A Comprehensive Survey on AI-based Methods for Patents

Homaira Huda Shomee **Zhu Wang** **Sathya N. Ravi** **Sourav Medya**
Department of Computer Science, University of Illinois at Chicago
{hshome2, zwang260, sathya, medya}@uic.edu

Abstract

Recent advancements in Artificial Intelligence (AI) and machine learning have demonstrated transformative capabilities across diverse domains. This progress extends to the field of patent analysis and innovation, where AI-based tools present opportunities to streamline and enhance important tasks in the patent cycle such as classification, retrieval, and valuation prediction. This not only accelerates the efficiency of patent researchers and applicants but also opens new avenues for technological innovation and discovery. Our survey provides a comprehensive summary of recent AI tools in patent analysis from more than 40 papers from 26 venues between 2017 and 2023. Unlike existing surveys, we include methods that work for patent image and text data. Furthermore, we introduce a novel taxonomy for the categorization based on the tasks in the patent life cycle as well as the specifics of the AI methods. This survey aims to serve as a resource for researchers, practitioners, and patent offices in the domain of AI-powered patent analysis.

1 Introduction

Recent progress in AI and machine learning has shown transformative capabilities across various domains including NLP, computer vision, and healthcare [11, 45]. The field of patents and technological innovation is not an exception. AI-based tools can streamline the complex patent related tasks such as classification, retrieval, and valuation prediction. For instance, for patent examination, patent offices often rely only on the examiner to judge whether a technology is innovative and thus, patentable. However, it is challenging for the human examiner to stay updated on various domains due to the exponential growth in technology. This intersection of AI and patent processes can accelerate the efficiency of the patent system—patent reviewers as well as applicants—and help in a faster technological innovation to benefit our society.

The patent application and granting process in the patent life cycle involves complex tasks that require significant human effort for both applicants and reviewers. To streamline these complex processes, AI can be helpful, particularly in patent classification, retrieval, and quality analysis [32]. Patent classification can benefit from the AI-based multi-label classification tools for the hierarchical schemes: International Patent Classification (IPC) and the Cooperative Patent Classification [50, 1]. To evaluate novelty and avoid infringement, the patent retrieval task becomes important while filing or reviewing a new patent application. On the other hand, quality analysis also requires a substantial amount of effort. AI-based representation learning methods can be useful in both tasks [9, 40]. Lastly, recent generative AI tools can generate accurate and technical language descriptions for patents and thus, are useful to optimize human resources and precision in patent writing [37].

In the literature, there is a lack of recent surveys on AI tools for patent analysis. The most recent survey [32] does not cover the recent studies in this area. Moreover, it focuses heavily on text-based approaches. Our survey aims to bridge this gap by providing a comprehensive and detailed summary of existing AI methods in more than 40 papers that appeared in 26 different venues from 2017 to

Preprint. Under review.

2023 for patent analysis. We include recent AI tools both for images and text data in patent analysis. Moreover, we introduce a novel taxonomy to categorize these methods based on the relevant tasks and the nature of the methods.

1.0.1 Overview

Figure 1 provides the hierarchical organization where the most important tasks and their related methods are presented. We organize the survey into the following sections: Section 2 provides background on both the relevant tasks in the patent life cycle as well as the patent datasets. Section 3 summarizes the methods for four individual tasks: patent classification, retrieval, quality analysis, and generation. These methods are further grouped based on their commonality. Finally, Section 4 provides a few important research directions including the use of generative AI and multimodal learning. Note, the frequently used AI methods in the papers covered by this survey are referred in Table 1.

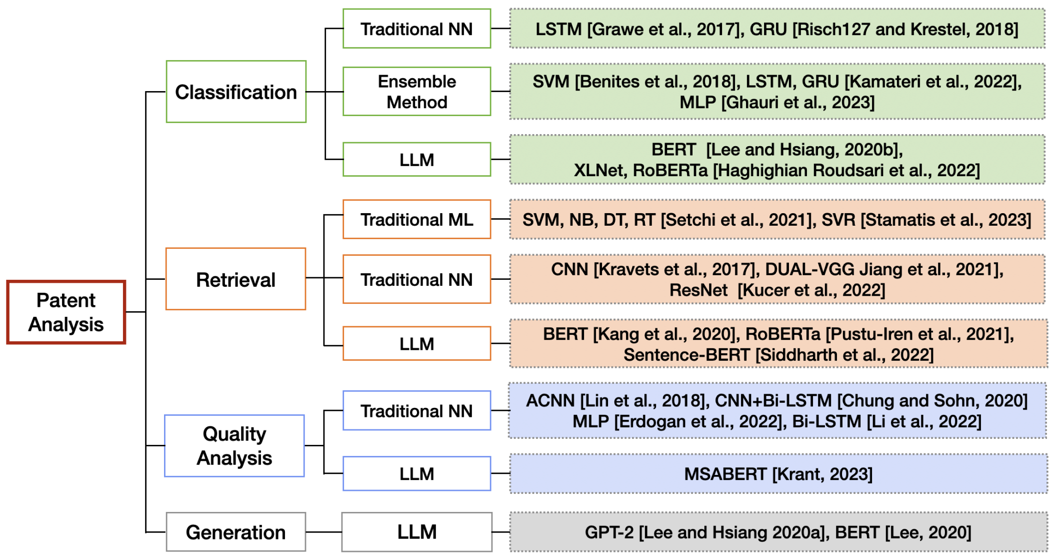

Figure 1: The schema of the main organization with a few examples of the methods in each patent related tasks. We summarize the methods for four individual tasks: patent classification, retrieval, quality analysis, and generation. “NN" and “LLM" denote neural network and large language model respectively.

2 Background

A patent grants the owner or holder exclusive rights to an invention, and can be a novel product or a process that usually offers a unique method or technical solution. In exchange for this right, inventors must publicly disclose detailed information about their invention in a patent application[1]. The United States Patent and Trademark Office (USPTO[2]) issues three types of patents: utility, design, and plant. Utility patents protect the rights related to how the invention works or is used. It provides the entitlement to the functionality of a product. On the other hand, design patents protect the right of the look of an invention and are intended to safeguard the form of a product. A plant patent is issued to an inventor who has discovered and invented a unique variety of a plant and asexually reproduced it. In this work, we focus on utility and design patents considering the relevance of the AI tools on the tasks related to these categories. Here, we describe the relevant tasks, the datasets, and their sources.

[1] `https://www.wipo.int/classifications/ipc/`
[2] `https://www.uspto.gov/`

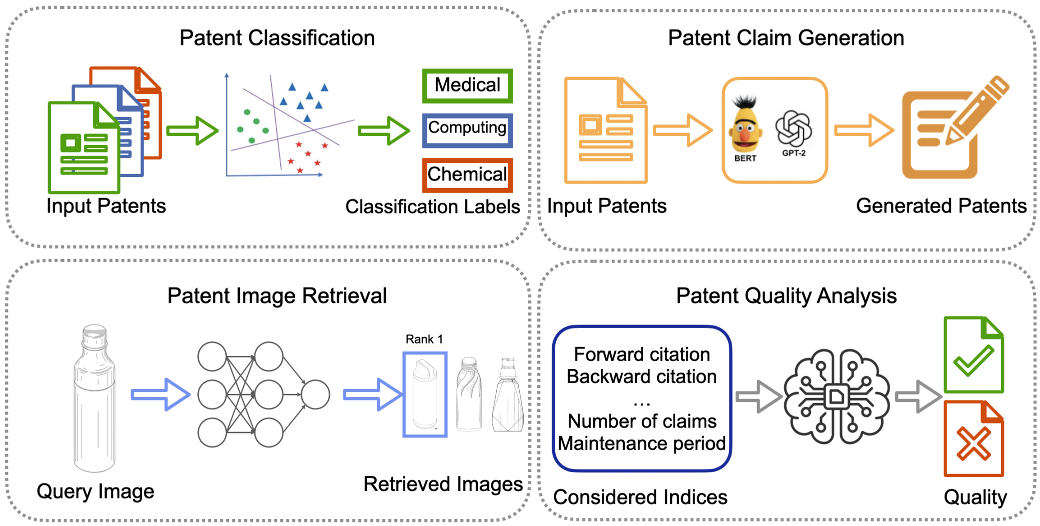

Figure 2: The overview of four major tasks of patent analysis. The patent retrieval task includes obtaining relevant patents (text and images). Please refer to the detailed descriptions of these tasks in Section 2.1.

2.1 Tasks

Patent application and granting processes are complex and involve many complex tasks that involve both applicants and patent reviewers. The AI tools can help in simplifying these tasks. Here, we outline the most relevant tasks that can exploit AI-based methods.

Patent Classification: Patent classification is an important and time-consuming task in the patent life cycle. This task involves a multi-label classification for patents where the classification scheme is hierarchical and a patent can get multiple labels. There are two widely used patent classification systems: the International Patent Classification (IPC) and the Cooperative Patent Classification (CPC). The IPC comprises 8 sections, 132 classes, 651 subclasses, 7590 groups, and 70788 subgroups in a hierarchical order (i.e., sections have classes and classes have subclasses, and so on). CPC is an expanded version of the IPC and is collaboratively administered by the European Patent Office (EPO) and the USPTO. It consists of around 250,000 classification entries and is divided into nine sections (A-H and Y), which are further broken down into classes, subclasses, groups, and subgroups[3].

Patent Retrieval: Patent Retrieval (PR) focuses on developing strategies and methods to effectively and efficiently retrieve relevant patent documents and images based on specific search queries [51]. PR plays a crucial role in identifying new patents related to new inventions. It is essential for evaluating novelty of a patent as well as ensuring that it does not infringe on existing patents. Moreover, patent image retrieval can serve as a source of design inspiration.

Quality analysis: Businesses have shown great interest in evaluating patent value due to its significant impact in generating revenue and investment [2]. Investors usually aim to predict the future value of a technological innovation from the target firm while making investment decisions. As a result, many companies hire professional patent analysts for quality analysis. This complex task demands substantial human effort as well as expertise in various domains [40]. The quality of a patent can be assessed using various measures, including the number of forward or backward citations, the number of claims, the grant lag, patent family size, the remaining lifetime of the patent [2, 12].

Patent Generation: Patents usually have a considerable amount of written text which requires significant human resources. The patent generation task involves generating specific sections of a patent such as abstract, independent claims, and dependent claims based on instructions for an AI tool. Patent documents require precise and technical language to accurately describe the invention and its claims [47]. AI assisted patent generation will help to automate the drafting process which

[3] `https://www.cooperativepatentclassification.org/`

involves time, effort, and legal requirements. This will also reduce the amount of patent attorney time which will be a substantial cost saver.

Table 1: Popular AI methods in the literature. We use the acronyms frequently in our survey.

Acronym	Full Name	Paper
LSTM	Long short-term memory	[21]
CNN	Convolutional Neural Networks	[34]
Bi-LSTM	Bidirectional Long Short-Term Memory	[16]
Word2Vec	–	[42]
GRU	Gated Recurrent Units	[7]
Bi-GRU	Bidirectional Gated Recurrent Units	[7]
DUAL-VGG	Dual Visual Geometry Group	[54]
FastText	–	[26]
BERT	Bidirectional Encoder Representations from Transformers	[11]
RoBERTa	Robustly Optimized BERT Pre-training Approach	[41]
SciBERT	Scientific BERT	[3]

3 Methods

There has been a surge in research interest in developing AI-based methods in patent analysis. We organize the popular and important patent tasks that can benefit from AI tools. An overview of the major categorization of the patent tasks is shown in Figure 2.

3.1 Classification

One of the major tasks of a patent reviewer is to assign a CPC or IPC code to the submitted patent. This task is time consuming due to the number of classification codes and their level of hierarchy. In the literature, several models have been used to automate this process. We organize them based on the nature of the method into three major categories. Table 2 represents a summary of the methods for patent classification.

Table 2: Summary of the papers for the patent classification task. Hierarchy levels for classification include Section, Class (white), Subclass (blue), Group, and Subgroup (grey). An example label “A61B 5/02" represents Section A, Class 61, Subclass B, Group 5, and Subgroup 02. The color green represents category of visualizations. The WIPO-alpha is a dataset for automated patent classification systems, and ALTA2018 is a dataset from Language Technology Programming Competition.

Papers	Embeddings	Methods	Components	Data
[17]	Word2Vec	Single layer LSTM	Description	USPTO
[52]	Fixed Hierarchy Vectors	LSTM	ADC	-
[48]	FastText	GRU	Full text	WIPO-alpha, USPTO
[4]	TF-IDF	SVM	Single Text Block	ALTA2018, WIPO
[49]	FastText	GRU	Full text	WIPO-alpha, USPTO
[38]	–	BERT-base	Claims	USPTO
[1]	–	BERT, SciBERT	Claims	USPTO
[55]	Word2Vec	Multiple LSTMs	Description	EPO, WIPO
[50]	Word2Vec, FastText	BERT, XLNet, RoBERTa	Title, abstract	USPTO, CLEF-IP 2011
[28]	FastText, Glove, Word2Vec	CNN, LSTM, GRU	TADC	CLEF-IP 2011
[14]	Vision Transformer	MLP	Image	CLEF-IP 2011, USPTO
[27]	FastText	Bi-LSTM, Bi-GRU, LSTM	Metadata	CLEFIP-0.54M

3.1.1 Traditional Neural Networks

The commonality among these methods is that they follow a two-step approach: generate initial features and then use a classifier for the final classification. One of the initial studies, [17] implements a single-layer LSTM to classify patents at the IPC subgroup level where the initial features are obtained by the Word2Vec method. Similarly, [52] use LSTM for IPC subclass level classification. For the initial document representation, the method uses fixed hierarchy vectors that utilize distinct models for various segments of the document. [48] and [49] focus on training fastText word embeddings on a corpus of 5 million patent documents and, then use Bi-GRU for classification. Similarly, [55] apply text mining techniques to extract key sections from patents, train Word2Vec, and then use multiple parallel LSTMs for the classification task. These collectively show the usefulness of neural networks in patent analysis.

3.1.2 Ensemble Models

The models in this category use ensembling different word embeddings and deep learning models. [4] use SVM as a baseline method and experiment with various datasets, the number of features, and semi-supervised learning approaches. Meanwhile, [27] and [28] both investigate ensemble models incorporating Bi-LSTM, Bi-GRU, LSTM, and GRU. More specifically, [28] experiment with different word embedding techniques, whereas [27] focus on applying various partitioning techniques to enhance the performance of the proposed framework. While the above methods heavily focus on texts, [14] classify patent images into distinct types of visualizations, such as graphs, block circuits, flowcharts, and technical drawings, along with various perspectives including side, top, left, and perspective views. The approach utilizes the CLIP model with Multi-layer Perceptron (MLP) and various CNN models.

3.1.3 Large Language Models (LLMs)

The first study [38] which involves LLMs, fine-tune the BERT model on the USPTO-2M dataset and introduces a new dataset, USPTO-3M at the subclass level to aid in future research. Concurrently, [50] also fine-tune BERT, along with XLNet [58], and RoBERTa on the USPTO-2M dataset. They establish XLNet as the new state-of-the-art in classification performance, achieving the highest precision, recall, and F1 measure. [1] implement domain adaptive pre-trained Linguistically Informed Masking and show that SciBERT based representations perform better than BERT-based representations in patent classification. SciBERT is pre-trained on scientific literature which helps the method to understand the technical language of patents.

Discussion. The evaluation measures for patent classification are accuracy, precision, recall, and the F1 score on the CPC or IPC. The earlier works on patent classification are mostly focused on simpler neural networks. Applying models such as LSTM can capture the sequence and context in the text which are suitable for patent domain since the context is critical. However, these are comparatively simple models that might be limited to capture complex technical structure in patent documentation. More advanced techniques have become popular over time, including the adoption of LLMs. LLMs could be powerful because of their pretraining step on a massive amount of data. Patent text is different from the usual text in scientific articles (e.g., research papers). Thus, fine-tuning LLMs on patent datasets might be able to address some of these concerns by providing context-aware representations for the patent domain.

Table 3: Works on Patent Retrieval: The papers have white, blue, and gray based on the data type of text, image, and both respectively. Freepatent and Findpatent are patent data websites, where Findpatent includes patents registered in Russia. WIPS is a patent information search system.

Work	Method	Data	Training
[31]	CNN	Freepatent, Findpatent	supervised
[29]	BERT	WIPS	pre-trained
[5]	BiLSTM-CRF, BiGRU-HAN	USPTO	supervised
[25]	DUAL-VGG	CLEF-IP, USPTO	supervised
[51]	SVM, Naive Bayes, Random Forest, MLP	IPO	supervised
[44]	RoBERTa, CLIP	EPO	pre-trained
[53]	Sentence-BERT, TransE	USPTO	pre-trained, unsupervised
[33]	(ImageNet, Sketchy) ResNet50	DeepPatent	supervised, finetuned
[20]	Deep Metric Learning	DeepPatent	self-supervised
[19]	InfoNCE and ArcFace	DeepPatent	self-supervised

3.2 Retrieval

We divide the relevant studies into three parts based on the deployed model for the retrieval task. Table 3 provides a concise overview of studies for patent retrieval.

Table 4: Summary of the methods on patent quality analysis: “Many" includes Linear regression, Ridge regression, Random Forest, XGBoost, CNN, and LSTM. “APR" stands for the measures accuracy, precision, and recall. IncoPat is a global patent database. We denote Attribute Network Embedding, Attention-based Convolutional Neural Network, European Telecommunications Standards Institute, Derwent Innovation by ANE, ACNN, ETSI, and DI respectively.

Papers	Indicators	Methods	Evaluation Metrics	Datasets
[40]	Citations, meta features	ANE, ACNN	RMSE	USPTO, OECD
[57]	PCA	DNN	Accuracy	ETSI and DI
[22]	Investor reaction, citations	Many	MAE	Patentsview
[9]	Abstract, claims, predefined	CNN, Bi-LSTM	Precision, recall	USPTO
[2]	12 patent indices	ANN	APR, F1, FNR, MAE	USPTO, OECD
[12]	9 patent indices	MLP	Accuracy, Kappa, MAE	USPTO
[39]	Maintenance period	BiLSTM-ATT-CRF	APR, F1	IncoPat
[30]	Patent text	MSABERT	MSE	USPTO, OECD

3.2.1 Traditional Machine Learning

Initial studies have used traditional machine learning methods for patent retrieval. [51] describe five technical requirements to investigate the feasibility of AI for the task. These requirements include query expansion and identification of semantically similar documents. The study uses machine learning algorithms such as SVMs, Naive Bayesian learning, decision tree induction, RF along with word embeddings to solve the prior art retrieval problem. Prior art usually implies the references which may be used to determine the novelty of a patent application. Patent data is searched through multiple resources and returns results based on the query and the database and these results need to be merged to create the final result. [56] employ techniques such as random forest, Support Vector Regression, and Decision Trees to effectively merge the search findings.

3.2.2 Traditional Neural Networks

The methods based on neural networks have been popular in recent years for patent retrieval. [31], [25], and [33] implement CNN, DUAL-VGG, and ResNet, respectively, to retrieve patent images based on a query image. [5] aim to solve entity identification and semantic relation extraction by BiLSTM-CRF [23] and BiGRU-HAN [18] respectively.

3.2.3 Large Language Models (LLMs)

LLMs are useful in many text-related tasks and patent retrieval is not an exception. [44] utilize CLIP for image embedding alongside RoBERTa for capturing textual features, and thus, enhances the search process by incorporating both visual and textual data. [29] use the BERT language model (without fine-tuning) which includes the combinations of title, abstract, and claim. [53] incorporate [46] for text embeddings as well as use the TransE method for the citation and inventor knowledge graph embeddings. They identify that the mean cosine similarity among the vector representations of the patents is effective in linking multiple existing patents to a target patent.

Among other techniques, [20], [19] employ a deep metric learning framework with cross-entropy methods such as InfoNCE [43] and ArcFace [10]. Multimodal techniques have also been used in the information retrieval [44]. Here, the visual features are extracted using vision transformers, while textual features are from sentence transformers.

Discussion. Patent retrieval process involves several subtasks such as defining technical requirements and merging search outcomes from various databases. The proposed methods often use traditional machine learning techniques like SVM, Naive Bayes, Decision trees, etc. While the image retrieval methods apply a variety of CNNs to effectively handle and analyze the visual data, the text retrieval methods have shifted towards LLMs such as BERT for advanced linguistic analysis. Clearly, traditional machine learning techniques are limited to capture the complexity of both patent image and text. While CNNs are popular for image retrieval tasks, the question remains in their effectiveness of patent image retrieval as patent images are non-traditional and technical. On the other hand, utilizing Vision Transformer alongside RoBERTa, Sentence-BERT, TransE shows a multimodal approach might be more suitable for handling the multimodal (e.g., text, images) aspect of patents.

3.3 Quality Analysis

We divide the studies into two categories based on the methods for quality analysis. A summary of research on patent quality analysis is given in Table 4.

3.3.1 Traditional Neural Network

[12] apply an MLP-based approach for quality analysis, utilizing nine indices such as claim counts, forward citations, backward citations, the patent family size to measure the value of a patent. [39] classifies patents based on their maintenance period in four categories. This study implements a Bi-LSTM along with the attention mechanism and Conditional Random Field (CRF) to classify the quality of a patent during the initial stages of its life cycle. The use of Deep Neural Networks is seen in [57] where 11 indicators were considered. [22] predict forward citation and investor reaction to patent announcements implementing CNN-LSTM neural networks and various ML models. [9], [40] and [2] apply a variety of neural networks such as CNN, Bi-LSTM, Attention-based CNN (ACNN), deep and wide Artificial Neural Networks (ANN), respectively.

3.3.2 Large Language Models (LLMs)

[30] propose a variation of BERT (MSABERT) to assess patent value based entirely on the textual data and use the OECD [13] quality indicators for evaluation. Building upon BERT, MSABERT is capable of processing the multi-section structure and longer texts of patent documents. The OECD index includes composite indicators and generality with other predominant indices.

Discussion. While numerous measures are commonly used in assessing the quality of a patent, the absence of universally accepted “gold standard" poses a challenge. Several indices have been considered for patent valuation, among which forward citations are directly associated with the value—both monetary and quality—of a patent. While applying different deep learning models has some success, the question of building a method to handle technical information, metadata, images together inside a patent document remains open. Though MSABERT on entire dataset will be computationally costly, building upon it might be useful for quality evaluation.

3.4 Generation

The generative models are becoming increasingly popular in many domains. The recent developments in AI also have led to the novel research area of generating patents and thus reducing significant human effort that is otherwise needed to write a long document. However, only a couple of studies have attempted to address this problem. [37] implement GPT-2 [45] models to generate the independent claims in patents. The researchers fine-tune the model on 555,890 patent claims of the granted utility patents in 2013 from USPTO. Providing a few words, the method generates the first independent claim of the patent. However, the study is limited towards providing quantitative metrics to evaluate the effectiveness or quality of the patent claims generated by the model. In a separate study, [35] utilizes an alternative methodology, focusing on personalized claim generation by fine-tuning a pre-trained GPT-2 model with inventor-centric data. They also provide a few words or context as input to generate claims. The measure of personalization in the generated claims has been assessed using a BERT model. The underlying hypothesis is that the generated patent claims would demonstrate greater relevance to the respective inventor.

Discussion. First, the problem of patent generation has not been addressed fully. While GPT-2 and BERT have been applied for this task, patents require precise terminology and often contain complex, interrelated concepts that extend beyond the token context window utilized by GPT-2 and BERT. Also, GPT-2 sometimes produces ambiguous or vague text which is not suitable for patents. We believe that more advanced models (e.g., GPT-3.5) along with a large dataset for fine-tuning could help to improve the quality of generated patents.

3.5 Others

In addition to the above-mentioned tasks, there are other interesting studies in the patent domain. Recent work focuses on patent infringement, such as [6] develop a model with different deep learning methods, such as CNN and LSTM, to predict the possibility of a patent application being granted and classify the reason for a failed application. Another work [8] applied a transformer and a Graph Neural Network (GNN) on patent classification for patent landscaping. [59] present an unsupervised method to identify the correlations between patent classification codes and search keywords using PCA and k-means. These studies provide advanced deep learning methods to avoid the risks in patent application. Moreover, there are various studies on generating new ideas and evaluating novelty, such as identifying the inventive process of novel patent using BERT [15], and an explainable AI (XAI) model for novelty analysis via [24]. [60] propose a new task to predict the trends of patents for the companies, and also provide a solution for the task by training an event-based GNN. These studies bring new insights and directions for patent ideas and developments.

4 Future Directions

Many researchers have employed AI to address a wide range of patent-related challenges. In spite of that, this field presents ample opportunities for further research. We strongly believe that a foundation Large Language/Vision-Language Model for patent data will provide a better comprehensive understanding and further improvement in performance on different tasks. Furthermore, it is beneficial to utilize AI for generating novel patent ideas and solutions as well as evaluating patent applications and their quality. We introduce the potential future directions in detail as follows.

4.1 Creating Multimodal Patent Datasets

The availability of multiple modalities (e.g., text, images) in patent documents offers a comprehensive understanding of the related patent tasks. For the various tasks in the patent domain that are mentioned earlier, there is a lack of comprehensive datasets covering multiple modalities. The field could be further advanced by compiling larger datasets that include both textual and visual information found in patents. One of the challenges is that the patent images are often more complex and use advanced domain related concepts compared to the natural (or RGB) images. Thus, it may be necessary to get the opinion of corresponding domain experts to verify and curate the patent datasets accordingly.

4.2 Multimodal Patent Data and Learning

Most of the AI-based patent analysis methods have used either text or image data but not both of them. Recent advances in multimodal learning would allow for more reliable and accurate AI-driven patent analysis. For example, consider the simplest task of patent classification. Here, multiple studies have incorporated patent text and metadata to some extent [38, 17]. However, these studies have largely overlooked the potential of patent images. Intuitively, drawings or sketches provide geometrical information about individual patents. So for classification task, images may be grounded with domain knowledge that is available in the form of text descriptions. In general, multimodal learning can be used to *align representations* derived from text descriptions with those derived from technical images. By doing so, we can guarantee that predictions (classification on unseen or test data) after training are more reliable for related tasks.

4.3 Generative AI for Patents

The domain of patent generation remains relatively unexplored despite its importance. Generating novel and innovative patent claims, along with abstracts and other important sections of the patent document using different inputs, is an exciting potential research field. However, the challenge is in the assessment of the text generated by the generative models such as LLMs. Additionally, machine-centric and token-based methods are used to evaluate generative language models [36]. As patents include jargons and many domain specific words, evaluating generated patent text in terms of only natural language will be limited. Thus, the important question becomes—*how to construct domain-specific evaluation measures for the synthetic or the generated text from LLMs?*

4.4 AI for Patent Assessment

In patent retrieval, one of the important tasks is to generate search queries. This often needs alternate search terms, related words, synonyms which require domain knowledge. The quality and structure of queries directly impact the relevance of the search results. The current AI methods are yet to automate this entire process. Thus, it brings challenges to obtain adequate retrieval accuracy and correctly assess patent's innovativeness and novelty. On the other hand, the generic quality analysis are based on some well-known measures [2, 12]. Nonetheless, it remains unclear which of these indices are associated with the actual value of the patent (e.g., generated revenue).

4.5 Building a Knowledge Graph

Patents are represented as nodes connected by edges such as citations in a citation network. This structured representation allows for detailed citation analysis which is considered a crucial metric in understanding a patent's value. One interesting future direction would be to build a knowledge graph using other important information such as meta-data, semantic similarity of patents, etc. This may lead to a more organized landscape of patents. This knowledge graph can help with prior art searches, the identification of related patents, and identify valuable patents (e.g., patents with high citations).

5 Conclusions

We have provided a comprehensive overview of various patent analysis tasks with AI models in this survey. We have presented a novel schema with a detailed organization of the research papers, analyzing the methodologies used, their advantages, limitations, and how they apply to patents. The process from filing a patent application to being it granted often requires a considerable amount of time. Given the crucial role of patents in economic development and the lengthy process from application to grant, the need for AI-driven approaches in this field is growing to automate different parts of this process. Although numerous studies have addressed this issue, the modern AI techniques have opened many directions for improvement. We have offered several insights into such potential future directions. This survey aims to be a useful guide for researchers, practitioners, and patent offices in the intersection of AI and patent systems.

References

[1] S. Althammer, M. Buckley, S. Hofstätter, and A. Hanbury. Linguistically informed masking for representation learning in the patent domain. *arXiv preprint arXiv:2106.05768*, 2021.

[2] L. Aristodemou. *Identifying valuable patents: A deep learning approach*. PhD thesis, 2021.

[3] I. Beltagy, K. Lo, and A. Cohan. Scibert: A pretrained language model for scientific text. In *EMNLP*, 2019.

[4] F. Benites, S. Malmasi, and M. Zampieri. Classifying patent applications with ensemble methods. *ALTA Workshop*, 2018.

[5] L. Chen, S. Xu, L. Zhu, J. Zhang, X. Lei, and G. Yang. A deep learning based method for extracting semantic information from patent documents. *Scientometrics*, 125:289–312, 2020.

[6] Y.-C. Chi and H.-C. Wang. Establish a patent risk prediction model for emerging technologies using deep learning and data augmentation. *Advanced Engineering Informatics*, 52:101509, 2022.

[7] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio. Learning phrase representations using rnn encoder-decoder for statistical machine translation. *ACL*, 2014.

[8] S. Choi, H. Lee, E. Park, and S. Choi. Deep learning for patent landscaping using transformer and graph embedding. *Technological Forecasting and Social Change*, 175, 2022.

[9] P. Chung and S. Y. Sohn. Early detection of valuable patents using a deep learning model: Case of semiconductor industry. *Technological Forecasting and Social Change*, 158:120146, 2020.

[10] J. Deng, J. Guo, N. Xue, and S. Zafeiriou. Arcface: Additive angular margin loss for deep face recognition. In *CVPR*, 2019.

[11] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *NAACL-HLT*, 2019.

[12] Z. Erdogan, S. Altuntas, and T. Dereli. Predicting patent quality based on machine learning approach. *IEEE Trans Eng Manag*, 2022.

[13] Eurostat, O. *Oslo Manual: Guidelines for Collecting and Interpreting Innovation Data*. OECD, Paris, 2005. A joint publication of OECD and Eurostat.

[14] J. A. Ghauri, E. Müller-Budack, and R. Ewerth. Classification of visualization types and perspectives in patents. In *TPDL*, 2023.

[15] V. Giordano, G. Puccetti, F. Chiarello, T. Pavanello, and G. Fantoni. Unveiling the inventive process from patents by extracting problems, solutions and advantages with natural language processing. *Expert Systems with Applications*, 229:120499, 2023.

[16] A. Graves and J. Schmidhuber. Framewise phoneme classification with bidirectional lstm and other neural network architectures. *Neural networks*, 18(5-6):602–610, 2005.

[17] M. F. Grawe, C. A. Martins, and A. G. Bonfante. Automated patent classification using word embedding. In *ICMLA*. IEEE, 2017.

[18] X. Han, T. Gao, Y. Yao, D. Ye, Z. Liu, and M. Sun. Opennre: An open and extensible toolkit for neural relation extraction. In *EMNLP*, 2019.

[19] K. Higuchi, Y. Honbu, and K. Yanai. Patent image retrieval using cross-entropy-based metric learning. In *IW-FCV*, 2023.

[20] K. Higuchi and K. Yanai. Patent image retrieval using transformer-based deep metric learning. *WPI*, 74:102217, 2023.

[21] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[22] P.-H. Hsu, D. Lee, P. Tambe, and D. H. Hsu. Deep learning, text, and patent valuation. *Text, and Patent Valuation*, 2020.

[23] Z. Huang, W. Xu, and K. Yu. Bidirectional lstm-crf models for sequence tagging. *arXiv preprint arXiv:1508.01991*, 2015.

[24] H. Jang, S. Kim, and B. Yoon. An explainable ai (xai) model for text-based patent novelty analysis. *Expert Systems with Applications*, 2023.

[25] S. Jiang, J. Luo, G. Ruiz-Pava, J. Hu, and C. L. Magee. Deriving design feature vectors for patent images using convolutional neural networks. *Journal of Mechanical Design*, 143(6):061405, 2021.

[26] A. Joulin, E. Grave, P. Bojanowski, and T. Mikolov. Bag of tricks for efficient text classification. In *EACL*, 2017.

[27] E. Kamateri, M. Salampasis, and K. Diamantaras. An ensemble framework for patent classification. *WPI*, 75:102233, 2023.

[28] E. Kamateri, V. Stamatis, K. Diamantaras, and M. Salampasis. Automated single-label patent classification using ensemble classifiers. In *ICMLC*, 2022.

[29] D. M. Kang, C. C. Lee, S. Lee, and W. Lee. Patent prior art search using deep learning language model. In *IDEAS*, 2020.

[30] X. Krant. *Text-based Patent-Quality Prediction Using Multi-Section Attention*. PhD thesis, 2023.

[31] A. Kravets, N. Lebedev, and M. Legenchenko. Patents images retrieval and convolutional neural network training dataset quality improvement. In *ITSMSSM*, 2017.

[32] R. Krestel, R. Chikkamath, C. Hewel, and J. Risch. A survey on deep learning for patent analysis. *WPI*, 65:102035, 2021.

[33] M. Kucer, D. Oyen, J. Castorena, and J. Wu. Deeppatent: Large scale patent drawing recognition and retrieval. In *WACV*, 2022.

[34] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

[35] J.-S. Lee. Patent transformer: A framework for personalized patent claim generation. In *CEUR Workshop Proceedings*, volume 2598. CEUR-WS, 2020.

[36] J.-S. Lee. Evaluating generative patent language models. *WPI*, 72, 2023.

[37] J.-S. Lee and J. Hsiang. Patent claim generation by fine-tuning openai gpt-2. *WPI*, 62:101983, 2020.

[38] J.-S. Lee and J. Hsiang. Patent classification grawe-tuning bert language model. *WPI*, 61:101965, 2020.

[39] R. Li, H. Zhan, Y. Lin, J. Yu, and R. Wang. A deep learning-based early patent quality recognition model. In *ICNC-FSKD*, 2022.

[40] H. Lin, H. Wang, D. Du, H. Wu, B. Chang, and E. Chen. Patent quality valuation with deep learning models. In *DASFAA*, 2018.

[41] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*, 2019.

[42] T. Mikolov, K. Chen, G. Corrado, and J. Dean. Efficient estimation of word representations in vector space. In *ICLR*, 2013.

[43] A. v. d. Oord, Y. Li, and O. Vinyals. Representation learning with contrastive predictive coding. *arXiv preprint arXiv:1807.03748*, 2018.

[44] K. Pustu-Iren, G. Bruns, and R. Ewerth. A multimodal approach for semantic patent image retrieval. In *PatentSemTech*, 2021.

[45] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, I. Sutskever, et al. Language models are unsupervised multitask learners. *OpenAI blog*, 1:9, 2019.

[46] N. Reimers and I. Gurevych. Sentence-bert: Sentence embeddings using siamese bert-networks. *EMNLP*, 2019.

[47] J. Risch, N. Alder, C. Hewel, and R. Krestel. Patentmatch: A dataset for matching patent claims & prior art. In *PatentSemTech@SIGIR*, 2021.

[48] J. Risch and R. Krestel. Learning patent speak: Investigating domain-specific word embeddings. In *ICDIM*, 2018.

[49] J. Risch and R. Krestel. Domain-specific word embeddings for patent classification. *Data Technologies and Applications*, 53(1), 2019.

[50] A. Roudsari, J. Afshar, W. Lee, and S. Lee. Patentnet: multi-label classification of patent documents using deep learning based language understanding. *Scientometrics*, pages 1–25, 2022.

[51] R. Setchi, I. Spasić, J. Morgan, C. Harrison, and R. Corken. Artificial intelligence for patent prior art searching. *WPI*, 2021.

[52] M. Shalaby, J. Stutzki, M. Schubert, and S. Günnemann. An lstm approach to patent classification based on fixed hierarchy vectors. In *SDM*, 2018.

[53] L. Siddharth, G. Li, and J. Luo. Enhancing patent retrieval using text and knowledge graph embeddings: a technical note. *Journal of Engineering Design*, 33:670–683, 2022.

[54] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. In *ICLR*, 2015.

[55] M. Sofean. Deep learning based pipeline with multichannel inputs for patent classification. *WPI*, 66:102060, 2021.

[56] V. Stamatis, M. Salampasis, and K. Diamantaras. Machine learning methods for results merging in patent retrieval. *Data Technologies and Applications*, 2023.

[57] A. J. Trappey, C. V. Trappey, U. H. Govindarajan, and J. J. Sun. Patent value analysis using deep learning models—the case of iot technology mining for the manufacturing industry. *IEEE-TEM*, 68(5):1334–1346, 2019.

[58] Z. Yang, Z. Dai, Y. Yang, J. Carbonell, R. R. Salakhutdinov, and Q. V. Le. Xlnet: Generalized autoregressive pretraining for language understanding. *NeurIPS*, 2019.

[59] W. M. F. Zaini, D. T. C. Lai, and R. C. Lim. Identifying patent classification codes associated with specific search keywords using machine learning. *WPI*, 71:102153, 2022.

[60] T. Zou, L. Yu, L. Sun, B. Du, D. Wang, and F. Zhuang. Event-based dynamic graph representation learning for patent application trend prediction. *IEEE TKDE*, 2023.