Abstract

In today’s data-driven world, quick access to scholarly info is vital. However, current academic search engines face challenges such as restricted text-based searching, uncertainties related to researcher names, absence of contact details, and lack of profile summaries. To mitigate these issues, we introduce ScholarFace, an innovative concept that could transform how we search for academic knowledge. It uses face recognition and language generation technology to spot scholars in photos effortlessly, giving us rich profiles and summaries of their work. It also offers an interactive chat element for users to get more insights about a scholar, reducing online search efforts. We take privacy and ethics issues seriously and ensure ScholarFace complies with rules and regulations. With ScholarFace, we hope to create a smarter, effortless, more connected scholarly world.
ScholarFace: Scanning Faces, Discovering Minds

Wong Yu Kang*, Toluwani Aremu*, Younes Balah*, Maryam Nadeem*, Ivo Gollini Navarette* and Abdulmotaleb El Saddik†
*Mohamed Bin Zayed University of Artificial Intelligence (MBZUAI) †University of Ottawa
{yukang.wong, toluwani.aremu, younes.balah, maryam.nadeem, ivo.navarrete, a.elsaddik}@mbzuai.ac.ae
elsaddik@uottawa.ca

Abstract—In today’s data-driven world, quick access to scholarly info is vital. However, current academic search engines face challenges such as restricted text-based searching, uncertainties related to researcher names, absence of contact details, and lack of profile summaries. To mitigate these issues, we introduce ScholarFace, an innovative concept that could transform how we search for academic knowledge. It uses face recognition and language generation technology to spot scholars in photos effortlessly, giving us rich profiles and summaries of their work. It also offers an interactive chat element for users to get more insights about a scholar, reducing online search efforts. We take privacy and ethics issues seriously and ensure ScholarFace complies with rules and regulations. With ScholarFace, we hope to create a smarter, effortless, more connected scholarly world.

Index Terms—Information Retrieval, Google Scholar, Language Model, Face Recognition

I. INTRODUCTION

In today’s rapidly evolving information landscape, navigating the vast sea of academic research and scholarly articles by numerous researchers requires a reliable compass. This process is where academic search engines emerge as indispensable tools for both students and researchers. Academic search engines are specialized online platforms designed to facilitate the discovery and retrieval of academic content. While various search engines such as ORCID, ResearchGate, and Semantic Scholar exist, the most widely recognized one remains Google Scholar.

While academic search engines have a longstanding presence, there has been minimal improvement in how they facilitate user searches. The current platforms present various challenges:

- Limited text-only search: Academic search engines offer exclusively text-based search capabilities. The users are constrained to conducting searches based solely on text input (e.g., names, published papers or manually sifting through the names in relevant fields).
- Ambiguity in researchers’ names: When users search for the researcher’s name, they might encounter multiple researchers sharing the same name. In the absence of additional details like affiliation, profile image, or field of expertise, pinpointing the desired researcher becomes challenging.
- Lack of contact information: Notably, widely used search engines like Google Scholar do not provide researchers’ email addresses or contact information. Users must take extra steps to navigate the respective university’s homepage to access the necessary contact information they aim to engage with, thereby elongating the process of establishing connections and collaborations.
- Absence of profile summaries: Current academic search engine lacks the capability to provide summaries. When users seek information about a specific researcher, they must sift through the researcher’s entire publication history to ascertain their past contributions. This task becomes even more challenging if the researcher is highly prolific and has amassed many publications, making the navigation process cumbersome.

To address these inefficiencies, we present ScholarFace, a pioneering concept harnessing the synergy of face recognition and language generation technologies for efficient scholarly information discovery. It swiftly spots scholars in images, like conferences, and utilizes AI to recognize them. Generative language models then provide comprehensive summaries. An interactive element offers users more insights into academic achievements and affiliations, reducing online search efforts. ScholarFace prioritizes ethics, following strict privacy measures for responsible facial recognition use in scholarly info retrieval. To the best of our knowledge, we are the first to attend to the issues surrounding information retrieval in academic search engines.

Here are some ways in which ScholarFace solves the limitations mentioned above:

- Utilizing image-based input for searches: ScholarFace employs sophisticated facial recognition technology to identify scholars precisely.
- Reduce researcher profile ambiguity: ScholarFace forces users to provide valid and up-to-date facial profiles and other information such as affiliations.
- Enhanced user experience: ScholarFace utilizes language generation technology to generate detailed summaries of scholar’s publications. Furthermore, ScholarFace offers user-friendly chat functionalities for supplementary insights.
- Facilitating Collaboration and Networking: ScholarFace encourages collaboration and networking among scholars by displaying up-to-date contact information to users.
II. RELATED WORK

A. Face Recognition

The history of Neural Network-based face recognition traces back to the early developments in the late 1980s and 1990s. These pioneering efforts laid the groundwork for applying neural network architectures to the challenging task of face recognition. One of the earliest successes was the "Eigenfaces" method, introduced by Sirovich and Kirby and used by Turk and Pentland in 1991 [1], [2], which employed Principal Component Analysis (PCA) [4] to extract facial features.

In the following years, advancements in neural network models, particularly the emergence of deep learning, led to substantial improvements in face recognition accuracy and performance [5], [6]. The seminal work of DeepFace [3] by Facebook in 2014 demonstrated the potential of deep convolutional neural networks (CNNs) for highly accurate and scalable face recognition. Since then, neural network-based face recognition has seen rapid progress, driven by the availability of large-scale datasets, powerful computational resources, and continuous innovations in network architectures. Today, this technology has become integral to various applications, from biometric authentication to intelligent surveillance systems. Today, this technology has become integral to various applications, from biometric authentication to intelligent surveillance systems, heralding a new era of efficient and robust face recognition systems.

B. Generative Language Model

In recent times, there has been significant advancement in generative language models, thanks to the emergence of Pre-trained Language Models (PLMs) and Large Language Models (LLMs). These models, which are trained on large quantities of text data, have revolutionized the field of natural language processing. An outstanding example of these LLMs is OpenAI’s ChatGPT, which exhibits remarkable conversational abilities with humans.

These technologies have attracted considerable interest from both individuals and businesses, who are exploring its diverse applications. Several applications have already entered the market, including ChatGPT, Perplexity AI, Microsoft Bing Chat, and Google Bard, each offering unique features and capabilities to users.

C. Scholar Search

In academic collaboration, several platforms have emerged as valuable tools for connecting scholars and researchers within specific domains. These platforms are crucial for academic discovery and collaboration.

Google Scholar is the most widely used platform, offering a vast database of scholarly articles, conference papers, and patents. While it provides a user-friendly interface for searching academic content, challenges such as naming ambiguity and profile picture omission necessitate users to sift through lengthy lists of articles, which can impede effective collaboration and technological advancements. Similar challenges are encountered in Academic Social Networks (ASNs) like Academia and Mendeley.

Semantic Scholar employs natural language processing and machine learning to enhance the search experience but does not fully address naming issues and the trouble of iterating through articles to search for relevant content. Meanwhile, ResearchGate was created as a social networking platform for researchers, fostering knowledge exchange, but the same issues persist.

ORCID introduced a unique identifier system, simplifying authorship recognition. However, it is effective only if users possess the exact name or ID of the scholar they seek. LinkedIn also gained significance in academia but scholarly search can be overwhelming due to its diverse user base.

Our platform addresses these challenges by leveraging facial recognition and language technology. Users can search for scholars using facial recognition, access well-written research summaries, and navigate publications through an integrated LLM-backed chatbot. Further details on this technology will be provided in subsequent sections.

III. SCHOLARFACE

The ScholarFace concept embodies a meticulously designed systematic flow, employing cutting-edge technologies to process facial data and retrieve scholar profiles. At its core, the process involves extracting facial embeddings from a scanned facial image, enabling precise identification and association with the corresponding scholar profile stored within our comprehensive database. Through this initial step, ScholarFace lays the groundwork for a refined and accurate representation of the individual in focus.

Subsequently, the acquired facial embeddings serve as the foundation for our advanced language model to weave a comprehensive tapestry of the scholar’s academic journey. With the prowess of natural language generation, the language model crafts a detailed and insightful summary encapsulating the scholar’s notable achievements, groundbreaking research contributions, and affiliations within the academic sphere. This generated summary stands as a testament to ScholarFace’s dedication to providing users with a holistic understanding of the scholar’s scholarly endeavors.

The culmination of this intricate process lies in the seamless presentation of the comprehensive academic summary as the final output. By orchestrating this seamless integration of facial recognition technology, advanced language models, and a vast repository of scholarly information, ScholarFace empowers users with an efficient and intuitive means to access rich scholarly knowledge. ScholarFace aims to revolutionize scholarly information retrieval as a transformative tool and foster a more interconnected and informed global academic community. Figure 1 shows the basic flow of ScholarFace.

A. Face Recognition

Face recognition is a pivotal component of ScholarFace, our innovative system designed to revolutionize the discovery and access of scholarly information. It is the first crucial
Fig. 1: The ScholarFace system comprises two main components: Sign Up and Search. New users must register by capturing a self-image and entering their scholar details. The captured image undergoes face recognition, producing a vector used as a database key, while the scholar details serve as the corresponding database value stored in the vector database. In the Search section, users start by capturing an image of an unknown researcher. This image is processed by the face recognition model, which identifies the person with the highest similarity in the vector database and returns their name. The system then retrieves the person’s profile. If the profile is cached, indicating a previous retrieval, the system promptly returns the profile summary. Otherwise, it gathers information from the internet based on the Scholar Profile link. Subsequently, the system combines the crawled data with a predetermined prompt and submits it to Large Language Models (LLM) to generate the final profile summary. If the user seeks more information, they can engage in a conversation with the LLM.

Privacy and ethical considerations are paramount in the development of ScholarFace. We are deeply committed to safeguarding individuals’ rights and personal data. To this end, our system will strictly adhere to robust privacy protocols when deployed, ensuring responsible usage of face recognition technology. We recognize the need to balance the convenience of accessing scholarly information and protecting users’ privacy. In later sections, we expand on more design choices to maximize user security.

B. Generative Language Technology

Once a scholar is accurately identified through face recognition, ScholarFace seamlessly retrieves comprehensive profiles containing vital academic information. In our proof-of-concept, to enhance the user experience, we utilized the Microsoft Bing Chat to provide detailed summaries and rel-
evant insights into scholars’ academic achievements, research contributions, and affiliations.

The Microsoft Bing Chat offers a conversational AI chatbot feature for Bing’s search engine, enabling users to interact with an AI chatbot instead of typing search queries. Powered by OpenAI’s GPT-4 language model and Bing’s search rank algorithm, this powerful tool can answer questions, generate text, and assist with creative tasks. It is important to note that we opted for Bing to create a Proof-of-Concept, and we will fine-tune our language models to use when deploying the system.

The integration of face recognition and generative language models in ScholarFace represents a synergistic approach toward achieving our mission of empowering researchers, academics, and knowledge seekers. By bridging facial identification and scholarly information retrieval, our system paves the way for seamless access to academic knowledge and fosters meaningful collaborations within the global scholarly community. We are committed to continuously improving and refining ScholarFace to uphold the highest efficiency, privacy, and user satisfaction standards.

C. Ethical Framework

As we continuously refine and advance ScholarFace, our unwavering commitment to upholding the highest standards of privacy, ethics, and technological innovation remains at the core of our vision. Our goal is to create a transformative tool that simplifies scholarly information discovery and embodies the responsible use of emerging technologies for the greater benefit of academia. To achieve this, we have meticulously curated an extensive set of ethical design policies that we will be following to ensure the utmost protection and integrity of user data:

- Data Collection Integrity: We prioritize the privacy and consent of our users. ScholarFace will never extract or crawl image data from the internet without explicit user consent. Facial data will solely be obtained from users who willingly sign up for the application, ensuring the ethical handling of personal information.
- User Control: Transparency and user agency are fundamental principles in our ethical framework. For users of ScholarFace, we will provide the option to control access to their information by allowing them to opt out of facial data-based profile sharing. Respecting user autonomy is paramount to building trust and maintaining ethical standards.
- Data Security: To safeguard the privacy of our users, we will employ state-of-the-art cryptographic algorithms to encrypt the database. ScholarFace will only store facial embeddings derived from the user’s live face image, preventing unauthorized access to raw data and promoting secure data handling practices.
- Regular Facial Profile Updates: To enhance the accuracy and validity of facial recognition, users will be prompted to update their live facial profiles every six months. This proactive measure ensures precise identification and prevents potential inaccuracies over extended periods of use.
- Rigorous User Validation: To mitigate the risk of impersonation and ensure the credibility of user profiles, ScholarFace will conduct both automatic and manual validation of users’ information. We fortify the system against potential fraudulent activities by employing a two-step verification process.

The above policies only represent a subset of the comprehensive measures we will adhere to and create to ensure ScholarFace adheres to the highest ethical standards. Our commitment to ethical considerations remains central to creating a safe and trustworthy platform for the global scholarly community as we progress.

IV. EVALUATION

To validate the feasibility of the core concept behind ScholarFace, we executed a proof-of-concept utilizing the ArcFace [7] face recognition model in conjunction with the Bing Chat for the prototype illustrated in Figure 2. Our dataset encompassed a modest collection comprising 149 MBZUAI researchers, along with a supplementary test dataset containing 79 distinct images of the same individuals drawn from the database. The evaluation process was divided into two phases. Firstly, we assessed the capability of the face recognition model to identify the researchers’ faces precisely. Subsequently, the correct outcomes were directed to a second evaluation phase to detect whether the Bing Chat could provide accurate person summaries when presented with a predefined prompt containing the individual’s name and affiliation.

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
<th>Number of test cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Inaccurate</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>Major Inaccuracies</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>Partial Inaccuracies</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>Accurate</td>
<td>60</td>
</tr>
</tbody>
</table>

For the first evaluation, out of 79 test cases, 65 were correctly identified. We then input the information obtained from the first step with the predefined prompt into the Bing Chat. We manually evaluate the summary generated by the Bing Chat to ensure its validity. We assign scores for different summaries generated based on Table 1. The results showed that most of the information in the summary is valid. However, there are a few cases that the Bing Chat misidentified different scholars as the same person due to the same name and Bing Chat cannot find anything related to the person.

In the initial assessment, among the 79 test cases, 65 were accurately identified. Following this, we fed the data acquired from this initial step and a predetermined prompt into Bing Chat. We then personally assessed the summaries produced by Bing Chat to verify their validity. We also assigned scores based on the criteria outlined in Table 1. There are 92% of the test cases fully accurate. Upon analyzing the results, it became evident that the majority of the information in the summaries was indeed accurate. However, a few instances emerged where
Bing Chat mistakenly identified different scholars’ achievements who shared the same name, or in some cases, it was unable to retrieve any relevant information pertaining to the individual in question.

We will explore novel face recognition ideas and language models to ensure the best results through our product. Besides, we will also investigate further the prompt used to obtain the result.

V. CONCLUSION

In this paper, we introduced the ScholarFace concept and shed light on the advantages that could come with the integration of face recognition algorithms and generative language models for academic/research purposes. We meticulously explore the system’s design choices, training methodologies, and fine-tuning processes, unveiling the robust foundation enabling ScholarFace to process diverse facial appearances and generate contextually rich scholarly information while ensuring that the highest ethical standards are met.

Moreover, we also demonstrated ScholarFace’s usability in real-world scenarios. Our design choices may substantiate ScholarFace as a potent tool with tremendous potential for students and researchers, both in academic and industrial settings. The responsible application of this technology can optimize networking and collaboration, fostering a more connected and informed global scholarly community.

In conclusion, ScholarFace represents a significant advancement in harnessing computer vision and language technologies to empower scholars, researchers, and knowledge seekers with a seamless approach to accessing scholarly information. As we continue to explore the intricacies of ScholarFace’s design, implementation, and evaluation, we remain optimistic about the transformative impact this system can have on academic search engines while upholding the highest ethical standards.

REFERENCES


