybrid color analysis and semantic keyword structuring. Egyptian Informatics Journal, 31, 100775. https://doi.org/https://doi.org/10.1016/j.eij.2025.100775

Vesalainen, A., Tolonen, M., & Ruotsalainen, L. (2024). Document Layout Error Rate (DLER) metric to evaluate image segmentation methods. Machine Learning with Applications, 18, 100606. https://doi.org/https://doi.org/10.1016/j.mlwa.2024.100606

Wang, H., He, M., Zhu, M., & Liu, G. (2025). WCG-VMamba: A multi-modal classification model for corn disease. Computers and Electronics in Agriculture, 230, 109835. https://doi.org/https://doi.org/10.1016/j.compag.2024.109835

Wang, Y., He, Y., Wang, J., Li, K., Sun, L., Yin, J., Zhang, M., & Wang, X. (2025). Enhancing intent understanding for ambiguous prompt: A human–machine co-adaption strategy. Neurocomputing, 646, 130415. https://doi.org/https://doi.org/10.1016/j.neucom.2025.130415

Wang, Z., & Chai, J. (2024). On Metaphor Translation into English Based on Artificial Intelligence. Procedia Computer Science, 247, 1359–1365. https://doi.org/https://doi.org/10.1016/j.procs.2024.10.162

Wu, F., & Tang, W. (2025). An Indoor Design Assistance System Integrating Depth Sensors and Computer Graphics. International Journal of E-Collaboration, 21(1). https://doi.org/https://doi.org/10.4018/IJeC.389057

Xia, K., Hu, Y., Cai, S., Lin, M., Lu, M., Lu, H., Ye, Y., Lin, F., Gao, L., Xia, Q., Tian, R., Lin, W., Xie, L., Tan, D., Lu, Y., Lin, X., Yang, X., Zhong, L., Xu, L., … Xu, H. (2025). GastritisMIL: An interpretable deep learning model for the comprehensive histological assessment of chronic gastritis. Patterns, 6(8), 101286. https://doi.org/https://doi.org/10.1016/j.patter.2025.101286

Xing, Y., Gan, W., Chen, Q., & Yu, P. S. (2025). AI-generated content in landscape architecture: A survey. AI Open. https://doi.org/https://doi.org/10.1016/j.aiopen.2025.10.002

Yang, X., Bahadur Bist, R., Paneru, B., Liu, T., Applegate, T., Ritz, C., Kim, W., Regmi, P., & Chai, L. (2024). Computer Vision-Based cybernetics systems for promoting modern poultry Farming: A critical review. Computers and Electronics in Agriculture, 225, 109339. https://doi.org/https://doi.org/10.1016/j.compag.2024.109339

Yuan, F., Wang, G., Huang, Q., & Li, X. (2025). A newton interpolation network for smoke semantic segmentation. Pattern Recognition, 159, 111119. https://doi.org/https://doi.org/10.1016/j.patcog.2024.111119

Zafar, M. B., Ali, H., & Yasin, T. (2025). Reimagining human creativity and learning in the age of generative AI: A multi-method meta-thematic synthesis. Next Research, 2(4), 100802. https://doi.org/https://doi.org/10.1016/j.nexres.2025.100802

Zeng, X., & Liu, S. (2024). Research on the application of knowledge mapping and knowledge structure construction based on adaptive learning model. Expert Systems with Applications, 249, 123400. https://doi.org/https://doi.org/10.1016/j.eswa.2024.123400

Zhai, Y. (2025). Research on library management paradigm in the AIGC era: Theoretical construction and practical exploration. The Journal of Academic Librarianship, 51(4), 103085. https://doi.org/https://doi.org/10.1016/j.acalib.2025.103085

Zhai, Z., Wang, Z., Xu, L., Zhang, L., Zhang, Y., Yin, J., Zeng, P., Li, C., Sun, T., & Jiang, T. (2025). A systematic review of computer-aided acupoint localization. IScience, 28(11), 113708. https://doi.org/https://doi.org/10.1016/j.isci.2025.113708

Zhang, C., Wang, Y., Zhao, Z., Chen, X., Ye, H., Liu, S., Yang, Y., & Peng, K. (2024). Performance-driven closed-loop optimization and control for smart manufacturing processes in the cloud-edge-device collaborative architecture: A review and new perspectives. Computers in Industry, 162, 104131. https://doi.org/https://doi.org/10.1016/j.compind.2024.104131

Zhang, L., Cao, H., & Yan, Q. (2024). Insights from cross-cultural memes: An empirical study on instagram and Douban. Telematics and Informatics, 94, 102186. https://doi.org/https://doi.org/10.1016/j.tele.2024.102186

Zhang, X., Hong, S., Luo, H., Jiang, Z., & Yuan, F. (2025). A dual-level part distillation network for fine-grained visual categorization. Signal Processing: Image Communication, 138, 117383. https://doi.org/https://doi.org/10.1016/j.image.2025.117383

Zhang, Y., Wang, Y., Sheng, Q. Z., Yao, L., Chen, H., Wang, K., Mahmood, A., Zhang, W. E., Zaib, M., Sagar, S., & Zhao, R. (2025). Deep learning meets bibliometrics: A survey of citation function classification. Journal of Informetrics, 19(1), 101608. https://doi.org/https://doi.org/10.1016/j.joi.2024.101608

Zhao, B., & Pan, D. (2025). Immersive e-learning application in intelligent teaching of English composition based on neural network algorithm. Entertainment Computing, 52, 100710. https://doi.org/https://doi.org/10.1016/j.entcom.2024.100710

Zhao, T., Wang, S., Ouyang, C., Chen, M., Liu, C., Zhang, J., Yu, L., Wang, F., Xie, Y., Li, J., Wang, F., Grunwald, S., Wong, B. M., Zhang, F., Qian, Z., Xu, Y., Yu, C., Han, W., Sun, T., … Wang, L. (2024). Artificial intelligence for geoscience: Progress, challenges, and perspectives. The Innovation, 5(5), 100691. https://doi.org/https://doi.org/10.1016/j.xinn.2024.100691

Zhu, W., & Yuan, F. (2025). Multimodal learning with feature fusion transformer for image captioning. Displays, 90, 103126. https://doi.org/https://doi.org/10.1016/j.displa.2025.103126

Zhu, Y. (2024). A knowledge graph and BiLSTM-CRF-enabled intelligent adaptive learning model and its potential application. Alexandria Engineering Journal, 91, 305–320. https://doi.org/https://doi.org/10.1016/j.aej.2024.02.011

Zinnen, M., Madhu, P., Leemans, I., Bell, P., Hussian, A., Tran, H., Hürriyetoğlu, A., Maier, A., & Christlein, V. (2024). Smelly, dense, and spreaded: The Object Detection for Olfactory References (ODOR) dataset. Expert Systems with Applications, 255, 124576. https://doi.org/https://doi.org/10.1016/j.eswa.2024.124576