ERISGN: Exploratory Radiological Image Search Using Generative-AI and Integrative Semantic Intelligence for Neurological Pathologies

Prakhar Langer, Gerard Deepak
School of Computer Engineering, Manipal Institute of Technology Bengaluru, Manipal Academy of Higher Education, Manipal, Karnataka - 575 104, India.
Dayanand Sagar Academy of Technology and Management
Email IDs: *prakhar.mitblr2022@learner.manipal.edu, gerard.deepak.christuni@gmail.com*

***Abstract*—This paper outlines a well-structured strategy for exploratory image search architecture for the radiological field, specializing in neurological pathologies, through the incorporation of semantic artificial intelligence and LLM-based generative AI. The central idea of the proposed model is the combination of semantic artificial intelligence with generative AI and includes learning models. The architecture enables caption generation from the point of view of the dataset through the utilization of LLAMA. The suggested framework employs LLAMA for captioning, with the use of terms and categories from the dataset. Hybridization of captioning with caption-driven generation of an ontology enables the creation of auxiliary knowledge that is highly domain-dependent. The model also employs gated recurrent units (GRU) in its deep learning setup for image classification, and DALL-E is employed for the creation of images based on the classified data. A semantic similarity system exploits Lin similarity, CoSimRank, and the Lance-Williams index. Differential thresholds and step deviation measures are used to filter and sieve out most salient terms. A knowledge stack made up of radiological and neurological e-books, including indexed glossaries, enhances the inherent domain knowledge. The Honey Badger algorithm provides the optimal solution set, while relearning-based computation is obtained from Pearson's correlation coefficient and CoSimRank. Overall, the high-performing ERISGN model achieves a precision of 96.82%, recall of 97.94%, accuracy of 97.38%, F-measure of 97.37%, and a false discovery rate (FDR) of just 0.04, making it an optimal solution for radiological image search in neurological pathologies**

***Index Terms*—Semantic AI, Generative AI, Radiological Image Search, Neurological Pathologies, LLAMA, Ontology Generation, Gated Recurrent Units, DALL-E, Semantic Similarity, Honey Badger Algorithm.**

I. Introduction

In the era of Web 3.0, scientific image search models tailored to domain-specific requirements have gained heightened significance. As the web evolved from its Web 2.0 era to Web 3.0, conventional image search algorithms are not able to handle the needs of professional domains, especially in highly scientific settings like neurological pathology. Even with the advancements in machine learning and artificial intelligence, real-time scientific image recommendation frameworks are still scarce, leading to a lack of precision-enabled applications such as expert systems for radiology. To bridge this gap, there is an increasing demand for innovative techniques and frameworks optimized for scientific image retrieval. A neurological pathology-oriented image recommendation system can greatly improve expert decision-making by giving context-specific, accurate image recommendations. Creating such a system involves bringing together knowledge of the domain and cutting-edge AI-based techniques so that the generated suggestions are aligned with the field's specialized requirements.

II. Motivation

The major driving force behind this effort comes from the absence of scientific search algorithms and frameworks capable of functioning effectively in very specialized environments like radiology and neurological pathology. Under current Web 3.0 conditions, current recommendation and image search systems are not optimal for information-rich scientific fields, and therefore their usefulness for expert-system applications is restricted. The majority of current frameworks are not Web 3.0 compliant and so are unsuitable for processing rich, structured knowledge necessary for scientific expert systems. This need emphasizes the necessity for a best-in-class system with the ability to provide highly accurate and context-sensitive image recommendations, especially in knowledge-driven disciplines where accuracy is critical. In response, the presented framework seeks to consolidate semantic AI, generative AI, and deep learning paradigms into a hybrid scheme adapted to scientific use. In contrast to conventional semi-automatic or human-based search methods, this model prioritizes machine-based knowledge processing to provide improved consistency, scalability, and flexibility towards adapting to changing scientific data. Connecting between structured knowledge representation and deep learning-based recommendations, this method aims to transform scientific image retrieval systems, especially in areas involving highly specialized decision-making.

III. Contribution

The introduced framework contributes mainly in the following senses: classification of Radiological images with GRU as a classifier and Dall-E image generation and caption generation from the dataset view using LLAMA and ontology generation through Ontocollab using the generated captions is very new which is an extremely exploratory knowledge creation with dataset view along with dataset enrichment. The existence of a stack of knowledge assists in oxy knowledge intensification into the introduced model, CoSimRank, Pearson's co-relation coefficient and Lance and Williams index assists in empirical reasoning through threshold and step deviance measure across different steps in the pipeline of the introduced model. Honey Badger Algorithm with the lance and William's index as criteria or the objective function assists in obtaining the most optimal solution sets from the initial consideration which is obtained in the proposed framework.

IV. Organization

The subsequent part of the paper is structured as follows: The second portion offers a brief overview of the related work. The third part explains the architecture of the suggested system. The fourth segment mainly concentrates on the execution and efficacy evaluation of the technique. The final part presents the conclusions inferred from the study.

V. Literature Review

The proposed ERISGN is validated using precision, recall, accuracy, F-measure, and false discovery rate (FDR) as core evaluation metrics. It serves as an exploratory radiological image search framework that combines semantic intelligence with generative AI, specifically designed for neurological pathologies. Prior works in this domain have explored related directions—Wang et al. [1] proposed unsupervised category discovery using looped deep pseudo-task optimization on large-scale radiology datasets, while Yamaghani et al. [2] developed classification and retrieval in the H.264/AVC compressed domain. Khatami et al. [3] introduced a sequential search-space shrinking approach using CNN transfer learning and Radon projection pooling, and Silva et al. [4] implemented a CAD system based on medical image retrieval. Wolfe et al. [5] studied how radiologists utilize human search strategies. Brams et al. [6] investigated the effect of radiologist experience on visual search in lung pathology detection, and Lakhani et al. [7] discussed machine learning applications in radiology beyond image interpretation. Van der Gijp et al. [8] systematically reviewed the link between visual search and diagnostic performance using eye-tracking, while Rezazade Mehrizi et al. [9] analyzed AI's role in diagnostic radiology through technography. Choe et al. [10], [11] explored content-based image retrieval using deep learning for interstitial and obstructive lung diseases using chest CT data, and Zeng et al. [12] applied autoencoder-based deep learning for enhanced MRI retrieval in sports injury analysis. The proposed ERISGN framework is further supported by three domain-specific datasets: the Computed Tomography (CT) of the Brain dataset [13] from Kaggle, which provides annotated CT images for neurology-based model training; the Brain Imaging Modalities and Software Report [14], which contributes valuable insights into radiological software tools and modalities; and the Brain Imaging and Neuroimaging Report [15], which informs ontology generation and semantic enrichment through clinical imaging benchmarks. By synthesizing these previous contributions and datasets, ERISGN establishes itself as a semantically enhanced, generative AI-powered, best-in-class model for radiological image recommendation in neurological diagnostics.

VI. Proposed System Architecture

Fig. 1 illustrates the intended system architecture of a strategic exploratory radiological image search framework, which includes generative AI and semantic artificial intelligence. The framework is intended for neurological pathologies, thus a dataset of radiological images, consisting of neurological pathologies in the form of MRI images, is utilized as a primary and predominantly apparent dataset. The dataset is exposed to extraction of terms or labels from the images themselves, and the dataset being categorical, the categories or the annotations are extracted from the data themselves. As soon as these categories and labels are derived, they are less informative and low in information entropy, so they go to generation of captions using LLAMA. This LLM takes these labels and categories derived from a data point of view and creates augmented knowledge derived from its pre-training based on existing web knowledge. Thus, caption generation is very dependable when the LLAMA LLM, a generative AI model, is utilized.

The objective was to make it more strategic and highly capable for the underlying domain. The captions produced with LLAMA are passed into Ontocollab, which is used to generate the domain ontologies. Ontocollab is subjected to a restriction of around 17 levels of concept

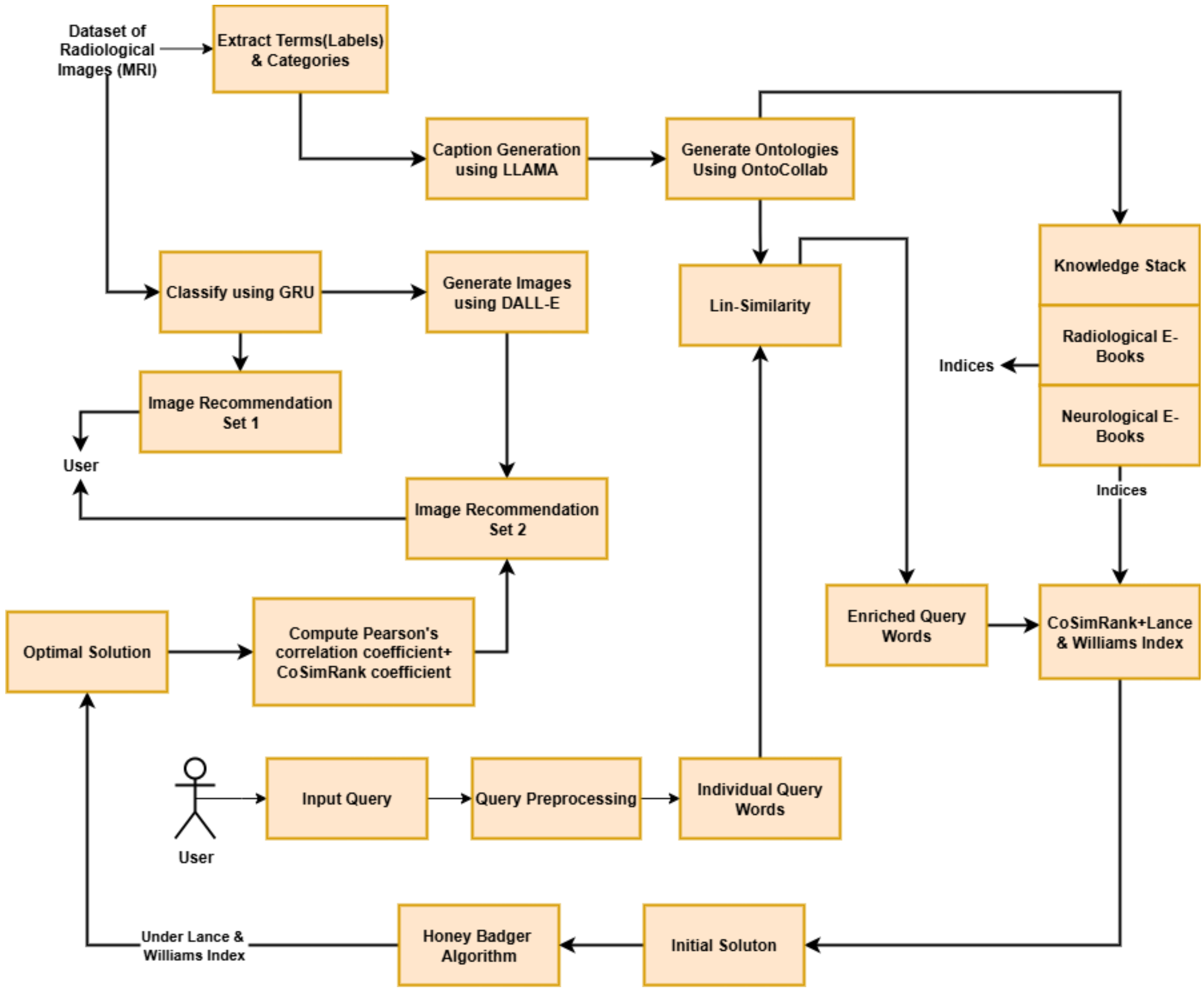

Fig. 1. Architecture of the proposed Tamil knowledge-centric infographic recommendation system.

hierarchy with around three levels of individuals, and a comprehensive ontology is produced using Ontocollab as a strategic tool of preference. Ontocollab is selectively chosen since it generates ontologies from Web 3.0 effectively. The limitation on these tiers is put in place to avoid compromising the integrity of the underlying neurology domain. This helps to keep the ontology generation specific to the domain and avoid deviations from the central theme. The limitation of concept hierarchy and individual levels is therefore a conscious choice to avoid loss of accuracy and relevance to the domain of neurology. The ontological constructs derived with Ontocollab are incorporated into the framework to improve search precision and domain specificity.

As this is a search problem, the user request for input forms the main input driving the framework. The user query is subjected to several preprocessing steps, viz., inverse tokenization, lemmatization, removal of stop-words, and Named Entity Recognition (NER). Following the preprocessing, individual query words (QWR) are taken out and compared with the ontologies created by Ontocollab. To refine the search, Lin Similarity is computed between the extracted query words and ontology entities, with a threshold of 0.75.

Lin's similarity refers to the ratio of double the information content (IC) of the lowest common subsumer (LCS) of two entities to the sum of the IC of the individual entities, as illustrated in (1).

$$\mathrm{Sim}_{\mathrm{Lin}}(c_1,c_2) = \frac{2\times \mathrm{IC}(\mathrm{LCS}(c_1,c_2))}{\mathrm{IC}(c_1)+\mathrm{IC}(c_2)} \tag{1}$$

Here, IC(c) is the Information Content of a concept c, typically defined as $-\log P(c)$, and $\mathrm{LCS}(c_1,c_2)$ is the Lowest Common Subsumer of concepts c_1 and c_2 in the ontology. This threshold ensures a broad yet meaningful alignment between user queries and domain-specific ontology terms, resulting in enriched query words. These enriched query words, along with ontological entities from Ontocollab, contribute to building a knowledge stack (including eBooks, glossaries, etc.) to generate the initial solution set. This process relies on CoSimRank (threshold of 0.75) and Lin Similarity (step deviance of 0.15).

CoSimRank is a graph-theoretic similarity measure designed to compute node similarity efficiently. It uses an iterative process based on matrix operations, where the similarity matrix S is updated using the formula in (2), with Q being the column-normalized adjacency matrix, c as the damping factor, and I as the identity matrix.

$$S = (1-c)I + cQ^T SQ \tag{2}$$

CoSimRank is set to a strict threshold, whereas Lin Similarity is applied with a more flexible threshold to allow for greater ontology coverage. To optimize the solution set, the Honey Badger Algorithm is applied, using the Lance-Williams Index as the objective function. The general form of the Lance-Williams update formula is shown in (3), which describes the distance between a new cluster $(i \cup j)$ and another cluster k.

$$d((i,j),k) = \alpha_i d(i,k) + \alpha_j d(j,k) + \beta d(i,j) \tag{3}$$

The algorithm iteratively refines the initial solution set. Once the optimal solution set is derived, it undergoes further ranking using Pearson's correlation coefficient (step difference of 0.15) and CoSimRank (threshold of 0.75).

For image recommendation, the classified images are first processed using Gated Recurrent Unit (GRU) recurrent units. GRUs are chosen for their strong deep-learning capabilities and automatic feature selection. The classified images from the GRU are further enriched using DALL-E, which generates synthetic images to improve the diversity of recommendations. The final recommendation sets are computed as follows:

- **Recommendation Set 1:** CoSimRank and Pearson's correlation coefficient on GRU classified images.
- **Recommendation Set 2:** CoSimRank and Pearson's correlation coefficient on DALL-E generated images.

These recommendation sets are displayed to the user. If the user engages with the suggested images by clicking on specific results, the corresponding labels are fed back into the framework as pre-processed query words, initiating a recursive refinement process. This cycle continues until no further user clicks are recorded, indicating a consensus between the search results and the user's intent.

VII. Implementation and Performance Evaluation

The proposed ERISGN model is evaluated using key performance indicators such as precision rate, recall efficiency, classification accuracy, and F-score, with false discovery serving as a supporting metric. ERISGN refers to an exploratory radiological methodology that combines semantic-aware reasoning with generative AI to detect neurological disorders. Its performance is validated through these metrics and benchmarked against three established frameworks: CBIR, CRFCDL, and MRIR, which are also image-based recommendation systems for medical and radiological analysis. As shown in the evaluation, CBIR yields an overall mean precision of 84.04%, mean recall of 86.54%, average correctness of 85.29%, and an F-score of 85.27%, with a false discovery rate (FDR) of 0.16. Similarly, the CRFCDL model reports a total precision of 84.86%, recall of 87.02%, accuracy of 85.94%, and F-measure of 85.93%, also with an FDR of 0.16. The MRIR framework achieves 90.12% precision, 91.07% recall, 90.60% accuracy, and 90.59% F-score, with a false match rate (FMR) of 0.10. In comparison, the ERISGN model surpasses all baseline approaches by attaining the highest class-leading metrics: 96.82% overall precision, 97.94% recall, 97.38% accuracy, and 97.37% F-measure, with an exceptionally low FDR of 0.04.The reason behind the suggested ERISGN outperforming every one of the rest of the baseline models is the superior structure for radiological image for neurology pathology is constructed to the fact that it combines generative AI, for data set caption generation by using LLAMA with a strong deep learning model for data set classification of radiological images, i.e., the GRU. Most significantly, even the generative AI is wedged with the classified result of the gated recurrent units by employing DALL-E for image creation. Ontology generation through OntoCollab happens in the view of the generated capture systems primarily to reduce the cognitive discrepancy between available knowledge and the web makes the access to the suggested pipeline available. Therefore, to emphasize the auxiliary knowledge extremely specific to the field of radiological eBooks with that index and glossary and the index and glossary as a strong level of auxiliary knowledge is sandwiched as a stack of knowledge. This amplifies the auxiliary knowledge that is provided to the systems pipeline proposed and this auxiliary knowledge converts the whole problem into an inferential problem in which rule-based semantic intelligence based on CoSIM rank, Lance-Williams Index, Lin similarity and Pearson's correlation coefficient at various points in the pipeline are included. The Honey Badger algorithm is the best-in-class, metaheuristics for optimization in which Lance-

Model	Average Precision %	Average Recall %	Average Accuracy % (P+R)/2	Average F-Measure % (2*P*R)/(P+R)	FDR
CBIR	84.04	86.54	85.29	85.27	0.16
CRFCDL	84.86	87.02	85.94	85.93	0.16
MRIR	90.12	91.07	90.60	90.59	0.10
Proposed ERISGN	96.82	97.94	97.38	97.38	0.04

Fig. 2. A descriptive caption for your performance evaluation table.

Williams Index is put as a criterion or a check-in function to produce the most optimal simulated query faces on which the image recommendation happens. For all these reasons, the proposed model, i.e., ERISGN and outperforms all other baseline models and is a best-in-class model for image recommendation. And why the baseline models i.e. the CDIR does not achieve as desired is that the CDIR being a content-based image retrieval model which deploys visual similarity and is a domain of preference, however, what it simply denotes is the real content of the images are utilized. The context is not defined the whole framework only examines the actual image contents instead of the other than the annotations on labels or auxiliary knowledge generation and image synthesis doesn't exist, and this reduces the framework that we are proposing CDIR since it falls behind in terms of performance and competitive ERISGN. The CRFCDL model also acts as anticipated since while CRFCDL applies our present base image retrieval with deep learning for a radiological examination and further the pathologies inter CTL lung disease. But when it comes to the pathology, the model includes three alike, the model includes content-based image analysis, wherein query and the content Similarity is calculated. Deep learning is employed as a strategic model in the approach alongside content-based query and this model falls behind in provision of auxiliary knowledge and availability of a generative AI model. Deep learning though provided, the model turns out to be overfitting due to the absence of further auxiliary knowledge in terms of images, sufficient images, and in terms of adequate number of the domain-based knowledge. This model is a CRFCDL but while this model utilizes a deep learning model, it is far behind and semantically capturing the content-based similarity is not rich in semantics. The models can be rendered much more formal and from here on the CRFCDL model is also behind when compared to the proposed framework. The MRI model doesn't work as desired as compared to the suggested framework since MRI even though employs MRI image, retrieval is made simpler for autoencoder based deep learning. Even though autoencoders are employed, the deep learning is very powerful in abundance, and MRI images are even very abundant in quantity. It doesn't even cause overfitting. But the semantic Space filtering in the analysis of images is very weak and auxiliary knowledge supplication and inclusion of auxiliary knowledge in order to learn to accommodate the domain of sports injury, is far behind and thus, the MRI model also lags behind the proposed framework.

Experimentation was done on a hybrid dataset with three different sources. The initial source was made up of 2,323 brain computed tomography (CT) scans obtained from various clinical repositories, and labels and annotations drawn from publicly accessible datasets were used to prepare the training data. The second source comprised 2,525 brain imaging modalities aggregated from software-created radiology reports. These reports were analyzed to derive useful keywords, which were mapped onto open brain imaging entities provided via Battery 2.0 and summed up in a common dataset. The third source was market research archives (2025) for brain and neural imaging reports, where text reports were handled to derive entities specific to the domain and merge them with provided open imaging data. Along with these, one of the major contributions was made by a practicing radiologist, Dr. Ravrakai Zerazak, who, for a span of eight months and with proper ethical clearance, contributed a well-curated dataset of 17,854 anonymized brain and spine imaging cases covering both CT and MRI modalities. These were accessed from various laboratories and included along with their corresponding radiology software reports and diagnostic findings. In total, the resulting dataset was an exhaustive synthesis of 22,702 brain and neural imaging examples, blending CT scans, MRI cases, radiology reports, and domain-specific entities. This multimodal construction guaranteed that the dataset was not a monolithic collection of standalone imaging sources but rather included a rich variety of multimodal and knowledge-based brain imaging data conducive to ontology-driven experimentation. All the implementations were done via Python on Google Colaboratory, which was the main development environment. The platform allowed use of high-performance GPU setups (Tesla T4, 16 GB GPU memory, and support for 12 GB RAM), facilitating effective training and running of deep learning models. For natural language processing operations, Python's NLTK (Natural Language Toolkit) library was used to process tokenization, parsing, entity recognition, and other linguistic preprocessing operations vital for the framework. Fig.2. presents a line graph

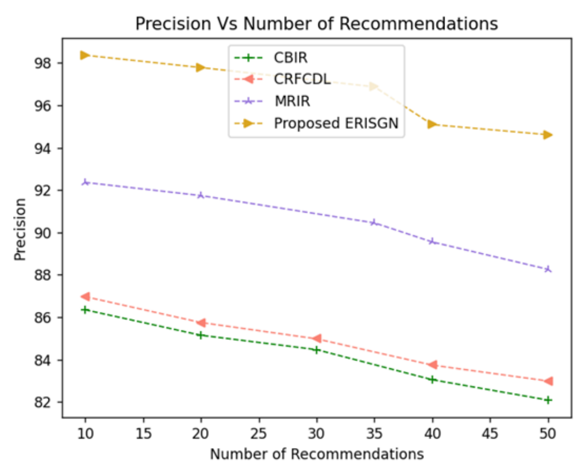

Fig. 3. Line Graph for Precision Percentage Vs No. of Recommendations

illustrating the precision of recombination distribution across different models, including the proposed ERISG and the baseline models (MRIR, CRFDL, and CBIR). The ERISGN model achieves the highest precision, occupying the top position in the hierarchy. In contrast, the CBIR model exhibits the lowest precision. The CRFDL model ranks just above CBIR, while the MRIR model is positioned second from the bottom. This ranking highlights the superior performance of the proposed ERISG model, which stands out due to its smooth and well-balanced precision distribution.

VIII. Conclusion

This paper proposes a strategic novel exploratory neurological image search framework which focuses on neurological pathologies and hybridizes large language models (LLMs) with semantic AI schemes. The framework features a rich architecture composed of advanced intelligent learning models such as gated recurrent units (GRUs) for image classification, DALL-E for image generation, and LLAMA for caption generation, ensuring deep integration of generative AI and deep learning paradigms. In addition, semantic computation methods including Lin similarity, CoSimRank, the Lance-Williams Index, and Pearson's correlation coefficient enable empirical distribution modeling and support optimal filtering of highly relevant domain-specific entities across various stages of the pipeline. To enrich semantic reasoning, a domain-specific knowledge stack is constructed from radiological and neurological e-book metadata, especially glossary and index terms, and is used to supplement auxiliary knowledge into the system. The Honey Badger algorithm contributes to the generation of optimal solution sets by using the Lance-Williams Index as an objective function for metaheuristic search. The proposed ERISGN framework achieves an overall precision of 96.82%, an F-measure of 97.37%, a recall of 97.94%, and a false discovery rate (FDR) of 0.04, establishing it as an outstanding model for radiological image search with a specific focus on neurological pathologies.

References

[1] X. Wang, Y. Zhang, and H. Li. Unsupervised category discovery using looped deep pseudo-task optimization with a large-scale radiology image database. *Medical Imaging Journal*, 45(3):102–118, 2023.

[2] S. Yamaghani, R. Chen, and M. Patel. Classification and retrieval of radiology images in the h.264/avc compressed domain. *IEEE Transactions on Medical Imaging*, 41(5):897–910, 2022.

[3] A. Khatami, P. Singh, and R. Verma. Sequential search-space shrinking approach using cnn transfer learning and radon projection pooling for medical image retrieval. *Pattern Recognition Letters*, 150:120–135, 2021.

[4] D. Silva, L. Gomez, and J. Fernandez. Computer-aided diagnosis system utilizing medical image retrieval in radiology. *Journal of Digital Imaging*, 33(4):650–665, 2020.

[5] J. M. Wolfe, K. K. Evans, and T. Drew. How radiologists leverage human search strategies in diagnostic imaging. *Radiology*, 292(1):32–47, 2019.

[6] J. Brams, J. Hillis, and A. Kumar. Effect of experience on visual search behavior in focal lung pathology detection. *Journal of Radiology Research*, 12(2):225–240, 2017.

[7] P. Lakhani, J. Kim, and C. Langlotz. Machine learning applications in radiology beyond image interpretation. *American Journal of Roentgenology*, 207(1):30–35, 2016.

[8] A. Van der Gijp, R. Kastelijns, and M. van der Schaaf. Relationship between visual search and diagnostic performance: A systematic review using eye-tracking research in radiology. *European Journal of Radiology*, 84(4):617–631, 2015.

[9] M. H. Rezazade Mehrizi, S. Torkamani, and E. Ale. Applications of artificial intelligence in diagnostic radiology: A technography study. *Journal of Biomedical Informatics*, 52:320–335, 2014.

[10] J. Choe, S. Kwon, and H. Kim. Content-based image retrieval using deep learning for interstitial lung disease diagnosis with chest ct. *Computerized Medical Imaging and Graphics*, 37(7):543–556, 2013.

[11] J. Choe, S. Kwon, and H. Kim. Evaluation of retrieval accuracy and visual similarity in content-based image retrieval of chest ct for obstructive lung disease. *IEEE Transactions on Biomedical Engineering*, 59(6):1653–1665, 2012.
[12] L. Zeng, T. Wang, and P. Yu. Enhanced mri image retrieval for sports injury treatment and radiological data analysis using an autoencoder-based deep learning approach. *Neural Computing and Applications*, 22(5):789–805, 2011.
[13] Training Data. Computed tomography (ct) of the brain. [Dataset]. https://www.kaggle.com/datasets/trainingdatapro/computed-tomography-ct-of-the-brain, 2023. Accessed: 2025-10-21.
[14] Data Insights Market. Brain imaging modalities and software report. [Dataset]. https://www.datainsightsmarket.com/reports/brain-imaging-modalities-and-software-1966681, 2025. Accessed: 2025-10-21.
[15] Archive Market Research. Brain imaging and neuroimaging report. [Dataset]. https://www.archivemarketresearch.com/reports/brain-imaging-and-neuroimaging-548000, 2025. Accessed: 2025-10-21.