**RESEARCH ARTICLE**

Tianjie FU, Shimin LIU, Peiyu LI

# Intelligent smelting process, management system: Efficient and intelligent management strategy by incorporating large language model

**Abstract** In the steelmaking industry, enhancing production cost-effectiveness and operational efficiency requires the integration of intelligent systems to support production activities. Thus, effectively integrating various production modules is crucial to enable collaborative operations throughout the entire production chain, reducing management costs and complexities. This paper proposes, for the first time, the integration of Vision-Language Model (VLM) and Large Language Model (LLM) technologies in the steel manufacturing domain, creating a novel steelmaking process management system. The system facilitates data collection, analysis, visualization, and intelligent dialogue for the steelmaking process. The VLM module provides textual descriptions for slab defect detection, while LLM technology supports the analysis of production data and intelligent question-answering. The feasibility, superiority, and effectiveness of the system are demonstrated through production data and comparative experiments. The system has significantly lowered costs and enhanced operational understanding, marking a critical step toward intelligent and cost-effective management in the steelmaking domain.

**Keywords** smelting steel, process management, large language models, intelligent Q & A, ChatGPT

Tianjie FU, Peiyu LI
The School of Mechanical Engineering, Zhejiang University, Hangzhou 310058, China

Shimin LIU (✉)
Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hong Kong, China; State Key Laboratory of Ultra-precision Machining Technology, Department of Industrial Systems and Engineering, The Hong Kong Polytechnic University, Hong Kong, China
E-mail: shimin.liu@polyu.edu.hk

## 1 Introduction

Large Language Models (LLMs) have experienced a remarkable surge in usage in the engineering sector in recent years, driven by advancements in Natural Language Processing (NLP) technology. This surge signifies a fundamental shift for LLMs, transitioning them from theoretical models to practical tools with substantial influence on various engineering disciplines (Xiao et al., 2023). This trend is driven by the urgent need for environmentally-friendly development and sustainable practices, as well as the requirement for a significant transformation in engineering processes. As cutting-edge research converges with technological maturity, pushing these models into the mainstream, their scope and applicability continue to expand, offering transformative solutions for longstanding challenges (Hein-Pensel et al., 2023). The engineering community is recognizing that incorporating sustainable development principles into engineering practices can be facilitated by leveraging advanced technology and fostering interdisciplinary collaboration. The increasing use of LLMs is ushering in a new era of interdisciplinary collaboration, underscoring the multifaceted nature of subjects such as computer science, engineering, and linguistics. These models are being employed not only for their inherent language capabilities, but also due to the availability of large datasets and sophisticated algorithmic advancements. Consequently, the engineering community finds itself at a pivotal juncture, where computational accuracy and nuanced comprehension of natural language are merging. The transformative potential of LLMs is becoming evident in their ability to revolutionize information retrieval systems and enhance human-machine interaction. Thus, we observe the significant role that sustainable and green development play, with LLMs supporting the objective of sustainable development while also opening up new avenues for engineering practice.

Due to varying levels of technological maturity across

industries, there are significant differences in the integration of intelligent systems in industrial production today. Although some industries have made significant progress in combining automation and intelligence, a considerable portion of the industrial ecosystem still faces challenges in achieving seamless intelligent production. In the field of industry, the lack of coordination among different systems not only leads to wasteful resource utilization but also poses a significant obstacle to realizing the primary goals of Industry 4.0 standards. One major contributing factor to this technological lag is the increasing costs associated with human resource management and communication in industrial environments. The effective management of complex automated processes and efficient operator communication are becoming increasingly crucial as companies adopt intelligent technologies. However, due to the growing time and resource costs of communication, integrating these factors has become difficult. Overcoming the challenge of bridging the gap between automated intelligence and manual operating systems is not only financially significant but also immensely challenging. The industrial production environment incurs substantial management expenses, further complicating the implementation of intelligent automation. A robust management infrastructure is required to oversee both automated and manual components, leading to substantial expenditures for maintenance, monitoring, and training. Consequently, the persistent intelligence gap in industrial output is greatly intensified by the lack of efficient and cost-effective management solutions.

The development of intelligent and efficient manufacturing processes is hindered by numerous obstacles encountered in modern industrial production (Shi, 2015; Shi et al., 2017; Zheng et al. 2024). Among these challenges, the most significant is the escalating management expenses associated with industrial output. This upward trend takes into account the expansive nature of the management domain and the tools utilized within, highlighting the complex interactions between manual operating systems and emerging automated intelligence domains. Additionally, a prevalent issue in industrial production is the absence of a cohesive management structure, resulting in the fragmentation of managerial responsibilities. The dispersion of administrative authority within industrial sectors exacerbates inefficiencies and hinders the seamless integration of intelligent technology. The absence of a unified management framework decreases the possibility of achieving maximum production by causing duplications and difficulties in coordinating various activities. One noteworthy aspect of the industrial production paradigm is the underutilization of intelligent automation. A weakness in scientific management methods is the widespread reliance on expert operators in the face of system abnormalities or deviations. The scalability of intelligent systems is impeded due to the absence of standard operating procedures for dealing with unforeseen

events, forcing the industry to rely more on specialized knowledge rather than scalable and standardized processes.

In summary, industrial production is faced with a range of challenges, including increased management costs, decentralized management systems, and inadequate levels of automation. LLMs, with their exceptional features and capabilities, offer innovative solutions to tackle these challenges.

LLMs possess robust semantic understanding and information extraction capabilities, enabling the automatic processing and analysis of vast amounts of managerial content. This aids in reducing management costs by providing more efficient and precise decision support within managerial means. Additionally, LLMs, through the establishment of a unified knowledge graph, can integrate and manage dispersed production information, thereby enhancing coordination and efficiency while mitigating issues associated with managerial decentralization.

Moreover, the intelligent characteristics of LLMs enable the automation of numerous managerial tasks in industrial production. In exceptional circumstances, LLMs can provide scientifically grounded management process recommendations by learning from experience and data, reducing the dependence on experienced workers. This presents a new avenue for the intelligent upgrade of industrial production, achieving more sustainable and efficient operations.

Therefore, the characteristics of LLMs in industrial production management, including semantic understanding, information integration, and intelligent processing, position them as potent tools for addressing challenges such as increased management costs, decentralized management systems, and inadequate levels of automation. Their application elevates the intelligence level of industrial production, offering novel possibilities for efficiency enhancement and cost reduction.

Confronted with the aforementioned complex challenges in industrial production, two pressing issues have emerged to drive the optimization and transformation of the industry:

1) The utilization of intelligent systems to assist production, leading to cost reduction and improved efficiency. The challenge lies in developing solutions that synergize with both human and automated systems, maximizing resources, increasing production efficiency, and creating a sustainable manufacturing process (Liu and Xie, 2024).

2) To reduce management costs and complexity, it is crucial to effectively integrate different production modules. This requires collaboration between experts in engineering, computer science, and management to create a comprehensive, integrated management system. Solving this issue will result in a unified management framework for the industry, promoting better teamwork and

propelling industrial production towards a more efficient, flexible, and controllable direction.

3) The utilization of technology based on LLMs can effectively reduce comprehension costs. Advanced algorithms for natural language processing can handle complex production information, enabling intelligent systems to accurately comprehend human-machine interactions, production reports, and anomaly handling. As a result, optimizing the use of LLM technology has become increasingly crucial for efficient management and monitoring of production environments.

Given the rising costs of industrial production management, decentralized management structures, and insufficient intelligent automation, a comprehensive analysis indicates the necessity for a paradigm shift. To tackle these challenges, it is crucial to establish an industrial production management system based on LLMs. This transformation aims to create a more integrated, cost-effective, and intelligent industrial production landscape. The management system will drive the industry to overcome technological limitations and foster innovation and upgrades in production modes, ultimately leading to the development of industrial intelligence.

The subsequent sections are structured as follows: Chapter 2, "Related works," presents the current state of LLMs and industrial management systems. Chapter 3, "LLM-based production management system," primarily outlines the system framework of industrial production processes and its relevant modules. Chapter 4, "Methodology," provides a detailed explanation of visual language modeling and the fine-tuning of LLMs. Chapter 5, "System implementation and validation," conducts experimental verification of the aforementioned content. Finally, Chapter 6 offers a comprehensive summary of the entire document.

## 2 Related works

This section conducts an in-depth analysis of the intersection between LLMs and management technologies within the industrial production process, with the aim of highlighting current research gaps in this domain.

### 2.1 LLM

ChatGPT, developed by OpenAI in San Francisco, California, serves as an example of the evolution of artificial intelligence (AI) technology that goes beyond mere transformative change and aligns with the ongoing digital transformation of our society. Generative AI systems have made significant progress since the release of OpenAI's ChatGPT. However, there are risks and limitations associated with the adoption of LLMs by businesses (O'Leary, 2023; Yu & Gong, 2024). As a recent disruptor

in the technological sphere, ChatGPT has attracted attention for its potential to transform various domains. Snoswell et al. (2023) proposed a fundamental shift in how patients and clinical practitioners acquire and receive information by integrating LLMs with medicine. Mallio et al. (2023) highlighted the significant potential of LLMs in generating structured radiology reports and compare four LLM models in terms of structured reporting knowledge and template suggestions. Thiébaut et al. (2023) emphasized the need to express queries to intelligent agents in natural language and provided recommendations for applying LLMs to enhance public health. As a popular LLM in recent years, ChatGPT has gained considerable attention in the field of clinical medicine. Li et al. (2023) explored the positive contributions of ChatGPT in medical device design, optimization, and improvement, while addressing its limited impact on clinical medical equipment manufacturing. Additionally, several scholars advocate for the application of LLMs in the healthcare sector to facilitate collaborative interactions between artificial intelligence and human participation (Borkowski, 2023).

Furthermore, De Curtò et al (2023) combine LLMs and visual language models (VLMs) to generate accurate textual descriptions of scenes captured by unmanned aerial vehicles. Pavlopoulos et al. (2023) proposed an ML-assisted workflow for predicting imminent failures in cars, effectively classifying textual symptom statements for large fleet companies, and streamlining automotive fault management. Demertzis et al. (2023) discussed the benefits and limitations of using computational intelligence in civil engineering, addressing challenges faced by researchers and practitioners in implementing these technologies. The advancements in LLMs (Wei et al., 2022; Gu et al., 2022; Alayrac et al., 2022) and VLMs (Radford et al., 2021) empower AI systems to interact with humans in unprecedented ways. These models demonstrate excellent performance across various tasks, including robot operations (Nair et al., 2022; Zeng et al., 2022a; Cui et al., 2022) and superior performance in navigation and guidance (Zeng et al., 2022b; Huang et al., 2022).

Despite these advancements, there is currently a lack of application and technology integration of LLMs and visual language models in industrial production management systems. Hence, there are industry-specific and technological gaps in the field of industrial production.

### 2.2 Industrial process management

The industrial sector has undergone significant transformation in recent years due to the rapid development of intelligent manufacturing technologies, Industry 4.0, and Digital Twins (DT) (Kouzapas et al., 2023, Bellavista et al., 2023, Peng et al., 2022, Sievers & Blank, 2023, Jadhav et al., 2023). These advancements have led to substantial changes in global manufacturing technology systems and paradigms through the extensive use of large-

scale intelligent devices in the manufacturing process.

Iwańkowicz & Rutkowski (2023) categorized the digitization of ship design and production processes into planning, monitoring, and process analysis activities. Huang et al. (2023) developed a Cyber-Physical Production System (CPPS) model based on 5G networks for multipart machining, enhancing interoperability in the domains of Information and Communication Technology (ICT) and Operational Technology (OT). Jaber et al. (2023) proposed a Complex Industrial Information System based on a Hybrid Machine Learning Model (CIIS-HMLM) to address the issue of sensor loss. Bessarabov et al. (2023) systematically analyzed water resource supply issues and proposed a comprehensive water management system covering all relevant processes and auxiliary operations. There have also been suggestions from scholars to use blockchain and various encryption tools to protect privacy in the Industrial Internet of Things (IoTs) (Bao et al., 2023).

The steel industry, as a critical sector supporting the economy, has also made notable advancements. Fang et al. (2023) developed temperature prediction models for the Basic Oxygen Furnace (BOF), Ladle Furnace (LF), and Ruhrstahl-Heraeus (RH) steelmaking processes using an enhanced BP neural network. Semenov et al. (2022) proposed three Decision Support System (DSS) models to enhance energy efficiency in blast furnace operations. Pilot tests were conducted to validate the effectiveness of the blast furnace workshop's automation control system. Zhu et al. (2023) developed ultra-low emission control technologies to address pollution emissions in the steelmaking industry, providing technical support for the industry's green development.

In summary, intelligent manufacturing technologies have been widely implemented in the industrial sector. However, in the context of the steelmaking industry, existing process management systems only provide partial information on the smelting process, lacking consolidation of the complex production data into a unified system.

### 2.3 Research gaps

In the domain of steel manufacturing, there are two primary challenges that persist: the singularity and fragmentation of management systems, along with a deficiency in automation. These issues contribute to difficulties and inefficiencies in production management. Expanding on the aforementioned analysis, the following problems are identified in the smelting process of the steel manufacturing industry:

1) The management systems in the steel manufacturing industry are relatively singular and fragmented, lacking a comprehensive macro-level management system. Currently, integrating management information across various stages proves challenging, leading to decentralized decision-making processes across different departments and systems. This fragmentation impedes overall managerial coordination, adversely affecting production efficiency.

2) The current level of automation in management systems within the steel manufacturing industry is suboptimal, failing to meet the rapidly evolving demands of production management. The absence of advanced intelligent decision support systems restricts aspects such as production planning, quality control, and equipment maintenance. This not only increases production risks but also hampers the industry's competitiveness.

3) A dedicated large language model designed specifically for steel manufacturing production is currently lacking. This absence results in a lack of targeted intelligent tools, preventing the full utilization of accumulated data in production. The introduction of large language models could offer highly tailored solutions for the steel manufacturing industry, better adapting to specific production environments and requirements.

Large language models' semantic comprehension, information integration, and intelligent processing skills will give the steel production sector a more integrated management system, enabling effective information integration and wise decision-making. Therefore, it is expected that this would improve production efficiency and support the steel manufacturing industry's competitiveness and sustainability. The following sections of this paper will provide further elaboration on these points.

## 3 LLM-based production management system

This section proposes an LLM-based steel manufacturing process management system tailored for the steel smelting process. It aims to provide a foundational model and experimental basis for the establishment of the system.

### 3.1 Steel smelting process

Due to its traditional nature, the steel production industry currently lacks adequate digitalization, leading to inefficiency and insufficient assurance of steel quality during the smelting process (Fu et al., 2024a; 2024b). The smelting process can be divided into four main steps, as illustrated in Fig. 1. These steps include raw material processing, blast furnace smelting, refining, continuous casting, and production of finished steel products.

1) **Raw material processing**: The main raw materials for smelting pig iron, including freshly extracted iron ore, limestone, coal, and other raw materials, must undergo procedures such as crushing, ore sorting, and ore washing.
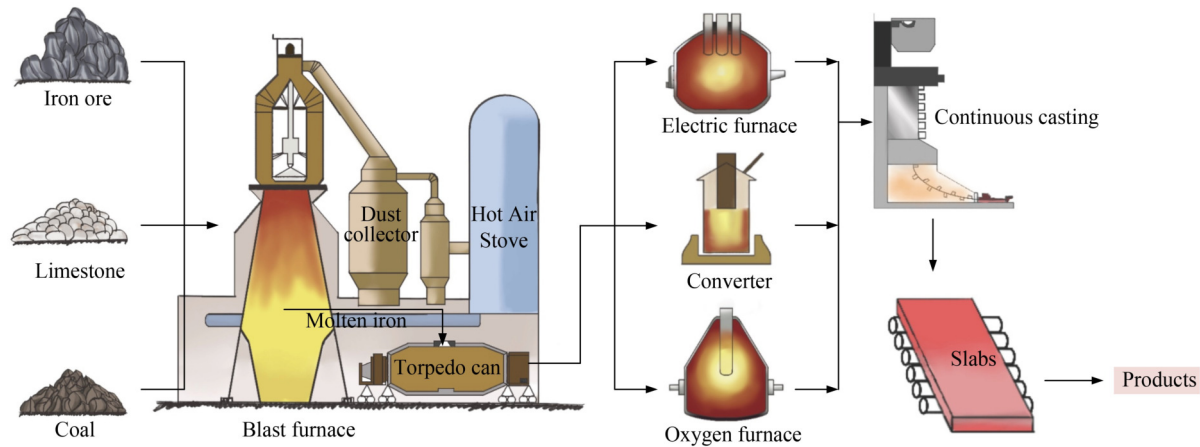
2) **Blast furnace smelting**: The treated raw materials

**Fig. 1**   Steel smelting process.

are melted in a blast furnace to obtain molten iron, which serves as the basic ingredient for steel production.

3) **Refining**: This step involves adding the appropriate alloy components to the molten iron while removing excess carbon, sulfur, phosphorus, and other impurities. The steel industry employs three types of furnaces: the converter, open-hearth furnace, and electric arc furnace. While the electric arc furnace is used to melt and refine scrap steel, the converter and open-hearth furnace blend molten iron from the blast furnace with scrap steel. However, due to their high energy consumption and lengthy production cycles, many steel mills have phased out the use of open-hearth furnaces.

4) **Continuous casting and discharging**: The molten steel is continuously cast to form slabs, which are then further processed to produce the desired steel products.

The steel smelting process involves various production elements, which often result in decentralized production locations for many steel plants. Consequently, managing the steel smelting process becomes complex and fragmented.

Ensuring product quality during the smelting process requires comprehensive control and management of multidimensional data. This paper introduces LLM into the steel smelting process to achieve control and management of the entire smelting data. Only by effectively managing data throughout the entire smelting process can the quality of the smelting process be enhanced while simultaneously reducing the associated management costs.

## 3.2   System architecture

The preceding analysis has provided a comprehensive overview of the crucial stages involved in steel smelting, representing a significant advancement towards the digitalization of steel manufacturing process management. This paper presents the LLM-based steel manufacturing process management system for steel smelting, which serves as the fundamental model and experimental basis for system development.

In Fig. 2, the architecture of the LLM-based steel smelting process management system is illustrated, consisting of four primary processes. Algorithms and data obtained from various sensors and scanners are transmitted to the cloud and stored in a database. Subsequently, the data in the database is processed and converted into easily understandable natural language. This processed data is then linked with the ChatGPT-3.5 API to enable visualization and interaction on a dialogue platform. The following sections will provide a detailed explanation of each component.

1) **Algorithms and sensors:** The detection of slab defects plays a significant role in ensuring the quality of the final steel product in the smelting process. Accurate prediction of key steelmaking temperatures can help reduce material splatter waste and achieve cost savings. In addition, various algorithms are employed in other smelting processes. A wide range of sensors, including infrared scanners, temperature sensors, weight sensors, incline sensors, rotation angle sensors, light sensors, humidity sensors, and other instruments, are utilized to collect relevant data on smelting conditions and equipment performance during the smelting process.

2) **Database and cloud:** The utilization of database and cloud technology in the management system for steel smelting processes greatly enhances managerial levels, data security, and overall production efficiency. The database is employed to store and maintain smelting status data and operating data for smelting equipment, creating a repository of historical records that facilitate accurate communication with the ChatGPT-3.5 API. By providing a highly organized platform for storing various industrial data, such as raw material information, equipment status, and production parameters, the database centralizes data administration and archiving, ensuring data correctness and consistency. Through the use of indexing and optimization techniques, the database
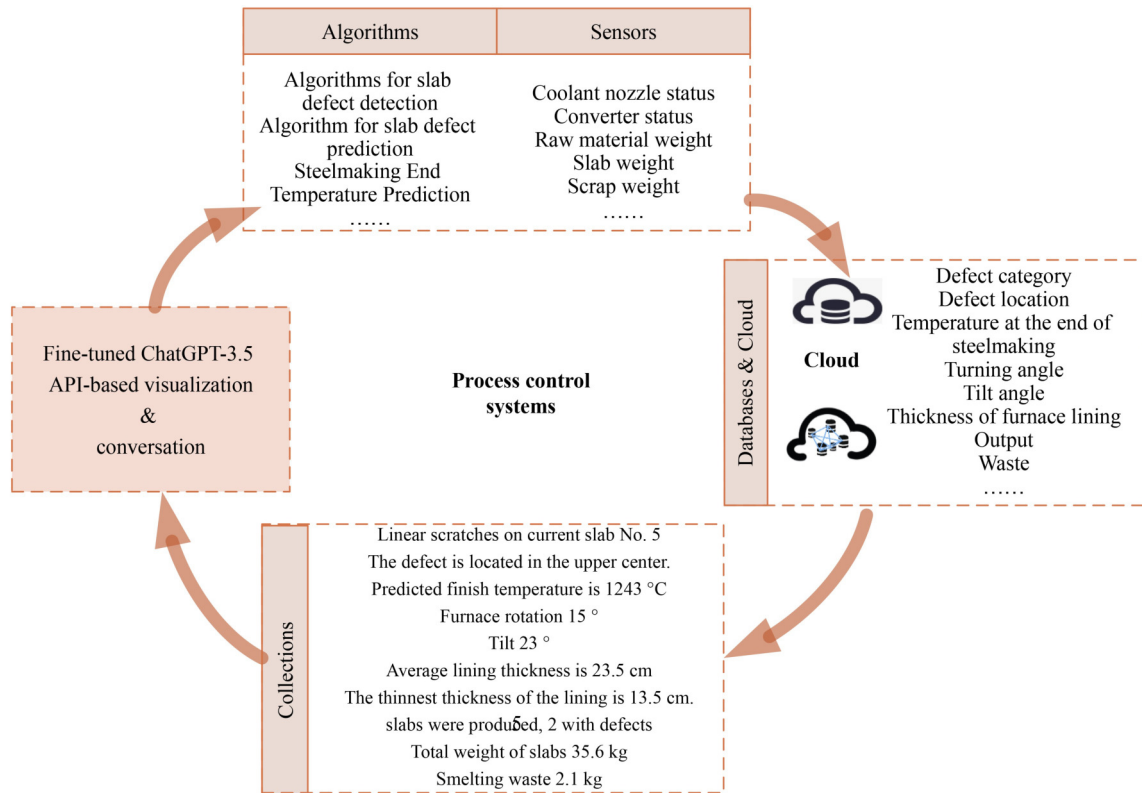
| Algorithms | Sensors |
|---|---|
| Algorithms for slab defect detection<br>Algorithm for slab defect prediction<br>Steelmaking End Temperature Prediction<br>…… | Coolant nozzle status<br>Converter status<br>Raw material weight<br>Slab weight<br>Scrap weight<br>…… |

Fine-tuned ChatGPT-3.5
API-based visualization
&
conversation

**Process control systems**

**Databases & Cloud**

**Cloud**

Defect category
Defect location
Temperature at the end of steelmaking
Turning angle
Tilt angle
Thickness of furnace lining
Output
Waste
……

**Collections**

Linear scratches on current slab No. 5
The defect is located in the upper center.
Predicted finish temperature is 1243 °C
Furnace rotation 15 °
Tilt 23 °
Average lining thickness is 23.5 cm
The thinnest thickness of the lining is 13.5 cm.
5 slabs were produced, 2 with defects
Total weight of slabs 35.6 kg
Smelting waste 2.1 kg

**Fig. 2**   Management system architecture.

system enables rapid retrieval of large amounts of data, providing operators with real-time monitoring information and decision assistance during the smelting production process. Cloud computing platforms allow for dynamic adjustment of processing and storage resources in response to real-time requirements, granting the system increased flexibility and scalability. This empowers steel smelting companies to adapt to fluctuations in demand and output, while simultaneously reducing hardware expenses through the consolidation of processing and storage capabilities on the cloud platform.

3) **Collections:** The system effectively supports data visualization and conversation platforms by representing database data in a natural language pattern that is easily understood by the API.

4) **Visualization and dialogue:** Users can seamlessly monitor and control the steel smelting process using simple language through various client interfaces, including PC, phone, pad, web, and VR. This simplifies management and comprehension, resulting in cost savings.

In conclusion, the LLM-based steel smelting process management system provides customers with comprehensive smelting and management data. Leveraging its exceptional data parsing and comprehension capabilities, the system offers customers an intelligent and efficient data support and management solution. Additionally, it offers management benefits, such as intelligent decision

support and automatic report generation.

### 3.3   Smelting data management

Accurate monitoring and management of smelting data are significant for the steel smelting process. Smelting enterprises must effectively address issues related to equipment status, environmental impact, and quality control to optimize production efficiency and sustainability. A diligent focus on these aspects enables the steel smelting industry to make well-informed decisions when confronted with complex challenges.

A comprehensive analysis of smelting data reveals key factors that are essential for achieving efficient production and maintaining high production quality.

1) **Data yield**: The collection of data is crucial in the steel smelting process as it plays an indispensable role. It has immense value in enhancing manufacturing effectiveness, implementing quality control measures, and directing strategic planning.

2) **Quality control**: The competitiveness of smelting companies in the market is directly affected by the quality of their products. Therefore, close attention must be paid to the production process to ensure that the final products meet consumer expectations and comply with relevant requirements.

3) **Temperature information**: High temperatures are essential throughout the smelting process to ensure

proper alloy formation and metal melting. This information is vital for predicting the endpoint temperature of the steelmaking process. By collecting and monitoring temperature data, the heating system can be adjusted in real-time to ensure that the alloy composition meets the necessary requirements.

4) **Pressure and flow data**: Monitoring the pressure and flow of gases and liquids during the smelting process is crucial in ensuring the correct delivery of fuel and coolant to the equipment, thereby preserving production stability.

5) **Energy consumption data**: Efficient energy usage is one of the key challenges in the smelting process. Keeping track of energy consumption data helps identify areas where production efficiency can be improved and energy waste can be reduced.

6) **Equipment failure and maintenance**: Proper functioning of smelting equipment is essential to maintain output. Preventive maintenance and routine monitoring of equipment status help minimize the risk of unexpected breakdowns.

In summary, the smelting of steel involves the documentation and evaluation of various types of data. This article discusses the use of sensors and algorithms to gather relevant data, which is then stored in a database for analysis and presentation.

### 3.4 LLM model

The LLM model holds great potential for application in industrial management systems. We propose integrating the LLM model into the process management system to address the following aspects and meet the requirements of steel production:

1) **Demand forecasting and planning optimization**: Leveraging historical production data analysis using the LLM model helps predict steel product demand and optimize production planning within the process management system.

2) **Scheduling optimization**: By utilizing the LLM model, scheduling arrangements within the smelting production line can be optimized based on factors such as furnace equipment status and personnel availability. This enhances production efficiency and resource utilization.

3) **Predictive equipment maintenance**: By analyzing equipment sensor data and operational status with the LLM model, potential equipment failures can be predicted. This enables proactive maintenance alerts within the system, minimizing production disruptions.

4) **Equipment fault diagnosis**: Through the analysis of equipment sensor data and expert knowledge, the identification of equipment faults and the inference of causative factors are conducted. This is accompanied by the provision of corresponding fault diagnosis and resolution strategies.

5) **Quality prediction**: The analysis of various parame-

ters and indicators during smelting production processes enables the prediction of quality. This facilitates the necessary adjustments to production processes to ensure adherence to product quality requirements.

6) **Process optimization**: By leveraging data and historical experience from production processes, optimization recommendations are provided to enhance process efficiency and product quality within the process management system.

In summary, the large language model plays a vital role within steel manufacturing process management systems. It facilitates diverse functionalities crucial for operational efficacy and product quality assurance. Based on OpenAI's GPT-3.5 model, the ChatGPT-3.5 API is a potent natural language processing tool that excels in producing excellent natural language writing. The API's contextual understanding, question-answering, and conversation creation capabilities have shown notable advancements over its predecessor, providing more adaptable and effective solutions for a variety of applications (Franco D'Souza, Amanullah, Mathew, Surapaneni, 2023, Massey et al., 2023, Stepanov et al., 2023).

This API is ideal for creating intelligent dialogue systems and producing emotionally charged, logical, and natural conversation. Its invocation adds a sophisticated and humane layer to the management of the steel smelting process by enabling smooth natural language contact with users. By utilizing the ChatGPT-3.5 API for natural language query responding, effective knowledge-based query systems can be built. This system can extract pertinent information from large amounts of text material and respond to user queries in natural language based on the questions asked.

With its robust language model and exceptional natural language comprehension capabilities, the ChatGPT-3.5 API performs well in tasks like question answering and dialogue production. It provides a wide range of customizable parameters and choices that can be adjusted to suit the specific requirements of the steel smelting procedure. As a deep learning-based API, ChatGPT-3.5 continuously gains knowledge from large datasets and adapts to new languages and subjects, gradually improving its performance. Therefore, the ChatGPT-3.5 API provides a strong foundation for the iterative optimization of solutions tailored to the challenges encountered throughout the steel smelting process.

In conclusion, the ChatGPT-3.5 API is a powerful natural language processing tool that offers customizable settings and parameters ideally suited to the steel smelting industry.

## 4 Methodology

This section primarily outlines the methodologies for fine-tuning VLM and LLM within the management

system, aiming to provide detailed insights into the model architecture and experimental foundations for system construction.

## 4.1  Visual language models

Based on the utilization of LLM and Contrastive Language-Image Pre-training (CLIP), we propose a textual description for detecting slab defects in the steel smelting process.

LLMs and VLMs are typically neural network-based, comprising interconnected processing units with the ability to learn and adapt through training. LLMs are machine learning models trained on extensive text datasets, capable of generating text that resembles human language. One of the key characteristics of large language models is their capacity to produce coherent and believable text that is similar to human language and difficult to distinguish. Essentially, LLMs serve as powerful and versatile tools for understanding and processing natural language data.

CLIP, which utilizes a transformer architecture, is a neural network particularly suited for tasks involving sequential data. After training, the model predicts the next word in a sentence based on the context of the preceding word and an image as additional context. CLIP can learn a continuous space for representing both images and text, enabling it to generate high-quality captions for a wide range of images (Mokady et al., 2021). With a dataset containing matched images and corresponding textual descriptions $\{x^i, c^i\}_{i=1}^N$, CLIP synthesizes appropriate and accurate textual descriptions for previously unseen sample images. Extracting visual information from image $x^i$ is accomplished using the pre-trained CLIP model's visual encoder (Radford et al., 2021). CLIP includes basic visual data, treating this condition as a prefix for the textual description (Zhou et al., 2020). Since the required semantic information is encapsulated in the prefix, an autoregressive language model can predict the next token. Eq. (1) represents the network's objective function, where $\theta$ denotes the trainable parameters of the model.

Each vector $p_j^i$ has the same dimension as word embeddings, and the obtained visual embeddings are concatenated to the textual description $c^i$. The textual description can be represented as $c^i = c_1^i, \ldots, c_l^i$, padded to the maximum length $l$. A mapping network $F$ is then employed to map CLIP embeddings to $k$ embedding vectors, and training of the mapping F is conducted using cross-entropy loss.

$$\max_{\theta} \sum_{i=1}^N \sum_{j=1}^l \log p_{\theta}\left(c_j^i \mid x^i, c_1^i, \ldots, c_{j-1}^i\right), \qquad (1)$$

$$p_1^i, \ldots, p_j^i = F\left(CLIP(x^i)\right). \qquad (2)$$

Figure 3 illustrates the proposed structure of the VLM. The key modules within the VLM's architecture comprise the CLIP prefix for generating textual descriptions, the YOLOv5 model for defect detection, and GPT-3.5. This framework relies on the CLIP technology's outputs to generate corresponding textual descriptions and combines sentences from Conceptual Captions with defect detection and classification output from YOLOv5 (Sharma et al., 2018).

The CLIP prefix, designed for generating textual descriptions, adopts a lightweight mapping network architecture based on transformers (Dosovitskiy et al., 2021). YOLOv5, known for its cost-effectiveness as a single-stage object detection model, employs basic data augmentation methods, the Mosaic data augmentation method, and Extended Efficient Layer Aggregation Networks (E-ELANs). You Only Look Once (YOLO) is a real-time object detection algorithm with a pyramid structure, utilized for the recognition and classification of target objects in images and videos. This algorithm predicts the category and location of target objects in images divided into cells. Utilizing anchors at multiple scales facilitates predictions for different-sized target objects, and confidence scores are utilized to filter out false detections (Redmon and Farhadi, 2017). The output of the object detection model provides the location and classification of slab defects in the image, forming a textual description comprehensible to the ChatGPT-3.5
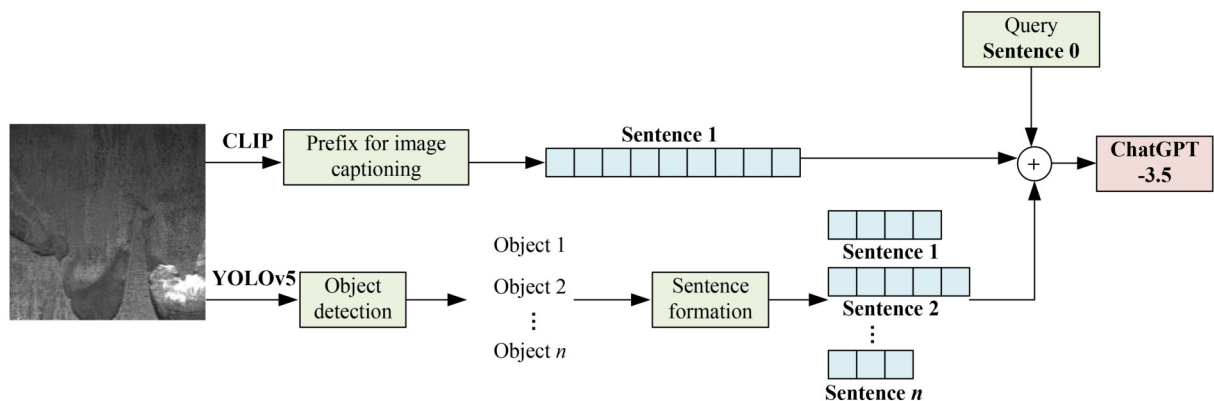


**Fig. 3**  Structure of the visual language model.

language module, ensuring accurate textual representation of slab defect detection. Additionally, user input is required to specify the task that GPT-3.5 should perform.

## 4.2 Fine-tuning of LLM

In addition to the VLM, ChatGPT-3.5 also offers specialized knowledge question-answering capabilities for the field of steelmaking, requiring fine-tuning of the system. To achieve robust linguistic performance by LLM in the domain of steel production, domain-specific fine-tuning on relevant data is necessary to ensure satisfactory generated answers.

Figure 4 shows the large-scale language model fine-tuning process. The fine-tuning module comprises two datasets: a labeled judgment dataset for supervised fine-tuning and a question-answering dataset for unsupervised learning. C represents the labeled dialogues, where each instance includes an input sequence and its label $y$, given by $x^1, x^2, \ldots, x^n$. The input dataset is represented as $U$, with each input $x$ containing $u_1, u_2, \ldots, u_m$. The output based on the transformer for each input $x$ is denoted as Eq. (3):

$$h_0 = UW_e + W_p, \tag{3}$$

where $W_e$ represents the description embedding matrix, and $W_p$ represents the position embedding matrix. When the number of layers in the Transformer exceeds 1, it is necessary to pass through the Transformer module. The Transformer module comprises self-attention and feedforward neural networks. Equations (4)–(6) represent the calculation formulas for the self-attention module:

$$\begin{aligned} Q &= W_q * h_{l-1}, \\ K &= W_k * h_{l-1}, \\ V &= W_v * h_{l-1}, \end{aligned} \tag{4}$$

$$\text{Attention}_{\text{Score}} = \frac{\text{softmax}(Q * K^T)}{\sqrt{d_k}}, \tag{5}$$
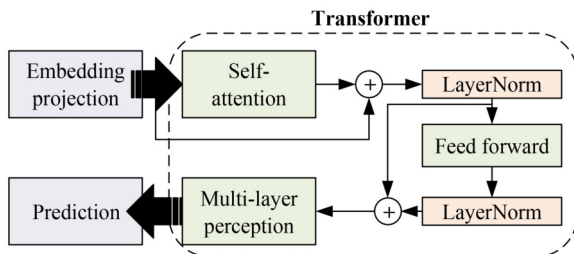
$$Z_{l-1} = \text{Attention}_{\text{Score}} * V, \tag{6}$$



**Transformer**

**Fig. 4** Fine-tuning large-scale language models for the manufacturing domain.

$$h_l = Relu(Z_{l-1}) * W_2 + b_2, \tag{7}$$

$$P(y|x^1, \ldots, x^n) = \text{softmax}(h_l W_y), \tag{8}$$

$$L_2(C) = \sum_{(x,y)} \log P(y|x^1, \ldots, x^n). \tag{9}$$

Within the context, $Q$ represents Query, $K$ represents Key, $V$ represents Value, and $W_q, W_k, W_v, W_y$ represent the corresponding weight matrices. $d_k$ is the dimension of the $K$ matrix, and $K^T$ is the transpose matrix of the $K$ matrix. The weight matrix $W_2$ and bias matrix $b_2$ are utilized as parameters in the feedforward neural network. $h_l$ represents the activation function. Eq. (8) represents the final result obtained by the fine-tuned model, and Eq. (9) represents the objective function.

We have carefully selected a comprehensive dataset pertaining to the iron and steel manufacturing industry. This dataset includes a wealth of knowledge including various aspects such as the smelting process, material properties, and equipment operations. It includes both common questions and specific domain-related issues, ensuring that ChatGPT-3.5 possesses a wide coverage when providing professional knowledge.

The following is a partial excerpt from the example dataset:

Question: What are the fundamental procedures involved in smelting iron in a blast furnace?

Answer: In a blast furnace, the process of smelting iron includes charging raw materials, reducing gas, and desulfurizing iron.

Question: What part does pellet ore play in the sintering process, please?

Answer: In order to improve the strength and wear resistance of the sinter, pellet ore helps in bonding and combustion during the sintering process.

Question: What are the fundamental operational procedures involved in the production of converter steel?

Answer: The method of creating converter steel involves blowing oxygen, adding alloy, and adjusting the basic slag. The furnace's oxygen concentration is controlled to achieve the steelmaking process.

With the utilization of such a dataset, ChatGPT-3.5 can be fine-tuned to provide more precise responses to expert inquiries regarding the smelting of iron and steel. It can offer detailed explanations and relevant facts, thereby enhancing the system's professionalism and comprehensiveness in addressing user inquiries in the field of steel manufacturing.

In the implementation of ChatGPT-3.5's Q & A functionality, the fine-tuning module undergoes a complex series of steps to ensure the model is appropriately optimized for the domain-specific knowledge of iron and steel smelting. The fine-tuning module leverages a dataset denoted as C, specially curated for the domain of

iron and steel smelting. This dataset includes a wide range of professional knowledge, with every query and potential response being closely connected to dataset *C*. As a result, the model possesses a comprehensive understanding of domain-specific facts.

During the fine-tuning stage, samples from dataset *C* are utilized to modify the model's parameters. This adjustment aims to enhance the model's adaptability to the body of knowledge in the iron and steel smelting domain by optimizing the loss function. The following outlines the specifics of the changes made to the model's parameters:

1) **Layer scaling adjustment**: The layer scale of the model has been modified to effectively address complex issues encountered in iron and steel smelting. This adjustment includes fine-tuning the number of attention heads to capture unique semantic links within the domain.

2) **Attention mechanism adjustment**: The attention mechanism parameters have been modified to enhance the model's ability to grasp long-range dependencies and comprehend technical terminology specific to the iron and steel smelting domain. This adjustment is crucial due to the involvement of complex technical terms and dependencies in the domain.

3) **Learning rate adjustment**: The learning rate for fine-tuning has been adjusted to strike a balance between the model's agility in learning new tasks and its ability to grasp general language. This adjustment enables the model to gain deeper insights into iron and steel smelting from dataset C when appropriately calibrated.

4) **Task-specific initialization**: The model's initial parameters have been modified using task-specific initialization procedures to better cater to the requirements of the iron and steel smelting task. Specific initialization techniques are employed for the model's embedding layer and other pivotal layers to better meet the demands of the task.

5) **Fine-tuning epochs**: The number of fine-tuning epochs has been adjusted to ensure thorough learning of dataset *C* by the model, enhancing its capacity to adapt to expert knowledge in iron and steel smelting.

By implementing these targeted adjustments to the model parameters, our objective is to elevate the performance and professionalism of ChatGPT-3.5 in question-answering tasks associated with iron and steel smelting. These adjustments align the model more closely with the specific requirements of the domain, bolstering its applicability within the field of professional knowledge. As a result, the model becomes better equipped to comprehend context accurately and provide professional responses to relevant inquiries.

Additionally, the fine-tuning module employs a SoftMax layer to normalize the model's output. This normalization step transforms the output into a probability distribution, facilitating the thoughtful generation of likely answers. Through this normalization process, the model can discern the most appropriate response, thereby improving the accuracy of its answers to user questions.

During the fine-tuning process, we incorporate task-specific prompt engineering to guide the model in generating domain-tailored outputs for iron and steel smelting. This involves explicitly mentioning "iron and steel smelting" and other relevant keywords in the prompt, guiding the model to acquire a deeper understanding and expertise in the domain. This approach ensures that the generated answers maintain a high level of professionalism.

With these refined details, ChatGPT-3.5 demonstrates a heightened level of professionalism in addressing question-answering tasks within the iron and steel smelting domain. It establishes itself as a dependable and professional resource, offering users accurate and comprehensive answers to their inquiries regarding professional knowledge.

Following the completion of the aforementioned refinements, the fine-tuned Transformer-based model is seamlessly integrated with ChatGPT-3.5. The fine-tuned model serves as a precursor, preceding ChatGPT-3.5 in the dialogue generation process. Specifically, during the application of the dialogue generation functionality, the input text initially passes through the fine-tuned model before being forwarded to ChatGPT-3.5 for text generation. This approach guarantees that the generated dialogue aligns with the steel production requirements specified by the fine-tuned model.

## 5    System implementation and validation

### 5.1    Experimental configuration

In pursuit of effectively managing the iron and steel smelting process while ensuring the quality of steel products, this paper establishes a Steel Smelting Process Management System based on LLM.

The experiments employ various sensors, including the SA-S6016 dual-color infrared thermometer temperature sensor, the ADXL345 piezoelectric accelerometer sensor, MEMSIC's MXC6232xMP tilt sensor, PCB Corporation's PCB 1102-05A rotary torque sensor, and a laser scanner, FARO SCENE, from Faro Technologies. The smelting process utilizes a converter for high-temperature steel smelting.

The constructed smelting process management system is developed on an 11th Gen Intel(R) Core (TM) i7-11700K @ 3.60GHz 3.60 GHz CPU, 32.0 GB RAM PC, and Visual Studio 2013.

Figure 5 illustrates the experimental verification process. Initially, sensors and algorithms with diverse functionalities are employed to collect, identify, and predict relevant smelting data during the iron and steel
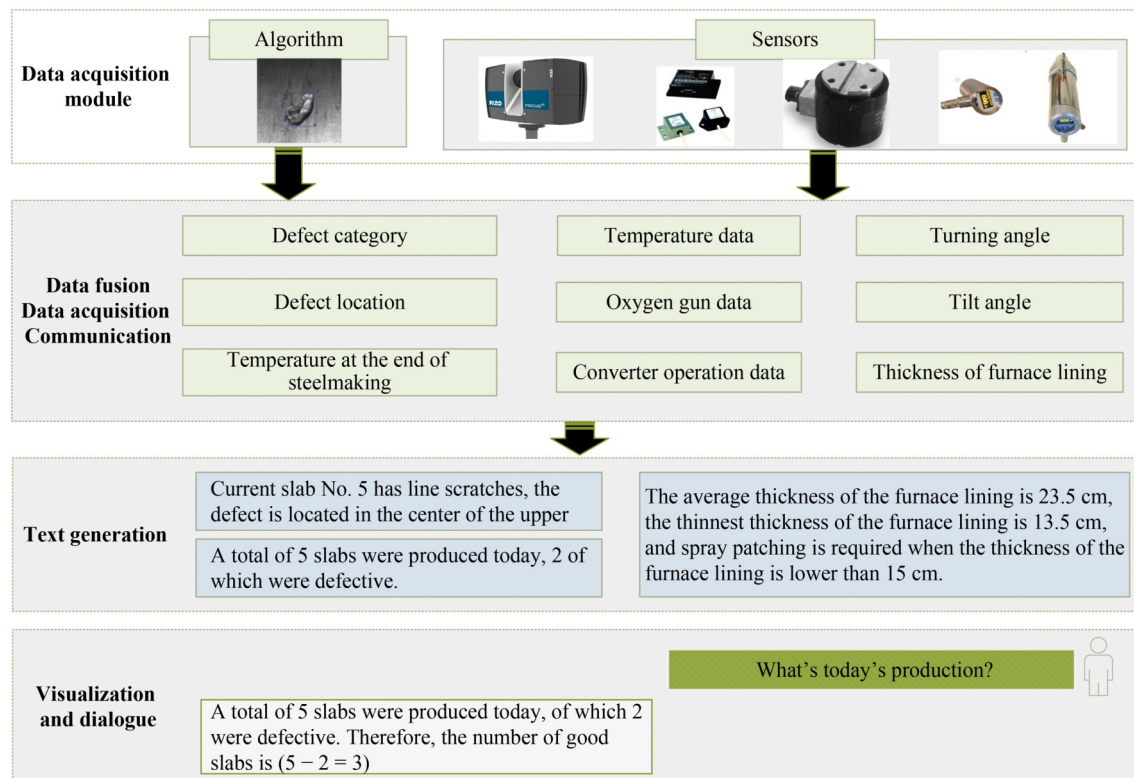
**Fig. 5** Experimental verification process.

smelting process. Subsequently, a combination of database and cloud computing technologies is utilized to gather data from various sources, facilitating the generation of natural language that can be comprehended by large language models. Finally, a visualization tool presents smelting data in user-friendly and easily comprehensible text formats.

To achieve the collection, identification, and prediction of relevant data in the iron and steel smelting process using sensors and algorithms, various types of sensors and algorithms will be employed to gather diverse information. Table 1 outlines the types and sources of the relevant data.

The aforementioned data types include various aspects, including temperature, pressure, flow rate, vibration, and images, providing a comprehensive representation of diverse information during the smelting process. The data size is dependent on the collection frequency and sensor sensitivity. The system utilizes appropriate data compression and storage techniques to ensure efficient data management and analysis.

Upon obtaining the extensive production data from the aforementioned manufacturing processes, integration into a LLM proceeds through the following steps:

1) **Data collection**: Primarily, there is a need to gather a significant volume of production data related to the manufacturing processes outlined in the dataset. This includes sensor data, production records, log files, operational instructions, and similar information.

2) **Data preprocessing**: The collected data is susceptible to issues such as noise, blank values, or inconsistency, necessitating cleaning and preprocessing. This involves procedures such as outlier removal, missing value imputation, data normalization, or standardization to ensure data quality and consistency.

3) **Corpus construction**: The cleaned and preprocessed data forms the corpus, which is essential for training the language model. The corpus should include diverse and representative data, covering various processes and scenarios within the production domain.

4) **Model training**: The corpus is input into the model for training purposes. During the training phase, the model learns patterns, regularities, and linguistic structures from the corpus, thereby enhancing its comprehension and text generation capabilities.

5) **Deployment and application**: Subsequently, the trained language model is deployed into the production environment and integrated into the relevant systems. The model is used for tasks such as natural language generation and dialogue generation, providing support for management and decision-making processes within the steel production industry.

5.2   Verification process

This paper describes the integration of the ChatGPT-3.5 API into an LLM-based steel smelting process management system. This integration allows users to access

**Table 1**  Type and source of data

| Data Type | Data Sources |
|---|---|
| Temperature data | Measurement of the temperature distribution in the converter by means of a two-color infrared pyrometer temperature sensor |
| Pressure Data | Monitoring of pressure changes during the smelting process by means of piezoelectric acceleration sensors. In addition, the weight of raw materials and finished products is measured by means of sensors |
| Turning Angle Data | Measurement of the angle of rotation in the converter by means of a rotational torque sensor |
| Tilt data | Measurement of the inclination angle in the converter by means of a tilt sensor |
| Vibration Data | Vibration monitoring of the machine and the furnace by means of vibration sensors |
| Lining Thickness Data | Measurement of the three-dimensional point cloud in the converter by means of an infrared laser scanner and obtaining of the furnace lining thicknesses |
| Defect Data | Defect detection by means of a vision sensor after taking an image of the slab to localize and classify the defects |

information about smelting processes and steel smelting technology. Users can make inquiries and receive relevant information, enabling them to perform management and control operations based on the provided data. The conversational readability of the system makes it accessible not only to professionals and researchers but also to personnel from various roles. This integration offers an open and user-friendly platform, promoting information sharing and collaborative operations in the smelting process. Accurate smelting data facilitates informed decision-making regarding equipment.

Before starting a steel smelting task, personnel can access the LLM-based steel smelting process management system platform by entering their account credentials, data addresses, and related information. The system includes the following functionalities:

**1) System visualization functionality**

As shown in Fig. 6, the system presents key parameters of the smelting process through intuitive charts and real-time data. The interface displays images of the smelting site, furnace lining thickness, and basic parameters of the smelting process. Users can also control smelting operations and access historical smelting data through interactive buttons on the right. The user-friendly data display and interactive interface facilitate browsing, analysis, and comprehensive understanding of the smelting process status. Users can also perform control and management operations based on different smelting states.

**2) Dialogue functionality**

The system incorporates the ChatGPT-3.5 API, allowing users to engage in natural language conversations. By clicking the intelligent dialogue interface button in the lower right corner of the process management system, users can ask questions about the smelting status, such as *"What is the current furnace temperature?"* or *"Have there been recent changes in production efficiency?"* Leveraging the robust language understanding and generation capabilities of the ChatGPT-3.5 API, the system provides clear and understandable answers to users' questions. This enables users to obtain detailed information about the smelting process without requiring specialized knowledge. Figure 7 illustrates users posing questions such as *"What are the current production results?",*

*"What is the current level of slab defects?"* and *"Does the current converter require spray patching?"* along with the intelligent responses. The dialogue functionality is also available on multiple independent client devices, including mobile phones and computers, making smelting process management more convenient.

**3) Domain-specific Q&A functionality**

The system facilitates user interaction with the Chat-GPT-3.5 API to address various technical and operational inquiries within the steel smelting domain. Users can seek information on specific alloy formulations, smelting reaction mechanisms, equipment maintenance, and other aspects, thereby deepening their understanding of the smelting process.

For instance, users can ask questions like, *"Could you provide a detailed description of the recommended alloy formulation for high-strength alloy production?"* or *"What is the oxidation reaction mechanism during the smelting process at high temperatures?"* or *"If vibrations occur during the operation of the smelting furnace, what problems might exist, and what equipment maintenance measures should be taken?"* The system, grounded in the steel smelting domain, delivers more professional responses to users.

The system's proficiency in domain-specific Q&A is evident through its use of accurate and professional terminology, comprehensive knowledge of the steel smelting domain, and provision of high-level technical and operational guidance to users.

Compared to other industrial process management systems, this model-based system offers several significant advantages:

1) User-friendliness: By utilizing dialogue and visualization, the system reduces communication barriers between users and the system, facilitating easy comprehension of the smelting process status for personnel at all levels.

2) Intelligent Q&A: The integration of the ChatGPT-3.5 API enhances system intelligence, allowing users to ask questions in natural language and receive high-quality answers, eliminating the need for specialized expertise.

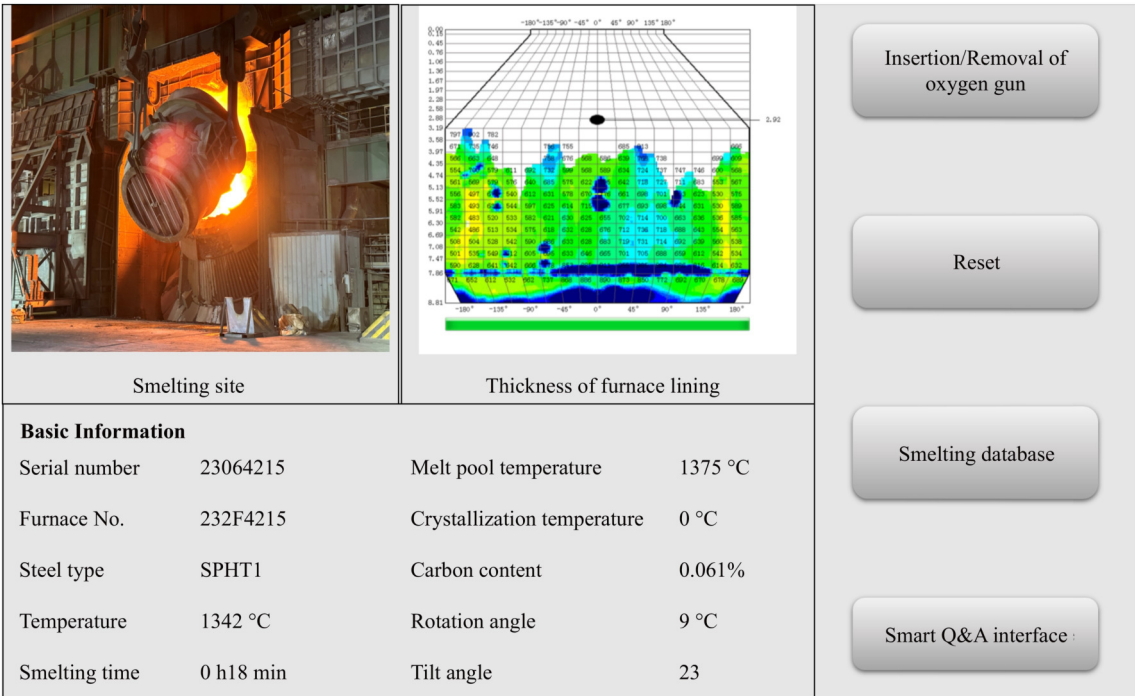3) Real-time monitoring: The system updates data in real-time, ensuring user engagement with the smelting

| Basic Information | | | |
| --- | --- | --- | --- |
| Serial number | 23064215 | Melt pool temperature | 1375 °C |
| Furnace No. | 232F4215 | Crystallization temperature | 0 °C |
| Steel type | SPHT1 | Carbon content | 0.061% |
| Temperature | 1342 °C | Rotation angle | 9 °C |
| Smelting time | 0 h18 min | Tilt angle | 23 |

Smelting site — Thickness of furnace lining

Insertion/Removal of oxygen gun

Reset

Smelting database

Smart Q&A interface

**Fig. 6**  Process Management System (PMS) interface.
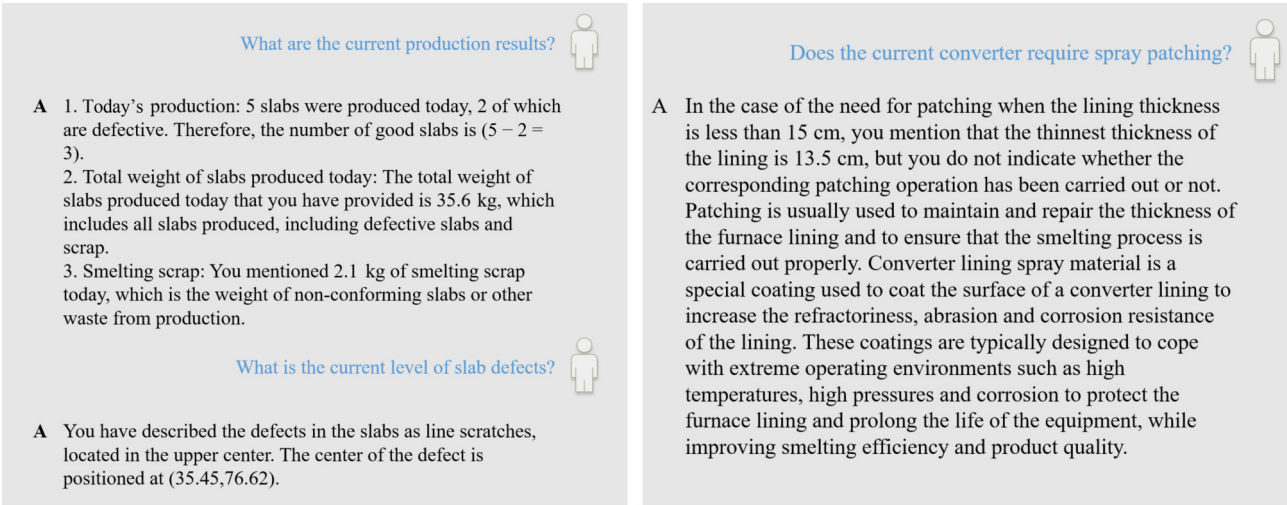


**Fig. 7**  Visualization and conversation.

process and enabling timely operations and adjustments, thereby enhancing real-time production and flexibility.

4) Wide-ranging applications: By supporting domain-specific Q&A, the system not only provides smelting process status information but also offers users comprehensive technical knowledge and operational advice, catering to the needs of users at different levels.

Q&A systems play a crucial role in the management of steel production processes. They assist managers by integrating data, analyzing information, and providing real-time feedback, enabling informed decision-making. The system continuously monitors various parameters and indicators throughout the production process and promptly generates alerts and notifications when anomalies are detected. Managers can engage in a dialogue with the system to address specific issues and take appropriate actions, thus preventing production interruptions or quality problems. Additionally, the system leverages historical data analysis to identify trends and patterns within the production process. This enables accurate forecasting of future production scenarios, supporting managers in making adjustments and decisions effectively. Furthermore, the system acts as a comprehensive knowledge repository, storing and managing valuable information and experiences related to production processes. When faced with production issues, personnel can query the

system, which provides solutions or recommendations based on existing knowledge and experience. Moreover, the system promotes knowledge sharing and collaboration among different departments and personnel, thereby enhancing the overall efficiency and quality of the team's work. Lastly, the system offers decision support and optimization suggestions to managers based on real-time data and historical insights.

In summary, this system enhances the monitoring, management, and technical exchange of the steel smelting process by providing intelligent, intuitive, and comprehensive visualization and dialogue. Its unique features distinguish it in the field of industrial process management, offering users a holistic and efficient solution for steel smelting process management.

This LLM-based steel smelting process management system, integrated with the ChatGPT-3.5 API, aims to deliver an intelligent and user-friendly interactive experience. Through visualization and dialogue, the system empowers users to gain in-depth insights into the real-time status of the smelting process, access relevant technical information, and perform necessary management and control operations.

The system utilizes a MongoDB-based software development approach for its databases. Real-time data pertaining to smelting processes and equipment resulting from the steel smelting process are stored in MongoDB database tables, each equipped with the appropriate fields. Additionally, the system has incorporated secondary development interfaces, facilitating interactions with big data and artificial intelligence models to achieve business optimization.

In order to establish the superiority of our proposed method, we have selected the Vision-Language Pre-training (VLP) model by Zhou et al. (2020), the Oscar model by Li et al. (2020), and the BUTD model by Anderson et al. (2018) for comparison. Our evaluation metrics include ROUGE-L (Lin, Och, 2004), CIDEr (Vedantam et al., 2015), and SPICE (Anderson et al., 2016), and the comparative effectiveness is presented in Table 2. The experimental results clearly demonstrate that our proposed model significantly outperforms the other models in terms of text generation quality, image description quality, and textual accuracy. Our model generates text that closely resembles the reference text, resulting in more vivid and enriched image descriptions, and achieves noticeably higher scores in comparison to the other models. Additionally, our model exhibits remarkable performance in terms of textual accuracy, thanks to its enrichment of semantic information and precision in description. Therefore, the effectiveness of our proposed method is convincingly demonstrated.

The monthly production capacity of a steel smelting plant is a critical measure of its operational efficiency. This measure is typically evaluated based on production yield data obtained throughout the manufacturing process.

To accurately assess the monthly production capacity, it is crucial to track and record the amount of crude steel produced each month. This involves collecting daily production data, including crude steel output, and ensuring the accuracy and completeness of the data. Statistical analysis is then performed on the daily production data to calculate the total crude steel production for each month by aggregating the daily figures. To validate the data quality, checks are conducted to ensure there are no omissions or errors. The total monthly crude steel production is considered an approximation of the monthly production capacity, representing the steel smelting plant's actual output capabilities over the course of a month.

When discussing steel smelting, wastage refers to the various factors that contribute to the consumption of energy and raw materials throughout the production process. To evaluate these losses, different factors are continuously monitored and measured. This involves categorizing the losses into various types, such as ore loss and slag loss, each with their own distinct causes and monitoring methods. Sensors and monitoring devices are utilized to track the production process, and data on losses for each category is recorded and synchronized with production data for further comparison and analysis.

Understanding these two assessment metrics and their specific evaluation processes is crucial for comprehending the operational status of a steel smelting plant. This understanding provides robust support for improvement and optimization. Table 3 presents a comparative result of output and waste before and after implementing our LLM-based steel smelting process management system, validating its effectiveness in reducing management and communication costs, enhancing production efficiency, minimizing material waste, and consequently lowering smelting costs.

**Table 2**  Comparison of model effect

| Model | ROUGE-L↑ | CIDEr↑ | SPICE↑ |
|---|---|---|---|
| VLP | 27.5 | 70.1 | 20.6 |
| Oscar | 28.1 | 73.4 | 25.4 |
| BUTD | 26.7 | 72.8 | 21.3 |
| Ours | 30.4 | 79.6 | 28.3 |

**Table 3**  Comparison of practical application effects

| | Monthly production capacity(t) | Wastage (t) |
|---|---|---|
| None | 5746 | 272 |
| With our system | 6762 | 241 |

# 6  Conclusions and future work

In this study, we integrate management data with intelligent Q&A to enhance the benefits of the LLM-based steel

smelting process management system. This system offers a way to improve production efficiency and optimize management in the steel smelting sector by leveraging the strong processing capabilities and clever decision support mechanisms associated with integrating massive language models. Additionally, it offers a potential pathway and opportunity for integrating massive language models with the steel smelting sector in the future.

The exceptional natural language processing skills of the LLM-based system enable it to parse and comprehend complex data, including sensor data and manufacturing parameters involved in the smelting process, with ease. This ensures the precision and thoroughness of the system's reaction to data input. The proposed LLM-based steel smelting process management system achieves the following goals:

1) Instant data feedback: The LLM system's ability to assess patterns in smelting data in real-time provides instant data feedback. This enables quick responses to changes in the manufacturing process, which is crucial for quality control, anomaly identification, and production monitoring.

2) Intelligent decision assistance: The solution offers intelligent decision assistance to smelting firms, utilizing deep learning techniques and massive language models. It aids managers in making more informed decisions by providing advice for resource scheduling and optimal production plans based on analysis of large amounts of historical data.

3) Seamless cross-platform data interchange: The LLM system facilitates seamless cross-platform data interchange across various departments within smelting firms. This promotes comprehensive data exchange and unified management, thereby enhancing internal organizational efficiency.

The practical viability of the LLM-based steel smelting process management system is achieved through the development of the aforementioned functionalities. The system's design confirms its efficacy and efficiency in managing the steel smelting process. Currently, the paper has established the framework system, provided access, and demonstrated significant characteristics of the steel smelting process. Additionally, the integration of ChatGPT-3.5 API into the system enhances its conversational capabilities. However, there are still challenges regarding data gathering diversification and feedback control in physical systems. This work aims to address the issues posed by low intelligence and high management and communication expenses in the steel smelting process by building an LLM-based process management system specifically designed for steel smelting. The system utilizes databases and the cloud, which offer benefits such as enhanced data security, improved collaboration, and increased data management effectiveness. These features significantly support the modernization of the steel smelting sector.

Nevertheless, constant research and development of these technologies' specific applications are necessary to better cater to the real-world requirements of smelting companies. The subsequent phases of this project will involve maximal utilization of big data analysis and cloud computing to achieve intelligent operation and maintenance management of the steel smelting process.

**Competing Interests**   The authors declare that they have no competing interests.

# References

Alayrac J B, Donahue J, Luc P, Miech A, Barr I, Hasson Y, Lenc K, Mensch A, Millican K, Reynolds M, Ring R (2022). Reynolds M. Flamingo: A visual language model for few-shot learning. Advances in Neural Information Processing Systems, 35: 23716–23736

Anderson P, Fernando B, Johnson M, Gould S (2016). Spice: Semantic propositional image caption evaluation. In: Proceedings of European Conference on Computer Vision (ECCV): 382–398

Anderson P, He X, Buehler C, Teney D, Johnson M, Gould S, Zhang L (2018). Bottom-up and top-down attention for image captioning and visual question answering. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 6077–6086

Bao Z, He D, Khan M K, Luo M, Xie Q (2023). PBidm: Privacy-preserving blockchain-based identity management system for industrial internet of things. IEEE Transactions on Industrial Informatics, 19(2): 1524–1534

Bellavista P, Fogli M, Giannelli C, Stefanelli C (2023). Application-aware network traffic management in MEC-integrated industrial environments. Future Internet, 15(2): 42

Bessarabov A M, Trokhin V E, Popov A K, Radetskaya A S (2023). CALS project: Hardware and technological design of a modular water management system for industrial applications. Chemical and Petroleum Engineering, 58(9–10): 855–864

Borkowski A A (2023). Applications of ChatGPT and large language models in medicine and health care: Benefits and pitfalls. Federal Practitioner, 40(6): 170–173

Cui Y, Niekum S, Gupta A, Kumar V, Rajeswaran A (2022). Can foundation models perform zero-shot task specification for robot manipulation? In: Proceedings of 4th Annual Learning for Dynamics and Control Conference, Stanford, USA

De Curtò J, De Zarzà I, Calafate C T (2023). Semantic scene under-

standing with large language models on unmanned aerial vehicles. Drones, 7(2): 114

Demertzis K, Demertzis S, Iliadis L (2023). A selective survey review of computational intelligence applications in the primary subdomains of civil engineering specializations. Applied Sciences-Basel, 13(6): 3380

Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T T, Dehghani M, Minderer M, Heigold G, Gelly S, Uszkoreit J (2021). An image is worth 16x16 words: Transformers for image recognition at scale. In: Proceedings of International Conference on Learning Representations 2021

Fang L, Su F, Kang Z, Zhu H (2023). Artificial neural network model for temperature prediction and regulation during molten steel transportation process. Processes, 11(6): 1629

Franco D'Souza R, Amanullah S, Mathew M, Surapaneni K M (2023). Appraising the performance of ChatGPT in psychiatry using 100 clinical case vignettes. Asian Journal of Psychiatry, 89: 103770

Fu T, Li P, Liu S (2024a). An imbalanced small sample slab defect recognition method based on image generation. Journal of Manufacturing Processes, 118: 376–388

Fu T, Liu S, Li P (2024b). Digital twin-driven smelting process management method for converter steelmaking. Journal of Intelligent Manufacturing, 2024: 1–17

Gu X, O'Leary T Y, Kuo W, Cui Y (2022). Open-vocabulary object detection via vision and language knowledge distillation. In: Proceedings of International Conference on Learning Representations 2022

Hein-Pensel F, Winkler H, Brückner A, Wölke M, Jabs I, Mayan I J, Kirschenbaum A, Friedrich J, Zinke-Wehlmann C (2023). Maturity assessment for Industry 5.0: A review of existing maturity models. Journal of Manufacturing Systems, 66: 200–210

Huang H C, Tsai C H, Lin H C (2023). Development of 5G cyber-physical production system. International Journal of Networked and Distributed Computing, 11(1): 9–19

Huang W, Abbeel P, Pathak D, Mordatch I (2022). Language models as zero-shot planners: Extracting actionable knowledge for embodied agents. In: Proceedings of 39th International Conference on Machine Learning (ICML), Baltimore, MA, USA

Iwańkowicz R, Rutkowski R (2023). Digital twin of shipbuilding process in Shipyard 4.0. Sustainability, 15(12): 9733

Jaber M M, Ali M H, Abd S K, Jassim M M, Alkhayyat A, Kadhim E H, Alkhuwaylidee A R, Alyousif S (2023). AHI: A hybrid machine learning model for complex industrial information systems. Journal of Combinatorial Optimization, 45(2): 58

Jadhav A, Shandilya S K, Izonin I, Gregus M (2023). Effective software effort estimation leveraging machine learning for digital transformation. IEEE Access: Practical Innovations, Open Solutions, 11: 83523–83536

Kouzapas D, Stylianidis N, Panayiotou C G, Eliades D G (2023). Ontology-based reasoning to reconFigure industrial processes for energy efficiency. In: Proceedings of 2023 31st Mediterranean Conference on Control and Automation (MED). 79–84

Li S, Guo Z, Zang X (2023). Advancing the production of clinical medical devices through ChatGPT. Annals of Biomedical Engineering, 52(3): 441–445

Li X J, Yin X, Li C Y, Zhang P C, Hu X W, Zhang L, Wang L, Hu H,

Dong L, Wei F, Choi Y (2020). Oscar: Object-semantics aligned pre-training for vision-language tasks. In: Proceedings of 16th European Conference on Computer Vision (ECCV 2020). 121–137

Lin C Y, Och F J (2004). Automatic evaluation of machine translation quality using longest common subsequence and skip-bigram statistics. In: Proceedings of 42nd Annual Meeting of the Association for Computational Linguistics (ACL-04), 605–612

Liu R, Xie X (2024). Improve the industrial digital transformation through Industrial Internet platforms. Frontiers of Engineering Management, 11(1): 167–174

Mallio C A, Sertorio A C, Bernetti C, Beomonte Zobel B (2023). Large language models for structured reporting in radiology: performance of GPT-4, ChatGPT-3.5, Perplexity and Bing. La Radiologia Medica, 128(7): 808–812

Massey P A, Montgomery C, Zhang A S (2023). Comparison of Chat-GPT-3.5, ChatGPT-4, and orthopaedic resident performance on orthopaedic assessment examinations. Journal of the American Academy of Orthopaedic Surgeons, 31(23): 1173–1179

Mokady R, Hertz A, Bermano A H (2021). ClipCap: CLIP prefix for image captioning. Computer Science. arXiv: 2111.09734

Nair S, Rajeswaran A, Kumar V, Finn C, Gupta A (2022). R3M: A universal visual representation for robot manipulation. arXiv: 2203.12601

O'Leary D E (2023). Enterprise large language models: Knowledge characteristics, risks, and organizational activities. Intelligent Systems in Accounting, Finance & Management, 30(3): 113–119

Pavlopoulos J, Romell A, Curman J, Steinert O, Lindgren T, Borg M, Randl K (2023). Automotive fault nowcasting with machine learning and natural language processing. Machine Learning, 113(2): 843–861

Peng G, Cheng Y, Zhang Y, Shao J, Wang H, Shen W (2022). Industrial big data-driven mechanical performance prediction for hot-rolling steel using lower upper bound estimation method. Journal of Manufacturing Systems, 65: 104–114

Radford A, Kim J W, Hallacy C, Ramesh A, Goh G, Agarwal S, Sastry G, Askell A, Mishkin P, Clark J, Krueger G, Sutskever I (2021). Learning transferable visual models from natural language supervision. In: Proceedings of 38th International Conference on Machine Learning, Virtual

Redmon J, Farhadi A (2017). YOLO9000: Better, faster, stronger. In: Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, USA, 6517–6525

Semenov Y S, Shumelchyk Y I, Horupakha V V, Semion I Y, Vashchenko S V, Khudyakov O Y, Chychov I V, Hulina I H, Zakharov R H (2022). Development and implementation of decision support systems for blast smelting control in the conditions of PrJSC "Kamet-Steel". Metals, 12(6): 985

Sharma P, Ding N, Goodman S, Soricut R (2018). Conceptual captions: A cleaned, hypernymed, image alt-text dataset for automatic image captioning. In: Proceedings of 56th Annual Meeting of the Association-for-Computational-Linguistics (ACL), Melbourne, Australia, 2556–2565

Shi J J, Zeng S, Meng X (2017). Intelligent data analytics is here to change engineering management. Frontiers of Engineering Management, 4(1): 41–48

Shi Y (2015). Challenges to engineering management in the big data

era. Frontiers of Engineering Management, 2(3): 293–303

Sievers J, Blank T (2023). A systematic literature review on data-driven residential and industrial energy management systems. Energies, 16(4): 1688

Snoswell C L, Snoswell A J, Kelly J T, Caffery L J, Smith A C (2023). Artificial intelligence: Augmenting telehealth with large language models. Journal of Telemedicine and Telecare: 1357633X2311690

Stepanov V K, Madzhumder M S, Begunova D D (2023). Exploring the potential of applying the artificial intelligence language model ChatGPT-3.5 in library and bibliographic activities. Scientific and Technical Information Processing, 50(3): 166–175

Thiébaut R, Hejblum B, Mougin F, Tzourio C, Richert L (2023). Chat-GPT and beyond with artificial intelligence (AI) in health: Lessons to be learned. Joint, Bone, Spine, 90(5): 105607

Vedantam R, Zitnick C L, Parikh D (2015). Cider: Consensus-based image description evaluation. In: Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 4566–4575

Wei J, Tay Y, Bommasani R, Raffel C, Zoph B, Borgeaud S, Yogatama D, Bosma M, Zhou D, Metzler D, Chi E H (2022). Emergent abilities of large language models. arXiv: 2206.07682

Xiao Y, Zheng S, Shi J, Du X, Hong J (2023). Knowledge graph-based manufacturing process planning: A state-of-the-art review. Journal of Manufacturing Systems, 70: 417–435

Yu Z, Gong Y (2024). ChatGPT, AI-generated content, and engineering management. Frontiers of Engineering Management, 11(1): 159–166

Zeng A, Attarian M, Ichter B, Choromanski K, Wong A, Welker S, Tombari F, Purohit A, Ryoo M, Sindhwani V, Lee J (2022b). Socratic models: Composing zero-shot multimodal reasoning with language. arXiv: 2204.00598

Zeng A, Florence P, Tompson J, Welker S, Chien J, Attarian M, Armstrong T, Krasin I, Duong D, Sindhwani V, Lee J (2022a). Transporter networks: Rearranging the visual world for robotic manipulation. arXiv: 2010.14406

Zheng H, Liu S, Zhang H, Yu J, Bao J (2024). Visual triggered contextual guidance for lithium battery disassembly: A multi-modal event knowledge graph approach. Journal of Engineering Design, 2024: 1–26

Zhou L, Palangi H, Zhang L, Hu H, Corso J, Gao J (2020). Unified vision-language pretraining for image captioning and VQA. Proceedings of the AAAI Conference on Artificial Intelligence, 34(7): 13041–13049

Zhu T, Wang X, Yu Y, Li C, Yao Q, Li Y (2023). Multi-process and multi-pollutant control technology for ultra-low emissions in the iron and steel industry. Journal of Environmental Sciences, 123: 83–95 (in Chinese)